Bioinformatics for Biochemistry and Molecular Biology

Hsueh-Fen Juan
(阮雪芬)
05/20/2010

Email: yukijuan@ntu.edu.tw

Outline

- Introduction
- Construction of gene/protein network
- Drug Discovery
- Mechanism study

Introduction

Systems Biology

- **Definition**: Quantitative study of biological processes as whole systems instead of isolate parts.
- **Goal**: Construction and experimental validation of models that explain and predict the behavior of biological systems -2001, 2nd International Conference on Systems Biology

Why Do We Need Systems Biology?

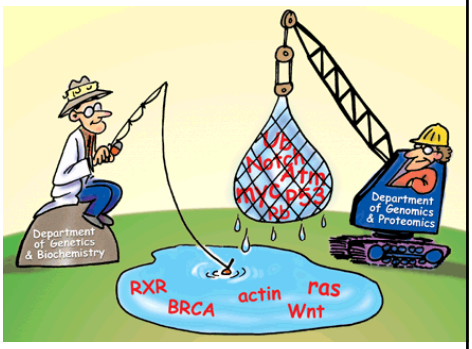
- The map of the genome is just the rule book; "systems biology" is the ball game.
  - by Alexandra Stikeman, Technology Review, 2002 March

Recent Developments of Systems Biology in US

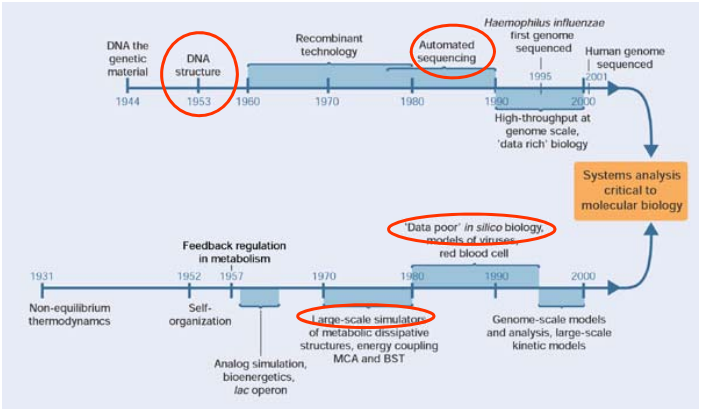
- In 2003, MIT received a 5-year 16 million research grant from National Institute of Health to study systems biology.
- Harvard University created "Department of Systems Biology" in 2003.

Traditional Biology & Systems Biology

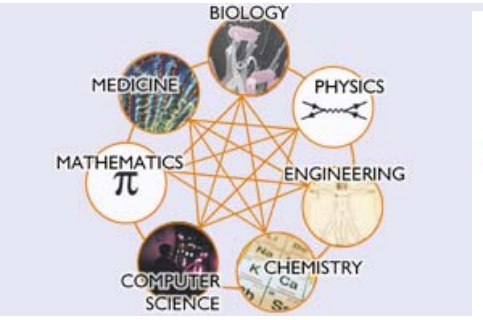
- Traditional biology :
  - Single gene or protein
- Systems biology:
  - Simultaneously study the complex interaction of many levels of biological information to understand how they work together



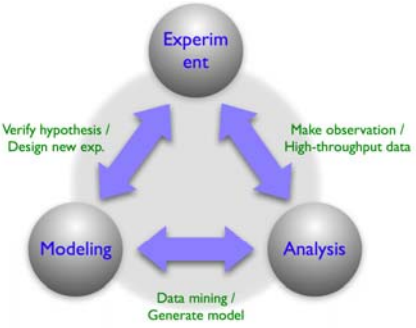
The Evolution of Molecular Biology into Systems Biology



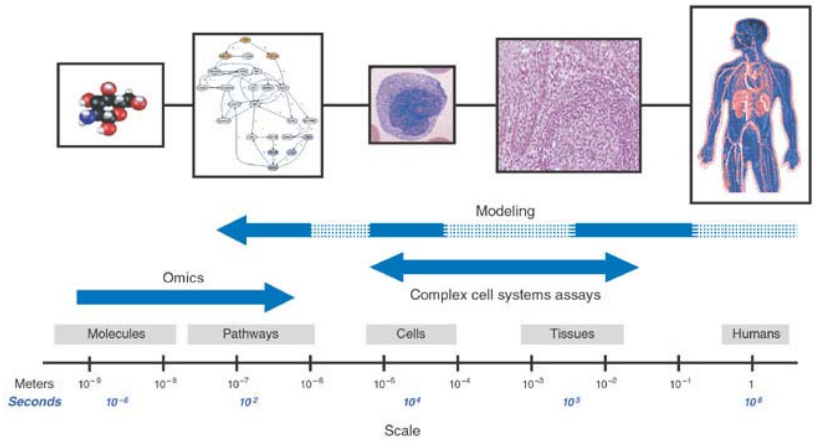
Systems Biology is Integrative Biology



Leroy Hood, ISB

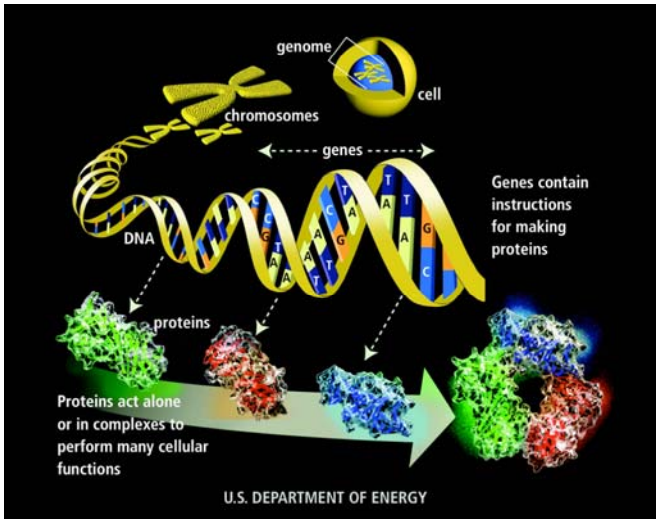


Approaches to Systems Biology in the Pharmaceutical Industry

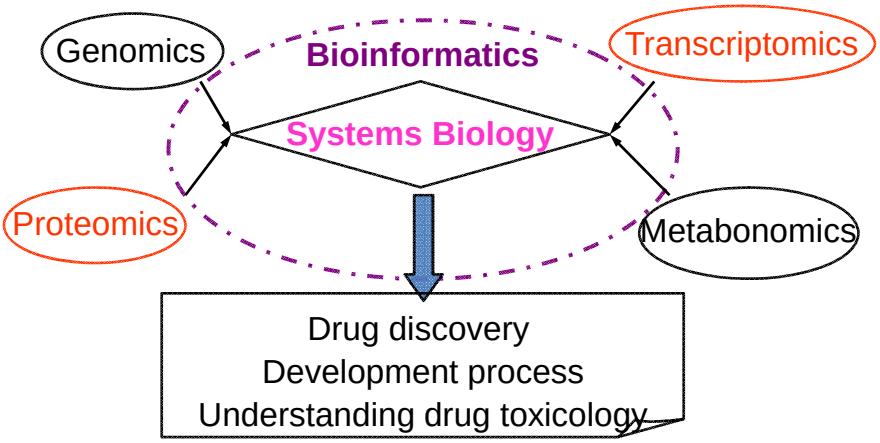


Nature Biotechnology 2004

Basics of Biology



Systems Biology & Omics Data



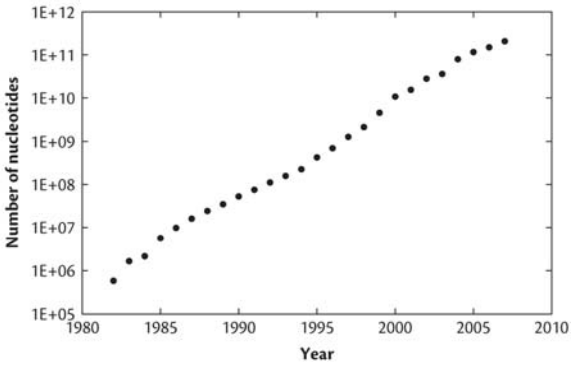
Traditional and Current Biology

- Traditionally, biology has been an observational science.
- Now, biology has been converted into deductive science.

The Data of Bioinformatics

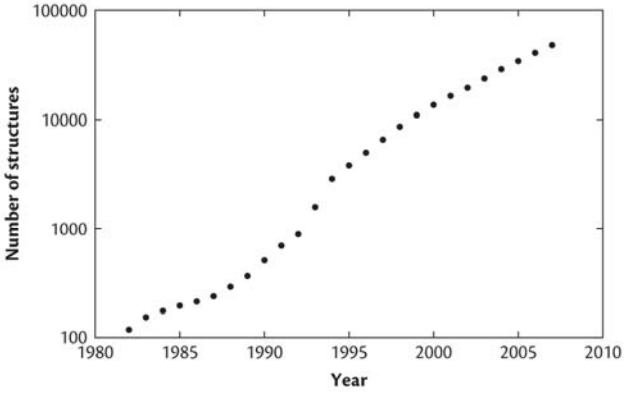
- Very very large amount
- Nucleotide sequence databanks contain 1.7×10^{12} bases
- The full three-dimensional coordinates of proteins of average length ~400 residues > 50000 entries
- Not only are the individual databanks large, but their sizes are increasing as a very high rate.

Growth of the International Nucleotide Sequence Database Collection



Log scale on y-axes

Growth of the world-wide Protein Data Bank



Log scale on y-axes

The History

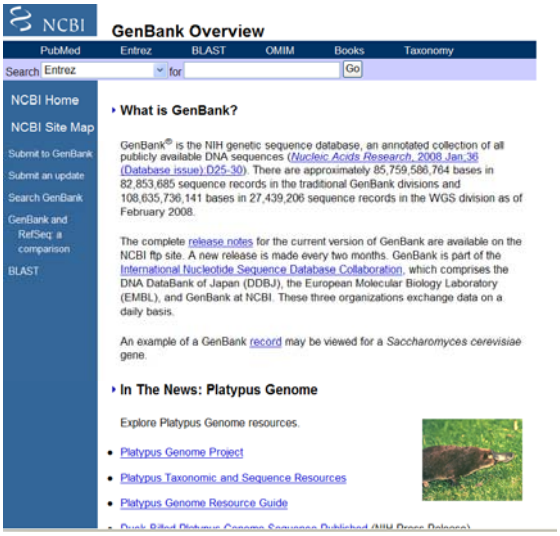
- Sequence database began in the early 1960s, when Margaret Dayhoff and colleagues at the Protein Information Resource (PIR) collected all of the protein sequences known at that time.
- *Atlase of Protein Sequence and Structure*

PIR: http://pir.georgetown.edu/

The History

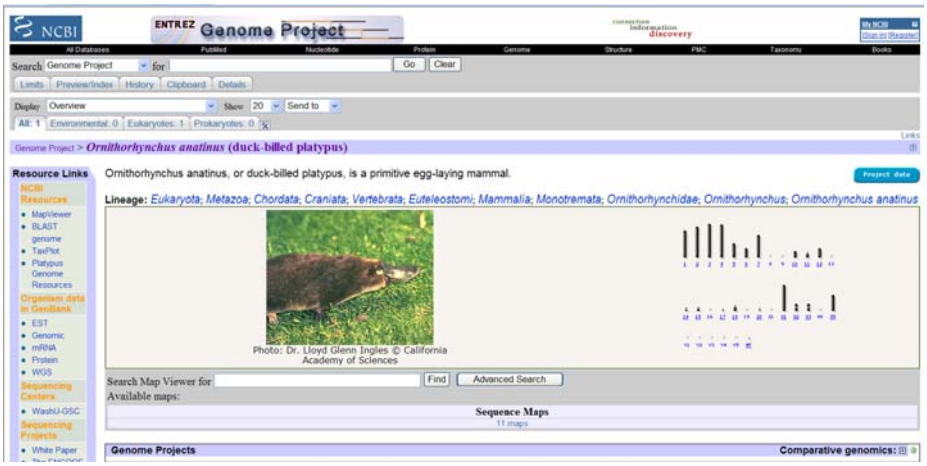
- The advent of DNA sequence databases in 1982, initiated by the European Molecular Biology Laboratory (EMBL) and joined shortly thereafter by GenBank, led to the next phase in the history of sequence databases

GenBank



http://www.ncbi.nlm.nih.gov/Genbank/

Platypus Geneome



NCBI **GenBank Overview**

PubMed Entrez BLAST OMIM Books Taxonomy Structure

Search Entrez for Go

NCBI
SITE MAP

Submit to GenBank

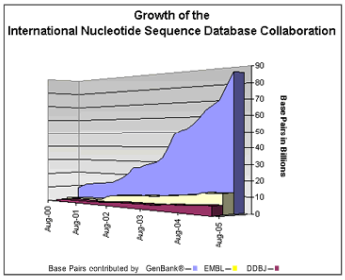
Updates

Search GenBank
Entrez Nucleotide

BLAST

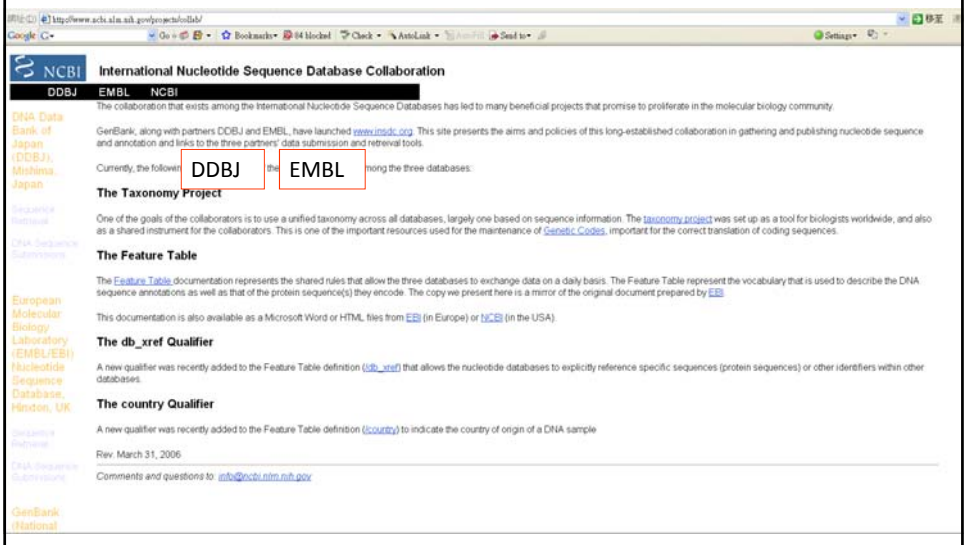
International sequence databases exceed 100 gigabases

In August 2005, the INSDC announced the DNA sequence database exceeded 100 gigabases. GenBank is proud of its contributions toward this milestone. We thank all the scientists who have worked through the submission process at GenBank and made their sequence data available to the world. See the related press release.



We look forward to working with you all in the future to continue this tradition as the database continues to grow exponentially.

International Nucleotide Sequence Database Collaboration



NCBI **International Nucleotide Sequence Database Collaboration**

DDBJ EMBL NCBI

The collaboration that exists among the International Nucleotide Sequence Databases has led to many beneficial projects that promise to proliferate in the molecular biology community.

DNA Data Bank of Japan (DDBJ), Mishima, Japan

GenBank, along with partners DDBJ and EMBL, have launched www.insdc.org. This site presents the aims and policies of this long-established collaboration in gathering and publishing nucleotide sequence and annotation and links to the three partners' data submission and retrieval tools.

Currently, the follower DDBJ the EMBL hong the three databases:

The Taxonomy Project

One of the goals of the collaborators is to use a unified taxonomy across all databases, largely one based on sequence information. The taxonomy project was set up as a tool for biologists worldwide, and also as a shared instrument for the collaborators. This is one of the important resources used for the maintenance of Genetic Codes, important for the correct translation of coding sequences.

The Feature Table

The Feature Table documentation represents the shared rules that allow the three databases to exchange data on a daily basis. The Feature Table represent the vocabulary that is used to describe the DNA sequence annotations as well as that of the protein sequence(s) they encode. The copy we present here is a mirror of the original document prepared by EBI.

European Molecular Biology Laboratory (EMBL/EBI) Nucleotide Sequence Database, Hinxton, UK

This documentation is also available as a Microsoft Word or HTML files from EBI (in Europe) or NCBI (in the USA).

The db_xref Qualifier

A new qualifier was recently added to the Feature Table definition (/db_xref) that allows the nucleotide databases to explicitly reference specific sequences (protein sequences) or other identifiers within other databases.

The country Qualifier

A new qualifier was recently added to the Feature Table definition (/country) to indicate the country of origin of a DNA sample.

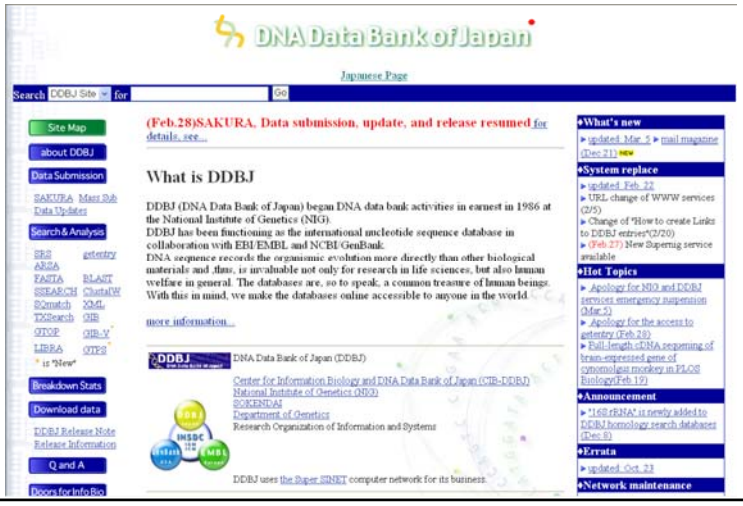
Rev. March 31, 2006

Comments and questions to: info@ncbi.nlm.nih.gov

GenBank (National

DDBJ

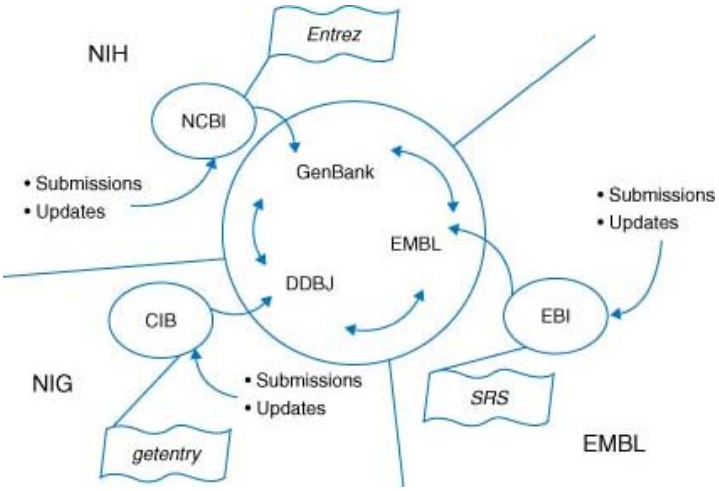
Joined the data-collecting collaboration a few years later



International Nucleotide Sequence Database Collaboration

- In 1988, there was an agreement to use a common format for data elements within a unit record and to have each database update only the records that were directly submitted to it.
- DDBJ/EMBL/GenBank records are updated automatically every 24 hours at all these sites.

Data Flow for New Submissions and Updates between the Three Databases



First Protein Sequence Database

- In 1980s, Swiss-Prot Protein Sequence Database was laid.

http://www.expasy.ch/sprot/

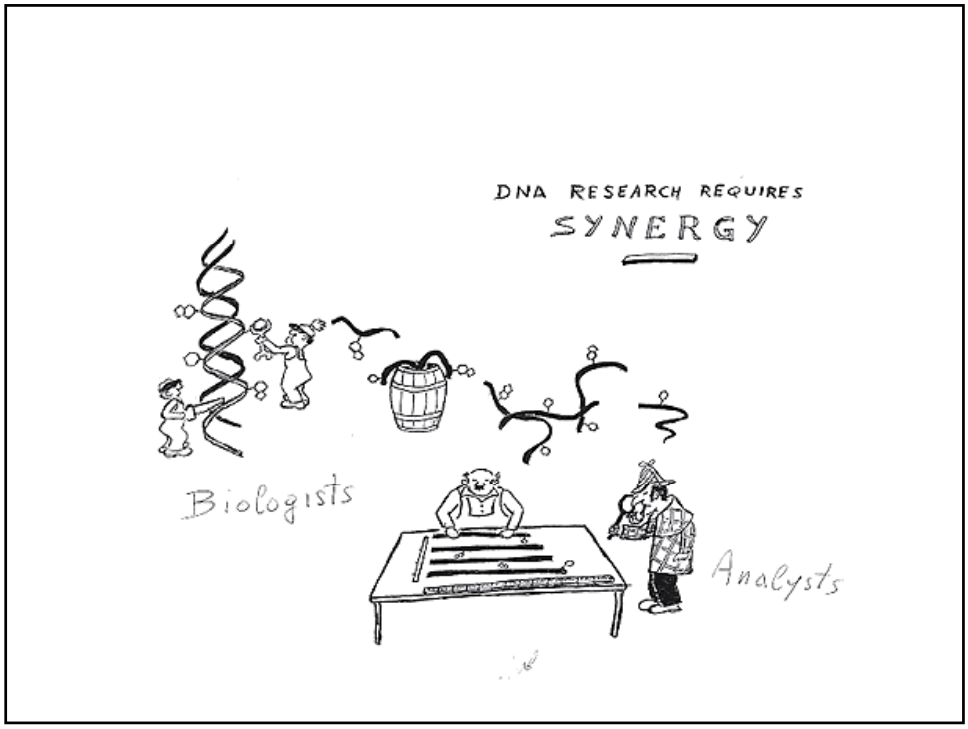
TrEMBL: Translation of EMBL nucleotide sequences

Goals

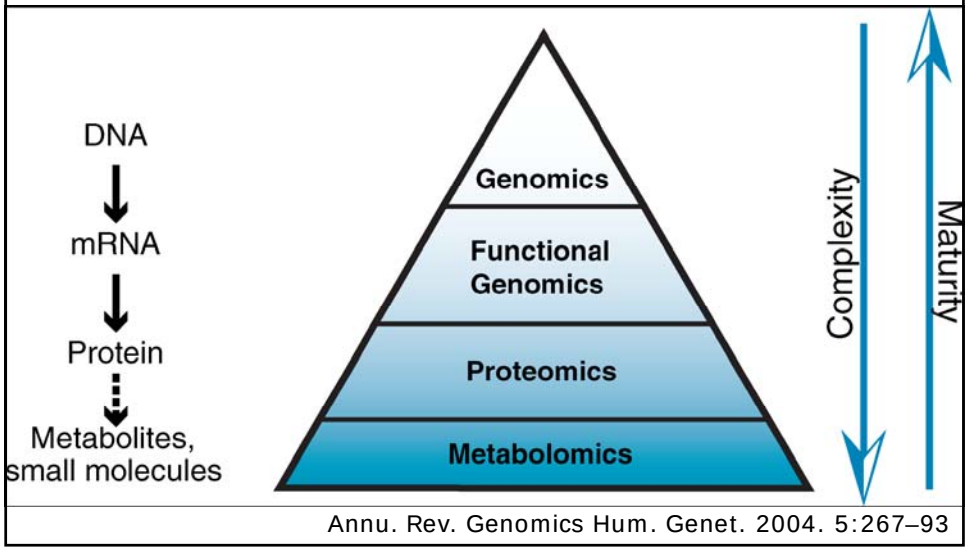
- "Saw life clearly and saw it whole"

→ Understand integrative aspects of the biology of organisms

- To interrelate sequence, three-dimensional structure, interactions, and function of individual proteins, nucleic acids and protein-nucleic acid complexes



Schematic of the Relationship between the Different "omics" Disciplines



Annu. Rev. Genomics Hum. Genet. 2004. 5:267–93

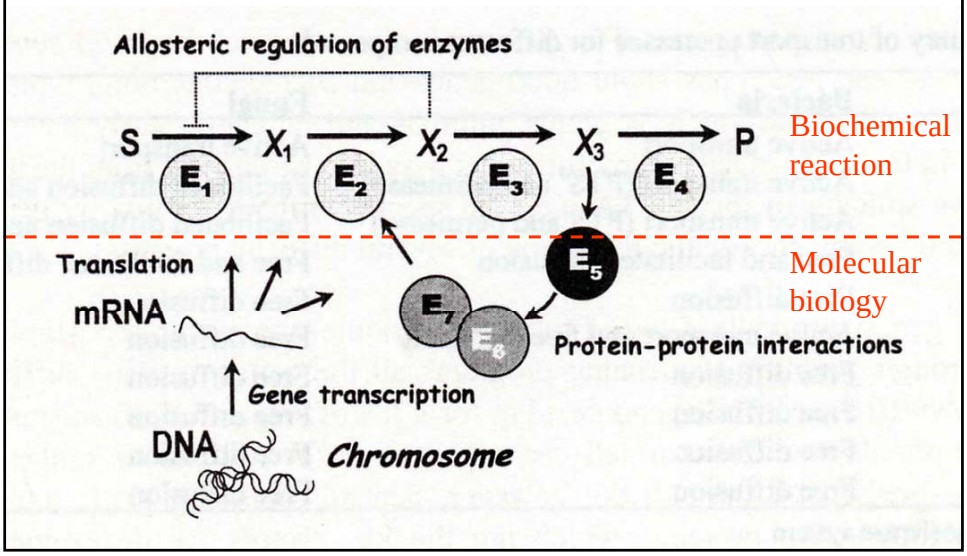
Protein-Protein Interactions & Biological Pathways

- Kanehisa (2000):

Interaction ⟶ Network ⟶ Function

Post-genome informatics (2000)

Processes Involved in Expressing a Specific Phenotype in a Cell

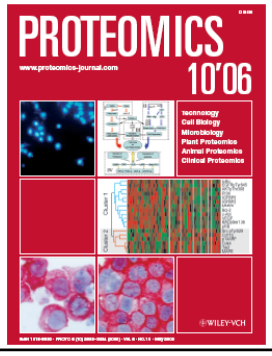


Protein-Protein Interactions

- Protein-protein interactions are intrinsic to every cellular process.
- Form the basis of phenomena
  - DNA replication and transcription
  - Metabolism
  - Signal transduction
  - Cell cycle control
  - Secretion

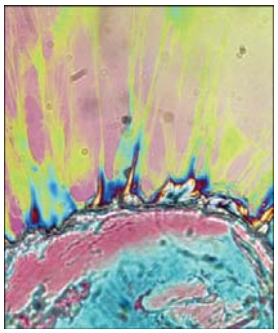
Construction of Gene/protein Network

- *Electrophoresis* (2002) **23**: 2490 -2504.
- *Bioinformatics (2004)* **20**: 3691-3693.
- *Proteomics* (2006) **6**, 2991-3000.



Cancer Robustness

- Viewing cancer as a robust system with potential points of fragility opens up new strategies for the development of drugs and therapies.

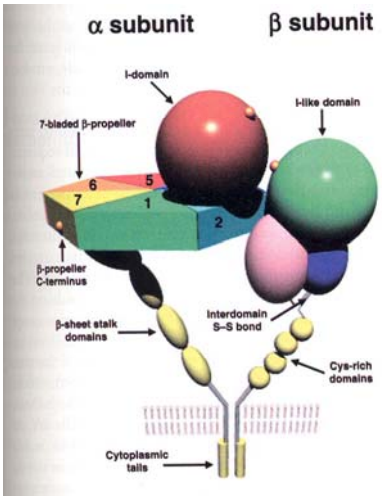


NATURE 2003, 426, 125.

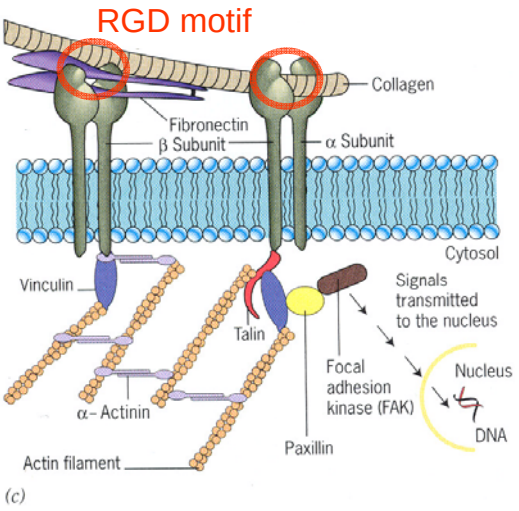
Hunt for fragility: weaknesses in tumor growth dynamics could yield new anti-cancer therapies.

Structure of Integrin

- Hynes in 1987 to emphasize the role of these RGD receptors in integrating the extracellular matrix outside the cell with the actin-containing cytoskeleton inside the cell.



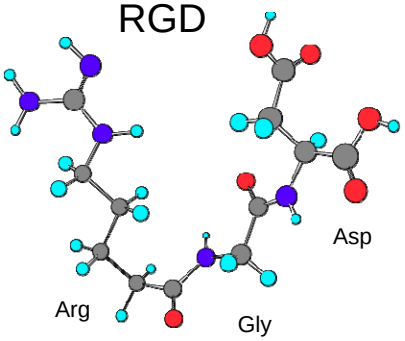
Interactions of Integrins w/ Other Proteins



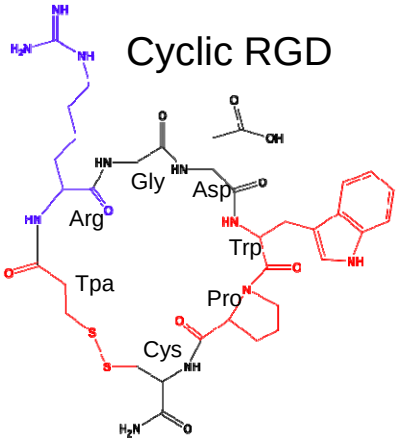
- Signal are presumably transmitted into the nucleus, where they stimulate the transcription of gene involved in cell growth and proliferation

A Novel Compound — Cyclic RGD

RGD

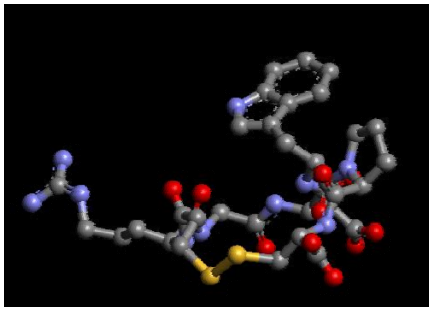


Cyclic RGD

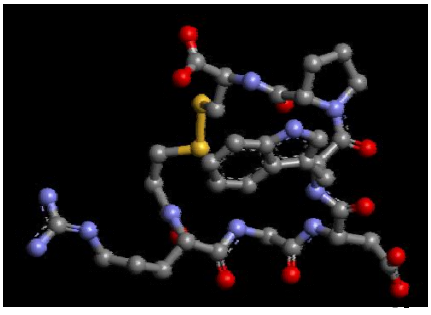


3D Structure of Cyclic-RGD

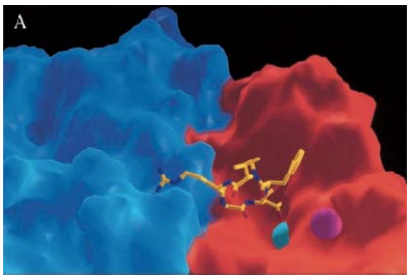
Cis form



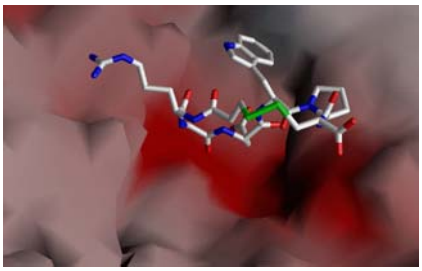
Trans form



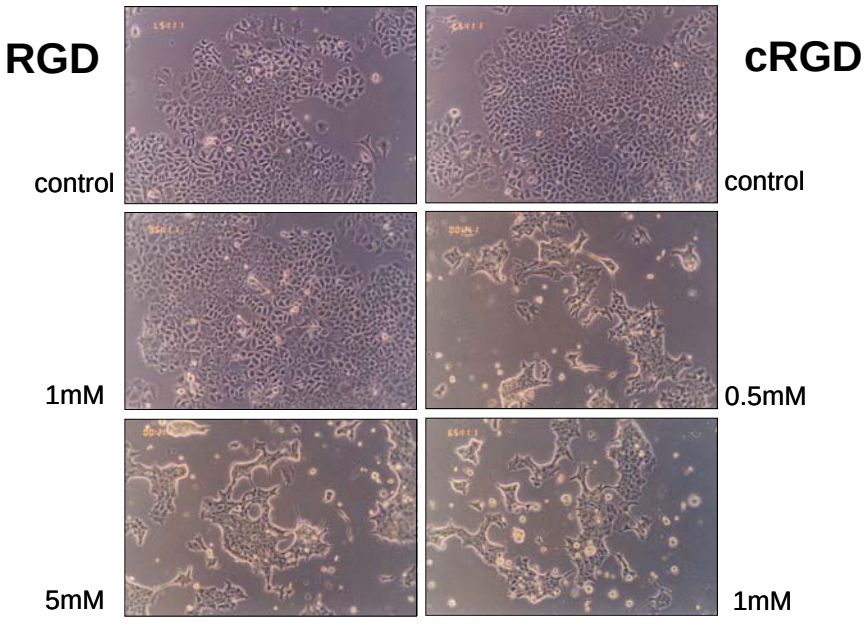
X-tal Structure: Integrin + RGD



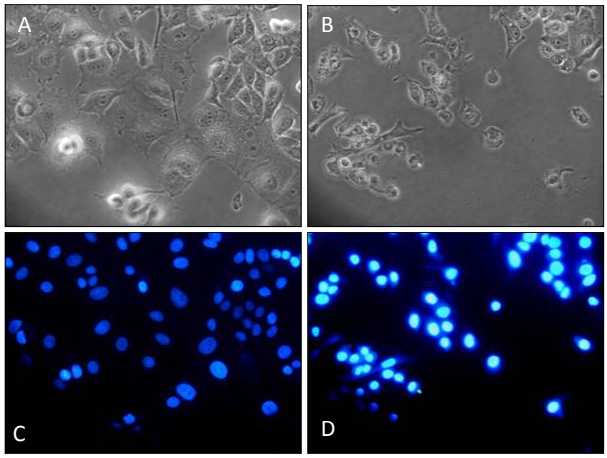
Science 296, 151 (2002)

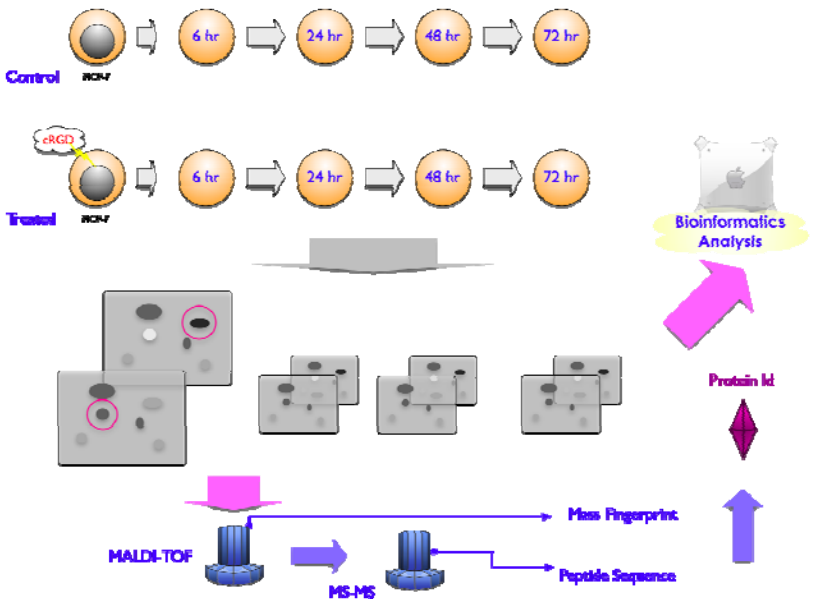


Only Trans-cRGD Fits

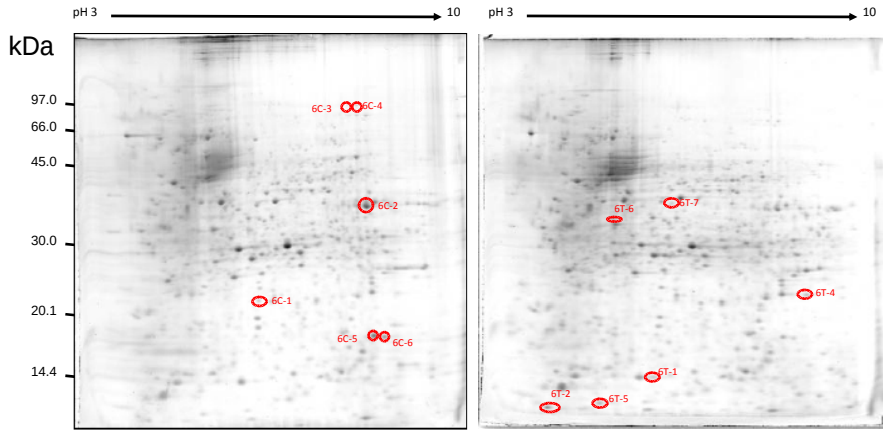


Characterization of the Cyclic RGD Induced Cell Death in Human MCF-7 Cells

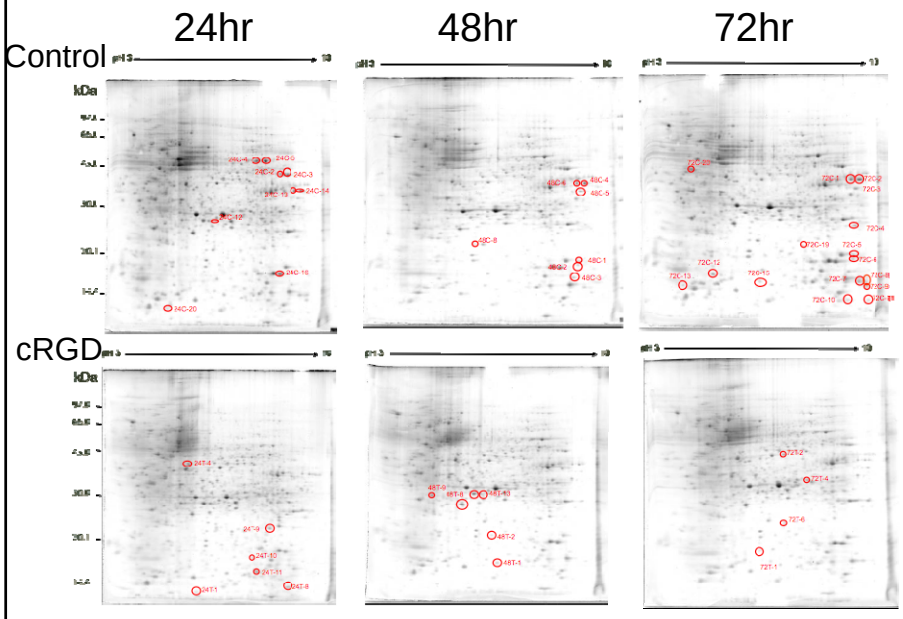




2DE Map: 6hr



Control cRGD



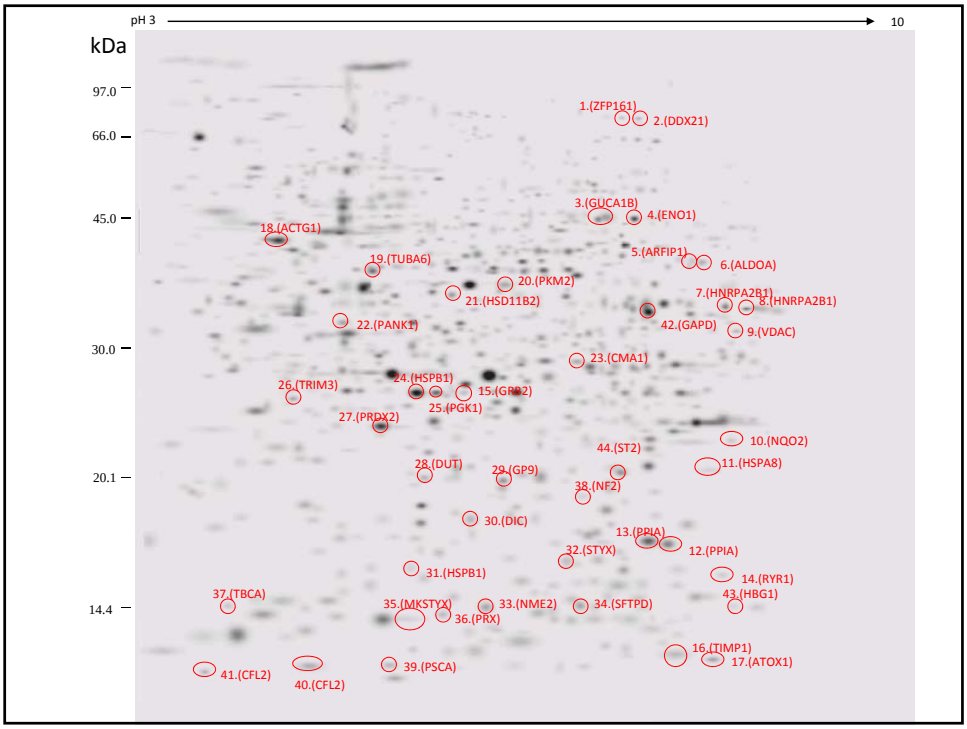
Protein Identification

Up-regulated proteins	
6T-1[a]	MKSTYX
6T-2[b]	CFL2
6T-4[b]	HSPA8
6T-5[b]	CFL2
6T-6[a]	PANK1
6T-7[a]	HSD11B2
24T-1[a]	PSCA
24T-4[b]	TUBA6
24T-8[a]	TIMP1
24T-9[a]	ST2
24T-10[a]	STYX
24T-11[a]	SFTPD
48T-1[b]	NME2
48T-2[b]	DIC
48T-8[b]	HSPB1
48T-9[b]	TRIM3
48T-13[b]	PGK1
48T-14[b]	PRDX2

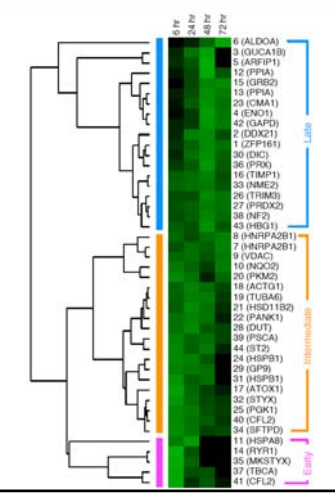
Down-regulated proteins	
6C-1[a]	DUT
6C-2[a]	GAPD
6C-3[a]	ZFP161
6C-4[a]	DDX21
6C-5[a]	PPIA
6C-6[a]	PPIA
24C-2[a]	ARFIP1
24C-3[b]	ALDOA
24C-4[a]	GUCA1B
24C-5[a]	ENO1
24C-12[b]	GRB2
24C-13[b]	HNRPA2B1
24C-14[b]	HNRPA2B1
24C-16[b]	PPIA
24C-20[a]	TIAF1
48C-1[b]	NME2
48C-2[b]	UBE2L3

48C-3[a]	RYR1
48C-4[b]	HNRPA2B1
48C-5[b]	VDAC
48C-6[b]	HNRPA2B1
48C-8[a]	DUT
72C-1[b]	HNRPA2B1
72C-2[b]	HNRPA2B1
72C-3[a]	VDAC
72C-4[a]	NQO2
72C-5[b]	NME1
72C-6[a]	TIMM8A
72C-7[b]	HRG1
72C-8[b]	NPM1
72C-9[a]	TIAF1
72C-10[a]	ATOX1
72C-11[a]	RNASE8
72C-12[a]	HSPE1
72C-13[a]	TBCA

72C-15[a]	PRX
72C-19[a]	NF2
72C-20[a]	ACTG1



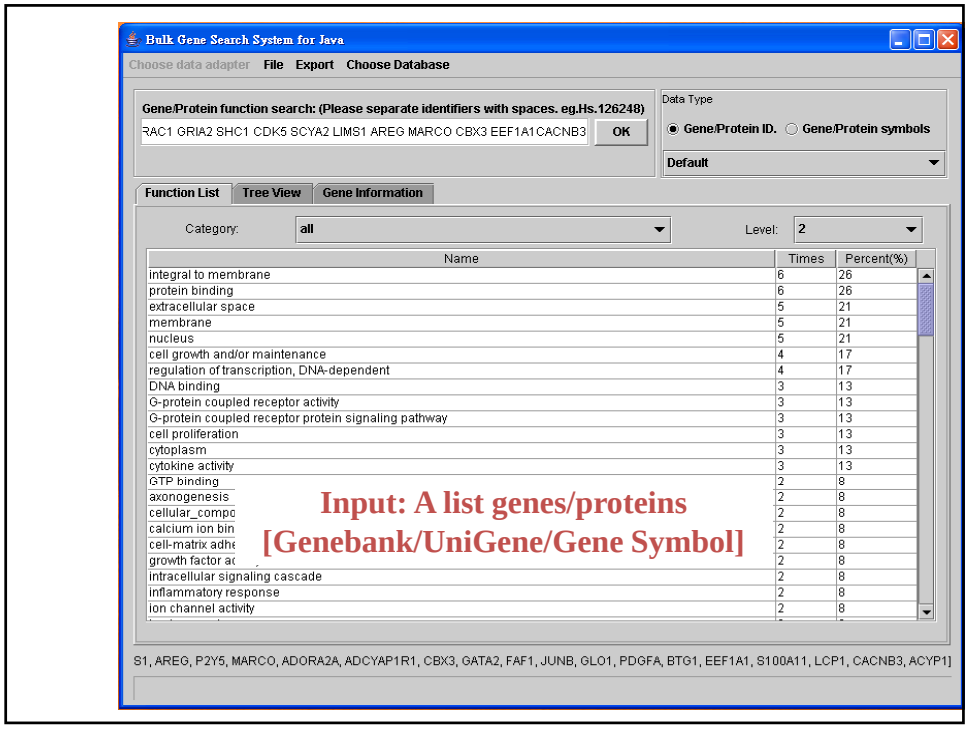
Hierarchical Clustering Analysis of Cyclic RGD-induced Protein Profile of MCF-7 cells

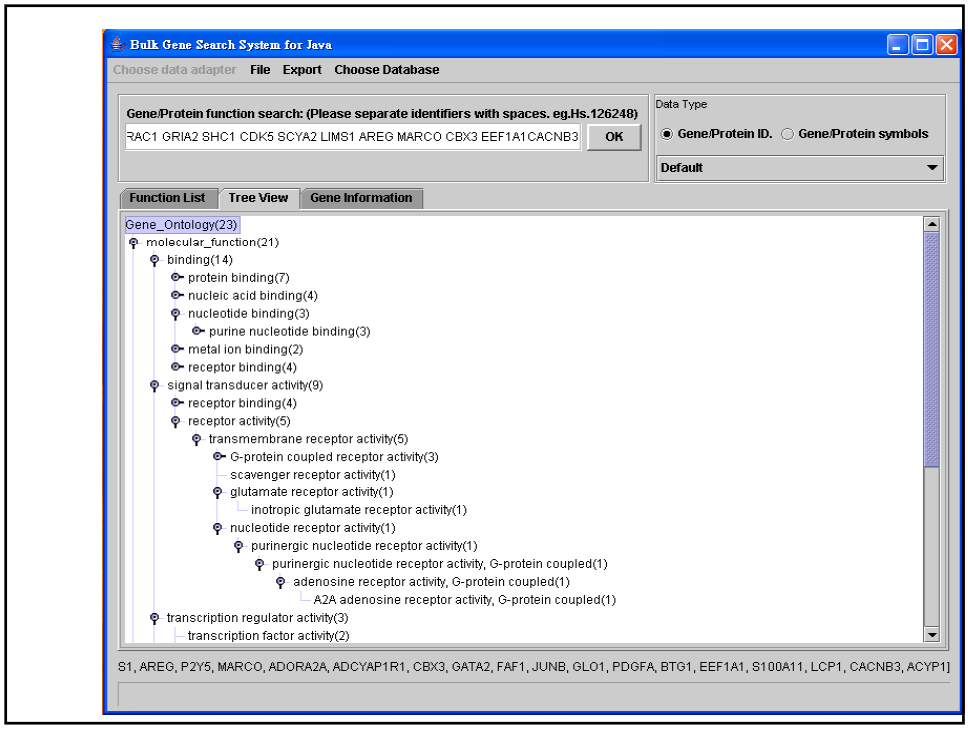


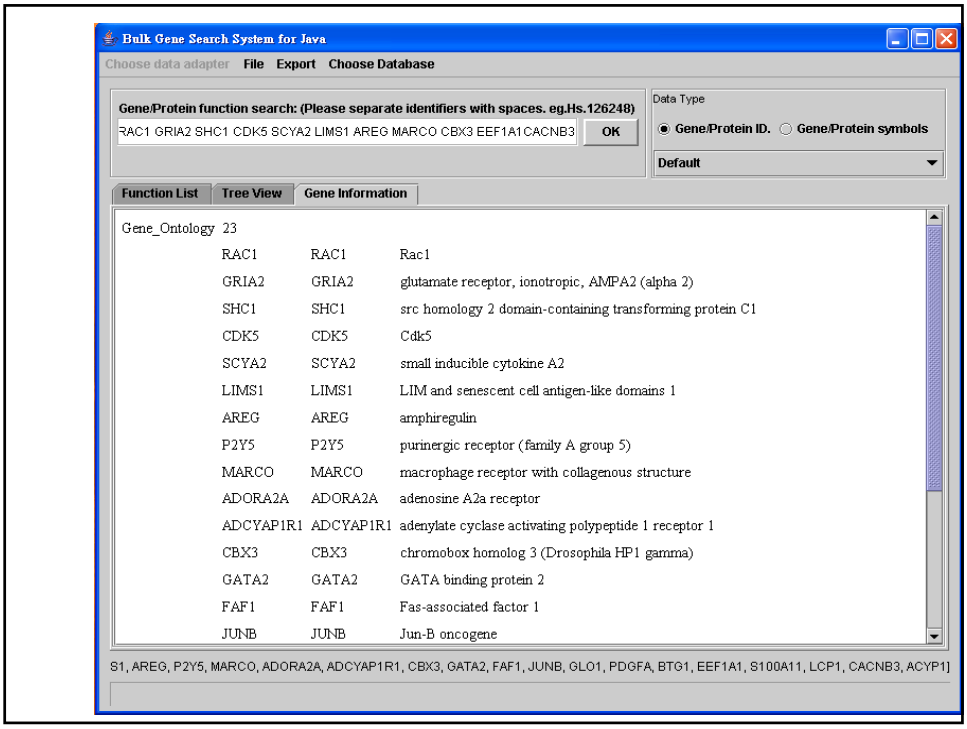
Gene Annotation

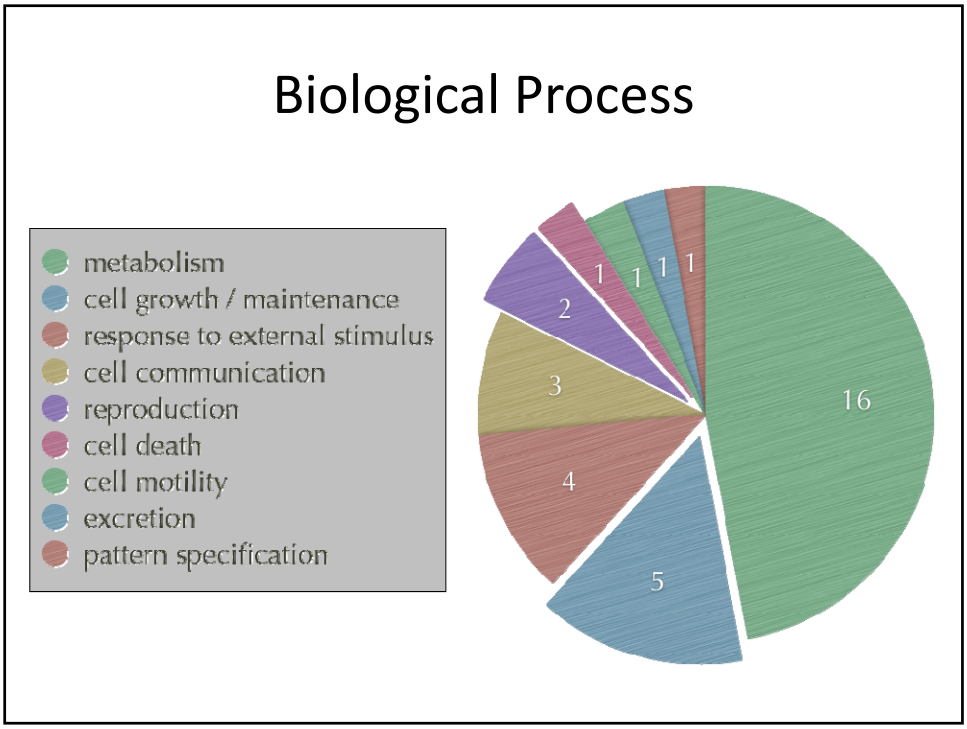
- Gene Ontology Consortium
  - Biological Process
  - Cellular Component
  - Molecular Function
- BGSSJ
  - *http://bgssj.sourceforge.net/*
  - Automatic functional classification
  - JAVA application (requires J2RE 1.3)

Microarrays: Methods and Protocols, The humana press inc. (2007)

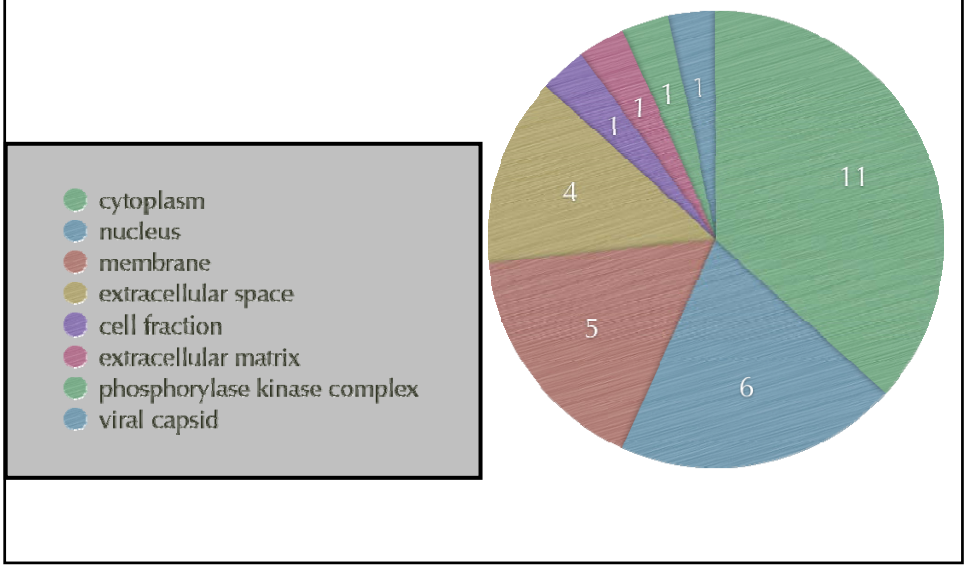

Bulk Gene Search System for Java
Choose data adapter File Export Choose Database
Gene/Protein function search: (Please separate identifiers with spaces. eg.Hs.126248)
RAC1 GRIA2 SHC1 CDK5 SCYA2 LIMS1 AREG MARCO CBX3 EEF1A1 CACNB3
OK
Data Type
Gene/Protein ID.
Gene/Protein symbols
Default
Function List
Tree View
Gene Information
Category:
all
Level:
2
Name
Times
Percent(%)
integral to membrane 6 26
protein binding 6 26
extracellular space 5 21
membrane 5 21
nucleus 5 21
cell growth and/or maintenance 4 17
regulation of transcription, DNA-dependent 4 17
DNA binding 3 13
G-protein coupled receptor activity 3 13
G-protein coupled receptor protein signaling pathway 3 13
cell proliferation 3 13
cytoplasm 3 13
cytokine activity 3 13
GTP binding 2 8
axonogenesis 2 8
cellular_compo 2 8
calcium ion bin 2 8
cell-matrix adhe 2 8
growth factor ac 2 8
intracellular signaling cascade 2 8
inflammatory response 2 8
ion channel activity 2 8
Input: A list genes/proteins
[Genebank/UniGene/Gene Symbol]
S1, AREG, P2Y5, MARCO, ADORA2A, ADCYAP1R1, CBX3, GATA2, FAF1, JUNB, GLO1, PDGFA, BTG1, EEF1A1, S100A11, LCP1, CACNB3, ACYP1]


Bulk Gene Search System for Java
Choose data adapter File Export Choose Database
Gene/Protein function search: (Please separate identifiers with spaces. eg.Hs.126248)
RAC1 GRIA2 SHC1 CDK5 SCYA2 LIMS1 AREG MARCO CBX3 EEF1A1 CACNB3
OK
Data Type
Gene/Protein ID.
Gene/Protein symbols
Default
Function List
Tree View
Gene Information
Gene_Ontology(23)
molecular_function(21)
binding(14)
protein binding(7)
nucleic acid binding(4)
nucleotide binding(3)
purine nucleotide binding(3)
metal ion binding(2)
receptor binding(4)
signal transducer activity(9)
receptor binding(4)
receptor activity(5)
transmembrane receptor activity(5)
G-protein coupled receptor activity(3)
scavenger receptor activity(1)
glutamate receptor activity(1)
inotropic glutamate receptor activity(1)
nucleotide receptor activity(1)
purinergic nucleotide receptor activity(1)
purinergic nucleotide receptor activity, G-protein coupled(1)
adenosine receptor activity, G-protein coupled(1)
A2A adenosine receptor activity, G-protein coupled(1)
transcription regulator activity(3)
transcription factor activity(2)
S1, AREG, P2Y5, MARCO, ADORA2A, ADCYAP1R1, CBX3, GATA2, FAF1, JUNB, GLO1, PDGFA, BTG1, EEF1A1, S100A11, LCP1, CACNB3, ACYP1]

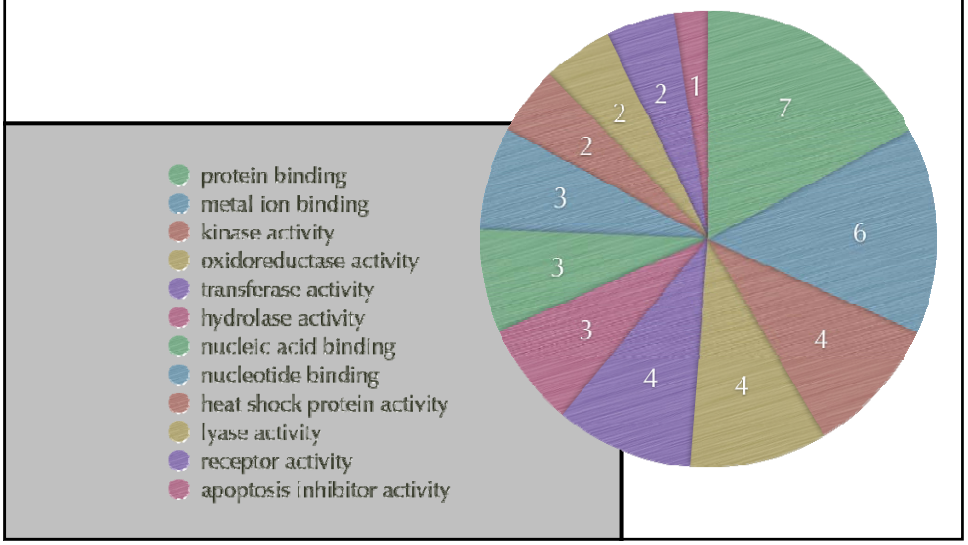

Bulk Gene Search System for Java
Choose data adapter
File
Export
Choose Database
Gene/Protein function search: (Please separate identifiers with spaces. eg.Hs.126248)
RAC1 GRIA2 SHC1 CDK5 SCYA2 LIMS1 AREG MARCO CBX3 EEF1A1 CACNB3
OK
Data Type
Gene/Protein ID.
Gene/Protein symbols
Default
Function List
Tree View
Gene Information
Gene_Ontology 23
RAC1 RAC1 Rac1
GRIA2 GRIA2 glutamate receptor, ionotropic, AMPA2 (alpha 2)
SHC1 SHC1 src homology 2 domain-containing transforming protein C1
CDK5 CDK5 Cdk5
SCYA2 SCYA2 small inducible cytokine A2
LIMS1 LIMS1 LIM and senescent cell antigen-like domains 1
AREG AREG amphiregulin
P2Y5 P2Y5 purinergic receptor (family A group 5)
MARCO MARCO macrophage receptor with collagenous structure
ADORA2A ADORA2A adenosine A2a receptor
ADCYAP1R1 ADCYAP1R1 adenylate cyclase activating polypeptide 1 receptor 1
CBX3 CBX3 chromobox homolog 3 (Drosophila HP1 gamma)
GATA2 GATA2 GATA binding protein 2
FAF1 FAF1 Fas-associated factor 1
JUNB JUNB Jun-B oncogene
S1, AREG, P2Y5, MARCO, ADORA2A, ADCYAP1R1, CBX3, GATA2, FAF1, JUNB, GLO1, PDGFA, BTG1, EEF1A1, S100A11, LCP1, CACNB3, ACYP1]


Biological Process
metabolism
cell growth / maintenance
response to external stimulus
cell communication
reproduction
cell death
cell motility
excretion
pattern specification
16
5
4
3
2
1
1
1
1

Cellular Component



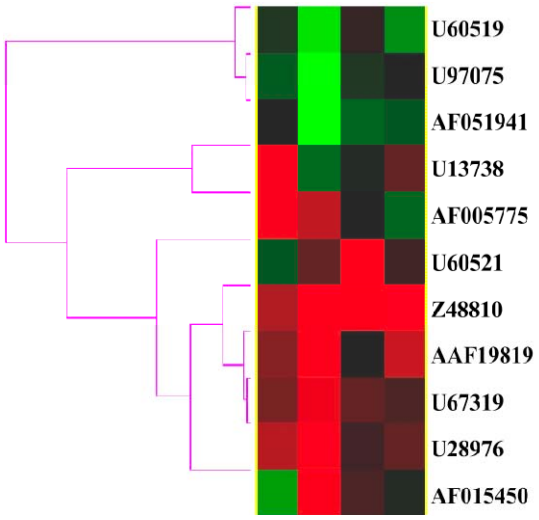
Molecular Function

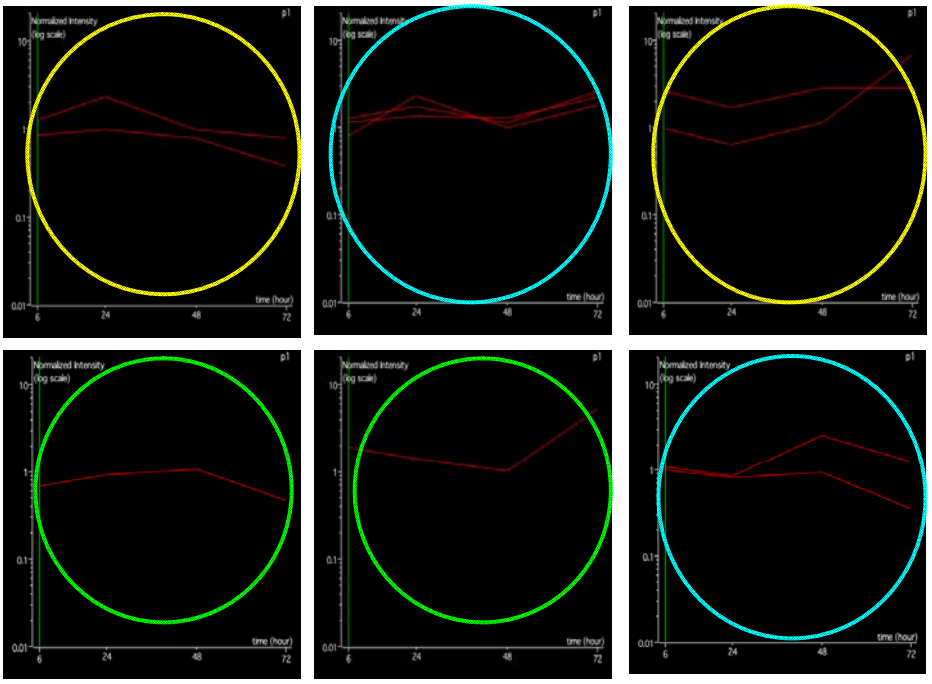


Apoptosis Regulator

Description	Genebank accession No.	6 h Fold Change	24 h Fold Change	48 h Fold Change	72 h Fold Change
Group 1					
caspase 10, apoptosis-related cysteine protease	U60519	-	-	-	0.471
CASP8 and FADD-like apoptosis regulator	U97075	-	-	-	0.355
nucleoside diphosphate kinase type 6 (inhibitor of p53-induced apoptosis-alpha)	AF051941	-	-	-	0.376
Group 2					
caspase 3, apoptosis-related cysteine protease	U13738	-	2.301	-	-
CASP8 and FADD-like apoptosis regulator	AF005775	-	2.272	-	-
Group 3					
caspase 9, apoptosis-related cysteine protease	U60521	-	-	2.519	-
Group 4					
caspase 4, apoptosis-related cysteine protease	Z48810	2.615	-	2.796	2.819
Group 5					
inhibitor of apoptosis protein	AAF19819	-	-	-	5.249
caspase 7, apoptosis-related cysteine protease	U67319	-	-	-	2.19
caspase 4, apoptosis-related cysteine protease	U28976	-	-	-	2.603
Group 6					
CASP8 and FADD-like apoptosis regulator	AF015450	-	-	-	6.912

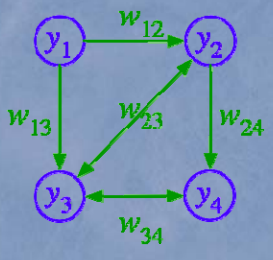
Apoptosis Regulator





Gene Network Inference

Linear Model



$$\frac{\Delta y_i(t)}{\Delta t} = \sum_j w_{ij} y_j(t) + b_i$$

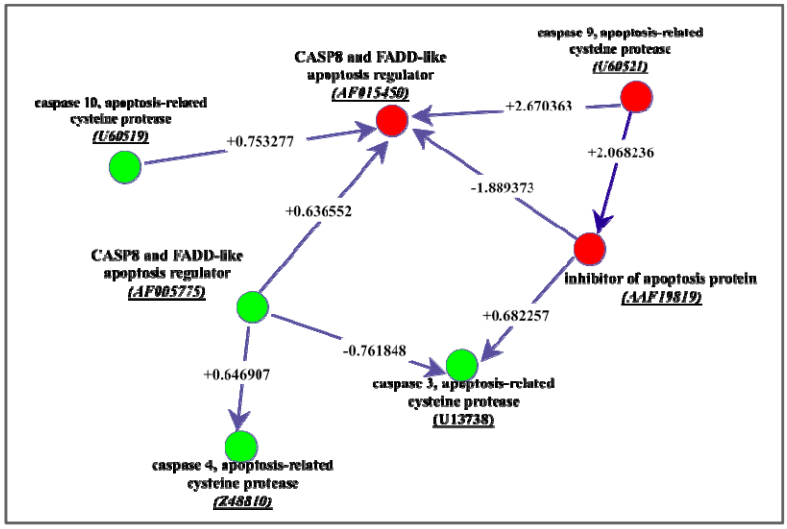
$$\frac{\Delta \mathbf{y}(t)}{\Delta t} = \mathbf{W} \times \mathbf{y}(t) + \mathbf{B}$$

$$\mathbf{W}^{\dagger} = \frac{\Delta \mathbf{Y}}{\Delta t} \times \widetilde{\mathbf{Y}}^T \times (\widetilde{\mathbf{Y}}^T \times \widetilde{\mathbf{Y}})^{-1}$$

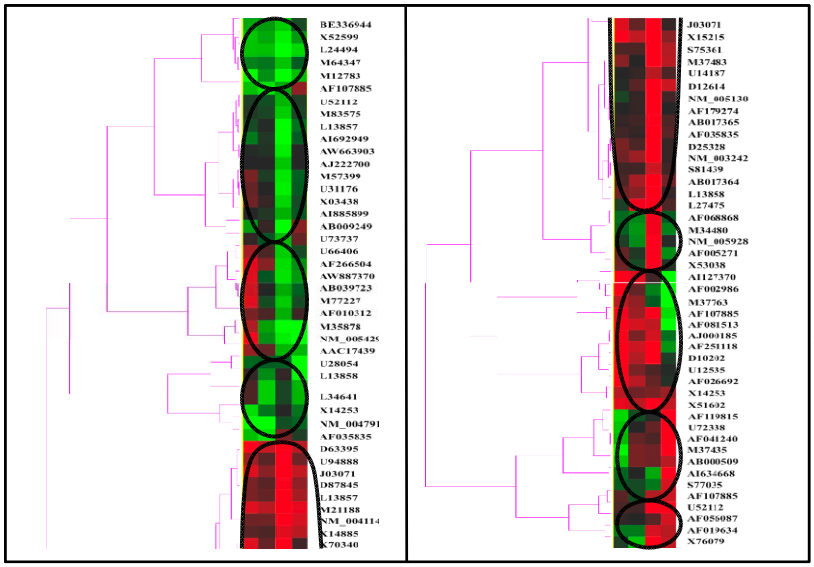
Weights Matrix
Apoptosis Regulator

Weights	Gene 1	Gene 2
2.670363	AF015450	U60521
2.068236	AAF19819	U60521
-1.889373	AF015450	AAF19819
-1.427408	AAF19819	AAF19819
-0.81632	AF005775	AF005775
-0.761848	U13738	AF005775
0.753277	AF015450	U60519
0.682257	U13738	AAF19819
0.646907	Z48810	AF005775
0.636552	AF015450	AF005775
0.632796	AF005775	Z48810
0.594627	AF005775	AAF19819
-0.55848	Z48810	AAF19819
0.543142	AAF19819	U60519
-0.527872	U60521	U60521
0.518056	U28976	U60521
0.508007	U60521	AF005775
0.499483	U13738	Z48810

Gene Regulatory Network:
Apoptosis Regulator



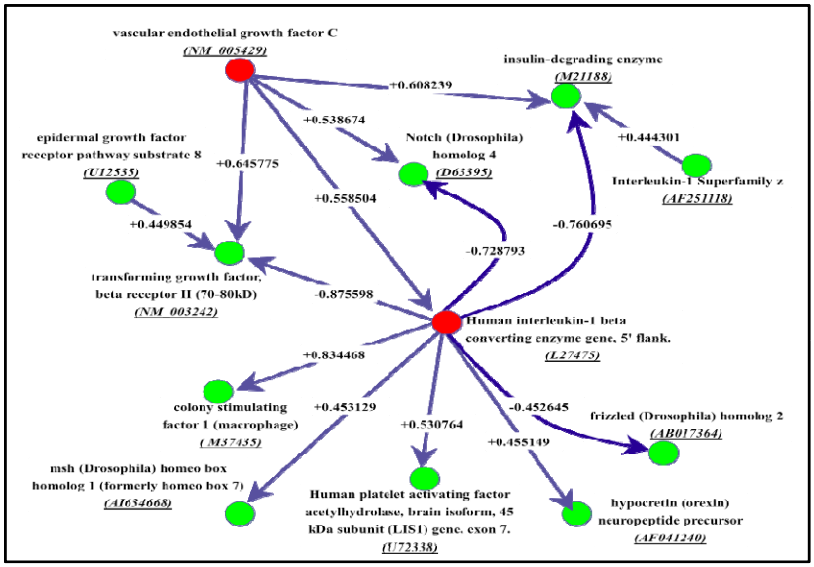
Signal Transducer



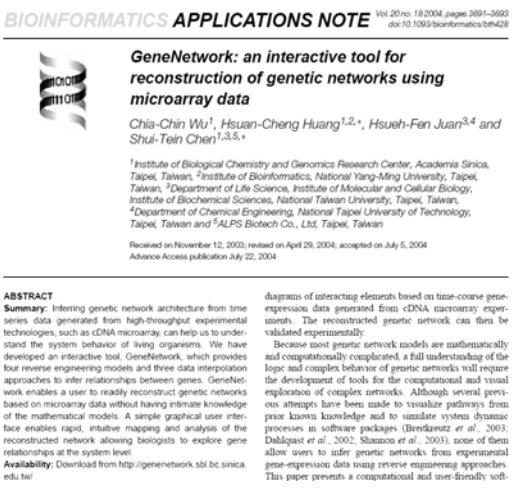
Weights Matrix
Signal Transducer

Weights	Gene1	Gene2
-0.875598	NM_003242	L27475
0.834468	M37435	L27476
-0.78655	L27475	L27477
-0.760695	M21188	L27478
-0.728793	D63395	L27479
0.645775	NM_003242	NM_005429
0.608239	NM_003242	NM_005430
0.558504	L27475	NM_005431
0.538674	D63395	NM_005432
0.530764	U72338	L27475
0.455149	AF041240	L27476
0.453129	AI634668	L27477
-0.452645	AB017364	L27478
0.449854	NM_003242	U12535
0.444301	M21188	AF251118
-0.442259	M37435	U52112
0.437687	M21188	U12535
0.429097	NM_003242	AF251118

Gene Network: Signal Transducer



Gene Network Modeling and Reconstruction



BIOINFORMATICS **APPLICATIONS NOTE** Vol. 20 no. 18 2004, pages 3691–3693 doi:10.1093/bioinformatics/bth428

GeneNetwork: an interactive tool for reconstruction of genetic networks using microarray data

Chia-Chin Wu[1], Hsuan-Cheng Huang[1,2,], Hsueh-Fen Juan[3,4] and Shui-Tein Chen[1,3,5,*]*

[1]Institute of Biological Chemistry and Genomics Research Center, Academia Sinica, Taipei, Taiwan, [2]Institute of Bioinformatics, National Yang-Ming University, Taipei, Taiwan, [3]Department of Life Science, Institute of Molecular and Cellular Biology, Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan, [4]Department of Chemical Engineering, National Taipei University of Technology, Taipei, Taiwan and [5]ALPS Biotech Co., Ltd, Taipei, Taiwan

Received on November 12, 2003; revised on April 29, 2004; accepted on July 5, 2004
Advance Access publication July 22, 2004

ABSTRACT

Summary: Inferring genetic network architecture from time series data generated from high-throughput experimental technologies, such as cDNA microarray, can help us to understand the system behavior of living organisms. We have developed an interactive tool, GeneNetwork, which provides four reverse engineering models and three data interpolation approaches to infer relationships between genes. GeneNetwork enables a user to readily reconstruct genetic networks based on microarray data without having intimate knowledge of the mathematical models. A simple graphical user interface enables rapid, intuitive mapping and analysis of the reconstructed network allowing biologists to explore gene relationships at the system level.

Availability: Download from http://genenetwork.sbl.bc.sinica.edu.tw/

diagrams of interacting elements based on time-course gene-expression data generated from cDNA microarray experiments. The reconstructed genetic network can then be validated experimentally.

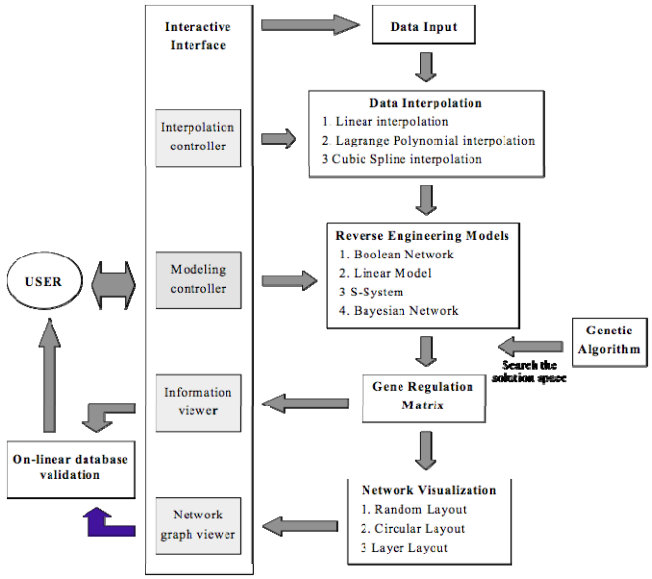
Because most genetic network models are mathematically and computationally complicated, a full understanding of the logic and complex behavior of genetic networks will require the development of tools for the computational and visual exploration of complex networks. Although several previous attempts have been made to visualize pathways from prior known knowledge and to simulate system dynamic processes in software packages (Brentkreutz *et al.*, 2003; Dahlquist *et al.*, 2002; Shannon *et al.*, 2003), none of them allow users to infer genetic networks from experimental gene-expression data using reverse engineering approaches. This paper presents a computational and user-friendly soft-

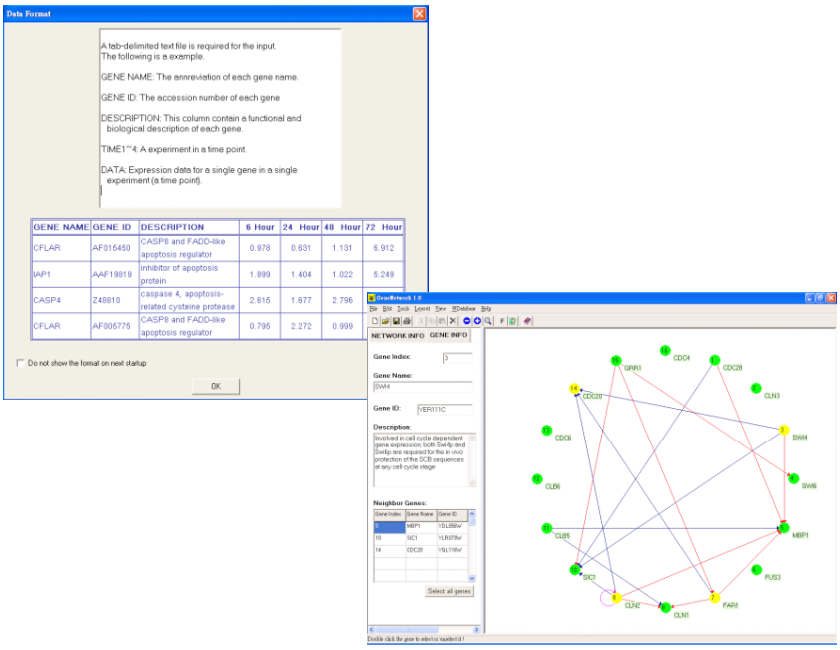
Bioinformatics (2004) **20**: 3691-3.

GeneNetwork Software

- Interactive interface for reconstructing gene network based on microarray data
- Reconstruction Model
  - Boolean Network
  - Baysian Network
  - Linear Model
  - S-System
- http://genenetwork.sbl.bc.sinica.edu.tw/

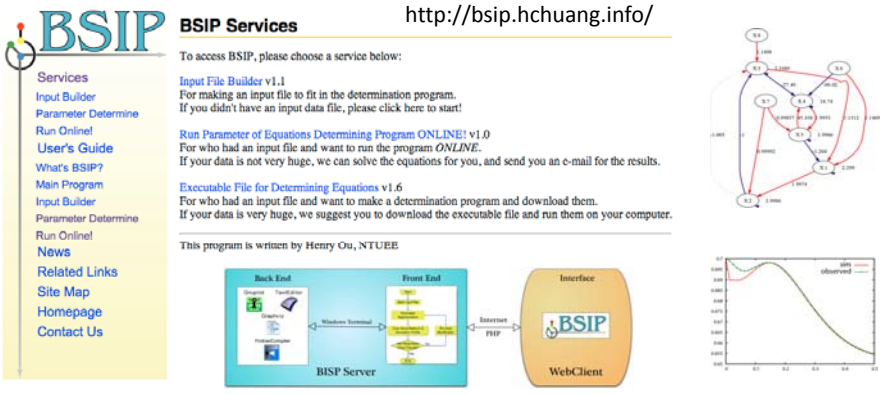
Bioinformatics (2004) 20: 3691-3





BSIP

- Biological network modeling based on S-system
- Network reconstruction from time-series data

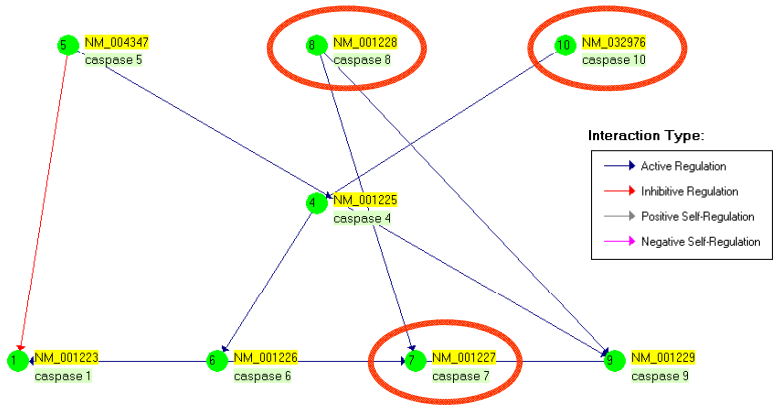


http://bsip.hchuang.info/

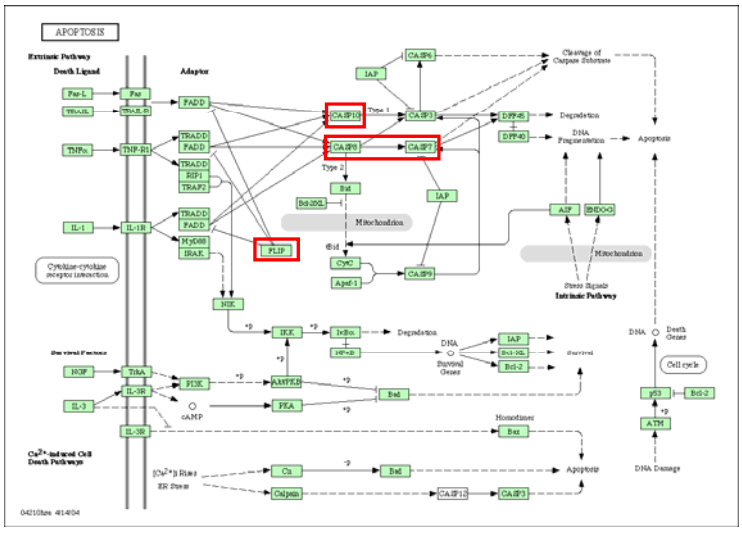
APBC2006 Poster Award

Collaborated w/ Prof. F.S.Wang (CCU)

Gene Network of Caspase Genes



Apoptosis pathway



http://kegg.com/dbget-bin/get_pathway?org_name=hsa&mapno=04210

Summary

- We have synthesized a novel cyclic-RGD peptide, which induces apoptosis of MCF-7
- These results provide a molecular explanation for the properties of cRGD in breast cancer cells and present a valuable in-depth description of their possible impact on breast cancer therapy

Drug Discovery:
ATP synthase, a new target for cancer therapy

75

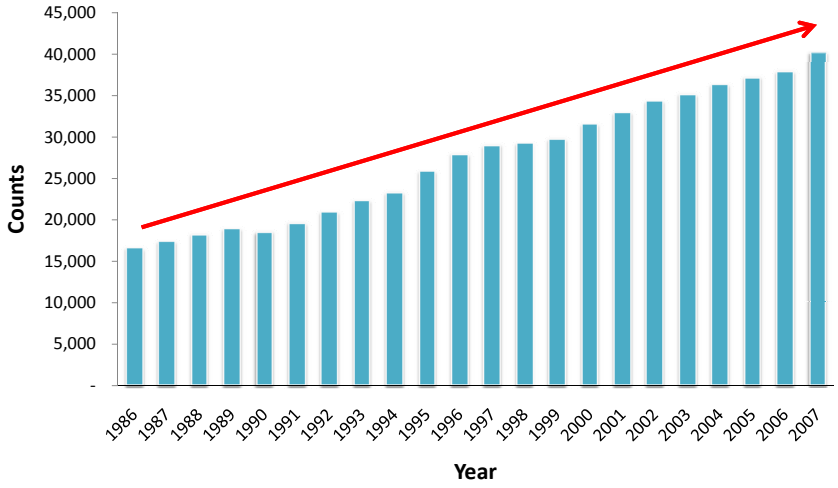
Rank of deaths from leading causes of death in 2007

Rank	Causes of death
1	Malignant neoplasms
2	Heart disease
3	Cerebrovascular disease
4	Diabetes mellitus
5	Accidents and adverse effects
6	Pneumonia
7	Chronic liver disease and cirrhosis
8	Nephritis, nephrotic syndrome and nephrosis
9	Suicide
10	Hypertensive disease

76

Statistics data, Department of Health, Executive Tuan, Taiwan. 2008.

Number of deaths of malignant neoplasm in the past two decades

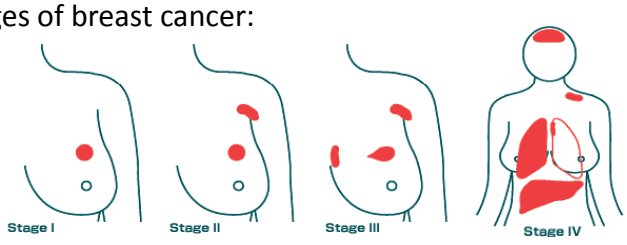


77

Statistics data, Department of Health, Executive Tuan, Taiwan. 2008.

Breast cancer

- The most common malignancy among women in developed regions of the world.
- In the United States, more than 200,000 women are diagnosed with breast cancer each year and nearly 41,000 patients die.
- Although breast cancer is primarily a disease of women, about 1% of breast cancers occur in men.
- stages of breast cancer:



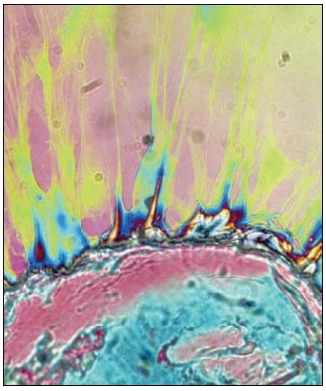
Harris JR. Staging and natural history of breast cancer. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 2000: 403-424.

Cserni G, Kulka J. [New TNM classification of breast tumors]. *Orv Hetil.* 2003 Aug 10;144(32):1563-8

78

Cancer Robustness

- Viewing cancer as a robust system with potential points of fragility opens up new strategies for the development of drugs and therapies.



NATURE 2003, 426, 125.

Hunt for fragility: weaknesses in tumor growth dynamics could yield new anti-cancer therapies.

Targeting Therapy

- Drugs or other substances which identify and attack specific cancer cells without harming normal cells.

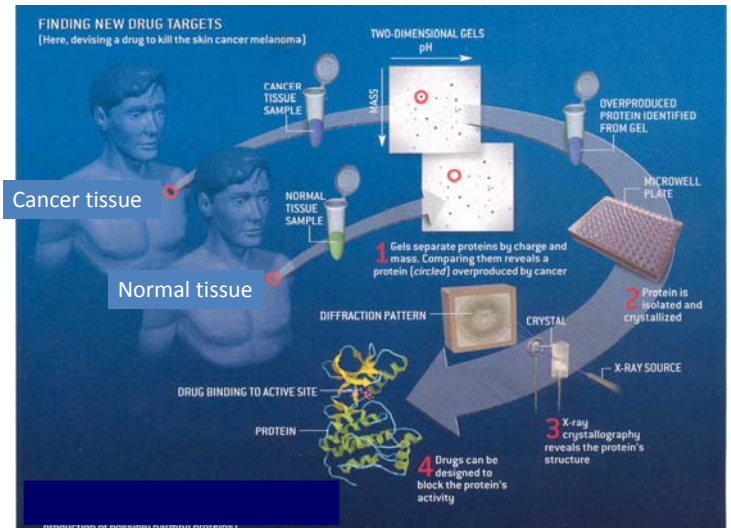
80

Goal

- To discover drug targets and antitumor agents.

81

How Proteomics can help drug development



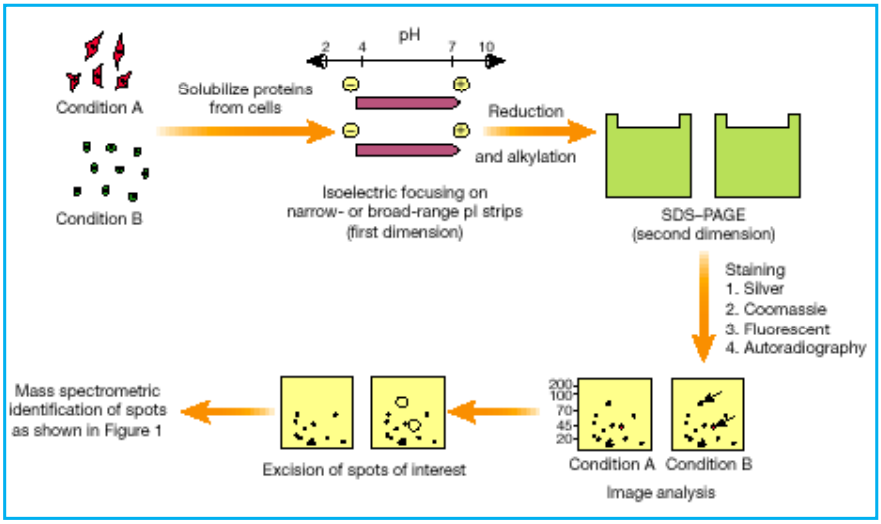
Cancer tissue

Normal tissue

82

Modified by Scientific American 2002

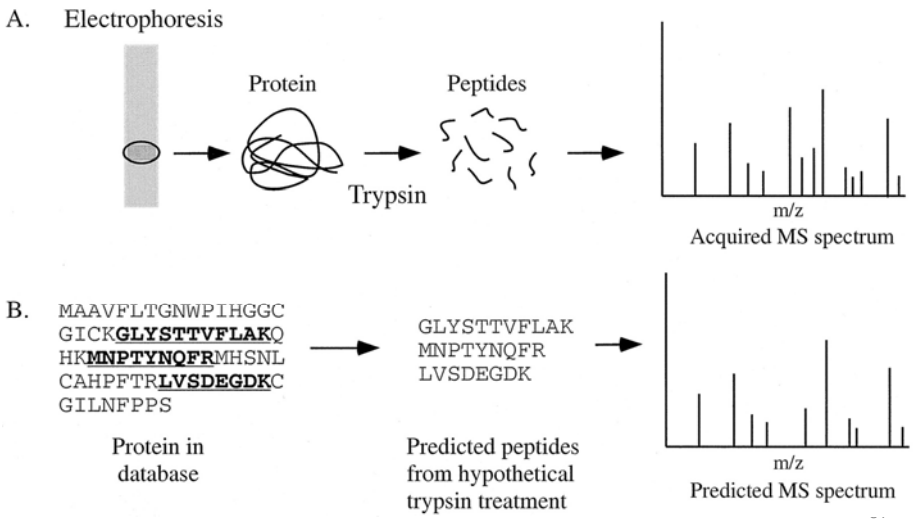
Proteomics: Two-dimensional Gel Approach



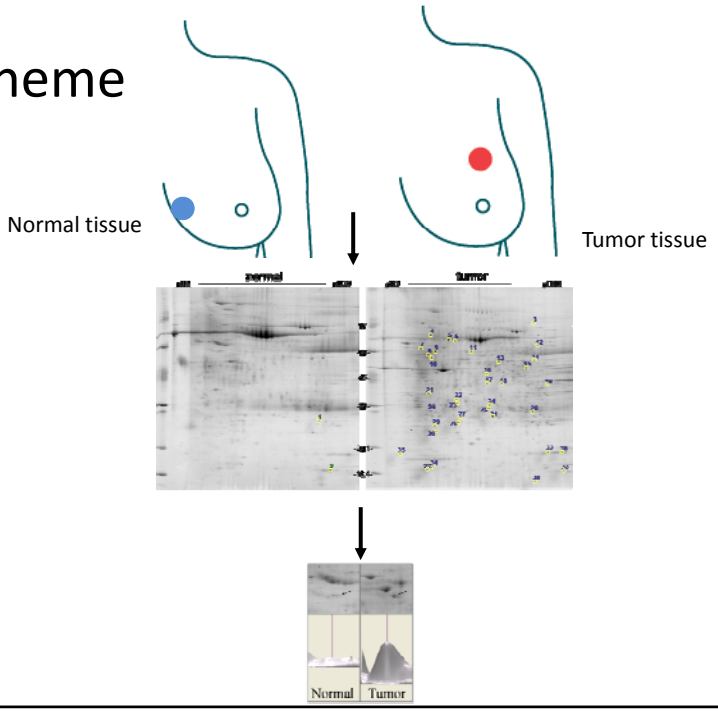
Nature 2000, 405, 837-846

83

Proteomics: Peptide Mass Fingerprinting

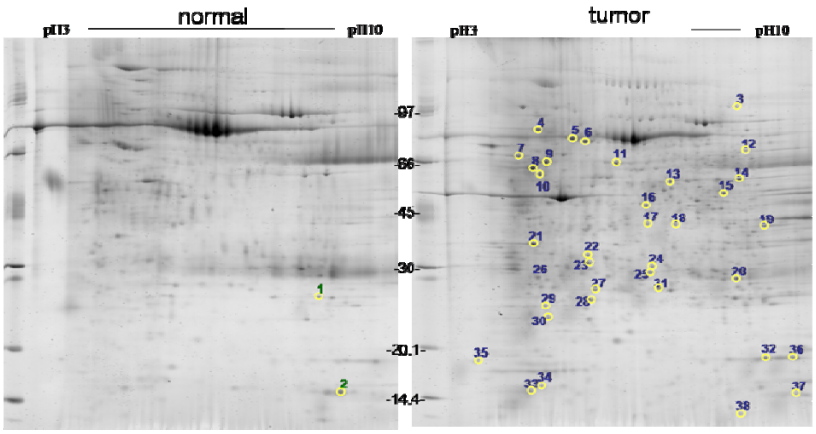


Scheme



85

2DE Map



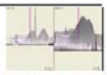
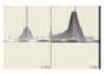
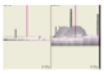
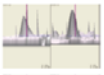
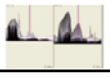
J Proteome Research 2008

86

ATP Synthase

Targeting Therapy for Breast Carcinoma research articles

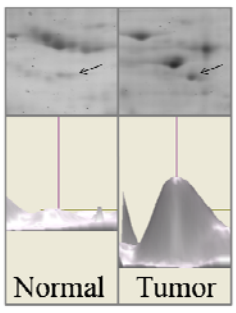
Table 1. List of Identified Proteins with Significant Differential Expression in Cancerous Tissues

Spot	Accession number	Protein name	Biological process	Score	Coverage	Queries matched	Theoretical MW (Da)	Theoretical PI	Expression Ratio	3D view
4	P11021	78 kDa glucose-regulated protein precursor	anti-apoptosis	153	0.35	19	72288	5.07	4.52	
6	P08107	Heat shock 70 kDa protein 1	anti-apoptosis	160	0.32	17	69995	5.48	1.58	
21	P08758	Annexin A5	anti-apoptosis	184	0.56	17	35914	4.94	2.99	
28	P32119	Peroxiredoxin-2	anti-apoptosis	100	0.36	8	21878	5.66	2.73	
29	P32119	Peroxiredoxin-2	anti-apoptosis	59	0.35	6	21878	5.66	3.46	
27	P09211	Glutathione S-transferase P	anti-apoptosis; metabolic process	56	0.3	5	23341	5.43	1.23	
10	P06576	ATP synthase subunit beta, mitochondrial precursor	catalytic activity	149	0.42	17	56525	5.26	4.46	

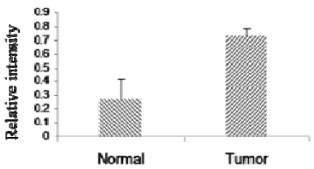
7

The Expression Levels of ATP Synthase β Subunit

2DE



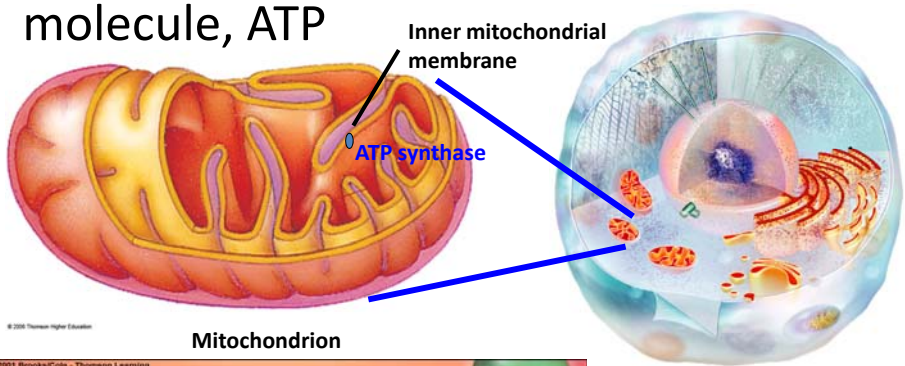
Western blot

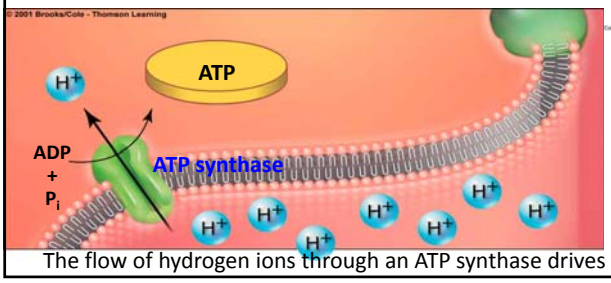


The expressions of ATP synthase β subunit in tumor tissues are higher than normal tissues.

88

ATP synthase: to produce energy molecule, ATP



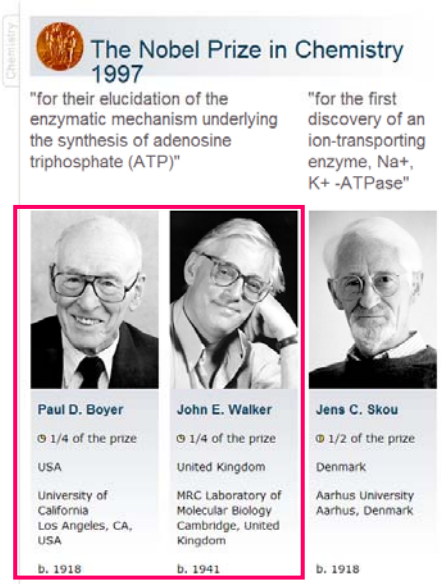


ADP→ ATP

The flow of hydrogen ions through an ATP synthase drives the formation of ATP.

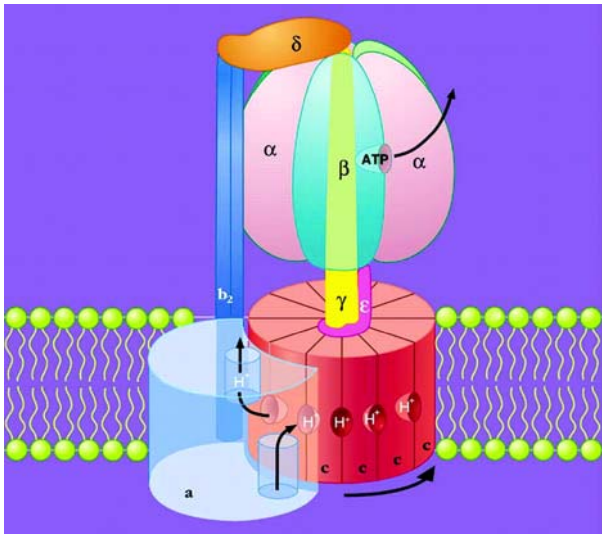
89

ATP synthase



90

ATP synthase

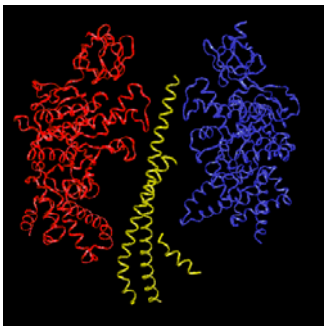


91

Proc. Natl. Acad. Sci. U. S. A. 98, 8519-8524 (2001)

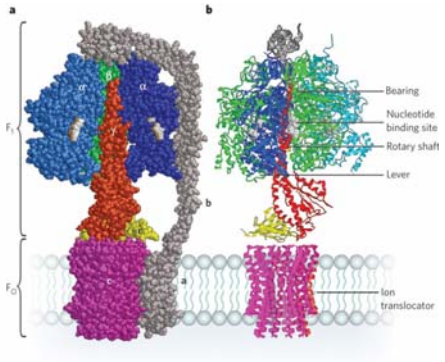
The action of subunit γ on the lever of subunit β

β



α

γ

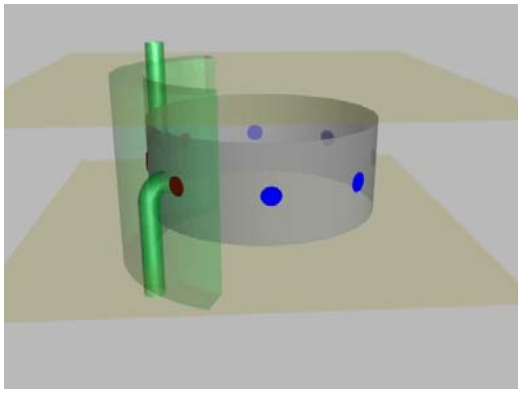


Nature 459, 364-370 (21 May 2009)

92

The function of the rotary electromotor

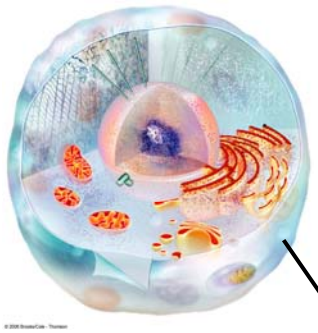
- Torque generation by Brownian rotary motion and directed ion flow



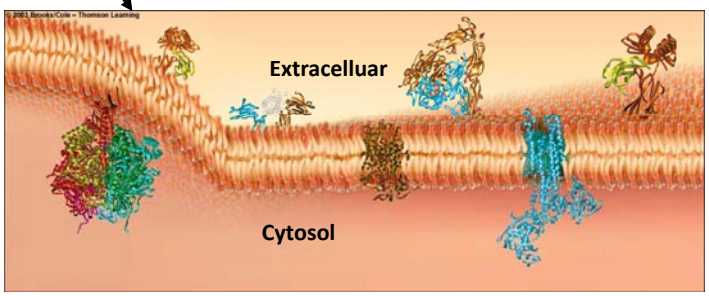
ATP

Nature 459, 364-370 (21 May 2009)

93



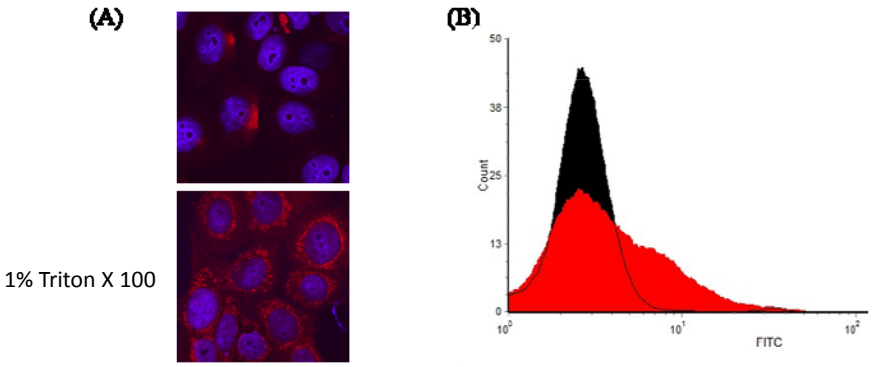
Is ATP synthase also located on cell membrane ??



For a long time, ATP synthase expression was believed to be found only in mitochondria where most cellular ATP synthase takes place.

94

Characterization of ATP synthase expressed on MCF-7 cell surface



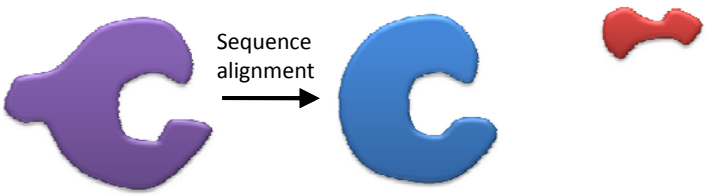
(A) Confocal microscopy image of the distribution of ATP synthase subunit on MCF-7 cell surface. MCF-7 were fix and with (lower)/without (upper) permeabilization. Red, ATP synthase subunit; blue, hoechst 33342. (B) Cells in red area represent the ones stained with anti-ATP synthase subunit Ab followed by FITC-labeled anti-mouse IgG antibodies. Cells in black area denote negative control cells.

J Proteome Research 2008

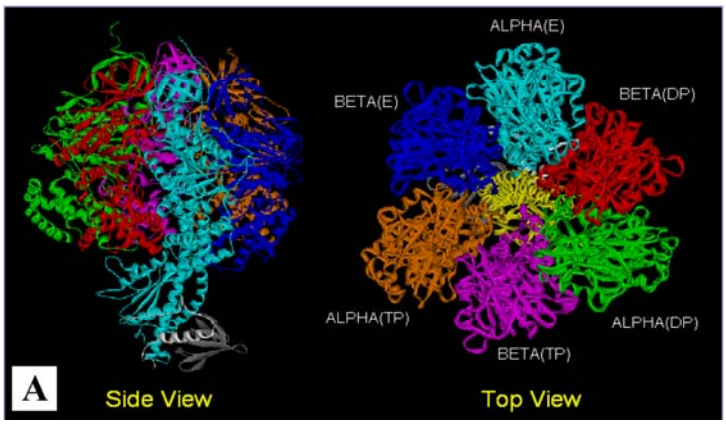
95

Molecular modeling of the inhibition on ATP synthase

Homology modeling and protein-ligand docking simulation



Homology modeling of human ATP synthase using bovine as the template



JPR2008

The sequence identity between human and bovine ATP synthase is 99%.
We used MODELER program encoded in InsightII and used bovine ATP synthase which has the 3D structure as the template to model human ATP synthase.
The structure with the lowest energy scores was chosen as the candidate.

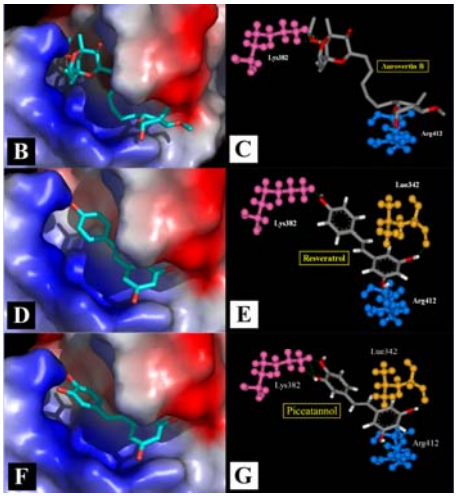
97

Three potent drugs were screened out by protein-ligand docking simulation

Aurovertin B

Resveratrol

Piceatannol

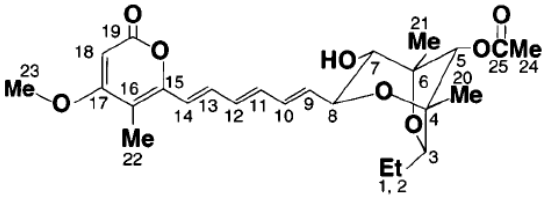


Docking simulation was done by the shape-based docking algorithm LigandFit.
All calculations were carried out in the Discovery Studio 1.2.
Aurovertin B can dock into ATP synthase β subunit.

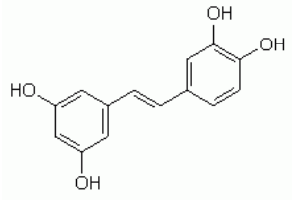
98

J Proteome Research 2008

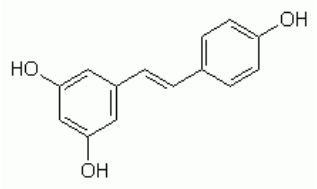
Structures of ATP synthase β subunit inhibitors



Aurovertin B



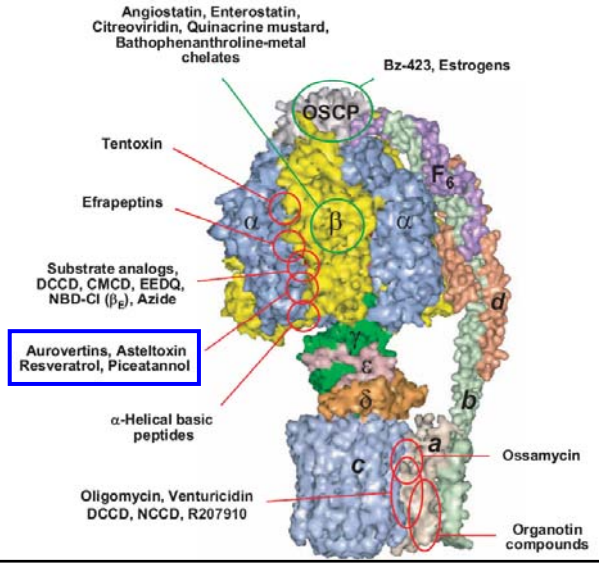
Piceatannol



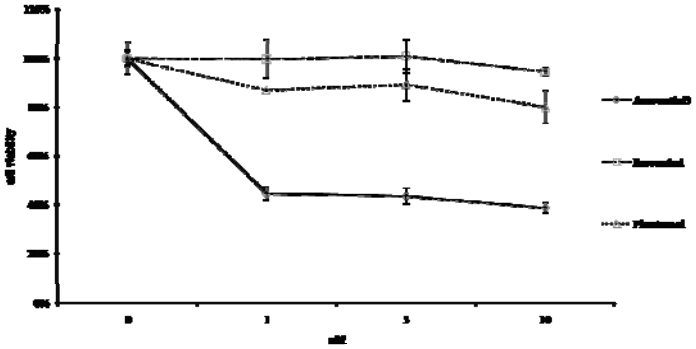
Resveratrol

99

Inhibitory sites of ATP synthase



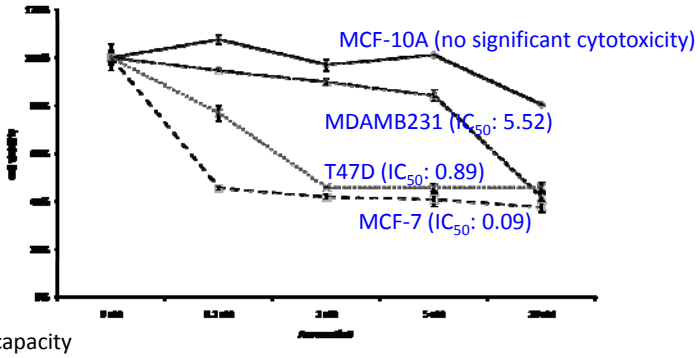
Effects of ATP synthase inhibitors on cell viability in MCF-7 breast cancer cells



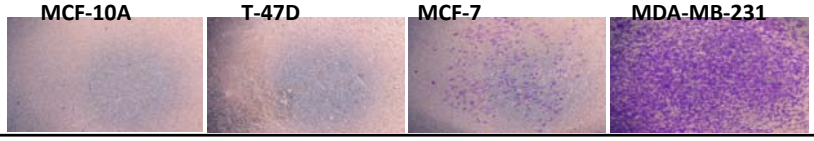
When treated with ATP synthase inhibitor, aurovertin B, breast cancer cells exhibited a significant decrease in cell density.
Aurovetin B inhibited the growth of MCF-7 cells in a dose-dependent manner.
The IC_{50} is 0.1 um.

101

Effects of aurovertin B on cell viability in breast normal cells (MCF-10A) and cancer cells *in vitro*

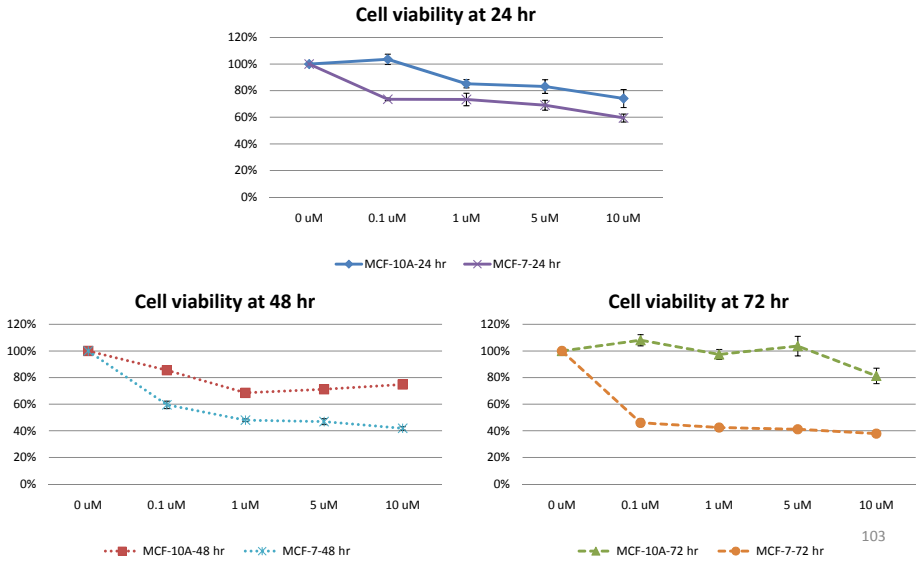


The invasion capacity



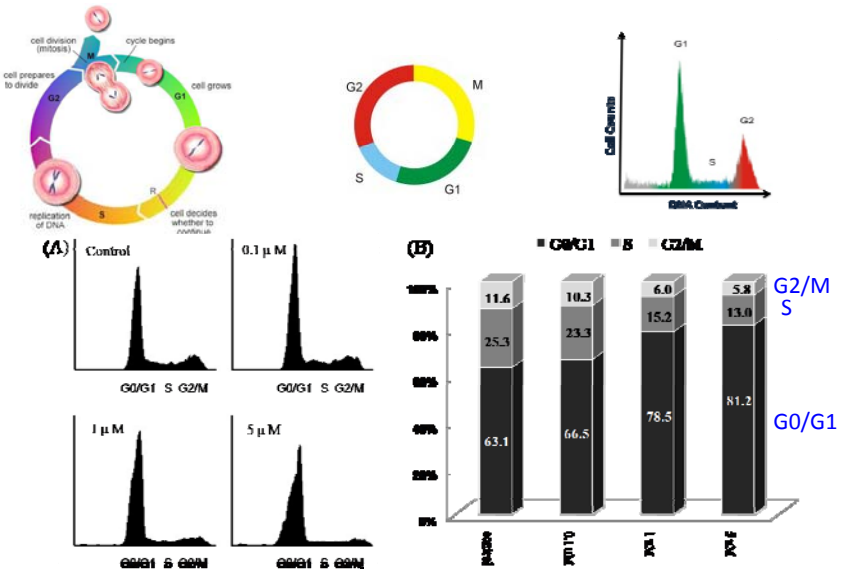
102

Cell viability of aurovertin B -treated MCF-7 and MCF-10A in different time



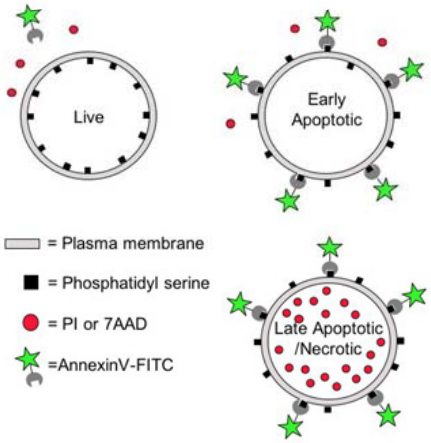
103

Aurovertin B induces cell cycle arrest



In aurovertin B treated cells, percentage of cells in G0/G1 phase was increased.[104]

Annexin V for apoptosis & PI for necrosis assay



Annexin V for apoptosis

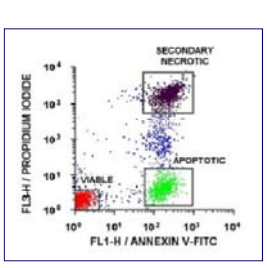
- ✓ As cells undergo apoptosis, PS is translocated to the outer leaflet of the plasma membrane and exposed to the extracellular environment.
- ✓ Annexin V, a 35-36 kD Ca^{2+}-dependent phospholipid-binding protein, could identify apoptotic cells by binding to PS exposed to the outer leaflet.

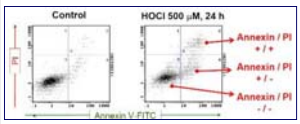
PI for necrosis

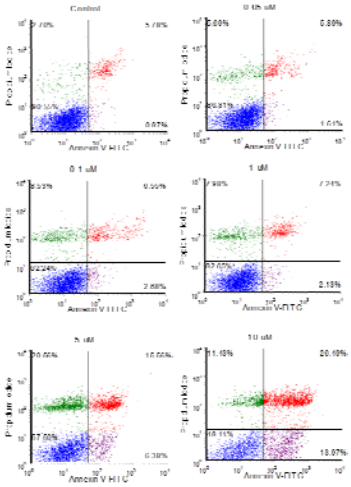
- ✓ PI, propidium iodine, a dye for DNA structure.
- ✓ PI stain ability provides evidence of the loss of plasma membrane integrity

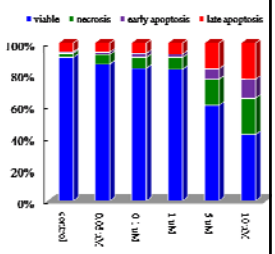
105

Aurovertin B induces apoptosis of MCF-7



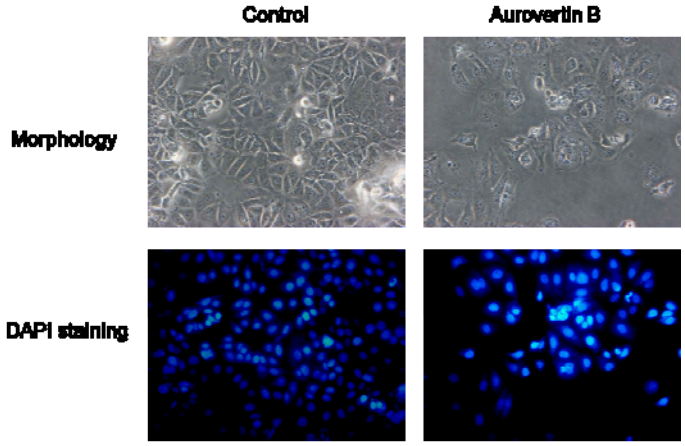






106

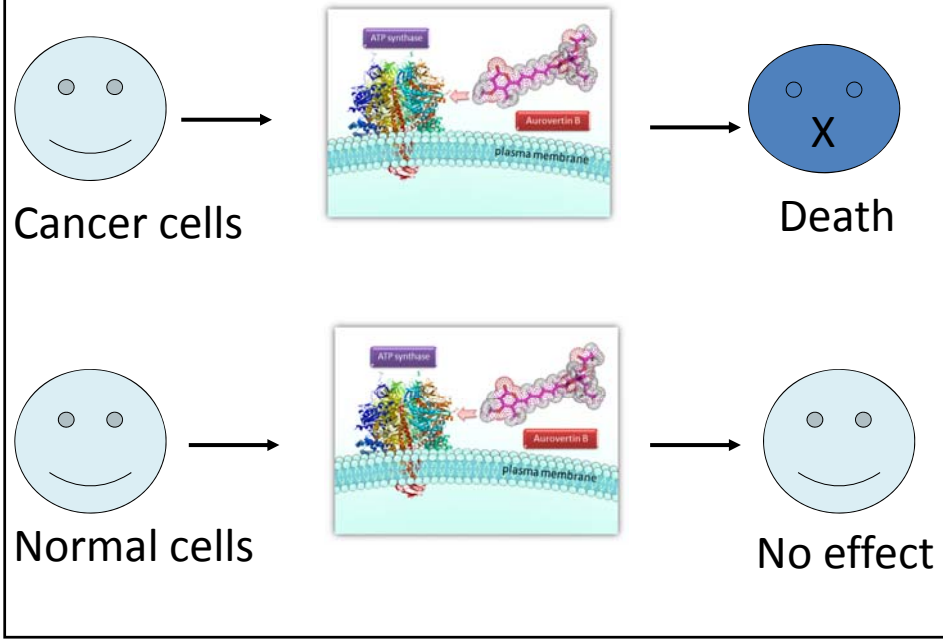
Characterization of aurovertin B induced cell death in human MCF-7 cells



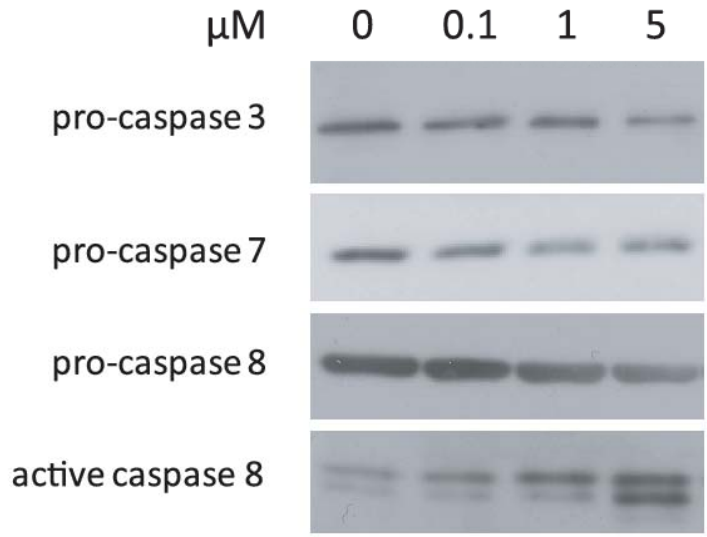
Phase contrast microscopy shows cell shrinkage, irregularity in shape, and cellular detachment in aurovertin B-treated cells.

MCF-7 cells stained with 4, 6-diamidino-2-phenylindole (DAPI). Condensation and shrinkage of nuclei in aurovertin B-treated cells

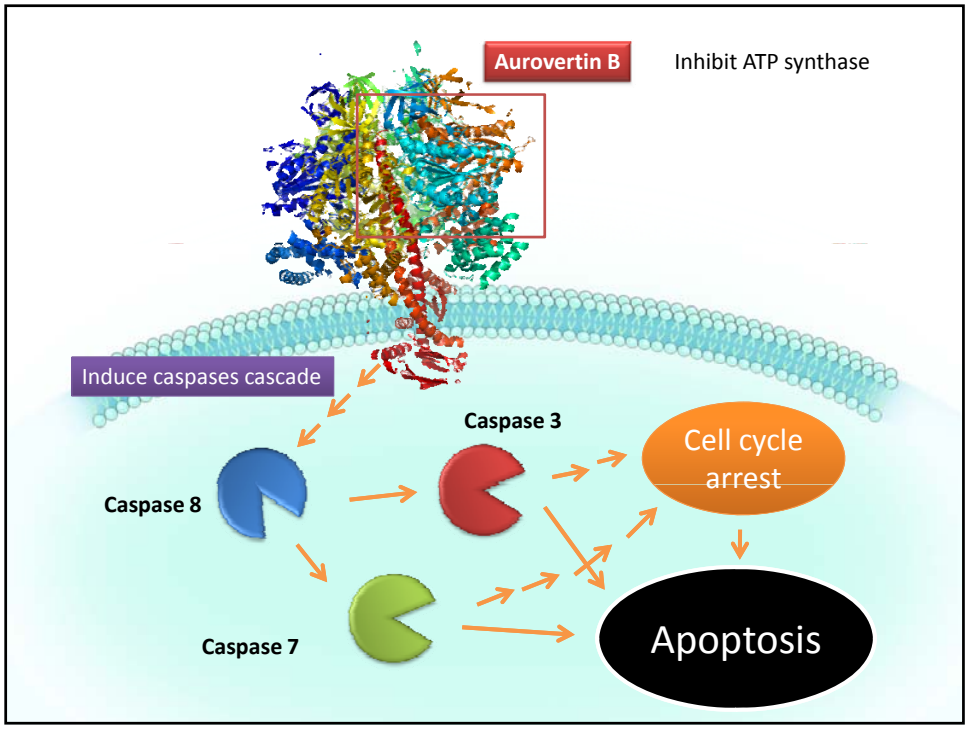
107



Aurovertin B activated caspase-dependent apoptotic pathway



109



Summary

- ATP synthase was upregulated in cancerous tissues and expressed on the surface of cancer cells.
- ATP synthase inhibitor aurovertin B could target on the cancer cells and leave normal unharmful.
- Aurovertin B inhibits proliferation of breast cancer cells by inducing apoptosis and arresting cell cycle at the G0/G1 phase.
- ATP synthase can be a good novel therapeutic target.

111

Mechanism study:
Systems analysis reveals the anti-leukemia molecular mechanism induced by *Ganoderma lucidum* polysaccharides

Outline

- Introduction
- Cell Death
- Cell Differentiation
- Summary

Access more than 5,000 times


BMC Genomics

Ganoderma lucidum polysaccharides in human monocytic leukemia cells: from gene expression to network construction

114

INTRODUCTION

115

Ganoderma lucidum (Reishi)

- *Ganoderma lucidum* has been widely used as an herbal medicine for promoting health and longevity in China and other Asian countries.

  Bioorg. Med. Chem. 10, 1057-1062 (2002).

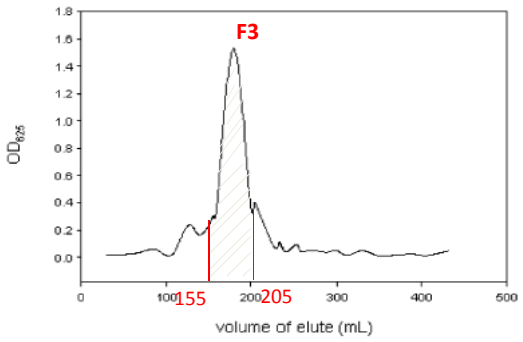
- Polysaccharide extracts from *Ganoderma lucidum* has been reported to exhibit immuno-modulating and anti-tumor activities.

  Bioorg. Med. Chem. 12, 5595-5601 (2004).

116

F3

- President Wong's group purified the active components of the polysaccharide extracts by gel filtration chromatography and designated it as F3.

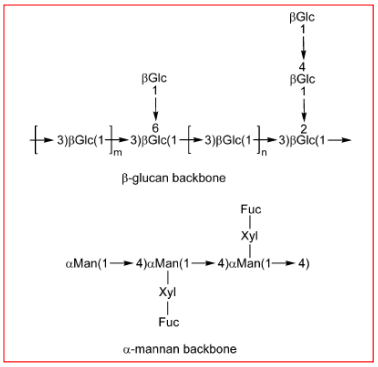


Bioorg. Med. Chem. 10, 1057-1062 (2002).

117

Reishi Polysaccharides

- The major carbohydrate components in F3 are glucose, mannose and galactose.



Sugar Components of F3	Percentage (%)
D-Glucose	58.1
D-Mannose	15.1
D-Galactose	13.5
L-Fucose	7.1
D-Xylose	3.1
D-N-acetylglucosamine	1.2
L-Rhamnose	0.7

Bioorg. Med. Chem. 10, 1057-1062 (2002).

118

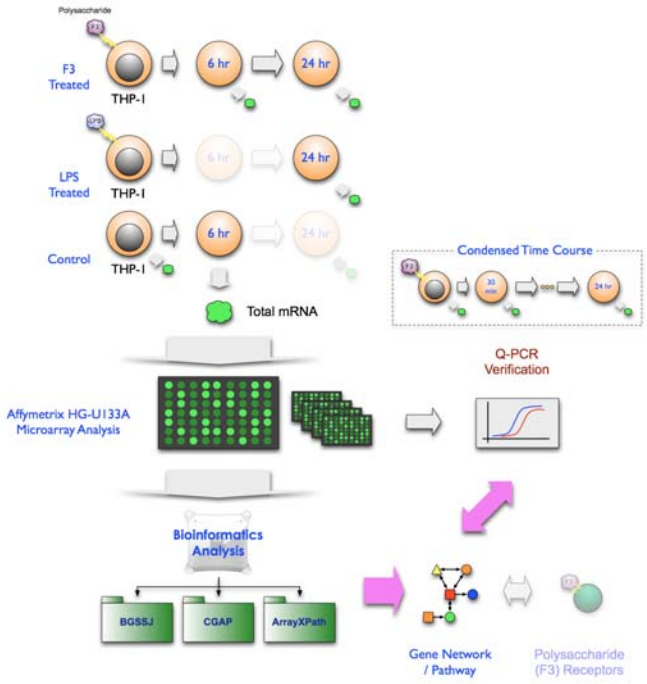
Goal

- How can F3 stimulate anti-tumor effects on cancer cells?
- To understand the molecular mechanism underlying the F3 exertion in THP-1 cells

119

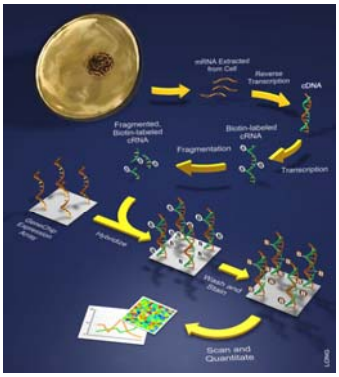
CELL DEATH

120

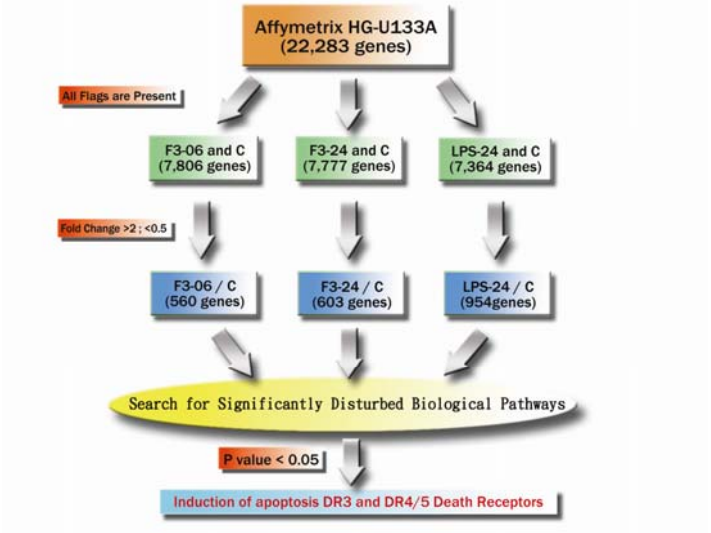


Microarray

- A powerful approach to accurately measure changes in global mRNA expression levels.
- Used to
  - Discover novel genes
  - Determine gene functions
  - Evaluate drugs
  - Dissect pathways
  - Classify clinical samples.



Flowchart for Microarray Data Analysis

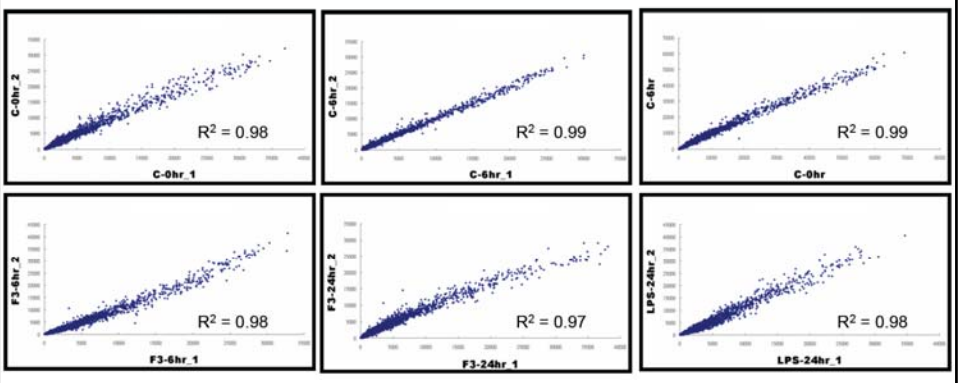


BMC Genomics 8:411 (2008)

123

A Quick Check for Data Validity

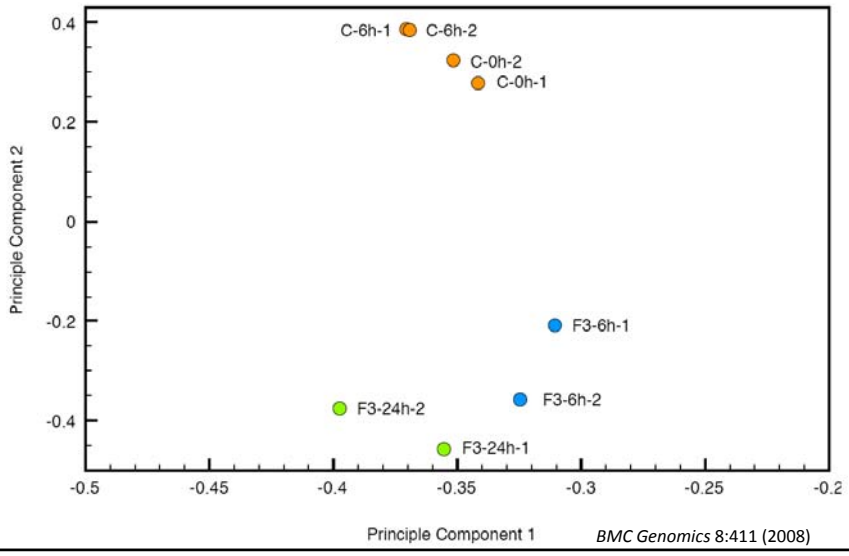
These results showed the consistency in duplicate microarray experiments.



BMC Genomics 8:411 (2008)

The scatter plot

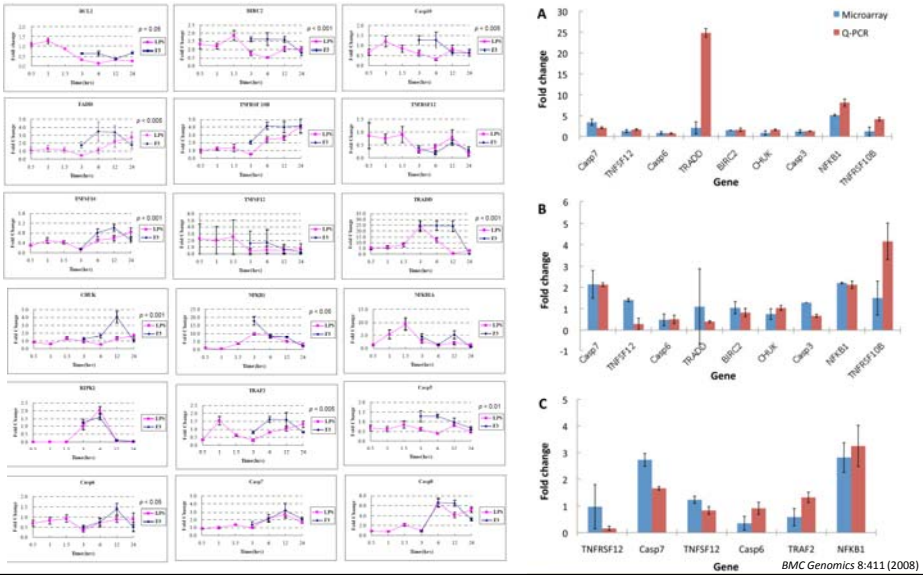
Clear separation of control and F3-treated samples.



25

BMC Genomics 8:411 (2008)

Fold Change in Time-course Gene Expression of F3- or LPS-induced THP-1 Cells



BMC Genomics 8:411 (2008)

Search for Significantly Disturbed Biological Pathways

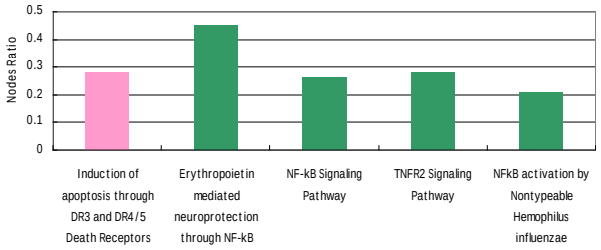
- Statistical methods based on Fisher's exact test and false discovery rate.

Nucleic Acids Research, 2004, Vol. 32,W460-464

Nucleic Acids Research, 2004, Vol. 32 W460-W464

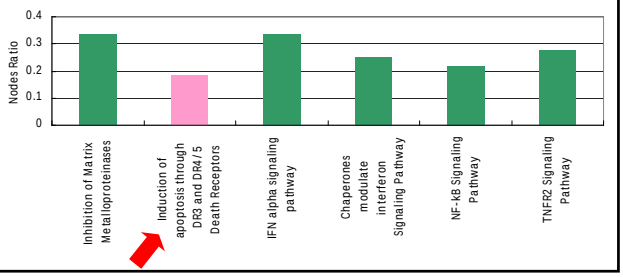
ArrayXPath

All p value of significant pathways < 0.05



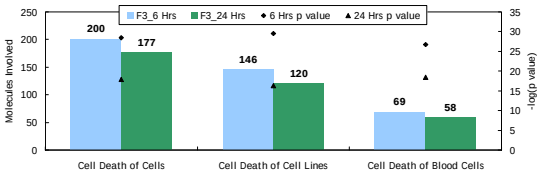
← F3 Treatment for 6 Hrs

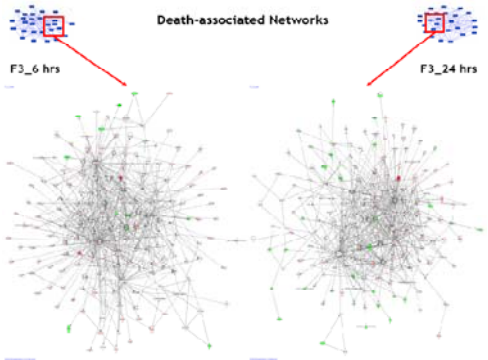
F3 Treatment for 24 Hrs →



Biofunction: Cell Death

- Genes associated with cell death were showed by IPA web tool.

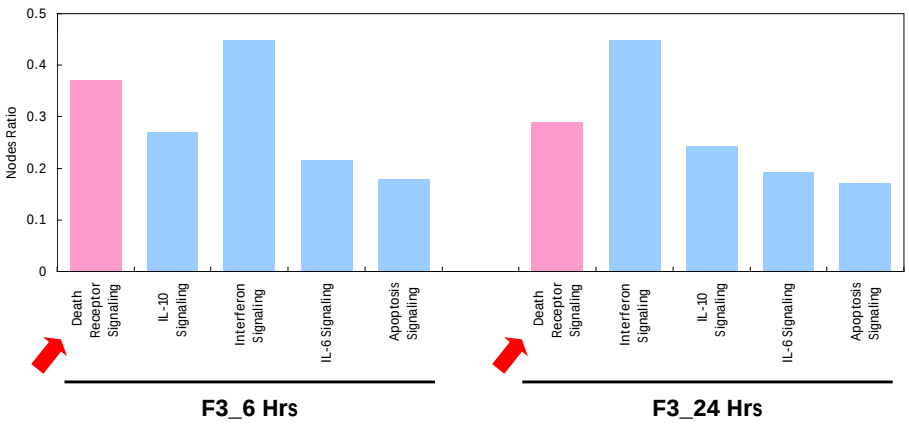




Protein-protein Interaction Network

Ingenuity Pathway Analysis (IPA)

All p value of significant pathways < 0.01

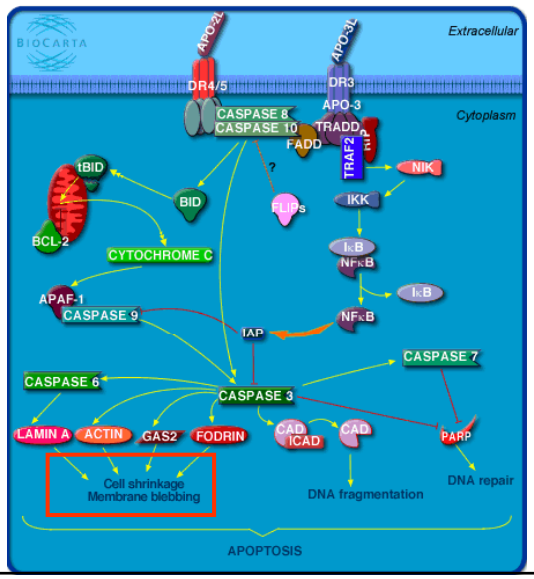


Comparison: F3 Treatment for 6 and 24 Hrs

- ArrayXPath
  - Induction of apoptosis through DR3 and DR4/5 death receptors
  - NF-κB signaling pathway
  - TNFR2 signaling pathway
- IPA
  - Death receptor signaling
  - IL-10 signaling
  - Interferon signaling
  - IL-6 signaling
  - Apoptosis signaling

131

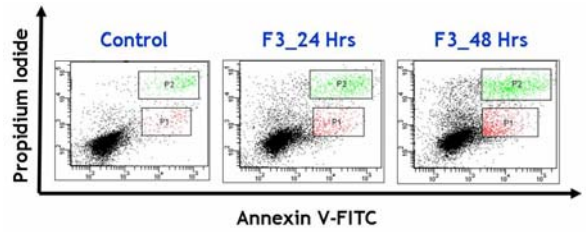
Induction of Apoptosis through DR3 and DR4/5 Death Receptors

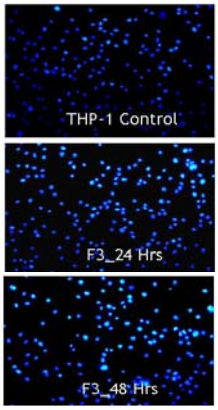


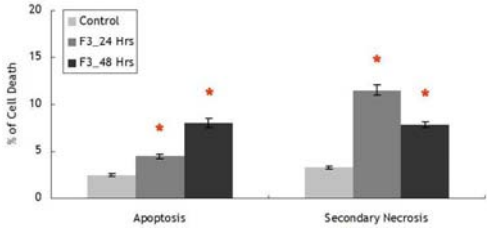
132

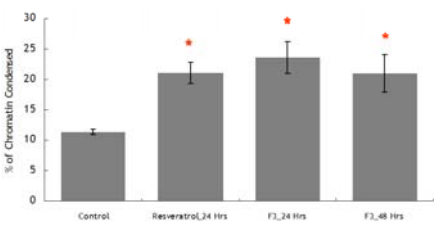
Apoptosis Assay

F3 Treatment: 30 μg/mL





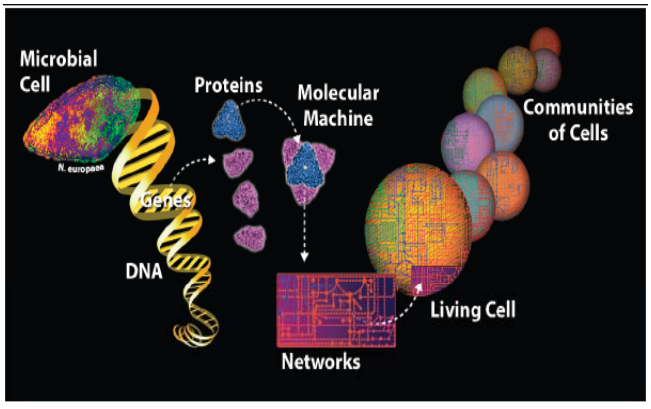




* p value < 0.01

133

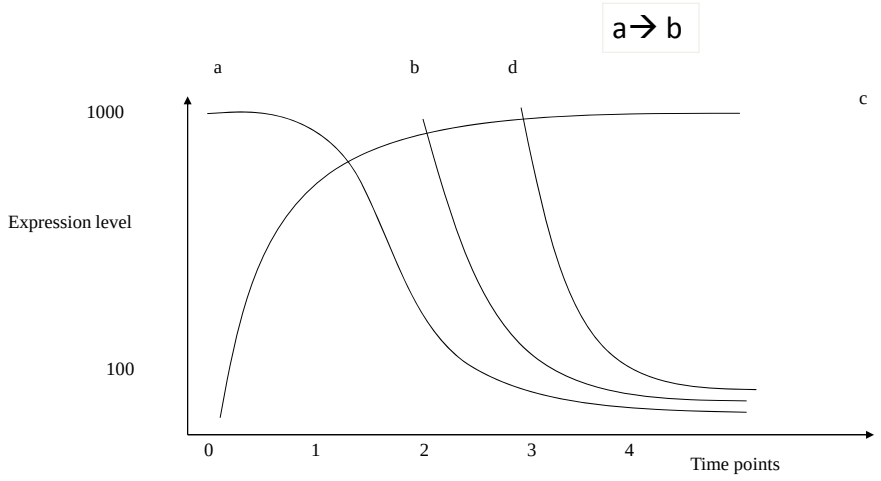
Network Plays a Central Role in Systems Biology



U.S. Department of Energy

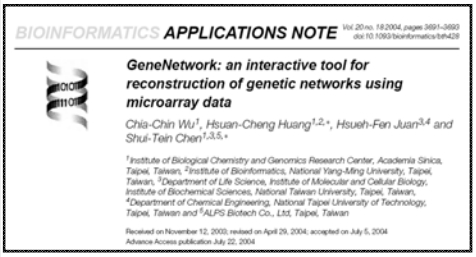
134

Time-course Data can be Used to Infer the Network Structure



135

Gene Network Modeling and Reconstruction


BIOINFORMATICS APPLICATIONS NOTE

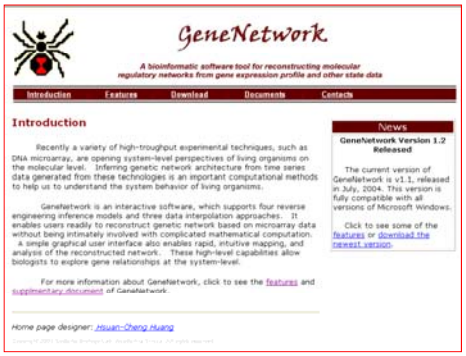
GeneNetwork: an interactive tool for reconstruction of genetic networks using microarray data

Chia-Chin Wu[1], Hsuan-Cheng Huang[1,2,*], Hsueh-Fen Juan[3,4] and Shui-Tein Chen[1,3,5,*]

[1]Institute of Biological Chemistry and Genomics Research Center, Academia Sinica, Taipei, Taiwan, [2]Institute of Bioinformatics, National Yang-Ming University, Taipei, Taiwan, [3]Department of Life Science, Institute of Molecular and Cellular Biology, Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan, [4]Department of Chemical Engineering, National Taipei University of Technology, Taipei, Taiwan and [5]ALPS Biotech Co., Ltd, Taipei, Taiwan

Received on November 12, 2003; revised on April 29, 2004; accepted on July 5, 2004

Bioinformatics 2004, 20, 3691.


GeneNetwork

A bioinformatic software tool for reconstructing molecular regulatory networks from gene expression profile and other state data

Introduction Features Download Documents Contacts

Introduction

Recently a variety of high-troughput experimental techniques, such as DNA microarray, are opening system-level perspectives of living organisms on the molecular level. Inferring genetic network architecture from time series data generated from these technologies is an important computational methods to help us to understand the system behavior of living organisms.

GeneNetwork is an interactive software, which supports four reverse engineering inference models and three data interpolation approaches. It enables users readily to reconstruct genetic network based on microarray data without being intimately involved with complicated mathematical computation. A simple graphical user interface also enables rapid, intuitive mapping, and analysis of the reconstructed network. These high-level capabilities allow biologists to explore gene relationships at the system-level.

For more information about GeneNetwork, click to see the features and supplementary document of GeneNetwork.

News

GeneNetwork Version 1.2 Released

The current version of GeneNetwork is v1.1, released in July, 2004. This version is fully compatible with all versions of Microsoft Windows.

Click to see some of the features or download the newest version.

Home page designer: Hsuan-Cheng Huang

http://genenetwork.sbl.bc.sinica.edu.tw/

136

GeneNetwork Software

- Interactive interface for reconstructing gene network based on microarray or proteomics data
- Reconstruction Model
  - Boolean Network
  - Baysian Network
  - Linear Model
  - Simplified S-System

S-system Formalism

- Dynamics of biological systems can be described mathematically by S-System

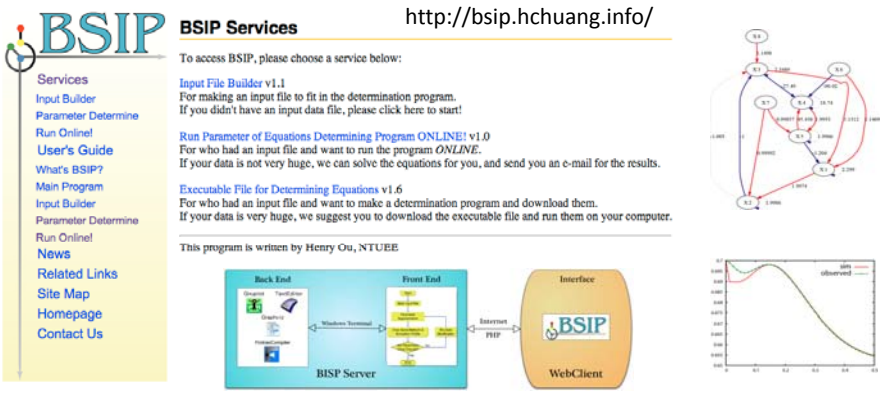
$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \mathbf{p}) = \begin{bmatrix} \alpha_1 \prod_{j=1}^{n+m} X_j^{g_{1j}} - \beta_1 \prod_{j=1}^{n+m} X_j^{h_{1j}} \\ \vdots \\ \alpha_n \prod_{j=1}^{n+m} X_j^{g_{nj}} - \beta_n \prod_{j=1}^{n+m} X_j^{h_{nj}} \end{bmatrix}, \mathbf{X}(0) = \mathbf{X_0}$$

Bioinformatics, 2005, 7, 1180-8.

138

BSIP

- Biological network modeling based on S-system
- Network reconstruction from time-series data

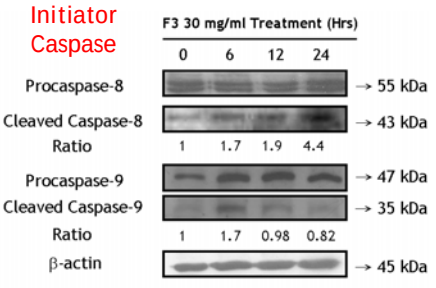


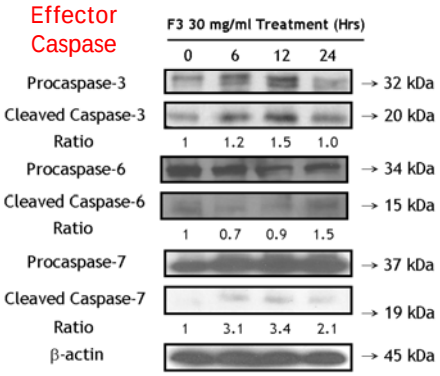
http://bsip.hchuang.info/

APBC2006 Poster Award
EITC 2006

139

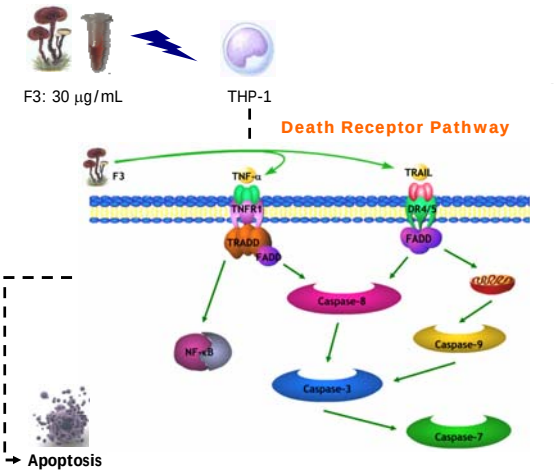
Caspase Cleavage





140

Death Receptor Pathway



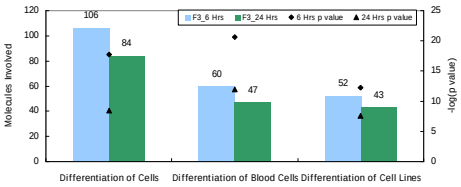
141

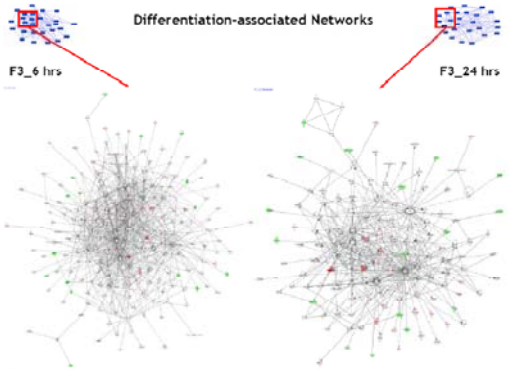
CELL DIFFERENTIATION

142

Biofunction: Cell Differentiation

- Genes associated with cell differentiation were showed by IPA web tool.

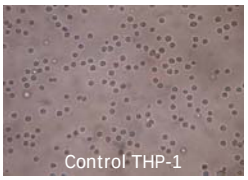


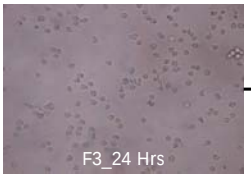


Protein-protein Interaction Network

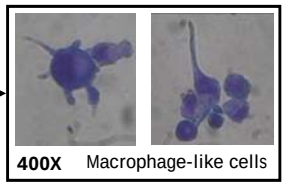
F3 Treatment: 30 μg/mL

Cell Morphology

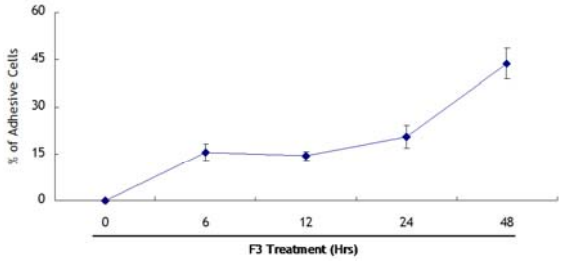




100X



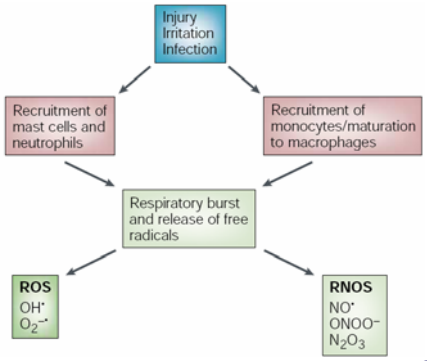
400X Macrophage-like cells

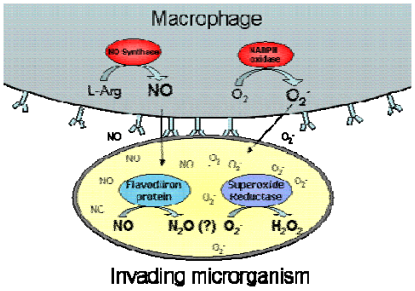


- Adhesion cells were increased in F3-treated THP-1 cells.
- **Macrophage-like cells** were shown under the microscope by Liu's staining.

Macrophage and Superoxide

- Superoxide and NO are toxic radicals produced by macrophages to kill micro-organisms upon infections.

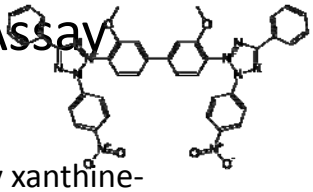




S. Perwez H. et al. (2003) *Nat. Rev. Cancer*[145]

NBT Reduction Assay

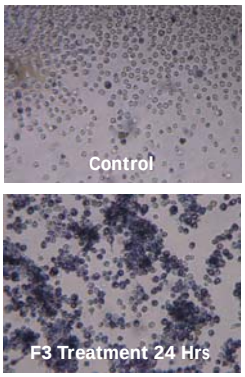
NBT Structure



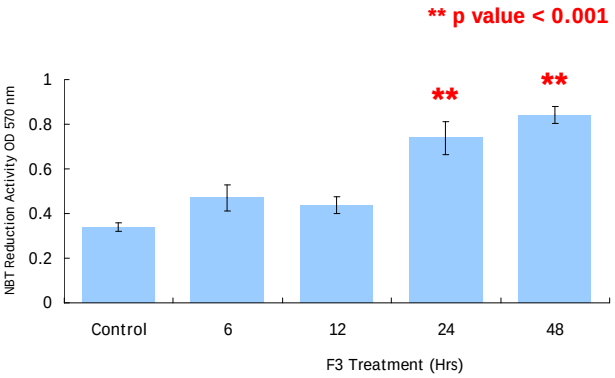
- Superoxide anion was generated by xanthine-xanthine oxidase and detected by the NBT reduction.
- NBT (nitroblue tetrazolium) is a reagent that can absorb superoxide and change it color to purple (absorbed at 570nm).
- The soluble yellow form of NBT was reduced by superoxide to form insoluble, **blue formazan** precipitates.

146

NBT Assay: Superoxide Anion Formation

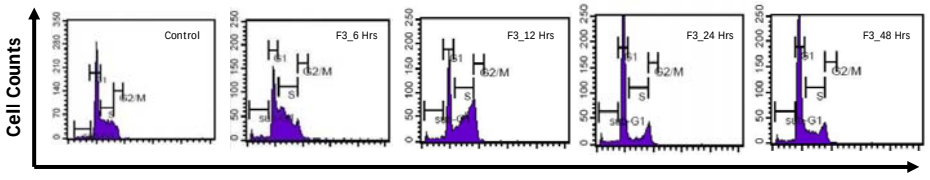


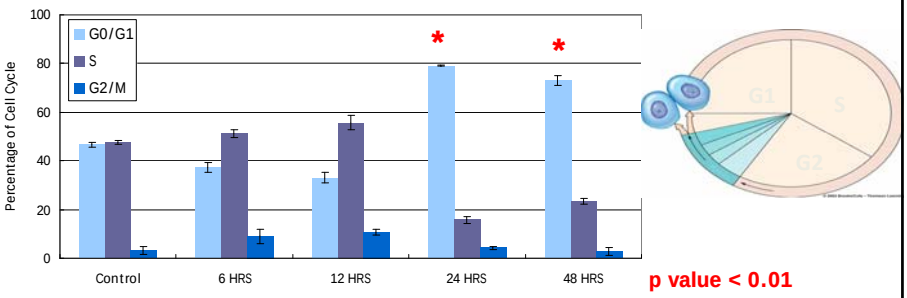
100X



- NBT assay showed insoluble and blue formazan precipitates obviously in F3-treated THP-1 cells.
- F3 treatment was able to increase NBT reduction significantly and might induced **differentiation** of THP-1 cells into macrophages.

Cell Cycle Analysis





p value < 0.01

- F3 could enhance **G0/G1 arrest** in F3-treated THP-1 cells for 24 and 48 hrs.

148

Macrophage Differentiation Markers

- Macrophage differentiation is associated with increased expression of several genes, which are critical for the functionality of macrophage.

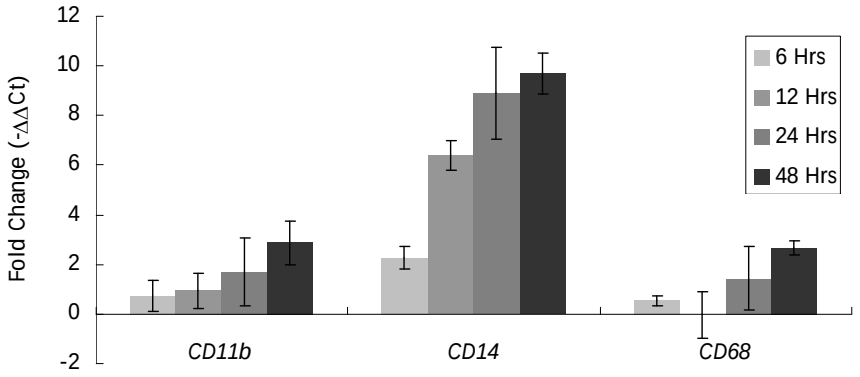
- CD11b, CD14, CD68, MPO, and MMP9 are classical markers of macrophage differentiation.

Aordet O. et al. *Blood* 15: 4446-53 (2002).
Ding Q. et al. *J. Leukoc. Biol*. 81:1568-76 (2007).
Jung H.S. et al. *Inflamm. Res.* 56:45-50 (2007).

149

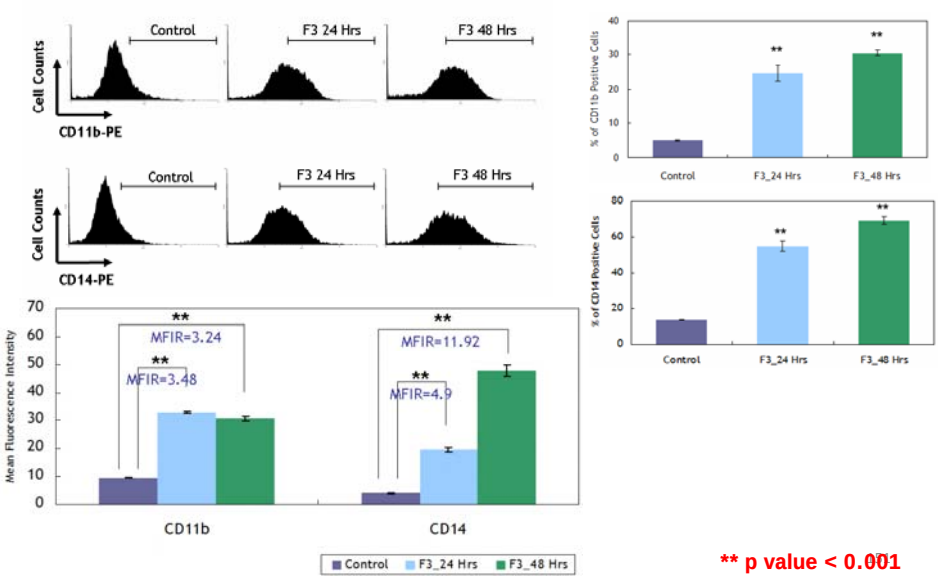
Macrophage Cell Markers

- **CD11b**, **CD14**, and **CD68** are classical markers of macrophage differentiation.



150

Macrophage Cell Surface Markers



MPO: Myeloperoxidase

- MPO is present in the granule, which is a major component of lysosome.
- The enzyme MPO is synthesized only in granulocytes and monocytic cells, making it an important marker of myeloid lineage.
- Decrease in the activity of MPO is a characteristic feature observed as monocytic cells differentiate into macrophages.

Koeffler H. P. et al. *J. Clin. Invest*. 66:1101-8 (1980).
Lin K.M. et al. *Leukemia*. 16:1143-53 (2002).
Shiney S. J. et al. *Mol Cel Biochem*. 233: 9-17 (2002).

152

MMP-9: Matrix Metalloproteinase-9

- MMP-9, a Zn^{2+}-dependent secreted type IV collagenase, can degrade extracellular mactrix components such as collagnes and elastins.
- MMP-9 expression increases as blood monocytes differentiate into macrophages.

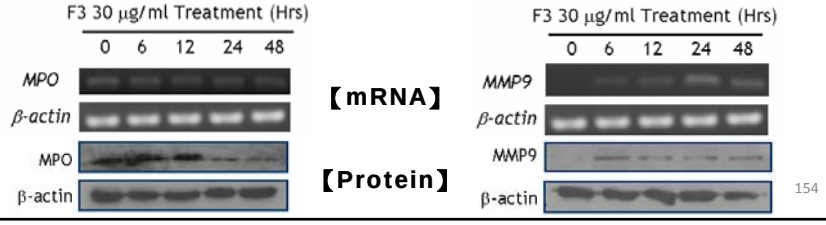
Michael H. et al. *Biol Chem*. 387:69-78 (2006)
Nicole. L. W. et al. *J. Leuko Biolo*. 80:1052-66 (2007).

153

Macrophage Specific Enzyme

- F3-induced macrophage differentiation is associated with increased expression of CD11b, CD14, and CD68, while intracellular MPO decreased.
- F3-induced macrophage differentiation is associated with slightly increased expression of MMP-9, an important mediator of macrophage chemotactic activity.

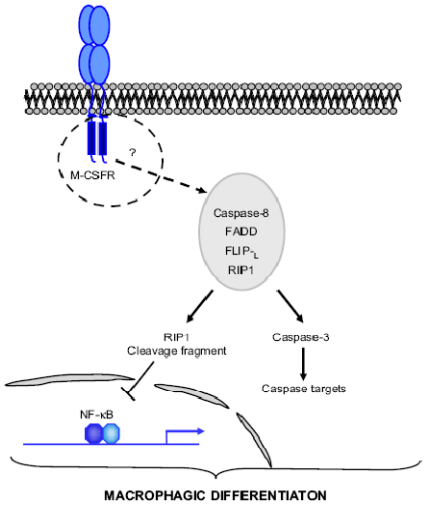
F3 Treatment Hrs	*MPO* ΔCt ± S.D	*MMP-9* ΔCt ± S.D
0	15.5 ± 0.99	N/A
6	15.2 ± 0.1	22.48 ± 0.1
12	17.83 ± 0.98	19.83 ± 1.12
24	16.66 ± 0.5	17.9 ± 0.53
48	16.1 ± 0.15	14.83 ± 0.47



154

Specific Involvement of Caspases in Differentiation into Macrophages

- Alteration of this pathway might account for the accumulation of monocytes in the bone marrow and lead to leukemia formation.
- Differentiation-associated caspase activation is highly specific of the cell type.

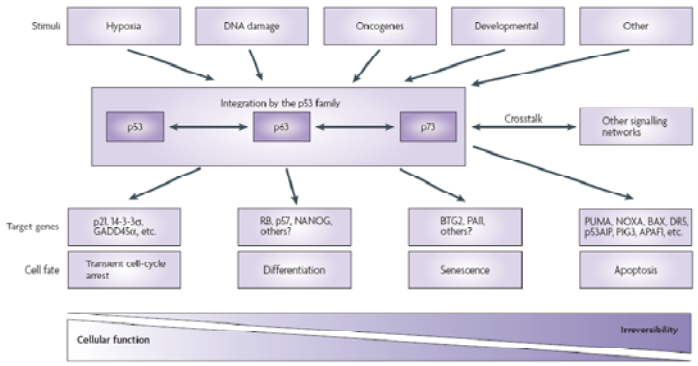


Nathalie Droin et al. *Biochimie* 1:1-7 (2007)
Olivier S. et al. *Blood* 15: 4446-53 (2002).

155

p53 Family

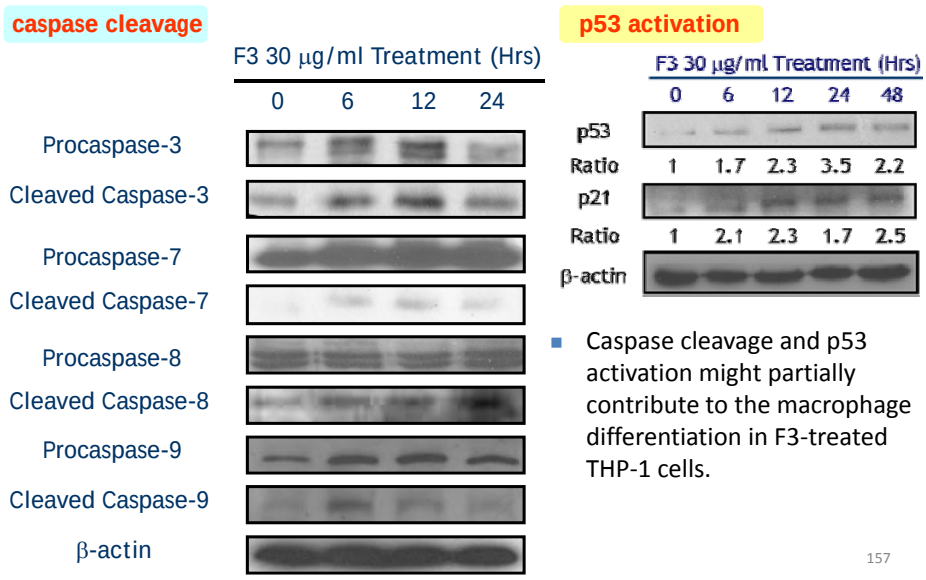
- The role of p53 as a tumor suppressor is generally attributed to its ability to stop the proliferation by inducing cell cycle arrest or apoptosis.
- The p53 also have essential functions in differentiation control.



Thurston S. *Nature Rev. Cancer* 7:165-168 (2007).

156

Regulation of Differentiation into Macrophage

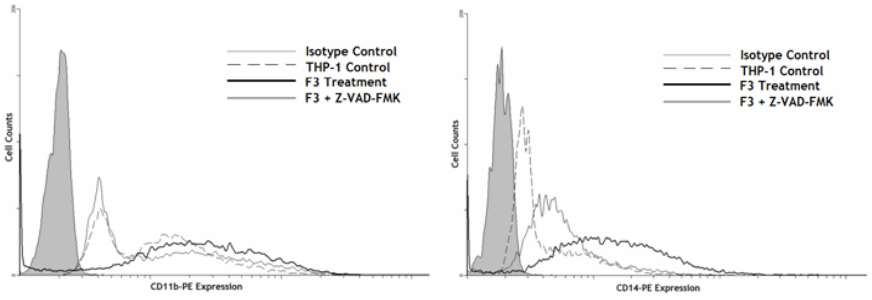


- Caspase cleavage and p53 activation might partially contribute to the macrophage differentiation in F3-treated THP-1 cells.

157

Caspase Inhibitor: Z-VAD-FMK

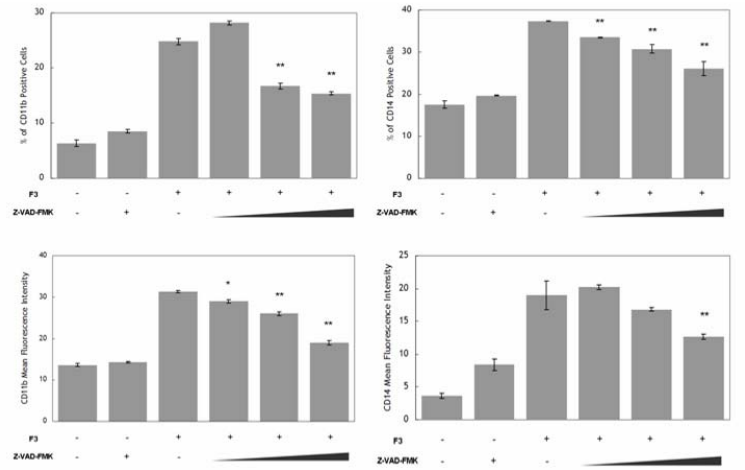
- The general caspase inhibitor, Z-VAD-FMK, prevented the differentiation of F3-induced THP-1 cells into macrophages in a dose-dependent manner and blocked the appearance of CD11b and CD14 on their plasma membrane.



Regulation of Differentiation into Macrophage: Caspase Cleavage

Caspase Inhibitor: Z-VAD-FMK

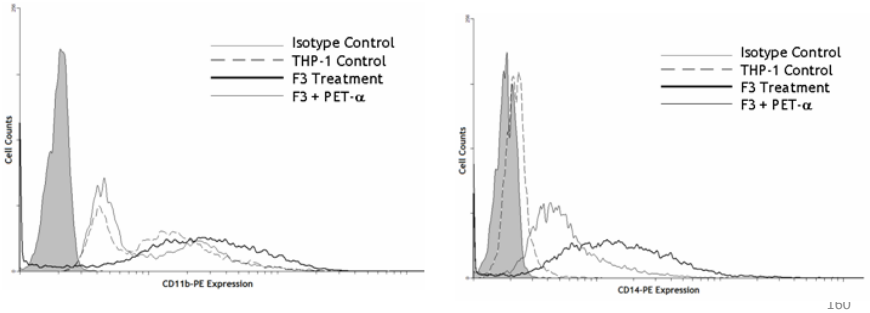
- The caspase cleavage is specifically involved in the macrophage differentiation process.



159

p53 Inhibitor: Pifithrin-α

- The p53 inhibitor, pifithrin-α, also prevented the differentiation of F3-induced THP-1 cells into macrophages in a dose-dependent manner and blocked the appearance of CD11b and CD14 on their plasma membrane.

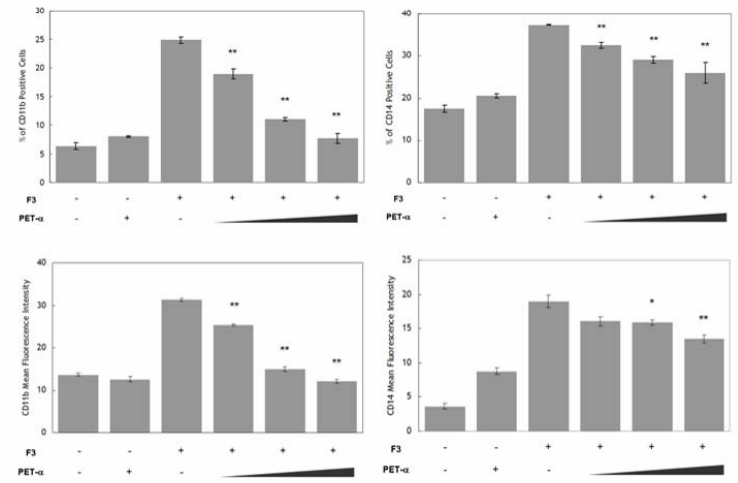


160

Regulation of Differentiation into Macrophage: p53 activation

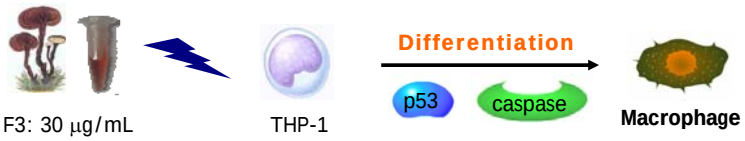
p53 Inhibitor: Pifithrin-α

- Expression of p53 is involved in the induction of differentiation specific markers.



161

Cell Differentiation to Macrophage

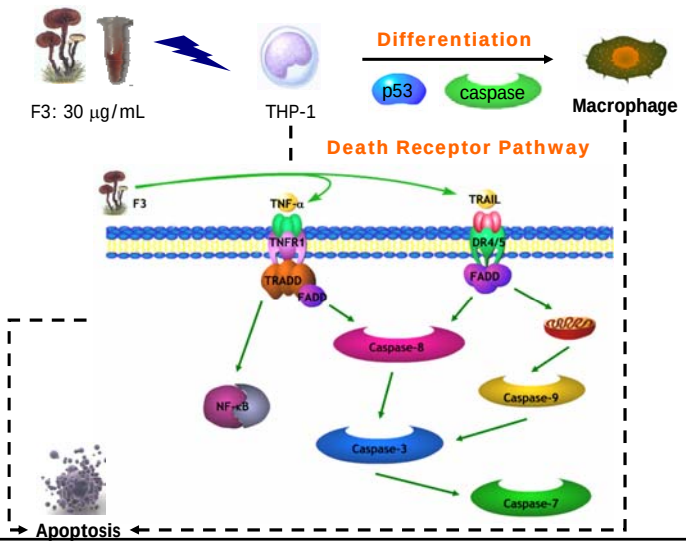


162

SUMMARY

163

The anti-leukemia molecular mechanism induced by F3



164