MicroRNA TARGET PREDICTION AND COMPARATIVE MiRNomics: STUDY IN DIFFERENT CANCERS- SPECIAL FOCUS ON BREAST CANCER METASTASIS

A Dissertation submitted to the Central University of Punjab

For the award of

Master of Philosophy

In

Biosciences

BY
Bibekananda Sarkar

Supervisor

Dr. Felix Bast

Centre for Biosciences
School of Basic and Applied Sciences
Central University of Punjab, Bathinda

August, 2012

CERTIFICATE

I declare that the dissertation entitled “MicroRNA TARGET PREDICTION AND COMPARATIVE MiRNomics: STUDY IN DIFFERENT CANCERS- SPECIAL FOCUS ON BREAST CANCER METASTASIS” has been prepared by me under the guidance of Dr. Felix Bast, Assistant Professor, Centre for Bioscience, School of Basic and Applied Sciences, Central University of Punjab. No part of this dissertation has formed the basis for the award of any degree or fellowship previously.

Bibekananda Sarkar
Centre for Biosciences
School of Basic and Applied Sciences
Central University of Punjab,
Bathinda - 151001.

Date:

CERTIFICATE

I certify that Bibekananda Sarkar has prepared his dissertation entitled “MicroRNA TARGET PREDICTION AND COMPARATIVE MiRNomics: STUDY IN DIFFERENT CANCERS- SPECIAL FOCUS ON BREAST CANCER METASTASIS”, for the award of M.Phil. degree of the Central University of Punjab, under my guidance. He has carried out this work at the Centre for Biosciences, School of Basic and Applied Sciences, Central University of Punjab.

Dr. Felix Bast
Centre for Biosciences,
School of Basic and Applied Sciences,
Central University of Punjab,
Bathinda - 151001.

DATE:

ABSTRACT

MicroRNA Target Prediction and Comparative MiRNomics: Study in Different Cancers- Special Focus on Breast Cancer Metastasis

Name of student	: Bibekananda Sarkar
Registration Number	: CUP/M.Ph-Ph.D/SBAS/BIO/2010-11/08
Degree for which submitted	: Master of Philosophy
Name of supervisor	: Dr. Felix Bast
Centre	: Biosciences
School of Studies	: Basic and Applied Sciences
Key words	: microRNA, MetastamiR, EMT, MiRBase, Precision, and Sensitivity.

Cancer or carcinoma is uncontrolled growth of abnormal cells. The transition of normal cell to cancerous cell is very complicated process and the cause of transition varies with situation. Metastasis is the main cause of death in cancer. There are genes which not only promote metastasis but also maintain microenvironment of tumor cells, and initiate the process of epithelium–mesenchymal transition (EMT). MicroRNA (micro RNA) is small, highly conserved noncoding RNAs that control gene expression post-transcriptionally. MicroRNA controls various cellular events like division, differentiation and apoptosis. Their deregulation may result in to cancerous growth. Most (50%) of the microRNA genes are located in the fragile chromosomal regions, which are more susceptible to amplification, deletion or translocation during tumor development. It is predicted that 30% of the all mRNA are directly or indirectly controlled by the microRNA. MetastamiR are that microRNA which have role in regulation of metastasis. MicroRNA expression profile in different cancers showed that it can act as Oncogene as well as Tumor suppressor gene. Till date there are 1921 mature human microRNA sequences registered in miRBase. MicroRNA target prediction is the first step in functional analysis of microRNA. Target prediction is complicated due to partial complementarity between microRNA and its target. There are many target prediction programs available, but the efficiency and sensitivity of these programs are not known. To enhance its efficiency, we need to know problems during prediction. Comparative analysis of different microRNA prediction tools provides an insight into the above parameter. In this study comparative analysis of seven prediction tools is carried out with help of validated microRNA targets of metastatic breast cancer.

Bibekananda Sarkar Dr. Felix Bast

ACKNOWLEDGEMENT

It gives me immense pleasure to pay my deep sense of gratitude and thanks to Revered Prof. P. Ramarao, Dean, School for Basic and Applied Sciences, Central University of Punjab for their valuable guidance, consistent inspiration, and support. Their continuous encouragement and inspiration made this work possible in the present form.

I am highly grateful to Dr. Felix Bast, Assistant Professor, Supervisor, for his guidance, supervision, continuous support and consistent deep insights on this project to make this work possible in the present form.

I offer my sincere respect to Prof. R.G. Saini, Coordinator of Centre for Biosciences, for their moral support.

All these studies were not possible, if required facilities are not provided to me, for this I would like to thank our esteemed Vice Chancellor.

My vocabulary utterly fails to express deep sense of gratitude to Dr. Sanjeev Thakur, Dr. Pankaj Bharadwaj and Dr. Sandeep Singh for their guidance and constant inspiration. I thankful to computer lab in charge Mr. Pritpal Singh for his support.

I pay my special thanks to all the friends, non-teaching staff in our Centre for Bioscience for their kind cooperation and support in the hour of need.

I would like to express my immeasurable appreciation and sincere thanks to My Parents and beloved ones for their love, trust, inspiration, encouragement and endless moral support during my work. This thesis is dedicated to them.

Bibekananda Sarkar

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Sr. No.	Full form	Abbreviation
1	B cell lymphoma	Bcl2
2	Tropomyosine1	TPM1
3	Programmed Cell Death 4	PDCD4
4	Sex determining region Y-box 4	Sox4
5	Zinc Finger E-Box Binding Homeobox	ZEB
6	HomeoboxD10	HOXD10
7	High Mobility Group AT-Hook-2	HMGA2
8	Intrigrin α3	ITGA3
9	Untranslated Region	UTR
10	MicroRNA	miR
11	Argonaute	Ago
12	RNA Induced Silencing Complex	RISC
13	DiGeorge Syndrome Critical Region	DICER
14	Chronic Lumphocytic Leukemia	CLL
15	Matrix Metalloproteinase	MMP16
16	Insulin Like Growth Factor Binding Protein	IGFBP2
17	True positive	TP
18	False Positive	FP
19	Untranslated Region	UTR

Chapter One

Introduction

Cancer or carcinoma is uncontrolled growth of cells with abnormal morphology. Cancer is break down of regulatory mechanisms that govern normal cell behavior. The proliferation, multiplication, differentiation and survival of individual cells in multicellular organisms are strictly regulated to meet the needs of the organism as a whole. The fundamental abnormality of resulting in development of cancer is the continual uncontrolled proliferation of cancer cells. There are six hallmarks which differentiate normal cells from cancerous cell (Hanahan & Weinberg, 2011): (i) Sustained proliferating signaling; (ii) Evading growth suppressors; (iii) Resisting cell death; (iv) Enabling replicative immortality; (v) Inducing angiogenesis, and (vi) Activating invasion and metastasis. These six hallmarks are only present in cancerous cells. Recent studies indicates genome unstability as a reason making these cells immortal (Negrini et al., 2010).

Since the last century, scientists have been conducting research on cancer to liberate humankind from one of the most deadly diseases. Global cancer statistics based on GLOBOCAN, 2008 estimates about 12.7 million cancer cases and 7.6 million cancer deaths have occurred in 2008. Breast cancer is the most frequently diagnosed cancer and lung cancer having highest mortality (DeSantis et al., 2011). In recent years death due to breast cancer decreased in North America and European countries, largely due to early detection and improved treatment. In some Asian and African countries including India, the incidence and mortality rate of breast cancer have been rising (Jemal et al., 2011). Seventy percent of deaths due to cancer are occurring in developing countries (WHO, 2011). Death in some cancers like breast, colorectal and cervical can be preventable to a large extent if early diagnosed and treated properly (Chaudhry, 2001).

Major steps in cancer development include initiation, promotion, malignant conversion, progression and metastasis (Valastyan, 2012). Metastasis is the principal cause of death in cancer and the biggest obstacle in cancer treatment. Metastasis involves multiple sequential steps; cells enter into circulatory system, evade the immune system and survive during transportation until it reaches in a secondary organ, where it starts growth under the influence of local growth factors.

1.1 MicroRNA and Breast Cancer

MicroRNA play important regulatory role in differentiation, apoptosis, proliferation and many cellular events. Downregulation or upregulation of microRNA may play important role in cancer development and progression (Dykxhoorn, 2010). The new term miRNomics is used to refer a novel sub-discipline of genomics that focuses on identification, target prediction and validation, and biological function of microRNA on genomics scale (Ghosh et al., 2007). MicroRNA in the circulatory system functions as potential and feasible biomarkers in early stage breast cancer detection. Classically, the cancer is considered as deregulation of different proteins and its study focuses only on different proteins and protein-protein interactions. Focus of cancer research transformed when Victor Ambros and his colleagues discovered the gene lin-4 that affects development in Caenorhabditis elegans and its product is a non-coding RNA (Lee et al., 1993). Non coding RNA in definition includes all functional RNAs that are not translated in to protein products. Examples are microRNA, long intergenic RNA, small interfering RNA (siRNA) and also rRNA and tRNAs. MicroRNAs are small, highly conserved non-coding RNA molecule (ncRNA) of length ~21-25 bases. Originally it was discovered in C. elegans, but, later it was found to be wide spread in most eukaryotes including humans. It account for 1-5% of human genome but regulate (directly or indirectly) 30% of protein coding genes (Drakaki & Iliopoulos, 2009). To date 16772 microRNA reported in 152 species and out of that 1921 mature microRNA sequences belongs to Homo sapiens (miRBase, 2012). Most (50%) of the microRNA genes are located in the fragile chromosomal

regions, which are more susceptible to amplification, deletion or translocation during tumor development (Rodriguez et al., 2004).

Breast cancer is a heterogeneous disease that encompasses a range of phenotypically distinct tumor types. Underlying this heterogeneity is a spectrum of molecular alterations. Breast cancer is one of the diverse diseases and its classification is based on treatment response and prognosis. Classification scheme of breast cancer typically includes histopathological type, TNM (Tumor, Lymph node and Metastasis) stage and receptor status. Role of microRNA in breast cancer was first demonstrated in 2005 (Iorio et al., 2005). These RNAs are involved in regulating gene expression at posttranscriptional level. Functional role of microRNA is determined by analyzing microRNA-mRNA interaction. MicroRNA directly acts on the mRNA so as to control various cellular activities. Researcher had initially determined the microRNA targets through classical genetic screens; however these experimental techniques are very expensive.

Early diagnosis of breast cancer is an important necessity of public health system, as if diagnosed early; most of the cancers are preventable. For this a reliable biomarker is needed. MicroRNA is a potential candidate biomarker in the early detection of various cancers, including breast cancer. MicroRNAs are more stable due to their small size compared to mRNA, allowing expression profiling from fixed tissues or other biological material, and thus supporting their possible use as novel, minimally invasive and robust biomarker (Corcoran et al., 2011). MicroRNAs are intimately linked with the regulation of development and cellular differentiation. Therefore, the dysregulation of microRNA expression is associated with oncogenic transformation. MicroRNAs involved in cancer may be broadly classified into oncogenes (e.g: miR-155 and miR-21) and tumor suppressor genes (e.g. miR-15a, miR-16-1). MicoRNA also regulates specific steps in metastatic pathways, known as “metastamiR”.

Cancer metastasis can be undetectable and can remain latent for many years after removal of primary tumor. The process of metastasis is unpredictable phenomenon and some tumors are fast metastatic while some are slow metastatic. In breast cancer early detection and prognosis after the tumor removal is very important. Ninety percent of death in cancer is due to the metastasis (Chaffer & Weinberg, 2011). Cancer metastasis can be divided into seven steps: Local Invasion, Intravasation, Survival in the circulation, Arrest at the distant organ site, Extravasation, Micrometastasis formation and Metastatic colonization (Valastyan, 2012).

Detection of microRNA expression levels in human breast cancer had been a challenging task. In absence of high-throughput techniques, it has become inevitable to utilize in-silico techniques for microRNA functional analysis (Witkos et al., 2011).

1.2 MicroRNA Target Prediction in Breast Cancer

There are now more than 1900 microRNA sequences registered in miRBase database, but out of that only a few microRNA targets are known to us. In this case computational target prediction is helpful. In animal system, microRNA target prediction is most complicated due to partial complementarity or incomplete base pairing.

Initial experimentally validated microRNA targets showed that there is enough similarity between all microRNA targets, and one clever computational technique might enable their discovery in-silico possible. Till now numerous programs have been independently developed by many laboratories for computational prediction. The rules for targeting transcripts by microRNA have not been fully examined yet and are based mainly on experimentally validated microRNA-mRNA interactions. There is lot of issues which hampers perfect target prediction. Some of these are mentioned below:

1. microRNA Size
2. Identification of 3'UTRs
3. Conservation analysis
4. Presence of multiple targets
5. Incomplete hybridization

These issues make target prediction a difficult task. It is difficult to assess accurately the performance of the target prediction methods. The main problem is very less number of validated targets as compared to predicted targets. So, the initial trials to assess the target prediction tools are not so effective. There are very few trials to compare the performance of different methods, but few comparative study compared these programs; miRanda, TargetScan, TargetScanS, PicTar, RNAhybrid, EiMMo and DIANAmicroT (Sethupathy et al., 2006; Alexiou et al., 2009; Zhang & Verbeek, 2010), All these algorithms predict large number of putative targets, but out of that only few are correct.

Main objective of this work is to study different microRNA target prediction programs and assess them on basis of validated microRNA target involved with the breast cancer metastasis. In order to determine the functional role of microRNA, analyzing microRNA-mRNA interaction is a prerequisite. Important aspect of microRNA regulation is its direct interaction with its target. microRNA:target interaction is best studied by determining its target gene. Assessment of microRNA target prediction programs is very necessary, as it will demonstrate the performance of current target prediction algorithms and help for functional analysis of microRNA.

Chapter Two

Review of Literature

Cancer is a major public health problem in the India and many other parts of the world. Hundreds of thousands of people develop cancer every year. Yet in spite of this, the picture is not clear. Today, most of the cancers can be completely cured if caught at an early stage. According to World Health Organization (WHO), Cancer is the uncontrolled growth and spread of cells. It can affect almost any part of the body. The growths often invade surrounding tissue and can metastasize to distant sites (WHO, 2012). Cancer is the second deadly disease after cardiovascular disease (Jemal et al., 2011).

2.1 Cancer Statistics

Cancer is group of diseases with similar characteristics. Cancer having worldwide distribution and occur in both sexes and also having no age limit. The main causes of cancer epidemics are urbanization, industrialization, change in lifestyle and population growth. The highest increase among females was for cancer of the breast and among males for cancer of the prostate. Breast cancer is the most frequently diagnosed cancer in female and also leading cause of death worldwide, accounting for 23% of the total cancer deaths and 14% of total death due to cancer in 2008 (GLOBOCAN, 2008). Lung cancer was the most commonly diagnosed cancer as well as leading cause of death due to cancer in male in 2008 globally. Among females lung cancer is fourth most commonly diagnosed cancer and second leading cause of death due to cancer. Prostate cancer is the second most commonly diagnosed cancer in male and sixth leading cause of cancer deaths in male. India has world's highest number (nearly 20%) of oral cancer patients. It is aptly labeled the oral cancer capital of the world with an estimated nearly 1% of the total population having oral premalignant lesions (Byakodi et al., 2011). According to Ministry of health and family welfare, nearly 184 million Indians are using tobacco and total number of

tobacco users in the world has been estimated at 1.2 billion, which is expected to rise to 1.6 billion during 2020 (Chaudhry, 2001). Today cancer became epidemic and it is spreading day by day in new locations. Most of the cancer deaths are occurring in developing countries due to low levels of medical facility and lack of early detection facility.

2.2 Early Diagnosis of Cancer

Most of cancer cases can be prevented if diagnosed early. For early detection major requirement is a reliable biomarker which helps us in screening, detection as well as in prognosis; i.e. Universal Biomarker. Circulating microRNAs can act as novel minimally invasive biomarkers for cancer (Heneghan et al., 2010). MicroRNAs are more stable due to their small size compared to mRNA, allows expression profiling from fixed tissues or other biological material. MicroRNA can be extracted and detected from frozen and paraffin embedded tissues, from blood (Mitchell et al., 2008; Roth et al., 2010), circulating exosomes, and different biological fluids like urine, saliva (Kroh et al., 2010; Park et al., 2009) and even sputum.

Since the discovery of the founding members of the microRNA (miRNA) family, lin-4 and let-7, (Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993), hundreds of microRNAs have been identified in plants, animals, and viruses by the use of different molecular cloning and bioinformatics approaches. These microRNAs have ubiquitous in distribution in all organisms. In last decade microRNA is one of the main focus of research and it being implicated in different diseases along with different cancers. Both loss and gain of cellular as well as circulatory microRNA contribute to cancer development through downstream regulations (Dombkowski et al., 2011). MicroRNA regulates the mRNA by incomplete hybridizations at 3’ UTR of target mRNA. This hybridization of microRNA-mRNA induces silencing complex cause degradation of the mRNA. When the hybridization is incomplete then the target mRNA is regulated by translational silencing. A genomic analysis of microRNA has revealed that more than 50% of human microRNA genes are located within intronic regions of annotated protein coding regions (Calin et al., 2004; Rodriguez et al.,

2004). MicroRNA has critical function in gene expression, and their dysregulation may cause tumor formation and progression. A single microRNA may target more than one mRNA; conversely one mRNA may be targeted by multiple microRNA with variable efficiencies. Thus, tumerogenic phenotype needs some significant changes in miRNome content of the cell (Ghosh et al., 2007)

2.3 Biogenesis of microRNA and Functional Relevance

Most of the microRNA transcribed from noncoding intronic part of genome and historically these regions were referred to as “junk DNA” because their function was not known to us. Primary transcript (pri-microRNAs) contains one or more 70 bases long hairpin microRNA precursor (pre-microRNA). Approximately 50% of mammalian microRNA loci are found in close proximity to other microRNAs. These clustered microRNAs are transcribed from a single polycistronic transcription unit (TU) (Lee et al., 2002). MicroRNA is important regulator in most of the cellular events, such as division, proliferation, differentiation and apoptosis. They regulate these events in post transcriptional level or in translational level. MicroRNA may act directly or indirectly via different target pathways.

In human microRNA biogenesis can be divided into two pathways, DGCR8 and riboendonuclease III dependent; i.e., canonical microRNA and DGCR8 and riboendonuclease III independent pathways, in second class we have mirtrons and tailed mirtrons. It releases their pre-microRNA by splicing and exonuclease trimming. MicroRNA gene transcribed to produce a primary microRNA and processed to precursor microRNA and finally releasing mature RNA (MacFarlane & Murphy, 2010). The microRNA biogenesis is not universal and it differs from one microRNA to other (Winter et al., 2009). The main process of microRNA biogenesis can be divided in to two broad categories.

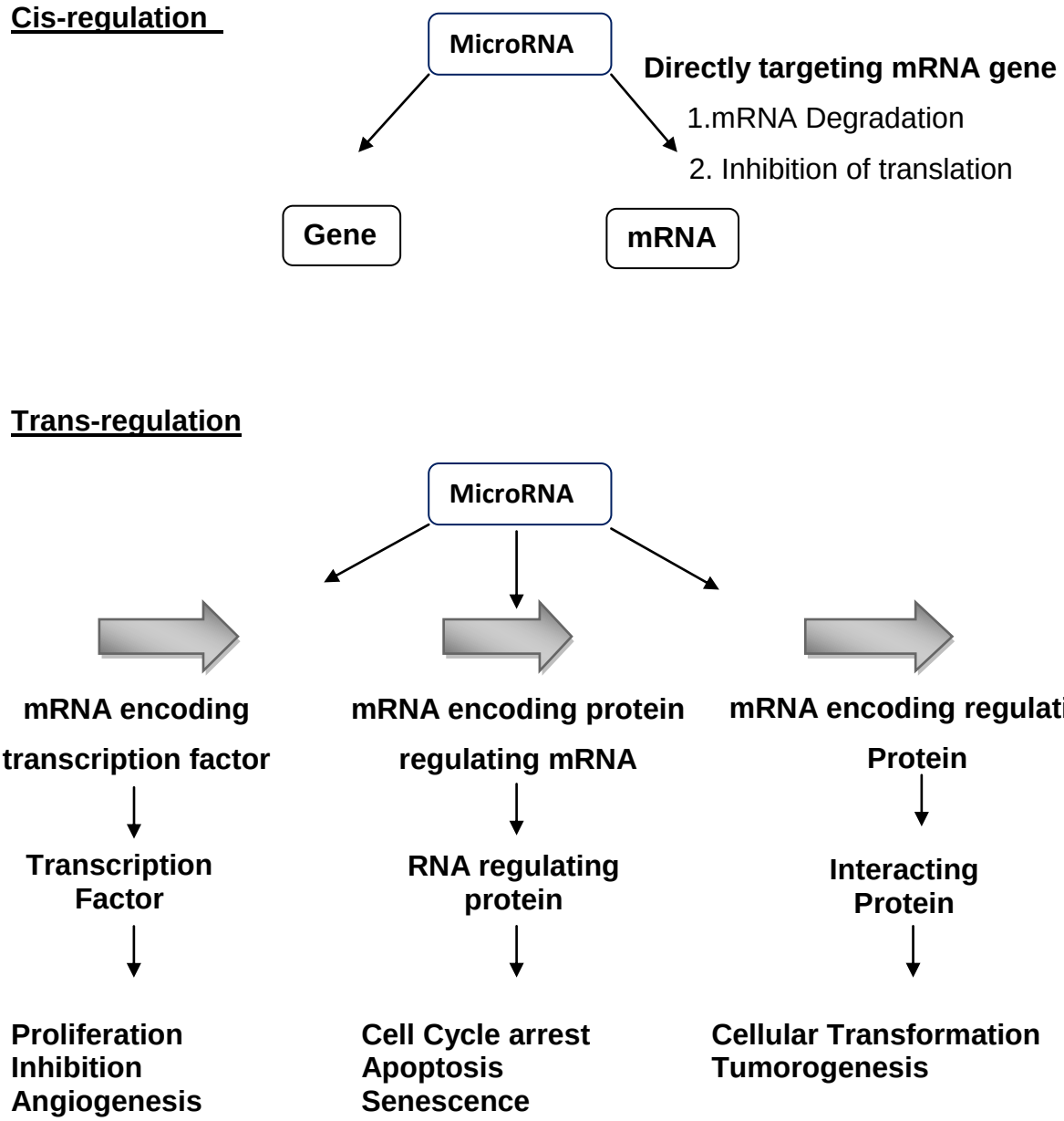

Figure 2.1: MicroRNA regulation in different cellular events. In Cis-regulation: microRNA directly regulates mRNA expression and regulates expression of target gene at post transcriptional level. Tran-regulation: The microRNA regulate level of target genes and regulate different phenomenon but in indirect fashion (Modified from Liu et al., 2009).

2.3.1 Intergenic MicroRNA

Intergenic microRNAs are transcribed by RNA polymerase II or RNA polymerase III. A primary transcript (pri-microRNA) is a large stem-loop structure with single stranded RNA extensions at both ends. Primary-microRNA folds itself 'imperfectly' and creates a stem-loop structure. The maturation process starts from the nuclear cleavage of pri-microRNA by a protein complex known as "**the microprocessor**". This is comprised of the RNase III endonuclease Drosha and DiGeorge Syndrome Critical Region Gene 8 (DGCR8). The auto and cross-regulations between Drosha and DGCR8 may help in maintaining the homeostatic control of microRNA biogenesis (Han et al., 2009). There are additional microprocessor associated proteins which are necessary for the processing a particular microRNA, such as miR-18a which requires the protein factor hnRNPA1.

2.3.2 Coding Intronic MicroRNA

MicroRNA located within an intron of a protein coding gene is transcribed by RNA polymerase-II as part of pre-microRNA (Rodriguez et al., 2004). The process of splicing occurs by two processes. The investigation determined that introns are excised out of the pre-microRNA and debranched by spliceosomal components. An additional hypothesis is that nuclear pathway possible for those small (~50-200 nt.) excised debranched introns containing microRNA, which have the structure to support hairpin formation, thus bypassing the Drosha-DGCR8 step.

2.3.3 MicroRNA Nuclear Export

The microRNA after its processing in nucleus transferred to cytoplasm via Exportin 5 (Exp5) and its Ran-guanosine triphosphate (Ran-GTP) factor (Lund et al., 2004). The Exp/Ran-GTP complex have high affinity towards pre-microRNA, protecting them from the moment they are generated in the nucleus until they are ready for the next cleavage step in the cytoplasm, where GTP hydrolysed to GDP and Exp/Ran-GDP complex release its cargo to cytoplasm.

2.3.4 Cytoplasmic Processing of MicroRNA

In the cytoplasm, the pre-microRNA is cleaved by another RNase III type class III enzyme, Dicer, which is a 200 KDa multidomain protein (Garofalo & Croce, 2011). These domain help in different functions, these domains includes an RNA helicase/ ATPase domain, DUF283 and PAZ signatures, two neighboring RIID (RIIIa and RIIIb), and a dsRBD. The dsRBD and RIIIDs are most certainly involved in the binding and cleavage of double stranded RNA. These all protein participate in selection of microRNAs to their final stop, the microRNA induced silencing complex which mediate in the degradation or translation inhibition of mRNA target gene.

At the core of RISCs are Ago family proteins (Garofalo & Croce, 2011). These are four subfamilies of Ago protein (Ago 1-4), only Ago 2 possesses target cleavage (silencer) activity. In flies, both Ago 1 and Ago 2 have silencer activity. In mammals, microRNAs guide the RISC to complementary target sites in mRNAs, where endonucleolytically active Ago proteins cleave the RNA (Martinez et al., 2002).

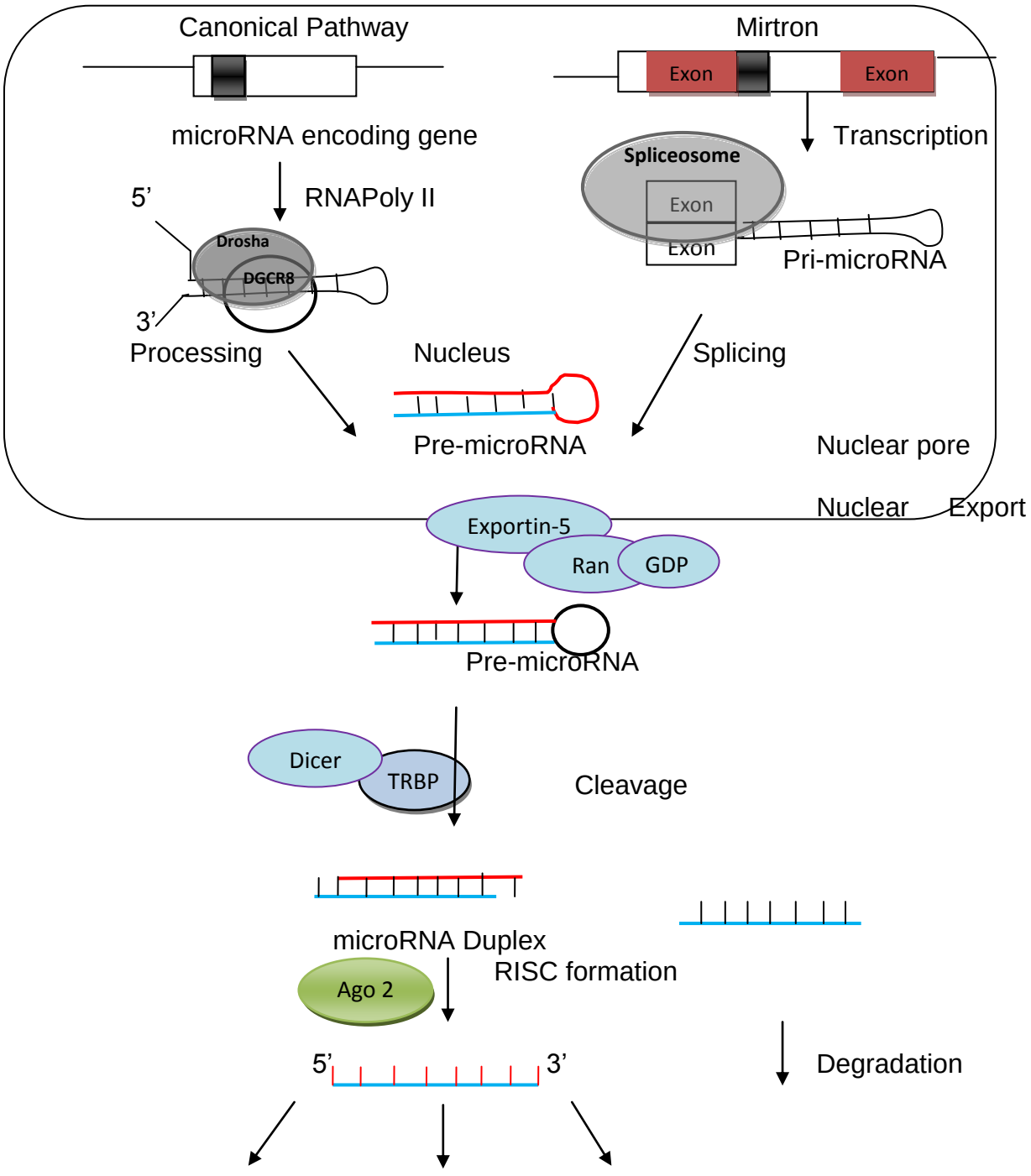

Figure 2.2: A schematic representation of the microRNA processing pathway, beginning at the transcription level as Canonical and Mirtron pathway. Processing starts from pri-microRNA stage and progressing to the pre-microRNA stage, and final processing in

cytoplasm, and incorporation into the RISC protein complex. Modified from (modified from Kim et al., 2009; Krol et al., 2010; Westholm et al., 2011).

2.4 MicroRNA and Cancer (oncomiR)

It is now clear that microRNA have predominant role in carcinogenesis (Iorio & Croce, 2012). To know exact role of microRNA in carcinogenesis we have to estimate exact expression of microRNA in normal and in cancer tissue. Sun et al. showed that microRNA regulates different cellular mechanisms such as cell growth and apoptosis (Sun et al., 2010). MicroRNA mostly originates from (60%) intronic part of genome. Heterochromatic region is noncoding in nature and they are defined as junk DNA and have no function. The noncoding RNA never codes for any protein product and have regulatory activity in cell cycle and cellular activities. The significance of microRNA and its role in cancer began to be revealed in 2002 when Croce and colleagues identified that a small genomic region in chromosome 13q14 that is commonly deleted in chronic lymphocytic leukemia (CLL) contained miR-15a and miR-16-1 genes (Calin et al., 2002). These suggested that there might be some link between microRNA and cancer occurrence. Research on microRNA and cancer showed aberrant expression of different microRNAs in cancer cells and it confirmed the role of microRNA in oncogenesis.

The microRNA in cancer having wide role and it can be used as diagnostic or prognostic biomarker. It can help us in early detection of cancer and act as reliable biomarker for all cancers. The microRNA expression profile can be used for the classification and stratification of cancers (Lu et al., 2005). The aberrant expression of microRNA in cancer sample can be examined when it is compared with normal tissue samples. The detection of microRNA is generally done by the high throughput techniques, such as microarray analysis, Northern blotting and Real Time PCR analysis. Cancer development is a multi-factorial change and these factors can be classified in to tumor suppressor and tumor inducer. The cancer formation is combination of these two factors. Several expressions showed that microRNA act as tumor suppressor as well as oncogene. MicroRNAs; whose expression increase in

tumor is considered as oncogene. These oncogenes (microRNA) are known as OncomiR and promote tumor development by negatively inhibiting tumor suppressor genes. In some cancer the expression of microRNA level goes down so they are called as tumor suppressor.

2.4.1 MicroRNA Modulation of Tumor Suppressor and Oncogenic Pathways

Many microRNAs have been shown to function as oncogene as well as tumor suppressor gene and they have strong influence on occurrence of disease. Over expression and amplification are the main criteria in defining microRNA as oncomiR.

2.4.1.1 MicroRNA as Oncogenes

The most abundant cancer related oncomiR is miR-21. miR-21 uprergulates in most of the cancers, such as pancreatic, lymphoma and lung cancers (Zhu et al., 2008). miR-21 is located on chromosome 17, and regulated by active STAT3, which is erroneously expressed and activated in many cancer types. miR-21 can act as an oncogene by regulating the tumor-suppressor genes PTEN (Meng et al., 2007) and PDCD (Frankel et al., 2008) by exerting their anti-apoptotic activity. Inactivation of miR-21 in several cell lines resulted in increased cell death by reactivating Caspases as well as an increased sensitivity to Gemcitabine via activation of PTEN.

miR-221/222 are located in chromosome X and one of the example oncomiR cluster. They get upregulate in different cancers especially in prostate cancer (Sun et al., 2009). The miR-17-92 cluster was among the first microRNA discovered to be deregulated in number of human malignant tumors, including lyphomas, lung and breast cancers. This cluster contains six microRNA and they are transcribed as single poly-cistron. The upregulation of these microRNA correlated with increased levels of cellular proliferation and decreased level of apoptosis.

2.4.1.2 MicroRNAs as Tumor-Suppressors

The let-7 family of 12 microRNA is often cited as the archetypal tumor-suppressing microRNA. Let-7 is essential for cell type determination during embryogenesis in C. elegans (Reinhart et al., 2000). Let-7 expression is lost at an early stage in breast cancer progression (Sempere et al., 2007), and continued expression is associated with low-grade, ER-positive, luminal A tumors (Blenkiron et al., 2007). Several functional targets of let-7 have been identified, including the classical proto-oncogene Ras (Johnson et al., 2005), and the oncofoetal proteins HMGA2 (Mayr et al., 2007) and IMP-1(Boyerinas et al., 2008).

miR-34 (a family of three miRs) is transcriptionally upregulated by p53, and targets a range of genes involved in cell proliferation and apoptosis, including Bcl-2 (Bommer et al., 2007; He et al., 2007). miR-125a and miR-125b were identified early in the development of the field as being downregulated in breast tumors (Iorio et al., 2005), itself an independent prognostic marker in breast cancer. miR-125 act as tumor suppressor and have wide spread role in cancer development.

The miR-200 family of 5 microRNA is suppressors of Epithelial-mesenchymal transformation (EMT). They directly modulate ZEB1 and ZEB2, indeed control transcription of numerous genes involved in maintaining epithelial cell polarity, down regulating E-cadherin and inducing vimentin. ZEB1 also suppresses the transcription of at least two members of the miR-200 family (Burk et al., 2008), establishing a feed-forward mechanism that would be expected to maintain the mesenchymal phenotype.

2.4.1.3 MicroRNA with Both Oncogenic and Tumor Suppressor Function

There are many microRNA which overexpress in one cancer while down regulate in other type of cancer. miR-205 up regulates in lung, bladder and pancreatic cancers, but down regulate in prostate cancer and breast cancer. There is long list of microRNA which differentially express in different cancers.

Table 2.1: MicroRNA Regulating Cancer

MicroRNA	Cancers		Reference
	Up regulation	Down regulation	
miR-1		Lung	(Nasser et al., 2008)
miR-7		Glioblastoma	(Kefas et al., 2008)
let-7		Liposercoma	(Yu et al., 2007)
let-7a		Breast, ovarian, CLL	(Sempere et al., 2007; Iorio et al., 2007),
let-7b		Ovarian	(Nam et al., 2008)
let-7c		Burkitt Lymphoma, ovarian, prostate	(Leucci et al., 2008; Iorio et al., 2007; Ozen et al., 2007)
let-7d	Pancreatic	Breast, ovarian	(Iorio et al., 2005; Iorio et al., 2007; Lee et al., 2007)
miR-9	Bladder	lung, ovarian	(Ichimi et al., 2009),(Yanaihara et al., 2006; Iorio et al., 2007)
miR-10a	Pancreatic, hepatocarcinoma		(Bloomston et al., 2007; Varnholt et al., 2008a)
miR-10b	Glioblastoma, pancreatic, Breast.	Breast, ovarian	(Bloomston et al., 2007; Ciafrè et al., 2005; Iorio et al., 2005; Nam et al., 2008)
miR-15/16		CLL, DLBCL, multiple myeloma, pituitary	(Bonci et al., 2008)
miR-17-3p	Lung		(Yanaihara, et al., 2006)
miR-17-5p	Bladder, neuroblastoma, colon,		(Díaz et al., 2008; Fontana et al., 2008; Gottardo et al., 2007)
miR-20a	Ovarian		(Nam et al., 2008)
miR-21	Breast, lung, prostate, gastric, cervical, head and neck, and colorectal cancer.		(Krichevsky & Gabriely, 2009)
miR-23a	Bladder, ovarian, pancreatic		(Bloomston et al., 2007; Gottardo et al., 2007; Nam et al., 2008)
miR-24-1	Pancreatic , thyroid		(Lee et al., 2007; He et al., 2005)
miR-24-2	Lung, pancreatic, thyroid		(He et al., 2005; Lee et al., 2007; Yanaihara, et al., 2006)
miR-25	Glioblastoma		(Ciafrè et al., 2005)
miR-26a		Ovarian, thyroid	(Nam et al., 2008; Visone et al., 2007)
miR-26b	Glioblastoma		(Ciafrè et al., 2005)
miR-29a		Lung, ovarian	(Fabbri et al., 2007;Nam et al., 2008)

miR-29b	Breast , thyroid		(He et al., 2005; Yan et al., 2008)
miR-29c	Breast , thyroid		(He et al., 2005; Yan et al., 2008)
miR-30a 3p	SCCT		(Wong et al., 2008)
miR-30d		CLL , thyroid	(Marton et al., 2007; Visone et al., 2007)
miR-31	Colorectal, hepatocarcinoma, SCCT	Breast	(Slaby et al., 2007; Wong et al., 2008)
miR-32	Prostate	Lung	(Ambs et al., 2008; Gottardo et al., 2007; Yanaihara, et al., 2006)
miR-33		Lung	(Yanaihara, et al., 2006)
miR-34a	Thyroid	Bladder, breast, colon, lung, melanoma, pancreatic, prostate, renal	(Gottardo et al., 2007; He et al., 2005; Lodygin et al., 2008)
miR34b	SCCT	Burkitt Lymphoma	(Leucci et al., 2008; Wong et al., 2009)
miR-34c	SCCT		(Wong et al., 2009)
miR-99a		Bladder , ovarian , SCCT	(Ichimi et al., 2009; Iorio et al., 2007)
miR-100	Pancreatic	Bladder , ovarian , SCCT	(Bloomston et al., 2007; Ichimi et al., 2009; Iorio et al., 2007; Wong et al., 2009)
miR-101-1		Breast, lung	(Iorio et al., 2005; Yanaihara, et al., 2006)
miR-103-1	Bladder		(Gottardo et al., 2007)
miR-103-2	Pancreatic		(Bloomston et al., 2007)
miR-105		Ovarian	(Iorio et al., 2007)
miR-106	Bladder , lung	Colon	(Díaz et al., 2008; Ichimi et al., 2009; Yanaihara, et al., 2006)
miR-107	Pancreatic	SCCT	(Bloomston et al., 2007; Wong et al., 2009)
miR-221 and 222	Hapetocarcinoma, CLL, melanoma and globlastoma, as well as lung, breast, thyroid, colon, pancreas and prostate cancer		(Volinia et al., 2006)
miR-122	Hepatocarcinoma (hepatitis C-virus associated)	Hepatocarcinoma	(Varnholt et al., 2008a; Wong et al., 2008)
miR-122a	Breast	Hepatocarcinoma	(Iorio et al., 2005; Wong et al., 2008)
miR-123	Glioblastoma		(Ciafrè et al., 2005)
miR-124		Anaplastic Astrocytoma, glioblastoma	(Silber et al., 2008)

miR-124a	SCCT	Lung , ovarian	(Iorio et al., 2007; Wong et al., 2009; Yanaihara, et al., 2006)
miR-125a	Pancreatic	Breast , lung , ovarian	(Bloomston et al., 2007; Iorio et al., 2005; Iorio et al., 2007; Yanaihara, et al., 2006)
miR-125b		Bladder, ovarian, prostate, SCCT, thyroid	(Gottardo et al., 2007; Wong et al., 2009; Yang et al., 2008)
miR-125b-1	Glioblastoma, pancreatic	Breast , ovarian	(Iorio et al., 2005; Iorio et al., 2007)
miR-125b-2	Glioblastoma	Breast	(Iorio et al., 2005)
miR-126		Cervical, colon, hepatocarcinoma , lung, ovarian	(Iorio et al., 2005; Wong et al., 2009; Yanaihara, et al., 2006)
miR-127	AML	Breast, ovarian, Prostate	(Iorio et al., 2005; Iorio et al., 2007)
miR-128a	SCCT	Classic Hodgkin Lymphoma, glioblastoma	(Ciafrè et al., 2005; Navarro et al., 2008; Wong et al., 2009)
miR-128b	Breast	Classic Hodgkin Lymphoma	(Iorio et al., 2005; Navarro et al., 2008)
miR-130a	Glioblastoma		(Ciafrè et al., 2005)
miR-133a	Bladder	Bladder, ovarian, SCCT	(Iorio et al., 2007; Wong et al., 2009)
miR-133b		Bladder, colorectal SCCT	(Wong et al., 2008)
miR-134	SCCT	Ovarian	(Iorio et al., 2007)
miR-135b	Colorectal		(Xi et al., 2006)
miR-143	Pancreatic	Pancreatic Breast, cervical, CLL, colon, colorectal, lung, ovarian	(Bandres et al., 2006; Iorio et al., 2005; Iorio et al., 2007; Wong et al., 2008; Yanaihara, et al., 2006)
miR-145		Bladder, breast, cervical, colon,colorectal, CLL , hepatocarcinoma, lung, ovarian	(Bandres et al., 2006; Iorio et al., 2005; Iorio et al., 2007; Wang et al., 2008)
miR-146	Lung, pancreatic, thyroid		(Bloomston et al., 2007; Visone et al., 2007; Yanaihara, et al., 2006)
miR-146a	Cervical	Prostate	(Lin et al., 2008; Wang et al., 2008)
miR-155	Periatratic Burkitt's lymphoma, Hodgkin's disease, non-hodgkin's lymphoma, and CLL,as well as breast, lung, colon and pancreatic cancers.		(Iorio et al., 2005; Volinia et al., 2006; Yanaihara, et al.,2006)
miR-190	Bladder		(Ichimi et al., 2009)
miR-198	Retinoblastoma, SCCT	Hepatocarcinoma, lung	(Wong et al., 2008;

			Yanaihara, et al., 2006)
miR-200a	Ovarian		(Iorio et al., 2007)
miR-200b	Ovarian		(Iorio et al., 2007)
miR-200c	Colon , colorectal, ovarian		(Nakajima et al., 2006; Xi et al., 2006)
miR-211	Ovarian		(Iorio et al., 2007)
miR-212	Lung , pancreatic		(Yanaihara et al, 2006)
miR-221	Bladder, glioblastom, hepatocarcinoma , lung, melanoma, ovarian, pancreatic, thyroid, Breast		(Bloomston et al., 2007; Garofalo & Croce, 2011; Gottardo et al., 2007; Sun et al., 2009; Varnholt et al., 2008b)
miR-373	Prostate, breast		(Iorio et al., 2005; Yang et al., 2009)
miR-520c	Prostate, breast		(Iorio et al., 2005; Yang et al., 2009)
miR-106B-93-25 cluster	Gastric, prostate, colon and pancreatic cancers. As well as in neuroblastoma and multiple myeloma		(Volinia et al., 2006; Zhang et al., 2006)
miR-17-92 cluster	Lung and colon cancer, as wellas lymphoma, medulloblastoma, and multiple myeloma		(Mendell, 2008; Volinia et al., 2006; Zhang et al., 2006)

2.5 MicroRNA and Cancer Metastasis (MetastamiR)

Metastasis is a most important hallmark of malignant tumors. Metastasis makes management of cancer a big problem in patients. The pathway of cancer metastasis is not clear till now. The movement of cancer cells to distant organ through blood or lymphatic system is controlled by various factors. According Stephen Paget cancer cells are like “Seed” and the distant organ act as “soil”; this hypothesis also called as Paget hypothesis (Fidler, 2003). The process of cancer metastasis is multivariate process controlled by various factors. Metastasis and invasion cascade starts when primary cells invade to the local extra cellular matrix (ECM).

The total metastatic cascade can divided in to two broad phases: (i) physical translocation of cancer cell from the primary tumor to the microenvironment of distant

tissue and the second is (ii) colonization at distant place (Chaffer & Weinberg, 2011). There are six major steps in metastasis and these are:

2.5.1 Local Invasion

Local invasiveness is comprised of the events that permit cancer cells that had previously resided within well encapsulated primary tumors. In order to do so these cells acquire the ability to migrate and invade. These characteristics enable the cells to degrade and move through the surrounding tissue toward blood and lymphatic vessels, which in turn provide highways for their migration in to distant secondary sites. First indication of migration comes from the detection of lymph node. Lymph nodes are the place where these cells stop for some time in the course of migration. The migration process starts with the degradation of basement membrane (BM). The proteolysis is typically enacted by matrix metalloproteinase (MMPs), whose actions are frequently hyper activated during the course of malignant progression (Noh et al., 2012). Degradation of basement membrane stimulates the cells to migrate as “collective invasion” or, can deploy either of two alternate single-cell invasion strategies known as “mesenchymal invasion” and “amoeboid invasion” (Symons & Segall, 2009).

Numerous microRNAs have been found to affect the local invasiveness of cancer cells. The most prominent microRNA which regulate invasion is miR-10b. miR-10b promotes the invasion of breast carcinoma cells via suppression of its downstream target Homeobox D10 (HOXD10) and consequent upregulation pro metastatic gene, RHOC (Ma et al., 2007). Similarly, miR-373/520 promotes tumor invasion and metastasis through pathways that control the adhesion and signal transduction molecule CD44 in breast cancer (Huang et al., 2008). There are many microRNA which regulate metastatic invasion through MMPs. miR-21 inhibits expression of the MMP antagonist tissue inhibitor of metalloproteinase-3 (TIMP3) (Zhu et al., 2008). Conversely, the miR-29b suppress invasion due to its ability to suppress MMP2, while miR-31 can post-transcriptionally suppress Matrix Metallopeptidase 16 (MMP16) levels (Valastyan et al., 2009).

Epithelial-mesenchymal transition (EMT) is an important pivotal program that induces rapid change in the shape and motility of epithelial cells (Polyak & Weinberg, 2009)**.** During this cell lose their cell-cell and cell-matrix contact polarity, gain the disseminated cell migration ability and become motile mesenchymal cells. MicroRNA regulate EMT pathway, most prominent of these are miR-200 family. They directly modulate ZEB1 and ZEB2, indeed control transcription of numerous genes involved in maintaining epithelial cell polarity, down regulating E-cadherin and inducing vimentin. ZEB1 regulate the transcription of two member of miR-200 family. This reciprocal relationship aids in explaining the apparent robustness of the miR-200 status of a cell for defining its epithelial versus mesenchymal nature, as this control the cell either in mesenchymal or in epithelial state. There are other microRNA which regulate EMT, miR-9 by targeting E-cadherin, miR-9 level were significantly elevated in primary tumors from patients with diagnosed metastases in comparison with those from metastasis free patients (Ma et al., 2010) and the ZEB1 and ZEB2 suppressing miR-205 (Gregory et al., 2008).

2.5.2 Intravasation

The term intravasation describes the cellular events whereby locally invasion tumor cells enter into the circulatory system of lymphatic or blood vessels. Carcinomas initially metastasize in local lymph nodes and only at a later stage in other organs. The hematogenous metastasis in carcinomas develops as a result of vascular invasion and penetration at primary tumor site or whether they are derived from cells that have localized lymph nodes. In the first case, the tumor cells need to degrade the vascular basement membrane and interrupt into the circulation. The structural anatomy of tumor associated blood vessels is often quite distinct from that of the normal healthy vasculature. Most of the cases tumor associated blood vessels are more tortuous, leaky, and continuously in a state of reconfiguration (Carmeliet & Jain, 2011). Intravasation is a rate limiting step and controlled by MMP9. Lymph nodes are first site for cancer metastasis.

Role microRNA in cancer intravasation is not clear, but microRNA-21 plays some role in cancer intravasation via programmed cell death-4 (PDCD-4) encoding mRNA.

2.5.3 Survival in Circulation

The blood stream is a harsh environment for metastatic tumor cells because of its velocity induced shear force, lack of substratum and presence of immune cells. Tumor cells grow highest when it attaches with substratum. In absence of substratum or cellular adhesion to ECM component, epithelial cells are susceptible to a form of apoptotic cell death known as anoikios. Circulating Tumor Cells (CTCs) travelling through the hematogenous circulation are become prey to host's innate immunity system.

There are evidences that prove that microRNA play pivotal role in the regulation of circulating cancer cell. The sensitivity of breast cancer cells to anoikios is heightened by the action of miR-31, owing to miR-31 conferred suppression of ITGA5, RDX and RhoA (Valastyan et al., 2009). Similarly miR-7 triggers anoikios responses due to the down regulation of EGFR. It clearly shows that microRNA play important role in the intravasation and metastasis of cancer.

2.5.4 Arrest at Distant Location

The circulating cancer tumor cells (CTC) travel through the circulation but the frequency of metastasis is not equal in all anatomical sites. There are many genes which express in metastasis but not in primary tumors. Tumor cell binding to coagulation factors include tissue factors, fibrinogen, fibrin and thrombin creates a facility for arrest in capillary bed. The endothelial cells E and P selectins also contribute to the tumor cell arrest, other potential mediators of tumor cell arrest are tumor derived, such as glycosylatin pattern and integrins.

miR-223 is found to have crucial roles in regulating granulocyte proliferation and activation by regulating myeloid ELF-1-like factor (Mef2c) (Johnnidis et al., 2008). miR-126 expression is often downregulated in cancers and is able to decrease

leukocyte and possibly cancer cell adherence to endothelial cells by targeting VCAM-1 (Harris et al., 2008).

2.5.5 Extravasation

Extravasion also called as homing, the tumor cells enter in to the parenchyma of distant tissue when it lodged in the capillary bed of foreign tissue. The tumor cell extravasate by inducing endothelial retraction, leading to the attachment of tumor cells to the sub endothelial retraction, leading to the attachment of tumor cells to the sub epithelial ECM and reformation of the capillary. It seems that extravasation process is opposite of intravasation but in true sense these two processes are totally different. Bone marrows are made of single layer of epithelial cells and devoid of supporting mural cell, this route become least resistant to.

The microRNA play vital role in the extravasation in cancer cell metastasis. It regulates extravasation by regulating ITGA5 and RhoA. miR-214 has also been proposed to act during the extravasation step of the metastasis cascade, perhaps doing so via modulation of integrin α3 (ITGA3) and transcription factors AP2γ (TFAP2C) (Penna et al., 2011).

2.5.6 Metastatic Colonization

Just the extravasation of tumor cells in to the parenchyma of tissue does not guarantee their subsequent viability and micrometastasis formation. The successful colonization is the ability to acquire mitogenic stimulation from growths and cytokines that are naturally present in the alien environment. The outcome of successful metastasis is rapidly expanding the micrometastasis that can now serve as the focus for disseminating a shower of secondary metastases. The colonization is complex phenomenon. Prostate cancer most preferentially metastasizes to bone, while CRCs metastasize to liver. The metastatic colonization is a highly organ specific event, that is uniquely dependent on the particular microenvironment of distant organ of where the disseminated tumor cells are endeavoring to colonize.

Several microRNAs have been implicated in the regulation of metastatic colonization. Such as, anti-metastatic microRNA, miR-31 antagonizing the metastatic

colonization of breast carcinoma cells in the lung, by controlling ITGA5 and RDX (Valastyan et al., 2009). Similarly, miR-126 opposes both lung and bone metastasis by breast cancer cells, this process is controlled by the Insulin like growth factor binding protein-2 (IGFBP2).

Table 2.2: Stereotypical Patterns of Tumors to Distant Organs. (Nguyen & Massagué, 2007).

Cancer Type	Site of Metastasis
Breast Cancer	Primarily bone and lung; less frequency liver and brain. ER-positive tumor preferentially spread to bone; ER-negative tumors metastasize more aggressively to visceral organs.
Lung Cancer	The two most common types of lung cancer have different etiologies. Small cell lung carcinoma (SCLC) disseminate rapidly to many organs while non-small cell lung carcinoma (NSCLC) often spread to brain and also to adrenal gland, liver and bones
Prostate Cancer	The carcinoma exclusively migrates to bone, forming osteoblastic lesions and filing the bone marrow cavity with mineralized bone matrix.
Pancreatic Cancer	Aggressively spread to liver and adjacent visceral organ.
Colon Cancer	The mesenteric circulation pattern favors dissemination to liver and peritoneal organ, but metastasis also occurs in the lung.
Ovarian cancer	Locally spread in to peritoneal organs

2.6 MicroRNA and Breast Cancer Metastasis

According to National Cancer Institute (NCI), Cancer that forms in tissues of the breast, usually the ducts (tubes that carry milk to the nipple) and lobules (glands that make milk) are called as breast cancer. Metastasis is the main cause of breast

cancer deaths. Metastasis represents a complex process by which primary solid tumor cells invade adjacent tissue and grow into secondary tumor. Most of the death in breast cancer is due to the development of metastasis, which is culmination of neoplastic progression. In breast cancer also metastasis can be divided into different steps. In each step of metastatic development requires spatio-temporal expression of different proteins. Although the mechanism in metastatic initiation is not known clearly, but only few cells in primary tumor are believed to complete all the steps and give rise metastases. Epithelial to mesenchymal transition is necessary for the progression of benign tumor cells to invasive and metastatic cells.

There are two prevailing models of breast cancer tumorigenesis and cancer cell origin: the stochastic model and the cancer stem cell model of carcinogenesis. The stochastic model is also called as clonal evolution model, postulates that transformation originates from random mutations in any breast epithelial cell and that neoplastic process further evolves through accumulation of random genetic events that drives uncontrolled proliferation and resistance to apoptosis. During the long process the neoplastic cells undergo genetic and epigenetic changes and co-evolve with their micro environment. According to cancer stem cells hypothesis, all tissues are derived from organ specific stem cells that are defined by their capacity to undergo self-renewal. The origin of cancer is not controlled by only one event and it is driven by these two theories (Shackleton et al., 2009).

Breast cancer can be classified into subcategories based on their estrogen (ER), progesterone (PR) and HER receptor status. Sub types were designated as Luminal A, which expessed ER or PR but not HER2; Luminal B, which were ER, PR and HER positive; Basal tumors which were ER, PR and HER negative, preferentially affecting young women and women of African origin.

Table 2.3: MicroRNA Regulating Breast Cancer Metastasis.

microRNA	Important Targets	References
has-let -7 family	HMGA2, Ras	(Mayr et al., 2007)
has-miR-9	E-cadherin	(Ma et al., 2010)
has-miR-10b	HOXD10	(Ma et al., 2007)
has-miR-21	BCL-2, TPM1, PDCD4, PTEN, MASPIN	(Frankel et al., 2008; Meng et al., 2007)
has-miR-101	EZH2	(Varambally et al., 2008)
has-miR-146 family	ROCK1, IRAK1, TRAF6 , BRMS1, NF-κβ	(Hurst et al., 2009)
has-miR-200 family	BMI1,ZEB1, ZAB2, Sox4, Klf4, Sec23a, Sox2	(Burk et al., 2008; Gregory et al., 2008)
has-miR-206	ESR1	(Iorio et al., 2005)
has-miR-221	$p27^{Kip1}$	(Miller et al., 2008)
has-miR-335	Sox4, PTPRN2, MERTK, TNC	(Tavazoie et al., 2008)
has-miR-373	CD44	(Huang et al., 2008)

2.7 MicroRNA Target Prediction

MicroRNA play an important role in the several cellular processes, such as development, cell proliferation, differentiation and apoptosis, and these are crucial in different diseases, including cancer (Garzon et al., 2009). MicroRNA genes are most frequently found in introns, but they can occur in exons or reverse strands of both non protein coding and protein coding genes. The number of genes coding for microRNA is increasing steadily.

To determine the functional role of microRNA, we have to analyze microRNA-mRNA interaction. The important aspect of microRNA regulation is its direct interaction with its target. Due to the limitation in current techniques, high throughput target validation via biological experiments is not practical. Given these circumstances, a lot of algorithms for computational target prediction have been developed. Recently the number of microRNA target prediction algorithms has been

significantly increased. Predictability of computational target prediction methods has also improved but main drawback is lack of overlap between these methods. It suggests that the field of target prediction methods still has a long way to go before converging on the set of rule necessary to identify all functional microRNA targets.

2.7.1 MicroRNA Target Feature

Identification of microRNA target in animal has been very challenging. This is mainly because of the limited complementarity between microRNAs and their targets, which might lead to the finding of hundred of target per microRNA. Therefore combination of experimental and computational approach is best choice. MicroRNA target prediction depend up on the target features, these can be categorized in to six categories, microRNA:mRNA pairing, site location, conservation, site accessibility, multiple site and expression profile.

a) microRNA:mRNA Pairing:-Seed Site is Most Important Feature For Target Recognition

A fundamental challenge to computational prediction of microRNA target is the incomplete base pairing of microRNA to their target sites. MicroRNA target site commonly have at least one region that has Watson-Crick pairing to the 5' part of microRNA. This 5' part, located at positions 2^{nd}- 8^{th} from the 5' end of microRNA, this is known as 'seed' as RISC uses these position for nucleation for recognizing target mRNA (Sturm et al., 2010). The corresponding region in mRNA called as 'seed sites'. A stringent seed site has perfect Watson-Crick pairing and can be divided in to four 'seed' types- 8mer, 7mer- m8, 7mer- A1, and 6mer depending on the combination of the nucleotide of position 1 and pairing at position 8. 8mer has both adenine at position at 1 of the target site and base pairing at position 8. 7mer-A1 has an adenine at position 1, while 7mer-m8 has base pairing at position 8, 6mer has neither an adenine at 1 position nor base pairing at position 8. The presence of adenine at 1 site on target site of microRNA increases the efficiency of target recognition. The

determination of optimal seed rule has become major contention within the field of target recognition.

The first published microRNA target recognition method TargetScan (Lewis et al., 2005) , PicTar (Krek et al., 2005), and MiRanda (John et al., 2004) are emphasized on the identification of conserved matches to microRNA 5'seeds in the 3'UTRs of target transcripts. TargetScan more emphasize on strict microRNA:mRNA match, after seed region.

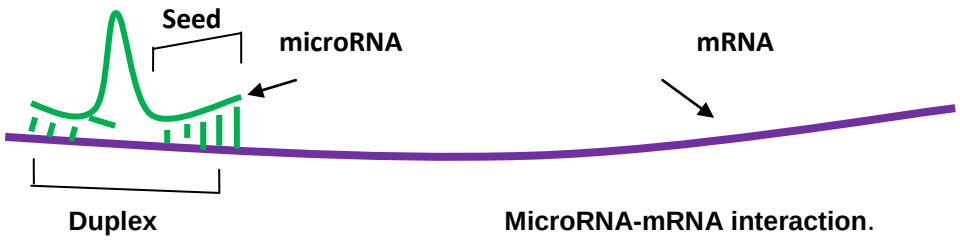

Figure 2.3: The diagram showing the microRNA:mRNA binding and seed matching site (Alexiou et al., 2009).

b) Conservation of binding site microRNA and their targets are conserved among related species

Matching the conservation of target site and microRNA are most important parameters in target prediction. Conservation in microRNA and its target is to identify functional microRNA:target relationships that are important to animals, such that positive selection is acting upon the microRNA target sequences (Hammell, 2010). MicroRNA families are comprised of microRNAs that have the same seed site, and are well conserved among related species. The microRNA targets in related species are also conserved. There are species specific microRNAs and targets. The use of this filter in the prediction algorithm reduce the number of false positive number, but these type of filter good for conserved microRNA and its target only.

c) Thermodynamics of binding site

The formation of stable microRNA:target duplex in vivo must be governed by thermodynamic considerations. Energetically favorable microRNA: target duplexes will remain paired longer, giving time for the RISC proteins to carry out their enzymatic activity. A highly stable RNA duplex is represented as having a minimum free energy of hybridization (MFE). Minimum free energy is usually used to estimate the secondary structure and RNA hybridization, but the amount of A: U s surrounding the site can also be used to estimate the site accessibility. This target prediction parameter is considered by RNA Hybrid (Rehsmeier et al., 2004) and DIANAMicroT.

d) Accessibility of target binding sites

In order to form a microRNA:target duplex. RISC-associated microRNAs must gain access to the binding sites on the microRNA transcript. The secondary structure is very necessary for microRNA targeting. An effective microRNA:mRNA interaction needs an open structure on the target site to begin hybridization reaction. It is reported that microRNA binding sites tend to reside in regions of 3' UTR that predicted to be structurally accessible (Hammell, 2010). MFE calculations that include the energetic cost of disrupting local pairing within the binding site can often distinguish the functional and non functional microRNA binding sites (Thadani & Tammi, 2006).

The target prediction methods that include structural accessibility as a microRNA binding site parameter vary widely in the size of the sequence window used for calculating the accessibility, in the way those calculations are performed, and in the way those calculations are included in the predictions. TargetScan (Lewis et al., 2005). TargetRank (Nielsen et al., 2007) look at the percent of A/U nucleotides adjacent to the seed match of the binding site, which correlates with the local accessibility. Other prediction methods calculate the accessibility into energy parameters referred to as $\Delta\Delta G_{total}$. Accessibility energy represents the open degree

of 3'UTR bounded by microRNA in the thermodynamic view. Lower the accessibility energy is, the more likely the 3' UTR is to be a target.

$$\Delta\Delta G = \Delta G_{duplex} - \Delta G_{open}$$

Where ΔG_{duplex} is the energy gained by the microRNA binding to its targets. ΔG_{open} is the energy required to make the target region accessible for microRNA (Yue et al., 2009)

$$\Delta G_{open} = \Delta G_{free} - \Delta G_{unpaired}$$

e) Multiple Site

Strong microRNA target have multiple target site instead of one single site. Considering the number of putative microRNA site per mRNA can therefore significantly enhance the prediction (John et al., 2004; Pillai, 2005). Although this multiple binding site only additive and enhance efficiency of microRNA target prediction.

f) Expression Profile

Recent development in microRNA target prediction involves analyzing the expression patterns of microRNA and their predicted targets. microRNA express differentially in different tissues. MicroRNA and mRNA are negatively correlated then differential expression level in different tissues can be examined, in this case the mRNA being targeted by microRNA. This idea of microRNA and target expression anti correlation, originally shown by Stark et al. (Stark et al., 2005) and this has been successfully incorporated in different prediction methods. By filtering putative targets on expression profile correlations is an approach to reduce the false positive rate.

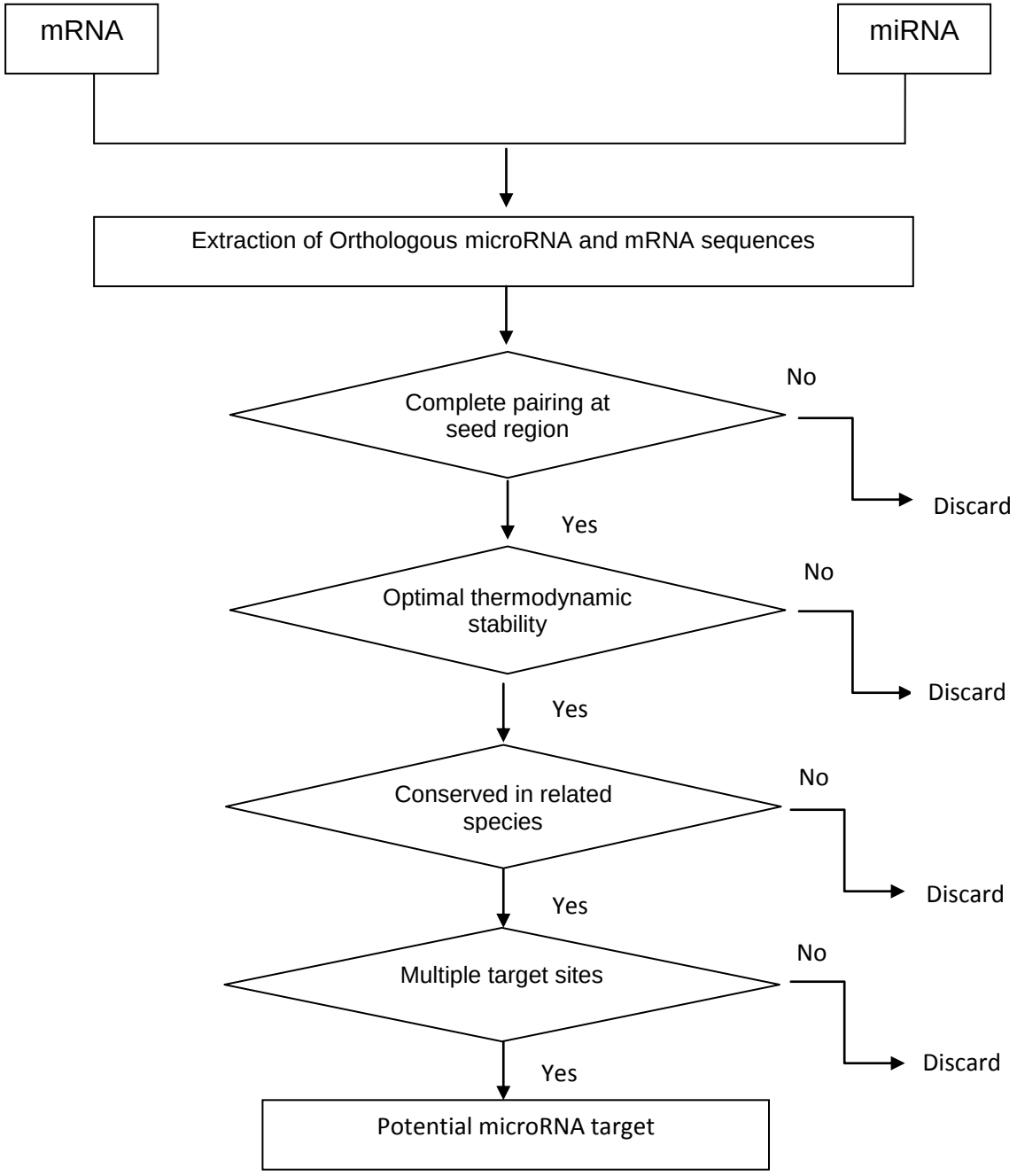

Figure 2.4: Most probable pipeline for microRNA target prediction (Watanabe et al., 2007)

2.8 Online Available MicroRNA Resources

2.8.1 miRBase

The rapid rate of microRNA gene discovery has led to two basic needs for the microRNA community. One the microRNA must have proper name and nomenclature and next one is an updated open source repository which can have all microRNA information. To fulfill this requirement microRNA registry program started. The miRBase database was created in 2002 for naming, annotation and dissemination of all published microRNA sequences (Griffiths-Jones et al., 2008).

miRBase is a primary repository for microRNA sequence and annotation. miRBase having two main functions, microRNA registry and microRNA sequence program. In microRNA registry program provides unique name to novel microRNA gene prior to the publication. The numbering of microRNA is simply sequential order and is numbered according to their publication. The name/ identifiers in database are of form “has-miR-21”. The first three or four letter denoted organism it came from. There is a difference between ‘miR’ and ‘mir’, as ‘miR’ denotes for ~22 nt long mature microRNA, whereas ‘mir’ for the ~70 nt precursor (Kozomara & Griffiths-Jones, 2011). In case of cloning studies often isolate two mature products, one from each arm of double stranded microRNA. In this case an asterisk has been used to denote less dominant form. When the data not sufficient to determine which microRNA is dominant then, miR-142-5p (for 5’arm) and miR-142-3p (for 3’ arm) used. The second aim of miRBase to provide a comprehensive and searchable database of all published microRNA sequences. The portal contains browsable list of microRNA entries, name, keyword and publication searches. The user able to search database for predicted hairpins and mature microRNA. Each database entry contains predicted stem loop containing microRNA. A brief description of genomic location, homologous sequence and possible targets by different algorithms is provided. There also having link for primary source and links to literature references for more information.

Typically, each microRNA entry in the miRBase sequence database contains:-

- An accession number.
- An ID number in the following format: [first letter of genus][first two letters of species]-mir-[sequential number]. For example, human (Homo sapiens) microRNA 143would be listed as “has-mir-21”. Orthologous microRNAs (from different organisms) are generally given the same number.
- A sequence for both the pre-microRNA and the mature microRNA.
- Sources of experimentally verification for each microRNA (with a literature reference).
- Location on chromosome, related microRNAs. and additional literature references.

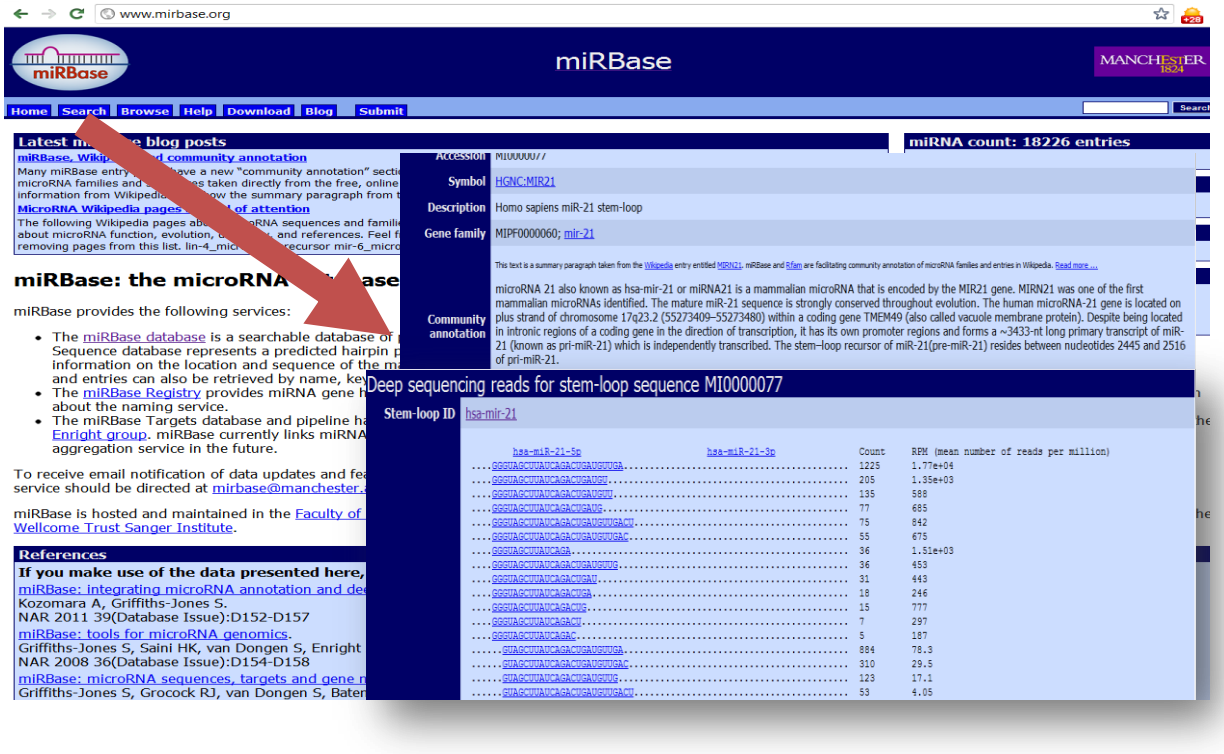

Figure 2.5: Screenshot of miRBase. Showing search result of has-miR-21 and its deep sequencing reads for stem loop sequences (miRBase, 2012)

2.8.2 Databases for Experimentally Supported Targets

miR2Disease

miR2Disease (Jiang et al., 2009) first released in 2008 and it is manually curated database, aim to provide microRNA related pathogenesis. miR2Disease curates 809 microRNA–gene interactions for Homo sapiens, coupled with related disease information derived from relevant literature. This database contains 3273 microRNA disease related entries, which is one of the strongest parts of this database.

MiRecords

MiRecords (Xiao et al., 2009) contains manually curated and predicted microRNA targets. The validated component of database contains 2286 interactions between 548 microRNAs and 1579 target gene in nine species. These all interactions are derived from gene specific interactions. The predicted targets component of miRecords is an integration of predicted microRNA targets produced by 11 established microRNA target prediction programs. The user does not have the ability to filter results based on any of the available field. The miRecords interface also enables the user to insert new microRNA:target interactions.

miRTarBase

miRTarBase (Hsu et al., 2010) was first released in 2010. It include miRTarBase currently curates 4,270 experimentally verified microRNA-Target interactions between 669 microRNAs and 2,533 target genes among 14 species (update 15 April 2011). It contains information related to the microRNA, the target gene and target site. In many cases, where the articles do not explicitly present target site information, miRTarBase can provide predicted regions by using a computational target prediction algorithm.

miRWalk

miRWalk (Dweep et al., 2011) provides experimentally supported microRNA targets identified solely from text-mined abstracts available in MedLine. The miRWalk validated targets module hosts text-mined interactions for 1572 microRNAs interacting with 5080 genes for three species (human, mouse and rat).

starBase

starBase (sRNA target Base) (Yang et al., 2010) started to facilitate the comprehensive exploration of microRNA–target interaction maps from CLIP-Seq and Degradome-Seq data. starBase currently includes tens of thousands of microRNA-target interactions among 6 species. starBase also integrates thousands of binding sites of other RNA-binding proteins (RBPs) to decode combinatorial effects of Ago and other RBPs. The developers utilized five prediction programs to locate putative targets, which were subsequently intersected with the previously analyzed high-throughput data. starBase provides the in-house developed DeepView Genome Browser, which enables access to mapped reads, predicted and known microRNA targets, ncRNAs, and protein coding genes, target clusters, target peaks and target plots.

TarBase 6.0

Tarbase (Vergoulis et al., 2012) one of the oldest database containing experimentally supported microRNA targets. The TarBase 6.0 includes 65814 experimentally validated microRNA gene interactions, which are extracted from relavant literature by TarBase. This is a 50 fold increase of entries from the latest TarBase version. Tarbase hosts data derived from 3 CLIP-Seq and Degradome_Seq studies. The database was populated with entries derived from manual manuscript curation. The curators noted the microRNA, the related target gene as well as information regarding the experiment such as the utilized cell line or tissue. The utilized methodology (gene specific or high throughput) was specifically mentioned: Reporter genes, qPCR, Western blotting, MicroArray, Proteomics (such as pSILAC),

Sequencing (RNA-Seq, HITS-CLIP, PAR-CLIP), Degradome-Seq Other (e.g. ELISA, RACE, immunohistochemistry etc.).

2.8.3 Target Prediction Algorithms

There are many microRNA target prediction software and databases available at present. These in-silco algorithms predict most probable microRNA targets based on probability scores. As these tools use different mathematical models and also different microRNA database for prediction so it produces drastic difference in result. These all microRNA target prediction algorithms can be classified in to three categories Sequence based or Rule based algorithm, Thermodynamic approach based algorithm and Data driven or Machine Learning prediction approach.

We can also classify microRNA target prediction tools into two generations. The first generation methods like miRanda (John et al., 2004), DIANA microT (Maragkakis et al., 2009), RNAHybrid (Rehmsmeier et al., 2004), MicroInspector (Rusinov et al., 2005) and TargetScan (Lewis et al., 2005) are mainly base on three characteristics: 1) The 5’ seed of the microRNA (nucleotide positions 2-8 of the microRNA) is is complementary to the 3’ UTR of the target mRNA. 2) The microRNA:mRNA duplex has a highest negative folding energy. 3) The binding site of microRNA:mRNA are highly conserved from species to species and mostly conserved in same kingdom. The second generation prediction tools uses machine learning base approach like support vector machine (SVM), neural network, Hidden Markov model (HMM), and Naïve Bayes (NB). PicTar (Lall et al., 2006) is the best example of second generation algorithm; this tool scan alignment of 3’ UTRs for those matches the microRNA and calculates the thermodynamic stability this binding. Each predicted targets scored by using a HMM maximum likelihood fit approach.

Rule Based Algorithms

Rule based algorithms generally consist of a set of rules to be satisfied by a testing 3' UTR. The majority of the sequence based target prediction tools come under rule based programs, these uses sequence complementarity as main

prediction criteria, miRanda and TargetScan uses sequence complementarity for microRNA target prediction

TargetScan 6.0

TargetScan (Lewis et al., 2005) represents one of the earliest attempts to employ the seed concept in determining microRNA targets. TargetScan (Friedman et al., 2009) seeks instances of seed's exact reverse compliment (i.e. no bulges or wobble are permitted in the seed region) in the 3' UTRs of putative target genes. TargetScan predicts microRNA targets based on the identification of aligned seed matches and their conservation in several species. The overall scoring of a microRNA target site depends on the level of conservation, whether it binds to the microRNA on position 8 and/or whether it has an A at position 1, the distance of the target from the 3'UTR end and the AU composition of the flanking area. The algorithm then use RNA Fold package for the calculation of the binding sites and multiple binding sites. Then Z score of the 3' UTR sites is calculated by using following equation.

$$\mathrm{Z} = \sum_{\mathrm{j}=1}^{n} e^{-\frac{G_j}{T}} \qquad (1)$$

In equation 1 n is the number of seed match in the 3' untranslated region, G_j is the calculated free energy (units: kcal/mol) of the interaction between the microRNA its target for the j^{th} target evaluated in the previous step. T is a parameter that influences the relative weighting of UTRs as a function of the affinity and abundance of their target sites.

ElMMo

ElMMo (Hausser et al., 2009) uses a general Bayesian method that scores the conservation of microRNA binding sites according to an evolutionary model that utilizes the assumed phylogenetic relationship among several species.

TargetRank

TargetRank (Nielsen et al., 2007) scores the seed matches in a UTR relative to given microRNA, and then calculates an overall score for the mRNA as whole by summing the score for all seed matches present in the 3' UTR. The score for each seed match is based on (a) its seed match type, (b) the base composition at position T9, (c) flanking AU content (of the 50 nt. immediately 3'of the seed match) and (d) flanking conservation (50 nt. immediately 5' of the seed match).

MiRanda

MiRanda (John et al., 2004) is one of the earliest developed large scale prediction algorithm for vertebrate. This selects target gene on the basis of three basic properties: sequence complementarity using position-weighted local alignment algorithm, free energy of microRNA:mRNA duplex using Vienna RNAfold package and conservation of targets in related genomes. This algorithm search co-occurrence of predicted target sites for multiple microRNAs on a particular mRNA and also microRNA expression profile in various mammalian tissues. When basic parameter settings are used, the approximated false positive rate was 24% and 39%. These values decrease significantly when multiple sites are considered.

Table-2.4: **Scoring Matrix used by miRanda**

	C	G	A	T	U	X
C	-3	+5	-3	-3	-3	-1
G	+5	-3	-3	+1	+1	-1
A	-3	-3	-3	+5	+5	-1
T	-3	+1	+5	-3	-3	-1
U	-3	+1	+5	-3	-3	-1
X	-1	-1	-1	-1	-1	-1

MicroCosm

microCosm Targets (Griffiths-Jones et al., 2008) uses miRanda algorithm to identify potential binding sites for a given microRNA. The current version of programme uses dynamic programming alignment to identify highly complementary sites. Strict complementority at 5' seed region is demanded. Thermodynamic stability is estimated for each target site by the use of Vienna RNAfold package. Every potential target site in a 3'UTR detected is checked to see whether the site is conserved in orthologous transcripts from other species. The conserved sites must be detected at the same position in a cross-species orthologous UTR alignment by an microRNA of the same family. Each target must be conserved in at least two species for inclusion in the database.

Thermodynamics Approach

In energy base algorithms free energy of microRNA-mRNA binding is calculated. In this category DIANA MicroT and RNAHybrid are main examples.

(DNA Intelligent Analysis) DIANA-microT 4.0

The DIANA-microT 4.0 (Maragkakis et al., 2009) algorithm is based on parameters calculated individually for each microRNA and each (microRNA Recognition Elements score (MRE score) depending on binding and conservation features. The prediction score of a miTG interaction is weighted sum of the scores of conserved and non-conserved MREs on a gene. Signal to noise ratio (SNR) and

precision score are calculated for each interaction to provide an estimate of the false positive rate of each predicted miTG scores.

The program identifies the highest scoring alignment between every nine nucleotide long window of the 3' UTR with the microRNA driver sequence using dynamic programming algorithm. Alignment is based on following rules:

1. Minimum six nucleotide match (Watson-Crick) or G:U is required . If the six matches are W-C and the binding is starting from position 1 or 2 of the microRNA deiver sequence, then the MRE considered as a 6 mer.
2. For the site less than 7 consecutive W-C matches an additional energy filter is applied.

The first version of DIANA-microT Web server was designed to support the functional analysis of human and mouse microRNAs. New version of DIANA-microT server upgraded to support two more species for prediction.

RNAHybrid

RNAHybrid (Rehsmeier et al., 2004) is a tool for finding minimum free energy hybridization of long (target) and short (query) RNA. This is an extension of classical RNA secondary structure prediction software tools. The microRNA is hybridized to the target in an energetically optimal way i.e. minimum free energy, forbidding intra molecular base pairings and branching structures (multiloops). Dynamic programming technique helps in calculation of MFE of hybridizations of all possible start positions in the microRNA and in the target. The microRNA can predict multiple targets in single site. RNAHybrid finds the energetically most favorable hybridization sites of a small RNA within a large target RNA sequence, and base pairings between target nucleotides or between microRNA nucleotides are not allowed.

Machine Learning Based Approach

Machine learning based approached is used by PicTar, NBmiRTar and miTarget, this approach uses machine learning technique for the pattern finding shared by microRNA-target interaction.

PicTar

PicTar (Lall et al., 2006) identifies two types of microRNA:target interactions: (i) those with perfect complementarity between the seed region of the microRNA (7 nt starting at position 1 or 2 of the microRNA's 5' end) and the 3' UTR target site and (ii) those for which the perfect complementarity is interrupted by at most one nucleotide bulge, mismatch, or G:U wobble. In both instances, the algorithm requires that the binding stability of the putative microRNA:target interaction, as measured by thermodynamic binding energy, exceeds a specified threshold. Once individual microRNA:target interactions are identified, the algorithm labels highly conserved (among 4 or 5 species) target sites as 'anchors' and filters out those 3' UTRs that do not harbor a specified number of anchors. A hidden Markov model is then used to score the likelihood of a 3' UTR being targeted by microRNAs in a combinatorial manner and several different microRNAs can act together to repress the same gene. These scores are computed for a set of species and combined to compute the final score.

Table-2.5: MicroRNA Target Prediction Software

Name	Input Data	Method	Date of Release	Reference
TargetScan	Sequence, conservation, energy	Scoring and Ranking	2003	(Lewis et al., 2005).
RNA Hybrid	Sequence, energy	Dynamic programming and statistical analysis	2004	(Rehmsmeier et al., 2004).
miRanda	Sequence, conservation, energy	Dynamic programming and statistical filtering	2004	(John et al., 2004).
PicTar	Sequence, conservation, energy	Hidden Markov Model	2005	(Lall et al., 2006)
TargetBoost	Sequence	Boosted genetic algorithm	2005	(Saetrom et al., 2005)
MovingTarget	Sequence, energy	Scoring and Ranking	2005	(Burgler & Macdonald, 2005)
TargetScanS	Sequence, conservation	Scoring and Ranking	2005	(Lewis et al., 2005)
MicroInspector	Sequence, energy	Scoring and Ranking	2005	(Rusinov et al., 2005)
RNA22	Sequence, energy	Markov chain	2006	(Miranda et al., 2006)
miTarget	Sequence, conservation, structure	Support Vector machine	2006	(Kim et al., 2006)
MicroTar	Sequence, energy	Scoring and Ranking	2006	(Thadani & Tammi, 2006)
EiMMO	Sequence, conservation	Bayesian model	2007	(Hausser et al., 2009)
Pita	Sequence, conservation, energy, structure	Scoring and Ranking	2007	(Kertesz et al., 2007)
TargetRank	Sequence, conservation	Scoring and Ranking	2007	(Nielsen et al., 2007)
NBmirTar	Sequence, conservation, energy	naive Bayesian model	2007	(Yousef et al., 2007)
GenMiR++	Other target predictions(TargetScans), microRNA expression, mRNA expression	Variational Bayesian learning	2007	(Huang et al., 2007)
MicroCosm		Support Vector Machine	2008	(Griffiths-Jones et

				al., 2008)
mirWIP	Sequence, conservation, energy, structure	Scoring and Ranking	2008	(Hammell et al., 2008)
miRTif	Sequence	Support Vector machine	2008	(Yang et al., 2008)
FindTar	Sequence, structure	Scoring and Ranking	2008	(Ye et al., 2008)
HuMiTar	Sequence	Scoring and Ranking	2008	(Ruan et al., 2008)
miRGator	Other target predictions(miRanda, PicTar, TargetScans), microRNA expression, mRNA expression	Gene set enrichment analysis, scoring and ranking	2008	(Nam et al., 2008)
MirTarget2	Sequence, conservation, energy, microRNA expression, mRNA expression	Support Vector machine	2008	(Wang & Naqa, 2008)
SigTerms	Other target predictions(miRanda, PicTar, TargetScans), microRNA expression, mRNA expression	Gene set enrichment analysis, scoring and ranking	2008	(Creighton et al., 2008)
DIANA micro-T	Sequence, conservation	Scoring and Ranking	2008	(Maragkakis et al., 2009)
TargetMiner	Sequence	Support Vector machine	2009	(Bandyopadhyay & Mitra, 2009)
FastH	Sequence, energy	Scoring and Ranking	2009	(Ragan et al., 2009)
SVN+SC	Sequence, microRNA expression	Sequence alignment, Support Vector machine	2009	(Joung & Fei, 2009)
HocTar	Other target predictions(miRanda, PicTar, TargetScans), microRNA expression, mRNA expression	Scoring and Ranking	2009	(Gennarino et al., 2009)
Mtar	Sequence, energy	Artificial neural network	2010	(Chandra et al., 2010)
PACMIT	Sequence, conservation, energy, structure	Markov Model	2010	(Marín & Vaníček, 2011)
SVMicro	Sequence,	Support Vector	2010	(Liu, et al.,

	conservation, energy, structure	machine		2010)
TargetSpy	Sequence, energy, structure	Support Vector machine	2010	(Sturm et al., 2010)
ExprTarget	Other target predictions(miRanda, PicTar, TargetScans), microRNA expression, mRNA expression	multivariate logestic regression	2010	(Gamazon et al., 2010)
ExpMicro	Other target predictions(SVMicro), microRNA expression, mRNA expression	Gaussian mixture model	2010	(Liu, et al., 2010)

2.9 Comparison of microRNA Target Prediction Algorithms

Comparative analysis of microRNA target prediction algorithms is difficult. The main problem is very less number of validated targets as compared to predicted targets. So, the initial trials to assess the target prediction tools are not so effective. Sethupathy et al. compared five microRNA target prediction tools (TargetScan, DIANA microT, PicTar,miRanda and TargetScanS). The programs have sensitivity of TargetScan 20.8%, DIANA microT 9.5%, miRanda 48.8%, TargetScanS 47.6% and PicTar 47.6% respectively (Sethupathy et al., 2006). Alexiou et al. compared experimental with the five prediction tools (TargetScan 5.0, DIANA microT 3.0, PicTar, EiMMo and TargetScanS). These five programs have a precision of ~50% and sensitivity of about 2-12%. All these programs rely heavily on evolutionary conservation of seed region (Alexiou et al., 2009). Zhang and Verbeek (Zhang & Verbeek, 2010) tried to compare three prediction programs (miRanda, TargetScanS and RNAhybrid). They compared these three with the help of positive as well as negative targets. In this they obtained negative sets in a different way as compared to previous studies. Current research focuses on finding the true targets, and consequently, only a small number of false target are identified as by products. The most common way to generate negative data is sequence shuffling.

Chapter Three

Objectives and Hypothesis

On the basis of above literature review we can say that assessment of microRNA target prediction tools is very necessary. MicroRNA analysis starts from its functional analysis as it is main cellular regulator. MicroRNA level fluctuates in different cellular condition, especially in cancerous cells. Before going for any experimental microRNA functional analysis in silico prediction is very necessary. It is very necessary to use proper microRNA target prediction program according to requirement of research. The main objectives of this work are:

1. Prediction of microRNA targets in metastatic breast cancer.
2. Assessment of microRNA prediction programs with help of experimentally validated microRNA targets of metastatic breast cancer.

Hypothesis

The comparative analysis of different software with the help of experimentally validated microRNA targets of metastatic breast cancer will provide the best prediction program out of all the software analyzed. This will highlight the drawback of different algorithm and it will help in advancement of new algorithms. This study will provide prediction program which will have highest sensitivity and precision. Improvement of these prediction programs is very necessary as most of the algorithm have very low precision and sensitivity. We can directly compare different programs by the use of experimentally validated microRNA targets as benchmark level. The comparison of different prediction tools provides which prediction approach is the best; rule based approach or data driven approach.

Chapter Four

Material and Methods

Computational microRNA target prediction algorithms provide putative binding site for microRNA. MicroRNA can bind more than one target gene and in multiple sites. In case of plant microRNA bind to their target mRNAs with perfect or near perfect complimentarity (Zhang, 2005). While in animals the microRNA-mRNA binding having limited complimentarity, making target prediction in animal a challenging task.

4.1 Material

In this study Eighteen (18) microRNAs were taken for study. All these microRNA have direct or indirect role in breast cancer metastasis and progression.

4.1.1 Experimentally Validated microRNA Target

Experimentally validated microRNA targets where collected from TarBase 6.0. The TarBase 6.0 hosts a significant amount of information for each microRNA gene interaction ranging from microRNA and gene-related facts, information specific to their interaction, the experimental validation methodologies and their subsequent outcomes. TarBase 6.0 is the main repository for experimentally validated targets. This database contains only experimentally validated targets of microRNA. The microRNA targets are shown according to their gene names. The microRNA responsible for breast cancer metastasis and its progression are taken in to consideration. All 18 microRNA's validated targets were searched in TarBase, except one (miR-211) all validated targets of microRNA are listed in database.

4.1.2 Predicted microRNA Targets

For the purpose of microRNA target prediction seven most popular software have been chosen. All these software are in regular use and are most successful. All

these prediction software show their prediction results in descending order according to the gene number and probability score obtained by them.

a) TargetScan

This is one of the old microRNA target prediction software developed by (Lewis et al., 2005) and available at http://www.targetscan.org/vert_61/. TargetScan program have four variants namely; TargetScan Human, Mouse, Worm and Fly. In total it can predicts more than 23 vertebrate species, fly and worms. It takes input as Entrez gene symbol or as microRNA name/ symbol. There are three types of dropdown input box, one contains broadly conserved microRNA families the next one have conserved microRNA and the last one have poorly conserved microRNA families. In newer version one can enter microRNA symbol/ name and submit the input for target prediction. The result shows total number of conserved target, conserved sites and poorly conserved sites. The predicted microRNA targets are arranged according to probability scores.

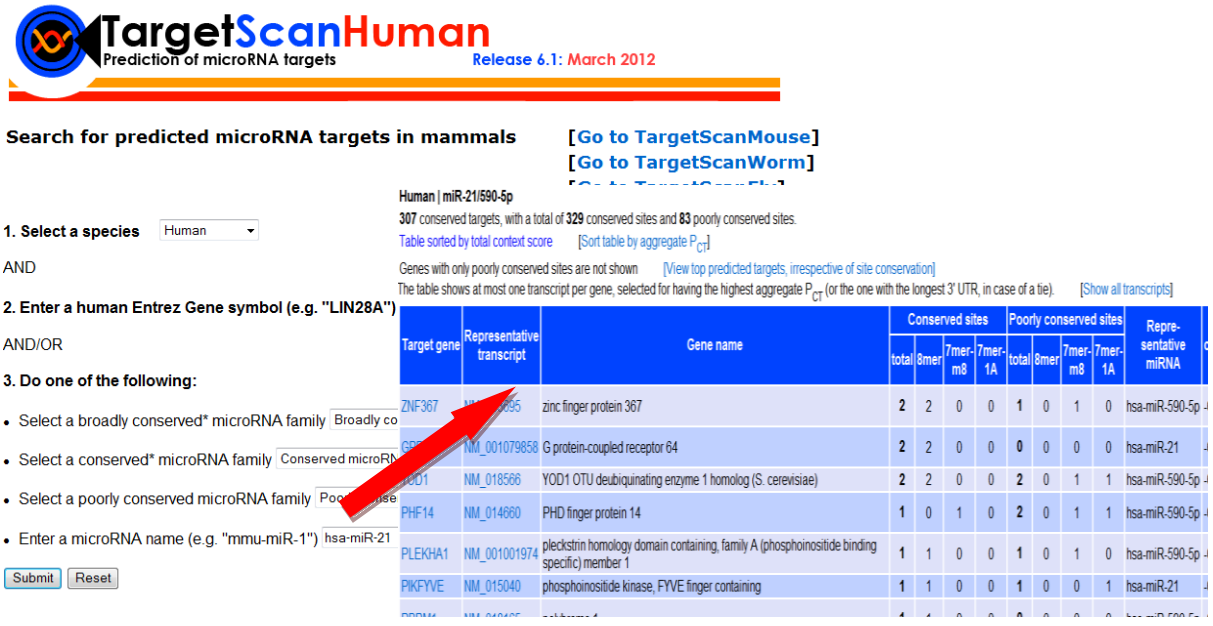

Figure 3.1: Screen shot of TargetScanHuman showing home page and result page. The arrow indicates result page of miR-21 prediction (TargetScan, 2012).

b) PicTar

PicTar is project of the Rajewsky lab ay NYU's Centre for Comparative Genomics. The program is available at http://PicTar.mdc-berlin.de/. There are three prediction programs available one developed by Krek et al, 2005 for vertebrates and flies Grün Et al., 2005 (Grün et al., 2005; Krek et al., 2005). Other one is Lall et al., 2006 for vertebrate, flies, nematodes (Lall et al., 2006). The last one is unpublished PicTar for mouse. PicTar vertebrate predicts only for vertebrate and fly. The dataset for microRNA target prediction is based on conservation in mammalian and chicken (human, mouse, chimps, rat, dog). The results are arranged according to the rank and it depends upon PicTar score obtained by individual gene.

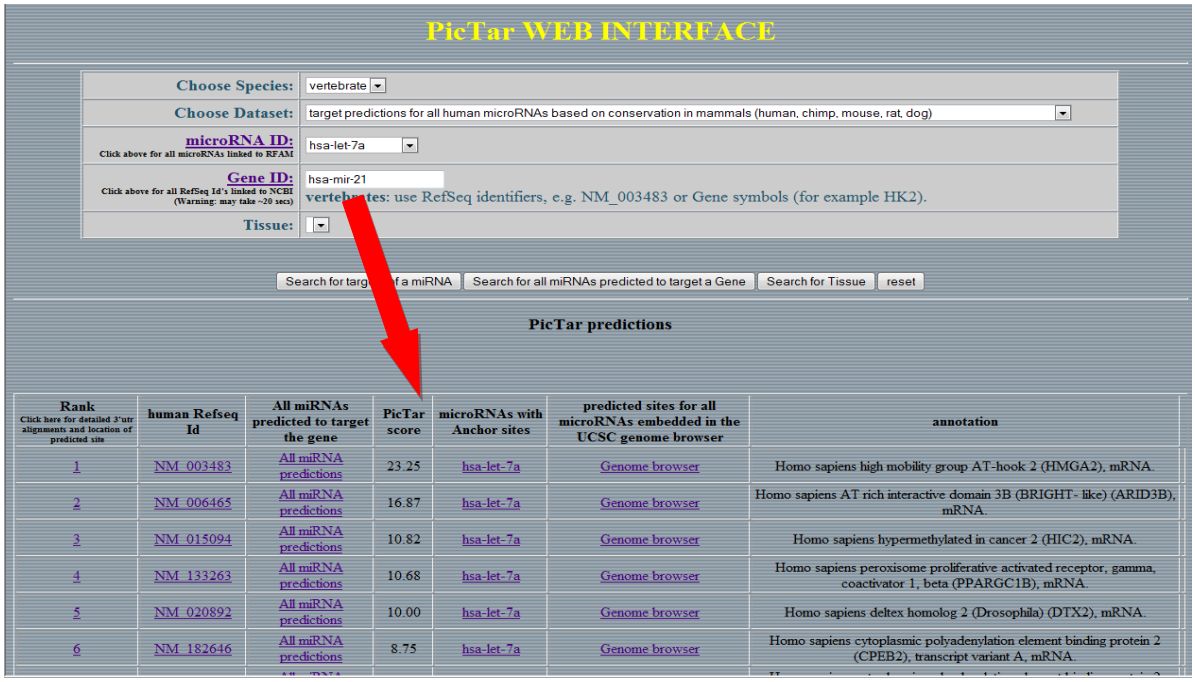

Figure 3.2: Screen shot of PicTar showing prediction page (PicTar, 2012).

c) DIANA microT v.4

DIANA microT is available at (DNA Intelligent Analysis) DIANA LAB portal http://diana.cslab.ece.ntua.gr/DianaTools/index.php?r=microtv4/index. The Prediction tool is able to predict human, mouse, fly and worm microRNA targets. Input can be given as microRNA name or as gene name. The new interface predicts the gene target on the basis of miTG score. DIANA microT have option to analyze new microRNA which is not listed in database also.

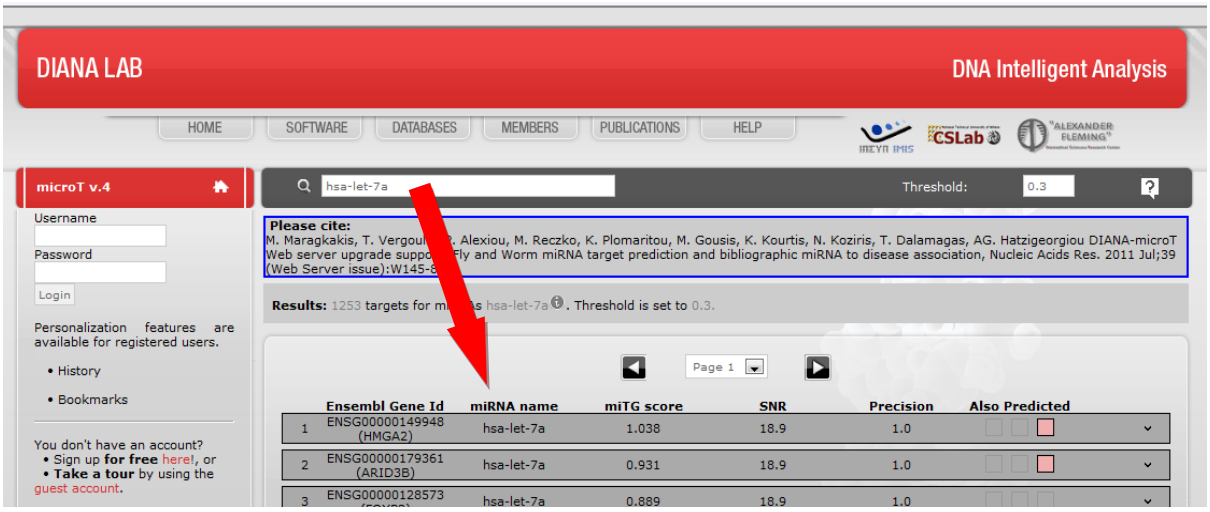

Figure 3.3: Screen shot of DIANA microT v.4 showing results output page of has-miR-21 (DIANA LAB, 2012)

d) EiMMo

EiMMo prediction program is maintained by Swiss Institute of Bioinformatics, Basel. EiMMo can predict microRNA target for six different species. There are total 692 human microRNAs listed in this database and required microRNA can be selected in checkbox. Search can be done by selecting more than one microRNA at a time. Only first 500 predictions are shown in the result.

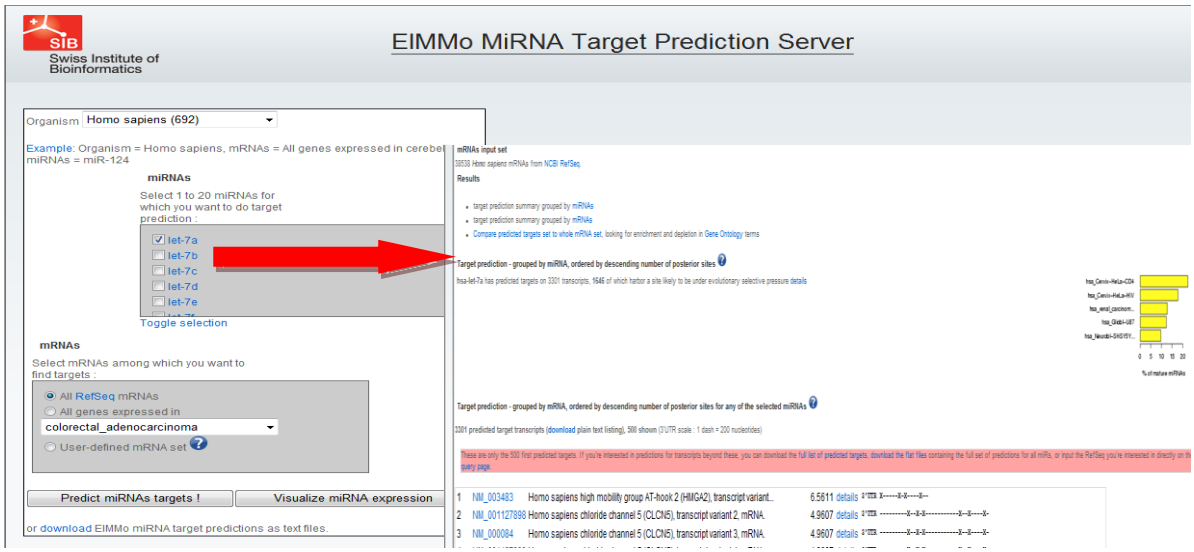

Figure 3.4: Screen shot of EiMMo program, showing results output page (EiMMo,2012)

e) TargetRank

It can predict only human and mouse microRNA targets. This program not only predicts microRNA listed in database but also new microRNA. MicroRNA sequence can be pasted in the box for target prediction. Search result show only first 100 prediction listed on the basis of Target Rank score.

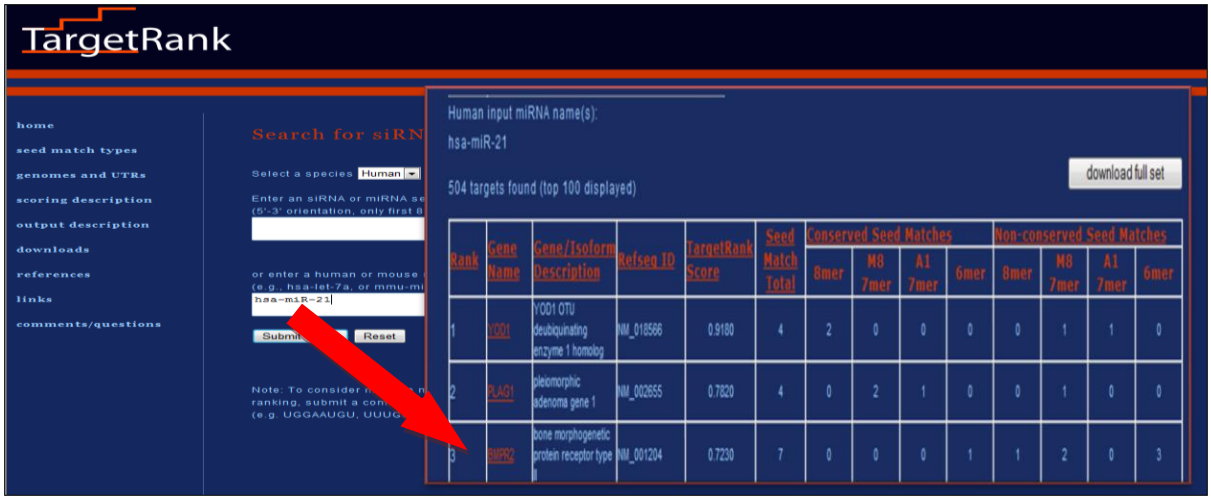

Figure 3.5: Screen shot of TargetRank showing web interface and results output page (TargetRank, 2012)

f) miRanda

miRanda can predict microRNA target for human, rat, mouse, c. elegans and Drosophila. Till now total 1100 microRNA listed in this database. Search can be done by entering required microRNA in the search box. The search show result on the basis of mirSVR score. Targets are linked with alignment details.

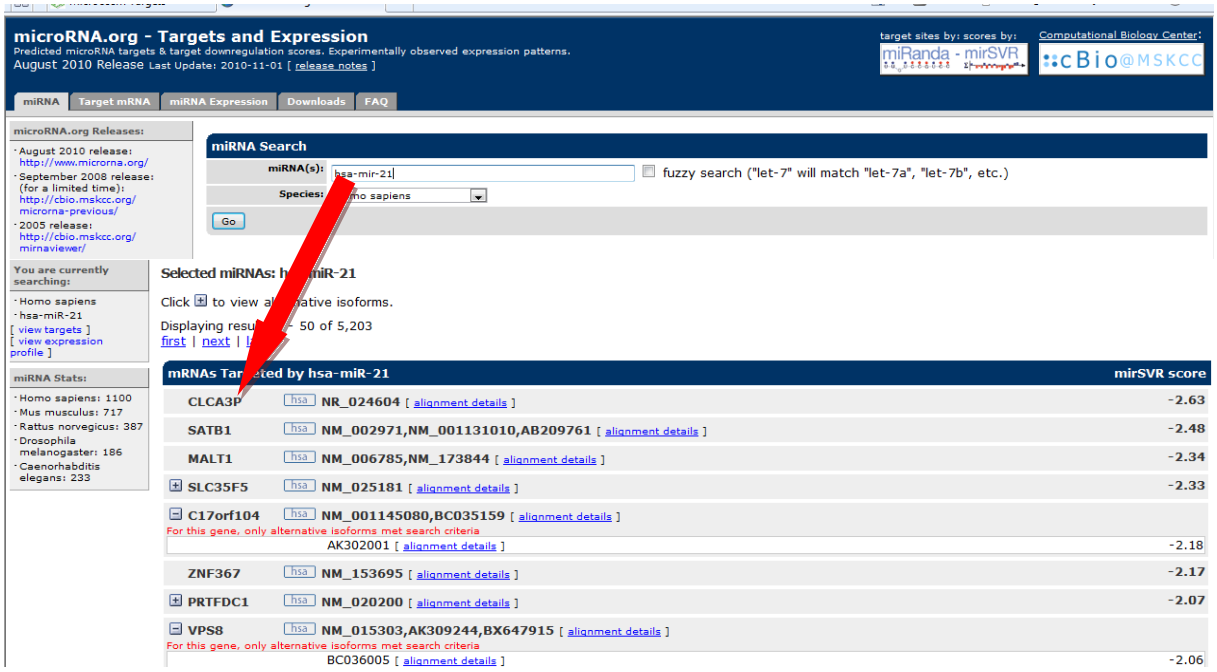

Figure 3.6: Screen shot of miRanda showing results output page (microRNA.org, 2012)

g) MicroCosm

This software uses miRanda algorithm for target prediction in genomic sequence. It can predict microRNA target from 22 species. In all these species number of microRNA, number of transcripts and number of targets also listed. MicroRNA targets can be viewed by selecting desired microRNA from different species. Result contains microRNA target on the basis of score obtained. Each target is linked with its transcriptome and GO term.

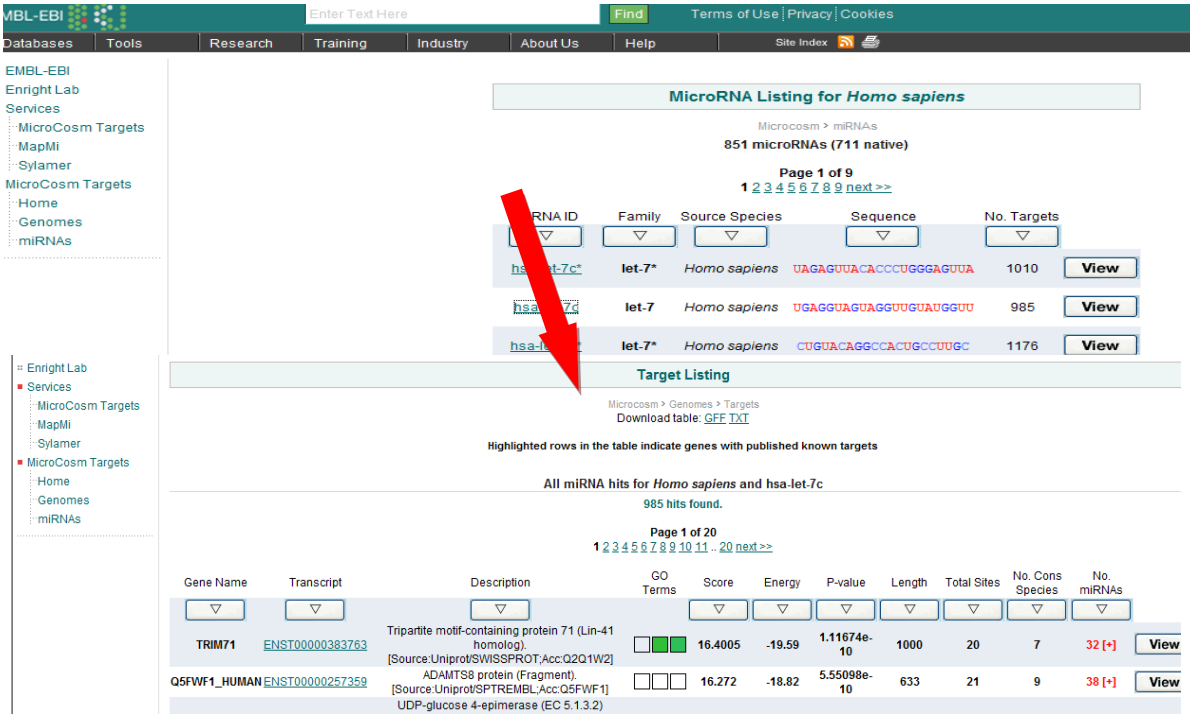

Figure 3.7: Screen shot of microCosm webpage, showing list miRNA and results output page (EBI,2012)

4.2 Methodology

Assessment of microRNA target prediction programs can be done with the validated targets. This methodology is used by the (Sethupathy et al., 2006) and (Alexiou et al., 2009). Different microRNA target prediction tolls uses different microRNA and mRNA data set for their prediction purpose. This is one of the important cause of dissimilarity of result. In this study, we have modified methodology for more accurate assessment.

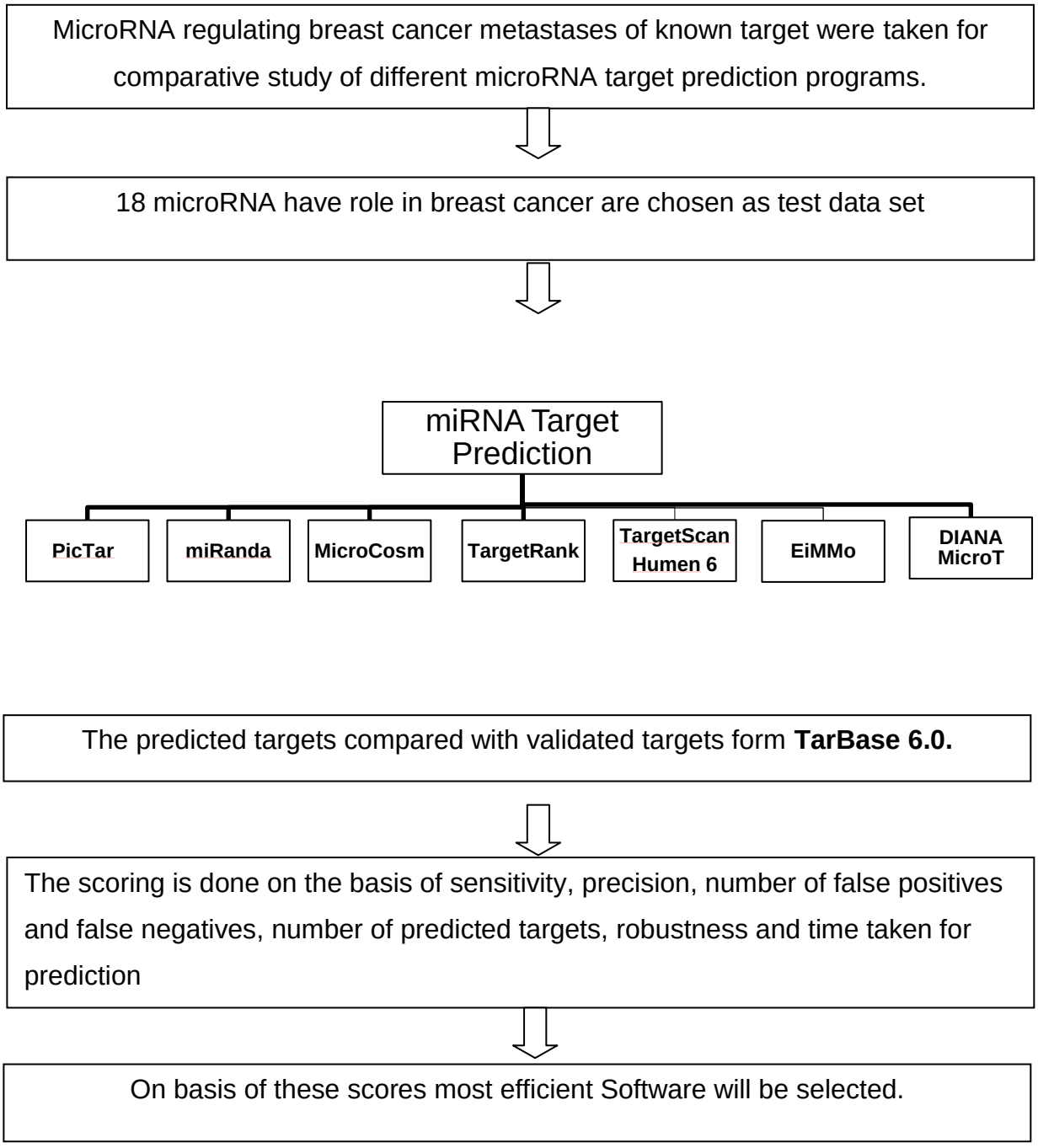

Figure 3.8: Flowchart showing methodology of assessment of different tools.

4.2.1 Comparative Study of microRNA Prediction Tools/Algorithms:-

There are many differences in result predicted by the prediction tools. So the accuracy of microRNA-mRNA interactions can be evaluated with the help of sensitivity and precision (Alexiou et al., 2009).

A. Sensitivity:

The performance of computational programs is measured by sensitivity. Sometimes sensitivity also called as true positive rate (TPR). True positive (TP) is the number of predicted microRNA-mRNA interactions that really exist. False negative (FN) is the number of microRNA-target interaction that do exist, but where not predicted.

$$\text{Sensitivity} = \frac{\text{True positive (TP)}}{\text{Total correct}} \quad (2)$$

Here in equation 2 total correct means experimentally validated targets of a particular microRNA listed in TarBase database. High sensitivity means more correct targets predicted by the algorithms.

B. Precision:

Precision gives us the number of predicted targets out of total target exists. This gives predictability of software.

$$\text{Precision} = \frac{\text{True positive (TP)}}{\text{Total predicted}} \quad (3)$$

In equation 3 total predicted means the target predicted by individual target prediction software. More precision means more chance of get validated. There is an

inclination of researcher to predict one miRNA by more than one target prediction algorithm as leading to higher prediction precision.

Chapter Five

Results and Discussion

Investigation of Microrna Target Prediction Algorithm and its Feature

There are many target prediction software today, but most importantly few programs are very important and used regularly by different researchers. The main problem of direct comparison of prediction performance among these tools is that predicted target genes from different tools do not overlap well. The comparative analyses of these programs are very necessary, as it will focus light on drawback of different programs.

5.1 Comparison between MicroRNA Target Prediction Tools

Table 5.1: Target Prediction Tools and Their Features

Prediction Tool	Species	Seed Match	Site location	Free Energy	Conservation	Clustering
Target Scan	23 vertebrate, fly and worm	Yes	Considered	Yes	With/Without Conservation filter	No, Multiple Site Considered
PicTar	Mouse, fly, worm,	Yes	No	Yes	With Conservation filter	No
DIANA microT	Human, Mouse	Yes	No	Yes	With Conservation filter	No
EiMMo	human, mouse, fly, zebrafish	Yes	No	No	With Conservation filter	No
Target Rank	Human and mouse	Yes	No	No	With/Without Conservation filter	Yes
MicroCosm	44 species	Yes	No	Yes	With Conservation filter	No
MiRanda	Human, mouse and rat.	Yes	No	Yes	With Conservation filter	No

Table 5.2: Total Number of MicroRNA in Different Target Prediction Tools

Software	TargetScan	PicTar	EiMMo	DIANA microT	TargetRank	MiRanda	MicroCosm
microRNA	1733	178	692	875	723	1100	851

MicroRNA target prediction tools mainly based on different target prediction features. These features are based on experimental data available. Most of the programs analyzed are able to predict more than one species, which shows the robustness of the prediction program. Initial algorithms are based on the strict complementarity in seed region. These algorithms does not consider free energy of microRNA:mRNA binding. Multiple binding sites are one of the important features which are used by different algorithms. Numbers of microRNA are increasing in very high pace; it is very necessary to update prediction programs. In table-VII microRNA available in different target prediction tools are listed. TargetScan and MiRanda having maximum microRNA listed in its database.

Table 5.3: Experimentally Validated MicroRNA Target Gene

Experimentally validated targets are available in different databases, but TarBase 6.0 is most updated and contains maximum microRNA.

microRNA	**Number of Validated Targets in TarBase 6.0**
hsa-let-7a	73
hsa-let-7b	481
hsa-let-7c	21
hsa-let-7d	100
hsa-miR-9	202
hsa-miR-10b	10
hsa-miR-21	581
hsa-miR-101	177
hsa-miR-146a	118

hsa-miR-146b	3
hsa-miR-200a	47
hsa-miR-200b	44
hsa-miR-203	13
hsa-miR-206	18
hsa-miR-211	n/a
hsa-miR-221	124
hsa-miR-335	2923
hsa-miR-373	165
Total	**5100**

C. MicroRNA Target Prediction Tools and Their Precision and Sensitivity:

Precision and sensitivity calculated for each microRNA and for each prediction tool and then average precision and sensitivity calculated of a particular program. These two statistical parameters can be used to characterize the performance of the prediction tool. In these tables, total correct column represents validated targets listed in Tarbase database. Only microRNA-221 is not listed in tarbase 6.0. Predicted targets of each microRNAs by different programs are listed in total predicted column. MicroRNA targets which is predicted by prediction tool and also present in the list of validated targets are listed in column correctly predicted.

Table 5.4: Precision and Sensitivity According To Target Prediction Tools of All miRNA

A. Target Scan:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	87	1	1.15	1.37
2	**microRNA-7b**	**481**	31	3	9.68	0.62
3	**microRNA-7c**	**21**	13	0	00	0
4	**microRNA-7d**	**100**	125	3	2.4	3
5	**microRNA-9**	**202**	0	0	0	0
6	**microRNA-10b**	**10**	51	3	5.88	30
7	**microRNA-21**	**581**	162	50	30.86	8.61
8	**microRNA-101**	**177**	500	60	12	33.89
9	**microRNA-146a**	**118**	203	3	1.47	2.54
10	**microRNA-146b**	**3**	22	1	4.54	33.33
11	**microRNA-200a**	**47**	0	0	0	0
12	**microRNA-200b**	**44**	790	7	.88	15.90
13	**microRNA-203**	**13**	878	7	.79	53.84
14	**microRNA-206**	**18**	0	0	0	0
15	**microRNA-211**		237	0	0	0
16	**microRNA-221**	**124**	255	17	7.17	13.70
17	**microRNA-335**	**2923**	0	27	10.58	.92
18	**microRNA-373**	**165**	0	0	0	0
Average					**4.85**	**10.98**

B. PicTar:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	724	19	2.62	26.0
2	**microRNA-7b**	**481**	527	53	10.05	11.01
3	**microRNA-7c**	**21**	500	8	1.6	38.1
4	**microRNA-7d**	**100**	500	15	3	3
5	**microRNA-9**	**202**	496	34	6.85	16.83
6	**microRNA-10b**	**10**	179	9	5.02	90
7	**microRNA-21**	**581**	189	74	39.15	12.73
8	**microRNA-101**	**177**	498	41	8.23	23.16
9	**microRNA-146a**	**118**	154	3	1.9	2.54
10	**microRNA-146b**	**3**	0	0	0	0
11	**microRNA-200a**	**47**	428	14	3.27	29.78
12	**microRNA-200b**	**44**	0	0	0	0
13	**microRNA-203**	**13**	439	4	.91	30.76
14	**microRNA-206**	**18**	502	7	1.39	.3809
15	**microRNA-211**		0		0	0
16	**microRNA-221**	**124**	312	22	7.05	17.74
17	**microRNA-335**	**2923**	182	16	8.79	.54
18	**microRNA-373**	**165**	345	11	3.18	6.66
Average					**5.728**	**19.32**

C.EiMMo:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	492	20	4.06	27.39
2	**microRNA-7b**	**481**	488	67	13.72	13.92
3	**microRNA-7c**	**21**	488	7	1.43	33.33
4	**microRNA-7d**	**100**	493	30	6.08	30
5	**microRNA-9**	**202**	495	28	5.65	13.86
6	**microRNA-10b**	**10**	499	7	1.40	70
7	**microRNA-21**	**581**	494	179	36.23	30.80
8	**microRNA-101**	**177**	498	60	12.04	33.89
9	**microRNA-146a**	**118**	498	5	1	4.23
10	**microRNA-146b**	**3**	0	0	0	0
11	**microRNA-200a**	**47**	498	8	1.60	17.02
12	**microRNA-200b**	**44**	499	16	3.20	36.36
13	**microRNA-203**	**13**	0	0	0	0
14	**microRNA-206**	**18**	497	9	1.81	50
15	**microRNA-211**		0	0	0	0
16	**microRNA-221**	**124**	495	37	7.47	29.83
17	**microRNA-335**	**2923**	494	37	7.48	1.2
18	**microRNA-373**	**165**	499	12	2.40	7.27
Average					**5.86**	**22.17**

D. DIANA microT:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	480	13	2.708	17.80
2	**microRNA-7b**	**481**	500	59	11.8	12.26
3	**microRNA-7c**	**21**	501	5	.998	23.80
4	**microRNA-7d**	**100**	500	38	7.6	38
5	**microRNA-9**	**202**	499	37	7.405	18.32
6	**microRNA-10b**	**10**	498	7	30.46	70
7	**microRNA-21**	**581**	499	152	12	26.16
8	**microRNA-101**	**177**	500	60	1	33.89
9	**microRNA-146a**	**118**	500	5	0	4.23
10	**microRNA-146b**	**3**	499	0	1.4	0
11	**microRNA-200a**	**47**	500	7	2.2	14.89
12	**microRNA-200b**	**44**	500	11	2.2	25
13	**microRNA-203**	**13**	500	3	.6	23.07
14	**microRNA-206**	**18**	500	4	.8	22.22
15	**microRNA-211**		499	0	0	0
16	**microRNA-221**	**124**	499	31	6.3265	25
17	**microRNA-335**	**2923**	330	31	9.39	1.06
18	**microRNA-373**	**165**	501	15	2.94	9.09
Average					**5.50**	**20.26**

E. Target Rank:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	100	5	5	6.84
2	**microRNA-7b**	**481**	100	18	18	3.74
3	**microRNA-7c**	**21**	100	4	4	19.04
4	**microRNA-7d**	**100**	100	9	9	9
5	**microRNA-9**	**202**	100	8	8	3.96
6	**microRNA-10b**	**10**	100	2	2	20
7	**microRNA-21**	**581**	100	42	42	7.22
8	**microRNA-101**	**177**	100	16	16	9.03
9	**microRNA-146a**	**118**	100	3	3	2.54
10	**microRNA-146b**	**3**	100	0	0	0
11	**microRNA-200a**	**47**	100	1	1	2.12
12	**microRNA-200b**	**44**	100	1	1	2.27
13	**microRNA-203**	**13**	100	0	0	0
14	**microRNA-206**	**18**	100	4	4	22.22
15	**microRNA-211**		100	0	0	0
16	**microRNA-221**	**124**	100	10	10	8.06
17	**microRNA-335**	**2923**	100	15	15	.51
18	**microRNA-373**	**165**	100	6	6	3.63
Average					**8**	**6.680**

F. MicroCosm:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	507	1	0.197	1.33
2	**microRNA-7b**	**481**	523	27	5.16	5.61
3	**microRNA-7c**	**21**	498	2	.4016	9.52
4	**microRNA-7d**	**100**	498	5	1	5
5	**microRNA-9**	**202**	502	5	.996	2.47
6	**microRNA-10b**	**10**	486	2	.411	20
7	**microRNA-21**	**581**	492	74	15.040	12.73
8	**microRNA-101**	**177**	494	15	3.03	8.474
9	**microRNA-146a**	**118**	491	12	2.44	10.16
10	**microRNA-146b**	**3**	0	0	0	0
11	**microRNA-200a**	**47**	497	6	1.20	12.76
12	**microRNA-200b**	**44**	493	1	.210	2.27
13	**microRNA-203**	**13**	0	0	0	0
14	**microRNA-206**	**18**	491	2	.40	11.11
15	**microRNA-211**		493	0	0	0
16	**microRNA-221**	**124**	496	12	2.41	9.677
17	**microRNA-335**	**2923**	495	32	6.464	1.09
18	**microRNA-373**	**165**	533	8	1.5	4.84
Average					**2.27**	**6.50**

G. MiRanda:-

S.No.	microRNA	Total Correct	Total Predicted	Correctly Predicted	Precision	Sensitivity
1	**microRNA-7a**	**73**	499	7	1.40	9.58
2	**microRNA-7b**	**481**	508	46	9.05	9.56
3	**microRNA-7c**	**21**	497	3	14.91	14.28
4	**microRNA-7d**	**100**	498	22	4.41	22
5	**microRNA-9**	**202**	497	25	5.03	12.38
6	**microRNA-10b**	**10**	498	1	.20	10
7	**microRNA-21**	**581**	498	91	18.27	15.66
8	**microRNA-101**	**177**	449	43	9.57	24.29
9	**microRNA-146a**	**118**	500	9	1.8	7.62
10	**microRNA-146b**	**3**	499	1	.200	33.33
11	**microRNA-200a**	**47**	499	5	1.00	10.63
12	**microRNA-200b**	**44**	499	6	1.2	13.63
13	**microRNA-203**	**13**	499	2	.400	15.38
14	**microRNA-206**	**18**	499	4	.802	22.22
15	**microRNA-211**		499	0	0	0
16	**microRNA-221**	**124**	499	19	3.80	15.323
17	**microRNA-335**	**2923**	499	49	9.81	1.67
18	**microRNA-373**	**165**	499	12	2.40	7.27
Average					**4.68**	**13.60**

Table 5.5: <u>Comparison of microRNA target prediction algorithms</u>

Software	Average Precision	Average Sensitivity
TargetScan	4.85	10.98
PicTar	5.72	19.32
DIANA MicroT	5.50	22.17
EiMMo	5.86	20.26
Target Rank	8.00	6.68
MicroCosm	2.27	6.50
MiRanda	4.68	13.60

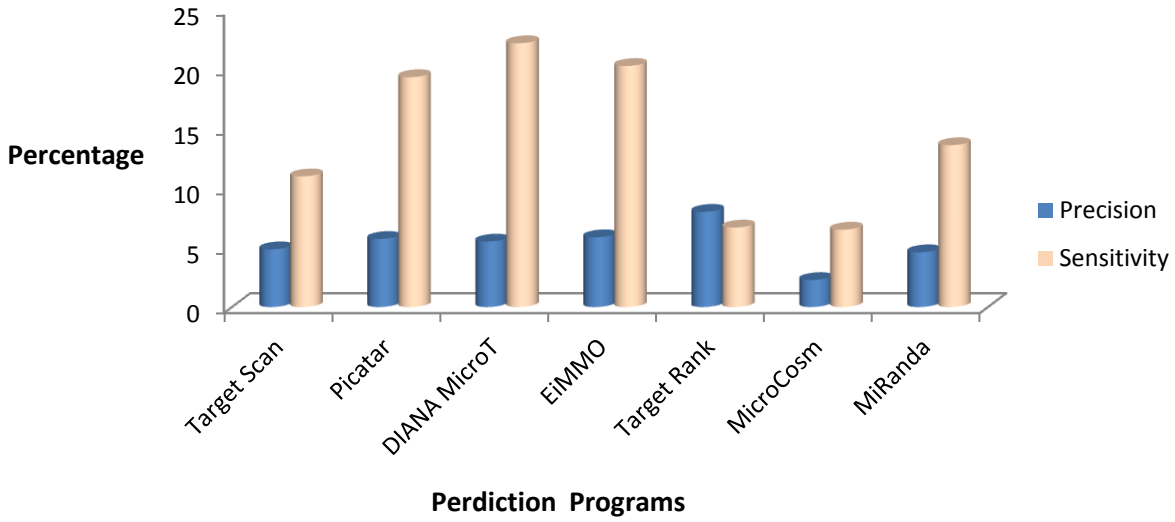

Figure 5.1: The comparative analysis of different tools, on the basis of Precision and Sensitivity. In this DIANA microT and EiMMO have highest sensitivity. All prediction tools have more or less similar precision.

Comparative analysis of different tools is showed that PicTar, DIANA microT and EiMMo having highest sensitivity. Most of the programs have similar precision in range of 5-8%.

D. Comparative Analysis of microRNA Prediction Tools on The Basis of Number of True Positives In Top Predictions

The target which is among the top predicted targets, they have more chance to get validated by experimental tools. The first 50 microRNA targets are most probable to get validated first.

Table 5.6: Number of True Positives among Top Ranked

Software	50	100	200	300	400	500
TargetScan	2.883	4.444	8.166	9.055	9.555	9.833
PicTar	3.33	6.111	11.11	14.38	16.333	18.111
EiMMo	4.44	8.44	14.277	19.722	24.55	29
DIANA MicroT	5.11	8.388	14.388	19.44	23	26.555
TargetRank	4.666	3.3				
MicroCosm	2.02	3.22	5.88	8.22	9.944	11.27
MiRanda	3.277	5.5	9.05	12.5	16.22	19.16

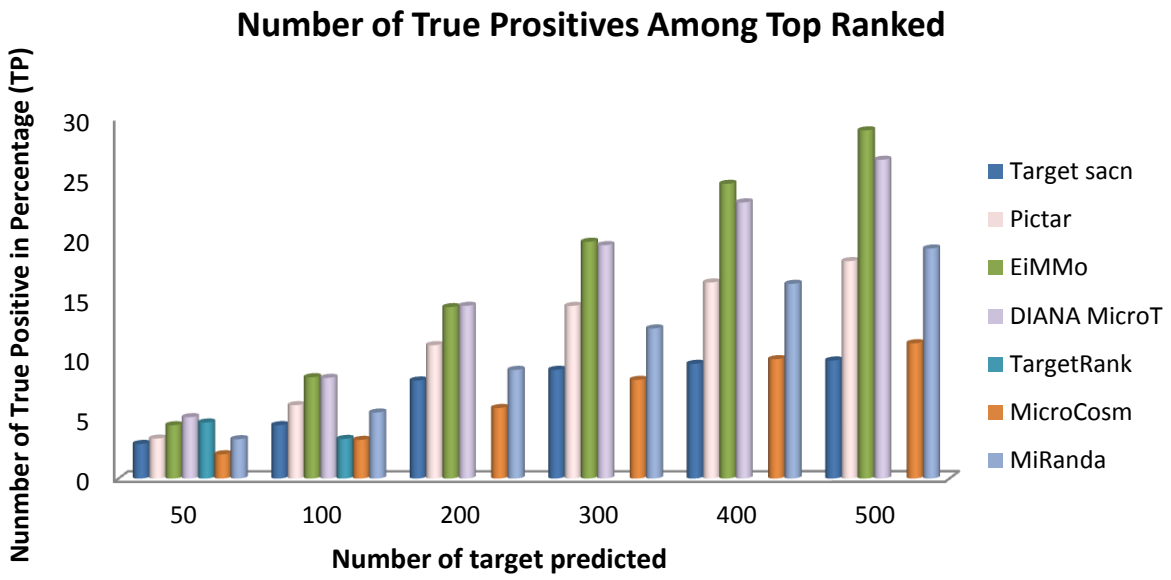

Figure 5.2: This graph shows the highest number of validated target in the top prediction. EiMMo have the highest number of true positive (TP) in its prediction.

In this result EiMMo and DIANA microT is the best performer among all prediction tools. EiMMo and DIANA MicroT are consistent from first 50 to 500 predictions. In up to 200 predictions DIANA MicroT predicted most true positives among all prediction tolls, but in 300, 400, and 500 EiMMo predicted highest true positives. This proves that initial predictions of DIANA MicroT having more chance to get validated than any other prediction tolls.

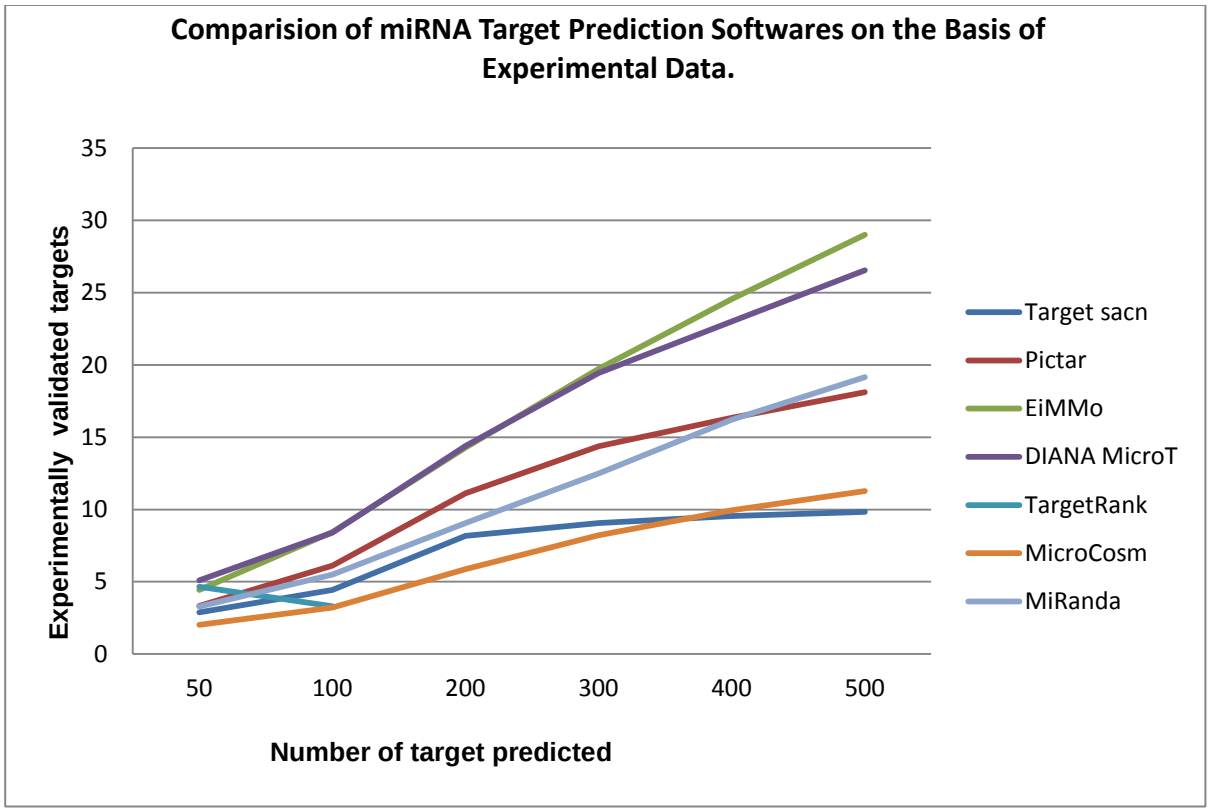

Figure 5.3: This graph showing number of true positive among first 50 predicted targets and so on, up to 500 predicted targets.

This graph is showing that EiMMO and DIANA microT are most consistent programs. PicTar achieved better result up to first 200 predictions but after that miRanda surpass PicTar.

Discussion

Biological function of microRNA in breast cancer is diverse, comprehensive and remains to be fully understood. The involvement of microRNAs in the initiation and progression of breast cancer holds great potential for new development in current diagnosis and therapeutic strategies in breast cancer management. High-throughput proteomics analysis allows researcher to obtain a wider view of microRNA function in cells. Such data may help in identification of new rules that govern microRNA function and serve as training sets for applications based on machine learning approach.

Many target prediction tools have been published, but which is best and accurate is the most difficult question to answer. They are very few comparative studies that independently compare target prediction programs. In this study, performance of microRNA target prediction tools have assessed based on predicted microRNA targets. All seven microRNA target prediction tools performance have been assessed on basis of validated targets from TarBase 6.0. Numbers of predictions predicted by different prediction tolls are different, for this reason, only first 500 predictions are taken for analysis. Sensitivity and precision of all seven programs have determined and compared. Average sensitivity is in the range of 15-20 % and average precision is in the range of 5-8 %. This result is opposite of the previous result. This may be due to the only selection of first 500 predictions of total predicted. Performance wise EiMMo, PicTar and DIANA microT performed well. This is similar with the previous results. Number of validated targets is very few as compared to the predicted targets.

In recent study Alexiou and colleagues (Alexiou et al., 2009) analysed different tools. They compared predicted targets of five programs DIANA-microT, TargetScan, PicTar, TargetScanS and EiMMO. These microRNA targets prediction programs were tested against genes proposed as targated in Selbach et.al. (Selbach et al., 2008). They noticed that these five programs has precision of ~50 percent with

sensitivity that of range from 6 to 12 %. These all algorithms are relying heavily on evolutionary conservation of the seed region. Previous studies done by Selbach et al., analyzed seven different tools. They compared the predicted targets of the programs TargetScanS, PicTar, rna22, PITA, miRBase, miRanda and DIANA-mircroT 3.0 with the results of a pSILAC analysis. All algorithms showed good result in test condition. Progress in high throughput experimental methods will lead to significant qualitative and quantitative improvements in the characterization of microRNA regulation. Huge Experimental data will enhance the level of benchmarking and will improve the assessment process. These results are supported by an older survey by Sethupathy and colleagues (Sethupathy et al., 2006). They analyzed TargetScan, PicTar, TargetScanS, PITA, miRanda and DIANA-mircroT performance by the use of experimentally supported microRNA:target interactions provided by TarBase.

Baek et al. compared the results of seven prediction tools with the results of their examination (Baek et al., 2008). They did a gene knockout concerning the microRNA mir-223 in mouse and monitored the changes of protein levels. Afterwards they tested the target predictions of miRBase:Targets, miRanda, PicTar, PITA and TargetScan with their data. TargetScan and PicTar proved to be the most effective tools among those methods, which considered conservation.

All these previous analysis are not very consistent in their results because of different microRNA sets analyzed and different software chosen for comparative analysis. Most importantly, those analyses, which have used experimental data from Selbach et al., (Selbach et al., 2008) have obtained similar result.

There is one important aspect of microRNA target prediction is the elucidation of the combinatorial effect of microRNAs (Vlachos et al., 2012). It is widely accepted that several microRNAs are co-regulated in microRNA gene cluster and are transcribed together. Additional levels of several microRNA may be correlated as marker of disease, indicate a co-regulation by more than one microRNA. Now the

main question arise that how do multiple microRNA affect single gene and how do multiple microRNA regulate biological pathway or disease. The second question needed complex computational approach that will precisely identify and predict microRNA regulatory network and model that interplay between microRNAs.

Metastatic microRNA have great role in breast cancer invasion, angiogenesis and metastasis. MicroRNA predicted targets compared with the validated targets available. This gives us an idea of microRNA involved in each step of cancer progression and metastasis. Animal models are in regular use in different cancer studies. Cancer metastasis is the main cause of most of the cancer death and most of the metastatic study based on animal model. In-silico microRNA target prediction gives us alternative to animal model system.

.

Conclusion

The result shows that microRNA have wide role in cancer development, progression and metastasis. Availability of experimentally validated data is one of the important aspects of microRNA target prediction.

In this study large numbers of existing computational algorithms for microRNA prediction are assessed. The prediction tools are of three categories. These all algorithms with their supported organism, website and approaches are listed in table-V and table- VI. To evaluate algorithms few representative algorithms are taken and assessed. A testing data set including experimentally validated positive microRNA targets was constructed. Each algorithm is evaluated on the basis of precision and sensitivity. Sensitivity and precision are estimated from the predicted targets and test data sets. The Sensitivity and precision gives us direct data for evaluation.

The analysis shows that utilizing more information makes the algorithm have better performance. Despite the recent advancement in algorithms on the miRNA target research, key problems still exist that prevent the computational approach from playing more active role in target prediction. These algorithms generally produce an excessively large number of false positives, but still far from any meaningful and workable hypotheses for subsequent experimental testing.

There is an urgent need to reassessment of prediction features which is used by different algorithms. Seed matching is one of the important feature which is used by different prediction tools, but still we need additional prediction features for more reliable prediction. One problem of using 5’ dominant site is that 3’ compensatory site containing a mismatch or wobble in the seed region cannot be detected by most target prediction methods. Although miRanda is sensitive for such targets (Sethupathy et al., 2006). More programs have to incorporate these features. Evolutionary conservation is another feature which reduces number of false positive and increase specificity. A highly stable RNA duplex is represented as having a

minimum free energy of hybridization (MFE). The calculation of MFE gives high number of true positive in prediction result.

This result is clearly highlighting role of miRNA in cancer metastasis and progression. These miRNA controls each steps of metastasis and maintain metastatic environment.

References

Alexiou, P., Maragkakis, M., Papadopoulos, G.L., Reczko, M., & Hatzigeorgiou, A.G. (2009). Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics, 25, 3049-3055.

Ambs, S., Prueitt, R.L., Yi, M., Hudson, R.S., Howe, T.M., Petrocca, F., Wallace, T.A., Liu, C.G., Volinia, S., & Calin, G.A. (2008). Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Research, 68, 6162-6170.

Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P., & Bartel, D. P. (2008). The impact of microRNAs on protein output. Nature, 455, 64-71. doi: 10.1038/nature07242

Bandres, E., Cubedo, E., Agirre, X., Malumbres, R., Zarate, R., Ramirez, N., Abajo, A., Navarro, A., Moreno, I., & Monzo, M. (2006). Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Molecular Cancer, 5, 29.

Bandyopadhyay, S., & Mitra, R. (2009). TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples. Bioinformatics, 25, 2625-2631. doi: 10.1093/bioinformatics/btp503

Blenkiron, C., Goldstein, L.D., Thorne, N.P., Spiteri, I., Chin, S.F., Dunning, M.J., Barbosa-Morais, N.L., Teschendorff, A.E., Green, A.R., & Ellis, I.O. (2007). MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. Genome Biology, 8, R214.

Bloomston, M., Frankel, W. L., Petrocca, F., Volinia, S., Alder, H., Hagan, J. P., Liu, C-G., Bhatt, D., Taccioli, C., & Croce, C. M. (2007). MicroRNA Expression Patterns to Differentiate Pancreatic Adenocarcinoma From Normal Pancreas and Chronic Pancreatitis. JAMA: The Journal of the American Medical Association, 297, 1901-1908. doi: 10.1001/jama.297.17.1901

Bommer, G.T., Gerin, I., Feng, Y., Kaczorowski, A.J., Kuick, R., Love, R.E., Zhai, Y., Giordano, T.J., Qin, Z.S., & Moore, B.B. (2007). p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Current Biology, 17, 1298-1307.

Bonci, D., Coppola, V., Musumeci, M., Addario, A., Giuffrida, R., Memeo, L., D'Urso, L., Pagliuca, A., Biffoni, M., Labbaye, C., Bartucci, M., Muto, G., Peschle, C., & De Maria, R. (2008). The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. Nature Medicine, 14, 1271-1277. doi: 10.1038/nm.1880

Boyerinas, B., Park, S.M., Shomron, N., Hedegaard, M.M., Vinther, J., Andersen, J.S., Feig, C., Xu, J., Burge, C.B., & Peter, M.E. (2008). Identification of let-7–regulated oncofetal genes. Cancer Research, 68, 2587-2591.

Burgler, C., & Macdonald, P. (2005). Prediction and verification of microRNA targets by MovingTargets, a highly adaptable prediction method. BMC Genomics, 6, 88.

Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S., & Brabletz, T. (2008). A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO Reports, 9, 582-589.

Byakodi, R., Byakodi, S., Hiremath, S., Byakodi, J., Adaki, S., Marathe, K., & Mahind, P. (2012). Oral Cancer in India: An Epidemiologic and Clinical Review. Journal of Community Health, 37,316-319.

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., & Croce, C. M. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proceedings of the National Academy of Sciences, 99, 15524-15529. doi: 10.1073/pnas.242606799

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., & Croce, C. M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proceedings of the National Academy of Sciences, 101, 2999-3004.

Carmeliet, P., & Jain, R.K. (2011). Molecular mechanisms and clinical applications of angiogenesis. Nature, 473, 298-307.

Chaffer, C. L., & Weinberg, R. A. (2011). A Perspective on Cancer Cell Metastasis. Science, 331, 1559-1564. doi: 10.1126/science.1203543

Chandra, V., Girijadevi, R., Nair, A., Pillai, S., & Pillai, R. (2010). MTar: a computational microRNA target prediction architecture for human transcriptome. BMC Bioinformatics, 11, S2.

Chaudhry, K. (2001). Tobacco Control in India. 50 Years of Cancer Control in India Retrieved March, 15, 2012, from http://mohfw.nic.in/WriteReadData/l892s/pg204to219-40207717.pdf

Ciafrè, S. A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C. G., Sabatino, G., Negrini, M., Maira, G., Croce, C. M., & Farace, M. G. (2005). Extensive modulation of a set of microRNAs in primary glioblastoma. Biochemical and Biophysical Research Communications, 334, 1351-1358. doi: 10.1016/j.bbrc.2005.07.030

Corcoran, C., Friel, A. M., Duffy, M. J., Crown, J., & O'Driscoll, L. (2011). Intracellular and extracellular microRNAs in breast cancer. Clinical Chemistry, 57, 18-32. doi: 10.1373/clinchem.2010.150730

Creighton, C. J., Nagaraja, A. K., Hanash, S. M., Matzuk, M. M., & Gunaratne, P. H. (2008). A bioinformatics tool for linking gene expression profiling results with public databases of microRNA target predictions. RNA, 14, 2290-2296. doi: 10.1261/rna.1188208

DeSantis, C., Siegel, R., Bandi, P., & Jemal, A. (2011). Breast cancer statistics, 2011. CA: A Cancer Journal for Clinicians, 61, 408-418. doi: 10.3322/caac.20134

Díaz, R., Silva, J., García, J.M., Lorenzo, Y., García, V., Peña, C., Rodríguez, R., Muñoz, C., García, F., & Bonilla, F. (2008). Deregulated expression of miR-106a predicts survival in human colon cancer patients. Genes, Chromosomes and Cancer, 47, 794-802.

Dombkowski, A. A., Sultana, Z., Craig, D. B., & Jamil, H. (2011). In silico Analysis of Combinatorial microRNA Activity Reveals Target Genes and Pathways Associated with Breast Cancer Metastasis. Cancer Informatics, 10, 13-29. doi: 10.4137/cin.s6631

Drakaki, A., & Iliopoulos, D. (2009). MicroRNA Gene Networks in Oncogenesis. Current Genomics, 10, 35-41.

Dweep, H., Sticht, C., Pandey, P., & Gretz, N. (2011). miRWalk – Database: Prediction of possible miRNA binding sites by “walking” the genes of three genomes. Journal of Biomedical Informatics, 44, 839-847. doi: 10.1016/j.jbi.2011.05.002

Dykxhoorn, D. M. (2010). MicroRNAs and metastasis: little RNAs go a long way. Cancer Research, 70, 6401-6406. doi: 10.1158/0008-5472.CAN-10-1346

Fabbri, M., Garzon, R., Cimmino, A., Liu, Z., Zanesi, N., Callegari, E., Liu, S., Alder, H., Costinean, S., & Fernandez-Cymering, C. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proceedings of the National Academy of Sciences, 104, 15805-15810.

Fidler. I.J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nature Reviews Cancer, 3, 1-6.

Fontana, L., Fiori, M.E., Albini, S., Cifaldi, L., Giovinazzi, S., Forloni, M., Boldrini, R., Donfrancesco, A., Federici, V., & Giacomini, P. (2008). Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. PLoS ONE, 3, e2236.

Frankel, L.B., Christoffersen, N.R., Jacobsen, A., Lindow, M., Krogh, A., & Lund, A.H. (2008). Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. Journal of Biological Chemistry, 283, 1026-1033.

Friedman, R.C., Farh, K.K.H., Burge, C.B., & Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Research, 19, 92-105.

Gamazon, E. R., Im, H. K., Duan, S., Lussier, Y. A., Cox, N. J., Dolan, M. E., & Zhang, W. (2010). ExprTarget: An Integrative Approach to Predicting Human MicroRNA Targets. PLoS ONE, 5, e13534. doi: 10.1371/journal.pone.0013534

Garofalo, M., & Croce, C. M. (2011). microRNAs: Master Regulators as Potential Therapeutics in Cancer. Annual Review of Pharmacology and Toxicology, 51, 25-43. doi: 10.1146/annurev-pharmtox-010510-100517

Garzon, R., Calin, G.A., & Croce, C.M. (2009). MicroRNAs in cancer. Annual Review of Medicine, 60, 167-179.

Gennarino, V. A., Sardiello, M., Avellino, R., Meola, N., Maselli, V., Anand, S., Cutillo, L., Ballabio, A., & Banfi, S. (2009). MicroRNA target prediction by expression analysis of host genes. Genome Research, 19, 481-490. doi: 10.1101/gr.084129.108

Ghosh, Z., Chakrabarti, J., & Mallick, B. (2007). miRNomics—The bioinformatics of microRNA genes. Biochemical and Biophysical Research Communications, 363, 6-11. doi: 10.1016/j.bbrc.2007.08.030

Gottardo, F., Liu, C. G., Ferracin, M., Calin, G. A., Fassan, M., Bassi, P., Sevignani, C., Byrne, D., Negrini, M., Pagano, F., Gomella, L. G., Croce, C. M., & Baffa, R. (2007). Micro-RNA profiling in kidney and bladder cancers. Urologic Oncology: Seminars and Original Investigations, 25, 387-392. doi: 10.1016/j.urolonc.2007.01.019

Gregory, P.A., Bracken, C.P., Bert, A.G., & Goodall, G.J. (2008). MicroRNAs as regulators of epithelial-mesenchymal transition. Cell Cycle, 7, 3112-3118.

Griffiths-Jones, S., Saini, H. K., van Dongen, S., & Enright, A. J. (2008). miRBase: tools for microRNA genomics. Nucleic Acids Research, 36, D154-D158. doi: 10.1093/nar/gkm952

Grün, D., Wang, Yi-Lu, Langenberger, D., Gunsalus, K. C., & Rajewsky, N. (2005). microRNA Target Predictions across Seven Drosophila Species and Comparison to Mammalian Targets. PLoS Computational Biology, 1, e13. doi: 10.1371/journal.pcbi.0010013

Hammell, M., Long, D., Zhang, L., Lee, A., Carmack, C.S., Han, M., Ding, Y., & Ambros, V. (2008). mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein–enriched transcripts. Nature Methods, 5, 813-819.

Hammell, Molly. (2010). Computational methods to identify miRNA targets. Seminars in Cell & Developmental Biology, 21, 738-744. doi: 10.1016/j.semcdb.2010.01.004

Han, J., Pedersen, J. S., Kwon, S. C., Belair, C. D., Kim, Y-K, Yeom, K-Hy, Yang, W-Y, Haussler, D., Blelloch, R., & Kim, V. N. (2009). Posttranscriptional Crossregulation between Drosha and DGCR8. Cell, 136, 75-84. doi: 10.1016/j.cell.2008.10.053

Hanahan, D., & Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell, 144, 646-674.

Harris, T. A., Yamakuchi, M., Ferlito, M., Mendell, J. T., & Lowenstein, C. J. (2008). MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. Proceedings of the National Academy of Sciences, 105, 1516-1521. doi: 10.1073/pnas.0707493105

Hausser, J., Berninger, P., Rodak, C., Jantscher, Y., Wirth, S., & Zavolan, M. (2009). MirZ: an integrated microRNA expression atlas and target prediction resource. Nucleic Acids Research, 37, W266-W272. doi: 10.1093/nar/gkp412

He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G. A., Liu, C-G., Franssila, K., Suster, S., Kloos, R. T., Croce, C. M., & de la Chapelle, A. (2005). The role of microRNA genes in papillary thyroid carcinoma. Proceedings of the National Academy of Sciences, 102, 19075-19080. doi: 10.1073/pnas.0509603102

He, L., He, X., Lim, L.P., De Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., & Ridzon, D. (2007). A microRNA component of the p53 tumour suppressor network. Nature, 447, 1130-1134.

Heneghan, H. M., Miller, N., Lowery, A. J., Sweeney, K. J., & Kerin, M. J. (2010). MicroRNAs as Novel Biomarkers for Breast Cancer. Journal of Oncology, 2010. doi: 10.1155/2010/950201

Hsu, S-D, Lin, F-M, Wu, W-Y, Liang, C., Huang, W-C, Chan, W-L, Tsai, W-T, Chen, G-Z, Lee, C-J, Chiu, C-M, Chien, C-H, Wu, M-C, Huang, C-Y, Tsou, A-P, & Huang, H-D. (2010). miRTarBase: a database curates experimentally validated microRNA–target interactions. Nucleic Acids Research, 39, D163-169. doi: 10.1093/nar/gkq1107

Huang, J.C., Morris, Q.D., & Frey, B.J. (2007). Bayesian inference of microRNA targets from sequence and expression data. Journal of Computational Biology, 14, 550-563.

Huang, Q., Gumireddy, K., Schrier, M., Le Sage, C., Nagel, R., Nair, S., Egan, D.A., Li, A., Huang, G., & Klein-Szanto, A.J. (2008). The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. Nature cell biology, 10, 202-210.

Hurst, D. R., Edmonds, M. D., Scott, G. K., Benz, C. C., Vaidya, K. S., & Welch, D. R. (2009). Breast Cancer Metastasis Suppressor 1 Up-regulates miR-146, Which Suppresses Breast Cancer Metastasis. Cancer Research, 69, 1279-1283. doi: 10.1158/0008-5472.can-08-3559

Ichimi, T., Enokida, H., Okuno, Y., Kunimoto, R., Chiyomaru, T., Kawamoto, K., Kawahara, K., Toki, K., Kawakami, K., Nishiyama, K., Tsujimoto, G., Nakagawa, M., & Seki, N. (2009). Identification of novel microRNA targets based on microRNA signatures in bladder cancer. International Journal of Cancer, 125, 345-352. doi: 10.1002/ijc.24390

Iorio, M. V., & Croce, C. M. (2012). microRNA involvement in human cancer. Carcinogenesis, 33, 1126-1133. doi: 10.1093/carcin/bgs140

Iorio, M. V., Ferracin, M., Liu, C-G, Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Ménard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, George A., Querzoli, Patrizia, Negrini, Massimo, & Croce, C. M. (2005). MicroRNA Gene Expression Deregulation in Human Breast Cancer. Cancer Research, 65, 7065-7070. doi: 10.1158/0008-5472.can-05-1783

Iorio, M. V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., Taccioli, C., Volinia, S., Liu, C-G, Alder, H., Calin, G. A., Ménard, S., & Croce, C. M. (2007). MicroRNA Signatures in Human Ovarian Cancer. Cancer Research, 67, 8699-8707. doi: 10.1158/0008-5472.can-07-1936

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. CA: A Cancer Journal for Clinicians, 61, 69-90. doi: 10.3322/caac.20107

Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., Li, M., Wang, G., & Liu, Y. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Research, 37, D98-D104.

John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., & Marks, D. S. (2004). Human MicroRNA targets. PLoS Biology, 2, e363. doi: 10.1371/journal.pbio.0020363

Johnnidis, J. B., Harris, M. H., Wheeler, R. T., Stehling-Sun, S., Lam, M. H., Kirak, O., Brummelkamp, T. R., Fleming, M. D., & Camargo, F. D. (2008). Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. Nature, 451, 1125-1129. doi: 10.1038/nature06607

Joung, Je-G., & Fei, Z. (2009). Computational identification of condition-specific miRNA targets based on gene expression profiles and sequence information. BMC Bioinformatics, 10, S34.

Kefas, B., Godlewski, J., Comeau, L., Li, Y., Abounader, R., Hawkinson, M., Lee, J., Fine, H., Chiocca, E. A., Lawler, S., & Purow, B. (2008). microRNA-7 Inhibits the Epidermal Growth Factor Receptor and the Akt Pathway and Is Down-regulated in Glioblastoma. Cancer Research, 68, 3566-3572. doi: 10.1158/0008-5472.can-07-6639

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007). The role of site accessibility in microRNA target recognition. Nature Genetics, 39, 1278-1284. doi: 10.1038/ng2135

Kim, S-K, Nam, J-W, Rhee, J-K, Lee, W-J, & Zhang, B-T. (2006). miTarget: microRNA target gene prediction using a support vector machine. BMC Bioinformatics, 7, 411.

Kim, V.N., Han, J., & Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nature Reviews Molecular Cell Biology, 10, 126-139.

Kozomara, A., & Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Research, 39, D152-D157.

Krek, A., Grün, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., Da Piedade, I., Gunsalus, K.C., & Stoffel, M. (2005). Combinatorial microRNA target predictions. Nature Genetics, 37, 495-500.

Krichevsky, A.M., & Gabriely, G. (2009). miR-21: a small multi faceted RNA. Journal of Cellular and Molecular Medicine, 13, 39-53.

Kroh, E.M., Parkin, R.K., Mitchell, P.S., & Tewari, M. (2010). Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods, 50, 298-301.

Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nature Reviews Genetics, 11, 597-610.

Lall, S., Grün, D., Krek, A., Chen, K., Wang, Y.L., Dewey, C.N., Sood, P., Colombo, T., Bray, N., & MacMenamin, P. (2006). A genome-wide map of conserved microRNA targets in C. elegans. Current Biology, 16, 460-471.

Lee, E. J, Gusev, Y., Jiang, J., Nuovo, G. J., Lerner, M. R., Frankel, W. L., Morgan, D. L., Postier, R. G., Brackett, D. J., & Schmittgen, T. D. (2007). Expression profiling identifies microRNA signature in pancreatic cancer. International Journal of Cancer, 120, 1046-1054. doi: 10.1002/ijc.22394

Lee, R.C., Feinbaum, R.L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell, 75, 843-854.

Lee, Y., Jeon, K., Lee, J-T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. [10.1093/emboj/cdf476]. EMBO J, 21, 4663-4670.

Leucci, E., Cocco, M., Onnis, A., De Falco, G., van Cleef, P., Bellan, C., van Rijk, A., Nyagol, J., Byakika, B., Lazzi, S., Tosi, P., van Krieken, H., & Leoncini, L. (2008). MYC translocation-negative classical Burkitt lymphoma cases: an alternative pathogenetic mechanism involving miRNA deregulation. The Journal of Pathology, 216, 440-450. doi: 10.1002/path.2410

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. Cell, 120, 15-20. doi: 10.1016/j.cell.2004.12.035

Lin, S.L., Chiang, A., Chang, D., & Ying, S.Y. (2008). Loss of mir-146a function in hormone-refractory prostate cancer. RNA, 14, 417-424.

Liu, H., Yue, D., Chen, Y., Gao, S-J, & Huang, Y. (2010). Improving performance of mammalian microRNA target prediction. BMC Bioinformatics, 11, 476.

Liu, H., Yue, D., Zhang, L., Chen, Y., Gao, S-J, & Huang, Y. (2010). A Bayesian approach for identifying miRNA targets by combining sequence prediction and gene expression profiling. BMC Genomics, 11, S12.

Liu, X., Chen, Z., Yu, J., Xia, J., & Zhou, X. (2009). MicroRNA profiling and head and neck cancer. Comparative and Functional Genomics, 2009, 11. doi: 10.1155/2009/837514

Lodygin, D., Tarasov, V., Epanchintsev, A., Berking, C., Knyazeva, T., Korner, H., Knyazev, P., Diebold, J., & Hermeking, H. (2008). Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. Cell Cycle, 7, 2591-2600.

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., & Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. Nature, 435, 834-838. doi: 10.1038/nature03702

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., & Kutay, U. (2004). Nuclear export of microRNA precursors. Science, 303, 95-98.

Ma, L., Teruya-Feldstein, J., & Weinberg, R.A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature, 449, 682-688.

Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T. T., Valastyan, S., Westermann, F., Speleman, F., Vandesompele, J., & Weinberg, R. A. (2010). miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nature Cell Biology, 12, 247-256. doi: 10.1038/ncb2024

MacFarlane, L-A., & R. Murphy, P. (2010). MicroRNA: Biogenesis, Function and Role in Cancer. Current Genomics, 11, 537-561. doi: 10.2174/138920210793175895

Maragkakis, M., Reczko, M., Simossis, VA, Alexiou, P., Papadopoulos, GL, Dalamagas, T., Giannopoulos, G., Goumas, G., Koukis, E., & Kourtis, K. (2009). DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Research, 37, W273-W276.

Marín, Ray M., & Vaníček, Jiří. (2011). Efficient use of accessibility in microRNA target prediction. Nucleic Acids Research, 39, 19-29. doi: 10.1093/nar/gkq768

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell, 110, 563-574.

Marton, S., Garcia, M. R., Robello, C., Persson, H., Trajtenberg, F., Pritsch, O., Rovira, C., Naya, H., Dighiero, G., & Cayota, A. (2007). Small RNAs analysis in CLL reveals a deregulation of miRNA expression and novel miRNA candidates of putative relevance in CLL pathogenesis. Leukemia, 22, 330-338.

Mayr, C., Hemann, M.T., & Bartel, D.P. (2007). Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science's STKE, 315, 1576-1579.

Mendell, J. T. (2008). miRiad Roles for the miR-17-92 Cluster in Development and Disease. Cell, 133, 217-222. doi: 10.1016/j.cell.2008.04.001

Meng, F., Henson, R., Wehbe-Janek, H., Ghoshal, K., Jacob, S.T., & Patel, T. (2007). MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. Gastroenterology, 133, 647-658.

Miller, T.E., Ghoshal, K., Ramaswamy, B., Roy, S., Datta, J., Shapiro, C.L., Jacob, S., & Majumder, S. (2008). MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1. Journal of Biological Chemistry, 283, 29897-29903.

Miranda, K. C., Huynh, T., Tay, Y., Ang, Y-S, Tam, W-L, Thomson, A. M., Lim, B., & Rigoutsos, I. (2006). A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes. Cell, 126, 1203-1217. doi: 10.1016/j.cell.2006.07.031

Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., & Allen, A. (2008). Circulating microRNAs as stable blood-based markers for cancer detection. Proceedings of the National Academy of Sciences, 105, 10513-10518.

Nakajima, G., Hayashi, K., Xi, Y., Kudo, K., Uchida, K., Takasaki, K., Yamamoto, M., & Ju, J. (2006). Non-coding microRNAs hsa-let-7g and hsa-miR-181b are associated with chemoresponse to S-1 in colon cancer. Cancer Genomics-Proteomics, 3, 317-324.

Nam, E. J., Yoon, H., Kim, S. W., Kim, H., Kim, Y. T., Kim, J. H., Kim, J. W., & Kim, S. (2008). MicroRNA Expression Profiles in Serous Ovarian Carcinoma. Clinical Cancer Research, 14, 2690-2695. doi: 10.1158/1078-0432.ccr-07-1731

Nam, S., Kim, B., Shin, S., & Lee, S. (2008). miRGator: an integrated system for functional annotation of microRNAs. Nucleic Acids Research, 36, D159-D164. doi: 10.1093/nar/gkm829

Nasser, Mohd. W., Datta, J., Nuovo, G., Kutay, H., Motiwala, T., Majumder, S., Wang, B., Suster, S., Jacob, S. T., & Ghoshal, K. (2008). Down-regulation of Micro-RNA-1 (miR-1) in Lung Cancer. Journal of Biological Chemistry, 283, 33394-33405. doi: 10.1074/jbc.M804788200

Navarro, A., Gaya, A., Martinez, A., Urbano-Ispizua, A., Pons, A., Balagué, O., Gel, B., Abrisqueta, P., Lopez-Guillermo, A., & Artells, R. (2008). MicroRNA expression profiling in classic Hodgkin lymphoma. Blood, 111, 2825-2832.

Negrini, S., Gorgoulis, V.G., & Halazonetis, T.D. (2010). Genomic instability—an evolving hallmark of cancer. Nature Reviews Molecular Cell Biology, 11, 220-228.

Nguyen, D.X., & Massagué, J. (2007). Genetic determinants of cancer metastasis. Nature Reviews Genetics, 8, 341-352.

Nielsen, C.B., Shomron, N., Sandberg, R., Hornstein, E., Kitzman, J., & Burge, C.B. (2007). Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. RNA, 13, 1894-1910.

Noh, S., Jung, J-J, Jung, M., Kim, K-H., Lee, H. Y., Wang, B., Cho, J., Kim, T. S., Jeung, H. C. , & Rha, S. Y. (2012). Body fluid MMP-2 as a putative biomarker in metastatic breast cancer. Oncology Letters, 3, 699-703.

Ozen, M., Creighton, C. J., Ozdemir, M., & Ittmann, M. (2007). Widespread deregulation of microRNA expression in human prostate cancer. Oncogene, 27, 1788-1793.

Park, N.J., Zhou, H., Elashoff, D., Henson, B.S., Kastratovic, D.A., Abemayor, E., & Wong, D.T. (2009). Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. Clinical Cancer Research, 15, 5473-5477.

Penna, E., Orso, F., Cimino, D., Tenaglia, E., Lembo, A., Quaglino, E., Poliseno, L., Haimovic, A., Osella-Abate, S., De Pitta, C., Pinatel, E., Stadler, M. B., Provero, P., Bernengo, M. G., Osman, I., & Taverna, D. (2011). microRNA-214 contributes to melanoma tumour progression through suppression of TFAP2C. EMBO Journal, 30, 1990-2007. doi: 10.1038/emboj.2011.102

Pillai, R.S. (2005). MicroRNA function: Multiple mechanisms for a tiny RNA?. RNA, 11, 1753-1761.

Polyak, K., & Weinberg, R.A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nature Reviews Cancer, 9, 265-273.

Ragan, C., Cloonan, N., Grimmond, S. M., Zuker, M., & Ragan, M.A. (2009). Transcriptome-Wide Prediction of miRNA Targets in Human and Mouse Using FASTH. PLoS ONE, 4, e5745. doi: 10.1371/journal.pone.0005745

Rehmsmeier, M., Steffen, P., Hochsmann, M., & Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. RNA, 10, 1507-1517. doi: 10.1261/rna.5248604

Rehsmeier, M., Steffen, P., Hochsmann, M., & Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. RNA, 10, 1507-1517. doi: 10.1261/rna.5248604

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature, 403, 901-906.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., & Bradley, A. (2004). Identification of Mammalian microRNA Host Genes and Transcription Units. Genome Research, 14, 1902-1910. doi: 10.1101/gr.2722704

Roth, C., Rack, B., Müller, V., Janni, W., Pantel, K., & Schwarzenbach, H. (2010). Circulating microRNAs as blood-based markers for patients with primary and metastatic breast cancer. Breast Cancer Research, 12, R90.

Ruan, J., Chen, H., Kurgan, L., Chen, K., Kang, C., & Pu, P. (2008). HuMiTar: A sequence-based method for prediction of human microRNA targets. Algorithms for Molecular Biology, 3, 16.

Rusinov, V., Baev, V., Minkov, I. N., & Tabler, M. (2005). MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. Nucleic Acids Research, 33, W696-W700. doi: 10.1093/nar/gki364

Saetrom, O., Snove, O., & Saetrom, P. (2005). Weighted sequence motifs as an improved seeding step in microRNA target prediction algorithms. RNA, 11, 995-1003. doi: 10.1261/rna.7290705

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature, 455, 58-63. doi: 10.1038/nature07228

Sempere, L. F., Christensen, M., Silahtaroglu, A., Bak, M., Heath, C. V., Schwartz, G., Wells, W., Kauppinen, S., & Cole, C. N. (2007). Altered MicroRNA Expression Confined to Specific Epithelial Cell Subpopulations in Breast Cancer. Cancer Research, 67, 11612-11620. doi: 10.1158/0008-5472.can-07-5019

Sethupathy, P., Megraw, M., & Hatzigeorgiou, A.G. (2006). A guide through present computational approaches for the identification of mammalian microRNA targets. Nature Methods, 3, 881-886.

Shackleton, M., Quintana, E., Fearon, E. R., & Morrison, S. J. (2009). Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. Cell, 138**,** 822-829. doi: 10.1016/j.cell.2009.08.017

Silber, J., Lim, D., Petritsch, C., Persson, A., Maunakea, A., Yu, M., Vandenberg, S., Ginzinger, D., James, C.D., & Costello, J. (2008). miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. BMC Medicine, 6, 14.

Slaby, O., Svoboda, M., Fabian, P., Smerdova, T., Knoflickova, D., Bednarikova, M., Nenutil, R., & Vyzula, R. (2007). Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer. Oncology, 72, 397-402.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B., & Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′ UTR evolution. Cell, 123, 1133-1146.

Sturm, M., Hackenberg, M., Langenberger, D., & Frishman, D. (2010). TargetSpy: a supervised machine learning approach for microRNA target prediction. BMC Bioinformatics, 11, 292.

Sun, T., Wang, Q., Balk, S., Brown, M., Lee, G.S.M., & Kantoff, P. (2009). The role of microRNA-221 and microRNA-222 in androgen-independent prostate cancer cell lines. Cancer Research, 69, 3356.

Sun, W., Julie Li, Y.S., Huang, H.D., Shyy, J.Y.J., & Chien, S. (2010). microRNA: a master regulator of cellular processes for bioengineering systems. Annual Review of Biomedical Engineering, 12, 1-27.

Symons, M., & Segall, J.E. (2009). Rac and Rho driving tumor invasion: who's at the wheel? Genome Biology, 10, 213.

Tavazoie, S.F., Alarcón, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P.D., Gerald, W.L., & Massagué, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. Nature, 451, 147-152.

Thadani, R., & Tammi, M.T. (2006). MicroTar: predicting microRNA targets from RNA duplexes. BMC Bioinformatics, 7, S20.

Valastyan, S. (2012). Roles of MicroRNAs and Other Non-coding RNAs in Breast Cancer Metastasis. Journal of Mammary Gland Biology and Neoplasia, 17, 23-32. doi: 10.1007/s10911-012-9241-9

Valastyan, S., Reinhardt, F., Benaich, N., Calogrias, D., Szász, A.M., Wang, Z.C., Brock, J.E., Richardson, A.L., & Weinberg, R.A. (2009). A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell, 137, 1032-1046.

Varambally, S., Cao, Q., Mani, R.S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., Brenner, J. C., Yu, J., Kim, J. H., Han, B., Tan, P., Sinha, C. K., Lonigro, R. J., Palanisamy, N., Maher, C. A., & Chinnaiyan, A. M. (2008). Genomic Loss of microRNA-101 Leads to Overexpression of Histone Methyltransferase EZH2 in Cancer. Science, 322, 1695-1699. doi: 10.1126/science.1165395

Varnholt, H., Drebber, U., Schulze, F., Wedemeyer, I., Schirmacher, P., Dienes, H-P, & Odenthal, M. (2008a). MicroRNA gene expression profile of hepatitis C virus–associated hepatocellular carcinoma. Hepatology, 47, 1223-1232. doi: 10.1002/hep.22158

Vergoulis, T., Vlachos, I.S., Alexiou, P., Georgakilas, G., Maragkakis, M., Reczko, M., Gerangelos, S., Koziris, N., Dalamagas, T., & Hatzigeorgiou, A.G. (2012). TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. Nucleic Acids Research, 40, D222-D229.

Visone, R., Pallante, P., Vecchione, A., Cirombella, R., Ferracin, M., Ferraro, A., Volinia, S., Coluzzi, S., Leone, V., & Borbone, E. (2007). Specific microRNAs are downregulated in human thyroid anaplastic carcinomas. Oncogene, 26, 7590-7595.

Vlachos, I. S., Kostoulas, N., Vergoulis, T., Georgakilas, G., Reczko, M., Maragkakis, M., Paraskevopoulou, M. D., Prionidis, K., Dalamagas, T., & Hatzigeorgiou, A. G. (2012). DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Research, 1-7. doi: 10.1093/nar/gks494

Volinia, S., Calin, G. A., Liu, C-G, Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G.,

Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C., & Croce, C. M. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. Proceedings of the National Academy of Sciences, 103, 2257-2261. doi: 10.1073/pnas.0510565103

Wang, X., & El Naqa, I. M. (2008). Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics, 24, 325-332. doi: 10.1093/bioinformatics/btm595

Wang, X., Tang, S., Le, S.Y., Lu, R., Rader, J.S., Meyers, C., & Zheng, Z.M. (2008). Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. PLoS ONE, 3, e2557.

Watanabe, Y., Tomita, M., & Kanai, A. (2007). Computational Methods for MicroRNA Target Prediction. In J. Rossi John & J. Hannon Gregory (Eds.), Methods in Enzymology (Vol. 427, pp. 65-86): Academic Press.

Westholm, J.O., & Lai, E.C. (2011). Mirtrons: microRNA biogenesis via splicing. Biochimie, 93, 1897-1904.

Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell, 75, 855-862.

Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation.. Nature Cell Biology, 11, 228-234. 10.1038/ncb0309-228

Witkos, TM, Koscianska, E., & Krzyzosiak, WJ. (2011). Practical aspects of microRNA target prediction. Current molecular medicine, 11, 93-109.

Wong, T.S., Ho, W.K., Chan, JY, Ng, RW, & Wei, W.I. (2009). Mature miR-184 and squamous cell carcinoma of the tongue. ScientificWorldJournal, 9, 130-132.

Wong, T.S., Liu, X.B., Wong, B.Y.H., Ng, R.W.M., Yuen, A.P.W., & Wei, W.I. (2008). Mature miR-184 as potential oncogenic microRNA of squamous cell carcinoma of tongue. Clinical Cancer Research, 14, 2588-2592.

Xi, Y., Formentini, A., Chien, M., Weir, D.B., Russo, J.J., Ju, J., Kornmann, M., & Ju, J. (2006). Prognostic values of microRNAs in colorectal cancer. Biomarker Insights, 1, 113.

Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X., & Li, T. (2009). miRecords: an integrated resource for microRNA–target interactions. Nucleic Acids Research, 37, D105-D110. doi: 10.1093/nar/gkn851

Yan, L.X., Huang, X.F., Shao, Q., Huang, M.A.Y., Deng, L., Wu, Q.L., Zeng, Y.X., & Shao, J.Y. (2008). MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. RNA, 14, 2348-2360.

Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R. M., Okamoto, A., Yokota, J., Tanaka, T., Calin, G. A., Liu, C. G., Croce, C. M., & Harris, C. C. (2006). Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer cell, 9, 189-198.

Yang, H., Kong, W., He, L., Zhao, J.J., O'Donnell, J.D., Wang, J., Wenham, R.M., Coppola, D., Kruk, P.A., & Nicosia, S.V. (2008). MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. Cancer Research, 68, 425-433.

Yang, J. H., Li, J. H., Shao, P., Zhou, H., Chen, Y. Q., & Qu, L. H. (2010). starBase: a database for exploring microRNA–mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Research, D202-D209. doi: 10.1093/nar/gkq1056

Yang, K., Handorean, A.M., & Iczkowski, K.A. (2009). MicroRNAs 373 and 520c are downregulated in prostate cancer, suppress CD44 translation and enhance invasion of prostate cancer cells in vitro. International Journal of Clinical and Experimental Pathology, 2, 361-369.

Yang, Y., Wang, Y. P., & Li, K. B. (2008). MiRTif: a support vector machine-based microRNA target interaction filter. BMC Bioinformatics, 9, S4.

Ye, Wenbin, Lv, Qing, Wong, Chung-Kwun Amy, Hu, Sean, Fu, Chao, Hua, Zhong, Cai, Guoping, Li, Guoxi, Yang, Burton B., & Zhang, Yaou. (2008). The Effect of Central Loops in miRNA:MRE Duplexes on the Efficiency of miRNA-Mediated Gene Regulation. PLoS ONE, 3, e1719. doi: 10.1371/journal.pone.0001719

Yousef, M., Jung, S., Showe, L., & Showe, M. (2007). Recursive Cluster Elimination (RCE) for classification and feature selection from gene expression data. BMC Bioinformatics, 8, 144.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., & Song, E. (2007). let-7 Regulates Self Renewal and Tumorigenicity of Breast Cancer Cells. Cell, 131, 1109-1123. doi: 10.1016/j.cell.2007.10.054

Yue,D., Liu, H. & Huang, Y. (2009). Survey of Computational Algorithms for MicroRNA Target Prediction. Current Genomics, 10, 478–492.

Zhang, L., Huang, J.., Yang, N., Greshock, J., Megraw, M. S., Giannakakis, A., Liang, S., Naylor, T. L., Barchetti, A., Ward, M. R., Yao, G., Medina, A., O'Brien-Jenkins, A., Katsaros, D., Hatzigeorgiou, A.., Gimotty, P. A., Weber, B. L., & Coukos, G. (2006). microRNAs exhibit high frequency genomic alterations in human cancer. Proceedings of the National Academy of Sciences, 103, 9136-9141. doi: 10.1073/pnas.0508889103

Zhang, Y. (2005). miRU: an automated plant miRNA target prediction server. Nucleic Acids Research, 33, W701-W704.

Zhang, Y., & Verbeek, F.J. (2010). Comparison and integration of target prediction algorithms for microRNA studies. Journal of Integrative Bioinformatics, 7.

Zhu, S., Wu, H., Wu, F., Nie, D., Sheng, S., & Mo, Y.Y. (2008). MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. Cell Research, 18, 350-359.

[DIANA LAB] DANA microT. (2012). DIANA LAB software page. < http://diana.cslab.ece.ntua.gr/DianaTools/index.php?r=microtv4/index>. Accessed on 2012 May 16.

[EIMMo] ElMMo miRNA target prediction server. (2012). homepage. < http://www.mirz.unibas.ch/ElMMo3/index.php>. Accessed on 2012 May 16

[EBI] microCosm. (2012). Enright lab microcosm targets. < http://www.ebi.ac.uk/enright-srv/microcosm/cgi-bin/targets/v5/genome.pl>. Accessed on 2012April 16.

[GLOBOCAN] GLOBOCAN, 2008. (2012). May, 22. GLOBOCAN facts sheets. < http://globocan.iarc.fr/factsheets/populations/factsheet.asp?uno=900>. Accessed 2012 May15.

[miRBase] miRBase. (2012). November. miRBase Release-18 readme page. < ftp://mirbase.org/pub/mirbase/CURRENT/README>. Accessed 2012 Feb 15.

[microRNA.org] miRanda. (2012). home page. <http://www.microrna.org/microrna/home.do>. Accessed on 2012 May 16.

[PicTar] PicTar (2012). March, 26, 2007. PicTar Web Interface. < http://pictar.mdc-berlin.de/cgi-bin/PicTar_vertebrate.cgi >. Accessed 2012 April 17.

[TargetScanHuman] TargetScan (2012). TargetScan home page. < http://www.targetscan.org/>. Accessed 2012 May 16.

[TargetRank] TargetRank. (2012). TargetRank home page.< http://genes.mit.edu/targetrank/>. Accessed on 2012April 16.

[WHO] World Health Organization. (2012). February. WHO Media Centre, < http://www.who.int/topics/cancer/en/index.html>. Accessed 2012 Feb 15.