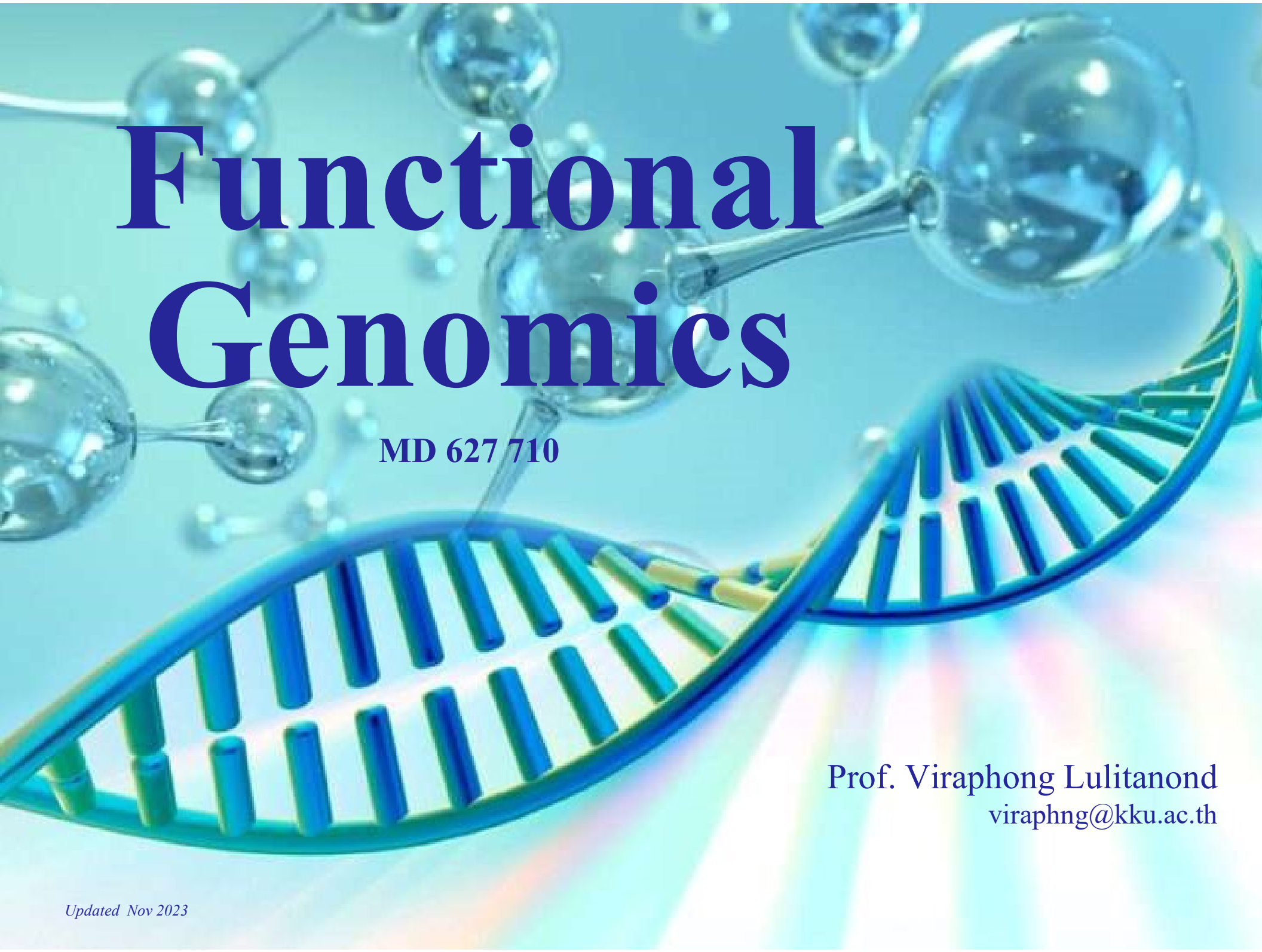


Functional Genomics

MD 627 710

Prof. Viraphong Lulitanond
viraphng@kku.ac.th

Updated Nov 2023

33 YEARS OF BIG BIOLOGY

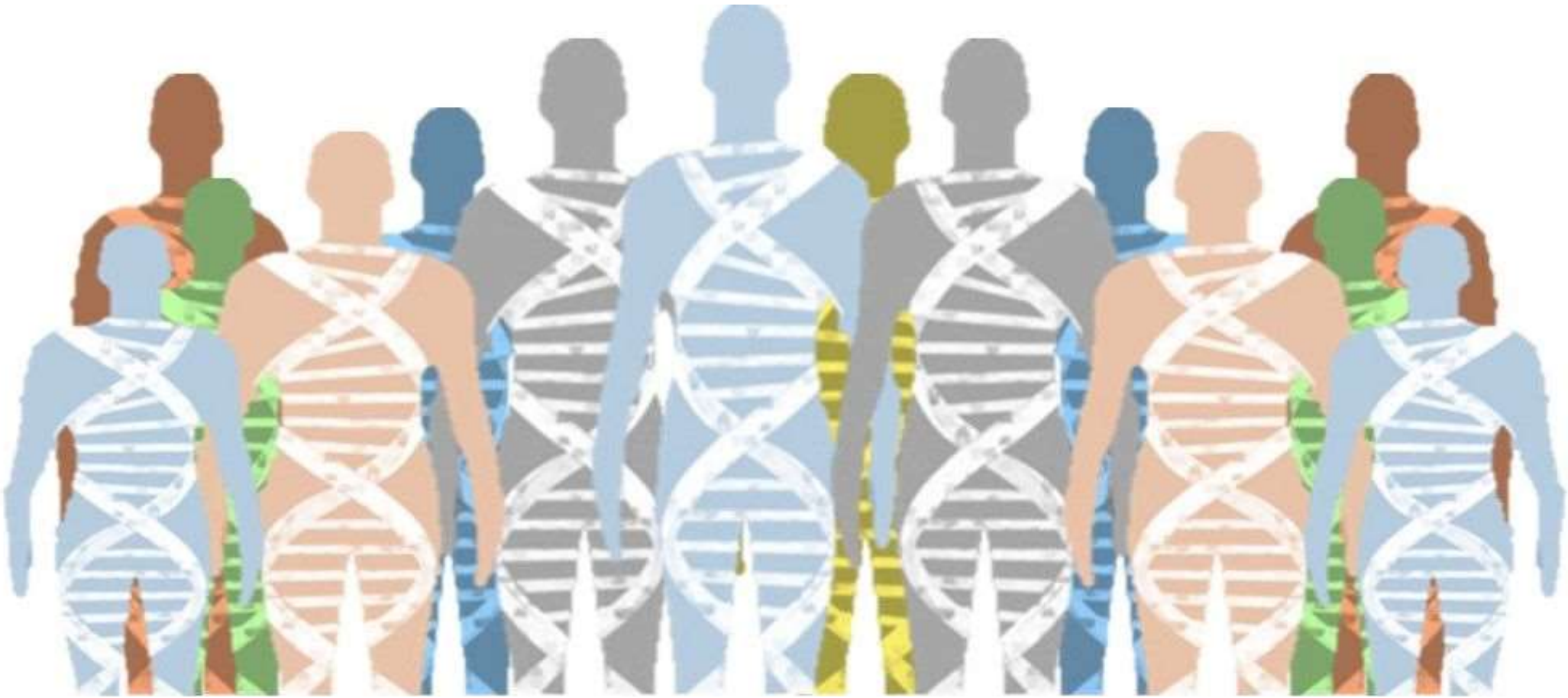
Nature 2015:526;29–31.

HGP
was the end of one journey,

But
the beginning of another!

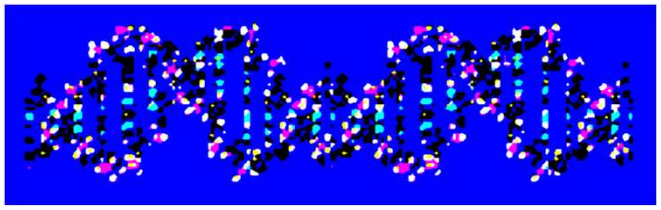
Post-genomics

A broad name given to activities made possible by **genomic information**



Omics Era

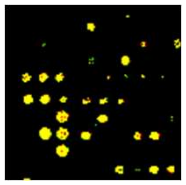
DNA



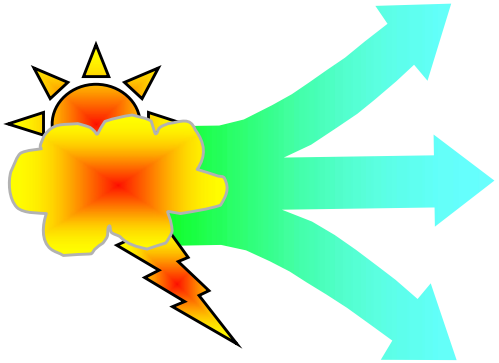
(storage)

Transcription

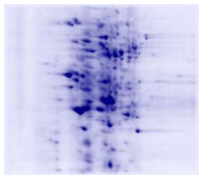
Translation



Gene Expression
(Transcriptome)

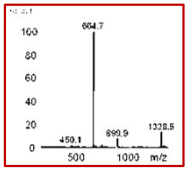


Proteins



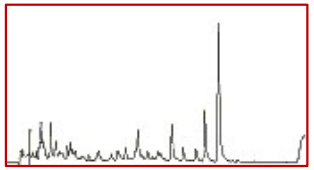
Proteomics

Biochemical Circuitry

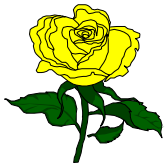


Metabolomics

Environment
Drug
Disease



Phenotypes (Traits)



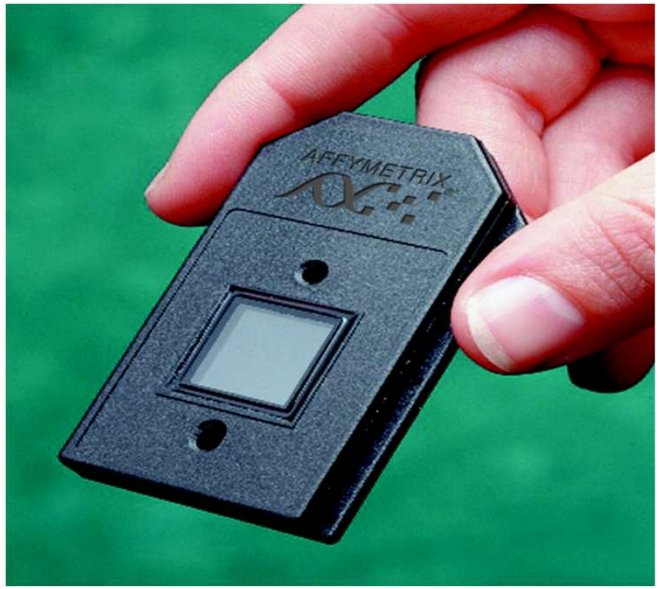
- **Functional genomics**: understanding the biological function of genes and other parts of the genome, e.g.
  - Study of all genes expressed by a specific cell type
  - Study of changes in gene expression pattern
  - Assignment of biological function to genes
- Analysis of gene expression is the cornerstone of functional genomics

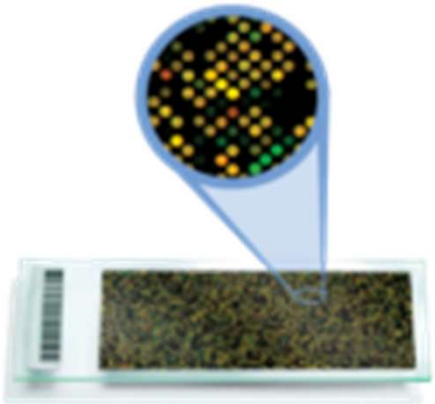
Functional Genomics Tools

- **Bioinformatics**
- **DNA microarrays**
- Differential display
- Substraction hybridization
- SAGE
- **Proteomic tools** e.g. 2-D gel electrophoresis, mass spectrophotometry, protein arrays etc.

DNA Microarray Technology
(Hybridization Array Technology)

High-throughput and large-scale technology to study all genes, their expression and/or their interaction simultaneously





What are DNA arrays or Gene Chips ?

- DNA arrays consist of large numbers of DNA molecules spotted in a systemic order on a solid substrate.
- Depending on the size of each DNA spot on the array, DNA arrays can be categorized as *microarrays* (dia <250 um) and *macroarrays* (dia >300 um).
- When the solid substrate used is small in size, arrays are also referred to as **DNA chips** or **Gene Chips**

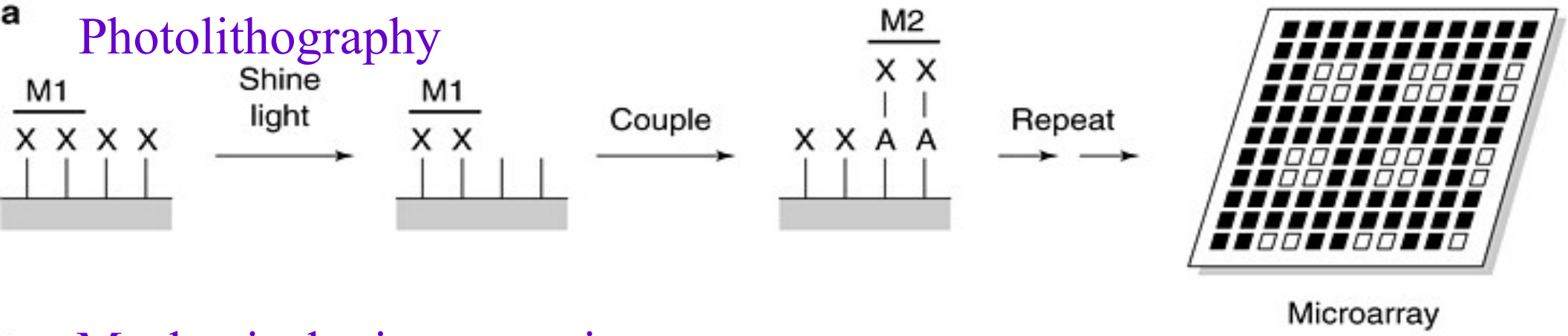
TYPES OF GENE CHIPS

A. Based on manufacturing

- Photolithography (synthesis on chip)
- Mechanical microspotting (delivery to chip)
- Ink jetting (delivery to chip)
- Pin and ring spotting (delivery to chip)

a

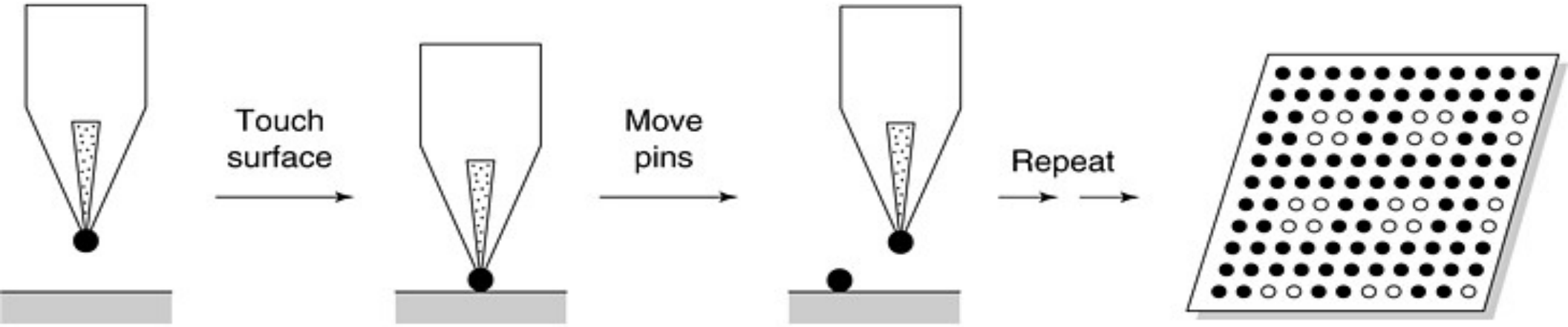
Photolithography



Microarray

b

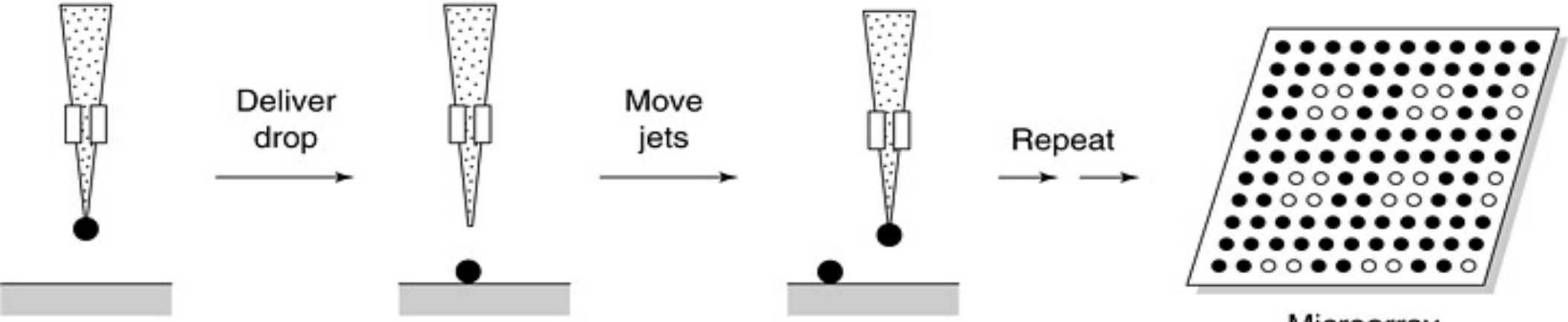
Mechanical miocrospotting



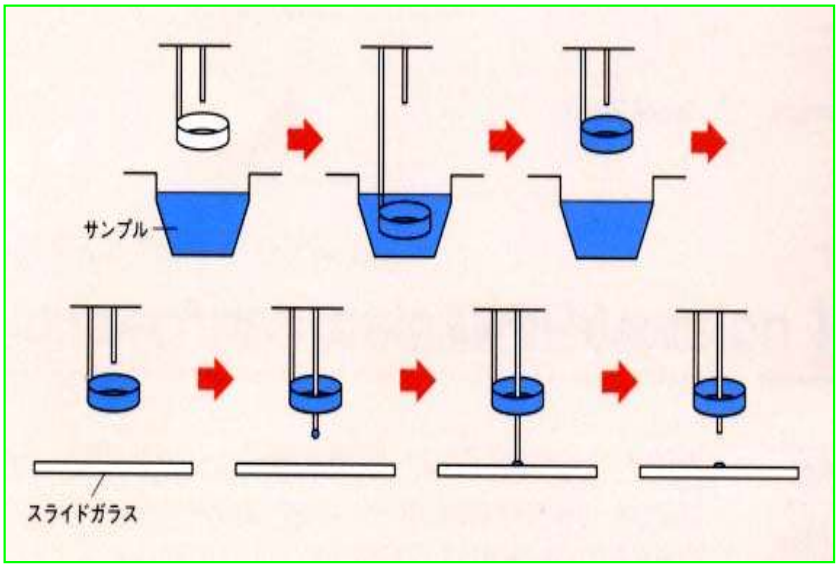
Microarray

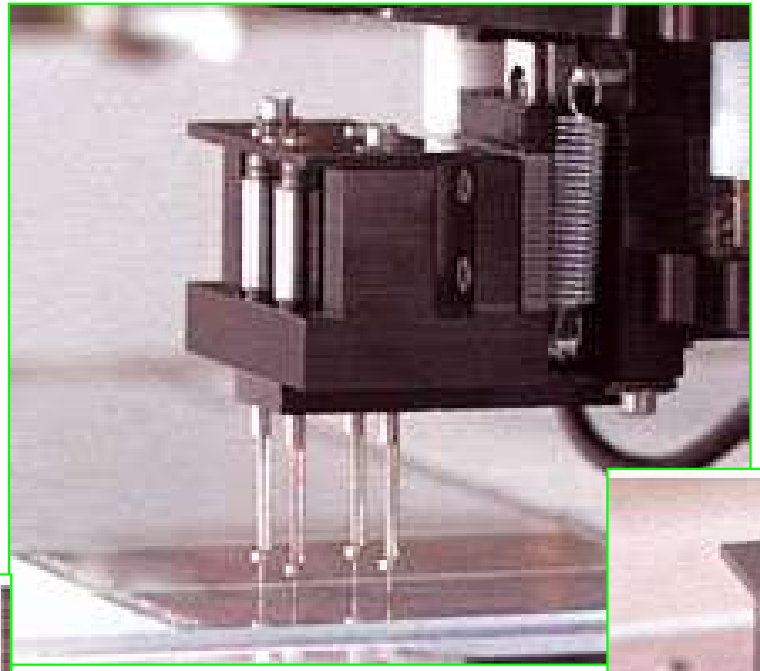
c

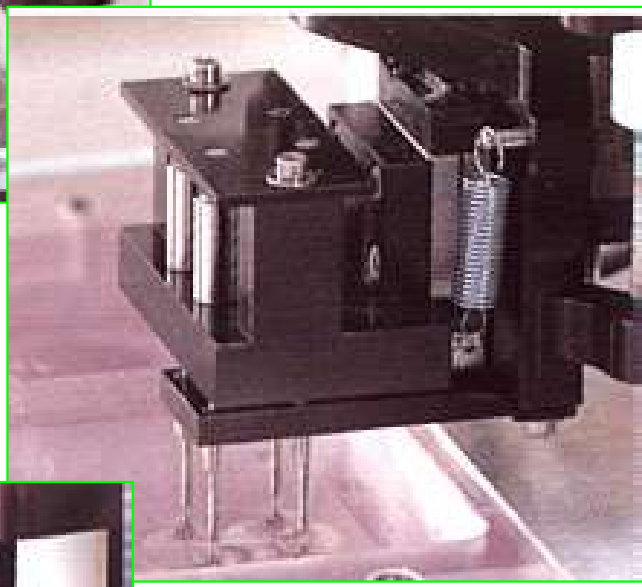
Ink Jetting

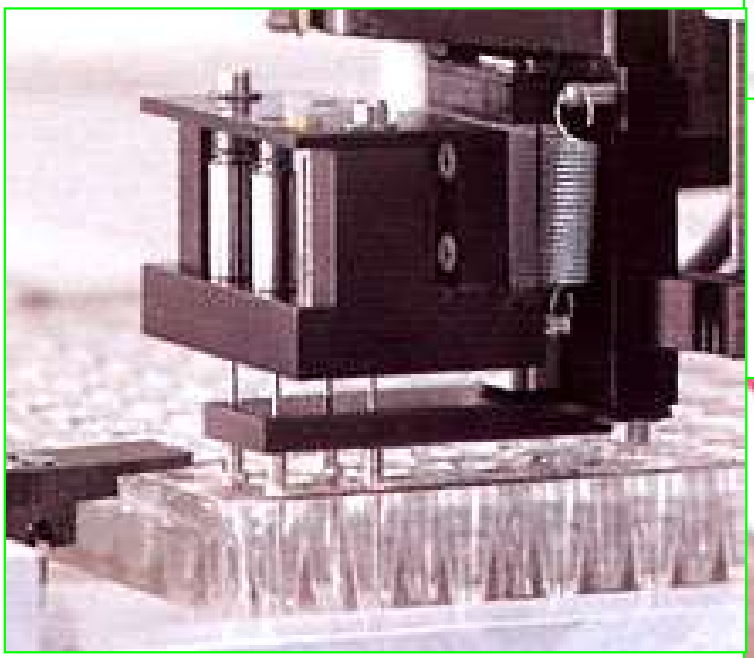


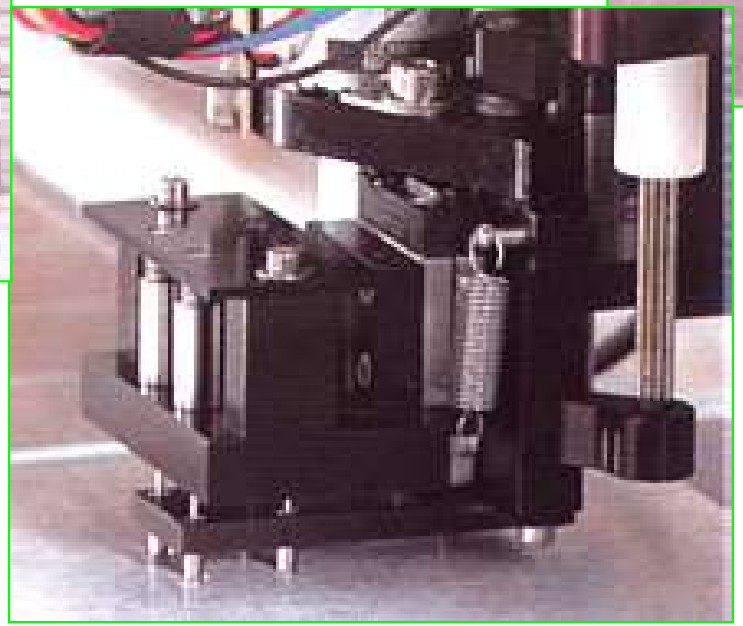
Microarray



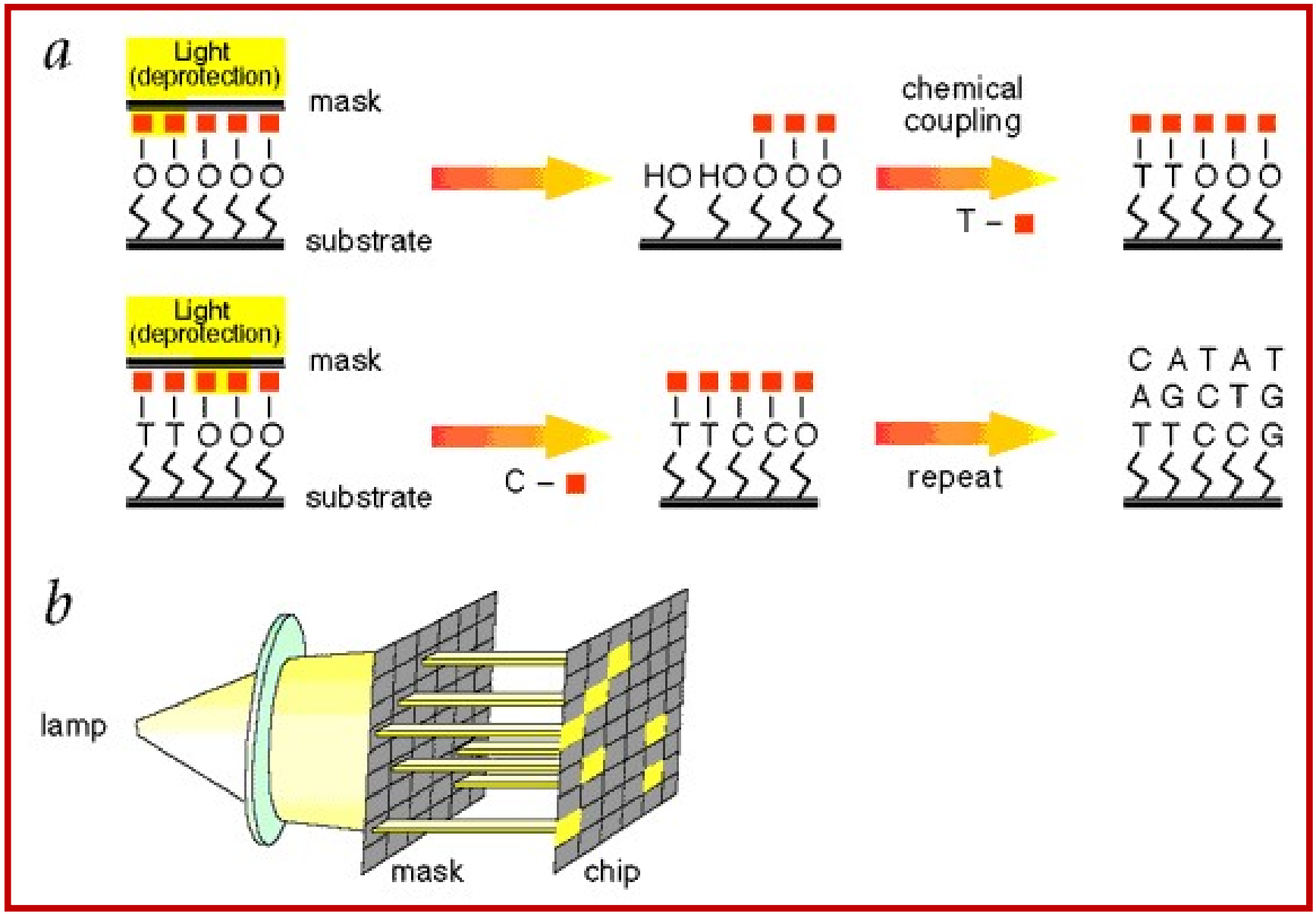


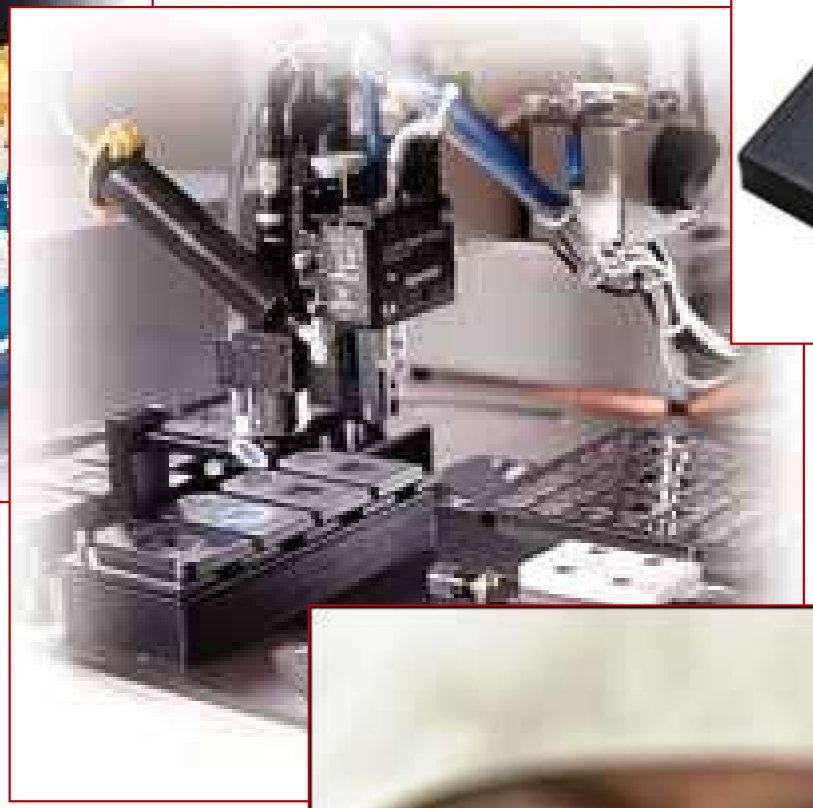




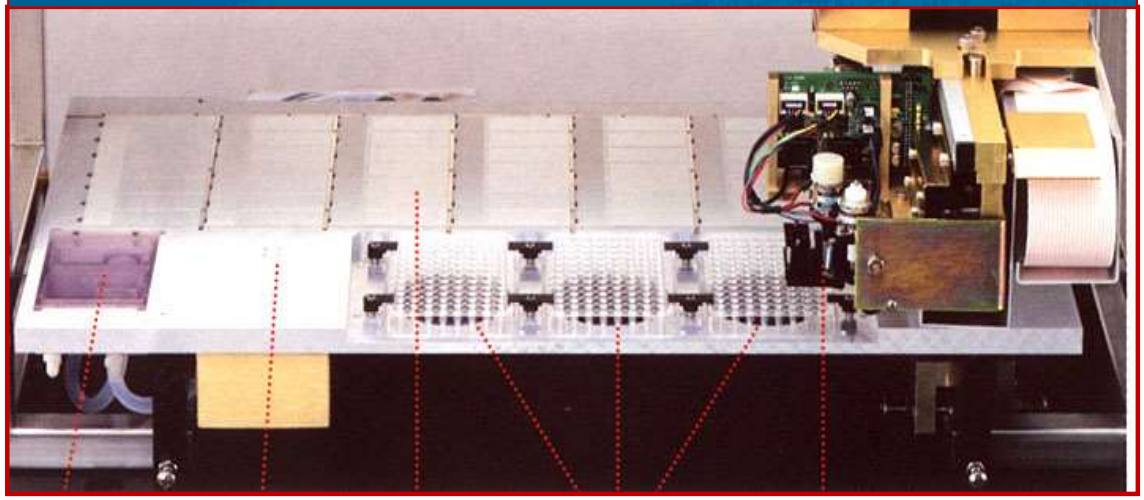


Pin and Ring Arrayer

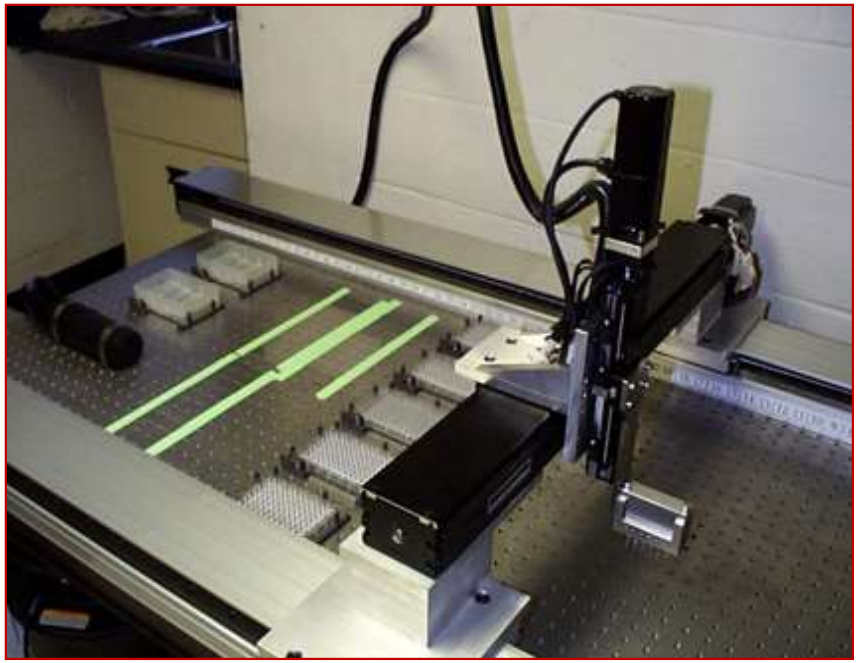

a
Light
(deprotection)
mask
substrate
chemical
coupling
T –
Light
(deprotection)
mask
substrate
C –
repeat
C A T A T
A G C T G
T T C C G
b
lamp
mask
chip

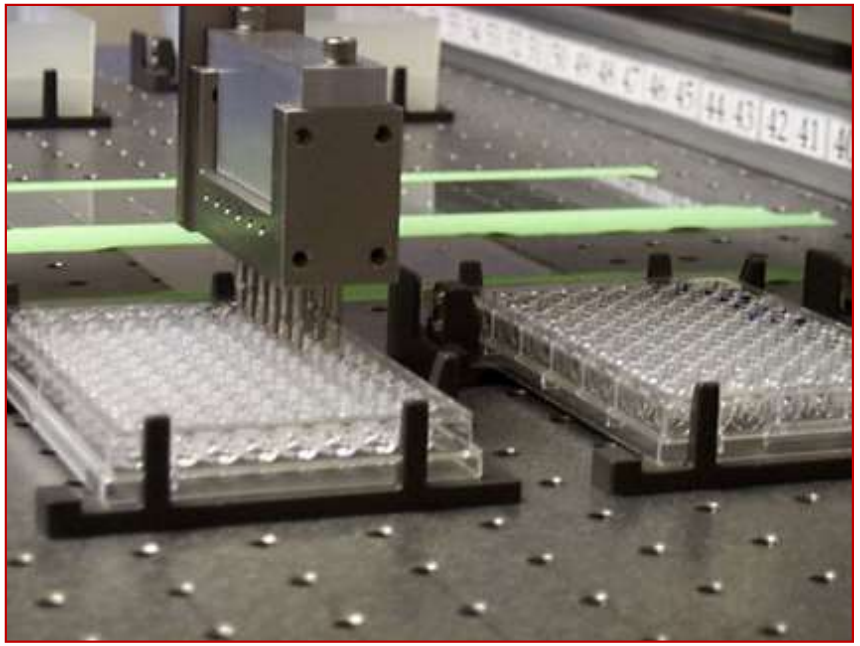


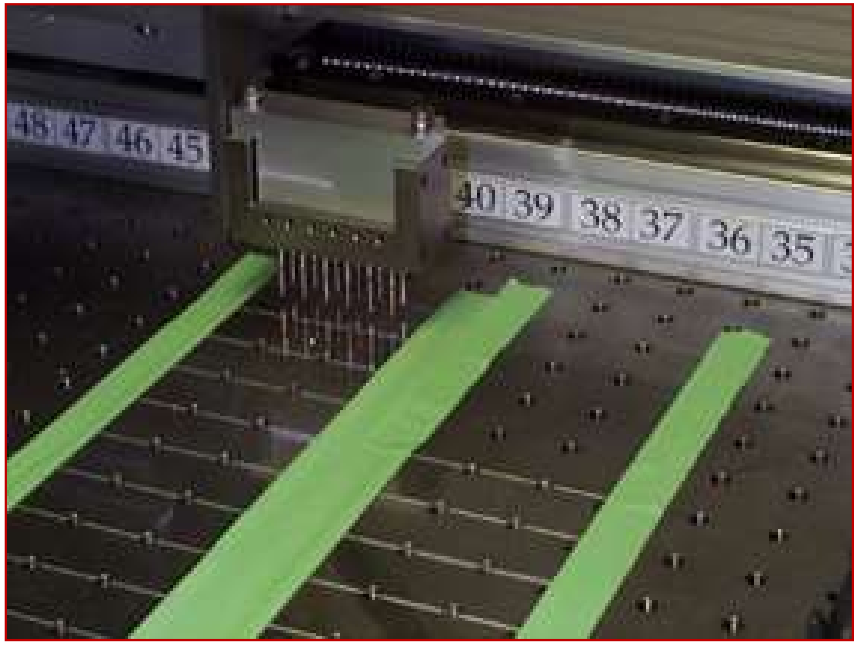
DNA microarray using photolithographic method

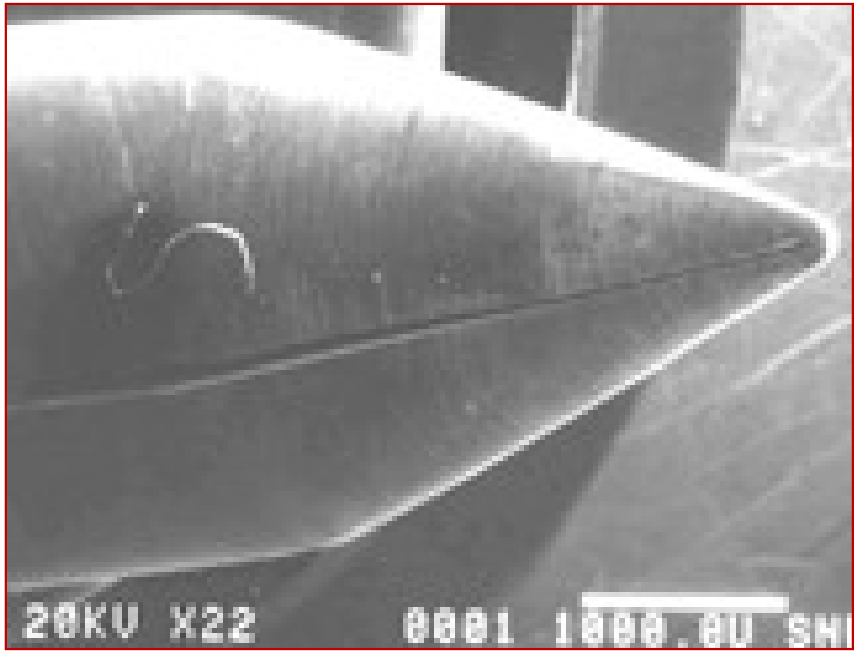


ARRAYER





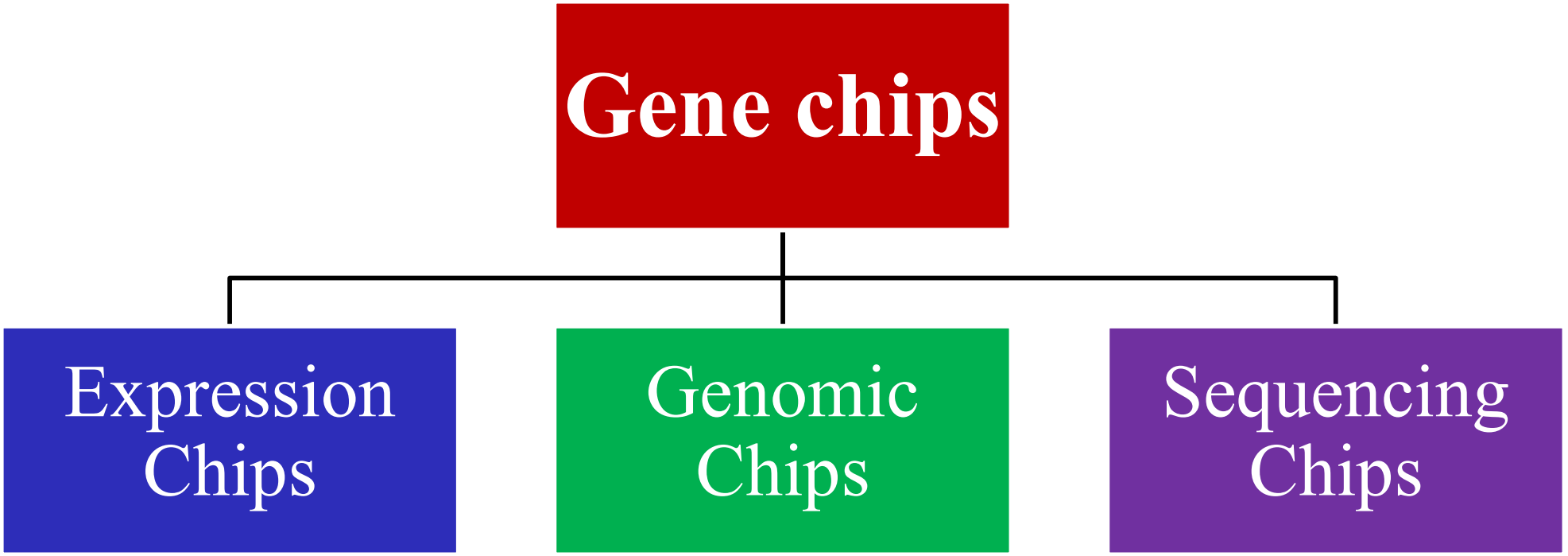

48 47 46 45
40 39 38 37 36 35


20KV X22
0001 1000.0U SM

Patrick O Brown

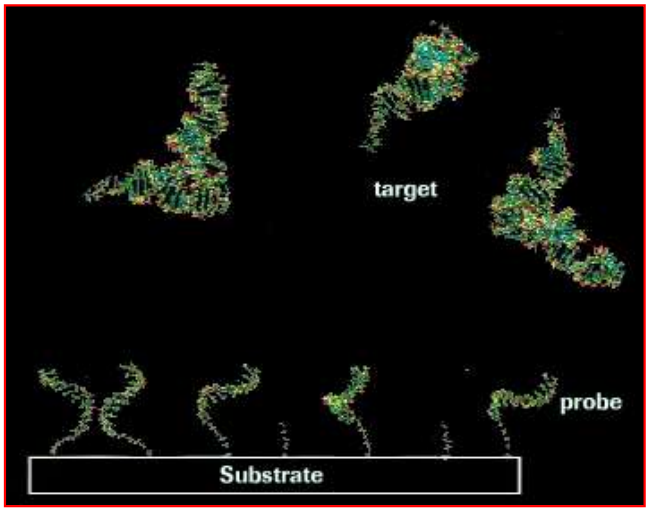
https://youtu.be/RdfXCopNseU

B. Based on function of microarrays



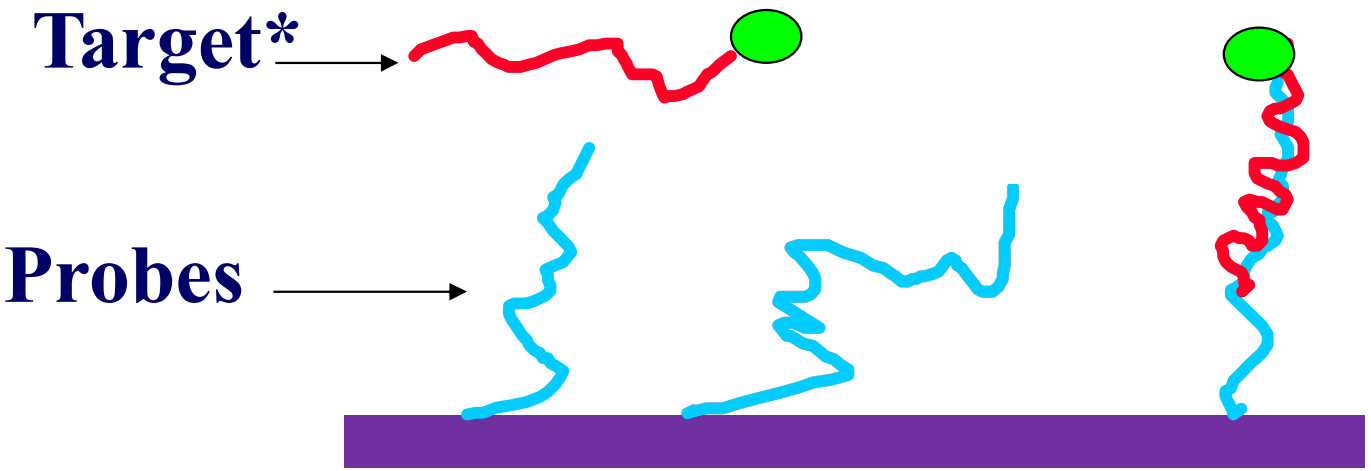
C. Based on types of probes

Oligo-Chip	**cDNA-Chip**	**Genomic Chip**
20 - 30 n	**< 2,000 n**	**> 50,000 n**
↓ ↓	↓	↓
Sequencing & Expression	Expression	Genomic analysis



Basic format of DNA microarrays

1. Spotting DNA of interest on solid substrate at defined positions
2. Sample isolation and labeling
3. Probe/target hybridization
4. Imaging of hybridization results and image analysis



Advantages of DNA microarrays

- Faster : Experiments can be performed at a faster rate
- Parallelism: Users can simultaneously screen hundreds to thousands of targets in a single experiment
- Automation: Analysis of gene chips can be automated through hardware/software
- Economic:

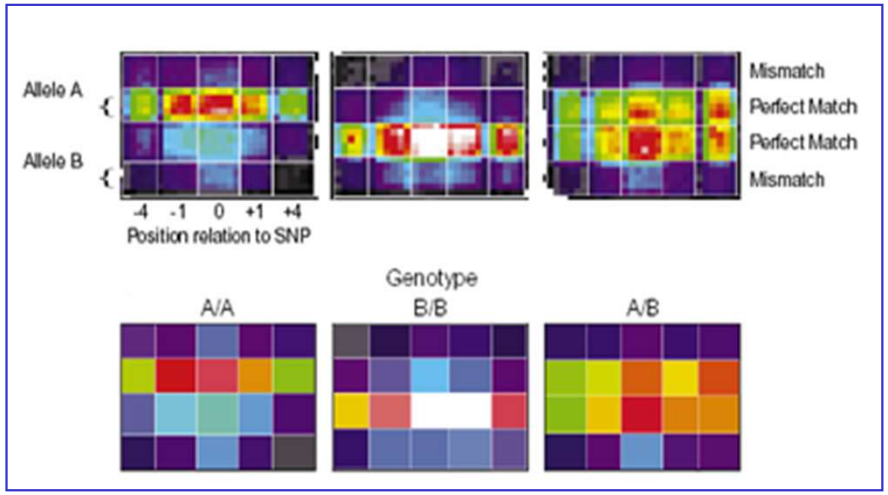
  Less reagents and samples are needed for experiments

  Low cost per target with microarray

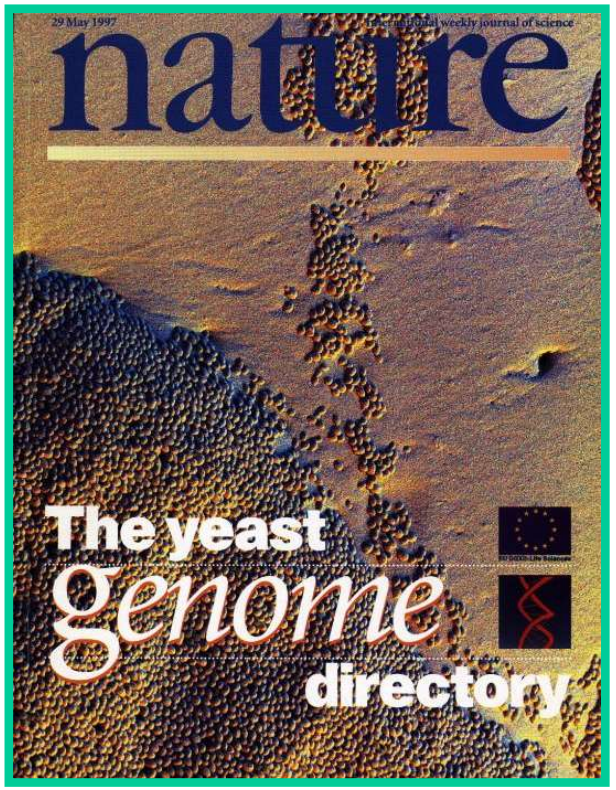
Microarrays - Applications

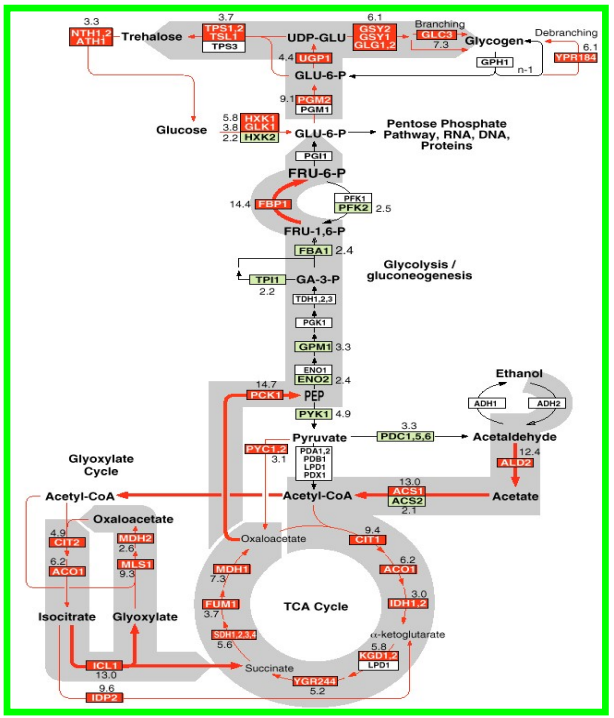
- **Gene expression**
  - Focus of most current array-based studies is monitoring of RNA expression levels
  - approx. 75 % of all publications

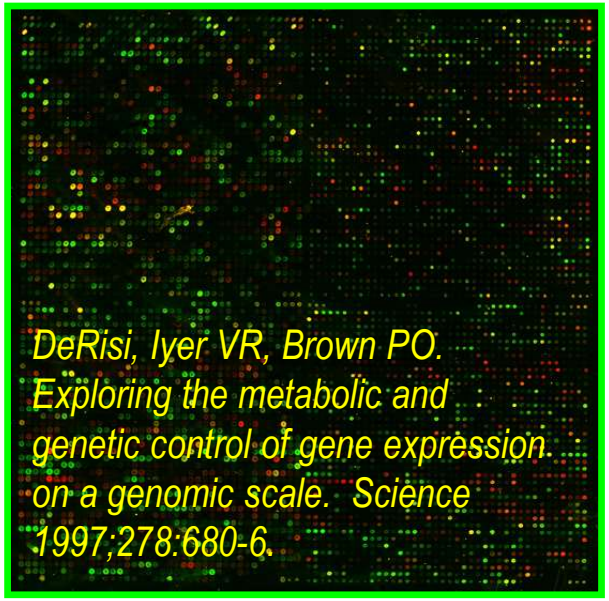
- **DNA variation**
  - identification and genotyping of mutations and polymorphisms (SNPs)
  - approx. 25 % of all publications



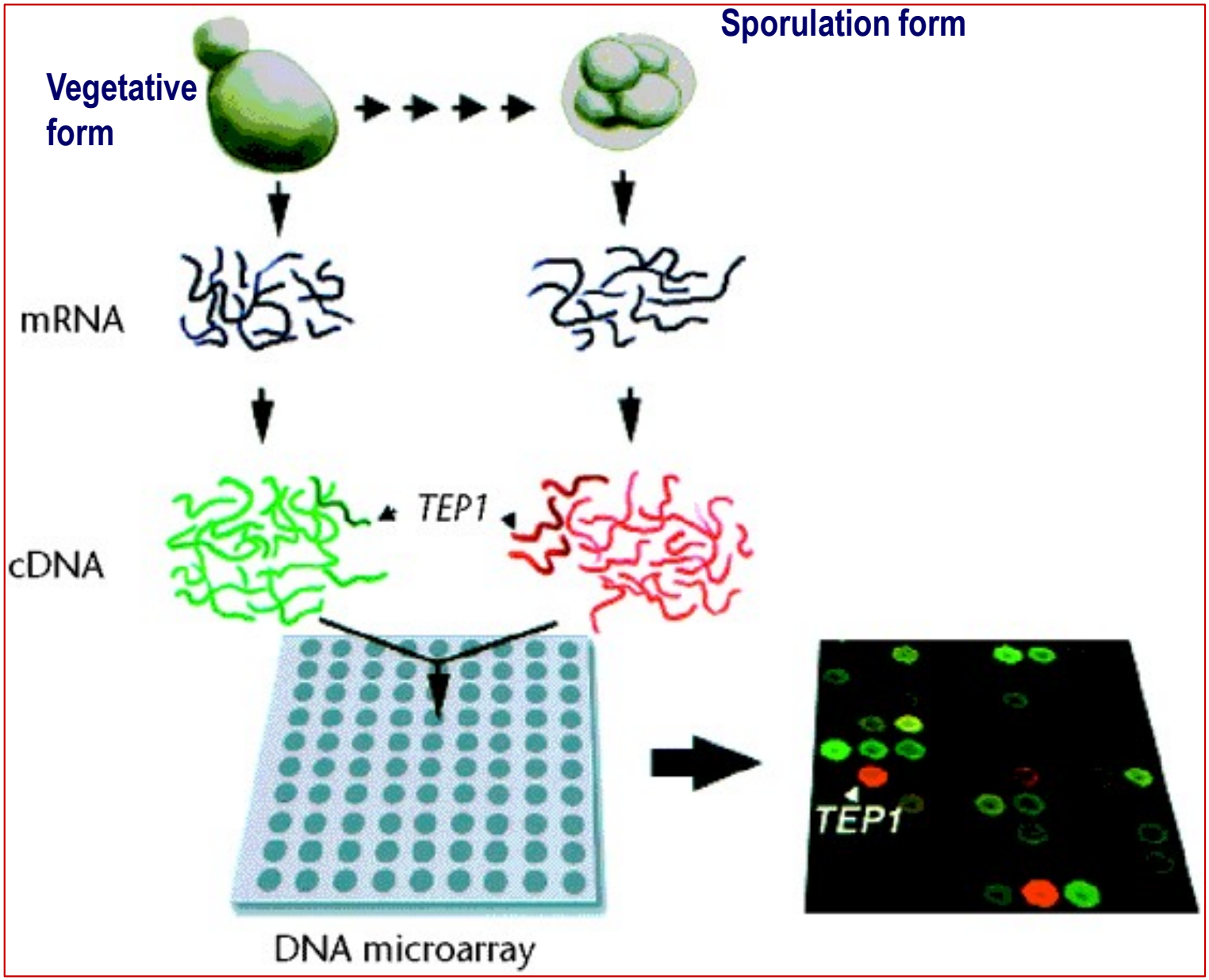
YEAST GENOME
COMPLETED
4/25/97
TOTAL TIME: 51 HRS.
SPOTS: 707,520


29 May 1997
international weekly journal of science
nature
The yeast
genome
directory

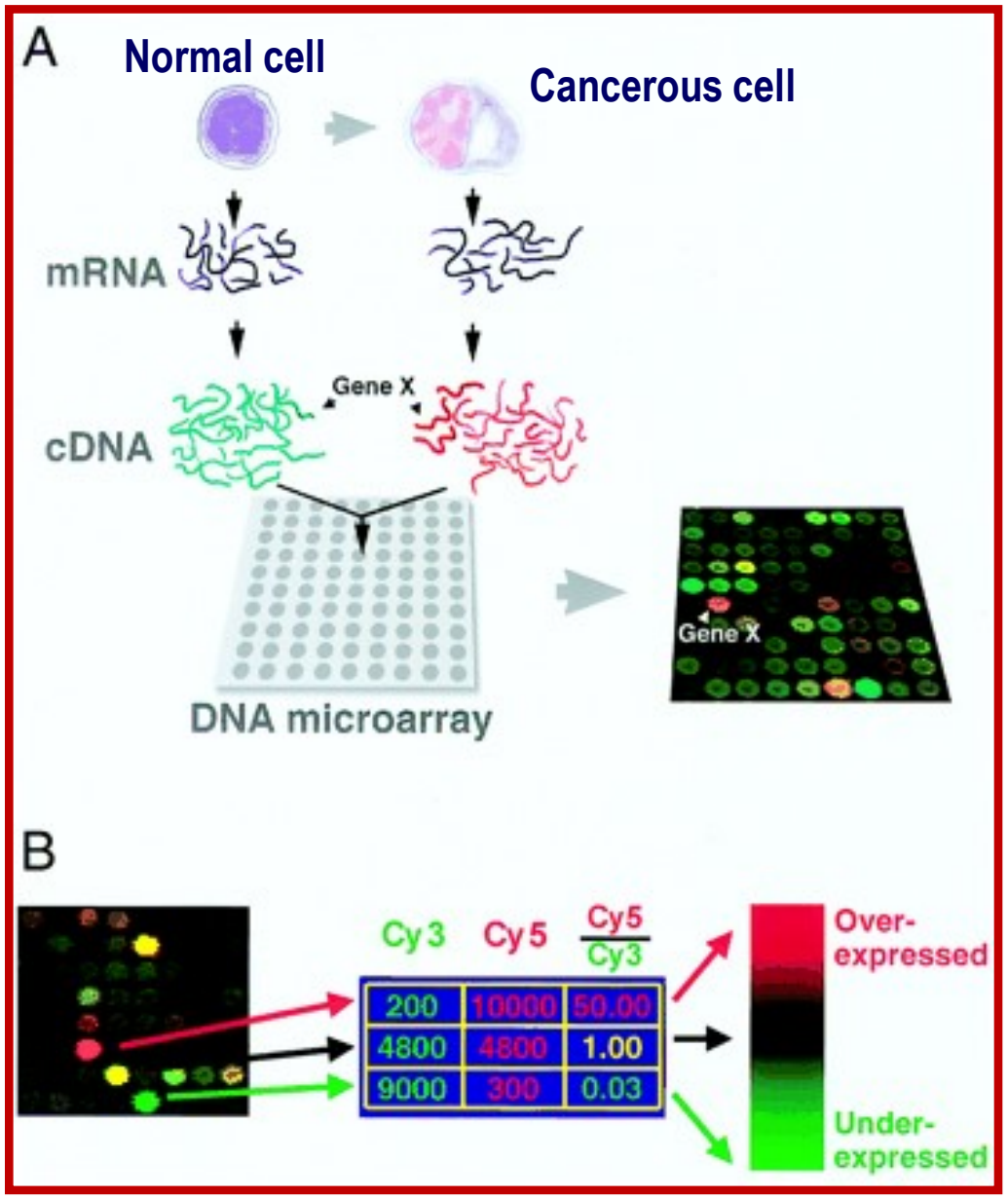

Trehalose
UDP-GLU
Glycogen
Branching
Debranching
GLU-6-P
Glucose
Pentose Phosphate Pathway, RNA, DNA, Proteins
FRU-6-P
FRU-1,6-P
GA-3-P
Glycolysis / gluconeogenesis
PEP
Pyruvate
Ethanol
Acetaldehyde
Acetate
Acetyl-CoA
Glyoxylate Cycle
Oxaloacetate
Isocitrate
Glyoxylate
TCA Cycle
Succinate
α-ketoglutarate


DeRisi, Iyer VR, Brown PO.
Exploring the metabolic and
genetic control of gene expression
on a genomic scale. Science
1997;278:680-6.

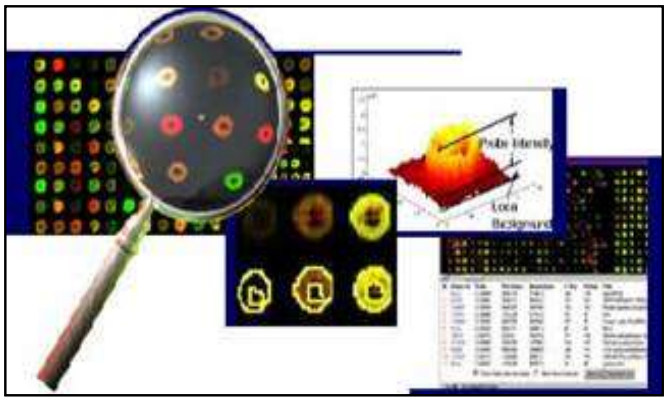
Using DNA Microarrays to Study Expression Profiles in Yeast



Using DNA Microarrays to Study Expression Profiles in Oncology



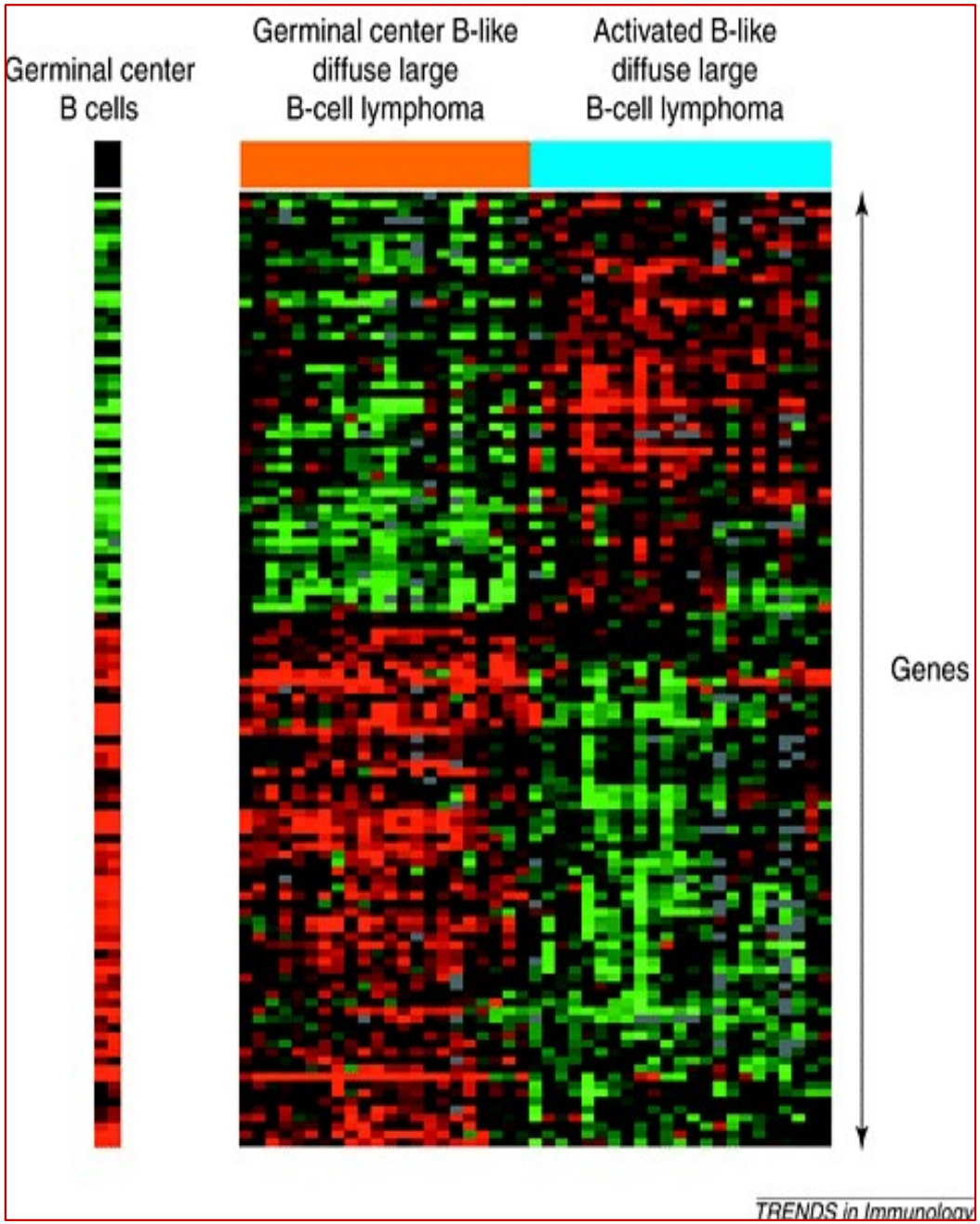
Bioinformatics

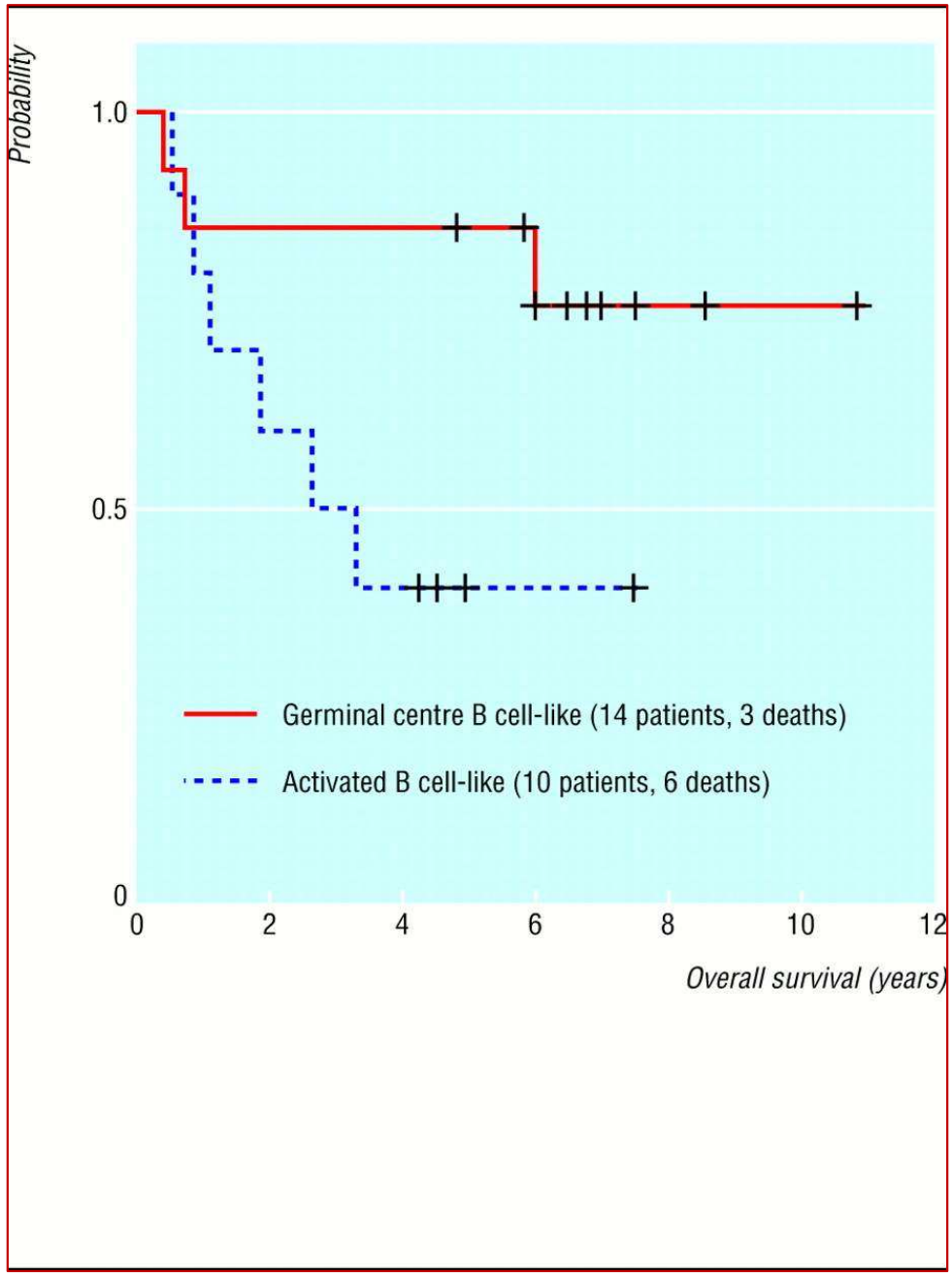


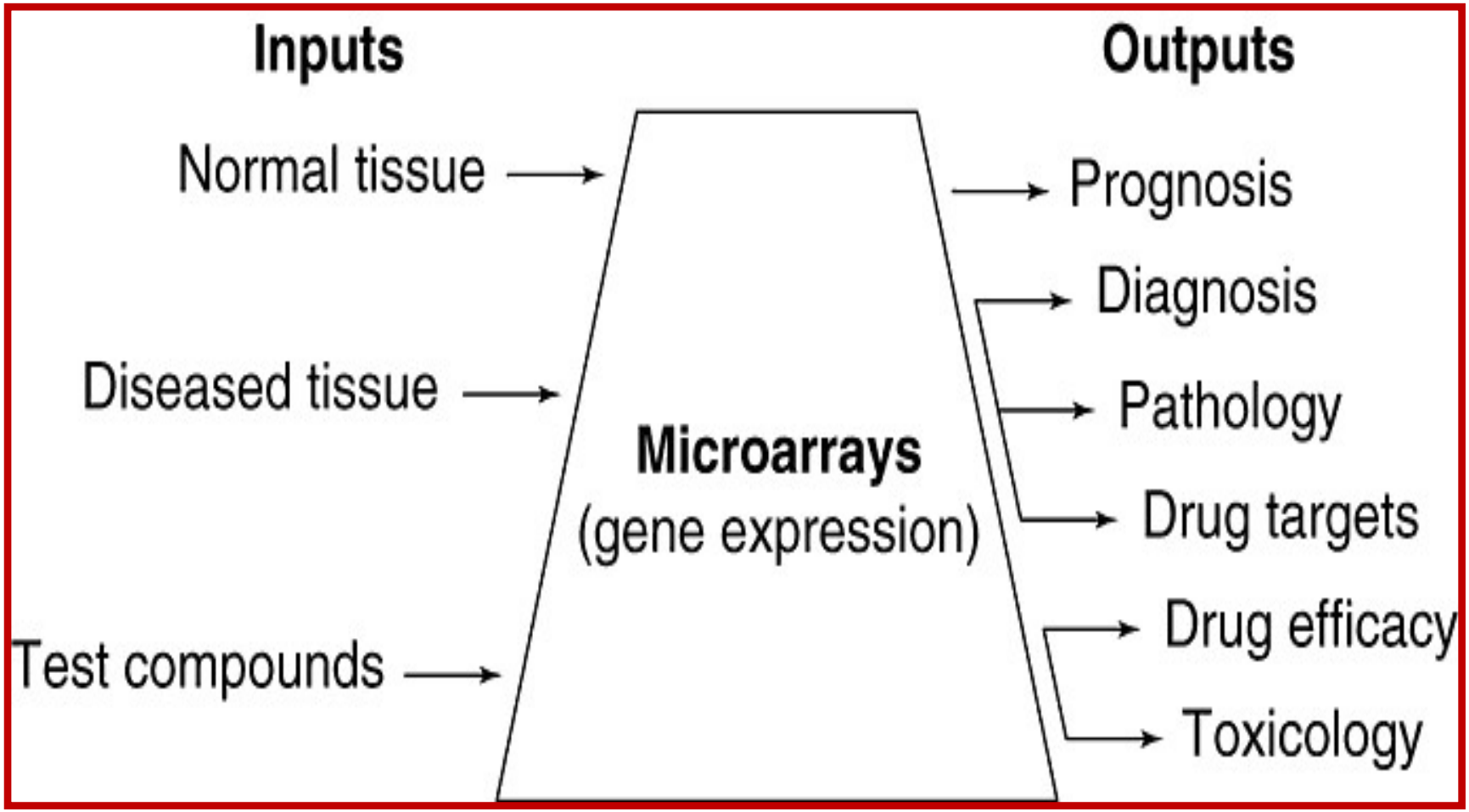
Data → **Information** → **Knowledge**

The science (and art) of converting large sets of raw data to meaningful information = **Knowledge**

Difference Expression Profiles in Diffuse Large B cell lymphoma

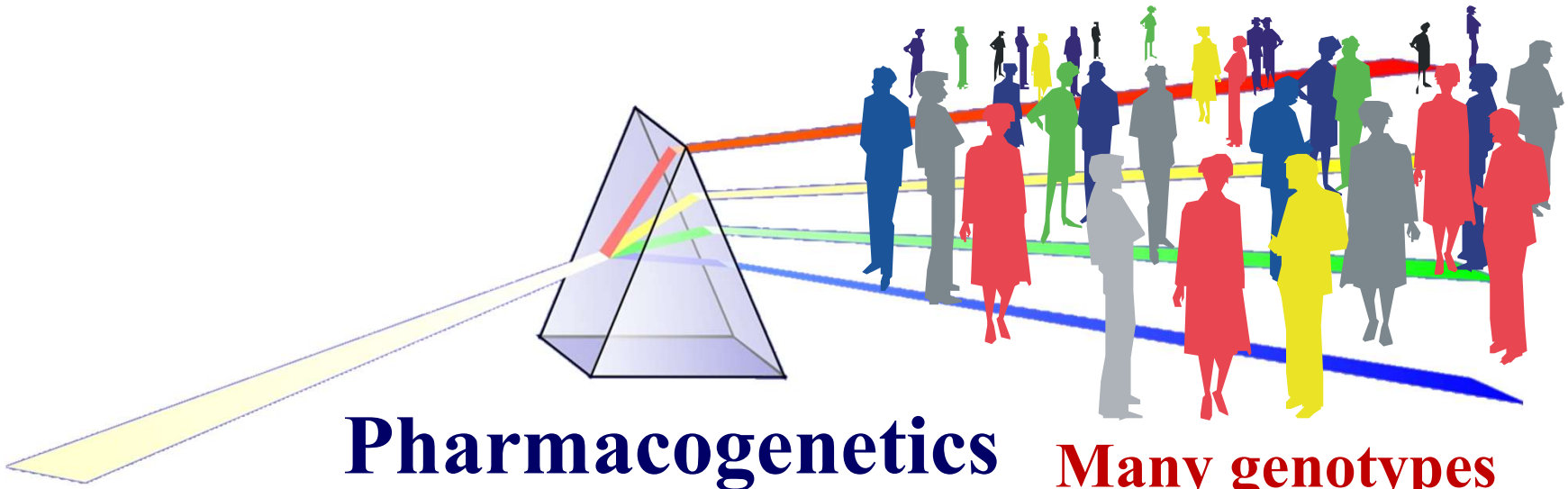






DNA microarrays approach to study expression profiles coupled with other type of information can be used in several biomedical research areas

Applications of DNA Microarrays in Pharmacogenetics



Pharmacogenetics

Many genotypes

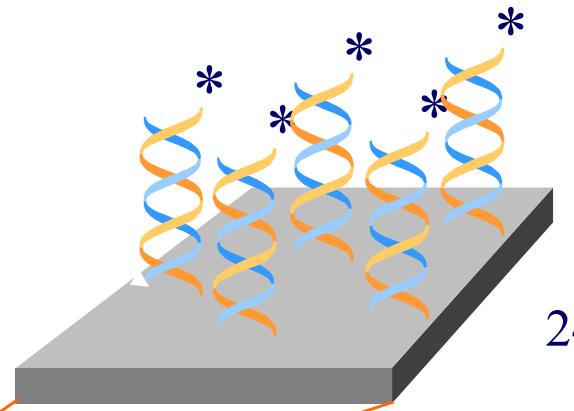
One phenotype

Patient stratification for Customized prescriptions

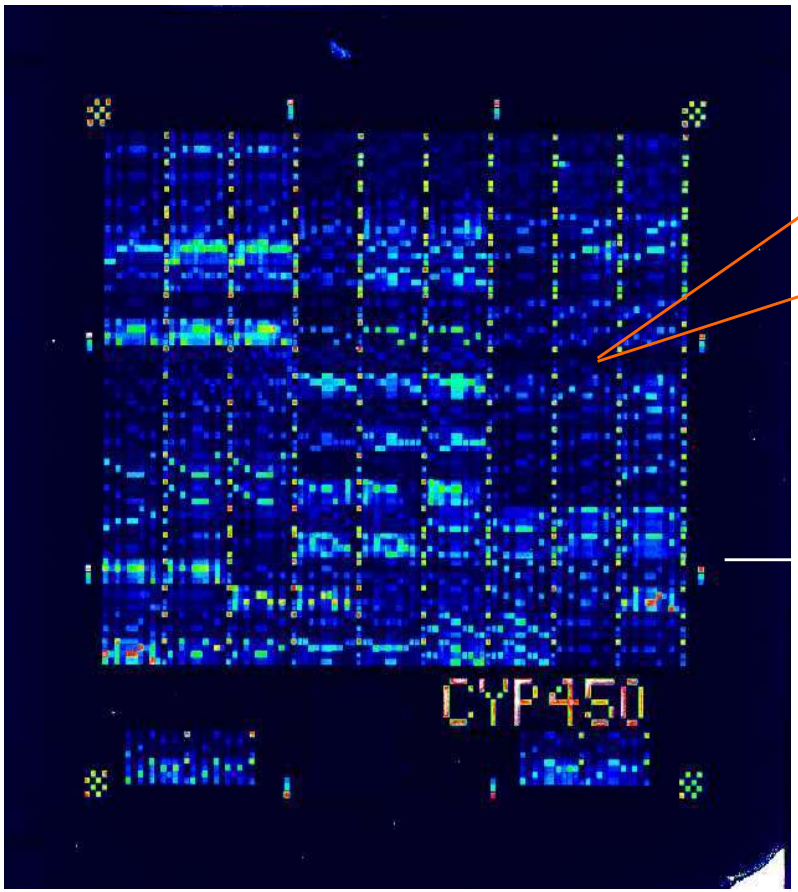
GeneChip CYP450 Probe Array

Fluorescent labeled DNA fragments (*)
hybridized to the Oligonucleotide probes on the chip

Oligonucleotide probes attached to the chip surface



24 µm



0.525 cm

Each probe cell contains 100 million
specific oligonucleotide probes

Each array contains over 8,000 different probes complementary to sense and anti-sense CYP450 genomic DNA. Probes range in length from 18mer to 22mer

Review on CYP450: http://aafp.org/afp/980101ap/cupp.html or Am Fam Physician 1998:57(1):1070

Proteomics

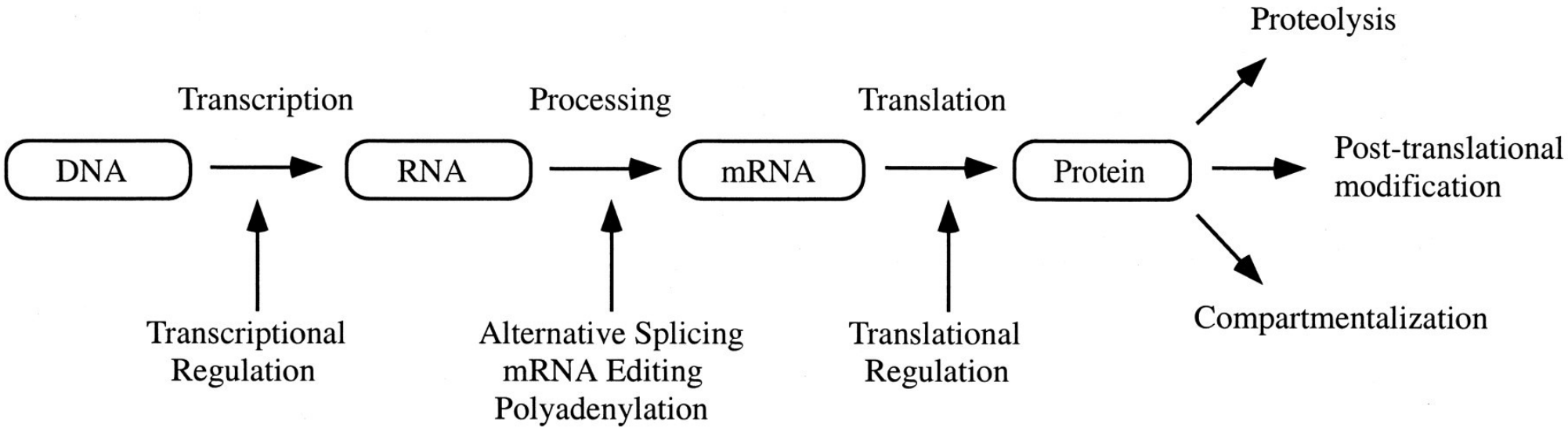
- The study of all the protein functions *(expression level and pattern,post translational modification etc.)* produced by cell type, tissue or organism on a large scale to obtain a global, integrated view of physiological and pathophysiological processes and networks at the protein level
- It involves the isolation, separation, identification and characterization of the proteins in the body

- **Proteome** refers to all the proteins expressed by a genome. (*proteome is dynamic*)
- **Differential or Expression proteomics:** The study of protein expression in comparative tissues samples.
- **Functional proteomics:** The study of how proteins interact with other cellular components in a single time frame.

Why proteomics?

- Many types of information cannot be obtained from the study of genes alone, *e.g. genomic information does not predict post-translational modifications*
- Annotation of the genome: genome information have to integrate with data obtained from protein studies to confirm the existence of a particular gene.
- mRNA expression (genomic scale) profile is not a direct reflection of the protein content in the cell.

- Protein modification
- Protein localization & compartmentalization
- Protein – protein interactions



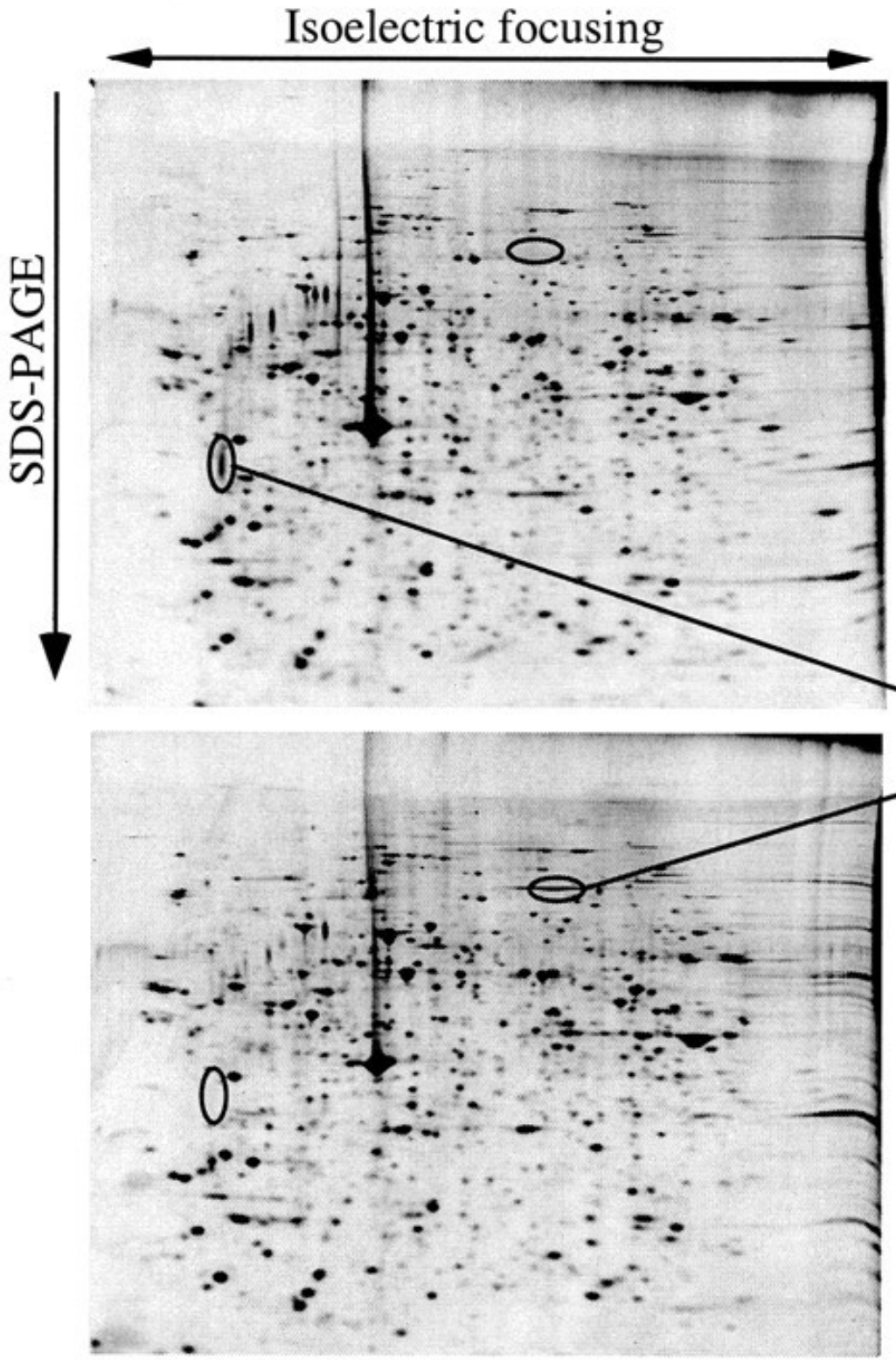
- A gene can give rise to multiple gene products.
- The average number of protein forms per gene was predicted to be : 1-2 in bacteria, 3 in yeast, more than three in humans.

Techonlogy of Proteomics

I. Separation and Isolation of Proteins
II. Acquisition of Protein Structure Information
III. Database Utilization

I. Separation and Isolation of Proteins

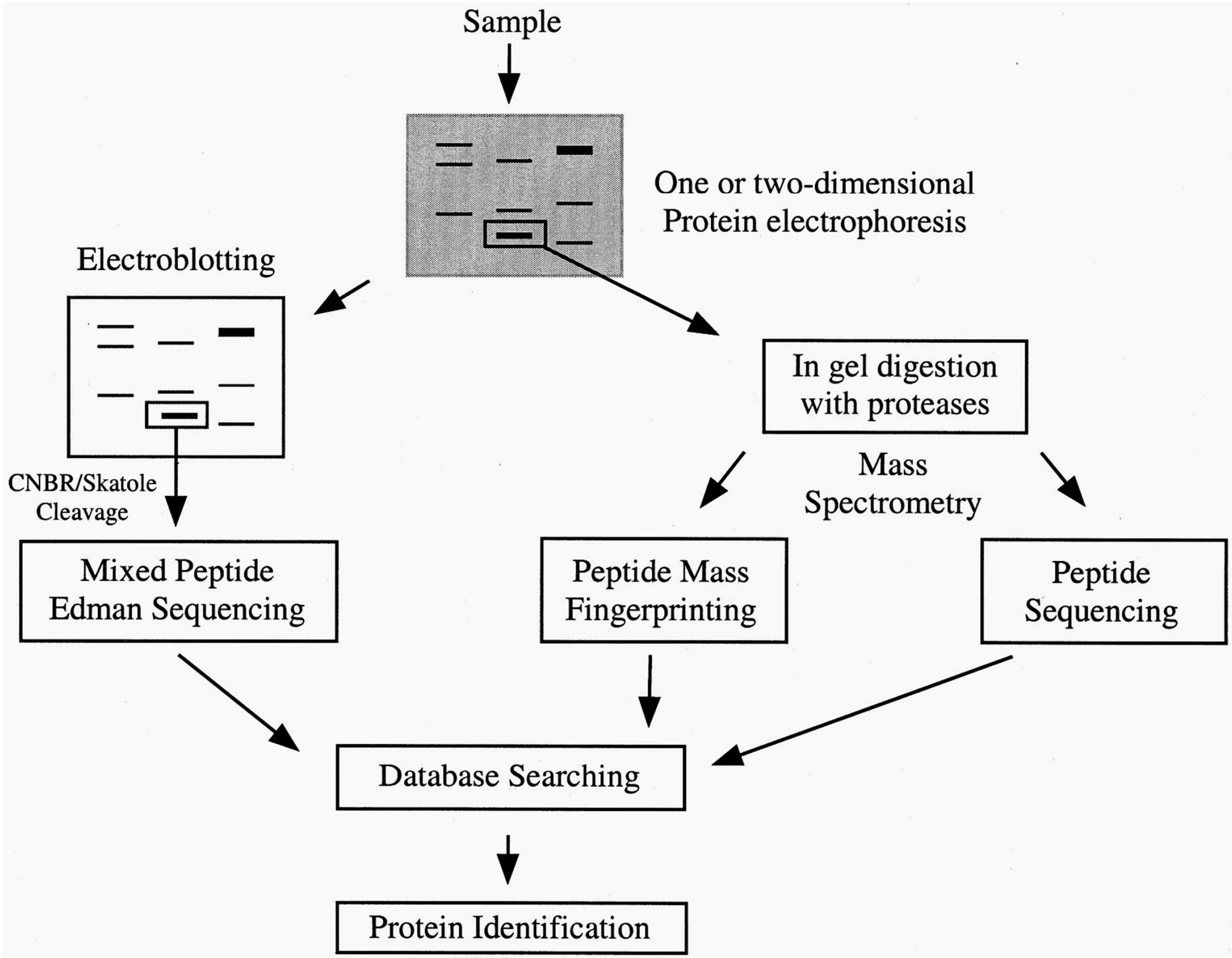
1. One – and two–dimensional gel electrophoresis
2. Alternatives to electrophoresis

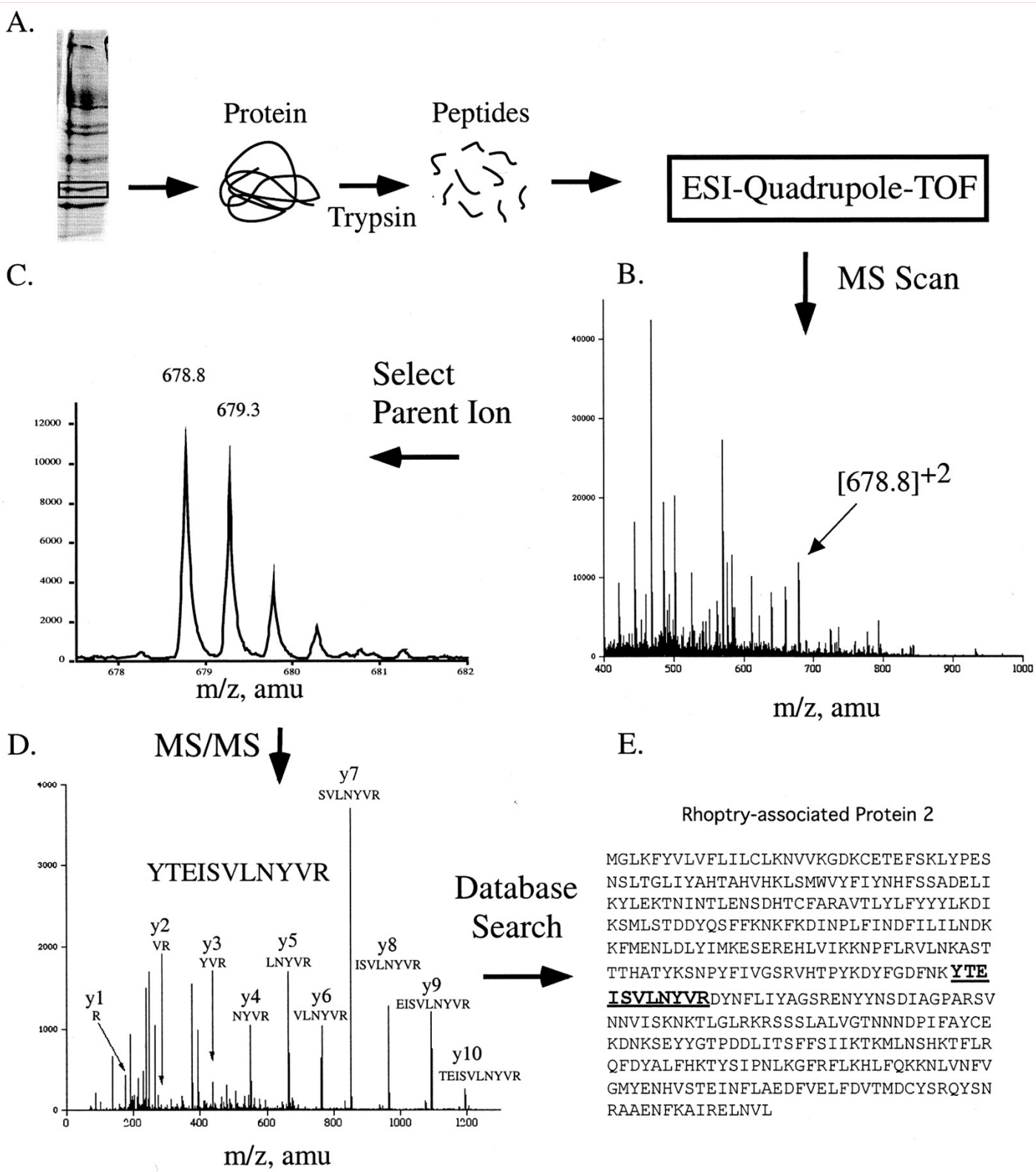


AMuLV - Abelson Murine Leukemia Virus

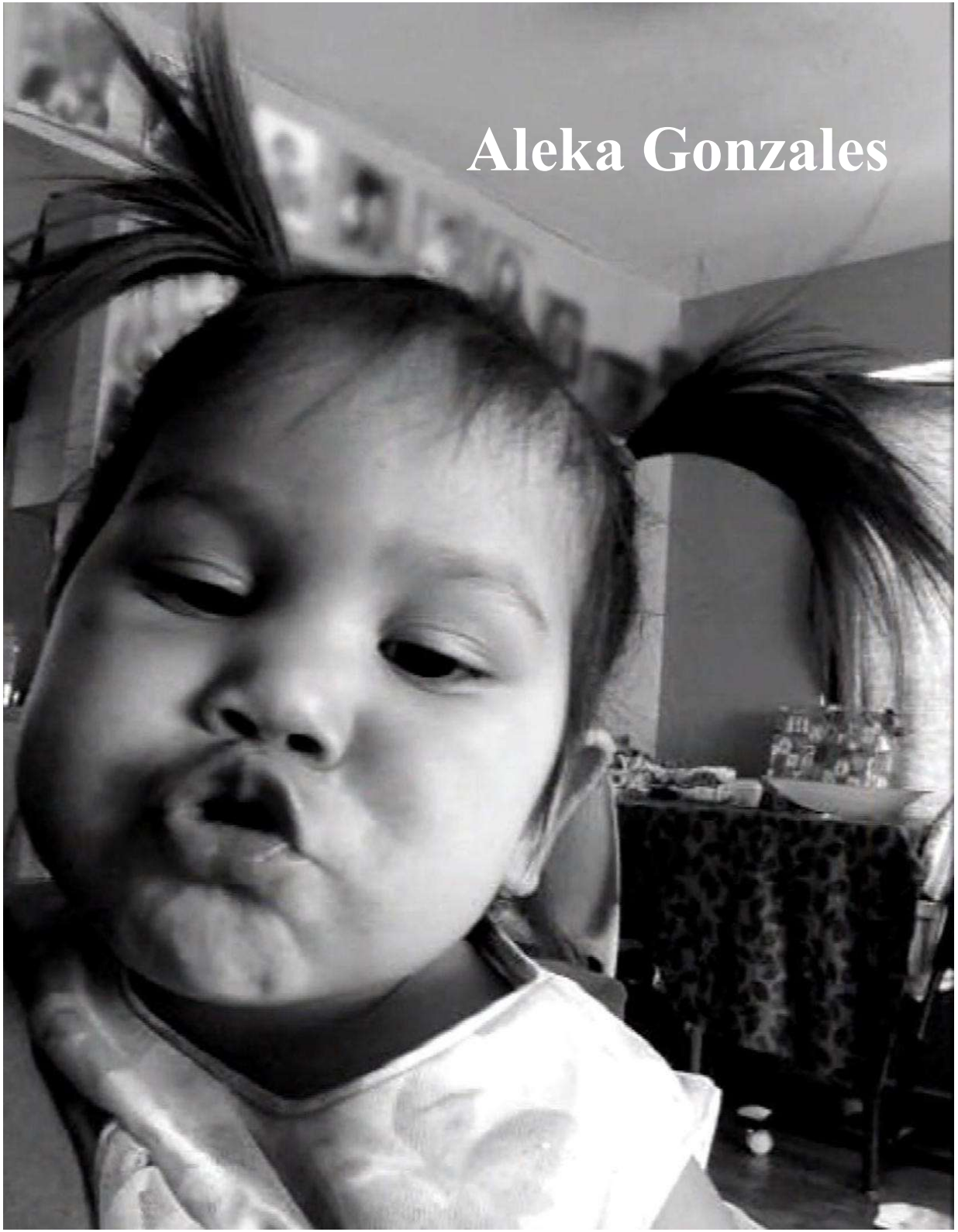
II. Acquisition of Protein Structure Information

- Edman sequencing
- Mass spectrometry


Sample
One or two-dimensional
Protein electrophoresis
Electroblotting
In gel digestion
with proteases
CNBR/Skatole
Cleavage
Mass
Spectrometry
Mixed Peptide
Edman Sequencing
Peptide Mass
Fingerprinting
Peptide
Sequencing
Database Searching
Protein Identification


A.
Protein
Peptides
Trypsin
ESI-Quadrupole-TOF
MS Scan
B.
[678.8]+2
m/z, amu
C.
678.8
679.3
Select
Parent Ion
m/z, amu
D.
MS/MS
y7
SVLNYVR
YTEISVLNYVR
y2
VR
y3
YVR
y5
LNYVR
y8
ISVLNYVR
y1
R
y4
NYVR
y6
VLNYVR
y9
EISVLNYVR
y10
TEISVLNYVR
m/z, amu
Database
Search
E.
Rhoptry-associated Protein 2
MGLKFYVLVFLILCLKNVVKGDKCETEFSKLYPES
NSLTGLIYAHTAHVHKLSMWVYFIYNHFSSADELI
KYLEKTNINTLENSDHTCFARAVTLYLFYYYLKDI
KSMLSTDDYQSFFKNKFKDINPLFINDFILILNDK
KFMENLDLYIMKESEREHLVIKKNPFLRVLNKAST
TTHATYKSNPYFIVGSRVHTPYKDYFGDFNKYTE
ISVLNYVRDYNFLIYAGSRENYYNSDIAGPARSV
NNVISKNKTLGLRKRSSSLALVGTNNNDPIFAYCE
KDNKSEYYGTPDDLITSFFSIIKTKMLNSHKTFLR
QFDYALFHKTYSIPNLKGFRFLKHLFQKKNLVNFV
GMYENHVSTEINFLAEDFVELFDVTMDCYSRQYSN
RAAENFKAIRELNVL

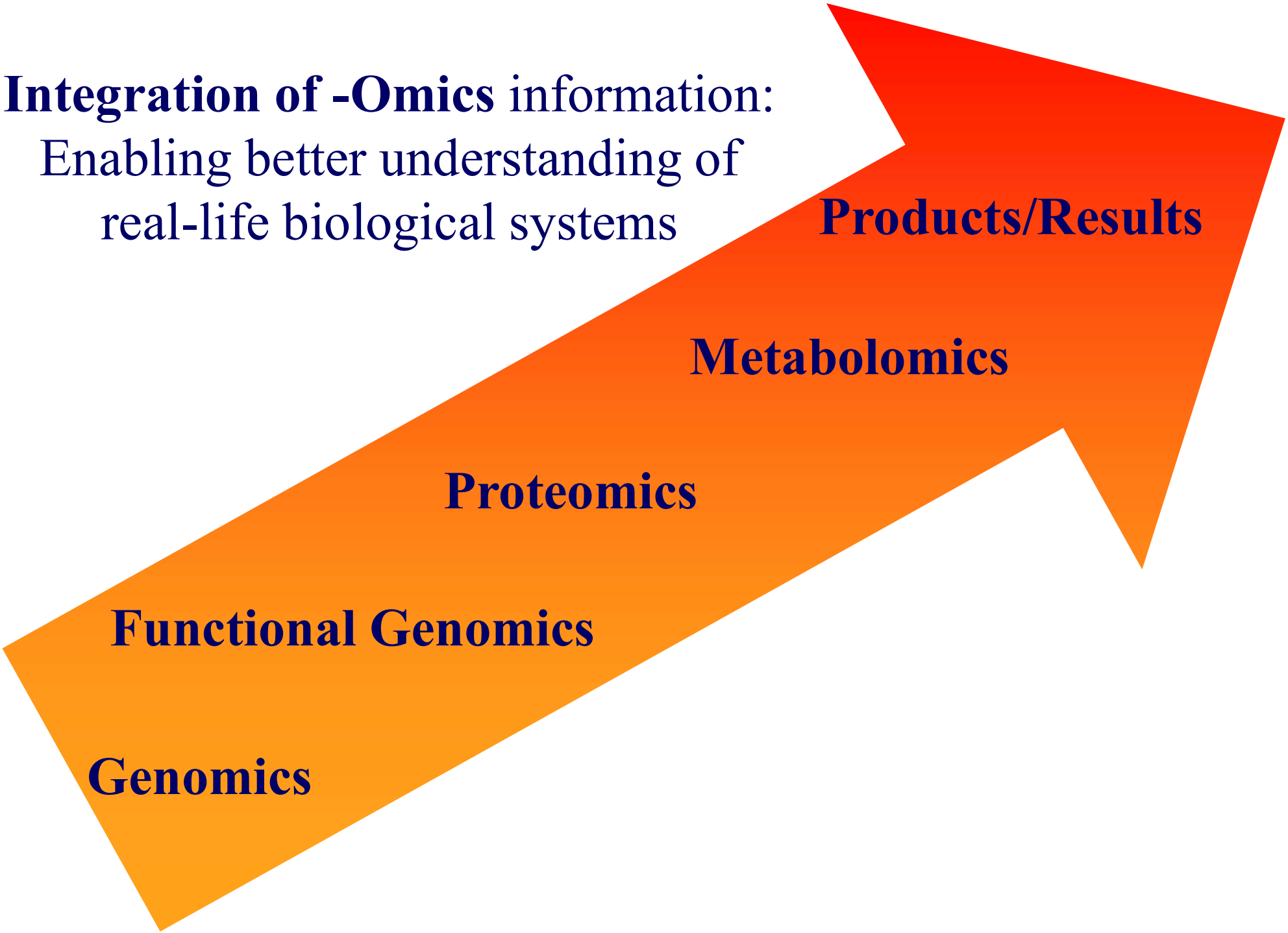
Forensic Proteomics


Aleka Gonzales

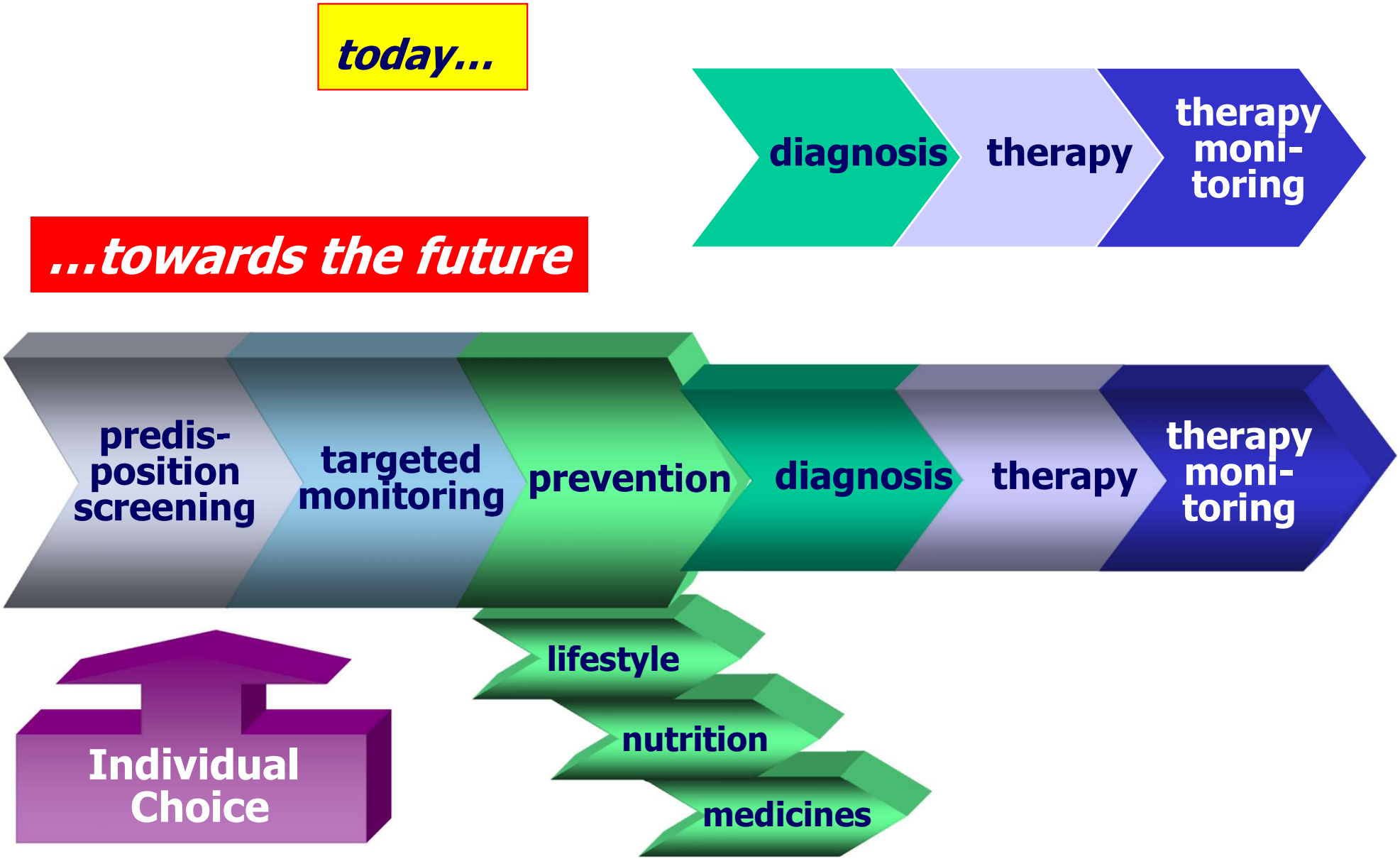
Prof. Steve Mackessy

Proteomics Offers New Clues for Forensic Investigations

Integration of -Omics information:
Enabling better understanding of
real-life biological systems



The Impact of the Human Genome Project in Medicine

- Once all genes are known, we will start understanding their function, their interactions and their pathways in various situations including pathological conditions.
- Thus, it is possible for us to correlate disease states to certain gene(s)
  - **Disease → Gene(s)**
  - **Gene(s) → Disease**
- Finally, it is possible to find ways for diagnosis, prognosis, rational treatments and monitoring

Integrated Health Care
an evolving paradigm