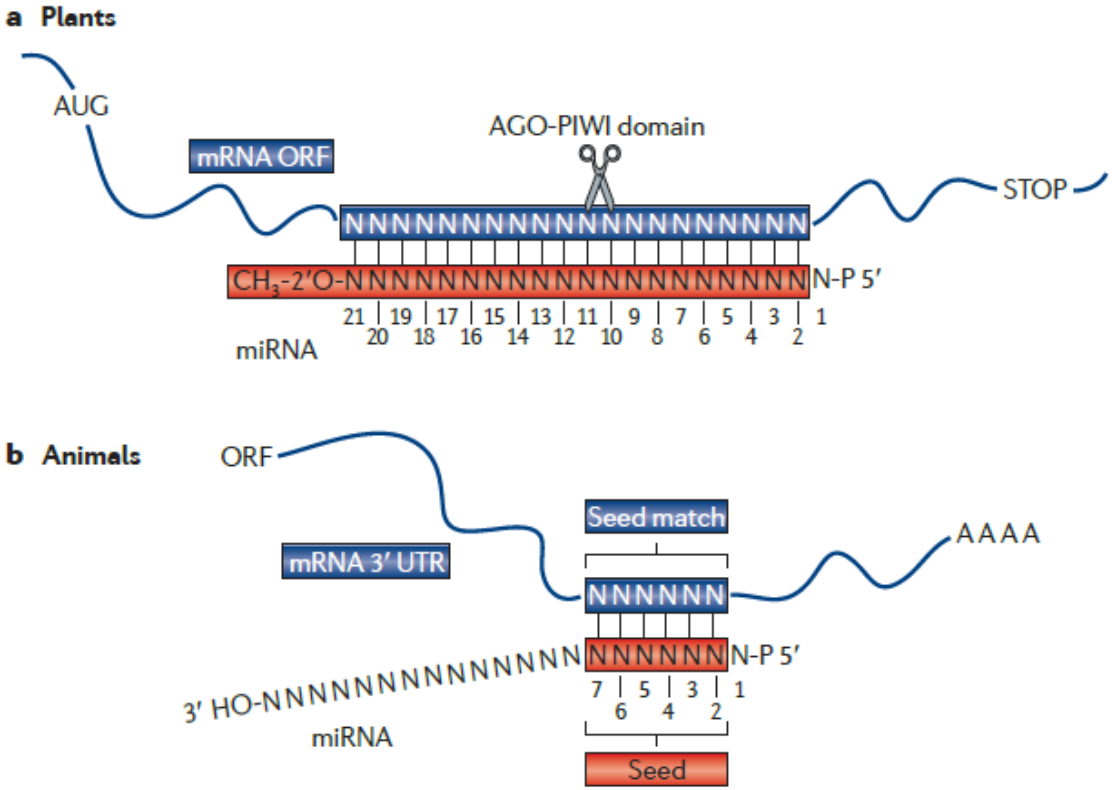
miRNAs recognize targets by Watson-Crick base pairing

(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.



$4^6 = (2^2)^6 = 2^{12} = 4096$ k-mers of length 6

On average, the 3'-UTR in humans is ca. 800 nt long (www.wikipedia.org)

20.000 genes x 800 nt / 4096 6-mers = 4000 binding sites for 1 miRNA 6-mer

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 to as '**seed region**', which includes positions 2–7 of miRNAs.

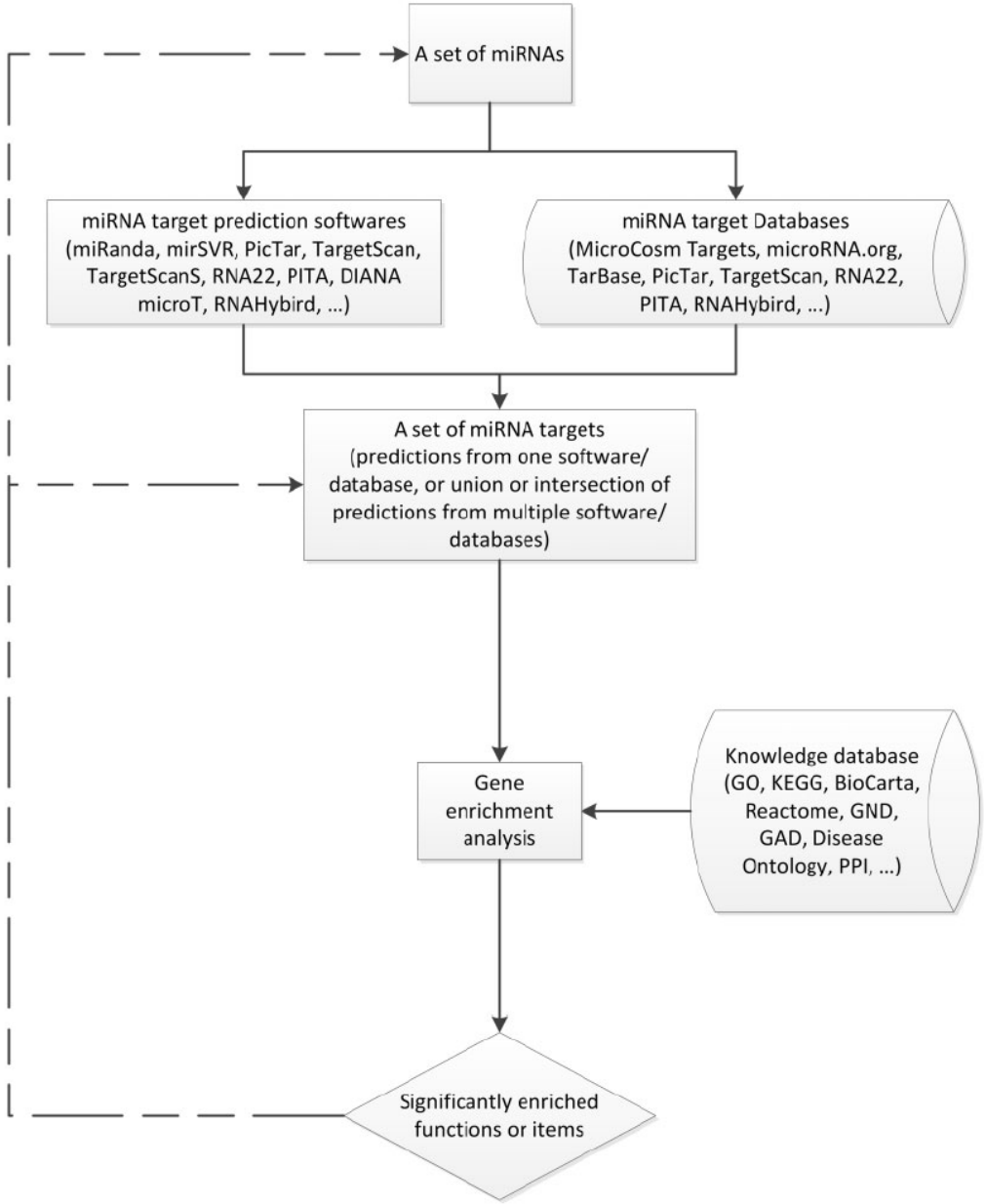
When bound in this way, miRNAs suppress their targets without requiring significant further base pairings at the 3'-end of the miRNAs.

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

Predicting miRNA function based on target genes

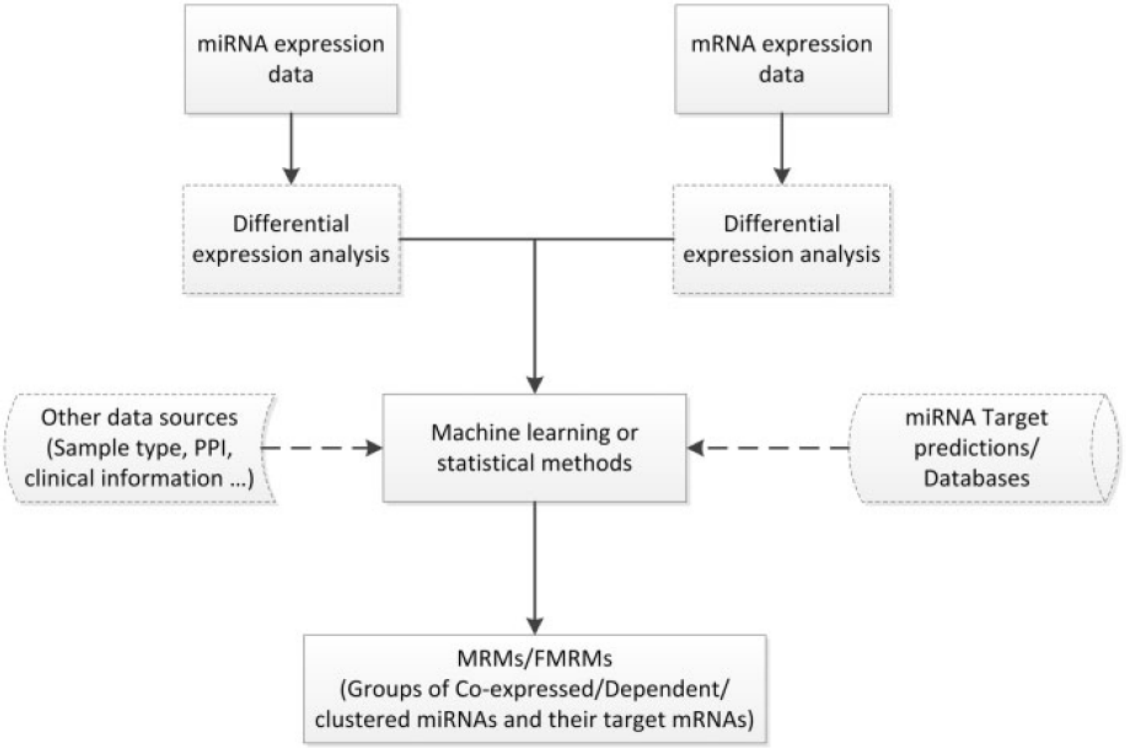


The most straight-forward approach for miRNA functional annotation is through **functional enrichment analysis** using the miRNA-target genes.

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

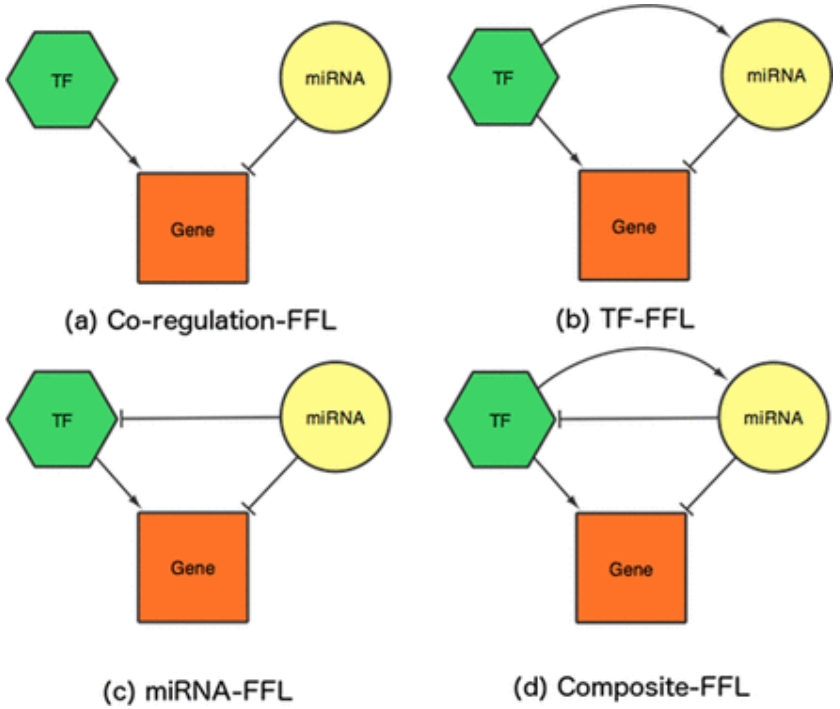


miRNA functional annotation heavily relies on the miRNA-target prediction.

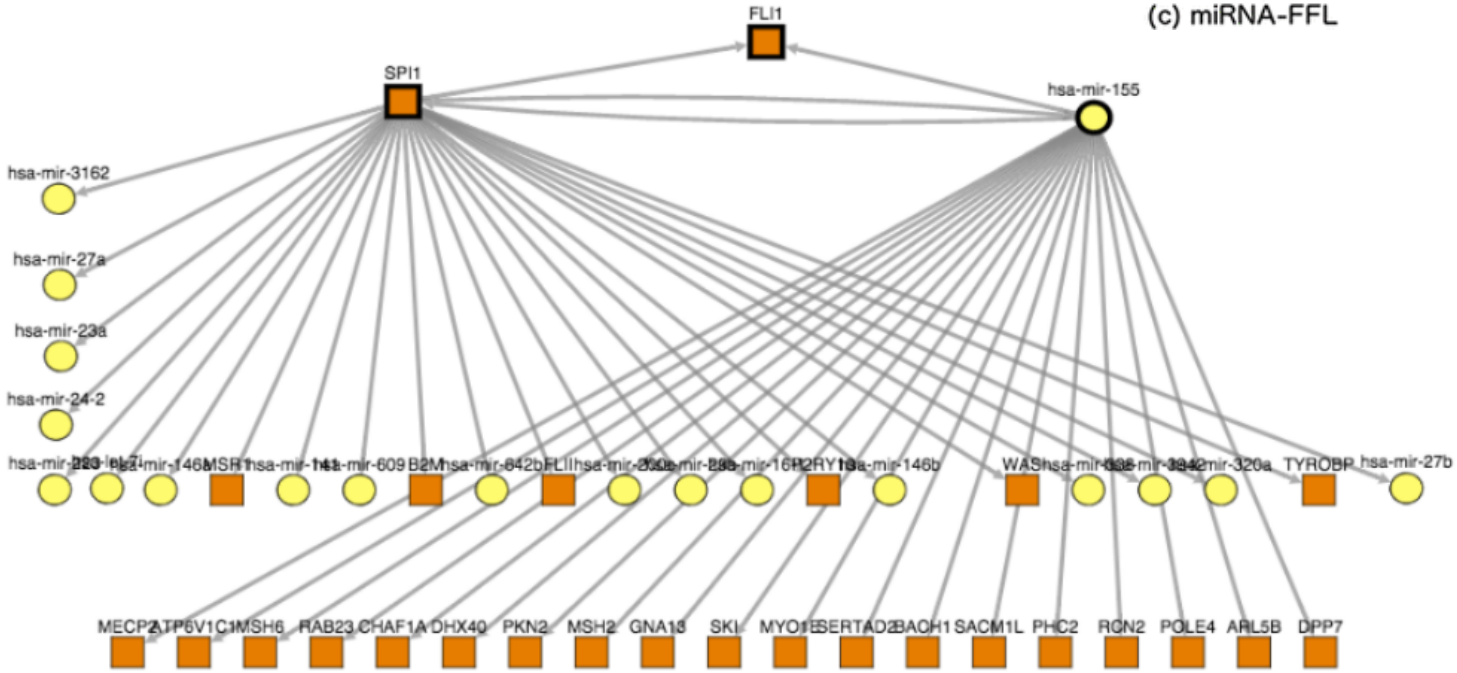
In the last few years, many studies have been conducted to infer the miRNA regulatory mechanisms by incorporating target prediction with other genomics data, such as the expression profiles of miRNAs and mRNAs.

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

TFmiR



Co-regulated subnetwork for TF: SPI1, miRNA: hsa-mir-155, Gene: FLI1



Hamed et al. Nucl Ac Res 43:W283-W288 (2015)

Significance of FFL motifs

Compare how often FFL motifs appear in the real network to the number of times they appear in randomized ensembles preserving the same node degrees.

Use degree preserving randomization algorithm.

For 2 × *L* steps, two edges *e*1 = (v1, v2) and *e*2 = (v3, v4) are randomly chosen from the network and rewired such that the start and end nodes are swapped, i.e. *e*3 = (v1, v4) and *e*4 = (v3, v2) if $\{e3, e4\} \in V$.

Construct 100 random networks. Compare motif frequencies to the real network. The *P-value* is calculated as

$$P\text{-value} = \frac{N_h}{N_r}$$

where N_h is the number of random times that a certain motif type is acquired more than or equal to its number in the real network, and N_r is 100.