

The logo for DIANA LAB features a red rectangular box on the left containing the text "DIANA LAB" in white. To the right of the box is a stylized illustration of a DNA double helix structure, with red lines representing the sugar-phosphate backbone and grey lines representing the nitrogenous base pairs.

DIANA LAB

The logo for the COST Conference "NGS: a look into the future" in Bratislava 2015. It features a stylized black silhouette of a telescope on a tripod. To the right of the telescope, the text "NGS: a look into the future" is written in a small font, followed by "COST Conference" and "BRATISLAVA 2015" in a larger font.

NGS: a look into the future
COST Conference
BRATISLAVA 2015

COST Conference "Next Generation Sequencing: a look into the future"

DIANA-TARBASE v7: INDEXING HUNDREDS OF THOUSANDS EXPERIMENTALLY SUPPORTED miRNA:mRNA INTERACTIONS

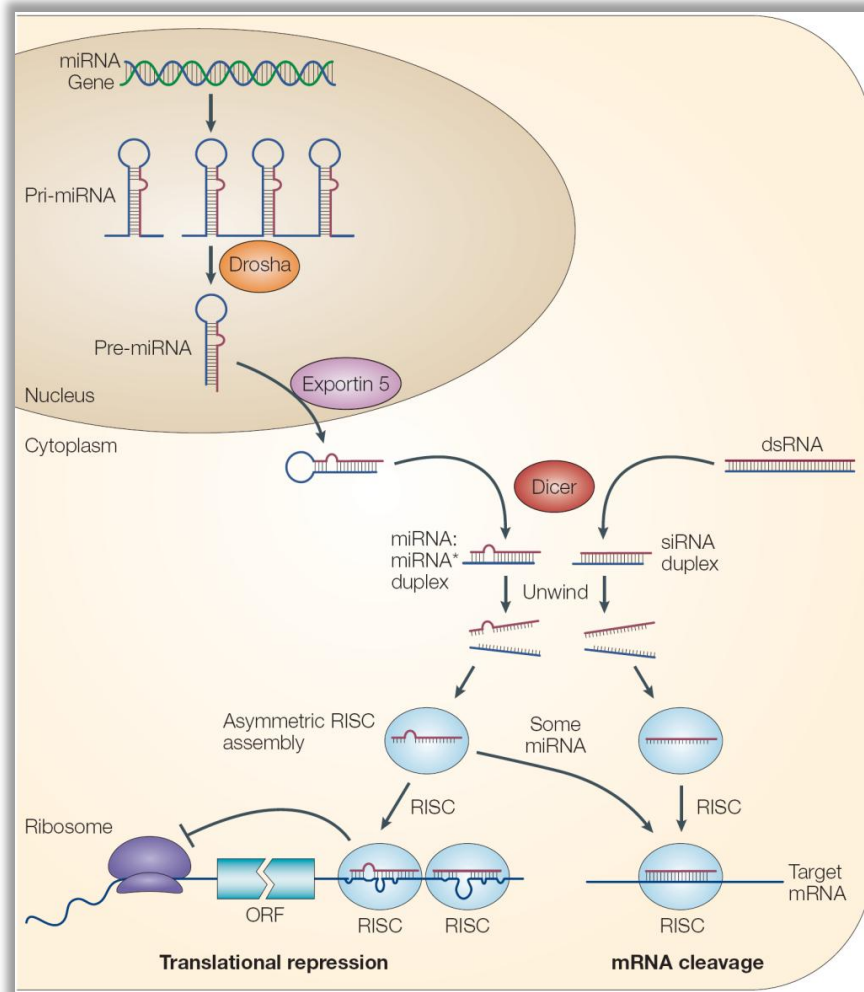
Dimitra Karagkouni

Principal Investigator
Dr. Artemis Hatzigeorgiou



University of Thessaly

microRNA Biogenesis



He, L., Hannon, G.J. **MicroRNAs: small RNAs with a big role in gene regulation.** Nat Rev Genet. 2004;5(7):522-31.

Role of microRNAs

Development

stem cell proliferation

Division

Differentiation

regulation of innate & adaptive immunity

apoptosis

cell signaling

metabolism

human pathologies

Cancer

viral infections

cardiovascular diseases

metabolic disorders neurological pathologies

psychiatric disorders renal disease hepatological conditions

autoimmune diseases gastroenterological conditions

obesity reproductive disorders

musculoskeletal disorders

periodontal pathologies

Identification of miRNA targets

- Computational methods
 - In silico target predictions programs
- Experimental methods
 - Essential to identify genuine miRNA:targets
 - Limitations exist in different experimental procedures

In Silico Determination of miRNA – Gene Interactions

Widely used implementations:

- TargetScan: www.targetscan.org
- DIANA-microT: www.microrna.gr
- miRanda: www.microrna.org
- PicTar: pictar.mdc-berlin.de
-

Even the most sophisticated implementations achieve 65% sensitivity and 32% specificity¹

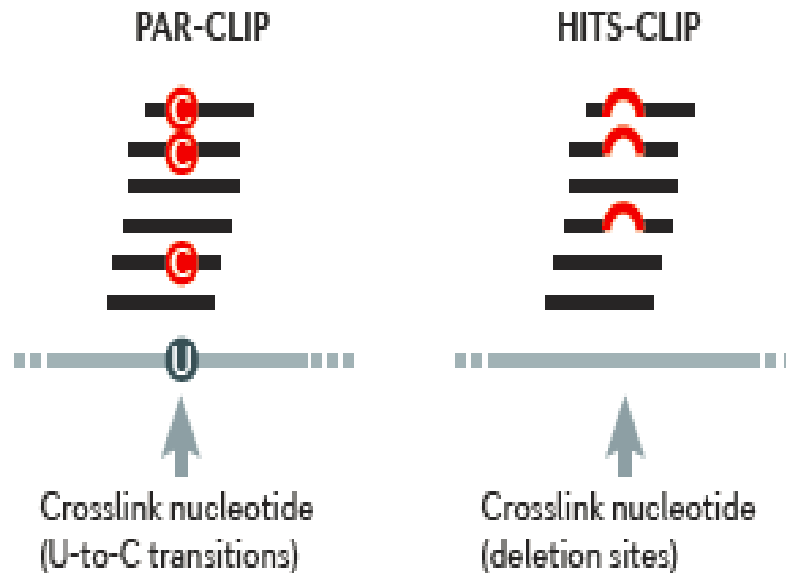
1: M. Reczko, M. Maragkakis, P. Alexiou, I. Grosse, and A. G. Hatzigeorgiou, "Functional microRNA targets in protein coding sequences," *Bioinformatics*, Jan 27 2012.

Experimental Determination of miRNA – Gene Interactions

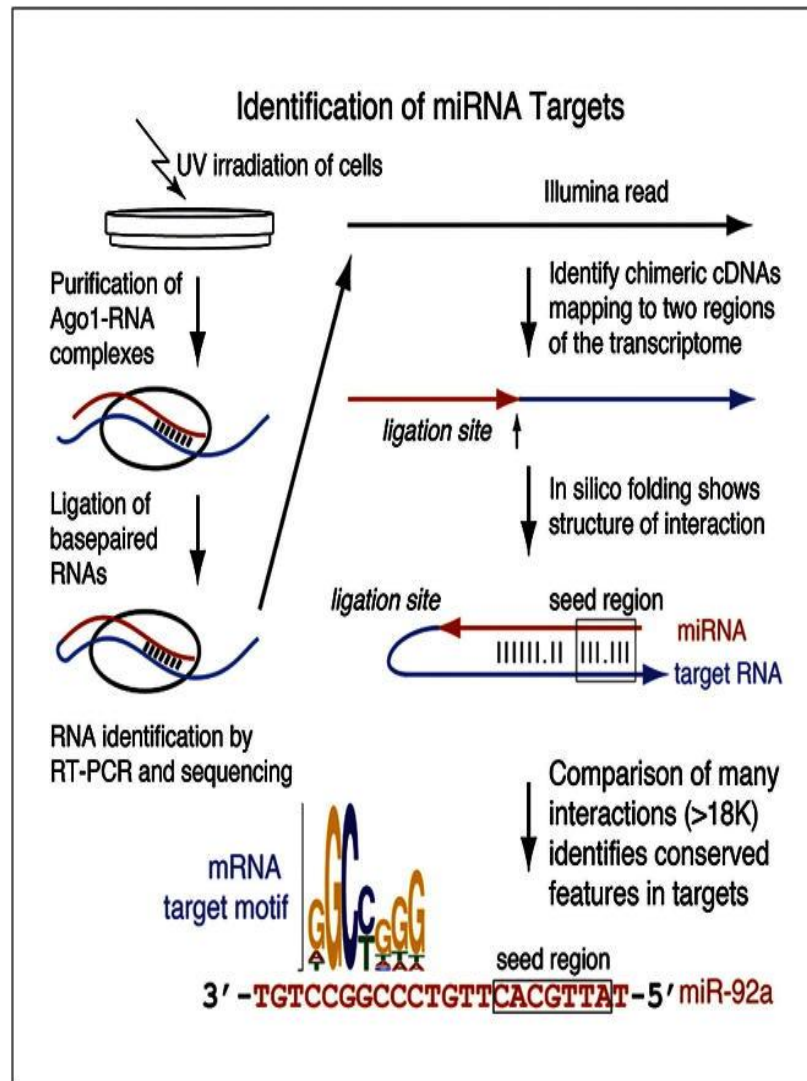
- **Gene Specific Techniques**
 - Reporter genes
 - Northern blotting
 - qPCR
 - Western blotting
 - ELISA
 - Immunohistochemistry
- **High Throughput Techniques**
 - MicroArrays
 - RNA-Seq
 - Proteomics (such as pSILAC)
 - CLIP-Seq (HITS-CLIP, PAR-CLIP, iCLIP)
 - CLASH
 - PARE-Seq
 - Degradome-Seq
 - Biotin tagged miRNA

CLIP experiments

- Ultraviolet Crosslinking and immunoprecipitation
- A method used to isolate and identify sequences that are bound by specific RNA-binding proteins



CLASH experiment



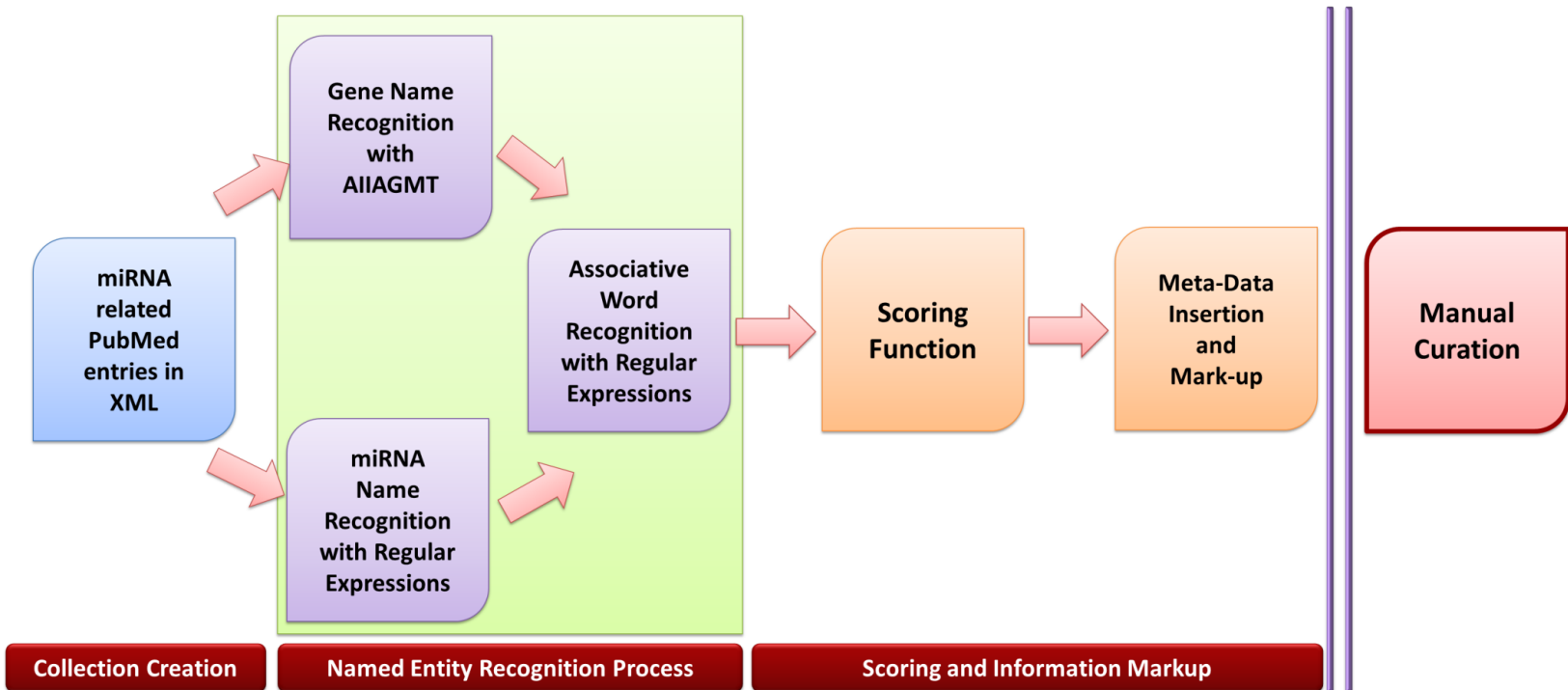
Aleksandra Helwak *et al*, Cell
153, 654–665, April 25, 2013

DIANA-TarBase

- Initially released in 2006
 - The first database aiming to catalog published experimentally validated miRNA:gene interactions
- The largest experimentally validated repository with miRNA:gene interactions
- Last update DIANA-TarBase v7 ¹
- <http://www.microrna.gr/tarbase>

¹I. S. Vlachos, M. D. Paraskevopoulou, D. Karagkouni, G. Georgakilas, T. Vergoulis, I. Kanellos, I-L. Anastasopoulos, S. Maniou, K. Karathanou, D. Kalfakakou, A. Fevgas, T. Dalamagas and A. G. Hatzigeorgiou. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. Nucl. Acids Res. (2014)

DIANA-TarBase Semi – Automatic Curation Pipeline



Named Entity Recognition Process

Abstract link: [20735361](#) **Score:** 11

Title: [Identification of microRNAs expressed highly in pancreatic islet-like cell clusters differentiated from human embryonic stem cells.](#)

Journal: Cell biology international (Cell Biol. Int.),

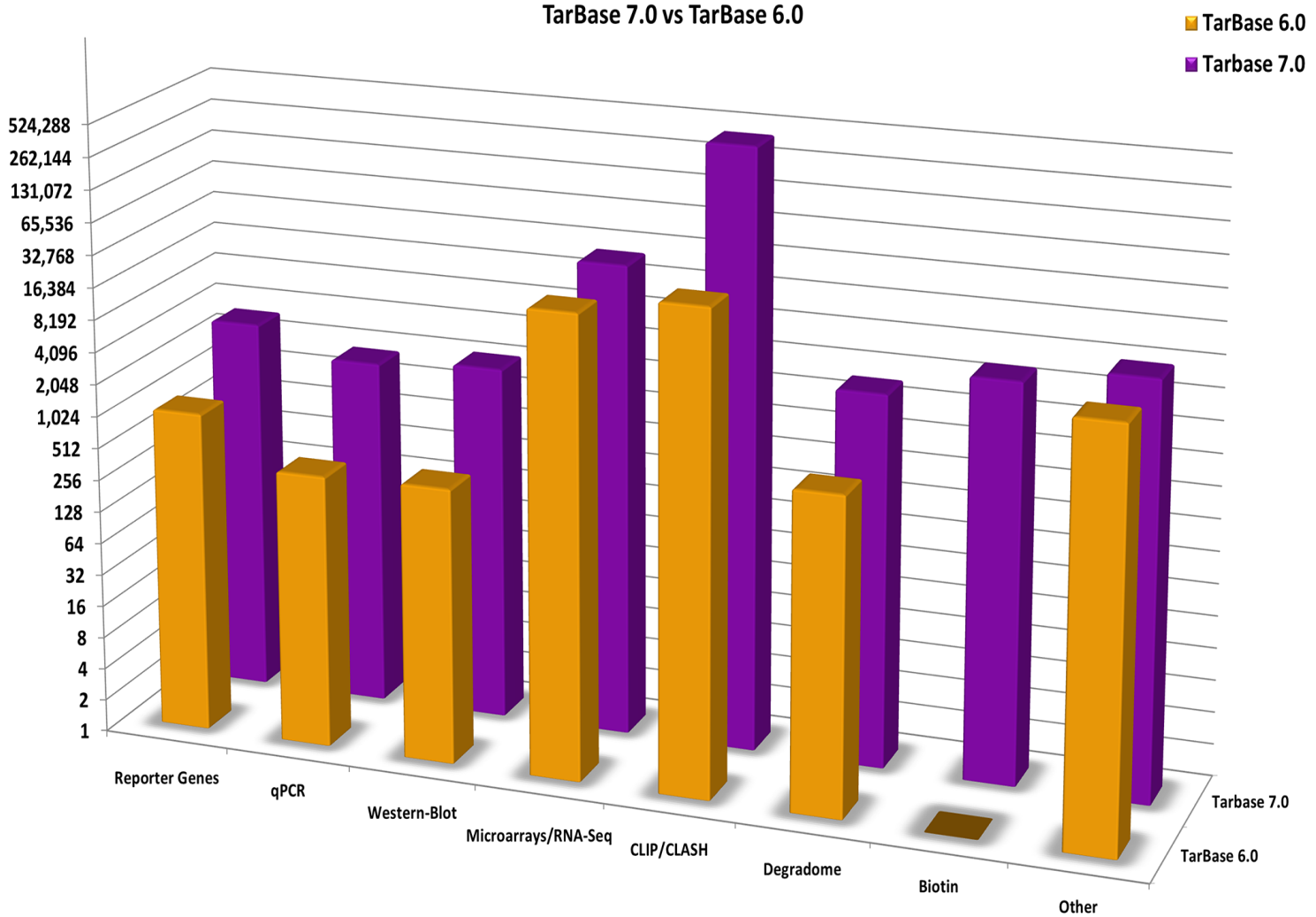
ISSN: 1065-6995, **Date:** 01 / 12 / 2010

Identification of microRNAs expressed highly in pancreatic islet-like cell clusters differentiated from human embryonic stem cells. Type 1 diabetes is an autoimmune destruction of pancreatic islet beta cell disease, making it important to find a new alternative source of the islet beta cells to replace the damaged cells. hES (human embryonic stem) cells possess unlimited self-renewal and pluripotency and thus have the potential to provide an unlimited supply of different cell types for tissue replacement. The hES-T3 cells with normal female karyotype were first differentiated into EBs (embryoid bodies) and then induced to generate the T3pi (pancreatic islet-like cell clusters derived from T3 cells), which expressed pancreatic islet cell-specific markers of [insulin](#), [glucagon](#) and [somatostatin](#). The expression profiles of microRNAs and mRNAs from the T3pi were analyzed and compared with those of undifferentiated hES-T3 cells and differentiated EBs. MicroRNAs negatively [regulate](#) the expression of protein-coding mRNAs. The T3pi showed very high expression of microRNAs, [miR-186](#), [miR-199a](#) and [miR-339](#), which [down-regulated](#) the expression of [LIN28](#), [PRDM1](#), [CALB1](#), [GCNT2](#), [RBM47](#), [PLEKHH1](#), [RBPMS2](#) and [PAK6](#). Therefore, these microRNAs and their [target](#) genes are very likely to play important [regulatory](#) roles in the development of pancreas and/or differentiation of islet cells, and they may be manipulated to [increase](#) the proportion of beta cells and [insulin](#) synthesis in the differentiated T3pi for cell therapy of type I diabetics.

DIANA-TarBase v7

- >500,000 interactions
 - 9 – to 250-fold increased compared to
 - Tarbase v6.0
 - other manually curated databases-miRTarBase,miRecords
- >7,500 interactions from specific techniques
- >500,000 derived from high-throughput techniques
- Data derived from
 - Hundreds of manually curated articles
 - Selected by our text mining pipeline

DIANA-TarBase v7



Experimental techniques

- ~3,000 Reporter Genes entries
- miRNA binding site identification
 - Information retrieved from manuscripts, figures, supplemental material
 - Genomic/transcript coordinates
 - Primer sequences
 - Mutation sites

CLIP-Seq data in Tarbase v7.0

~20-fold increase compared to Tarbase v6

CLIP-Seq	TarBase v6.0	TarBase v7.0
Studies	3	23
Conditions	6	68
Libraries	10	154

Analysis of raw data

- 1 CLASH
- 31 PAR-CLIP
- 122 HITS-CLIP libraries

DIANA-TarBase v7.0 – Statistics

- **24** different species
- **28** different experimental techniques
- **356** cell types
- **59** different tissues

TarBase v7.0 Interface

4. Methods

1. Database Search Terms

hsa-miR-34a-5p

Gene name	miRNA name	Methods	Pred.Score
ACSL4 (hsa)	hsa-miR-34a-5p	IP RS Pr	1.000

Related Pathways

Filters

- Species
- Method Type
- Method
- Regulation type
- Validation type
- Validated as
- Source
- Publication year: 1900

Only publications published after the selected year will be presented.

Apply Filter!

Publication	Methods	Tissue	Cell line	Tested cell line	Exp. condition
Skalsky Rebecca L. et al. 2012	IP	NA	EF3DAGO2	N/A	N/A
Balakrishnan I et al 2014	IP	Bone Marrow	HS5	N/A	N/A
Kaller M et al. 2011	Pr	NA	NA	N/A	N/A
Kaller M et al. 2011	RS	Intestine	SW480	H1299	N/A

Location	Method	Result	Regulation	Valid. type	Source
chrX:109641508-109644016 (3UTR)	Luciferase Reporter Assay	POSITIVE	↓	DIRECT	Tarbase 7.0
Kishore S et al. 2011	IP	Kidney	HEK293	N/A	mild MNase digestion

Location	Method	Result	Regulation	Valid. type	Source
chrX:109642109-109642134 (UNKNOWN)	HITS-CLIP	POSITIVE	↓	DIRECT	Tarbase 7.0

3. Filters

5. Click (i) for further information

2. Interaction info

TarBase v7.0 Interface

Gene information/expression

miRNA information

Q hsa-miR-34a-5p

Gene name	miRNA name	Methods	Pred.Score
ACSL4 (hsa)	hsa-miR-34a-5p	IP RS Pr	1.000

Publication	Methods	Tissue	Cell line	Tested cell line	Exp. condition
Skalsky Rebecca L. et al. 2012	IP	NA	EF3DAGO2	N/A	N/A
Balakrishnan I et al 2014	IP	Bone Marrow	HS5	N/A	N/A
Kaller M et al. 2011	Pr	NA	NA	N/A	N/A
Kaller M et al. 2011	RS	Intestine	SW480	H1299	N/A

Location	Method	Result	Regulation	Valid. type	Source
chrX:109641508-109644016 (3UTR)	Luciferase Reporter Assay	POSITIVE	↓	DIRECT	Tarbase 7.0

TarBase v7.0 Interface

- Interconnected to other DIANA-Tools
 - DIANA-miRPath v2.0
 - www.microrna.gr/miRPathv2
 - microT-CDS
 - www.microrna.gr/microT-CDS
 - LncBase
 - www.microrna.gr/LncBase



TarBase v7.0 Interface

Connection with
microT-CDS

Connection
with miRPath
v2.0

hsa-miR-34a-5p

Gene name	miRNA name	Methods	Prevalence
ACSL4 (hsa)	hsa-miR-34a-5p	IP RS Pr	1,000

Publication	Methods	Tissue	Cell line	Tested cell line	Exp. condition
Skalsky Rebecca L. et al. 2012	IP	NA	EF3DAGO2	N/A	N/A
Balakrishnan I et al 2014	IP	Bone Marrow	HS5	N/A	N/A
Kaller M et al. 2011	Pr	NA	NA	N/A	N/A
Kaller M et al. 2011	RS	Intestine	SW480	H1299	N/A
Kishore S et al. 2011	IP	Kidney	HEK293	N/A	mild MNase digestion

Location	Method	Result	Regulation	Valid. type	Source
chrX:109642109-109642134 (UNKNOWN)	HITS-CLIP	POSITIVE	↓	DIRECT	Tarbase 7.0

[Related Pathways](#)

Filters

Selected:
Species: KSHV, Homo sapiens
Regulation: -
Validation: -
Validated: -
Sources: -
Methods: 3LIFE, AGO-IP, Biotin-Microarrays, Biotin-Seq, CLASH, Degradome, HITS-CLIP, iCLIP, IMPACT-Seq, Microarrays, PAR-CLIP, pSILAC, RNA-Seq

Remove all X

Species (2)

Diana microT-CDS Interface

The screenshot displays the Diana microT-CDS web interface. On the left, a sidebar contains a 'Personalized search space' with user information and a 'Tool-specific history panel' listing recent searches for hsa-miR-1, hsa-miR-424-5p, and hsa-let-7a-5p. The main content area features a search bar with 'hsa-miR-1' entered, a 'Results summary' indicating 1041 targets, and a table of results. Callouts highlight 'Search terms', 'Advanced filtering options', 'Gene & miRNA details' (including a table of binding areas), and 'MRE-specific information' (showing a sequence alignment).

Search terms

Advanced filtering options

Results summary

Personalized search space

Tool-specific history panel

Visualization of MREs in UCSC

Gene & miRNA details

MRE-specific information

Ensembl Gene Id	miRNA name	miTG score	Also Predicted	
1 ENSG00000059728 (MXD1)	hsa-miR-1	1.000	<input type="checkbox"/>	
Gene details miRNA details pubMed links: miRNA gene both UCSC graphic				
Region	Binding Type	Transcript position	Score	Conservation
UTR3	9mer	205-233	0.128346633186705	9
Position on chromosome: 2:70165622-70165650 Conserved species: panTro2,rheMac2,oryCun2,bosTau4,dasNov2,loxAfr3,echTel1,monDom5,gaiGal3 (Transcript) 5' AAUUUGUCUCUGAGAGA A 3'				
Binding area: (miRNA) 3' CU UACAUUCCA GA AUGUAGGU A A 5'				
UTR3	6mer	536-564	0.00245594937370441	
UTR3	6mer	1319-1347	0.00443240062000479	7
UTR3	6mer	1745-1773	0.00249186205041115	6
2 ENSG00000127124 (HIVEP3)	hsa-miR-1	1.000	<input type="checkbox"/>	
3 ENSG00000070214 (SLC44A1)	hsa-miR-1	1.000	<input type="checkbox"/>	

miRPath v2.0 Interface

DIANA LAB
DNA Intelligent Analysis

[HOME](#) [SOFTWARE](#) [DATABASES](#) [MEMBERS](#) [PUBLICATIONS](#)

Mirpath
🏠

DIANA miRPath v2.0 is a web-based computational tool that identifies potentially altered molecular pathways by the expression of a single or multiple microRNAs.

Species: Human

Gene filter: [determine a list of genes \(optional\)](#)

Add miRNAs: Tarbase + [or upload a file](#)

[Reverse Search](#) [New Search](#) [Help](#)

[Run example](#)

Selected miRNAs (lists)

hsa-let-7a-5p	microT-CDS	disable	see genes	(496)
hsa-miR-21-5p	Tarbase	disable	see genes	(451)
hsa-miR-140-5p	Tarbase	disable	see genes	(0)

Disable the list

Add extra miRNAs (lists)

Show genes for this list

Show/Hide miRNAs (lists)

[Hide lists added](#) ^

Select the way to merge results: [genes union](#) [genes intersection](#) 3 + [pathways union](#) [pathways intersection](#)

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

P-value threshold: [Apply](#) [default](#)

MicroT threshold: [Apply](#) [default](#)

Statistical analysis options

FDR Correction:

Conservative Stats:

Merging and meta-analysis options

Heatmap & clustering options

Show Heatmap

Show microRNA/Pathway Clusters

Significance Clusters/Heatmap
 Targeted Pathways Clusters/Heatmap

Download results in file

	KEGG pathway	p-value	#genes	#miRNAs	download results
1.	ECM-receptor interaction (hsa04512)	<1e-16	8 see genes	2	details
	hsa-let-7a-5p/microT-CDS	2.915217e-20	6 see genes		
	hsa-miR-21-5p/Tarbase	0.04394519	4 see genes		
2.	Glycosaminoglycan biosynthesis - heparan sulfate (hsa00534)	0.0006455011	3 see genes	1	details

List of related miRNAs

TarBase v7.0 Interface - Advanced filter options

Filters

Selected:
Species: Homo sapiens
Regulation: -
Validation: -
Validated: -
Sources: -
Methods: -

Remove all X

Species (1) ^

- Arabidopsis thaliana
- Bombyx mori
- Bos taurus
- Caenorhabditis elegans
- Canis familiaris
- Danio rerio
- Dasyopus novemcinctus
- Drosophila melanogaster
- Epstein Barr virus
- Equus caballus
- Gallus gallus
- Glycine max
- HIV 1
- Homo sapiens
- KSHV
- Medicago truncatula
- Mus musculus
- Oryctolagus cuniculus
- Oryza sativa japonica
- Ovis aries
- Pan troglodytes
- Physcomitrella patens
- Prunus persica
- Rattus norvegicus
- Vitis vinifera
- Xenopus laevis
- Xenopus tropicalis
- Zea mays

Species (1) v

Method Type (1) ^

- Normal
- High Throughput

Method (15) ^

- 2D-DIGE
- 3LIFE
- AGO-IP
- Biotin-Microarrays
- Biotin-qPCR
- Biotin-Seq
- CLASH
- Degradome
- Electrophoretic Mobility Shift Assay
- ELISA
- Flow Cytometry
- Genetic Testing
- HITS-CLIP
- iCLIP
- Immunofluorescence
- Immunohistochemistry
- IMPACT-Seq
- In Situ Hybridization
- LAMP
- Luciferase Reporter Assay
- Microarrays
- Northern Blot
- Other
- PAR-CLIP
- PARE
- pSILAC
- qPCR
- RNA-Seq
- SILAC
- TRAP
- Western Blot

Regulation type (1) ^

- UP
- DOWN
- UNKNOWN

Validation type (1) ^

- DIRECT
- INDIRECT
- UNKNOWN

Validated as (1) ^

- POSITIVE
- NEGATIVE
- UNKNOWN

Source (4) ^

- TarBase 6.0
- TarBase 5.0
- MiRecords
- miR2Disease
- miRTarBase
- Tarbase 7.0

Publication year

1900

Only publications published after the selected year will be presented.

Apply Filter!

Prediction score

0

Only those having prediction score greater than the above will be presented.

Apply Filter!

ENSEMBL Integration

ENSEMBL integration:

- Integrated in ENSEMBL since February 2014!
- TarBase substituted *in silico* predicted targets track

ENSEMBL Integration

- Resequencing (20)
- Linkage Data
- Markers
- Other genome browsers
 - UCSC
 - NCBI
 - Vega

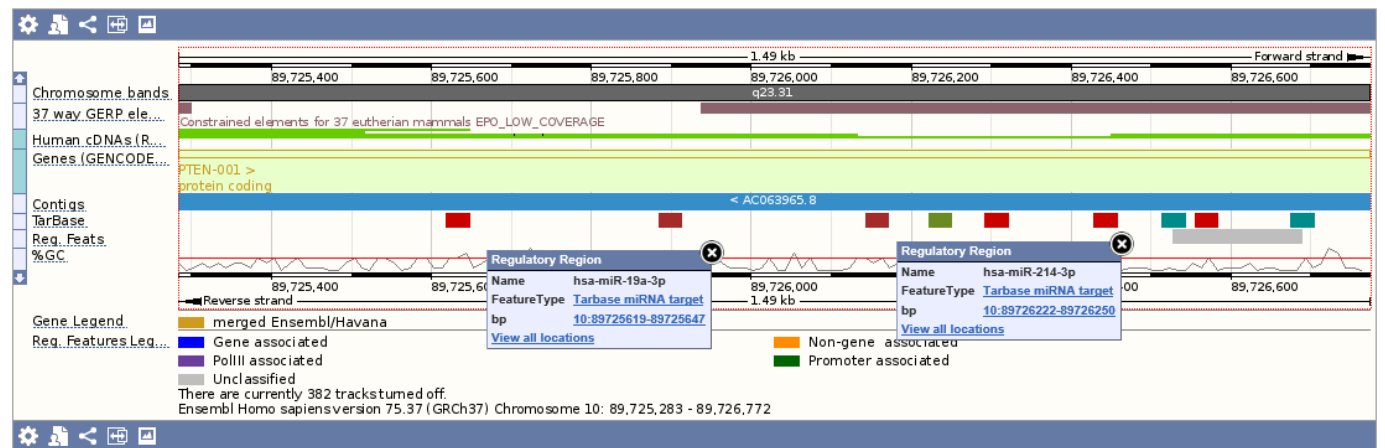
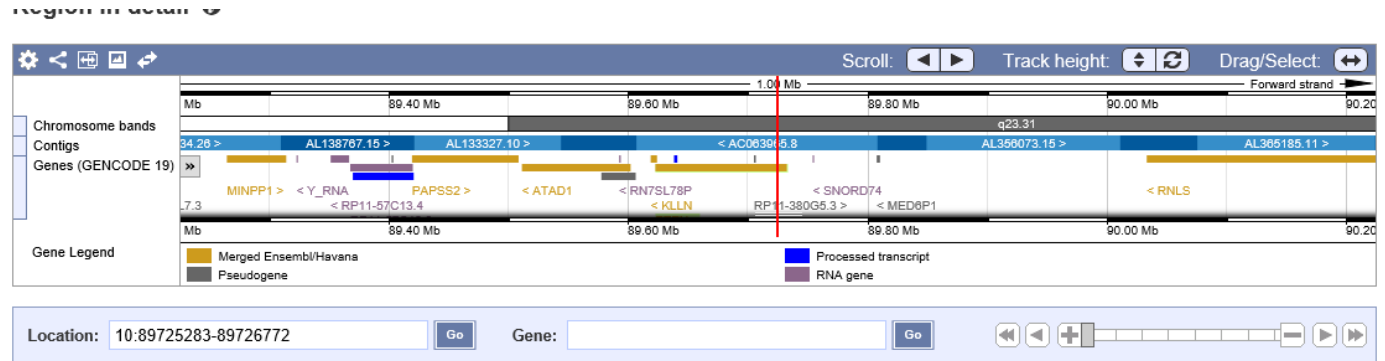
Configure this page

Add your data

Export data

Bookmark this page

Share this page

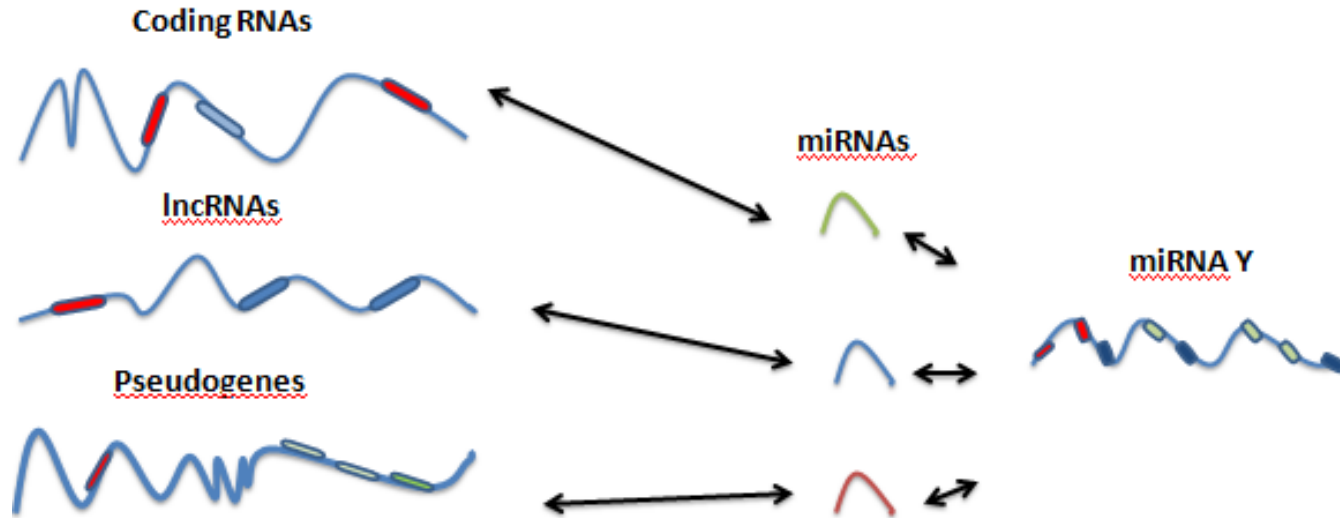


DIANA-TarBase v7.0

Conclusion

- > 500,000 entries
- Analysis >150 raw NGS data
- Hundreds of manually curated articles
- Completely redesigned
- Detailed information for each interaction
- Tissue specific results
- Overcomes Target prediction bottleneck
 - Enrich or substitute in silico predicted interactions
- Utilization in exploratory studies

LncBase Update (upcoming)



Competing endogenous RNA (ceRNA) network

LncBase Update (upcoming)

- Experimental Module
- >5,000 interactions
- 2,958 lncRNAs and 120 miRNAs
- Prediction Module
- >10 million interactions
- 56,097 lncRNAs and 3,078 miRNAs

- Currently working on an extension of LncBase database
- Data from Ensembl/NCBI
- CLIP-Seq Data analysis

LncBase Update (upcoming)

DIANA LAB DNA Intelligent Analysis

HOME SOFTWARE **Search field** MEMBERS PUBLICATIONS

logged in as **georgaki**

Need for help? Check out our FAQ!

Personal info
 Username: georgaki
 Email: geo2mandos@gmail.com
 Queries performed: 9
 Last query on: 2012-09-26 13:47:11
 Stored bookmarks: 0

Personalized search space

Advanced options:

Select species:
 Homo Sapiens
 Mus Musculus

Select ver. method:
 Reporter Gene Assay
 Western Blot
 PCR
 Proteomics
 microarray
 Sequencing
 Other
 Degradome
 Northern Blot

Select publ. year: 1900
 Only publications published after the selected year will be presented.

Select pred. score thr.: 0
 Only those having prediction score greater than the above will be presented.

Advance search options

Please cite: LncBase

	Gene name	miRNA name	Methods	Pred. score
1	hsaLOGG410001027 (XLOC008185)	hsa-miR-424-5p	R N W Q P M A D O	0.986

Interaction related information

Gene details
 Gene ID: hsaLOGG410001027
 Gene Name: XLOC008185
 Description: intergenic
 Chromosome: X
 Transcripts: hsaLOCT410001027

miRNA details
 Name: hsa-miR-424-5p
 Alternative description: MIMAT0001341
 Related names: There are no related names for this entry.
 miRNA sequence: CAGCAGCAAUUCAGUUUUGAA
 External links: miRBase
 Related MeSH terms: **Anoxia** Chronic Disease Colonic Neoplasms Disease Models, Animal Endometriosis Heart Diseases Hypertension, Pulmonary Leukemia, Lymphocytic, Chronic, B-Cell Neoplasms Neovascularization, Pathologic Neurodegenerative Diseases Precursor B-Cell Lymphoblastic Leukemia-Lymphoma Precursor T-Cell Lymphoblastic Leukemia-Lymphoma

Gene and miRNA details

Publication related information

Authors	Year	Methods	Regulation	Valid. type	Region
Hafner M. et al	2010	R N W Q P M A D O	↓	DIRECT	

Cell types: HEK293
 Original sources: LncBase
 Comments: PAR-CLIP

Acknowledgements

Lab Head

- **Dr. Artemis Hatzigeorgiou**

Lab members

- **Dr. Ioannis Vlachos**
- **Maria Paraskevopoulou**
- **Georgios Georgakilas**

Partners

- **Thanasis Vergoulis**
- **Ilias Kanellos**
- **Dr. Theodore Dalamagas**