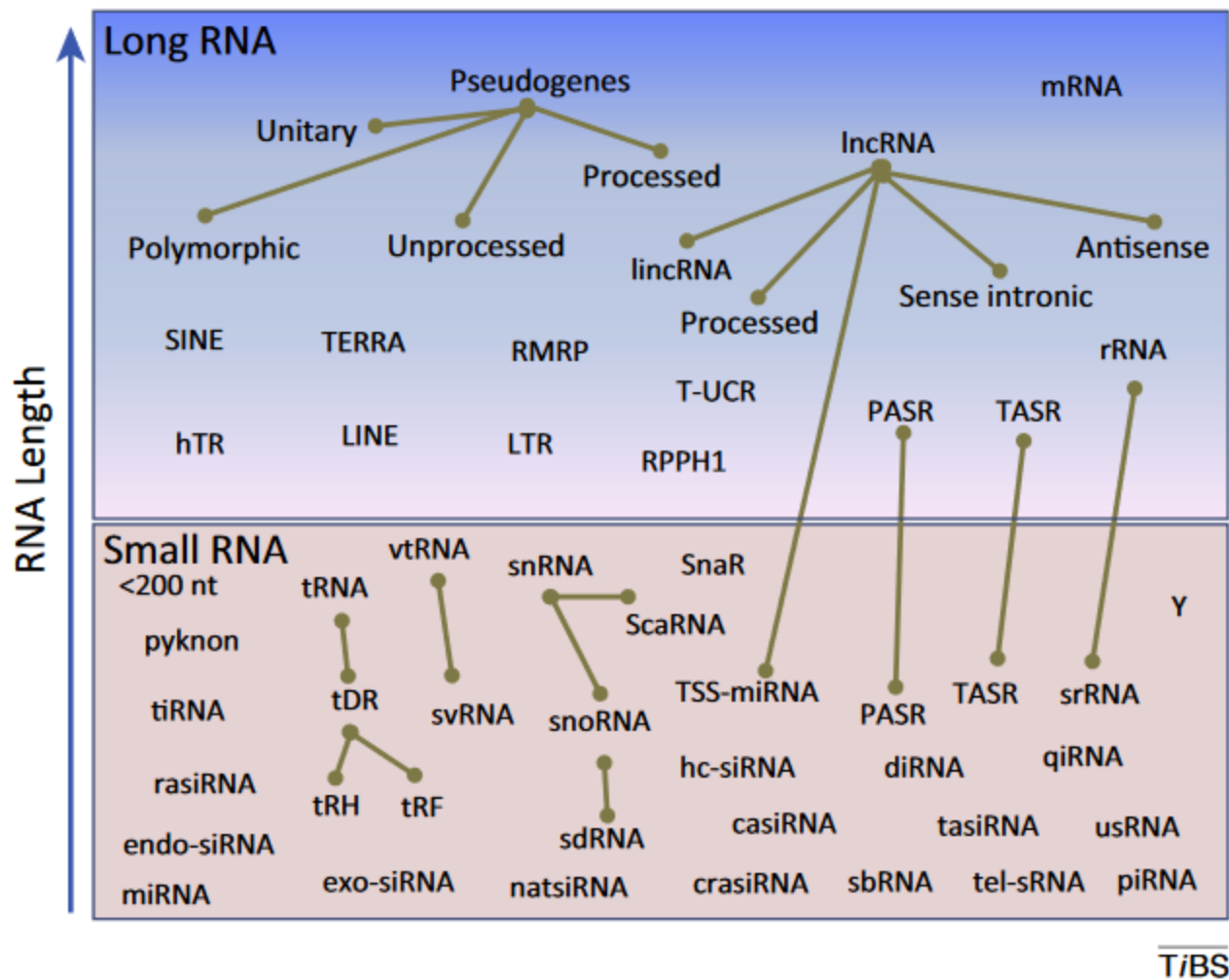


Tuesday – part 6

Small regulatory RNAs

Michał Szcześniak, PhD

Faculty of Biology, Adam Mickiewicz University, Poznań
ideas4biology Ltd.



Source: Vickers KC, Roteta LA, Hucheson-Dilks H, Han L, Guo Y. Mining diverse small RNA species in the deep transcriptome. Trends Biochem Sci. 2015 Jan;40(1):4-7.

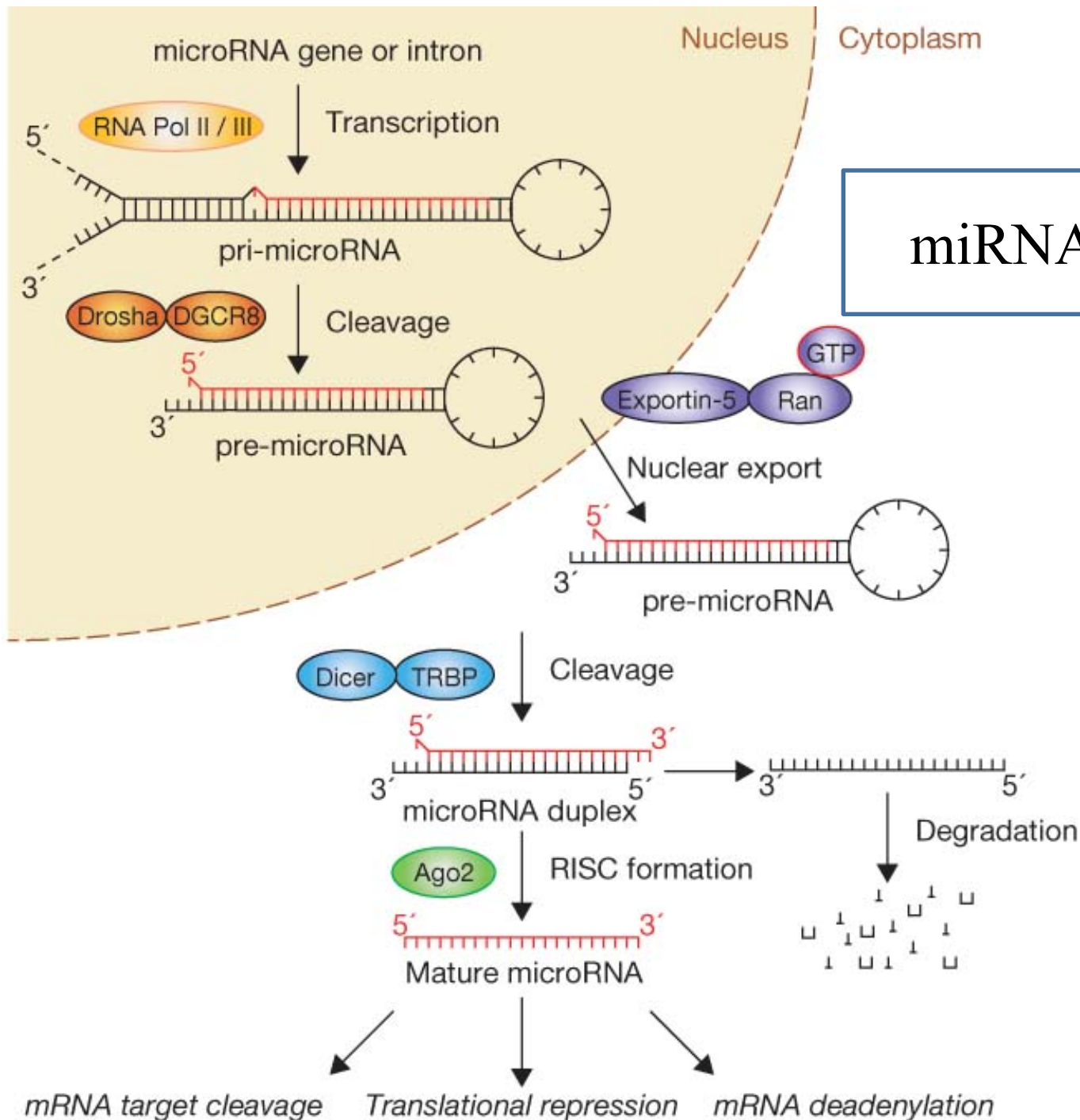
sRNAs originating from longer, noncoding precursors

Abbreviation	Name	Class	Size (nt)	Function	Database (website)
miRNA	microRNA	miRNA	19–22	Post-transcriptional gene repression	miRBase.org
TSS-miRNA	Transcriptional start-site-miRNA	miRNA	20–30	Post-transcriptional gene regulation	N/A
moRNA	miRNA-offset RNAs	miRNA	19–22	N/A	N/A
usRNA	Unusually small RNA	miRNA	15–17	Post-transcriptional gene regulation	N/A
endo-siRNA	Endogenous small interfering RNA	siRNA	21	Somatic inhibitor of retrotransposition, post-transcriptional gene repression	N/A
exo-siRNA	Exogenous small interfering RNA	siRNA	21	Gene-targeted silencing, anti-viral	N/A
natsiRNA	Natural antisense siRNA	siRNA	21–24	Pathogen resistance, post-transcriptional gene regulation	bis.zju.edu.cn/pnatdb
casRNA	Cis-acting siRNA	siRNA	24	Transposon methylation, chromatin modification	N/A
tasiRNA	Trans-acting siRNA	siRNA	21	Post-transcriptional gene repression	bioinfo.jit.edu.cn
rasiRNA	Repeat-associated siRNA	siRNA	26–31	Transposon methylation, chromatin modification	deepbase.sysu.edu.cn
hc-siRNA	Heterochromatic small interfering RNA	siRNA	24	Genome maintenance, DNA methylation	N/A
3' U tRF	tRNA-derived fragment	tDR	20	Post-transcriptional gene repression	genome.ucsc.edu
5' tRF	tRNA-derived fragment	tDR	20	Post-transcriptional gene repression, translational repression	genome.ucsc.edu
3'CCA tRF	tRNA-derived fragment	tDR	23	Post-transcriptional gene repression, RNA metabolism	genome.ucsc.edu
tRH	tRNA-derived half	tDR	30–35	Protein synthesis repression, post-transcriptional gene repression	genome.ucsc.edu

Source: Vickers KC, Roteta LA, Hucheson-Dilks H, Han L, Guo Y. Mining diverse small RNA species in the deep transcriptome. Trends Biochem Sci. 2015 Jan;40(1):4-7.

sRNAs originating from longer, noncoding precursors

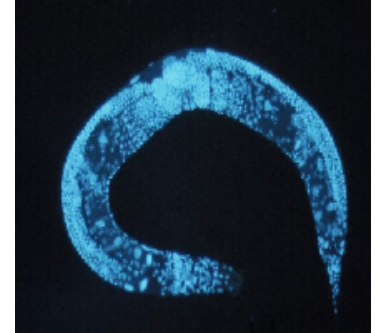
Abbreviation	Name	Class	Size (nt)	Function	Database (website)
piRNA	Piwi-interacting RNA	piRNA	25–33	Germline post-transcription gene repression, transposon regulation, chromatin modification	pirnabank.ibab.ac.in
crasiRNA	Centromere repeat-associated small RNA	chrRNA	34–42	Cell cycle, centromere formation	N/A
tel-sRNAs	Telomere-specific small RNA	chrRNA	24	Epigenetic regulation	N/A
PASR	Promoter-associated small RNA	CAsRNA	20–200	Transcriptional regulation of mRNAs, epigenetic regulation	deepbase.sysu.edu.cn
tiRNA	Transcription initiation smRNA	CAsRNA	18	Transcriptional regulation of mRNAs, epigenetic regulation	N/A
TSSaRNA	Transcription start site-associated RNAs	CAsRNA	20–90	Transcriptional regulation of mRNAs, epigenetic regulation	N/A
spliRNA	Splice site-associated small RNA	CAsRNA	17–18	N/A	N/A
H/ACA sdRNA	H/ACA snoRNA-derived RNA	snRNA	20–24	Post-transcriptional gene repression, alternative splicing	ensembl.org
C/D sdRNA	C/D snoRNA-derived RNA	snRNA	17–19, 30	Post-transcriptional gene repression, alternative splicing	ensembl.org
svRNA	Small vault RNAs	vRNA	22–37	Post-transcriptional gene repression	ensembl.org
srRNA	smRNAs-derived from rRNA	rRNA	24	Post-transcriptional gene regulation	ensembl.org
yDR	Y RNA-derived small RNAs	Y RNA	24–25, 30	N/A	ensembl.org
sbRNA	Stem-bulge RNA	Y RNA	67–133	RNA quality control, chromosomal replication	N/A
diRNA	Double-strand break-induced small RNAs	DSB	21	DNA double-strand break repair	N/A
qiRNA	QDE-2-interacting small RNAs	DSB	20–21	DNA double-strand break repair, protein translation	N/A



Julia Winter *et al.*, 2009

microRNAs participate in:

- morphogenesis
- development
- stress response
- signalling
- apoptosis
- and many more processes...



miRNAs have been implicated in a number of diseases (miR2Disease):

- 349 miRNA
- 163 choroby

miR2Disease
Base

Evolution of miRNAs

- ✓ Independently evolved in plants and animals
- ✓ Lost in some groups (*Placozoa*)
- ✓ Small number of miRNAs in some groups (*Cnidaria*, *Porifera*)
- ✓ Origin of novel miRNAs (*de novo*)
- ✓ Duplication of miRNA genes



siRNAs in plants

casRNA

Transcribed from transposons and tandem repeats.

Typically 24 nt long.

Transcribed with Pol IV.

Function: DNA methylation and histone modifications

tasiRNA

Secondary siRNA: other small RNAs required.

Cut in phase, 21 nt long.

nat-siRNA

Activated in stress response.

Typically 21 or 24 nt long.

Transcribed with Pol IV.

lsiRNA

Activated in stress response.

30-40 nt long.

Transcribed with Pol IV.



siRNA in animals

Transcribed from transposons (e.g. retrotransposon L1 in animals), intergenic regions, and long RNAs, including mRNAs.

L1 can be transcribed bidirectionally, generating two overlapping transcripts, able to interact and give rise to siRNA molecules.

Typically 21 nt long.

piRNA

Expressed in germline (spermatocytes). They provide stability for the germline by silencing the transposones, such as L1.

Pre-pachytene: silencing the transposons

Pachytene: functions unknown

Pachytene: a phase of meiosis, during which *crossing-over* takes place.

Databases of small RNAs

miRNA sequences

Name	Species (number)	Type of data	Methods and source of data	PMID
miRBase	Animals, plants, viruses (206)	Published data	Literature, user submissions	20205188
PMRD	Plants (123)	In silico miRNA predictions, expression and target data	Literature, microarray experiments	19808935
microPC	Plants (125)	In silico miRNA predictions	Algorithm for miRNA search in ESTs	19660144
miROrtho	Animals (46)	In silico miRNA predictions	Programs: R-COFFEE, RNAplfold, RNAalifold	18927110
Vir-Mir db	Viruses (1491)	In silico miRNA predictions	Program: Srnalloop; NCBI	17702763
miRNAMap	Animals (13)	Experimental miRNAs and their targets	Programs: miRanda, RNAhybrid, TargetScan; qPCR experiments	16381831
GrapeMiRNA	Vitis vinifera	In silico miRNA predictions	Program: FindMiRNA	19563653
miRNEST	Animals, plants, viruses (544)	In silico miRNA predictions and experimental miRNAs, targets, polymorphism and expression regulation	Literature, algorithm for miRNA search in ESTs and from RNA-Seq data, 13 databases	22135287

miRNA targets

Name	Species(number)	Type of data	Methods and source of data	PMID
miRWalk (previously: Argonaute)	Human, mouse, rat	In silico and experimental targets	Databases: GenBank, Ensembl, miRBase, programs: DIANA-microT, miRanda, miRDB, PicTar, PITA, RNA22, TargetScan/TargetScanS, miRWalk	21605702
HOCTAR	Human	mRNA targets	Programs: miRanda, TargetScan, PicTar.	21435384
RepTar	Human, mouse	In silico mRNA targets	A new algorithm that allows for more than one target site for a miRNA in a given 3'UTR	21149264
miRTarBase	Plants, animals,, viruses (14)	mRNA targets	Literature	21071411
miRGator	Human, mouse	mRNA targets and their expression associated with diseases	Databases: PhenomiR, GEO, ArrayExpress, programy: targetScan, PITA, miRanda, miRbridge	21062822
starBase	Human, mouse, C. elegans, A. thaliana, rice, Vitis vinifera	mRNA targets	CLIP-Seq and Degradome-Seq	21037263
miRSEL	Human, mouse, rat	mRNA targets	Databases: HGNC, MGD, Entrez Gene, Swiss-Prot Protein Database, miRGen, miRBase	20233441
miRecords	Animals (9)	mRNA targets	Literature, programs: DIANA-microT, MicroInspector, miRanda, miTarget, MirTarget2, NbmirTar, PicTar, PITA, RNA 22, RNA Hybrid, TargetScan/TargetScanS	18996891
TarBase	Animals (6)	Experimental mRNA targets	Literature	18957447
miRDB	Human, mouse, rat, dog, chicken	mRNA targets and functional annotation of miRNAs	Databases miRBase, new algorithm for target search	18426918
MicroRNA.org	Human, mouse, rat, fruitfly, C. elegans	mRNA targets and expression of miRNAs	Literature, miRanda, databases: miRBase, UCSC	18158296
MiRonTop	Human, mouse, rat	mRNA targets	Databases: miRBase, NCBI, programs: Targetscan, MicroCosm Targets, Miranda, PicTar	20959382

Expression of miRNAs(1)

Name	Species(number)	Type of data	Methods and source of data	PMID
CIRCUITSdb	Human, mouse	Regulation of miRNA expression by transcription factors	Literature, databases: TransmiR, TarBase, Myc Target Gene	20731828
mESAdb	Human, mouse, Danio rerio	Expression of miRNAs and their mRNA targets	Databases: Ensembl, miRBase, microCosm, HUGE, KEGG, GO	21177657
miRNeye	Mouse	Expression of miRNAs in mouse eye	Experiments: in situ RNA hybridization with use of LNA	21171988
dbDEMC	Human	Expression of miRNAs in cancer tissues	Literature	21143814
miReg	Human	Regulation of miRNA expression	Literature	20693604
PuTmiR	Human	Regulation of miRNA expression by transcription factors	Databases: miRBase, UCSC	20398296
S-MED	Human	Expression of miRNAs in sarcoma	Experiments with the use of BeadArrays system	20212452
PhenomiR	Human	Expression of miRNAs in diseases and biological processes	Literature, databases: OMIM Morbid Map, Gene Ontology, BRENDA Tissue Ontology	20089154
miRGen	Animals (11)	Regulation of miRNA expression, polymorphsm, mRNA targets	Literature, program mathTM tool (identification of TFBSs), databases: mammalian miRNA expression atlas, UCSC, dbSNP	19850714
TransmiR	Animals	Regulation of miRNA expression by transcription factors	Literature, database: UCbase & miRfunc	19786497

Expression of miRNAs (2)

Name	Species(number)	Type of data	Methods and source of data	PMID
miR2Disease	Human	Expression of miRNAs in diseases	Literature, database: TarBase	18927107
GenomeTraFaC	Human, mouse	Regulation of miRNA expression by transcription factors	Databases: Homologene, NCBI, MGI, miRBase	17178752
miSolRNA	Tomato, A. thaliana	Expression of miRNAs and their functions in biological processes	Literature	21059227
MirZ (previously: mammalian miRNA expression atlas)	Human, mouse, rat	Expression of miRNAs	NGS sequencing	17604727
mirEX	A. thaliana	Expression of miRNAs	Real-time PCR	22013167
mimiRNA	Human	Expression of miRNAs	Literature, programs: TargetScan, RNA22, PicTar, algorytm ExParser, databases: Hypertext cell line database, mammalian miRNA expression atlas, GEO	19933167
mirConnX	Human, mouse	Regulation of miRNA expression	Databases: TarBase, miRBase, DBTSS, UCSC, The Eukaryotic Promoter Database, programs: CoreBoost_HM, PITA, miRANDA, TargetScan, RNAhybrid, Pictar	21558324
PMTED	Plants (9)	microRNA target expression	Manually curated microarray data	23725466
TUMIR	Human	microRNA deregulation in various cancers	Literature	23594715

Polymorphism

Name	Species(number)	Type of data	Methods and source of data	PMID
miRvar	Human	Polymorphism and its functional consequences	Literature, databases: SNPdb, UCSC Genome Browser, miRBase, programs: PHDcleav, RISCbinder	21618345
Patrocles	Animals (7)	Polymorphism in miRNAs and their targets	Literature, databases: miRBase, Ensembl, program: RNAfold	19906729
PolymiRTS	Human, mouse	Polimorphism in target sequences	Databases: dbSNP, miRBase	17099235
dPORE-miRNA	Human	Polymorphism and regulation of miRNA expression	Databases: UCSC, PhenomiR, Tarbase, KEGG, program: BIOBASE MATCH	21326606
dbSMR	Human	Polymorphism in miRNAs	Databases: miRBase, Ensembl, programs: miRanda, RNAHybrid, TargetScan	19371411

Functions of miRNAs

Name	Species(number)	Type of data	Methods and source of data	PMID
UCbase & miRfunc	Human, mouse, rat	Functions of miRNAs; conservation of miRNA sequences	Databases: miRBase, UCSC, NCBI	18945703
miRNApath	Human, mouse, rat, chicken	Roles of miRNAs in methabolic pathways	Databases: miRBase, miRGen, miRGen, KEGG	18058708
miRò	Human	miRNA-phenotype associations	Databases: miRBase, mammalian miRNA expression atlas, miRecords, NCBI, GO, Genetic Association Database, programs: TargetScan, PicTar, miRanda	20157481
miREnvironment	Animals, plants (17)	miRNA-phenotype associations	Literature	21984757
miTALOS	Human, mouse	Roles of miRNAs in signalling pathways	Programs: TargetScan, TargetScan, PicTar, Pita, RNA22, bazy danych: KEGG, NCBI	21441347

Other databases

Name	Species(number)	Type of data	Methods and source of data	PMID
IntmiR	Human, mouse	Intronic miRNAs, their mRNA targets and deregulation in diseases	Not specified	21423893
CoGemiR	Animals (36)	Genomics and conservation of miRNA sequences	Databases: miRBase, Ensembl, SymAtlas, CoGemiR, program: miRNAMiner	18837977
AntagomirBase	Human	Antagomirs (molecules used for silencing of miRNAs)	Programs: Sfold, mfold	21904438
HNOCDDB	Human	miRNAs associated with head, neck and oral cancers	Literature	22024348
EpimiR	Human, mouse and five other species	Curated mutual regulation between miRNAs and epigenetic modifications	Literature	24682734
OncomiR	Human	Experimentally verified oncogenic and tumor-suppressive microRNAs	Literature	24651967
OncomiRdbB	Human	microRNAs and their targets in breast cancer	Program: miRanda, external databases	24428888
miR-SynthDB	Human	Design of multi-site multi-target synthetic miRNAs	A new algorithm	24627222
PASmiR	Plants (33)	miRNA molecular regulation in plant response to abiotic stress	Literature	23448274
miRCancer	Human	microRNA-cancer associations	Literature	23325619

miRBase

<http://mirbase.org/>



The screenshot shows the miRBase website interface. At the top, there is a dark blue header with the miRBase logo on the left, the text "miRBase" in the center, and the University of Manchester logo on the right. Below the header is a navigation bar with links: Home, Search, Browse, Help, Download, Blog, and Submit. A search box with a "Search" button is also present. The main content area is divided into several sections. On the left, there is a "Latest miRBase blog posts" section with a message "Error establishing a database connection". Below this is the "miRBase: the microRNA database" section, which states that miRBase provides the following services: a searchable database of published miRNA sequences and annotation, and a miRNA gene hunters service. On the right, there is a "miRNA count: 28645 entries" section with a "Release 21: June 2014" link. Below this is a "Search by miRNA name or keyword" section with a search box and "Go" and "Example" buttons. Further down is a "Download published miRNA data" section with "Download page" and "FTP site" links. At the bottom, there is a "References" section with a list of articles cited in the database.

miRBase

Home Search Browse Help Download Blog Submit

Latest miRBase blog posts
Error establishing a database connection

miRBase: the microRNA database

miRBase provides the following services:

- The [miRBase database](#) is a searchable database of published miRNA sequences and annotation. Each entry in the miRBase Sequence database represents a predicted hairpin portion of a miRNA transcript (termed mir in the database), with information on the location and sequence of the mature miRNA sequence (termed miR). Both hairpin and mature sequences are available for [searching](#) and [browsing](#), and entries can also be retrieved by name, keyword, references and annotation. All sequence and annotation data are also [available for download](#).
- The [miRBase Registry](#) provides miRNA gene hunters with unique names for novel miRNA genes prior to publication of results. Visit the [help pages](#) for more information about the naming service.

To receive email notification of data updates and feature changes please subscribe to the [miRBase announcements mailing list](#). Any queries about the website or naming service should be directed at mirbase@manchester.ac.uk.

miRBase is hosted and maintained in the [Faculty of Life Sciences](#) at the [University of Manchester](#) with funding from the [BBSRC](#), and was previously hosted and supported by the [Wellcome Trust Sanger Institute](#).

References

If you make use of the data presented here, please cite the following articles in addition to the primary data sources:

[miRBase: annotating high confidence microRNAs using deep sequencing data.](#)
Kozomara A, Griffiths-Jones S.
NAR 2014 42:D68-D73

[miRBase: integrating microRNA annotation and deep-sequencing data.](#)
Kozomara A, Griffiths-Jones S.
NAR 2011 39:D152-D157

[miRBase: tools for microRNA genomics.](#)
Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ.
NAR 2008 36:D154-D158

[miRBase: microRNA sequences, targets and gene nomenclature.](#)
Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ.
NAR 2006 34:D140-D144

miRNA count: 28645 entries
[Release 21: June 2014](#)

Search by miRNA name or keyword

Download published miRNA data
[Download page](#) | [FTP site](#)

miRBase

Stem-loop sequence hsa-mir-548j

Accession MI0006345

Symbol [HGNC:MIR548J](#)

Description Homo sapiens miR-548j stem-loop

Gene family MIPF0000317; [mir-548](#)

Stem-loop

```
--   agc   -aa   ua               u   c   g   ac
gggc  cagug  uagu  gcuggugcaaaguaau  gcggu  uuug  uauu  u
||||  |||||  |||  ||||| ||||| ||||| ||||| ||||| |||||
cuug  gucgc  auca  cgaccacguuuucaua  cguca  aaac  guga  u
ga    --a    aag   uc               -   a   g   cu
```

[Get sequence](#)

Deep sequencing [470](#) reads, [35.1](#) reads per million, 35 experiments

Coordinates (GRCh37.p5)

[chr22: 26951178-26951289 \[-\]](#)

Overlapping transcripts

sense [OTTHUMT00000320820](#); TPST2-001; intron 1
[OTTHUMT00000320822](#); TPST2-003; intron 1
[OTTHUMT00000320821](#); TPST2-002; intron 1
[OTTHUMT00000320825](#); TPST2-006; intron 1
[OTTHUMT00000320824](#); TPST2-005; intron 1
[OTTHUMT00000320826](#); TPST2-007; intron 1
[OTTHUMT00000320823](#); TPST2-004; intron 1
[ENST00000338754](#); TPST2-001; intron 1
[ENST00000398110](#); TPST2-003; intron 1

Genome context

miRBase

Deep sequencing reads for stem-loop sequence MI0006345

Stem-loop ID [hsa-mir-548j](#)

Reads

[illegible]

	Accession	Read count	Tissue	Link	Comment	Reference
✓	ER0000000001	4	embryonic stem cell	GEO : GSM541796		[Bar M et al.]
✓	ER0000000002	1	embryonic stem cell	GEO : GSM541797		[Bar M et al.]
✓	ER0000000103	4	melanoblast	GEO : GSM458535		[Stark MS et al.]
✓	ER0000000104	31	melanocyte	GEO : GSM458536		[Stark MS et al.]
✓	ER0000000106	6	melanoma	GEO : GSM458538	cell line: secondary acral melanoma	[Stark MS et al.]
✓	ER0000000107	62	melanoma	GEO : GSM458539	cell line: secondary mucosal melanoma	[Stark MS et al.]
✓	ER0000000108	380	melanoma	GEO : GSM458540	cell line: secondary cutaneous melanoma	[Stark MS et al.]

miRNEST

<http://mirnest.amu.edu.pl>

miRNEST 2.0 an integrative microRNA resource

[Home](#)[Browse miRNAs](#)[Deep-seq predictions](#)[Mirtrons](#)[miRNA genes](#)[Degradomes](#)[Download](#)[Upload](#)[Contact](#)

miRNEST is an integrative collection of animal, plant and virus microRNA data.

The database provides you with:

- a) microRNAs from our high-throughput predictions as well as from external databases
- b) predicted targets for plant candidates and experimental target support
- c) integrated data from 15 external databases, which includes e.g. sequences, polymorphism, expression, promoters.
- d) mirtrons, miRNA gene structures, degradome data and more!

miRNEST is being gradually developed to create an integrative resource of miRNA-associated data. The data comes from our computational predictions (new miRNAs, targets, mirtrons, miRNA gene structures) as well as from other databases and publications.

In current version, **miRNEST 2.0**, it contains miRNAs from 522 animal and plant species and 22 viruses.

What's new in miRNEST 2.0?

- i. [Animal and plant miRNAs identified using small RNA deep sequencing data](#)
- ii. [degradome support for plant miRNA targets](#)
- iii. [mirtrons](#)
- iv. [miRNA gene structures](#)
- v. HuntMi predictions for stored pre-miRNAs
- vi. [download page](#)
- vii. updated user interface

Read more ➔

miRNEST: tasks

Perform these four **independent** searches:

1. Find miR167 in *Zea mays*
2. Find all miRNAs from miRBase that have experimental support and show E-value $> 1e-5$ against UniProt.
3. Find the record *MNEST051503* and look into *deep sequencing data evidence*
4. What can be found under these tabs:
 - Deep-seq predictions
 - Mirtrons
 - miRNA genes
 - Degradomes

Other databases of sRNAs

piRNABank: <http://pirnabank.ibab.ac.in/index.shtml>

piRNA cluster database

<http://www.smallrnagroup.uni-mainz.de/piRNAclusterDB.html>

PlantNATsDB

<http://bis.zju.edu.cn/pnatdb/>

ta-siRNA Database

<http://bioinfo.jit.edu.cn/tasiRNADatabase/>

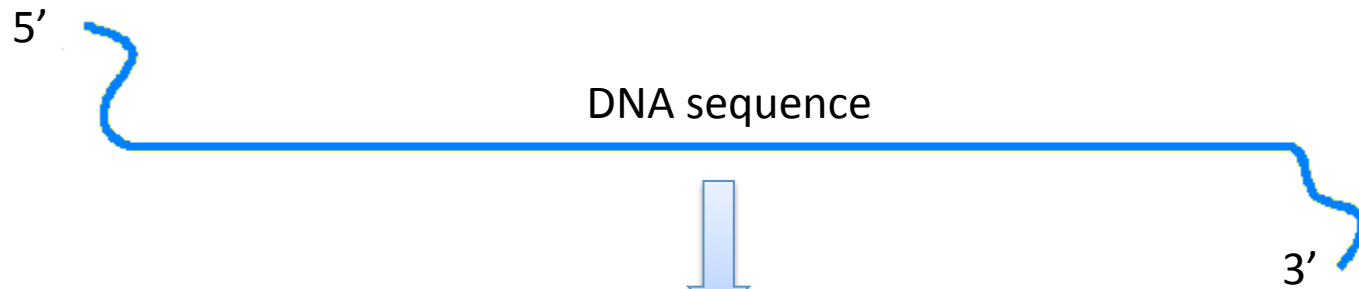
The MIT/ICBP siRNA Database

<http://web.mit.edu/sirna/>

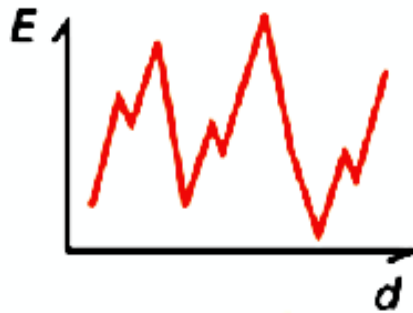
sRNATarBase 3.0

<http://ccb1.bmi.ac.cn/srnatarbase/>

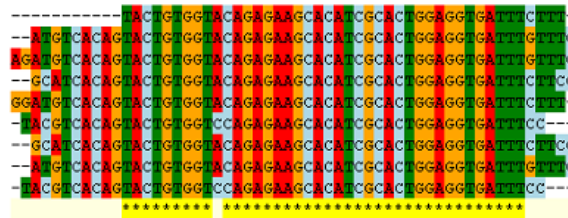
Identification of miRNAs and other sRNAs



Statistical features



Homology



Experiments

CTCCATGTACTTCTCCGTGCCACAGTACTGGGAACG

GTACTTCTCCGTGCCACAGTA 54.7

GTACTTCTCCGTGCCACAGTAC 27.3

TACTTCTCCGTGCCACAGTACT 355.3

TACTTCTCCGTGCCACAGTA 18939.6

TACTTCTCCGTGCCACA 82

TACTTCTCCGTGCCACAGT 573.9

TACTTCTCCGTGCCACAGTAC 11123.3

CTTCTCCGTGCCACAGTA 27.3



pre-miRNA

miRDeep2: preparing the data

Reads download: we already have it

```
wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR158/006/SRR1586016/SRR1586016.fastq.gz  
gunzip SRR1586016.fastq.gz
```

Study: microRNA sequencing of HEK293 and H9c2 cells

Sample: HEK293_microRNAs

Instrument: Illumina HiSeq 2000

Small RNA sequencing libraries were prepared using Illumina TruSeq Small RNA Sequencing kits

Downloading the genome: we already have it

```
wget  
ftp://ftp.ensembl.org/pub/release-88/fasta/homo\_sapiens/dna/  
Homo\_sapiens.GRCh38.dna.chromosome.22.fa.gz  
gunzip Homo_sapiens.GRCh38.dna.chromosome.22.fa.gz
```

Building a genome index: we already have it

```
bowtie-build Homo_sapiens.GRCh38.dna.chromosome.22.fa index/chr22
```

miRDeep2: analysis

```
mapper.pl SRR1586016.fastq -e -h -j -k TGGAATTCTCGGGTGCCAAGG -l 18 -m -p  
index/chr22 -s SRR1586016_collapsed.fa -t SRR1586016_collapsed_vs_genome.arf -  
v
```

- e: input data in FASTQ format
- h: convert FASTQ to FASTA
- j: discard reads with symbols other than a,c,g,t,u,n,A,C,G,T,U,N
- k: adapter to be removed
- l: min. Length of the read after adapter clipping
- m: remove sequence redundancy
- p: genome index
- s: the file with processed reads
- t: the file with mapping results
- v: verbose

miRDeep2: analysis

Downloading mature miRNAs from miRBase

```
wget ftp://mirbase.org/pub/mirbase/CURRENT/mature.fa.gz  
gunzip mature.fa.gz
```

Human mature miRNAs and pre-miRNAs from miRBase: **mature_mirbase_human.fa** and **premirna_mirbase_human.fa** (\r symbols have been removed and spaces are replaced with underscores)

Additionally, to avoid this error:

Error: *Genome file Homo_sapiens.GRCh38.dna.chromosome.22.fa has not allowed whitespaces in its first identifier*

the genome file needs to be edited:

```
nano Homo_sapiens.GRCh38.dna.chromosome.22.fa
```

Replace:

```
>22 dna:chromosome chromosome:GRCh38:22:1:50818468:1 REF  
with  
>22
```

Then, type: *CTRL+x*, then *Y*, then *Enter*.

This saves the changes.

miRDeep2: analysis

Identification of microRNAs

```
miRDeep2.pl SRR1586016_collapsed.fa Homo_sapiens.GRCh38.dna.chromosome.  
22.fa SRR1586016_collapsed_vs_genome.arf mature_mirbase_human.fa mature.fa  
premirna_mirbase_human.fa -t human 2> report.log
```

```
#Starting miRDeep2  
#testing input files  
#Quantitation of known miRNAs in data  
#parsing genome mappings  
#excising precursors  
#preparing signature  
#folding precursors  
#computing randfold p-values  
#running miRDeep core algorithm  
#running permuted controls  
#doing survey of accuracy  
#producing graphic results
```

novel miRNAs predicted by miRDeep2

[illegible]

Mapping sRNAs

Example analysis

Input data: SRR1586016_nonredundant.fasta

Let's map the reads to human pre-miRNAs.

Building the index

```
bowtie-build precursors.fasta index/precursors
```

Mapping

```
bowtie -t -p 2 -v 0 -a -f index/precursors SRR1586016_nonredundant.fasta  
SRR1586016.out
```

Visualizing

```
python alignment.py SRR1586016.out precursors.fasta # → SRR1586016.out.html
```


Chimira

Analysis of miRNA modifications

<http://wwwdev.ebi.ac.uk/enright-dev/chimira/>

The screenshot displays the Chimira web application interface. At the top, the word "chimira" is written in a red, lowercase, sans-serif font. Below this, on the left side, there is a vertical stack of five grey buttons with white text: "Home", "About", "Example", "Help", and "Contact". To the right of these buttons, there is a list of three bullet points describing the tool's capabilities. Below the list, a line of text states: "You can upload your FASTA/FASTQ files by dragging them here, or clicking on the upload button." At the bottom of the page, there is a white form titled "Run" with a tabbed interface. The "Run" tab is selected, and it contains a blue "Upload files" button, a section for "Options" with a "Load example files" link, and three numbered steps for configuring the analysis.

chimira

- Chimira allows you to upload compressed FASTA/FASTQ files containing adapter/barcode stripped or raw small RNA-Seq data.
- All sequences will be mapped against **miRBase** hairpin sequences and assigned a match (allowing up to two mismatches).
- Any modifications (3', 5', internal) in the input sequences will be identified.

You can upload your FASTA/FASTQ files by dragging them here, or clicking on the upload button.

Run Clean & Run Other tools

Identify miRNA counts & modifications from adapter/barcode trimmed data.

Upload files

Options [Load example files](#)

1. Select genome:
2. Split counts from paralogs: ☐ (?)
3. Send results to (e-mail):

mRNA targets

- i. It is generally easier to identify miRNA targets in plants because there is higher level of complementarity between miRNAs and target mRNAs
- ii. In animals miRNAs usually pair with 3'UTRs but they can also regulate gene expression by pairing with CDS or 5'UTR
- iii. There are *in silico* and experimental approaches for identification of miRNA targets
- iv. CLIP-Seq and Degradome-seq experiments make it possible to identify miRNA targets in a large scale and provide more reliable results than in silico predictions.

Questions

- i. What genes are under control of a given miRNA?
- ii. What miRNAs regulate a given gene?
- iii. What is the mechanism of miRNA action?
- iv. What is the expression of miRNAs?
- v. Is there any correlation between expression of miRNA and targeted mRNA?
- vi. What is the phenotype effect?

In silico search for miRNA targets and databases

Tools:

- psRNATarget

Databases:

- starBase
- DIANA mirPath
- miRTarBase
- miRNEST

Tools

psRNATarget

<http://plantgrn.noble.org/psRNATarget/>

User-submitted small RNAs / preloaded transcripts

Preloaded small RNAs / user-submitted transcripts

User-submitted small RNAs / user-submitted transcripts

Upload small RNA sequence(s) in FASTA format: [\[Load demo data\]](#) ?

Nie wybrano pliku.
or paste sequences below:

- file / input sequence size limit: 15M
- invalid small RNAs will be ignored during analysis. ?

Select a preloaded transcript/genomic library for target search:

Allium cepa (Onion), unigene, DFCI Gene Index (ONGI), version 2, released on 2008_07_17
Arabidopsis thaliana, transcript, removed miRNA gene, TAIR, version 10, released on 2010_12_14
Arabidopsis thaliana, unigene, DFCI Gene Index (AGI), version 15, released on 2010_04_08
Arabidopsis thaliana, genomic DNA, 3.4K segments from strand with 0.4K overlapped region, TAIR, released on 2.....
Aquilegia (columbine), unigene, DFCI Gene Index (AQCI), version 2.1, released on 2008_06_06
Beta vulgaris (beet), unigene, DFCI Gene Index (BVGI), version 4, released on 2011_03_17
Brachypodium distachyon (purple false brome), transcript, JGI genomic project, Phytozome, phytozome v8.0, inter.....
Brachypodium distachyon (purple false brome), unigene, DFCI Gene Index (BDGI), version 1, released on 2010_05_2.....
Brassica napus (rape), unigene, DFCI Gene Index (ENGI), version 5, released on 2011_03_18
Brassica rapa (turnip, turnip rape, fast plants, field mustard, or turnip mustard), cds, de novo scaffolds ass.....

Selected library:

- Request to add / update a transcript library.

Maximum expectation (* Prefer lower false positive prediction rate? Please set a more stringent cut-off threshold [0-2.0]; Prefer higher prediction coverage? Please set a more relaxed cut-off threshold [4.0-5.0]): (range: 0-5.0) ?

Length for complementarity scoring (hpsize): (range: 15-30bp) ?

Target accessibility - allowed maximum energy to unpair the target site (UPE): (range: 0-100, less is better) ?

Flanking length around target site for target accessibility analysis: bp in upstream / bp in downstream ?

Range of central mismatch leading to translational inhibition: - nt ?

psRNATarget

keywords:
 Expectation:
 UPE:
 Search

e.g. AT1G27360, miR156, transcription factor ...
 Range: 0.0 - 3.0
 Range: 0.0 - 25.0

Sort by:

List of Predicted miRNA/Target Pairs [#Session ID: 1382540492451780]

[Batch Download](#)

[Prev Page](#)
[Next Page](#)
 Page No. 1 / Total 48 Pages , 1414 Records

miRNA Acc.	Target Acc.	Expectation (E)	Target Accessibility (UPE)	Alignment	Target Description	Inhibition	Multiplicity
mdm-miR1511	TC65106	2.5	22.294	miRNA 20 AGUACCAUAGUCUCGA-UCCA 1 : : : : : : : : : : : : : : : : : : Target 96 UCAUGGUACAGAGCUAGGU 116		Cleavage	1
mdm-miR1511	TC69999	3.0	16.387	miRNA 21 AAGUACCAUAGUCUGAUCCA 1 : : : : : : : : : : : : : : : : : : Target 114 UUCUUGGUACCAAGCUAGGU 134	weakly similar to UniRef100_Q107W9 Cluster: Small basic intrinsic protein 1; n=2; Vitis viniferaRep: Small basic intrinsic protein 1 - Vitis vinifera (Grape), partial (78%)	Translation	1
mdm-miR1511	CN933425	3.0	22.889	miRNA 20 AGUACCAUAGUCUGAUCCA 1 : : : : : : : : : : : : : : : : : : Target 106 UCAUGGUACAGAGCCAAGU 125	homologue to UniRef100_A7PET8 Cluster: Chromosome chr11 scaffold_13, whole genome shotgun sequence; n=1; Vitis viniferaRep: Chromosome chr11 scaffold_13, whole genome shotgun sequence - Vitis vinifera (Grape), partial (9%)	Cleavage	1
mdm-miR156a	CO051680	0.5	16.758	miRNA 20 CACGAGUGAGAGAAGACAGU 1 : : : : : : : : : : : : : : : : : : Target 19 GUGCUCGUCUCUUCUGUCA 38	similar to UniRef100_Q94JW8 Cluster: Squamosa promoter-binding-like protein 6; n=1; Arabidopsis thalianaRep: Squamosa promoter-binding-like protein 6 - Arabidopsis thaliana (Mouse-ear cress), partial (4%)	Cleavage	1
mdm-miR156a	TC60068	1.0	12.258	miRNA 20 CACGAGUGAGAGAAGACAGU 1 : : : : : : : : : : : : : : : : : : Target 323 GUGCUCUCUCUCUUCUGUCA 342	homologue to UniRef100_Q5VK50 Cluster: Bcl-2/adenovirus E1B 19kD interaction protein 3C; n=1; Danio rerioRep: Bcl-2/adenovirus E1B 19kD interaction protein 3C - Danio rerio (Zebrafish) (Brachydanio rerio), partial (8%)	Cleavage	1
mdm-miR156a	TC92994	1.0	14.827	miRNA 20 CACGAGUGAGAGAAGACAGU 1 : : : : : : : : : : : : : : : : : : Target 540 GUGCUCUCUCUCUUCUGUCA 559	similar to UniRef100_Q9S840 Cluster: Squamosa promoter-binding-like protein 2; n=1; Arabidopsis thalianaRep: Squamosa promoter-binding-like protein 2 - Arabidopsis thaliana (Mouse-ear cress), partial (17%)	Cleavage	1
mdm-miR156a	TC74079	1.0	14.857	miRNA 20 CACGAGUGAGAGAAGACAGU 1 : : : : : : : : : : : : : : : : : : Target 416 GUGCUCUCUCUCUUCUGUCA 435	similar to UniRef100_Q653Z5 Cluster: Squamosa promoter-binding-like protein 11; n=1; Oryza sativa Japonica GroupRep: Squamosa promoter-binding-like protein 11 - Oryza sativa subsp. japonica (Rice), partial (11%)	Cleavage	1
mdm-miR156a	CO868185	1.0	21.256	miRNA 20 CACGAGUGAGAGAAGACAGU 1 : : : : : : : : : : : : : : : : : : Target 526 GUGCUCUCUCUCUUCUGUCA 545	similar to UniRef100_A7PVU5 Cluster: Chromosome chr8 scaffold_34, whole genome shotgun sequence; n=2; Vitis viniferaRep: Chromosome chr8 scaffold_34, whole genome shotgun sequence - Vitis vinifera (Grape), partial (17%)	Cleavage	1

TargetScan

<http://www.targetscan.org/>



Search for predicted microRNA targets in mammals

[\[Go to TargetScanMouse\]](#)
[\[Go to TargetScanWorm\]](#)
[\[Go to TargetScanFly\]](#)
[\[Go to TargetScanFish\]](#)

1. Select a species

AND

2. Enter a human Entrez Gene symbol (e.g. "LIN28A")

AND/OR

3. Do one of the following:

- Select a broadly conserved* microRNA family
- Select a conserved* microRNA family
- Select a poorly conserved microRNA family Note that these families also include small RNAs that
- Enter a microRNA name (e.g. "mmu-miR-1")

Databases

DIANA mirPath

<http://snf-515788.vm.okeanos.grnet.gr/>

DIANA TOOLS

Software » Mirpath

Username *

Password *

☐ Remember me next time

Login

Forgot your password?

- Sign up for free!
- or take a tour

Available features for registered users:

- Download Databases
- History
- Bookmarks

Login is not required to access the site!

Created by:
ALEXANDER FLEMING
Institute of Informatics and Robotics
IMEIR IPHIS

HOME SOFTWARE PUBLICATIONS CONTACT

Species: Human Reverse Search New Search Help

Gene filter: determine a list of genes (optional) Run example

Add miRNAs: Przeglądaj... Upload ? or Determine miRNAs manually

Hide lists added ^

hsa-miR-125b-5p	microT-CDS	disable	see genes	(602)
hsa-miR-145-5p	microT-CDS	disable	see genes	(169)
hsa-miR-21-5p	microT-CDS	disable	see genes	(242)
hsa-miR-155-5p	microT-CDS	disable	see genes	(529)

Select the way to merge results:

genes union genes intersection - 4 + pathways union pathways intersection

FDR Correction: ☒ Conservative Stats: ☐

P-value threshold: 0.05 Apply

default

MicroT threshold: 0.8 Apply default

In order to see HeatMap select pathway intersection or pathway union.

Show Heatmap

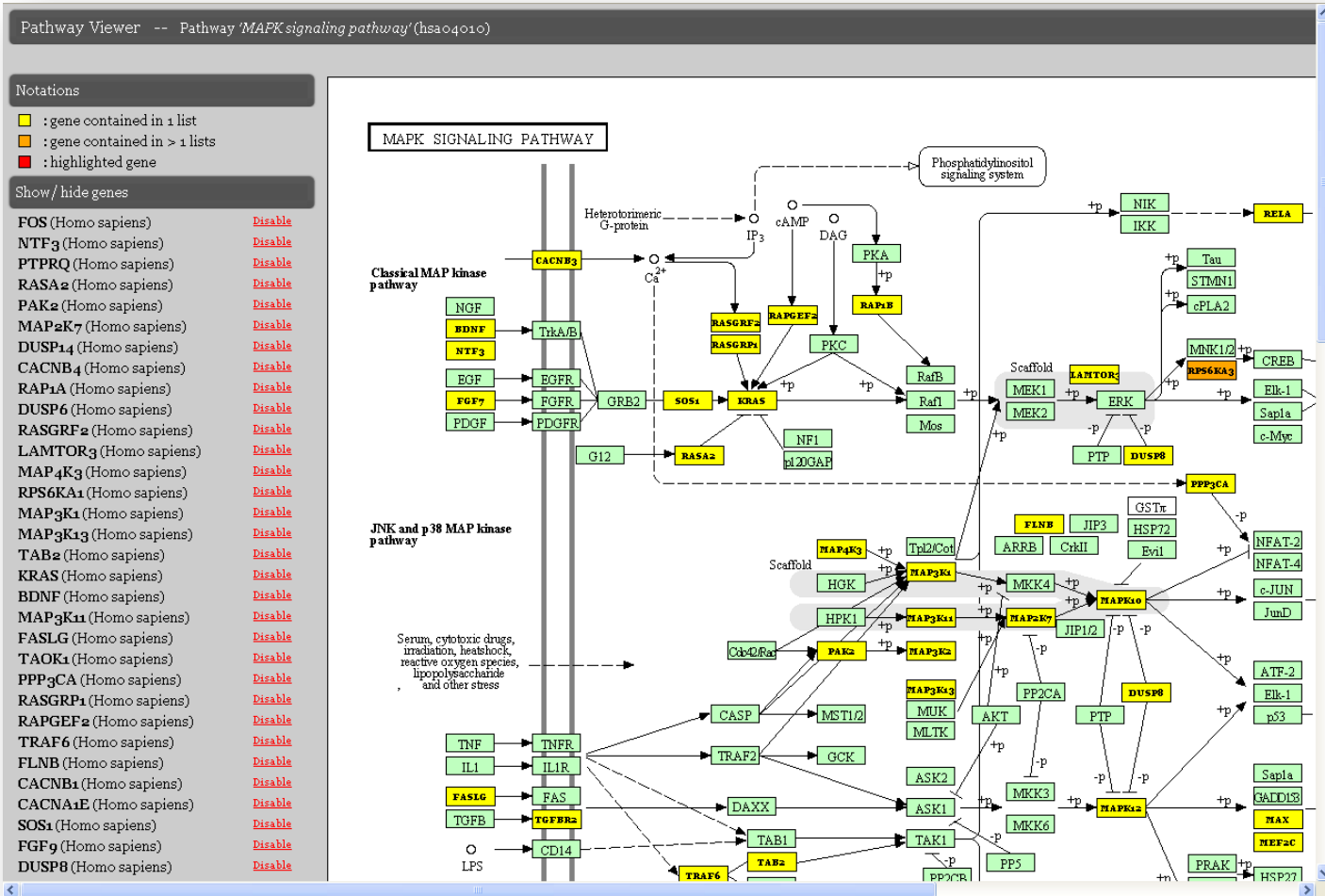
Show microRNA/Pathway Clusters

☒ Significance Clusters/Heatmap

☐ Targeted Pathways Clusters/Heatmap

#	KEGG pathway	p-value	#genes	#miRNAs	download results
1.	MAPK signaling pathway (hsa04010)	2.154379e-16	46 see genes	4	details
2.	Neurotrophin signaling pathway (hsa04722)	9.473359e-11	24 see genes	4	details
3.	B cell receptor signaling pathway (hsa04662)	1.125235e-06	15 see genes	4	details
4.	Hepatitis B (hsa05161)	1.32383e-06	21 see genes	4	details
5.	T cell receptor signaling pathway (hsa04660)	1.41182e-06	19 see genes	4	details
6.	Toll-like receptor signaling pathway (hsa04620)	2.215146e-05	17 see genes	3	details

DIANA mirPath



starBase

<http://starbase.sysu.edu.cn/>



Database, Visualize, Analyze, Discover.

[starBase v1.0](#)



[Home](#) [Browser](#) [miRNA-mRNA](#) [miRNA-ncRNA](#) [ceRNA Network](#) [Protein-RNA](#) [Functional Annotation](#) [Download](#) [Contact Us](#)

Welcome to starBase

starBase v2.0 for decoding **miRNA-mRNA**, **miRNA-ceRNA**, **miRNA-lncRNA**, **miRNA-circRNA**, **miRNA-pseudogene** and **protein-RNA** interaction networks from CLIP-Seq (HITS-CLIP, PAR-CLIP, iCLIP, CLASH) data.

starBase v2.0 also provides **visualization**, **analysis**, **discovery** and **downloading** of above-mentioned large-scale **functional genomics** data.

How to cite:

starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data.

Yang JH, Li JH, Shao P, Zhou H, Chen YQ and Qu LH.

Nucleic Acids Res. 2011;39:D202-D209. First published online: October 30, 2010

starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data.

Li JH, Liu S, Zhou H, Qu LH* and Yang JH*

Submitted.

What's New

starBase v2.0 release at 2013

- 108 CLIP-Seq datasets
- ~500,000 miRNA-mRNA interactions
- ~10,000 miRNA-lncRNA interactions
- ~16,000 miRNA-pseudogene interactions
- ~9,000 miRNA-circRNA interactions
- ~10,000 ceRNA-ceRNA interactions
- ~300,000 protein-RNA interactions
- two tools for functional annotation

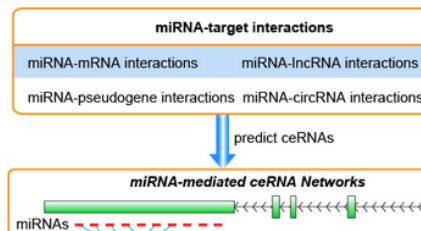
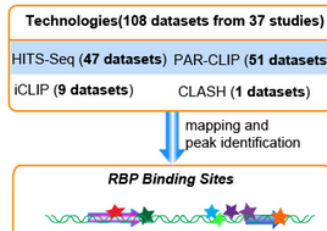
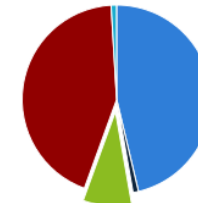
[starBase v2.0 Tutorial](#)

starBase v1.0 release at 2011

Data Sets

starBase contains **108** CLIP-Seq data from **37** studies. [Details](#)

CLIP-Seq datasets



miRTarBase

<http://mirtarbase.mbc.nctu.edu.tw/>

miRTarBase

HomeSearchBrowseStatisticsHelpDownloadContact Us

PTENSearchExample

miRTarBase: the experimentally validated microRNA-target interactions database

As a database, miRTarBase has accumulated more than thirty 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, microarray and next-generation sequencing experiments. While containing the largest amount of validated MTIs, the miRTarBase provides the most updated collection by comparing with other similar, previously developed databases.

[Browse the database !!](#)

Major improvements

Features	miRTarBase 1.0	miRTarBase 4.0
Known miRNA entry	miRBase (version 14)	miRBase (version 20)
Species	17 species	18 species
Curated miRNA-target interactions	3,576 MTIs	34,317 MTIs
MTIs were validated by experimental technology	Reporter assay, western blot, northern blot, qRT-PCR, microarray, pSILAC	Reporter assay, western blot, northern blot, qRT-PCR, microarray, pSILAC, NGS (CLIP-seq and Degradome-seq)
Number of MTIs validated by "reporter assay" or "western blot"	2,207 MTIs	4,560 MTIs

Current curation

Release 4.4: Sep. 15, 2013

Number of articles: 2,628

Number of species: 18

Number of target genes: 15,157

Number of miRNAs: 1,029

Number of miRNA-target interactions: 34,317

Related resources from ISBLAB

[miRTar](#) - An integrated web server for identifying miRNA-target interactions

[miRTarCLIP](#) - A computational approach for identifying microRNA-target interactions using high-throughput CLIP and PAR-CLIP sequencing

[miRStart](#) - A Database of microRNA transcription start sites

miRTarBase

miRTarBase

Home

Search

Browse

Statistics

Help

Download

Contact Us

Search...

Search

Example

Page 1 of 2

1

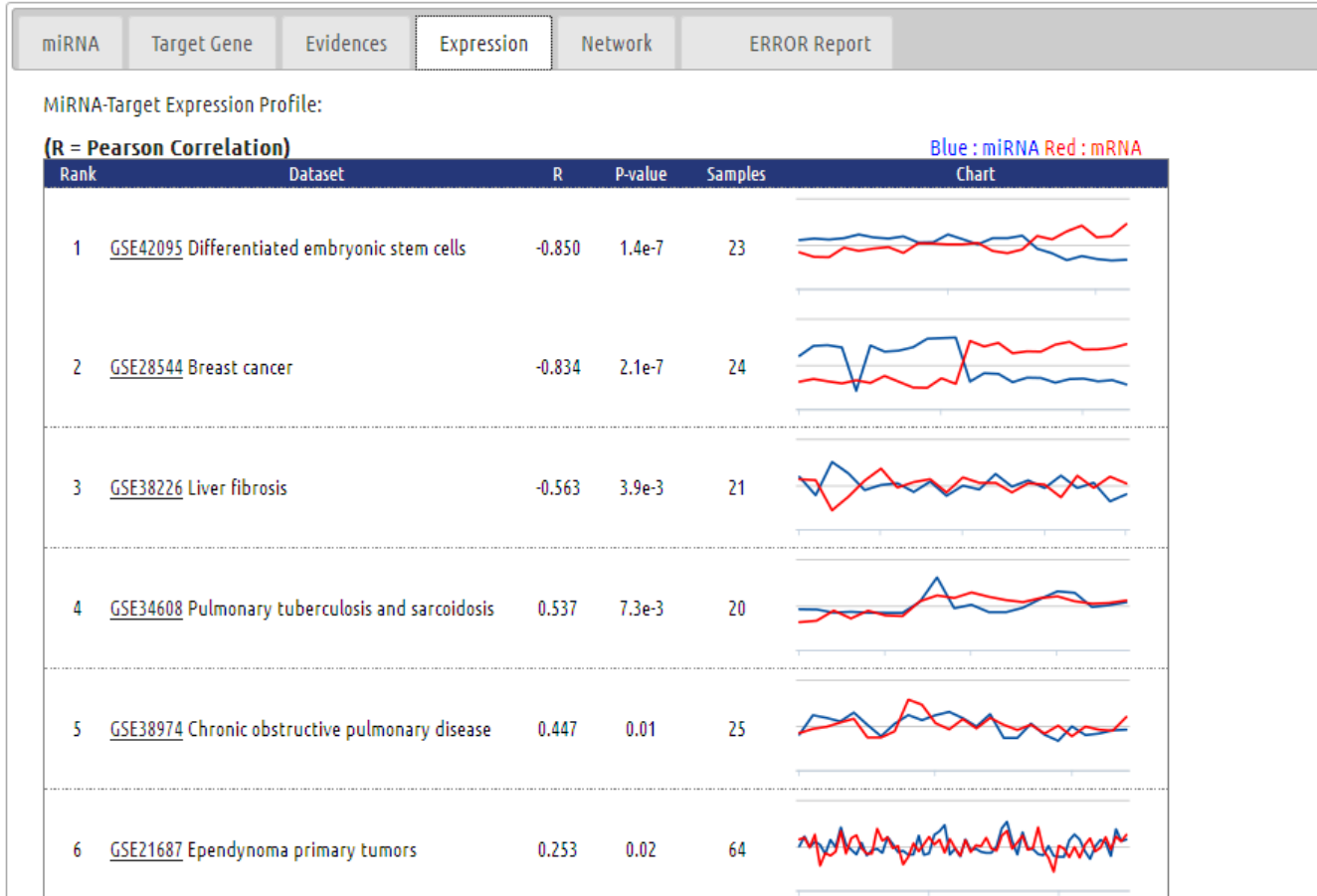
2

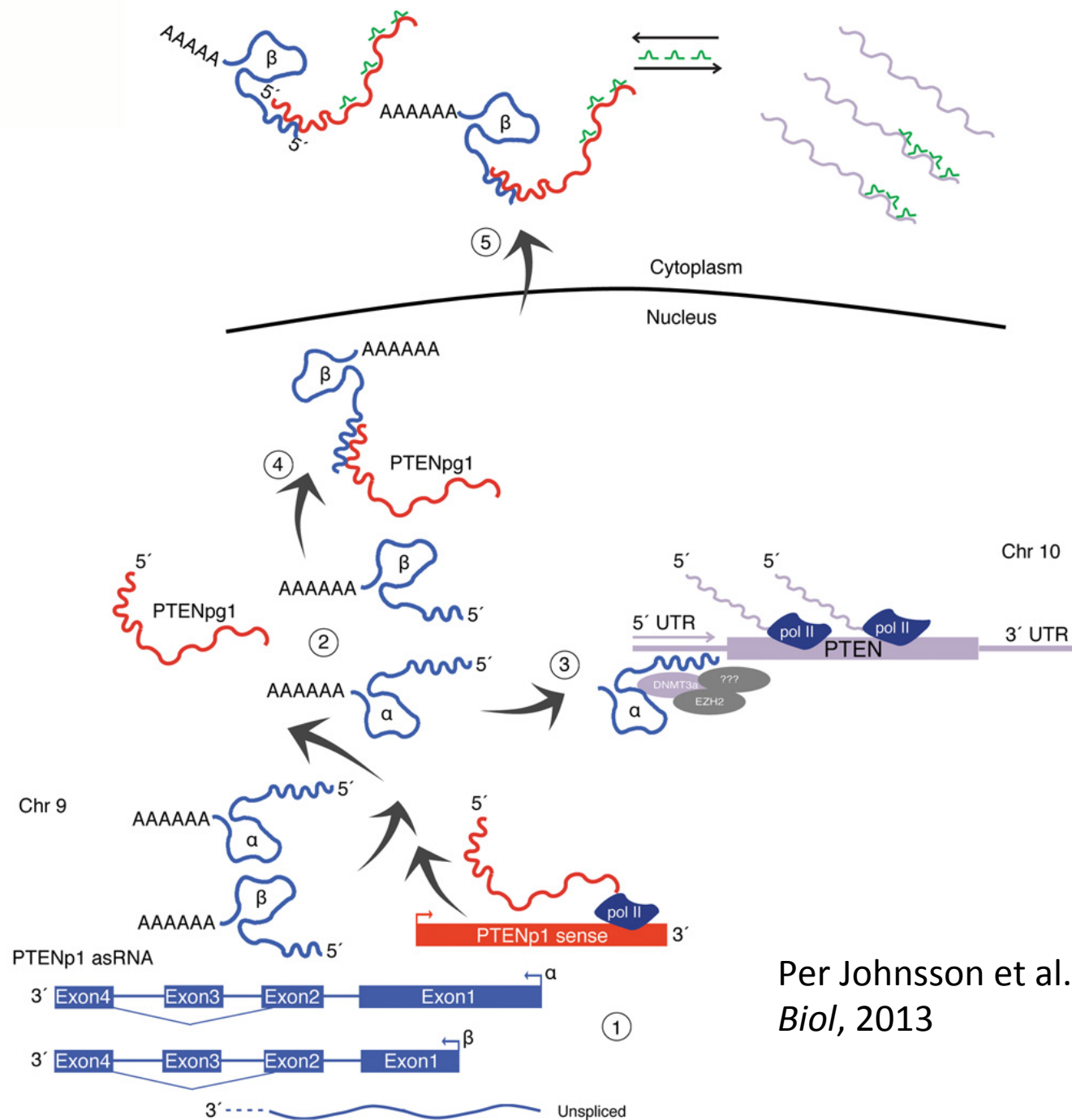
Next >

ID	Species (miRNA)	Species (Target)	miRNA	Target	Validation methods						Sum	# of papers	
					Strong evidence			Less strong evidence					
					Reporter assay	Western blot	qPCR	Microarray	NGS	pSILAC			Other
MIRT000499	Homo sapiens	Homo sapiens	hsa-miR-17-5p	PTEN	✓	✓	✓				✓	4	5
MIRT000533	Homo sapiens	Homo sapiens	hsa-miR-217	PTEN	✓	✓					✓	3	3
MIRT000534	Homo sapiens	Homo sapiens	hsa-miR-216a-5p	PTEN	✓	✓					✓	3	3
MIRT000799	Homo sapiens	Homo sapiens	hsa-miR-214-3p	PTEN	✓	✓	✓	✓			✓	5	5
MIRT001095	Homo sapiens	Homo sapiens	hsa-miR-26a-5p	PTEN	✓	✓					✓	3	5
MIRT001179	Mus musculus	Mus musculus	mmu-miR-21a-5p	Pten	✓	✓					✓	3	3
MIRT001190	Homo sapiens	Homo sapiens	hsa-miR-21-5p	PTEN	✓	✓	✓	✓			✓	5	33
MIRT001209	Homo sapiens	Homo sapiens	hsa-miR-494-3p	PTEN	✓	✓	✓	✓			✓	5	2

miRTarBase

Accession ID: MIRT000499 [miRNA, hsa-miR-17-5p :: PTEN, target gene]





Per Johnsson et al., *Nat Struct Mol Biol*, 2013

Thank you for your attention