Hellenic Pasteur Institute

Table of Contents

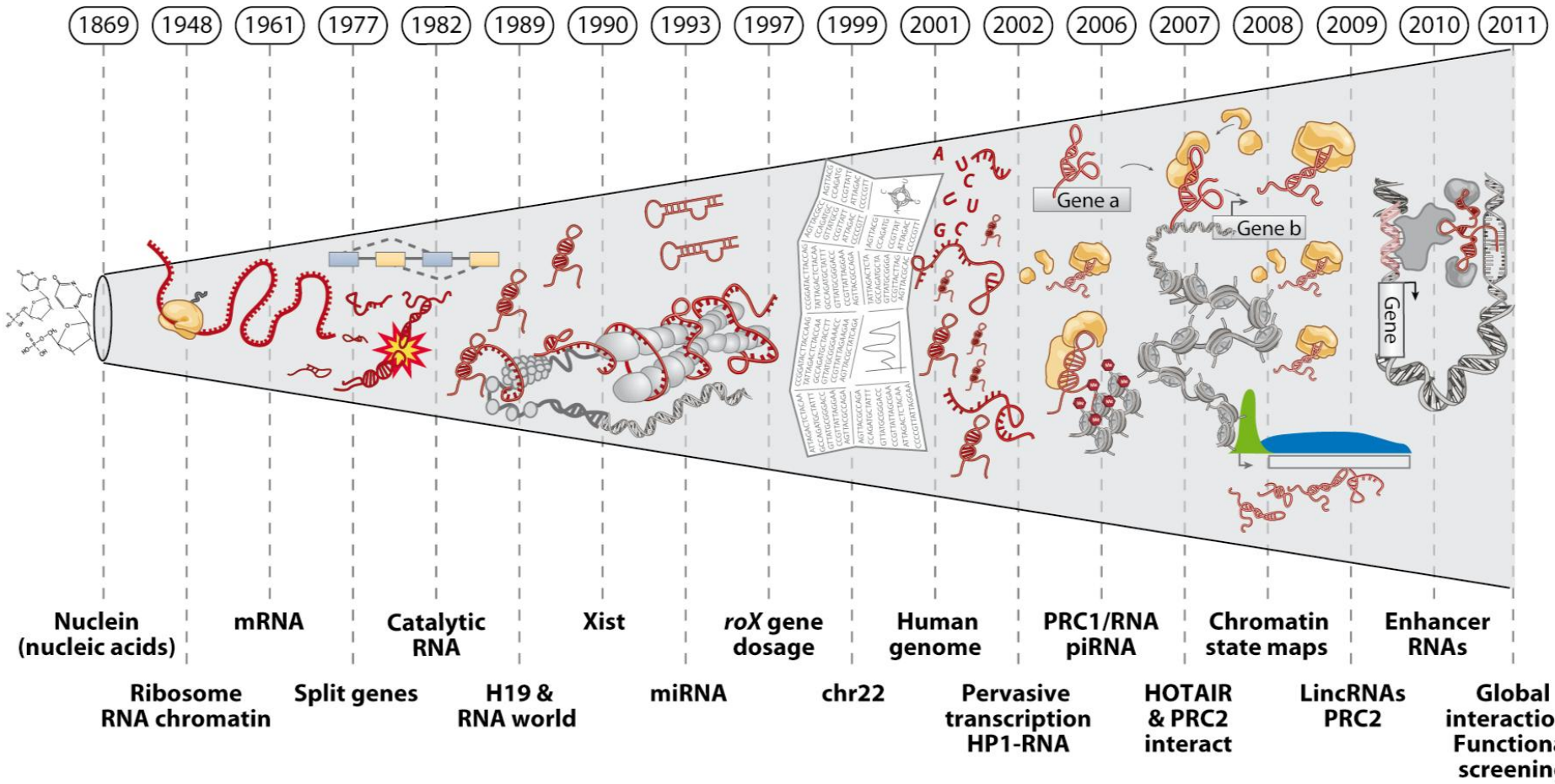
- Introduction to miRNA computational analysis (2-48)
- microRNA target identification (49-60)
- Functional characterization of miRNAs – Pathway analyses (61-69)
- Characterization of miRNA promoters and their TFs (70-78)
- Identification of miRNAs with important regulatory roles combining sRNA-Seq and RNA-Seq datasets (79-92)

Workshop material available in:
https://tinyurl.com/n3fjljb

Hellenic Pasteur Institute

Introduction to microRNA computational analysis

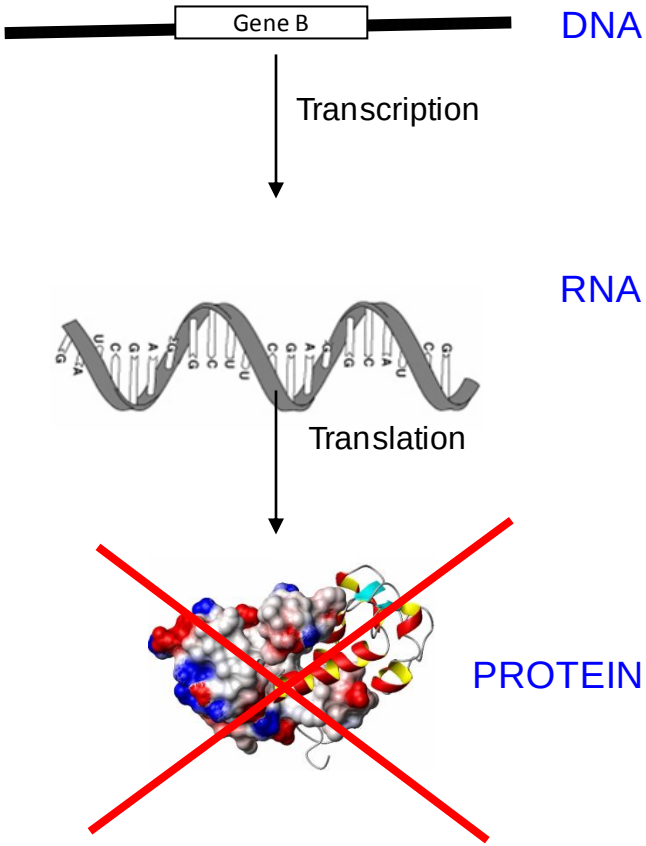
Hatzigeorgiou Artemis



Rinn JL, Chang HY. 2012.
Annu. Rev. Biochem. 81:145–66

Annual Reviews

What are microRNAs (miRNAs)?



miRNAs are about 22 nt long RNAs.

They post-transcriptionally regulate *protein coding* gene expression

MicroRNAs are involved in …

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

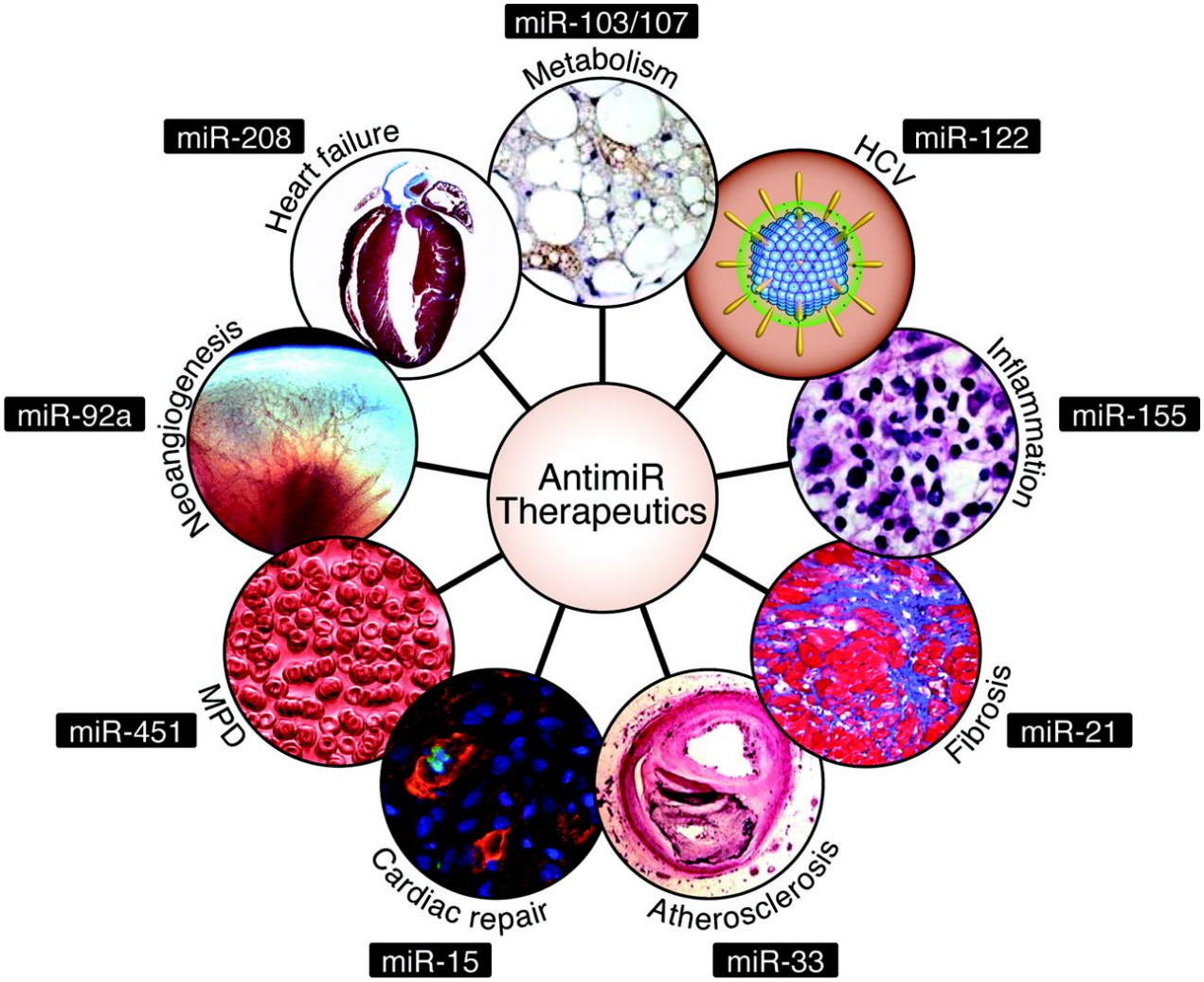
obesity

reproductive disorders

musculoskeletal disorders

periodontal pathologies

Specific miRNAs that are currently being pursued as clinical candidates.

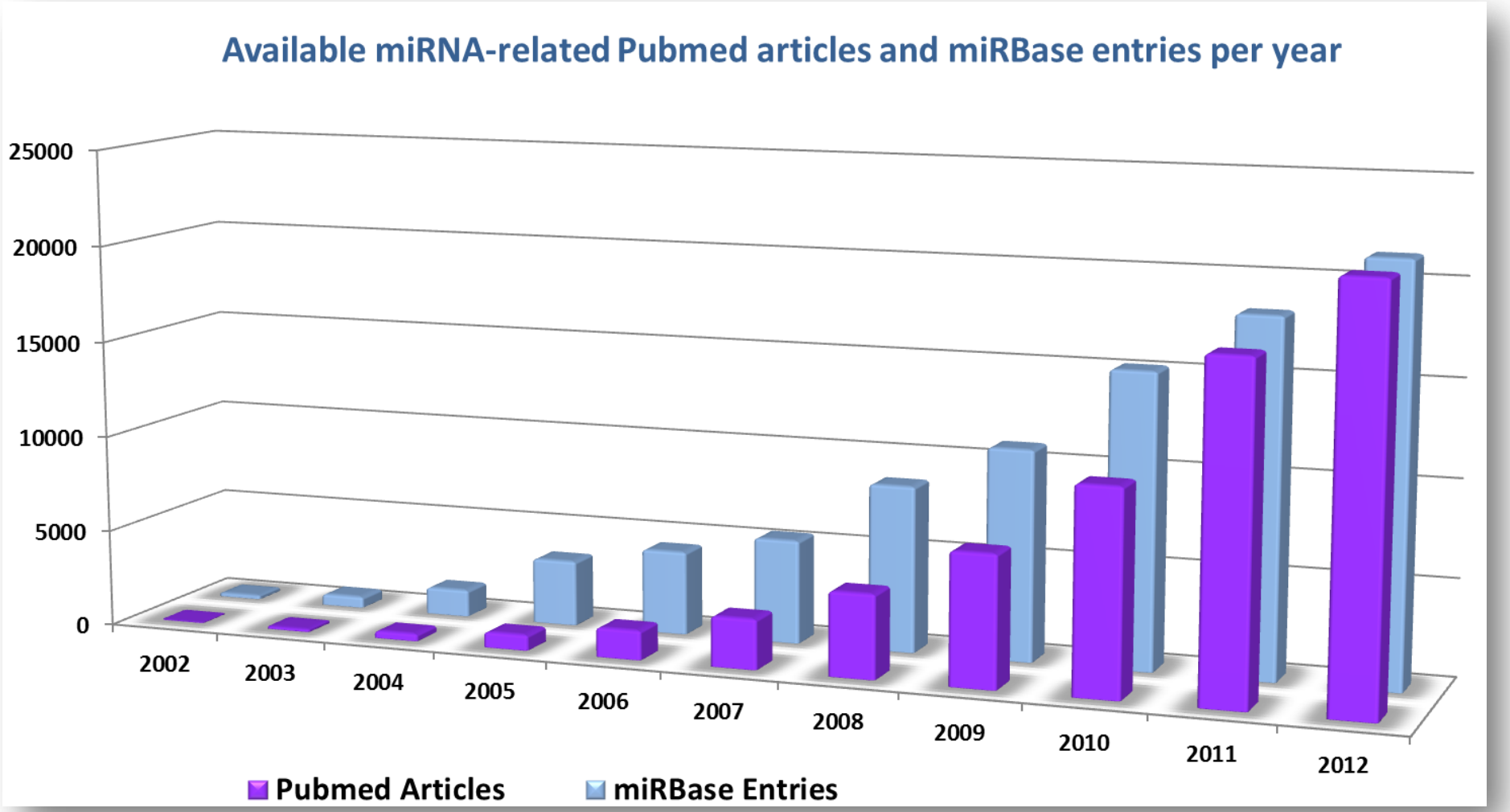


Eva van Rooij et al. Circ Res. 2012;110:496-507

American Heart Association.

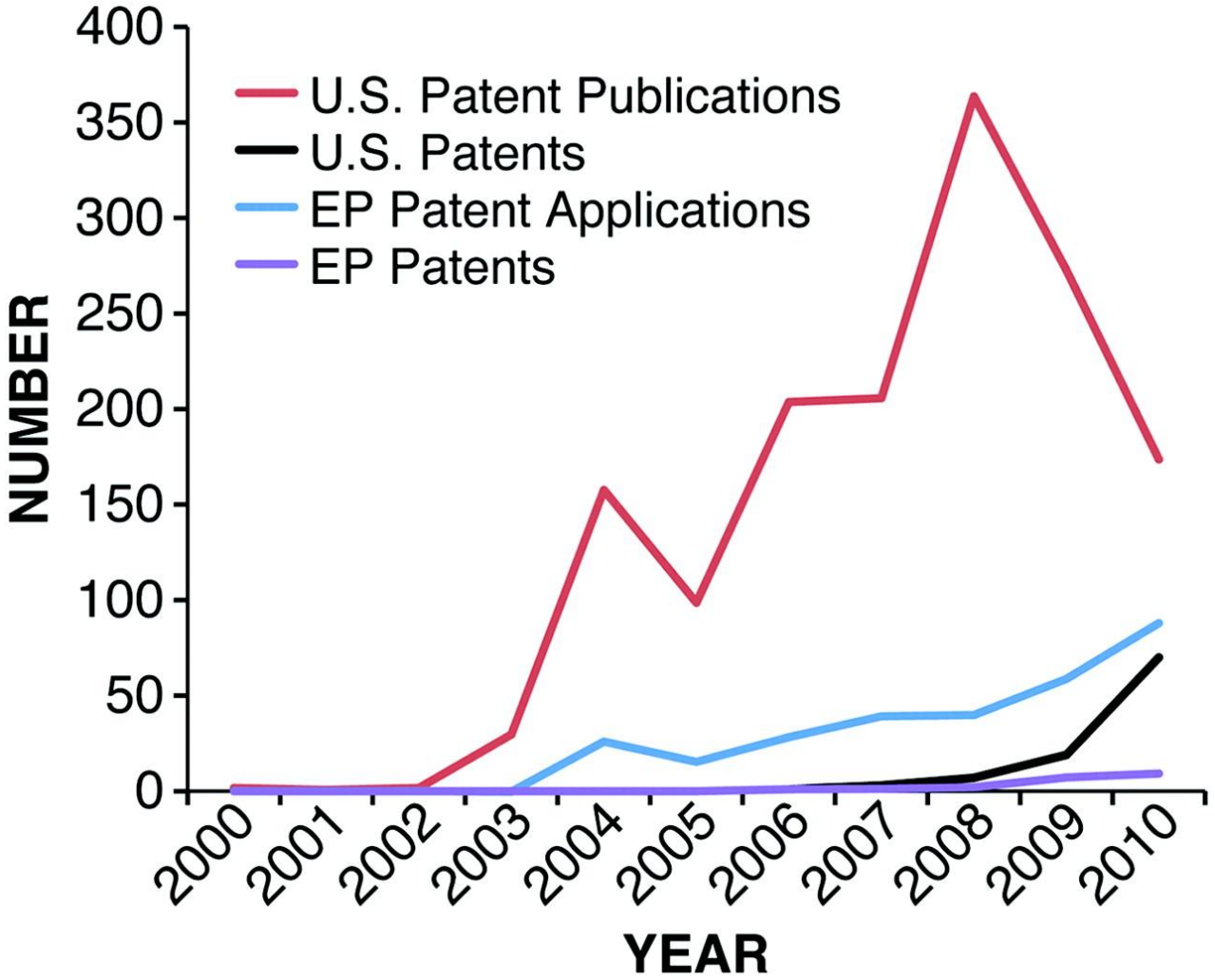
Copyright © American Heart Association, Inc. All rights reserved.

The microRNA world



***Vergoulis T, Vlachos IS et al. Nucleic Acids Res. 2012 January**; 40(D1)*

Annual number of US and European published patent applications and issued patents related to miRNAs and their applications.

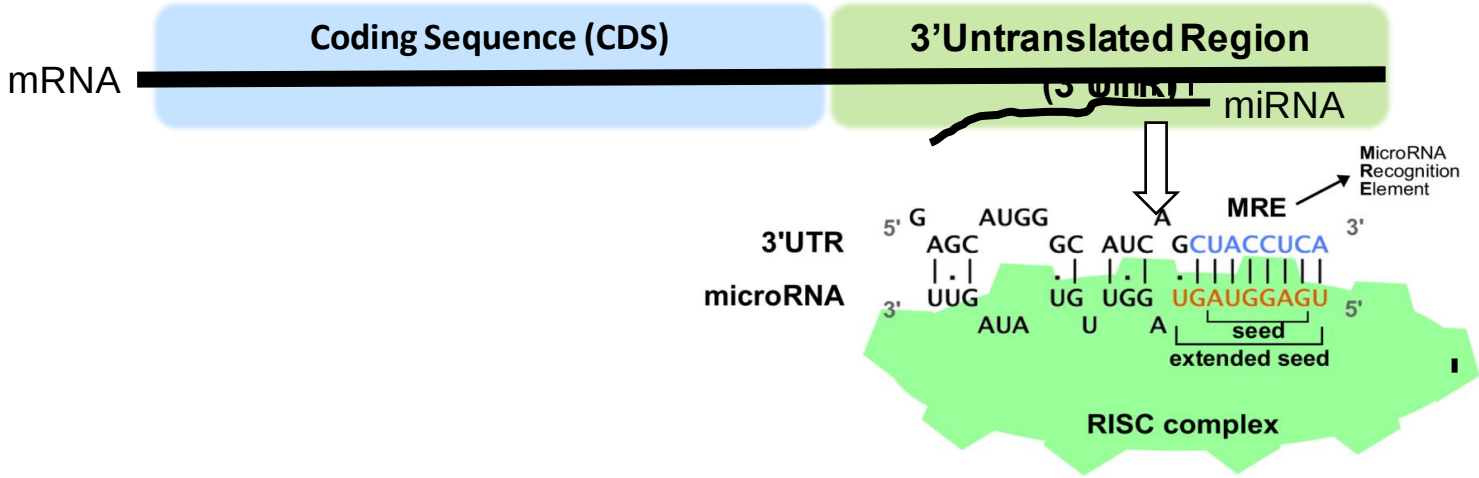


Eva van Rooij et al. Circ Res. 2012;110:496-507

American Heart Association.

Copyright © American Heart Association, Inc. All rights reserved.

MiRNA target prediction



Identifying a microRNA-recognition element by comparing the degree of complementarity of a microRNA nucleotide sequence to an mRNA sequence

Print | E-mail | Add to Marked List | Save to EndNote Web | Save to EndNote, RefMan, ProCite | more options

Patent Number(s): WO2005017111-A2; US2007026403-A1

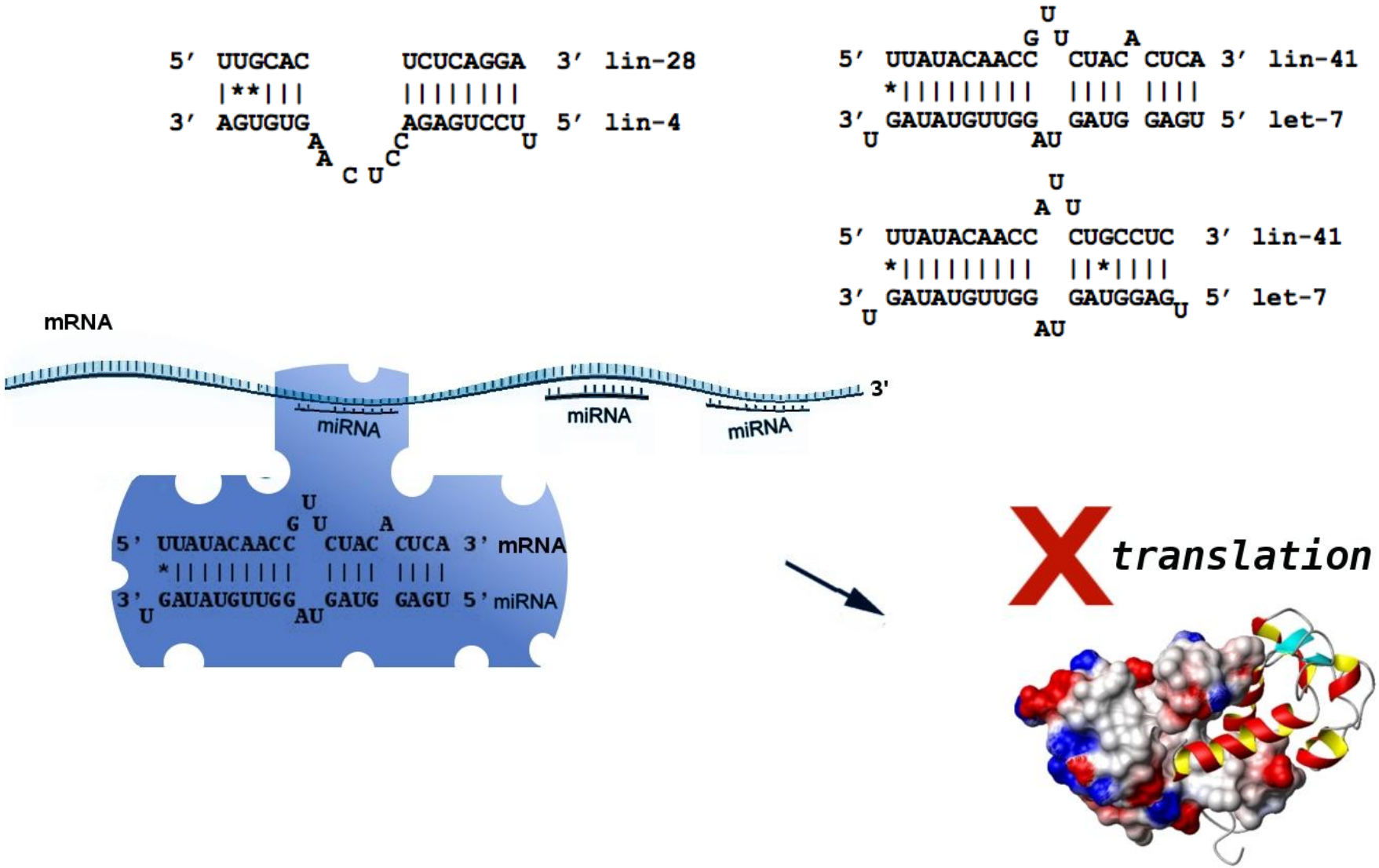
Inventor(s): HATZIGEORGIOU A G, MOURELATOS Z

Patent Assignee(s) and Codes(s):UNIV PENNSYLVANIA(UPEN-C)
HATZIGEORGIOU A G(HATZ-Individual)
MOURELATOS Z(MOUR-Individual)

Derwent Primary Accession Number: 2005-182352 [65]

Citing Patents: 2 **Articles Cited by Examiner:** 8

Computational identification of microRNA targets .



microRNAs bind the 3'UTR of mRNAs and repress translation

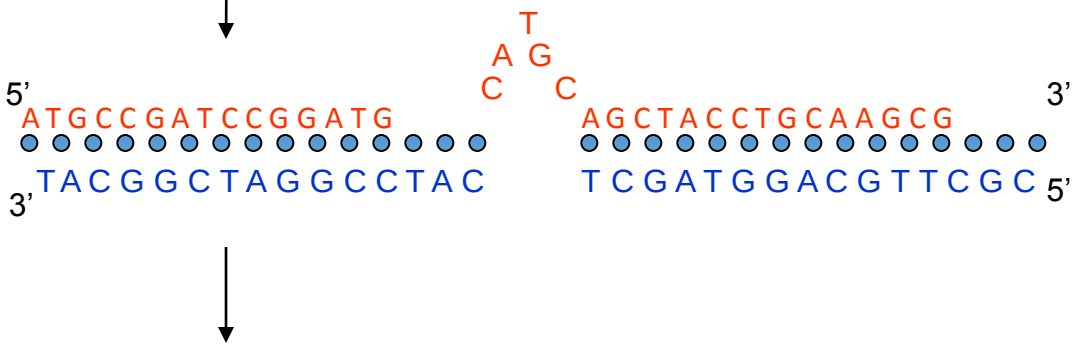
Computational identification of miRNA targets

…….ATGCCGATCCGGATGCATGCAGCTACCGCTAAGCGAATCGAACCG……. mRNA

1 35

ATGCCGATCCGGATGCATGCAGCTACCTGCAAGCG

Dynamic programming using free energies of dinucleotide pairs as scoring matrix.



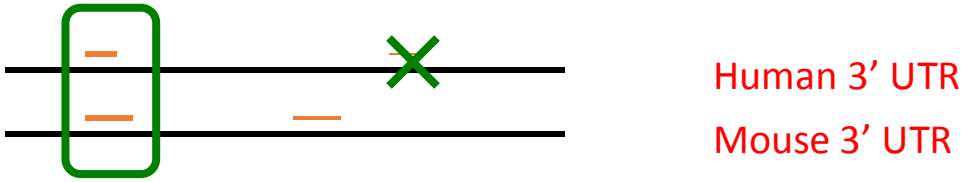
mRNA
microRNA

Minimum binding free energy score for mRNA window

Adding conservation

For each miRNA:

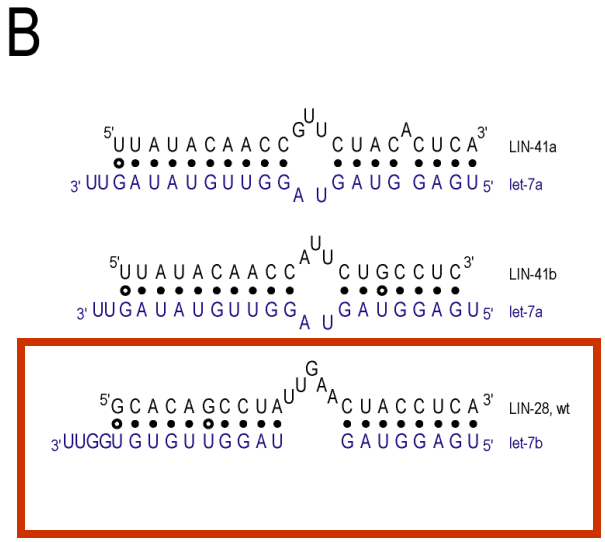
A) Keep targets conserved in human / mouse orthologs.

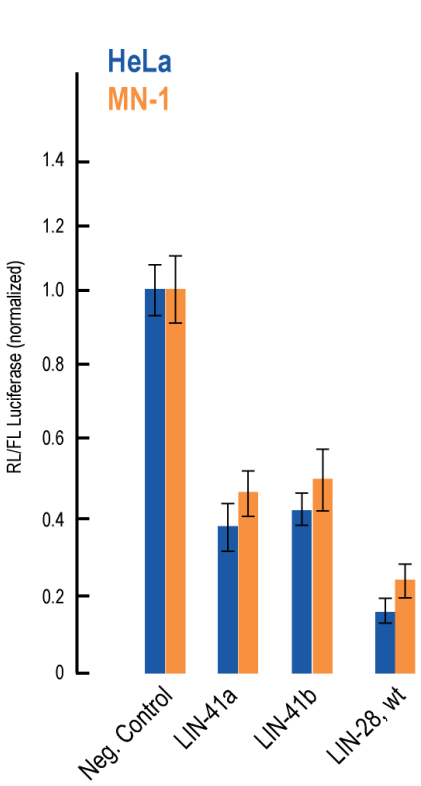


B) Sort all targets based on the minimum free energy binding score

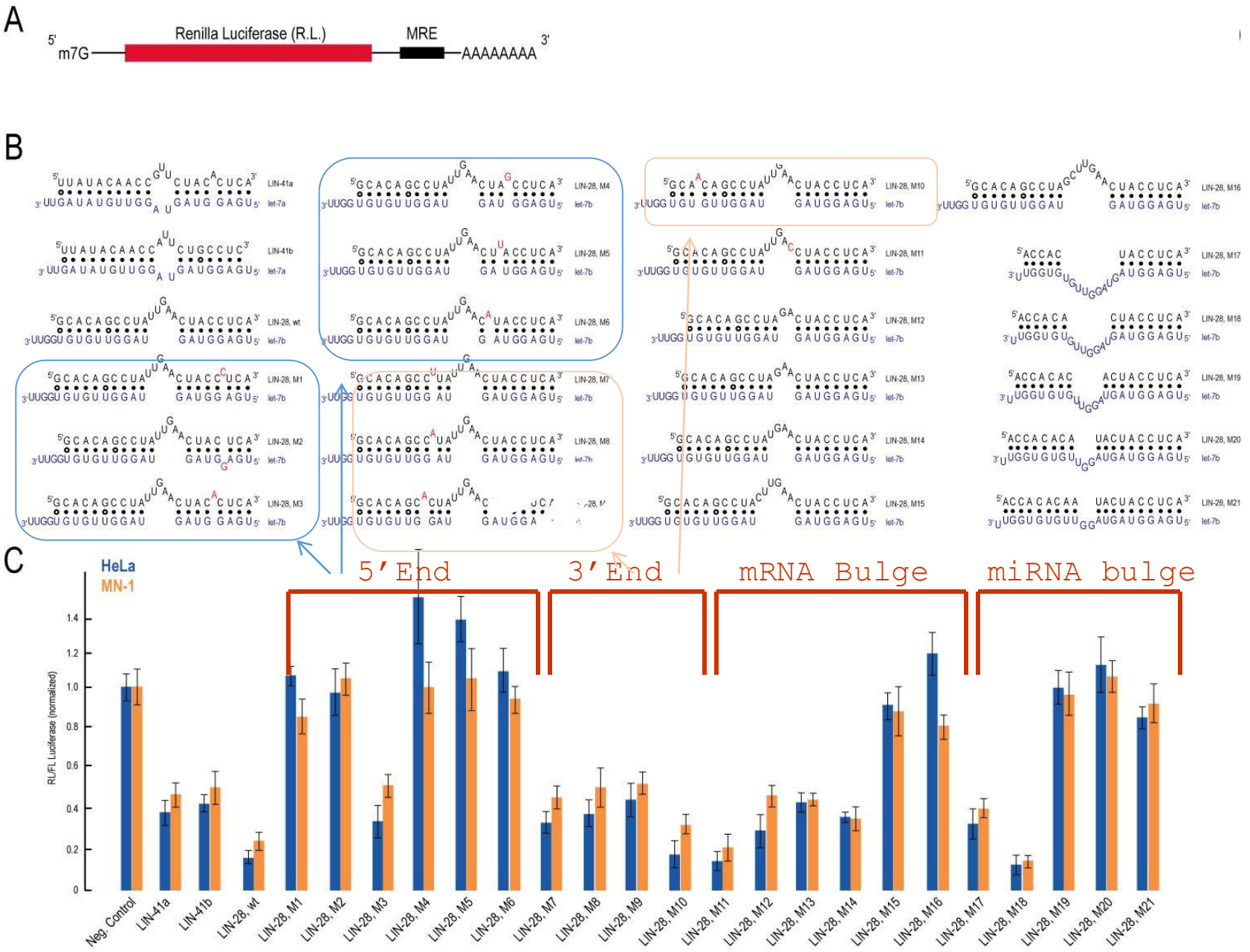
Top 13 targets selected for experiments.

Experimental identification





Experimental identification of miRNA binding rules



First Binding rules

- Binding on the 5’ prime of miRNA important
- Length of central bulge can change
- Binding on the 3’ prime of miRNA can be loose

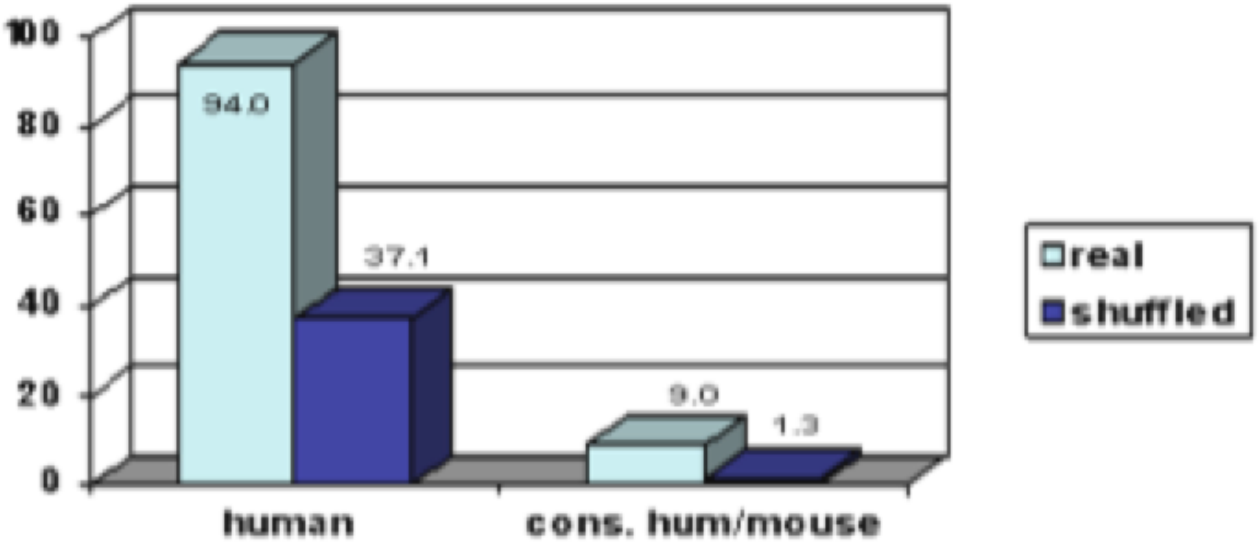
Statistical evaluation of miRNA binding rules

Signal : noise ratio

1. X = # total target site predictions for a set of real miRNAs
2. Y = # total target site predictions for a set of randomized miRNAs
3. X:Y is the signal:noise ratio which provides a measure of specificity

Example

1. X = 2000
2. Y = 1000
3. Signal:noise ratio is 2:1. For every 2 predicted targets, 1 is likely to be false (50% FPR).



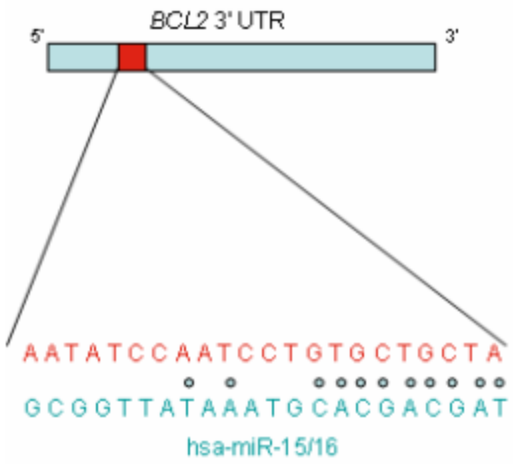
A combined computational-experimental approach predicts human microRNA targets

Marianthi Kiriakidou,[1,2] Peter T. Nelson,[1] Andrei Kouranov,[3] Petko Fitziev,[3,6] Costas Bouyioukos,[3] Zissimos Mourelatos,[1,7] and Artemis Hatzigeorgiou[3,4,5,8]

Departments of [1]Pathology, [2]Medicine, and [3]Genetics, School of Medicine, [4]Center for Bioinformatics, and [5]Computer and Information Science, School of Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

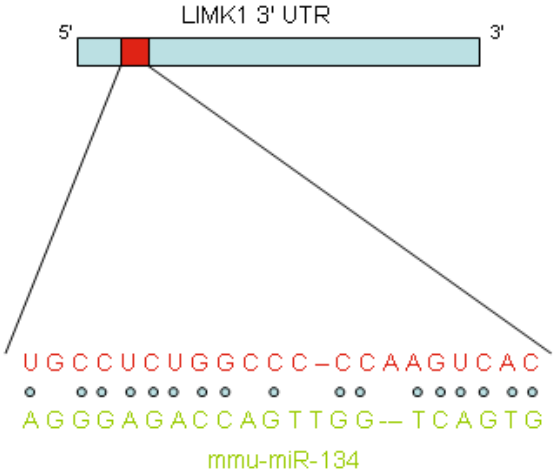
miRNA:target interaction categories

5'-dominant site



- Perfect base pairing to at least 7 nucleotides starting from the first or second nucleotide at the 5’-end of the miRNA

3’-compensatory site

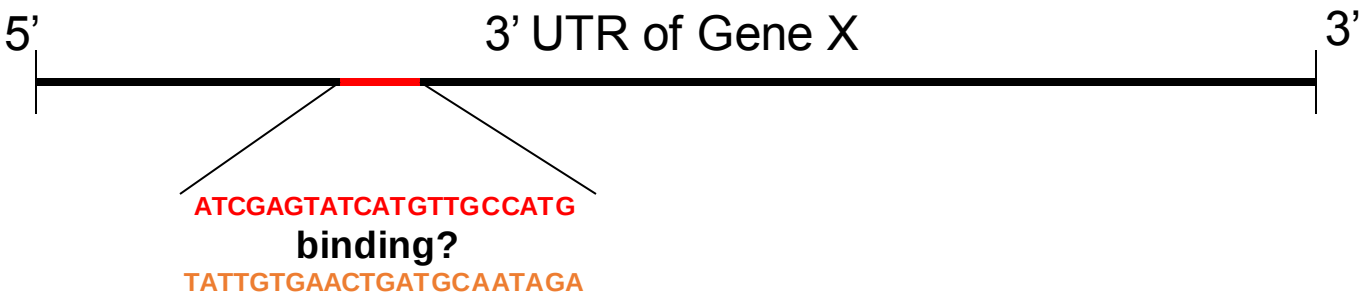


- Imperfect or shorter stretch of base pairing to at least 7 nucleotides starting from the first or second nucleotide at the 5’-end of the miRNA

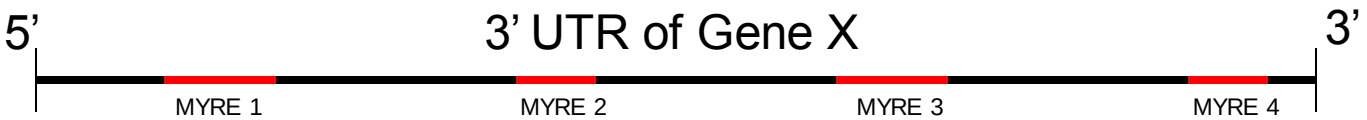
- Extensive binding to the 3’-end of the miRNA in order to compensate for the weaker binding to the miRNA 5’-end

miRNA target sites & miRNA:mRNA

- *Predicting individual miRNA:site interactions*



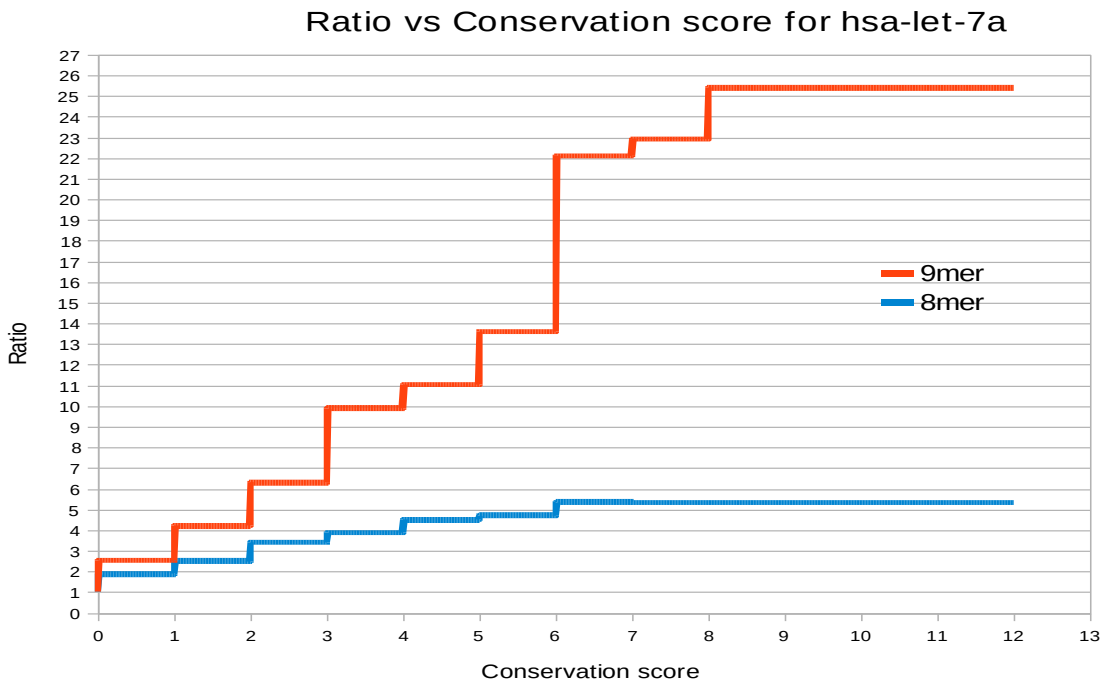
- *Providing a list of miRNA:gene interactions*



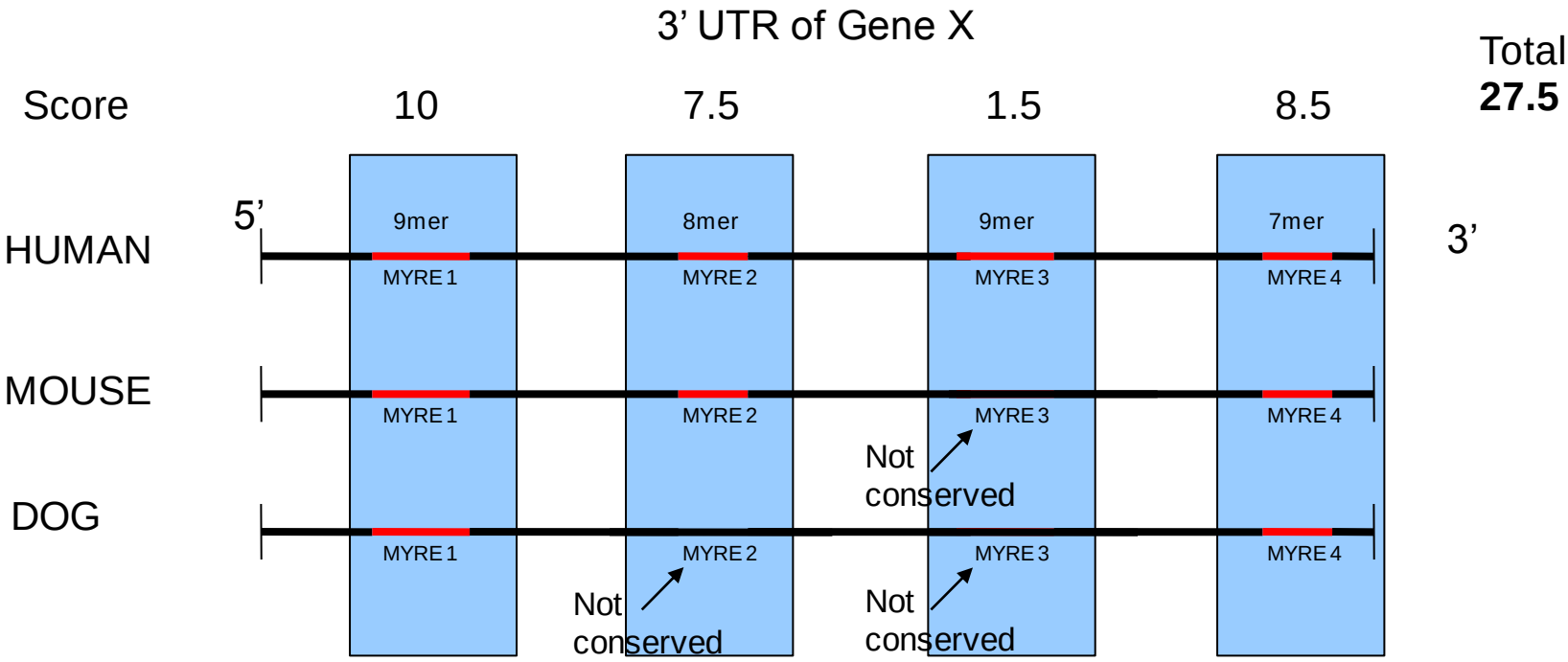
How likely is it that Gene X is a target of miRNA Y?

New microRNA target prediction : DIANA-microT 3.0

Calculating for each possible binding category conservation scores including 17 species.



Combined 3'UTR score



Weighted sum over of all putative target scores per 3'UTR.

BMC Bioinformatics

BioMed Central

Research article

Open Access

Accurate microRNA target prediction correlates with protein repression levels

Manolis Maragkakis*†[1,2], Panagiotis Alexiou†[1,3], Giorgio L Papadopoulos[1], Martin Reczko[1,4], Theodore Dalamagas[5], George Giannopoulos[5,6], George Goumas[7], Evangelos Koukis[7], Kornilios Kourtis[7], Victor A Simossis[1],

Evaluation of miRNA target prediction programs

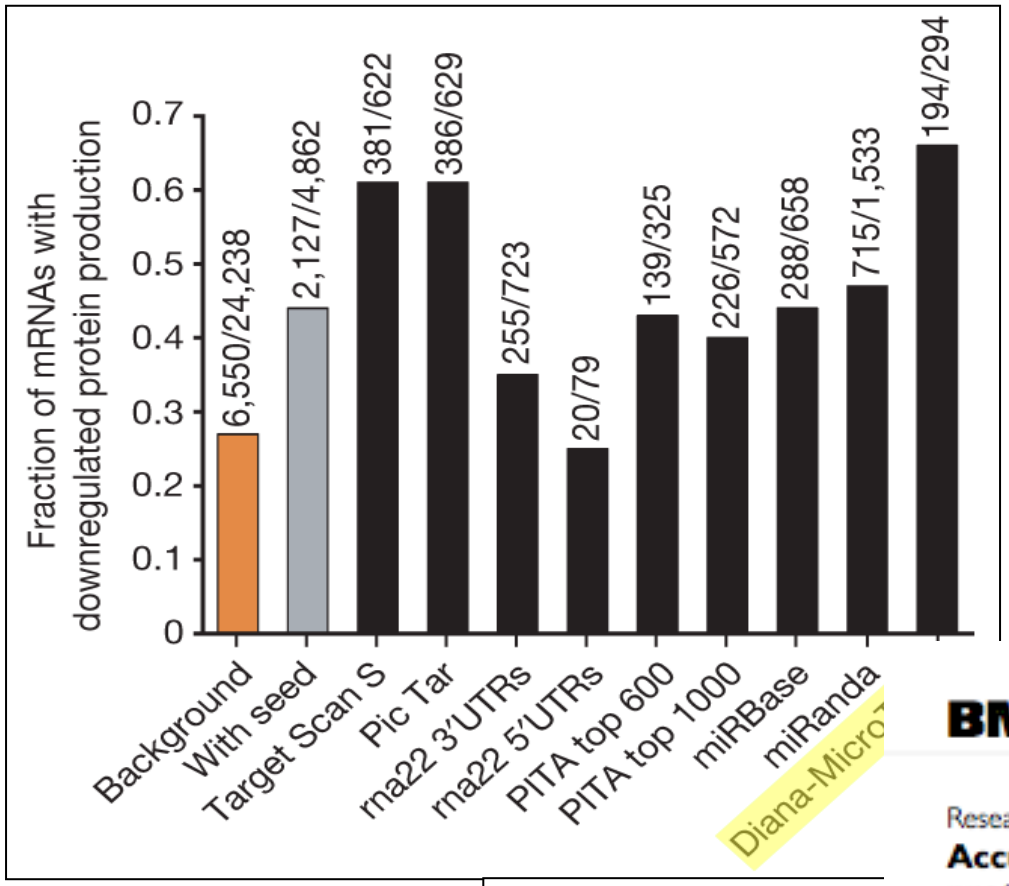
until 2008:

Statistically based on the ratio of real and artificial miRNA targets
Experimental verified targets

2008 :
pSilac and Silac approach:
Measured the effect of miRNAs in the cell on protein levels.
(Selbach et. al. & Baek D., Nature, 2008)

Widespread changes in protein synthesis induced by microRNAs

Matthias Selbach[1], Björn Schwanhäusser[1]*, Nadine Thierfelder[1]*, Zhuo Fang[1], Raya Khanin[2] & Nikolaus Rajewsky[1]



NATURE, 2008

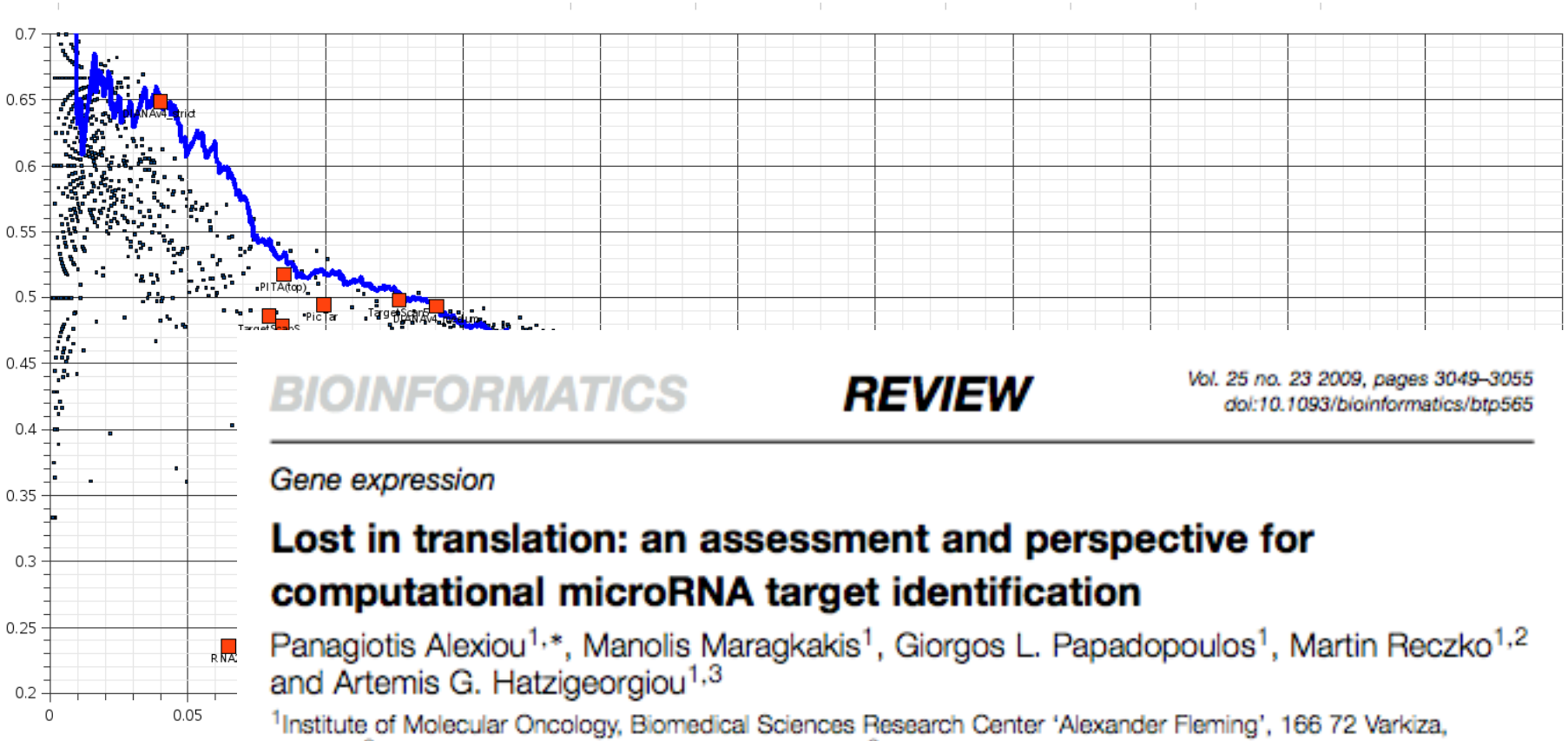
BMC Bioinformatics

Research article

Accurate microRNA target prediction correlates with protein repression levels

Manolis Maragkakis*†[1,2], Panagiotis Alexiou†[1,3], Giorgio L Papado
Martin Reczko[1,4], Theodore Dalamagas[5], George Giannopoulos[5,6],
George Goumas[7], Evangelos Koukis[7], Kornilios Kourtis[7], Victor A
Praveen Sethupathy[8], Thanasis Vergoulis[5,6], Nectarios Koziris[7],
Timos Sellis[5,6], Panagiotis Tsanakas[7] and Artemis G Hatzigeorgiou

Does more means also better ?



BIOINFORMATICS REVIEW

Vol. 25 no. 23 2009, pages 3049–3055
doi:10.1093/bioinformatics/btp565

Gene expression

Lost in translation: an assessment and perspective for computational microRNA target identification

Panagiotis Alexiou[1,*], Manolis Maragkakis[1], Giorgos L. Papadopoulos[1], Martin Reczko[1,2] and Artemis G. Hatzigeorgiou[1,3]

[1]Institute of Molecular Oncology, Biomedical Sciences Research Center 'Alexander Fleming', 166 72 Varkiza, Greece, [2]Synaptic Ltd., 700 13 Heraklion, Greece and [3]Computer and Information Sciences, University of Pennsylvania, 19104-6391 Philadelphia, USA

Received on June 16, 2009; revised on September 24, 2009; accepted on September 27, 2009

Advance Access publication September 29, 2009

Associate Editor: Jonathan Wren

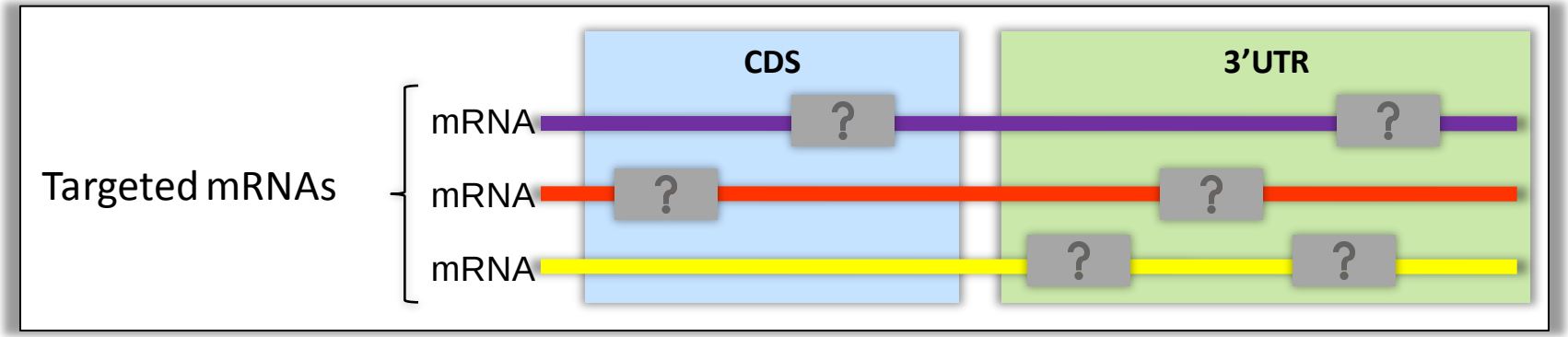
Until now

Experimentally identified binding sites

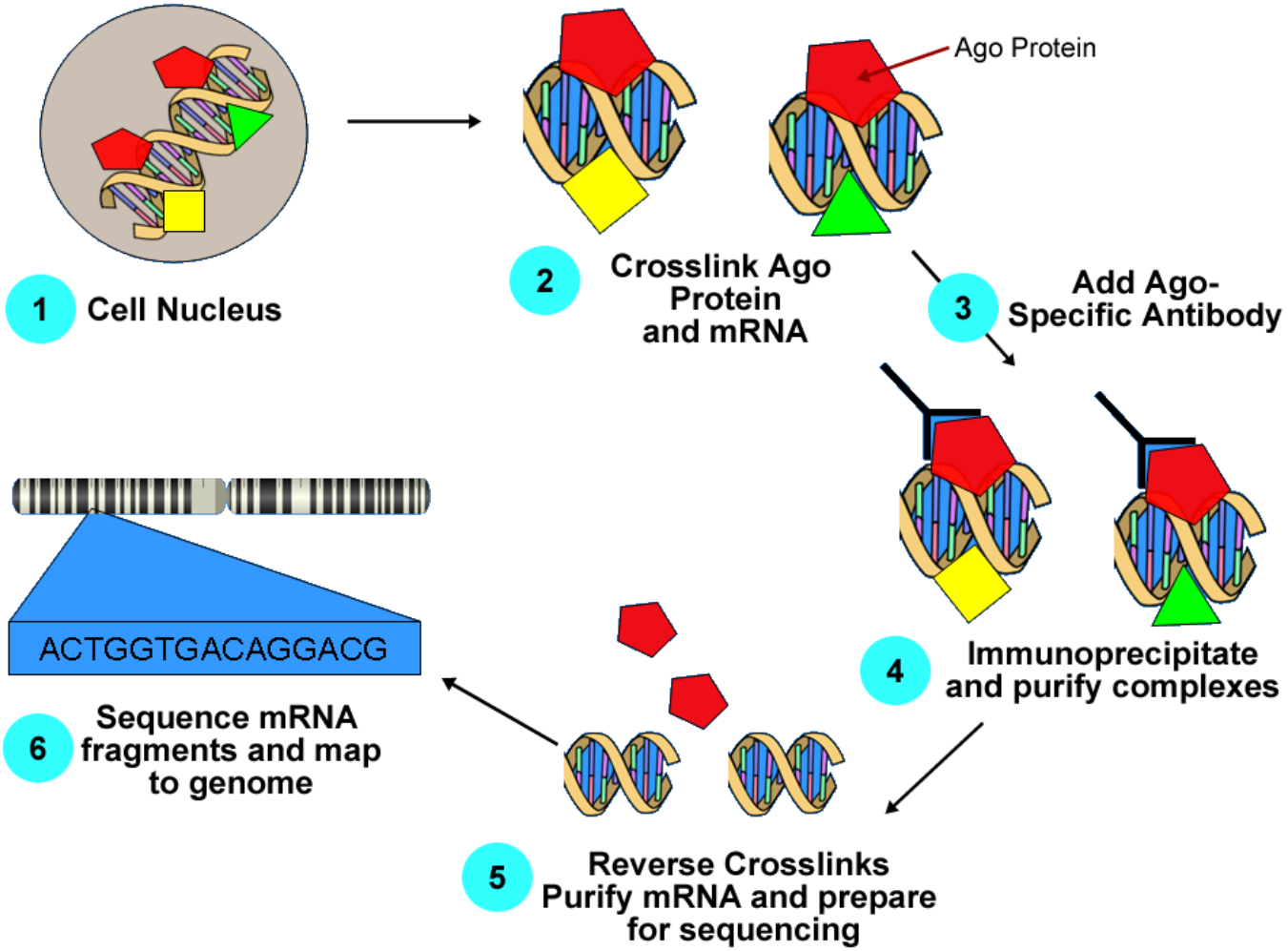
- Limited number -> cannot extract statistically significant features
- Specific for few miRNAs

Microarrays (mRNA level) or proteomics (protein level)

- Target genes are specified but binding sites remain unknown



CLIP data (Chi *et al.* 2009, Hafner *et al*. 2010)



Note

PAR-CLIP (Hafner et al. 2010) , HITS-CLIP (Chi et al. 2009).

T->C mutation on the tags specifies binding sites within a region of 5 nts.

Before

Experimentally identified binding sites

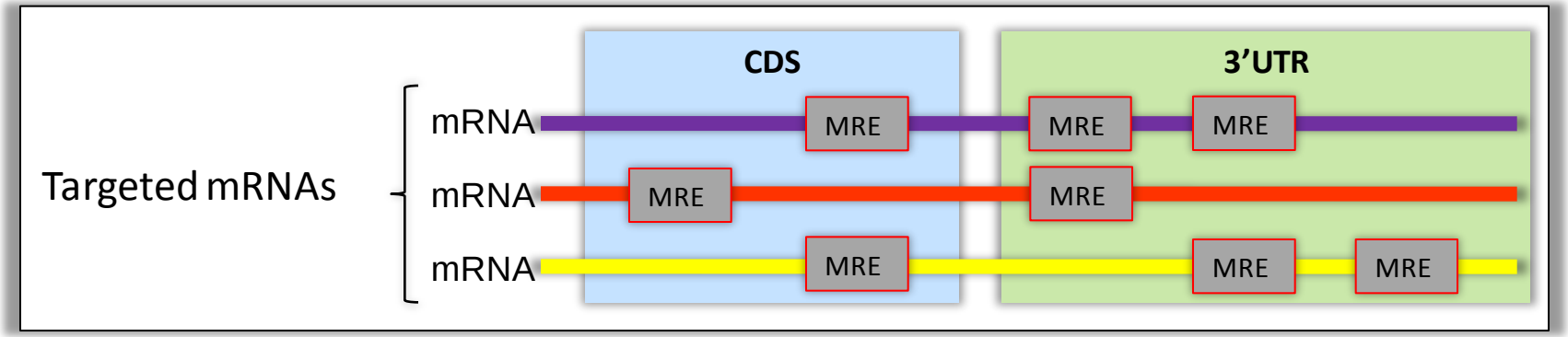
- Limited number -> cannot extract statistically significant features
- Specific for few miRNAs

Microarrays (mRNA level) or proteomics (protein level)

- Target genes are specified but binding sites remain unknown

Now - Sequencing data (PAR-CLIP, HITS-CLIP)

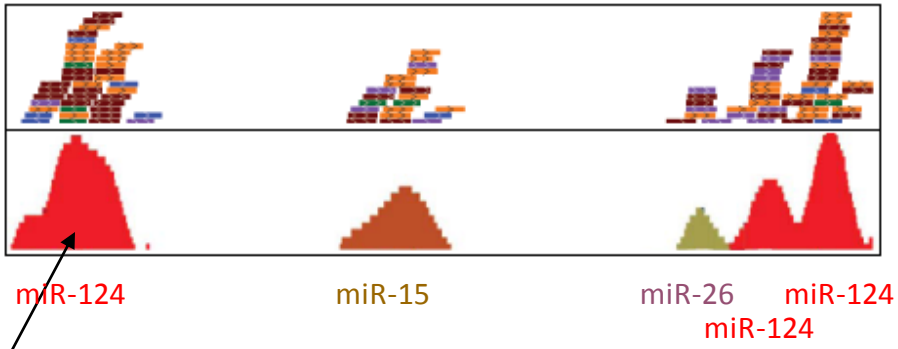
- Specify location of thousands of binding sites
- Enable statistical evaluation of several features



How CLIP data are processed

Sequence reads are mapped to genome

Peaks of reads are identified



Peaks indicate binding sites BUT do not specify by which miRNA

For this, the genomic location of the peak is aligned against all known miRNAs and the **best matching miRNA is chosen**

Out of 17310 peaks 5057 overlap with an MRE at the UTR and 6057 overlap with an MRE at the CDS.

Feature Extraction and Analysis

Identified binding sites are divided into 2 categories

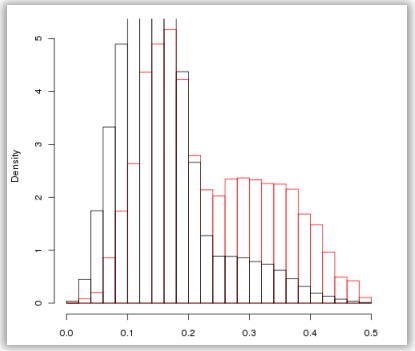
1. **Positive** sites - overlap with PAR-CLIP data
2. **Negative** sites - do not overlap with PAR-CLIP data

More than 150 features are tested to distinguish:

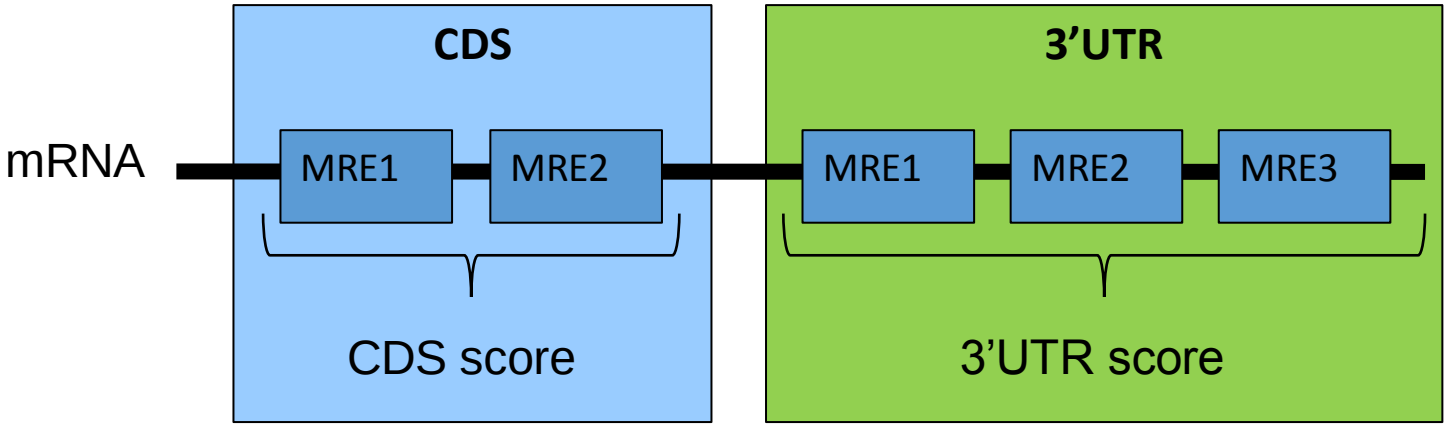
Positive binding sites (CLIP data) vs **Negative** binding sites

e.g.

Flanking AU content



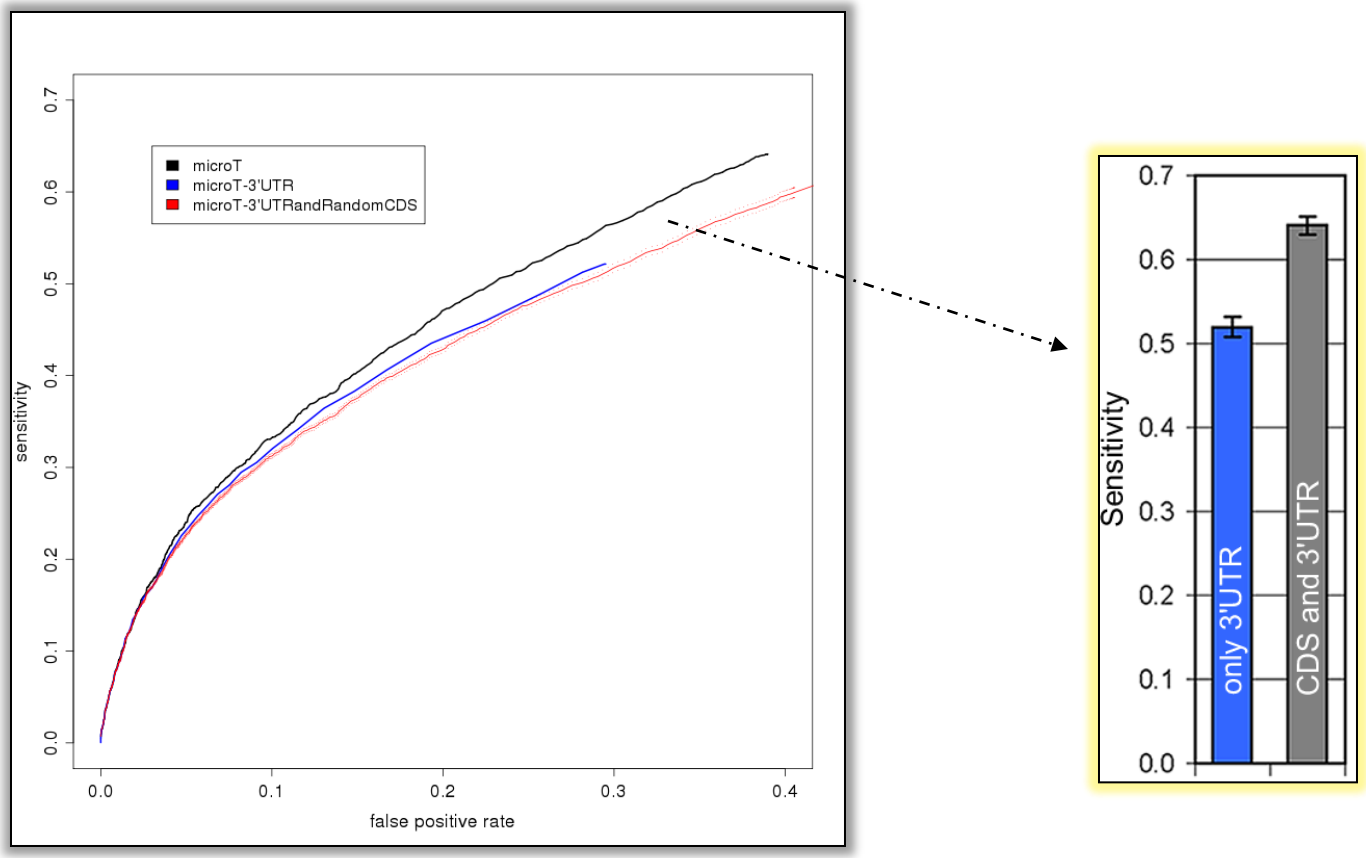
MicroT – CDS
A combination of MRE scores in CDS and 3'UTR regions



MRE scores in each region are combined to a region score

Integration of CDS score and 3 UTR score with a generalized linear model

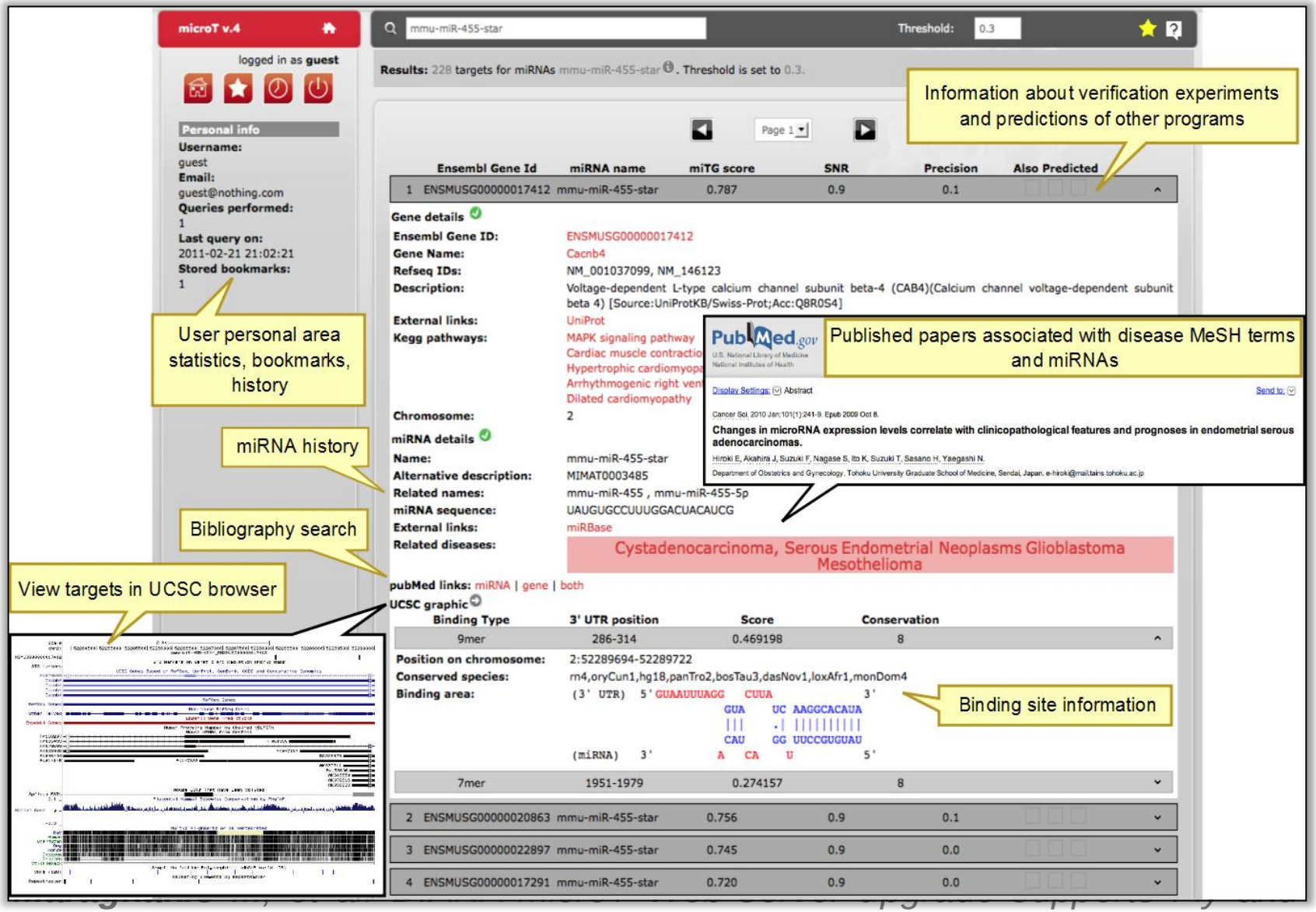
Does targeting in CDS improve prediction performance?



Including targeting in CDS <u>improves sensitivity by more than12%</u>

- 300 targets per miRNA that would be lost otherwise

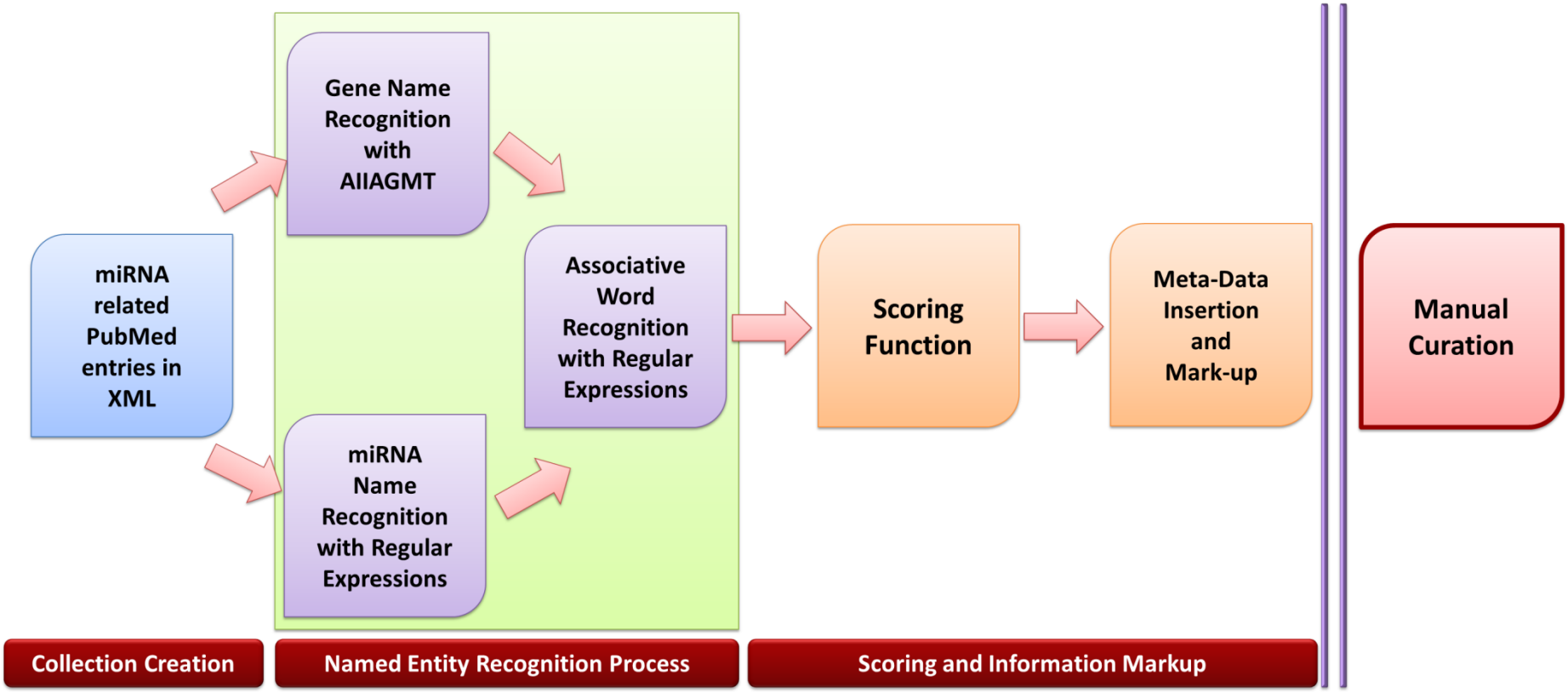
www.microrna.gr/microT-CDS



Worm miRNA target prediction and bibliographic miRNA to disease association. ***Nucleic Acids Research****, 2011.*

Experimental supported microRNA targets.

DIANA-TarBase
Semi – Automatic Curation Pipeline



Vergoulis, T. I. Vlachos *et al.* **Tarbase 6.0**: Capturing the Exponential Growth of miRNA Targets with Experimental Support. *Nucl. Acids Res.* 2012

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

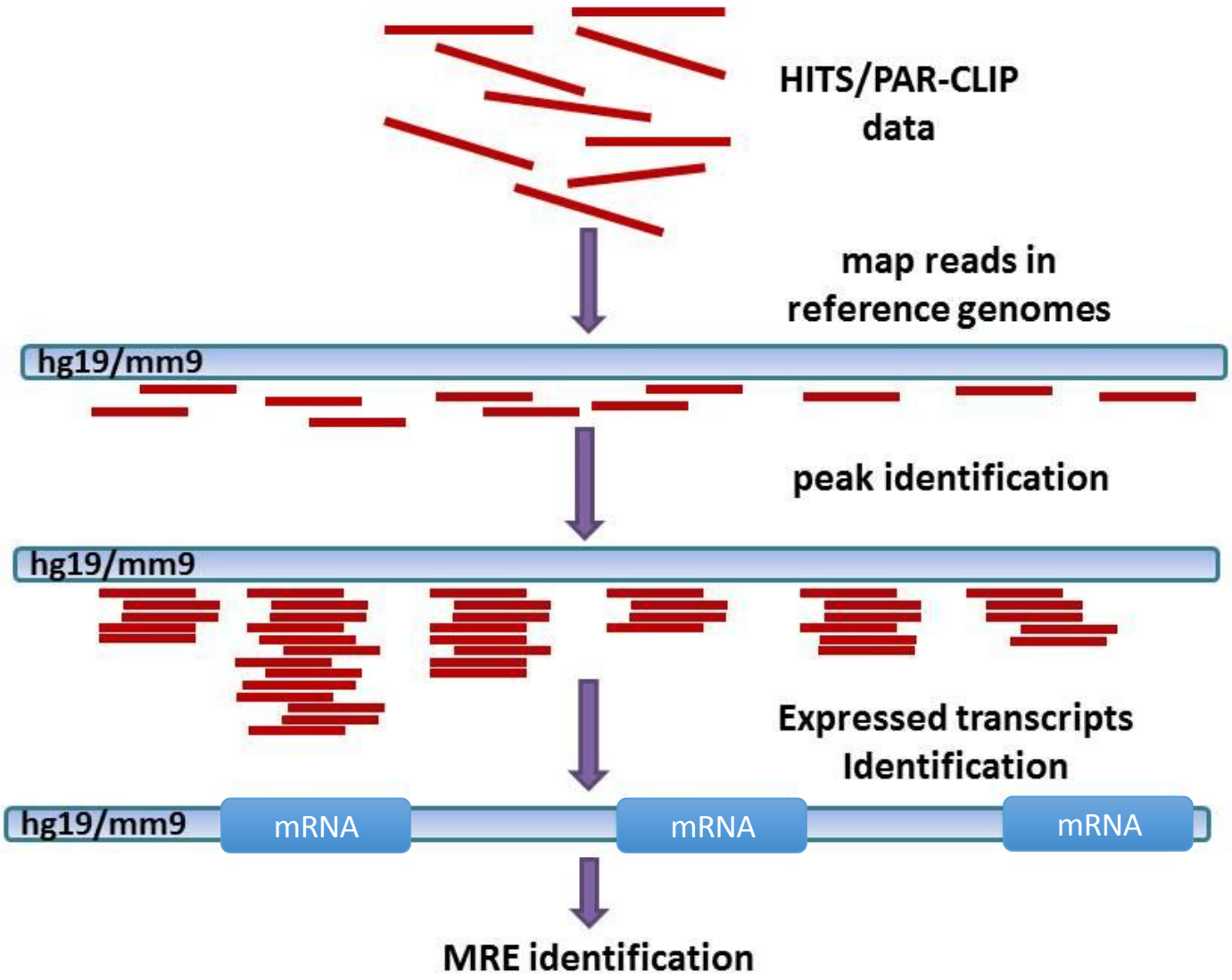
Identification of microRNAs expressed highly in pancreatic islet-like cell clusters differentiated from human embryonic stem cells. Type 1 diabetes is an autoimmune destruction of pancreatic islet beta cell disease, making it important to find a new alternative source of the islet beta cells to replace 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.

Wet Lab Determination of miRNA – Gene Interactions

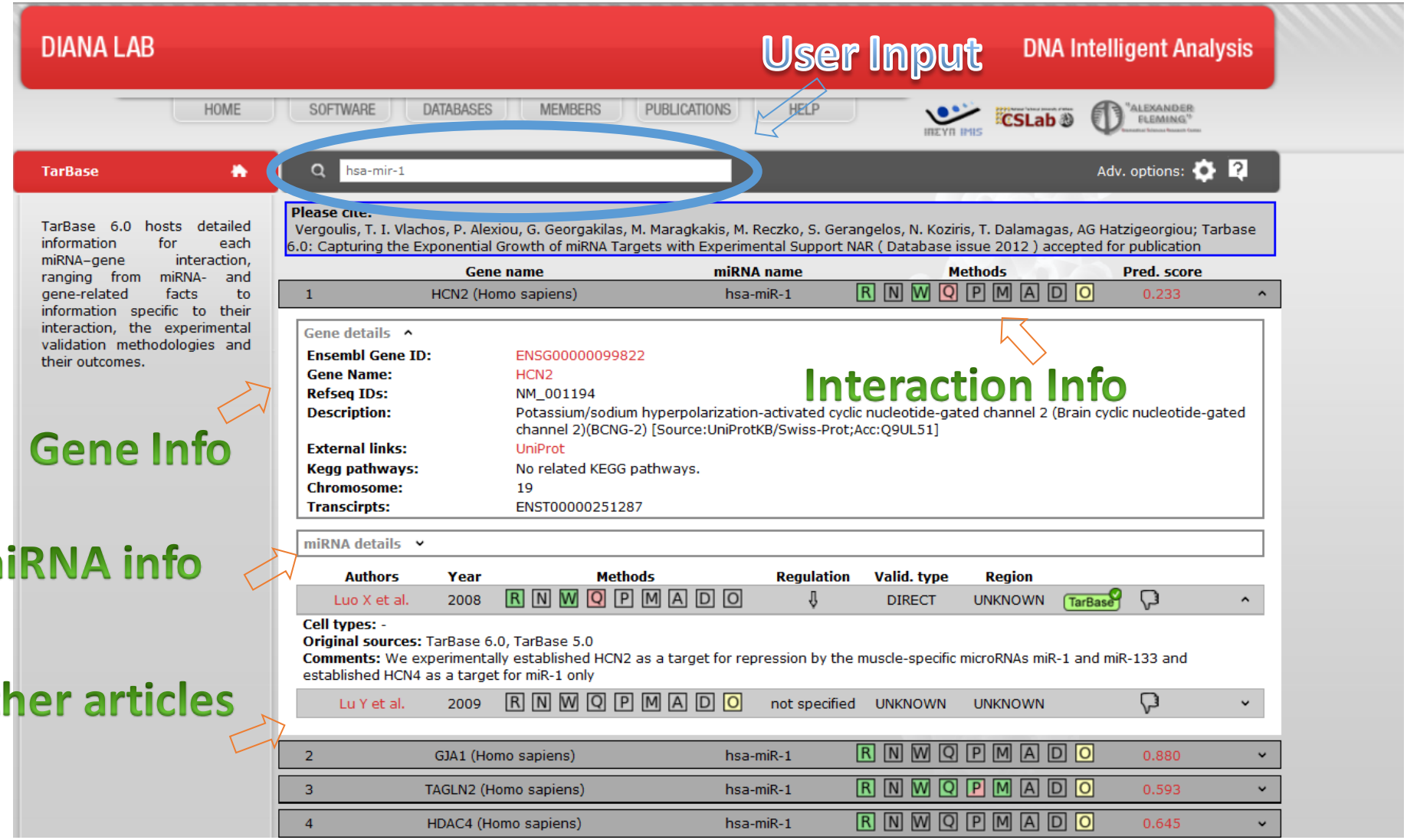
- Specific Techniques
  - Reporter genes
  - Northern blotting
  - qPCR
  - Western blotting
  - ELISA
  - Immunohistochemistry

- High Throughput Techniques
  - Microarrays
  - RNA-Seq
  - Proteomics (such as pSILAC)
  - CLIP-Seq (HITS-CLIP, PAR-CLIP, iCLIP)
  - CLASH
  - PARE-Seq
  - Degradome-Seq

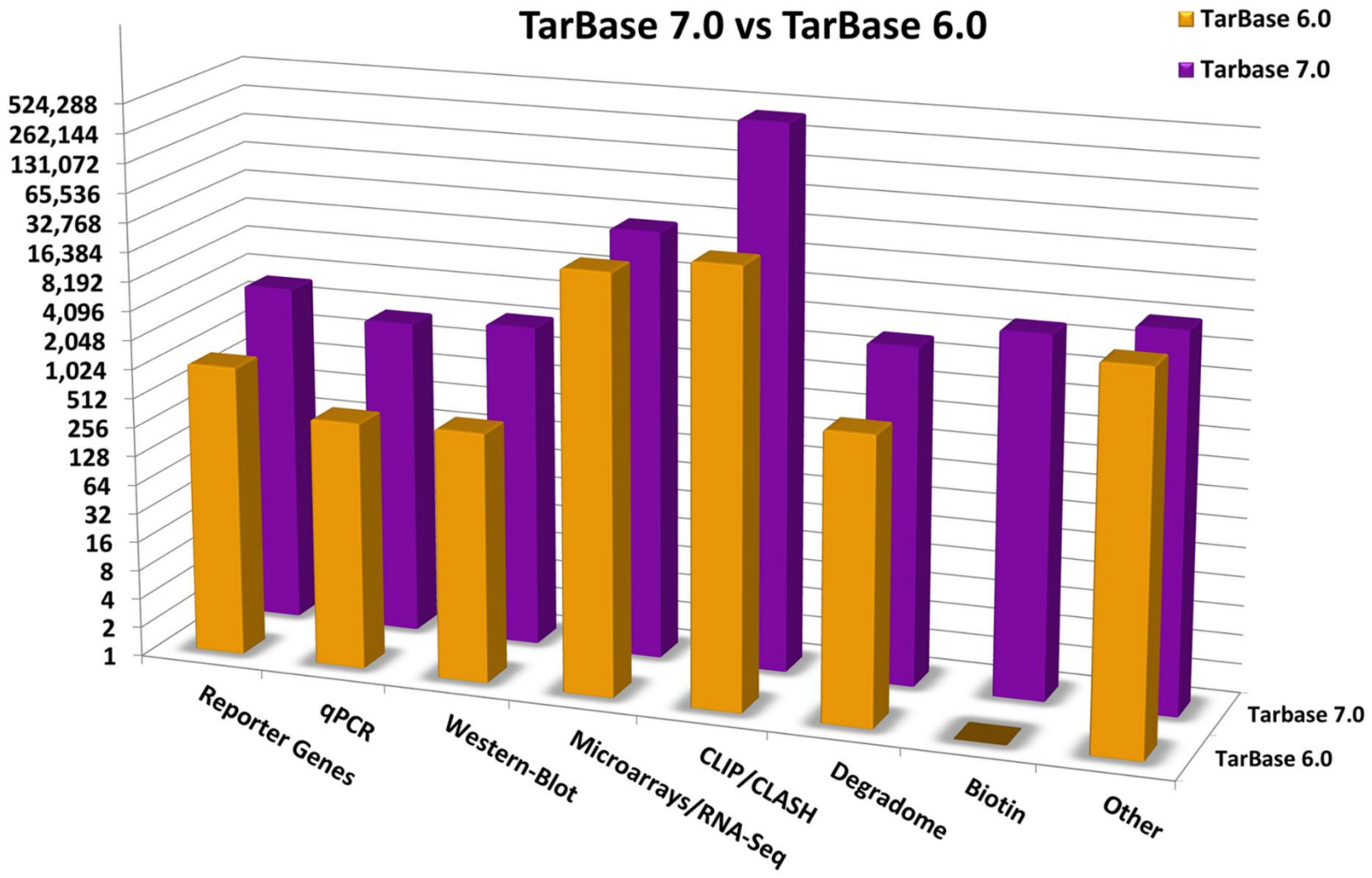
Analysis of PAR/HITS-CLIP data



TarBase v6.0 Interface



Growth of interactions per method



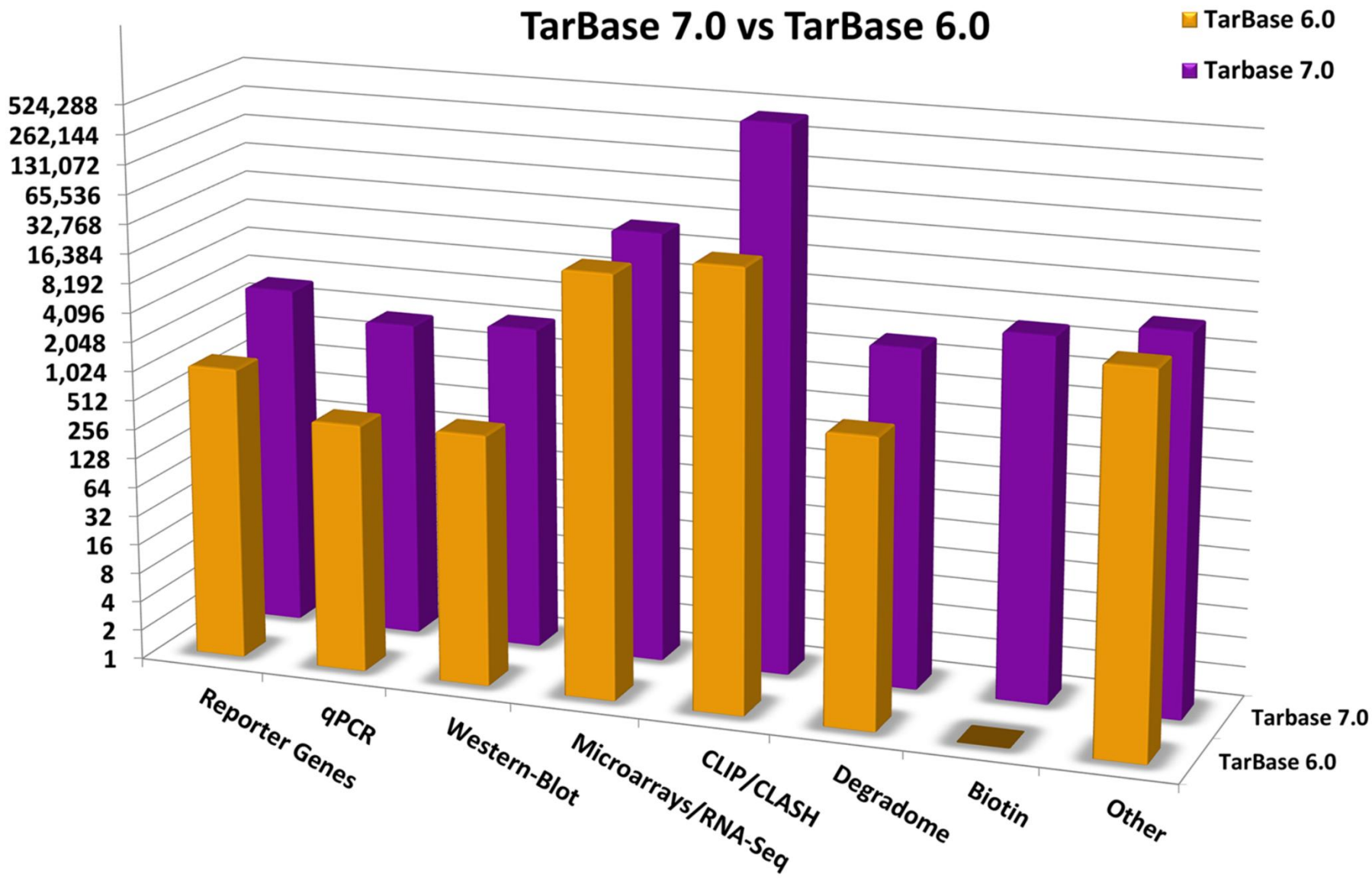
Database of experimentally supported targets: DIANA-TarBase

- Initially released in 2006
  - The first database to catalog published experimentally validates miRNA:gene interactions
- With more than 500,000 entries, the largest experimentally validated repository with miRNA:gene interactions
- Last update DIANA-TarBase v7 http://www.microrna.gr/tarbase

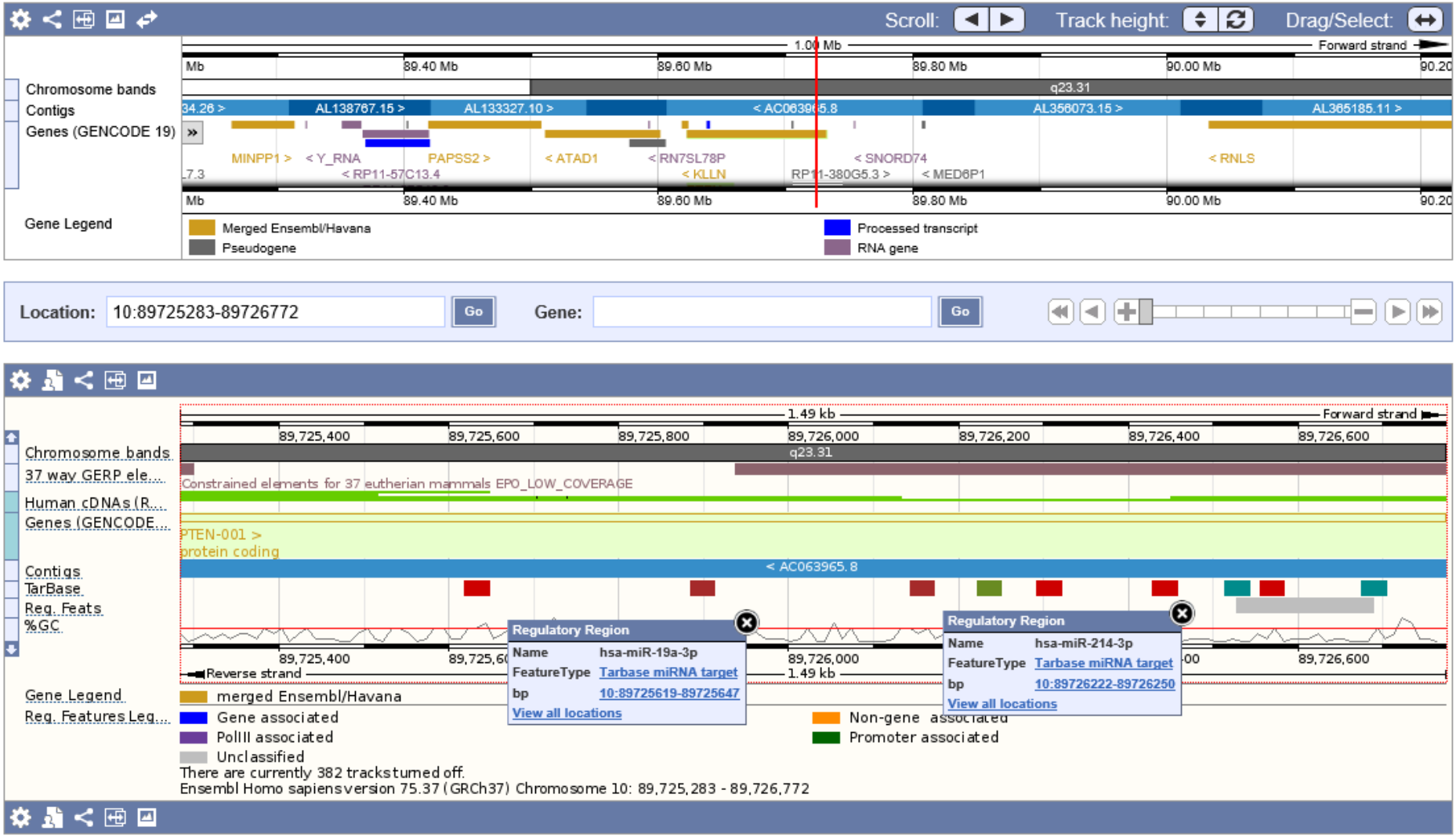
S. Vlachos, M. D. Paraskevopoulou, D. Karagkouni, G. Georgakilas, T. Vergoulis, I. Kanellos, I-L. Anastasopoulos, S. Maniou, K. Karathanou, D. Kalfakakou, A. Fevgas, T. Dalamagas and A. G. Hatzigeorgiou.

DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. ***Nucl. Acids Res****. (2014)*

Growth of interactions per method

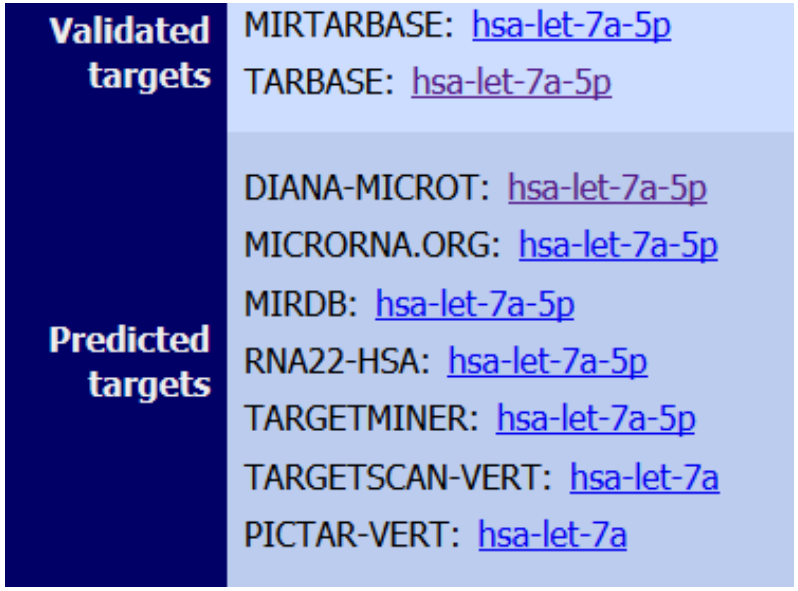


Integration in ENSEMBL, the European Browser for Genomes in EBI



miRBase

- Interconnects also entries with external resources:



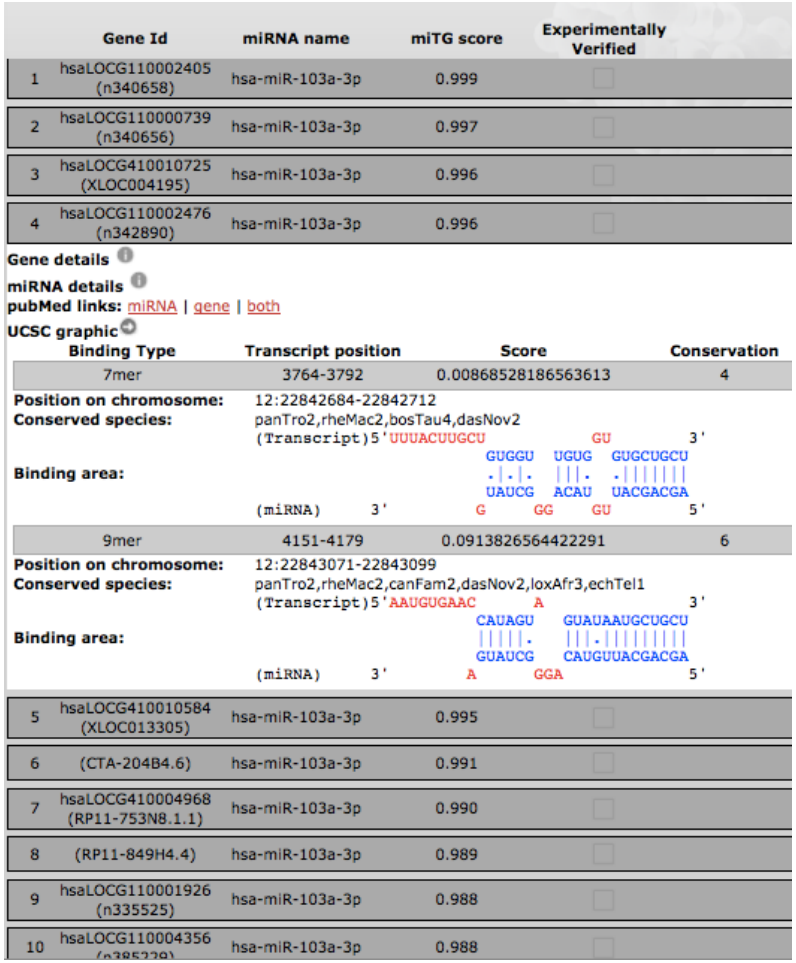
Long Non Coding RNAs

LncBase is the largest available repository of miRNA LNC RNA interactions

- The **Experimental Module** contains more than 5,000 interactions between 2,958 lncRNAs and 120 miRNAs.

- The **Prediction Module** contains detailed information between 56,097 lncRNAs and 3,078 miRNAs.

Integration into RNAcentral (EBI)



	Gene Id	miRNA name	miTG score	Experimentally Verified
1	hsaLOCG110002405 (n340658)	hsa-miR-103a-3p	0.999	
2	hsaLOCG110000739 (n340656)	hsa-miR-103a-3p	0.997	
3	hsaLOCG410010725 (XLOC004195)	hsa-miR-103a-3p	0.996	
4	hsaLOCG110002476 (n342890)	hsa-miR-103a-3p	0.996	

Gene details
miRNA details
pubMed links: miRNA | gene | both
UCSC graphic

Binding Type	Transcript position	Score	Conservation
7mer	3764-3792	0.00868528186563613	4

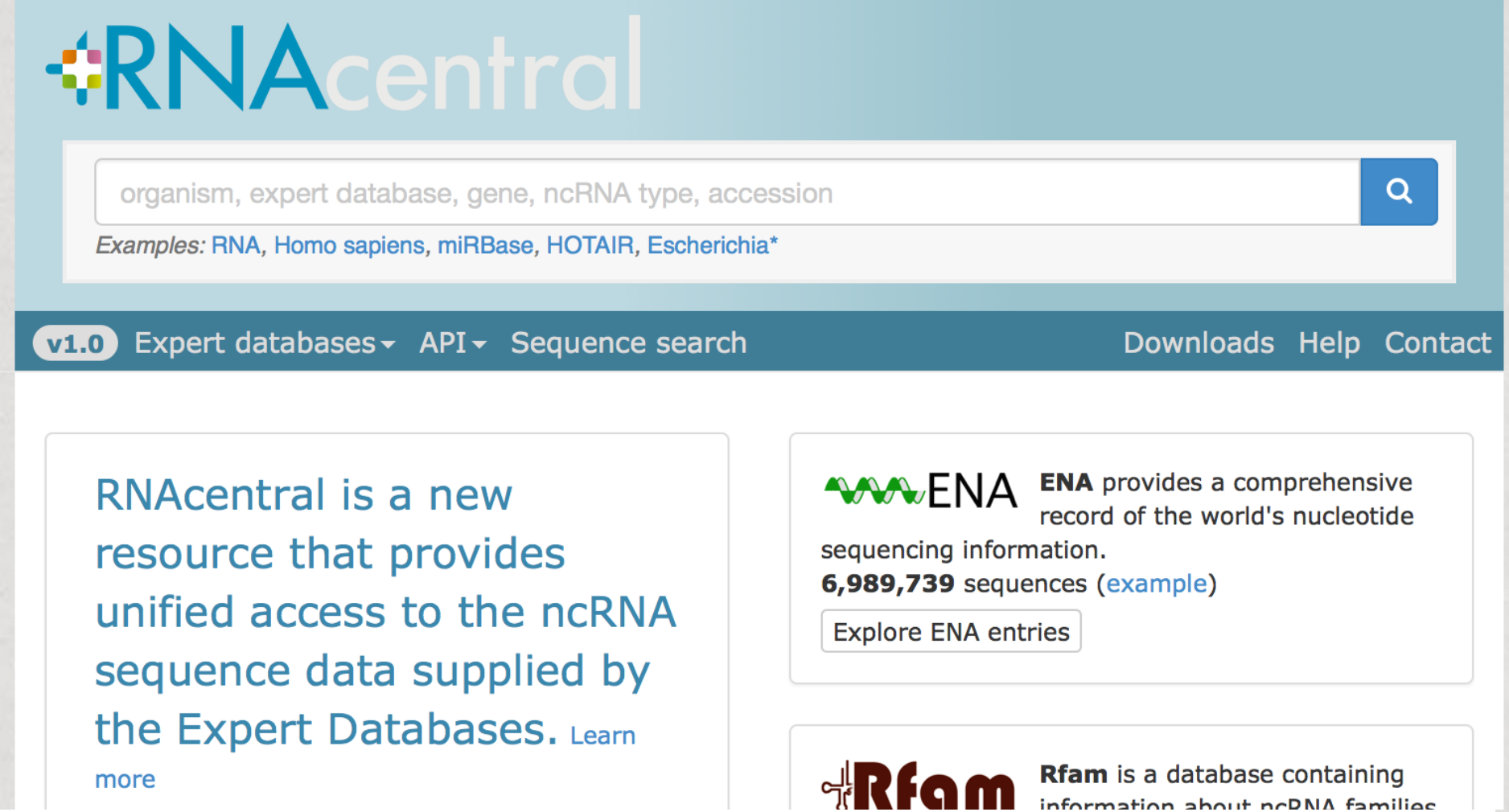
Position on chromosome: 12:22842684-22842712
Conserved species: panTro2,rheMac2,bosTau4,dasNov2
Binding area:
(Transcript)5' UUUACUUGCU GU 3'
GUGGU UGUG GUGCUGCU
UAUCG ACAU UACGACGA
(miRNA) 3' G GG GU 5'

Binding Type	Transcript position	Score	Conservation
9mer	4151-4179	0.0913826564422291	6

Position on chromosome: 12:22843071-22843099
Conserved species: panTro2,rheMac2,canFam2,dasNov2,loxAfr3,echTel1
Binding area:
(Transcript)5' AAUGUGAAC A 3'
CAUAGU GUAUAAUGCUGCU
GUAUCG CAUGUUACGACGA
(miRNA) 3' A GGA 5'

5	hsaLOCG410010584 (XLOC013305)	hsa-miR-103a-3p	0.995	
6	(CTA-204B4.6)	hsa-miR-103a-3p	0.991	
7	hsaLOCG410004968 (RP11-753N8.1.1)	hsa-miR-103a-3p	0.990	
8	(RP11-849H4.4)	hsa-miR-103a-3p	0.989	
9	hsaLOCG110001926 (n335525)	hsa-miR-103a-3p	0.988	
10	hsaLOCG110004356	hsa-miR-103a-3p	0.988	

Paraskevopoulou, MD, Vlachos, IS, Karagkouni D, Georgakilas, G, Kanellos I, N., Vergoulis, Tsanakas P, Floros E, Dalamagas T, Hatzigeorgiou AG. DIANA-LncBase v2: Indexing microRNA targets on non-coding transcripts ***Nucleic Acids Res***. 2016 Jan 4;44(D1):D231-8. doi: 10.1093/nar/gkv1270.



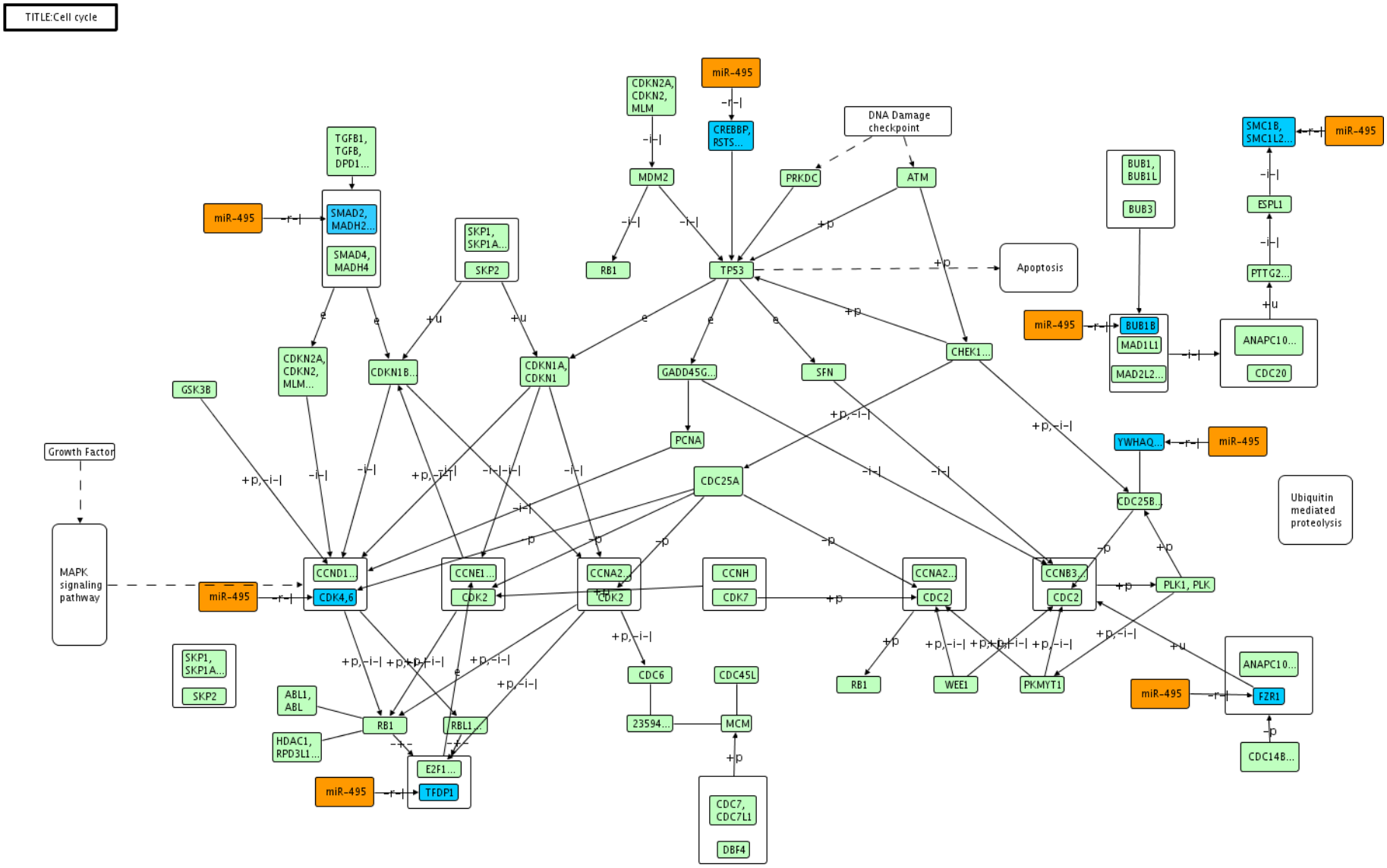
RNAcentral Expert Databases

Currently the RNAcentral Consortium is formed by **32** Expert Databases, **10** of which have already been integrated into RNAcentral (marked with a ✓ below). If you run an ncRNA database and would like to join RNAcentral, please contact us.

CRW Site	comparative sequence and structure information for ribosomal, intron, and

Connecting microRNAs to pathways

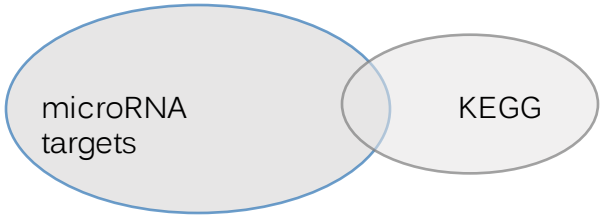
KEGG cell cycle pathways with miRNA targets of miR-495



www.microrna.gr/mirpath/

DIANA miRPath

Integrating human and mouse microRNAs in pathways



Small overlap – Not significant

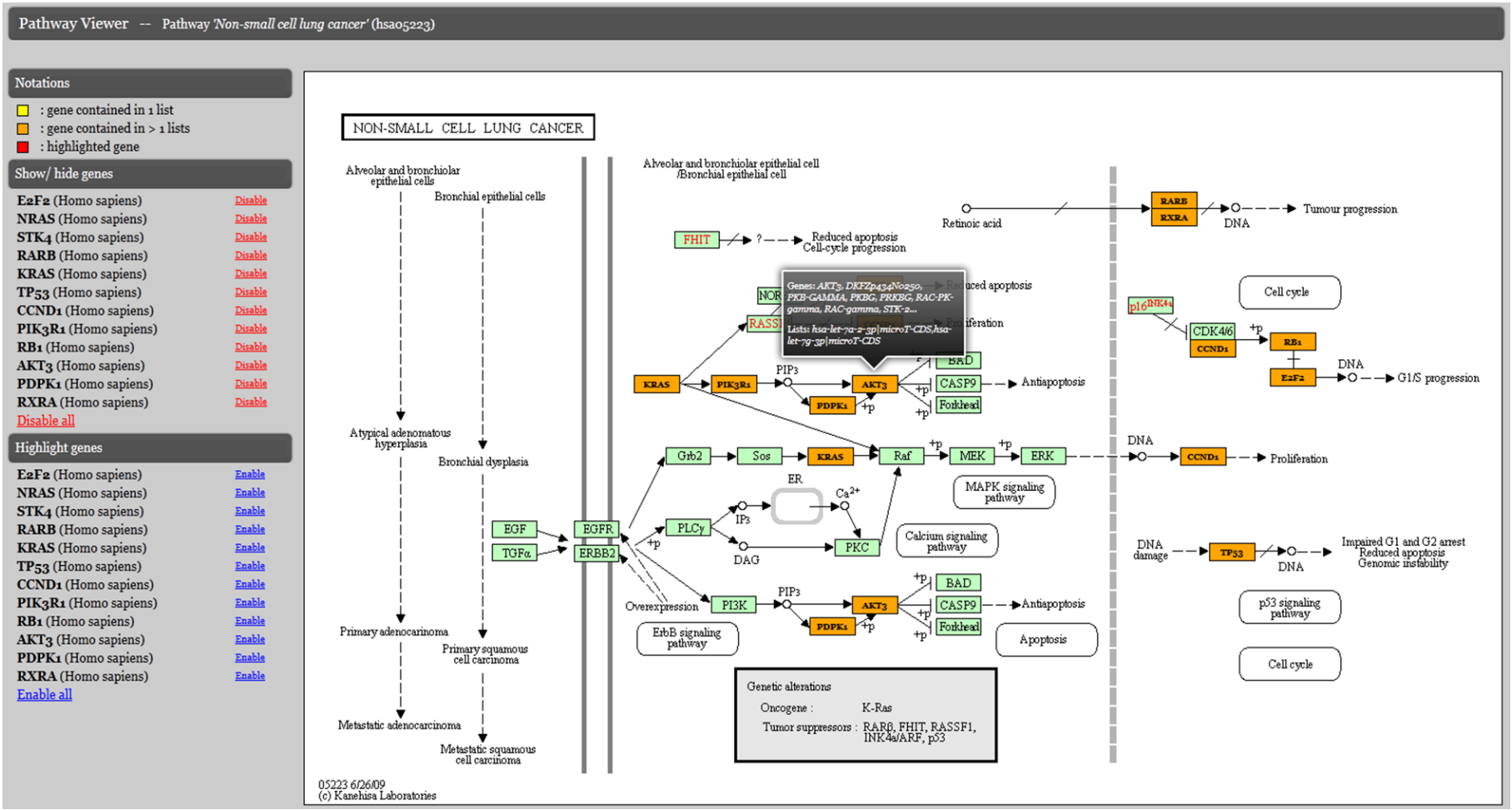
Large overlap – Significant

Input List Name	Number of Genes	Number of Genes in Pathways
Union	1250	306
let-7c_microT_4	723	166
miR-100_microT_4	35	11
miR-1_microT_4	562	147
Intersection	N/A	N/A

DOWNLOAD RESULTS

KEGG Pathway	Pathway ID	# of Genes (Union)	-ln(p-value) (Union)	# of Genes (let-7c_microT_4)	-ln(p-value) (let-7c_microT_4)	# of Genes (miR-100_microT_4)	-ln(p-value) (miR-100_microT_4)	# of Genes (miR-1_microT_4)	-ln(p-value) (miR-1_microT_4)	# of Genes (Intersection)	-ln(p-value) (Intersection)
Adherens junction	hsa04520	19	19.24	6	2.06	1	0.71	13	21.79	0	-
Glioma	hsa05214	14	10.23	7	4.28	2	6.62	8	7.62	0	-
Type II diabetes mellitus	hsa04930	10	9.38	6	6.32	1	1.47	3	1	0	-
mTOR signaling pathway	hsa04150	11	8.83	5	2.78	1	1.2	7	8.48	0	-
Colorectal cancer	hsa05210	16	8.63	7	2.32	3	13.19	8	4.6	0	-
MAPK signaling pathway	hsa04010	34	8.61	22	8.89	2	1.23	13	1.59	0	-
Bladder cancer	hsa05219	10	8.27	6	5.63	1	1.36	5	4.19	0	-
Focal adhesion	hsa04510	27	7.59	16	5.71	1	0.01	16	7.54	0	-
Wnt signaling pathway	hsa04310	22	7.44	9	1.34	3	7.48	13	7.01	0	-
Prostate cancer	hsa05215	15	6.53	7	2.13	2	4.68	9	6.05	0	-
Melanoma	hsa05218	13	6.48	8	5.04	1	0.71	7	4.29	0	-
Calcium signaling pathway	hsa04020	23	6.34	15	6.56	2	2.28	7	0.24	0	-
Huntington's disease	hsa05040	7	5.88	2	0.24	0	-	5	7.27	0	-
Chronic myeloid leukemia	hsa05220	13	5.75	8	4.54	0	-	7	3.86	0	-
Pancreatic cancer	hsa05212	12	4.87	7	3.29	0	-	6	2.62	0	-
Amyotrophic lateral sclerosis (ALS)	hsa05030	5	4.75	4	6.3	1	3	1	0.21	0	-
p53 signaling pathway	hsa04115	11	4.32	9	7.75	0	-	6	3.04	0	-

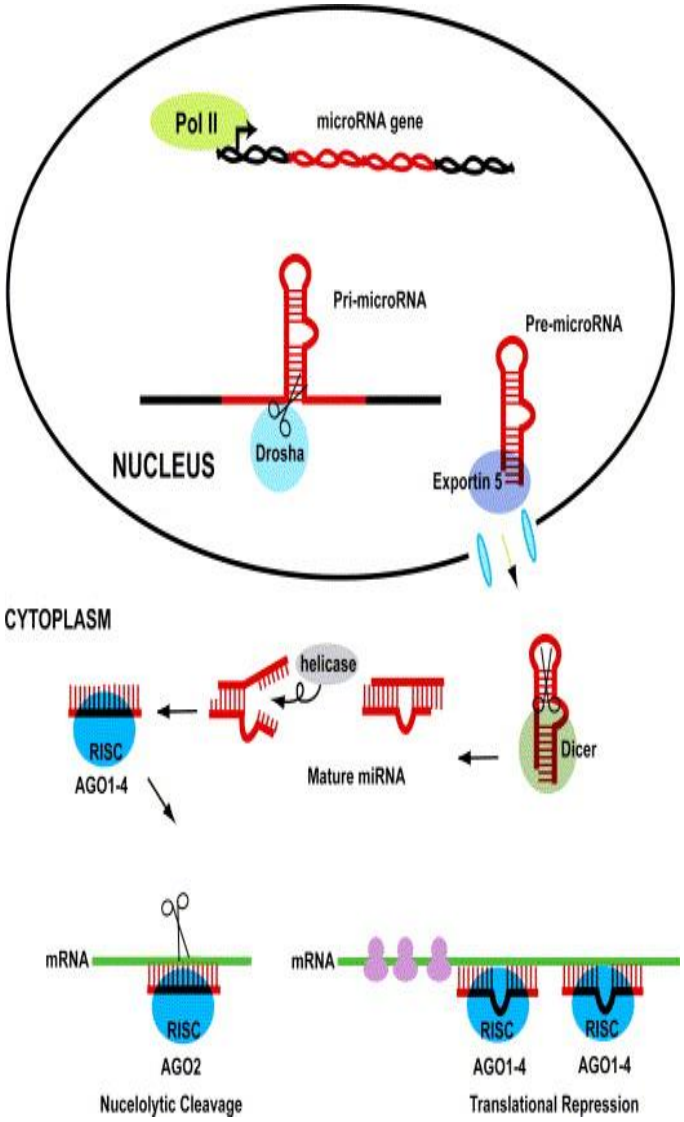
DIANA-miRPath v3.0
www.microrna.gr/miRPathv3

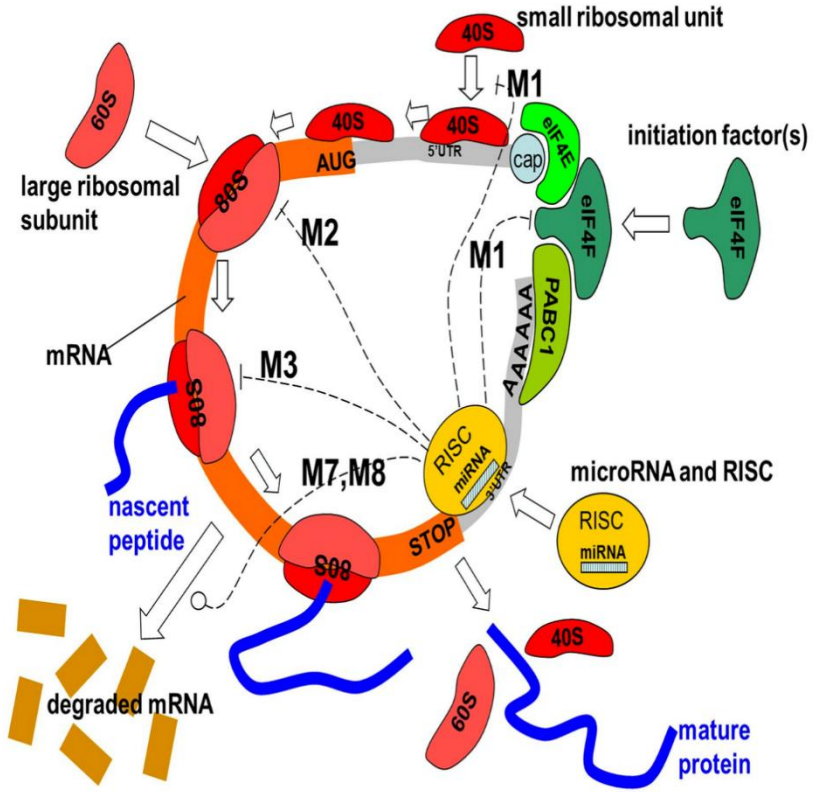


Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. ***Nucleic Acids Res***. 2015 Jul 1;43(W1):W460-6.

REGULATION

microRNA Biogenesis & Function



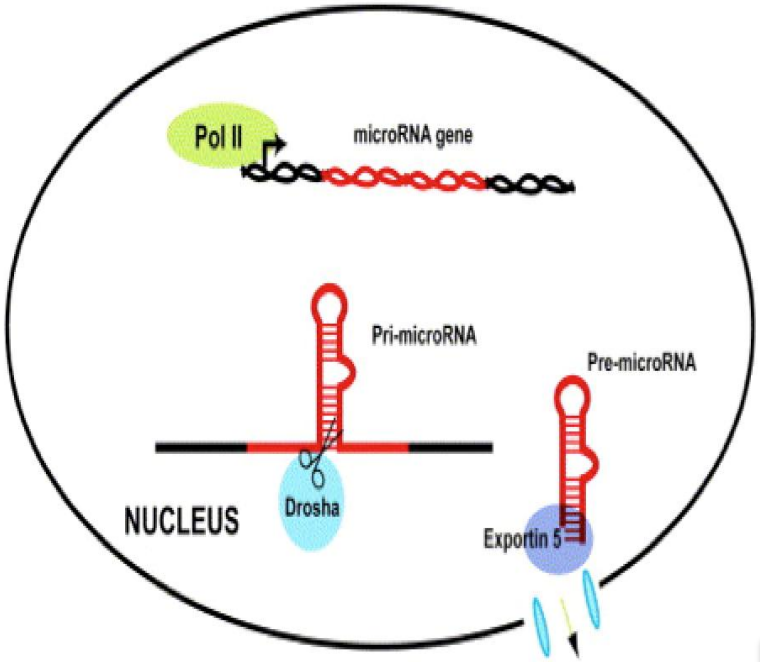


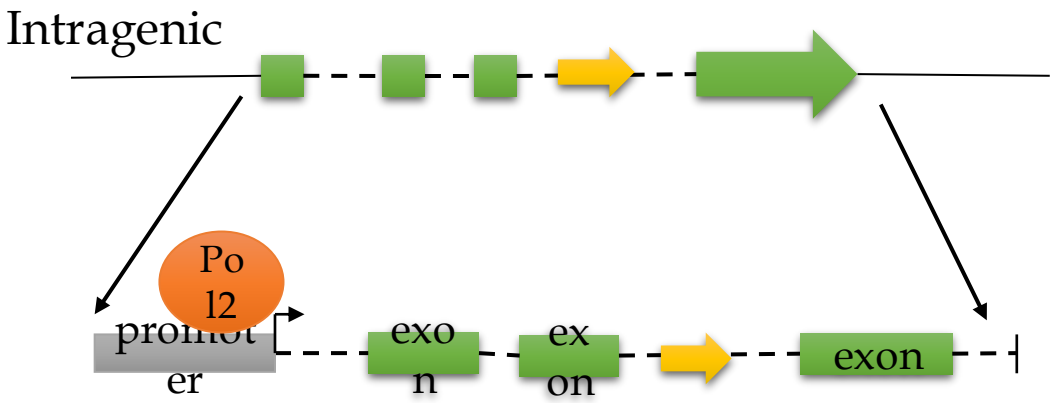
Hammond et al. **Dicing and slicing: The core machinery of the RNA interference pathway. Febs Letters,** Volume 579, Issue 26, 2005, Pages 5822–5829

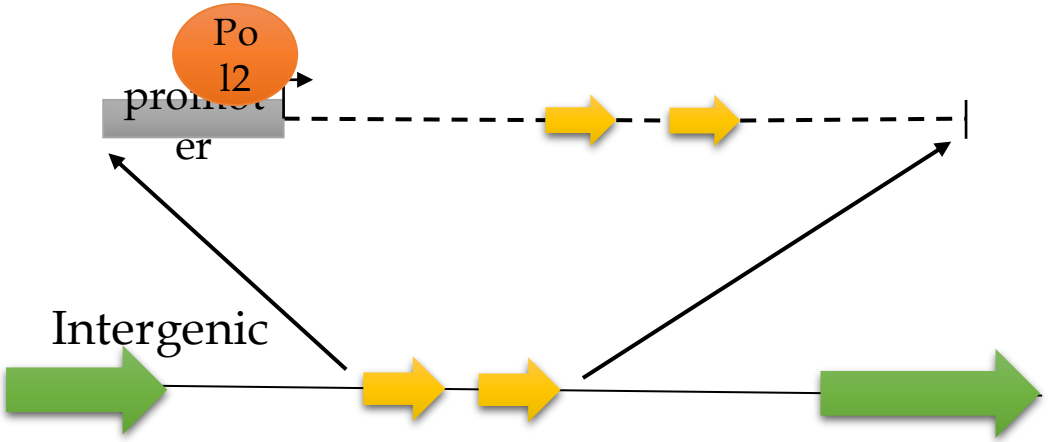
miRNA classification

Coding

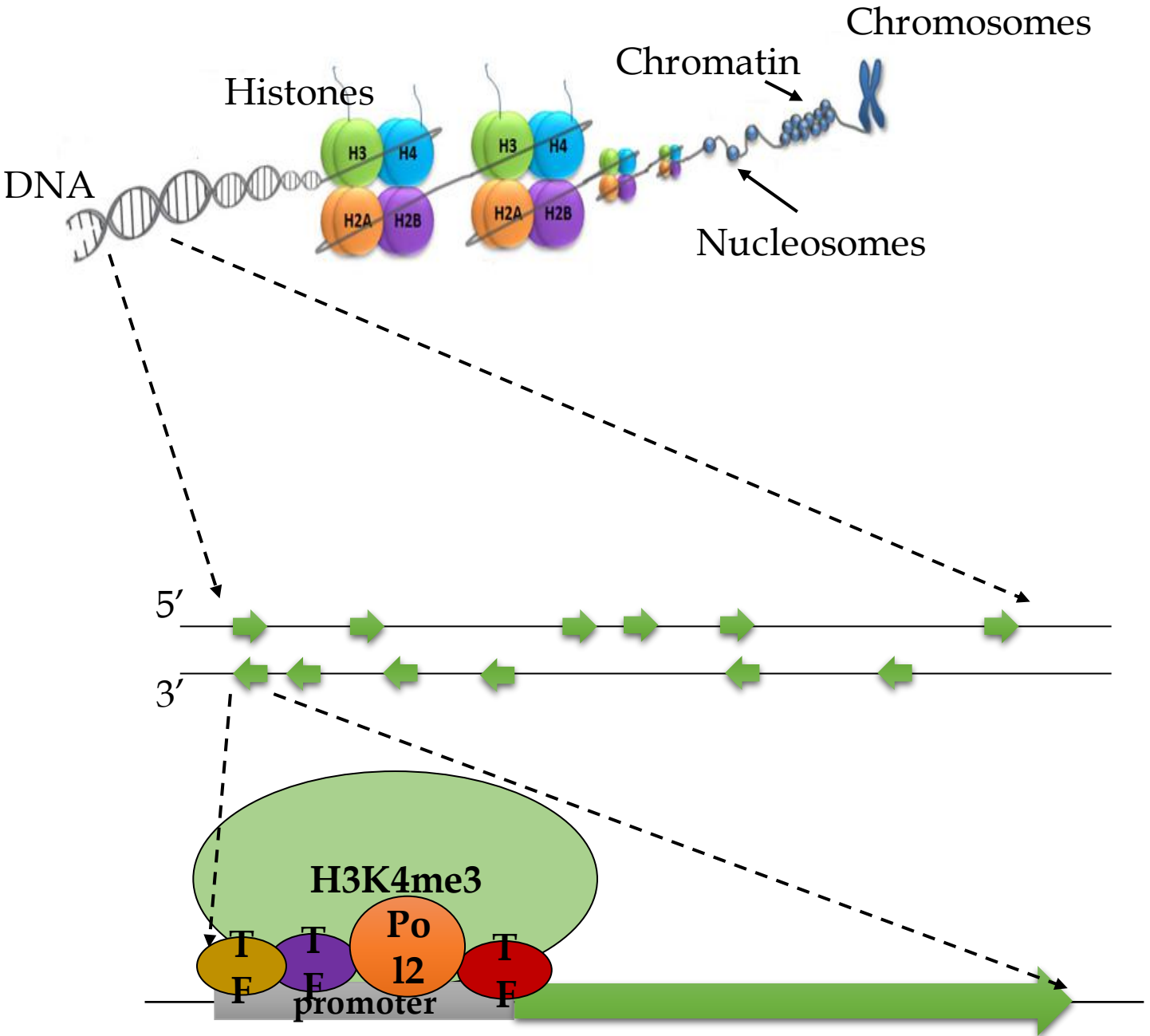
miRNA



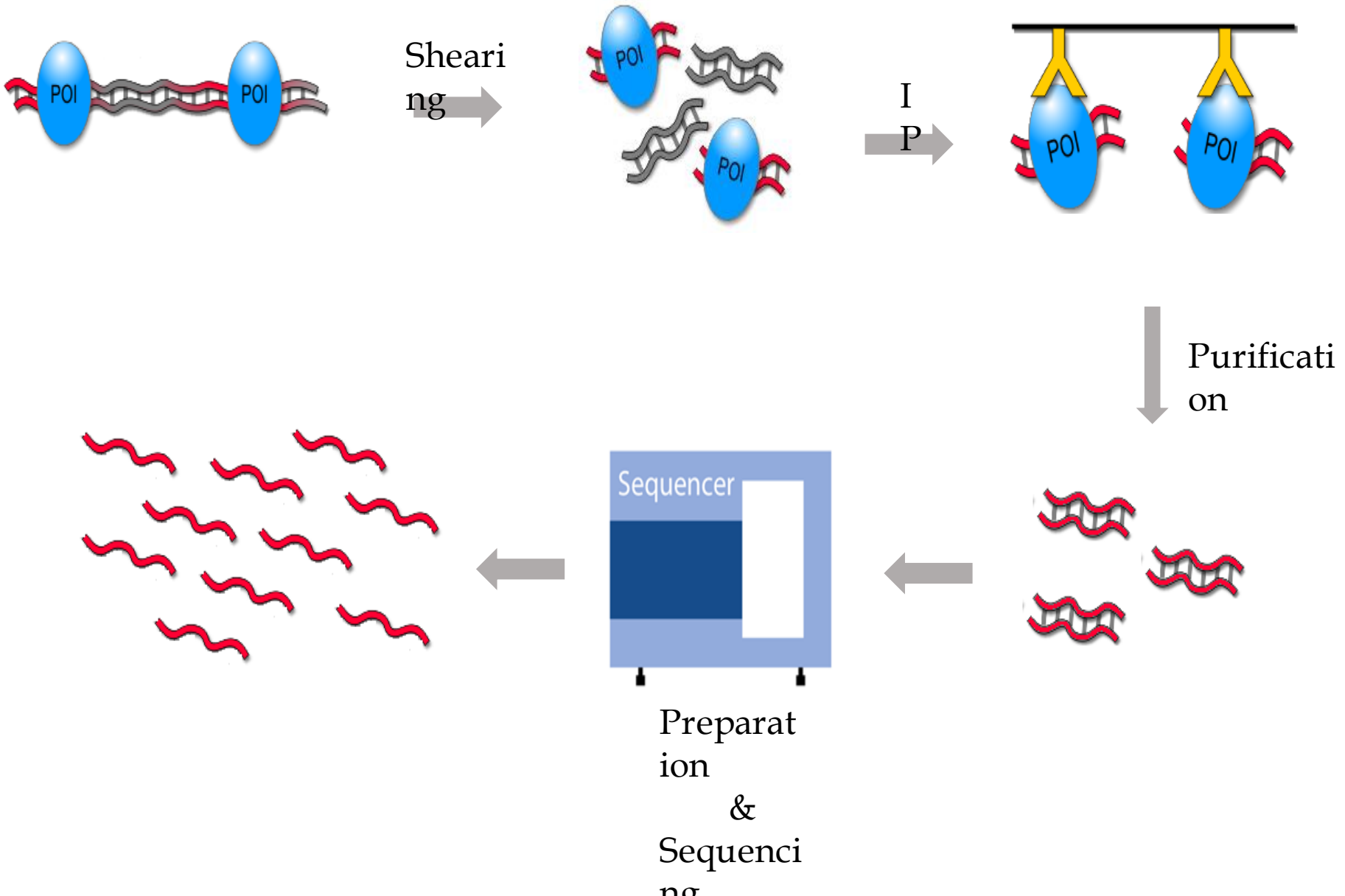




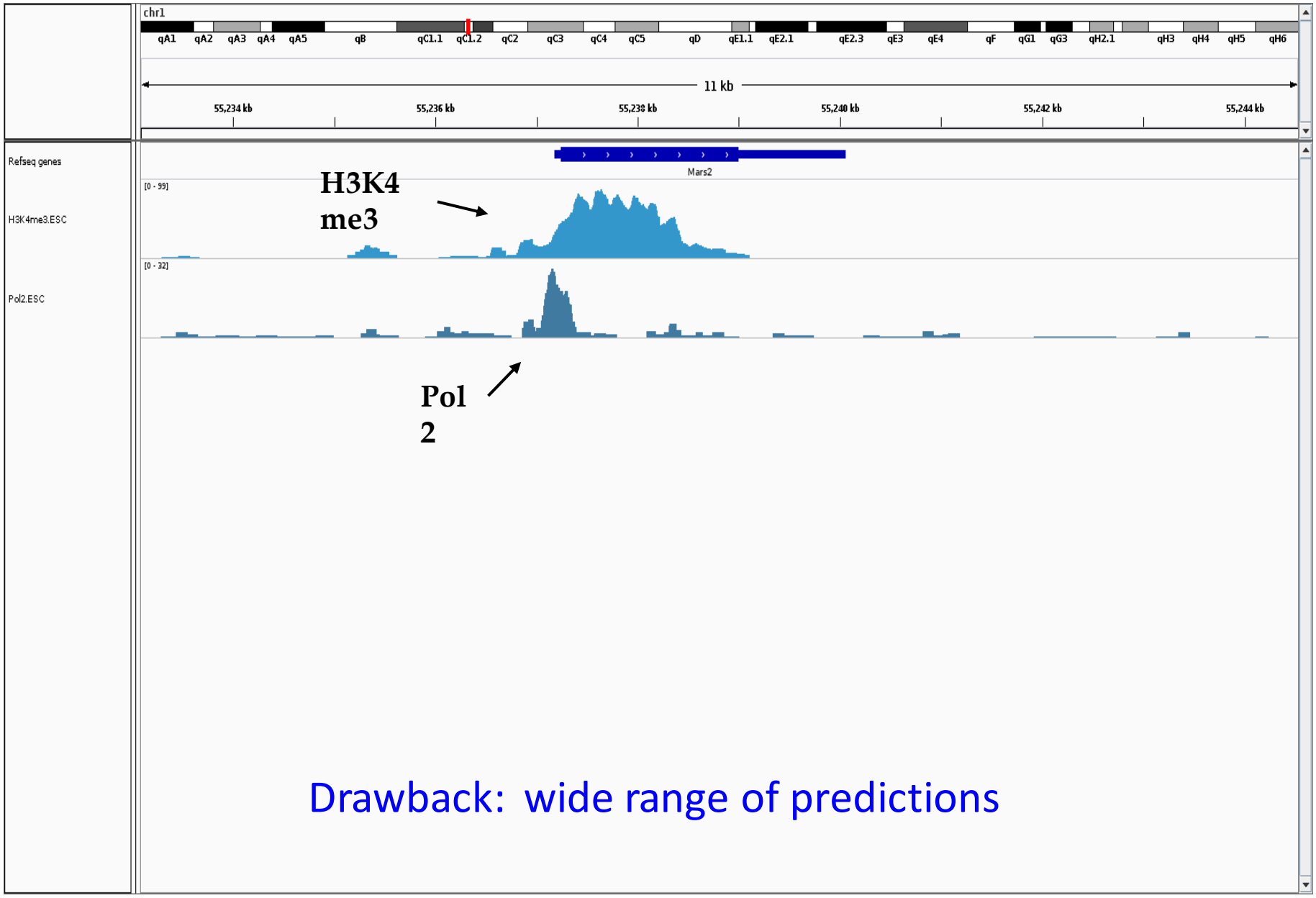
Chromatin structure and transcription



Chromatin ImmunoPrecipitation Sequencing (ChIP-Seq)

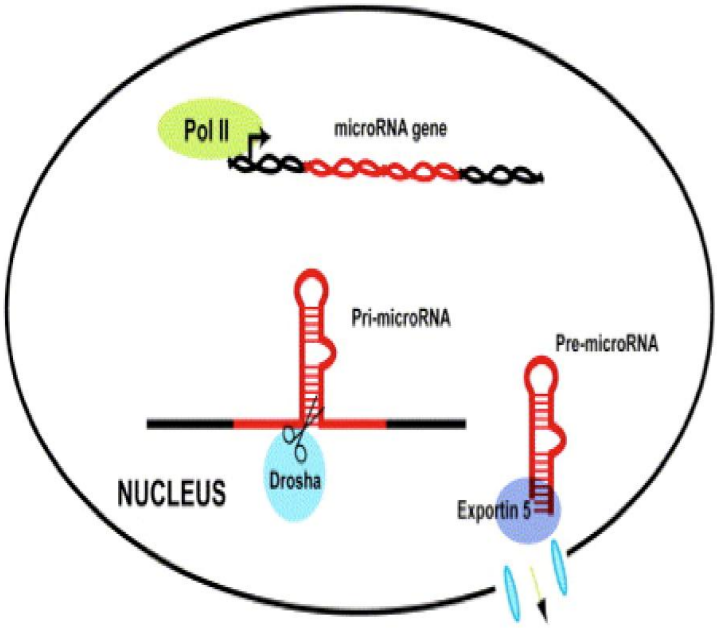


ChIP Sequencing Visualization



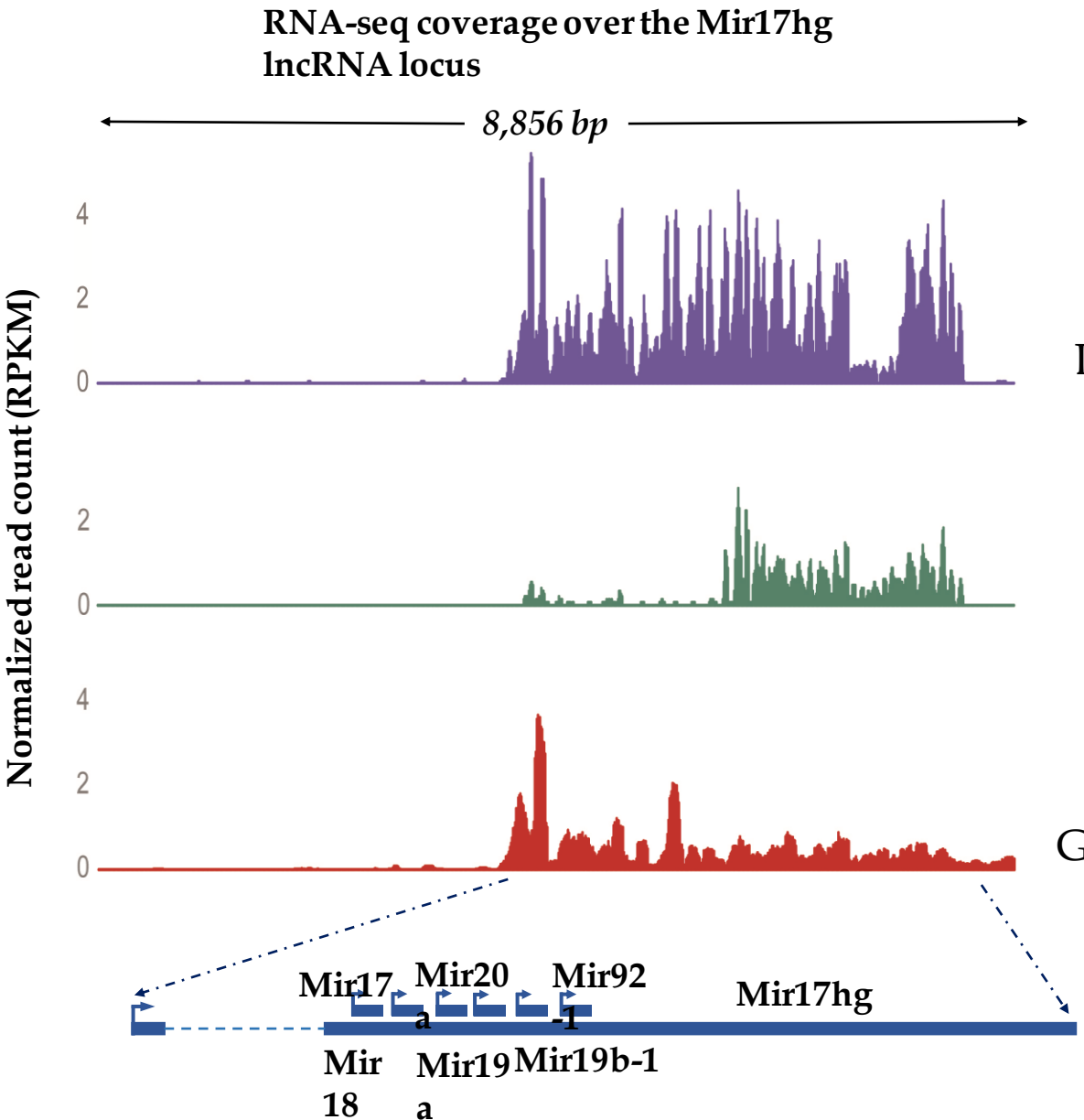
Drawback: wide range of predictions

How do we validate a miRNA TSS prediction ?



Drosha null/conditional-null (*Drosha*$^{LacZ/e4COIN}$) mouse model that has been generated using the conditional by inversion (COIN) methodology from Aris Economides @ REGENERON Pharmaceuticals

Comparison other RNA-Seq experiments



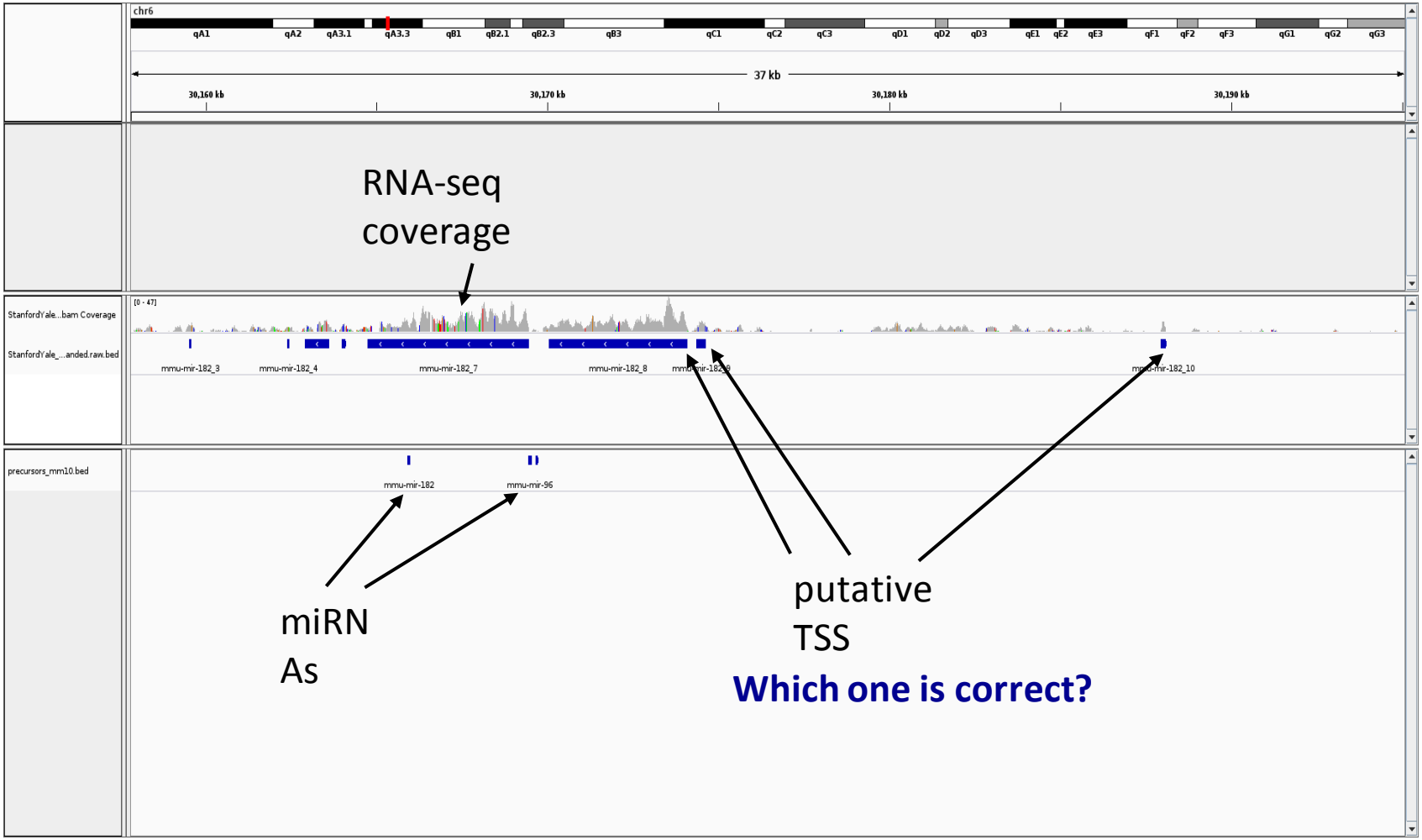
Drosha -/- mESCs with **27M reads**

Drosha +/+ mESCs with **19M reads**

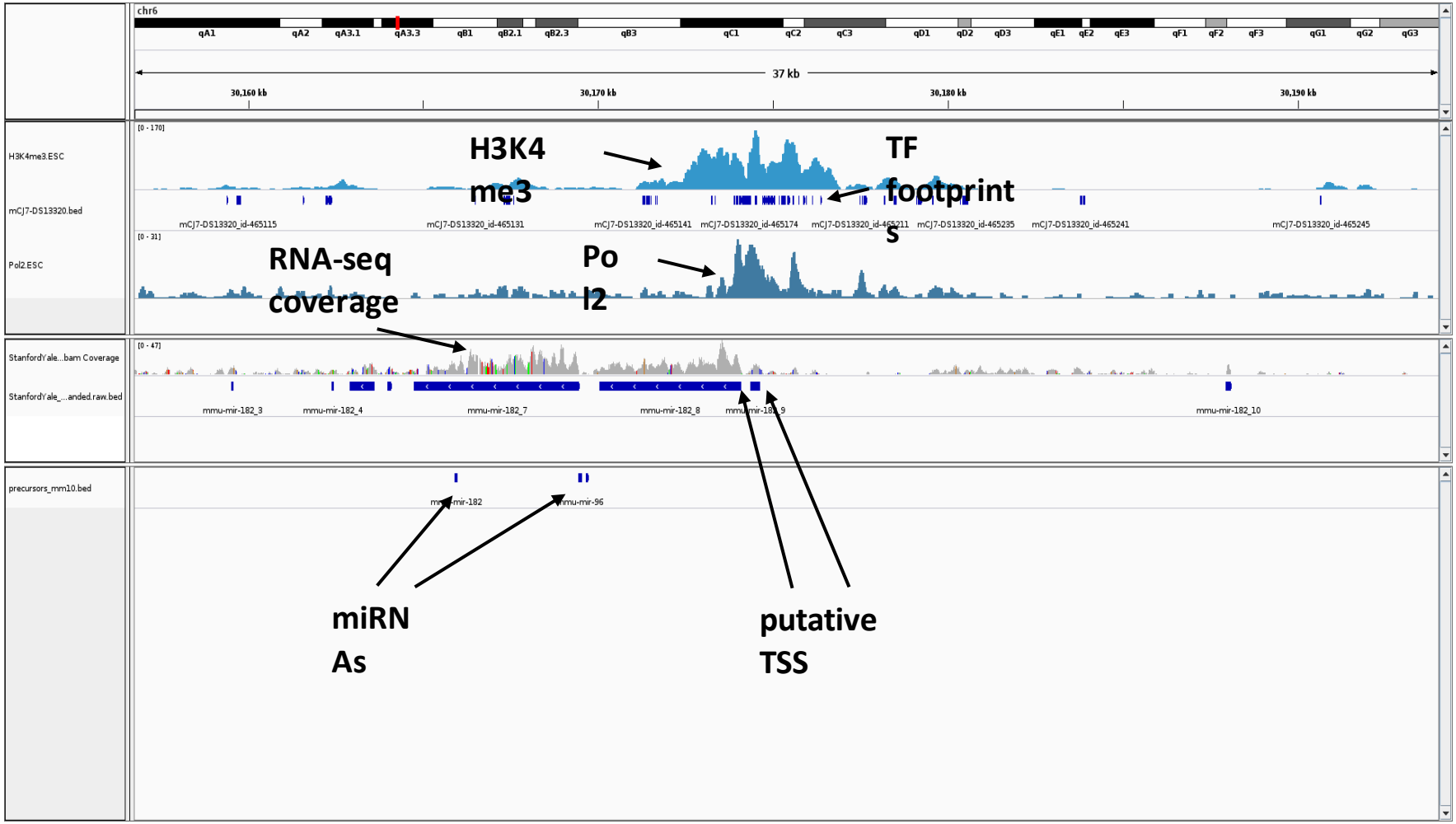
GSM973235 WT mESCs **180M reads**

RNA-seq read depth is essential!

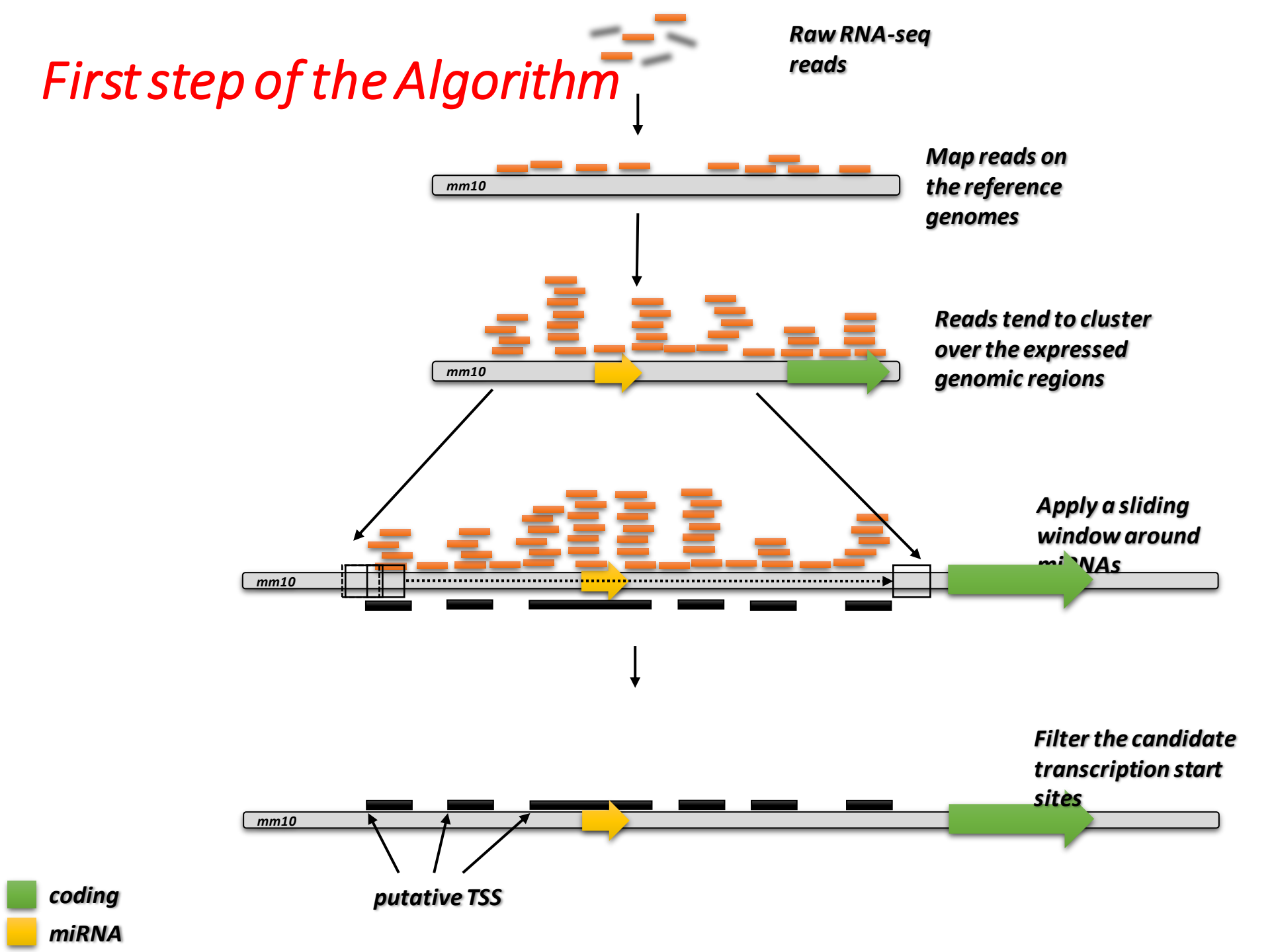
Deep RNA seq is not enough



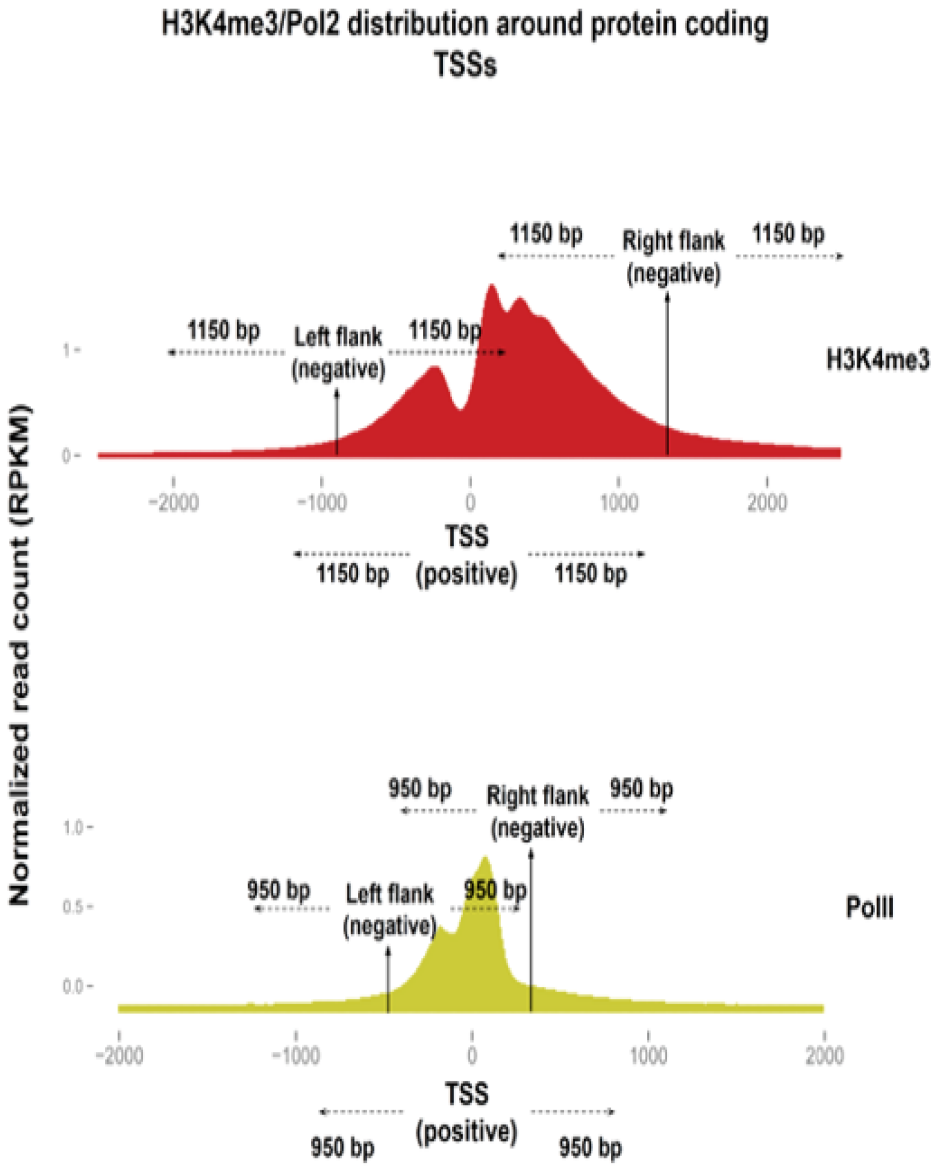
ChIP-seq information can effectively reduce putative TSS's



First step of the Algorithm



Algorithm - second step



An algorithm than can learn from data : machine learning

Here we used **Support Vector Machines**
A supervised machine learning algorithm
for each chip seq set of data.

Learns from

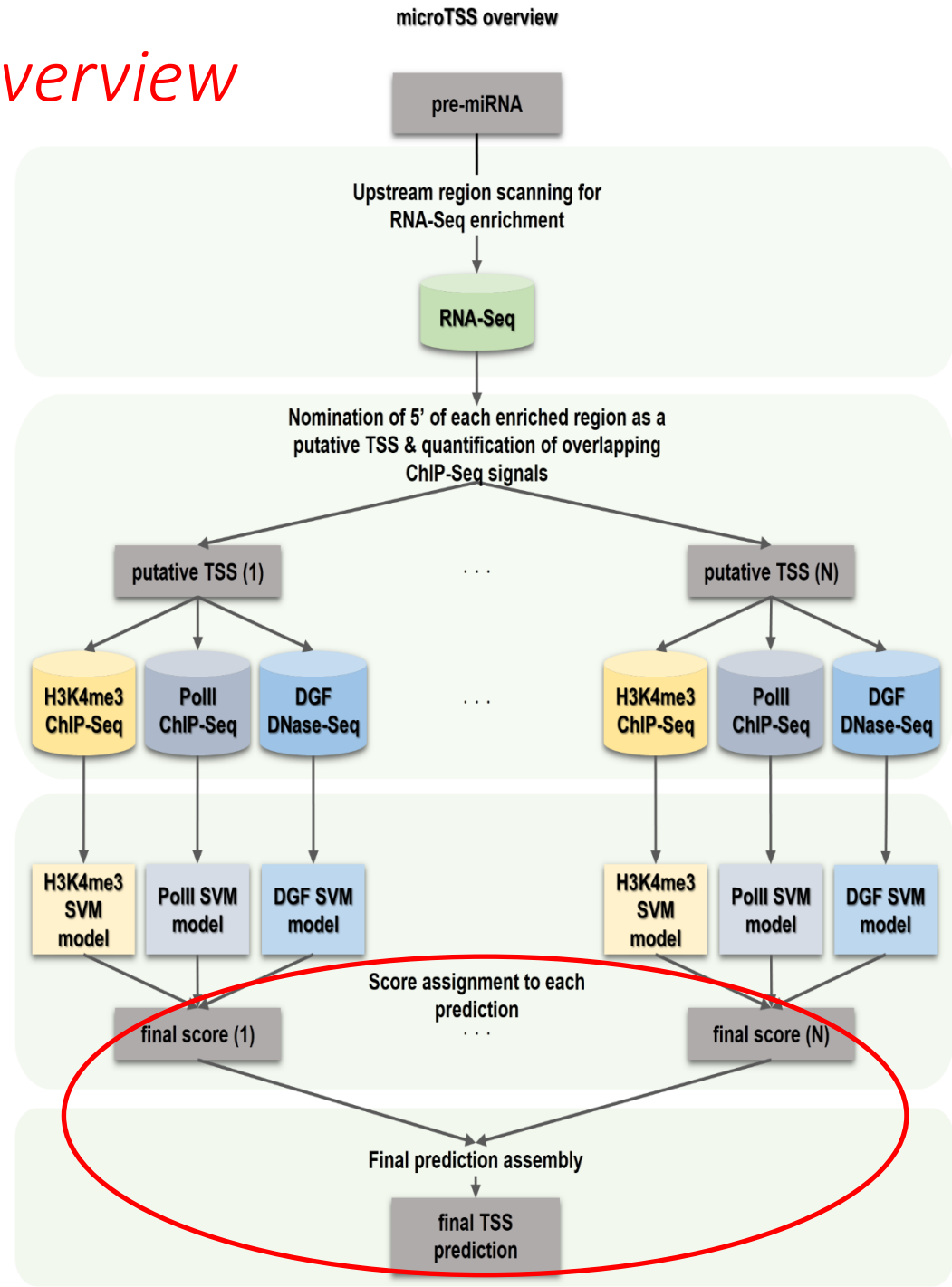
- positive examples (known TSS)
- negative examples (random intergenic locations, flanking positions)

Algorithm overview

First step

Second step

Final step



Comparison with existing methods

Performance on test set:
47 miRNA TSS derived from Drosha depleted mice.

Predictions < 1,000 bp from the validated TSS are considered True

Precision = TP / (TP+FP)

Sensitivity = Correct Predictions / Total Correct

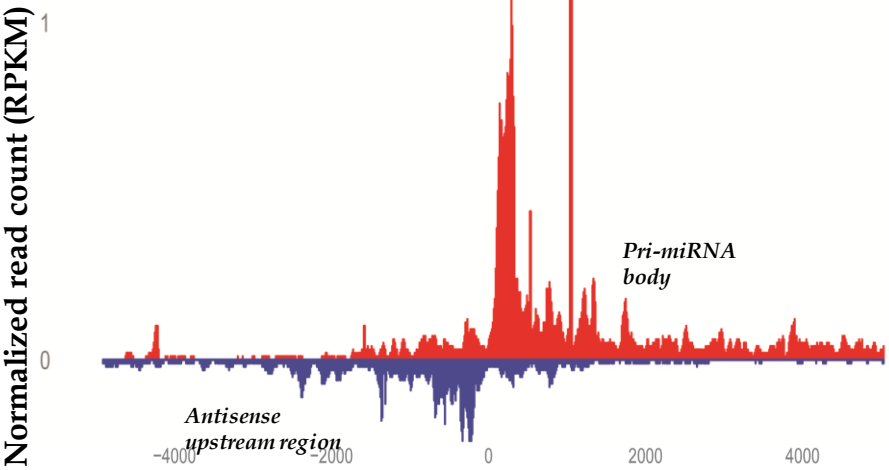
Algorithms Precision and Sensitivity at 1kbp distance threshold from validated TSSs in mESC		
	mESCs (N=47)	
	Sensitivity	Precision
Marson et al	54% (20/37)	64.5% (20/31)
PROmiRNA	78.7% (37/47)	25.4% (95/373)
S-Peaker	76.5% (36/47)	18.8% (77/409)
microTSS	93.6% (44/47)	100% (44/44)

RO-Seq analysis

GRO-Seq analysis revealed

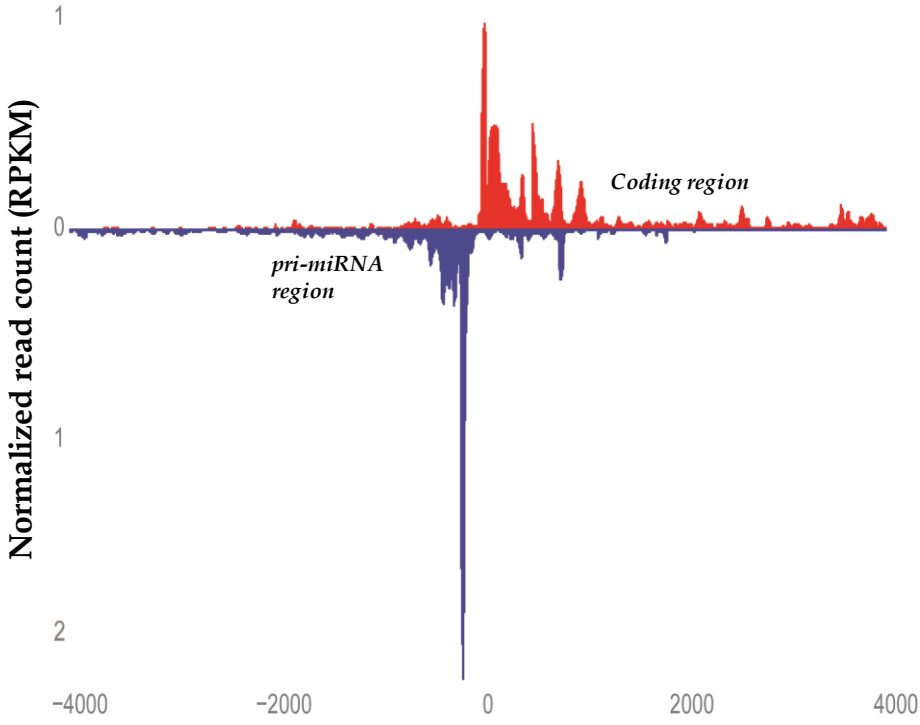
- 72 TSSs in hESCs
- 81 TSSs in IMR90 cells

GRO-Seq distribution on validated hESC miRNA TSSs



Distance from miRNA gene TSS

GRO-Seq distribution over protein coding TSSs with divergent pri-miRNAs



Distance from protein coding TSS

Refining non-coding gene annotation

Nespas non-coding locus

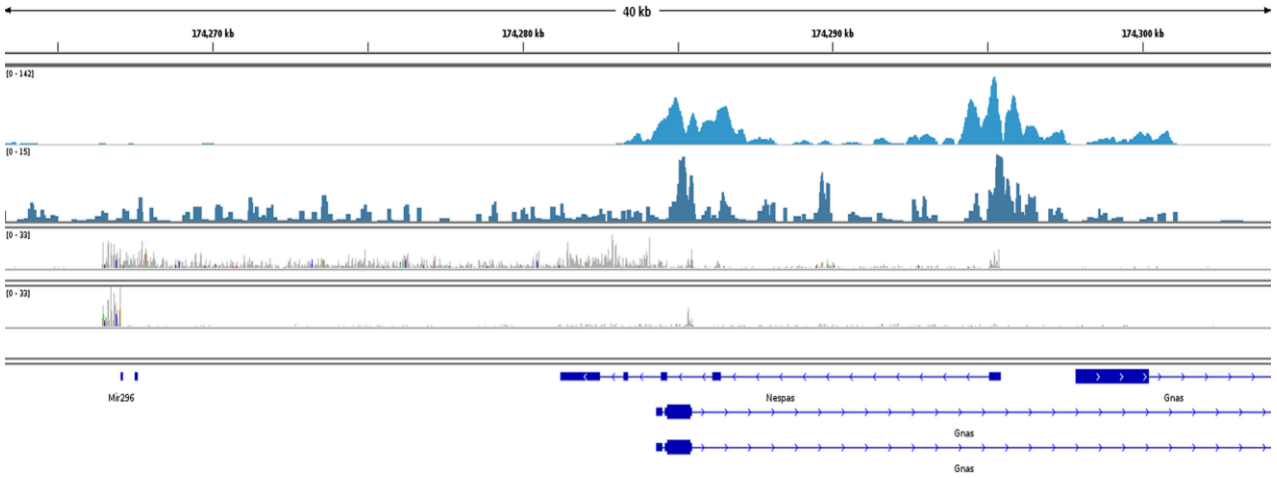
H3K4me3

Pol2

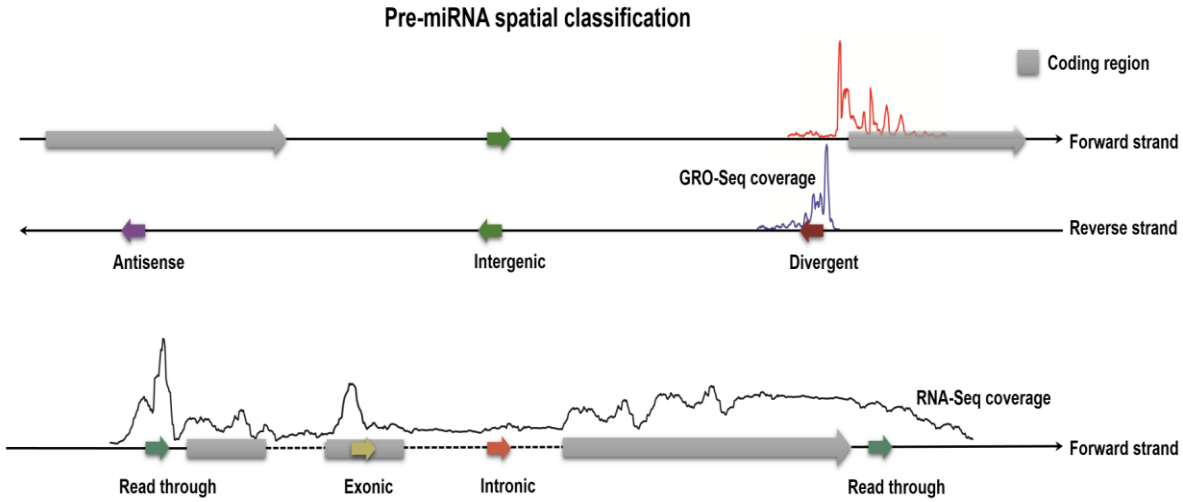
Drosha -/-

Drosha +/+

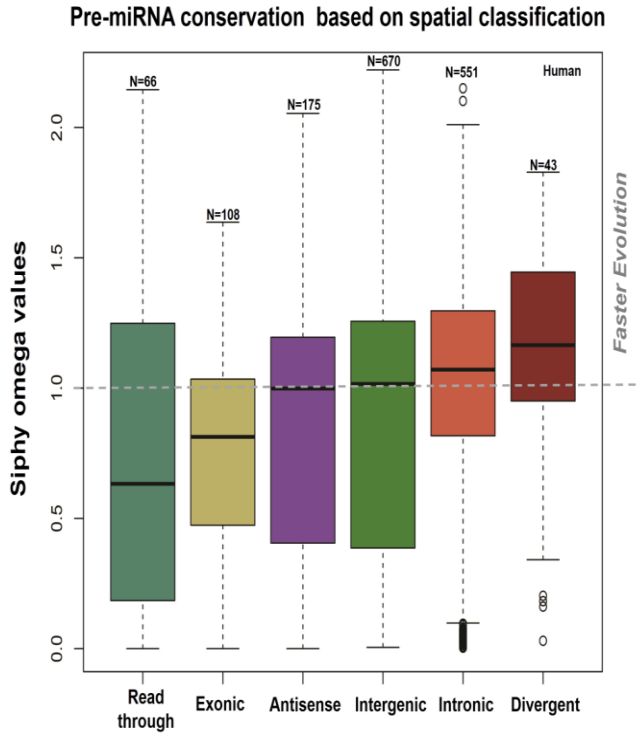
RefSeq annotation



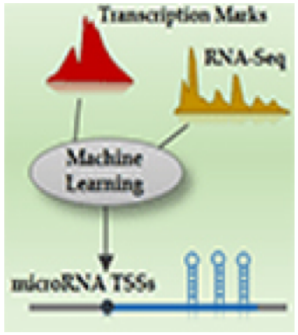
a


Pre-miRNA spatial classification
Coding region
Forward strand
GRO-Seq coverage
Reverse strand
Antisense
Intergenic
Divergent
RNA-Seq coverage
Forward strand
Read through
Exonic
Intronic
Read through

b


Pre-miRNA conservation based on spatial classification
Human
N=66
N=108
N=175
N=670
N=551
N=43
Siphy omega values
0.0
0.5
1.0
1.5
2.0
Faster Evolution
Read through
Exonic
Antisense
Intergenic
Intronic
Divergent

10 December 2014



ARTICLE

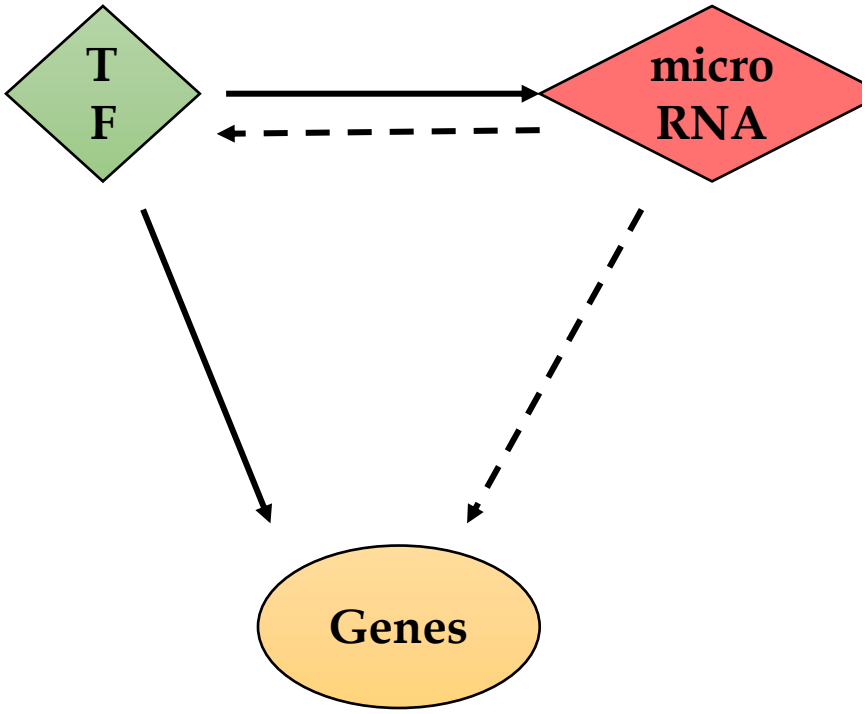
microTSS: accurate microRNA transcription start site identification reveals a significant number of divergent pri-miRNAs

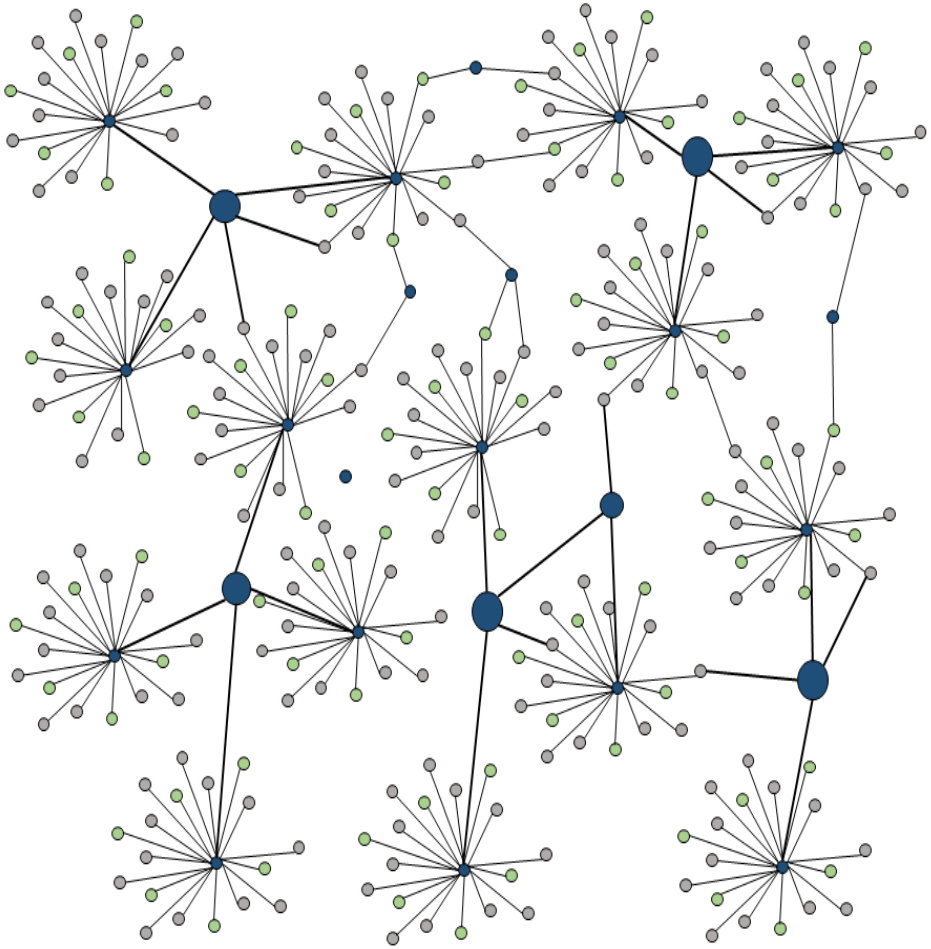
Georgios Georgakilas, Ioannis S. Vlachos, Maria D. Paraskevopoulou, Peter Yang, Yuhong Zhang, Aris N. Economides, Artemis G. Hatzigeorgiou

My account | Submit manuscript | Register | Subscribe
Login | Cart
nature COMMUNICATIONS

Refining Gene Regulatory Networks

Evolution of Gene Regulatory Networks





Identifying biomarkers

microRNAs and Epithelial Ovarian Cancer

Identify cause / markers for ovarian cancer progression and malignancy

mRNA expression by microarray

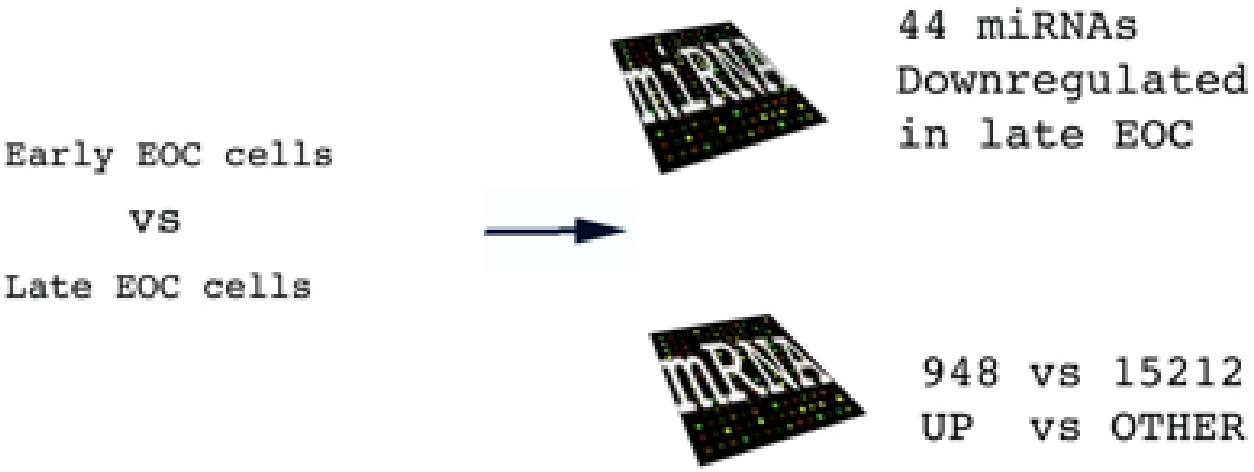
microRNA expression by microarray

PNAS

Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer

Lin Zhang[a,b,c], Stefano Volinia[d], Tomas Bonome[e], George Adrian Calin[d], Joel Greshock[f,g], Nuo Yang[a], Chang-Gong Liu[d], Antonis Giannakakis[a,h], Pangiotis Alexiou[i], Kosei Hasegawa[a], Cameron N. Johnstone[j], Molly S. Megraw[k], Sarah Adams[a,b], Heini Lassus[l], Jia Huang[f], Sippy Kaur[l], Shun Liang[a], Praveen Sethupathy[k], Arto Leminen[l], Victor A. Simossis[i], Raphael Sandaltzopoulos[h], Yoshio Naomoto[m], Dionyssios Katsaros[n], Phyllis A. Gimotty[o], Angela DeMichele[j], Qihong Huang[p], Ralf Bützow[l], Anil K. Rustgi[j], Barbara L. Weber[f,g], Michael J. Birrer[e], Artemis G. Hatzigeorgiou[c,f,i,k], Carlo M. Croce[c,d], and George Coukos[a,b,c,f]

Numerous miRNAs and protein coding genes are downregulated in late stage ovarian cancer.



MiRNA downregulation affects mRNA transcripts?

(miRNA = DOWN & targets = UP)

Calculating the hexamer distribution in the UTR's of genes that gain expression(UP) and genes that do not change(OTHER)

```
AAAAAA : 0,0,1,5,2,0,1 ...
AAAAAT : 1,1,3,0,0,1,2 ...

  ...

CCCCCG : 1,3,2,0,0,0,2 ...
CCCCCC : 1,0,0,0,1,0,2 ...
```

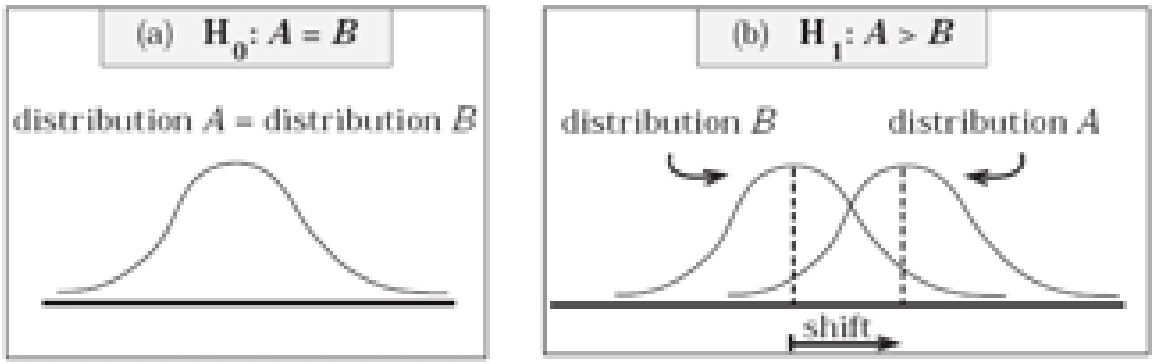
Normalize
divide by 3'UTR length

```
AAAAAA : 0.02343,0,0,0,0.00021,0,0.007462 ...
AAAAAT : 0,0.0001207,0.00072,0.1,0,0 ...

  ...

CCCCCG : 0.004,0.0667,0,0.1,0,0.0004 ...
CCCCCC : 0.12,0,0,0.00031,0.21109,0.005301 ...
```

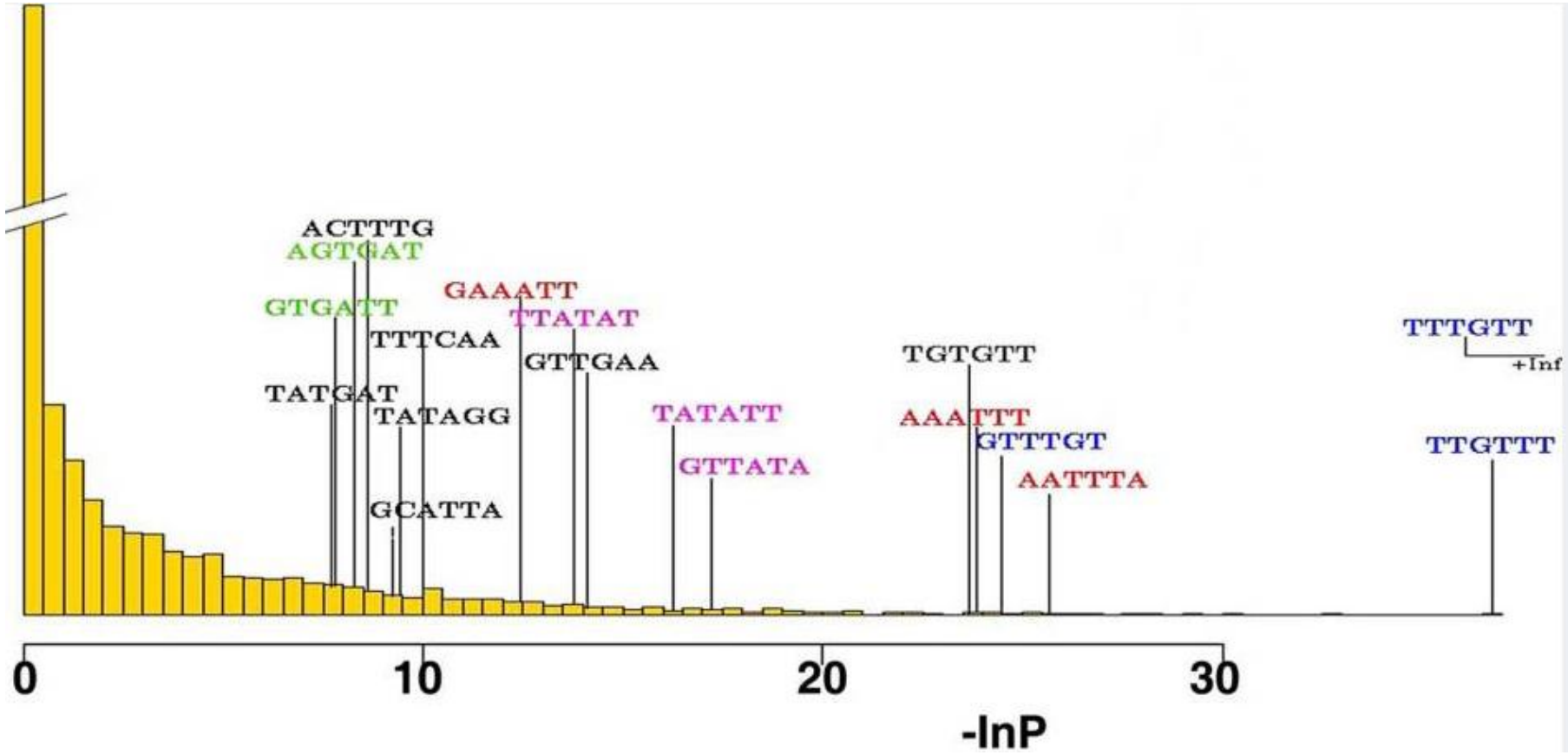
UP values	OTHER values
0.02401	0.34251
0.00054	0.00543
0.00022	0.00432
0.00322	0.00935
0.71533	0.00043
0.00640	0.04540
0.06422	0.07462
0.00001	0.00432
0.00242	0.00245
0.00535	0.05540
	0.60432
	0.04355
	0.05343
	0.04320
	0.01162
	0.00112
	0.02450
	0.01333

p - value

Calculating P-Values of hexamer distribution in two groups of genes (UP vs OTHER) through Wilcoxon Rank Sum test.

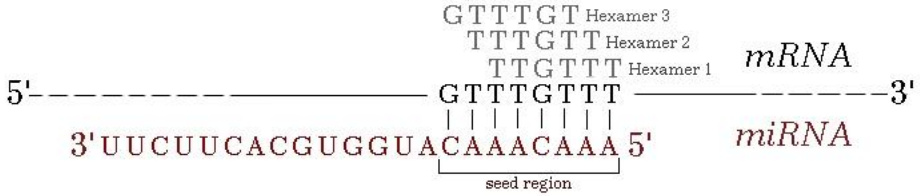


microRNA και Επιθηλιακός Καρκίνος των Ωοθηκών



Zhang et al, PNAS 2008

Linking hexamers to downregulated miRNAs.

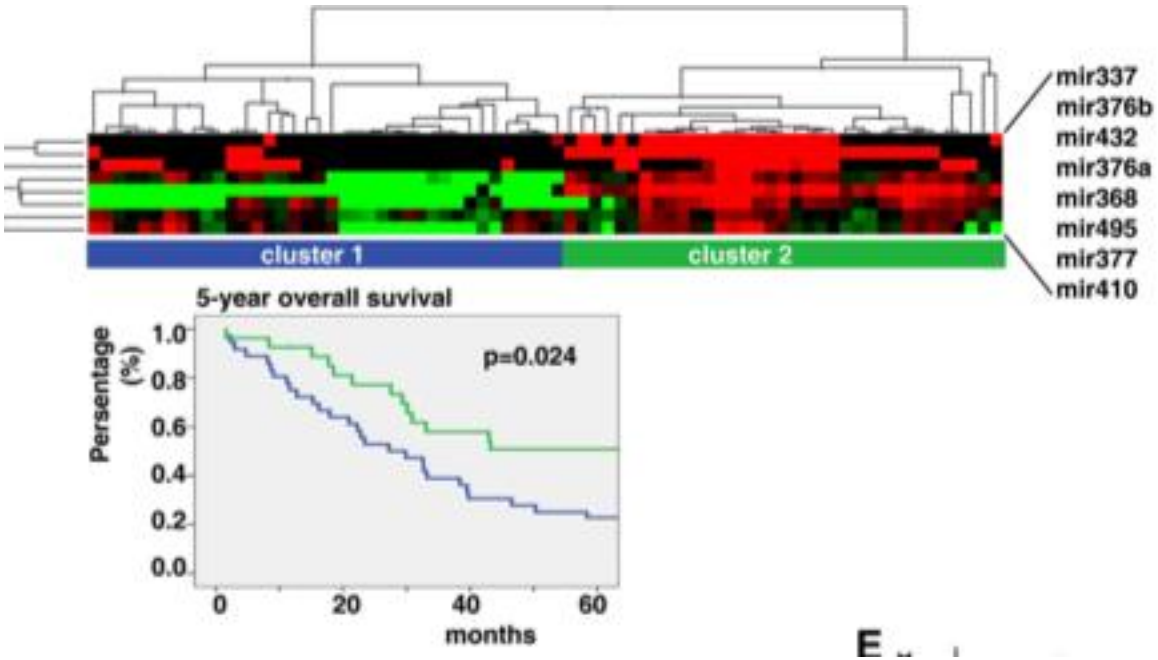


Sheet1

hexamer	-lnP	number of UTRs (UP)	hsa-miR	start position
TTTGTT	Inf	566	hsa-miR-495	2
TTGTTT	36.74	556	hsa-miR-495	1
AATTTA	26.66	432	hsa-miR-513-3p	1
GTTTGT	24.31	400	hsa-miR-495	3
AAATTT	23.86	500	hsa-miR-513-3p	2
TGTGTT	23.79	475	hsa-miR-362-3p	1
GTTATA	17.11	235	hsa-miR-410	3
TATATT	16.15	431	hsa-miR-410	1
GTTGAA	14.06	294	hsa-miR-95	1
TTATAT	13.94	395	hsa-miR-410	2
GAAATT	12.67	383	hsa-miR-513-3p	3
TTTCAA	9.72	417	hsa-miR-488	1
TATAGG	8.78	179	hsa-miR-337-3p	3
GCATTA	8.51	215	hsa-miR-365	1
ACTTTG	8.15	400	hsa-miR-519d	1
AGTGAT	7.89	291	hsa-miR-34b	3
GTGATT	7.53	298	hsa-miR-34b	2
TATGAT	7.52	248	hsa-miR-376b	1
TATGAT	7.52	248	hsa-miR-376a	1

The majority of the downregulated miRNAs are located at a big miRNA cluster (< 36) at the Dlk1-Gtl2 domain of chr. 14.

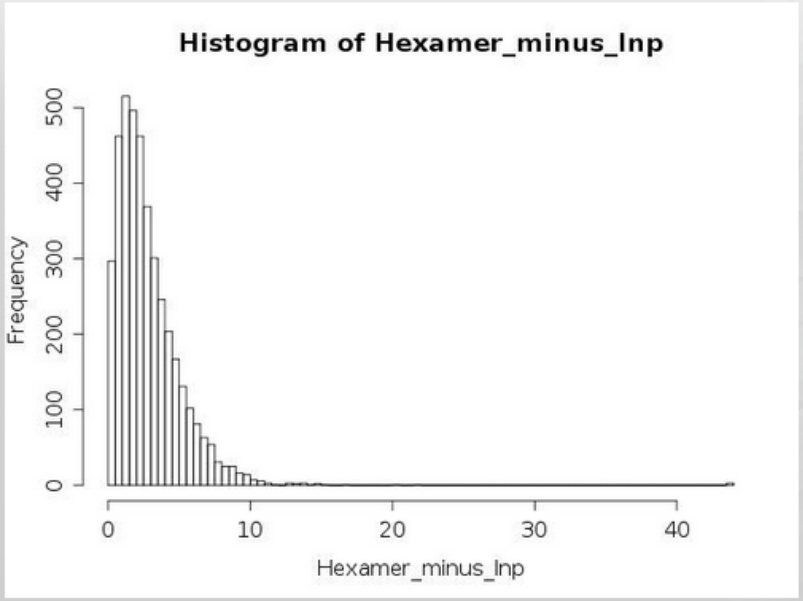
Downregulation of miRNA cluster at *DLK1-GTL2* domain is associated with poor survival.



DIANA-mirExTra

identifying miRNA involvement in disease through high throughput experimental data.

Hexamer Results Page (mmu_miR_122a)

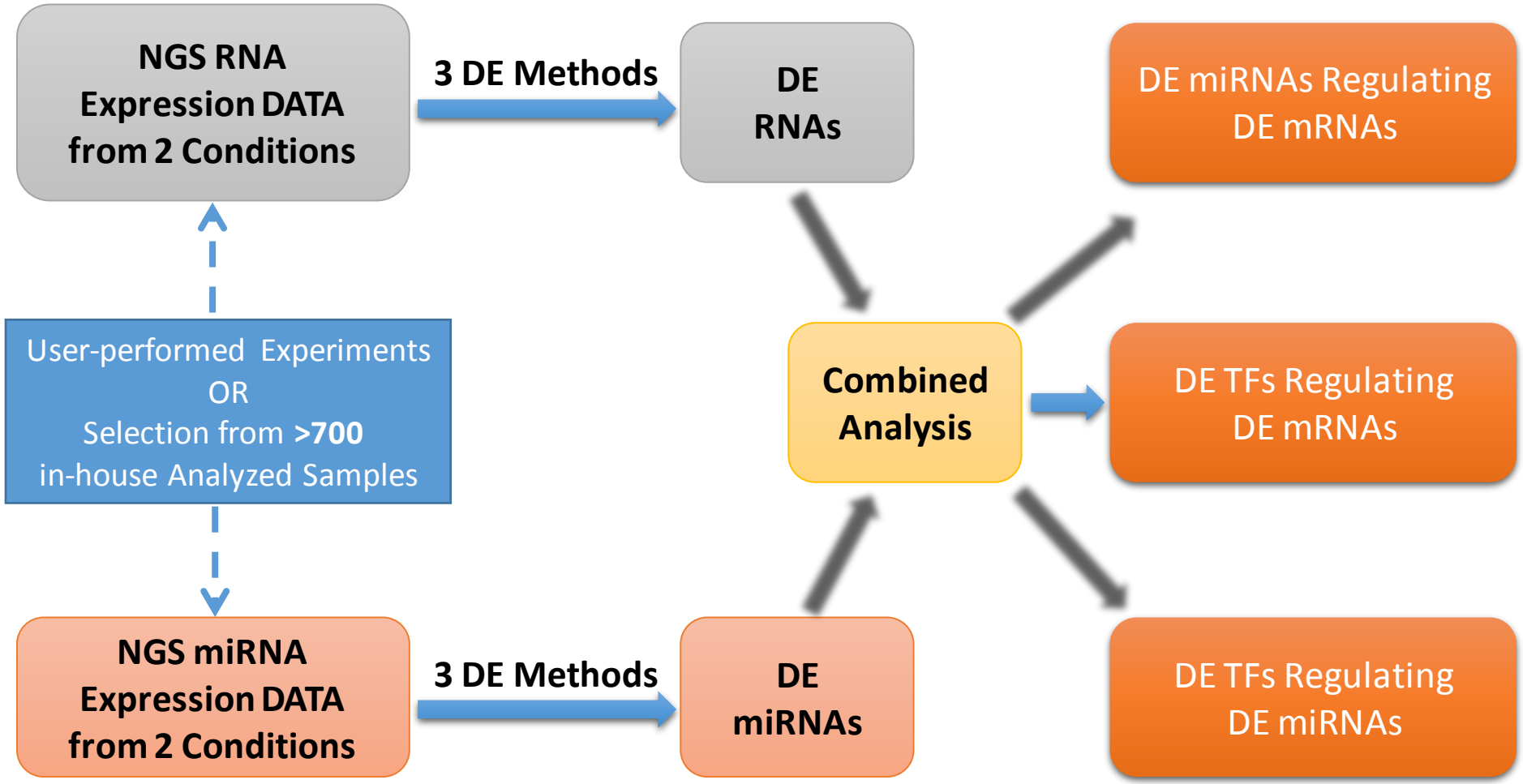


View Results per microRNA

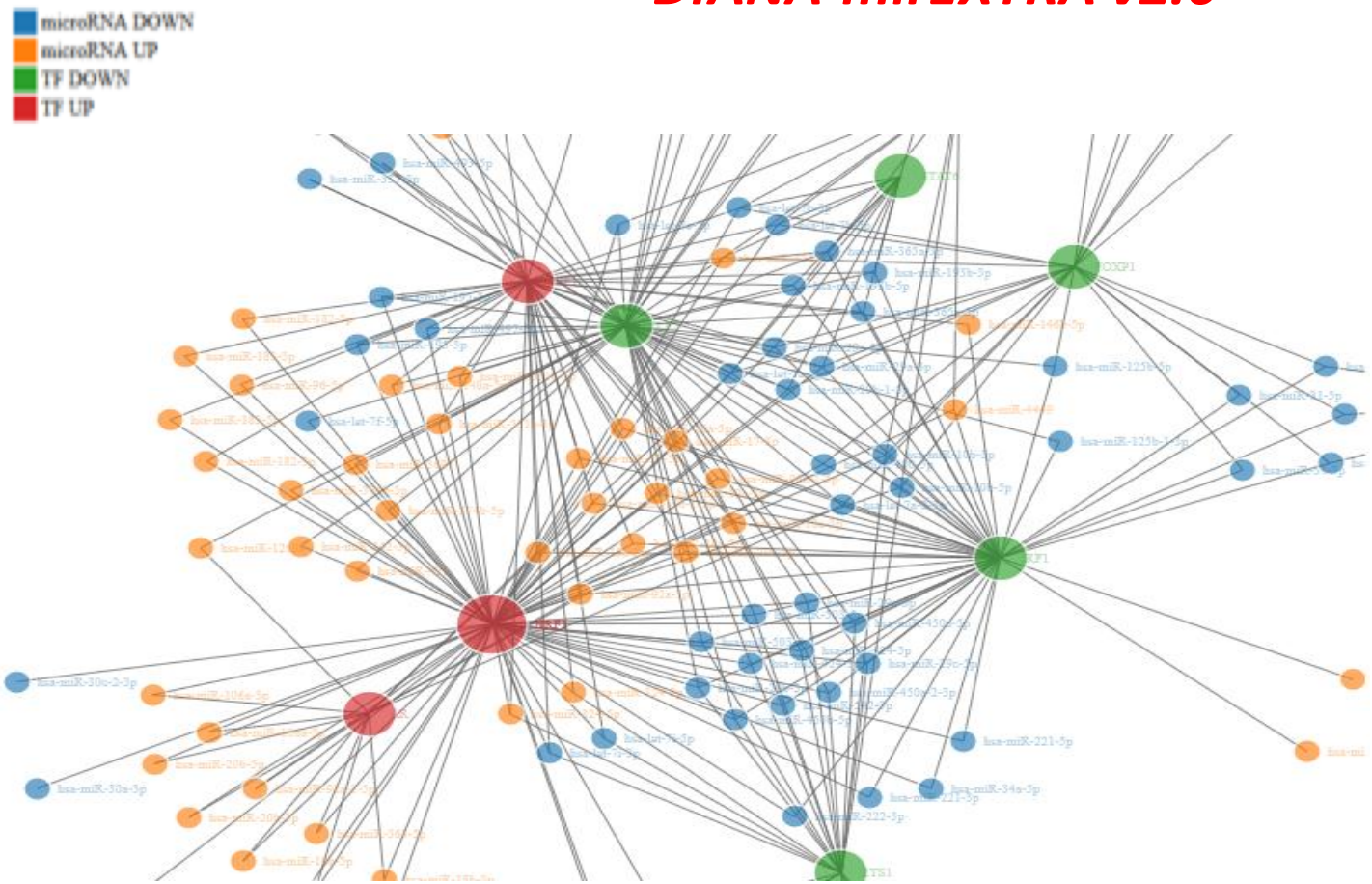
Hexamer	Start Position	miRNA	-ln(p value)	p value	Number of genes with hexamer	
ACACTC	3	mmu-miR-122a	43.74	1e-19	120	DIANA-Pathways
ACTCCA	1	mmu-miR-122a	43.74	1e-19	143	DIANA-Pathways
CACTCC	2	mmu-miR-122a	43.74	1e-19	156	DIANA-Pathways
AGCACT	2	mmu-miR-302b	15.15	2.61572454030023e-07	128	DIANA-Pathways
AGCACT	3	mmu-miR-294	15.15	2.61572454030023e-07	128	DIANA-Pathways
AGCACT	3	mmu-miR-302	15.15	2.61572454030023e-07	128	DIANA-Pathways
AGCACT	3	mmu-miR-302c	15.15	2.61572454030023e-07	128	DIANA-Pathways
AGCACT	3	mmu-miR-295	15.15	2.61572454030023e-07	128	DIANA-Pathways
AGCACT	3	mmu-miR-302d	15.15	2.61572454030023e-07	128	DIANA-Pathways

Alexiou et al, PLOS ONE 2010

miRExtra V2.0

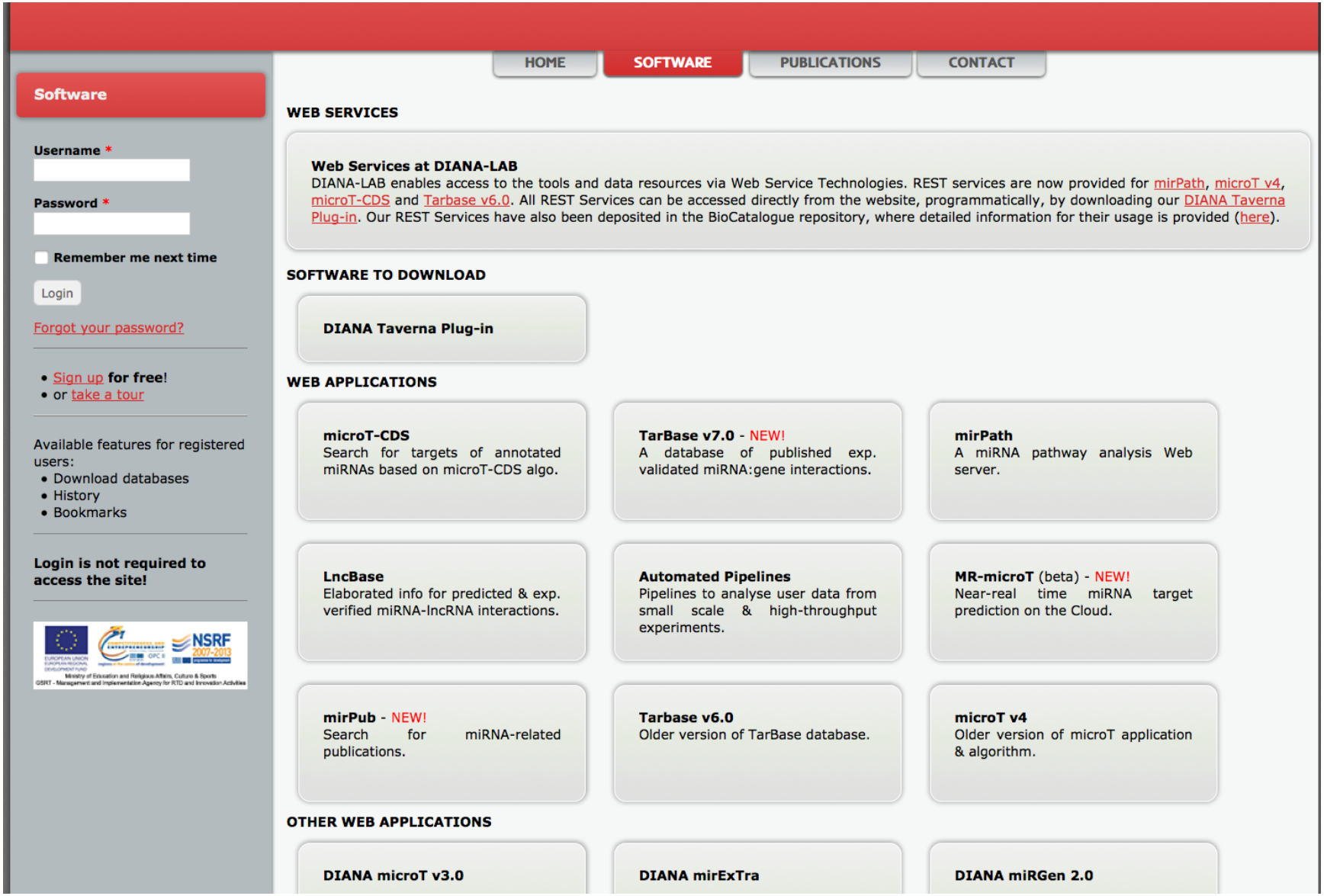


DIANA-mirEXTRA v2.0



Vlachos et al *DIANA-mirExTra v2.0: Uncovering microRNAs and Transcription Factors with crucial roles in NGS expression data.* ***NAR*** *2016*

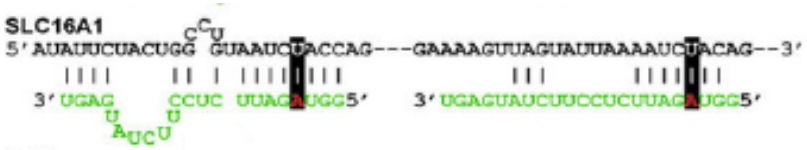
Software on microRNA.gr



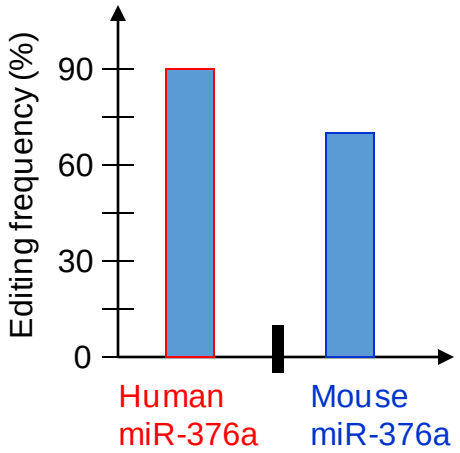
EDITING

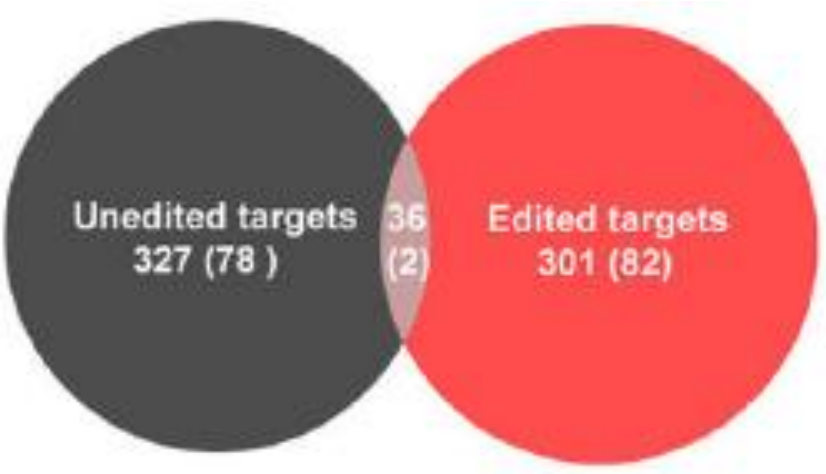
Editing of tissue specific adenosine to inosine (A->I) of mir-376 cluster.

Editing by ADAR can almost completely alter miRNA targeting activity (I pairs with G)



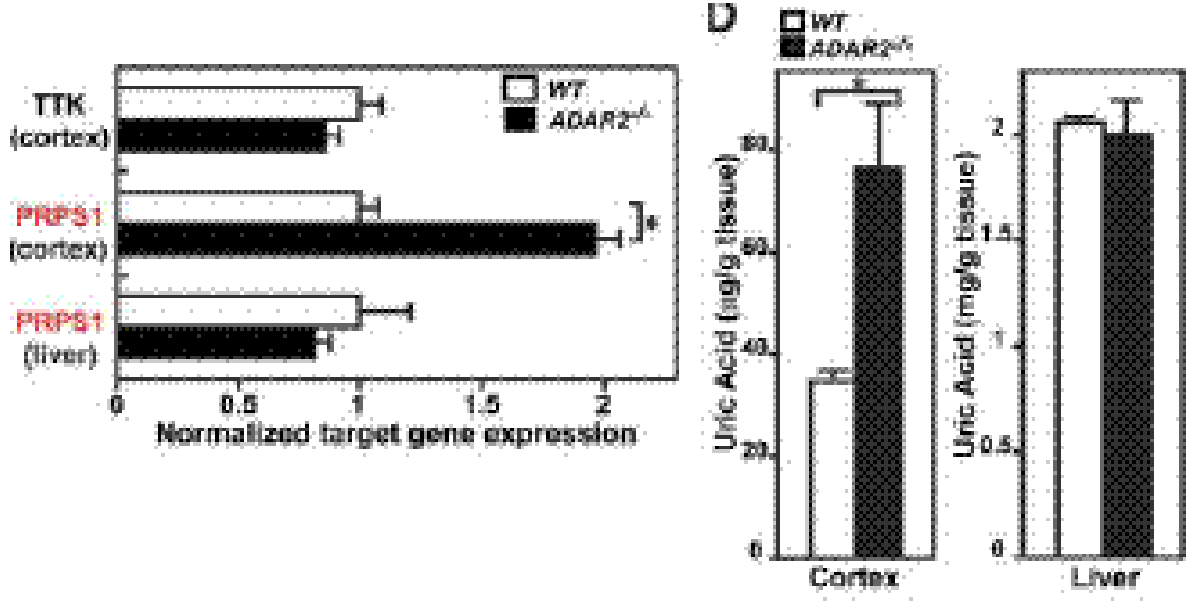
pre-miR-376a





Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs.***Science*** Feb 23;315(5815):1137-40.

Analysis of expression levels of mir-376a-5p targets in the WT and ADAR2-/- mouse.



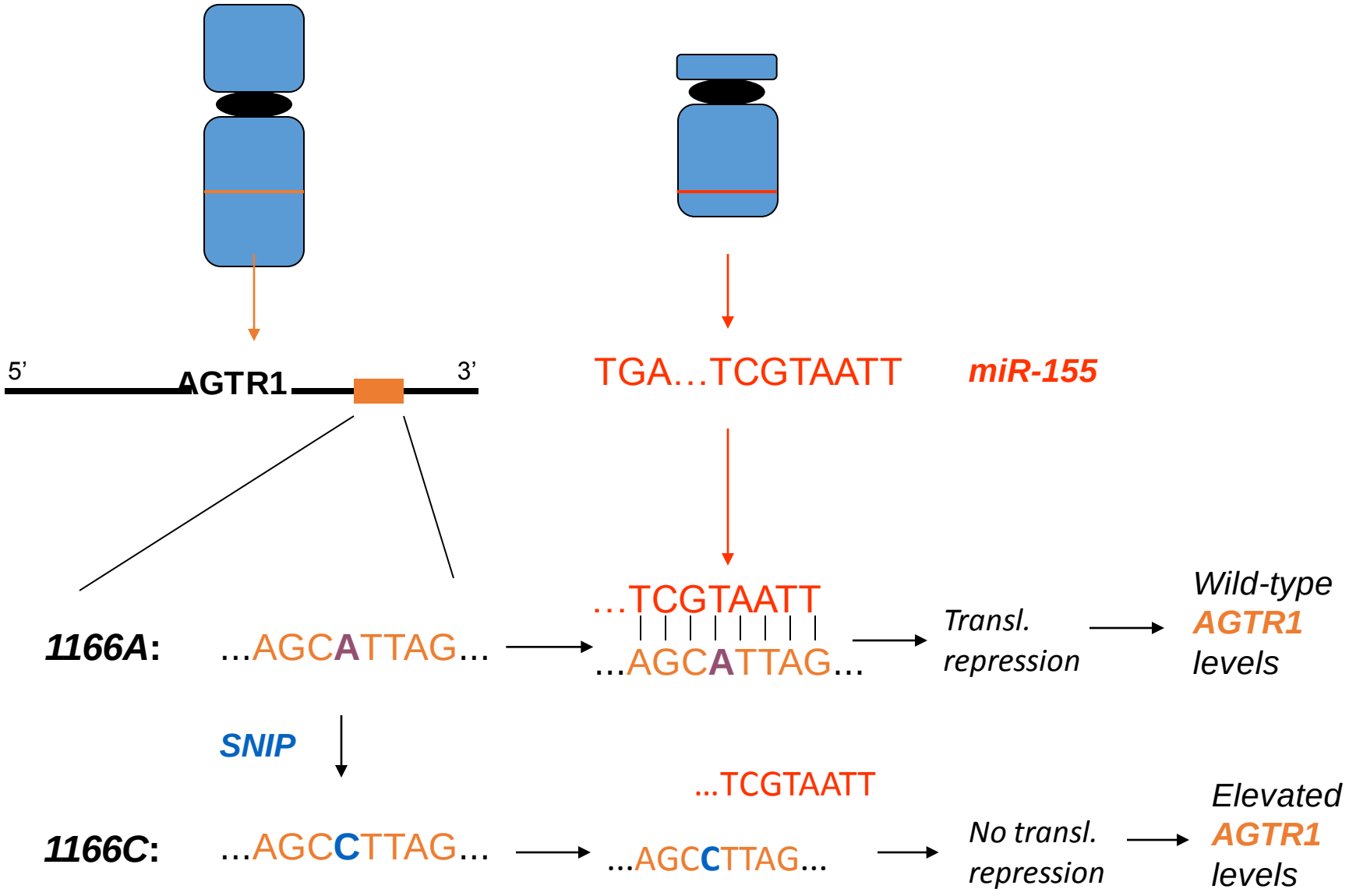
- PRSP1 is involved in purine metabolism and the Uric Acis synthesis pathway.

- Increase of PRSP1 levels cause a human disorder characterized by gout and neuro developmental impairment with hyperuricemia.

SNPs & miRNAs

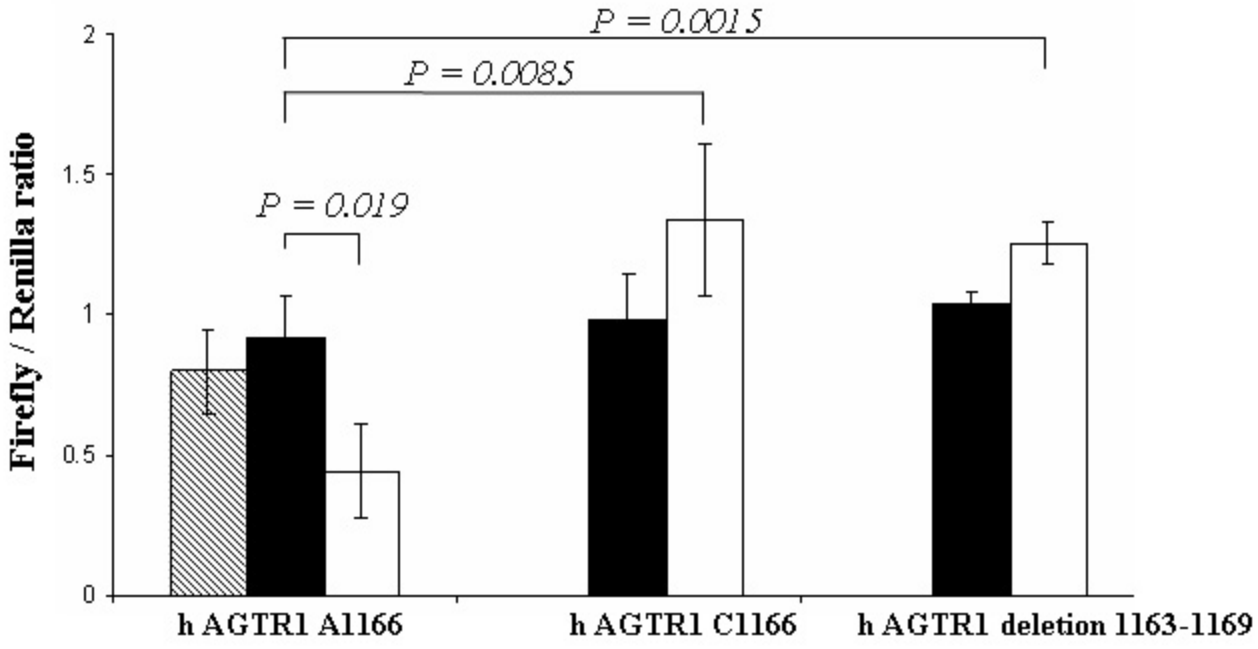
Polymorphic disease associations and microRNAs.

- SNPs that occur in functional miRNA target sites could affect miRNA binding
- Map all annotated SNPs from dbSNP onto all experimentally supported target sites from TarBase
- 2 of the 5 SNPs occur in a region that disrupts the 5'-dominant binding
- 1 of these 2 SNPs is genotyped according to ALFRED (ALlele FREquency Database)
- Does this SNP impair miR-155 binding and silencing of *AGTR1*?

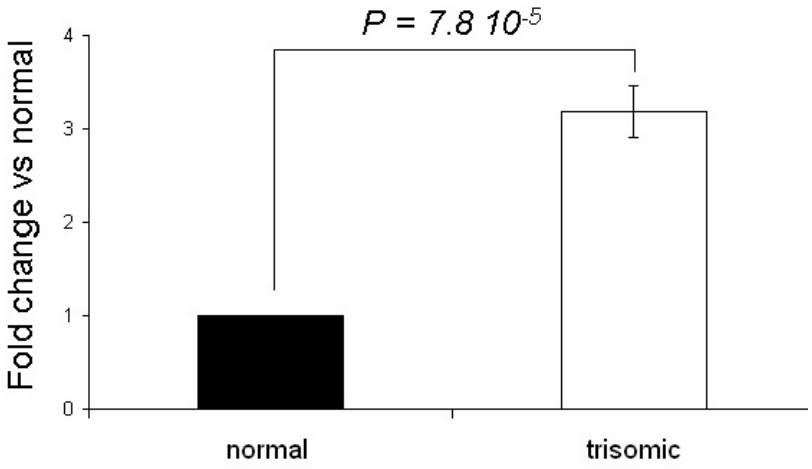

5’
AGTR1
3’
TGA…TCGTAATT
miR-155
…TCGTAATT
1166A:
…AGCATTAG…
…AGCATTAG…
Transl. repression
Wild-type AGTR1 levels
SNIP
…TCGTAATT
1166C:
…AGCCTTAG…
…AGCCTTAG…
No transl. repression
Elevated AGTR1 levels

Experimental validation in vitro.

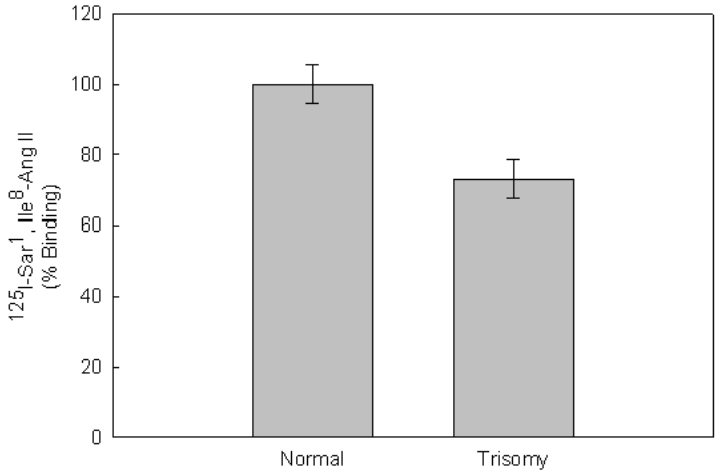
- In vitro luciferase assay to test the prediction



Fibroblast cells from monozygotic twins discordant for trisomy 21. In vivo evidence of mir-155 and ATGR1.



qRT-PCR for mature miR-155 expression

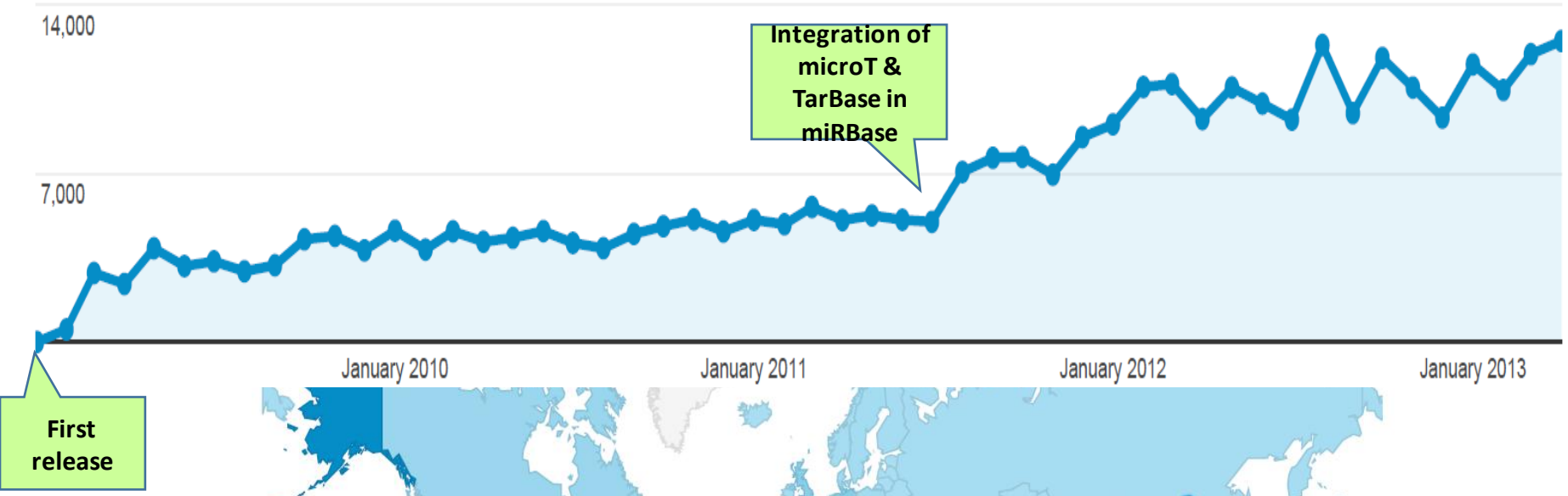


Whole cell AGTR1 binding assays

Sethupathy, P., Borel, C., Gagnebin, M., Grant, G.R, Deutsch S, Eltion TS, Hatzigeorgiou*, A.G, and Antonarakis, S.E. (2006) Human microRNA-155 on chromosome 21 differentially interacts with its polymorphic target in the AGTR1 3' untranslated region: a mechanism for functional single-nucleotide polymorphisms related to phenotypes. ***Am J Hum Genet.*** 2007 Aug;81(2):405-13.

DIANA-Tools

Visit us @
www.microrna.gr!



More than 130,000 visits per year, based on Google Analytics!

Home Reporting Customization Admin

artemis.hatzigeorgiou@gmail.com
DIANA LAB TOOLS INTEGRATED - h...
Diana lab tools

Overview

Sessions vs. Select a metric

Hourly Day Week Month

Sessions

10,000

5,000

Aug 1, 2014 - Aug 31, 2014
Sessions: 5,626

April 2014 July 2014 October 2014

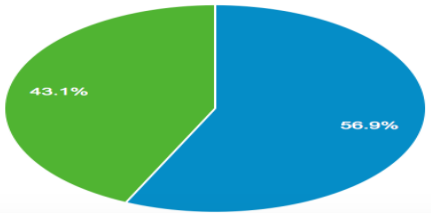
Sessions
78,824

Users
35,298

Pageviews
501,618

Pages / Session
6.36

Returning Visitor New Visitor



Home Reporting Customization Admin

artemis.hatzigeorgiou@gmail.com
http://diana.cslab.ece.ntua.gr/ - http://d...
diana.cslab.ece.ntua.gr

Overview

Sessions vs. Select a metric

Hourly Day Week Month

Sessions

6,000

3,000

Jul 1, 2014 - Jul 31, 2014
Sessions: 5,123

April 2014 July 2014 October 2014

Sessions
63,121

Users
34,539

Pageviews
254,736

Pages / Session
4.04

New Visitor Returning Visitor

- 1. Maragkakis, M., Vergoulis, T., Alexiou, P., Reczko, M., Plomaritou, K., Gousis, M., Kourtis, K., Koziris, N., Dalamagas, T., Hatzigeorgiou, A.G. DIANA-microT Web server upgrade supports Fly and Worm miRNA target prediction and bibliographic miRNA to disease association (2011) *Nucleic Acids Res.*, 39 (SUPPL. 2), pp. W145-W148. (**67 citations)**

- 2. Vergoulis, T., Vlachos, I.S., Alexiou, P., Georgakilas, G., Maragkakis, M., Reczko, M., Gerangelos, S., Koziris, N., Dalamangas T., Hatzigeorgiou, A.G. TarBase 6.0: Capturing the exponential growth of miRNA targets with experimental support (2012) *Nucleic Acids Res.*, 40 (D1), pp. D222-D229 **Highly Cited Paper**. **(345 citations)**

- 3. Reczko, M., Maragkakis, M., Alexiou, P., Grosse, I., Hatzigeorgiou, A.G. Functional microRNA targets in protein coding sequences (2012) *Bioinformatics*, 28 (6), art. no. bts043, pp. 771-776 **Highly Cited Paper**.**(142 citations)**

- 4. Vlachos, I.S., Kostoulas, N., Vergoulis, T., Georgakilas, G., Reczko, M., Maragkakis, M., Paraskevopoulou, M.D., Prionidis, K., Dalamagas, T., Hatzigeorgiou, A.G. DIANA miRPath v.2.0: Investigating the combinatorial effect of microRNAs in pathways (2012) *Nucleic Acids Res.*, 40 (W1), pp. W498-W504 **Highly Cited Paper**. (**250 citations)**

- 5. Reczko, M., Maragkakis, M., Alexiou, P., Papadopoulos, G.L., Hatzigeorgiou, A.G. Accurate microRNA target prediction using detailed binding site accessibility and machine learning on proteomics data (2012) Front Genet., 2012 Jan 18;2:103. doi: 10.3389/fgene.2011.00103.

- 6. Paraskevopoulou, M.D., Georgakilas, G., Kostoulas, N., Reczko, M., Maragkakis, M., Dalamagas, T.M., Hatzigeorgiou, A.G. DIANA-LncBase: Experimentally verified and computationally predicted microRNA targets on long non-coding RNAs (2013) *Nucleic Acids Res.*, 41 (D1), pp. D239-D245 **Highly Cited Paper**. **(102 citations)**

- 7. Paraskevopoulou, M.D., Georgakilas, G., Kostoulas, N., Vlachos, I.S., Vergoulis, T., Reczko, M., Filippidis, C., Dalamagas, T., Hatzigeorgiou, A.G. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. (2013) *Nucleic Acids Res.*, 41 (Web Server issue), pp. W169-173 **Highly Cited Paper . (170 citations)**

- 8. Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos IL, Maniou S, Karathanou K, Kalfakakou D, Fevgas A, Dalamagas T, Hatzigeorgiou AG. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res*. 2015 Jan 28;43 (Database issue):D153-9. doi: 10.1093/nar/gku1215. Epub 2014 Nov 21, **Highly Cited Paper.(70 citations)**

- *9. Georgakilas G, Vlachos IS, Paraskevopoulou MD, Yang P, Zhang Y, Economides AN, Hatzigeorgiou AG. microTSS: accurate microRNA transcription start site identification reveals a significant number of divergent pri-miRNAs. Nature Communications. 2014 Dec 10;5:5700. doi: 10.1038/ncomms6700.*

- *10. Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res. 2015 Jul 1;43(W1):W460-6.*

- *11. Georgakilas G, Vlachos IS, Zagganas K, Vergoulis T, Paraskevopoulou MD, Kanellos I, Tsanakas P, Dellis D. , FevgasA. Dalamagas T, Hatzigeorgiou AG. miRGen v3.0: accurate characterization of microRNA promoters and their regulators. Nucleic Acids Res. 2016 Jan 4;44(D1):D190-5. doi: 10.1093/nar/gkv1254.*

- *12. Paraskevopoulou, MD, Vlachos, IS, Karagkouni D, Georgakilas, G, Kanellos I, N., Vergoulis, Tsanakas P, Floros E, Dalamagas T, Hatzigeorgiou AG. DIANA-LncBase v2: Indexing microRNA targets on non-coding transcripts Nucleic Acids Res. 2016 Jan 4;44(D1):D231-8. doi: 10.1093/nar/gkv1270.*

- *13. Vlachos IS, Vergoulis T, Paraskevopoulou MD, Lykokanellos F, Georgakilas G, Georgiou P, Chatzopoulos S, Karagkouni D, Christodoulou F, Dalamagas T, Hatzigeorgiou AG. DIANA-mirExTra v2.0: Uncovering microRNAs and transcription factors with crucial roles in NGS expression data. Nucleic Acids Res. 2016 May 20. pii: gkw455.*