微陣列分析
Microarray Analysis

Shin, Jyh-wei

Microarray Center,
Department of Parasitology,
Institute of Microbiology and Immunology,
Medical College, National Cheng Kung University
Tel: 886.6.235.35.35 ext 5586
e-mail: z8308010@email.ncku.edu.tw

Microarray Definition

Micro

adjective

Very small or microscopic.
Basic or small-scale.

Array

noun

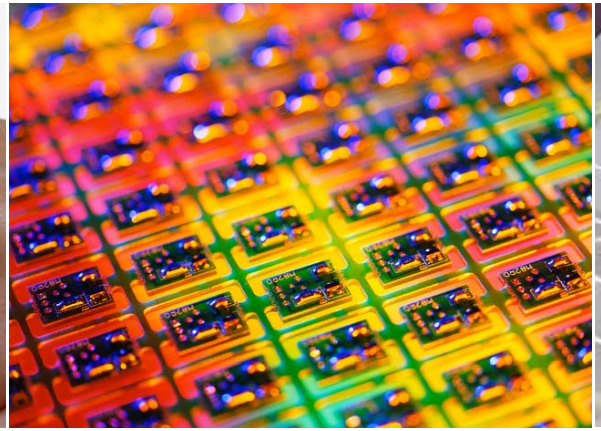
An orderly, often imposing arrangement.
An impressively large number, as of persons or objects.
Splendid attire; finery.

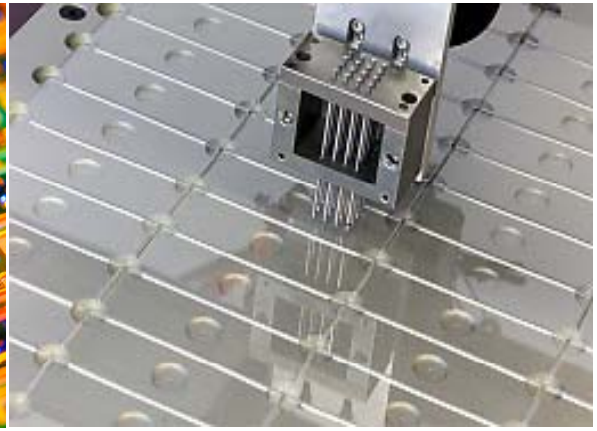
Microarray

noun

A collection of microscopic spots attached to a solid surface, such as glass, plastic or silicon chip forming an array.

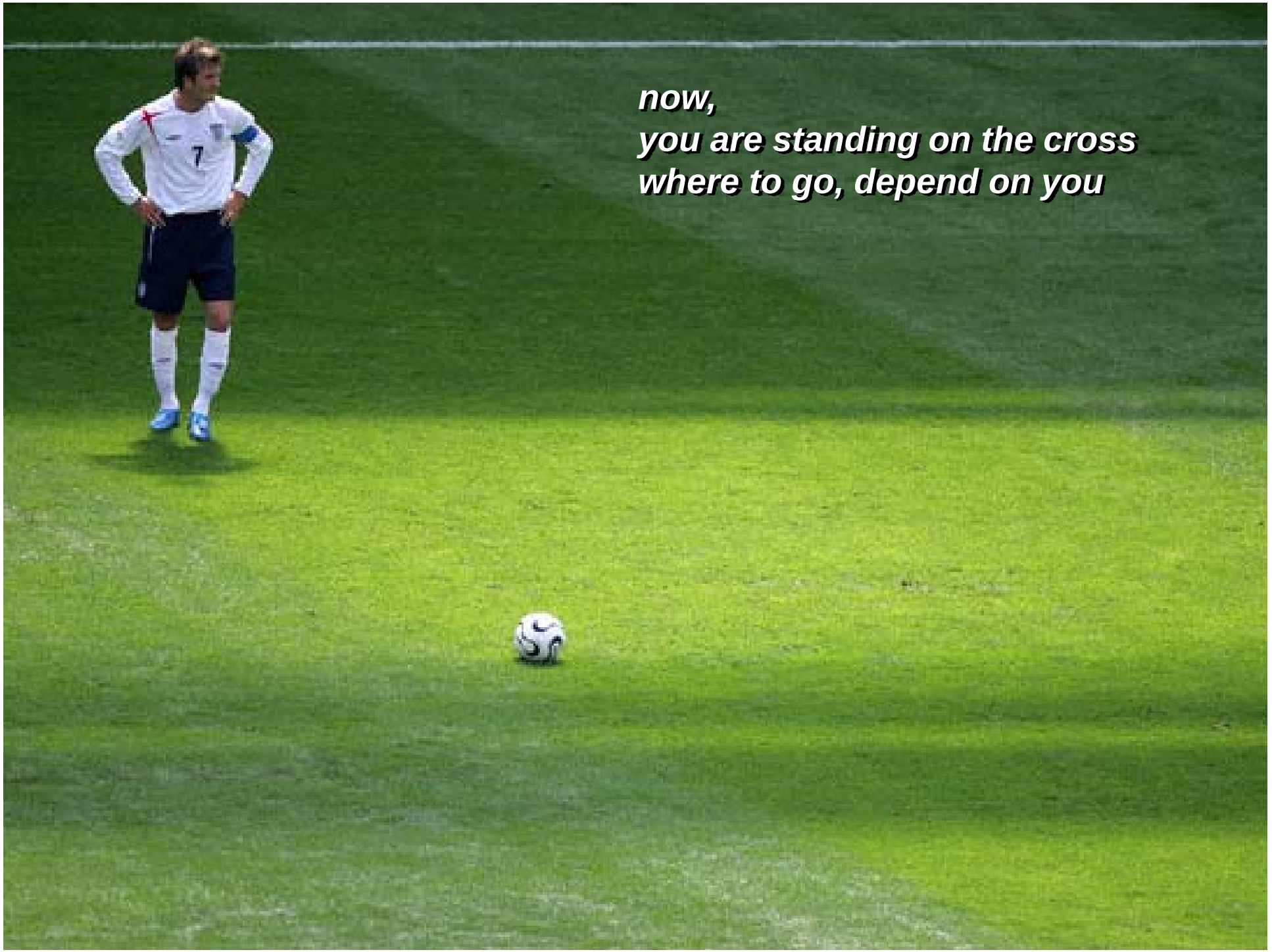




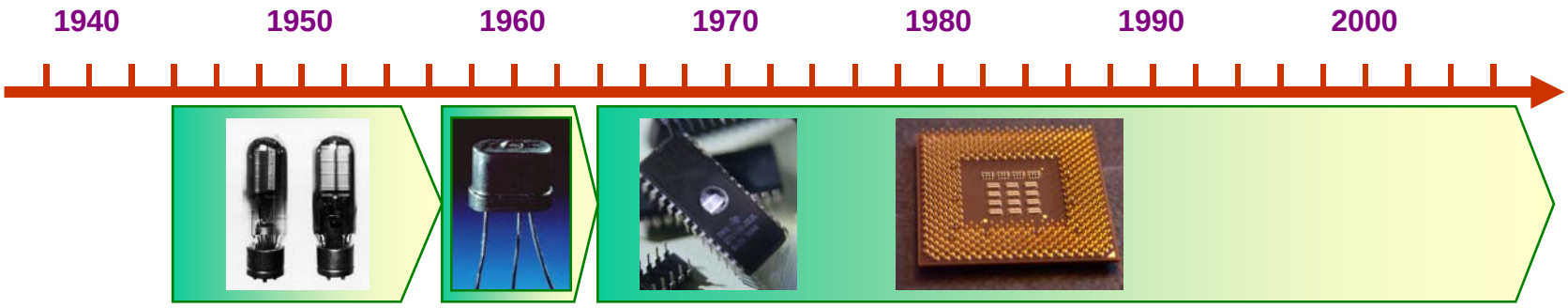


改變二十世紀的大事




now,
you are standing on the cross
where to go, depend on you

Generations of Computer



Vacuum tube

Transistor

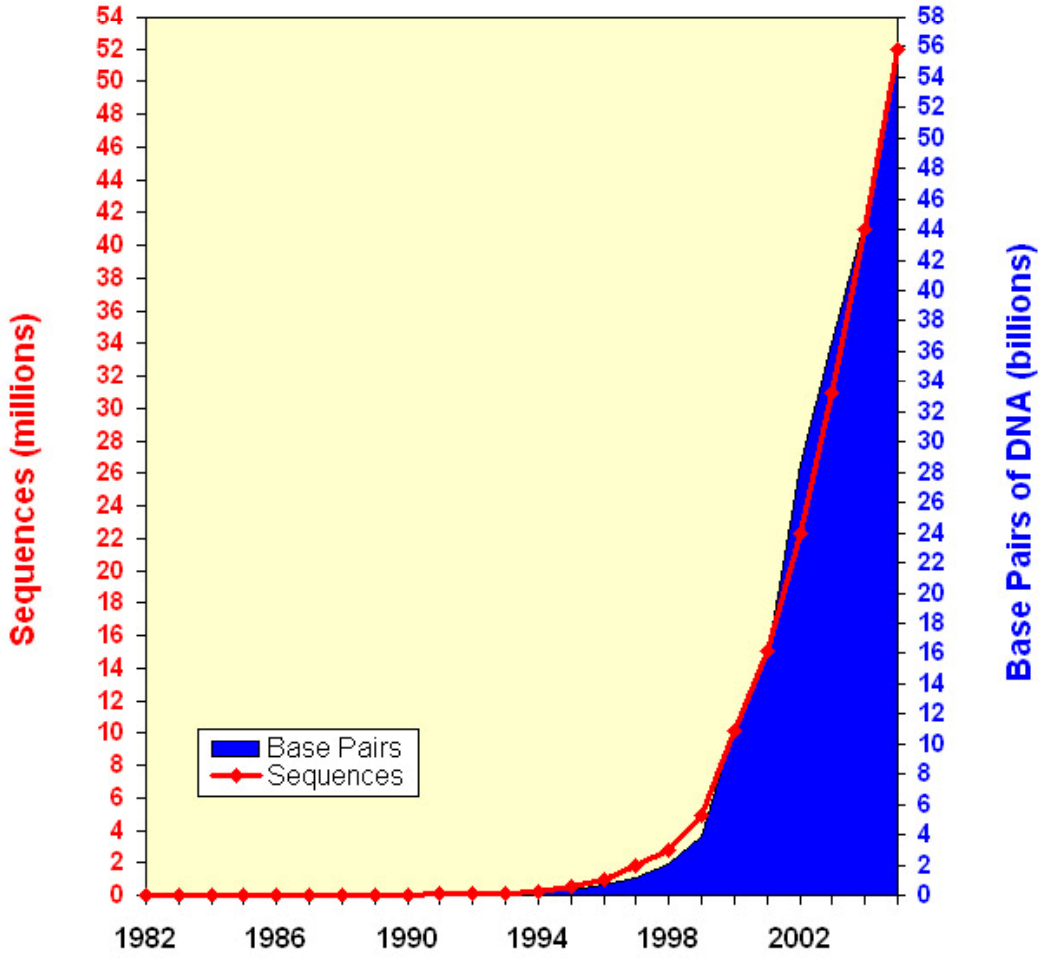
Small scale integration - 1965 on
Up to 100 devices on a chip
Medium scale integration - to 1971
100-3,000 devices on a chip

Large scale integration - 1971-1977
3 – 100K devices on a chip, semi conductor memory
Very large scale integration - 1978 to date
100K – 100M devices on a chip
Ultra large scale integration
Over 100M devices on a chip



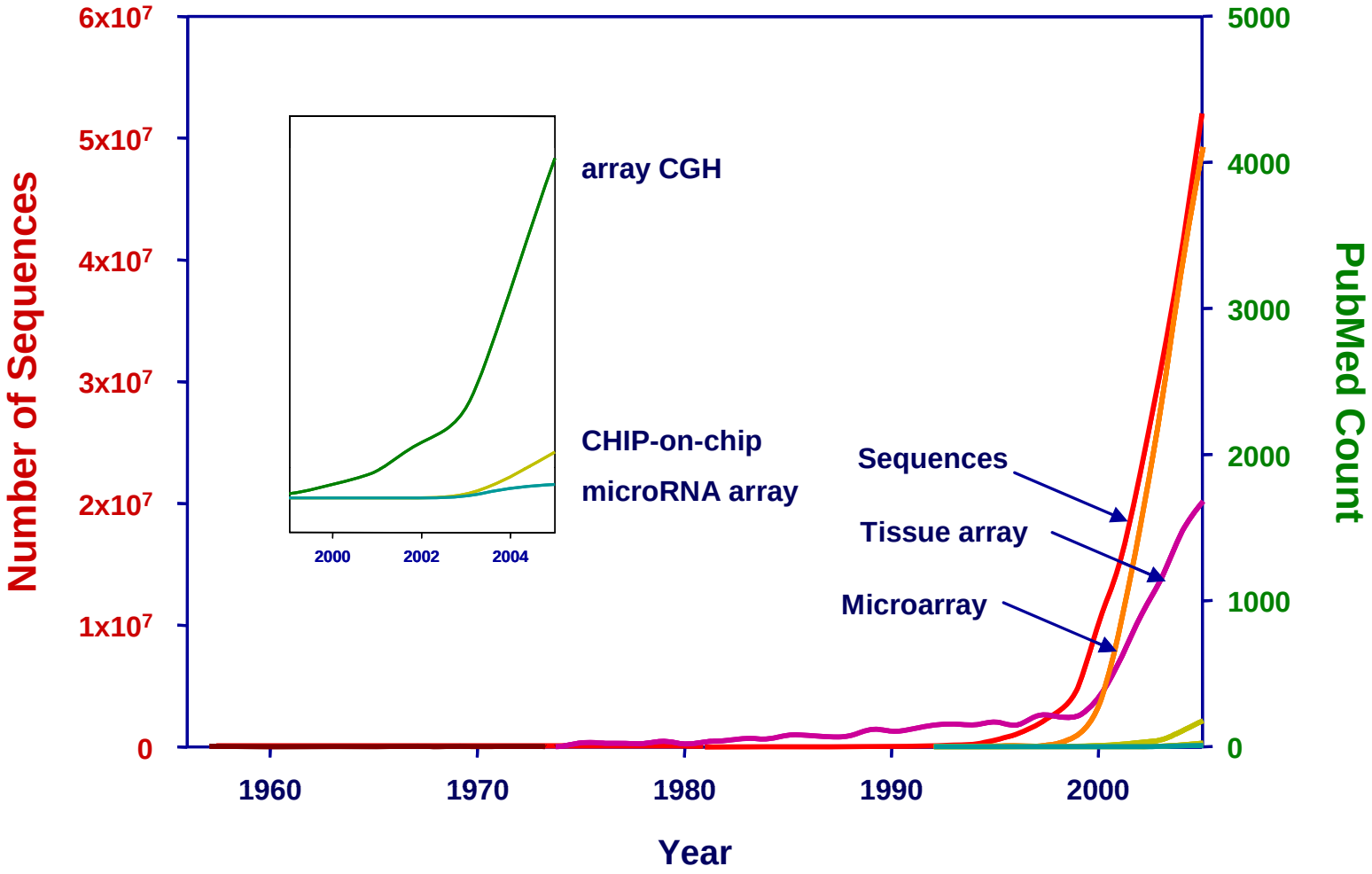
Growth of GenBank

(1982 – 2005)



Evolution of microarray (PubMed base)

up to 2005



Microarray Historical Perspective



1st Sales







nature biotechnology

1999 | 2000 | 2001 | 2002 | 2003 | 2004

THE UNIVERSITY of WISCONSIN MADISON

CHEMOGENIX







>1000 Array Designs

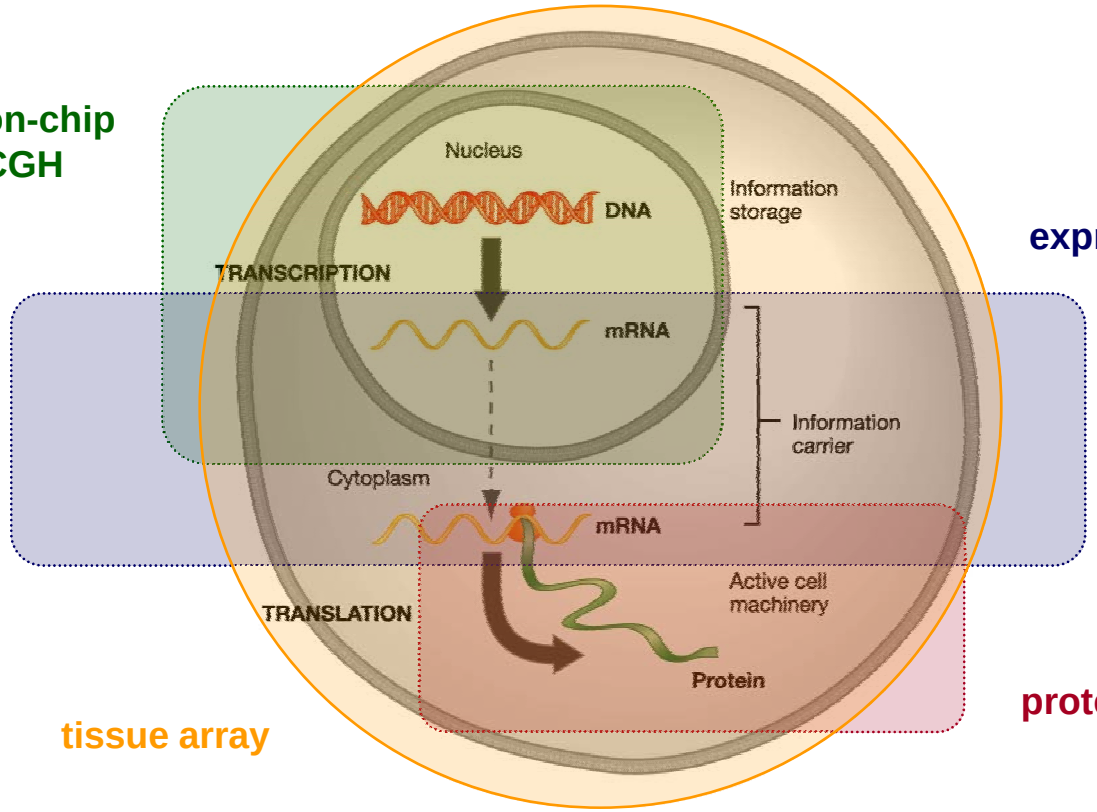
Xeotron purchased by Invitrogen, Febit bankruptcy, NimbleGen acquires Light Biology

$13M financing in Oct '04, bringing total to >$40M: Investors include Baird, Cargill, Tactics II, Venture Investors, Topspin, SWIB, WARF

Organization → Technology and Applications Development → Commercialization →

The Central Dogma

DNA → Transcription → RNA → Translation → Protein



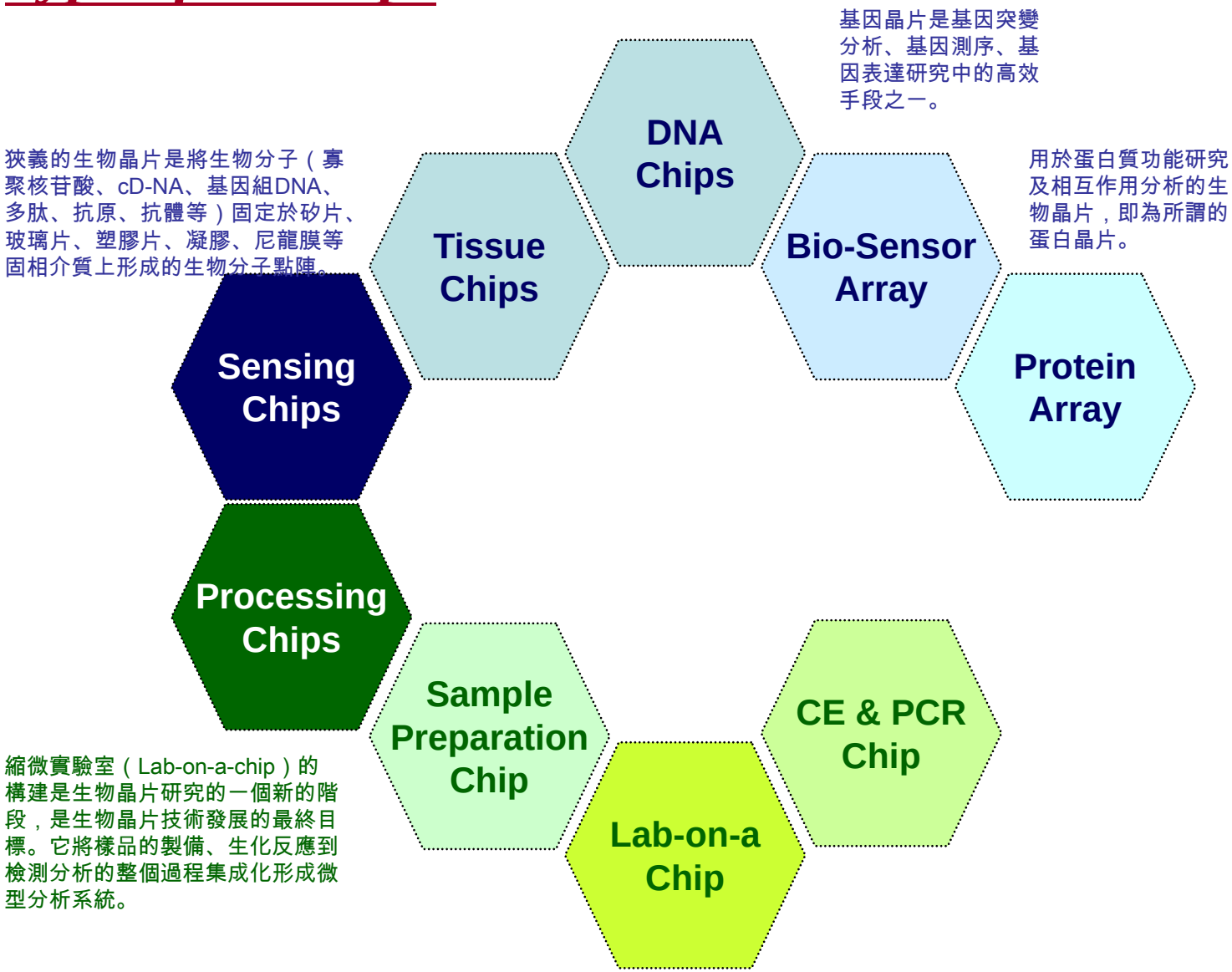
CHIP-on-chip
array CGH

expression array

tissue array

protein chip

Type of Biochips



Lab-on-a-Chip





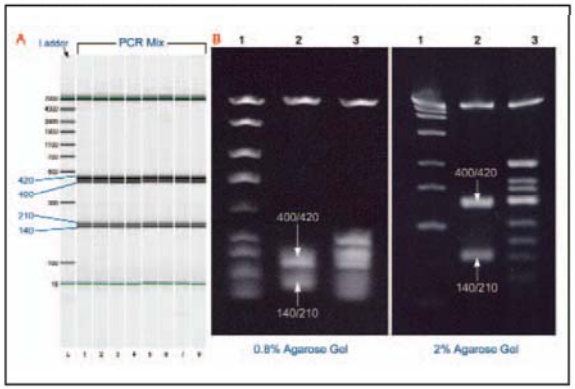
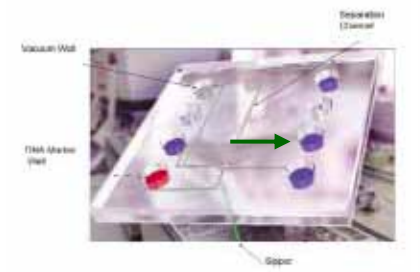
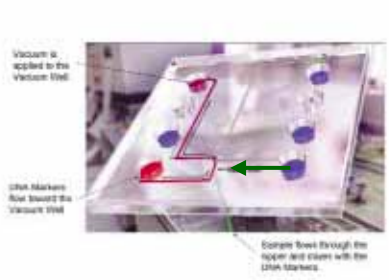


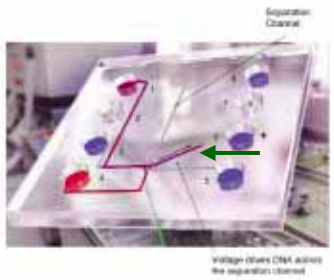
Figure 1. A) A LabChip 90 virtual gel showing PCR fragment resolution. B) Agarose gels - lane 1: ladder; lane 2: PCR fragments; lane 3: PhiX174/ *Hinf*I.



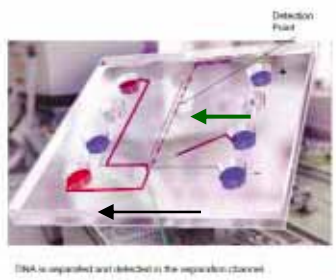
Sip reagents into storage wells



Mix and react sample with reagent



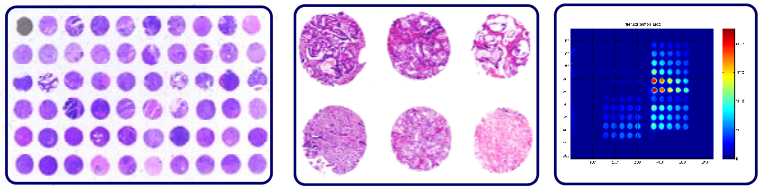
Flow products to separation column



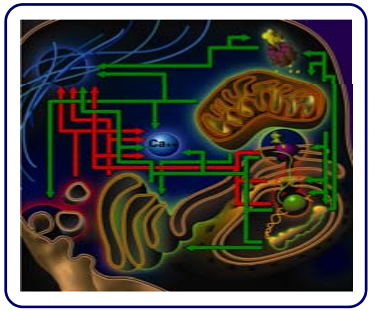
Run electrophoretic separation

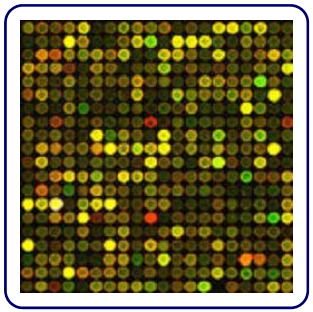
Types of Microarray

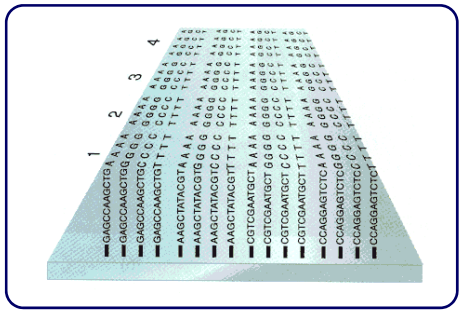
- **tissue chip**
  - tissue array



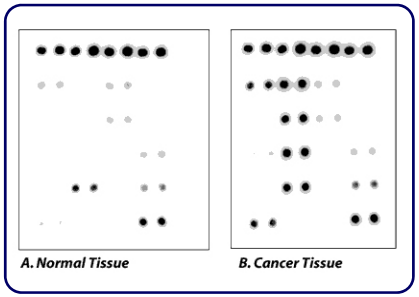
- **DNA chip**
  - expression array
    - cDNA array
    - oligo DNA array
  - CHIP-on-chip array
  - microRNA array
  - CGH array

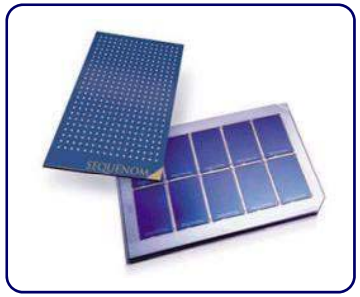




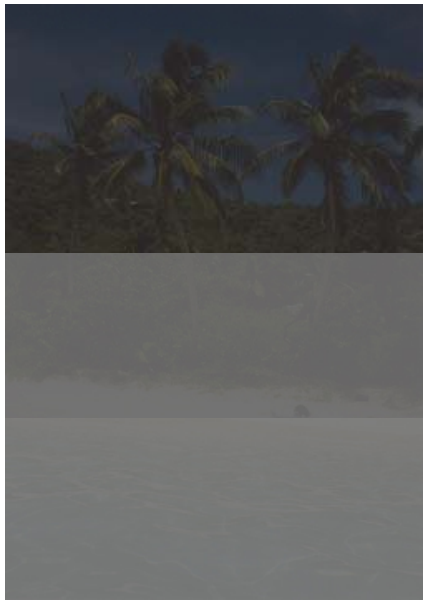


- **protein chip**
  - antibody array
  - MALDI-TOF array





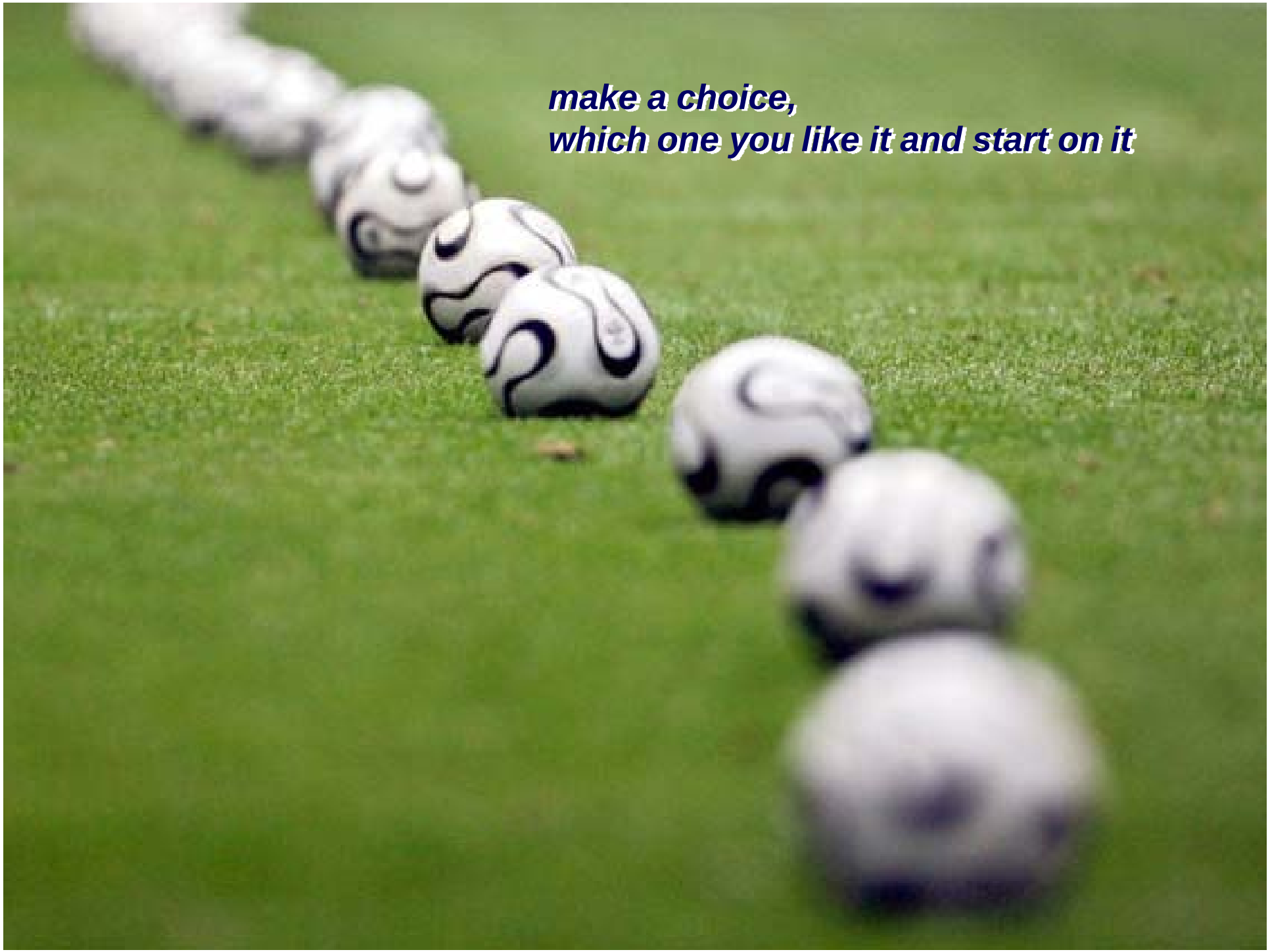
Digital Image Analogy

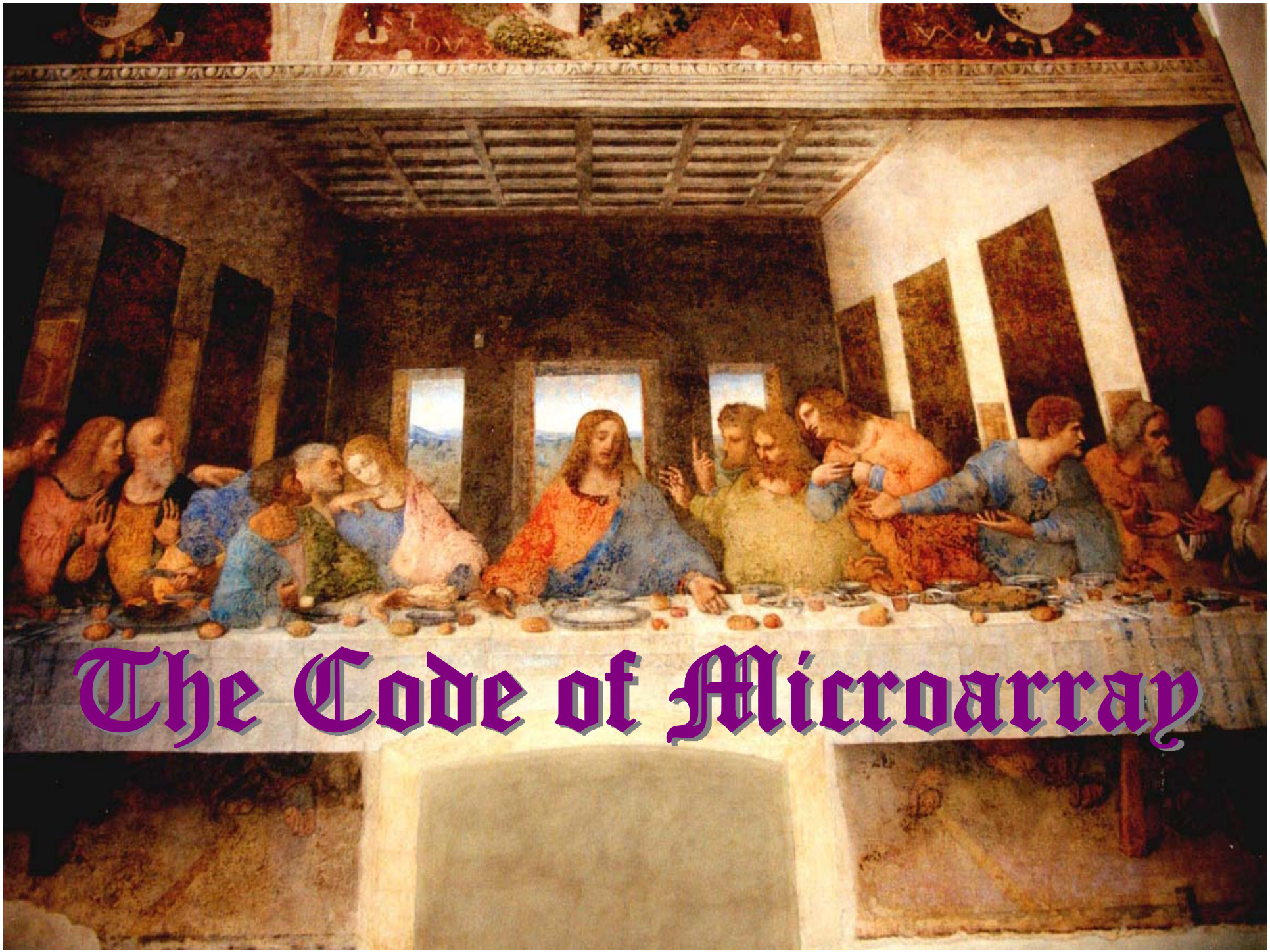
	Northern blot	Customized	Commercial
Expression			
CGH (LOH)			
ChIP, CGR, Exp Tiling, Methylation, …			
	Low density ▪ low content	**High density** ▪ high content **Short probes** ▪ limited contrast ▪ limited channels	**High density** ▪ high content **Long, isothermal probes** ▪ high contrast ▪ multiple channels

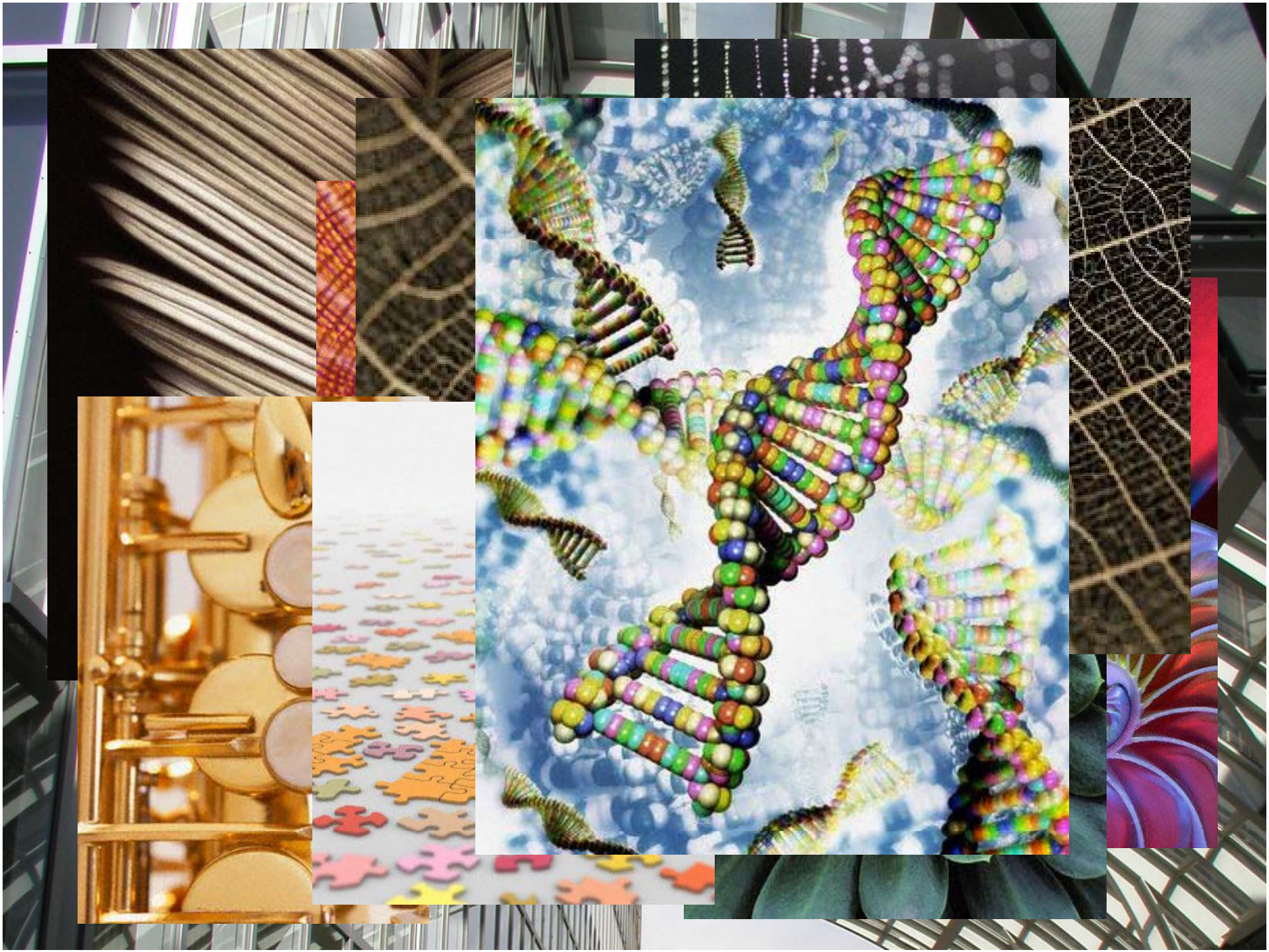
Advantages of Microarray

1. *Amplification of a scarce resource*
2. *Experimental uniformity*
3. *Decreased assay volume*
4. *Does not destroy original block for diagnosis*



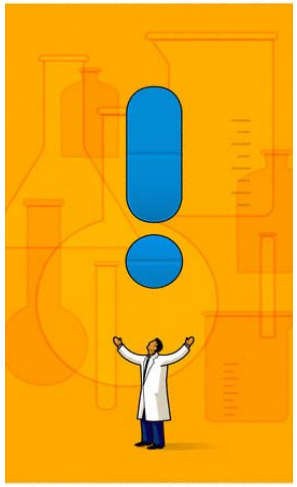

make a choice,
which one you like it and start on it


The Code of Microarray

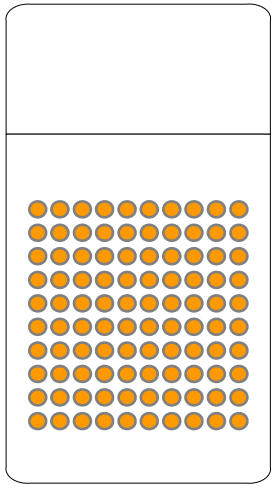


Key Steps in Applications of Array Data

1. **Experimental Design**



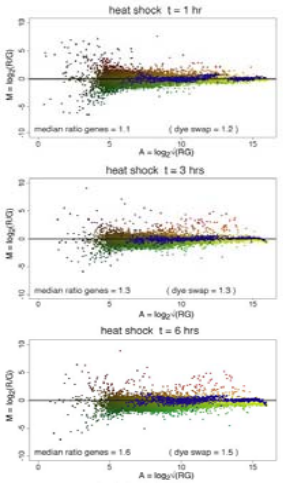
2. **Microarray Experiments**



3. **Image Processing**



4. **Data Normalization**



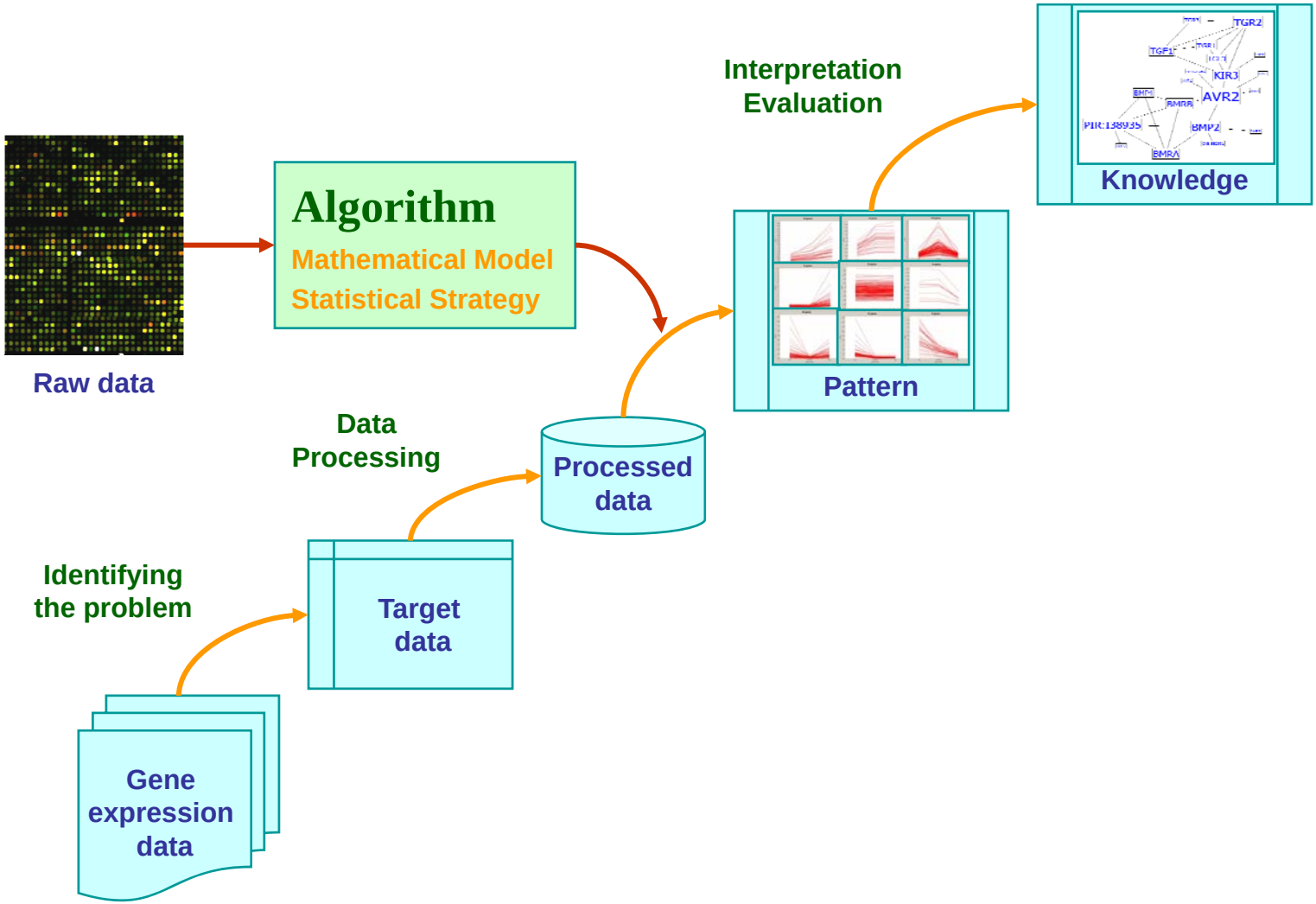
5. **Data Analysis**

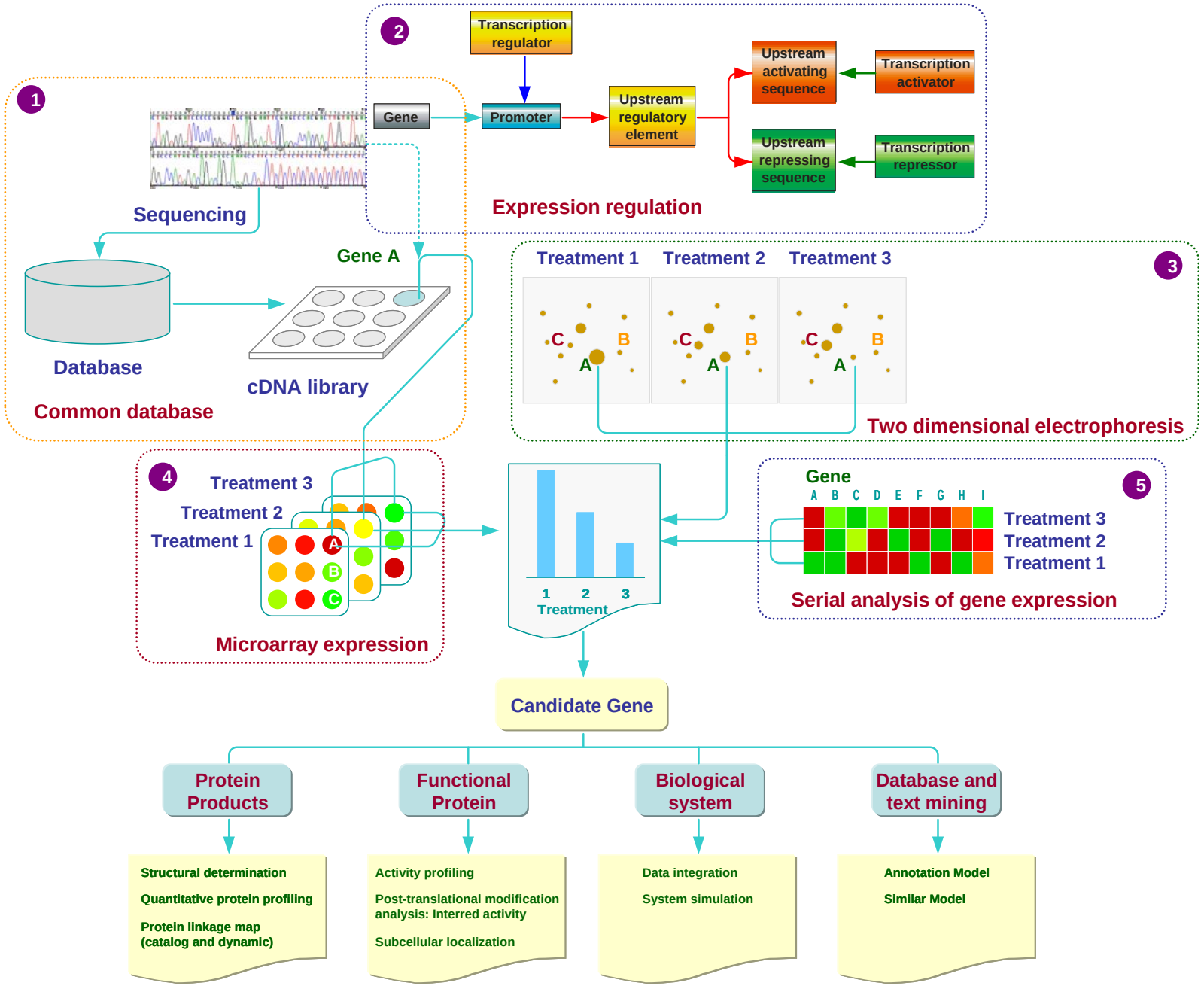


Bioinformatics Issues in Microarray Experiment

1. How to manage data ?
2. Which data should be stored ?
3. Validated database or just data storage ?
4. How to integrate in-house data with public repositories ?
5. Which databases ?
6. Local mirror DBs versus web link ?
7. How to analyze and visualize data ?
8. Which analysis should be standardized ?
9. Buy versus develop ?
10. How to integrated microarray data with other omics data ?

An Overview of Data Mining Process

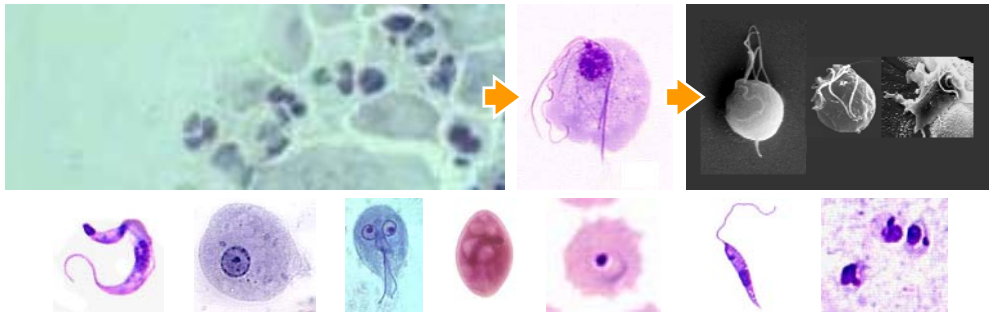


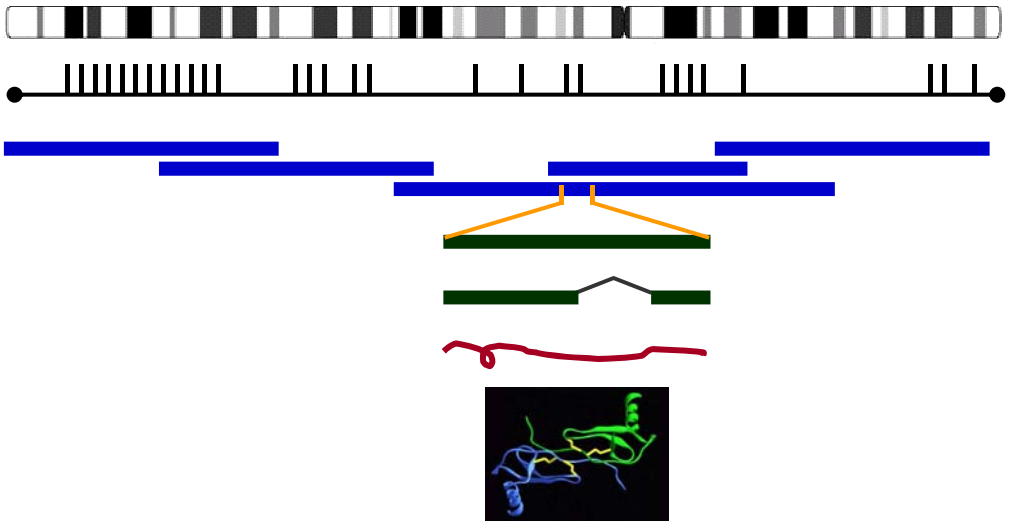

1
2
3
4
5
Transcription regulator
Promoter
Gene
Upstream regulatory element
Upstream activating sequence
Transcription activator
Upstream repressing sequence
Transcription repressor
Expression regulation
Sequencing
Gene A
Database
cDNA library
Common database
Treatment 1
Treatment 2
Treatment 3
C
B
A
Two dimensional electrophoresis
Treatment 3
Treatment 2
Treatment 1
Microarray expression
1 2 3
Treatment
Gene
A B C D E F G H I
Treatment 3
Treatment 2
Treatment 1
Serial analysis of gene expression
Candidate Gene
Protein Products
Functional Protein
Biological system
Database and text mining
Structural determination
Quantitative protein profiling
Protein linkage map (catalog and dynamic)
Activity profiling
Post-translational modification analysis: Interred activity
Subcellular localization
Data integration
System simulation
Annotation Model
Similar Model

Data organization: Basic Model

Research

- Population
- Organism
- Development
- Phenotype
- Genetics
- Chromosomes
- Scaffolds
- Gene
- mRNA
- ProProtein
- Mature peptide
- Structure
- Domains
- Function
- Disease
- Phenotype





OMIM
Online Mendelian Inheritance in Man

OMIA
Online Mendelian Inheritance in Animals

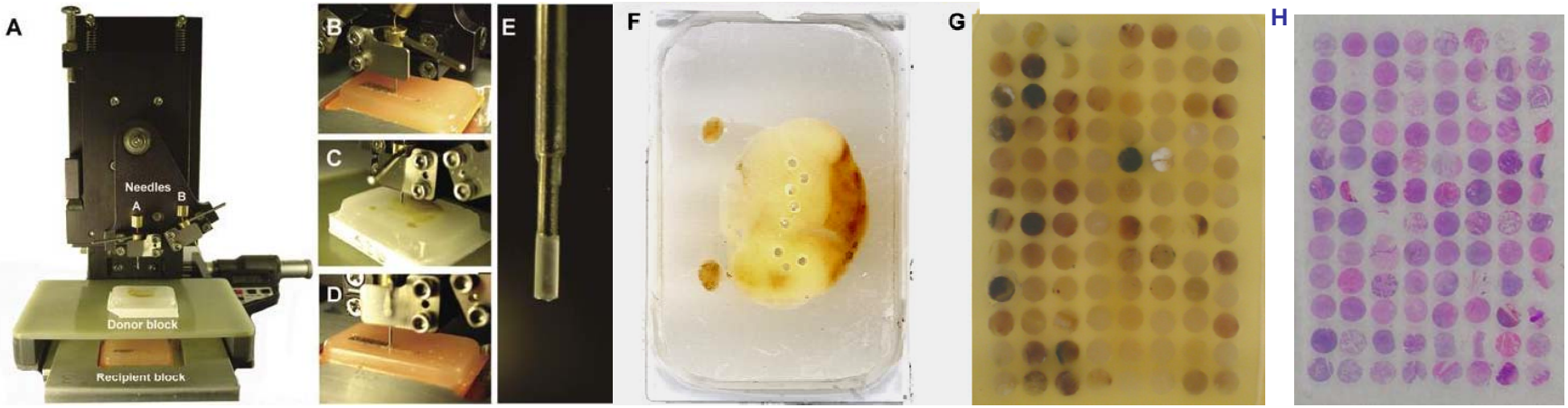
Bookshelf

PubMed
A service of
www.pubmed.gov

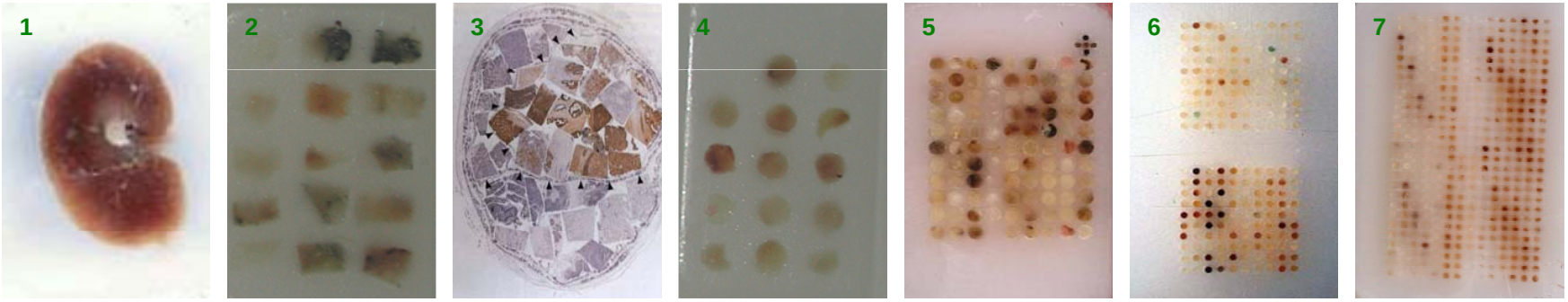
Resource

- Taxonomy
- Genome Project
- Genome Resources
- Map Viewer
- OMIM, OMIA
- Genome
- Trace
- Gene
- Nucleotide
- Protein
- RefSeq
- MMDB, Cn3D
- CDD
- Books
- OMIM, OMIA
- PubMed
- PubMed Central

Tissue Array Evolution



donor block array block slide



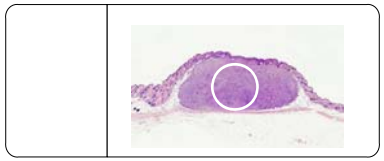
whole tissue

tissue cocktail

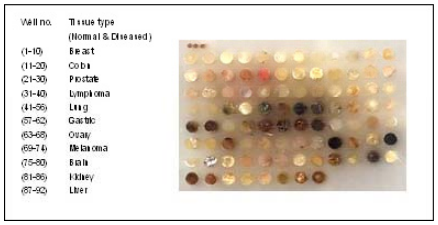
tissue sausage

tissue array

Tissue Array Section Preparation



Tumor area identified and marked on slide



Tissue array product





Tumor area marked on block and microdissected with tissue arrayer

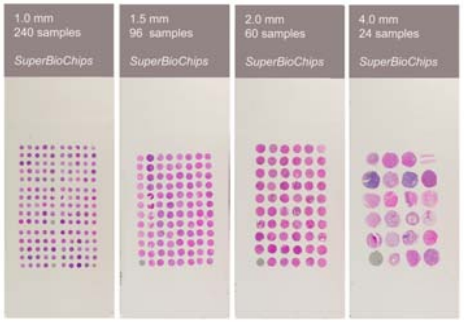


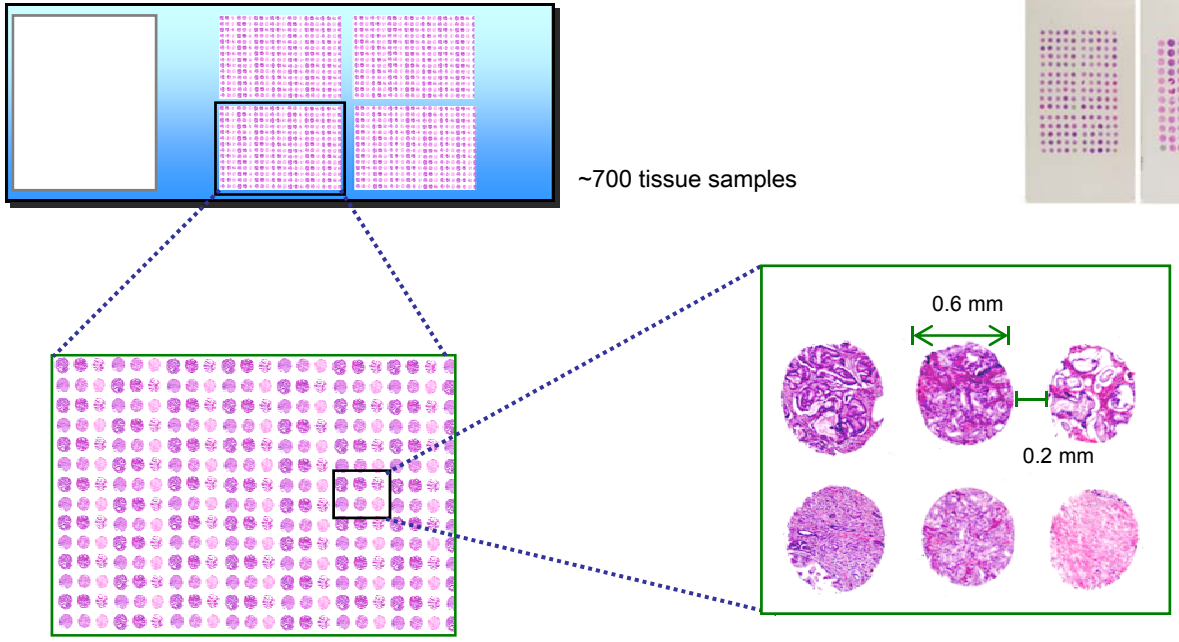


Microarray assembly

Tissue Microarray (TMA) *Technology*

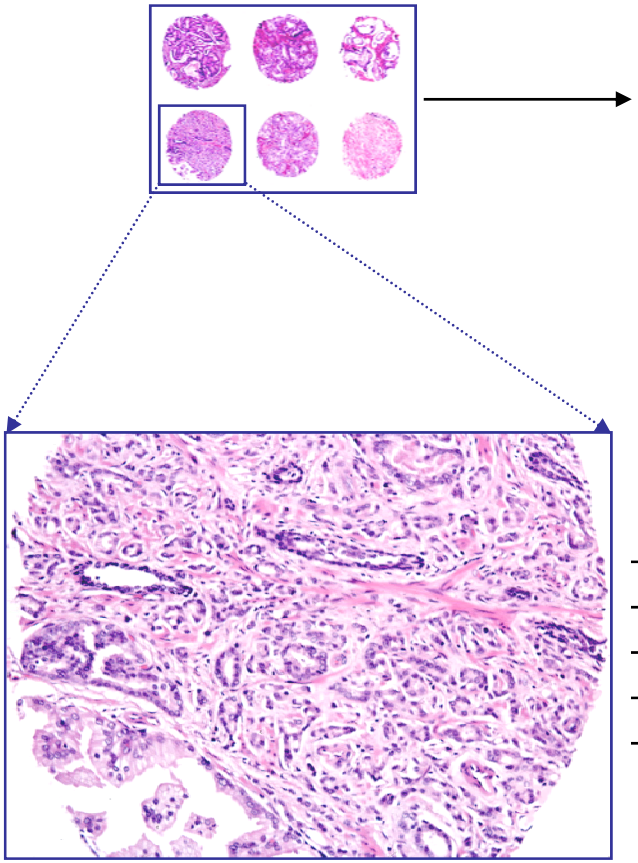
- Hundreds of tiny (typically 0.6 mm diameter) cylindrical tissue cores (densely and precisely arrayed into a single histologic paraffin block)
- From this new array block, up to 300 serial 4 - 8 µm thick sections may be produced.
- Targets for protein expression by immunohistochemical studies and fluorescence *in situ* hybridization (FISH).





Multiple measurements per patient:

Several spots per tumor sample and several "scores" per spot

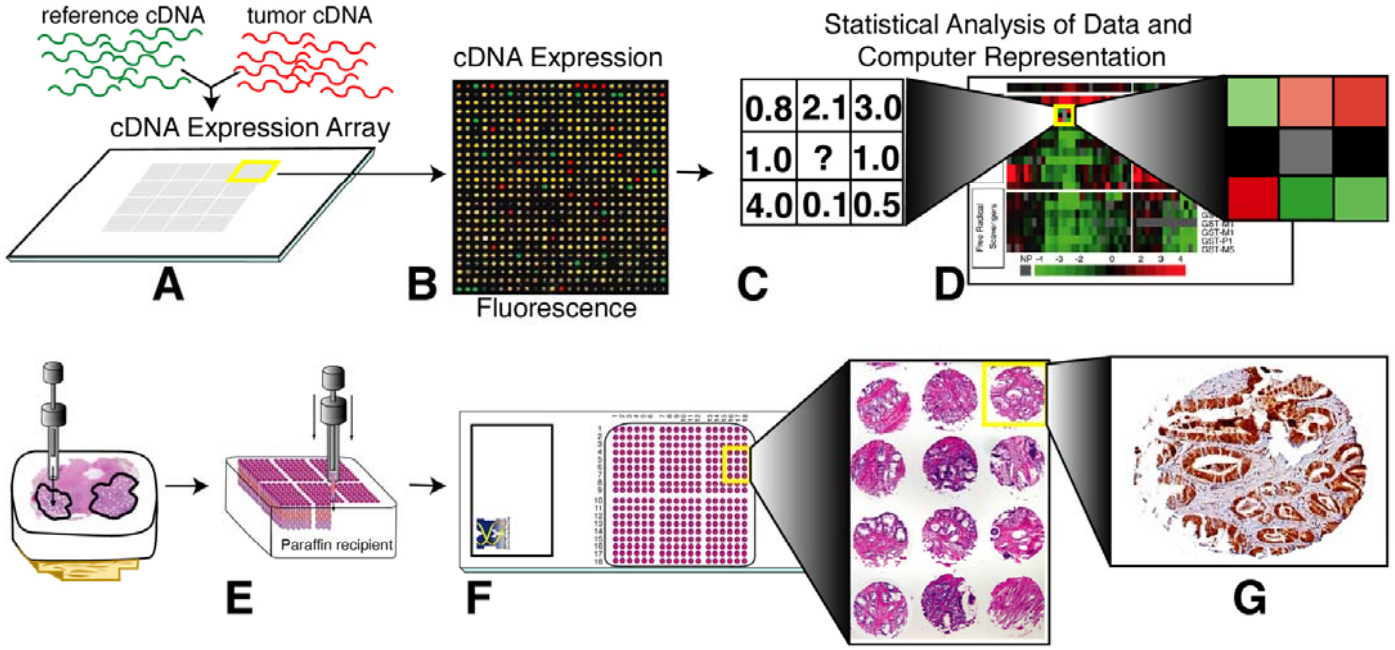


Each patients (tumor sample) is usually represented by multiple spots

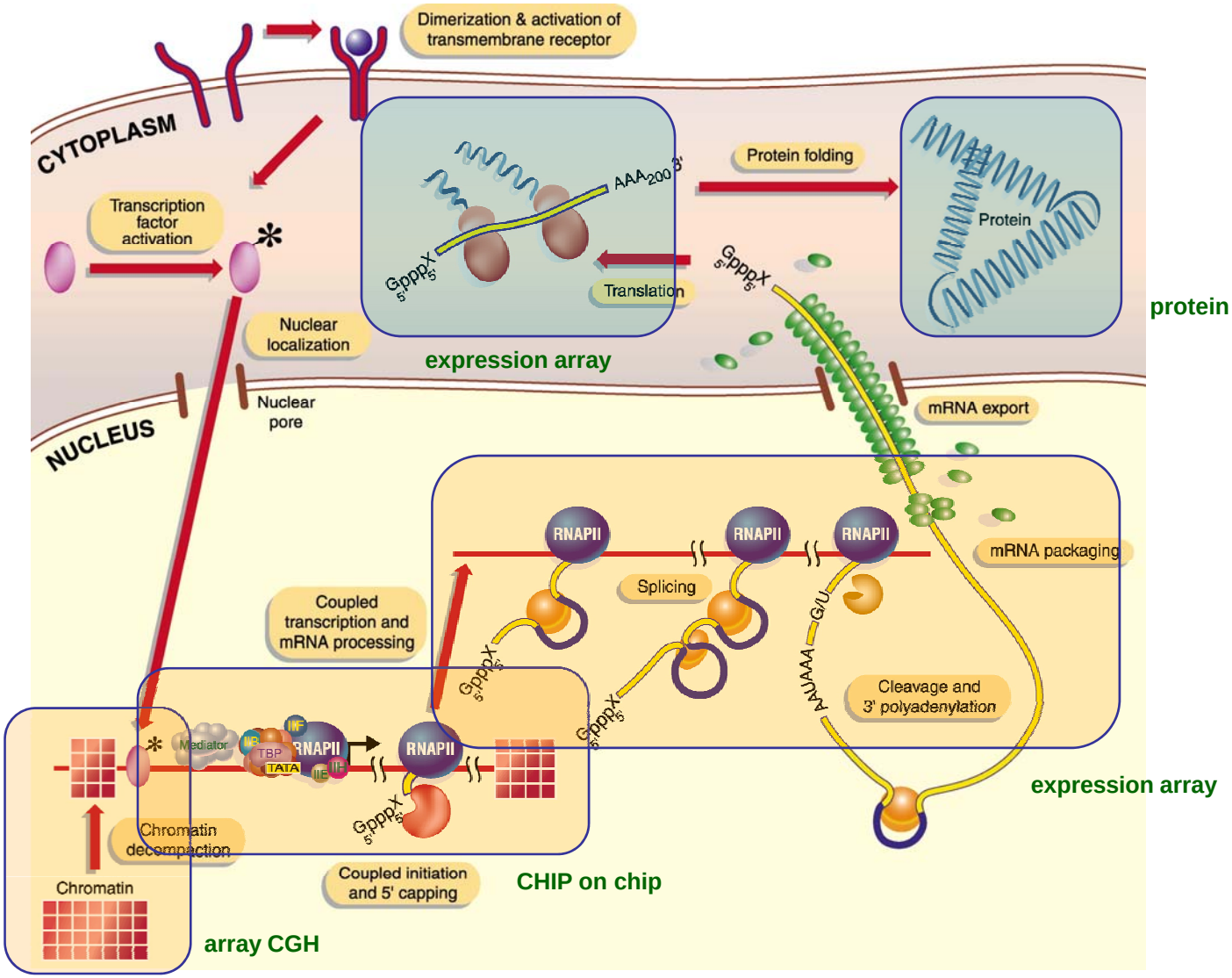
- **3 tumor spots**
- **1 matched normal spot**

- Maximum intensity = Max
- Percent of cells staining = Pos
- Percent of cells staining with the maximum intensity = PosMax
- Spots have a spot grade: NL,1,2,..
- Indicator of missingness

cDNA and Tissue Microarray Paradigm

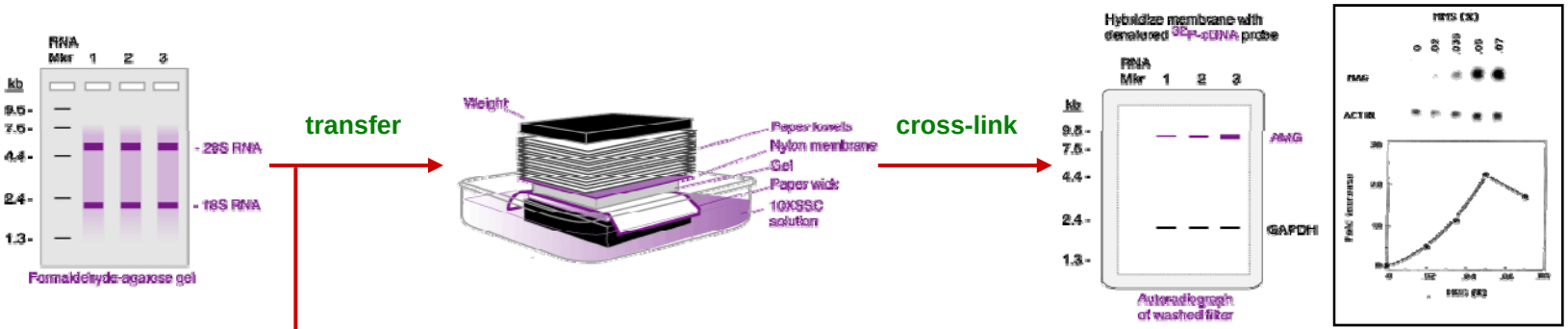


Central Dogma and Application of Microarray



Northern blotting versa *DNA microarray*

One gene (clone)!

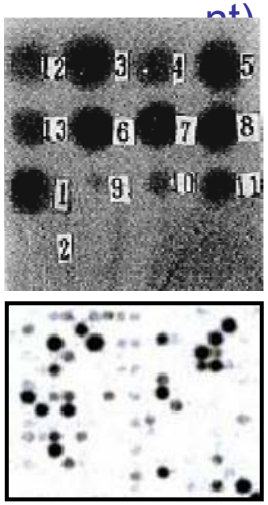


We could just 35,000 Northern to monitor expression of all genes!!!

DNA Microarray Evolution (Platforms and Formats)

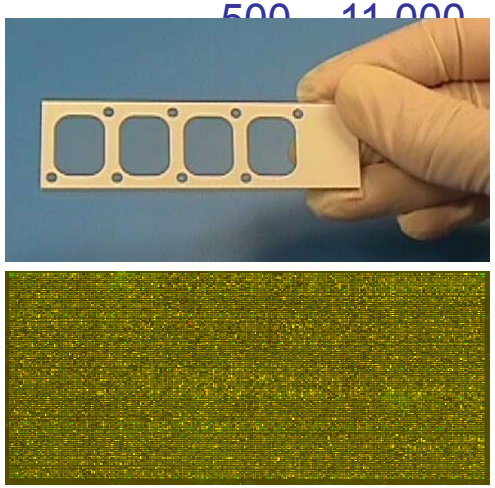
Isotope

Nylon
cDNA
(300-900
nt)



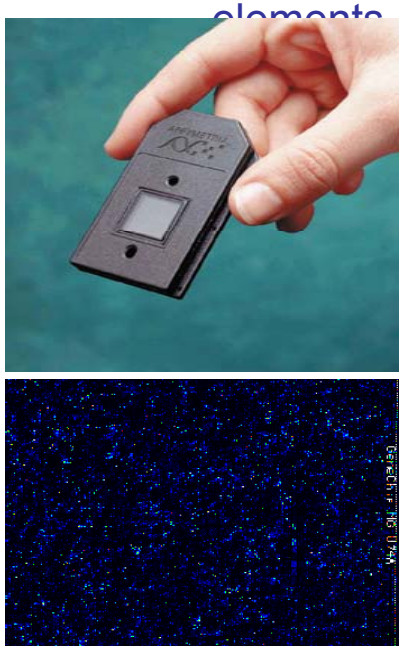
Two-colour

Glass
cDNA or
Oligo (80 nt)
500 – 11,000



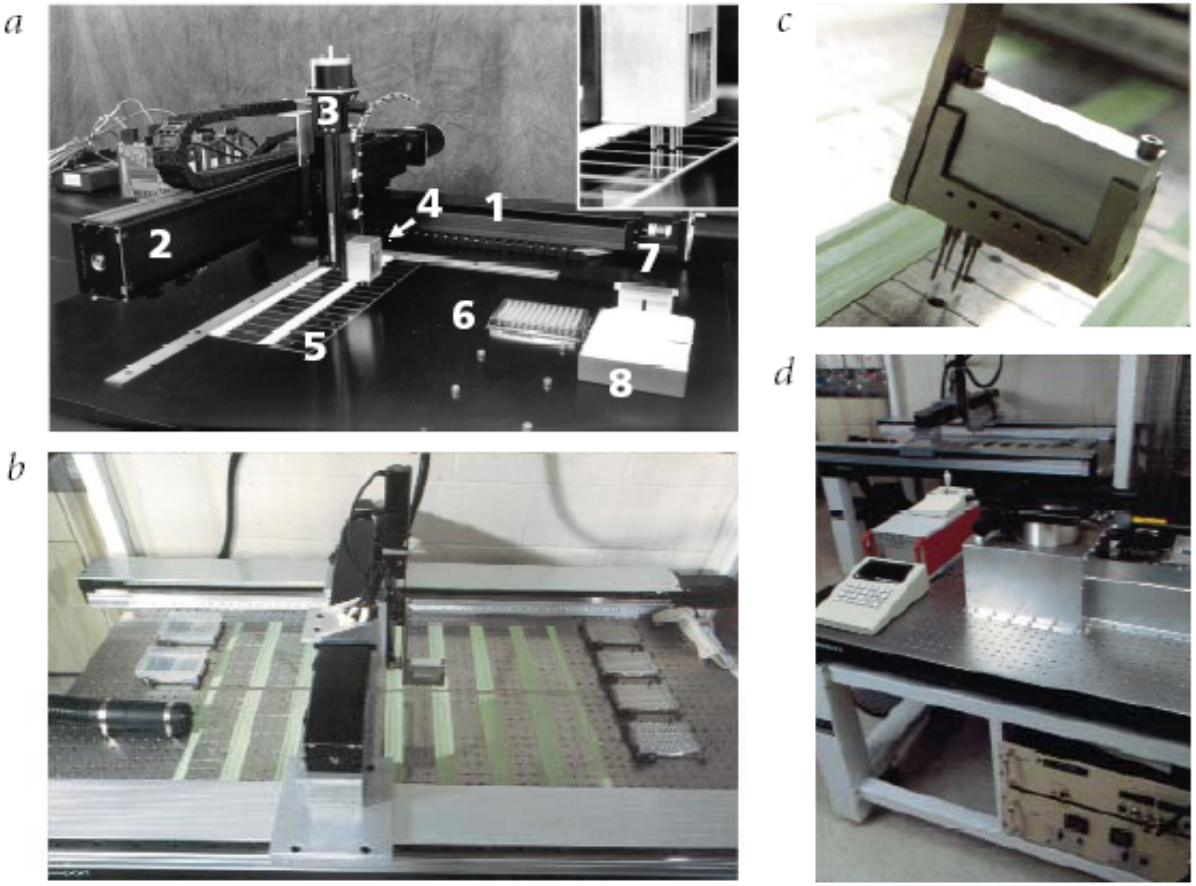
Affymetrix

Silicone –
oligo (20 nt)
22 ,000
elements



How do we manufacture a microarray ?

Fig. 1 ***a,*** Penn microarray robot. The X-, Y-, Z- axes are labeled 1, 2, and 3, respectively. The key component of the arrayer is the print-head, containing pens (4). Microscope glass slides are placed on the slide station (5). Samples are prepared and arrayed from 96-well sample plates (6). The pins are cleaned between sample acquisitions at the washing (7) and drying (8) stations. ***b,*** AECOM microarray robot. The table configuration shown contains 160 slides with four microtitre plates, two wash stations and the dryer. The print-head (***c***) shows four of the possible twelve pen tips in use. ***d,*** AECOM laser scanner. Visible are the optical table, power supplies for lasers and PMT cooling, the Ludl stage, and lasers. The 20× microscope objective is inside the ludl stage while lenses, mirrors and other optics are enclosed in the metal casing. PMTs are to the right and outside the photo.



Microarray Scale

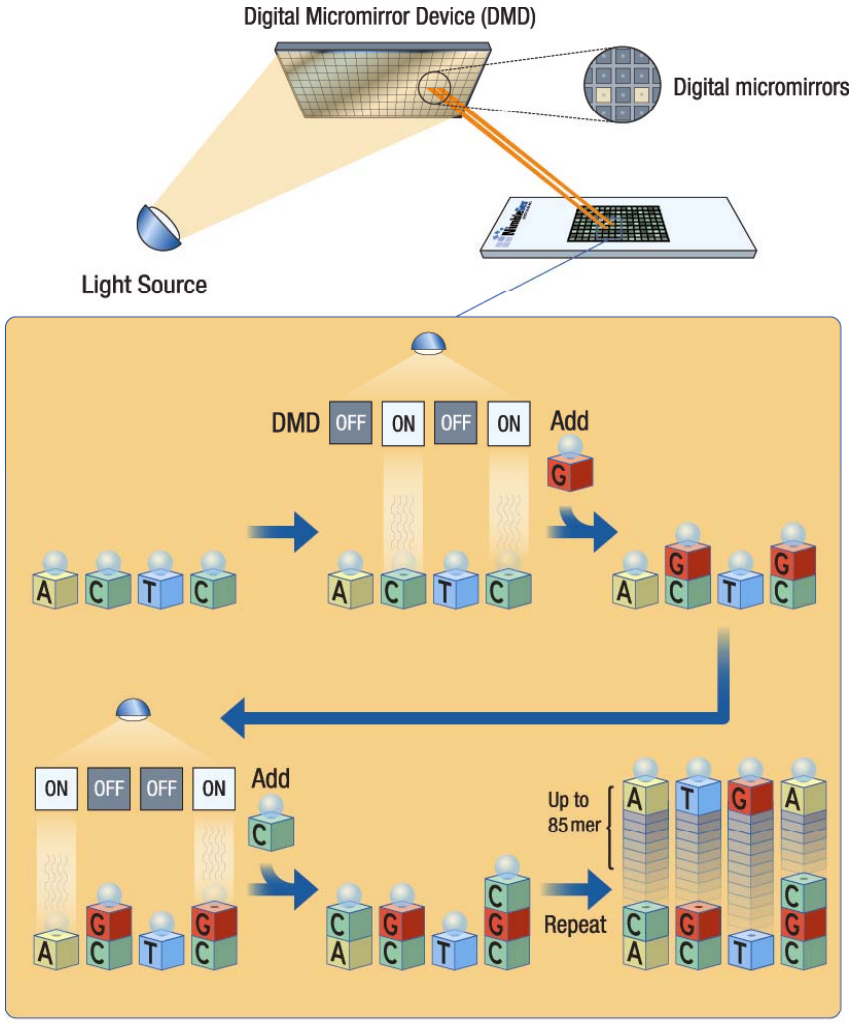
Customized Scale Inkjet Printer

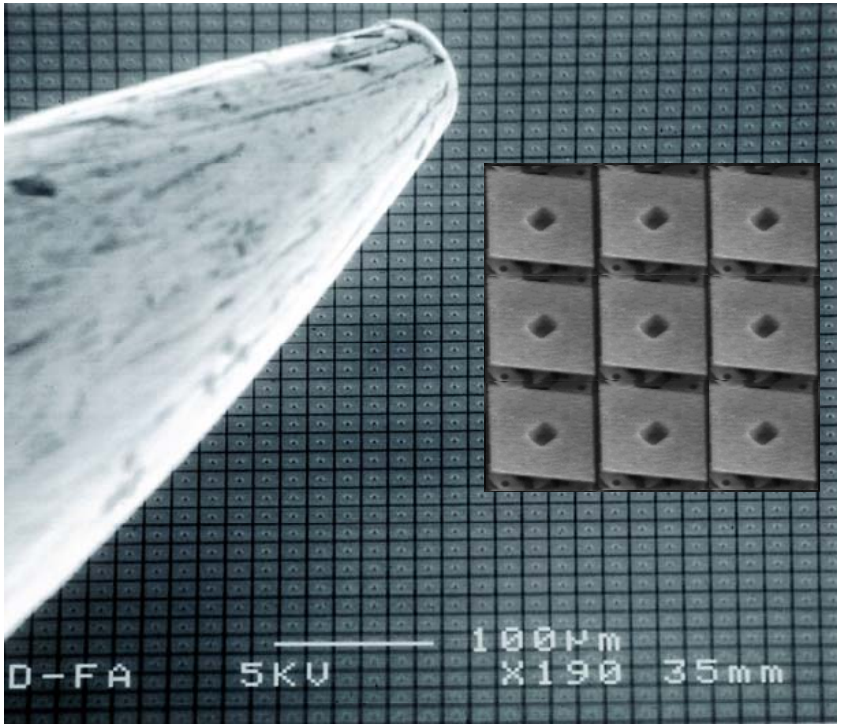


Industrial Scale Inkjet Printer



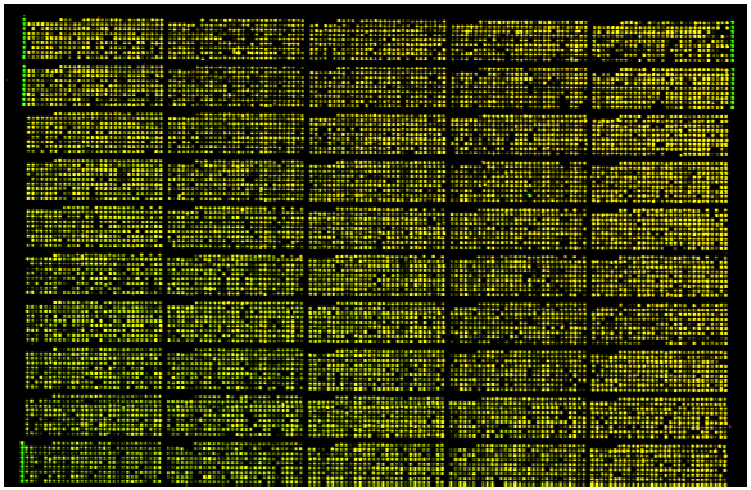
Digital Oligo DNA Microarray Synthesis





Array Quality Control





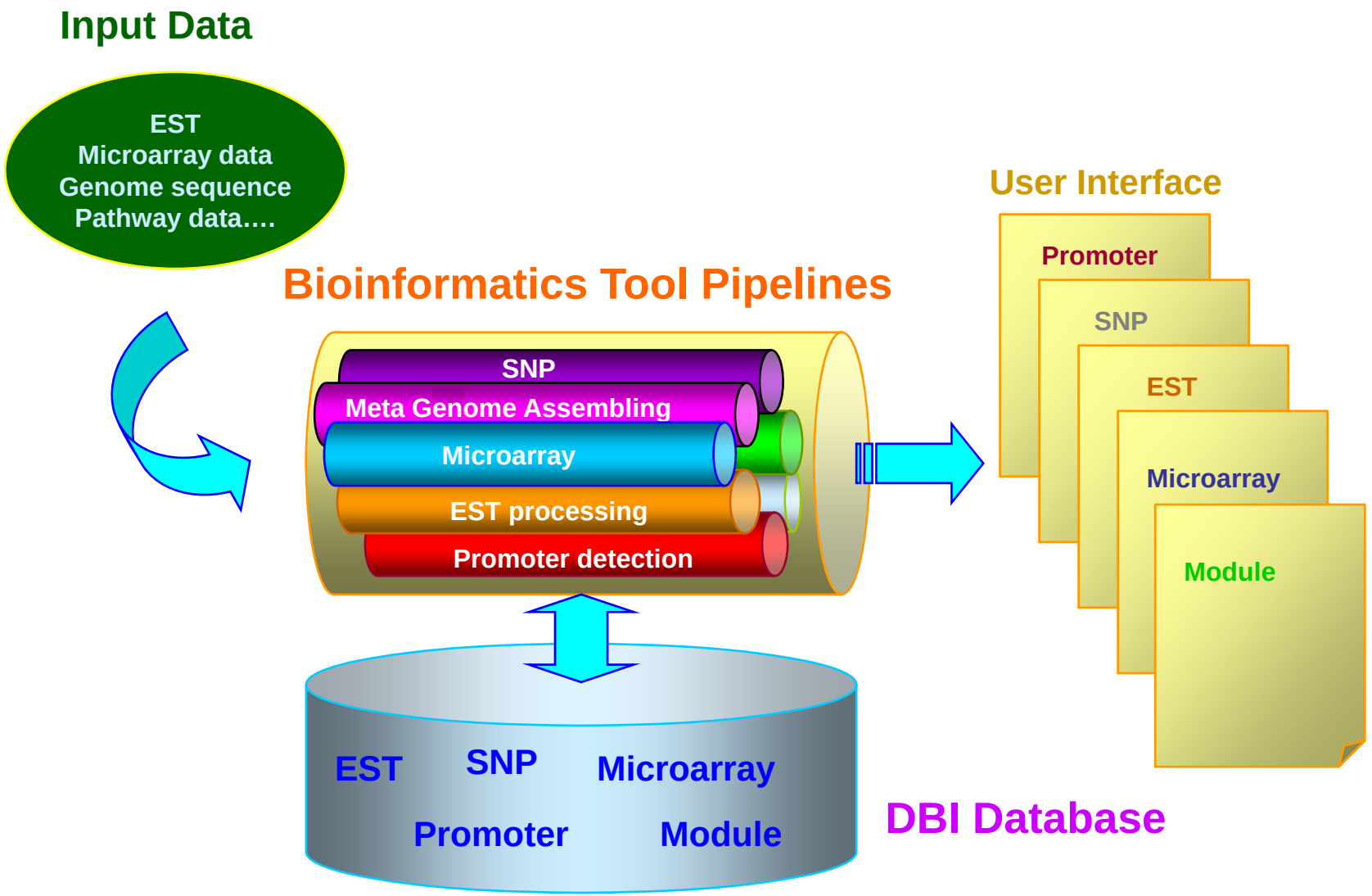
Process	QC Test
Glass	Test 1. Flatness / thickness 2. Low fluorescence
Substrate	Batch Test 1. Surface Integrity 2. Attachment Chemistry
Array	Writer QC 1. Vision System 2. Sensitivity 3. Specificity 4. Microarray Uniformity
Quarantine	Array Lot QC 1. Synthesis 2. Functional Test 3. Defects
Package	


thanks God,
i make it.

Bioinformatics Issues in Microarray Experiment

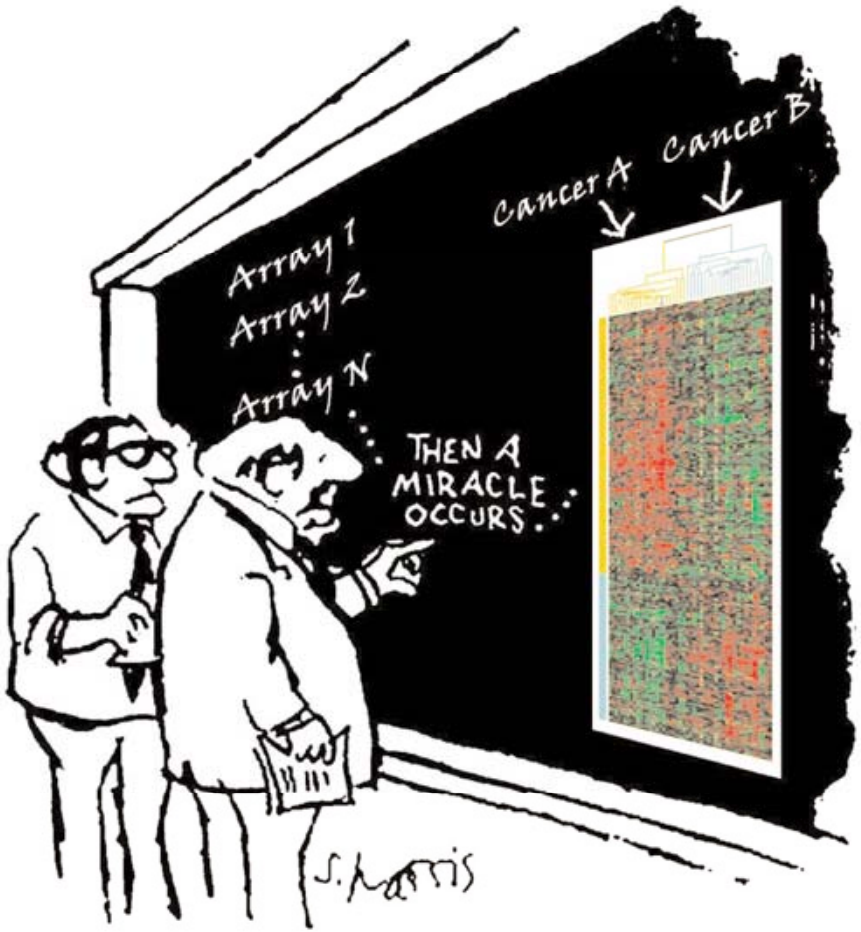
1. How to manage data ?
2. Which data should be stored ?
3. Validated database or just data storage ?
4. How to integrate in-house data with public repositories ?
5. Which databases ?
6. Local mirror DBs versus web link ?
7. How to analyze and visualize data ?
8. Which analysis should be standardized ?
9. Buy versus develop ?
10. How to integrated microarray data with other omics data ?

Bioinformatics Tool and Data Network




$16x^2+40xy+25y^2=(4x+5y)^2$
$16x^2+20xy+25y^2=(4x+5y)^2-20xy$
Microarray Algorithm

The miracle of microarray data analysis



"I think you should be more explicit here in step two."

Problems:

- Experimental design
- Normalization
- QC
- Ratios
- Appropriate controls
- Data interpretation in a biological context
- Data communication

Microarray – Transcriptomics Solution

Microarray Content → Sample Preparation → Microarray Processing → Informatics → Gene Expression Results

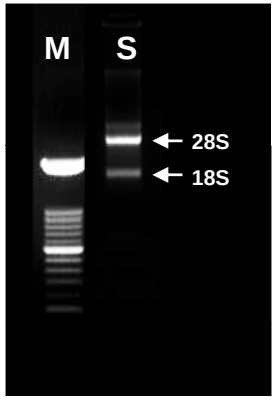
AFFYMETRIX

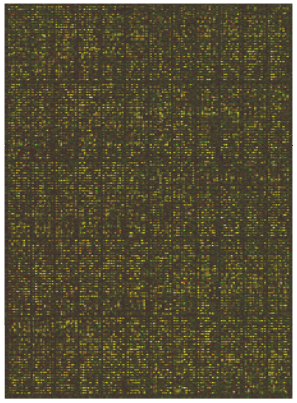
Agilent Technologies

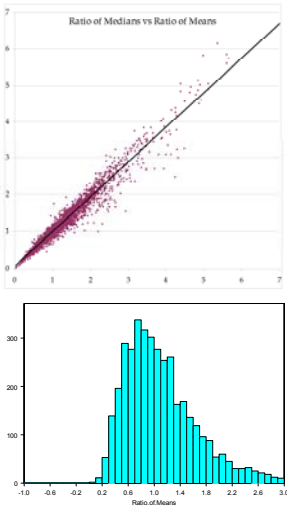
Amersham Biosciences

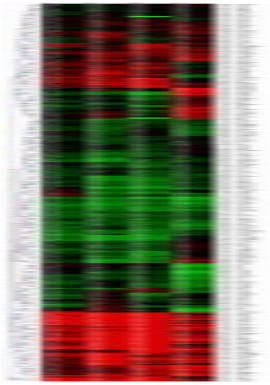


Custom array



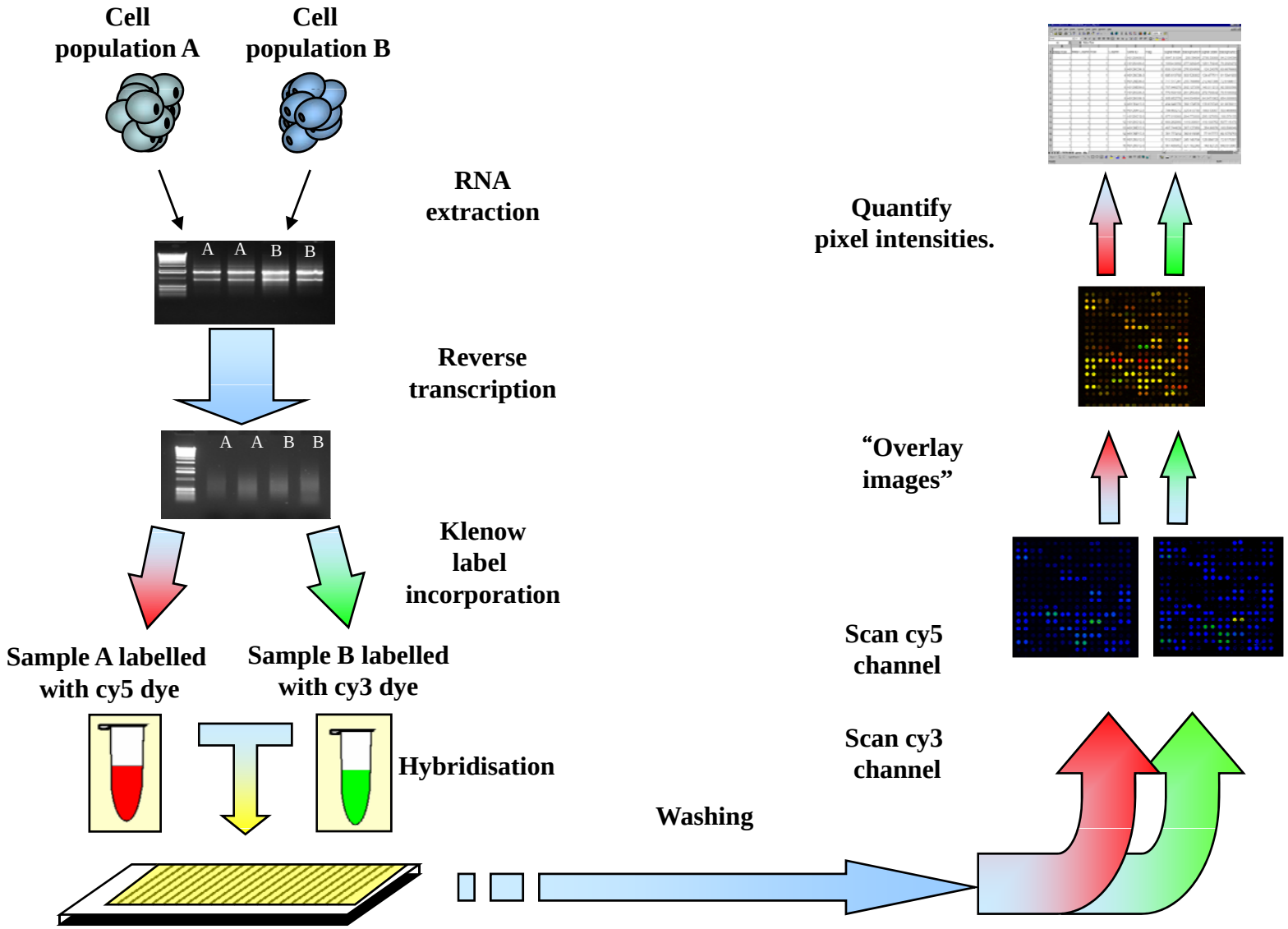


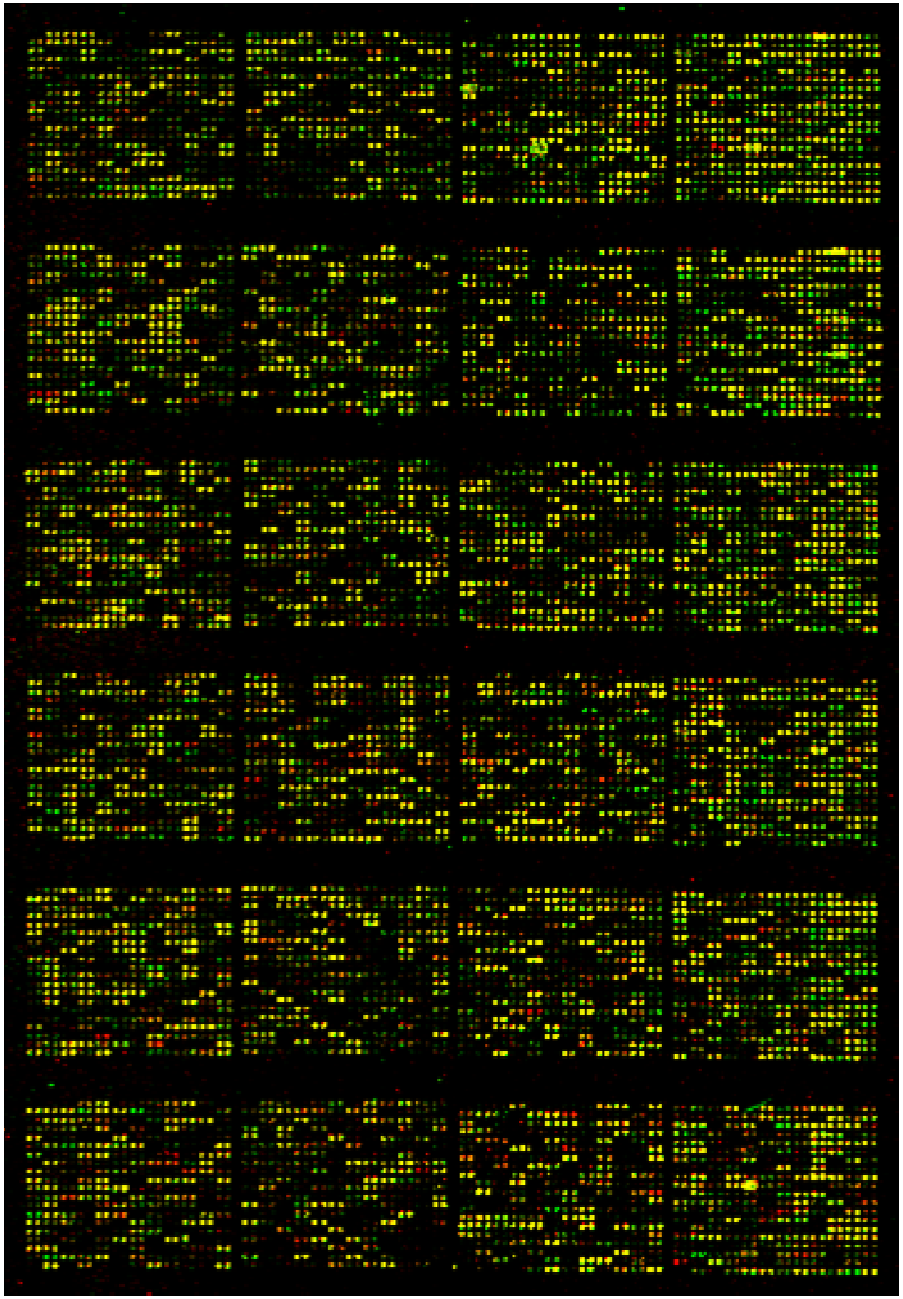


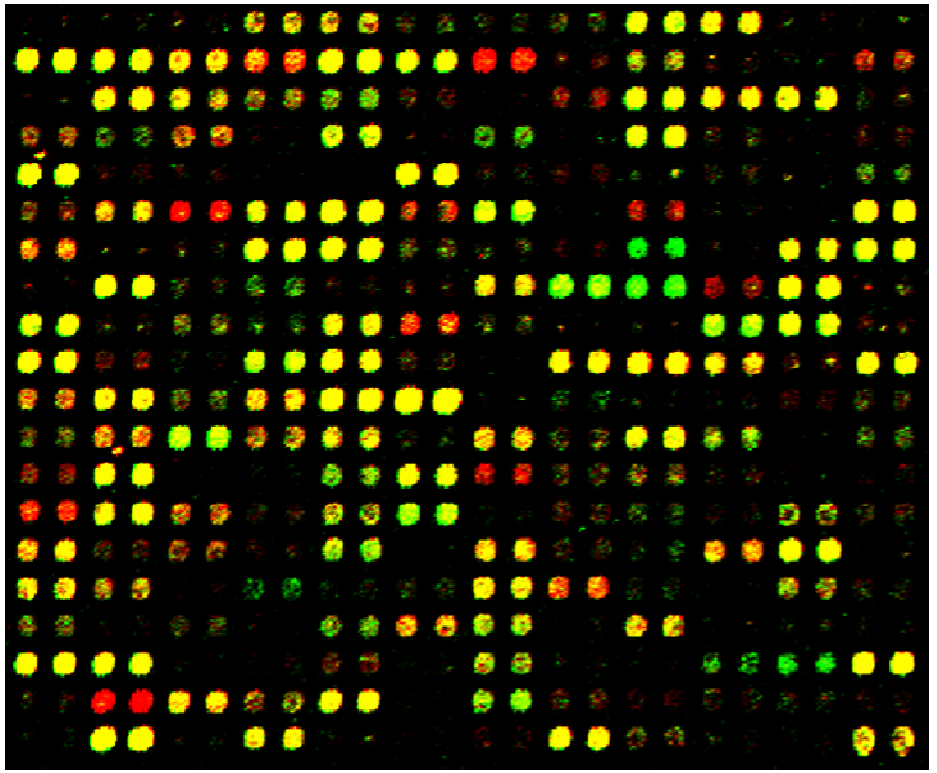


Modified from Agilent Technologies

Expression Microarray Experimental overview







Red – increase of Cy5 sample transcripts

Green – increase of Cy3 sample transcripts

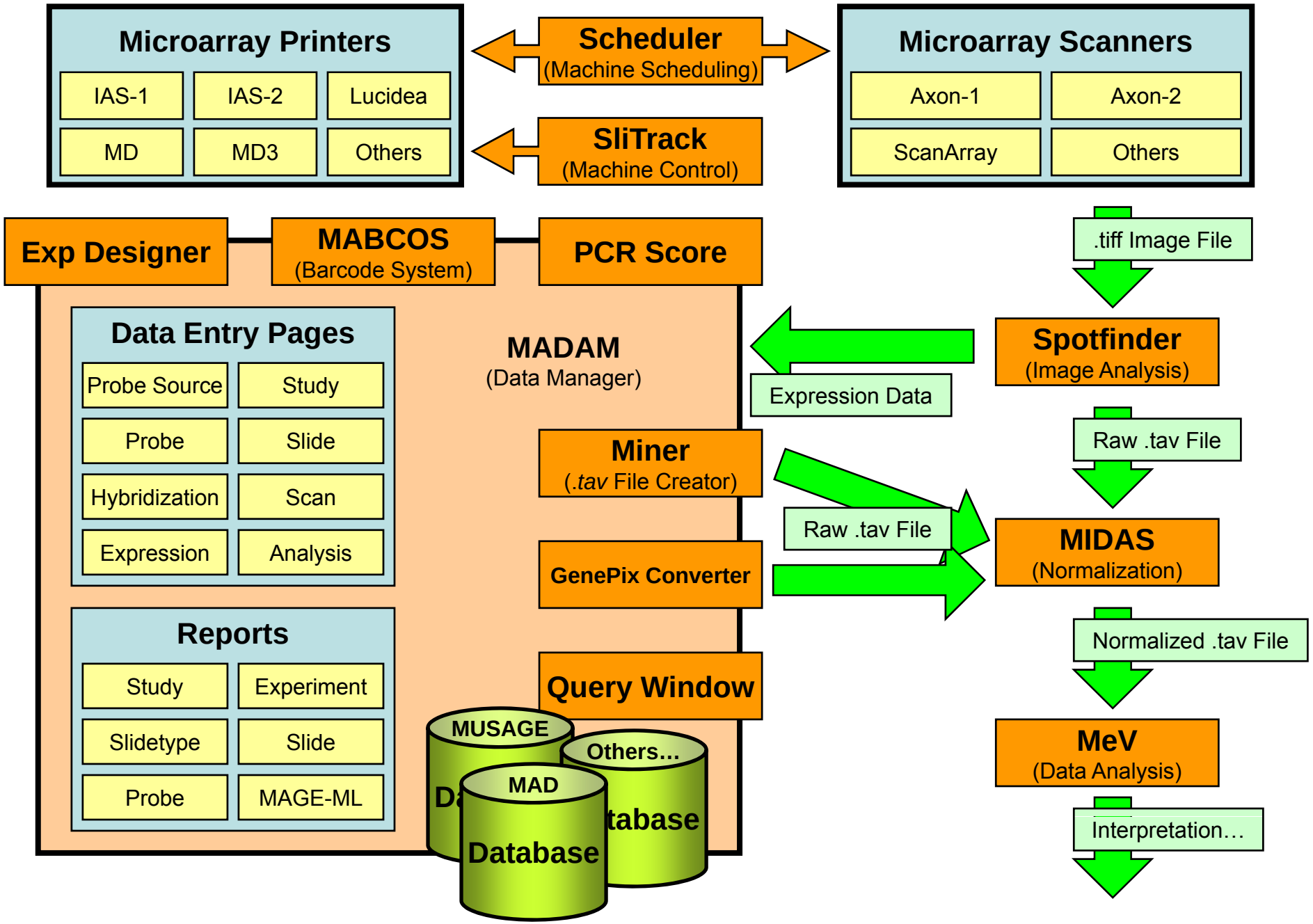
Yellow – equal abundance

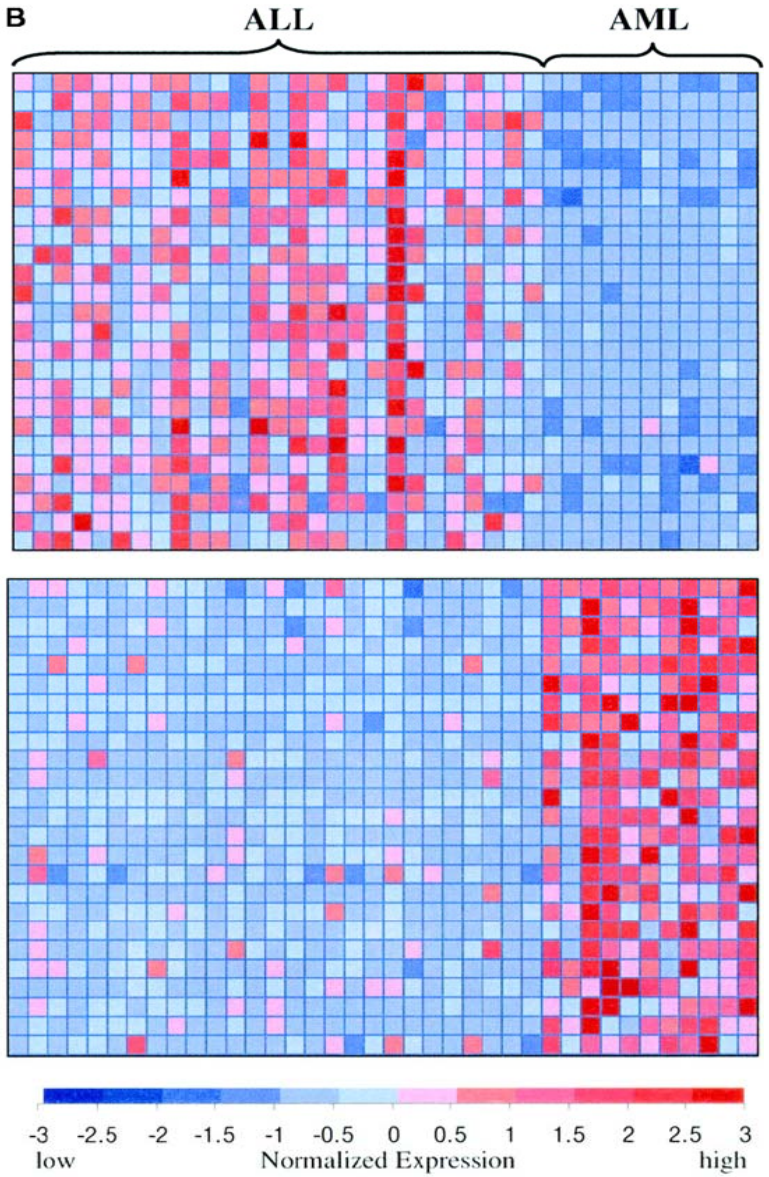
Limit of Detection: 1 in 30,000 transcripts

~ 20 transcripts/cell

Microarray Analysis Strategy

- **Microarray data processing**
  - Normalization
  - Visualization
  - QA/QC
- **Data analysis**
  - Contrast comparison
  - Series experiment
- **Database**
  - Storage and maintenance
  - Data integration and transform
  - User friendly interface
- **Data mining and function prediction**
  - Gene annotation
  - Gene function prediction

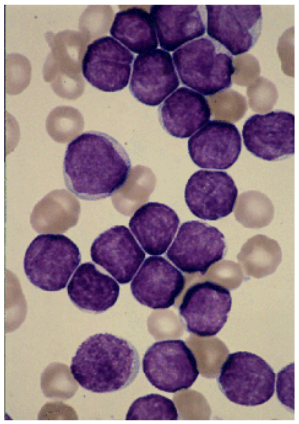

Microarray Printers
IAS-1
IAS-2
Lucidea
MD
MD3
Others
Scheduler
(Machine Scheduling)
SliTrack
(Machine Control)
Microarray Scanners
Axon-1
Axon-2
ScanArray
Others
Exp Designer
MABCOS
(Barcode System)
PCR Score
Data Entry Pages
Probe Source
Study
Probe
Slide
Hybridization
Scan
Expression
Analysis
MADAM
(Data Manager)
Reports
Study
Experiment
Slidetype
Slide
Probe
MAGE-ML
Miner
(.tav File Creator)
GenePix Converter
Query Window
MUSAGE
Others…
MAD
Database
.tiff Image File
Spotfinder
(Image Analysis)
Expression Data
Raw .tav File
Raw .tav File
MIDAS
(Normalization)
Normalized .tav File
MeV
(Data Analysis)
Interpretation…



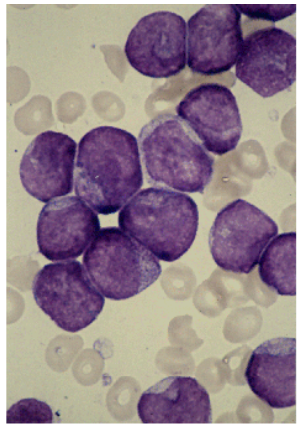
Classical study of cancer subtypes

Golub et al. (1999)

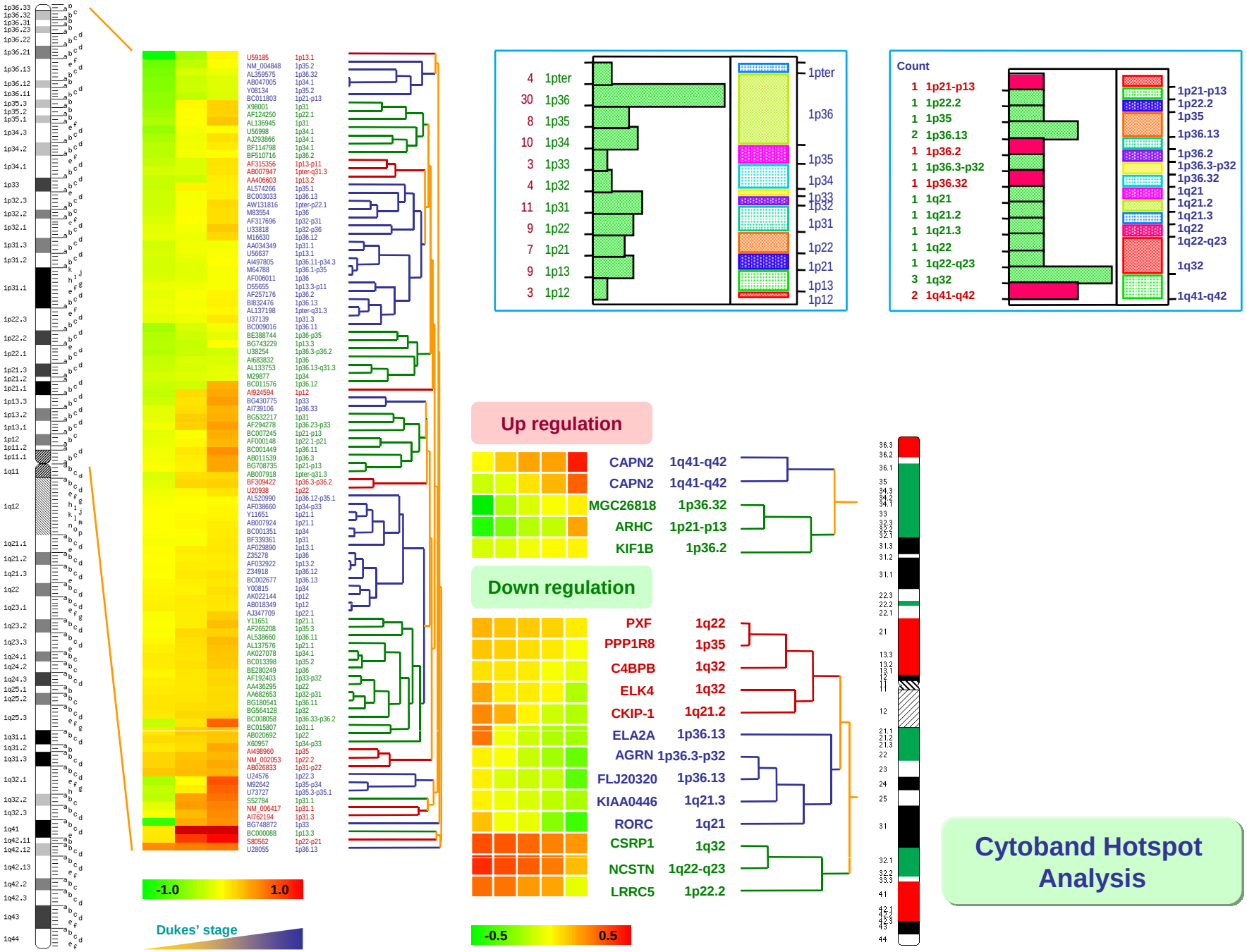
identification of diagnostic genes



ALL
acute lymphoblastic leukemia
(lymphoid precursors)



AML
acute myeloid leukemia
(myeloid precursor)


-1.0
1.0
Dukes' stage
4 1pter
30 1p36
8 1p35
10 1p34
3 1p33
4 1p32
11 1p31
9 1p22
7 1p21
9 1p13
3 1p12
Count
1 1p21-p13
1 1p22.2
1 1p35
2 1p36.13
1 1p36.2
1 1p36.3-p32
1 1p36.32
1 1q21
1 1q21.2
1 1q21.3
1 1q22
1 1q22-q23
3 1q32
2 1q41-q42
Up regulation
CAPN2 1q41-q42
CAPN2 1q41-q42
MGC26818 1p36.32
ARHC 1p21-p13
KIF1B 1p36.2
Down regulation
PXF 1q22
PPP1R8 1p35
C4BPB 1q32
ELK4 1q32
CKIP-1 1q21.2
ELA2A 1p36.13
AGRN 1p36.3-p32
FLJ20320 1p36.13
KIAA0446 1q21.3
RORC 1q21
CSRP1 1q32
NCSTN 1q22-q23
LRRC5 1p22.2
-0.5
0.5
Cytoband Hotspot Analysis

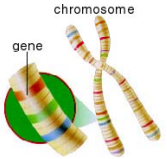
Data Analysis Strategy

3 Pathway



http://cgap.nci.nih.gov/Pathways

2 Cytoband

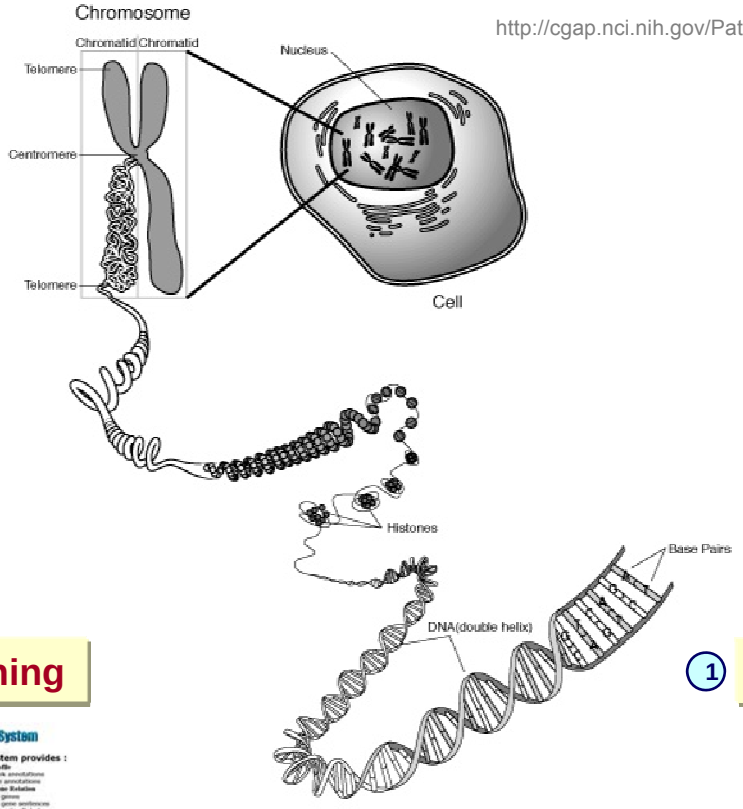


4 GO term



Molecular function
Biological process
Cellular component

http://www.ebi.ac.uk/GO/index.html



5 Text mining



http://140.116.247.48:8080/IGIS/index.jsp

1 Expression



Up regulation
Down regulation
Yoyo expression

http://biosys.csie.ncku.edu.tw/genefilter2/index.jsp

FIGURE SOURCE: http://www.accessexcellence.org/AB/GG/chromosome.html

Comparative Genomic Hybridization

Map Chromosomal DNA Changes

- amplifications
- deletions
- unbalanced translocations
- breakpoint mapping

Important for the Study of:

Cancer

(tumor suppressor genes, oncogenes)

Developmental disorders

(trisomies, sex chromosomes, microdeletions)

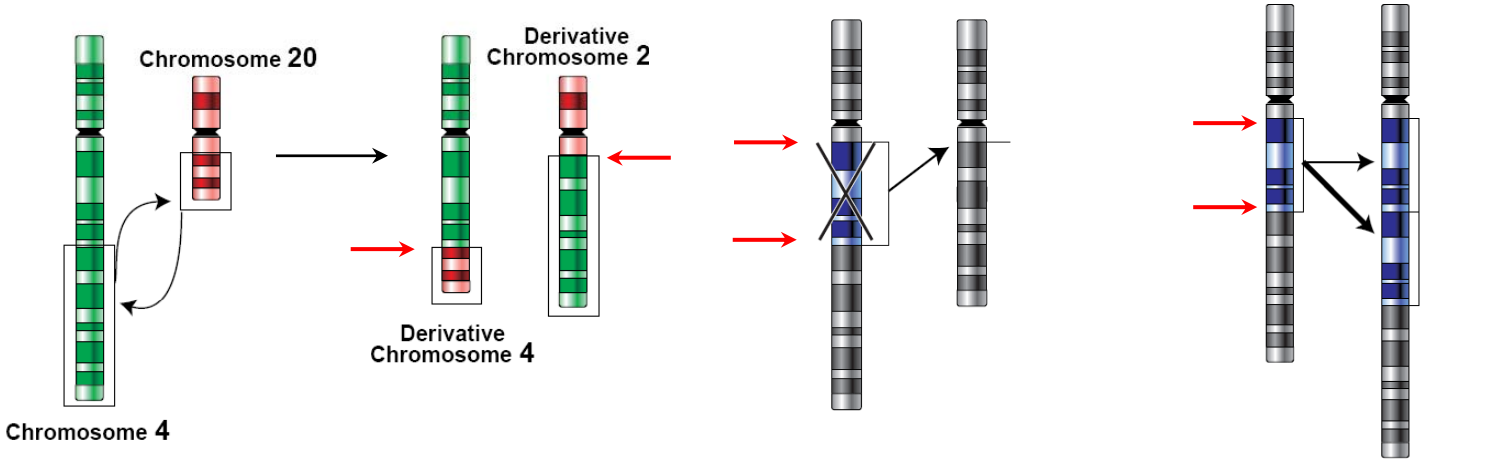
Role of CNPs in complex diseases

(psychiatric disorders, diabetes, cardiovascular disease…)

Array CGH Applications

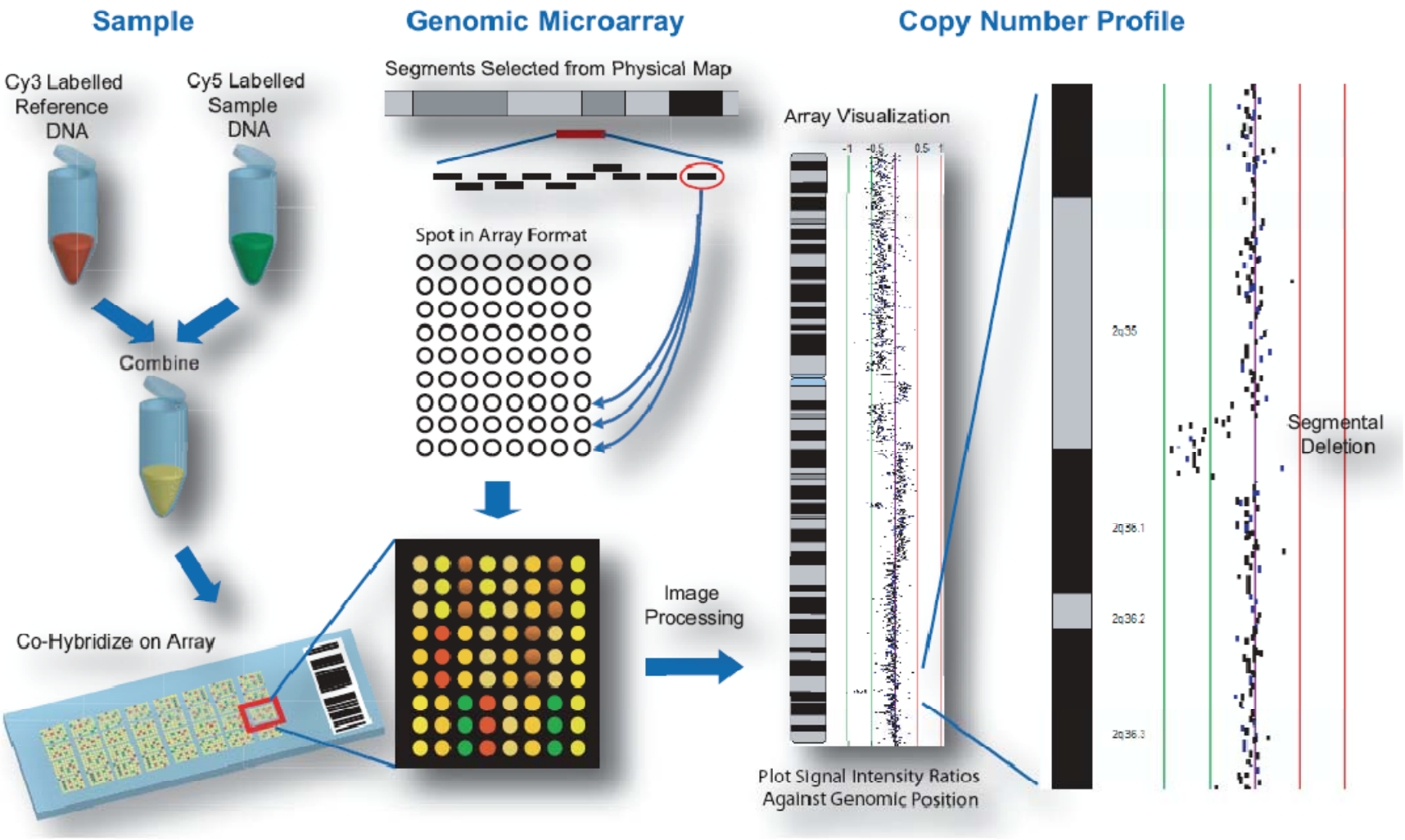
Detect chromosomal gains and losses
(e.g. tumor vs. normal DNA)

Map associated breakpoints

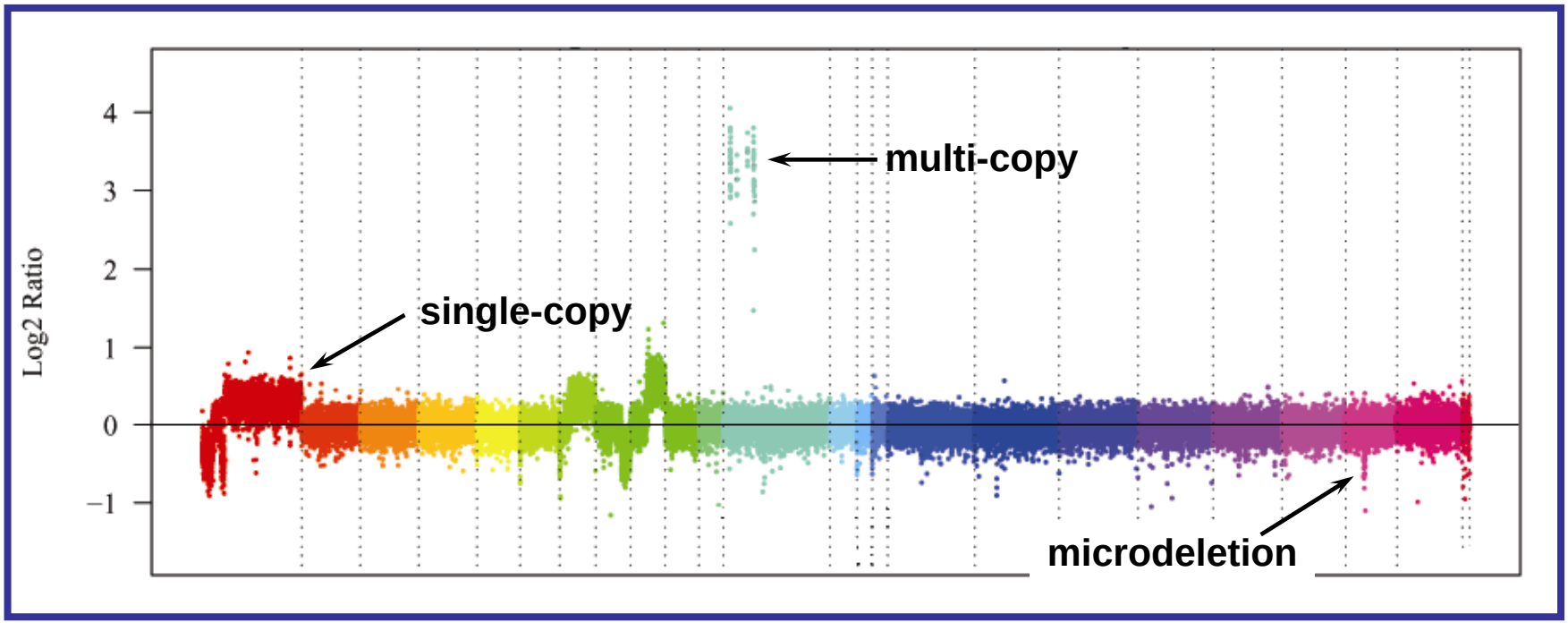


Translocations **Deletions** **Amplifications**

Array CGH Protocol

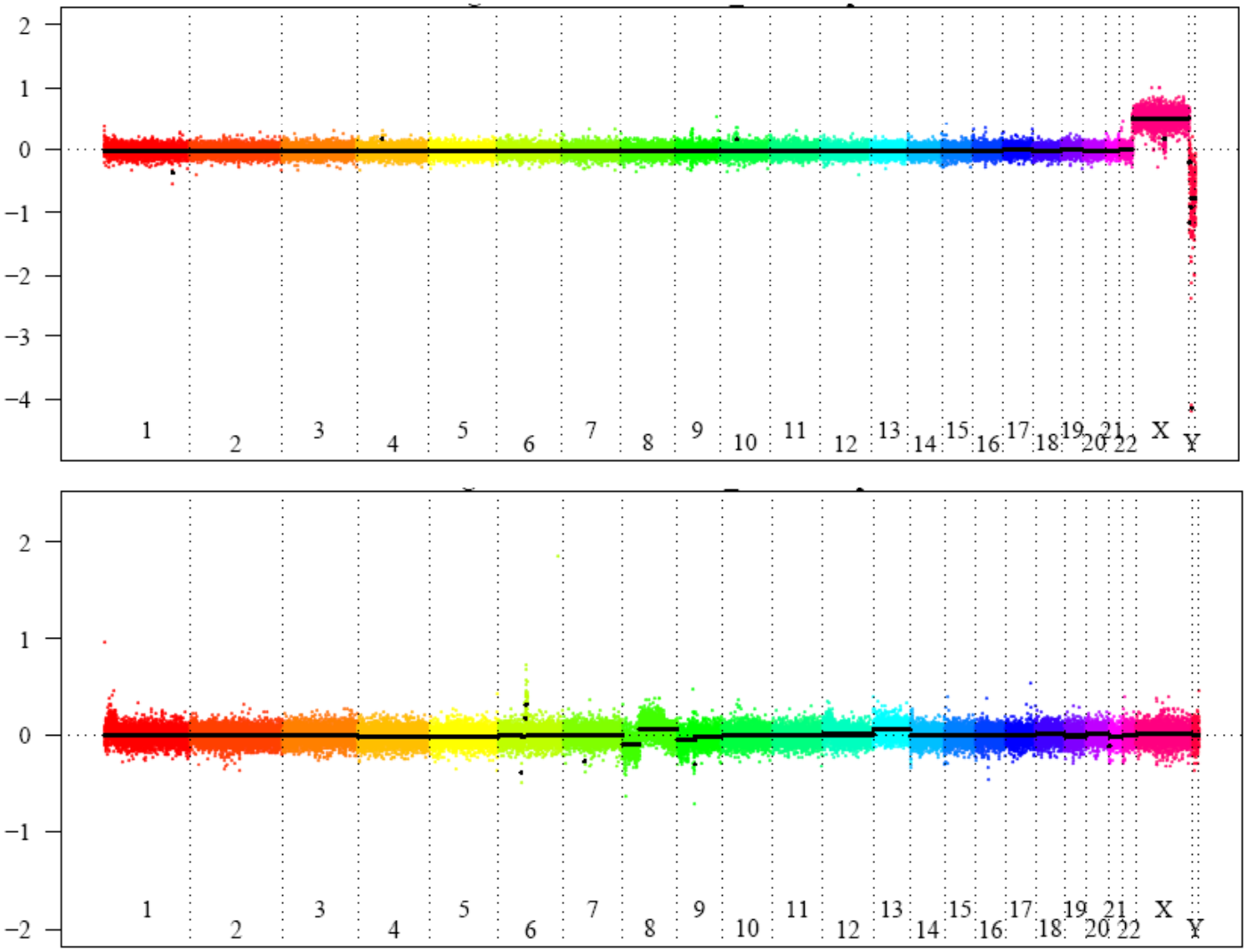


Human whole genome overview



- No complexity reduction of sample
- ~50kb “resolution” for CNP detection
- Unbiased genome coverage
- Single-copy amplifications detected

Chromosomes 1-22, X, Y



Female

Male Ref

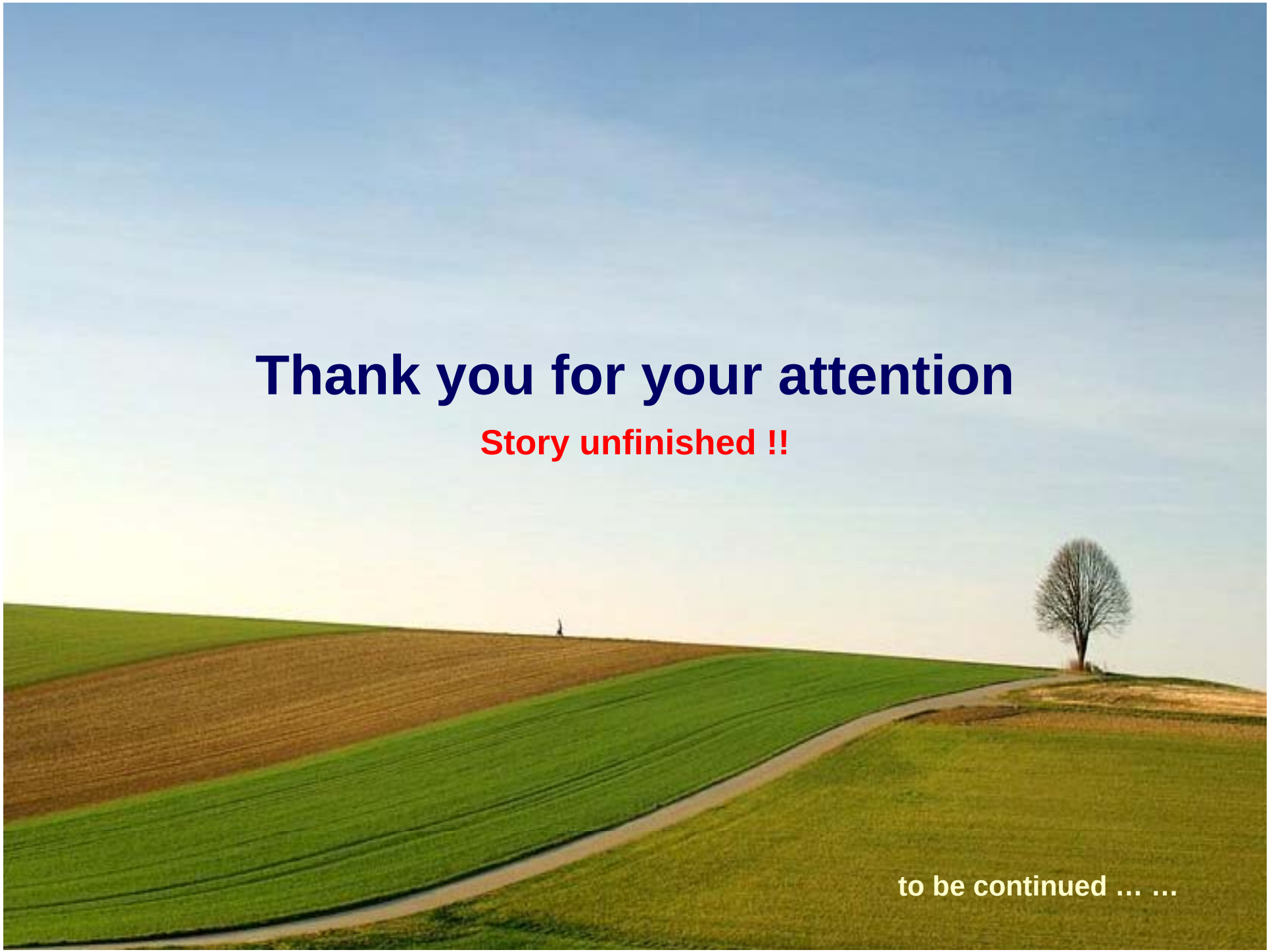
Burkitt lymphoma

Male

Male Ref


No more,
please


no,
you are up regulated


Thank you for your attention
Story unfinished !!
to be continued … …