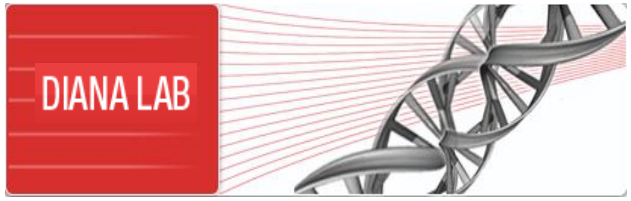


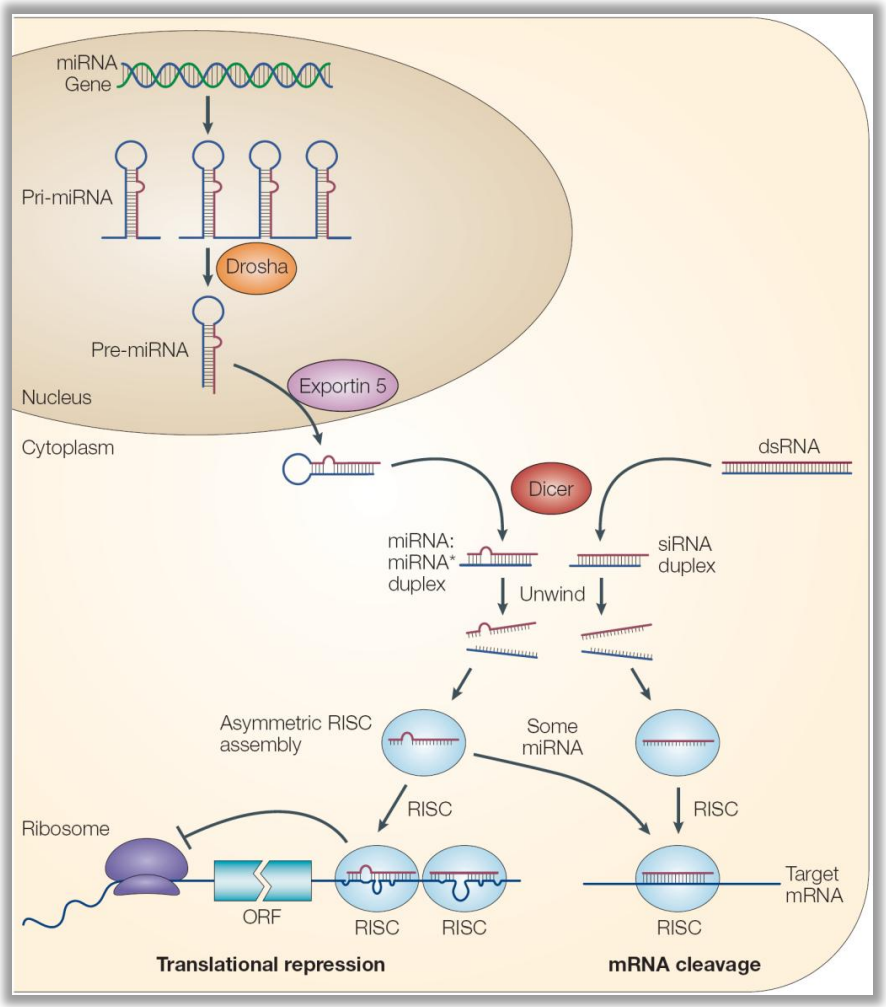
COST Conference "Next Generation Sequencing: a look into the future

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

Dimitra Karagkouni

Principal Investigator
Dr. Artemis Hatzigeorgiou

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% specificity[1]

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

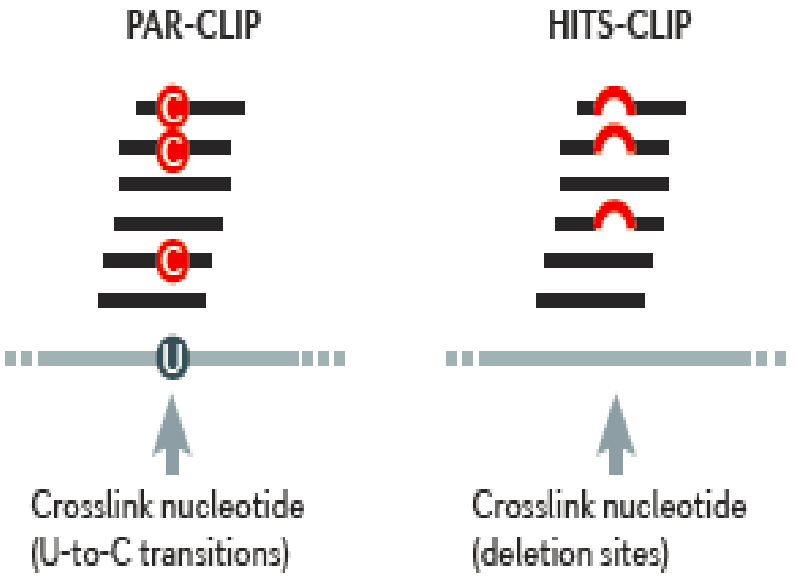
Experimental Determination of miRNA – Gene Interactions

- **Gene Specific Techniques**
  - Reporter genes
  - Northern blotting
  - qPCR
  - Western blotting
  - ELISA
  - Immunihistochemistry

- **High Throughput Technique**s
  - MicroArrays
  - RNA-Seq
  - Proteomics (such as pSILAC)
  - CLIP-Seq (HITS-CLIP, PAR-CLIP, iCLIP)
  - CLASH
  - PARE-Seq
  - Degradome-Seq
  - Biotin tagged miRNA

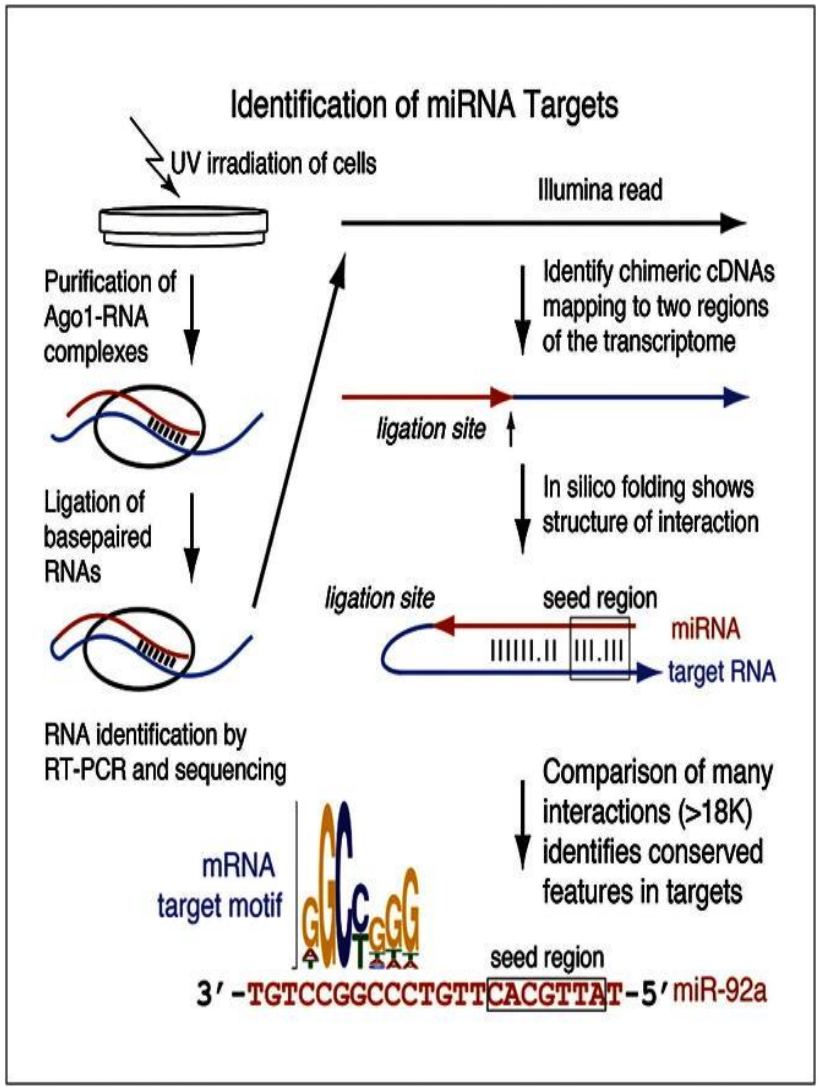
CLIP experiments

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



König J et al Nat Rev Genet. 2012 Jan 18;13(2)

CLASH experiment



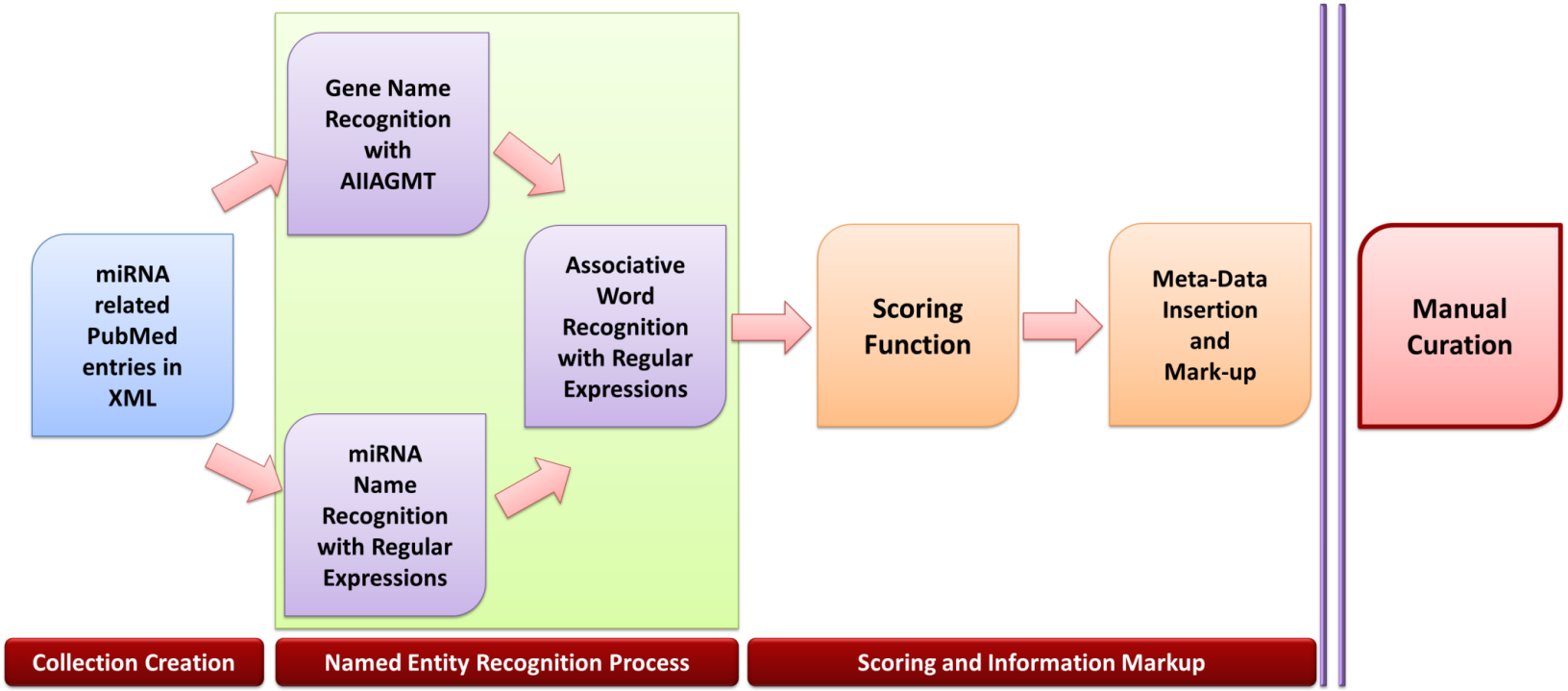
Aleksandra Helwak *et al*, Cell 153, 654–665, April 25, 2013

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

1:I. 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.),

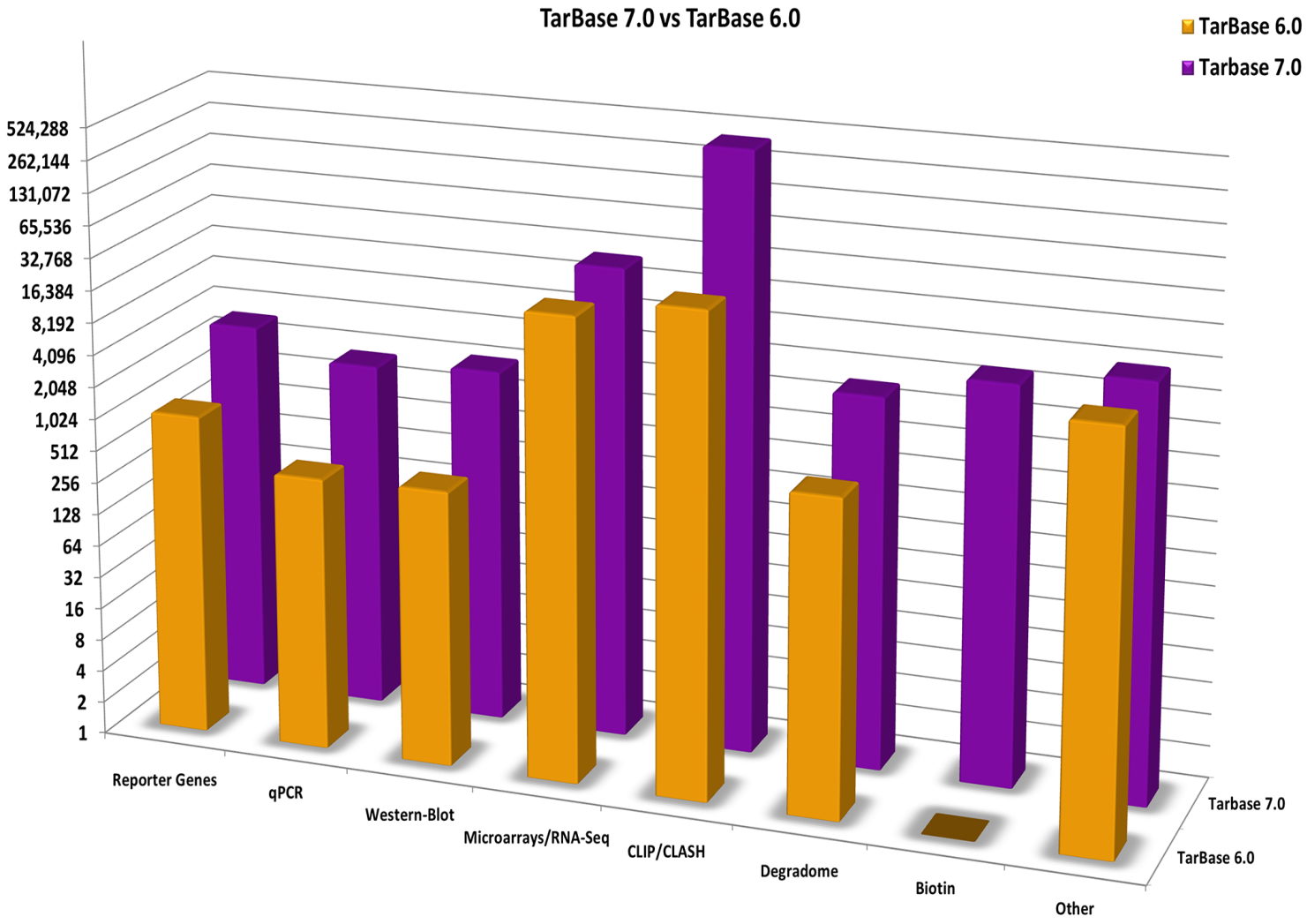
ISSN: 1065-6995, **Date:** 01 / 12 / 2010

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

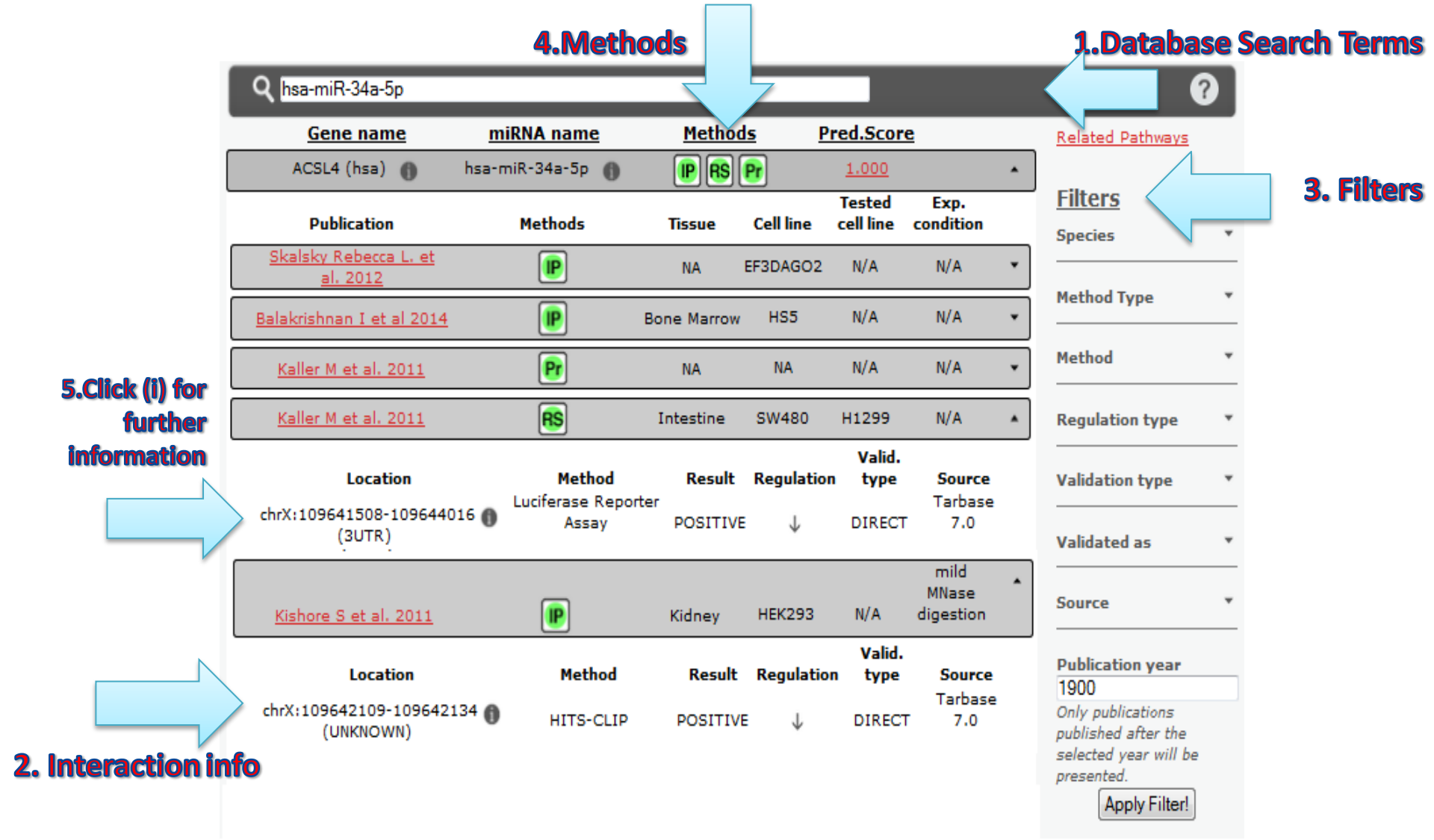
Analysis of raw data

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

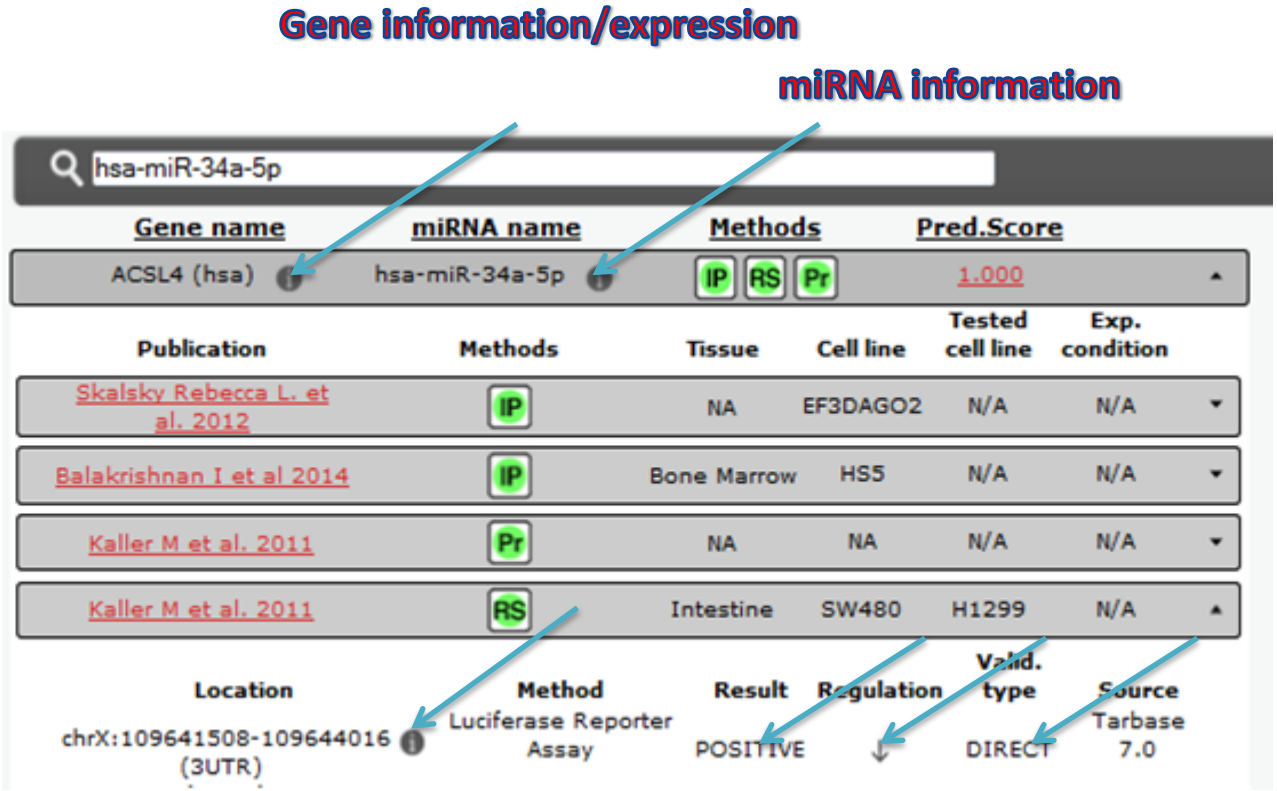
DIANA-TarBase v7.0 – Statistics

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

TarBase v7.0 Interface



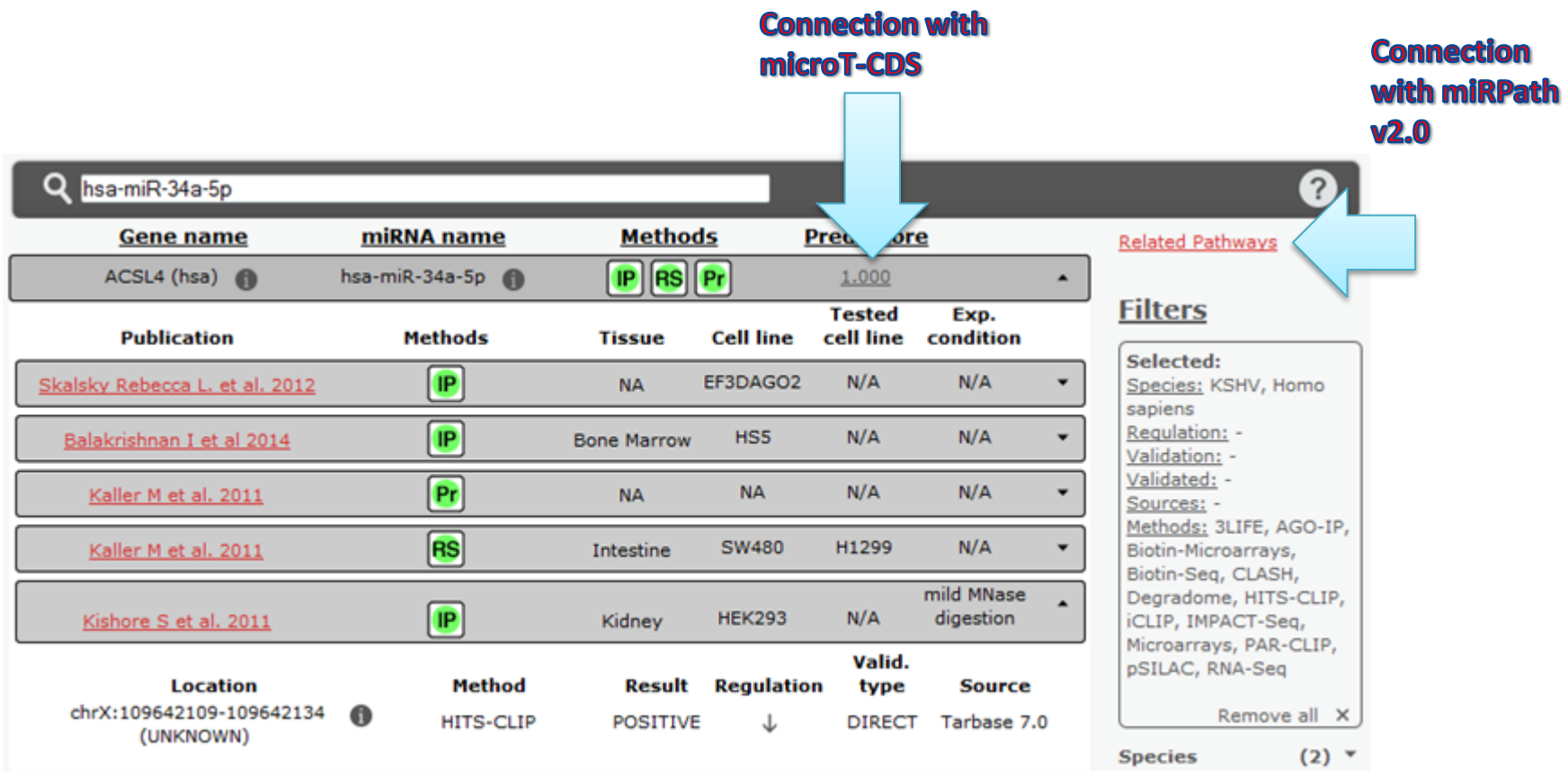
TarBase v7.0 Interface



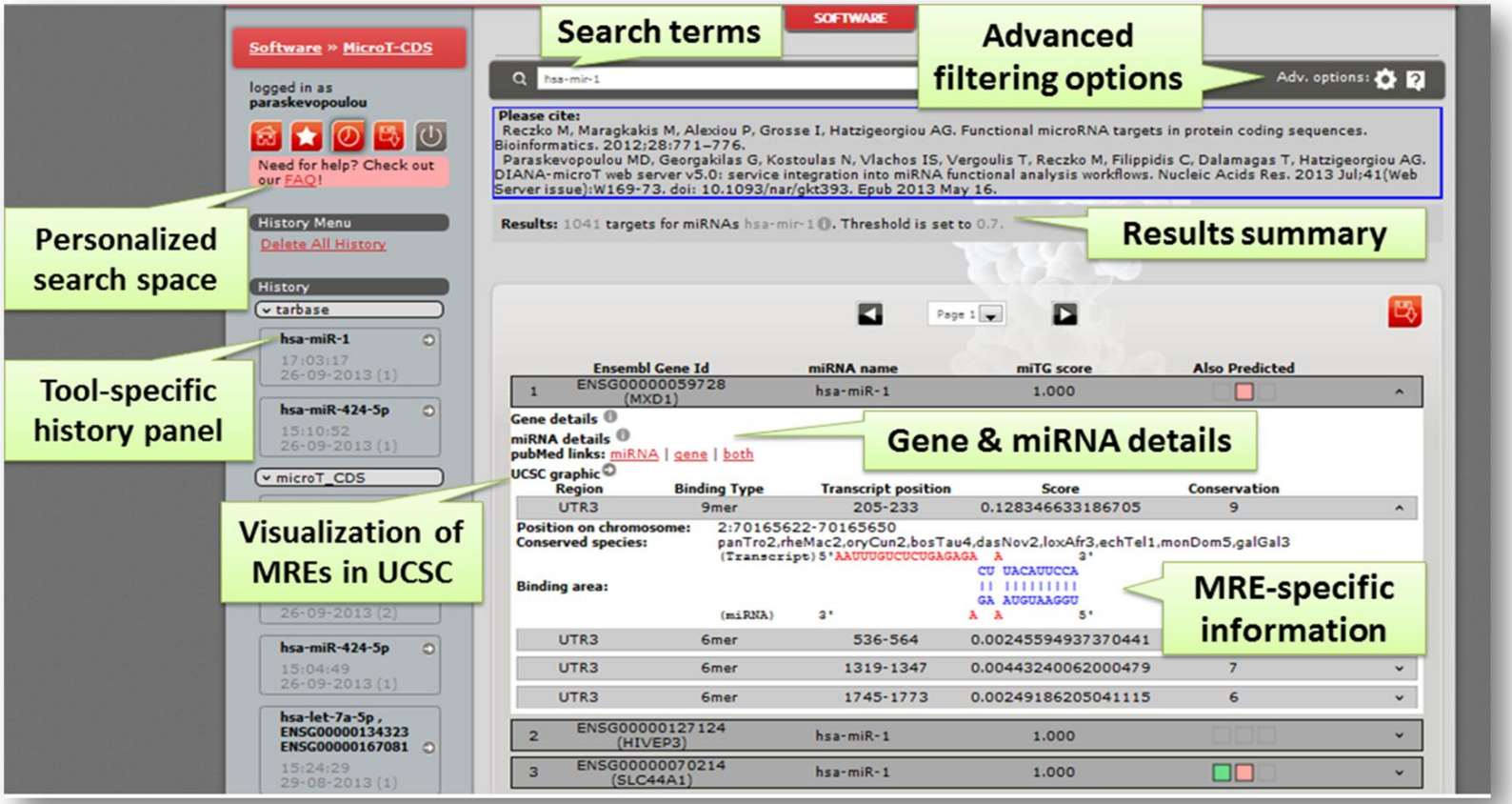
TarBase v7.0 Interface

- 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

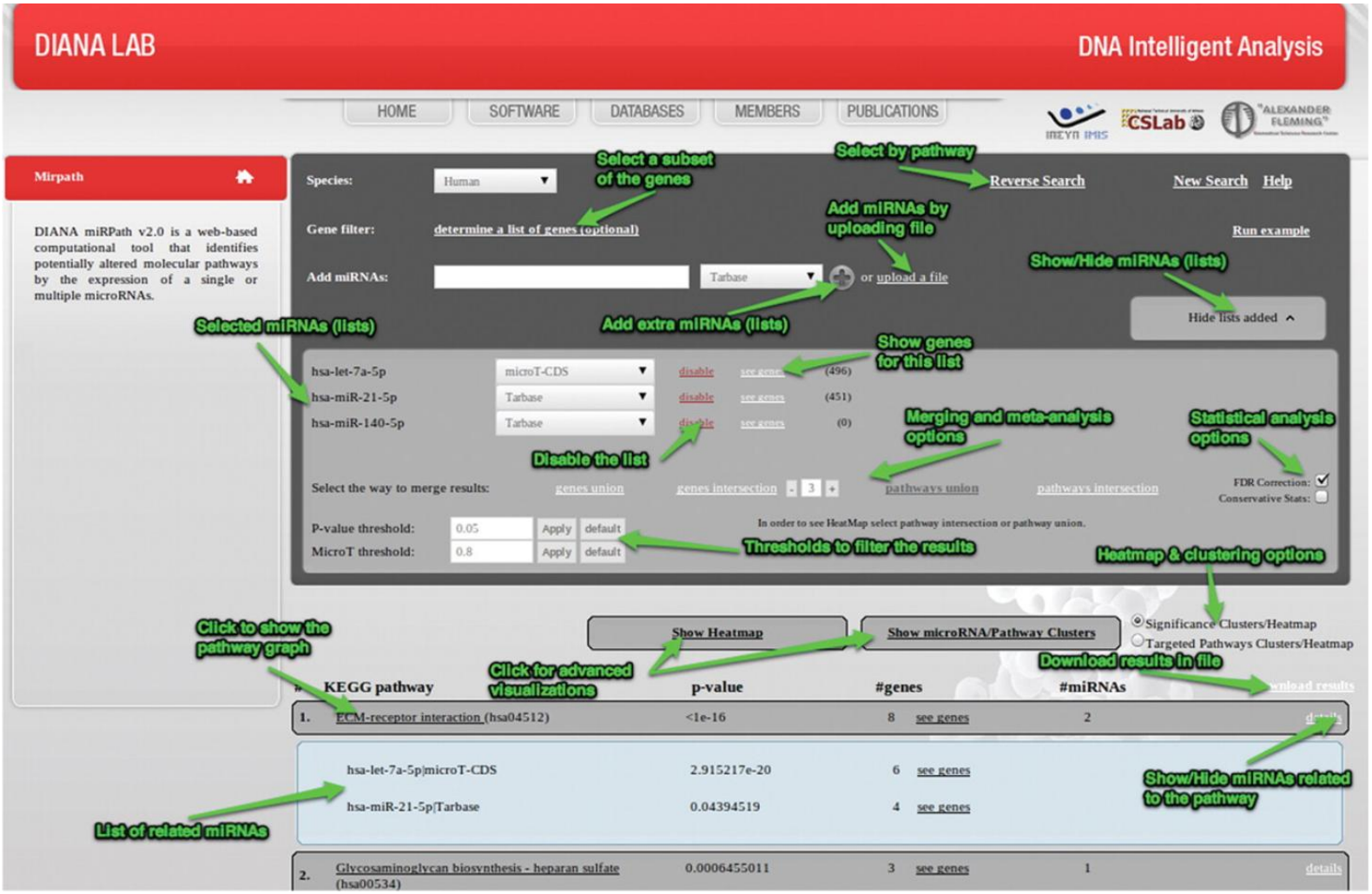
TarBase v7.0 Interface



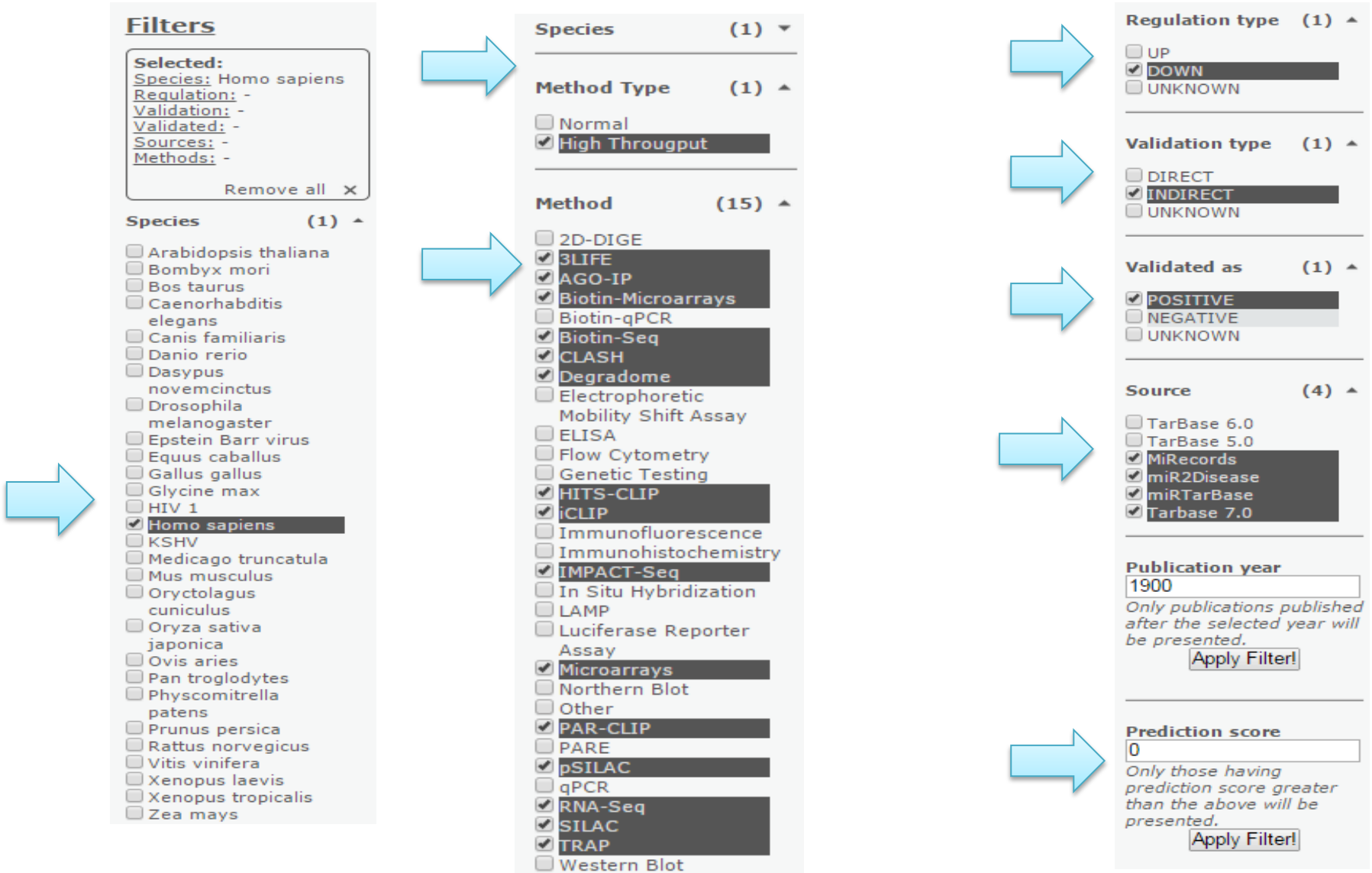
Diana microT-CDS Interface



miRPath v2.0 Interface



TarBase v7.0 Interface - Advanced filter options

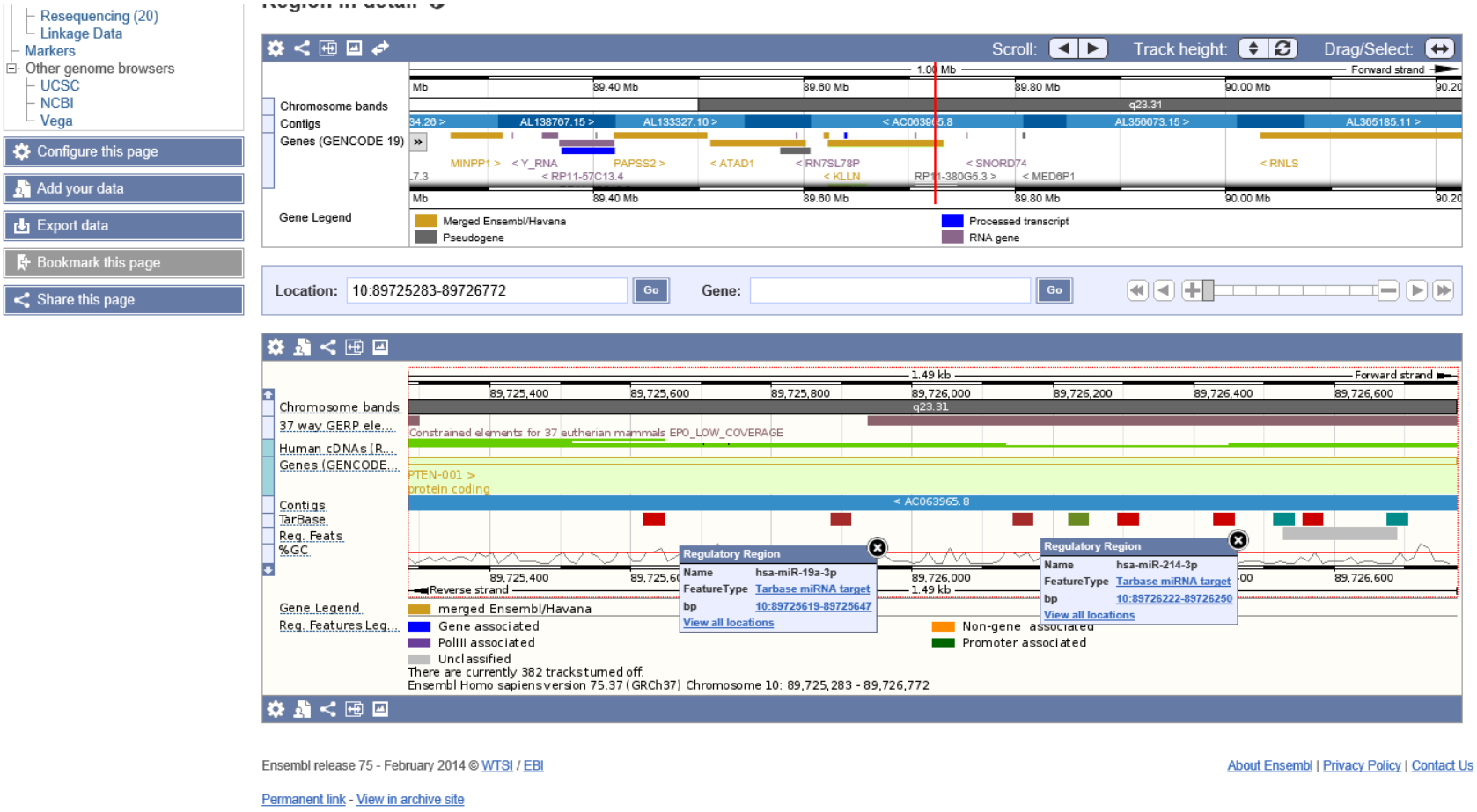


ENSEMBL Integration

ENSEMBL integration:

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

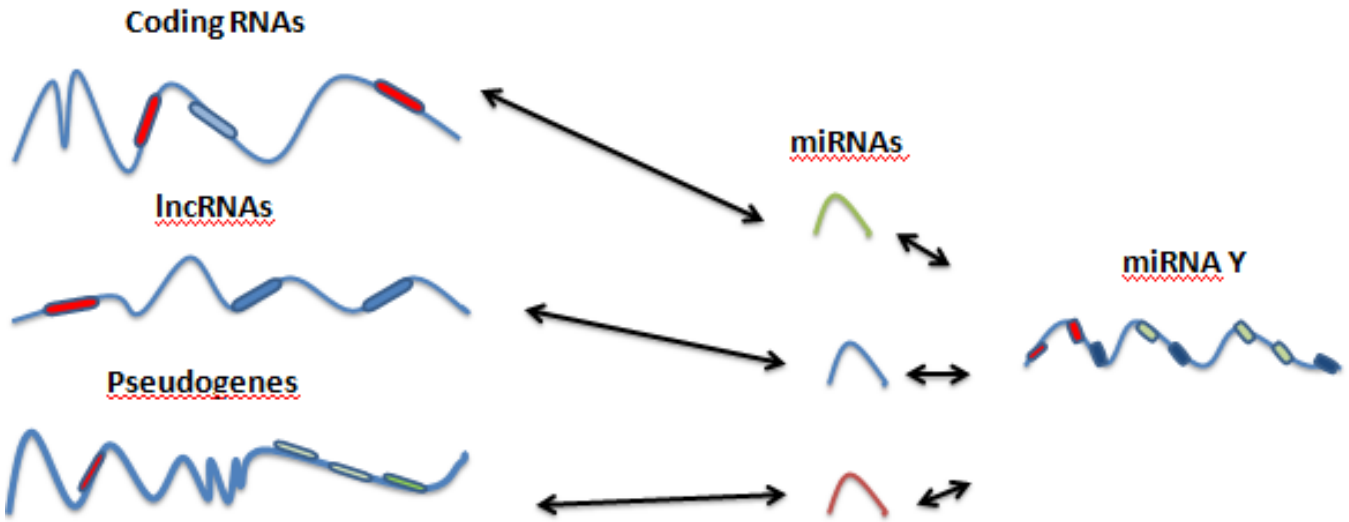
ENSEMBL Integration



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)*



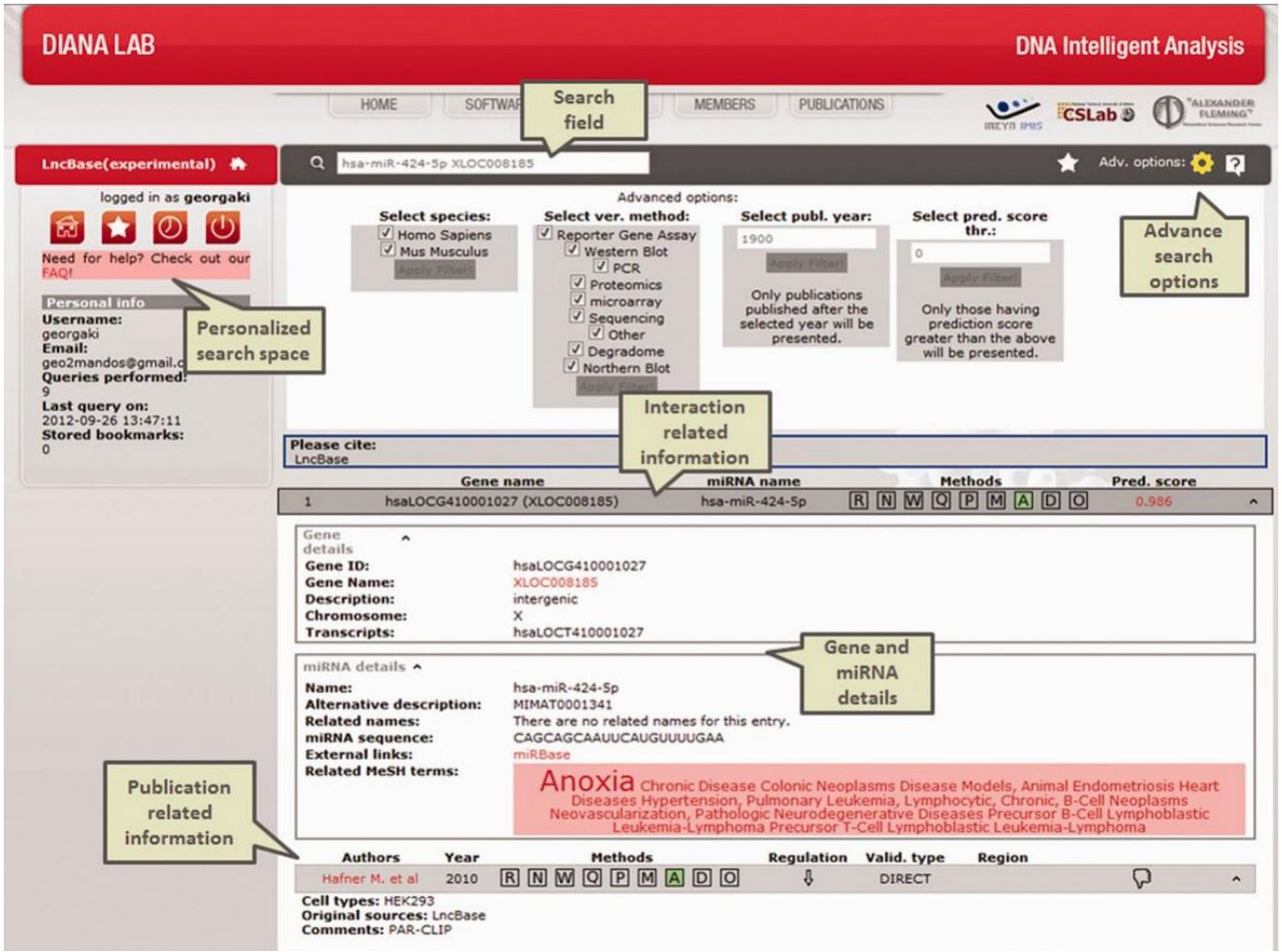
Competing endogenous RNA (ceRNA) network

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)*



Acknowledgements

Lab Head

- **Dr. Artemis Hatzigeorgiou**

Lab members

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

Partners

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

www.microrna.gr