BIOINFORMATICS APPROACH FOR PREDICTION OF TARGET GENES IN A SET OF DEREGULATED MICRORNAS IN HEPATOCELLULAR CARCINOMA: STUDY ON A CHEMICAL-INDUCED HCC MOUSE MODEL

L' Aquila, 1° March 2016

Dr. Filippo Del Vecchio
PhD student
Molecular Oncopathology Lab
Department of Biotechnological
and Applied Clinical Sciences
University of L'Aquila

HEPATOCELLULAR CARCINOMA (HCC) EPIDEMIOLOGY

Main hepatic primary tumour

Extremely aggressive with high rates of recurrence

Limited therapeutic options

Multi-phase pathogenesis

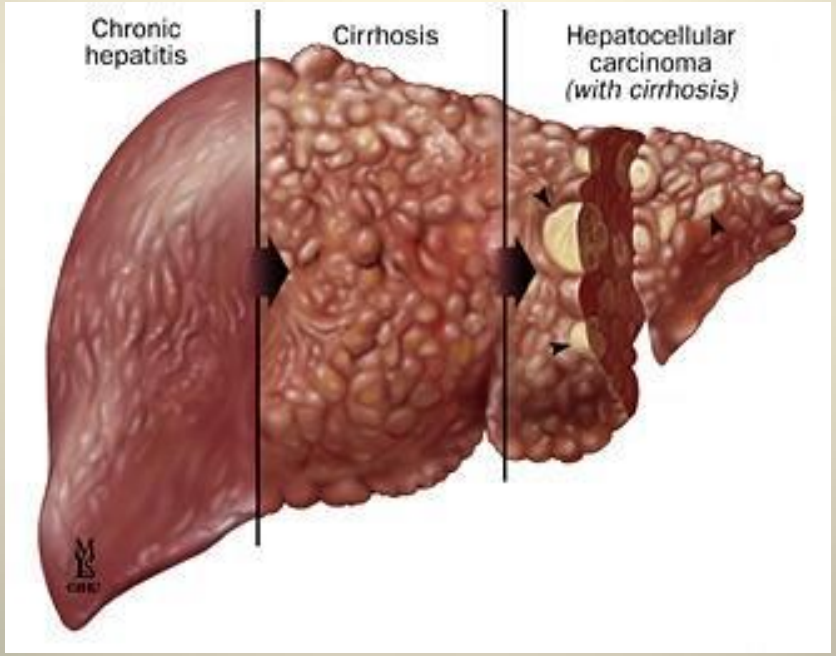
4.9 per 100.000 persons/year

Viral infections (HBV, HCV), metabolic disorders,
Toxic insults, autoimmune reactions

HEPATOCELLULAR CARCINOMA (HCC) EPIDEMIOLOGY

Significant correlation between metabolic disorders and Non Alcoholic Fatty Liver Disease (NAFLD) or HCC.

Liver damage progression → HCC development

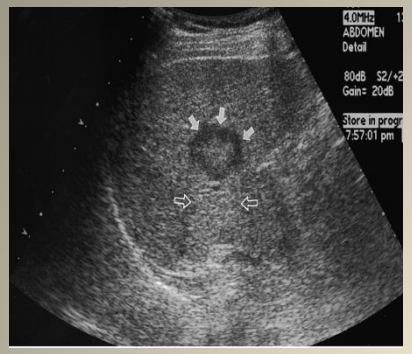


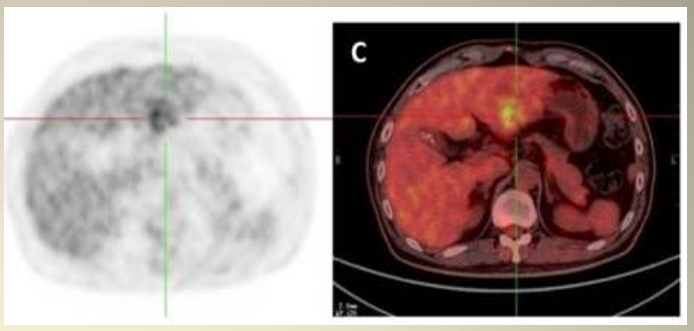
HCC STAGING

European systems	*Asian systems*
Barcelona-Clinic Liver Cancer (BCLC)	Okuda staging
Cancer of the Liver Italian Program (CLIP)	Japan Integrated Staging (JIS) score
tumour-lymph nodes-metastasis staging TNM	Chinese University Prognostic Index (CUPI)

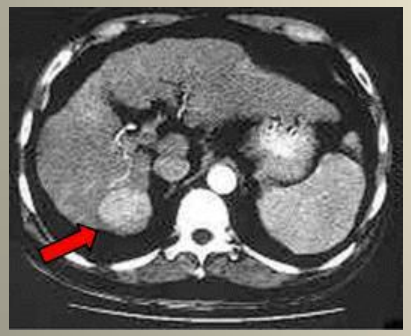
Median survival rates : 6-20 months from diagnosis

Early diagnosis is essential!



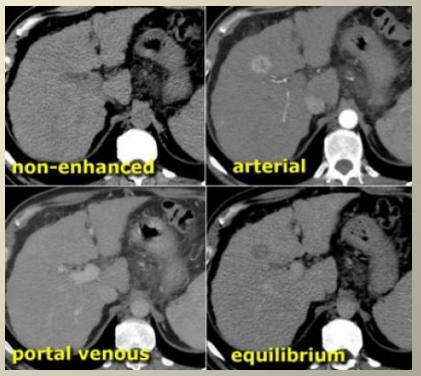


HCC DIAGNOSIS



BIOMARKERS

AFP, AFP-L3, OPN,
GPC3, GP73



HCC THERAPEUTIC MANAGEMENT

Not only HCC needs to be treated but also underlying liver disease require medical attention.

Before therapeutics, stratify the patients, usually BCLC, with a staging system is an effective strategy.

THERAPEUTIC CHOICES:

Hepatic resection

Liver transplantation

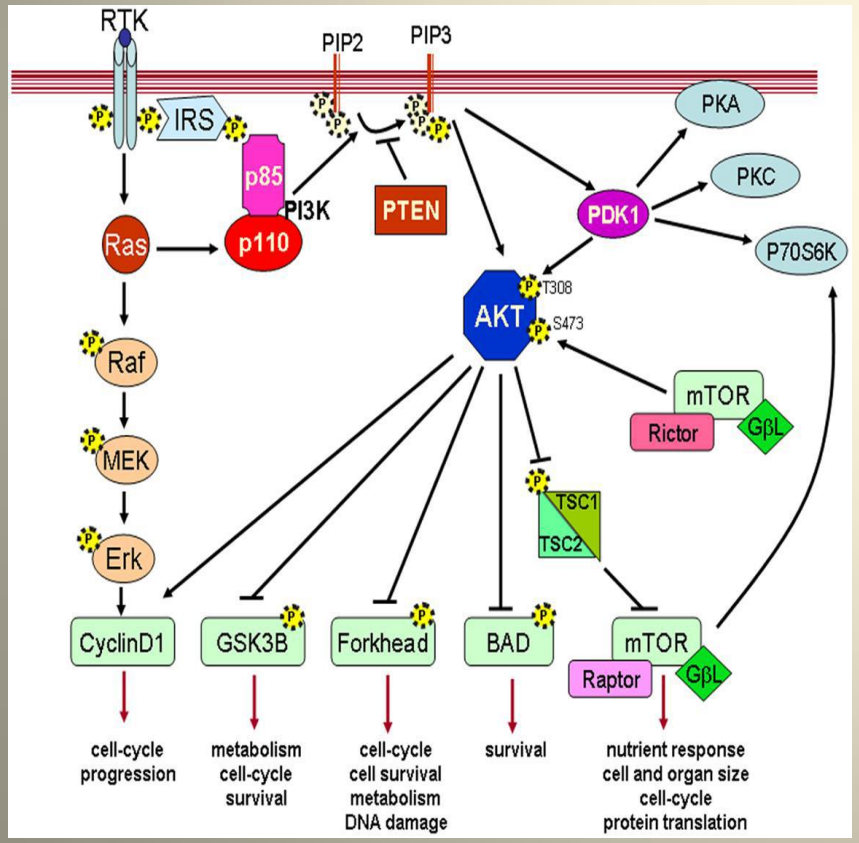
Percutaneous therapies (PEI, RFA)

Transarterial chemo/radioembolization (TACE, TARE)

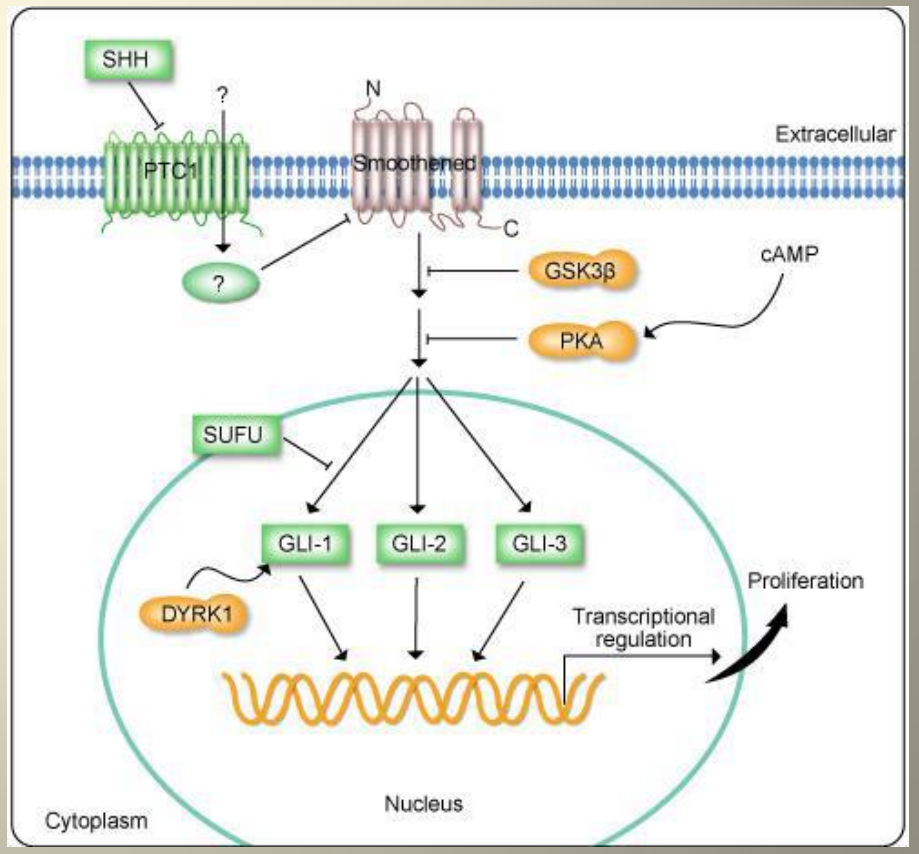
Kinase inhibitors (Sorafenib, Tivantinib, Cabozantinib)

DEREGULATED PATHWAYS IN HCC

PI3K/AKT signaling pathway

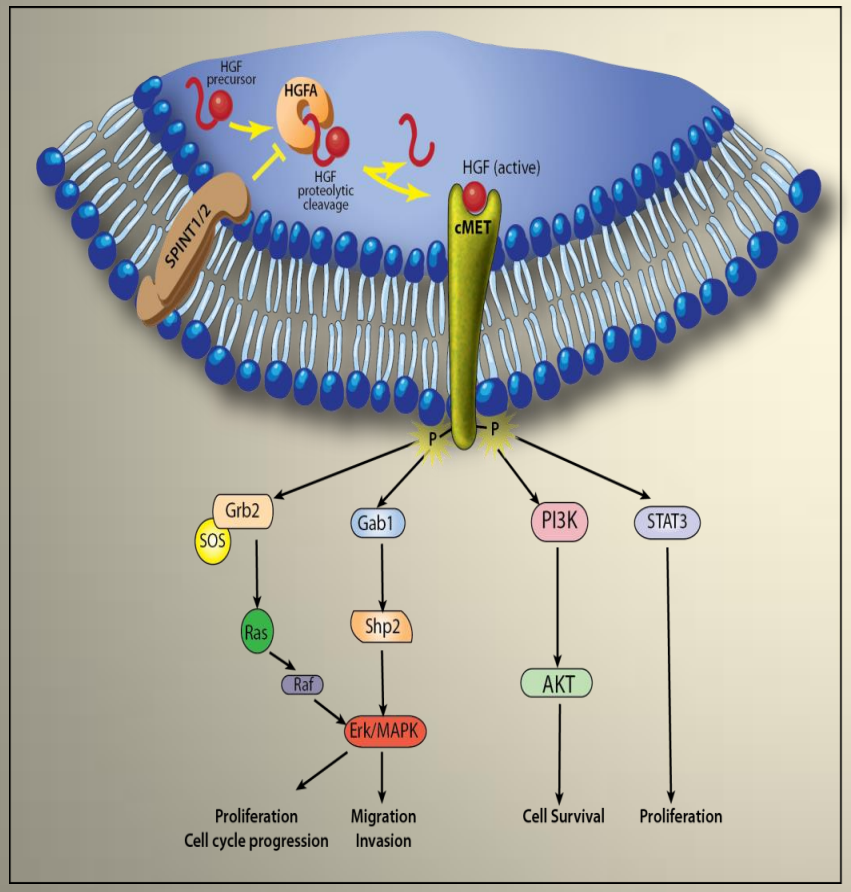


Sonic-hedgehog pathway

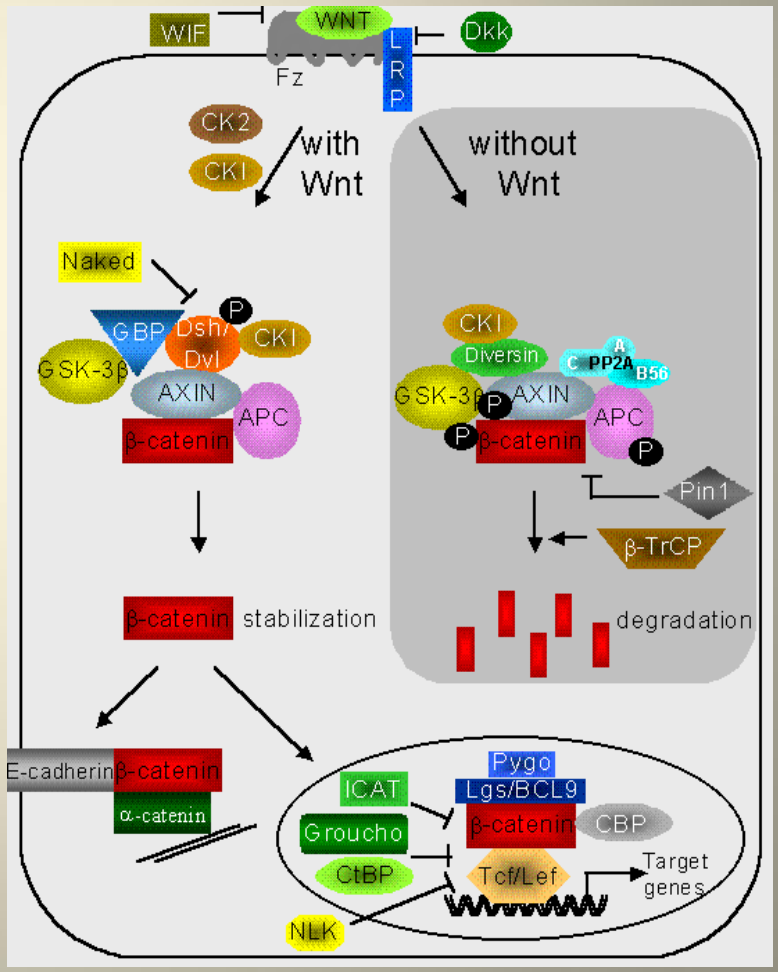


DEREGULATED PATHWAYS IN HCC

HGF/MET signaling pathway



WNT/β-catenin pathway



MicroRNAs

Non-coding RNAs

Tiny molecules (18-25 nt in length)

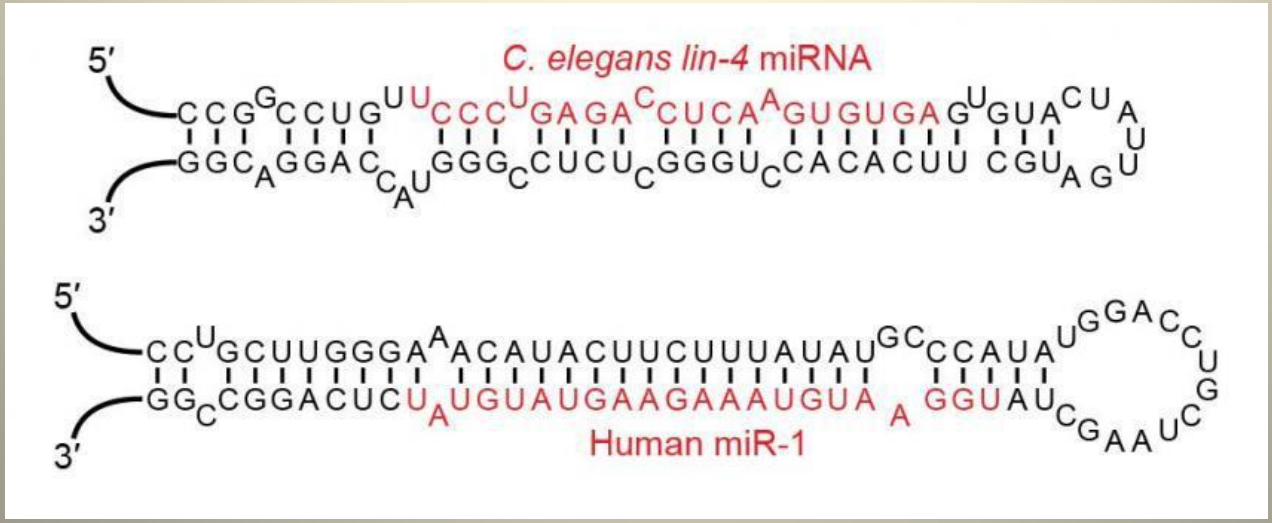
Single stranded

Post-transcriptional negative regulators of mRNA translation

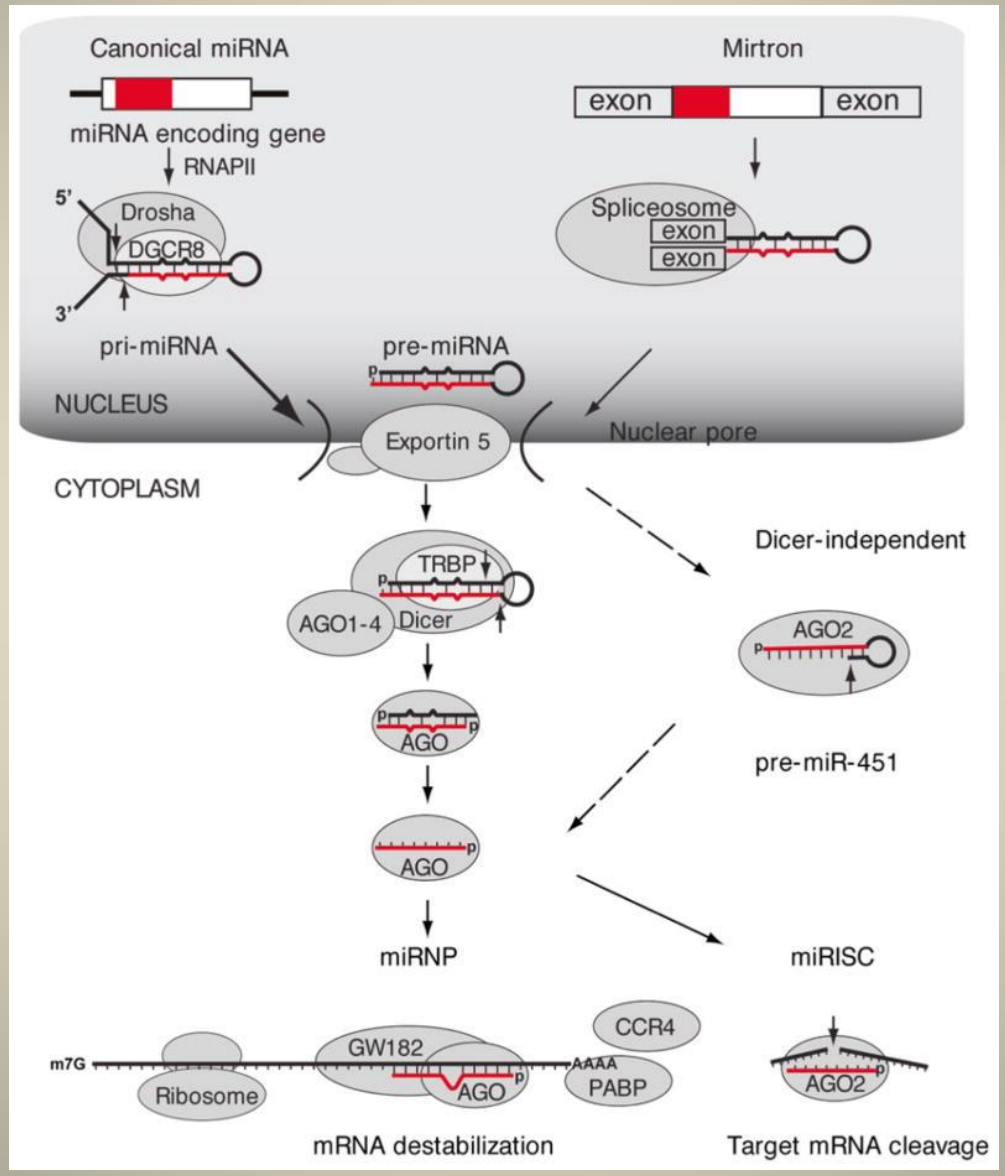
The first one, lin-4, discovered in *C.elegans* in 1993

Actually, almost 2600 mature sequences and 1880 precursors are deposited for human species (MiRbase, February 2016)

Regulation of cell proliferation, apoptosis, neuronal patterning, developmental processes and many more



MicroRNAs BIOGENESIS



MIRNAs-mRNAs INTERACTIONS

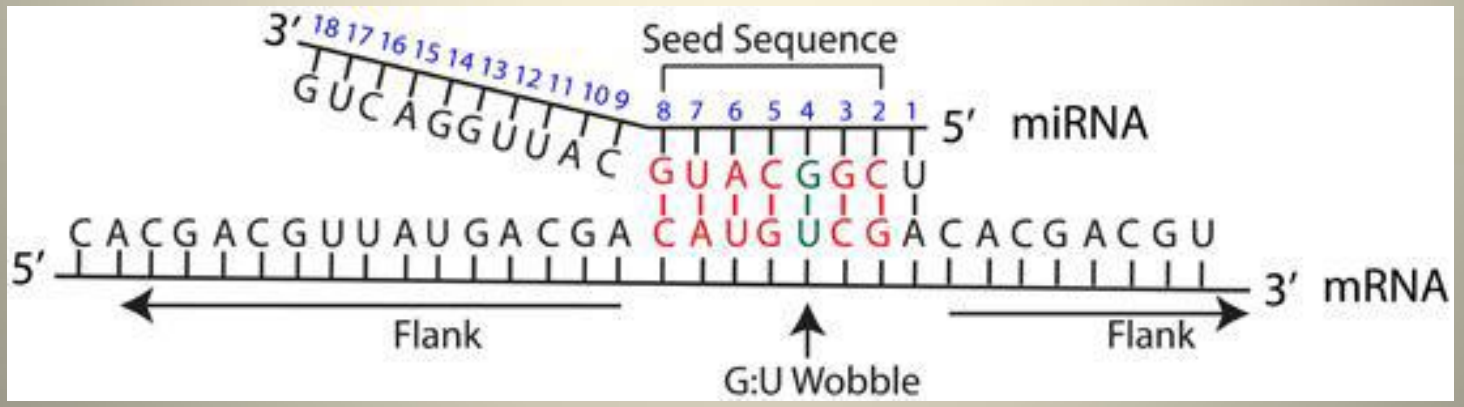
Generally it is an asymmetric interaction

It occurs between the “seed region” on the 5’ end of miRNA’s sequence and the “microRNA recognition elements” (MRE) on the 3’UTR of mRNA’s sequence.

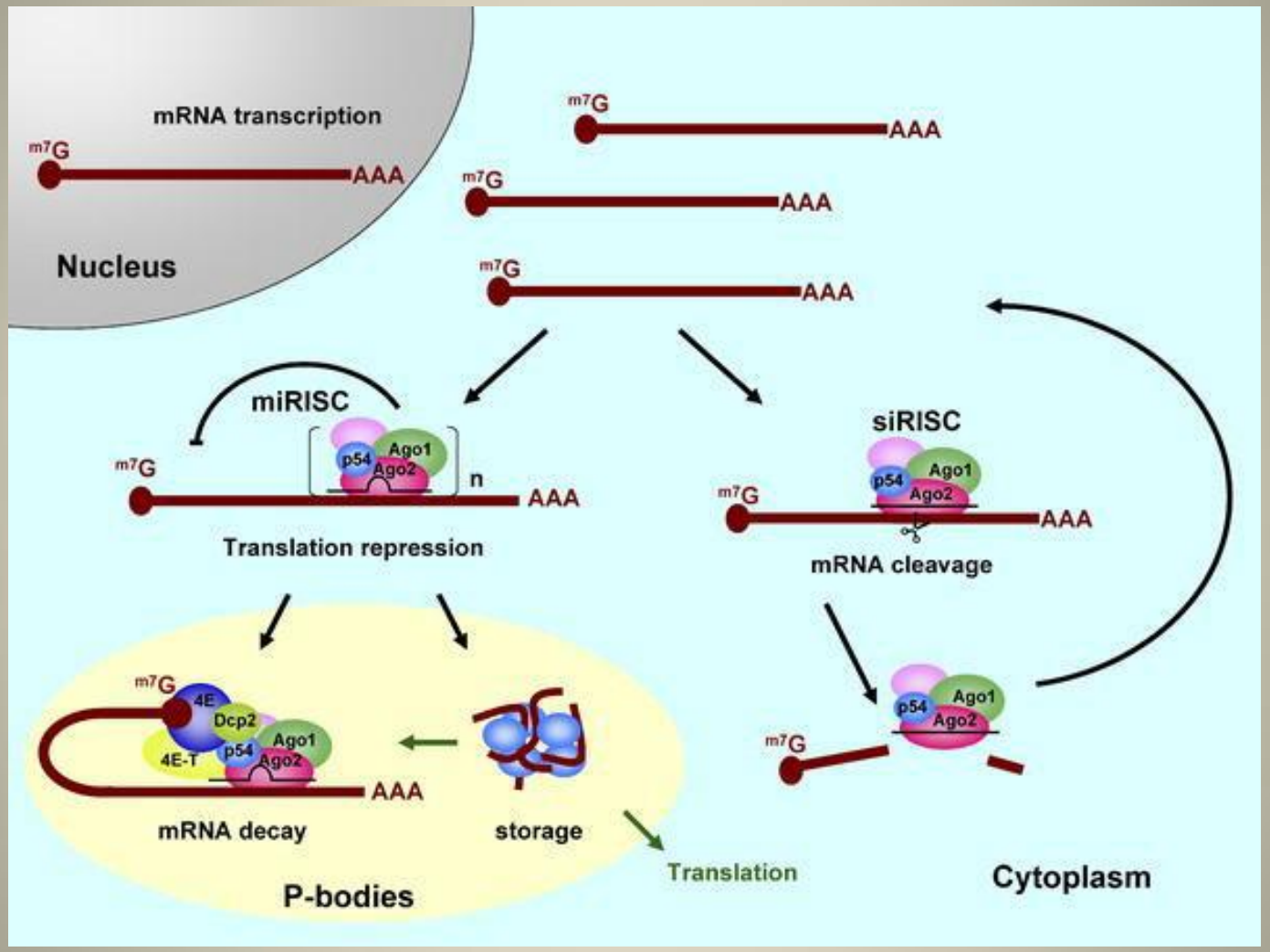
Alternatively, miRNAs-mRNAs interactions could happen in a “seedless way” (3’ end miRNA ⟷ 5’end mRNA).

Multiple binding sites can enhance the global interaction.

Complex mRNAs secondary structures could hide potential binding sites.



MiRNAs MECHANISMS OF ACTION



MiRNAs AND CANCER

Profound relationships with cancer events

One miR → multiple target genes → multiple signaling pathways

ONCOMIRS

Favours cellular oncogenic properties

Up-regulated in cancer

TUMOR SUPPRESSORS

Oppose to cellular oncogenic activities

Down-regulated in cancer

What is their utility?

-Potential cancer biomarkers

- miRNA-based prognostic approach

-Therapeutic molecules

BIOINFORMATICS

The term bioinformatics refers to a discipline aiming to handle and analyze biomedical information through the use of computers.

Great innovation in:

-Global gene expression profiles

-Homology and alignment studies

-Interaction studies

-Prediction of protein structures

MiRNAs and bioinformatics, why we can associate them?

In the identification of potential target genes for one or more miRNAs, applying some bioinformatics techniques avoid an impossible manual search between thousands of candidates, allowing to identify the most probable of them with high probabilities of success

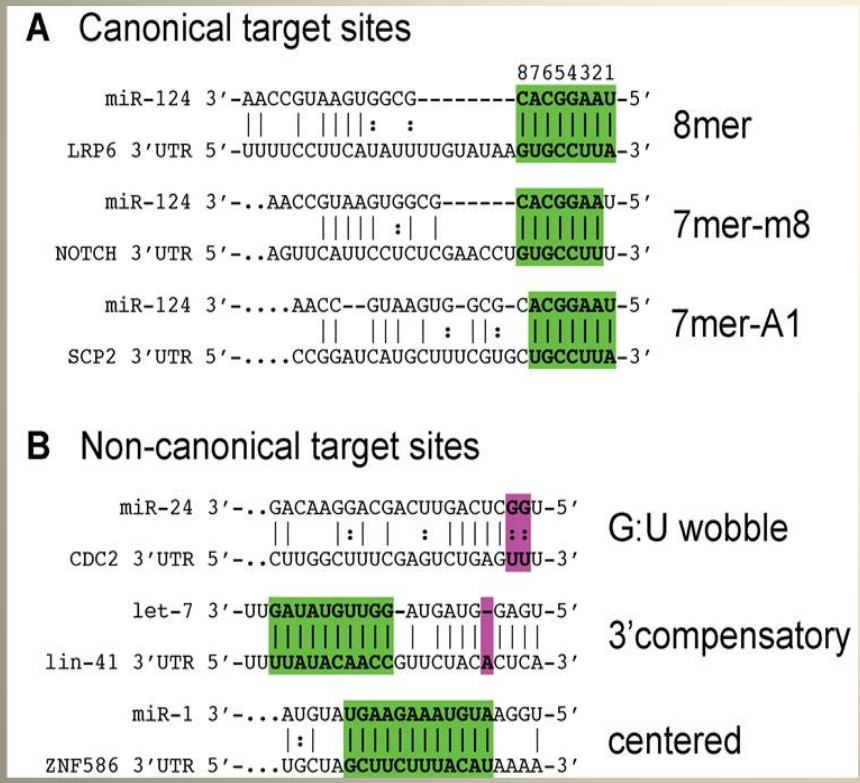
MiRNAs DATABASES

miRTarBase

miR2Disease Base

MiRNAs TARGET PREDICTION TOOLS

Complementarity between miRNA and mRNA



Presence of multiple binding sites on mRNA sequence

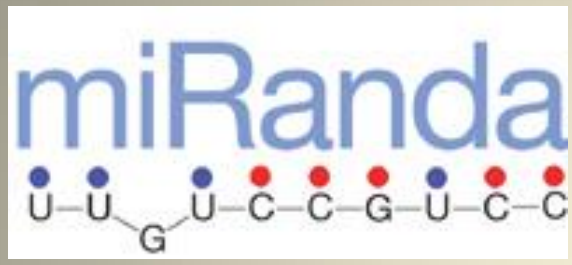
3’UTR orthologous sequence conservation

Site accessibility

Sequence patterns repeats

MiRNAs TARGET PREDICTION TOOLS

ALGORITHMS USING CONSERVATION FILTERS



Alignment score = sum of the scores of every single potential binding sites

Penalties for mismatch or gaps

Also site accessibility and multiple binding sites are considered into the final score

Identify non-canonical and non-conserved sites through mirSVR score

Searches for miRNA families

Double scoring system

Context+score for the interaction's efficacy between miRNA and mRNA

PcT score reflects probability of conserved sequences

Also site accessibility and multiple binding sites are considered into the final score

MiRNAs TARGET PREDICTION TOOLS

ALGORITHMS WITHOUT CONSERVATION FILTERS

PITA Filters measure site accessibility

Double-phase mechanism: first check of seed complementarity and then estimation of free energy interactions values

RNA-22 Detection of repeated sequence patterns

Score assignment for “hot spot” sequences

Estimation through free energy values of possible duplexes formation

GENE ONTOLOGY

GENE ENRICHMENT ANNOTATION

Biological raw data need to be refined to extract significant informations

This approach permits to take out what is actually known about analyzed genes and link them to specific biological processes related to the conditions under investigation

Gene-centric → group-of-genes perspective

Proper databases (DAVID, AmiGO, Reactome, Mouse Genome Informatics)

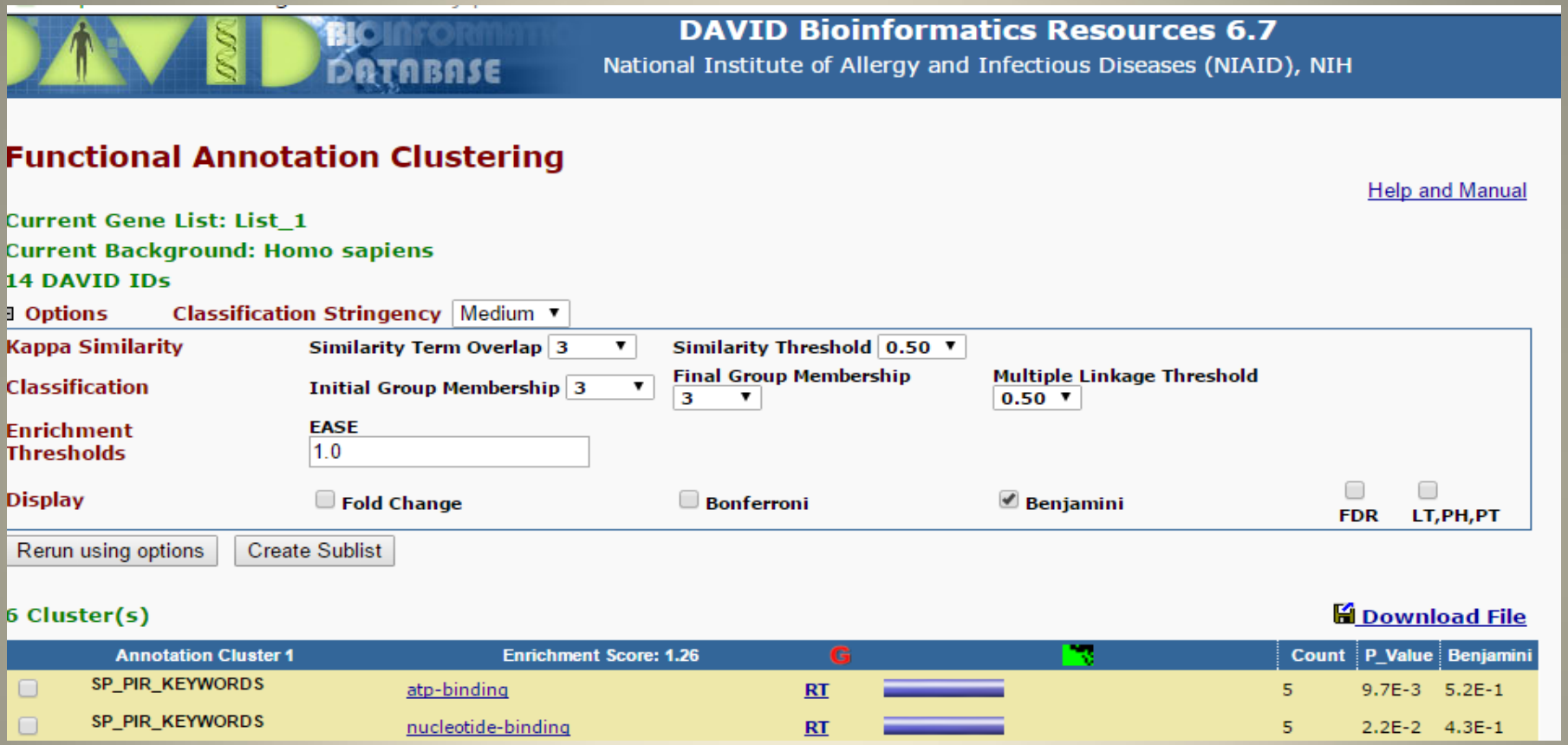
INTERACTION NETWORKS

Literature data mining

Construction of molecular networks and graphical visualization (Cytoscape, Ingenuity, Genemania)

Connection to biological pathways

GENE ONTOLOGY



DAVID Bioinformatics Resources 6.7
National Institute of Allergy and Infectious Diseases (NIAID), NIH

Functional Annotation Clustering

Help and Manual

Current Gene List: List_1
Current Background: Homo sapiens
14 DAVID IDs

Options Classification Stringency: Medium

Kappa Similarity Similarity Term Overlap: 3 Similarity Threshold: 0.50

Classification Initial Group Membership: 3 Final Group Membership: 3 Multiple Linkage Threshold: 0.50

Enrichment Thresholds EASE: 1.0

Display Fold Change Bonferroni Benjamini FDR LT,PH,PT

Rerun using options Create Sublist

6 Cluster(s) Download File

Annotation Cluster 1	Enrichment Score: 1.26	G		Count	P_Value	Benjamini
SP_PIR_KEYWORDS	atp-binding	RT		5	9.7E-3	5.2E-1
SP_PIR_KEYWORDS	nucleotide-binding	RT		5	2.2E-2	4.3E-1

MAIN OBJECTIVES

Investigate the role of microRNAs during the progression of liver damage until the onset of HCC

Predict potential target genes for a limited number of selected miRNAs through bioinformatic analyses

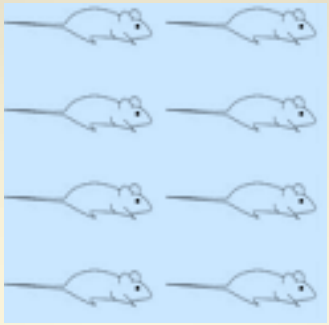
Provide interesting results for further investigations through the integration of *in silico* prediction and *in vivo* data

EXPERIMENTAL CONDITIONS

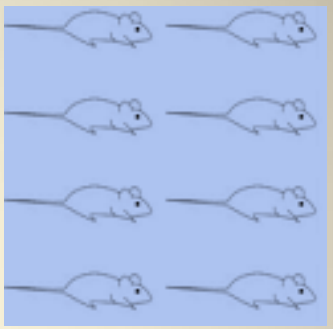
SALINE + DEN (25mg/kg)

Males C57BL/6J
24 mice

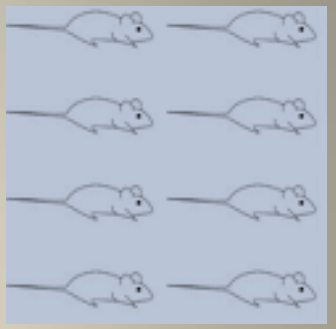
3 DEN
8 mice



6 DEN
8 mice



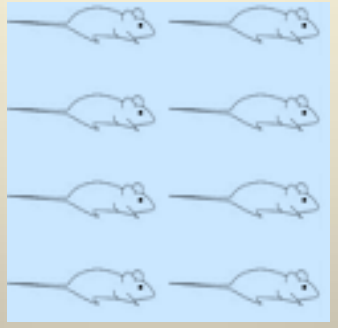
11 DEN
8 mice



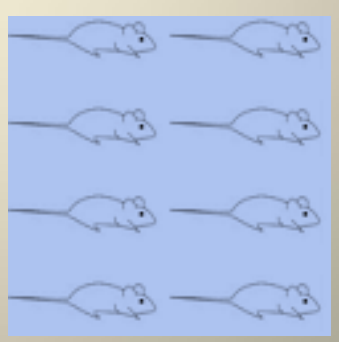
SALINE SOLUTION

Males C57BL/6J
24 mice

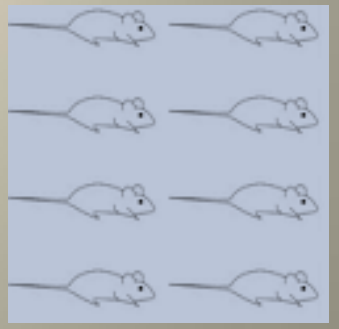
3 DEN
8 mice



6 DEN
8 mice



11 DEN
8 mice



EXPERIMENTAL CONDITIONS

TISSUE COLLECTION

Livers immediately explanted after sacrifice, weighed and sectioned.

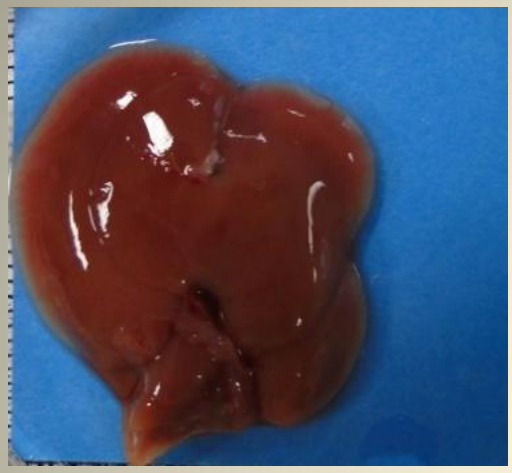
Samples for histological analyses → FFPE → H & E

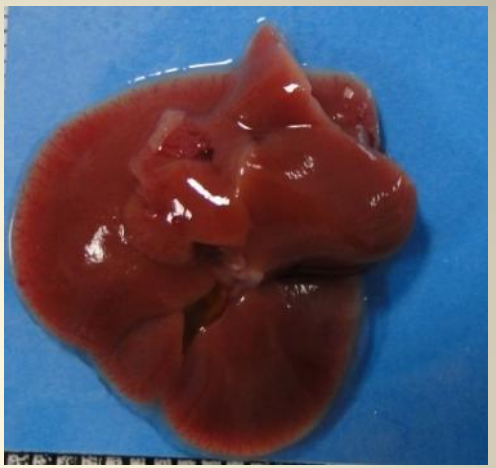
Optical microscope examination

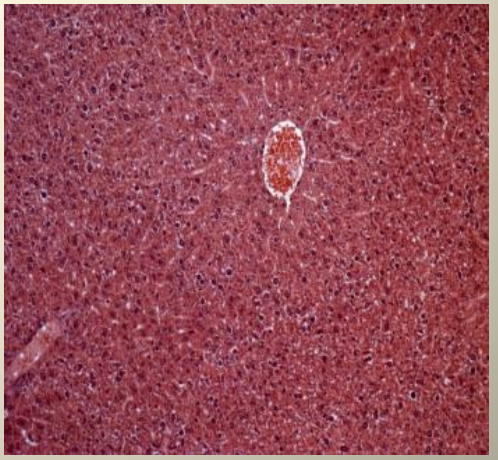
3-MONTHS SAMPLES EVALUATION

DEN-TREATED

CONTROL



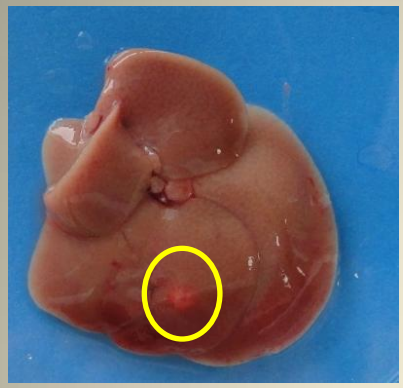


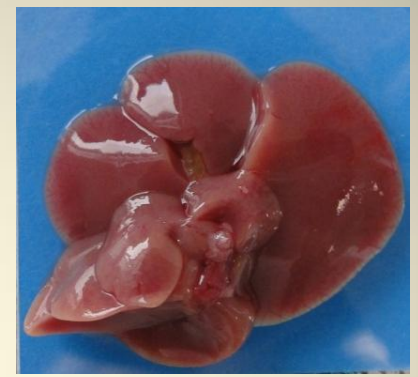


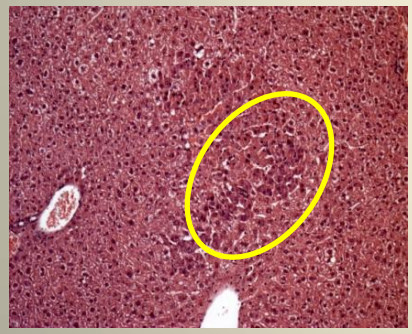
6-MONTHS SAMPLES EVALUATION

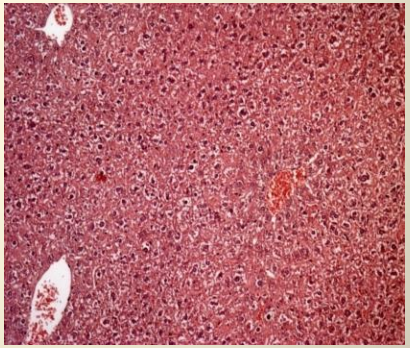
DEN-TREATED

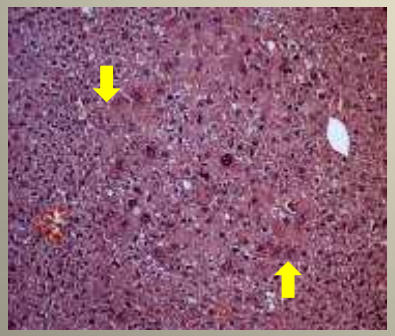
CONTROL







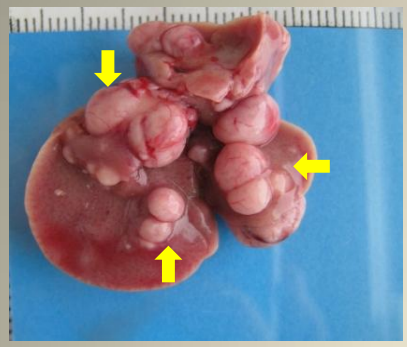


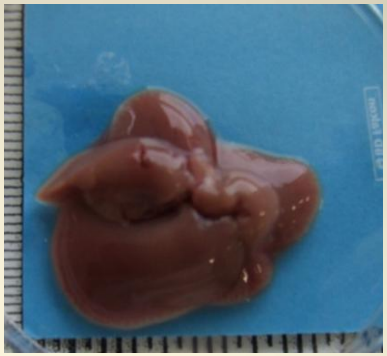


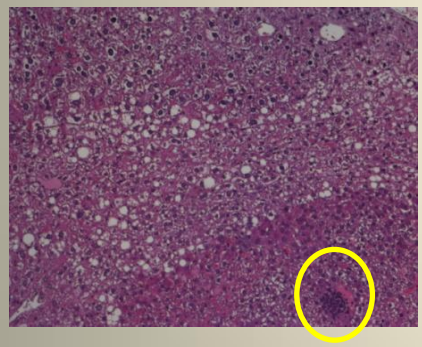
11-MONTHS SAMPLES EVALUATION

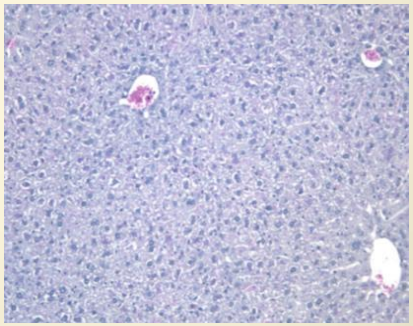
DEN-TREATED

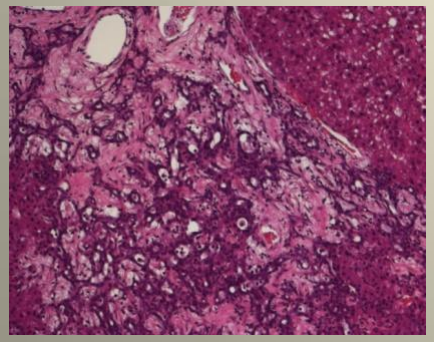
CONTROL











OVERALL EVALUATIONS

NODULAR STRUCTURES PRESENCE					
3-MONTHS		6-MONTHS		11-MONTHS	
DEN-TREATED	CONTROL	DEN-TREATED	CONTROL	DEN-TREATED	CONTROL
absent	absent	20%	absent	100%	absent

INFILTRATING LYMPHOCYTES PRESENCE					
3-MONTHS		6-MONTHS		11-MONTHS	
DEN-TREATED	CONTROL	DEN-TREATED	CONTROL	DEN-TREATED	CONTROL
50%	absent	70%	absent	100%	absent

Dysplastic alterations reported for hepatic tissue sections from mice belonging to the 6- and 11-months DEN-treated group. Histological alterations such as hyperaemia, portal vein congestion, neo-angiogenesis, multiple micronodules and wide fibrotic branches were evidenced

MiRNAs ANALYSIS

RNA extraction

Check for concentration, purity and integrity

RNAs extracted by each mouse were mixed to obtain a RNA pool for DEN-treated animals and a pool for controls. These two pools were obtained for animals belonging to the 3- and 6-months groups.

About 11-months groups, a RNA pool containing RNAs extracted by 7 randomly chosen tumors from DEN-treated animals was added to the established couple of pools.

miRNAs expression analysis → TaqMan rodent miRNA array microfluidic cards (Life Technologies)

Data collection → ViiA-7 system (Life Technologies)

↓

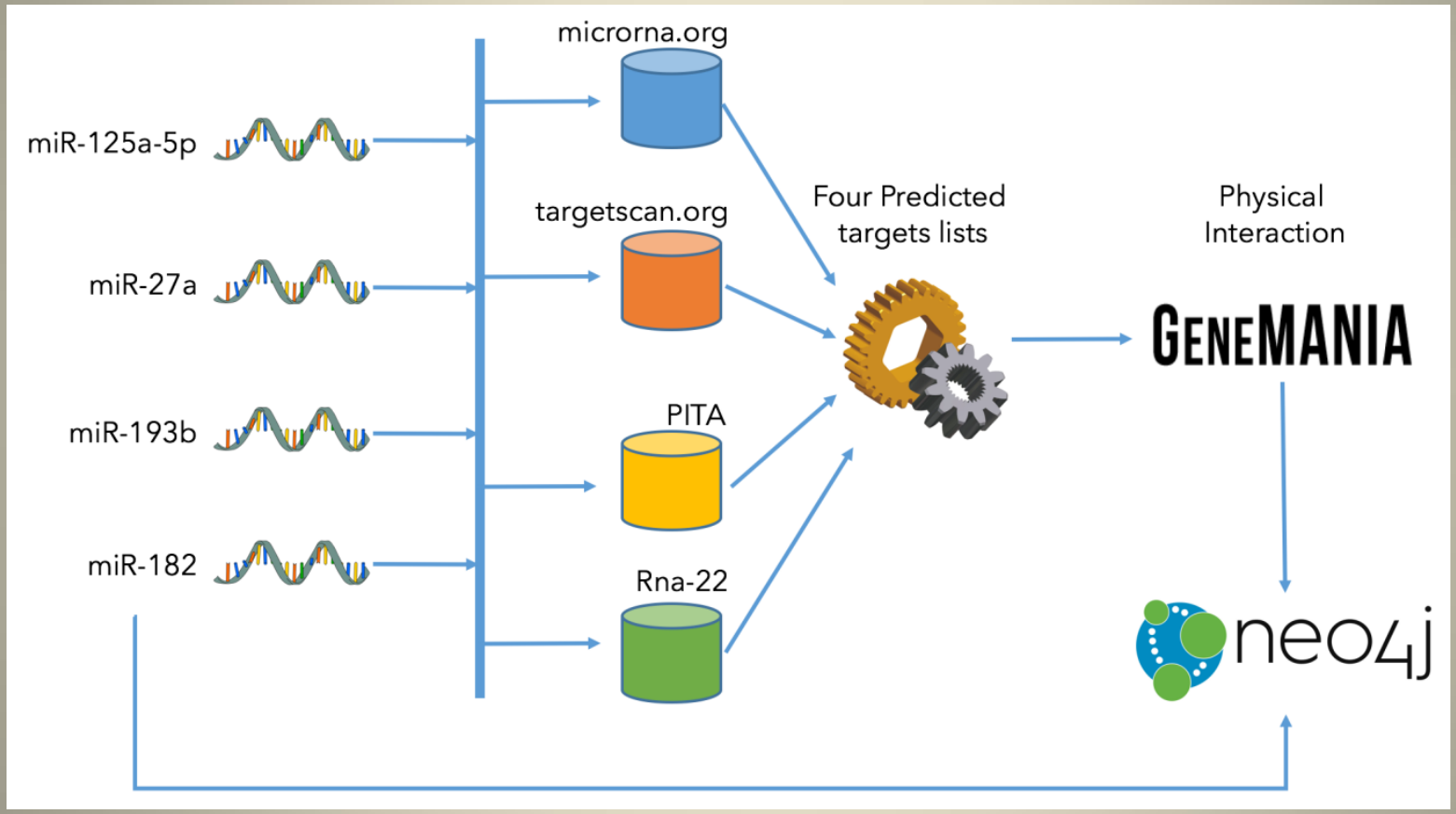
Relative quantification, ΔΔCt method

MiRNAs ANALYSIS

	RQ 3 M (DEN/CTR)	RQ 6 M (DEN/CTR)	RQ 11 M (DEN/CTR)	RQ 11 M (DEN-T/CTR)	RQ 11 M (DEN-T/DEN)
125 a-5p	0,75	0,92	2,2	3,34	1,52
182	0,97	1,05	3,52	10,09	2,86
193 b	1	1,65	1,33	4,89	3,67
27 a	0,17	15,72	4,27	4,89	1,11

Relative expression of miRNAs considered for the analysis in livers and tumors. RQ value is the relative quantification of miRNA expression, depending on the treatment's length (3, 6, 11 months), obtained by comparing hepatic samples from DEN-treated (DEN), controls (CTR), and tumor samples (DEN-T) from DEN-treated mice, as indicated in the round brackets. Results are mean of 3 iterations.

WORKFLOW OF THE miRNAs ANALYSIS



TARGET PREDICTION

MIRANDA

mirSVR score ≥ -1.2
"Good miRSVR score, non-conserved miRNA" table, 29 genes

TARGETSCAN

PcT > 0.1
"Conserved Family Info Results" table, 148 genes

PITA

ΔG < 10kcal/mol
Data from "3/15 flank All"table, "Mouse" column in "PITA Targets Catalog" ,Final table, 91 genes

RNA-22

ΔG < -25kcal/mol
Final table, 178 genes

PREDICTION TOOLS RESULTS

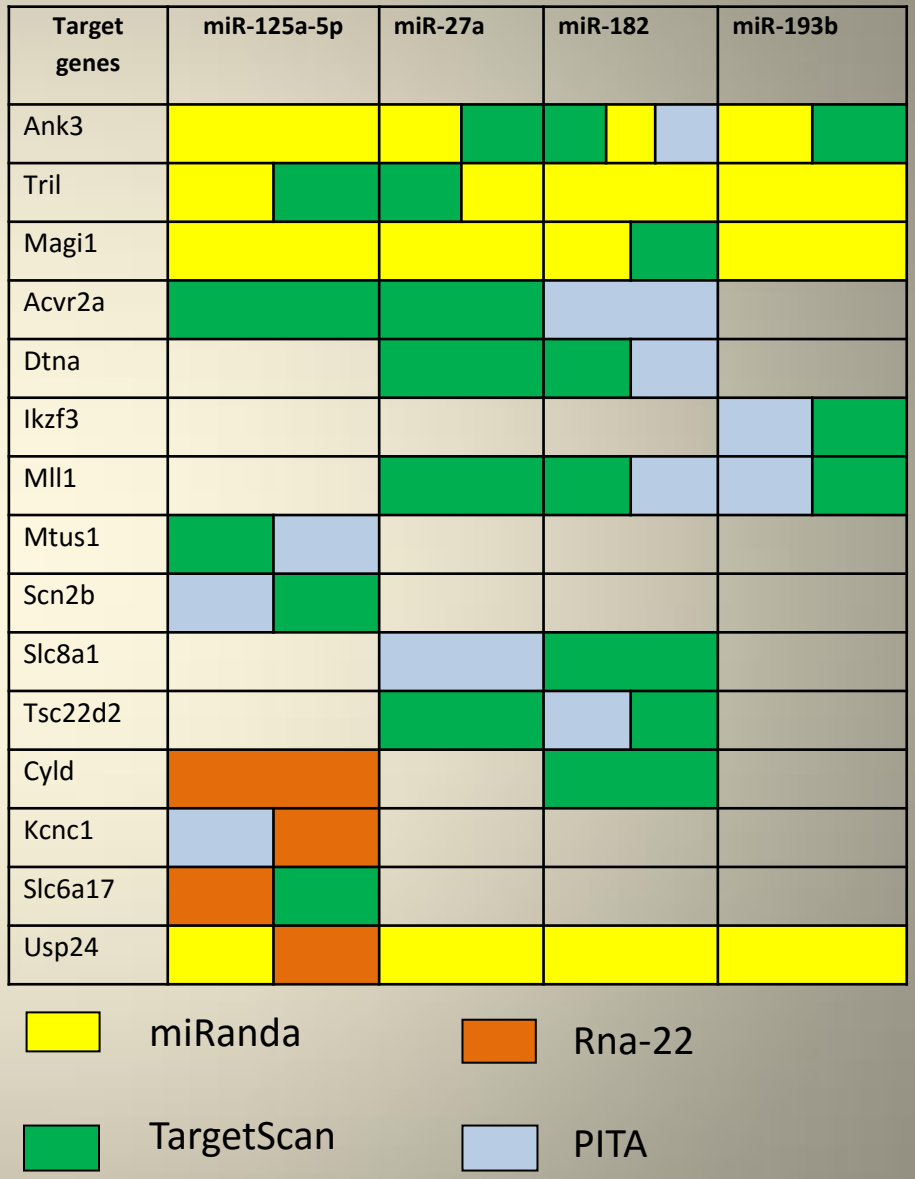
- Only genes predicted

by 2 out of the 4 used tools

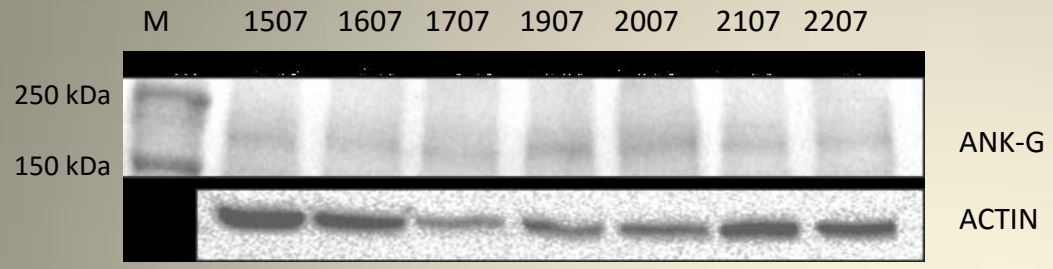
Totally 15 genes, “top-target genes”

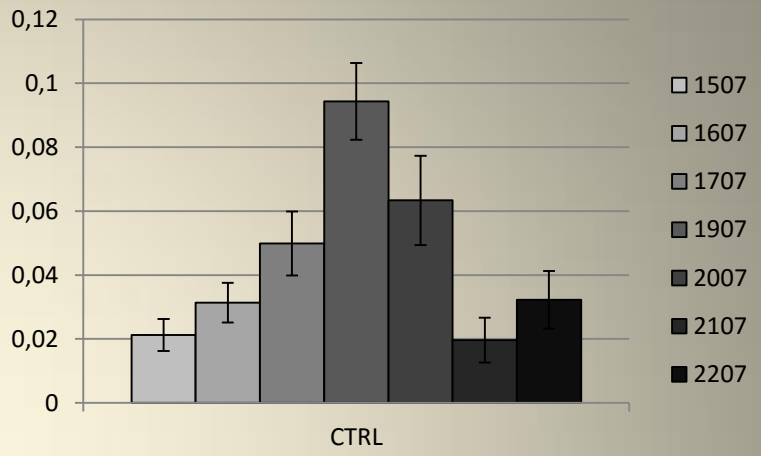
ANK3 was the only gene predicted by three algorithms (MiRanda, TargetScan, PITA)

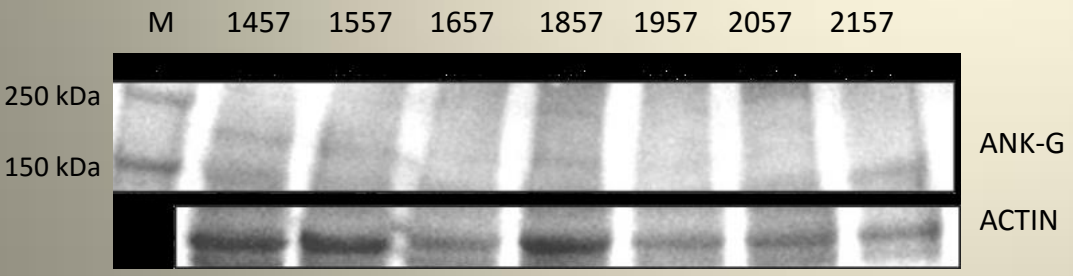
14 genes predicted by two algorithms

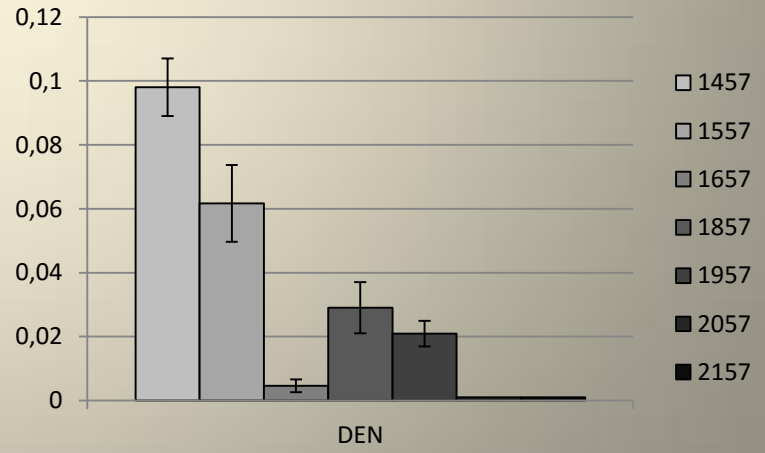


ANK-G EXPRESSION IN LIVER TISSUES

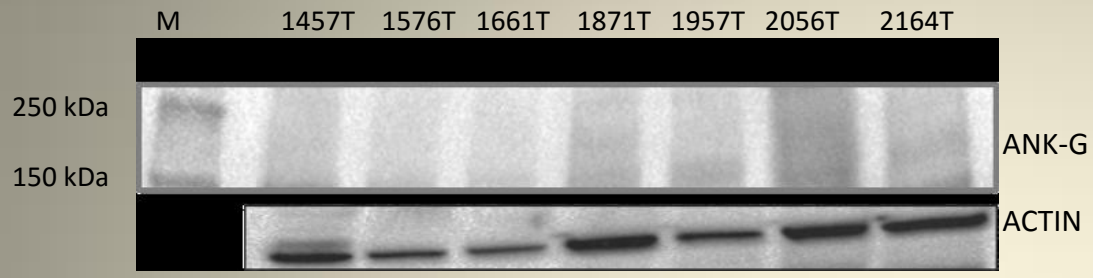


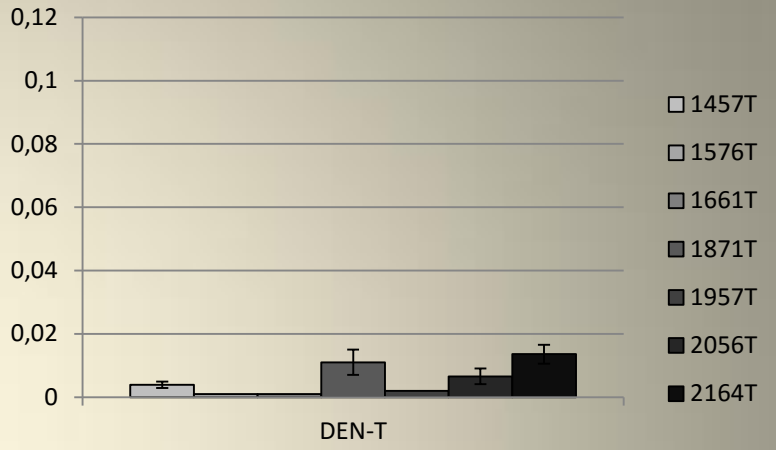






ANK-G EXPRESSION IN LIVER TISSUES





4 isoforms of ANK-G in mouse: 200, 170, 120 and 105 kDa

170 kDa ANK-G expression levels in 11-months control > 11-months DEN-treated group > 11-months tumour group.

These data correlates with the expression profiles of selected miRNAs, they were up-regulated in 11-months DEN-treated group and paired tumors compared with control group.

Validation of *in silico* data

ENRICHMENT ANNOTATION ANALYSES

Genemania v 3.1.2

Selected parameters:

1) Weighing method = molecular functions

2) "Number of genes" field
   - correlated genes = 100
   - correlated attributes = 20

"Interactions" field = only physical

15 top target genes → 26 significant clusters (FDR ≤ 0.05)

ion transporter activity, regulation of receptor protein serine/threonine kinase signaling pathway, protein import into nucleus, regulation of intracellular protein transport, regulation of cell adhesion, growth factor binding, positive regulation of pathway-restricted SMAD protein phosphorylation

ENRICHMENT ANNOTATION ANALYSES

After Genemania analysis, a set of 59 “secondary genes” interacting with top target genes was retrieved.

Selected parameters:

1) Weighing method = molecular functions

2) “Number of genes” field
   - correlated genes = 10
   - correlated attributes = 10

“Interactions” field = only physical

59 target genes → 20 significant clusters (FDR ≤ 0.05)

transmembrane receptor protein serine/threonine kinase signaling pathway, positive regulation of pathway-restricted SMAD protein phosphorylation, BMP signaling pathway, sequence specific DNA-binding RNA polymerase II transcription factor activity, occluding junctions, extrinsic apoptotic signaling pathway, regulation of necroptotic process

ENRICHMENT ANNOTATION ANALYSES

33 genes found to be described in hepatocarcinogenesis after a literature search were submitted to Genemania analysis. They derived from the set of 59 secondary genes.

Selected parameters:

1) Weighing method = molecular functions

2) “Number of genes” field → correlated genes = 10

2) “Number of genes” field → correlated attributes = 10

“Interactions” field = only physical

33 target genes → 14 significant clusters (FDR ≤ 0.05)

transmembrane receptor protein serine/threonine kinase signaling pathway,
positive regulation of pathway-restricted SMAD protein phosphorylation.

ENRICHMENT ANNOTATION ANALYSES

“Regulation of TGF-beta/SMAD signaling pathway” was among the most important and recurrent cluster after analyses.

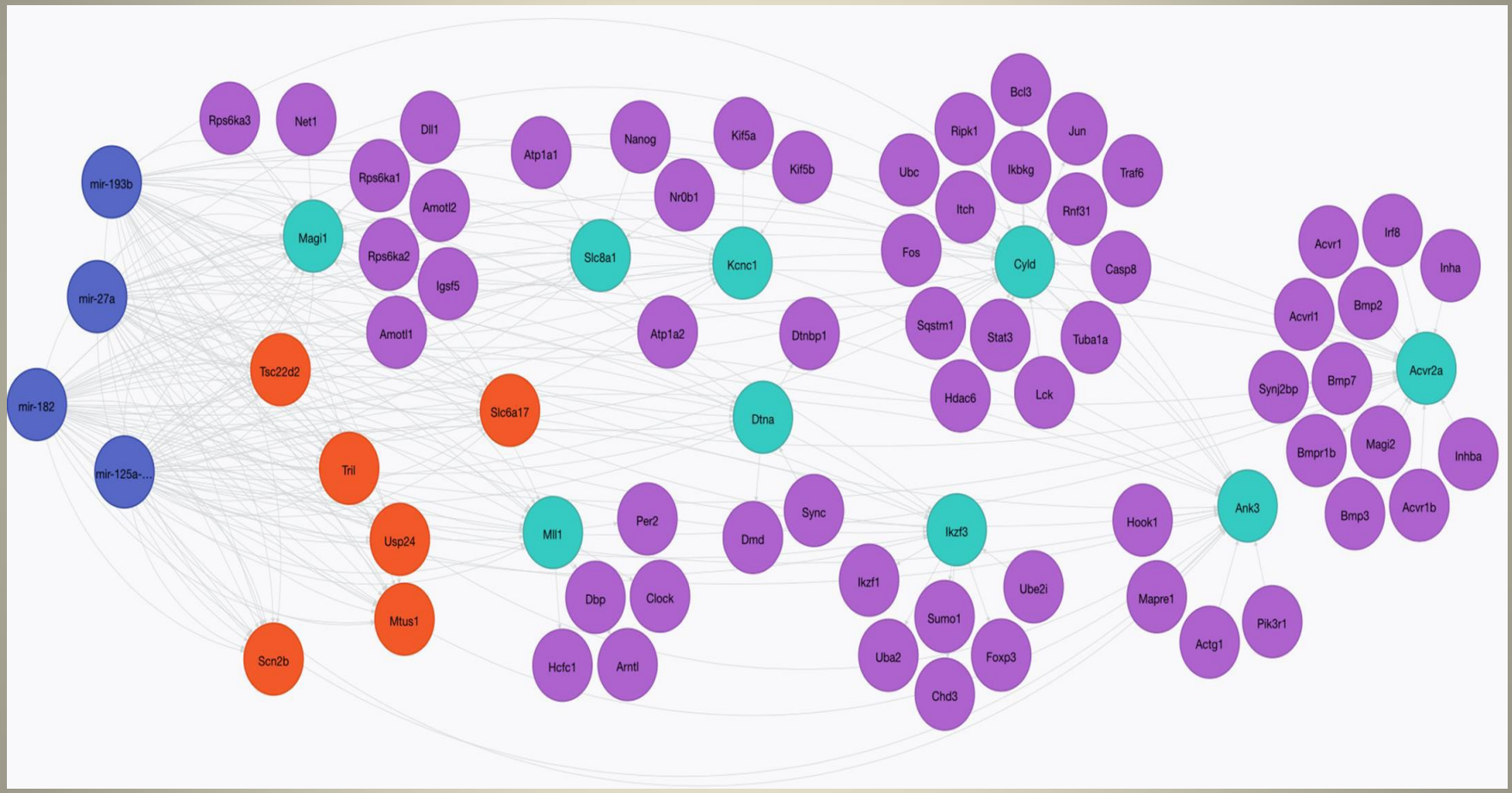
It was present in:

- 1 cluster in the 15 top target genes analyses
- 6 clusters in the 59 secondary genes analyses
- 7 clusters in the 33 liver-described genes analyses

Creation of an interaction network including selected miRNAs and putative target genes (top target and secondary genes)

Graphical visualization through Neo4J technology

INTERACTION NETWORK



DISCUSSIONS

ANK-3 expression is found to be down-regulated in multiple tumours (Glinski GV,Berezovska O, Glinskii , *J. Clin. Invest.* 2005).

Involved in the regulation of “anoikis” sensitivity and epithelial-to-mesenchymal transition (EMT) through direct link with HOOK1. (Kumar S, *Molecular and Cellular Biology*, 2011; Weimer JM, Biochem Biophys Res Commun, 2005)

Predicted target genes for miR-182 (miRanda, TargetScan, PITA), its protein, ANK-G is progressively reduced from control tissues to peritumour and tumours in our samples

miR-182 up-regulated in 11-months tumoral tissues compared to control

miR-182 up-regulation in tumoural tissue (Wang C, *Chin J Cancer res*, 2014; Wang J, *BMC Cancer*, 2012)

MiR-182 ⊣ ANK3 mRNA in HCC possible interaction!

DISCUSSIONS

CYLD is a tumour suppressor gene responsible for the onset of cylindromatosis.

It is reported to be under-expressed in HCC. (Kinoshita H, *Mol Clin Oncol*, 2013)

It is involved in the regulation of proliferation and survival pathways through deubiquitination of key molecules such as TNF receptor-associated factor 2/6 (TRAF2/6) and B-cell limphoma-3 (BCL-3).

Candidate gene for miR-125a-5p (RNA-22) and miR-182 (TargetScan) in our system

Validated as a target for human miR-182 (miRTarBase database)

Both these miR were up-regulated in 11-months tumour compared to controls.

MiR-125a-5p ⊣ CYLD mRNA

miR-182 ⊣ CYLD mRNA

mir-125a-5p + miR-182 ⊥ CYLD mRNA

DISCUSSIONS

SLC8A1 → NCX1 → regulation of intracellular calcium levels.

Its expression is reported to be decreased in some human cancers (Munoz JJ, *J Urol*, 2015) and in multidrug resistance phenomena (Januchowski R, *Biomed Pharmacother*, 2014).

Predicted target gene for miR-182 (TargetScan) and miR-27a (PITA).

miR-182 and miR-27a are up-regulated in 11-months tumour tissues compared to control in our system and they are described up-regulated in HCC (Wang J, *BMC Cancer*, 2012; He XX, *Mol Biosyst*, 2015).

miR-182 ⊣ SLC8A1 mRNA in HCC

miR-27a ⊣ SLC8A1 mRNA in HCC

miR-182 + miR-27a ⊣ SLC8A1 mRNA in HCC

DISCUSSIONS

MAGI1 is a tumour suppressor gene, involved in regulation of cell-cell contacts through PTEN action.

It is described to be under-expressed in HCC (Zhang G, *J Invest Surg,* 2012).

Predicted target for all of the selected miRNAs, in particular for miR-182.

MiR-125a-5p, -182,-193b and -27a are all up-regulated in 11-months tumour group with respect to controls in our system.

They could act singularly or in combination between them to decrease MAGI1 mRNA expression in HCC tissues.

MTUS1 is reported to be down-regulated in multiple tumour types (Zhao T, *BMC Cancer*, 2015) and frequently mutated in HCC (Di Benedetto M, *Mol Cell Endocrinol*, 2006).

Remaining genes (TRIL, ACVR2A, DTNA, IKZF3, MLL1, SCN2B, TSC22D2, KCNC1, SLC6A17 and USP24) are not described to be related to HCC in literature.

CONCLUSIVE REMARKS

MicroRNAs remain misterious yet unknowable tools but they hold a great potential for their use in medicine.

Their use in clinical environments needs to be promoted.

The work here presented pointed to the integration between *in vivo* data and *in silico* prediction through a multidisciplinary approach.

The followed approach revealed to be productive and it led to the generation of results that can be considered a starting point for further studies in the field.

TAKE HOME MESSAGES

The application of a combined approach of bioinformatic and molecular techniques resulted in the production of significant data to further deepen in the context of liver carcinogenesis.

Through the use of multiple target prediction tools, it was possible to obtain a list of target genes related to a starting set of selected miRNAs. Validation studies on one of the targets evidenced good applicability potential for the procedure.

Enrichment annotation tools provided significant clusters in which canalize additional analyses.

Sequential application of target prediction and enrichment annotation tools represents a reliable approach in the study of microRNAs set.

Molecular Oncopathology Lab

Prof. Edoardo Alesse

Dr. Alessandra Tessitore

Prof. Francesca Zazzeroni

Dr. Agata Gaggiano

Dr. Germana Cicciarelli

Dr. Davide Vecchiotti

Dr. Maria Fausta Fischietti

Dr. Daniela Verzella

Dr. Valentina Mastroiaco

Dr. Barbara Di Francesco

Department of Computer Engineering and Science and Mathematics

Dr. Francesco Gallo

Prof. Antinisca Di Marco

Dr. Pasquale Caianiello

Thank you for the attention!