# Artemis G. Hatzigeorgiou

*Professor of Bioinformatics*
*Department of Electrical and Computer Engineering, University of Thessaly*

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

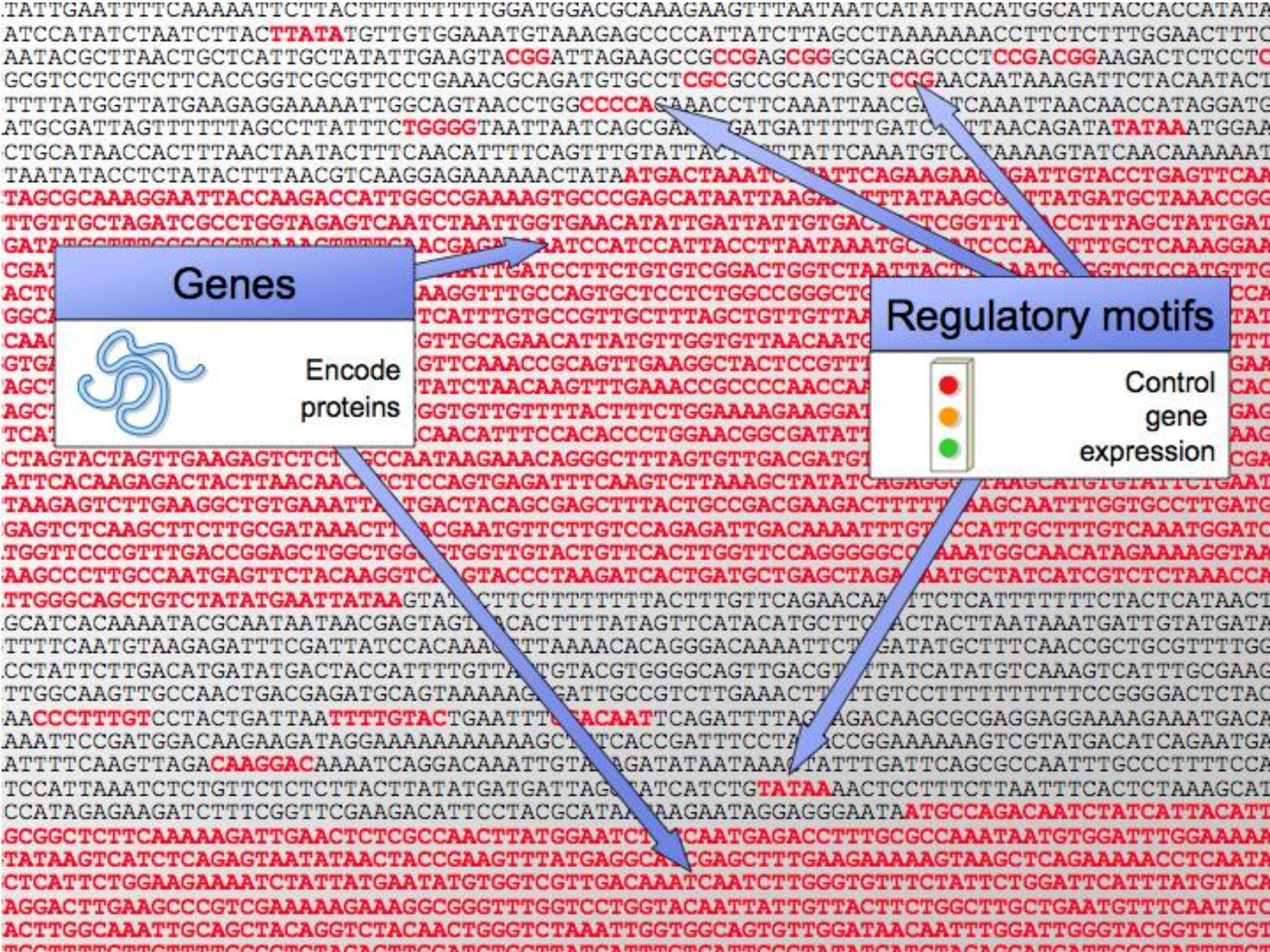
ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

# What is Bioinformatics?

- Bioinformatics is generally defined as the analysis, prediction, modeling and storage of biological data with the help of computers

```
TATTGAATTTTCAAAAATTCTTACTTTTTTTTTGGATGGACGCAAAGAAGTTTAATAATCATATTACATGGCATTACCACCATATA
ATCCATATCTAATCTTACTTATATGTTGTGGAAATGTAAAGAGCCCCATTATCTTAGCCTAAAAAAACCTTCTCTTTGGAACTTTC
AATACGCTTAACTGCTCATTGCTATATTGAAGTACGGATTAGAAGCCGCCGAGCGGGCGACAGCCCTCCGACGGAAGACTCTCCTC
GCGTCCTCGTCTTCACCGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATACT
TTTTATGGTTATGAAGAGGAAAAATTGGCAGTAACCTGGCCCCACAAACCTTCAAATTAACGAATCAAATTAACAACCATAGGATG
ATGCGATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTAACAGATATATAAATGGAA
CTGCATAACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTACTTCTTATTCAAATGTCATAAAAGTATCAACAAAAAAT
TAATATACCTCTATACTTTAACGTCAAGGAGAAAAAACTATAATGACTAAATCTCATTCAGAAGAAGTGATTGTACCTGAGTTCAA
TAGCGCAAAGGAATTACCAAGACCATTGGCCGAAAAGTGCCCGAGCATAATTAAGAAATTTATAAGCGCTTATGATGCTAAACCGG
TTGTTGCTAGATCGCCTGGTAGAGTCAATCTAATTGGTGAACATATTGATTATTGTGACTTCTCGGTTTTACCTTTAGCTATTGAT
GATATGCTTTGCGCCGTCAAAGTTTTGAACGAGAAAAATCCATCCATTACCTTAATAAATGCTGATCCCAAATTTGCTCAAAGGAA
CGATTTGCCGTTGGACGGTTCTTATGTCACAATTGATCCTTCTGTGTCGGACTGGTCTAATTACTTTAAATGTGGTCTCCATGTTG
ACTCTTTTCTAAAGAAACTTGCACCGGAAAGGTTTGCCAGTGCTCCTCTGGCCGGGCTGCAAGTCTTCTGTGAGGGTGATGTACCA
GGCAGTGGATTGTCTTCTTCGGCCGCATTCATTTGTGCCGTTGCTTTAGCTGTTGTTAAAGCGAATATGGGCCCTGGTTATCATAT
CAAGCAAAATTTAATGCGTATTACGGTCGTTGCAGAACATTATGTTGGTGTTAACAATGGCGGTATGGATCAGGCTGCCTCTGTTT
GTGAGGAAGATCATGCTCTATACGTTGAGTTCAAACCGCAGTTGAAGGCTACTCCGTTTAAATTTCCGCAATTAAAAAACCATGAA
AGCTTTGTTATTGCGAACACCCTTGTTGTATCTAACAAGTTTGAAACCGCCCCAACCAACTATAATTTAAGAGTGGTAGAAGTCAC
AGCTGCAAATGTTTTAGCTGCCACGTACGGTGTTGTTTTACTTTCTGGAAAAGAAGGATCGAGCACGAATAAAGGTAATCTAAGAG
TCATGAACGTTTATTATGCCAGATATCACAACATTTCCACACCCTGGAACGGCGATATTGAATCCGGCATCGAACGGTTAACAAAG
CTAGTACTAGTTGAAGAGTCTCTCGCCAATAAGAAACAGGGCTTTAGTGTTGACGATGTCGCACAATCCTTGAATTGTTCTCGCGA
ATTCACAAGAGACTACTTAACAACATCTCCAGTGAGATTTCAAGTCTTAAAGCTATATCAGAGGGCTAAGCATGTGTATTCTGAAT
TAAGAGTCTTGAAGGCTGTGAAATTAATGACTACAGCGAGCTTTACTGCCGACGAAGACTTTTTCAAGCAATTTGGTGCCTTGATG
GAGTCTCAAGCTTCTTGCGATAAACTTTACGAATGTTCTTGTCCAGAGATTGACAAAATTTGTTCCATTGCTTTGTCAAATGGATC
TGGTTCCCGTTTGACCGGAGCTGGCTGGGGTGGTTGTACTGTTCACTTGGTTCCAGGGGGCCCAAATGGCAACATAGAAAAGGTAA
AAGCCCTTGCCAATGAGTTCTACAAGGTCAAGTACCCTAAGATCACTGATGCTGAGCTAGAAAATGCTATCATCGTCTCTAAACCA
TTGGGCAGCTGTCTATATGAATTATAAGTATACTTCTTTTTTTTACTTTGTTCAGAACAACTTCTCATTTTTTTCTACTCATAACT
GCATCACAAAATACGCAATAATAACGAGTAGTAACACTTTTATAGTTCATACATGCTTCAACTACTTAATAAATGATTGTATGATA
TTTTCAATGTAAGAGATTTCGATTATCCACAAACTTTAAAACACAGGGACAAAATTCTTGATATGCTTTCAACCGCTGCGTTTTGG
CCTATTCTTGACATGATATGACTACCATTTTGTTATTGTACGTGGGGCAGTTGACGTCTTATCATATGTCAAAGTCATTTGCGAAG
TTGGCAAGTTGCCAACTGACGAGATGCAGTAAAAAGAGATTGCCGTCTTGAAACTTTTTGTCCTTTTTTTTTTCCGGGGACTCTAC
AACCCTTTGTCCTACTGATTAATTTTGTACTGAATTTGGACAATTCAGATTTTAGTAGACAAGCGCGAGGAGGAAAAGAAATGACA
AAATTCCGATGGACAAGAAGATAGGAAAAAAAAAAAAGCTTTCACCGATTTCCTAGACCGGAAAAAAGTCGTATGACATCAGAATGA
ATTTTCAAGTTAGACAAGGACAAAATCAGGACAAATTGTAAAGATATAATAAACTATTTGATTCAGCGCCAATTTGCCCTTTTCCA
TCCATTAAATCTCTGTTCTCTCTTACTTATATGATGATTAGGTATCATCTGTATAAAACTCCTTTCTTAATTTCACTCTAAAGCAT
CCATAGAGAAGATCTTTCGGTTCGAAGACATTCCTACGCATAATAAGAATAGGAGGGAATAATGCCAGACAATCTATCATTACATT
GCGGCTCTTCAAAAAGATTGAACTCTCGCCAACTTATGGAATCTTCCAATGAGACCTTTGCGCCAAATAATGTGGATTTGGAAAAA
TATAAGTCATCTCAGAGTAATATAACTACCGAAGTTTATGAGGCATCGAGCTTTGAAGAAAAAGTAAGCTCAGAAAAACCTCAATA
CTCATTCTGGAAGAAAATCTATTATGAATATGTGGTCGTTGACAAATCAATCTTGGGTGTTTCTATTCTGGATTCATTTATGTACA
AGGACTTGAAGCCCGTCGAAAAAGAAAGGCGGGTTTGGTCCTGGTACAATTATTGTTACTTCTGGCTTGCTGAATGTTTCAATATC
ACTTGGCAAATTGCAGCTACAGGTCTACAACTGGGTCTAAATTGGTGGCAGTGTTGGATAACAATTTGGATTGGGTACGGTTTCGT
```



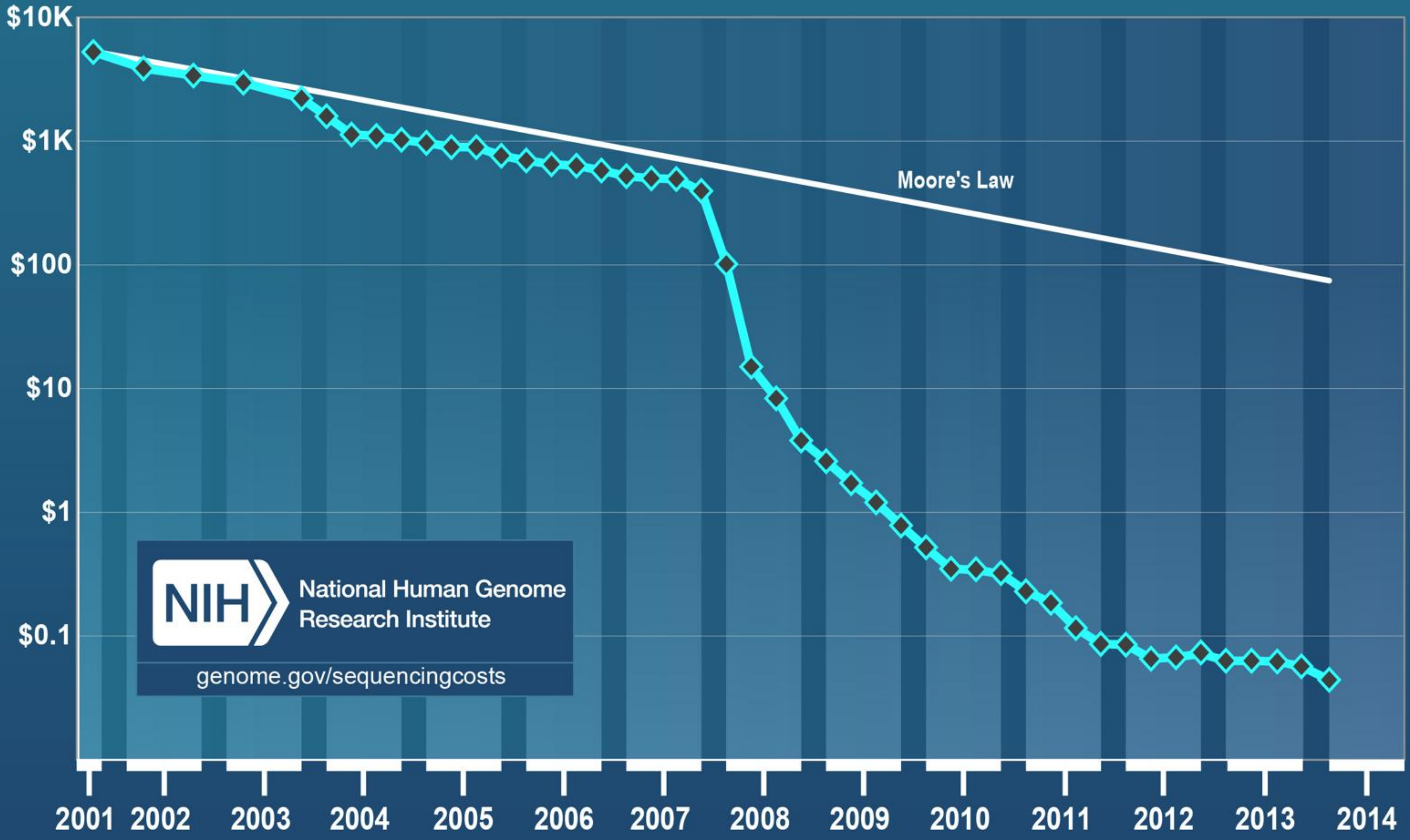
**Genes**

Encode proteins

**Regulatory motifs**

Control gene expression

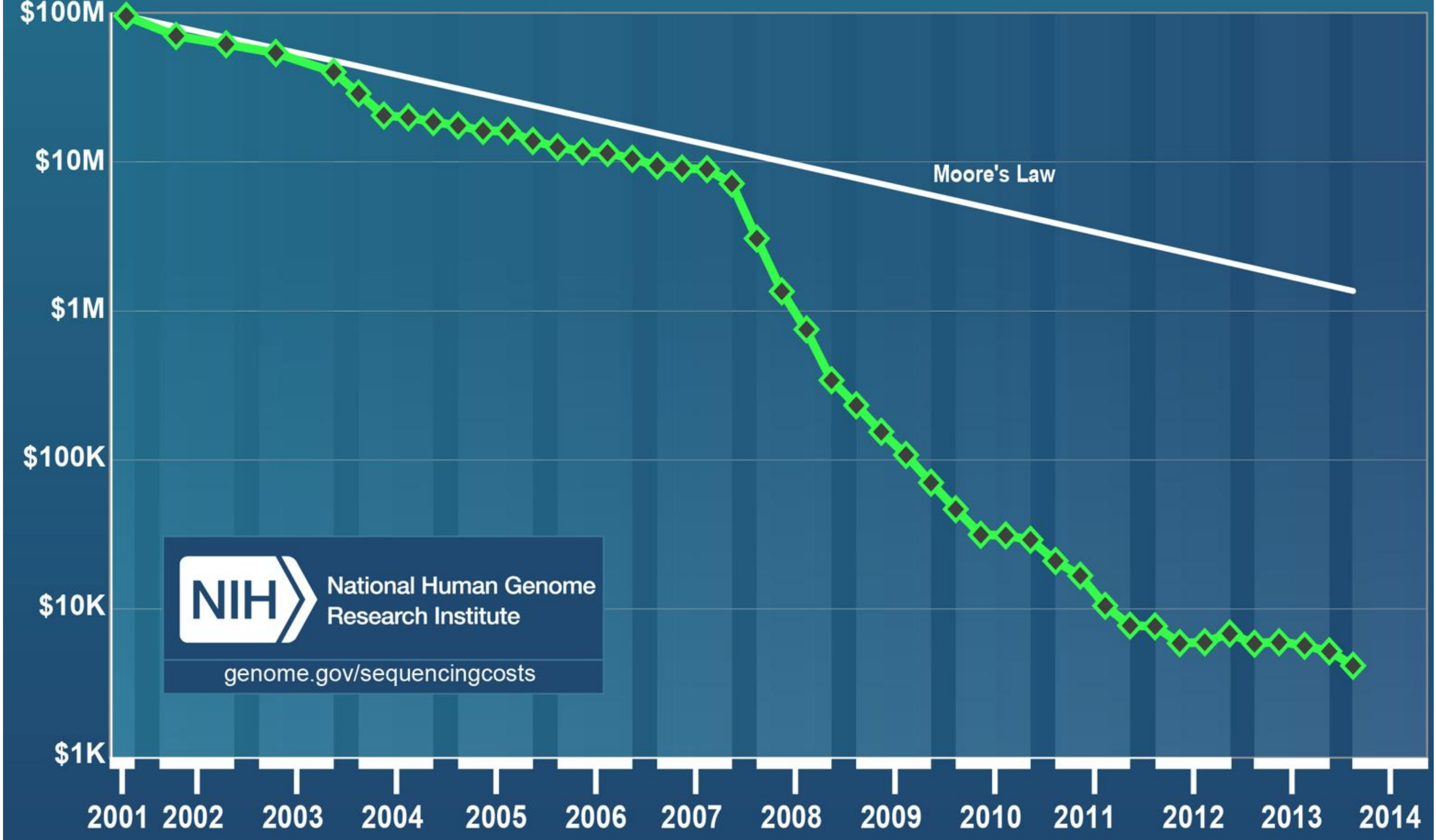
# Cost per Raw Megabase of DNA Sequence



NIH National Human Genome Research Institute

genome.gov/sequencingcosts

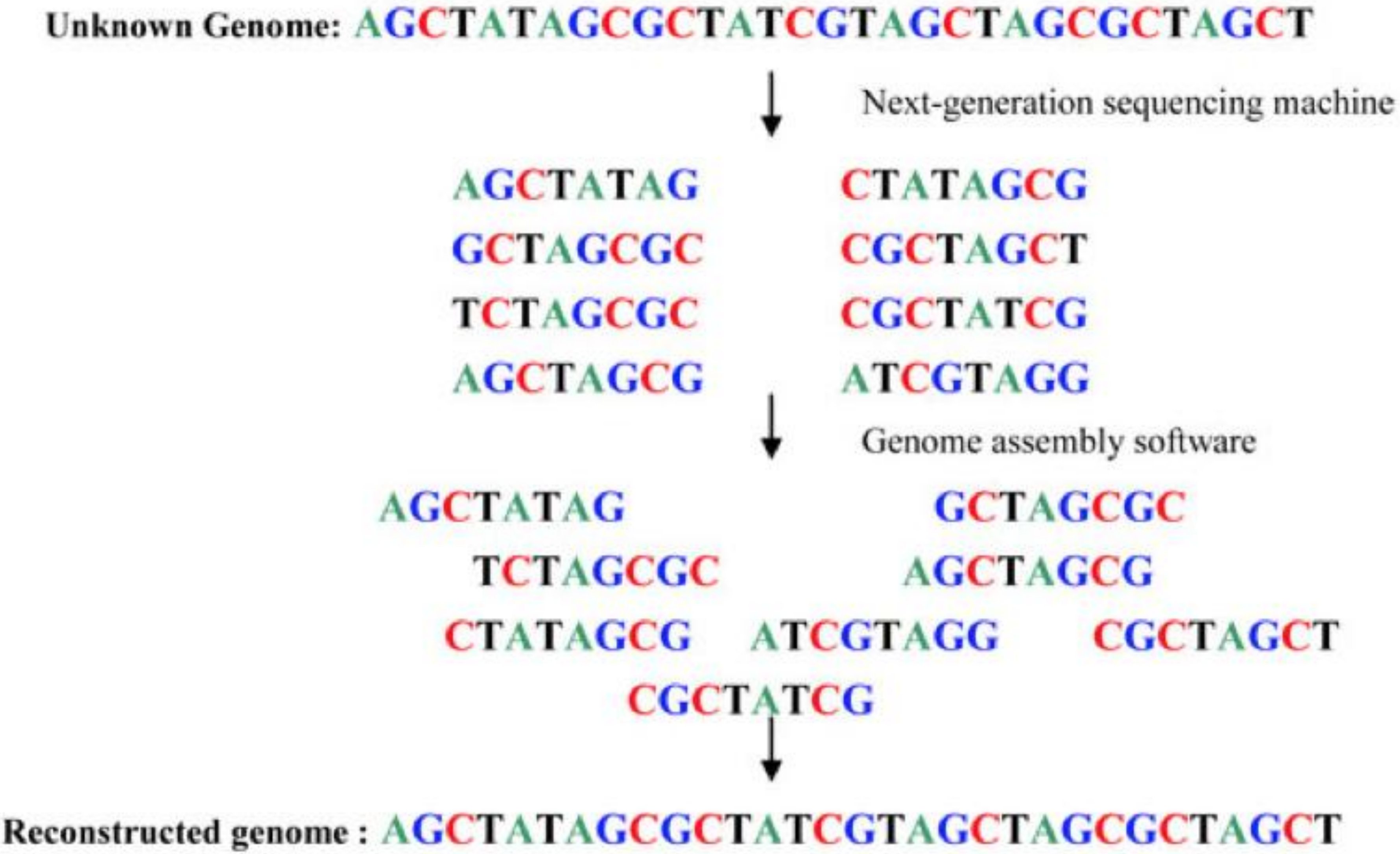
# Cost per Genome



National Human Genome Research Institute

genome.gov/sequencingcosts

## Next Generation Sequencing (NGS)



**Unknown Genome:** AGCTATAGCGCTATCGTAGCTAGCGCTAGCT

Next-generation sequencing machine

AGCTATAG CTATAGCG

GCTAGCGC CGCTAGCT

TCTAGCGC CGCTATCG

AGCTAGCG ATCGTAGG

Genome assembly software

AGCTATAG GCTAGCGC

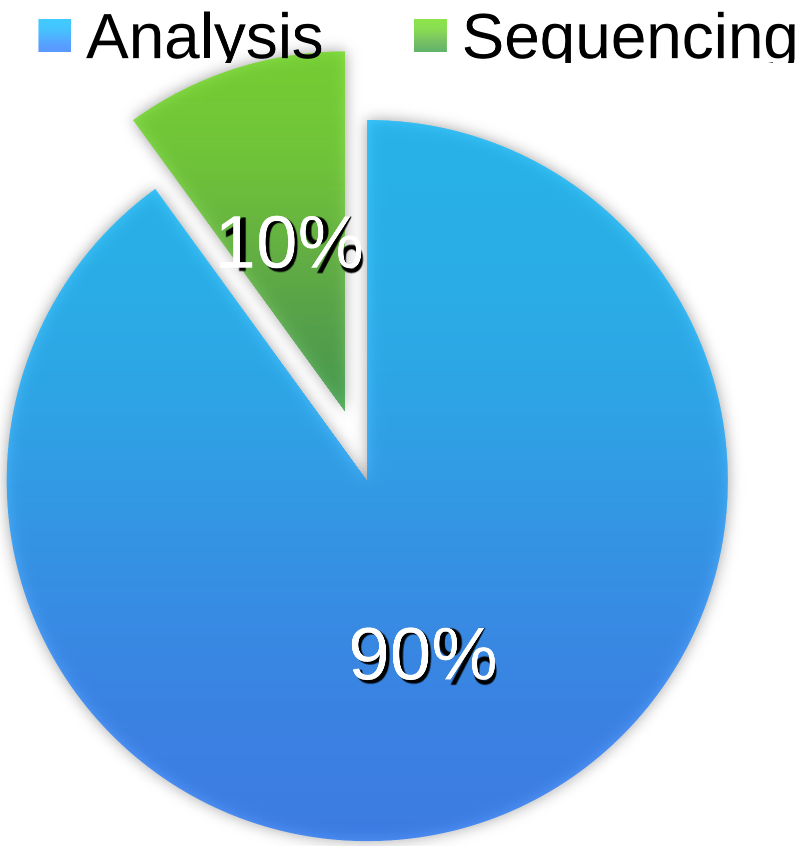
TCTAGCGC AGCTAGCG

CTATAGCG ATCGTAGG CGCTAGCT

CGCTATCG

**Reconstructed genome :** AGCTATAGCGCTATCGTAGCTAGCGCTAGCT

# *COSTS*




PHASE TWO: INTERPRETATION
I THINK I FOUND A CORNER PIECE.
3 BILLION PIECES
GENOME
SHENEMAN The Star-Ledger

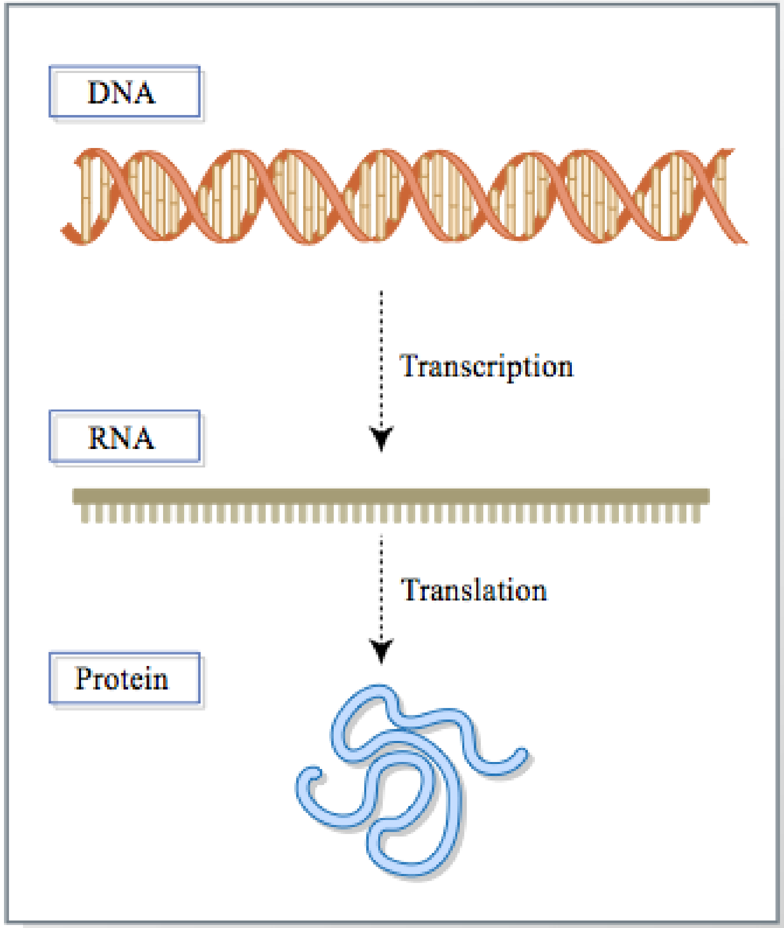
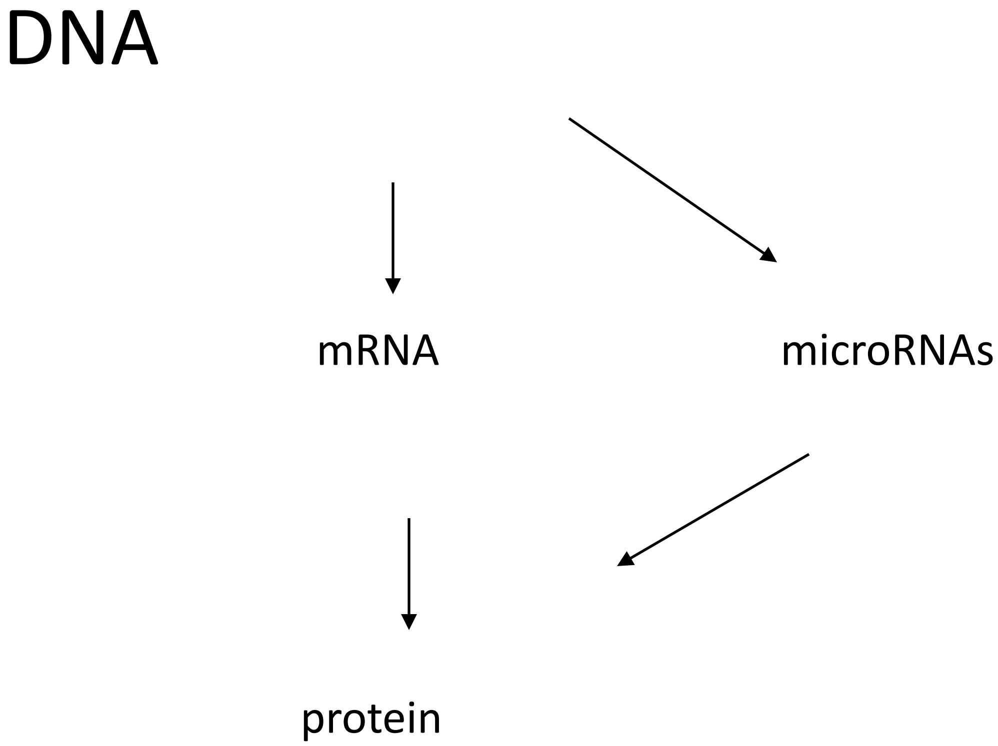


Figure by MIT OCW.

# *Central Dogma changes #1*

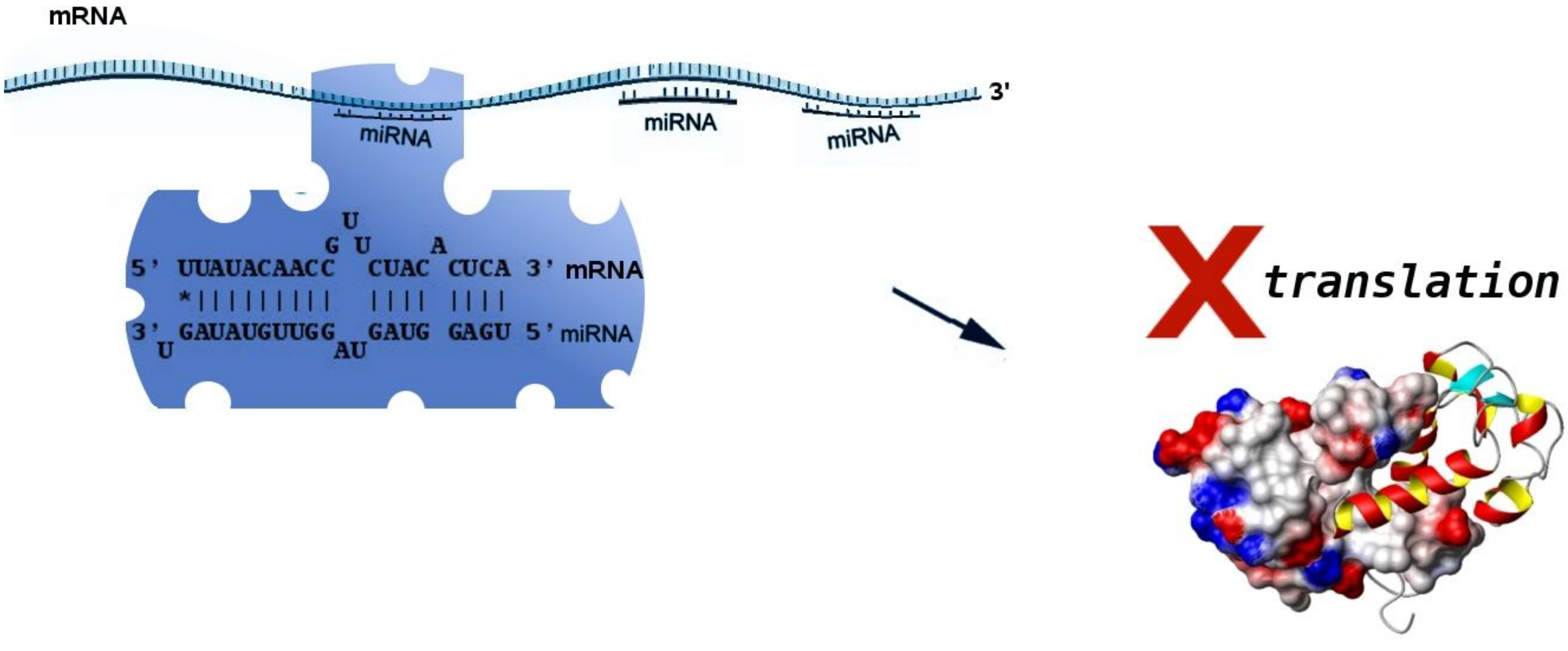


DNA

mRNA

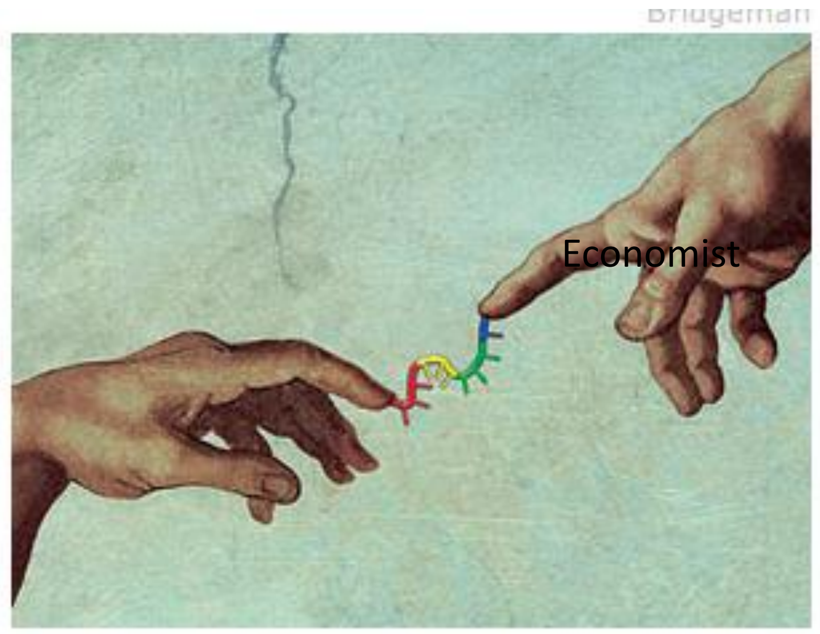
microRNAs

protein



microRNAs bind the 3'UTR of mRNAs and repress translation

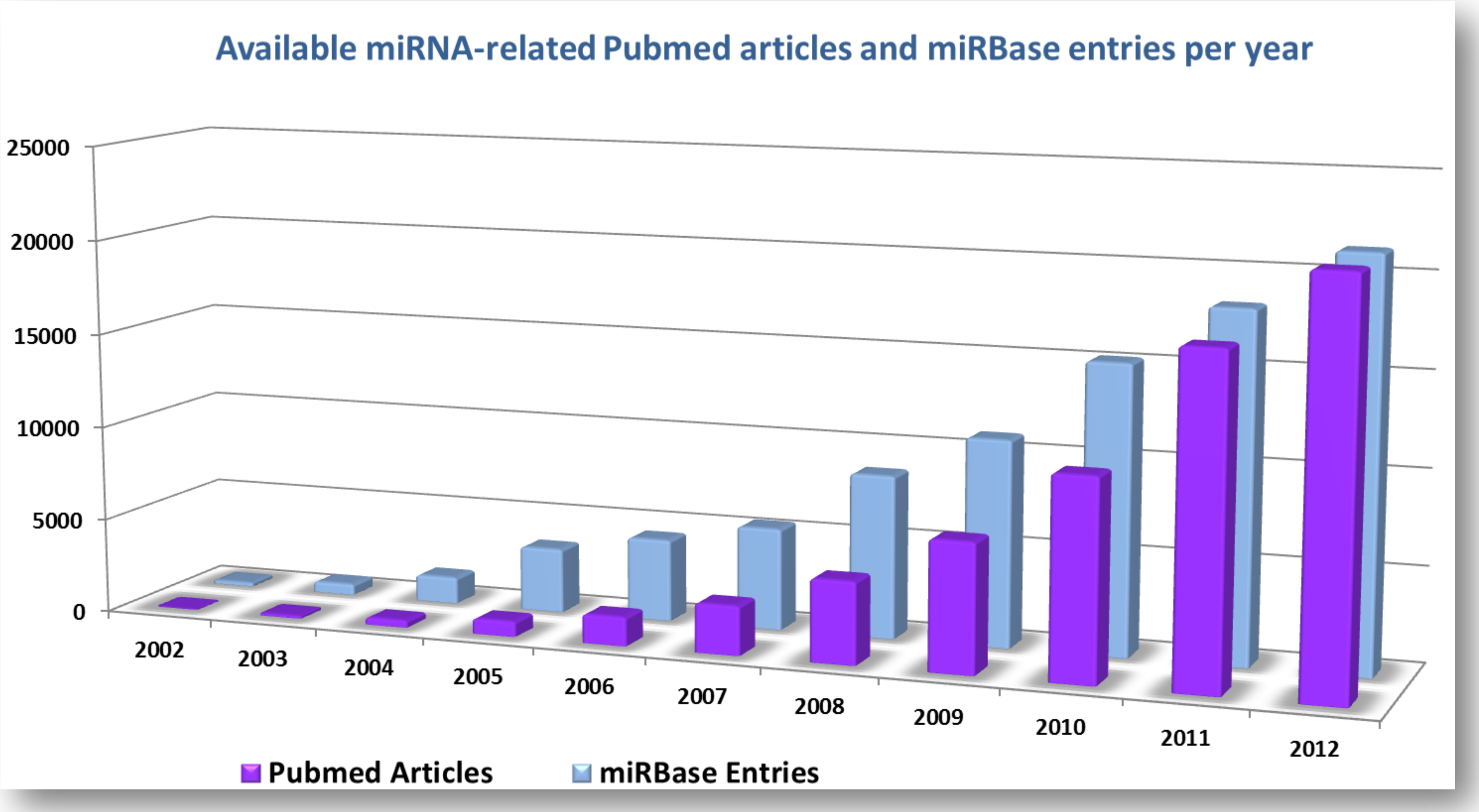
# The RNA revolution: Biology's Big Bang

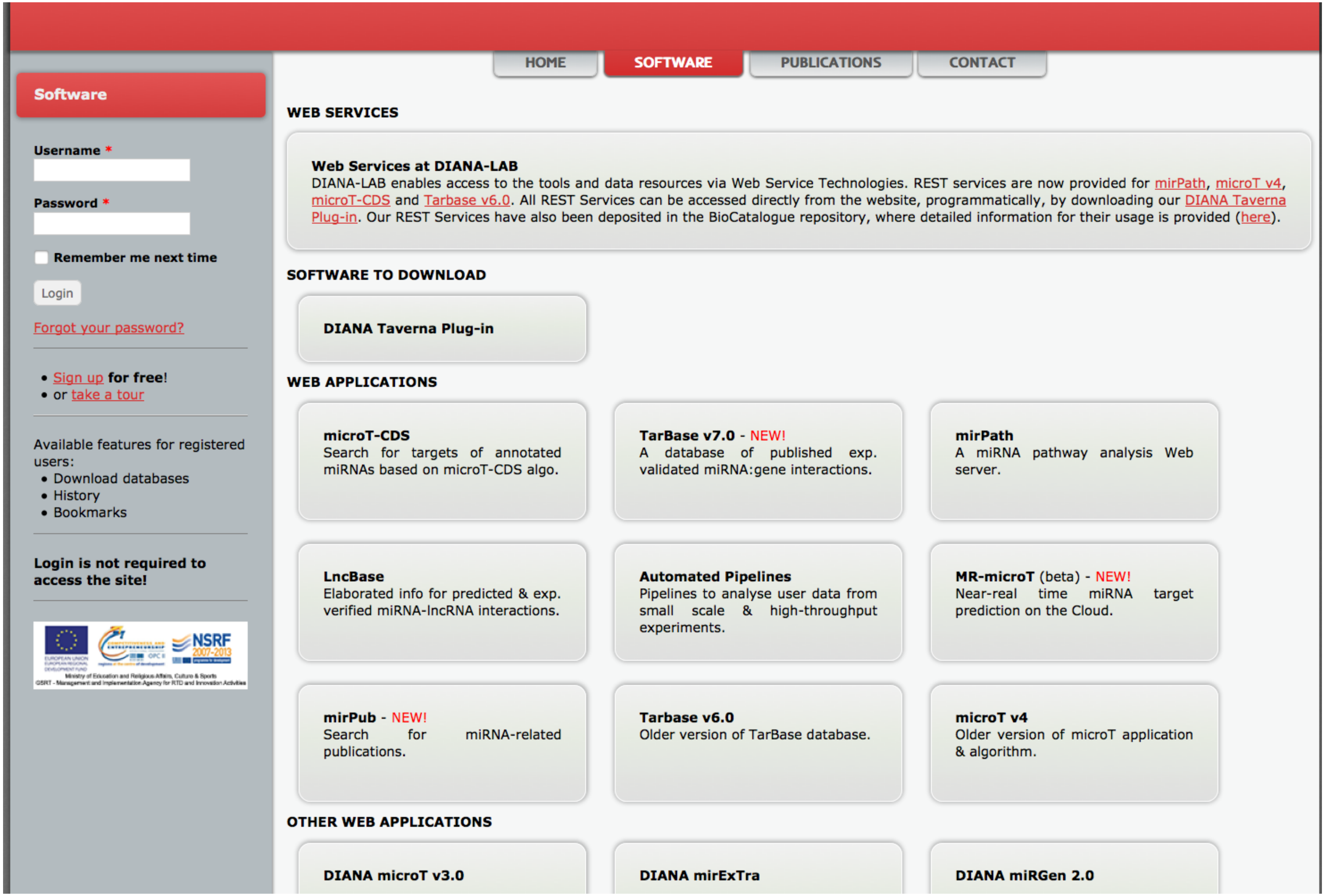


Jun 14th 2007 From The Economist :

What physics was to the 20th century, biology will be to the 21st —and RNA will be a vital part of it

# Superlinear Increase of known miRNAs and relevant Research





**Software**

**Username** *

**Password** *

☐ **Remember me next time**

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!**

WEB SERVICES

**Web Services at DIANA-LAB**

DIANA-LAB enables access to the tools and data resources via Web Service Technologies. REST services are now provided for mirPath, microT v4, microT-CDS and Tarbase v6.0. All REST Services can be accessed directly from the website, programmatically, by downloading our DIANA Taverna Plug-in. Our REST Services have also been deposited in the BioCatalogue repository, where detailed information for their usage is provided (here).

SOFTWARE TO DOWNLOAD

**DIANA Taverna Plug-in**

WEB APPLICATIONS

**microT-CDS**
Search for targets of annotated miRNAs based on microT-CDS algo.

**TarBase v7.0** - NEW!
A database of published exp. validated miRNA:gene interactions.

**mirPath**
A miRNA pathway analysis Web server.

**LncBase**
Elaborated info for predicted & exp. verified miRNA-lncRNA interactions.

**Automated Pipelines**
Pipelines to analyse user data from small scale & high-throughput experiments.

**MR-microT** (beta) - NEW!
Near-real time miRNA target prediction on the Cloud.

**mirPub** - NEW!
Search for miRNA-related publications.

**Tarbase v6.0**
Older version of TarBase database.

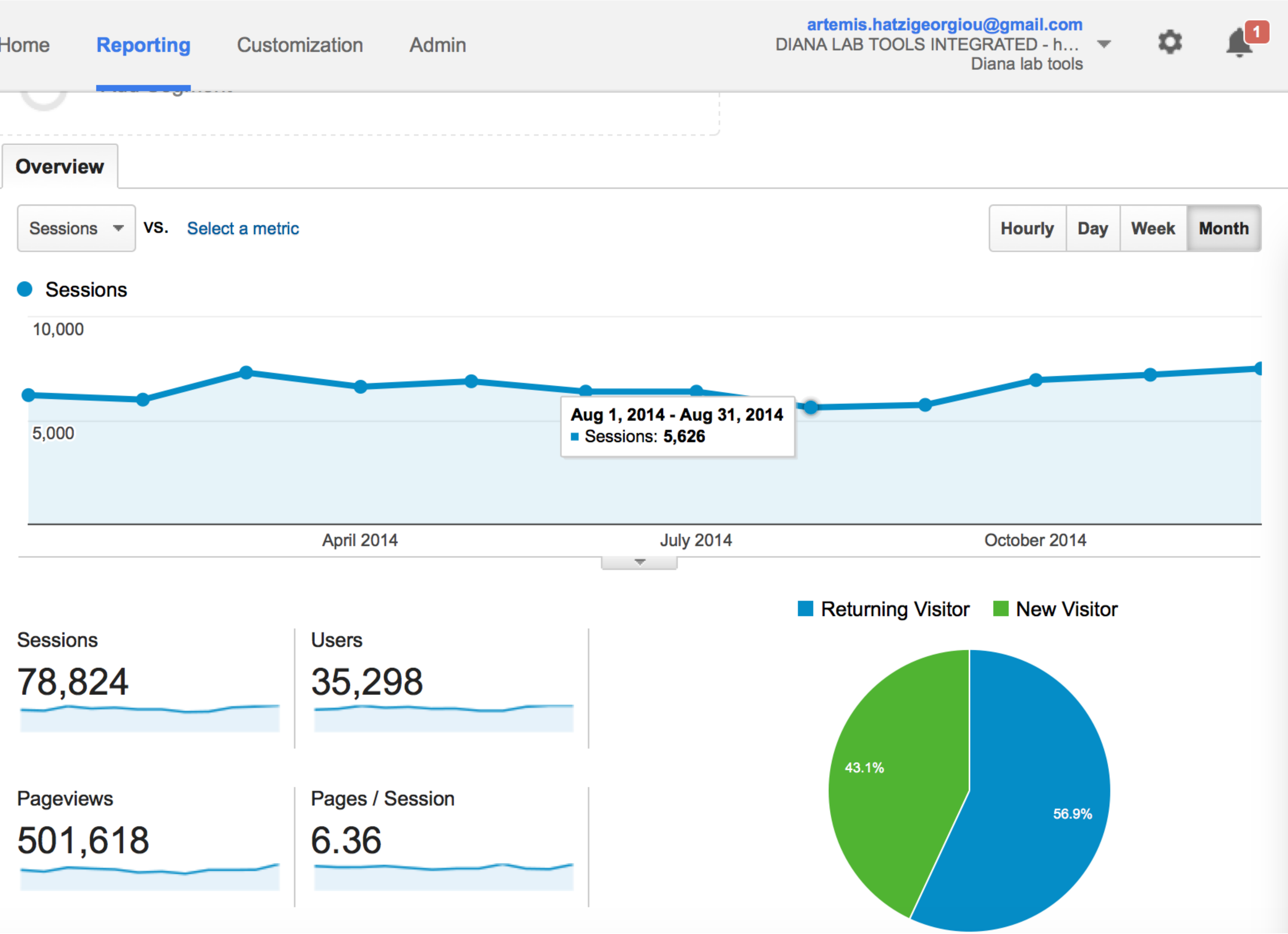
**microT v4**
Older version of microT application & algorithm.

OTHER WEB APPLICATIONS

**DIANA microT v3.0**

**DIANA mirExTra**

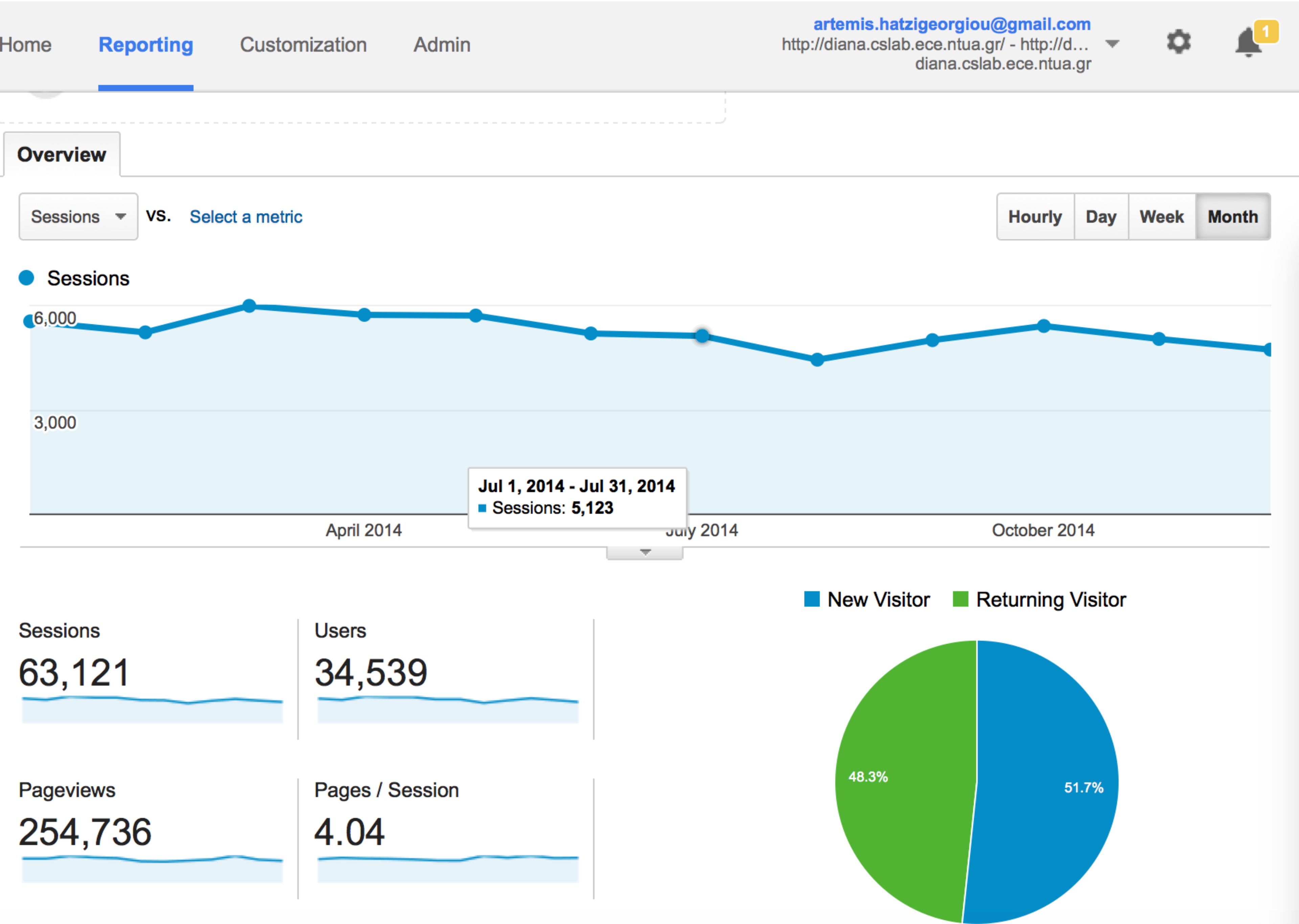
**DIANA miRGen 2.0**



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

| Sessions | Users |
|---|---|
| 78,824 | 35,298 |

| Pageviews | Pages / Session |
|---|---|
| 501,618 | 6.36 |



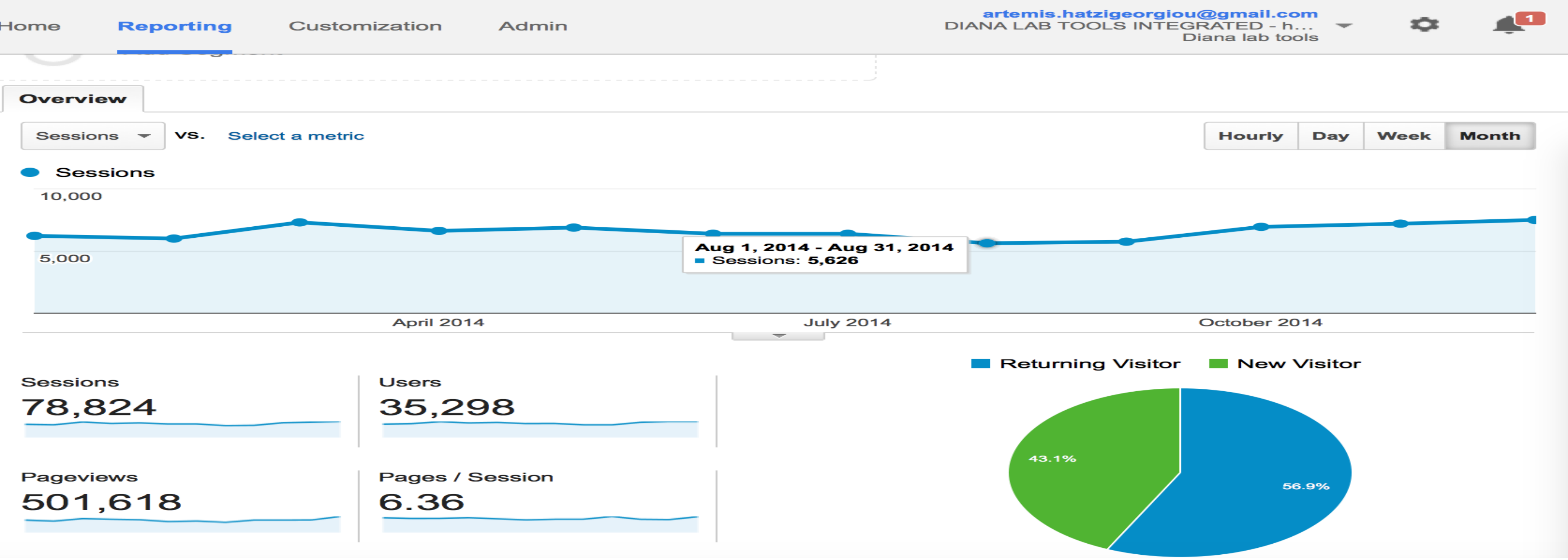
Jul 1, 2014 - Jul 31, 2014
Sessions: **5,123**

Sessions
63,121

Users
34,539

Pageviews
254,736

Pages / Session
4.04



Home Reporting Customization Admin

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

**Overview**

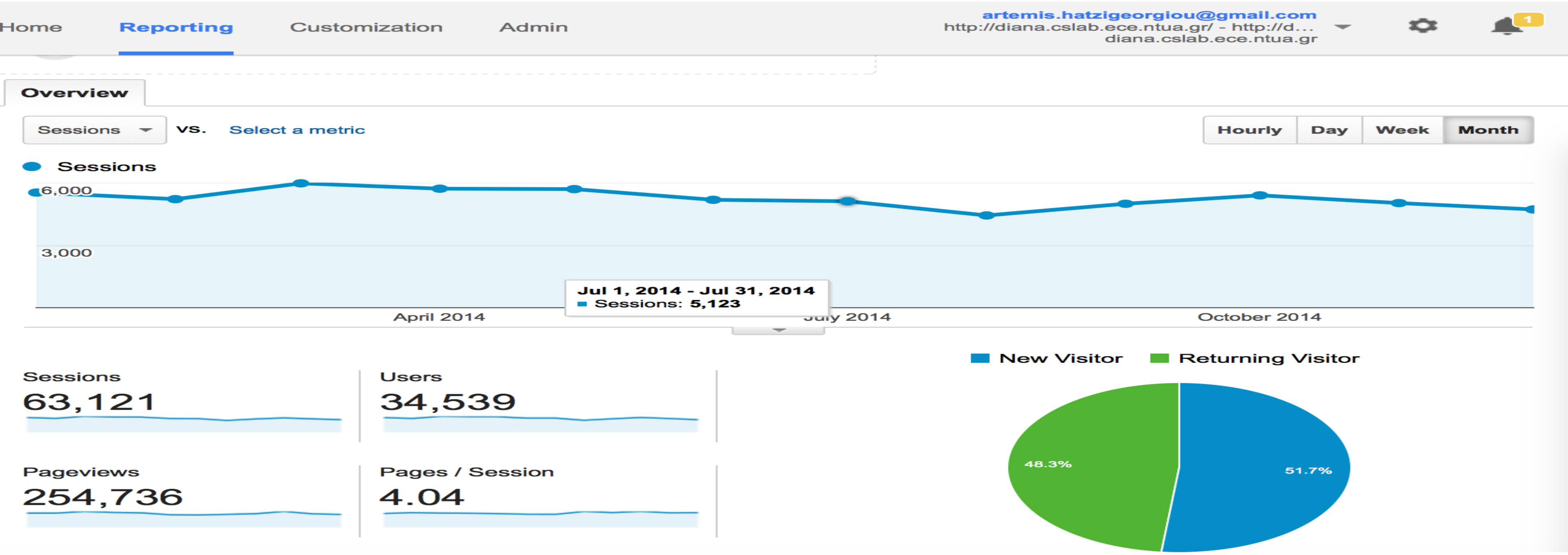
Sessions vs. Select a metric

Hourly Day Week **Month**

● Sessions

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

| | |
|---|---|
| Sessions<br>78,824 | Users<br>35,298 |
| Pageviews<br>501,618 | Pages / Session<br>6.36 |



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

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

| | |
|---|---|
| Sessions<br>63,121 | Users<br>34,539 |
| Pageviews<br>254,736 | Pages / Session<br>4.04 |

# *How much are we cited each year ?*

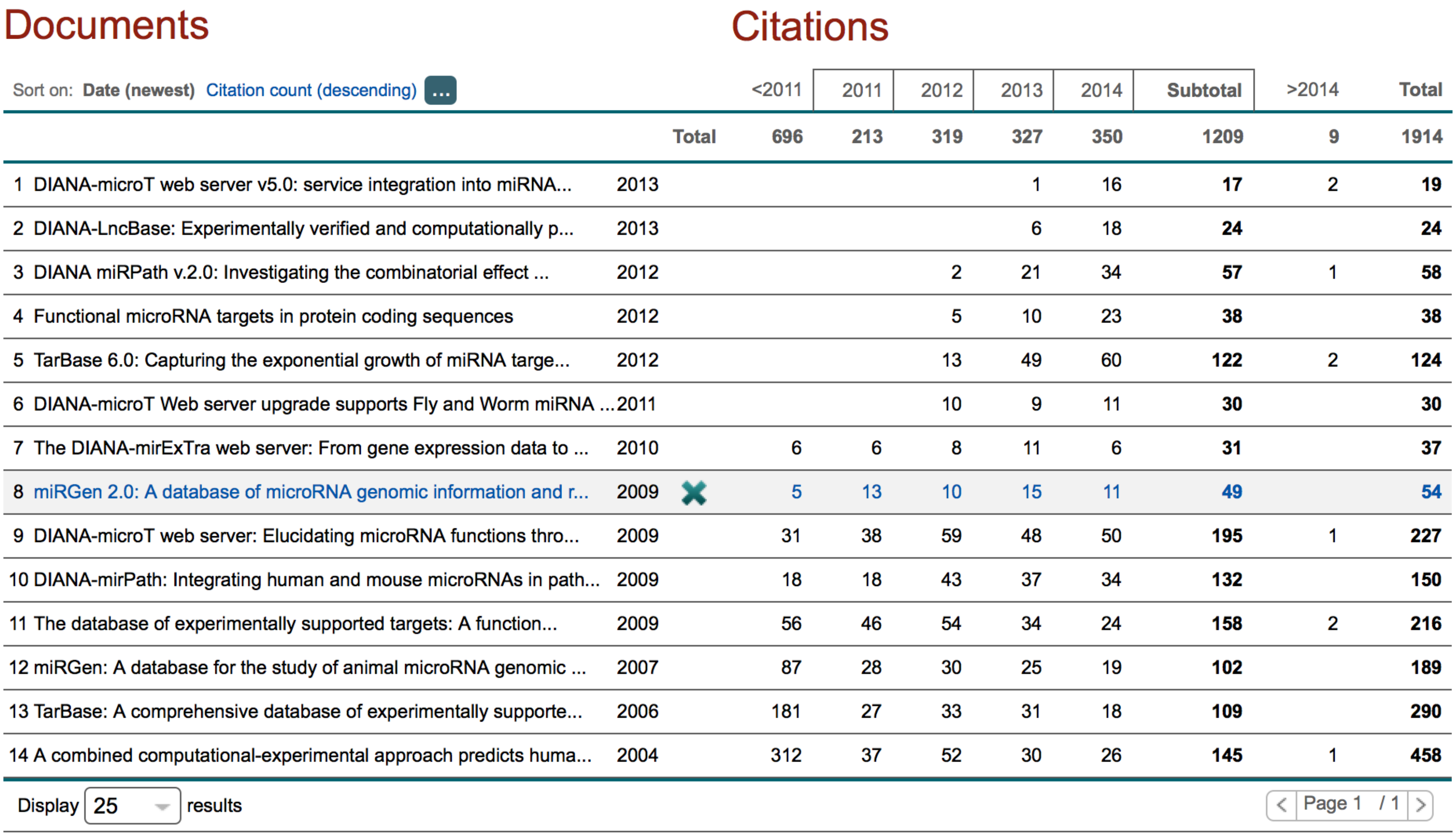
Around 40,000 users visited

**Microrna.gr**

in 2013

How much are the relevant papers cited in 2014 ?

a) 40 times
b) 350 times
c) 2000 times



## Documents

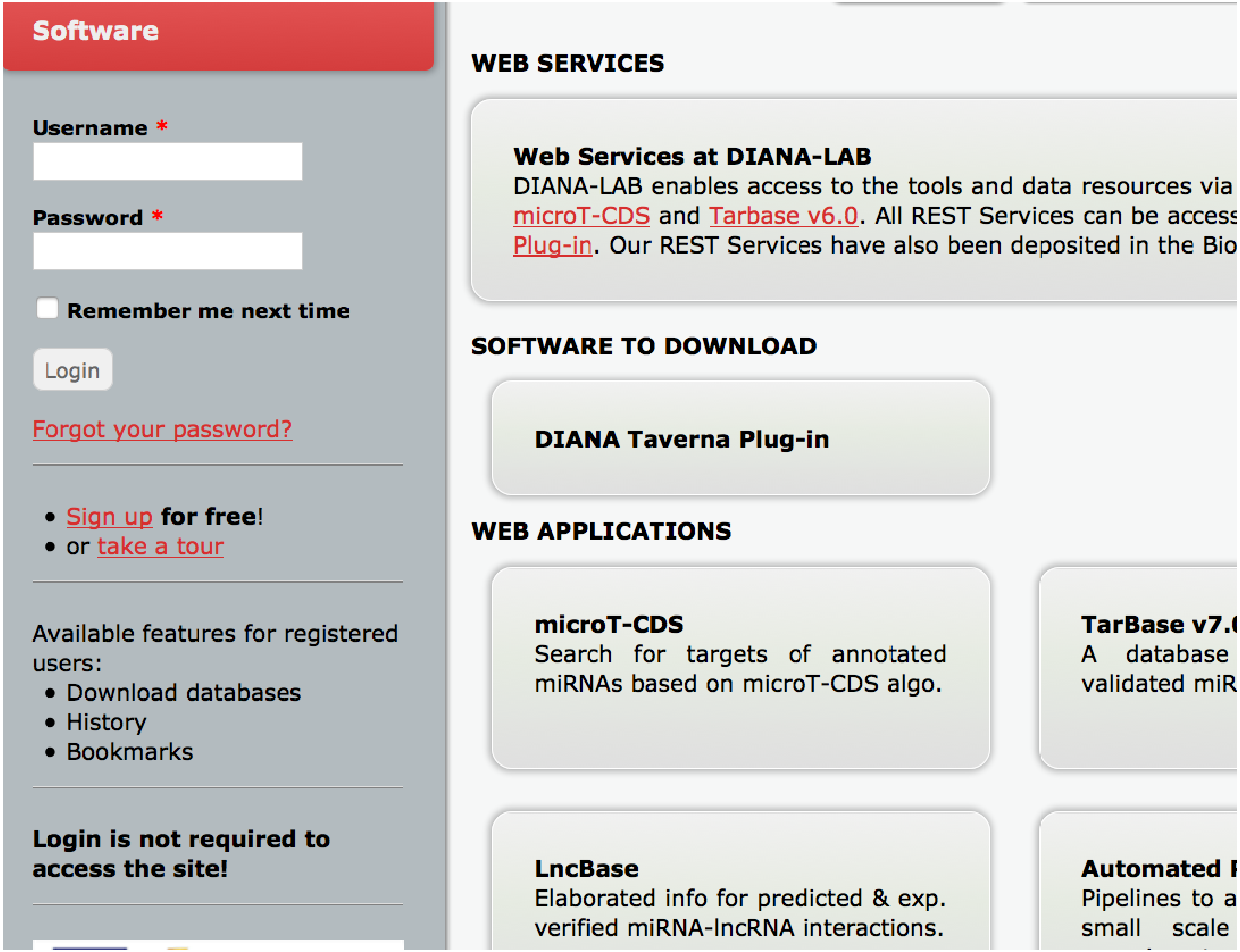
## Citations

Sort on: **Date (newest)** Citation count (descending) ...

| | | | | <2011 | 2011 | 2012 | 2013 | 2014 | Subtotal | >2014 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Total** | **696** | **213** | **319** | **327** | **350** | **1209** | **9** | **1914** |
| 1 | DIANA-microT web server v5.0: service integration into miRNA... | 2013 | | | | | 1 | 16 | **17** | 2 | **19** |
| 2 | DIANA-LncBase: Experimentally verified and computationally p... | 2013 | | | | | 6 | 18 | **24** | | **24** |
| 3 | DIANA miRPath v.2.0: Investigating the combinatorial effect ... | 2012 | | | | 2 | 21 | 34 | **57** | 1 | **58** |
| 4 | Functional microRNA targets in protein coding sequences | 2012 | | | | 5 | 10 | 23 | **38** | | **38** |
| 5 | TarBase 6.0: Capturing the exponential growth of miRNA targe... | 2012 | | | | 13 | 49 | 60 | **122** | 2 | **124** |
| 6 | DIANA-microT Web server upgrade supports Fly and Worm miRNA ... | 2011 | | | | 10 | 9 | 11 | **30** | | **30** |
| 7 | The DIANA-mirExTra web server: From gene expression data to ... | 2010 | | 6 | 6 | 8 | 11 | 6 | **31** | | **37** |
| 8 | miRGen 2.0: A database of microRNA genomic information and r... | 2009 | ✖ | 5 | 13 | 10 | 15 | 11 | **49** | | **54** |
| 9 | DIANA-microT web server: Elucidating microRNA functions thro... | 2009 | | 31 | 38 | 59 | 48 | 50 | **195** | 1 | **227** |
| 10 | DIANA-mirPath: Integrating human and mouse microRNAs in path... | 2009 | | 18 | 18 | 43 | 37 | 34 | **132** | | **150** |
| 11 | The database of experimentally supported targets: A function... | 2009 | | 56 | 46 | 54 | 34 | 24 | **158** | 2 | **216** |
| 12 | miRGen: A database for the study of animal microRNA genomic ... | 2007 | | 87 | 28 | 30 | 25 | 19 | **102** | | **189** |
| 13 | TarBase: A comprehensive database of experimentally supporte... | 2006 | | 181 | 27 | 33 | 31 | 18 | **109** | | **290** |
| 14 | A combined computational-experimental approach predicts huma... | 2004 | | 312 | 37 | 52 | 30 | 26 | **145** | 1 | **458** |

Display 25 results

Page 1 / 1



**Username** *

**Password** *

☐ **Remember me next time**

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!**

**WEB SERVICES**

**Web Services at DIANA-LAB**
DIANA-LAB enables access to the tools and data resources via
microT-CDS and Tarbase v6.0. All REST Services can be access
Plug-in. Our REST Services have also been deposited in the Bio

**SOFTWARE TO DOWNLOAD**

**DIANA Taverna Plug-in**

**WEB APPLICATIONS**

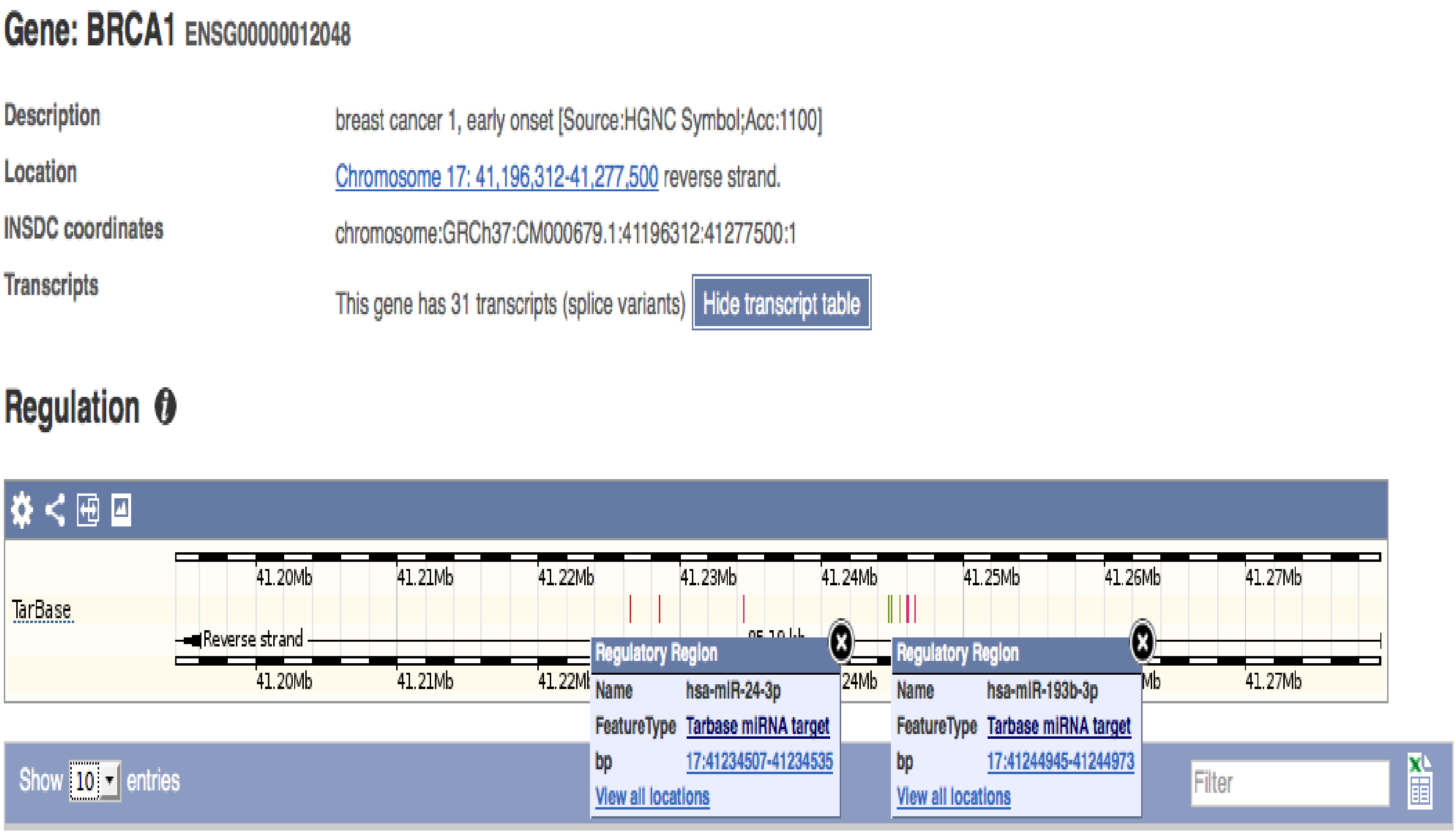
**microT-CDS**
Search for targets of annotated miRNAs based on microT-CDS algo.

**TarBase v7.**
A database
validated miR

**LncBase**
Elaborated info for predicted & exp. verified miRNA-lncRNA interactions.

**Automated**
Pipelines to a
small scale

# *Integration of Tarbase in **Ensembl***



## Gene: BRCA1 ENSG0000012048

| | |
|---|---|
| Description | breast cancer 1, early onset [Source:HGNC Symbol;Acc:1100] |
| Location | Chromosome 17: 41,196,312-41,277,500 reverse strand. |
| INSDC coordinates | chromosome:GRCh37:CM000679.1:41196312:41277500:1 |
| Transcripts | This gene has 31 transcripts (splice variants) Hide transcript table |

## Regulation

**Regulatory Region**

| | |
|---|---|
| Name | hsa-miR-24-3p |
| FeatureType | Tarbase miRNA target |
| bp | 17:41234507-41234535 |

View all locations

**Regulatory Region**

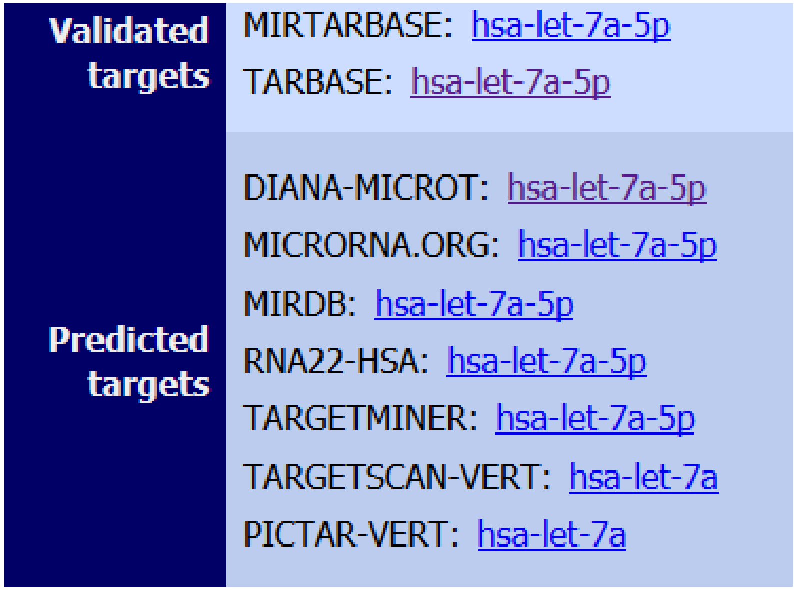
| | |
|---|---|
| Name | hsa-miR-193b-3p |
| FeatureType | Tarbase miRNA target |
| bp | 17:41244945-41244973 |

View all locations

Show 10 entries

# miRBase

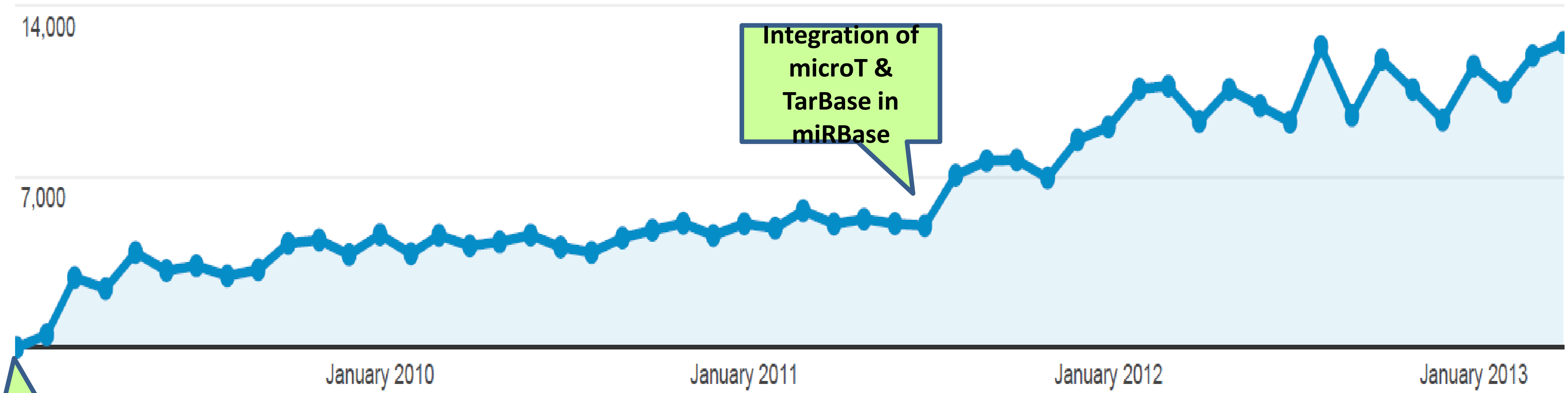
- Interconnects also entries with external resources:



| | |
|---|---|
| **Validated targets** | MIRTARBASE: hsa-let-7a-5p |
| | TARBASE: hsa-let-7a-5p |
| **Predicted targets** | DIANA-MICROT: hsa-let-7a-5p |
| | MICRORNA.ORG: hsa-let-7a-5p |
| | MIRDB: hsa-let-7a-5p |
| | RNA22-HSA: hsa-let-7a-5p |
| | TARGETMINER: hsa-let-7a-5p |
| | TARGETSCAN-VERT: hsa-let-7a |
| | PICTAR-VERT: hsa-let-7a |

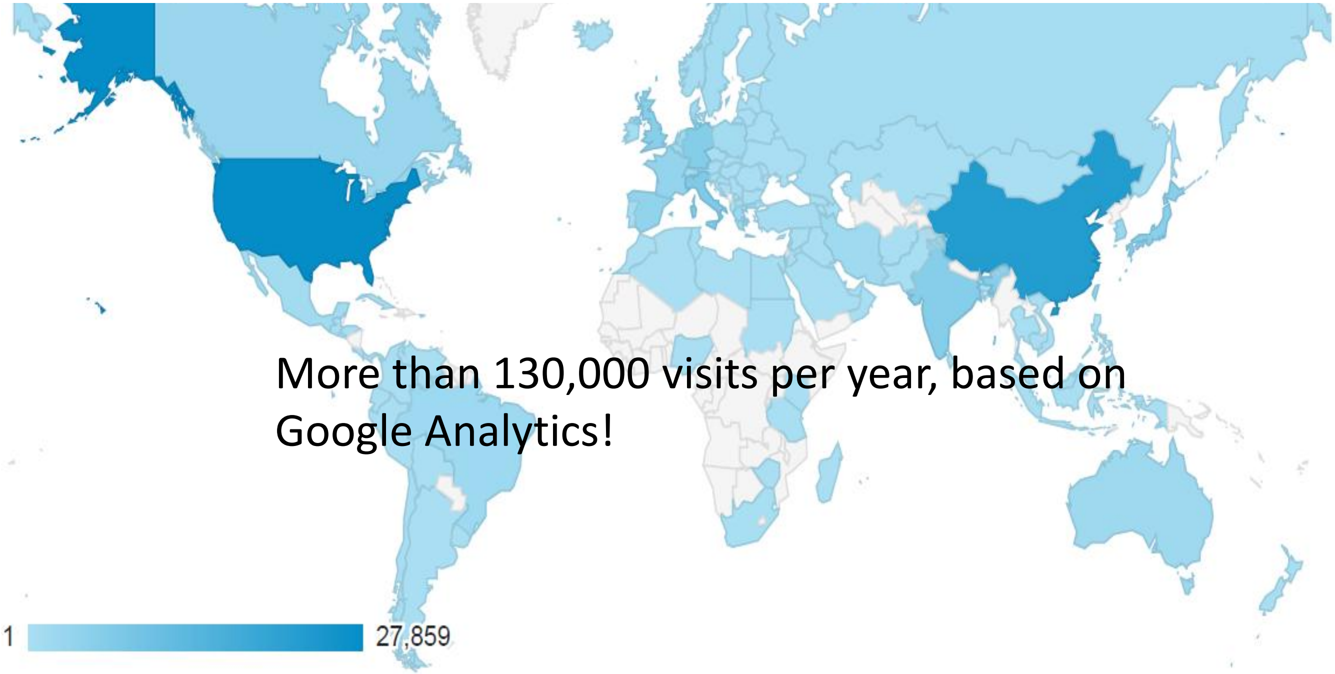
# ***DIANA-Tools***

Visit us @
***www.microrna.gr***!



**Integration of microT & TarBase in miRBase**

**First release**

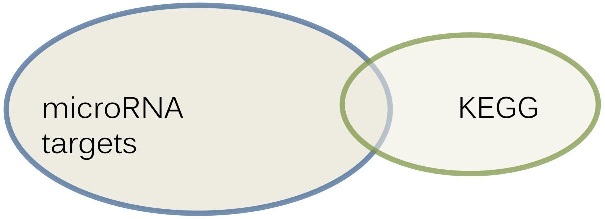


More than 130,000 visits per year, based on Google Analytics!

www.microrna.gr/mirpath/

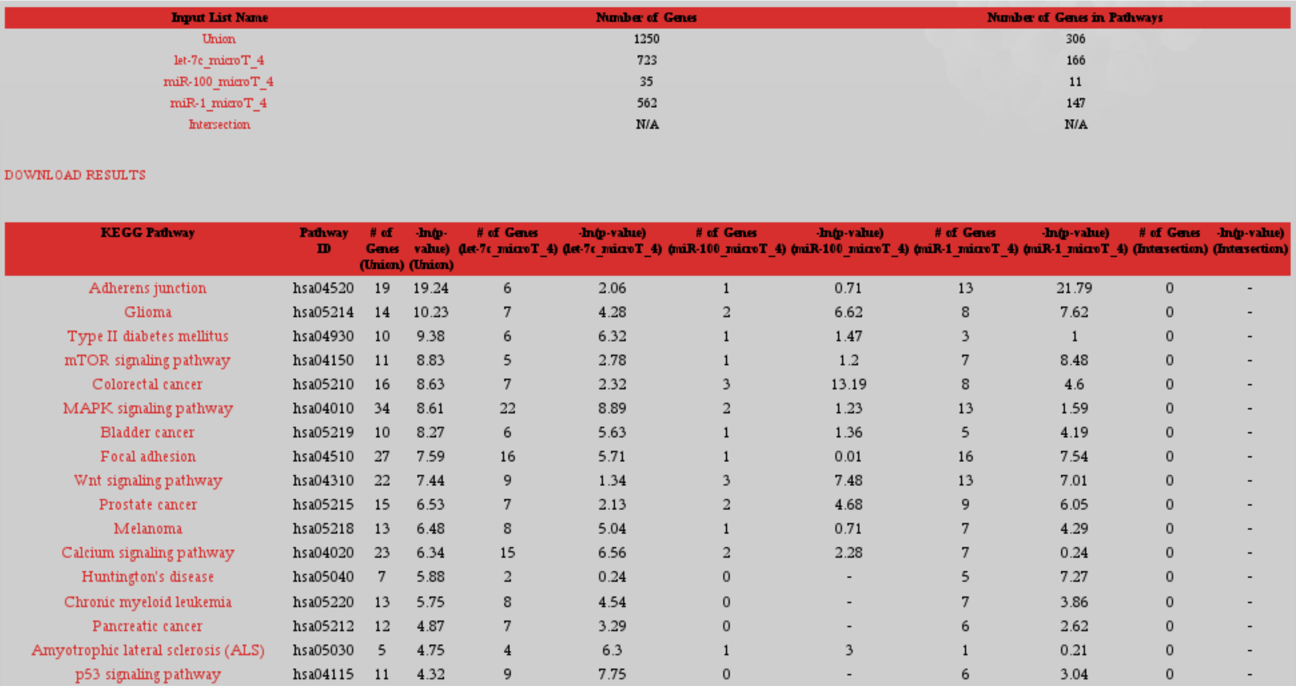
# DIANA-mirPath

Integrating human and mouse microRNAs in pathways



Small overlap – Not significant

Large overlap – Significant

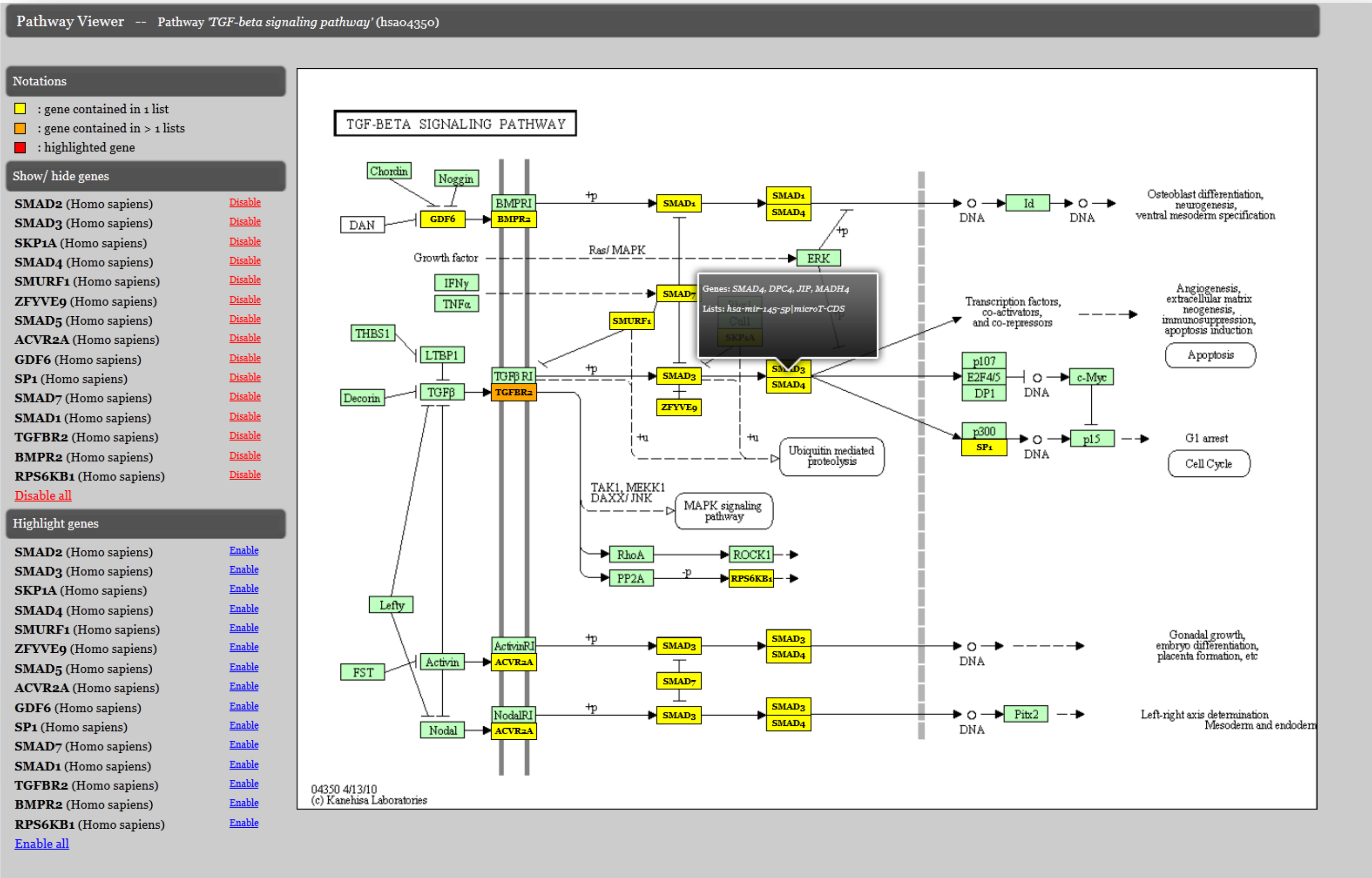


| Input List Name | Number of Genes | Number of Genes in Pathways |
|---|---|---|
| Union | 1250 | 306 |
| let-7c_microT_4 | 723 | 166 |
| miR-100_microT_4 | 35 | 11 |
| miR-1_microT_4 | 562 | 147 |
| Intersection | N/A | N/A |

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 | - |

# Active Pathway Visualization



**Pathway Viewer** -- Pathway *'TGF-beta signaling pathway'* (hsa04350)

**Notations**

🟨 : gene contained in 1 list
🟧 : gene contained in > 1 lists
🟥 : highlighted gene

**Show/ hide genes**

| | |
|---|---|
| **SMAD2** (Homo sapiens) | Disable |
| **SMAD3** (Homo sapiens) | Disable |
| **SKP1A** (Homo sapiens) | Disable |
| **SMAD4** (Homo sapiens) | Disable |
| **SMURF1** (Homo sapiens) | Disable |
| **ZFYVE9** (Homo sapiens) | Disable |
| **SMAD5** (Homo sapiens) | Disable |
| **ACVR2A** (Homo sapiens) | Disable |
| **GDF6** (Homo sapiens) | Disable |
| **SP1** (Homo sapiens) | Disable |
| **SMAD7** (Homo sapiens) | Disable |
| **SMAD1** (Homo sapiens) | Disable |
| **TGFBR2** (Homo sapiens) | Disable |
| **BMPR2** (Homo sapiens) | Disable |
| **RPS6KB1** (Homo sapiens) | Disable |

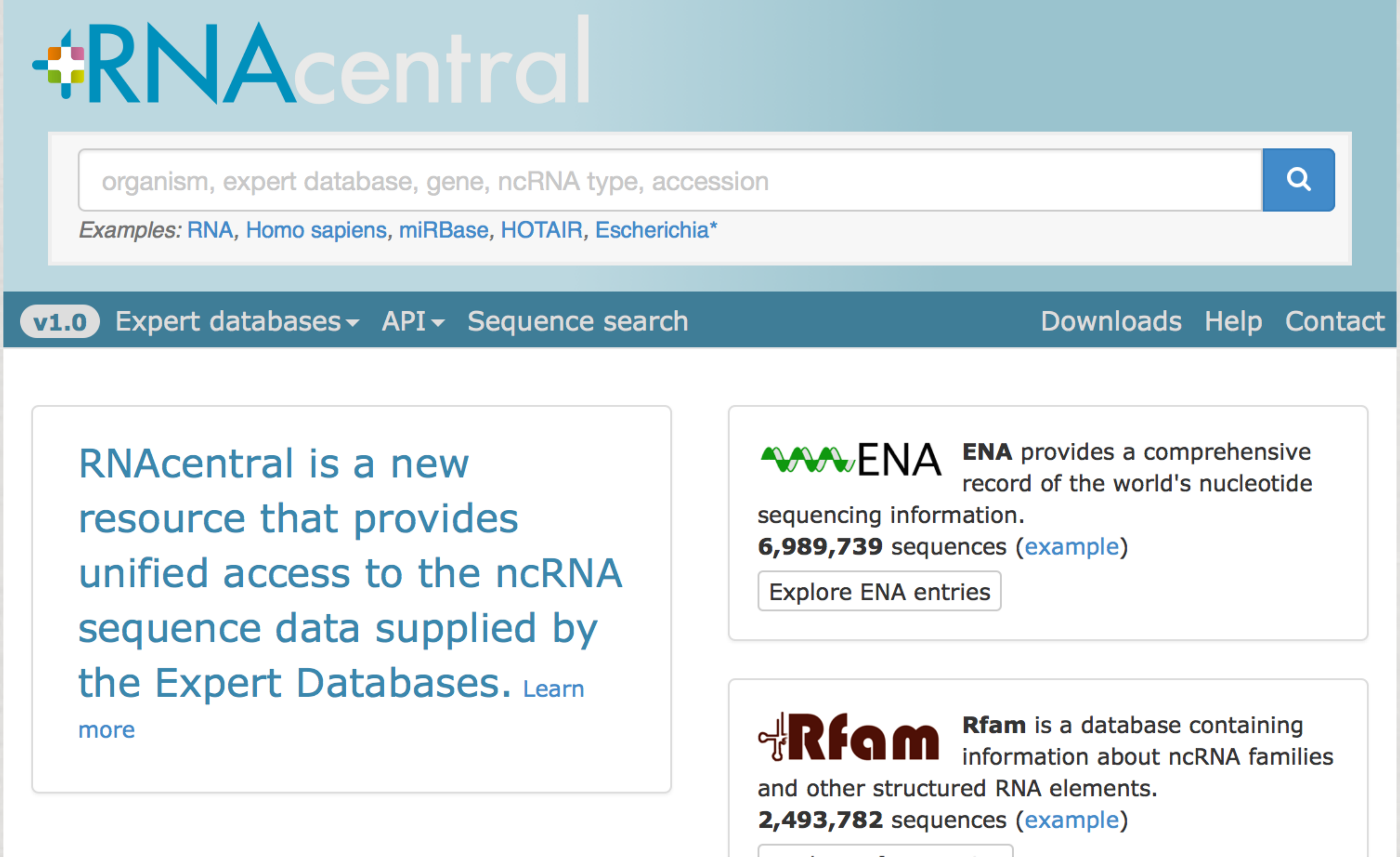
Disable all

**Highlight genes**

| | |
|---|---|
| **SMAD2** (Homo sapiens) | Enable |
| **SMAD3** (Homo sapiens) | Enable |
| **SKP1A** (Homo sapiens) | Enable |
| **SMAD4** (Homo sapiens) | Enable |
| **SMURF1** (Homo sapiens) | Enable |
| **ZFYVE9** (Homo sapiens) | Enable |
| **SMAD5** (Homo sapiens) | Enable |
| **ACVR2A** (Homo sapiens) | Enable |
| **GDF6** (Homo sapiens) | Enable |
| **SP1** (Homo sapiens) | Enable |
| **SMAD7** (Homo sapiens) | Enable |
| **SMAD1** (Homo sapiens) | Enable |
| **TGFBR2** (Homo sapiens) | Enable |
| **BMPR2** (Homo sapiens) | Enable |
| **RPS6KB1** (Homo sapiens) | Enable |

Enable all

Genes: *SMAD4, DPC4, JIP, MADH4*
Lists: *hsa-mir-145-5p|microT-CDS*

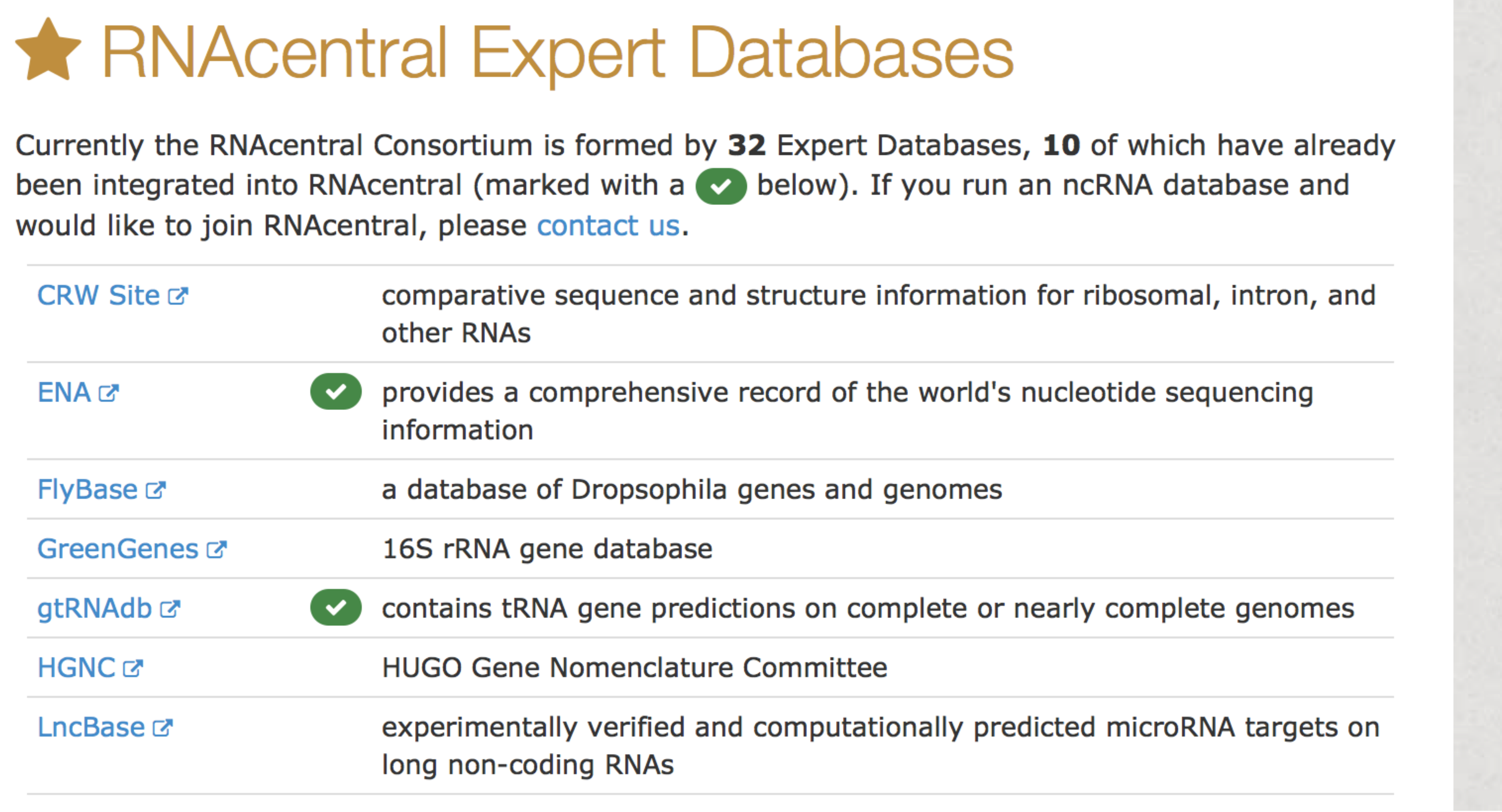


*Examples:* RNA, Homo sapiens, miRBase, HOTAIR, Escherichia*

RNAcentral is a new resource that provides unified access to the ncRNA sequence data supplied by the Expert Databases. Learn more

**ENA** provides a comprehensive record of the world's nucleotide sequencing information.
**6,989,739** sequences (example)

**Rfam** is a database containing information about ncRNA families and other structured RNA elements.
**2,493,782** sequences (example)



# 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 other RNAs |
| ENA | ✓ | provides a comprehensive record of the world's nucleotide sequencing information |
| FlyBase | | a database of Dropsophila genes and genomes |
| GreenGenes | | 16S rRNA gene database |
| gtRNAdb | ✓ | contains tRNA gene predictions on complete or nearly complete genomes |
| HGNC | | HUGO Gene Nomenclature Committee |
| LncBase | | experimentally verified and computationally predicted microRNA targets on long non-coding RNAs |

# *Few … wishes*

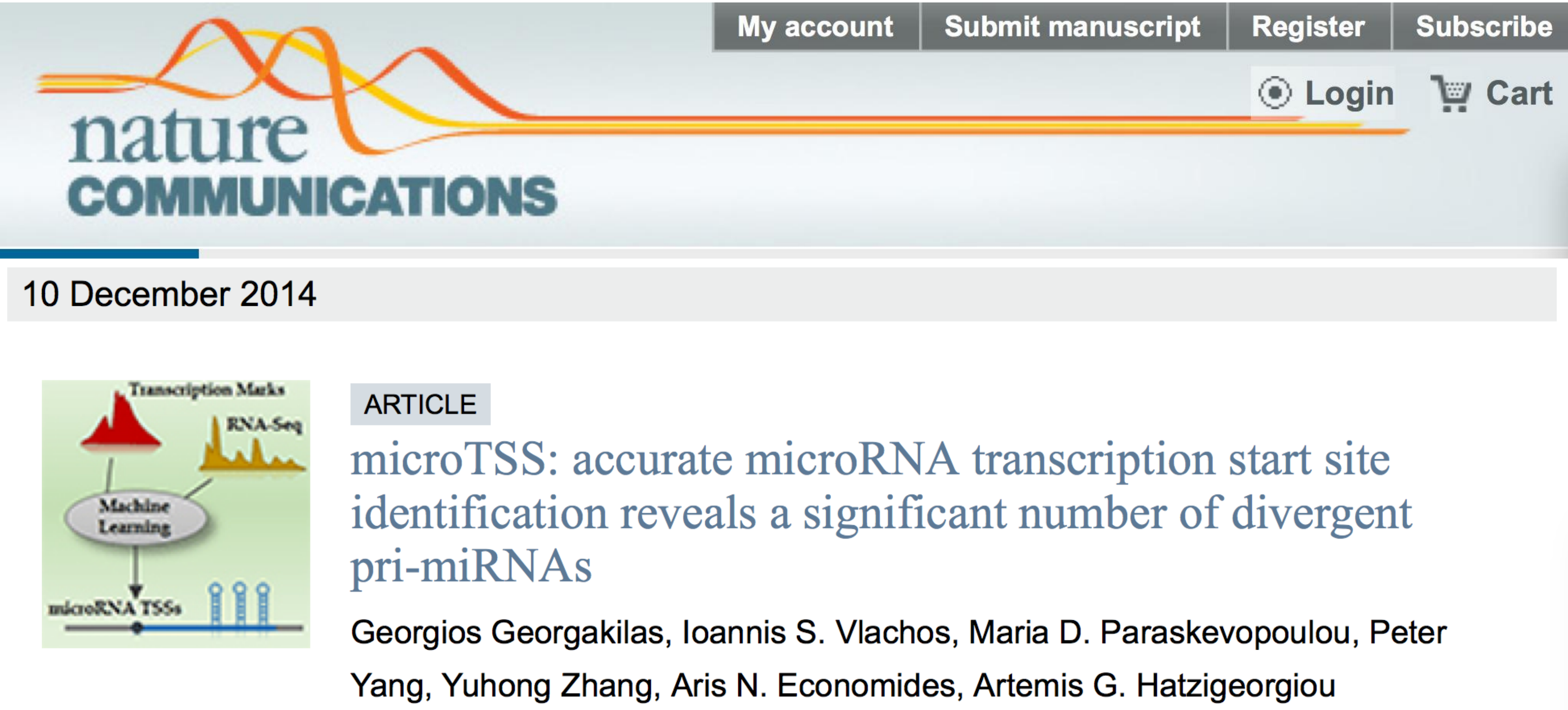
Check the citations of databases / webservers before publishing
For example could be a question added to reviewers :
Have the researcher cited properly the data used ?

Are the data used for training – testing available ?
Can the data be reproduced ?

Availability of databases through time – diachronic data
Credibility for diachronic databases/web services
Let's make database funding … legal

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

***Funding:*** *Project "TOM" that is implemented under the "ARISTEIA" Action of the "OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING" and is co-funded by the European Social Fund (ESF) and National Resources.*



nature COMMUNICATIONS

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