miRTar.Human

miRTarBase

RNALogo

miRStart
2010 bps

miRNAMap

ViTa
Virus' miRNA Target

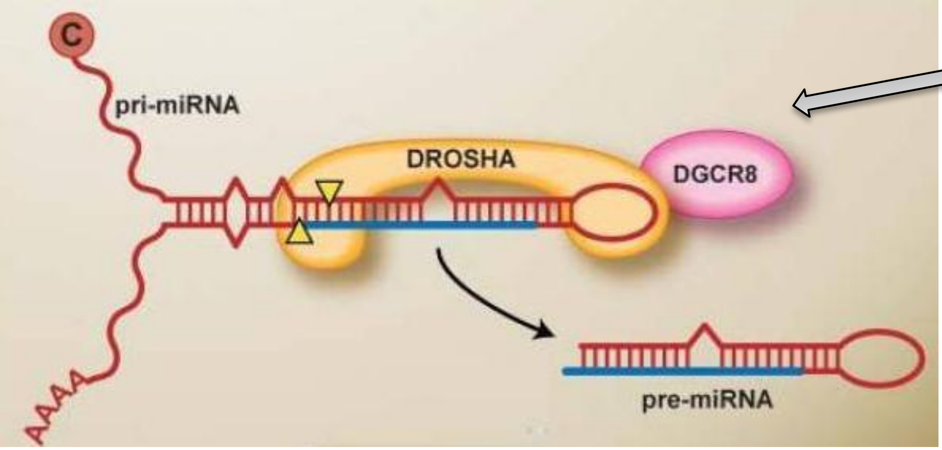
miRExpress

Easy to Learn Bioinformatics Analysis for MicroRNA Studies

簡單學 microRNA生物資訊分析

Dr. Hsien-Da Huang (黃憲達)
Professor and Chairman,
Institute of Bioinformatics and Systems Biology,
Department of Biological Science and Technology,
National Chiao Tung University, Taiwan

MicroRNA Biogenesis

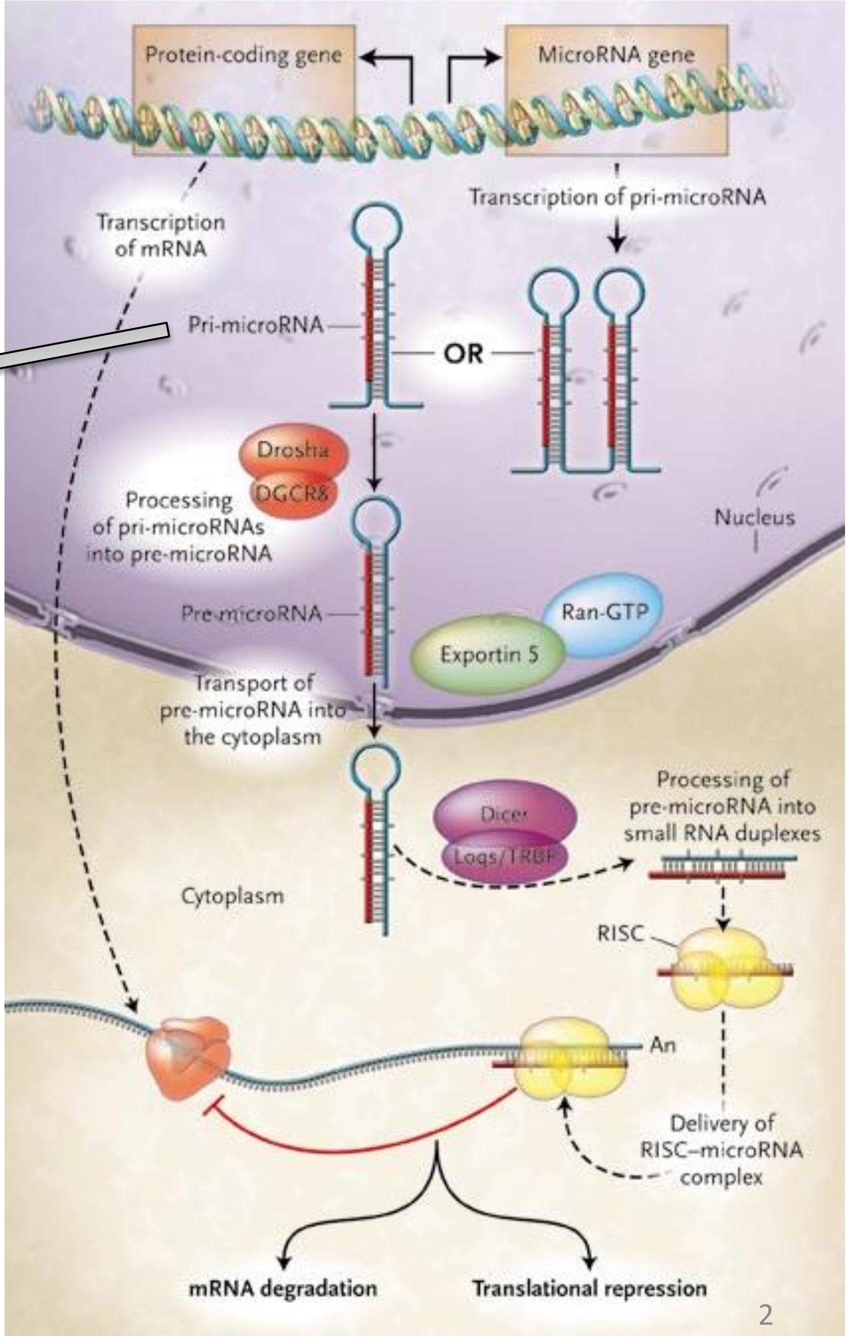


Key Steps in miRNA Biogenesis (Cullen BR., 2004)

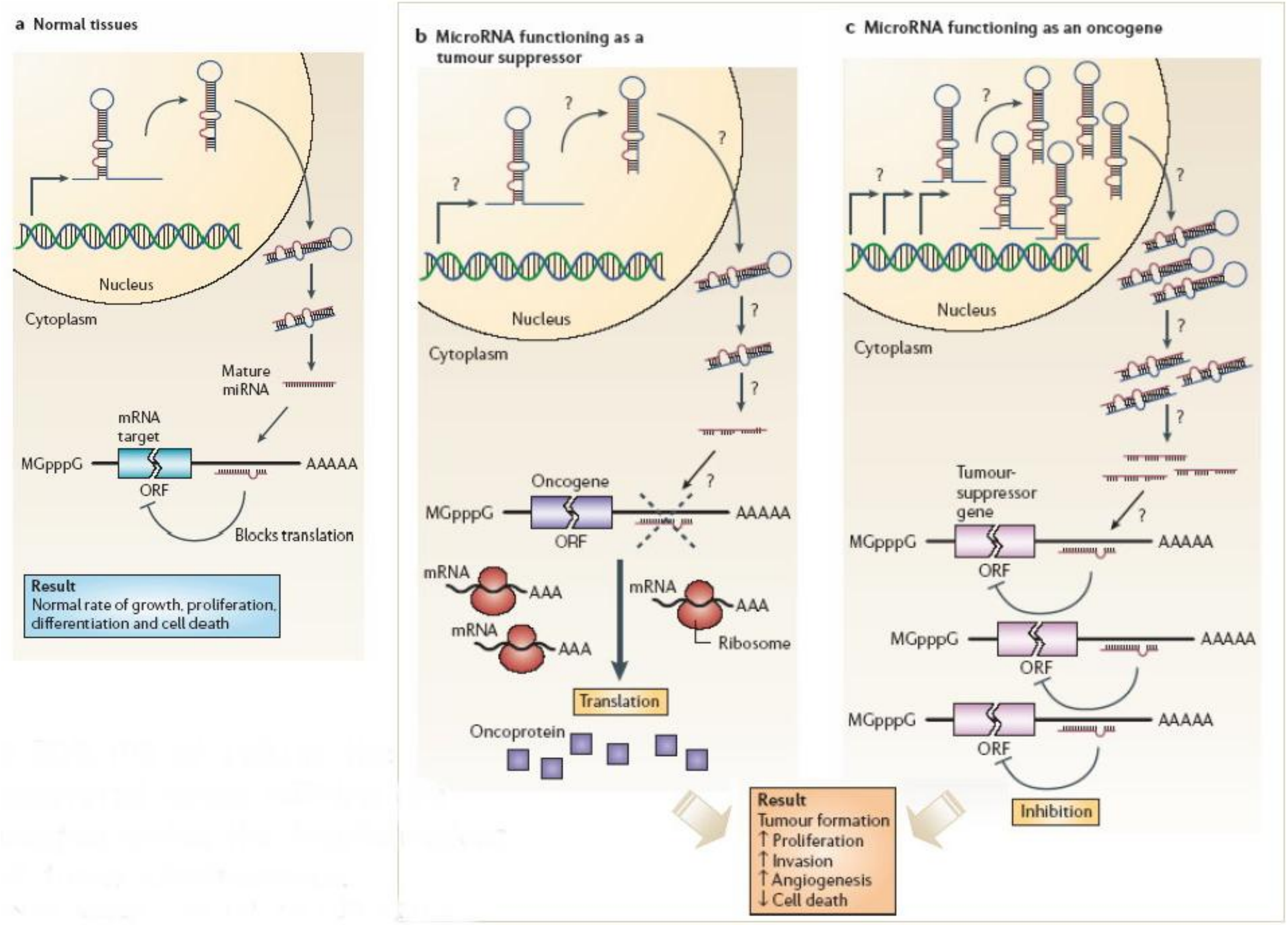
MicroRNAs (miRNAs) are important genetic regulators of development, differentiation, growth, and metabolism

The mammalian genome encodes ≈500 known miRNA genes

Approximately 50% are expressed from non-protein-coding transcripts, whereas the rest are located mostly in the introns of coding genes



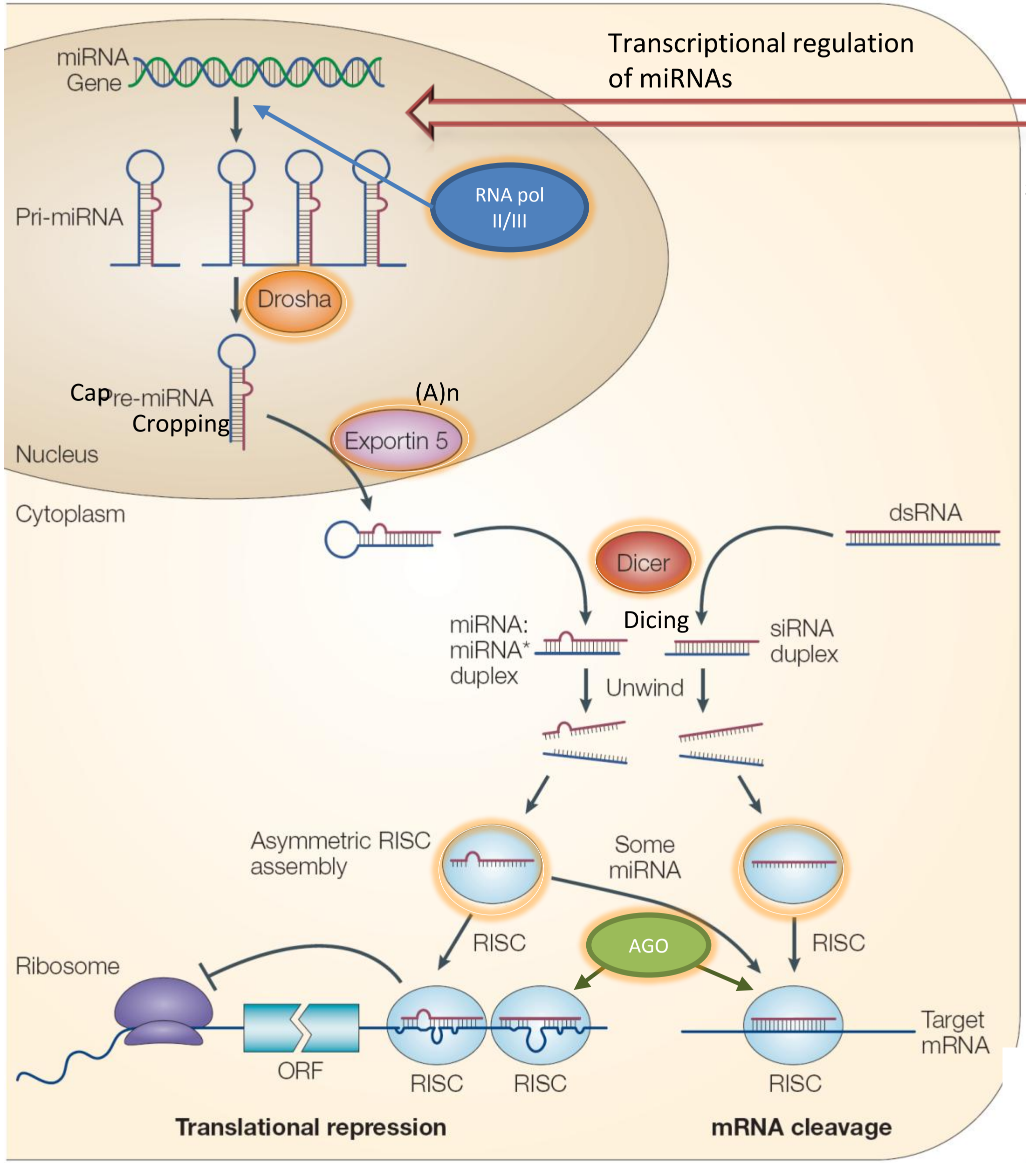
Oncomir: miRNA functions as a tumor suppressor or an oncogene



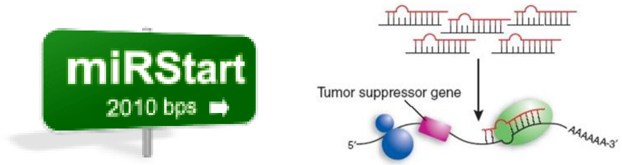
Nature Reviews 6: 259-267, 2006.

transcription
miRNA gene ------> pri-miRNA -> pre-miRNA -> miRNA ----down-regulation-----> target genes

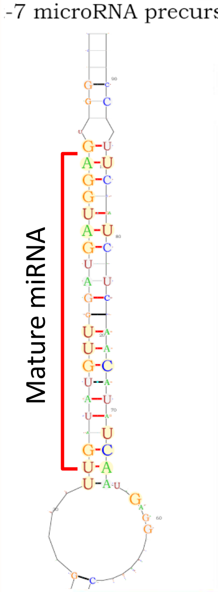
Bioinformatics Research in MicroRNA Regulation: Databases and Tools

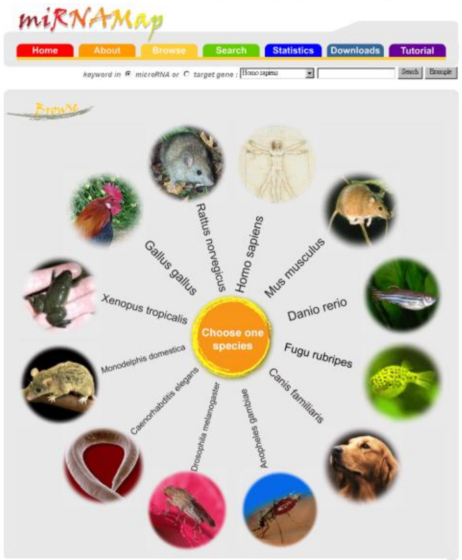


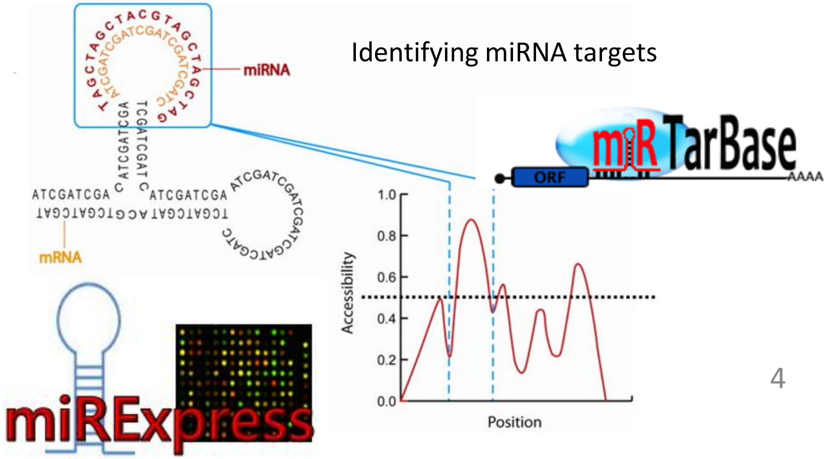
(The figure was modified from the graphic by Lin He, Nat Rev Genetics, 2004)



↓Tumor suppressor gene expression







Bioinformatics Resources for MicroRNA Research

- Gene prediction and annotation:
  - Identification of miRNA genes (intergenic and intragenic miRNAs)
    - miRNAMap (*NAR, 2006, 2008*), miRExpress (*BMC Bioinformatics, 2009*)
  - Comparative analysis of miRNAs
    - RNALogo (*NAR, 2009*)
- Functions of miRNAs:
  - Identification of miRNA targets
    - miRNAMap (*NAR, 2006, 2008*), miRTarBase (*NAR, 2011*), miRTar (*BMC Bioinformatics, 2011*) and ViTa (*NAR, 2009*)
  - Tumor suppressors or oncogenes
- Regulation of miRNAs:
  - Transcriptional regulation of miRNA
    - miRStart (*NAR, 2011*)
  - Expression profiles of miRNAs
    - miRExpress (*BMC Bioinformatics, 2009*)

miRNA/small RNA Resources

miRNA databases: miRNAMap, ViTa

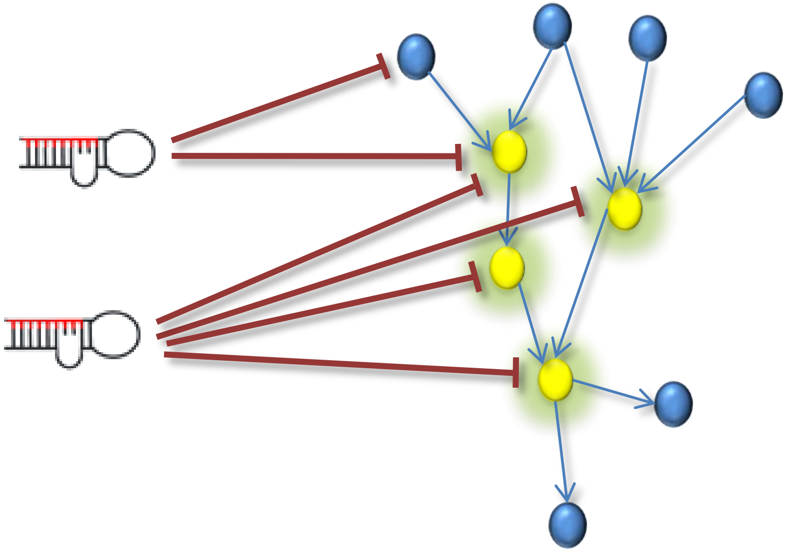
miRNA-target interaction database: miRTarBase

Promoters of miRNAs: miRStart

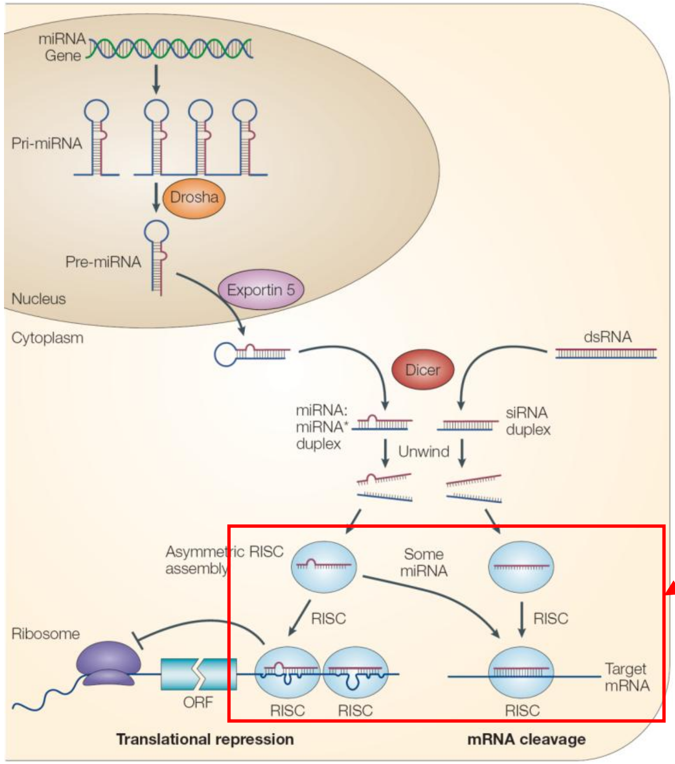
Tools for identifying miRNA-target interactions: miRTar, miRExpress

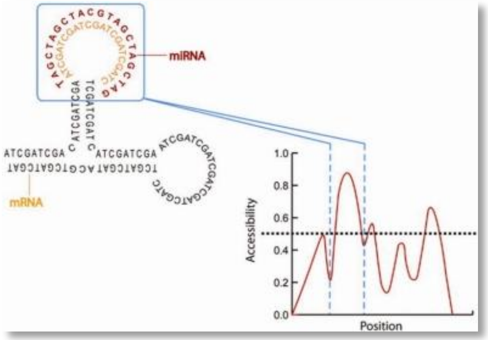
Tools for RNA analysis: RNALogo, RegRNA, sRNAMap, RNAMST, RiboSW

miRNAs/miRNA precursors



miRNA-target Interactions





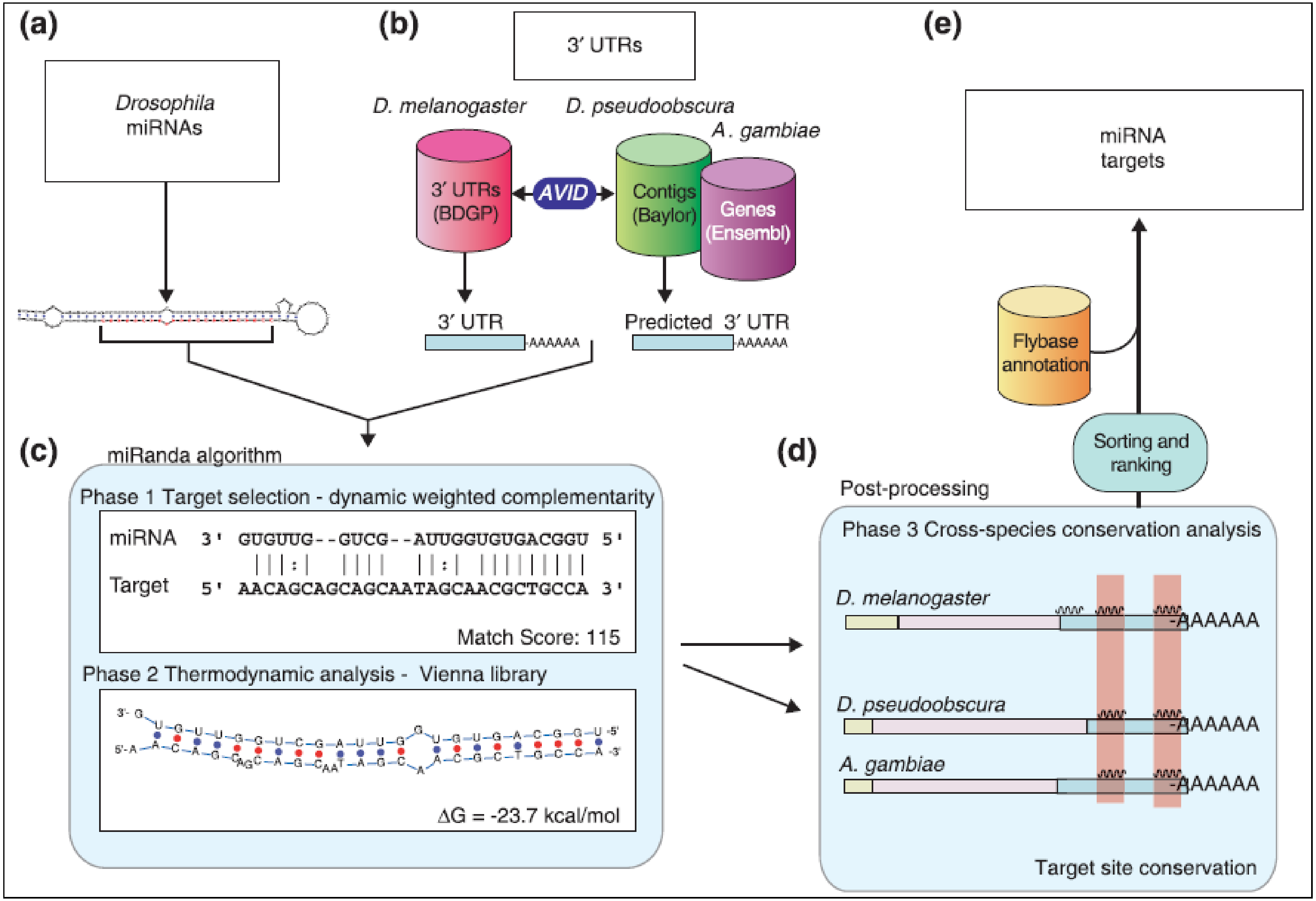
Computational identification of miRNA targets

Tools for identifying miRNA targets

miRanda

- microRNA sequences (in 2003) (*Genome Biol. 5*, R1, 2003)
  - miRBase
  - 73 unique miRNA sequences
- 3'UTR sequences
  - BDGP (Berkeley *Drosophila* Genome Project)
- Looking for high-complementarity regions on the 3'UTRs.
  - dynamic programming algorithm
- Evaluate thermodynamic of potential binding sites
  - Vienna RNA folding package

miRanda - Algorithms



Dr. Hsien-Da Huang, professor

An integrated web server for identifying miRNA-target interactions in human

miRTar

MicroRNA Target prediction (miRTar) is a tool that enables biologists easily to identify the **biological functions** and **regulatory relationships** between a group of known/putative **miRNAs** and **protein coding genes**. It also provides perspective of information on the miRNA targets on **alternatively spliced transcripts**.

miRTar supports four major features:

- Identifying miRNA Targets in Human
- Multiple scenarios
- Viewpoint on the regulation between miRNA and RNA alternative splicing
- Gene enrichment in KEGG pathway maps

Examples

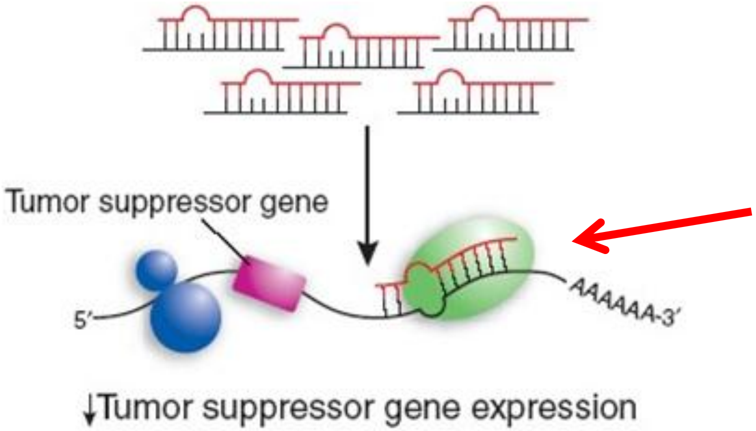
❶ Mature miRNA sequence:

Targets To

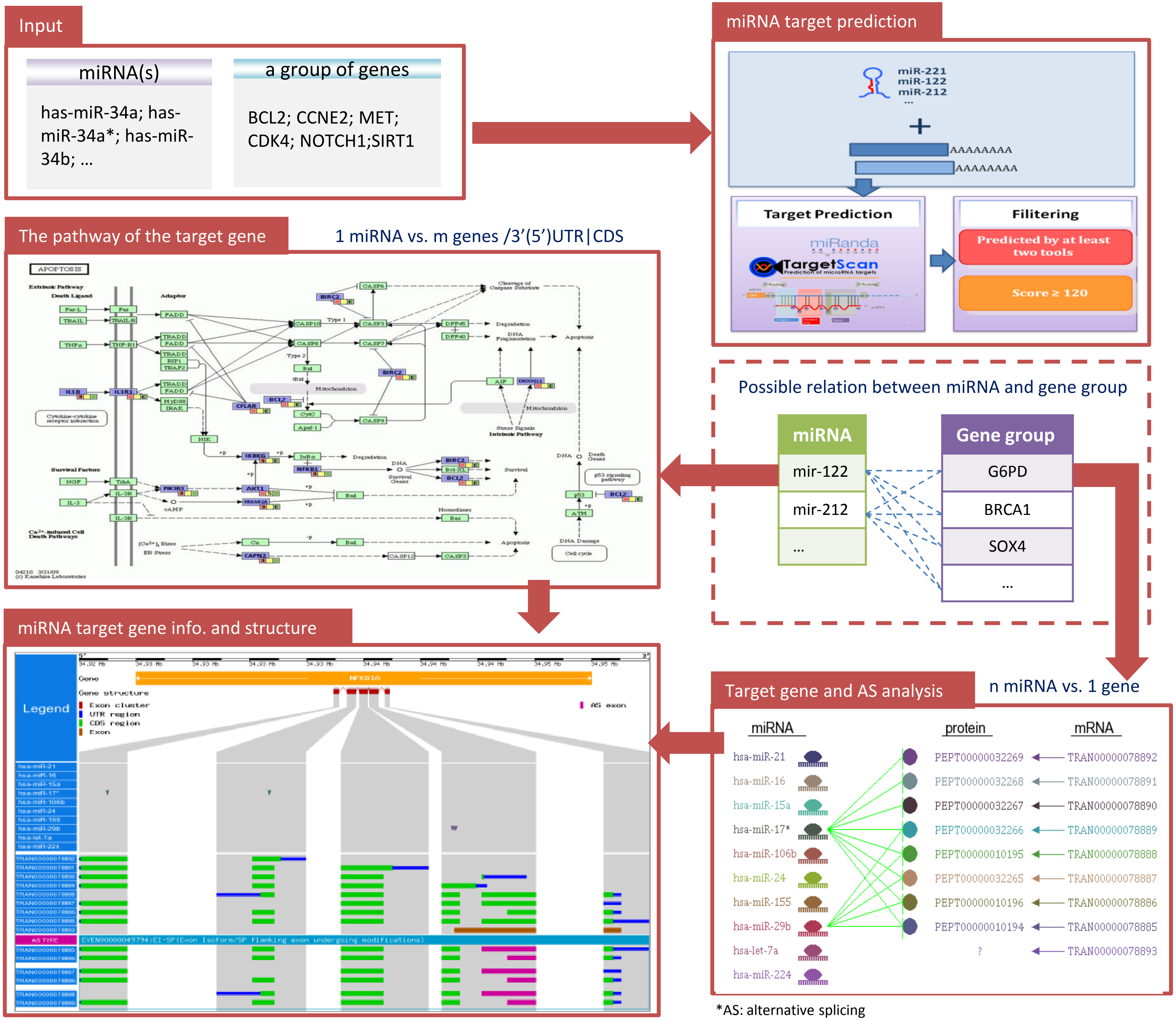
❷ Target Gene:

miRTar provides multiple functions for identifying miRNA-target interactions

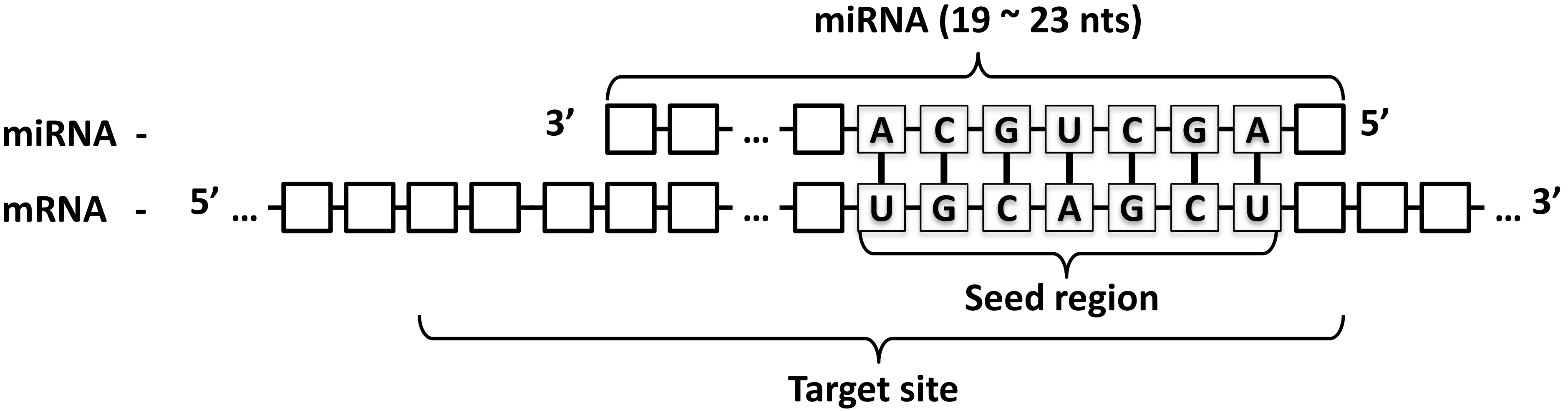
- Single miRNA to single gene (1:1)
- Single miRNA to multiple genes (1:N)
- Multiple miRNAs to single gene (N:1)
- Multiple miRNAs to multiple genes (N:N)



miRNA-target interactions


Input
miRNA(s)
has-miR-34a; has-miR-34a*; has-miR-34b; …
a group of genes
BCL2; CCNE2; MET; CDK4; NOTCH1;SIRT1
miRNA target prediction
miR-221
miR-122
miR-212
…
+
AAAAAAAA
AAAAAAAA
Target Prediction
miRanda
TargetScan
Filitering
Predicted by at least two tools
Score ≥ 120
The pathway of the target gene
1 miRNA vs. m genes /3'(5')UTR|CDS
APOPTOSIS
Possible relation between miRNA and gene group
miRNA
mir-122
mir-212
…
Gene group
G6PD
BRCA1
SOX4
…
miRNA target gene info. and structure
Legend
Gene
Gene structure
Exon cluster
UTR region
CDS region
Exon
AS exon
NFKBIA
hsa-miR-21
hsa-miR-16
hsa-miR-15a
hsa-miR-17*
hsa-miR-106b
hsa-miR-24
hsa-miR-155
hsa-miR-29b
hsa-let-7a
hsa-miR-224
AS TYPE
EVEN00000049794:EI-5F(Exon Isoform/5F flanking exon undergoing modifications)
Target gene and AS analysis
n miRNA vs. 1 gene
miRNA
protein
mRNA
hsa-miR-21
hsa-miR-16
hsa-miR-15a
hsa-miR-17*
hsa-miR-106b
hsa-miR-24
hsa-miR-155
hsa-miR-29b
hsa-let-7a
hsa-miR-224
PEPT00000032269
PEPT00000032268
PEPT00000032267
PEPT00000032266
PEPT00000010195
PEPT00000032265
PEPT00000010196
PEPT00000010194
?
TRAN00000078892
TRAN00000078891
TRAN00000078890
TRAN00000078889
TRAN00000078888
TRAN00000078887
TRAN00000078886
TRAN00000078885
TRAN00000078893
*AS: alternative splicing

miRNA-target interactions (MTIs)

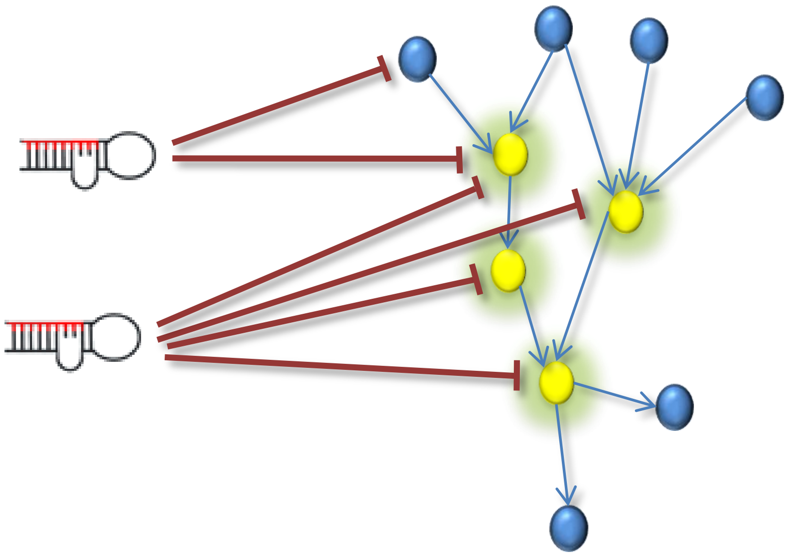

miRNA (19 ~ 23 nts)

miRNA - 3' … A C G U C G A 5'

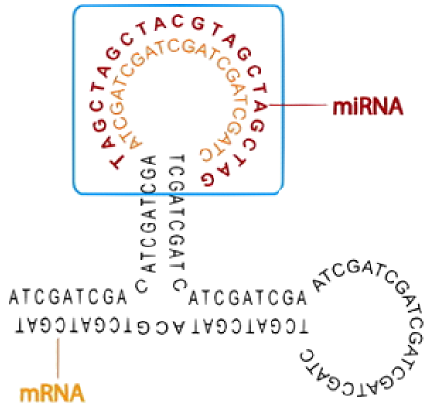
mRNA - 5' … … U G C A G C U … 3'

Seed region

Target site



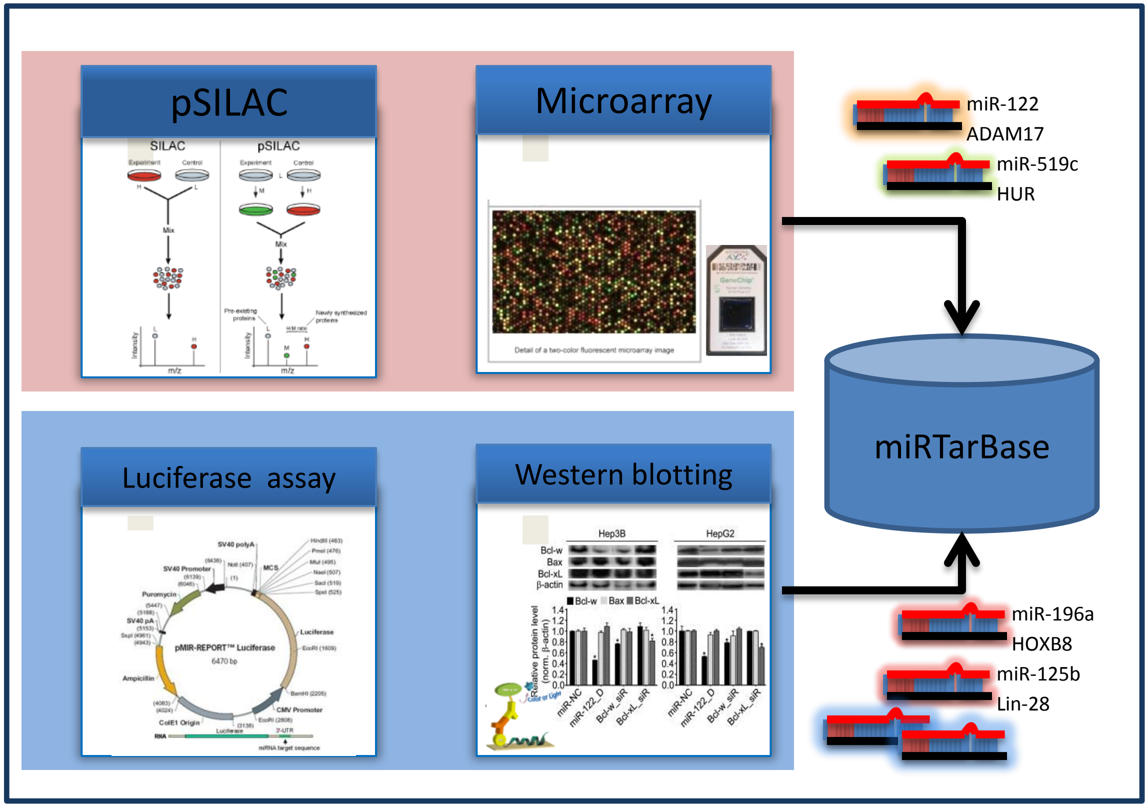
miRNA-target Interactions

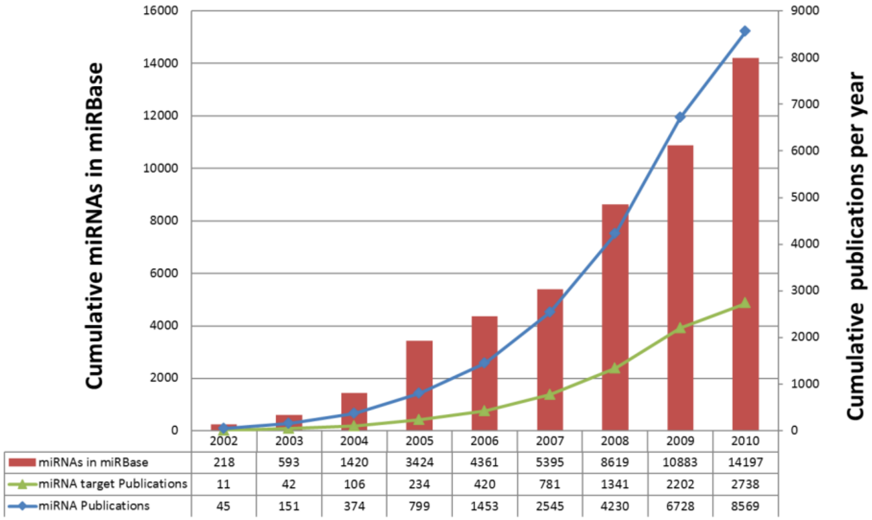


miRTarBase

a database curates experimentally validated miRNA-target interactions (NAR database issue, 2011)

http://miRTarBase.mbc.nctu.edu.tw/

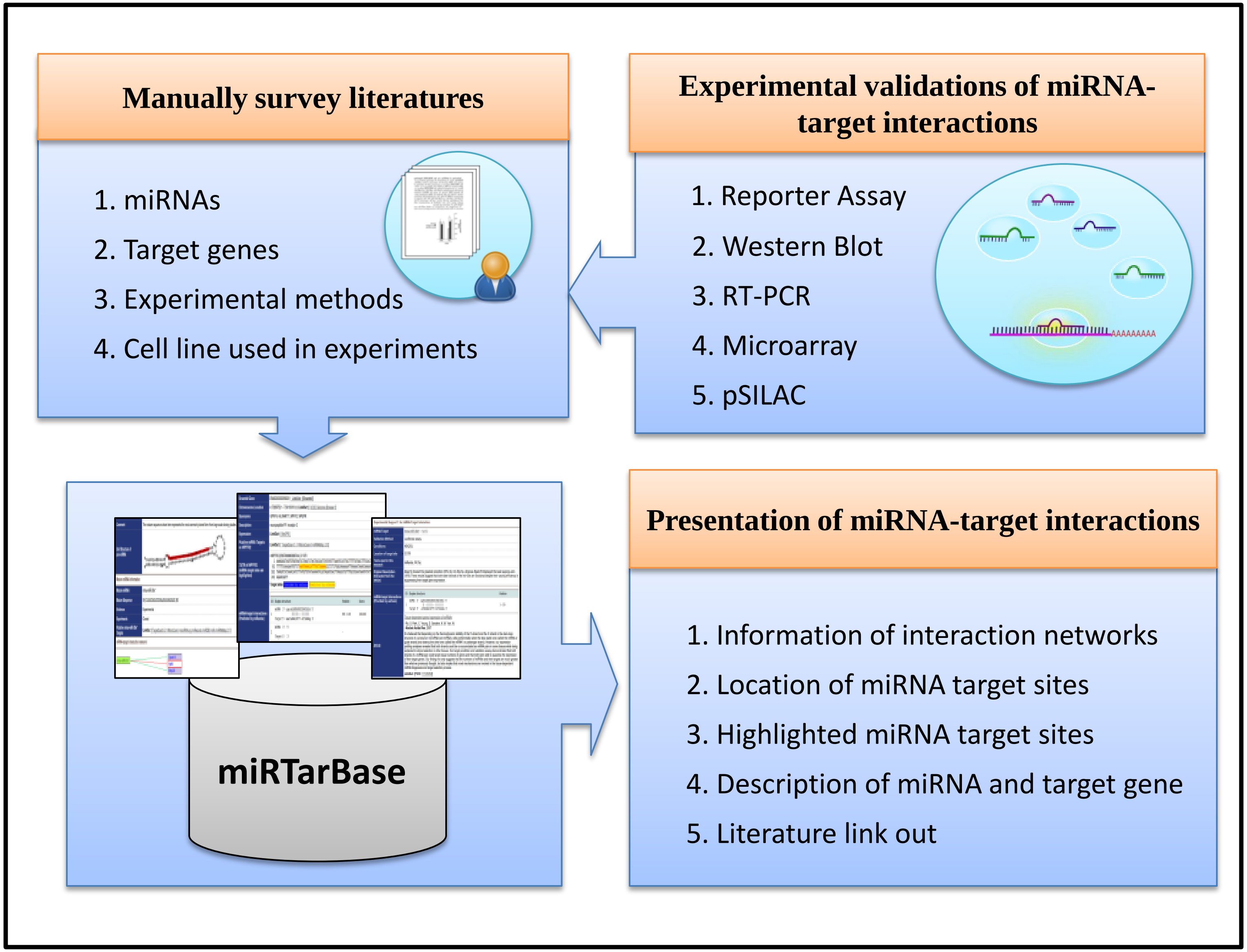




	2002	2003	2004	2005	2006	2007	2008	2009	2010
miRNAs in miRBase	218	593	1420	3424	4361	5395	8619	10883	14197
miRNA target Publications	11	42	106	234	420	781	1341	2202	2738
miRNA Publications	45	151	374	799	1453	2545	4230	6728	8569

1. By surveying literature
2. To collect updated information of miRNA-target interactions
3. 4270 entries now, will be more than 10,000 entries in 2013, expectedly
4. Biggest collection now
4. Discover miRNA-target interactions in mouse based on the extension of human miRNA-target interaction (Evolutionary conservation)

System flow of miRTarBase



Manually survey literatures

1. miRNAs
2. Target genes
3. Experimental methods
4. Cell line used in experiments

Experimental validations of miRNA-target interactions

1. Reporter Assay
2. Western Blot
3. RT-PCR
4. Microarray
5. pSILAC

Presentation of miRNA-target interactions

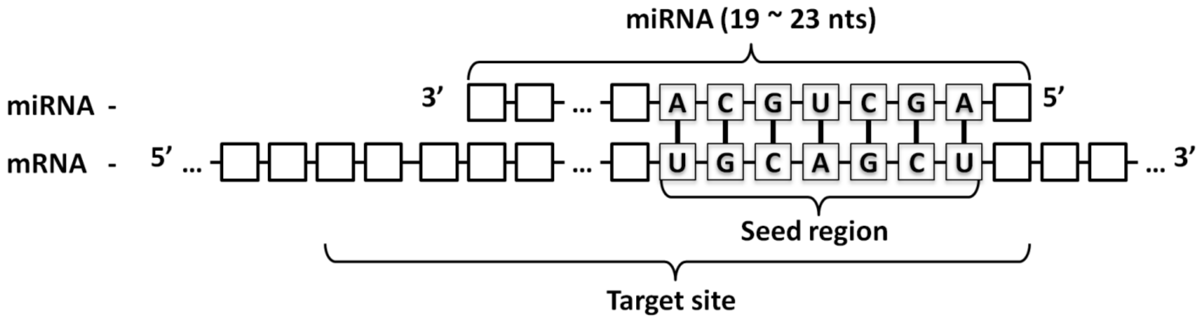
1. Information of interaction networks
2. Location of miRNA target sites
3. Highlighted miRNA target sites
4. Description of miRNA and target gene
5. Literature link out

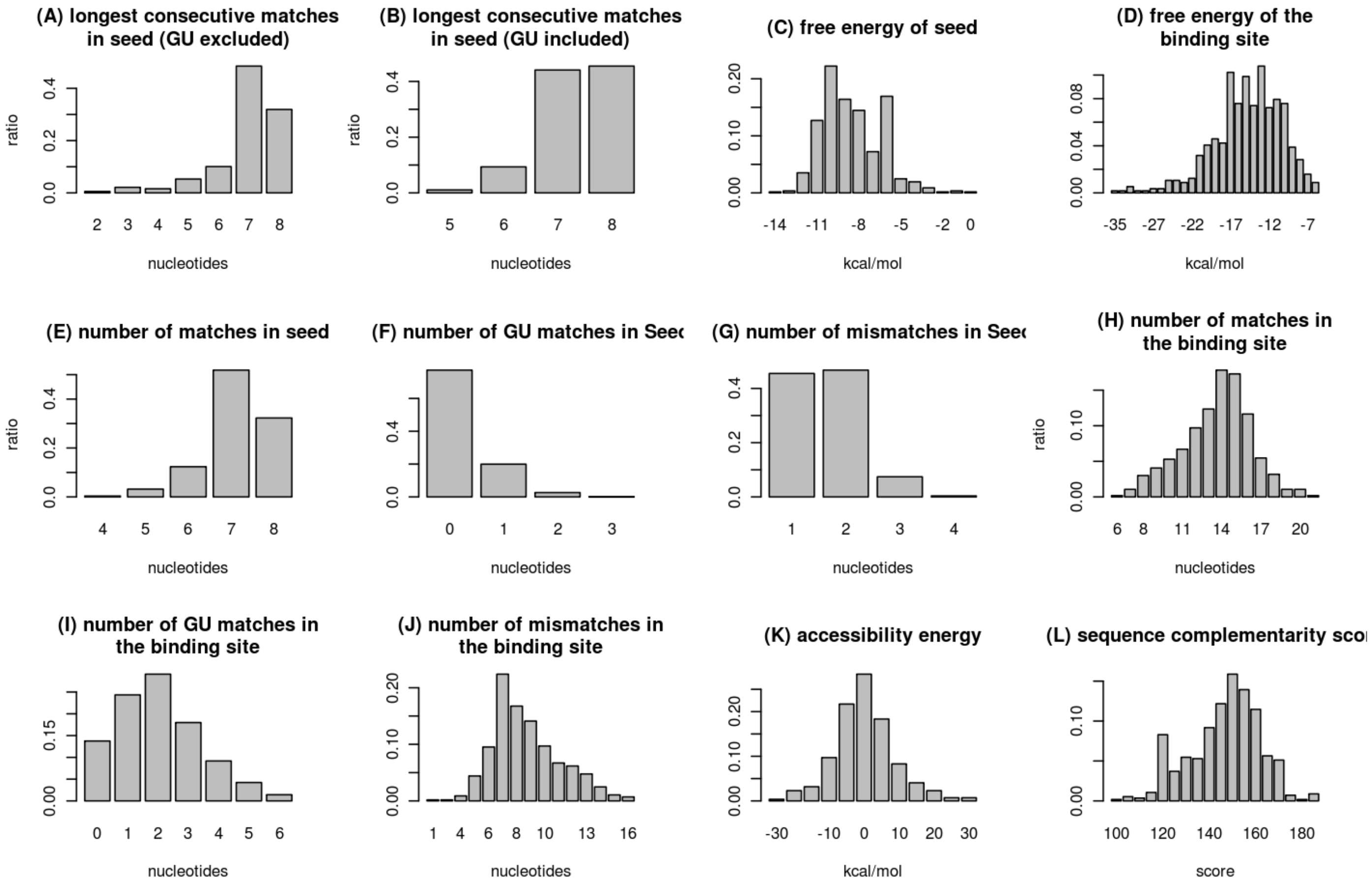
Comparison with other resources

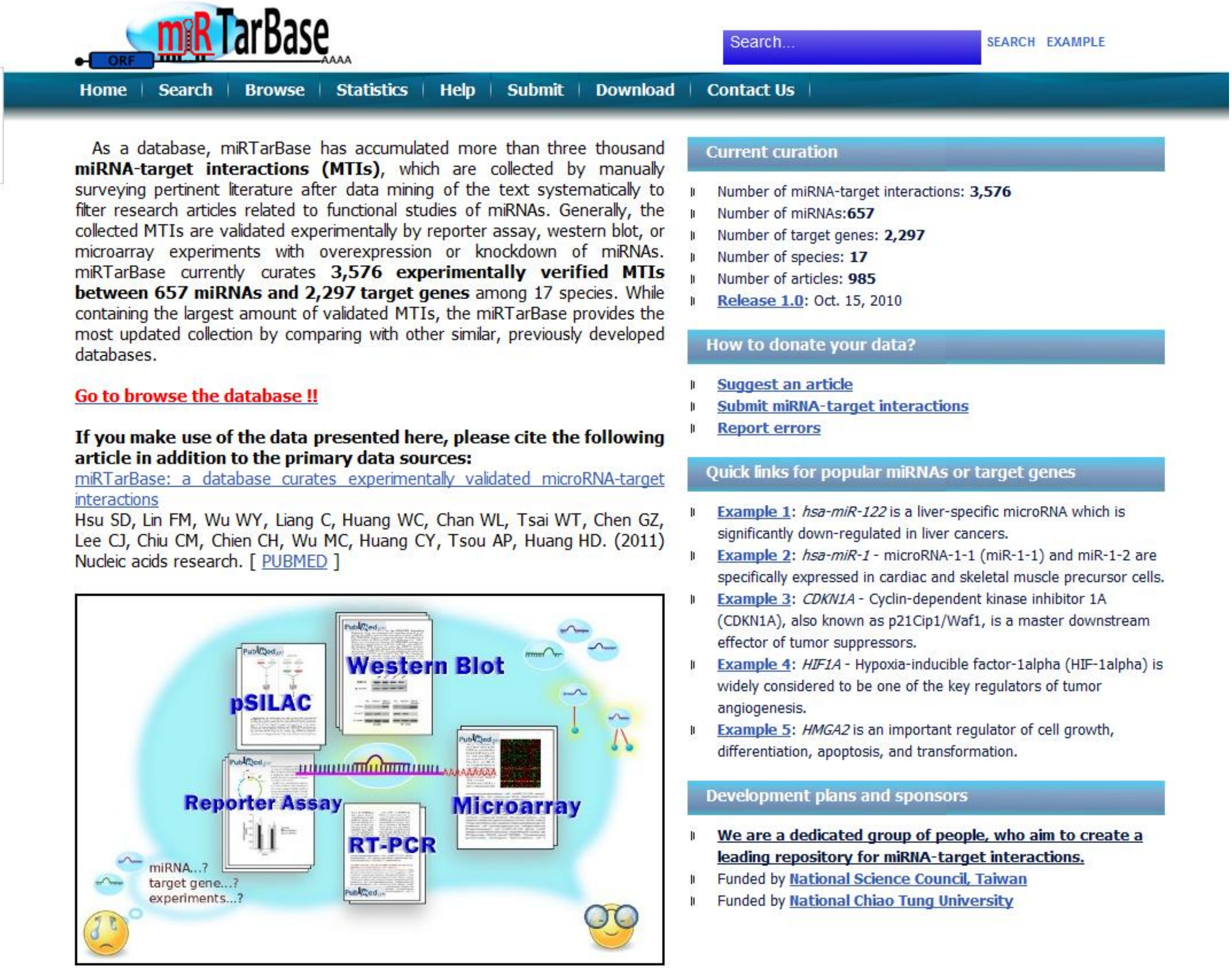
	TarBase	**miRecords**	**miR2Disease**	**miRTarBase**	**Number of records added**
Publications	RNA (2006), Nucleic Acids Res. Database Issue (2009)	Nucleic Acids Res. Database Issue (2009)	Nucleic Acids Res. Database Issue (2009)	NAR Database Issue (2011)	
Release version	V5	V1			
Last update	2008/6	2010/5/5	2010/06/02	2011/10/15	
Support species	Metazoa x 6 Viridiplantae Viruses	Metazoa x 11 Viruses x 2	Human	Metazoa x 8 Viridiplantae x 3 Viruses x 4	
Number of miRNAs	223	381*	179	657	+ 276
Number of target genes	1028	1058*	394	2297	+ 1240
Number of **articles**	155	410	430*	985	+ 564
Number of **miRNA-target interactions**	1264	1513*	635	3576	+ 2063
No. of miRNA-target interactions validated by "Luciferase reporter assay"	305*	256	0	2017	+ 1345
No. of miRNA-target interactions validated by "Western blotting"	27	290*	0	901	+ 606
No. of miRNA-target interactions validated by "Luciferase reporter assay AND Western blotting"	25	123	0	711	+ 588
No. of miRNA-target interactions validated by "Luciferase reporter assay or Western blotting"	307	636*	635	2207	+ 1460
No. of miRNA-target interactions validated by "pSILAC experiments"	455	0	0	455	+ 0
No. of miRNA-target interactions validated by "Microarray experiments"	343	380*	0	861	+ 481

Histogram of various features of experimental proven miRNA-target sites

Data were obtained from 527 human miRNA-target interactions curated in miRTarBase







As a database, miRTarBase has accumulated more than three thousand **miRNA-target interactions (MTIs)**, which are collected by manually surveying pertinent literature after data mining of the text systematically to filter research articles related to functional studies of miRNAs. Generally, the collected MTIs are validated experimentally by reporter assay, western blot, or microarray experiments with overexpression or knockdown of miRNAs. miRTarBase currently curates **3,576 experimentally verified MTIs between 657 miRNAs and 2,297 target genes** among 17 species. While containing the largest amount of validated MTIs, the miRTarBase provides the most updated collection by comparing with other similar, previously developed databases.

Go to browse the database !!

If you make use of the data presented here, please cite the following article in addition to the primary data sources:
miRTarBase: a database curates experimentally validated microRNA-target interactions
Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, Tsai WT, Chen GZ, Lee CJ, Chiu CM, Chien CH, Wu MC, Huang CY, Tsou AP, Huang HD. (2011) Nucleic acids research. [PUBMED]

Current curation

- Number of miRNA-target interactions: **3,576**
- Number of miRNAs:**657**
- Number of target genes: **2,297**
- Number of species: **17**
- Number of articles: **985**
- Release 1.0: Oct. 15, 2010

How to donate your data?

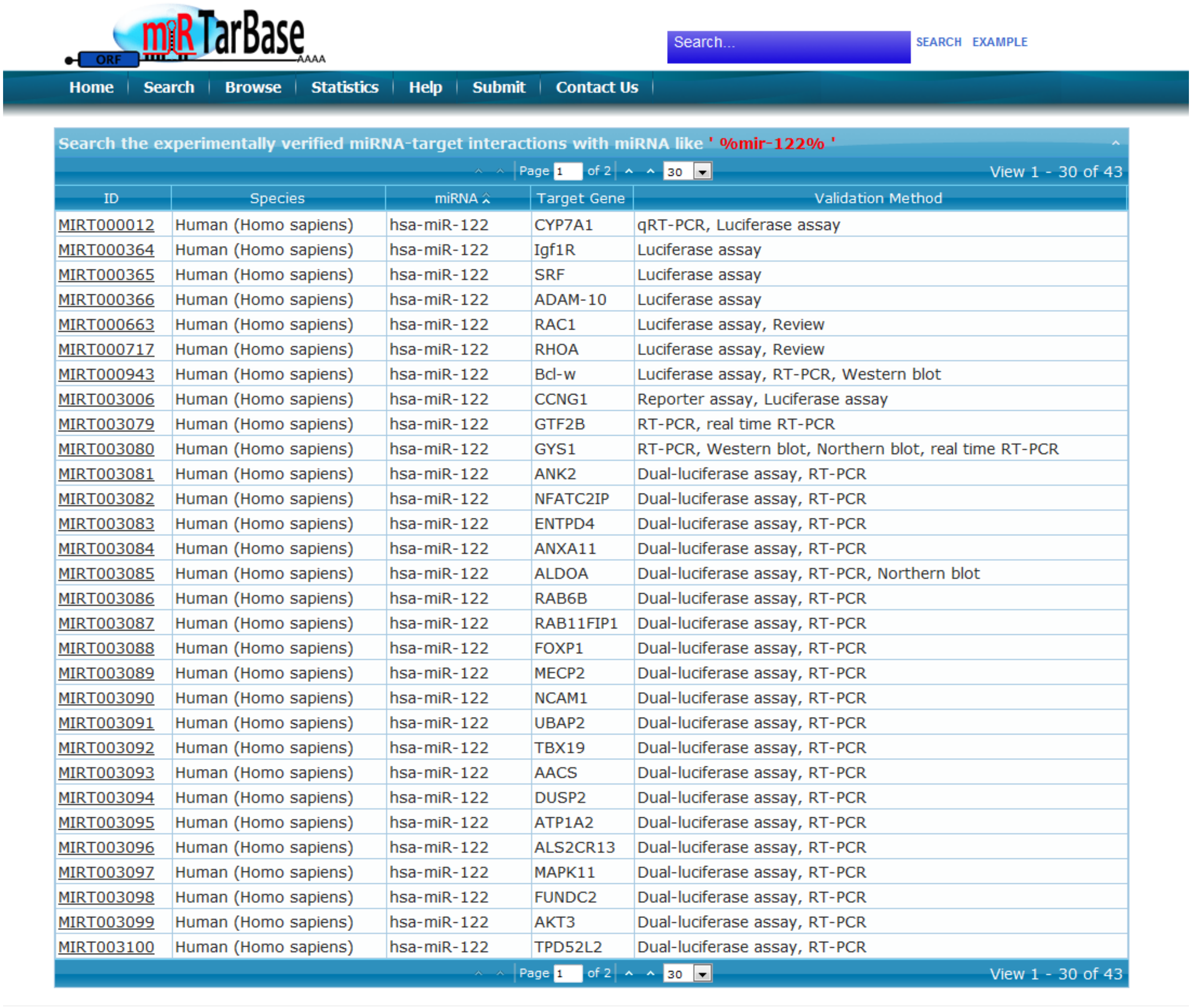
- Suggest an article
- Submit miRNA-target interactions
- Report errors

Quick links for popular miRNAs or target genes

- Example 1: *hsa-miR-122* is a liver-specific microRNA which is significantly down-regulated in liver cancers.
- Example 2: *hsa-miR-1* - microRNA-1-1 (miR-1-1) and miR-1-2 are specifically expressed in cardiac and skeletal muscle precursor cells.
- Example 3: *CDKN1A* - Cyclin-dependent kinase inhibitor 1A (CDKN1A), also known as p21Cip1/Waf1, is a master downstream effector of tumor suppressors.
- Example 4: *HIF1A* - Hypoxia-inducible factor-1alpha (HIF-1alpha) is widely considered to be one of the key regulators of tumor angiogenesis.
- Example 5: *HMGA2* is an important regulator of cell growth, differentiation, apoptosis, and transformation.

Development plans and sponsors

- **We are a dedicated group of people, who aim to create a leading repository for miRNA-target interactions.**
- Funded by **National Science Council, Taiwan**
- Funded by **National Chiao Tung University**



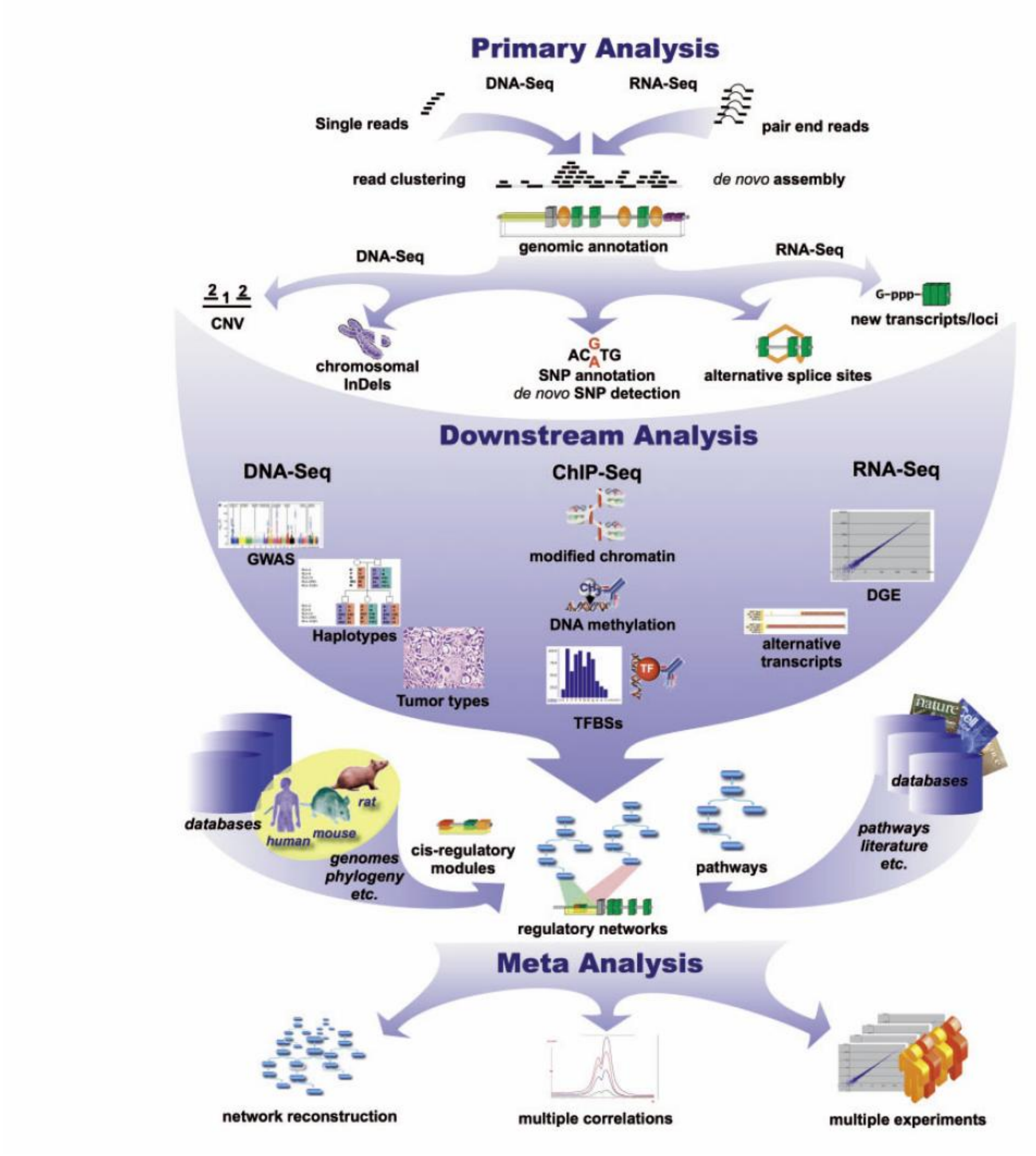
Search the experimentally verified miRNA-target interactions with miRNA like ' %mir-122% '

View 1 - 30 of 43

ID	Species	miRNA	Target Gene	Validation Method
MIRT000012	Human (Homo sapiens)	hsa-miR-122	CYP7A1	qRT-PCR, Luciferase assay
MIRT000364	Human (Homo sapiens)	hsa-miR-122	Igf1R	Luciferase assay
MIRT000365	Human (Homo sapiens)	hsa-miR-122	SRF	Luciferase assay
MIRT000366	Human (Homo sapiens)	hsa-miR-122	ADAM-10	Luciferase assay
MIRT000663	Human (Homo sapiens)	hsa-miR-122	RAC1	Luciferase assay, Review
MIRT000717	Human (Homo sapiens)	hsa-miR-122	RHOA	Luciferase assay, Review
MIRT000943	Human (Homo sapiens)	hsa-miR-122	Bcl-w	Luciferase assay, RT-PCR, Western blot
MIRT003006	Human (Homo sapiens)	hsa-miR-122	CCNG1	Reporter assay, Luciferase assay
MIRT003079	Human (Homo sapiens)	hsa-miR-122	GTF2B	RT-PCR, real time RT-PCR
MIRT003080	Human (Homo sapiens)	hsa-miR-122	GYS1	RT-PCR, Western blot, Northern blot, real time RT-PCR
MIRT003081	Human (Homo sapiens)	hsa-miR-122	ANK2	Dual-luciferase assay, RT-PCR
MIRT003082	Human (Homo sapiens)	hsa-miR-122	NFATC2IP	Dual-luciferase assay, RT-PCR
MIRT003083	Human (Homo sapiens)	hsa-miR-122	ENTPD4	Dual-luciferase assay, RT-PCR
MIRT003084	Human (Homo sapiens)	hsa-miR-122	ANXA11	Dual-luciferase assay, RT-PCR
MIRT003085	Human (Homo sapiens)	hsa-miR-122	ALDOA	Dual-luciferase assay, RT-PCR, Northern blot
MIRT003086	Human (Homo sapiens)	hsa-miR-122	RAB6B	Dual-luciferase assay, RT-PCR
MIRT003087	Human (Homo sapiens)	hsa-miR-122	RAB11FIP1	Dual-luciferase assay, RT-PCR
MIRT003088	Human (Homo sapiens)	hsa-miR-122	FOXP1	Dual-luciferase assay, RT-PCR
MIRT003089	Human (Homo sapiens)	hsa-miR-122	MECP2	Dual-luciferase assay, RT-PCR
MIRT003090	Human (Homo sapiens)	hsa-miR-122	NCAM1	Dual-luciferase assay, RT-PCR
MIRT003091	Human (Homo sapiens)	hsa-miR-122	UBAP2	Dual-luciferase assay, RT-PCR
MIRT003092	Human (Homo sapiens)	hsa-miR-122	TBX19	Dual-luciferase assay, RT-PCR
MIRT003093	Human (Homo sapiens)	hsa-miR-122	AACS	Dual-luciferase assay, RT-PCR
MIRT003094	Human (Homo sapiens)	hsa-miR-122	DUSP2	Dual-luciferase assay, RT-PCR
MIRT003095	Human (Homo sapiens)	hsa-miR-122	ATP1A2	Dual-luciferase assay, RT-PCR
MIRT003096	Human (Homo sapiens)	hsa-miR-122	ALS2CR13	Dual-luciferase assay, RT-PCR
MIRT003097	Human (Homo sapiens)	hsa-miR-122	MAPK11	Dual-luciferase assay, RT-PCR
MIRT003098	Human (Homo sapiens)	hsa-miR-122	FUNDC2	Dual-luciferase assay, RT-PCR
MIRT003099	Human (Homo sapiens)	hsa-miR-122	AKT3	Dual-luciferase assay, RT-PCR
MIRT003100	Human (Homo sapiens)	hsa-miR-122	TPD52L2	Dual-luciferase assay, RT-PCR

View 1 - 30 of 43

Next Generation Sequencing (NGS) in Functional Genomics

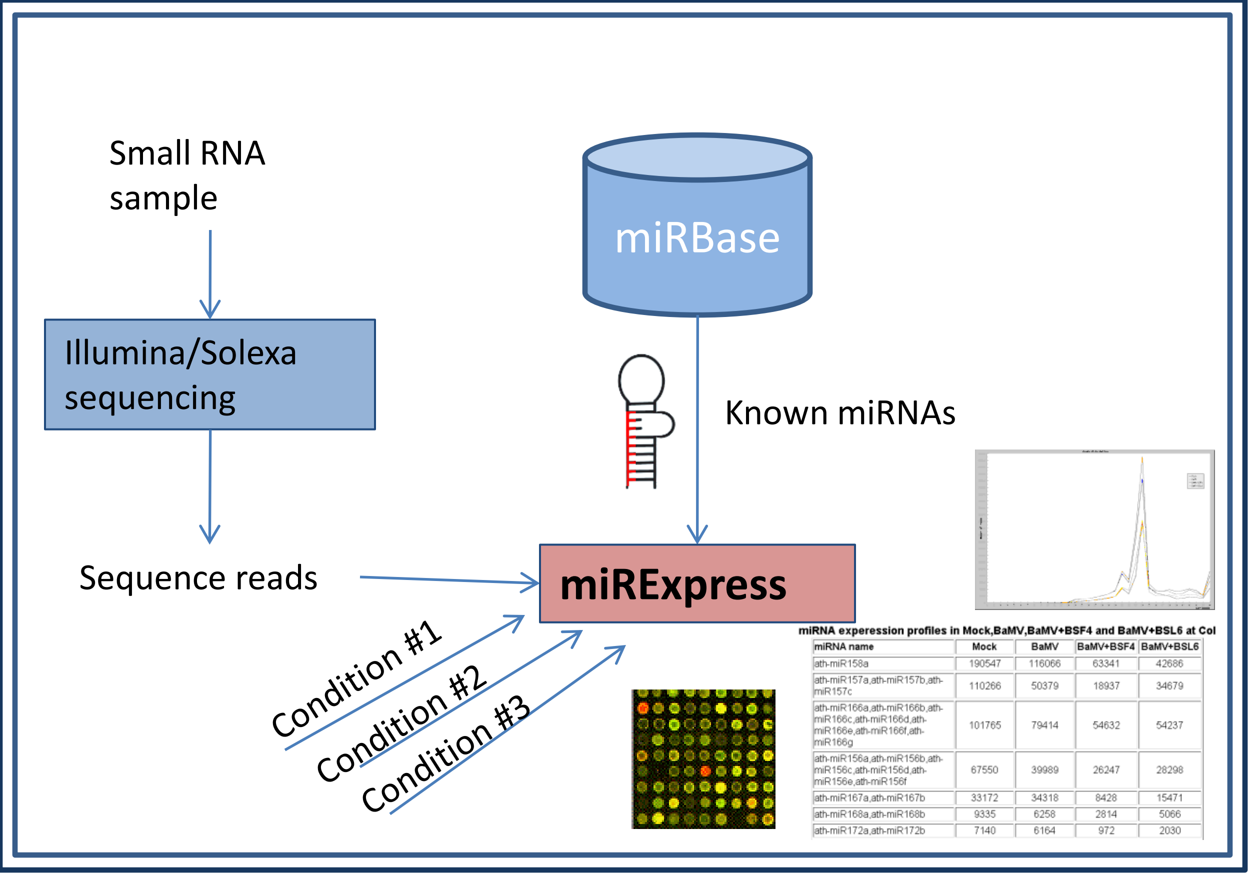


- Pipeline ready
  - RNA-seq
  - Small RNA-seq
  - ChIP-seq
  - Genome sequencing
  - Human resequencing
  - DNA methylation sequencing
- In developing
  - Amplicons sequencing
  - Metagenomics
    - Ribosomal RNAs sequencing only
    - Multiple genome sequencing

Figure from Briefings in Bioinformatics, 2010, Vol. 11, No.5, pp. 499-511

Next-Generation Sequencing (NGS) - small RNA sequencing

- Identification of miRNAs
- miRNA expression profiling


Small RNA sample

Illumina/Solexa sequencing

Sequence reads

miRBase

Known miRNAs

miRExpress

Condition #1

Condition #2

Condition #3

miRNA experession profiles in Mock,BaMV,BaMV+BSF4 and BaMV+BSL6 at Col

miRNA name	Mock	BaMV	BaMV+BSF4	BaMV+BSL6
ath-miR158a	190547	116066	63341	42686
ath-miR157a,ath-miR157b,ath-miR157c	110266	50379	18937	34679
ath-miR166a,ath-miR166b,ath-miR166c,ath-miR166d,ath-miR166e,ath-miR166f,ath-miR166g	101765	79414	54632	54237
ath-miR156a,ath-miR156b,ath-miR156c,ath-miR156d,ath-miR156e,ath-miR156f	67550	39989	26247	28298
ath-miR167a,ath-miR167b	33172	34318	8428	15471
ath-miR168a,ath-miR168b	9335	6258	2814	5066
ath-miR172a,ath-miR172b	7140	6164	972	2030

miRExpress

An effective tool to generate **miRNA Expression** profiles from high-throughput sequencing data (***BMC Bioinformatics, 2009***)

a. High-throughput
b. Sequencing short sequences
c. Inexpensive

Next-generation Sequencing Technology

(Elaine R. Mardis 2008)

	Platform		
	Roche(454)	**Illumina**	**SOLiD**
Sequencing chemistry	Pyrosequencing	Polymerase-based sequencing-by-synthesis	Ligation-based sequencing
Amplification approach	Emulsion PCR	Bridge amplification	Emulsion PCR
Paired ends/separation	Yes/3 kb	yes/200 bp	Yes/3 kb
Mb/run	100 Mb	1300 Mb	3000 Mb
Time/run (paired ends)	7 h	4 days	5 days
Read length	250 bp	32–40 bp	35 bp
Cost per run (total direct[a])	$8439	$8950	$17 447
Cost per Mb	$84.39	$5.97	$5.81

The example of next-generation sequencing (NGS) for miRNAs

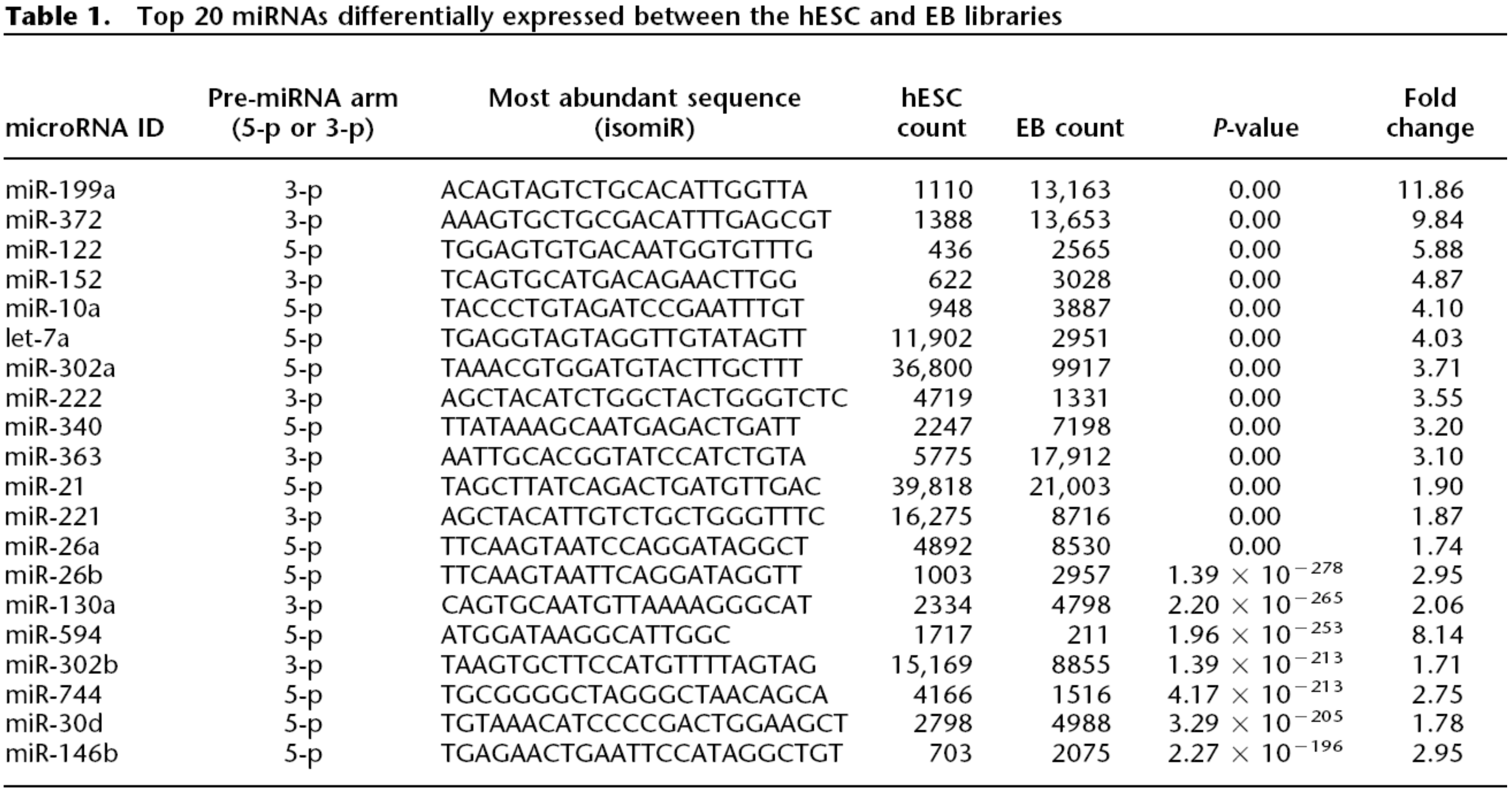


Table 1. Top 20 miRNAs differentially expressed between the hESC and EB libraries

microRNA ID	Pre-miRNA arm (5-p or 3-p)	Most abundant sequence (isomiR)	hESC count	EB count	*P*-value	Fold change
miR-199a	3-p	ACAGTAGTCTGCACATTGGTTA	1110	13,163	0.00	11.86
miR-372	3-p	AAAGTGCTGCGACATTTGAGCGT	1388	13,653	0.00	9.84
miR-122	5-p	TGGAGTGTGACAATGGTGTTTG	436	2565	0.00	5.88
miR-152	3-p	TCAGTGCATGACAGAACTTGG	622	3028	0.00	4.87
miR-10a	5-p	TACCCTGTAGATCCGAATTTGT	948	3887	0.00	4.10
let-7a	5-p	TGAGGTAGTAGGTTGTATAGTT	11,902	2951	0.00	4.03
miR-302a	5-p	TAAACGTGGATGTACTTGCTTT	36,800	9917	0.00	3.71
miR-222	3-p	AGCTACATCTGGCTACTGGGTCTC	4719	1331	0.00	3.55
miR-340	5-p	TTATAAAGCAATGAGACTGATT	2247	7198	0.00	3.20
miR-363	3-p	AATTGCACGGTATCCATCTGTA	5775	17,912	0.00	3.10
miR-21	5-p	TAGCTTATCAGACTGATGTTGAC	39,818	21,003	0.00	1.90
miR-221	3-p	AGCTACATTGTCTGCTGGGTTTC	16,275	8716	0.00	1.87
miR-26a	5-p	TTCAAGTAATCCAGGATAGGCT	4892	8530	0.00	1.74
miR-26b	5-p	TTCAAGTAATTCAGGATAGGTT	1003	2957	1.39×10^{-278}	2.95
miR-130a	3-p	CAGTGCAATGTTAAAAGGGCAT	2334	4798	2.20×10^{-265}	2.06
miR-594	5-p	ATGGATAAGGCATTGGC	1717	211	1.96×10^{-253}	8.14
miR-302b	3-p	TAAGTGCTTCCATGTTTTAGTAG	15,169	8855	1.39×10^{-213}	1.71
miR-744	5-p	TGCGGGGCTAGGGCTAACAGCA	4166	1516	4.17×10^{-213}	2.75
miR-30d	5-p	TGTAAACATCCCCGACTGGAAGCT	2798	4988	3.29×10^{-205}	1.78
miR-146b	5-p	TGAGAACTGAATTCCATAGGCTGT	703	2075	2.27×10^{-196}	2.95

Morin, R.D., et al., *Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells.* Genome Res, 2008. 18(4): p. 610-21.

Solexa reads with adaptor

Read 1 TGACAGAAGAAAGAGAGCACTCGTATGCCGTCTT
Read 2 TGACAGAAGAAAGAGAGCACCTTCGTATGCCGTC
Read 3 TCGTATGCCGTCTTCTGCTTGTCTTACCCCGTCT
Read 4 TGACAGAAGTCGTATGCCGTCTTCTGCTTGAAAG
⋮

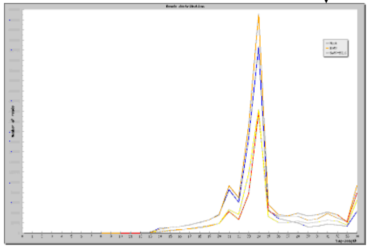
Removing adaptors

Adaptor sequence may appear in the start or end or middle of Solexa reads

Read 1 TGACAGAAGAAAGAGAGCACTCGTATGCCGTCTT
Read 2 TGACAGAAGAAAGAGAGCACCTTCGTATGCCGTC
Read 3 TCGTATGCCGTCTTCTGCTTGTCTTACCCCGTCT
Read 4 TGACAGAAGTCGTATGCCGTCTTCTGCTTGAAAG
⋮

Solexa reads without adaptor

Read 1 TGACAGAAGAAAGAGAGCAC
Read 2 TGACAGAAGAAAGAGAGCACC
Read 3 TCTTACCCCGTCT
⋮

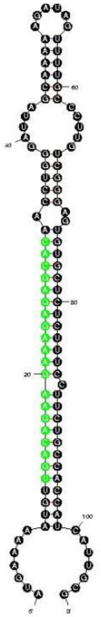


Distribution of solexa read number and count according to read length

Each reads are aligned to known miRNAs

http://microrna.sanger.ac.uk/sequences/

Each reads align with all known miRNAs and is assigned to the highest alignment score miRNA

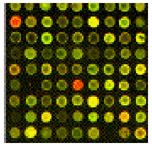


ath-miR156h ..AAAAUGUUGACAGAAGAAAGAGAGCACAACCUGG..

Reads 1 TGACAGAAGAAAGAGAGCAC
Reads 2 TGACAGAATAAAGAGAGCAC
Reads 3 TGACAGCAGAAAGAGAGCAC

Each known miRNA compute total solexa reads counts and build miRNA expression profile

miRNA expression profile



miRNA name	Mock	BaMV	BaMV+BSF4	BaMV+BSL6
ath-miR158a	193319	144993	72515	135784
ath-miR158b	150	111	61	93
ath-miR159a	580	876	170	619
ath-miR159b	124	240	48	138

Discovery of novel miRNAs in Human

Step 1

To cluster sequencing reads which can’t be mapped to known miRNAs

Constraints of miRNAs

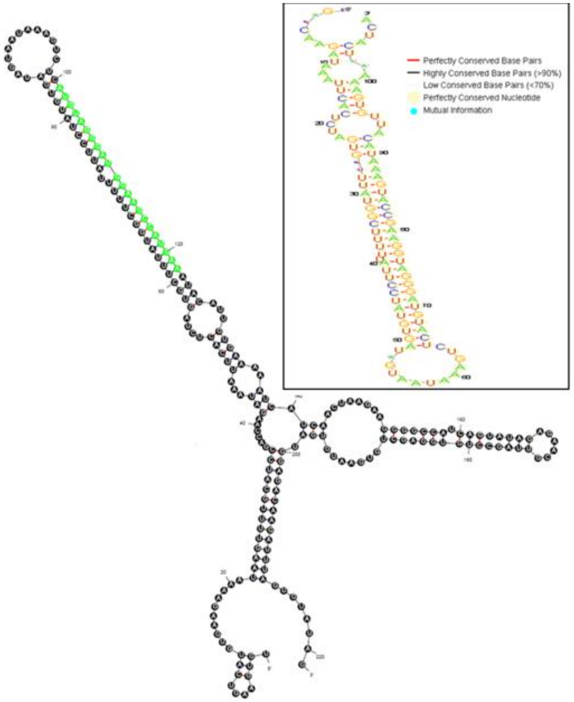
Step 2

To map sequencing read to genome and to extract its flanking upstream and downstream

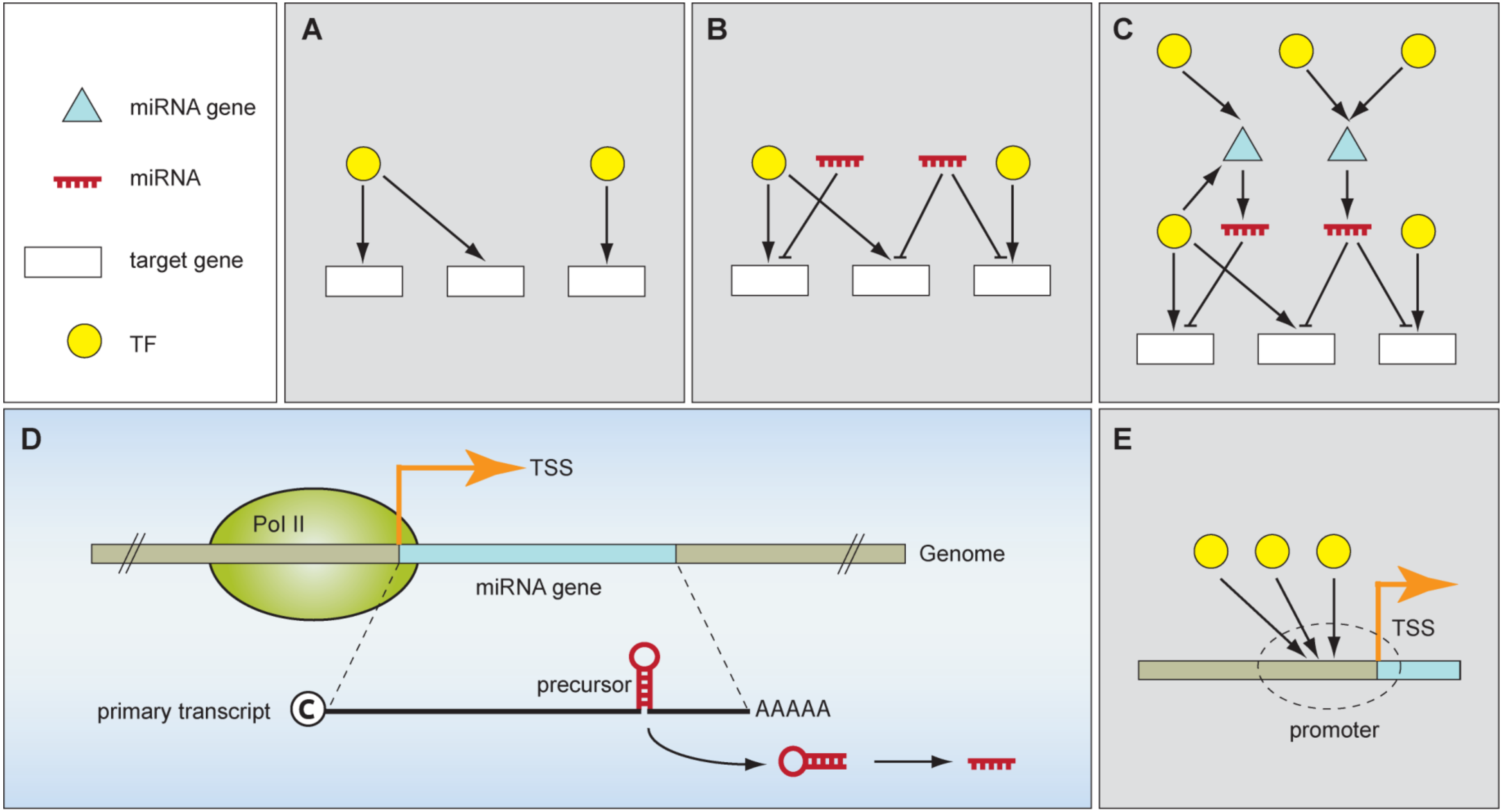
Step 3

To fold the RNA secondary structure of the candidate
[mfold]

To observe the conservation of the RNA secondary structure
[UCSC genome browser, RNALogo]



Transcriptional regulation of miRNA genes

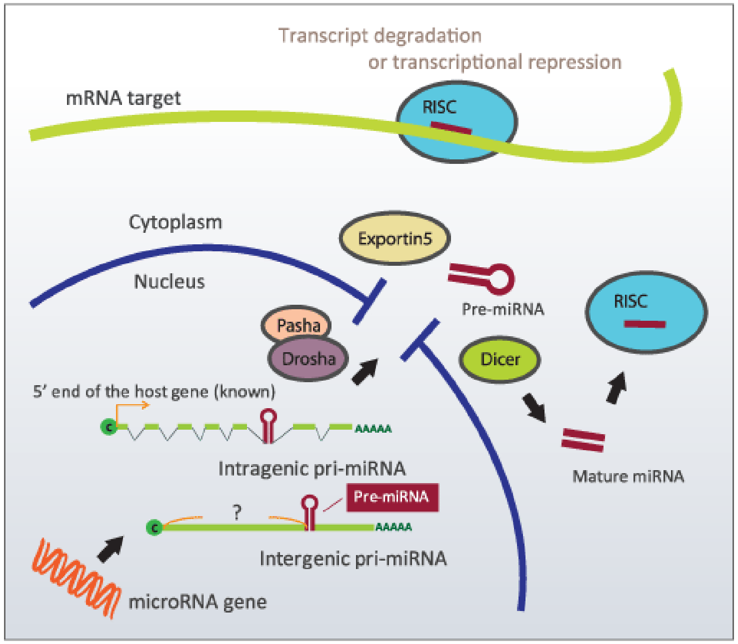


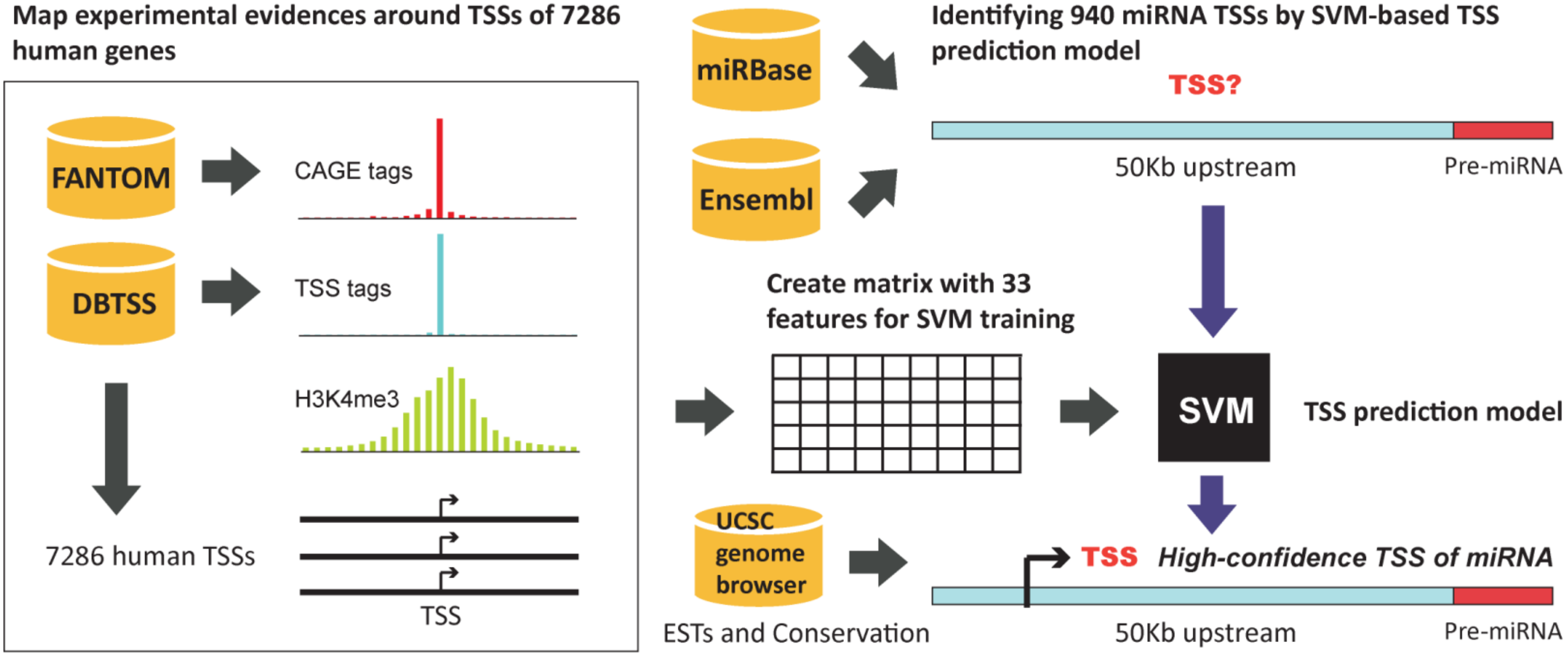
Deciphering transcriptional regulation of miRNAs

miRStart: An experimental-based resource defines transcriptional start sites of human microRNAs

(Nucleic Acids Research, 2011)

http://miRStart.mbc.nctu.edu.tw/




Map experimental evidences around TSSs of 7286 human genes

Identifying 940 miRNA TSSs by SVM-based TSS prediction model

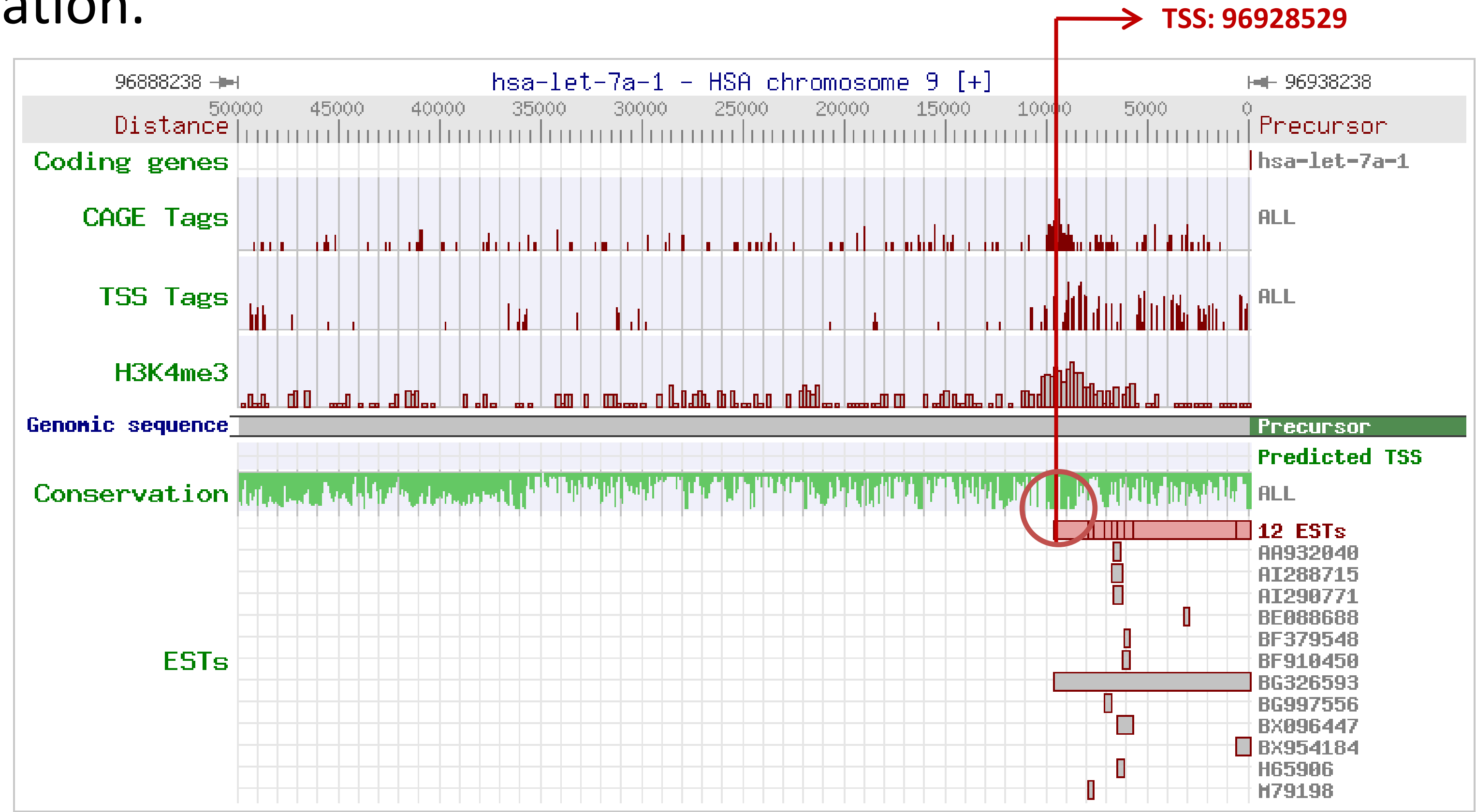
Create matrix with 33 features for SVM training

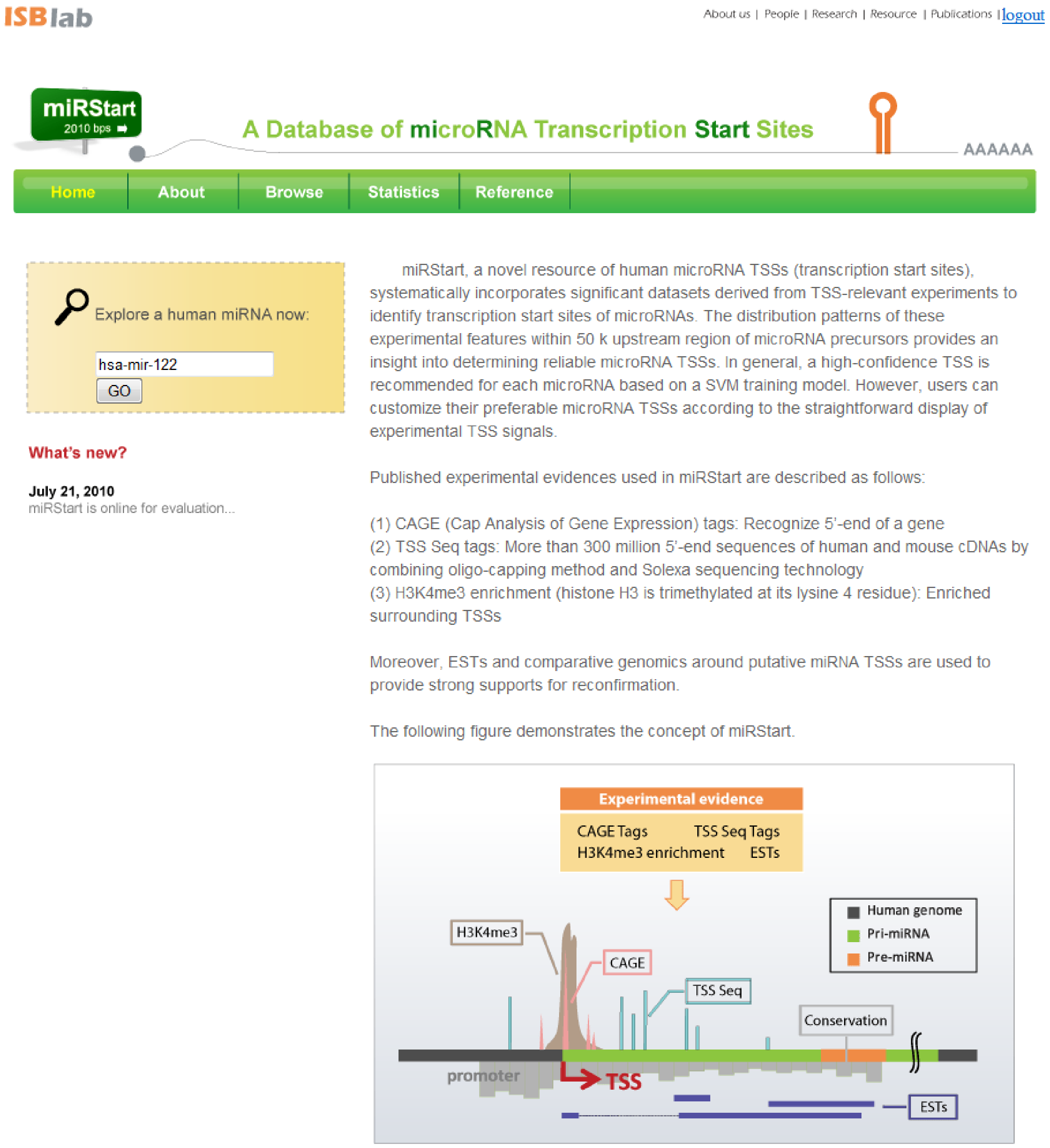
An example

5-fold cross-validation:

Sensitivity = 90.36%, Specificity = 90.05%, Accuracy = 90.21% and Precision = 90.08%, using promoters of protein-coding genes as training data

The human miRNA let-7a-1 demonstrates a good example of TSS identification.





Explore a human miRNA now:

hsa-mir-122

GO

What's new?

July 21, 2010
miRStart is online for evaluation...

miRStart, a novel resource of human microRNA TSSs (transcription start sites), systematically incorporates significant datasets derived from TSS-relevant experiments to identify transcription start sites of microRNAs. The distribution patterns of these experimental features within 50 k upstream region of microRNA precursors provides an insight into determining reliable microRNA TSSs. In general, a high-confidence TSS is recommended for each microRNA based on a SVM training model. However, users can customize their preferable microRNA TSSs according to the straightforward display of experimental TSS signals.

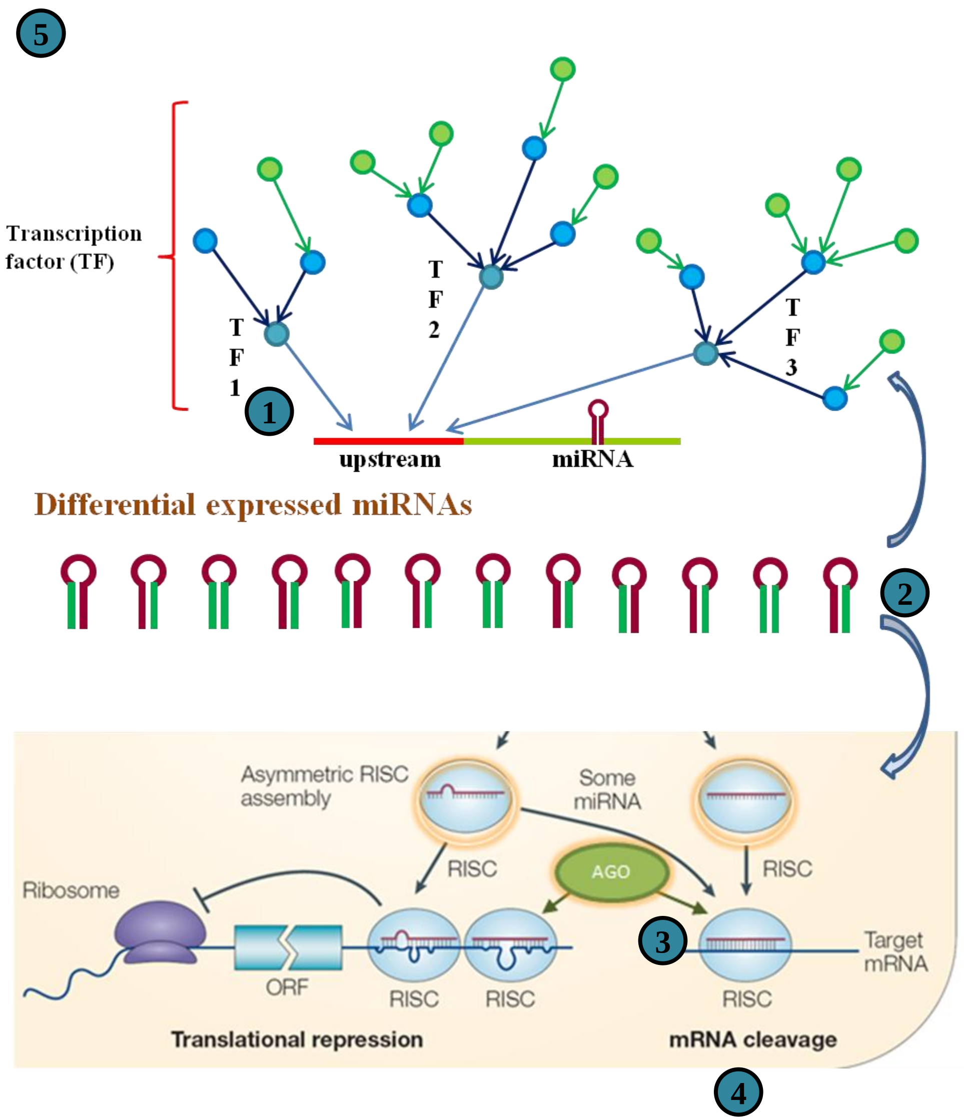
Published experimental evidences used in miRStart are described as follows:

(1) CAGE (Cap Analysis of Gene Expression) tags: Recognize 5'-end of a gene
(2) TSS Seq tags: More than 300 million 5'-end sequences of human and mouse cDNAs by combining oligo-capping method and Solexa sequencing technology
(3) H3K4me3 enrichment (histone H3 is trimethylated at its lysine 4 residue): Enriched surrounding TSSs

Moreover, ESTs and comparative genomics around putative miRNA TSSs are used to provide strong supports for reconfirmation.

The following figure demonstrates the concept of miRStart.

Using NGS for analyzing small non-coding RNAs



① **Identifying TFs which regulate miRNAs**

Chip (chromatin immunoprecipitation**)-seq**

② **Profiling miRNA expression**

Small RNA NGS sequencing

③ **Identifying miRNA target interaction**

CLIP (crosslinking immunoprecipitation**)-seq**

Degradome-seq

④ **mRNA expression**

RNA-seq

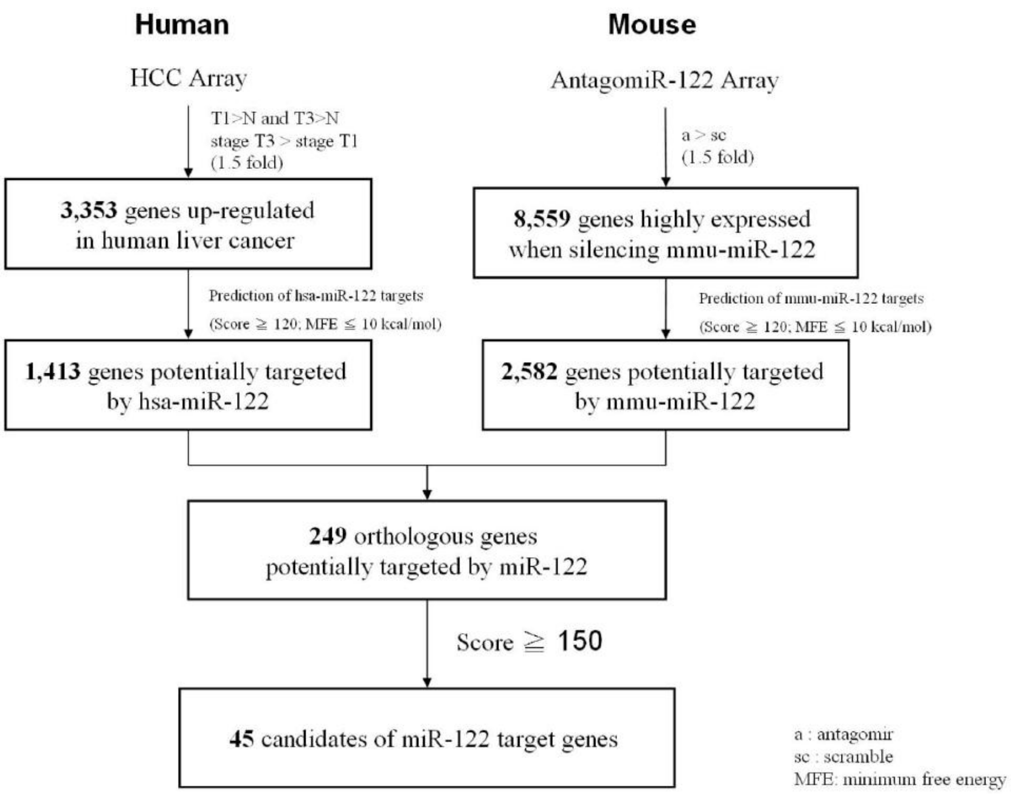
⑤ **DNA methylation**

BS (bisulfate**)-**seq
MeDIP (methylated DNA immunoprecipitation **)-**seq
MBD (methyl-binding protein **)-**seq

An example of
utilizing our miRNA databases and tools

Investigation roles of miR-122a in human hepatocellular carcinoma (HCC)

W.C. Tsai *et al.* (2009) "**MicoRNA-122, a tumor suppressor microRNA that regulates intra-hepatic metastasis of hepatocellular carcinoma**" ***Hepatology,*** Vol. 49, No. 5, pp. 1571-1582. (SCI IF=10.885, Rank = 2/72, GASTROENTEROLOGY & HEPATOLOGY)



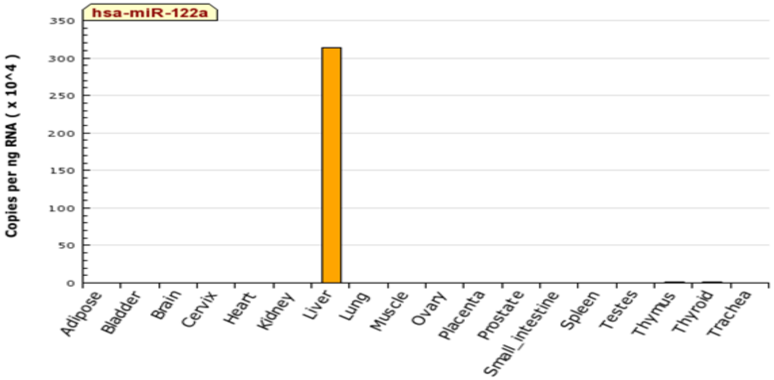
Collaborations with Prof. A.P. Tsou, National Yang Ming University

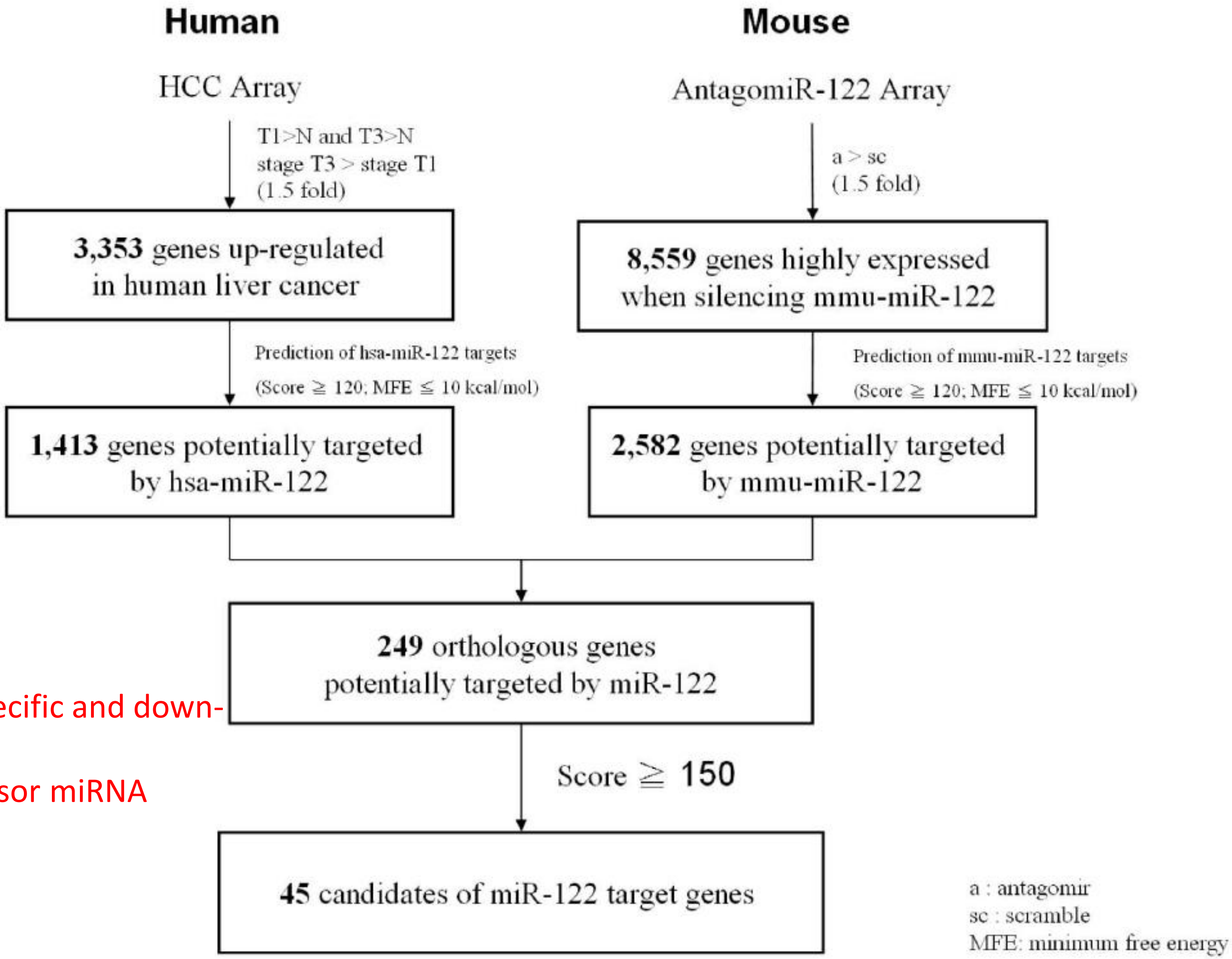
SCI Citations ~ 80

Highly-cited paper selected by Essential Science Indicator

Identifying miR-122 targets in human hepatocellular carcinoma

- miR-122 is liver-specific and down-regulated in HCC
- As a tumor suppressor microRNA
- Genes, which are potentially down-regulated by miR-122, were derived from HCC microarray and AntigomiR-122 microarray experiments
  - Hepatocellular carcinoma (HCC) microarray - 3353 up-regulated genes in liver cancer. (1.5 fold change) human
  - AntigomiR-122 microarray – 8859 genes up-regulated genes when silencing mmu-mir-122. (1.5 fold change) mouse





Human

HCC Array

T1>N and T3>N
stage T3 > stage T1
(1.5 fold)

3,353 genes up-regulated in human liver cancer

Prediction of hsa-miR-122 targets
(Score ≧ 120; MFE ≦ 10 kcal/mol)

1,413 genes potentially targeted by hsa-miR-122

Mouse

AntagomiR-122 Array

a > sc
(1.5 fold)

8,559 genes highly expressed when silencing mmu-miR-122

Prediction of mmu-miR-122 targets
(Score ≧ 120; MFE ≦ 10 kcal/mol)

2,582 genes potentially targeted by mmu-miR-122

249 orthologous genes potentially targeted by miR-122

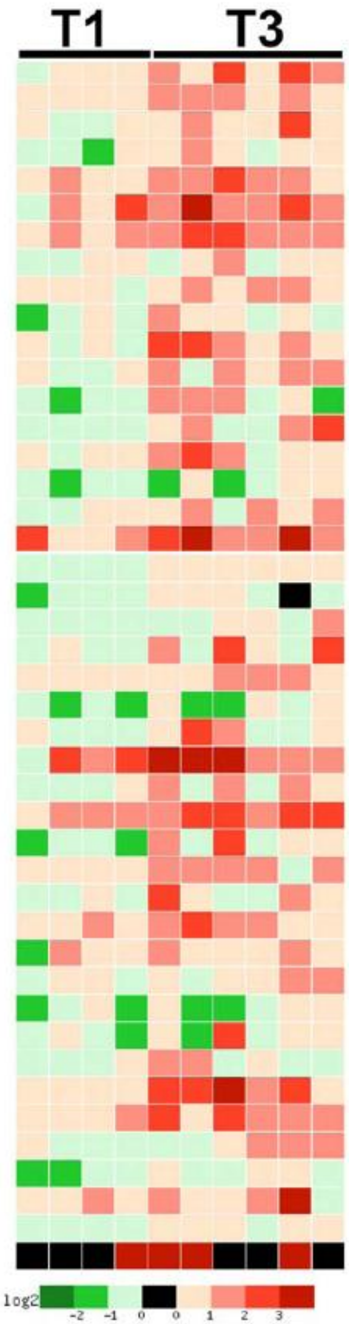
Score ≧ 150

45 candidates of miR-122 target genes

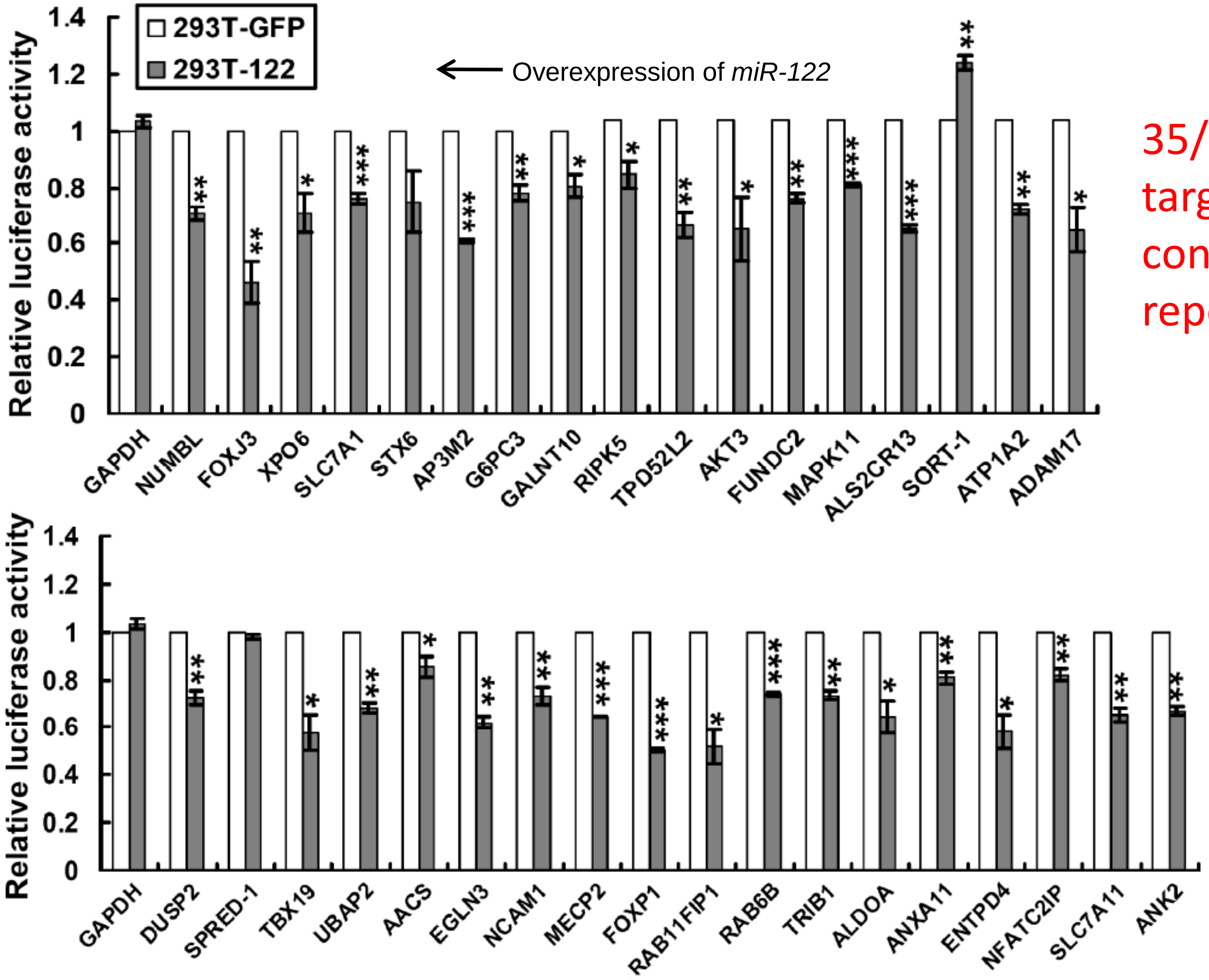
a : antagomir
sc : scramble
MFE: minimum free energy

- miR-122 is liver-specific and down-regulated in HCC
- As a tumor suppressor miRNA

Candidate target genes of *miR-122*. Relative gene expression detected in the microarray assays is shown as mean ratios of T3/T1 for HCC samples and an/c for mice treated with antagomiR-122 (an) compared to control mm-antagomiR122 (c).



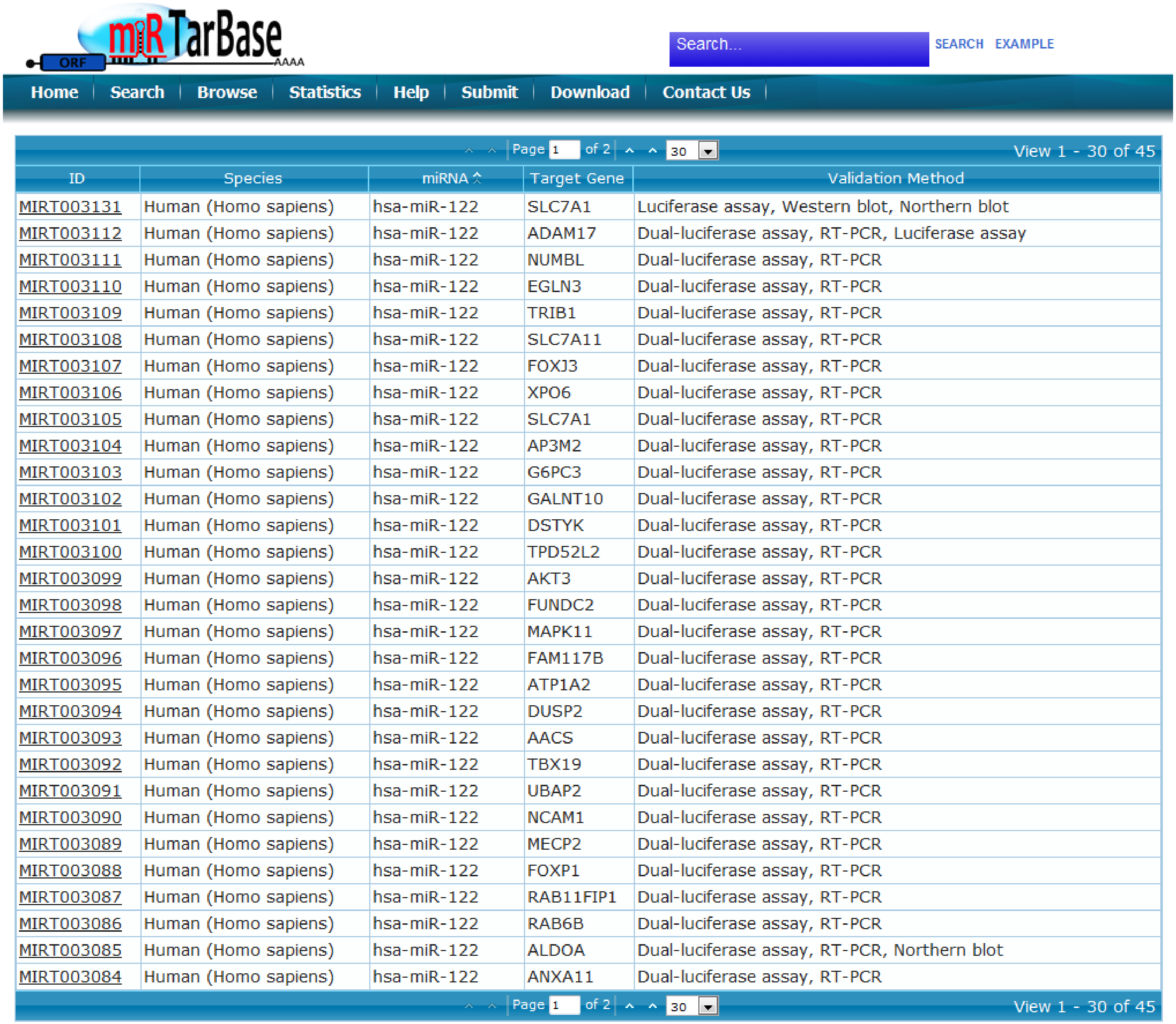
	Human			Mouse				
Gene	T3/T1	Sc-H	MFE-H	an/c	Sc-H	MFE-M	Molecular Function	# Binding sites
NUMBL	3.1	181	-20.5	4.5	131	-14.5	Protein binding	7
FOXJ3	1.6	176	-21	1.7	133	-10.6	Transcription factor	3
XPO6	2.2	174	-22.3	1.7	125	-16.4	Nuclear protein transport	2
SLC7A1	2.6	173	-22.4	1.7	173	-23.6	Amino acid transport	4
STX6	1.6	169	-20.9	2.5	135	-17.2	Protein transport	3
AP3M2	1.6	168	-19.9	3.0	129	-15.9	Protein trafficking	3
G6PC3	2.1	167	-18.4	6.1	158	-13.1	Glucose-6-phospgatase	1
GALNT10	1.6	167	-20.4	1.9	152	-20.5	Calcium binding	6
ARHGAP19	1.7	166	-21.5	1.7	143	-17.2	GTPase	4
RIPK5	1.8	166	-18	1.6	150	-16.6	Kinase activity	4
TPD52L2	3.0	166	-23.8	16.3	140	-16.5	Cell proliferation	6
AKT3	2.0	165	-18.7	2.1	162	-15.6	Cell proliferation, apoptosis	2
FUNDC2	2.5	165	-18	2.1	161	-15.8	HCV core binding protein	4
MAPK11	3.2	165	-20.2	1.6	123	-19.1	MAPK activity	2
ALS2CR13	2.0	162	-17.3	6.7	163	-17	Unknown	3
BACH2	1.8	162	-15	2.7	140	-13.5	Transcription factor	3
ATP11A	2.0	161	-12.9	1.9	162	-18.1	Transport ions	7
SORT1	2.2	161	-16.2	2.0	154	-15.3	Cell differentiation	2
ATP1A2	3.5	160	-17	1.5	132	-11.7	Ion concentration balance	7
ADAM17	1.7	159	-19	2.0	122	-17	Cell cell interaction	1
DUSP2	1.8	159	-13.4	1.6	123	-11.1	MAPK phosphatase	2
OSMR	2.5	159	-15.8	2.9	142	-15.5	Cell proliferation	4
RABIF	1.7	159	-20.1	2.4	139	-15.1	Small GPT regulator activity	2
PALM	2.3	156	-15.9	16.1	161	-18.7	Cell mobility and cell shape	3
SPRED1	2.2	156	-21.2	1.6	145	-19.8	Activate MAPK kinase	1
AACS	1.8	155	-13.8	1.8	163	-17.8	isoprenoid biogenesis	2
TBX19	2.1	155	-17.2	4.3	131	-13.1	Transcription factor	5
UBAP2	2.2	155	-18.8	2.7	131	-23.3	Ubiquitin associated protein	2
EGLN3	3.9	154	-17	1.6	137	-15.7	Apoptosis	1
NCAM1	1.9	154	-15.2	3.8	154	-15.2	Cell differentiation	4
MECP2	2.2	153	-12.1	3.1	160	-26.2	Transcription	6
CS	1.6	152	-12.6	1.8	148	-11.9	Catalyze synthesis of citrate	3
FOXP1	1.6	152	-10.8	1.8	167	-17.7	Transcription factor	1
RAB11FIP1	1.5	152	-20.1	1.9	133	-13.3	Protein transport	4
RAB6B	1.5	152	-13.3	1.6	164	-19.6	GTPase	5
TRIB1	2.1	152	-17.4	1.8	134	-12.8	Kinase activity	4
TTYH3	1.8	152	-17.1	1.9	160	-16.6	Chlorode anion channel	2
ALDOA	3.7	151	-13.2	2.2	157	-16.7	Amino acid transport	2
ANXA11	2.0	151	-18.6	4.3	139	-17.2	Calcium binding	7
CLDN18	1.7	151	-18.5	1.9	130	-15.1	Cell and cell adhesion	4
ENTPD4	2.0	151	-18.5	2.0	131	-13.5	Calcium binding	2
NFATC2IP	2.7	151	-19.9	2.2	135	-15.2	Protein modification	4
ANK2	2.1	150	-12	3.1	137	-17.9	Cell proliferation	5
MEP1A	4.3	150	-17.8	1.6	132	-12.8	Peptidase	3
NFATC1	1.8	150	-13.1	2.3	148	-14.9	Transcription factor	3
SLC7A11	9.5	150	-19.5	1.6	164	-14.4	Amino acid transport	4



← Overexpression of *miR-122*

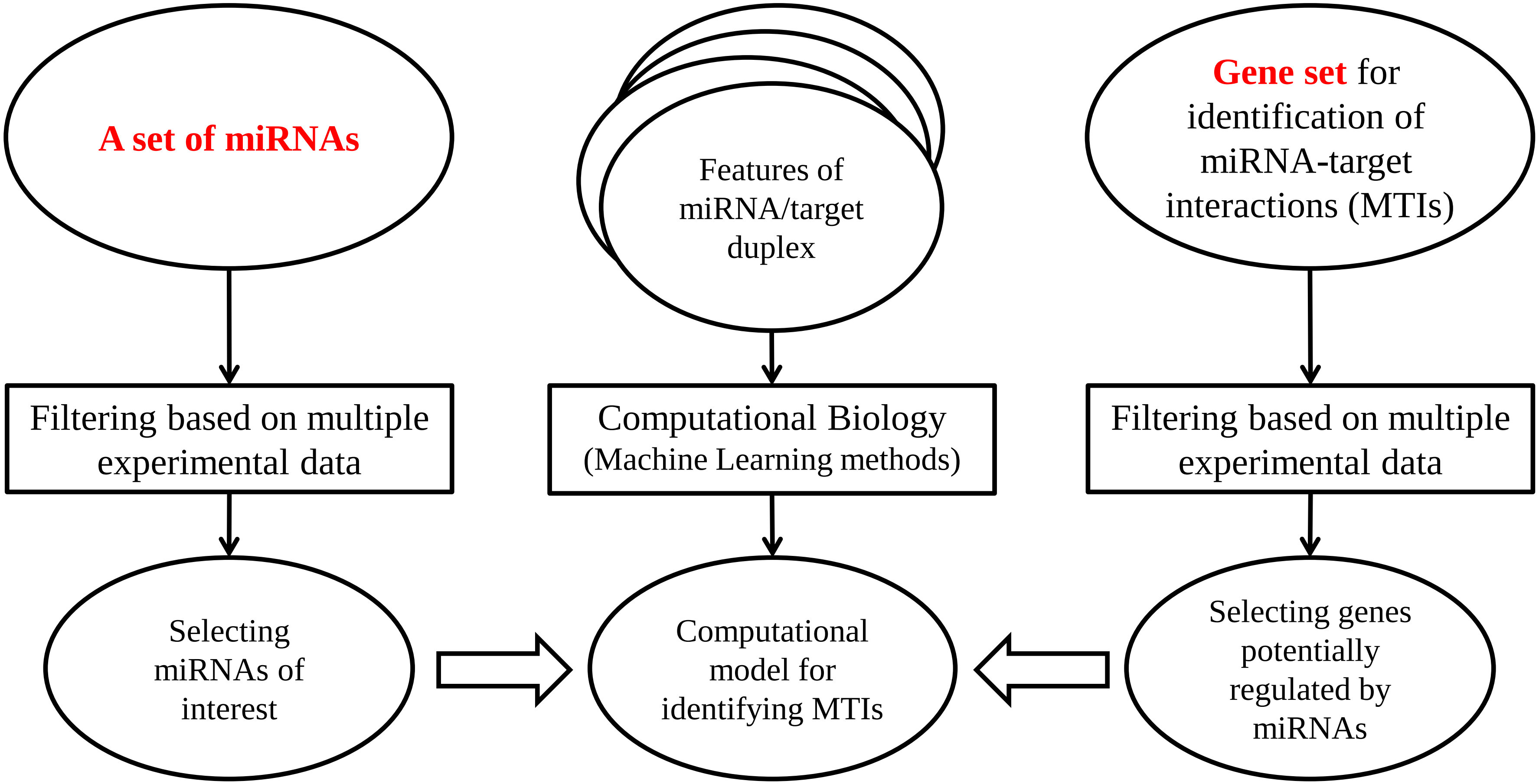
35/45 (78%) miR-122 target genes were confirmed by luciferase reporter assay.

Confirmation of *miR-122* target genes. *A,* The 3’UTR reporter assay. Putative *miR-122* binding sites in the 3’UTR of the predicted target genes were sublconed downstream of the luciferase gene in pGL3-Control vector. Target reporter assay was carried out in 293T cells overexpressing *miR-122* from lenti-122 virus infection (293T-122).



ID	Species	miRNA	Target Gene	Validation Method
MIRT003131	Human (Homo sapiens)	hsa-miR-122	SLC7A1	Luciferase assay, Western blot, Northern blot
MIRT003112	Human (Homo sapiens)	hsa-miR-122	ADAM17	Dual-luciferase assay, RT-PCR, Luciferase assay
MIRT003111	Human (Homo sapiens)	hsa-miR-122	NUMBL	Dual-luciferase assay, RT-PCR
MIRT003110	Human (Homo sapiens)	hsa-miR-122	EGLN3	Dual-luciferase assay, RT-PCR
MIRT003109	Human (Homo sapiens)	hsa-miR-122	TRIB1	Dual-luciferase assay, RT-PCR
MIRT003108	Human (Homo sapiens)	hsa-miR-122	SLC7A11	Dual-luciferase assay, RT-PCR
MIRT003107	Human (Homo sapiens)	hsa-miR-122	FOXJ3	Dual-luciferase assay, RT-PCR
MIRT003106	Human (Homo sapiens)	hsa-miR-122	XPO6	Dual-luciferase assay, RT-PCR
MIRT003105	Human (Homo sapiens)	hsa-miR-122	SLC7A1	Dual-luciferase assay, RT-PCR
MIRT003104	Human (Homo sapiens)	hsa-miR-122	AP3M2	Dual-luciferase assay, RT-PCR
MIRT003103	Human (Homo sapiens)	hsa-miR-122	G6PC3	Dual-luciferase assay, RT-PCR
MIRT003102	Human (Homo sapiens)	hsa-miR-122	GALNT10	Dual-luciferase assay, RT-PCR
MIRT003101	Human (Homo sapiens)	hsa-miR-122	DSTYK	Dual-luciferase assay, RT-PCR
MIRT003100	Human (Homo sapiens)	hsa-miR-122	TPD52L2	Dual-luciferase assay, RT-PCR
MIRT003099	Human (Homo sapiens)	hsa-miR-122	AKT3	Dual-luciferase assay, RT-PCR
MIRT003098	Human (Homo sapiens)	hsa-miR-122	FUNDC2	Dual-luciferase assay, RT-PCR
MIRT003097	Human (Homo sapiens)	hsa-miR-122	MAPK11	Dual-luciferase assay, RT-PCR
MIRT003096	Human (Homo sapiens)	hsa-miR-122	FAM117B	Dual-luciferase assay, RT-PCR
MIRT003095	Human (Homo sapiens)	hsa-miR-122	ATP1A2	Dual-luciferase assay, RT-PCR
MIRT003094	Human (Homo sapiens)	hsa-miR-122	DUSP2	Dual-luciferase assay, RT-PCR
MIRT003093	Human (Homo sapiens)	hsa-miR-122	AACS	Dual-luciferase assay, RT-PCR
MIRT003092	Human (Homo sapiens)	hsa-miR-122	TBX19	Dual-luciferase assay, RT-PCR
MIRT003091	Human (Homo sapiens)	hsa-miR-122	UBAP2	Dual-luciferase assay, RT-PCR
MIRT003090	Human (Homo sapiens)	hsa-miR-122	NCAM1	Dual-luciferase assay, RT-PCR
MIRT003089	Human (Homo sapiens)	hsa-miR-122	MECP2	Dual-luciferase assay, RT-PCR
MIRT003088	Human (Homo sapiens)	hsa-miR-122	FOXP1	Dual-luciferase assay, RT-PCR
MIRT003087	Human (Homo sapiens)	hsa-miR-122	RAB11FIP1	Dual-luciferase assay, RT-PCR
MIRT003086	Human (Homo sapiens)	hsa-miR-122	RAB6B	Dual-luciferase assay, RT-PCR
MIRT003085	Human (Homo sapiens)	hsa-miR-122	ALDOA	Dual-luciferase assay, RT-PCR, Northern blot
MIRT003084	Human (Homo sapiens)	hsa-miR-122	ANXA11	Dual-luciferase assay, RT-PCR

Strategies to achieve both high sensitivity and high specificity



A set of miRNAs

Filtering based on multiple experimental data

Selecting miRNAs of interest

Features of miRNA/target duplex

Computational Biology (Machine Learning methods)

Computational model for identifying MTIs

Gene set for identification of miRNA-target interactions (MTIs)

Filtering based on multiple experimental data

Selecting genes potentially regulated by miRNAs

Ways to achieve higher prediction specificity (lower false positive rate)

- More specific miRNAs
- More experimental data
- Selecting miRNAs by biologists' knowledge

- Species dependent
- More positive samples
- More reasonable negative samples
- More effective biological features or properties

- More experimental data
- Multi-dimensional data
- Multiple screening criteria
- Selecting genes by biologists' knowledge
- Smaller gene set

http://www.cc.nctu.edu.tw/~bryan

Or Google Search by “Hsien-Da Huang”

Thank you for your attention!!