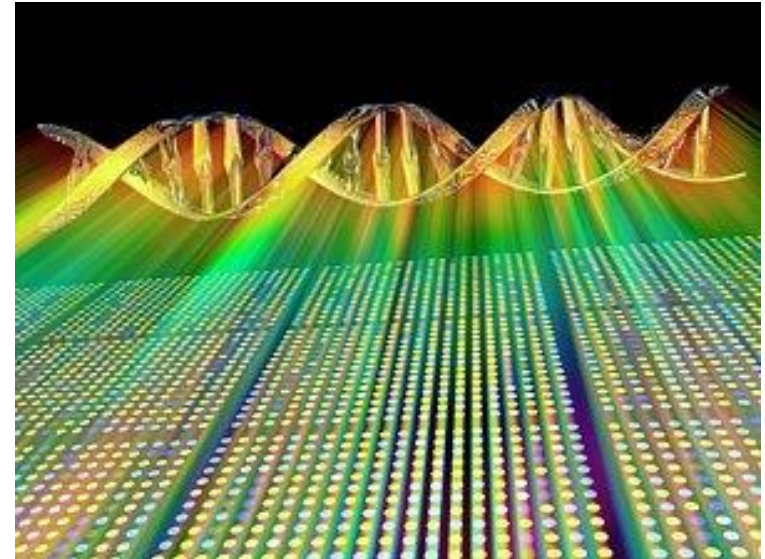
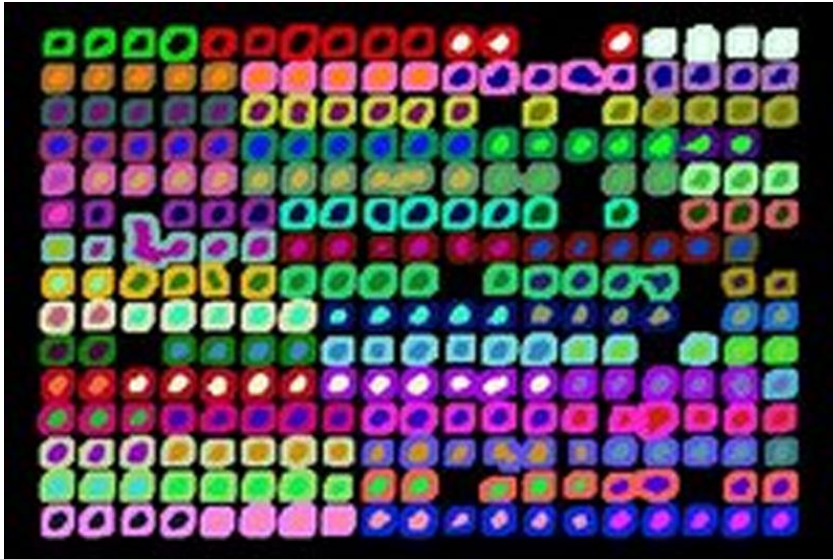




# DNA MICROARRAY

“Gene Expression Profiling in Health and Disease”



Hans Bluysen  
04-11-2020




Classes in period "Academic year 2022/2023, winter semester" (in progress)		
Time span:	2022-10-01 - 2023-02-26	
Type of class:	classes, 20 hours <a href="#">more information →</a> Sanaz Hasani discussion seminar, 10 hours <a href="#">more information →</a> Hans Bluysen lecture, 15 hours <a href="#">more information →</a> Hans Bluysen Joanna Wesoly	 
Coordinators:	Johannes Bluijssen	
Group instructors:	Johannes Bluijssen, Mohammad Emad Olya, Joanna Wesoly	
Students list:	<a href="#">view the list/edit groups →</a>	
Examination:	Course - Graded credit classes - Graded credit discussion seminar - Graded credit lecture - Graded credit	

**Chosen plan division:**

☐ this week

☒ course term

	MO	TU	WE	TH	FR
7					
8					
9					
10					
11					
12					
13					

 [see course schedule](#)



<https://dhmg.amu.edu.pl/teaching/>

## Methods in Molecular Diagnostics

**Coordinator:** Hans Bluysen

**Exercises:** [Sanaz Hassani](#)

**Journal Club:** Hans Bluysen

**Language:** English

**Exam:**

**Lectures:**

1. Gene expression profiling technologies (Microarrays, RNAseq)
2. Gene expression signatures in cancer Diagnostics, prognostics and predictive medicine.
3. Chromatin binding technologies (ChIPseq)
4. Diagnostics and Prognostics of Kidney Cancer
5. Diagnostics and Therapeutics of Cardiovascular Disease
6. [Pre-natal Diagnostics \(Genetics and Targeted Gene Sequencing-based techniques\)](#)
7. [Clinical Diagnostics: Multiplex Real-time PCR and NGS-based assays](#)

Hans Bluysen

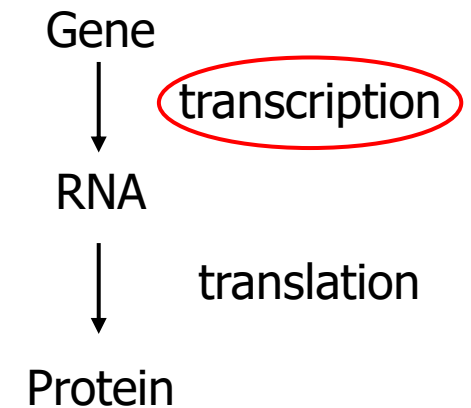
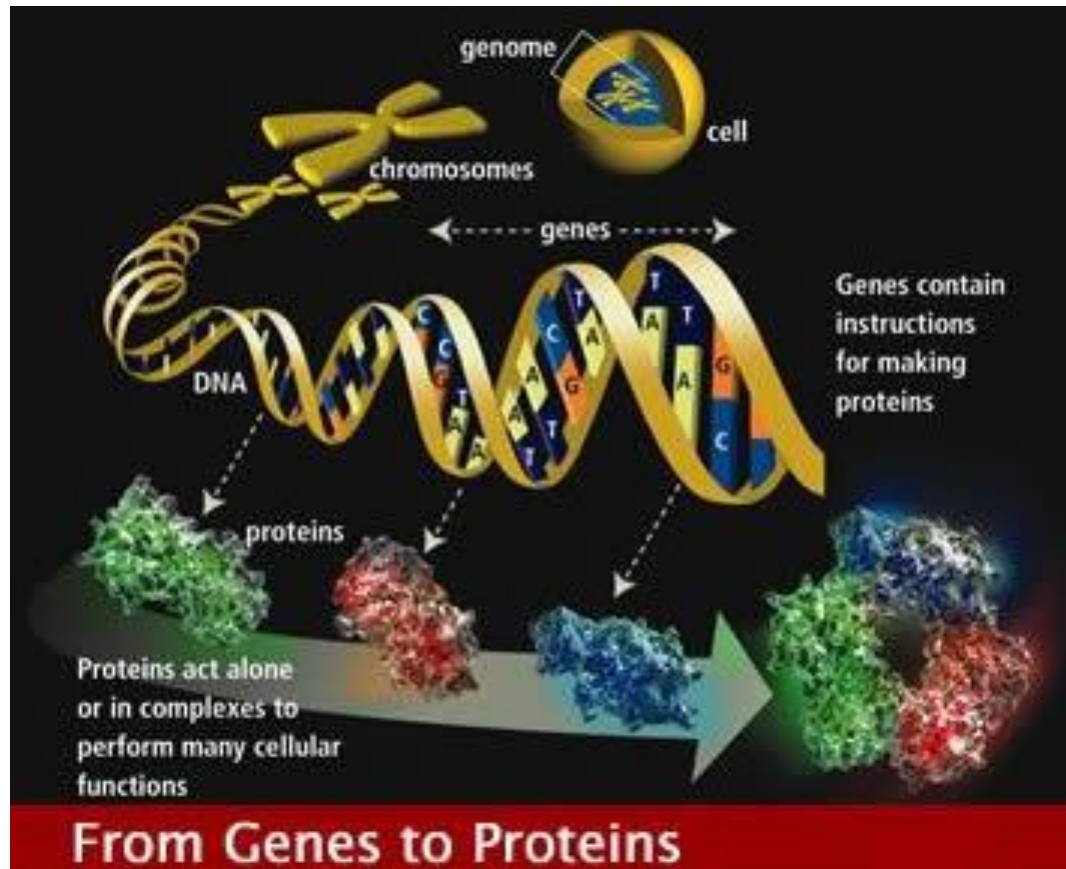
[Joanna Wesoly](#)



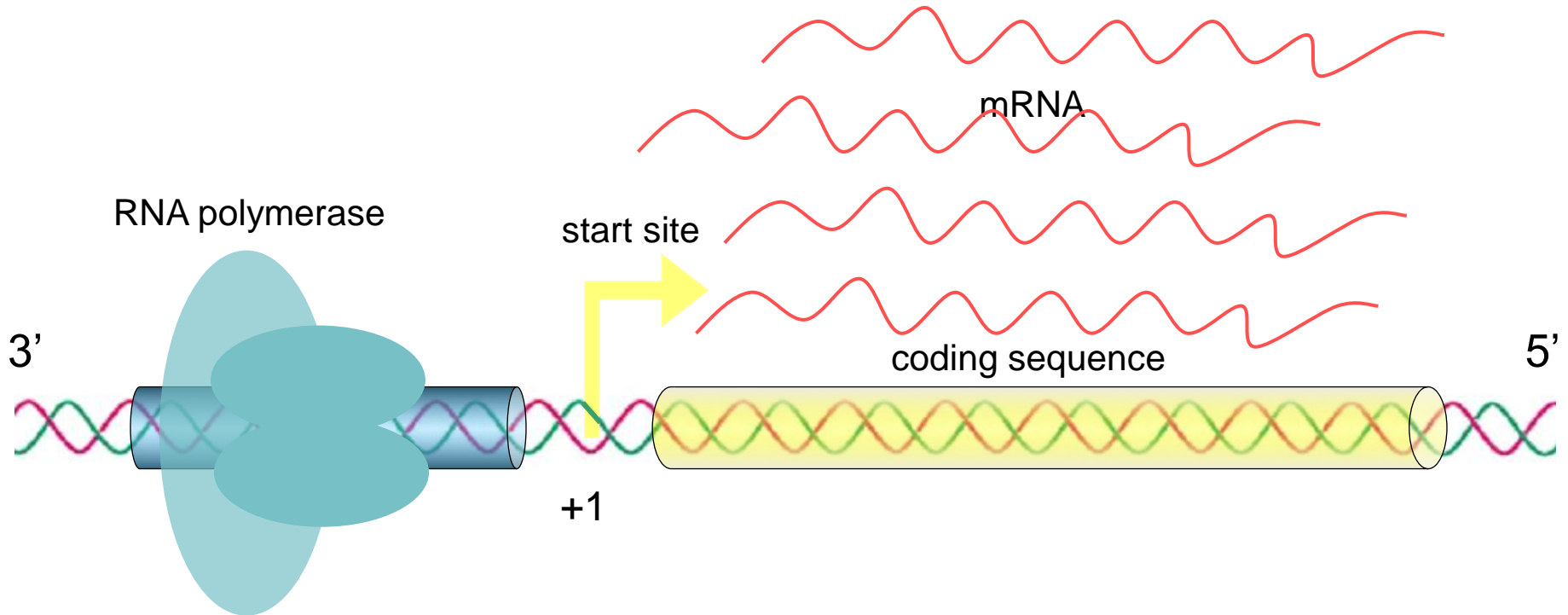
## Students

Name ↓	Surname	E-mail	Index	A
Anastasiia	Kompaniets	<a href="mailto:anakom@st.amu.edu.pl">anakom@st.amu.edu.pl</a>	553435	
Anastasiia	Romanenko	<a href="mailto:anarom1@st.amu.edu.pl">anarom1@st.amu.edu.pl</a>	553636	
Anna	Dudnik	<a href="mailto:anndud12@st.amu.edu.pl">anndud12@st.amu.edu.pl</a>	555438	
Antonina	Gonet	<a href="mailto:antgon3@st.amu.edu.pl">antgon3@st.amu.edu.pl</a>	553439	
Arzu	Abbasova	<a href="mailto:arzabb@st.amu.edu.pl">arzabb@st.amu.edu.pl</a>	555421	
Dawid	Kaszuba	<a href="mailto:dawkas2@st.amu.edu.pl">dawkas2@st.amu.edu.pl</a>	521858	
Marianna	Mirecka	<a href="mailto:marmir2@st.amu.edu.pl">marmir2@st.amu.edu.pl</a>	500685	
Mariia	Sokulska	<a href="mailto:marsok9@st.amu.edu.pl">marsok9@st.amu.edu.pl</a>	552224	
Olena	Nahimova	<a href="mailto:olenah@st.amu.edu.pl">olenah@st.amu.edu.pl</a>	552221	
Safa	Öksüz	<a href="mailto:safoks@st.amu.edu.pl">safoks@st.amu.edu.pl</a>	520358	
Yeva	Kovalyk	<a href="mailto:yevkov@st.amu.edu.pl">yevkov@st.amu.edu.pl</a>	552398	

# Genome & Genes



# RNA Transcription

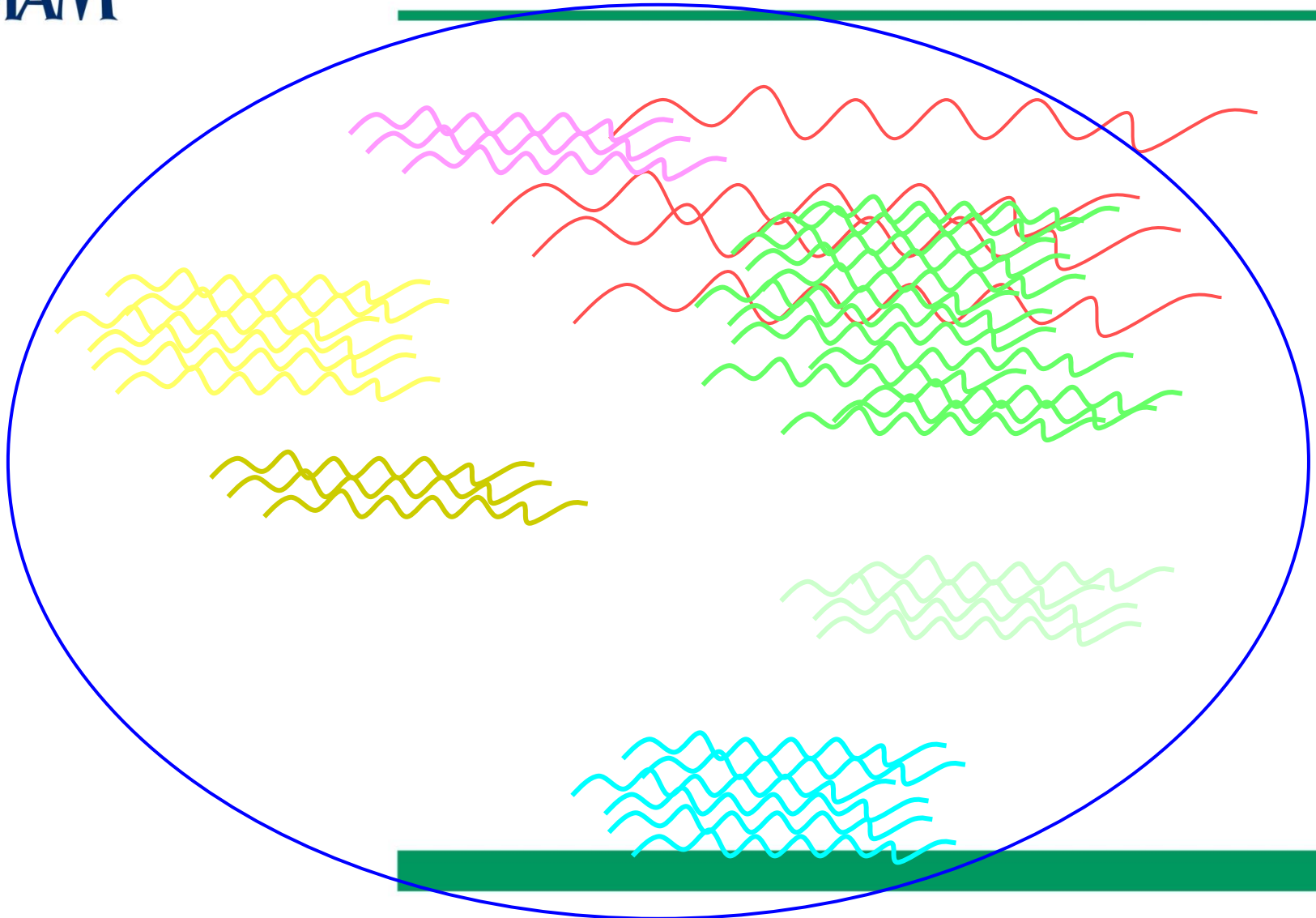


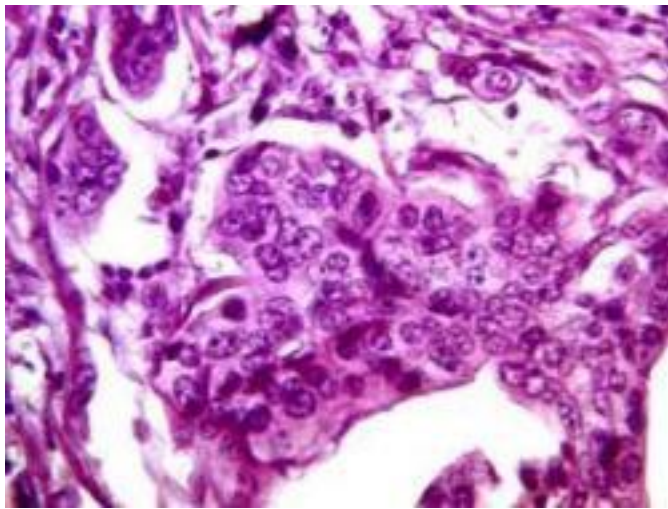
generation of mRNA from genomic DNA





# The Transcriptome

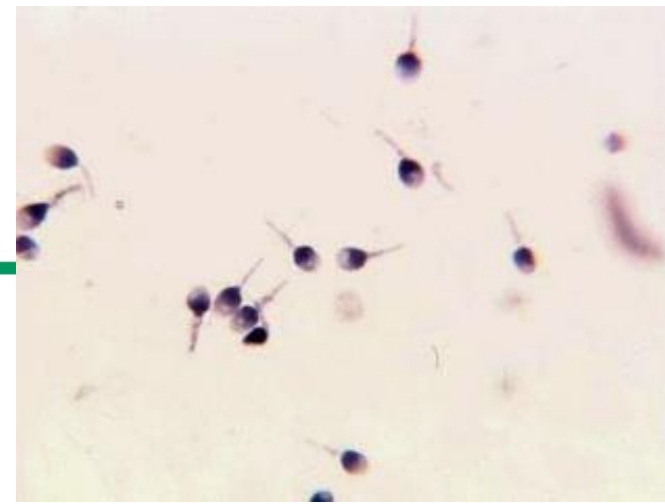




Breast Cancer cells



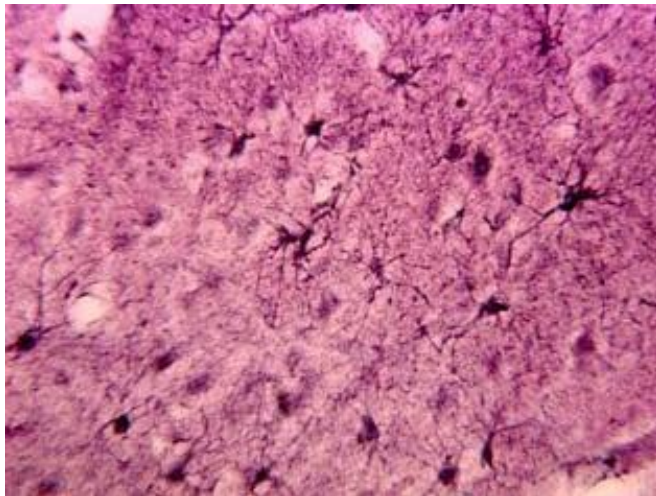
Red Blood cells



Sperm cells

## Transcriptional Program & Cellular Fate

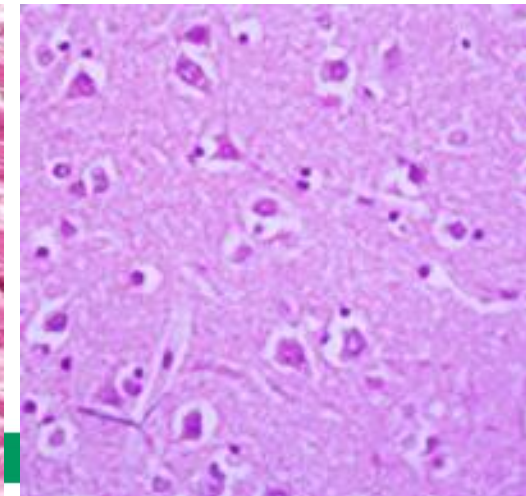
Astrocytes



Smooth Muscle cells



Neuronal cells



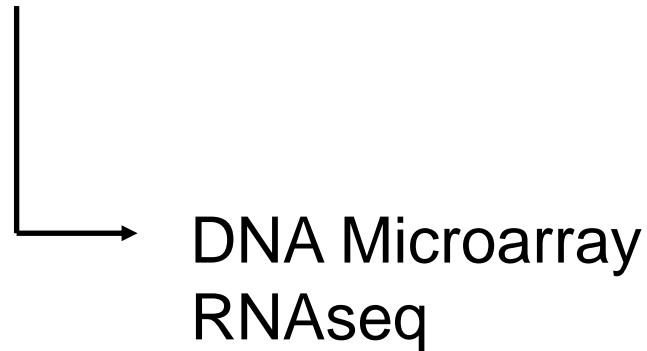




# Transcriptomics

---

Parallel monitoring of relative levels of thousands of mRNA species at one time point or condition:  
expression profiling





# DNA Microarrays: Basics

---

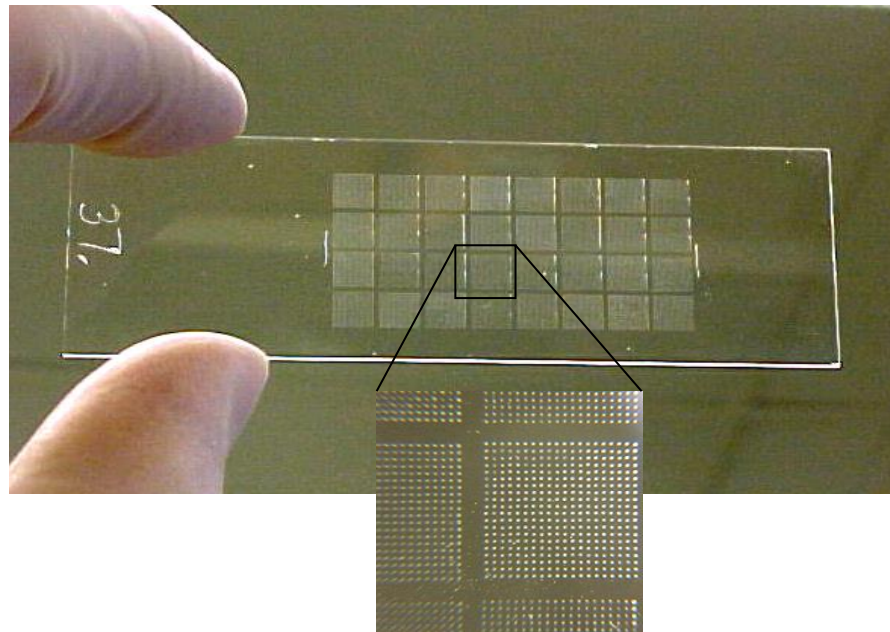
- Put a large number (~30K) of cDNA sequences or synthetic DNA oligomers onto a glass slide (or other substrate) in known locations on a grid.
  - Label an RNA sample and hybridize
  - Measure amounts of RNA bound to each square in the grid
  - Make comparisons
    - Cancerous vs. normal tissue
    - Treated vs. untreated
    - Time course
  - Many applications in both basic and clinical research
-

# What is a DNA microarray?

~2000

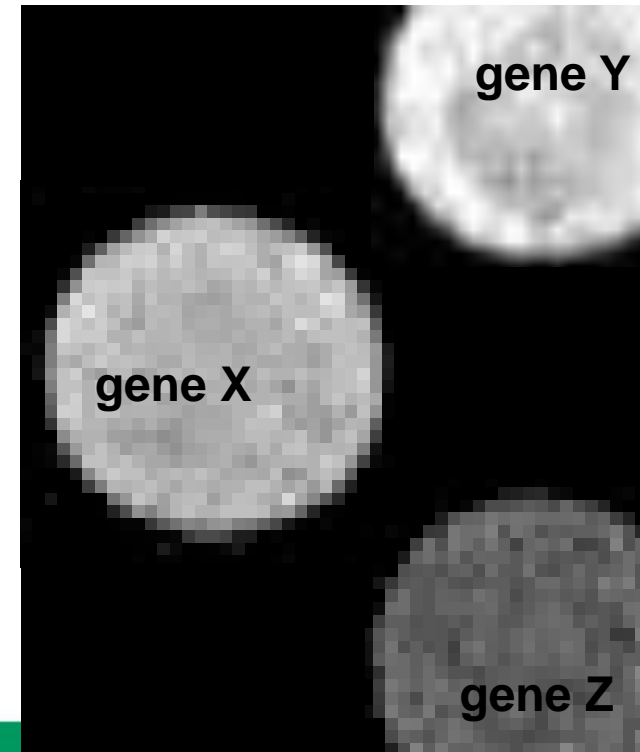
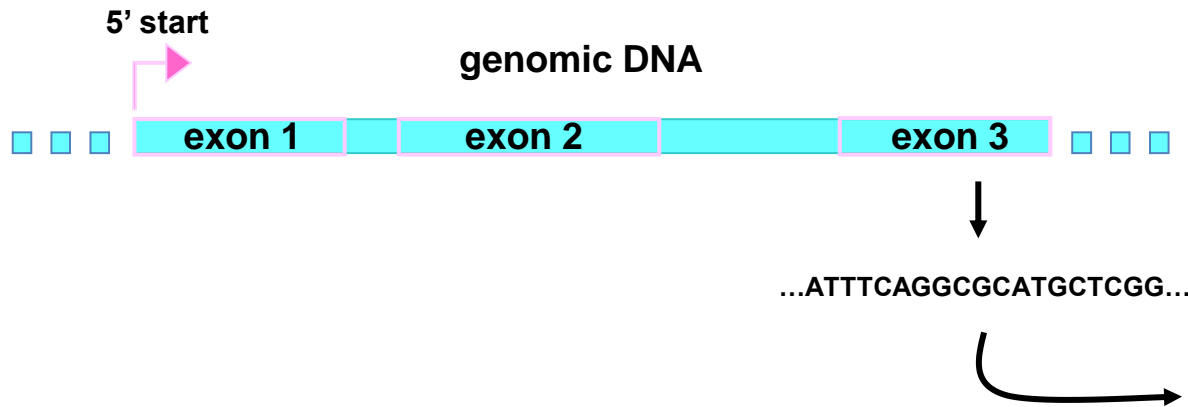
---

DNA microarrays are ordered assemblies of DNA sequences immobilized on a solid support (such as chemically modified glass).



# What is a DNA microarray?

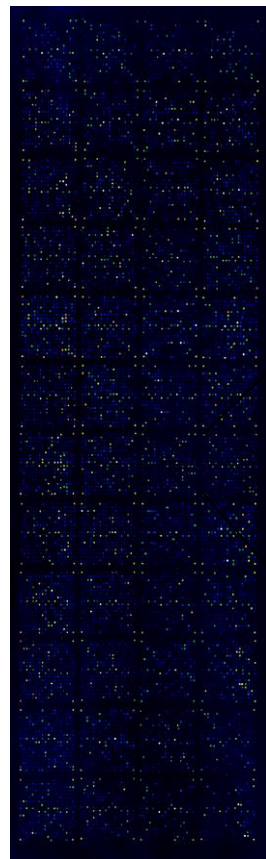
The DNA sequences (e.g. PCR products or oligos) correspond to the transcribed regions of genes.



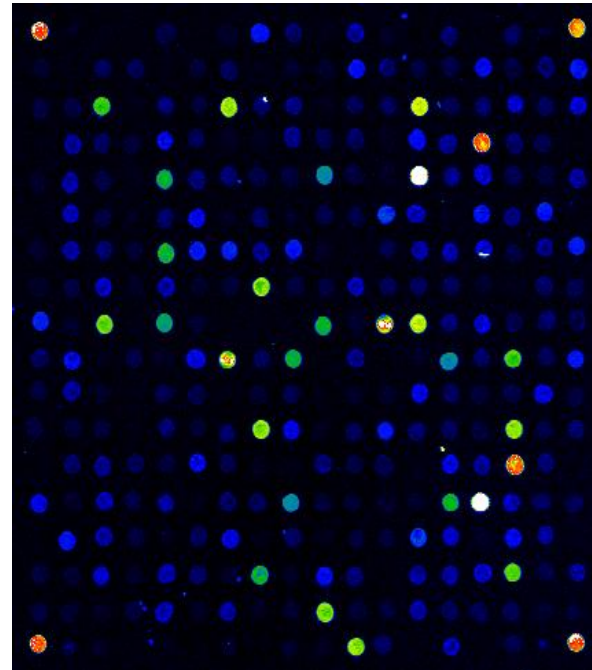
# Long oligo arrays

---

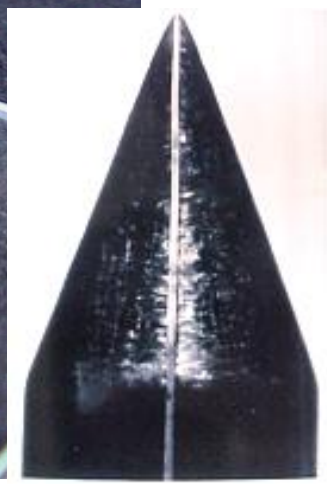
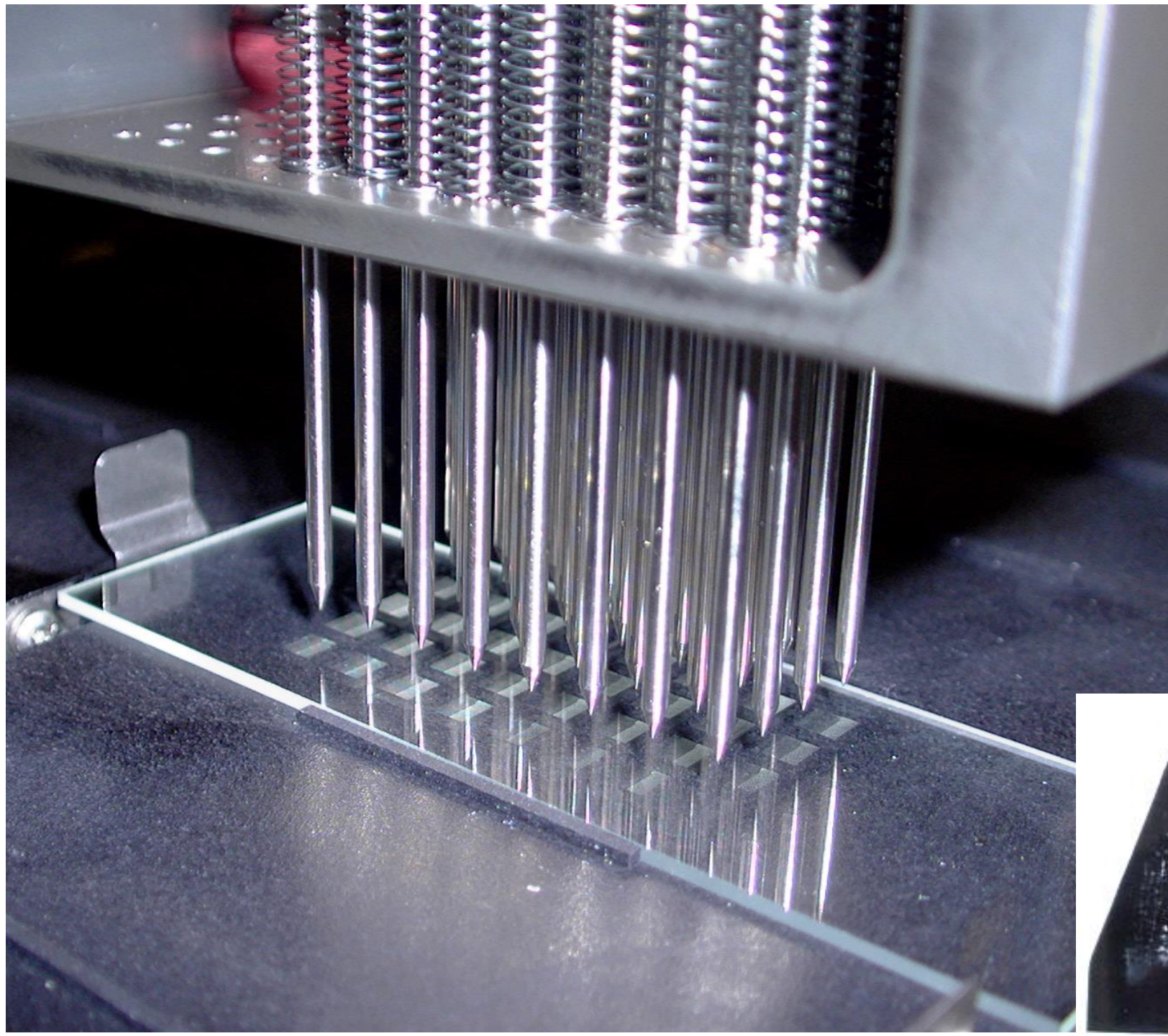
Each gene represented by single long oligo (60 - 70-mer)



Oligo's spotted by robot



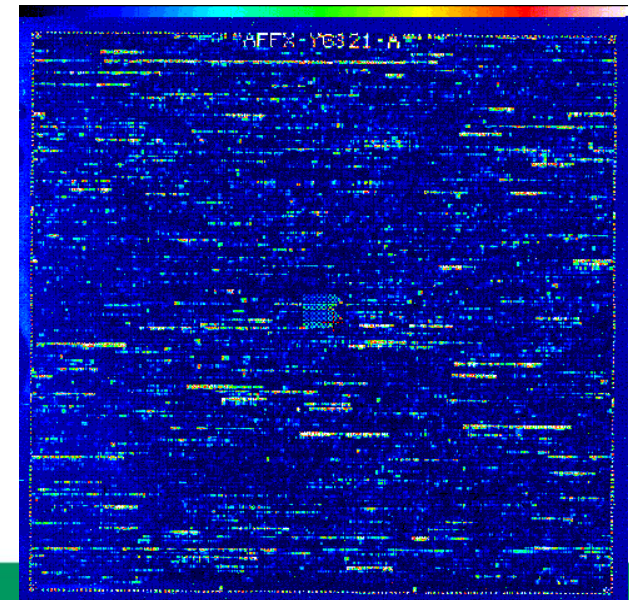
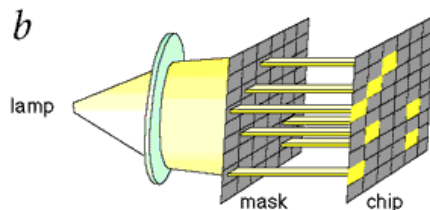
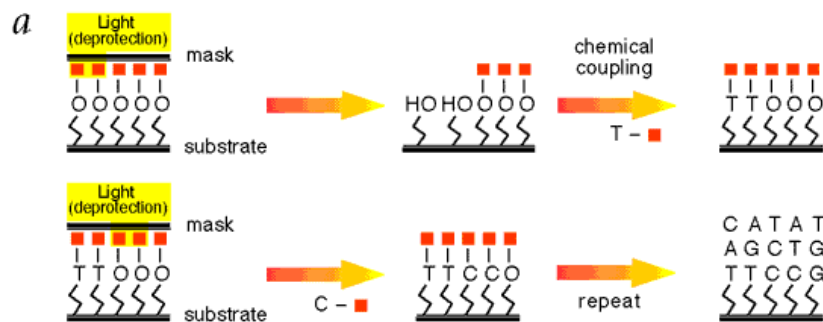




# Affymetrix chips

Each gene represented by 20 different short (25-mer) oligonucleotides and 20 mismatch controls

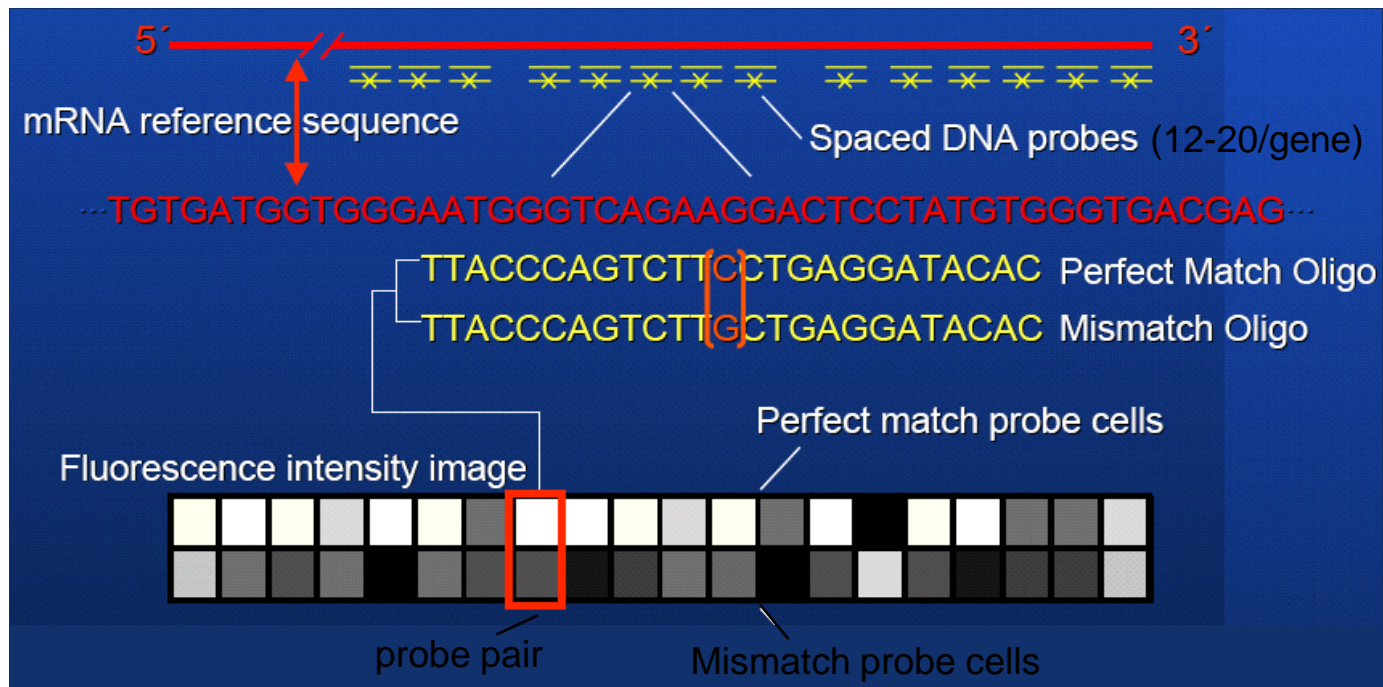
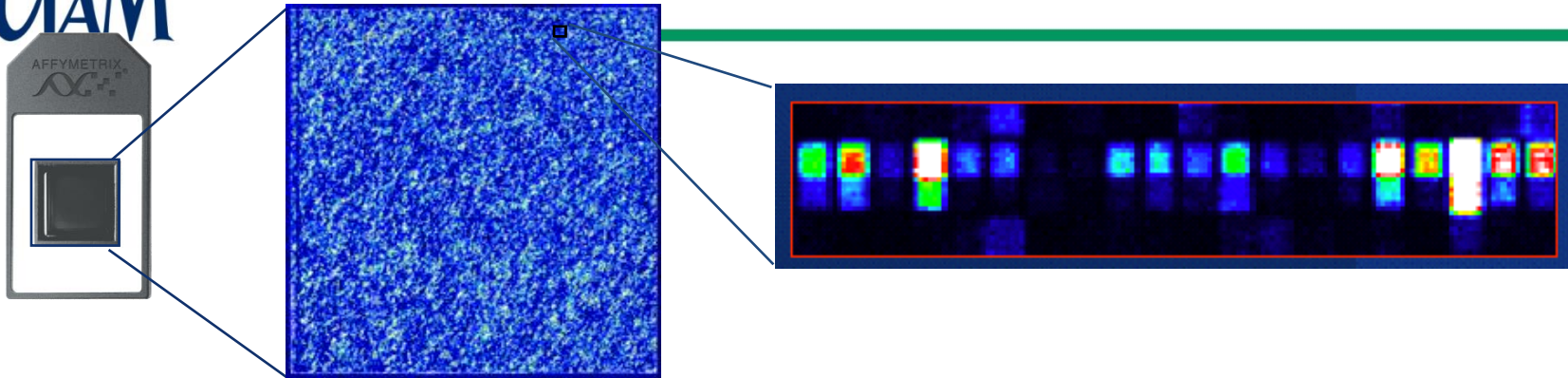
Oligonucleotides synthesized on chip by photolithographic masking







# Affymetrix chips





# Affymetrix chips

ThermoFisher  
SCIENTIFIC

Search All

Search



[Home](#) › [Life Sciences](#) › [Microarray Analysis](#) › Transcriptome Profiling with Microarrays

Re-evaluate what gene expression microarrays can bring to your research

## Transcriptome Profiling with Microarrays

‹ [Microarray Analysis](#)

### Transcriptome Profiling with Microarrays

- Arrays or RNA-Seq?
- Clariom Assays
- MyGeneChip Custom Array Program

[Microarray Data Analysis](#)

[Microarray Instruments, Software & Services](#)

[Microarray Analysis Partners & Programs](#)



Phenotypic abnormalities are rarely a result of expression changes in single genes, so generating a comprehensive expression profile is critical when studying normal biology and disease processes. Additionally, important expression changes, such as differential exon usage resulting from alternative splicing events, may be masked when profiling at the gene-level. Microarrays provide the distinct advantage of assaying millions of distinct sequences in parallel which makes the technique immune to issues detecting and measuring low abundance transcripts, or rare alternative splicing events.

[Request transcriptome profiling project costs ›](#)

For fast RNA expression analysis, we offer a complete range of arrays for whole-transcriptome-, gene-, exon-, or short noncoding (snc)RNA-level analysis. All of our expression arrays are compatible with a wide variety of sample types and accommodate low RNA input. They are available in single-sample array cartridge and multi-sample array plate formats for different throughput needs. They all include our fast, flexible analysis software at no additional cost.

For fast RNA expression analysis, we offer a complete range of arrays for whole-transcriptome-, gene-, exon-, or short noncoding (snc)RNA-level analysis. All of our expression arrays are compatible with a wide variety of sample types and accommodate low RNA input. They are available in single-sample array cartridge and multi-sample array plate formats for different throughput needs. They all include our fast, flexible analysis software at no additional cost.



## Transcriptome profiling solutions

### Clariom Assays

Quickly reveal critical biomarker signatures from coding and long noncoding (lnc)RNA to yield key insights into the complexity of biology with whole-transcriptome array analysis. Clariom D and Clariom S assays (for human, mouse, and rat) are designed for whole-transcriptome expression profiling and biomarker discovery. Built using the latest transcriptomic

Re-evaluate what gene expression microarrays can bring to your research

## Clariom Assays

◀ [Transcriptome Profiling with Microarrays](#)

[Arrays or RNA-Seq?](#)

**Clariom Assays**

[MyGeneChip Custom Array Program](#)

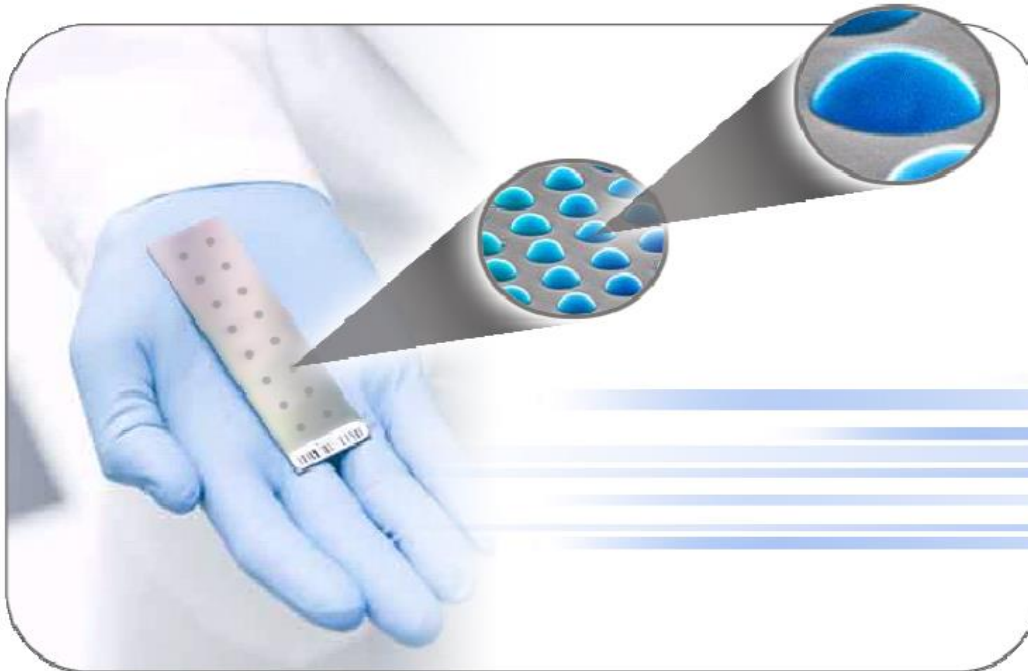


### The power of Clariom assays

Recent advanced transcriptome analyses have uncovered thousands of splice variants and long non-coding (lnc)RNAs, providing new sources for biomarker discovery. Given the complexity of the transcriptome, however, finding informative expression biomarkers is challenging, time-consuming, and costly. Clariom assays, built using the latest transcriptome knowledge from multiple databases, are simple and fast tools for finding high-fidelity expression biomarkers. They are compatible with clinical research samples, available in scalable formats for different throughput needs, and include flexible, intuitive software for fast and simple analysis.



## BeadArray™ Technology

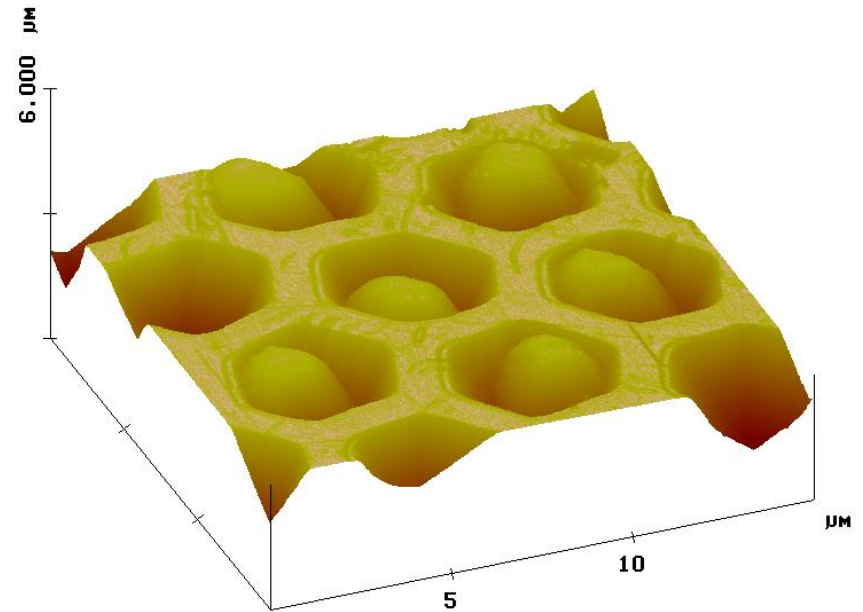
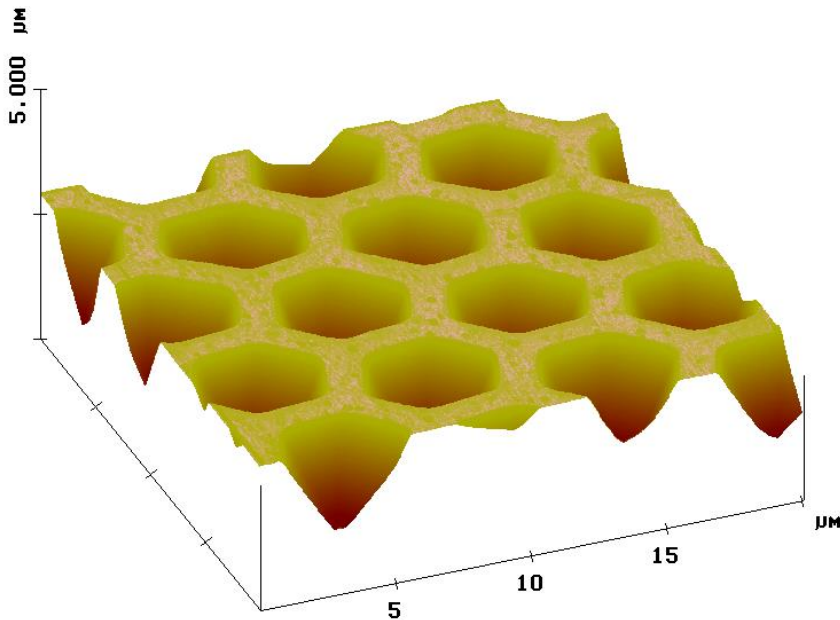


3-micron silica beads that self assemble in microwells on fiber optic bundles or planar silica slides.

When randomly assembled the beads have a uniform spacing of ~5.7 microns.

# Optical fiber with wells...

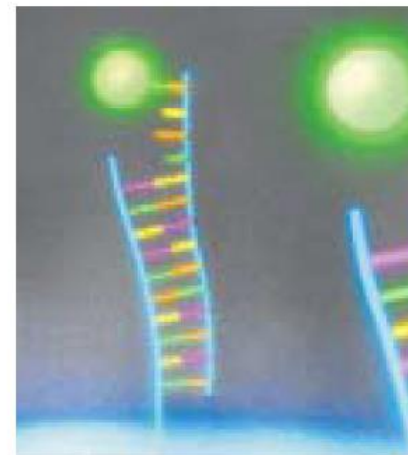
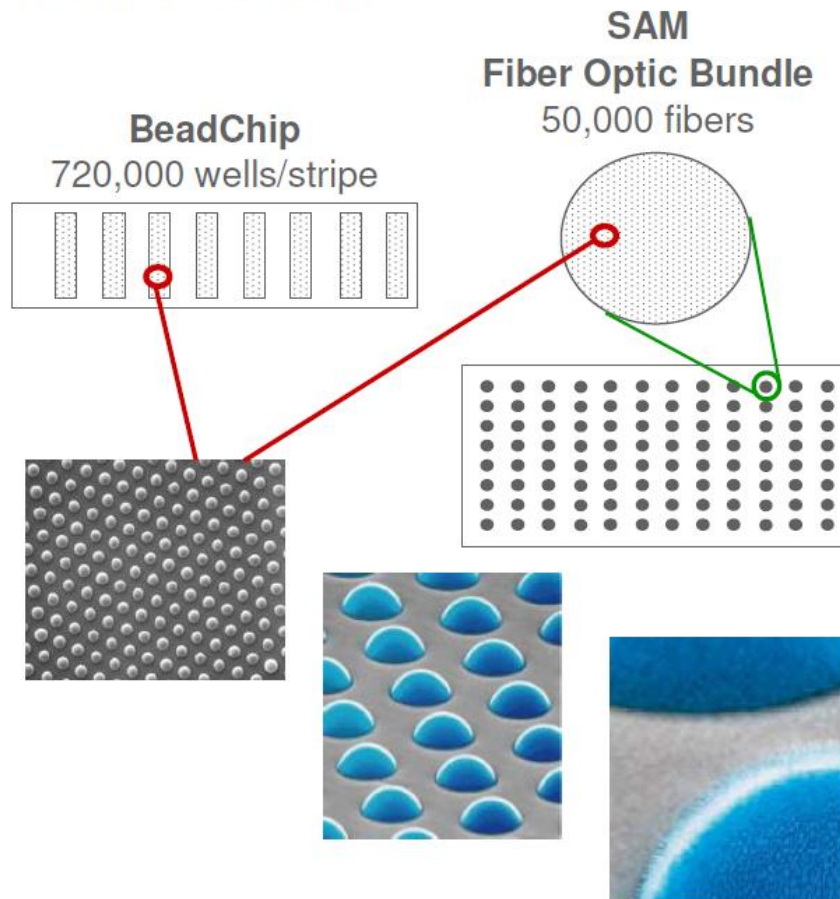
---



... and BEADS!

---

## Sentrix® Beads

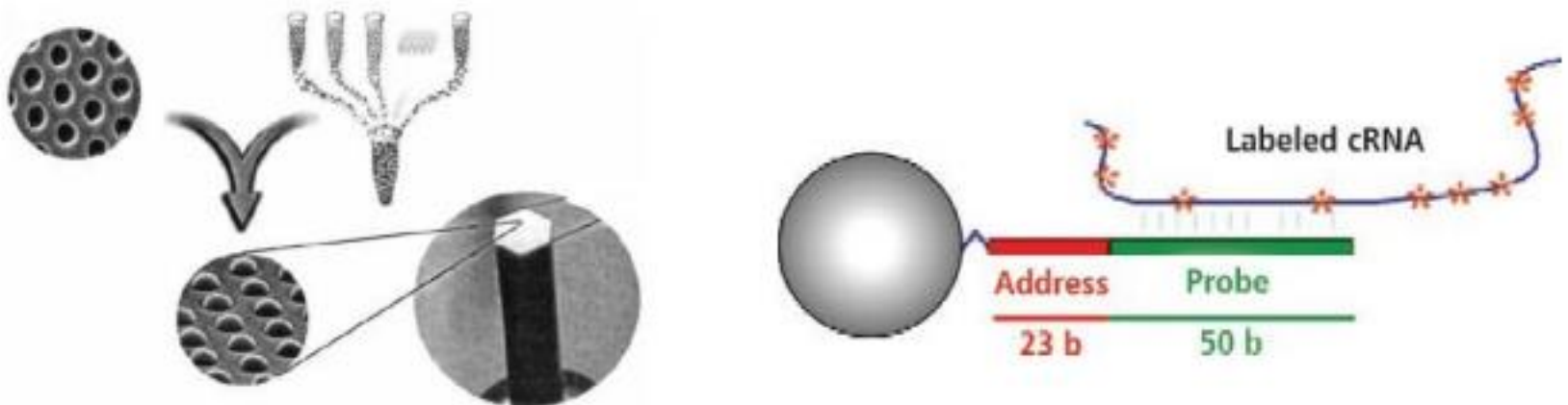


~800,000 copies of  
specific oligo  
per bead

N=20

## Manufacturing the arrays

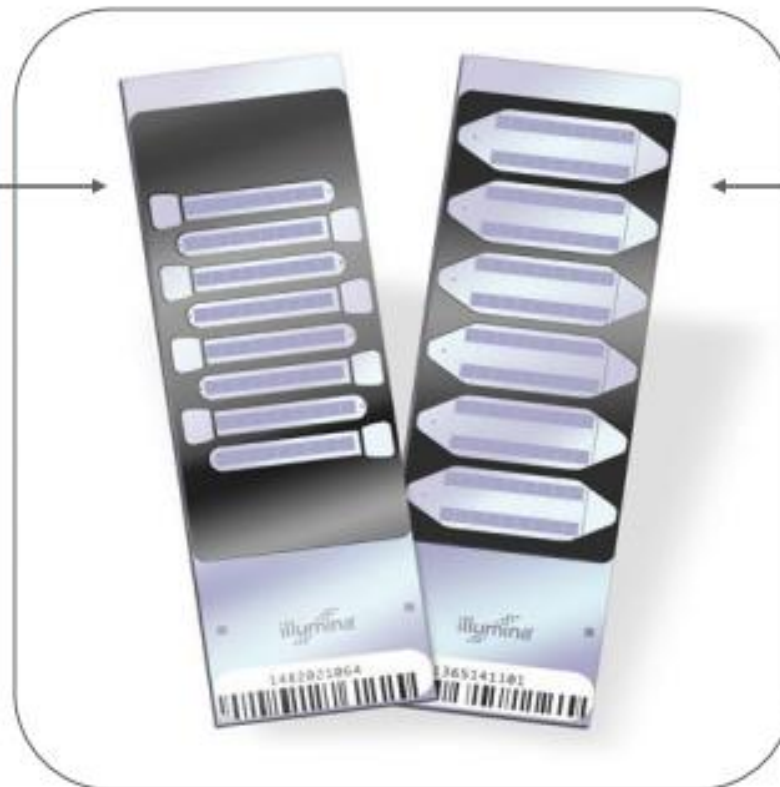
- Long oligonucleotides attached to glass beads
- 50 nucleotide probe
- 23 nucleotide address (bead ID)
- Pooled together in beadpools
- Random assembly in etched wells



## Direct Hyb: Whole Genome Expression Arrays

**HumanRef-8  
Mouse-Ref8  
RatRef-12**

- RefSeq BeadChip
- 8 or 12 samples per slide
- >22,000 sequences



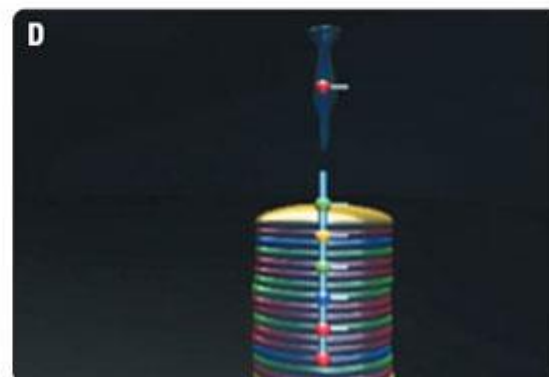
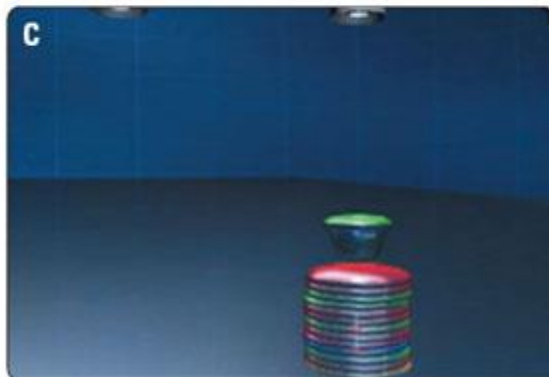
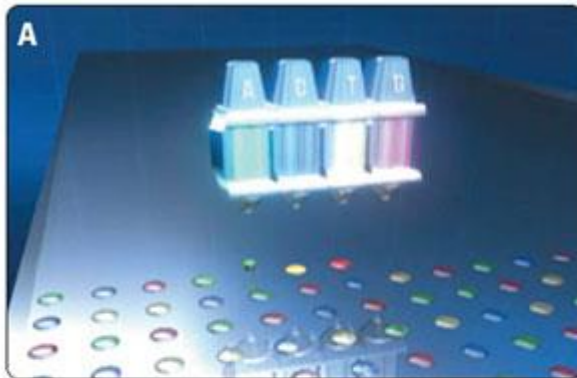
**HumanWG-6  
MouseWG-6**

- Whole Genome BeadChip
- 6 samples per slide
- >46,000 sequences
- >10 million features



# Agilent arrays

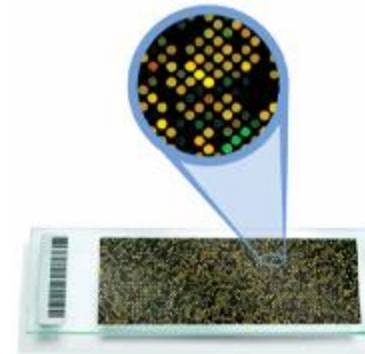
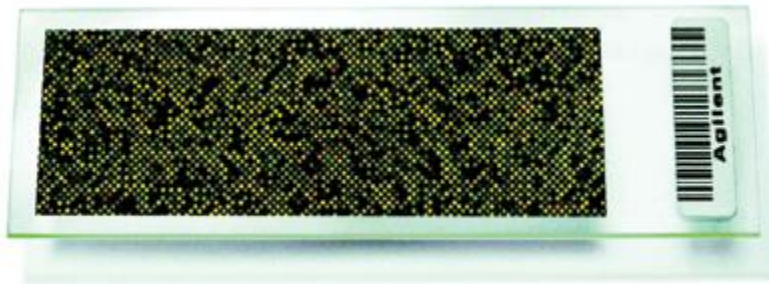
Oligo's synthesized on chip by ink-jet printing



# Agilent arrays

---

Each gene represented by single long oligo  
(60 - 70-mer)





# Agilent arrays

## Gene Expression & Exon Microarrays

Discover our large selection of Gene Expression & Exon Microarrays. This includes whole transcriptome gene expression for almost 30 different species, Exon microarrays to analyze splicing variants and gene expression microarrays with comprehensive content, including full LNCipedia databases to provide full coverage of the transcriptome in a single experiment.

### Gene Expression & Exon Microarrays

### Products

Refine By

#### — Products

Animal Microarrays (10)

Bacteria & Fungi  
Microarrays (3)

Custom Microarrays (2)

Human Microarrays (2)

Model Organism  
Microarrays (11)

Plant Microarrays (8)



#### Animal Microarrays (10) >

Learn about our selection of high-quality Animal Microarrays for your research and workflows. Choose between a wide range of species of interest, including livestock, pets and other species of interest for genetic ...



#### Bacteria & Fungi Microarrays (3) >

Learn about our high performing bacterial and fungi microarrays for your research and workflow needs, including the E. coli Gene Expression Microarrays and the Magnaporthe Gene Expression ...



#### Custom Microarrays (2) >

Discover the flexibility of the Agilent platform: multiple array formats, design gene expression or splicing variants slides, unlimited customization possibilities, and no minimum order. Boost your research ...



#### Human Microarrays (2) >



#### Model Organism Microarrays (11) >



#### Plant Microarrays (8) >

# ILLUMINA: RNAseq

~2005>>

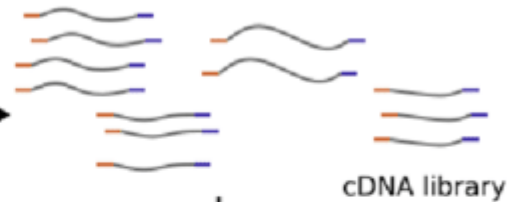
Sample of interest



Extract total RNA  
and enrich targets



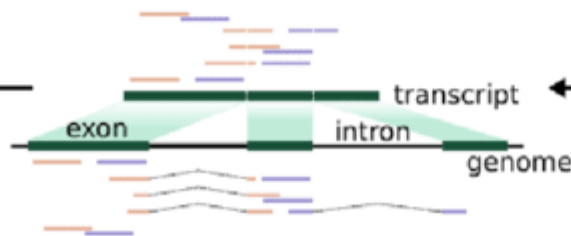
Fragment, reverse transcribe,  
ligate adapters, amplify



Data analysis

- differential expression
- variant calling
- annotation
- novel transcript discovery
- RNA editing
- ...

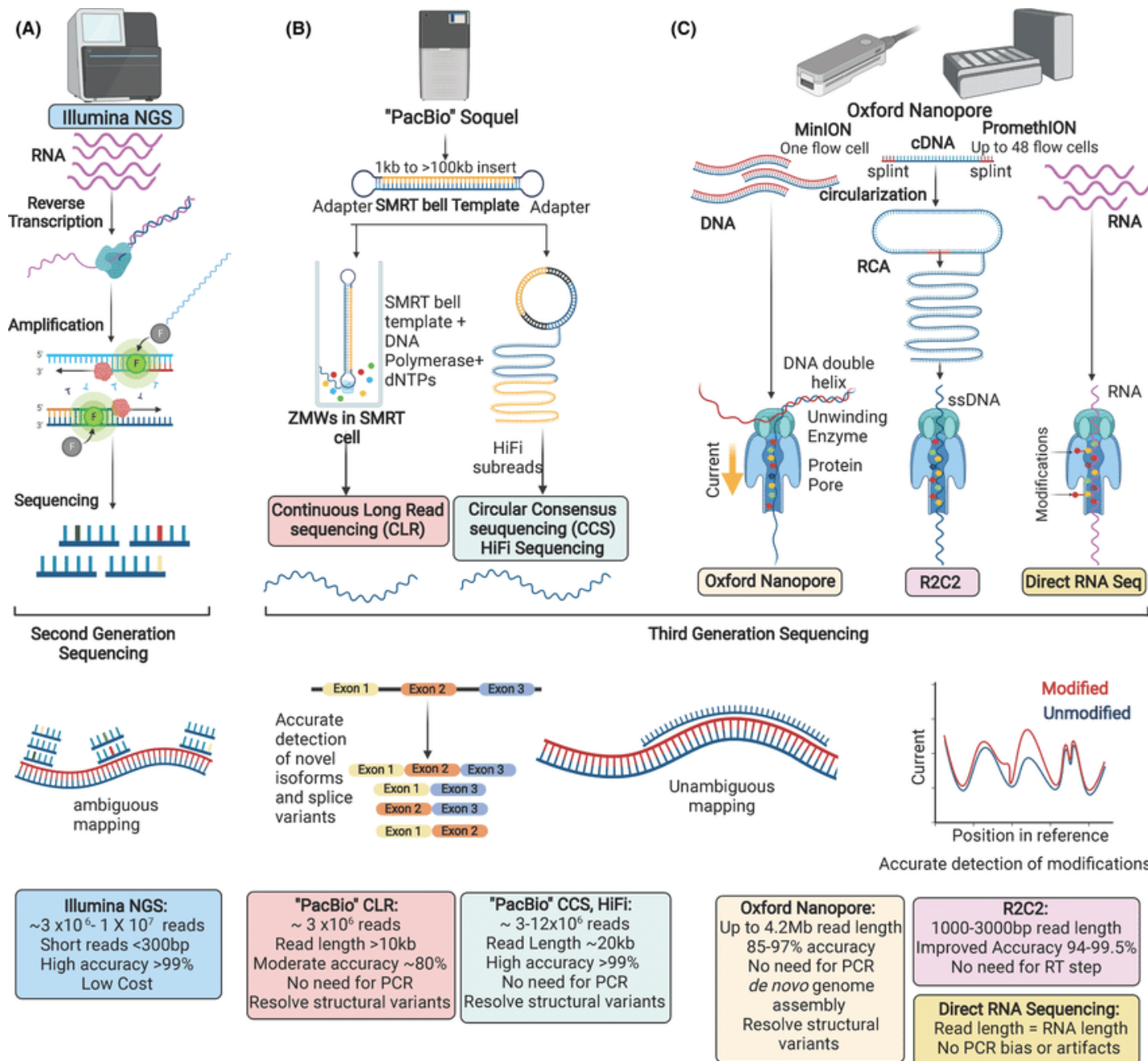
Transcriptome/genome  
mapping



Sequencing

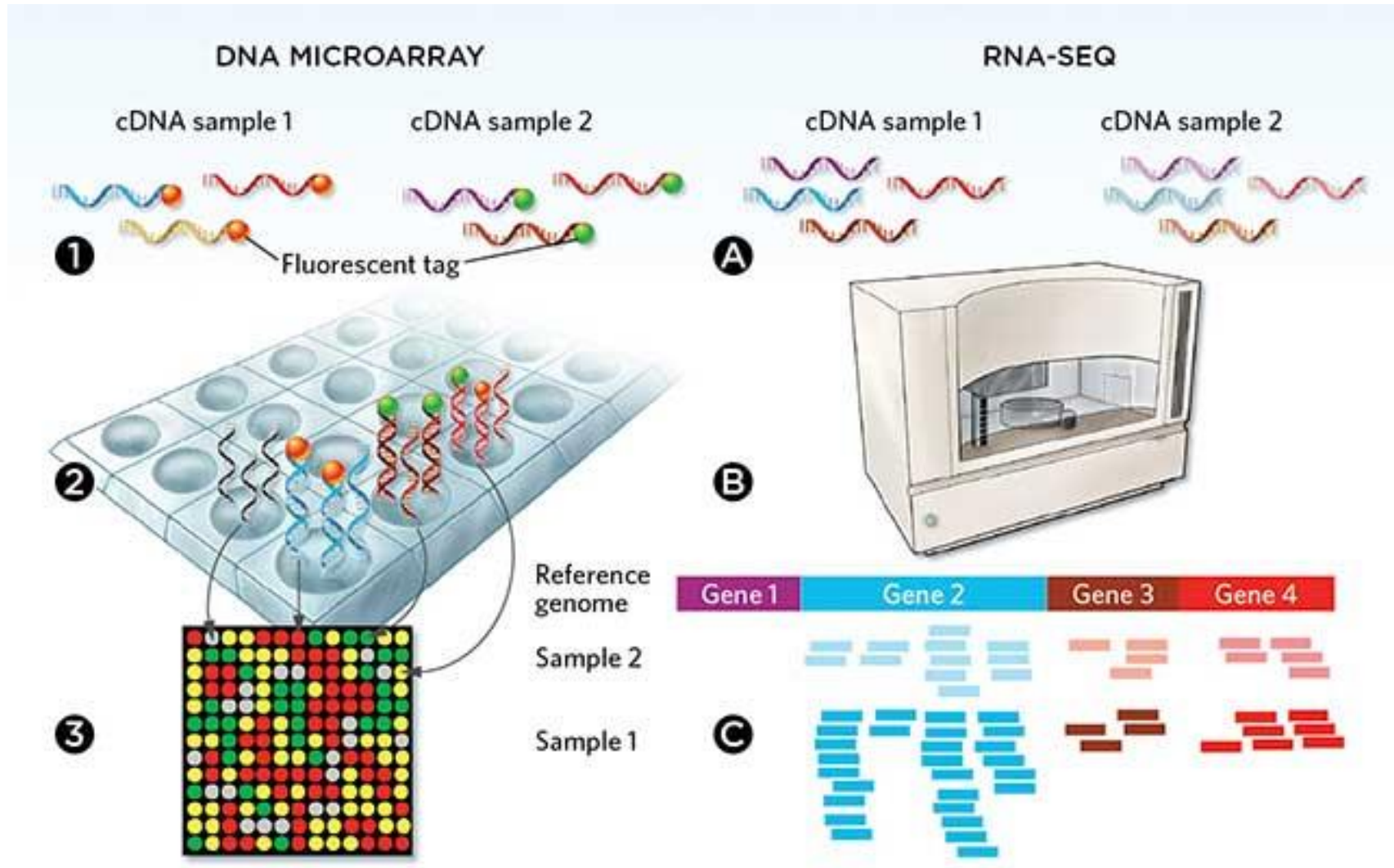








# Microarray vs NG-Sequencing





# Goals of a Microarray Experiment

---

1. **Find the genes** that change expression between experimental and control samples
  2. **Classify samples** based on a gene expression profile
  3. **Find patterns:** Groups of biologically related genes that change expression together across samples/treatments
  4. **Correlate expression profile** to disease state, diagnosis/prognosis or treatment
-



# Microarray Experiment Design

---

- **Type I:** ( $n = 2$ )
    - How is this gene expressed in target 1 as compared to target 2?
    - Which genes show up/down regulation between the two targets?
  - **Type II:** ( $n > 2$ )
    - How does the expression of gene A vary over time, tissues, or treatments?
    - Do any of the expression profiles exhibit similar patterns of expression?
-

# Viral Infection

---



Small pox

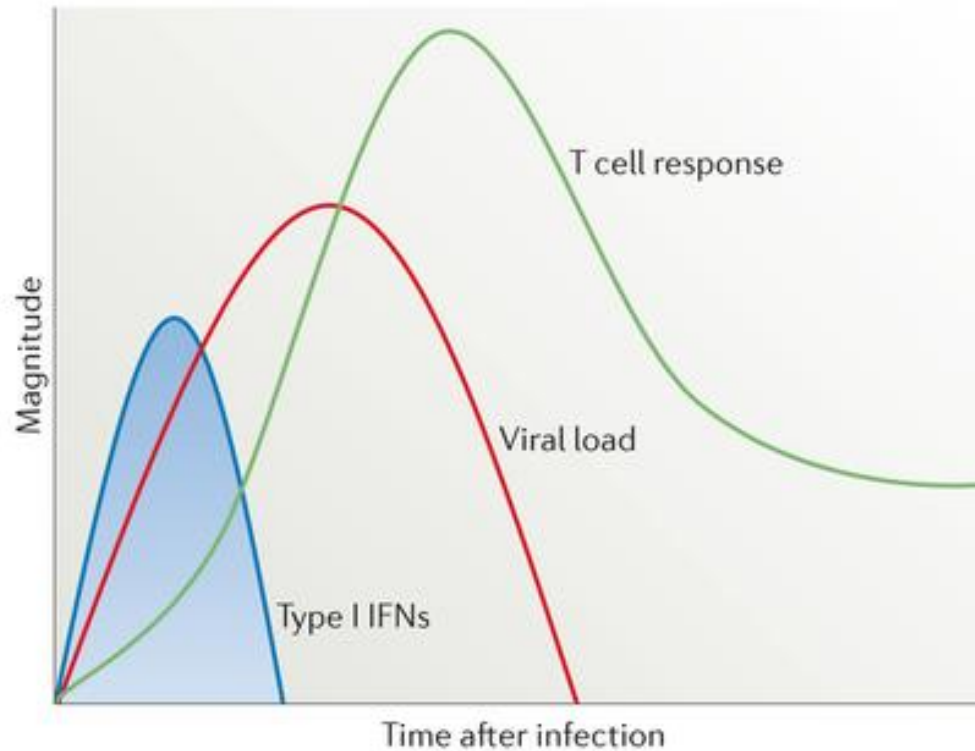


Herpes

Influenza

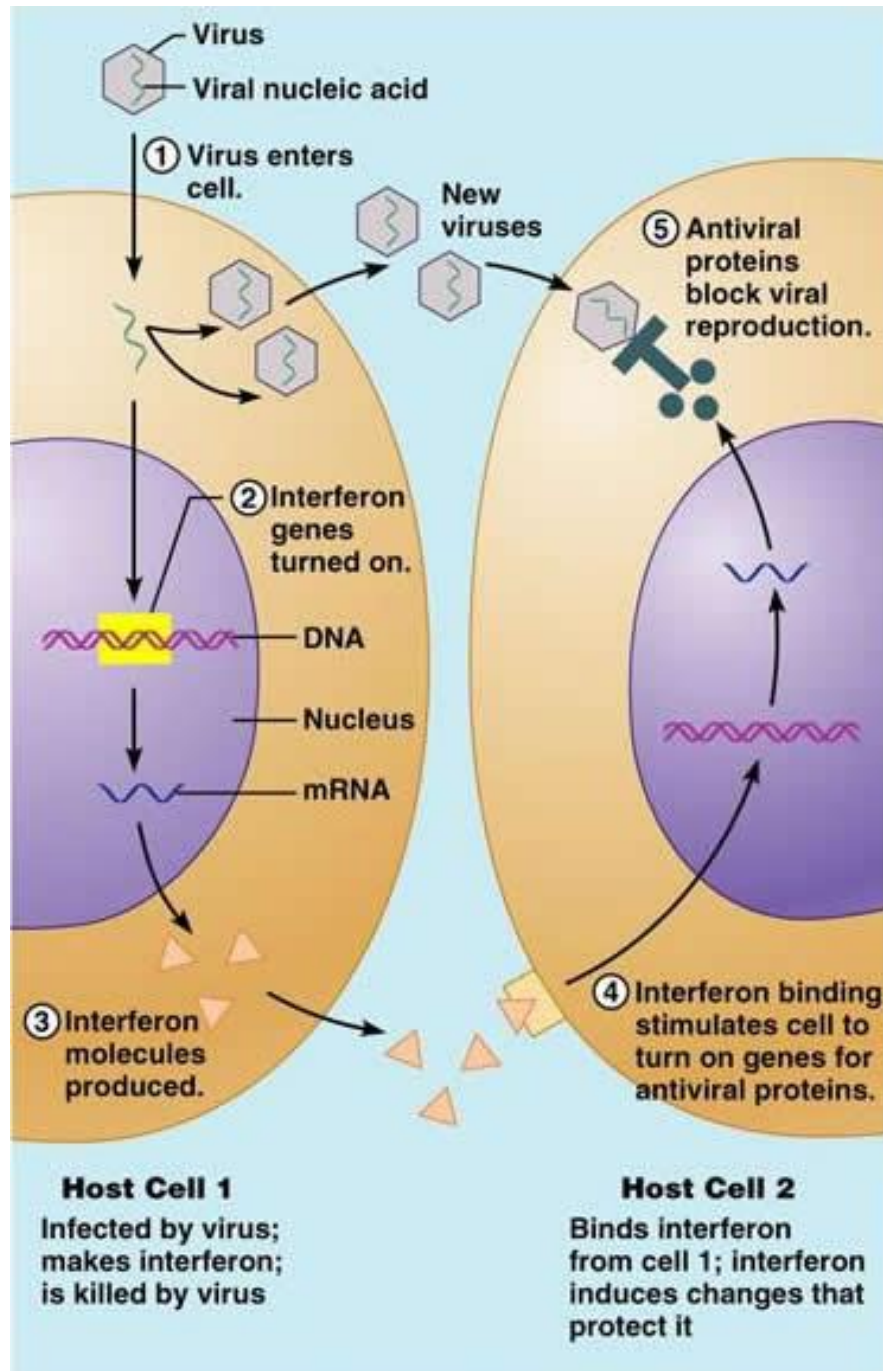


# Anti-viral Response





# Type I IFN Production & action



IFN

Inhibition viral replication

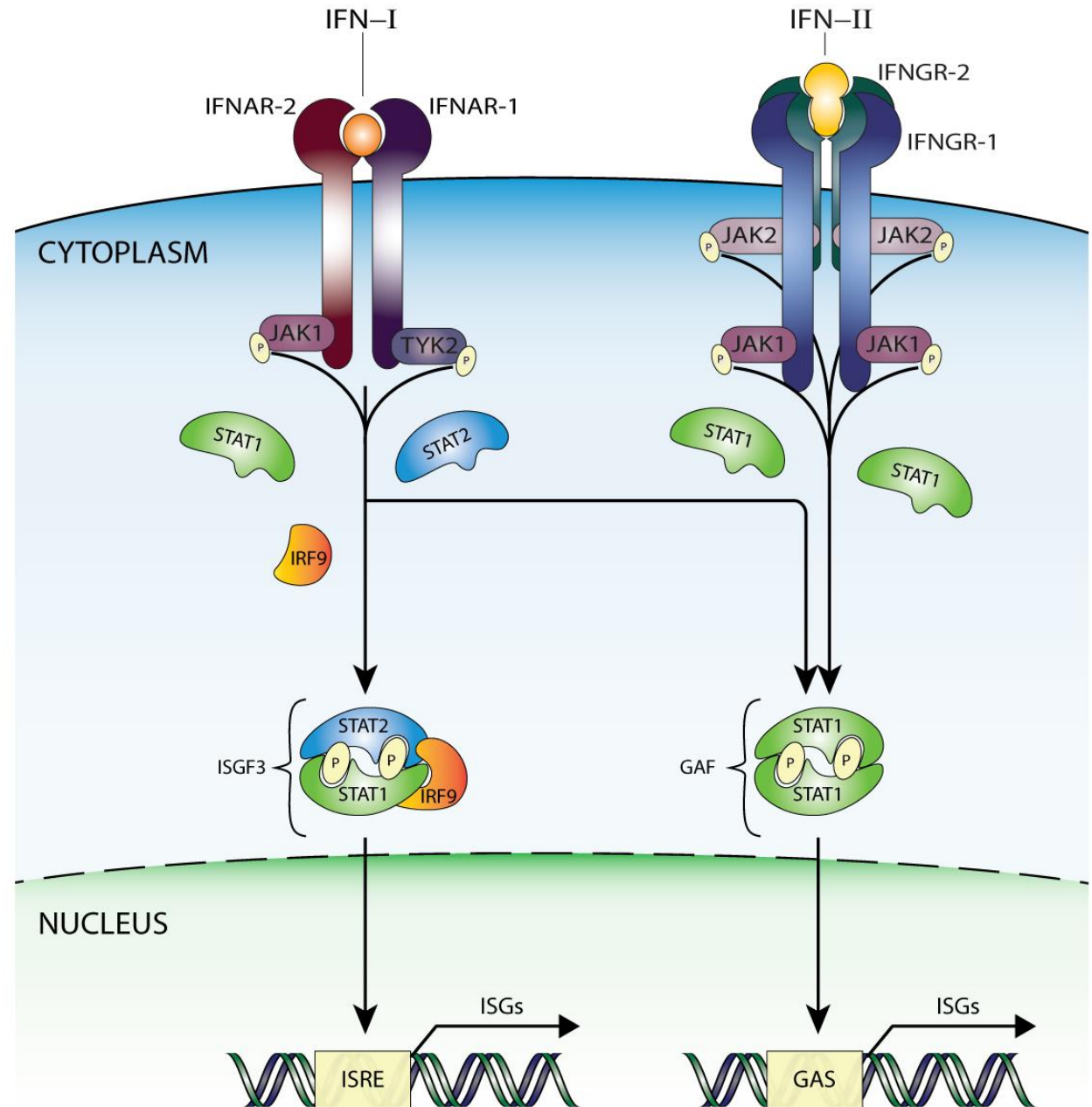
Inhibition cell growth

Activation immune system

Anti-viral State  
Adaptive immune response

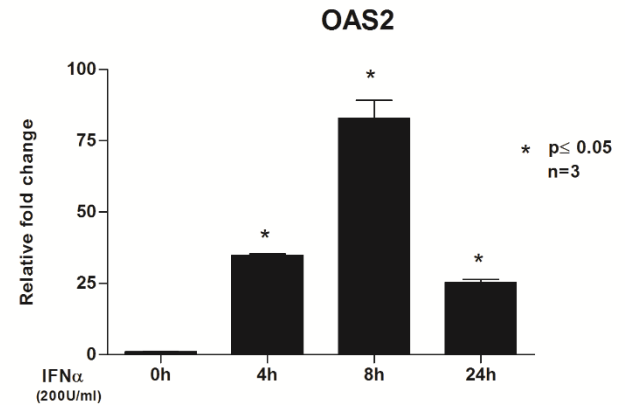
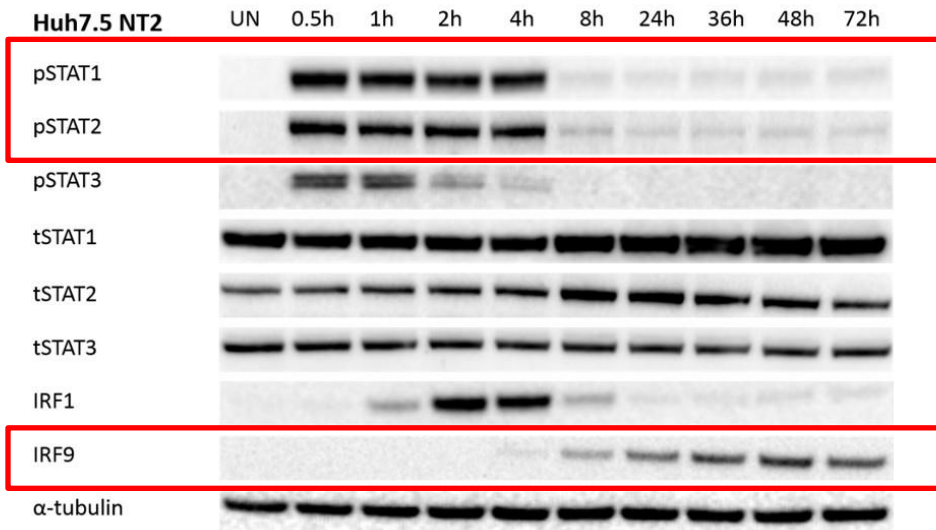


# Canonical IFN- signaling (1990)

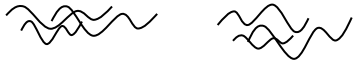


# IFN-signaling: pSTAT1, pSTAT2 & IRF9

## IFN-I



Sample 1    Sample 2



Total RNA

AAAA



cDNA

AAAA

TTTT



dsDNA

TTTT

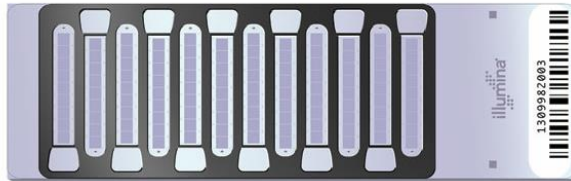
AAAA

cRNA

UUUU

biotin

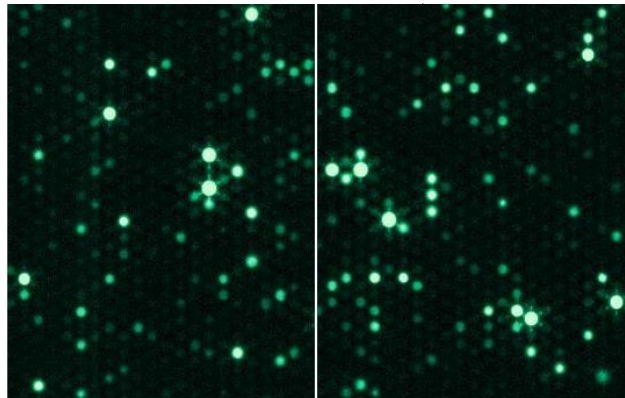
Cy-3-Streptavidin



Scanner



24 min/BeadChip



GenomeStudio



Differential gene expression



**Illumina  
DirectHyb**

**Microarray**



# Microarray on MEF WT treated with IFNA

UAM

SYMBOL	Ratio 0 vs 4	Ratio 0 vs 8	Ratio 0 vs 24	t-test 0 vs 4h	t-test 0 vs 8h	t-test 0 vs 24h	AVG detection p_value 4h	AVG detection p_value 8h	AVG detection p_value 24h
Cxcl9	262.44	136.17	30.09	0.00	0.00	0.01	0.00	0.00	0.03
Batf2	150.80	71.53	30.59	0.00	0.00	0.01	0.00	0.00	0.03
Oas1a	131.77	92.41	135.83	0.00	0.00	0.00	0.00	0.00	0.00
Oas1l	109.70	58.69	44.61	0.01	0.01	0.01	0.00	0.00	0.00
Rsd2	109.01	71.63	53.57	0.00	0.00	0.00	0.00	0.00	0.00
Serpina3f	84.20	10.51	1.89	0.04	0.11	0.85	0.00	0.01	0.31
Ifi47	79.59	32.72	21.20	0.00	0.01	0.01	0.00	0.00	0.00
Oas2	68.25	112.57	147.72	0.01	0.01	0.00	0.00	0.00	0.00
Gbp10	65.69	45.31	36.51	0.03	0.03	0.04	0.00	0.00	0.00
Tgtp	64.77	21.92	10.95	0.05	0.08	0.11	0.00	0.00	0.00
Gbp5	58.87	17.03	2.46	0.00	0.00	0.42	0.00	0.08	0.40
Cxcl10	57.24	22.72	13.14	0.00	0.00	0.01	0.00	0.00	0.00
Gbp6	51.22	27.93	26.31	0.02	0.04	0.03	0.00	0.00	0.00
LOC435565	48.21	43.05	44.27	0.01	0.01	0.01	0.00	0.00	0.00
Tyki	42.93	25.75	17.71	0.01	0.01	0.01	0.00	0.00	0.00
D14Ertd668e	42.56	35.27	29.67	0.00	0.00	0.00	0.00	0.00	0.00
Mx2	40.47	31.18	30.51	0.00	0.01	0.01	0.00	0.00	0.00
Ifi203	40.37	26.68	15.41	0.00	0.00	0.00	0.01	0.03	0.09
Oas1b	30.35	27.09	22.41	0.03	0.03	0.03	0.00	0.00	0.00

Anti-viral ISRE containing genes





Insert Gene/Sequence ID list: ([help](#)) **Pscan**

Select Organism:

Select Region:

Select Descriptors: ☐ Jaspar 2016 ☐ Jaspar 2014 ☐ Jaspar\_Fam ☒ Transfac ☐ User Defined

Messages:

  **Pscan Web Interface**  
 NEW-> [Ver. 1.4 \(Last update: 01 Dec 2015\)](#)

Use the input form on the left to set up your query. The results will be displayed in this window.

[If you need HELP please click here.](#)

**Source:**  
[Download Pscan source code](#)

**Sample data**

List of Human NFkB target genes, collected from literature. NFkBxx indicates that xx percent of the genes in the list are NFkB targets, while the others are random genes added to the set to assess the performance of the algorithm.

[NFkB100](#) [NFkB90](#) [NFkB80](#) [NFkB70](#) [NFkB60](#) [NFkB50](#) [NFkB40](#)

List of Human NRF1 target genes. NRFxx should be read as in the NFkB dataset.

[NRF1\\_100](#) [NRF1\\_90](#) [NRF1\\_80](#) [NRF1\\_70](#) [NRF1\\_60](#) [NRF1\\_50](#) [NRF1\\_40](#)

ISRE

Insert Gene/Sequence ID list: ([help](#)) **Pscan**

NM\_001547  
 NM\_001031  
 NM\_030641  
 NM\_207315  
 NM\_001548  
 NM\_022147  
 NM\_138456  
 NM\_152649  
 NM\_012420

Select Organism:

Select Region:

Select Descriptors: ☐ Jaspar 2016 ☐ Jaspar 2014 ☐ Jaspar\_Fam ☒ Transfac ☐ User Defined

XM\_378350  
 XM\_496974  
 XM\_497423  
 HS\_569921  
 XM\_937050  
 XM\_942259

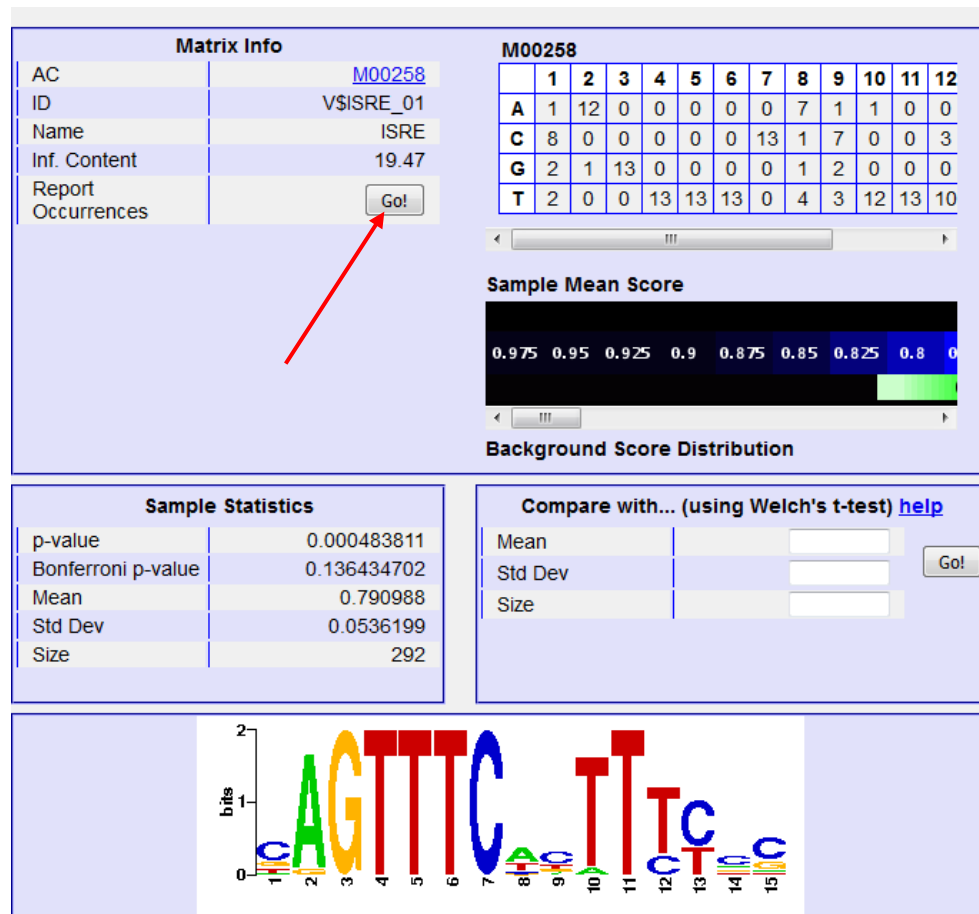
Working on 292 gene promoter(s).  
 Pscan running, please wait.  
 Done.]

[View Text Results](#)

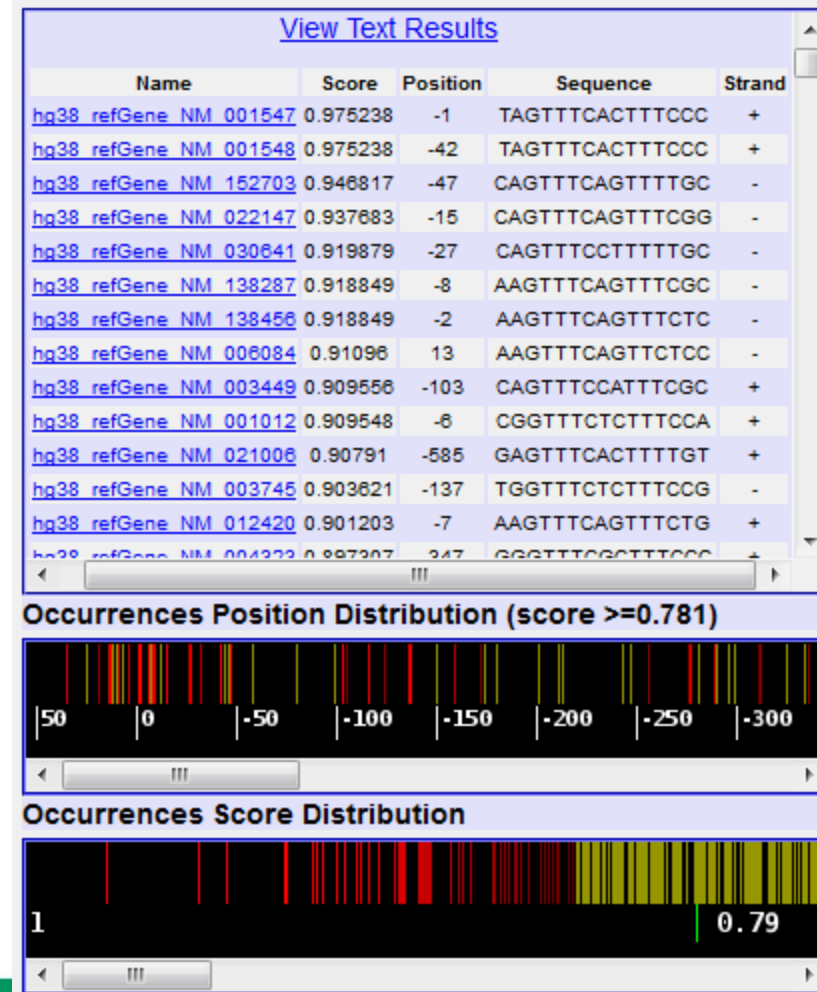
282 TF profiles used

Matrix ID	Matrix Name	P-value
M00063	<a href="#">V\$IRF2_01</a>	1.27589e-06
M00062	<a href="#">V\$IRF1_01</a>	1.78279e-05
M00196	<a href="#">V\$SP1_Q6</a>	0.000174368
M00223	<a href="#">V\$STAT_01</a>	0.000469797
M00258	<a href="#">V\$ISRE_01</a>	0.000483811
M00453	<a href="#">V\$IRF7_01</a>	0.000631963
M00224	<a href="#">V\$STAT1_01</a>	0.00600153
M00088	<a href="#">V\$IK3_01</a>	0.00866677
M00189	<a href="#">V\$AP2_Q6</a>	0.00868479
M00130	<a href="#">V\$FOXO3_01</a>	0.0169
M00497	<a href="#">V\$STAT3_02</a>	0.0188426
M00083	<a href="#">V\$SMZF1_01</a>	0.0211891
M00033	<a href="#">V\$P300_01</a>	0.0268041
M00396	<a href="#">V\$EN1_01</a>	0.0292041
M00008	<a href="#">V\$SP1_01</a>	0.0296641
M00113	<a href="#">V\$CREB_02</a>	0.0333268
M00076	<a href="#">V\$GATA2_01</a>	0.0392267
M00245	<a href="#">V\$EGR3_01</a>	0.044253
M00141	<a href="#">V\$LYF1_01</a>	0.0512761
M00025	<a href="#">V\$ELK1_02</a>	0.0533307
M00108	<a href="#">V\$NRF2_01</a>	0.0542658
M00244	<a href="#">V\$NGFIC_01</a>	0.0560242
M00459	<a href="#">V\$STAT5B_01</a>	0.058159
M00373	<a href="#">V\$PAX4_01</a>	0.0590223
M00235	<a href="#">V\$AHHRARNT_01</a>	0.0598517
M00460	<a href="#">V\$STAT5A_02</a>	0.067307
M00517	<a href="#">V\$AP1_01</a>	0.0690688
M00433	<a href="#">V\$HMX1_01</a>	0.0698115

Promoter  
Analysis

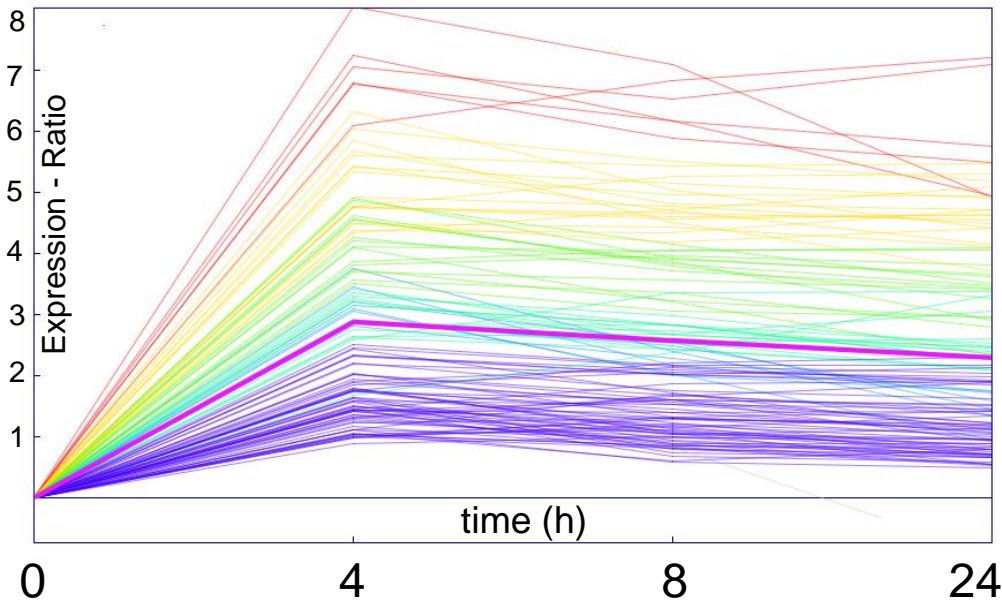


## ISRE containing genes





# ISG Expression in IFN-I treated cells



Expression view – 124 up regulated genes

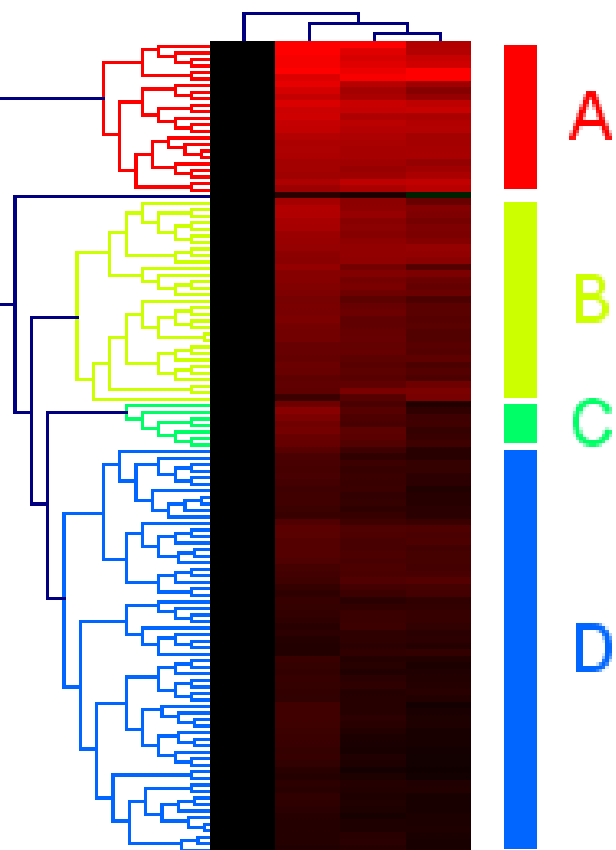
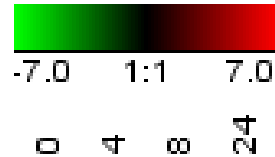
## Gene Ontology

ter_ID	Description	log <sub>10</sub> p-value
GO:0002376	immune system process	-29.12
GO:0051607	defense response to virus	-28.38
GO:0002252	immune effector process	-25.61
GO:0006955	immune response	-21.24
GO:0045087	innate immune response	-19.97
GO:0019882	antigen processing and presentation	-10.24
GO:0002682	regulation of immune system process	-10.01
GO:0042089	cytokine biosynthetic process	-5.04
GO:0042107	cytokine metabolic process	-4.91
GO:0009617	response to bacterium	-4.44
GO:0032608	interferon-beta production	-4
GO:0032606	type I interferon production	-4
GO:0045343	regulation of MHC class I biosynthetic process	-3.23

Anti viral response



# ISG Expression in MEF WT after IFNA treatment



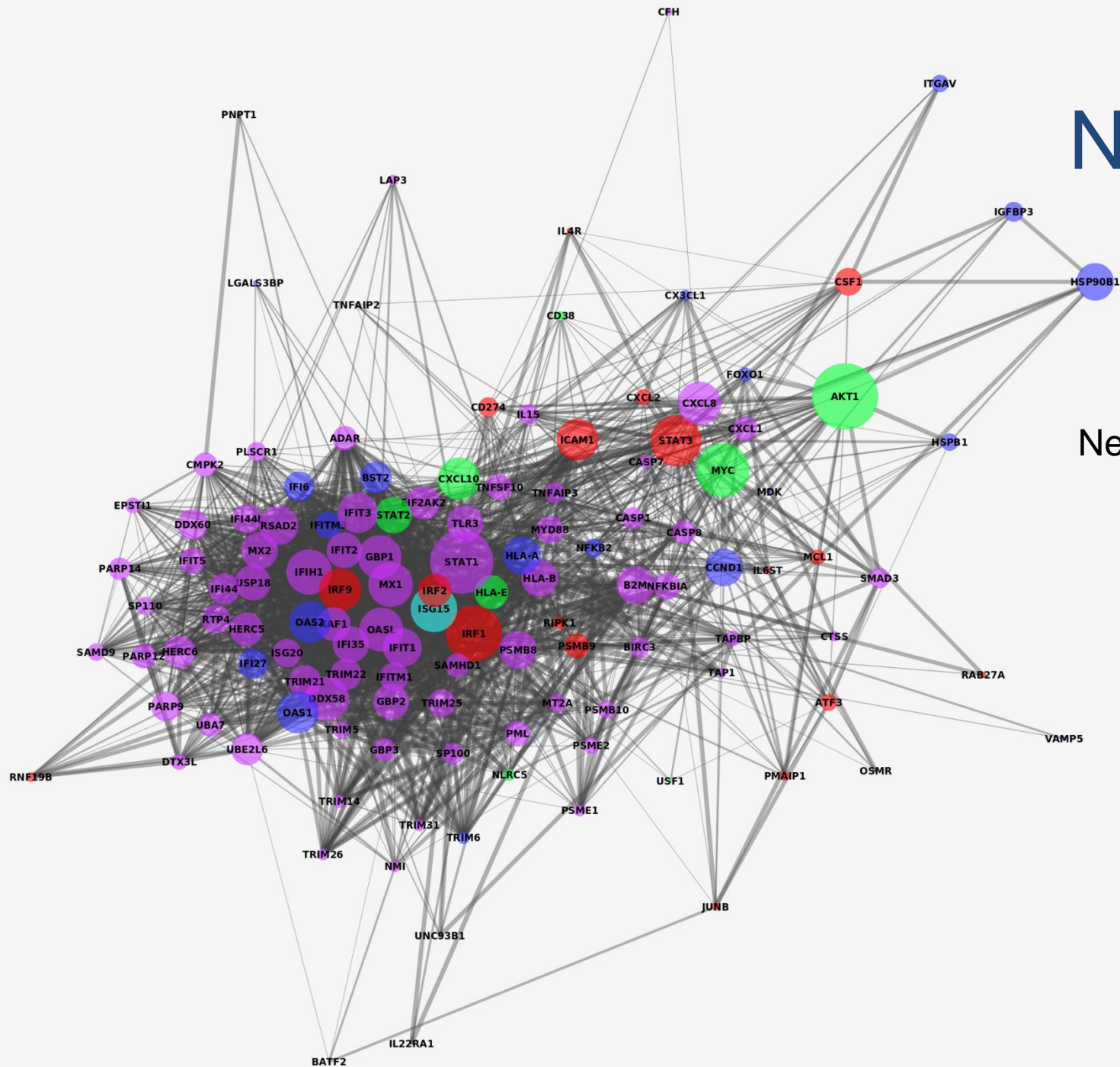
Cluster	Genes	GO Term	p-value
A	Oas1	defense response	4.72E-17
	Oas2	defense response to virus	1.63E-15
	Oas1a	response to virus	1.39E-14
	Oas1b	cellular response to interferon-beta	5.73E-13
B	Oas1c	response to interferon-beta	1.86E-12
	Oas1d	response to interferon	4.85E-07
C	Oas1e	immune system process	3.76E-06
	Oas1f	immune response	3.36E-04
	Oas1g	negative regulation of type I interferon mediated signaling pathway	8.17E-04
	Oas1h	antigen processing and presentation of endogenous peptide antigen	2.44E-09
D	Oas1i	antigen processing and presentation of peptide antigen via MHC class I	2.44E-09
	Oas1j	regulation of cell mediated cytotoxicity	7.26E-07
	Oas1k		
	Oas1l		

Protection against viral infection

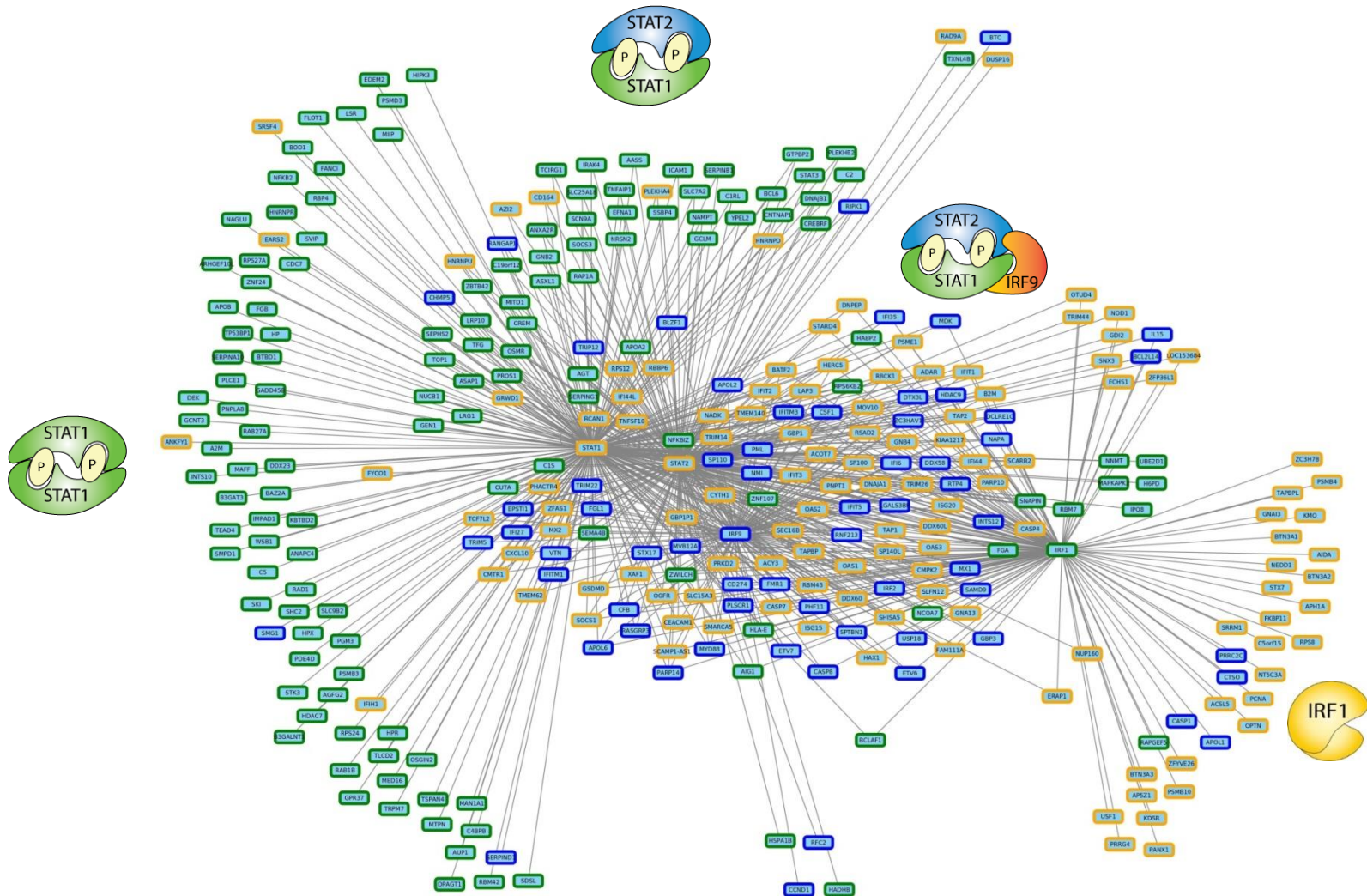


\_\_\_\_\_

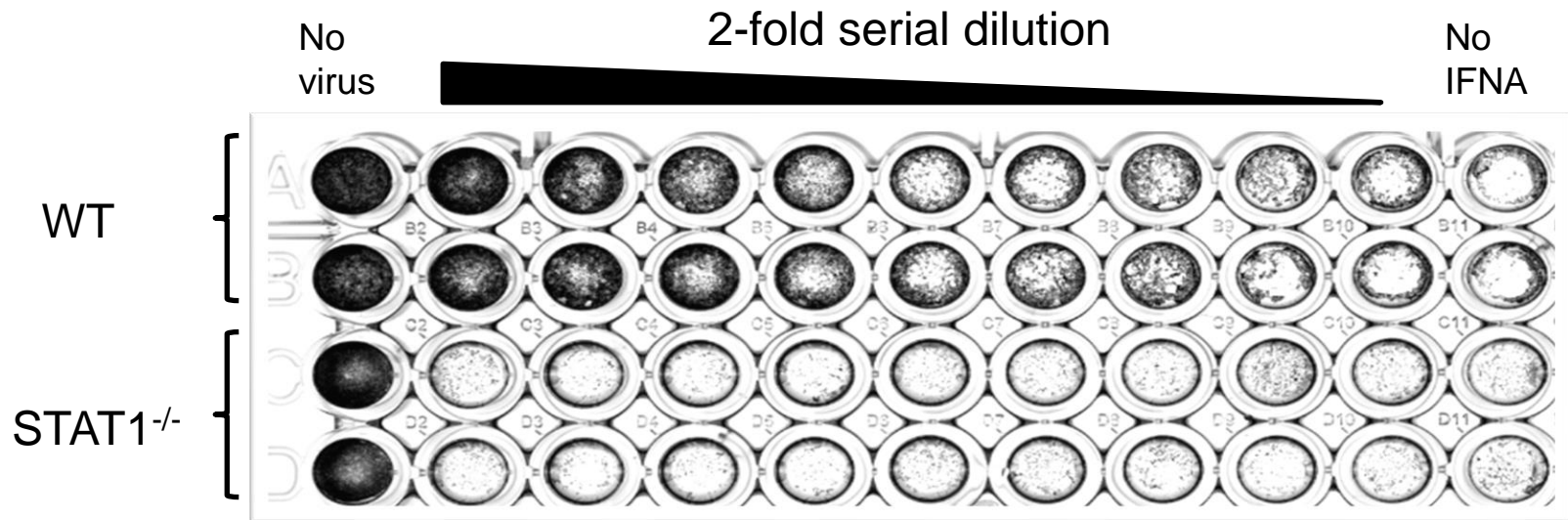
\_\_\_\_\_



# IFN $\alpha$ -induced Gene Network



# Anti-viral assay



Cell lines treated with 200U/ml mIFNA 2-fold dilution (24h)  
**EMCV** MOI=0.1 (18h)

ENCEPHALOMYOCARDITIS VIRUS - EMCV



# The transcriptional program in the response of human fibroblasts to serum

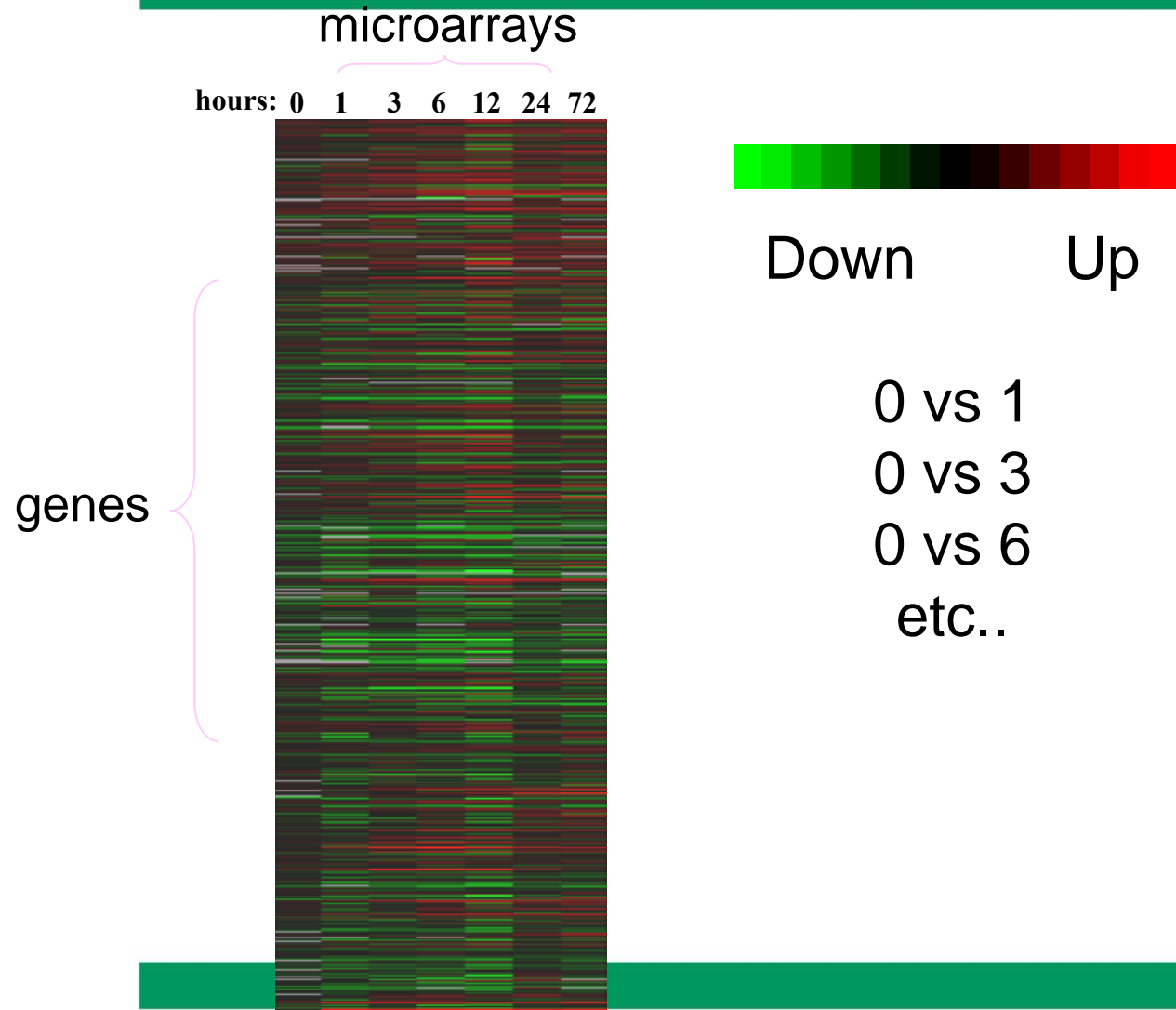
Iyer et al. (1999) *Science*, 283: 83

- Identify genes with similar expression
- Grouping unknown genes with known genes may provide insight into function of unknown genes
- Cluster genes by similar changes - only really meaningful across multiple treatments or time points
- Cluster samples by similar gene expression profiles





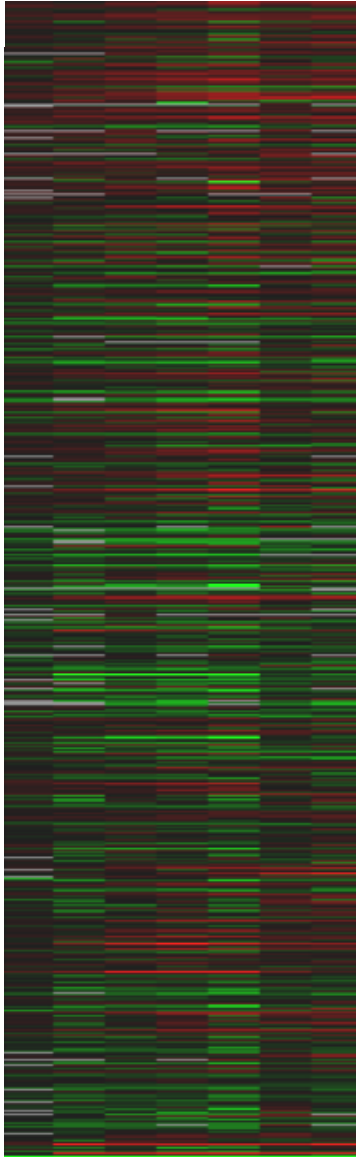
# Basics of Data Filtering and Visualization





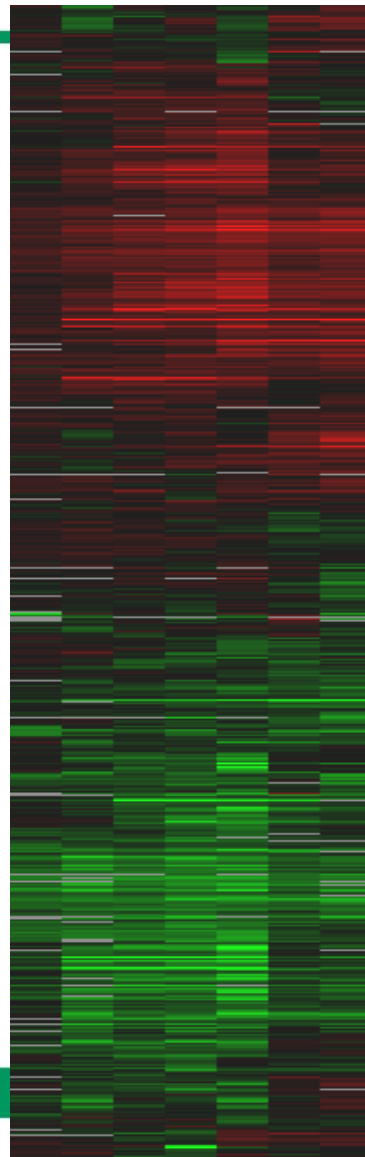
# Grouping genes: clustering

hours: 0 1 3 6 12 24 72



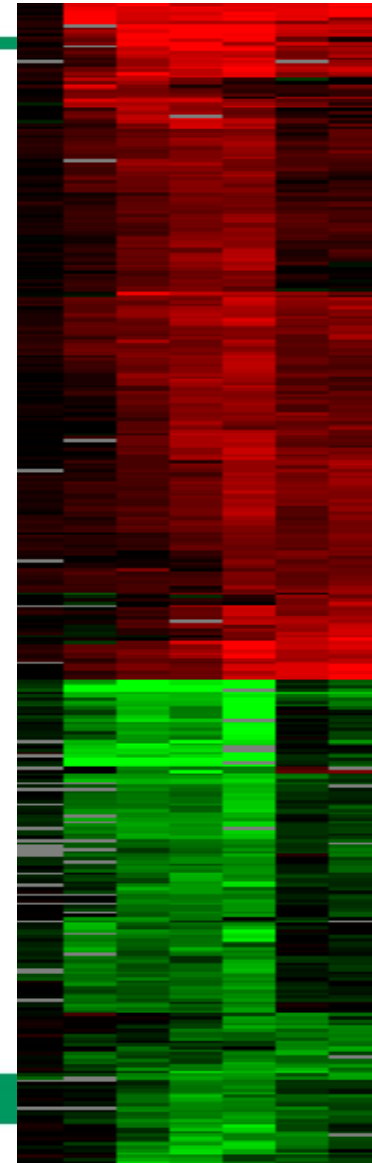
4,416 genes

hours: 0 1 3 6 12 24 72



4,124 genes  
spot-filtered  
clustered

hours: 0 1 3 6 12 24 72



464 genes  
spot-filtered  
ratio-filtered  
clustered

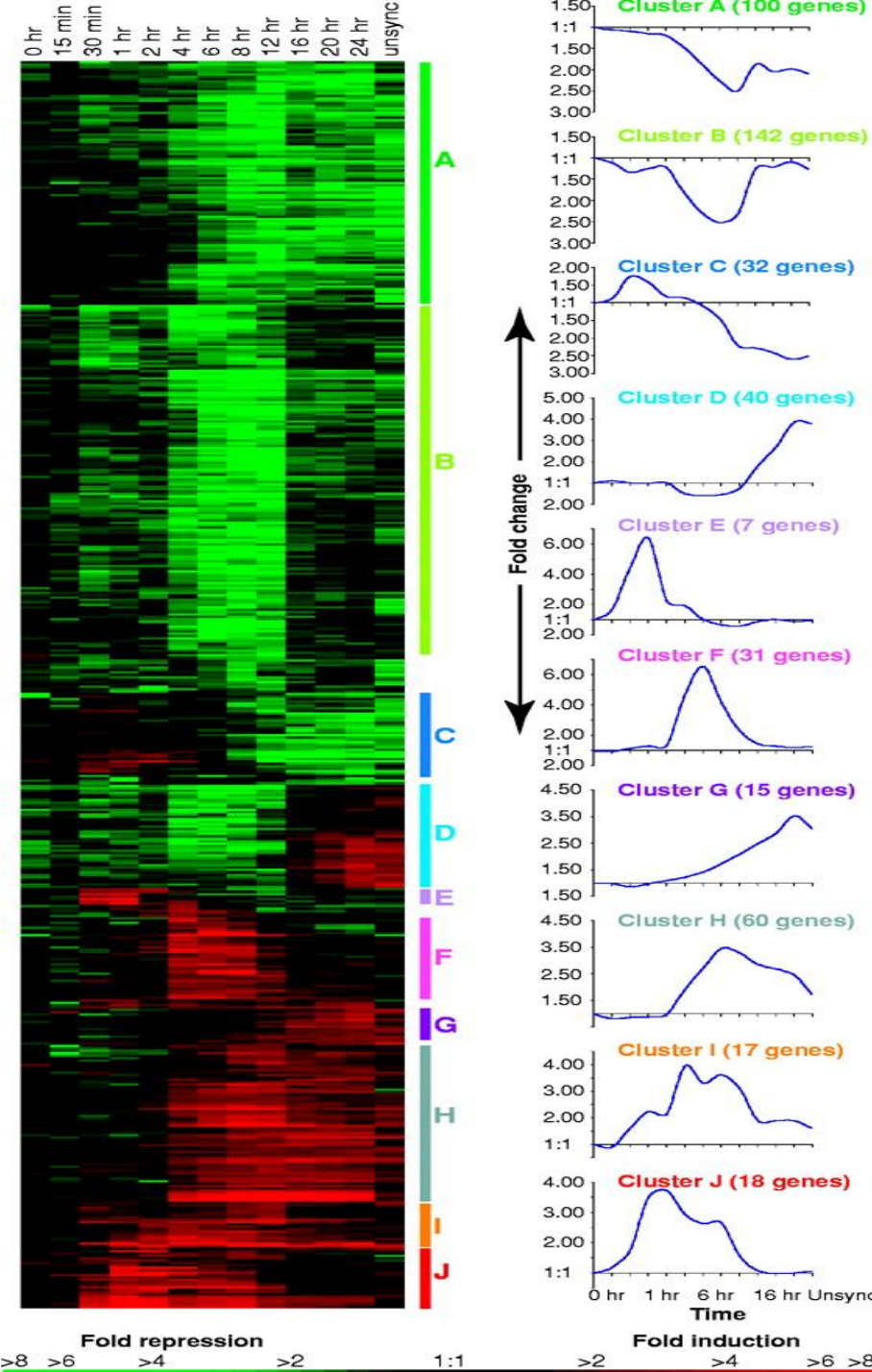
# Microarray (Clustering)

8600 cDNA clones

- Coordinated gene expression
- Differential gene expression



Biological information



Iyer et al. (1999) *Science*, 283: 83

# Expression signatures

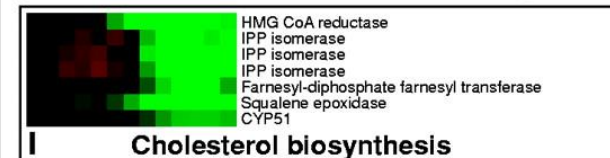
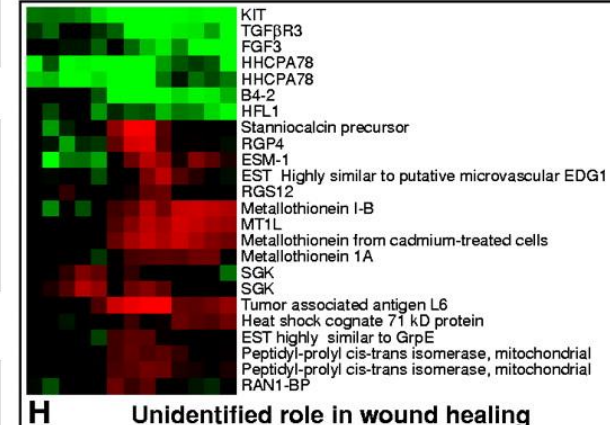
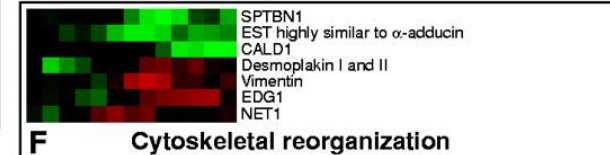
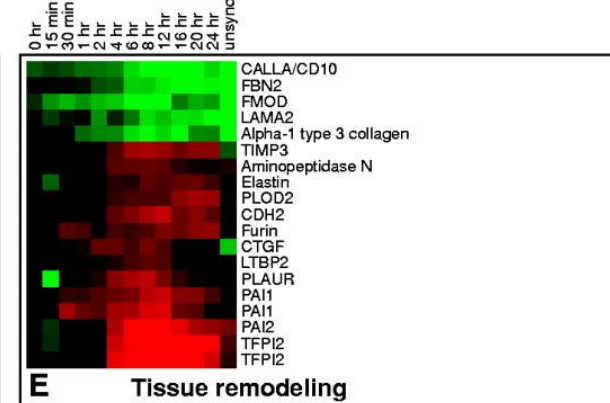
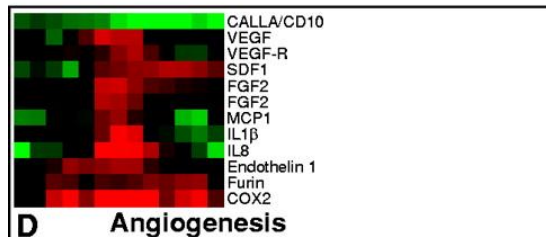
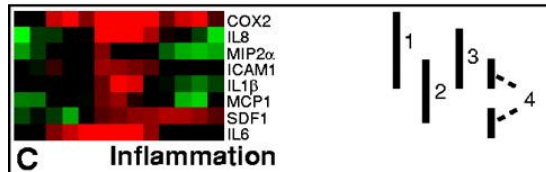
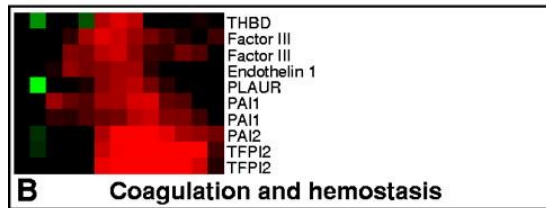
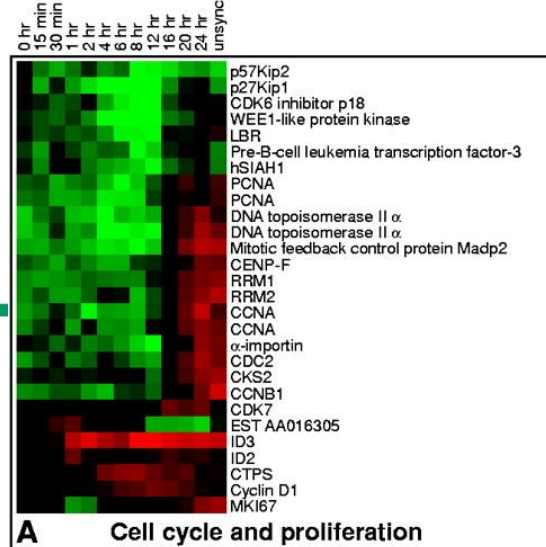
Couple expression to GO



Serum treatment '*in vitro*'



Wound healing '*in vivo*'

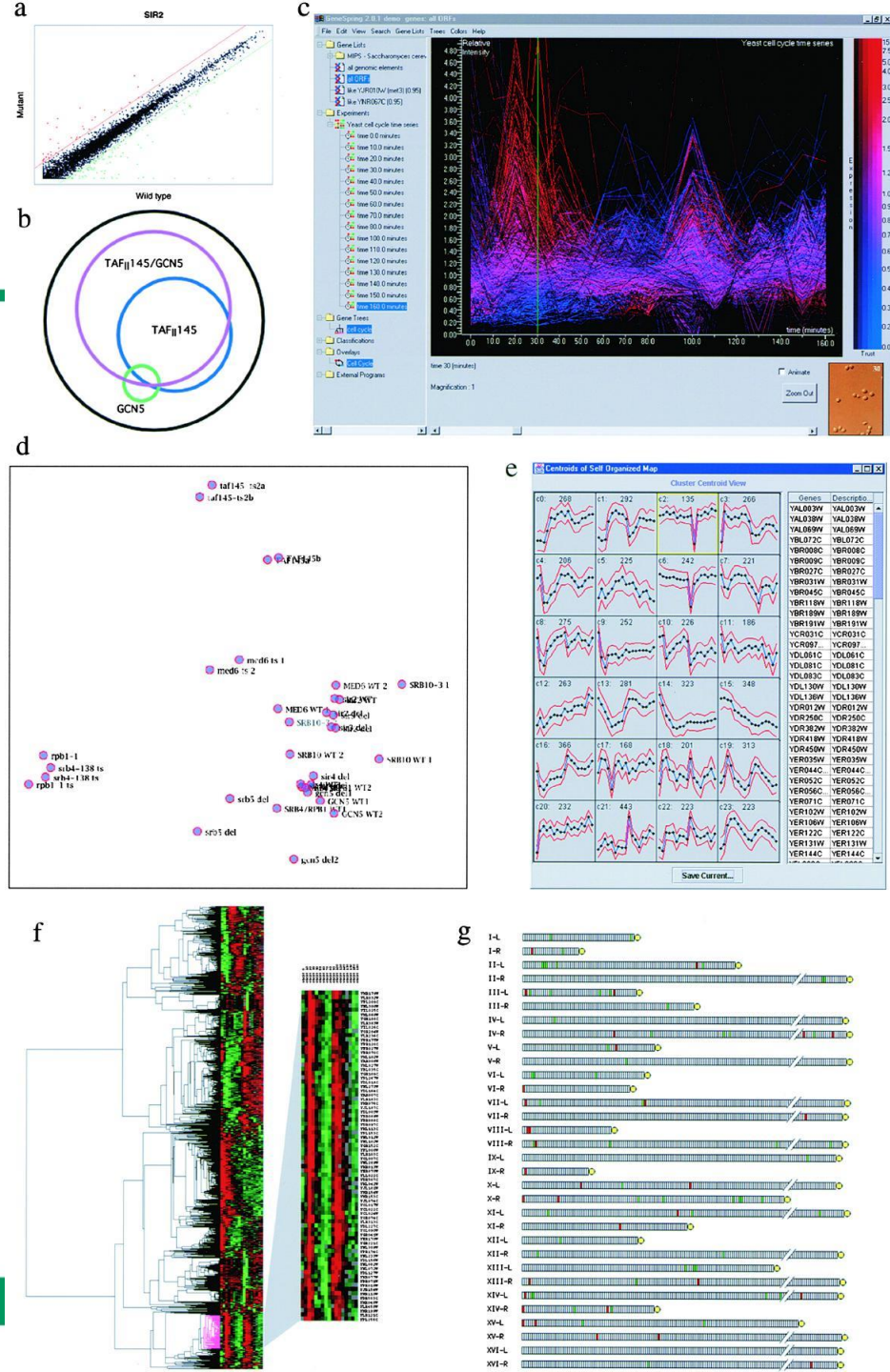




# Visualisation

# UAM

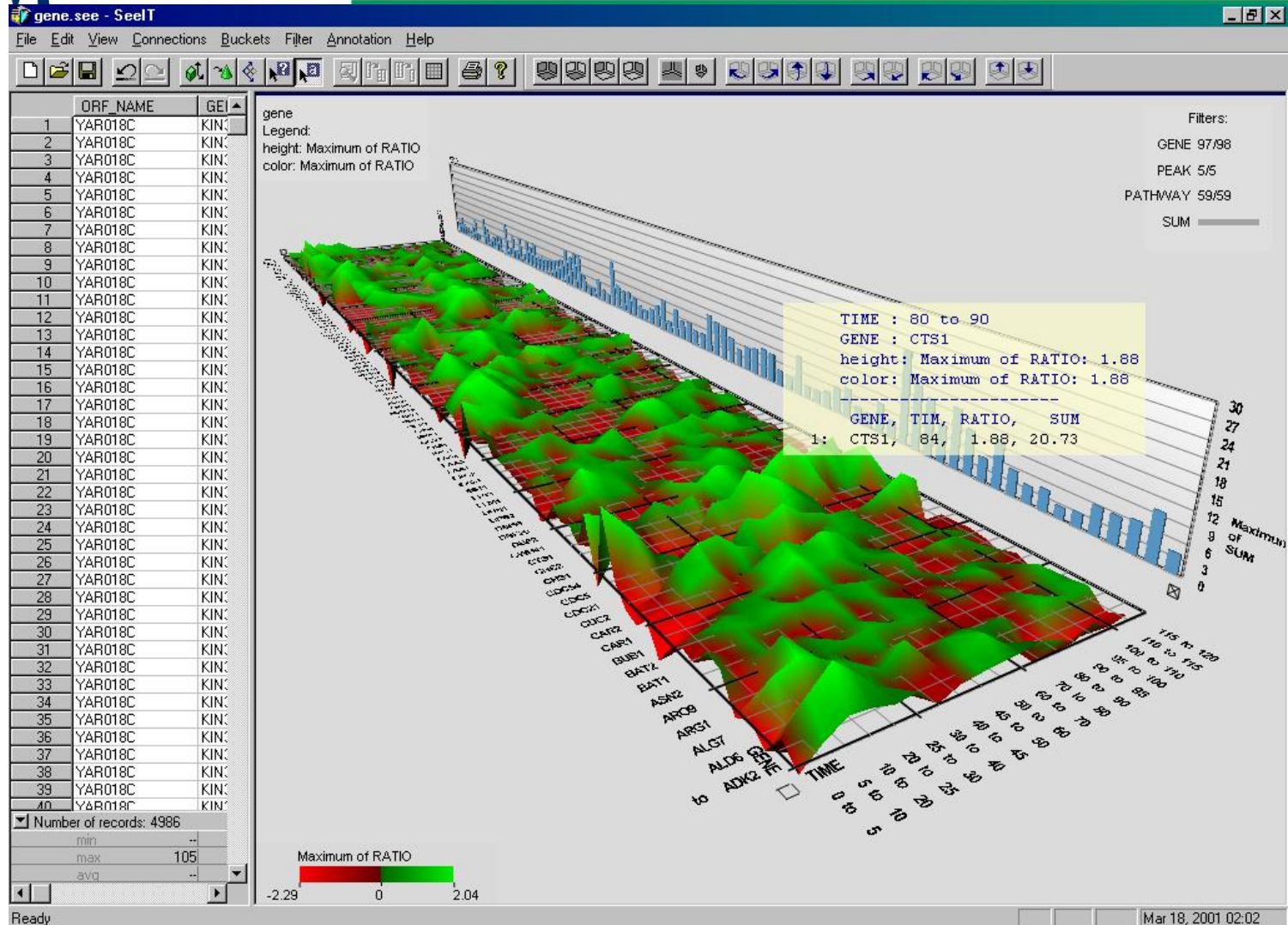
## GeneSpring



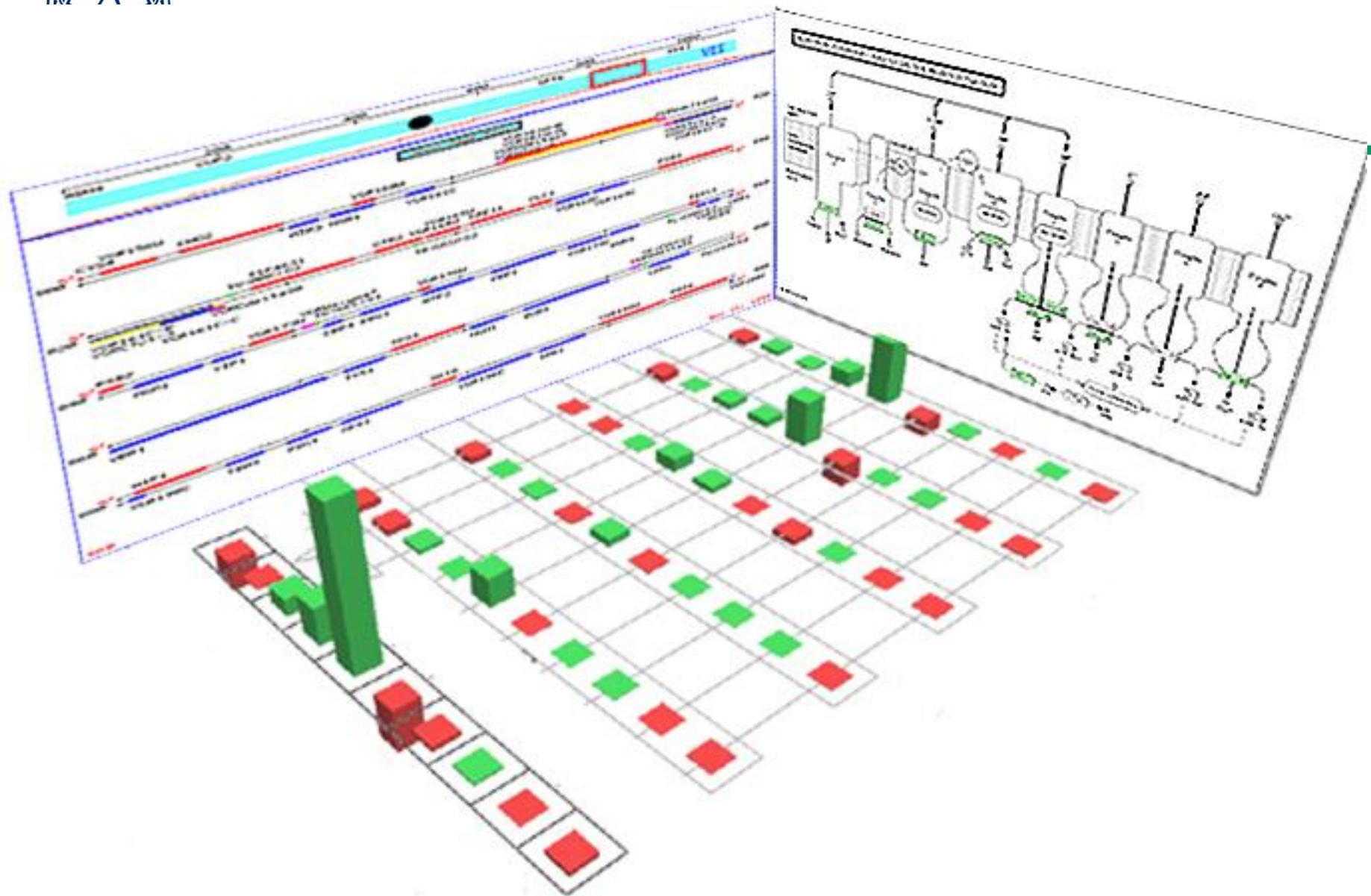




# Expression Landscape of cell-cycle regulated genes in yeast









# Microarray and cancer

---

- Identification of prognostic biomarkers specific to onset and progression
  - Disease classification
  - Development of drug resistance
  - Risk of relapse assessment
  - Metastasis
  - Response to treatment
  - Survival
-





# Variation in Gene Expression Patterns in Human Cancer Cell Lines

---

Ross DT, et. al., *Nature Genetics*, (24): 2000, 227-235.

Aim:

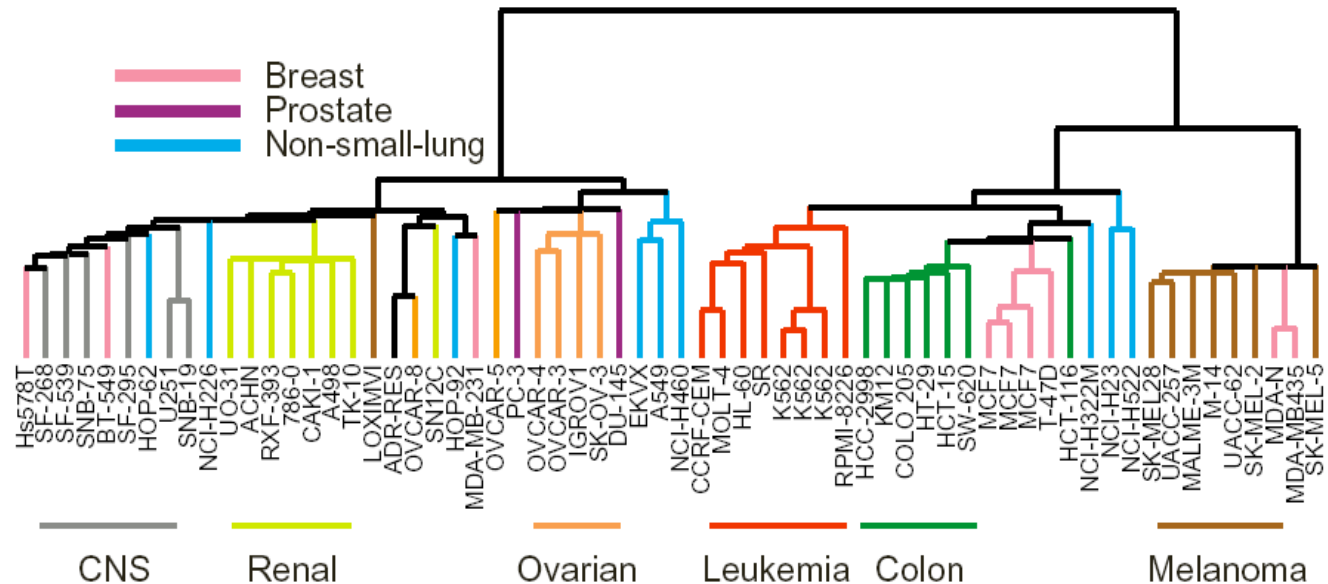
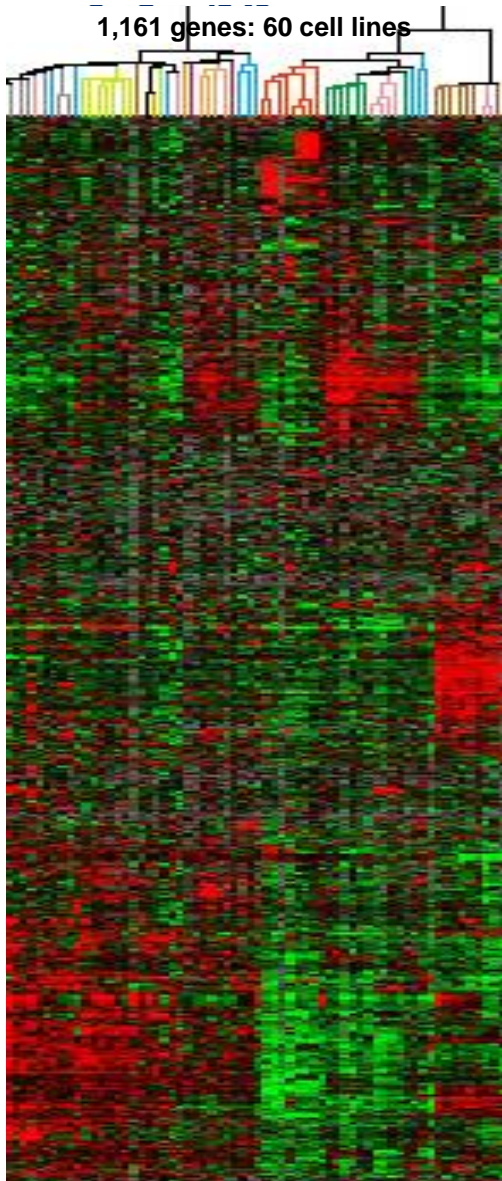
to explore the variation in gene expression  
of ~ 8,000 genes among 60 human cancer  
cell lines (spanning 9 distinct tissues)

---



# Hierarchical Clustering of Gene Expression Patterns

## Groups Cell Lines According to Tissue of Origin



- Relationship between expression profile and tissue of origin
- Recognize previously incorrect classified outliers
- Recognize relationships to tumors *in vivo*



# Distinct Types of Diffuse Large B-cell Lymphoma (DLBCL)

---

Alizadeh AA, et. al., *Nature*, (403): 2000,503-511.

Aim:

to determine whether gene expression profiling could subdivide DLBCL – a clinically heterogeneous diagnostic category – into molecularly distinct diseases with more homogeneous clinical behaviors

Only 40% of patients respond well to therapy

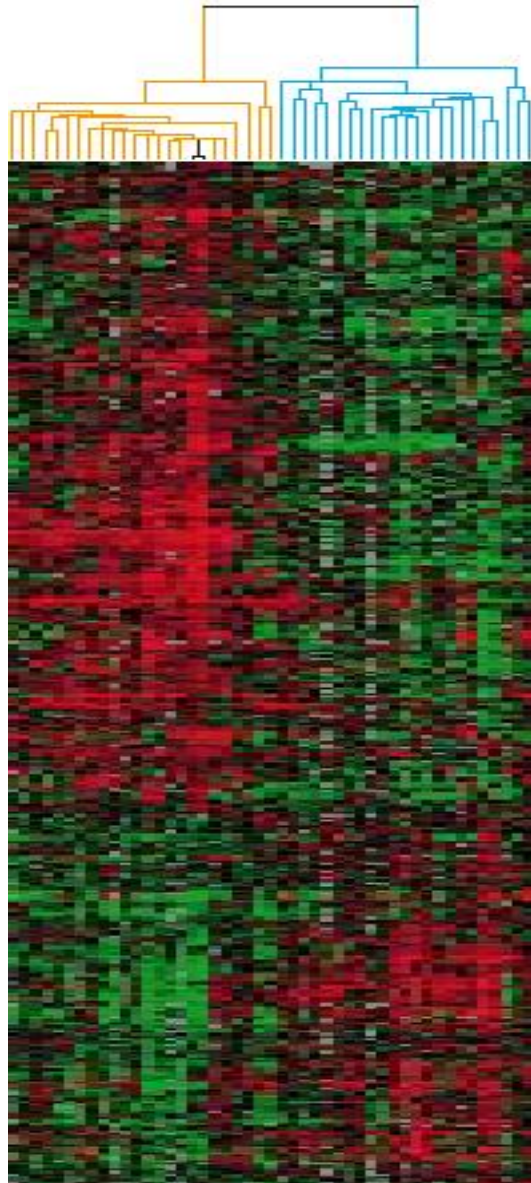
“Lymphochip”:

- 17.856 cDNA clones
- lymphoid cell origin
- cancer + immunology

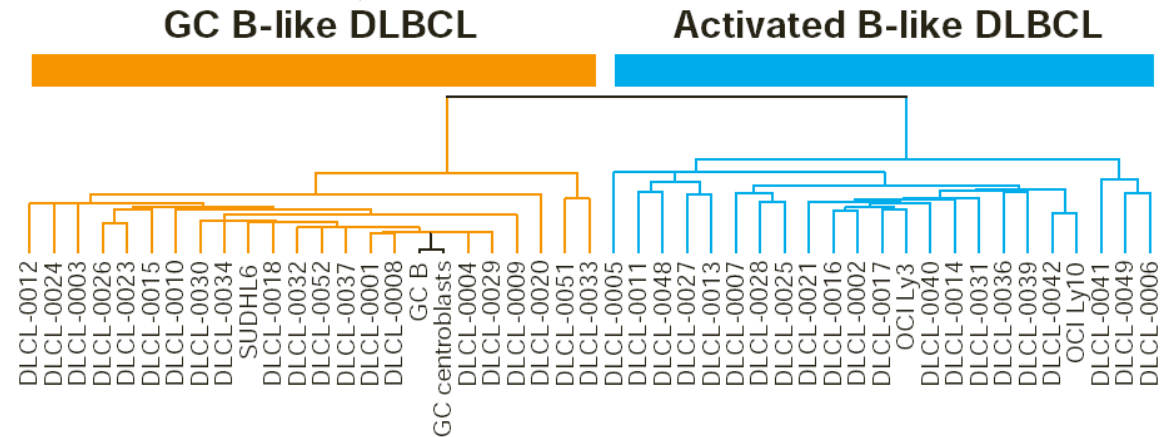
---



45 DLBCL biopsies



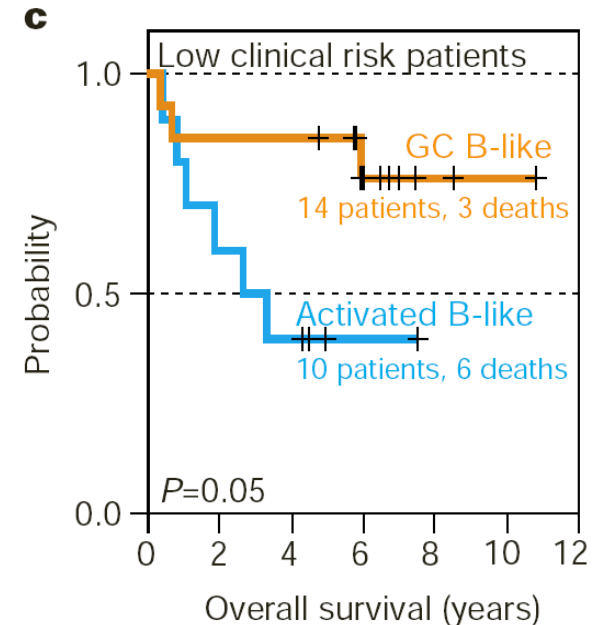
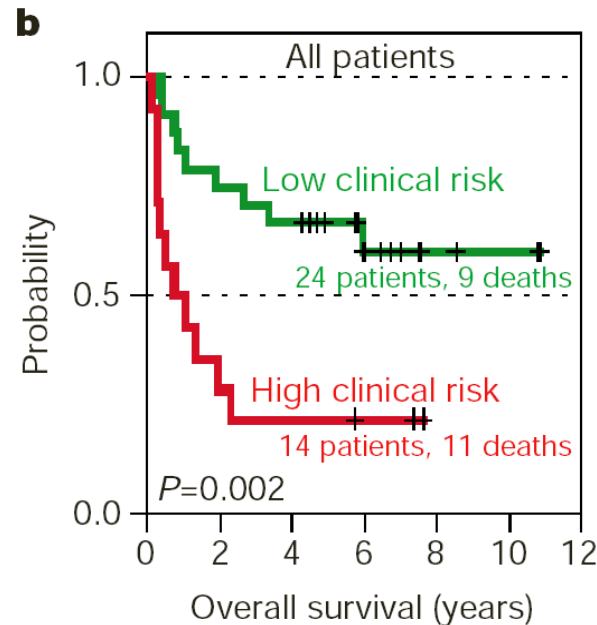
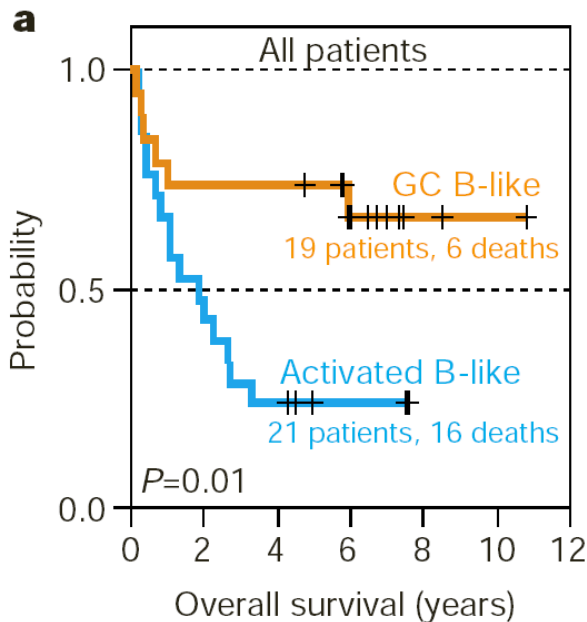
# Clustering Identifies 2 Major Subgroups of DLBCL



Different B-cell differentiation stage

Set of ~3000 genes

# DLBCL Subgroups Define Prognostic Categories



a. Kaplan-Meier plot of overall survival of DLBCL patients grouped on the basis of gene expression profiling.

b. Kaplan-Meier plot of overall survival of DLBCL patients grouped according to the International Prognostic Index.

c. Kaplan-Meier plot of overall survival of low clinical risk DLBCL patients grouped on the basis of gene expression profiles.



# Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications

Therese Sørli<sup>a,b,c</sup>, Charles M. Perou<sup>a,d</sup>, Robert Tibshirani<sup>e</sup>, Turid Aas<sup>f</sup>, Stephanie Geisler<sup>g</sup>, Hilde Johnsen<sup>b</sup>, Trevor Hastie<sup>e</sup>, Michael B. Eisen<sup>h</sup>, Matt van de Rijn<sup>i</sup>, Stefanie S. Jeffrey<sup>j</sup>, Thor Thorsen<sup>k</sup>, Hanne Quist<sup>l</sup>, John C. Matese<sup>c</sup>, Patrick O. Brown<sup>m</sup>, David Botstein<sup>c</sup>, Per Eystein Lønning<sup>g</sup>, and Anne-Lise Børresen-Dale<sup>b,n</sup>

2001, *PNAS*

Aim:

To classify breast carcinoma's based on expression profiling and to correlate these to clinical outcome



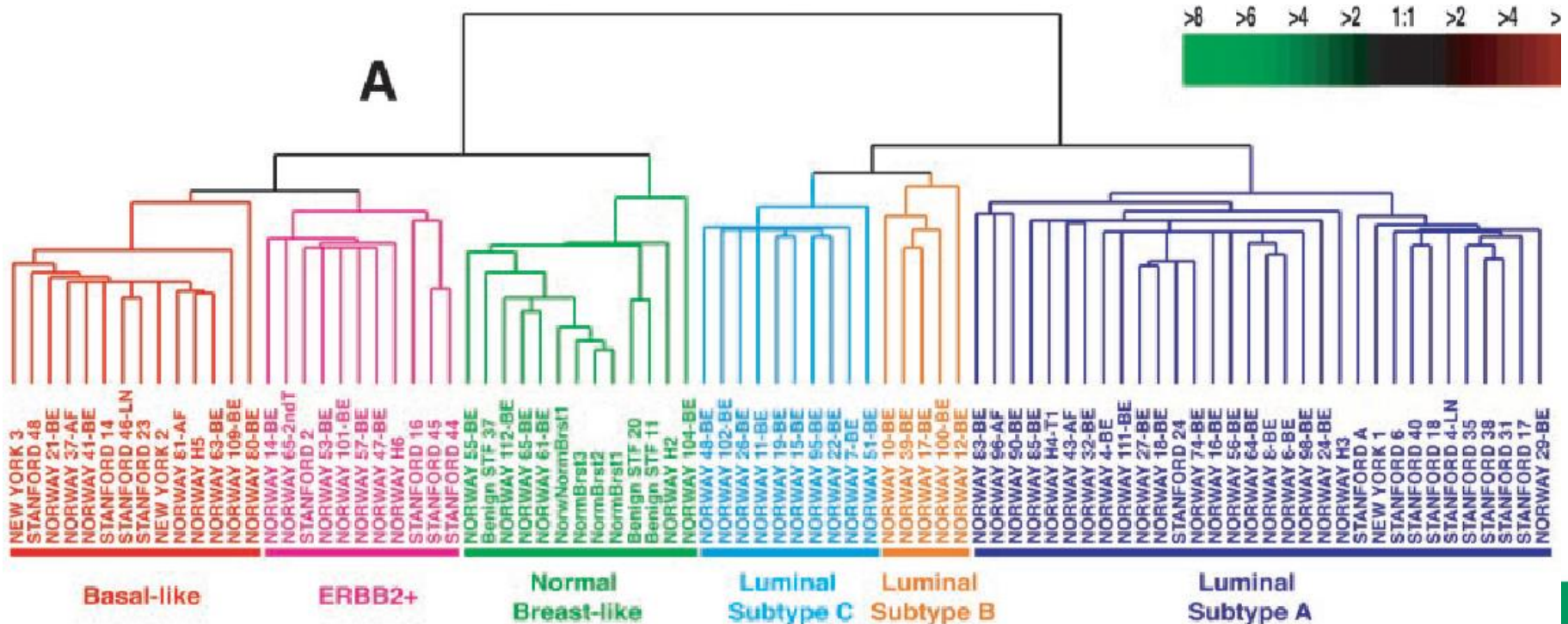


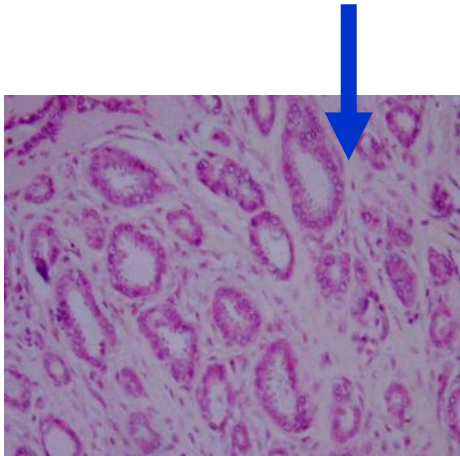
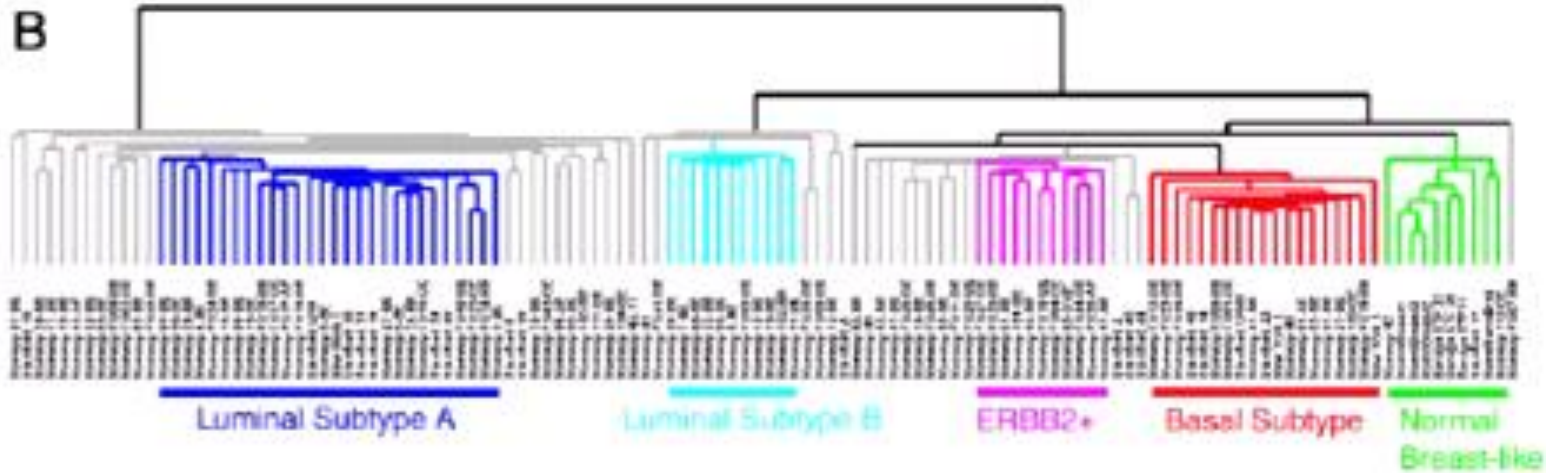
# Clustering Identifies novel and existing Subgroups of Breast cancer

Differential expressed genes: 476

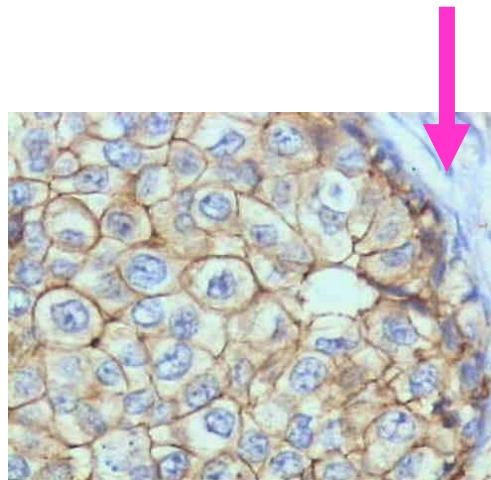
- tumor properties
- patient outcome

85 biopsy samples

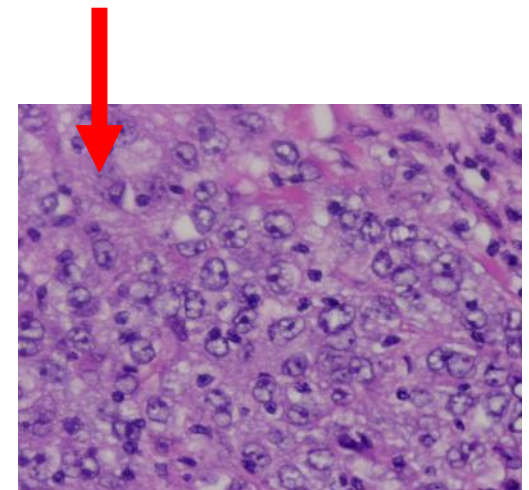




ER++, PR++, G1,2



HER2 ISH pos

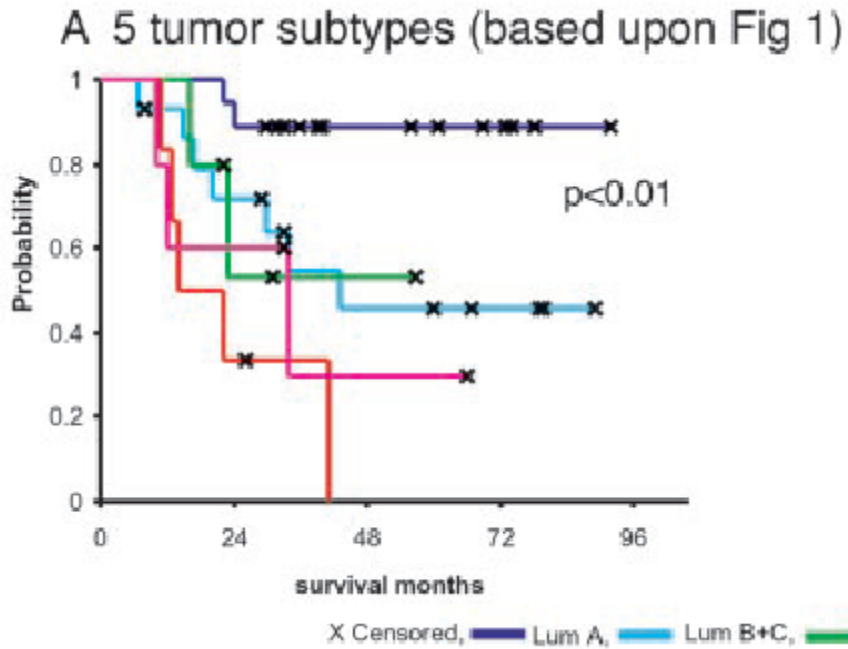


“triple neg,” CK5/6+

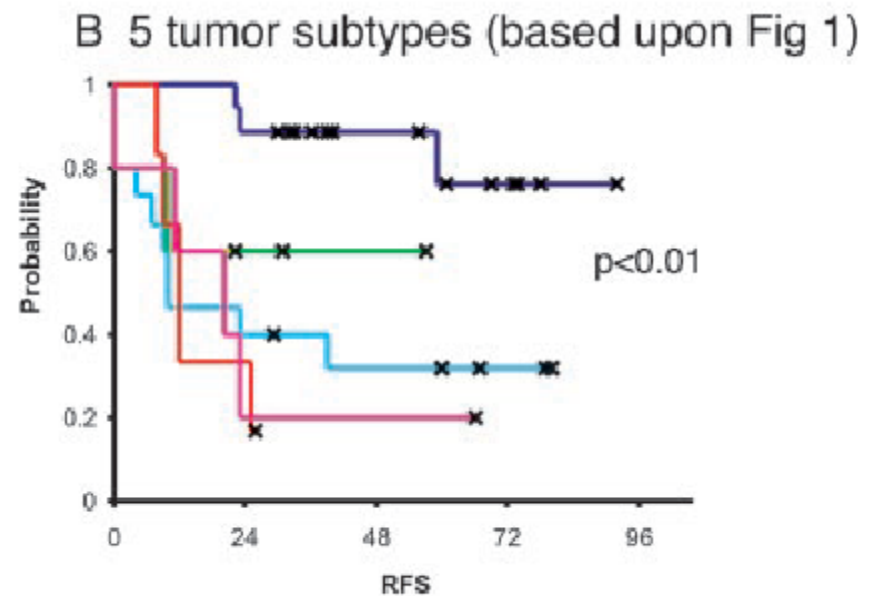


# Molecular classes are predictive of outcome

overall survival:



relapse-free survival:





## BioArray News (2, no. **35**, 2002)

### Arrays Hold Promise for Cancer Diagnostics

---

Oncologists would like to use arrays to predict whether or not a cancer is going to spread in the body, how likely it will respond to a certain type of treatment, and how long the patient will probably survive.

It would be useful if the gene expression signatures could distinguish between subtypes of tumours that standard methods, such as histological pathology from a biopsy, fail to discriminate, and that require different treatments.

---



# Gene expression profiling predicts clinical outcome of breast cancer

---

Van 't Veer, et. al., *Nature*, (415): 2002,530-536.

Aim:

to determine whether gene expression profiling could predict disease outcome and provide a strategy to select patients who would benefit from adjuvant therapy (metastasis)

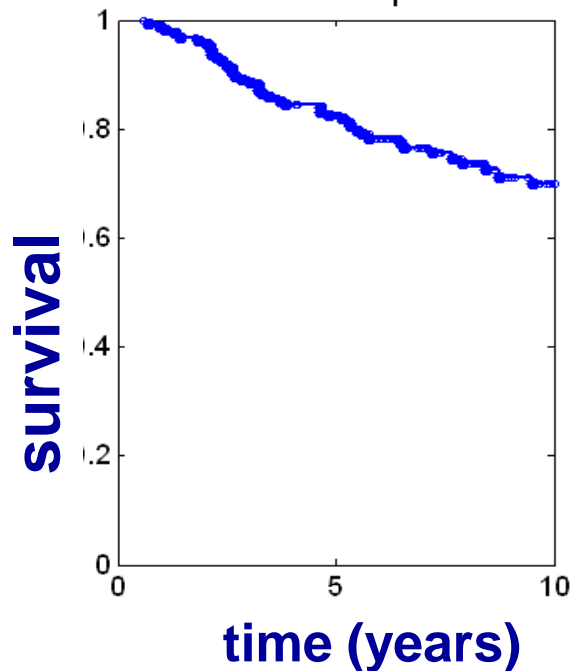
---





# Breast Cancer – Survival Pre-menopausal patients, lymph node negative

traditional diagnostics



**~30% die <10 year**

**~70% survive >10 year**

***Everyone receives chemotherapy...!***



# Breast Cancer – Survival Pre-menopausal patients, lymph node negative

---

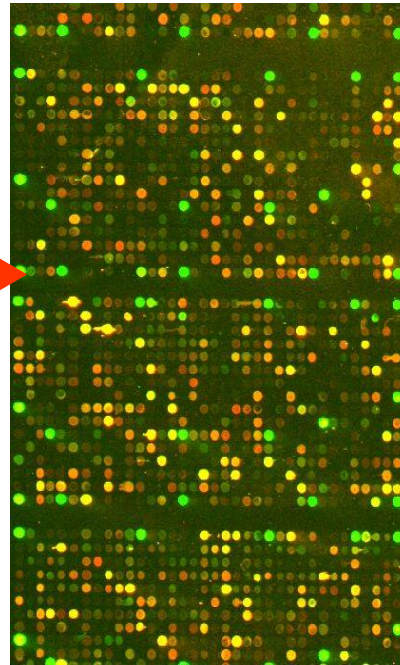
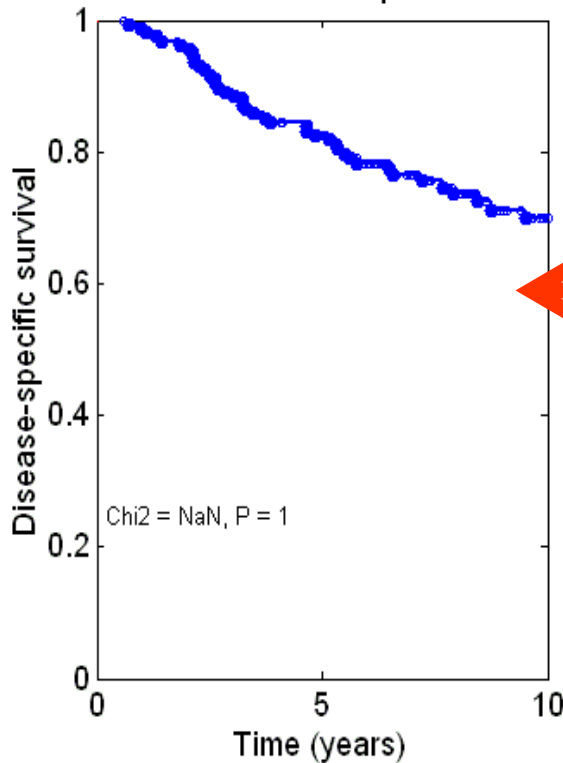
Current adjuvant treatment selection criteria:

- NIH (US) consensus criteria: > 95%
- St Gallen (EU) consensus criteria: > 80%  
receive adjuvant chemo- and hormonal therapy

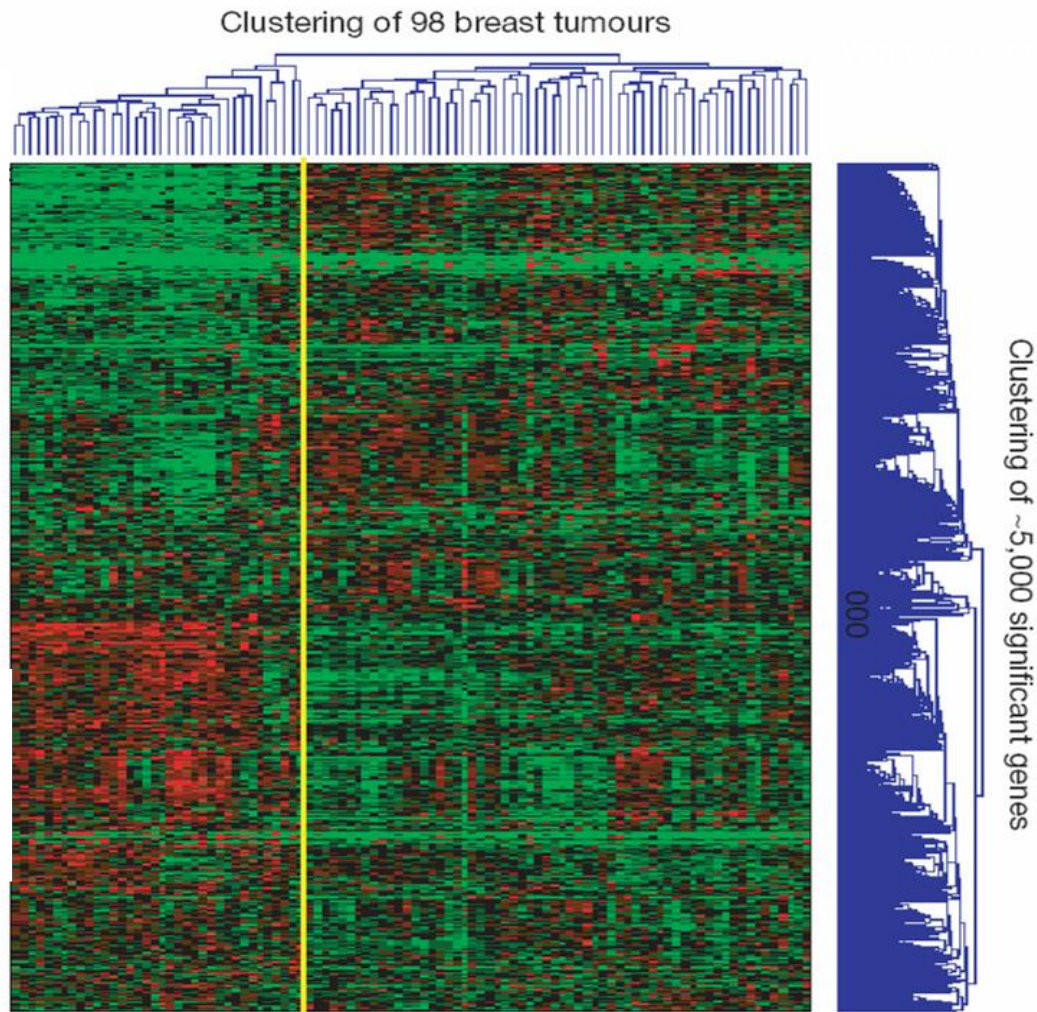
As only 30% of these patients develop distant metastases, some 50-65% of patients are over-treated with adjuvant (chemo)therapy

---

# Identification of gene expression changes in breast cancer



- analyse 98 breast tumors
  - 34 metastases-positive <5 year
    - bad prognosis
  - 44 metastases-negative >5 year
    - good prognosis
  - 18 BRCA1 +
  - 2 BRCA2 +
- } 'sporadic'



- 98 breast tumors analysed
- 
- 34 'bad' vs. 44 'good'
  - 18 BRCA1 +
  - 2 BRCA2 +
  - microarray with 24,000 genes
  - 5,000 genes showed expressional changes in tumors



*Different classes of breast tumors...!*



# 70-gene prognosis classifier for predicting risk of distant metastasis within 5 years

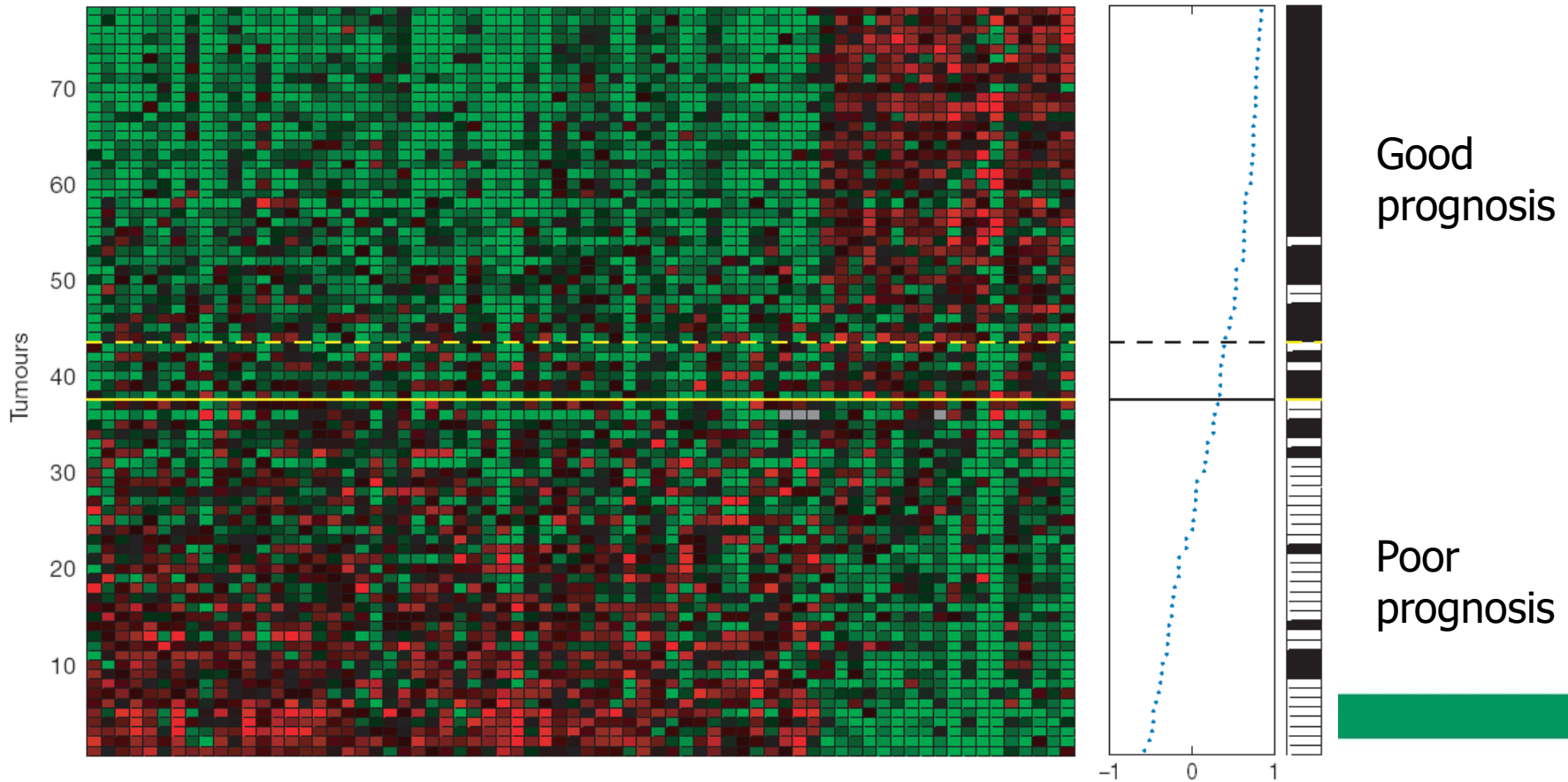
Sporadic breast tumours  
patients <55 years  
tumour size <5 cm  
lymph node negative (LN0)

Prognosis reporter genes

Distant metastases  
<5 years

No distant metastases  
>5 years

Supervised  
clustering

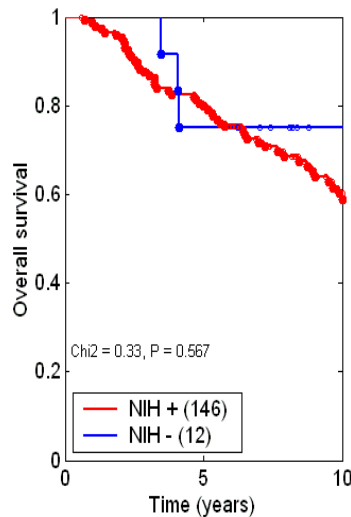






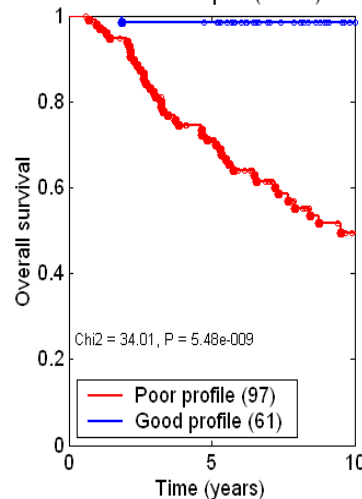
# Microarray classification vs. NIH classification

5 % low risk  
95 % high risk



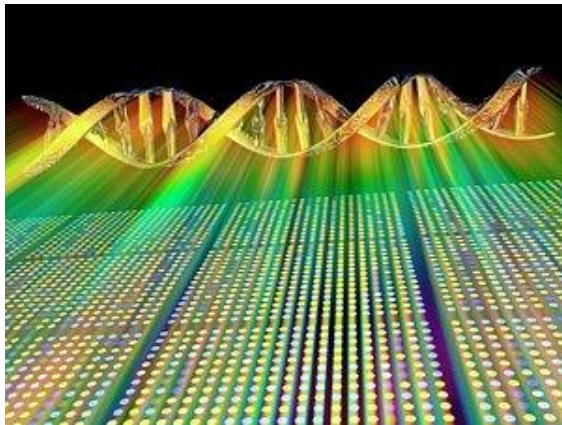
Classical  
NIH classification

39 % low risk  
61 % high risk



Classification based  
on microarray

- Classification of 158 breast cancer tumors
- Less unnecessary chemo-therapy
- Identification of genes playing a role in breast cancer



# Expression profiling & clinical application

---

“Though each tumor is molecularly unique, there exist common transcriptional cassettes that underlie biological and clinical properties of tumors that may be of diagnostic, prognostic and therapeutic significance”.

→ Also true for other complex diseases

---