V22: involvement of microRNAs in GRNs

What are **microRNAs**?

How can one **identify** microRNAs?

What is the **function** of microRNAs?

diseases such as cancer and metabolic disorders[3,4]. The number of miRNAs encoded by the genomes of the organisms that have been studied so far varies considerably from a handful to up to 500 in mammals[1,2]. Computational predictions and genome-wide identification of miRNA targets estimate that each animal miRNA regulates hundreds of different mRNAs, suggesting that a remarkably large proportion of the transcriptome (about 50% in humans) is subject to miRNA regulation[1,2].

Elisa Izaurralde,
MPI Tübingen

Huntzinger, Izaurralde, Nat. Rev. Genet. 12, 99 (2011)

Laird, Hum Mol Gen 14, R65 (2005)

RNA world

short name	full name	function	oligomerization
mRNA, rRNA, tRNA,	you know them well ...		Single-stranded
snRNA	small nuclear RNA	splicing and other functions	
snoRNA	small nucleolar RNA	nucleotide modification of RNAs	
Long ncRNA	Long noncoding RNA	various	
miRNA	microRNA	gene regulation	single-stranded
siRNA	small interfering RNA	gene regulation	double-stranded

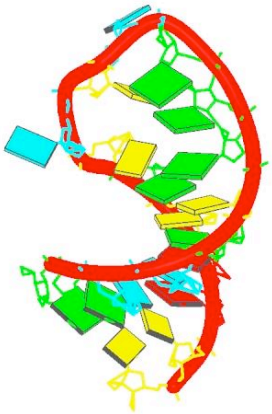
RNA structure

Also single stranded RNA molecules frequently adopt a specific **tertiary structure**.

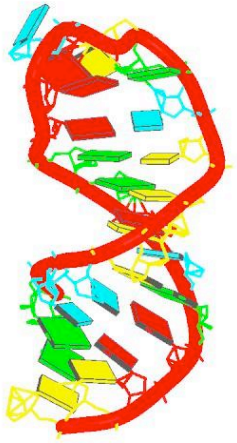
The scaffold for this structure is provided by secondary structural elements which are **H-bonds** within the molecule.

This leads to several recognizable structural "domain“ types of secondary structure such as **hairpin loops**, **bulges** and **internal loops**.

RNA hairpin 2RLU



Stem loop 1NZ1



www.rcsb.org

snRNAs

Small nuclear RNA (**snRNA**) are found within the **nucleus** of eukaryotic cells.

They are transcribed by RNA polymerase II or RNA polymerase III and are involved in a variety of important processes such as
- RNA splicing,
- regulation of transcription factors or RNA polymerase II, and
- maintaining the telomeres.

snRNAs are always associated with specific proteins.

The snRNA:protein complexes are referred to as
small nuclear ribonucleoproteins (**snRNP**) or sometimes as **snurps**.

www.wikipedia.org

snoRNAs

A large **subgroup** of snRNAs are known as small nucleolar RNAs (**snoRNAs**).

These are small RNA molecules that play an essential role in **RNA biogenesis** and guide chemical modifications of rRNAs, tRNAs and snRNAs.

They are located in the nucleolus and the cajal bodies of eukaryotic cells.

www.wikipedia.org

RNA interference

RNA interference may involve siRNAs or miRNAs.

Nobel prize in Physiology or Medicine **2006**
for their discovery of RNAi in *C. elegans* in 1998.

Andrew Fire

Craig Mello

Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*

Andrew Fire*, SiQun Xu*, Mary K. Montgomery*, Steven A. Kostas*†, Samuel E. Driver‡ & Craig C. Mello‡

* *Carnegie Institution of Washington, Department of Embryology, 115 West University Parkway, Baltimore, Maryland 21210, USA*
† *Biology Graduate Program, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA*
‡ *Program in Molecular Medicine, Department of Cell Biology, University of Massachusetts Cancer Center, Two Biotech Suite 213, 373 Plantation Street, Worcester, Massachusetts 01605, USA*

www.wikipedia.org

siRNAs

Small interfering RNA (**siRNA**), sometimes known as
short interfering RNA or silencing RNA, is a class of
- **double-stranded RNA molecules**,
- that are 20-25 nucleotides in length (often precisely 21 nt) and
play a variety of roles in biology.

Most notably, siRNA is involved in the **RNA interference (RNAi)** pathway,
where it interferes with the expression of a specific gene.

In addition to their role in the RNAi pathway,
siRNAs also act in RNAi-related pathways,
e.g., as an antiviral mechanism or in
shaping the chromatin structure of a genome.

www.wikipedia.org

miRNAs

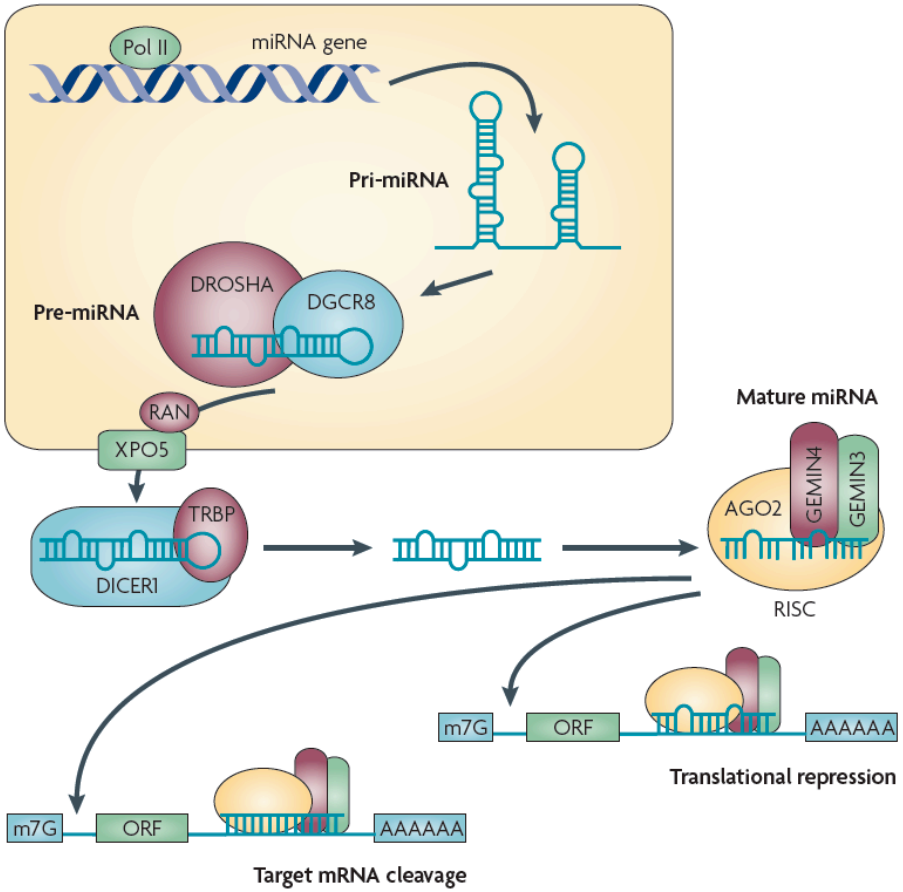
In contrast to double-stranded siRNA,
microRNAs (miRNA) are **single-stranded RNA** molecules
of 21-23 nucleotides in length.

miRNAs have a crucial role in regulating gene expression.

Remember: miRNAs are encoded by DNA but not translated into protein (non-coding RNA).

www.wikipedia.org

Overview of the miRNA network



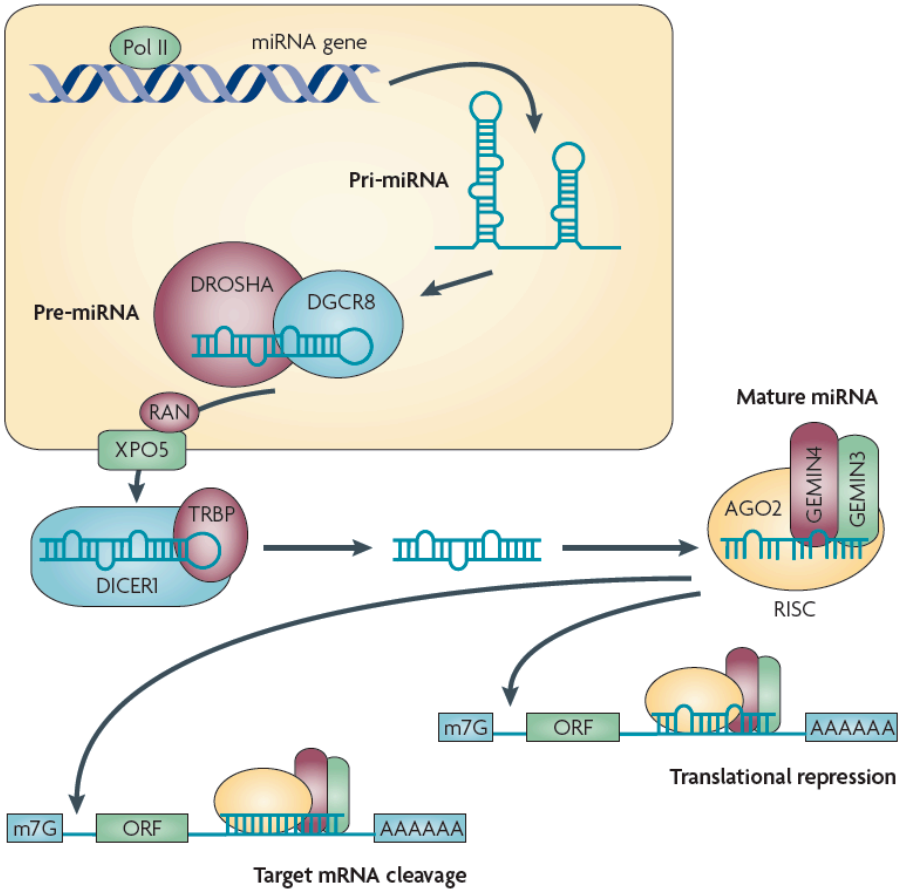
RNA polymerase II (Pol II) produces a 500–3,000 nucleotide transcript, called the primary microRNA **(pri-miRNA)**.

This is then cropped to form a **pre-miRNA** hairpin by a multi-protein complex that includes **DROSHA** (~60–100 nucleotides).

AA, poly A tail;
m7G, 7-methylguanosine cap;
ORF, open reading frame.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

Overview of the miRNA network



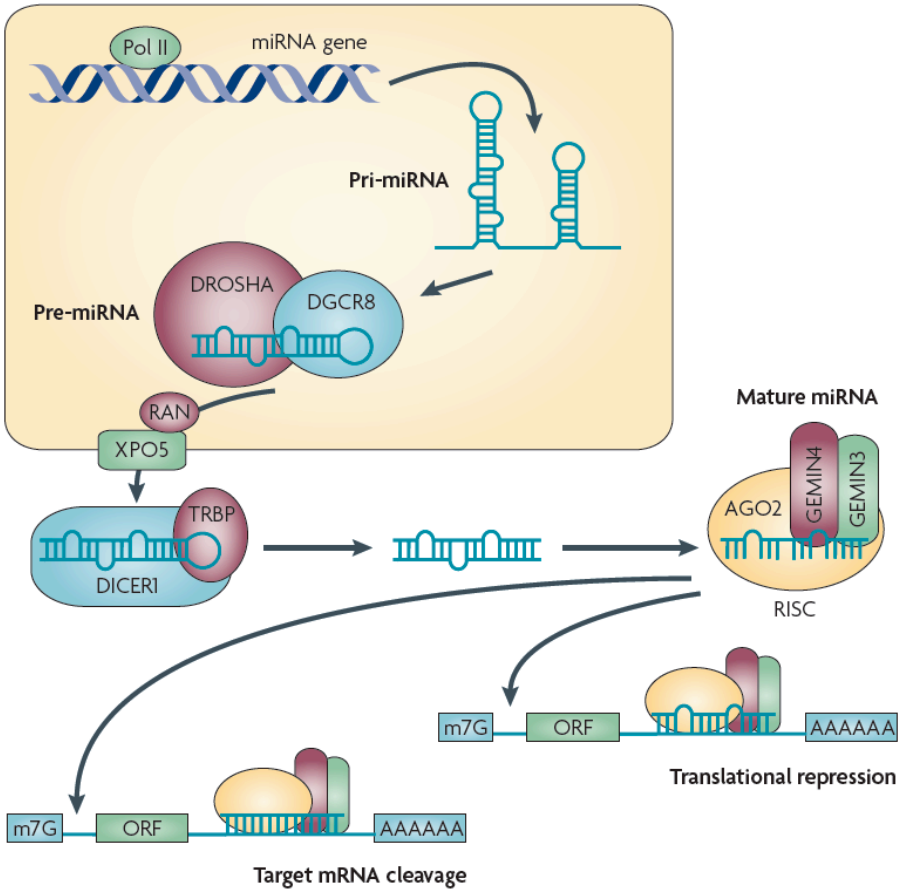
This double-stranded hairpin structure is **exported** from the nucleus by RAN GTPase and exportin 5 (XPO5).

Finally, the pre-miRNA is cleaved by **DICER1** to produce two miRNA strands, a mature miRNA sequence, approximately 20 nt in length, and its short-lived complementary sequence, which is denoted miR.

AA, poly A tail;
m7G, 7-methylguanosine cap;
ORF, open reading frame.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

Overview of the miRNA network



The thermodynamic stability of the miRNA duplex termini and the identity of the nucleotides in the 3′ overhang determines which of the strands is incorporated into the RNA-inducing silencing complex (**RISC**).

The single stranded miRNA is incorporated into RISC.

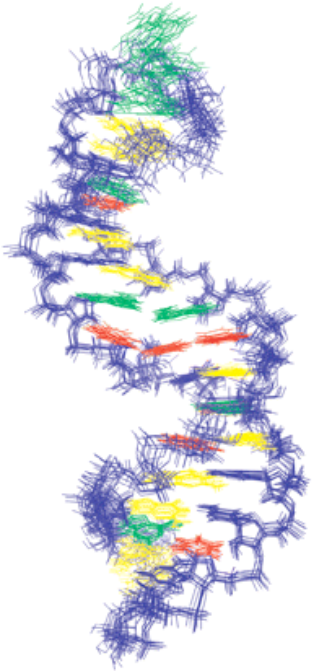
This complex then targets it e.g. to the target 3′ untranslated region of a mRNA sequence to facilitate repression and cleavage.

AA, poly A tail;
m7G, 7-methylguanosine cap;
ORF, open reading frame.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

miRNAs

Mature miRNA molecules are partially complementary to one or more mRNA molecules.



solution NMR-structure of *let-7* miRNA:*lin-41* mRNA complex from *C. elegans*
Cevec et al. *Nucl. Acids Res. (2008) 36: 2330.*

The main function of miRNAs is to down-regulate gene expression of their target mRNAs.

miRNAs typically have **incomplete base pairing** to a target and inhibit the translation of **many different mRNAs** with similar sequences.

In contrast, **siRNAs** typically **base-pair perfectly** and induce mRNA cleavage only in a single, specific target.

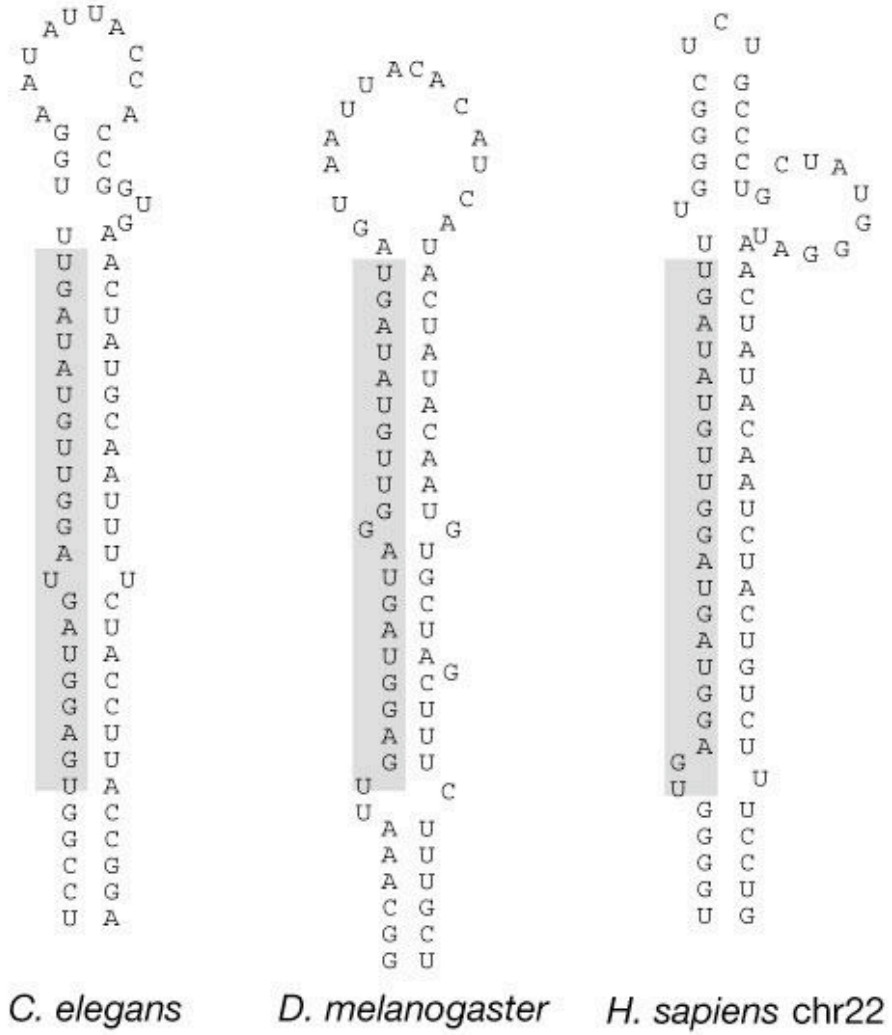
www.wikipedia.org

discovery of let7

The first two known microRNAs, lin-4 and let-7, were originally discovered in the nematode *C. elegans.*

They control the timing of stem-cell division and differentiation.
let-7 was subsequently found as the first known human miRNA.

let-7 and its family members are **highly conserved** across species in sequence and function.
Misregulation of let-7 leads to a less differentiated cellular state and the development of cell-based diseases such as cancer.



Pasquinelli et al. Nature (2000) 408, 86
www.wikipedia.org

Action of let7

Let-7 directly down-regulates the expression of the **oncogene *RAS*** in human cells.

All the three *RAS* genes in human, *K-, N-*, and *H-*,
have the predicted *let-7* binding sequences in their 3'UTRs.

In lung cancer patient samples, expression of *RAS* and *let-7* is anticorrelated.
Cancerous cells have low *let-7* and high *RAS*,
normal cells have high *let-7* and low *RAS*.

Another oncogene, *high mobility group A2* (***HMGA2***),
has also been identified as a target of *let-7*.

Let-7 directly inhibits *HMGA2* by binding to its 3'UTR.
Removal of the *let-7* binding site by 3'UTR deletion causes
overexpression of *HMGA2* and formation of tumor.

MYC is also considered as a oncogenic target of *let-7*.

www.wikipedia.org

miRNA discovery

miRNA discovery approaches, both biological and bioinformatics, have now yielded many thousands of miRNAs.

This process continues with new miRNA appearing daily in various databases.

miRNA sequences and annotations are compiled in the online repository **miRBase** (http://www.mirbase.org/).

Each entry in the database represents a predicted hairpin portion of a miRNA transcript with information on the **location** and **sequence** of the **mature miRNA sequence**

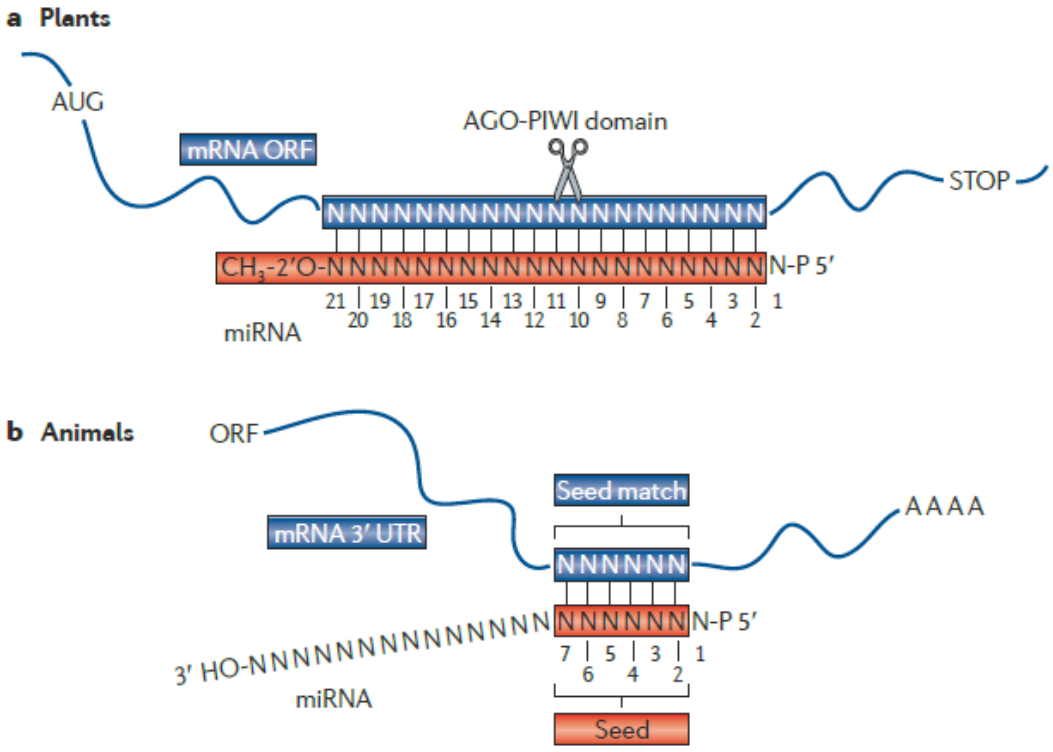
Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

miRNAs recognize targets by Watson-Crick base pairing

(a) **Plant miRNAs** recognize **fully** or **nearly complementary binding sites**.

(b) **Animal miRNAs** recognize **partially complementary** binding sites which are generally located in 3' UTRs of mRNA.

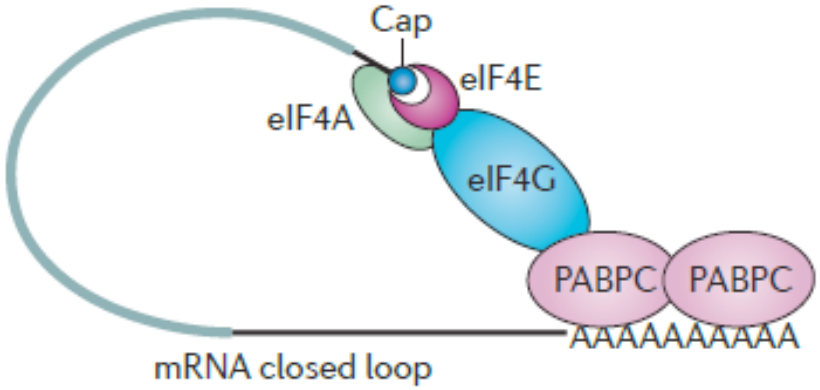
Complementarity to the 5' end of the miRNA – the "**seed**" sequence containing nucleotides 2-7 – is a major determinant in target recognition and is sufficient to trigger silencing.



Huntzinger, Izaurralde, Nat. Rev. Genet. 12, 99 (2011)

Mechanism of miRNA-mediated gene silencing

mRNAs are **competent for translation** if they possess a **5'cap structure** and a **3'-poly(A) tail**



mRNAs could, in principle, work by **translational repression** or by **target degradation**.

This has not been fully answered yet.

Current view: degradation of target mRNA dominates.

Huntzinger, Izaurralde, Nat. Rev. Genet. 12, 99 (2011)

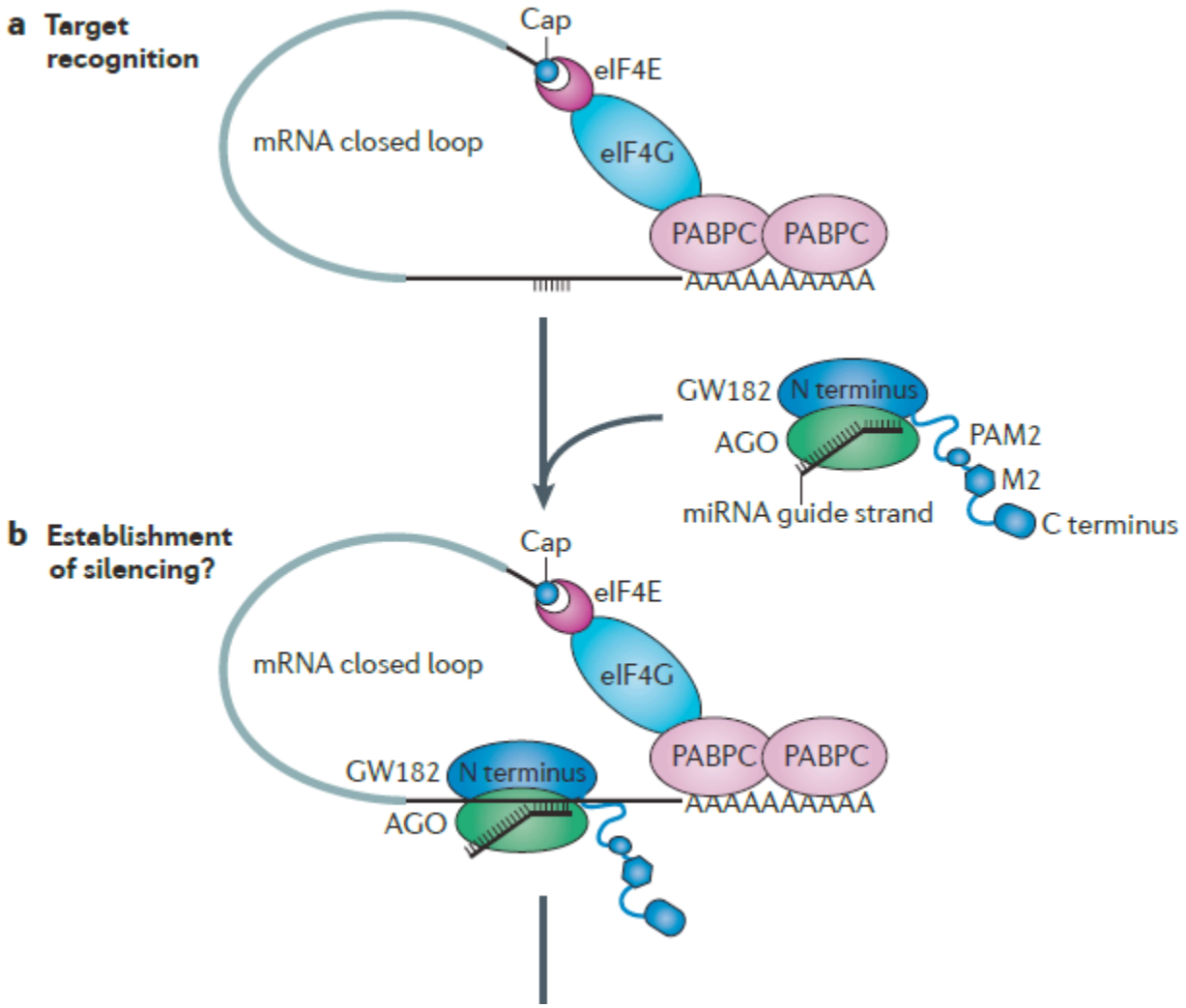
Mechanism of miRNA-mediated gene silencing

(a) The mRNA target is presented in a closed-loop conformation.

eIF: eukaryotic translation initiation factor

PABPC: poly(A)-binding protein

(b) Animal miRNAs bound to the argonaute protein AGO and to a GW182 protein recognize their mRNA targets by base-pairing to partially complementary binding sites.



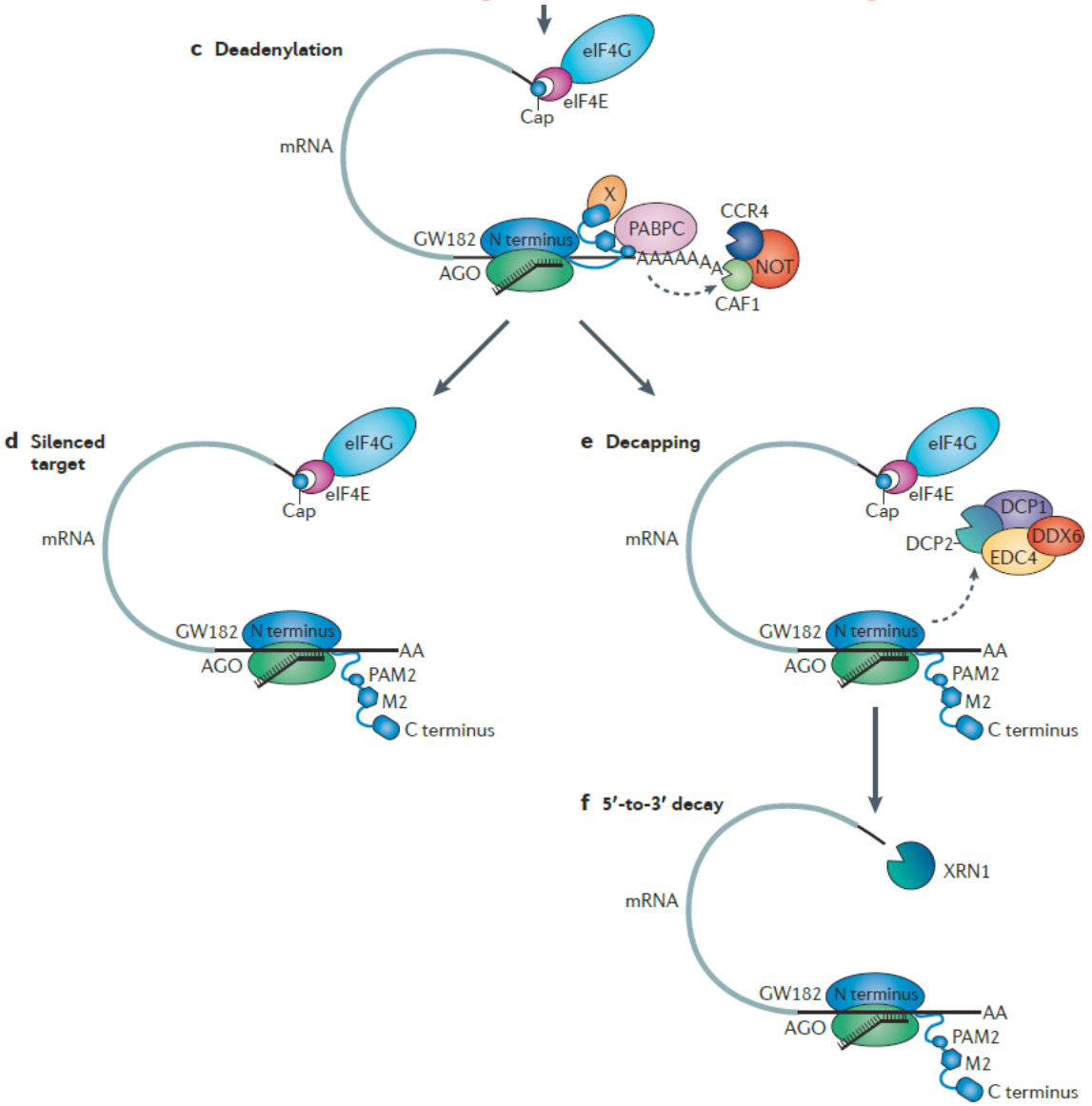
Huntzinger, Izaurralde, Nat. Rev. Genet. 12, 99 (2011)

Mechanism of miRNA-mediated gene silencing

(c)´The AGO-GW182 complex targets the mRNA to **deadenylation** by the deadenylation protein complex CCR4-CAF1-NOT.

(e) The mRNA is decapped by the protein DCP2 and then degraded (in f).

Alternatively (d), the deadenylated mRNA remains silenced.



Huntzinger, Izaurralde, Nat. Rev. Genet. 12, 99 (2011)

Bioinformatics prediction of miRNAs

With bioinformatics methods, putative miRNAs are first predicted in genome sequences based on the **structural features** of miRNA.

These algorithms essentially **identify hairpin structures** in **non-coding** and **non-repetitive** regions of the genome that are characteristic of miRNA precursor sequences.

The candidate miRNAs are then **filtered** by their **evolutionary conservation** in different species.

Known miRNA precursors play important roles in searching algorithms because structures of known miRNA are used to train the learning processes to discriminate between true predictions and false positives.

Many algorithms exist such as miRScan, miRSeeker, miRank, miRDeep, miRDeep2 and miRanalyzer.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

Recognition of miRNA targets

There seem to be two classes of binding patterns.

One class of miRNA target sites has **perfect Watson–Crick complementarity** to the 5'-end of the miRNAs, referred as '**seed region**', which includes positions 2–7 of miRNAs.

When bound in this way, miRNAs suppress their targets without requiring significant further base pairings at the 3'-end of the miRNAs.

The second class of target sites has **imperfect complementary base pairing** at the 5'-end of the miRNAs, but it is compensated via **additional base pairings** in the 3'-end of the miRNAs.

The multiple-to-multiple relations between miRNAs and mRNAs lead to complex miRNA regulatory mechanisms.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

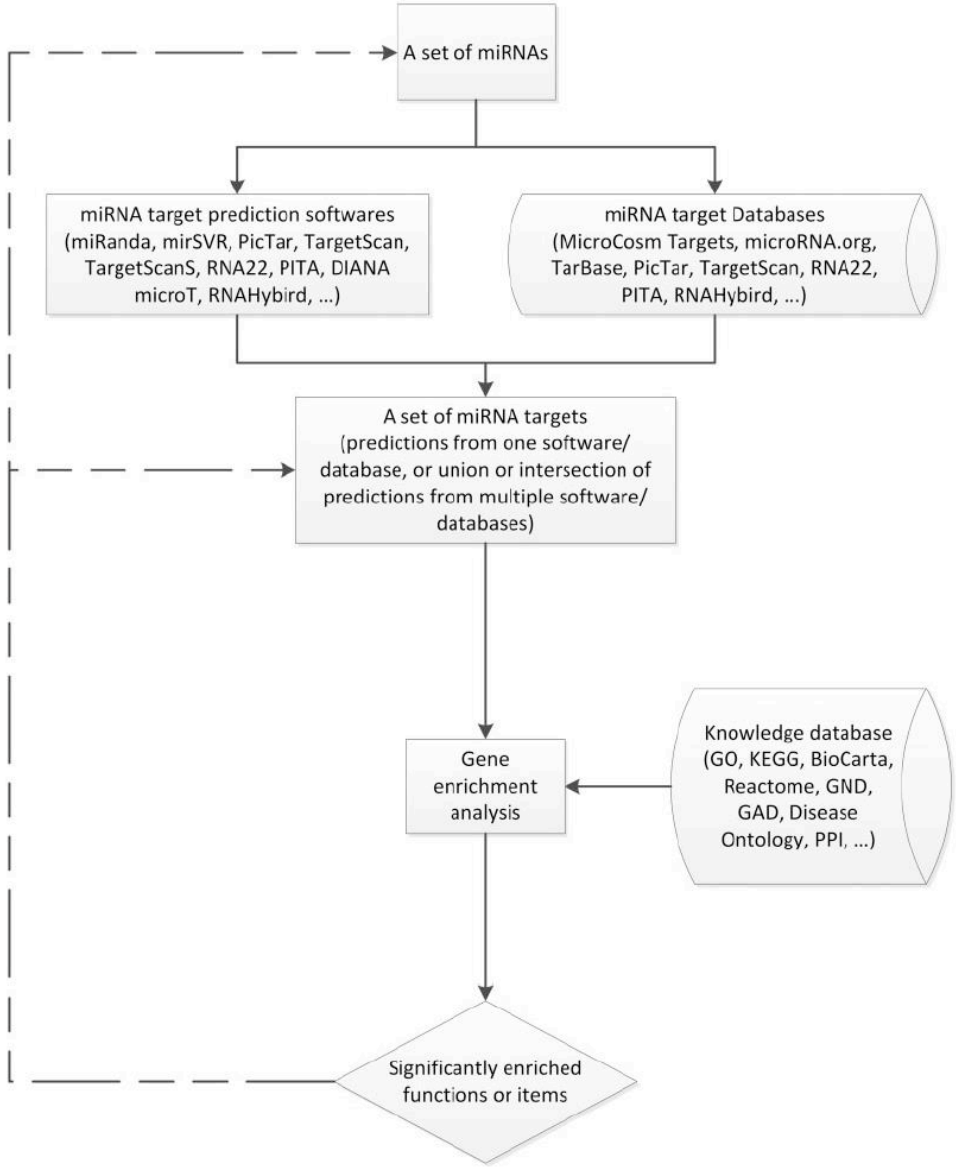
miRNA-target prediction algorithms

Table I: miRNA-target prediction algorithm

Algorithm	Regions scanned	Species conservation	Species	Brief description of the prediction method
miRanda	3′-UTR	Yes	Human, mouse, rat, fly and worm	Predict targets based on rules: (i) sequence complementarity, (ii) binding energy and (iii) evolutionary conservation.
mirSVR	No restriction	Yes	Human, mouse, rat, fly and worm	To score and rank miRanda-predicted miRNA-target sites with a supervised vector regression (SVR) model for features including secondary structure accessibility of the site and conservation.
PicTar	3′-UTR	Yes	Vertebrates, fly and worm	Filter alignments according to the thermodynamic stability, then score and rank the predicted target by hidden Markov model maximum-likelihood fit approach.
TargetScan	8mer and 7mer sites, and open reading frames	Yes	Human, mouse, rate, dog and chicken	Predict targets by searching for the presence of conserved 8mer and 7mer sites that match the seed region. Predictions are ranked by a combinatorial score based on site number, site type and site context.
TargetScanS	3′-UTR	Yes	Human, mouse, rate, dog and chicken	Predict targets that have a conserved 6 nt seed match flanked by either a m8 match or a tIA anchor.
RNA22	No restriction	No restriction	Any	Use the patterns discovered from the known mature miRNAs for predicting candidate miRNA-target sites in a sequence.
PITA	3′-UTR	Yes	Human, mouse, worm and fly	Predict miRNA targets using a non-parameter model that computes the difference between the free energy gained from the formation of the miRNA-target duplex and the energetic cost of unpairing the target to make it accessible to the miRNA.
RNAhybird	3′-UTR and coding sequence	No restriction	Any	A tool to identify mRNA secondary structure and energetically favourable hybridization between miRNA and target mRNA.
DIANA-microT	3′-UTR and CDS	No restriction	Human and mouse	The fifth version of microT algorithm which is specifically trained on a positive and negative set of miRNA recognition elements located in both the 3′-UTR and CDS region. The conserved and non-conserved miRNA recognition elements are combined into a final prediction score.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

Predicting miRNA function based on target genes

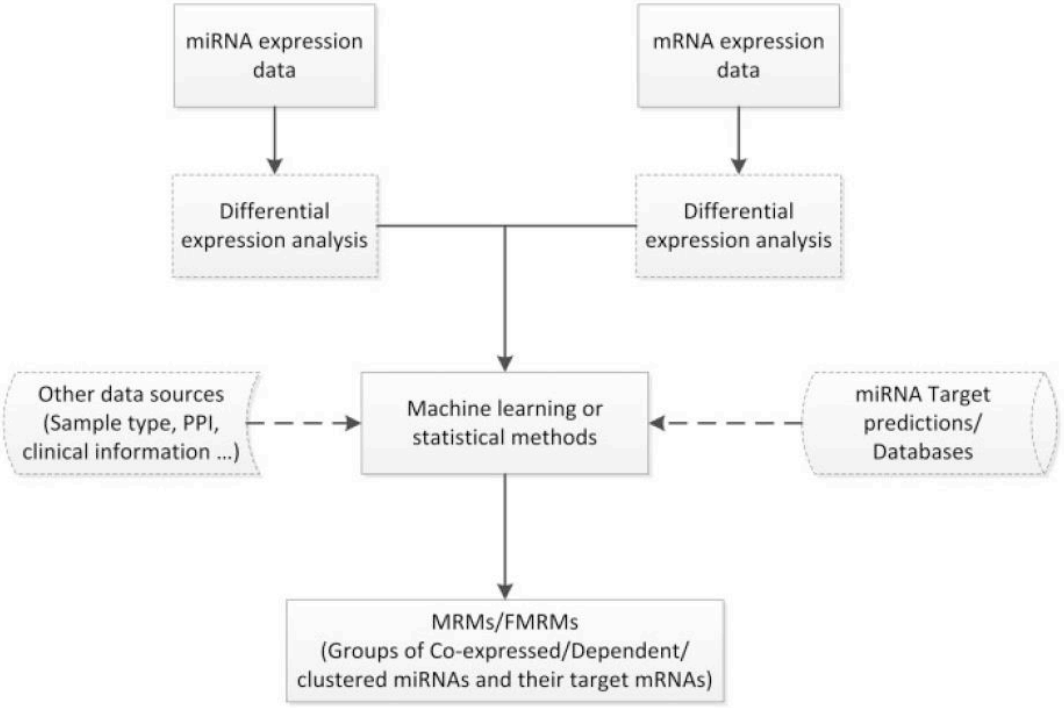


The most straight-forward approach for miRNA functional annotation is through **functional enrichment analysis** using the miRNA-target genes.

This approach assumes that miRNAs have similar functions as their target genes.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

Predicting miRNA function based on correlated expression



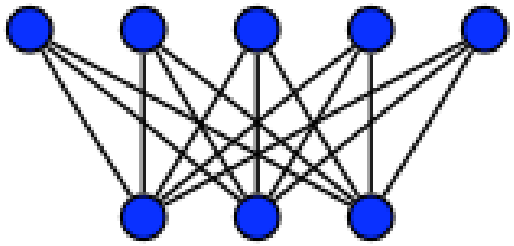
miRNA functional annotation heavily relies on the miRNA-target prediction.

In the last few years, many studies have been conducted to infer the miRNA regulatory mechanisms by incorporating target prediction with other genomics data, such as the expression profiles of miRNAs and mRNAs.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

Discovering MRMs

A MRM (group of co-expressed miRNAs and mRNAs) may be defined as a special bipartite graph, named **biclique**, where two sets of nodes are connected by edges.



Every node of the first set representing miRNA is connected to every node of the second set representing mRNAs.

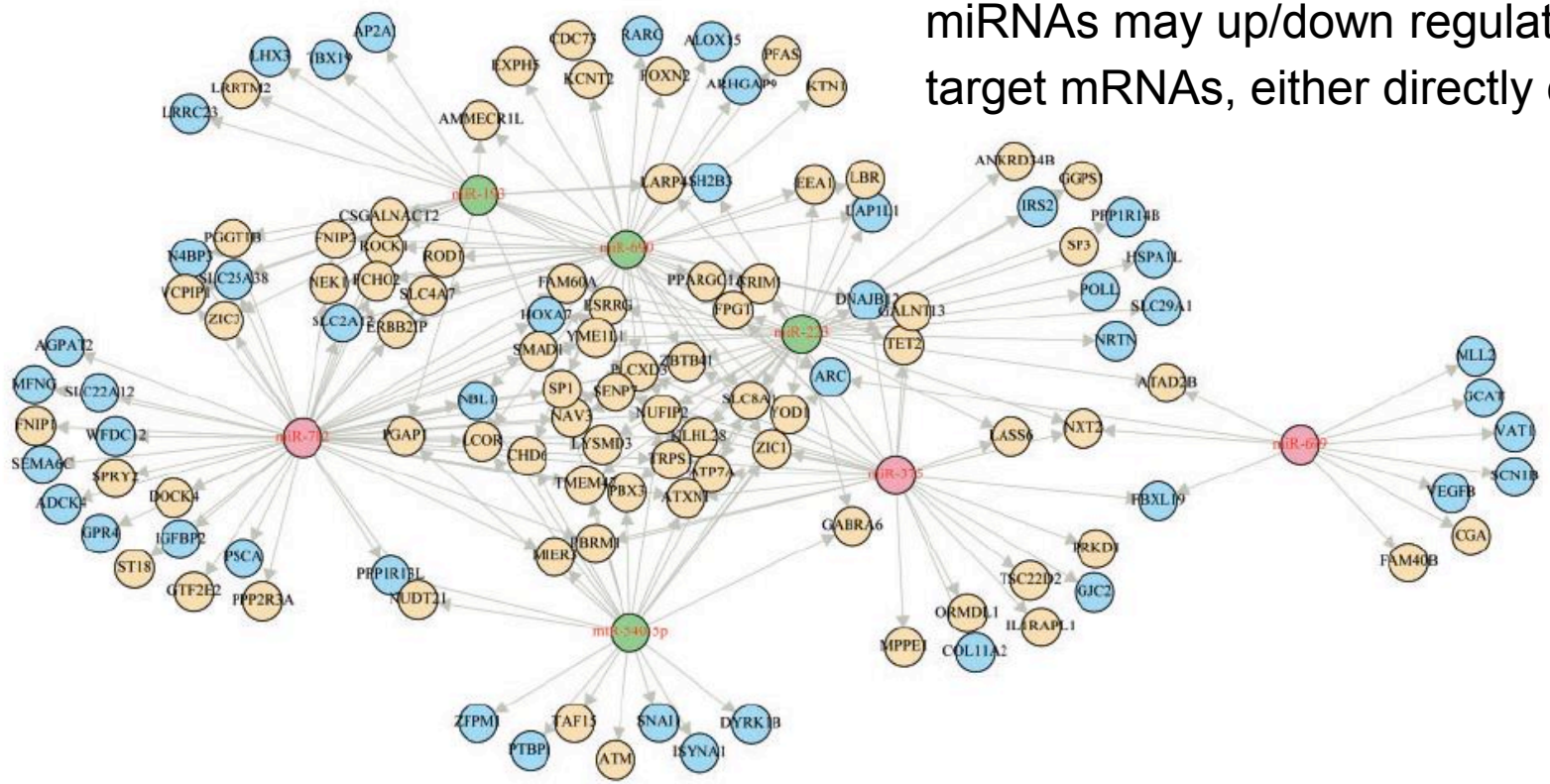
The weights of edges correspond to the miRNA–mRNA binding strength inferred from target prediction algorithms

Most of the integrative methods of MRM discovery are based on the **assumption** that miRNA **negatively regulate** their target mRNAs so that the expression of a specific miRNA and its targets should be anti-correlated.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

miRNA-mRNA network

A FMRM identified from analysis of **schizophrenia** patients. It shows that miRNAs may up/down regulate their target mRNAs, either directly or indirectly.



Up-regulated miRNAs are coloured in red and down-regulated miRNAs are coloured in green. Up-regulated mRNAs are coloured in yellow, while down-regulated mRNAs are coloured in blue.

Liu et al. Brief Bioinf. (2012) doi: 10.1093/bib/bbs075

SNPs in miRNA may lead to diseases

miRNAs can have dual oncogenic and tumour suppressive roles in cancer depending on the cell type and pattern of gene expression.

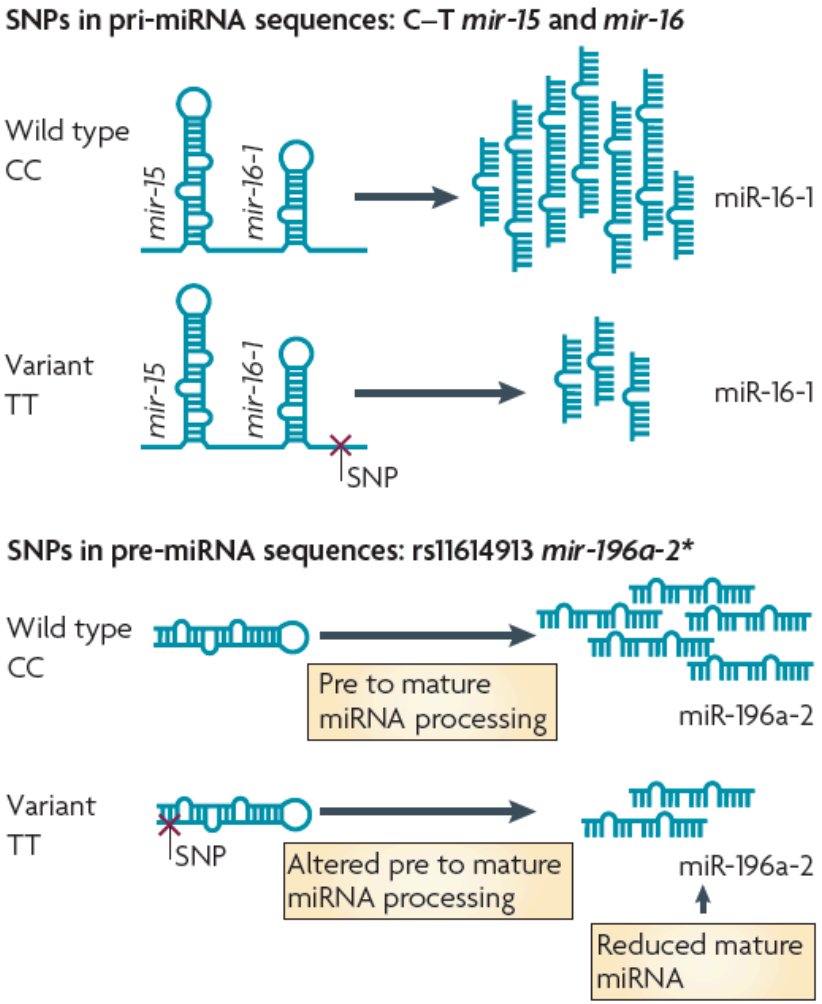
Approximately **50%** of all annotated human miRNA genes are located in **fragile sites** or areas of the genome that are associated with cancer.

E.g. Abelson *et al.* found that a mutation in the miR-189 binding site of *SLITRK1* was associated with Tourette's syndrome.

SNPs in miRNA genes are thought to affect function in one of three ways:
(1) through the transcription of the primary transcript;
(2) through pri-miRNA and pre-miRNA processing; and
(3) through effects on miRNA–mRNA interactions

Volinia et al. PNAS (2013) 110, 7413

SNPs in pri-miRNA and pre-miRNA sequences

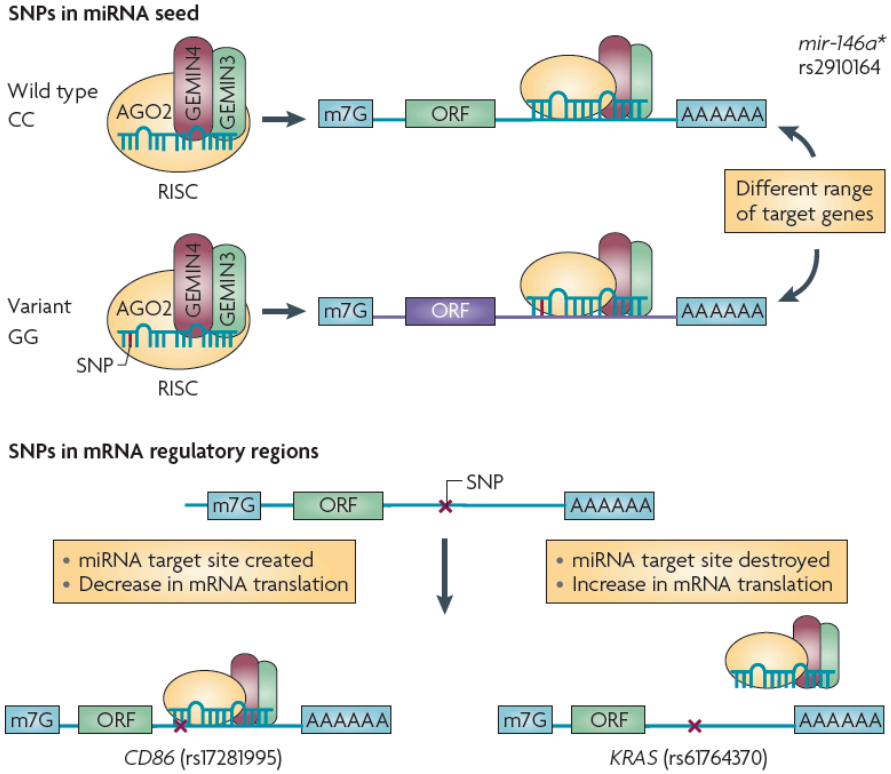


SNPs can occur in the pri-miRNA and pre-miRNA strands and are likely to affect miRNA processing and subsequent mature miRNA levels.

Such SNPs can lead to either an increase or decrease in processing.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

SNPs in miRNA seed and regulatory regions



SNPs in mature microRNAs (miRNAs) within the seed sequence can strengthen or reduce binding between the miRNA and its mRNA target.

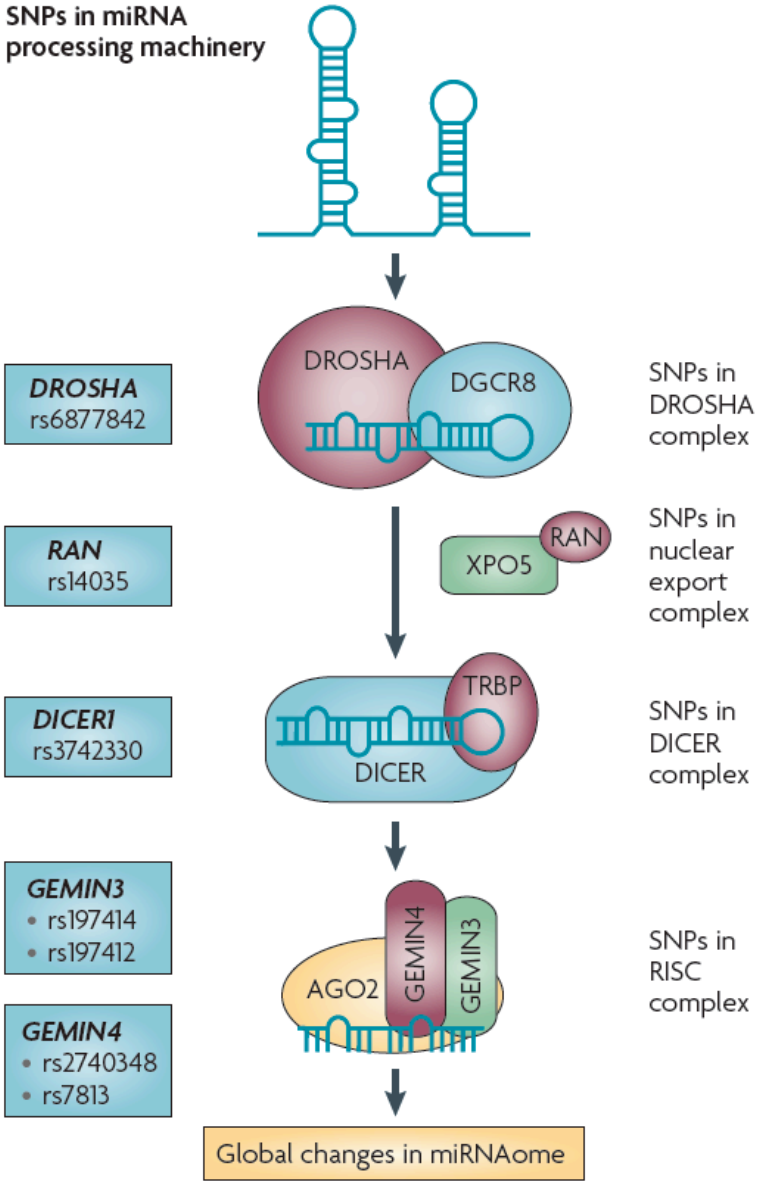
Moreover, such SNPs can create or destroy target binding sites, as is the case for mir-146a*.

SNPs located within the 3′ untranslated region of miRNA binding sites function analogously to seed region SNPs and modulate the miRNA–mRNA interaction.

They can create or destroy miRNA binding sites and affect subsequent mRNA translation.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

SnPs in miRNA processing machinery



SNPs can also occur within the processing machinery.

These SNPs are likely to affect the microRNAome (miRNAome) as a whole, possibly leading to the overall suppression of miRNA output.

In addition, SNPs in cofactors of miRNA processing, such as p53, may indirectly affect miRNA maturation.

Ryan et al. Nature Rev. Cancer (2010) 10, 389

microRNAs as biomarkers for cancer

miRNAs can be used for sensitive classification of cancer risks or cancer progression (e.g. 95%), see research in the Keller and Lenhof groups.

Various companies market such tools.

www.exiqon.com

Briefings in Bioinformatics Advance Access published December 4, 2013
BRIEFINGS IN BIOINFORMATICS. page 1 of 14 doi:10.1093/bib/bbt085

Transcription factor and microRNA co-regulatory loops: important regulatory motifs in biological processes and diseases

Hong-Mei Zhang, Shuzhen Kuang, Xushen Xiong, Tianliuyun Gao, Chenglin Liu and An-Yuan Guo

Key Points

- TFs and miRNAs can jointly regulate gene expression in the forms of FFLs and FBLs, which influence many aspects of normal cells and diseases.
- FFLs and FBLs can be classified into different types based on the master regulator or the regulation effects of two paths on target. Different types of loops have different mechanisms in gene regulation.
- The identification of TF and miRNA targets is a key step for detecting FFLs and FBLs. It is better to combine the experimentally verified targets with predicted targets by different methods.
- FFLs and FBLs are popular regulatory models and critical for biological processes and diseases. FFL has a specific function in noise buffering effect. It can minimize the cell response to stochastic signaling noise and maintain steady-state levels of targets. FBL can act as a toggle switch between two different fates in cell differentiation.

FFL: feed-forward loop (see lecture V8)

FBL: feedback loop

Volinia et al. PNAS (2013) 110, 7413

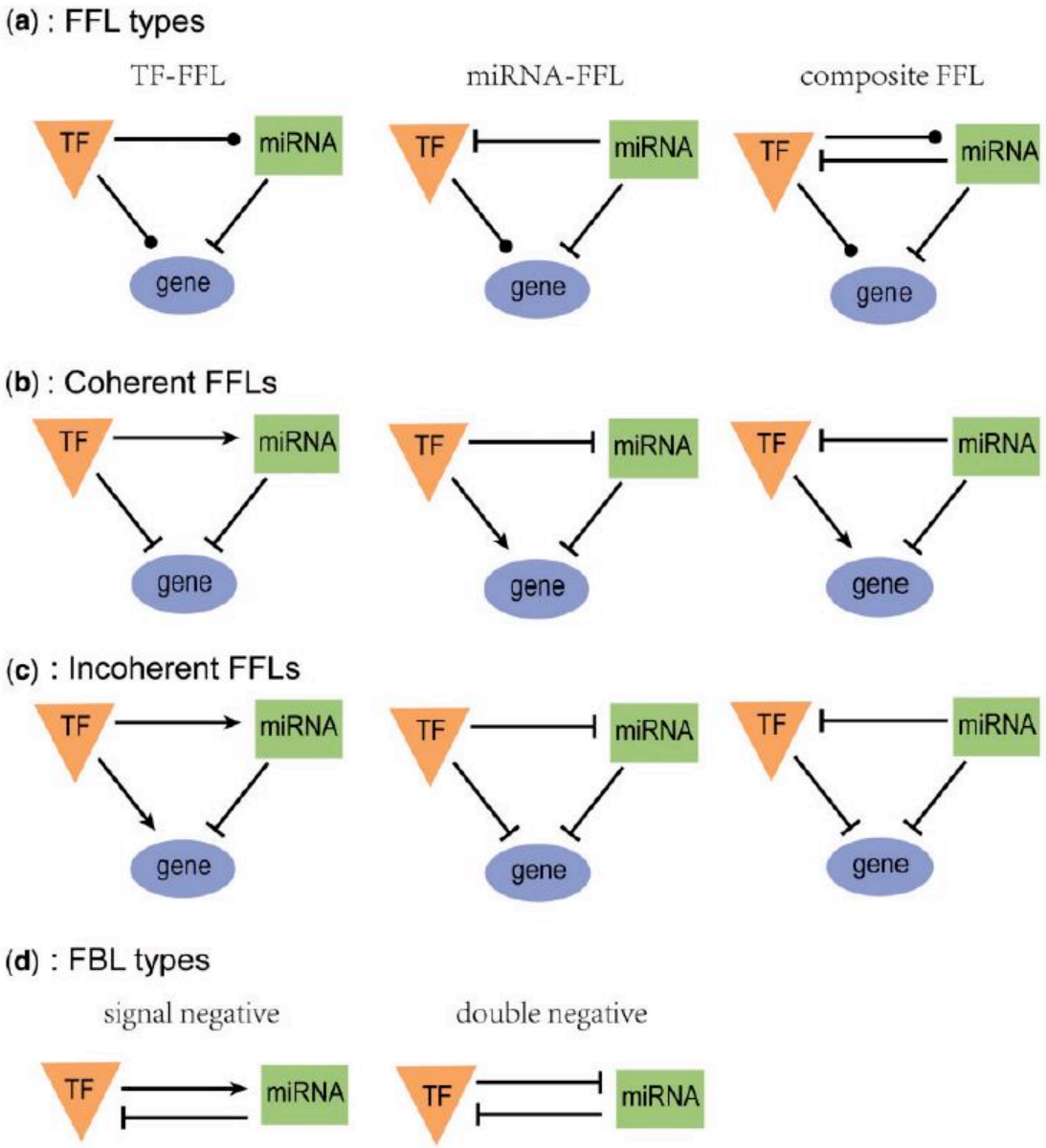
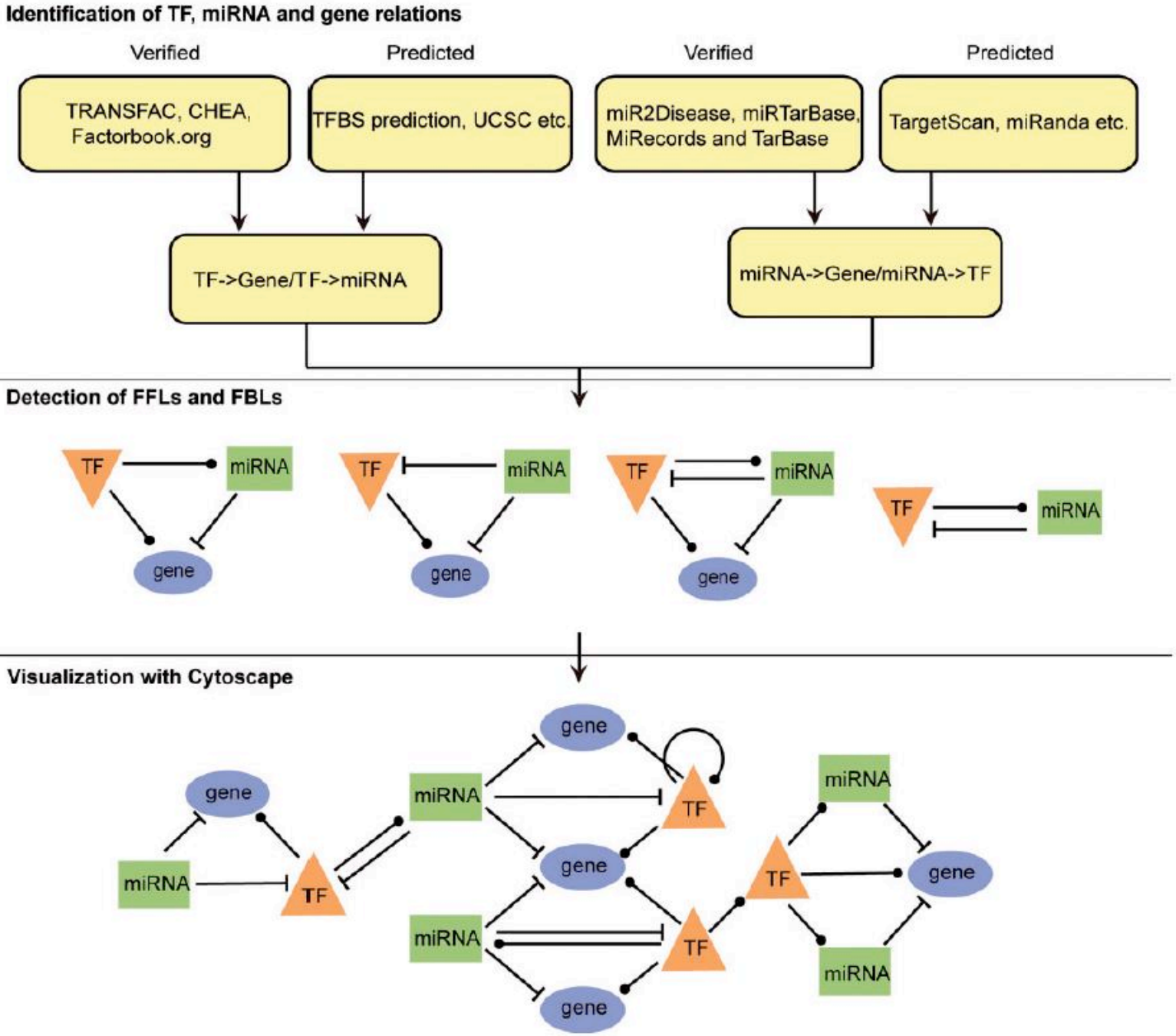


Figure 1: FFL and FBL types. (**a**) Three types of FFLs classified by the master regulator. Blunt arrows with dot end represent transcriptional activation or repression. (**b**) Coherent FFLs. In this kind of FFLs, two paths that regulate target gene have the same effects (either activation or repression). (**c**) Incoherent FFLs. The target gene is regulated by two opposite paths. (**d**) FBL types. Nodes: triangles are TFs; rectangles are miRNAs; ovals are genes; Edges: sharp arrow means activation; T-shaped arrow represents repression.


Identification of TF, miRNA and gene relations
Verified
Predicted
Verified
Predicted
TRANSFAC, CHEA, Factorbook.org
TFBS prediction, UCSC etc.
miR2Disease, miRTarBase, MiRecords and TarBase
TargetScan, miRanda etc.
TF->Gene/TF->miRNA
miRNA->Gene/miRNA->TF
Detection of FFLs and FBLs
TF
miRNA
gene
Visualization with Cytoscape

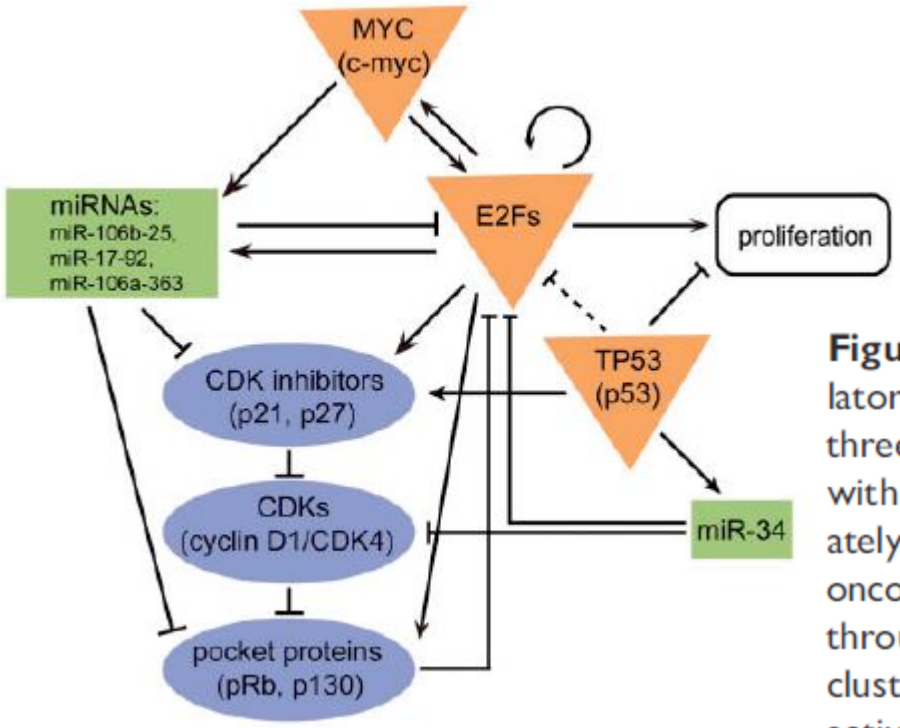


Figure 3: A schematic model for TF-miRNA co-regulatory network in cell proliferation. The E2F family and three miRNA clusters form several composite FFLs with CDK inhibitors and pocket proteins. They corporately control the progression of the cell cycle. The oncogene c-Myc can promote cell cycle progress through directly activating the E2F family and miRNA clusters, while the tumor repressor p53 represses E2Fs activity in an indirect way. The meanings of sharp arrows and T-shaped solid arrows are same as Figure I. T-shaped dotted arrow indicates the indirect repression of P53 to E2Fs. This figure is drawn based on two previous articles [55,56].

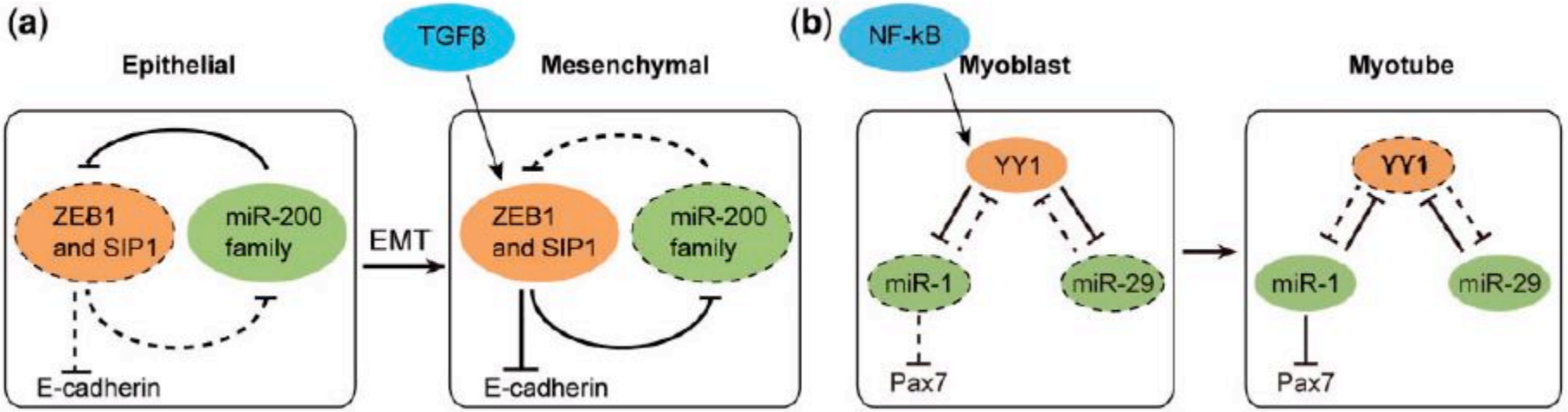


Figure 4: FFLs and FBLs in cell differentiation. Orange ovals are TFs; green ovals are miRNAs; light blue ovals are upstream signals. Dotted line means the activation or repression is inactive; dotted oval means the gene or miRNA is repressed or in a low expression. (**a**) The FBL between TFs ZEBI/SIPI and miR-200 family in EMT. In epithelial cells, ZEBI and SIPI are repressed by miR-200 family. EMT is induced when ZEBI and SIPI are activated by the TGFβ signal and miR-200 family is repressed. (**b**) The FBLs in skeletal myogenesis. The high expression of TF YYI activated by NF-κB signal maintains the undifferentiated states of myoblast cells. At the onset of myogenesis, the down-regulation of the NF-κB-YYI pathway leads to an upregulation of miR-I and miR-29, which ensures myoblast cells properly differentiate into myotubes.

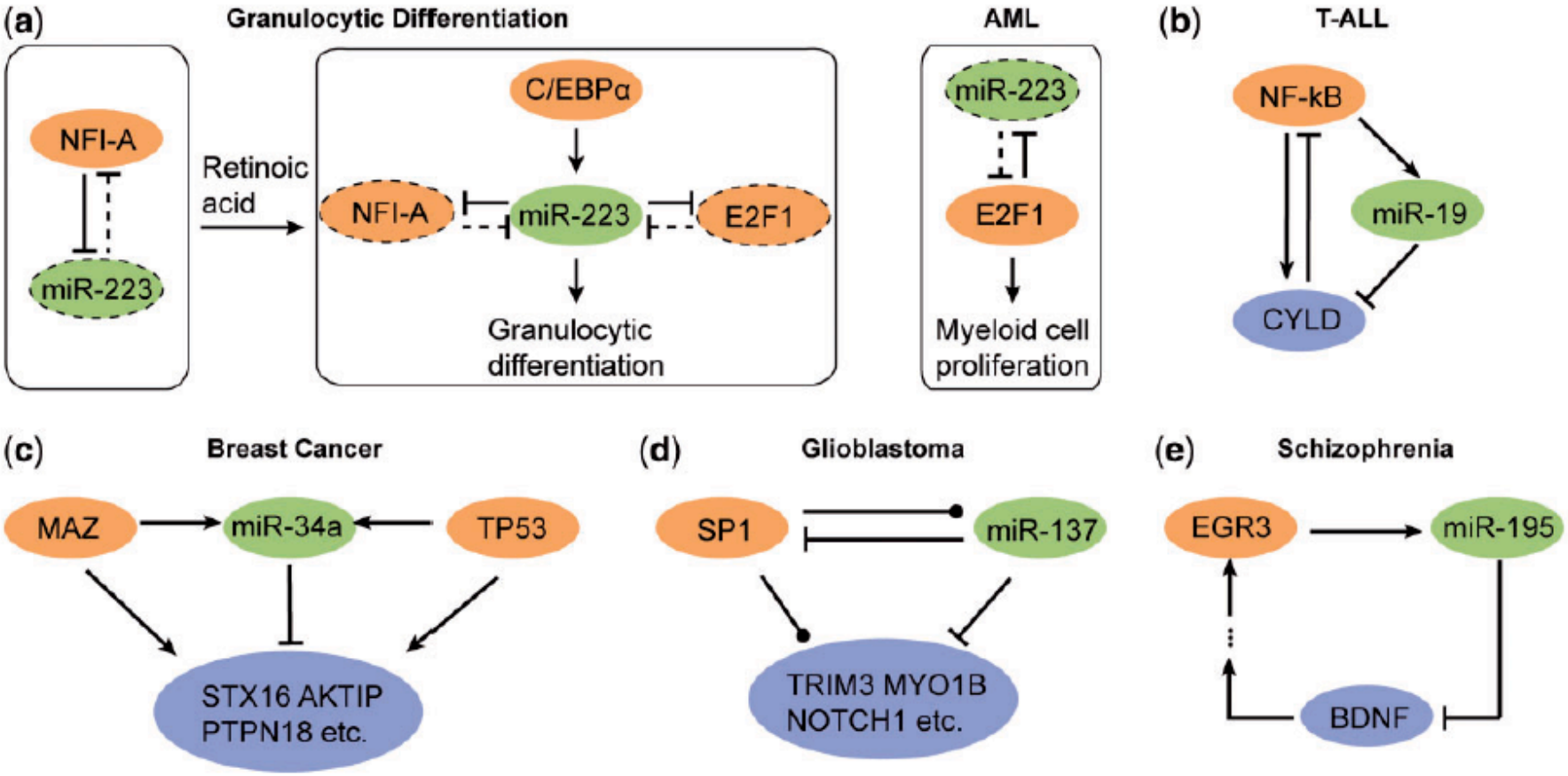


Figure 5: FFLs and FBLs in diseases. (**a**) The FBL in granulocytic differentiation and myeloid cell proliferation. In the undifferentiated cells, TF NFI-A maintains the miR-223 at low level. The TF C/EBPα is activated by retinoic acid and upregulates miR-223 expression, which in turn represses TFs NFI-A and E2FI, resulting in inhibition of cell cycle and advance of granulocytic differentiation (left). C/EBPα is deregulated in AML and overexpressed E2FI inhibits miR-223 transcription, thus promoting myeloid cell proliferation and blocking granulocytic differentiation (right). (**b**) A FFL in T-ALL. (**c**) The predicted FFLs in breast cancer. (**d**) A predicted FFL in glioblastoma. (**e**) A FFL in schizophrenia. TF EGR3 activates the transcription of miR-195, and in turn miR-195 indirectly reduces the expression of EGR3 by repressing gene BDNF.

Summary

The discovery of microRNAs has led to an additional layer of complexity in understanding cellular networks.

Prediction of miRNA-mRNA networks is challenging due to the often non-perfect base matching of miRNAs to their targets.

Individual SNPs may alter network properties, and may be associated with cancerogenesis.

miRNAs can be exploited as sensitive biomarkers.

miRNAs are becoming important elements of GRNs
-> new hierarchical layer, novel types of network motifs …

Bioinformaticians do not run out of work ☺

Volinia et al. PNAS (2013) 110, 7413