



# iMir: An innovative and complete pipeline for smallRNA-Seq data analysis

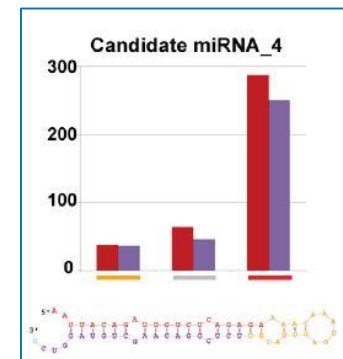
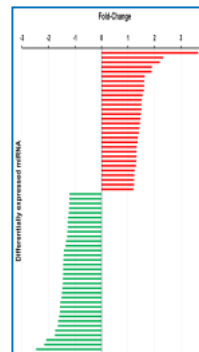
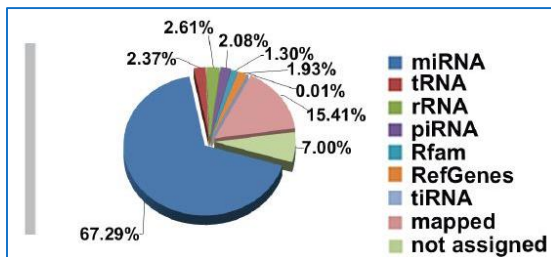
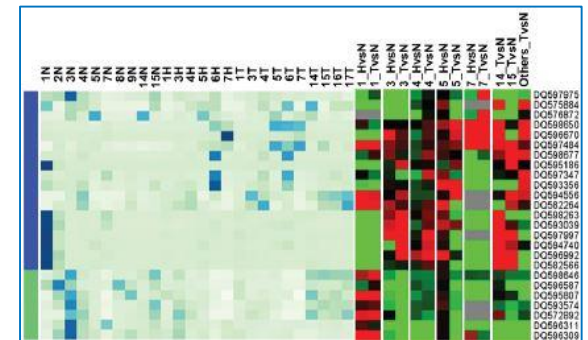
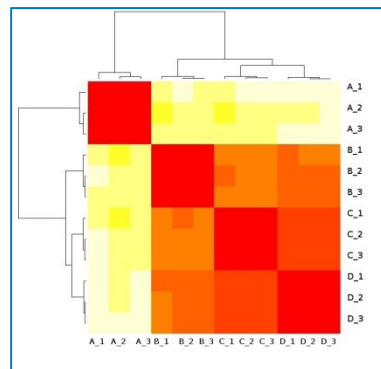
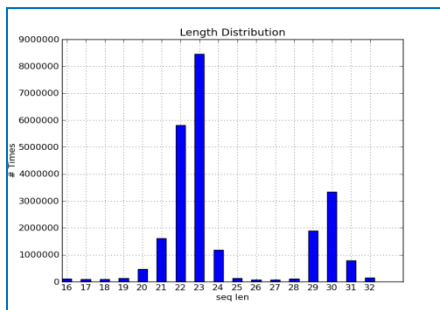
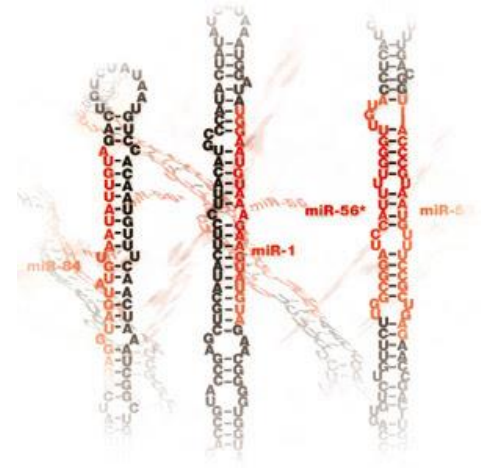
Giorgio Giurato, PhD  
ggiurato@unisa.it

Laboratory of Molecular Medicine and Genomics  
Department of Medicine and Surgery  
University of Salerno

[www.labmedmolge.unisa.it](http://www.labmedmolge.unisa.it) / [www.genomix4life.com](http://www.genomix4life.com)

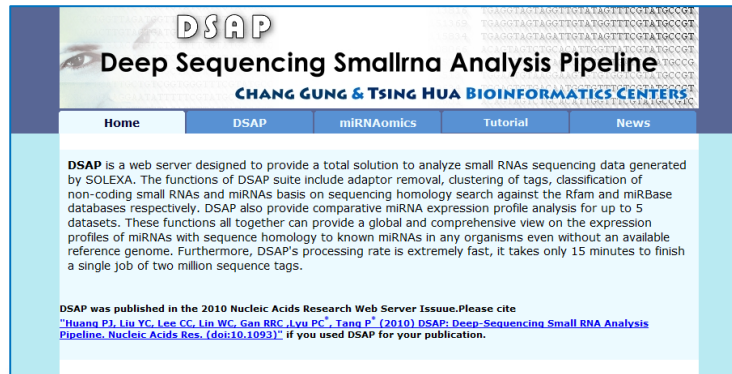


# smallRNA-Seq



# Bioinformatics Tools for sncRNA-Seq data analysis

## Web-based tools



**DSAP**  
Deep Sequencing SmallRNA Analysis Pipeline  
CHANG GUNG & TSING HUA BIOINFORMATICS CENTERS

Home DSAP miRNAomics Tutorial News

DSAP is a web server designed to provide a total solution to analyze small RNAs sequencing data generated by SOLEXA. The functions of DSAP suite include adaptor removal, clustering of tags, classification of non-coding small RNAs and miRNAs basis on sequencing homology search against the Rfam and miRBase databases respectively. DSAP also provide comparative miRNA expression profile analysis for up to 5 datasets. These functions all together can provide a global and comprehensive view on the expression profiles of miRNAs with sequence homology to known miRNAs in any organisms even without an available reference genome. Furthermore, DSAP's processing rate is extremely fast, it takes only 15 minutes to finish a single job of two million sequence tags.

DSAP was published in the 2010 Nucleic Acids Research Web Server Issue. Please cite  
"Huang PJ, Liu YC, Lee CC, Lin WC, Gan RRC, Lyu PC, Tang P (2010) DSAP: Deep-Sequencing Small RNA Analysis Pipeline, Nucleic Acids Res. [doi:10.1093]" if you used DSAP for your publication.



**mirTools**  
microRNA Profiling and Discovery Based on High-throughput Sequencing

Main Menu  
Home  
Filter  
Single  
Multiple  
Group  
Retrieve  
Help  
Feedback  
About Us

Useful Links  
miRanalyzer  
miRBase  
Rfam  
UCSC

MicroRNAs (miRNAs) are small, non-coding RNA (~20-22 nucleotides) that negatively regulate gene expression at post-transcriptional level. Accumulating evidence indicates that miRNAs play crucial roles in various physiological and pathological processes, such as development and tumorigenesis. Therefore, identification of comprehensive sets and differentially expressed miRNAs across tissues and cell lines becomes quite a hot area of study right now in solving a lot of the problems of gene regulation.

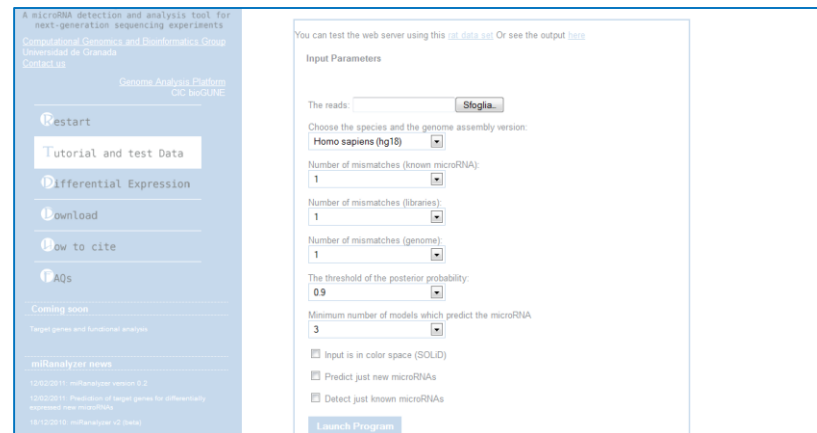
To comprehensive characterize the small RNA transcripts, the appearance of 'deep-sequencing' technologies, such as Roche 454 and Illumina Solexa, have a number of significant advantages in comparison with previous hybridization-based methodologies, such as microarray or PCR-based assays. Firstly, it provides a more integrated view of the miRNAs transcripts. With the added sequencing depth, high-throughput sequencing have an ability to identify modest or even low abundance miRNAs exhibiting expression differences between different samples, which can not be detected previously. Secondly, direct sequencing also offers the potential to detect variation in mature miRNA length and enzymatic modification of miRNA. Thirdly, high-throughput sequencing allows the successful discovery of novel miRNAs, which need not rely on querying candidate regions of the genome but rather can be achieved by direct observation and validation of the folding potential of flanking genomic sequence. Taken together, next-generation sequencing technologies offer a highly robust, accurate and scalable system that sets a new standard for productively, cost-effectively investigation of small RNA transcripts.

Although deep-sequencing schemes can generate millions of short sequences, they also present substantial informatics challenges for lack of efficient and flexible tools to handle and analysis a huge scale of short sequences. Therefore, a comprehensive web server mirTools was developed to comprehensive characterizes the small RNA transcripts.

Notice:  
miRNA database updated to miRBase Release 16.

How to cite:  
Zhu EL, Zhao FQ, Zhou LL, Hou HB, Xu G, Li XK, Bao QY, Sun ZS and Wu JY (2010) mirTools: microRNA profiling and discovery based on high-throughput sequencing, Nucleic Acids Research, 2010

The intention of mirTools:



A microRNA detection and analysis tool for next-generation sequencing experiments  
Computational Genomics and Bioinformatics Group  
Universidad de Granada  
Contact us

Genome Analysis Platform  
CIC bioCINE

Restart  
Tutorial and test Data  
Differential Expression  
Download  
How to cite  
FAQs

Coming soon  
Target genes and functional analysis

miRanalyzer news  
12/02/2011: miRanalyzer version 0.2  
12/02/2011: Prediction of target genes for differentially expressed miRNAs  
16/02/2010: Collaboration (2010)

You can test the web server using this [get data set](#) Or see the output [here](#)

Input Parameters

The reads:

Choose the species and the genome assembly version:

Number of mismatches (known microRNA):

Number of mismatches (libraries):

Number of mismatches (genome):

The threshold of the posterior probability:

Minimum number of models which predict the microRNA:

☐ Input is in color space (SOLID)  
☐ Predict just new microRNAs  
☐ Detect just known microRNAs

Launch Program

# Bioinformatics Tools for sncRNA-Seq data analysis

## Stand-alone tools



# *iMir*

Integrated pipeline for HT miRNA-Seq data analysis

Modular pipeline for comprehensive analysis of sncRNA-Seq data.

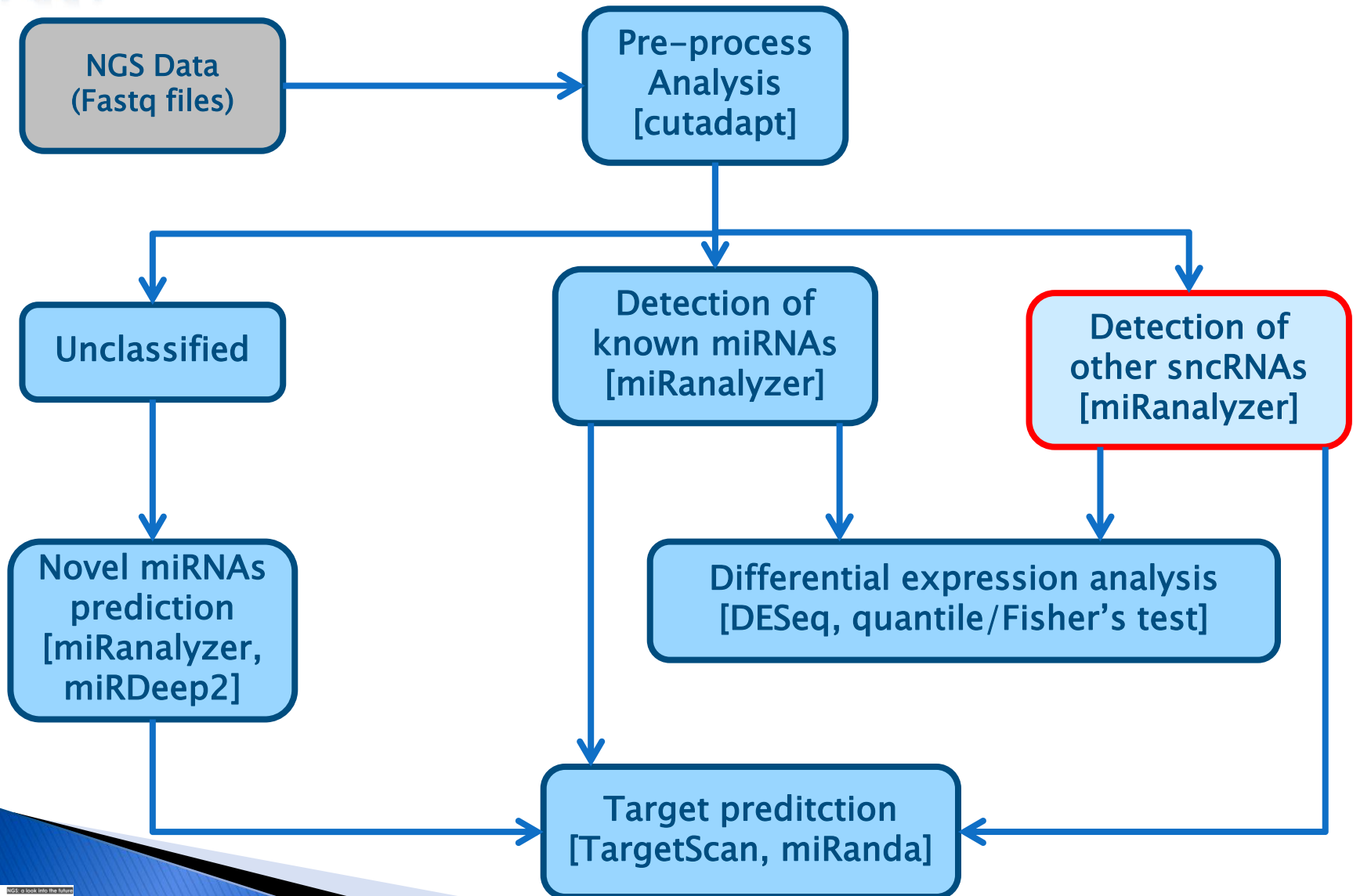
Multiple open source modules and resources linked together in automated flow

Graphical User Interface (GUI) to create projects and specify parameters

The pipeline output includes graphics and text files that are useful for a better interpretation of the results



# Workflow scheme



# iMir

Integrated pipeline for HT miRNA-Seq data analysis

The image displays the iMir software interface, which is designed for the integrated pipeline of HT miRNA-Seq data analysis. The main window, titled "iMir", features a "Run Complete Analysis" button and a menu of analysis options: "Remove Adapter", "Identification of sncRNAs/Novel miRNAs Prediction", "Diff Exp Analysis", "Target Prediction", and "Prediction novel miRNAs by miRDeep2". The footer identifies the "Laboratory of Molecular Medicine and Genomics, University of Salerno, Italy".

A secondary window, "iMir - Complete Analysis", provides detailed configuration options. It includes an "Output Directory" field with a "source..." button. The "Adapter Removal" section contains fields for "Adapter Sequence 3'" (TGGAATTCTCGGGTGCCAAGG), "Adapter Sequence 5'" (GTTCAGAGTTCTACAGTCCGACGATC), "Minimum Read Length" (16), "Error Rate in Adapter Sequence" (0.1), and "Quality CutOff in Adapter Sequence" (0). It also has checkboxes for "Colospace", "Double Encode", "Trim Primer", and "Strip F3". The "Other Parameters" section includes a "Minimum Read Count" set to 3. The "WorkFlow" section has checkboxes for "Length Distribution Analysis", "Identification of sncRNAs/Novel miRNAs Prediction", "Diff Exp Analysis", "Target Prediction", "Samples Cluster", and "Target Prediction Diff Exp miRNAs". The footer of this window also identifies the "Laboratory of Molecular Medicine and Genomics, University of Salerno, Italy". Navigation buttons "Main Menu", "Exit", and "Next >" are located at the bottom right.

# piRNAs biogenesis

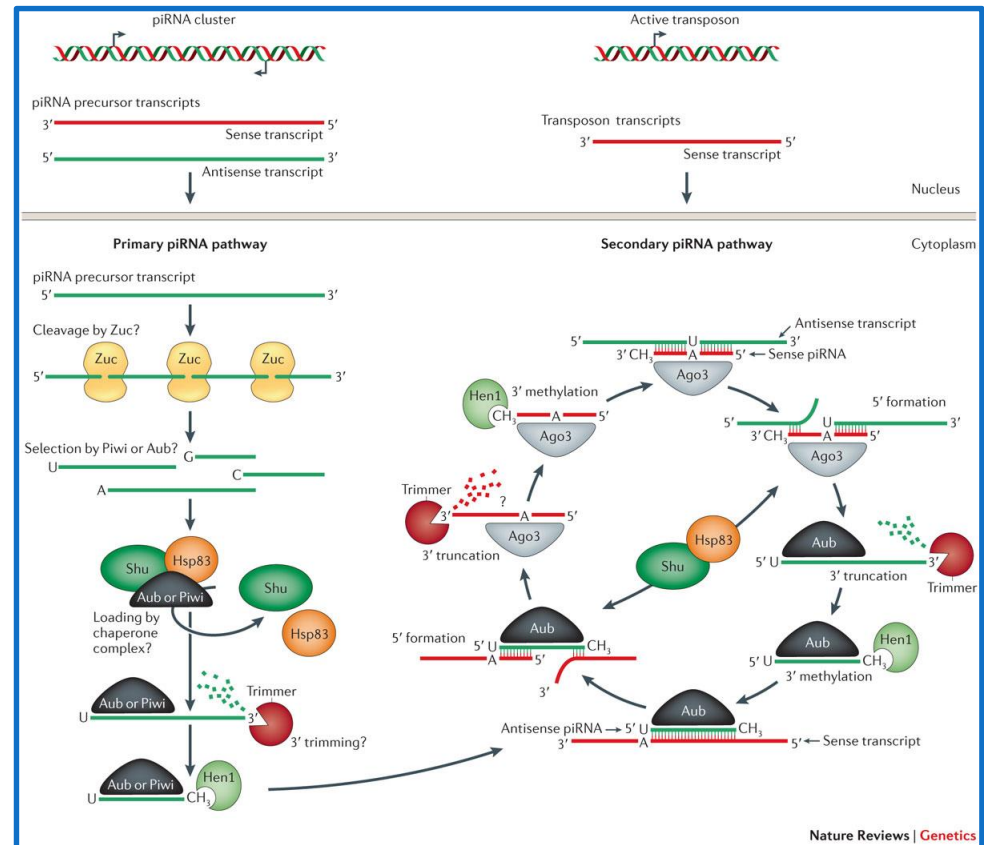
•The recent discovery of major class of sncRNAs initially identified in the germline of a variety of species, called PIWI-interacting RNA (piRNA), raised questions concerning their biogenesis and functions.

• piRNAs are frequently encoded by genes clustered in intergenic regions of the genome and have defined characteristics:

- an average length of 24-35 nt,
- a strong preference for uracil at the 5'-end and 2'-O-methylation of the 3' base,
- high sequence diversity.

•They are generated by an intricate pathway conserved in evolution, comprising primary processing and amplification loops ('ping-pong' cycles) and are known to be involved in germline development, silencing of selfish DNA elements and maintenance of germline DNA integrity.

• Recent evidences suggest that the Piwi/piRNA pathway may be functionally active also in somatic tissues, but this possibility has not yet been fully explored.





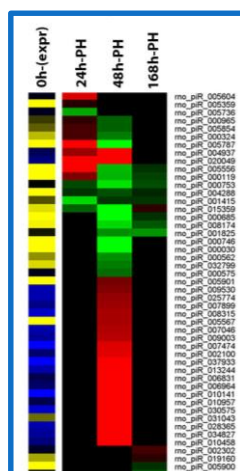
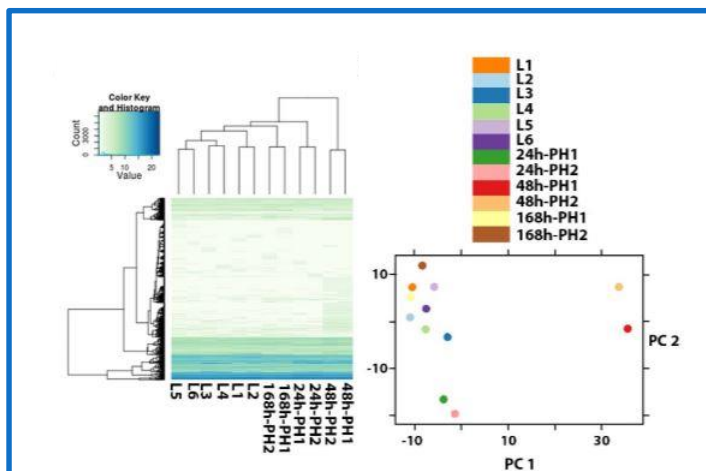
## RAPID COMMUNICATION

# Timed Regulation of P-Element-Induced Wimpy Testis-Interacting RNA Expression During Rat Liver Regeneration

Francesca Rizzo,<sup>1\*</sup> Adnan Hashim,<sup>1\*</sup> Giovanna Marchese,<sup>1</sup> Maria Ravo,<sup>1</sup> Roberta Tarallo,<sup>1</sup> Giovanni Nassa,<sup>1</sup> Giorgio Giurato,<sup>1</sup> Antonio Rinaldi,<sup>1</sup> Angela Cordella,<sup>2</sup> Marcello Persico,<sup>3</sup> Pia Sulas,<sup>4</sup> Andrea Perra,<sup>4</sup> Giovanna M. Ledda-Columbano,<sup>4</sup> Amedeo Columbano,<sup>4</sup> and Alessandro Weisz<sup>1</sup>

•We investigated piRNAs expression in rat liver and its response to the stimuli exerted by regenerative proliferation of this organ.

•RNA sequencing before, during and after the wave of cell proliferation that follows partial hepatectomy identified ~1400 mammalian germline piRNAs expressed, including 72 showing timed changes in expression 24-48 hours post-PH.



## RNA sequencing identifies specific PIWI-interacting small non-coding RNA expression patterns in breast cancer

Adnan Hashim<sup>1,\*</sup>, Francesca Rizzo<sup>1,\*</sup>, Giovanna Marchese<sup>1</sup>, Maria Ravo<sup>1</sup>, Roberta Tarallo<sup>1</sup>, Giovanni Nassa<sup>1</sup>, Giorgio Giurato<sup>1</sup>, Gianluca Santamaria<sup>1</sup>, Angela Cordella<sup>2</sup>, Concita Cantarella<sup>3</sup> and Alessandro Weisz<sup>1,4</sup>

<sup>1</sup> Laboratory of Molecular Medicine and Genomics, Faculty of Medicine and Surgery, University of Salerno, Baronissi, SA, Italy

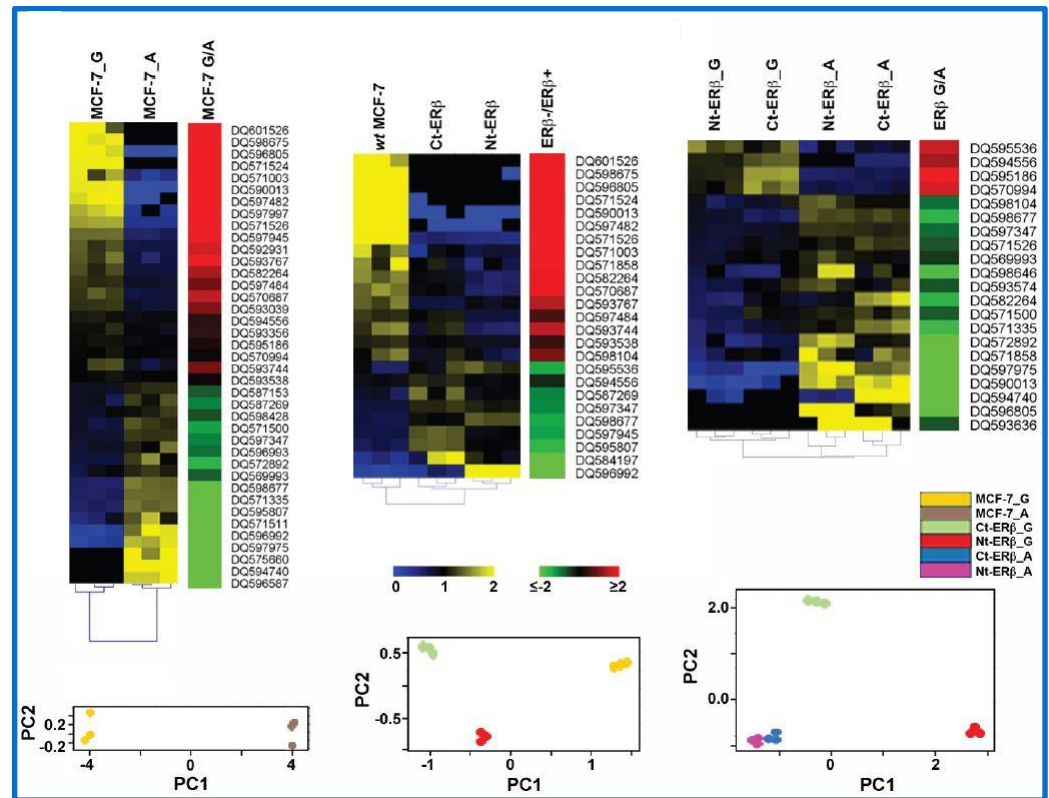
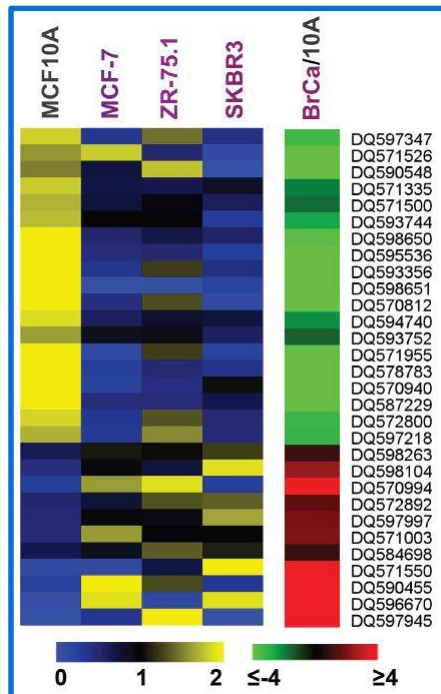
<sup>2</sup> Fondazione IRCCS SDN, Napoli, Italy

<sup>3</sup> Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Centro di Ricerca per l'Orticoltura, Pontecagnano, SA, Italy

<sup>4</sup> Division of Molecular Pathology and Medical Genomics, 'SS. Giovanni di Dio e Ruggi d'Aragona' Hospital, University of Salerno, Salerno, Italy

• We found that piRNA biogenesis and effector pathway are present in human breast cancer cell lines and tumor biopsies.

• We identified > 100 BC piRNAs, including some very abundant and/or differentially expressed in mammary epithelial compared to BC cells, where these were influenced by estrogen or estrogen receptor  $\beta$ .



## Small non-coding RNA deregulation in endometrial carcinogenesis

Maria Ravo<sup>1</sup>, Angela Cordella<sup>2</sup>, Antonio Rinaldi<sup>1</sup>, Giuseppina Bruno<sup>1</sup>, Elena Alexandrova<sup>1,3</sup>, Pasquale Saggese<sup>1</sup>, Giovanni Nassa<sup>1</sup>, Giorgio Giurato<sup>1</sup>, Roberta Tarallo<sup>1</sup>, Giovanna Marchese<sup>1,3</sup>, Francesca Rizzo<sup>1</sup>, Claudia Stellato<sup>1</sup>, Rossella Biancardi<sup>4</sup>, Jacopo Troisi<sup>4</sup>, Attilio Di Spiezio Sardo<sup>5</sup>, Fulvio Zullo<sup>5</sup>, Alessandro Weisz<sup>1,6</sup>, Maurizio Guida<sup>4</sup>

<sup>1</sup>Laboratory of Molecular Medicine and Genomics, Department of Medicine and Surgery, University of Salerno, Baronissi, Italy

<sup>2</sup>Fondazione IRCCS SDN, Napoli, Italy

<sup>3</sup>Genomix4Life Srl, Spin-Off of the Laboratory of Molecular Medicine and Genomics, University of Salerno, Baronissi, Italy

<sup>4</sup>Department of Medicine and Surgery and Division of Gynecology and Obstetrics, "SS. Giovanni di Dio e Ruggi d'Aragona - Schola Medica Salernitana", University of Salerno Hospital, Salerno, Italy

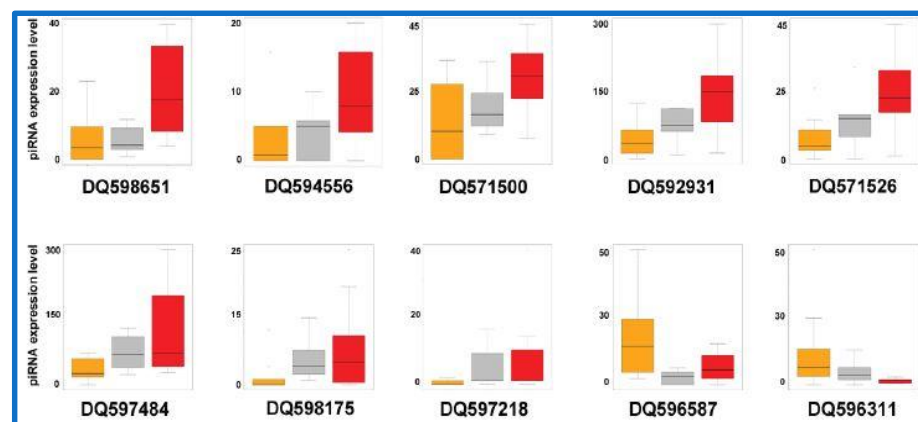
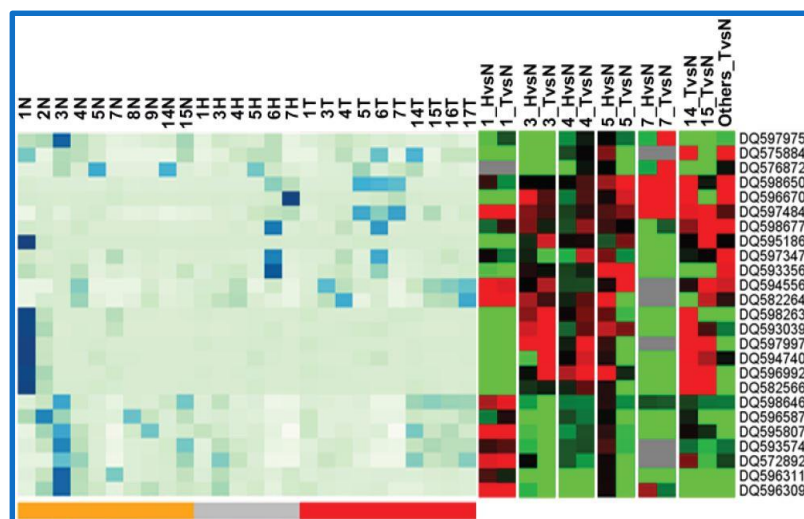
<sup>5</sup>Department of Gynecology and Obstetrics and Pathophysiology of Human Reproduction, University of Naples "Federico II", Napoli, Italy

<sup>6</sup>Division of Molecular Pathology and Medical Genomics, "SS. Giovanni di Dio e Ruggi d'Aragona - Schola Medica Salernitana", University of Salerno Hospital, Salerno, Italy

• Changes in sncRNA expression were identified by high-throughput genomic analysis of paired normal, hyperplastic and cancerous endometrial tissues,

• This led to the definition of a sncRNA signature of neoplastic transformation.

• Considering the regulatory role of sncRNAs, this newly identified signature is likely to reflect the events leading to endometrial cancer



Normal (N)  
Hyperplastic (H)  
Tumoral (T)





# iMir: An innovative and complete pipeline for smallRNA-Seq data analysis



<http://www.labmedmolge.unisa.it/inglese/research/imir>



# Acknowledgements



NGS, microarray and bioinformatics  
services for research and industry

*Laboratorio di Medicina Molecolare e Genomica*

Facoltà di Medicina e Chirurgia

Università degli Studi di Salerno

Via S. Allende, 1 84081 Baronissi (Sa)

[www.labmedmolge.unisa.it](http://www.labmedmolge.unisa.it)

**Prof. Alessandro Weisz**

## *Bioinformatics core:*

Antonio Rinaldi  
Domenico Memoli  
Riccardo Panero  
Giacomo Corleone

## *Genomics core:*

Roberta Tarallo  
Maria Ravo  
Giovanni Nassa  
Francesca Rizzo  
Angela Cordella  
Giovanna Marchese



MINISTERO DELL'ISTRUZIONE,  
DELL'UNIVERSITÀ E DELLA RICERCA