

Genomica, proteomica,
genomica strutturale, banche
dati

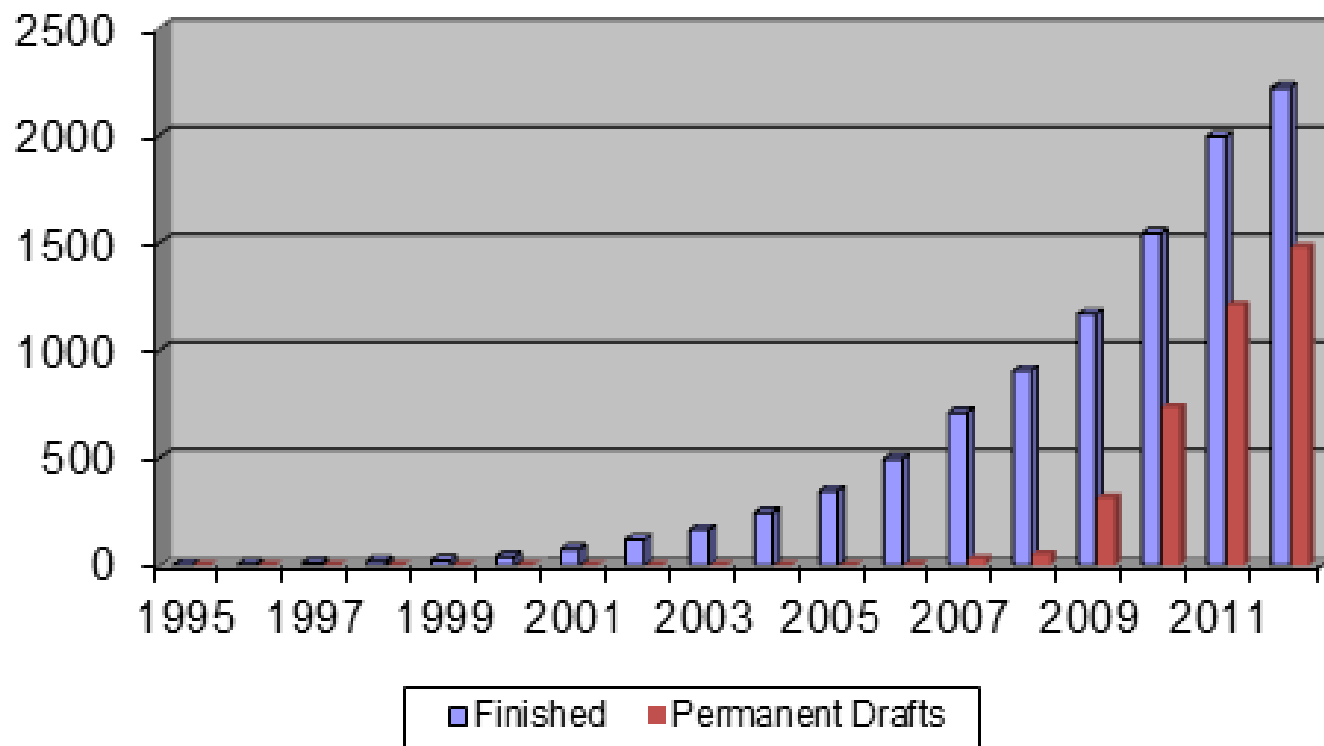
Alcune pietre miliari della biologia

<u>anno</u>	<u>risultato</u>
1866	Mendel scopre i geni
1944	il DNA è il materiale genetico
1951	prima sequenza di una proteina (insulina)
1953	struttura del DNA
1959	struttura della mioglobina
1960s	delucidazione del codice genetico
1977	avvento del sequenziamento del DNA
1975-79	primi clonaggi di geni umani
1986	sviluppo di un sistema di seq. aut. del DNA
1995	primo genoma completo (<i>H. influenzae</i>)
1997	genoma di <i>E. coli</i>
1999	primo cromosoma umano (Chr #22)
2000	genomi <i>Drosophila</i> e <i>Arabidopsis</i>
2001	genomi dell'uomo e di topo

I genomi sequenziati
(<http://www.genomesonline.org>)

Complete Genome Projects ©

September 2012: 3699 Projects



La quantità di informazione genetica già disponibile è impressionante e cresce a ritmo vertiginoso.

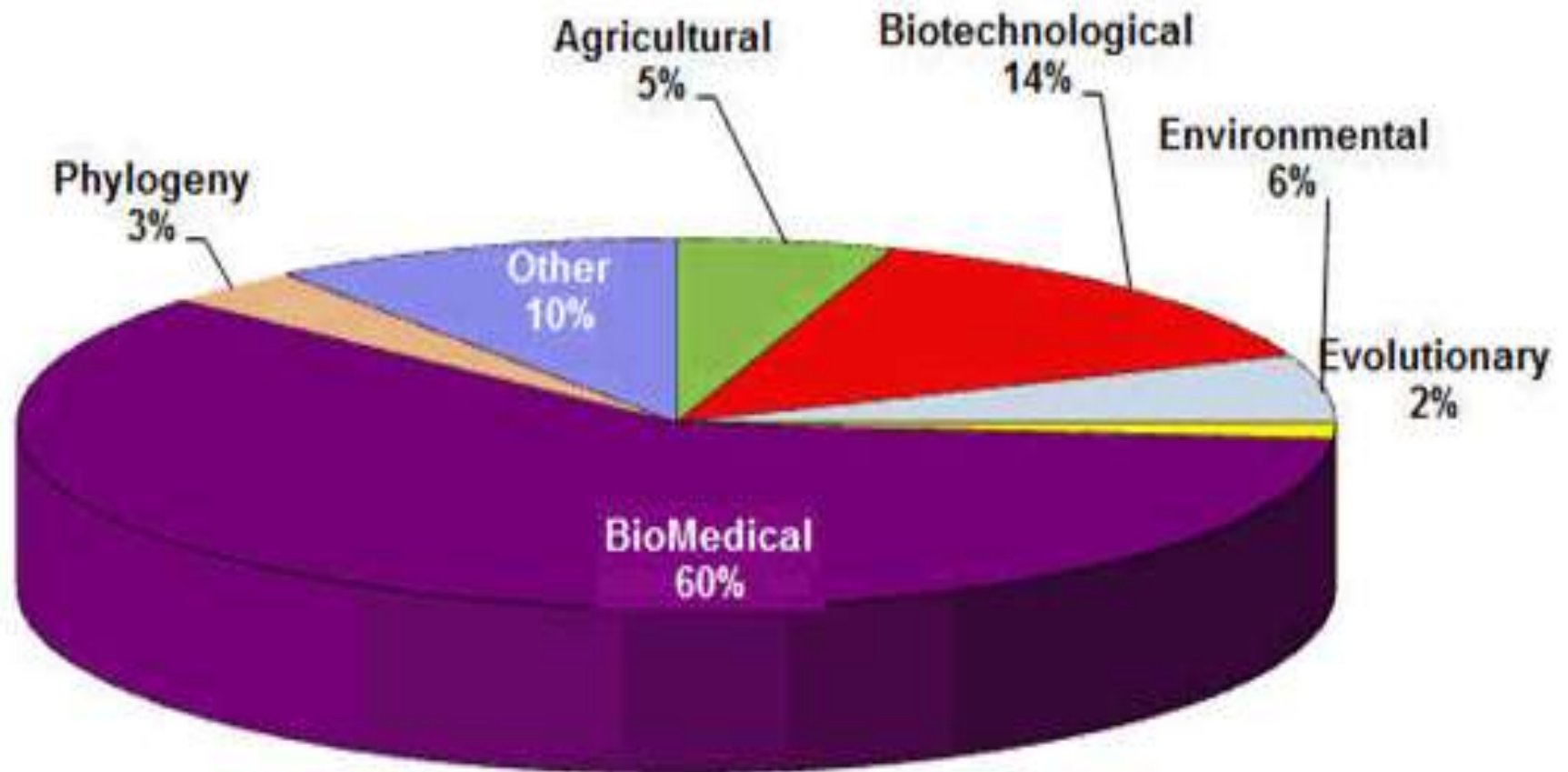
La disponibilità di questa massa di informazioni sta cambiando la ricerca bio(tecno)logica

Metagenomica

La Metagenomica (detta anche genomica ambientale, ecogenomica o genomica delle comunità) è lo studio dei genomi recuperati da ambienti piuttosto che da singoli organismi

Comunità intestinali, comunità marine (es. i batteri del mar dei Sargassi), biofilm

Funding Relevance of Bacterial Genome Projects



Banche dati di acidi nucleici

- Tre consorzi che scambiano informazioni (International Nucleotide Sequence Database Collaboration) :

- [GenBank](#) (americana)
- [EMBL](#) (europea)
- [DDBJ](#) (giapponese)

DATABASE

Una collezione di informazioni organizzata in modo che un programma al computer possa velocemente accedere a determinate porzioni di dati

Banche dati di proteine

- **SWISS-PROT**

sequenze di proteine (<http://www.expasy.ch/>)

- **PDB**

strutture 3D di proteine (<http://www.rcsb.org>)

Bioinformatica e analisi dei genomi

- Bioinformatica : l'analisi tramite strumenti informatici delle informazioni biologiche.

Generalmente ci si riferisce all'analisi computazionale di grandi set di dati su DNA, proteine e dati strutturali

Analisi di singoli geni

- Mappe di restrizione
- Mappe di plasmidi
- ORF e sequenze codificati
- Ricerche in database
- Confronto di sequenze
- Allineamenti multipli di proteine
-

Analisi più complesse

- Assemblaggio di genomi
- Predizione degli ORF
- Identificatione di domini
- Confronto di strutture
- Predizioni strutturali
- Predizioni di promotori e giunzioni di splicing
- Analisi di genomi

Le analisi “-omiche”

Genomica

Trascrittomica

Proteomica

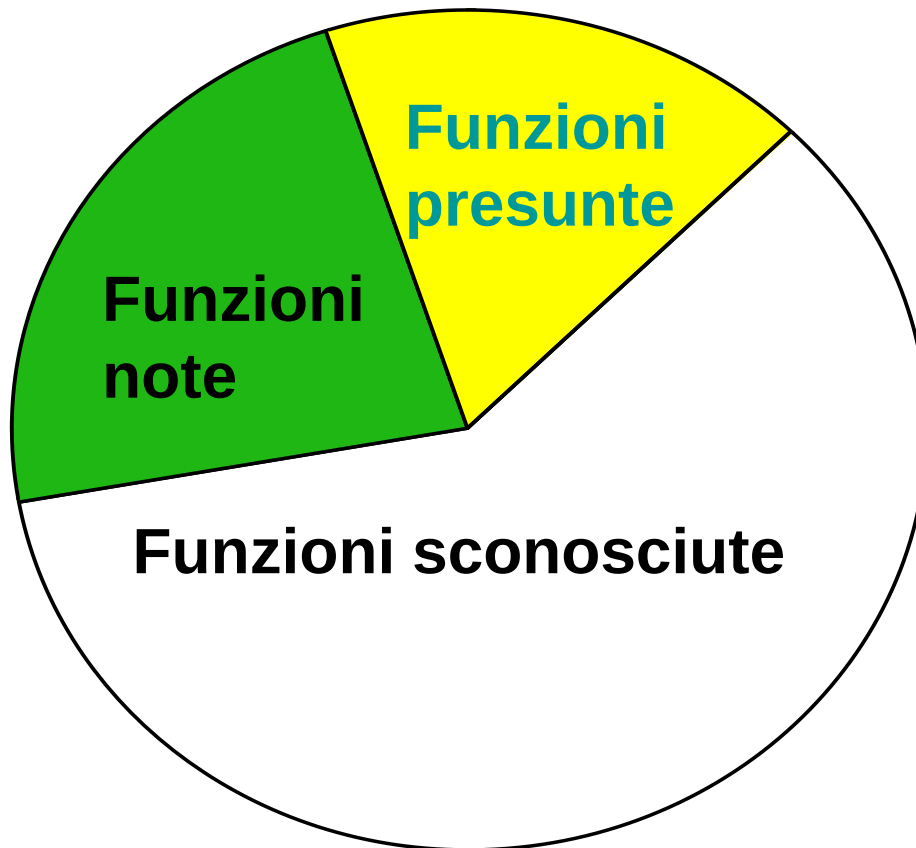
Metabolomica

Level of analysis	Definition	Status	Method of analysis
Genome	Complete set of genes of an organism or its organelles	Context-independent (modifications to the yeast genome may be made with exquisite precision)	Systematic DNA sequencing
Transcriptome	Complete set of messenger RNA molecules present in a cell, tissue or organ	Context-dependent (the complement of messenger RNAs varies with changes in physiology, development or pathology)	Hybridization arrays ⁸ SAGE ¹³ High-throughput Northern analysis ¹⁴
Proteome	Complete set of protein molecules present in a cell, tissue or organ	Context-dependent	Two-dimensional gel electrophoresis, peptide mass fingerprinting ^{9,10} Two-hybrid analysis ⁵
Metabolome	Complete set of metabolites (low-molecular-weight intermediates) in a cell, tissue or organ	Context-dependent	Infra-red spectroscopy ¹⁵ Mass spectrometry ¹⁵ Nuclear magnetic resonance spectrometry ¹⁶

Alcune sfide della genomica e della proteomica

- Qual è la funzione di ciascun gene (e di ciascuna proteina) ?
- Come viene regolata l'espressione dei geni ?
- Come rispondono i geni agli stimoli ambientali ?
- Quali geni sono coinvolti nelle diverse malattie ?
- In che modo le diverse proteine interagiscono tra loro ?
-

A che servono tutte queste proteine ?



Analisi del genoma/proteoma

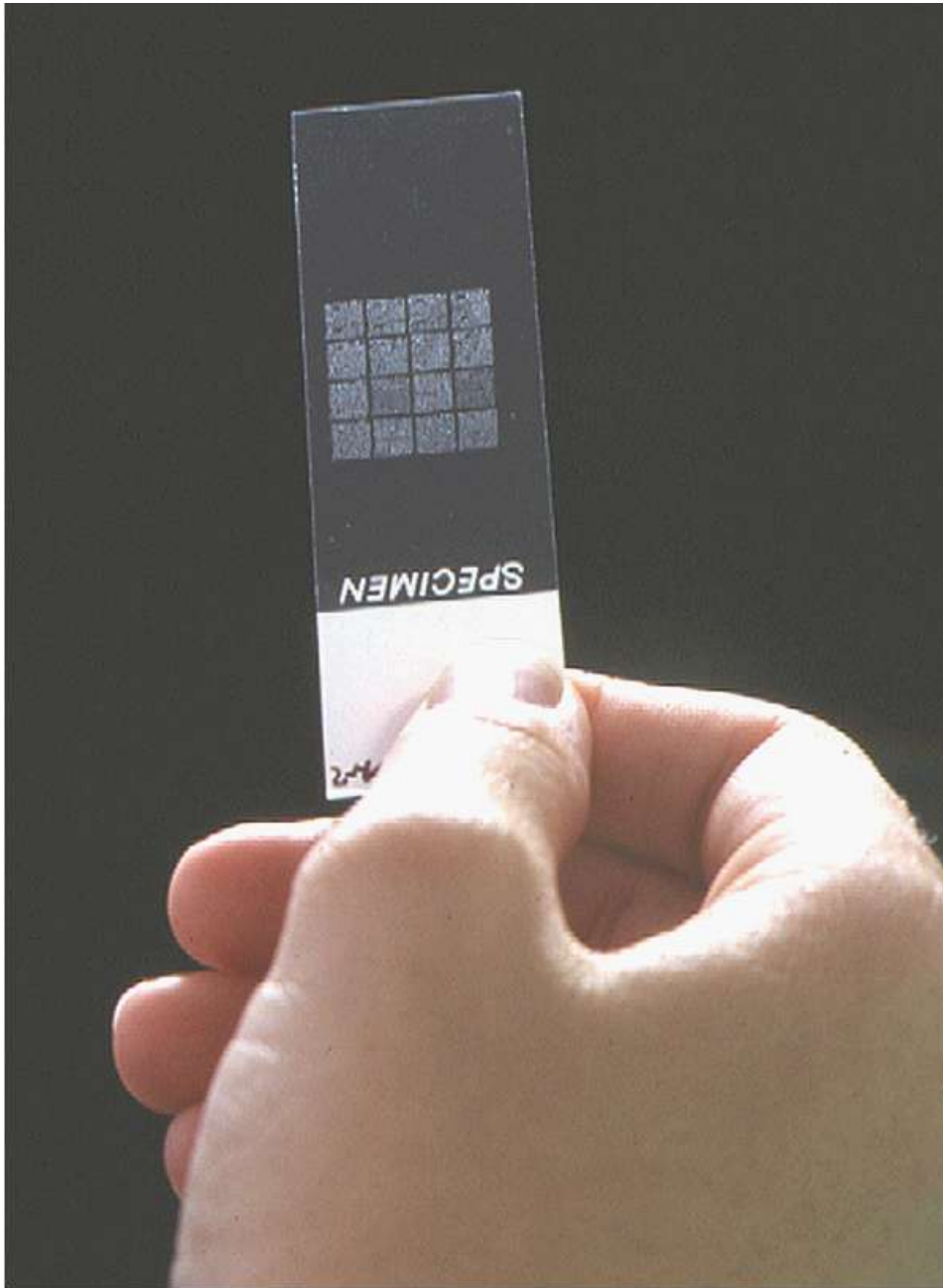
- Fino poco tempo fa i ricercatori studiavano l'espressione di un singolo gene
- Ora è possibile studiare simultaneamente l'espressione di tutti i geni di un organismo (questo può aiutare a capire meglio la funzione dei singoli geni nel contesto cellulare)

Perchè analizzare così tanti geni?

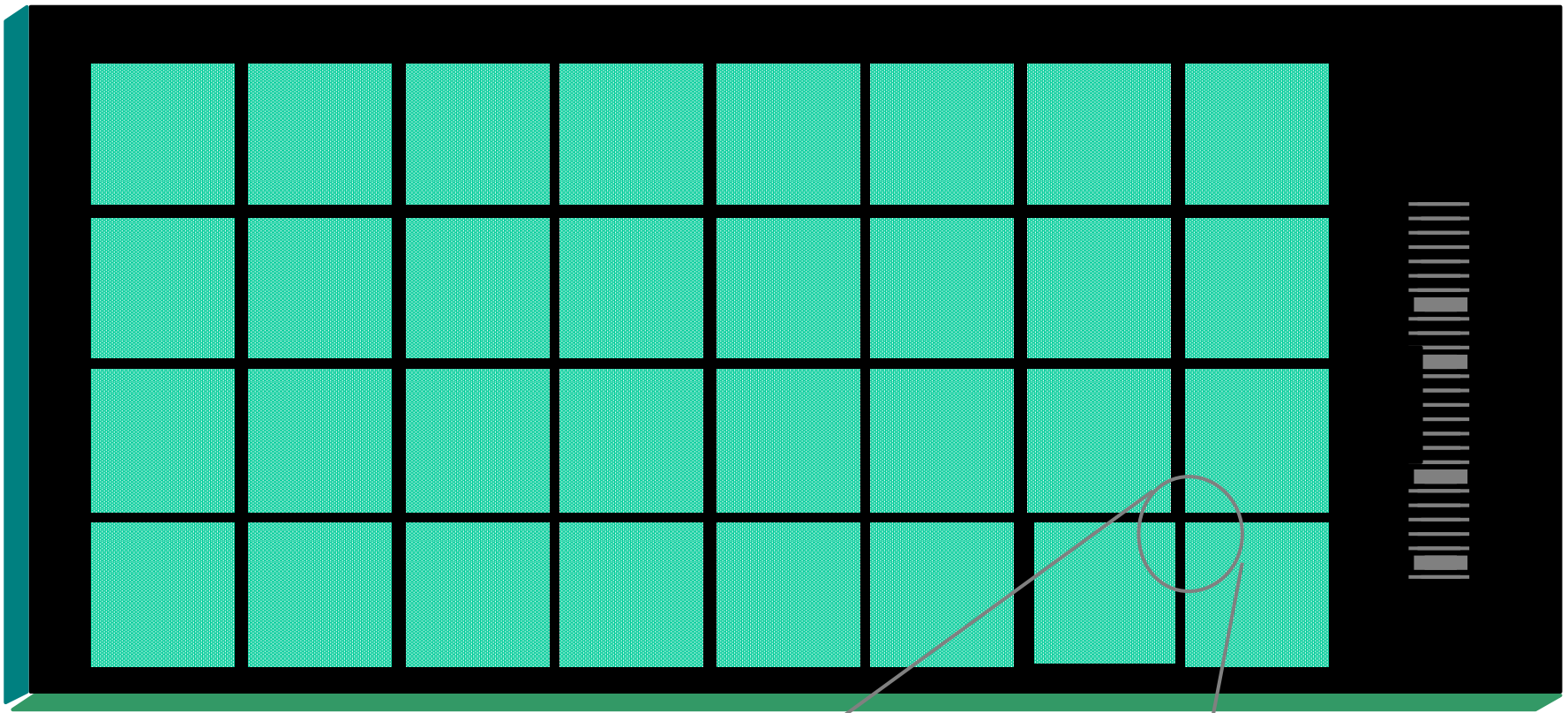
- <10% dei geni umani sono stati studiati a livello della loro funzione individuale. Ma i geni totali sono circa 40,000
- I pannelli di espressione globale forniscono molte più informazioni e soprattutto si ottengono informazioni non attese !

Cosa sono i Microarrays?

- I microarrays sono semplicemente delle lastrine di vetro o silice sulla cui superficie sono disposti in modo ordinato migliaia di geni (tra 500-20,000)
- Tramite una convenzionale ibridazione con una sonda marcata (fluorescente), viene misurato il livello di espressione di tutti questi geni
- I dati vengono rilevati tramite opportuni lettori
- Si confrontano i risultati con campioni di controllo



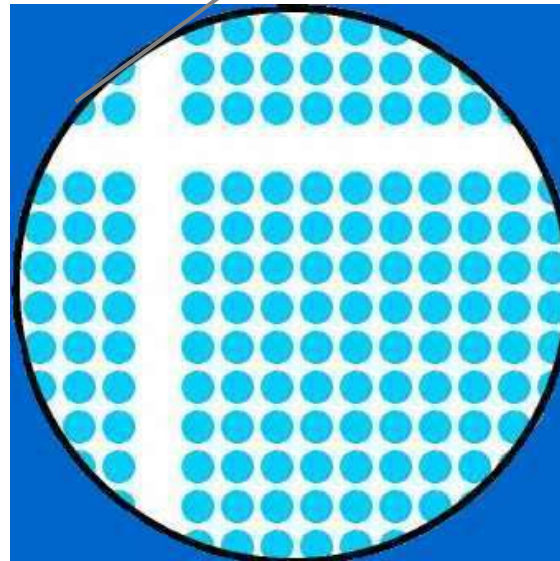
Un microarray



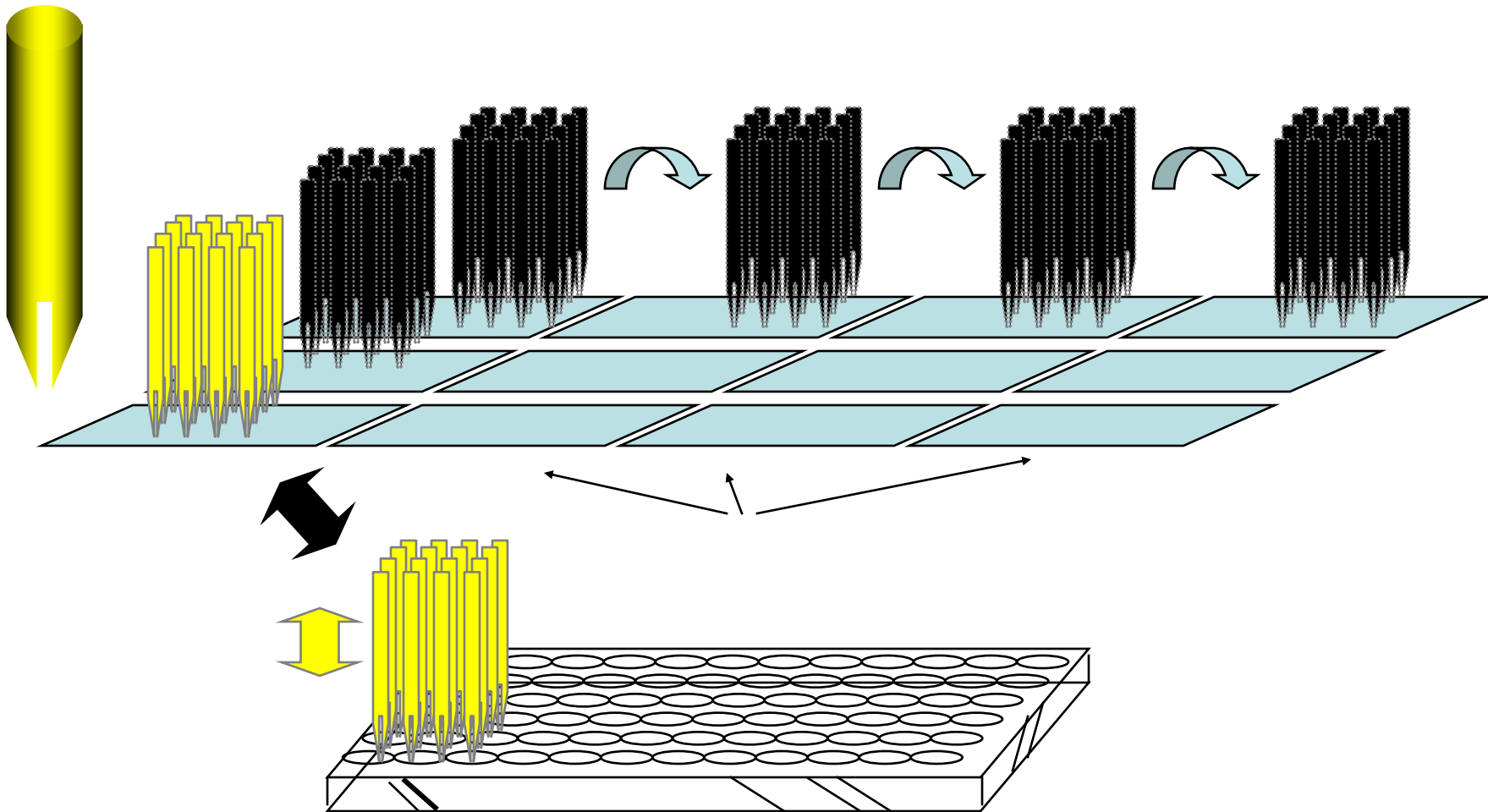
Yeast genome: 12,800 points

Diameter: 120 microns

**Slide size: 170 mm
x 340 mm**



Preparare un microarray



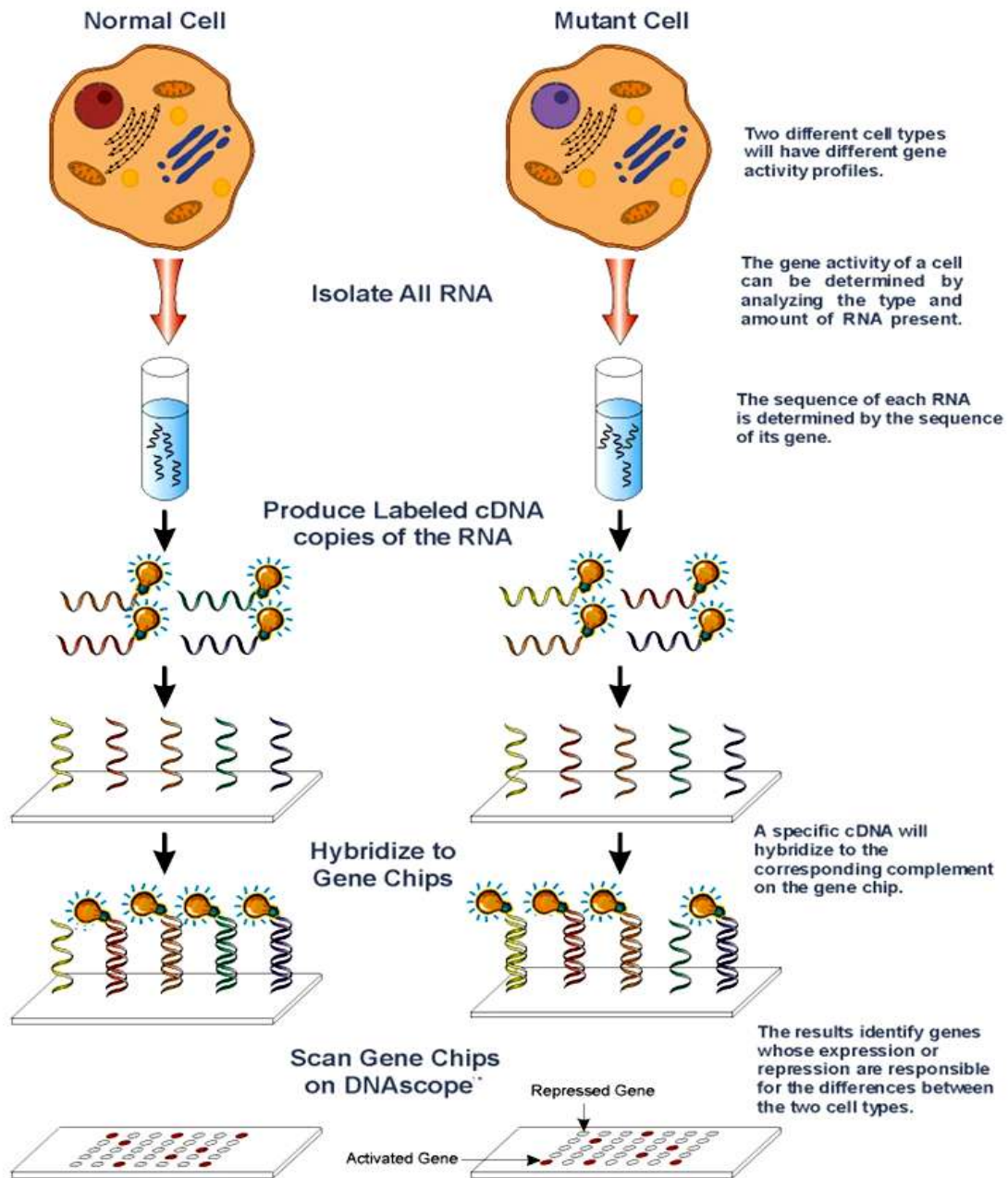
1 goccia di un nanolitro
90-120 μm diametro



Microarray at OCI

I diversi passaggi di un esperimento

- 1 - disegno sperimentale: cosa paragonare a cosa?
- 2 - ottenere un microarray
- 3 - preparazione della sonda e ibridazione
- 4 - acquisizione delle immagini e quantificazione
- 5 - costruzione di un database
- 6 - analisi statistica - Risultati

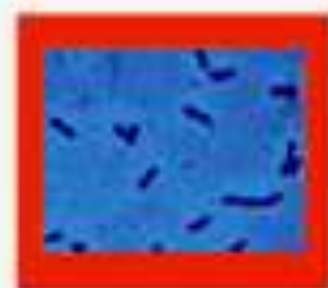


- La popolazione di messaggeri (mRNA) proveniente dal campione biologico di interesse viene usata come stampo per la retrotrascrizione. Uno dei nucleotidi è marcato con un colorante fluorescente che può essere:

Cy3, che emette fluorescenza di colore verde

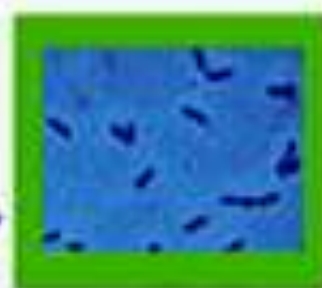
Cy5, che emette fluorescenza di colore rosso.

- I due diversi fluorocromi vengono usati per marcare RNA provenienti da due campioni diversi. In questo modo su uno stesso vetrino si possono saggiare due diverse popolazioni di mRNA e si possono confrontare una con l'altra



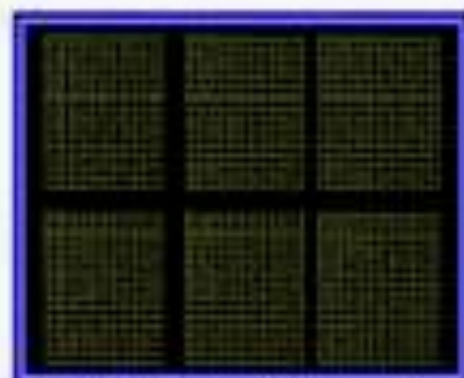
Control

Test



Isolate mRNA
& label cDNA

Isolate mRNA
& label cDNA



cDNA hybridised to microarray



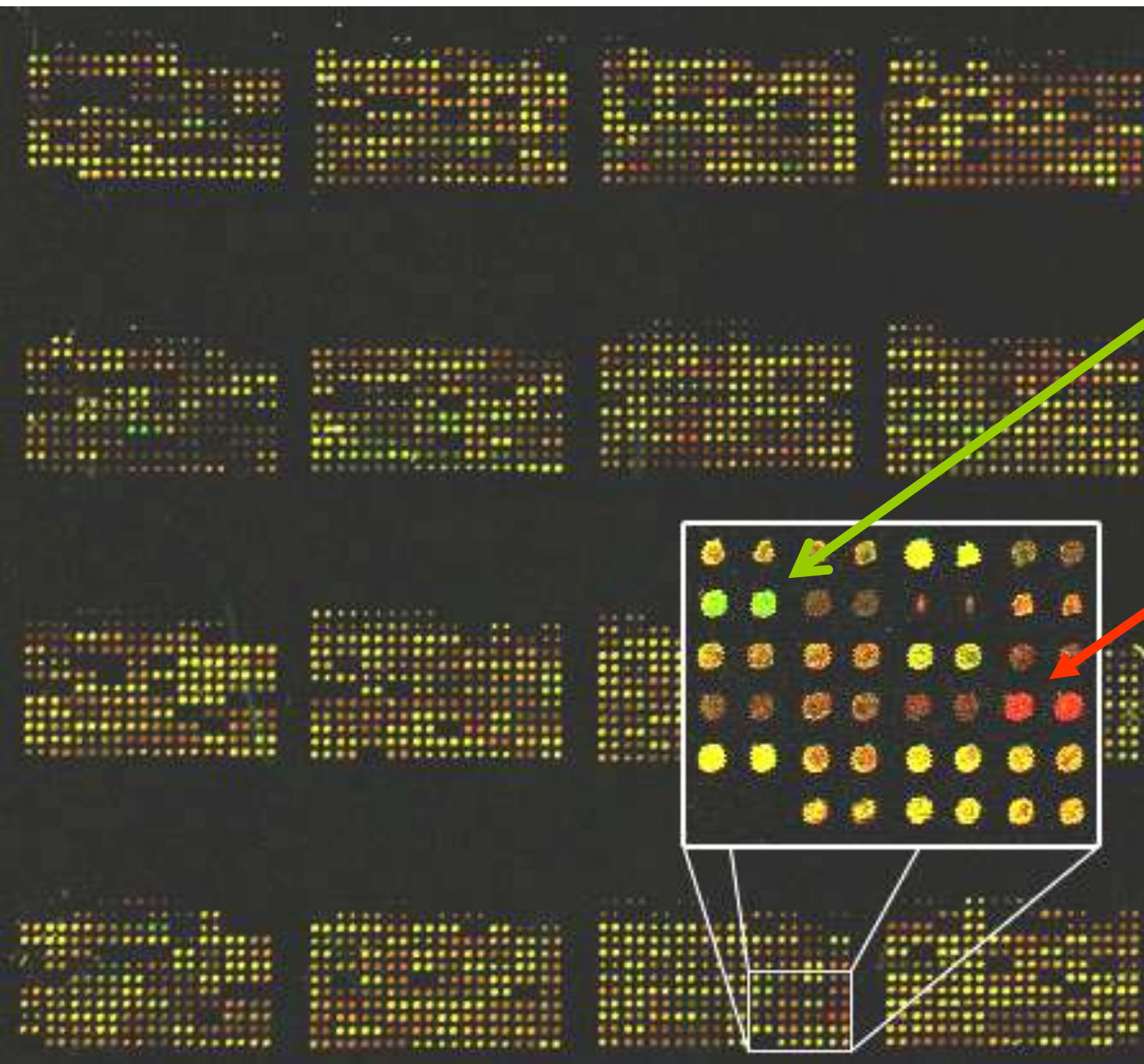
cDNA binds to
corresponding gene

Expressed only
in control

Expressed in
both conditions

Gene
induced

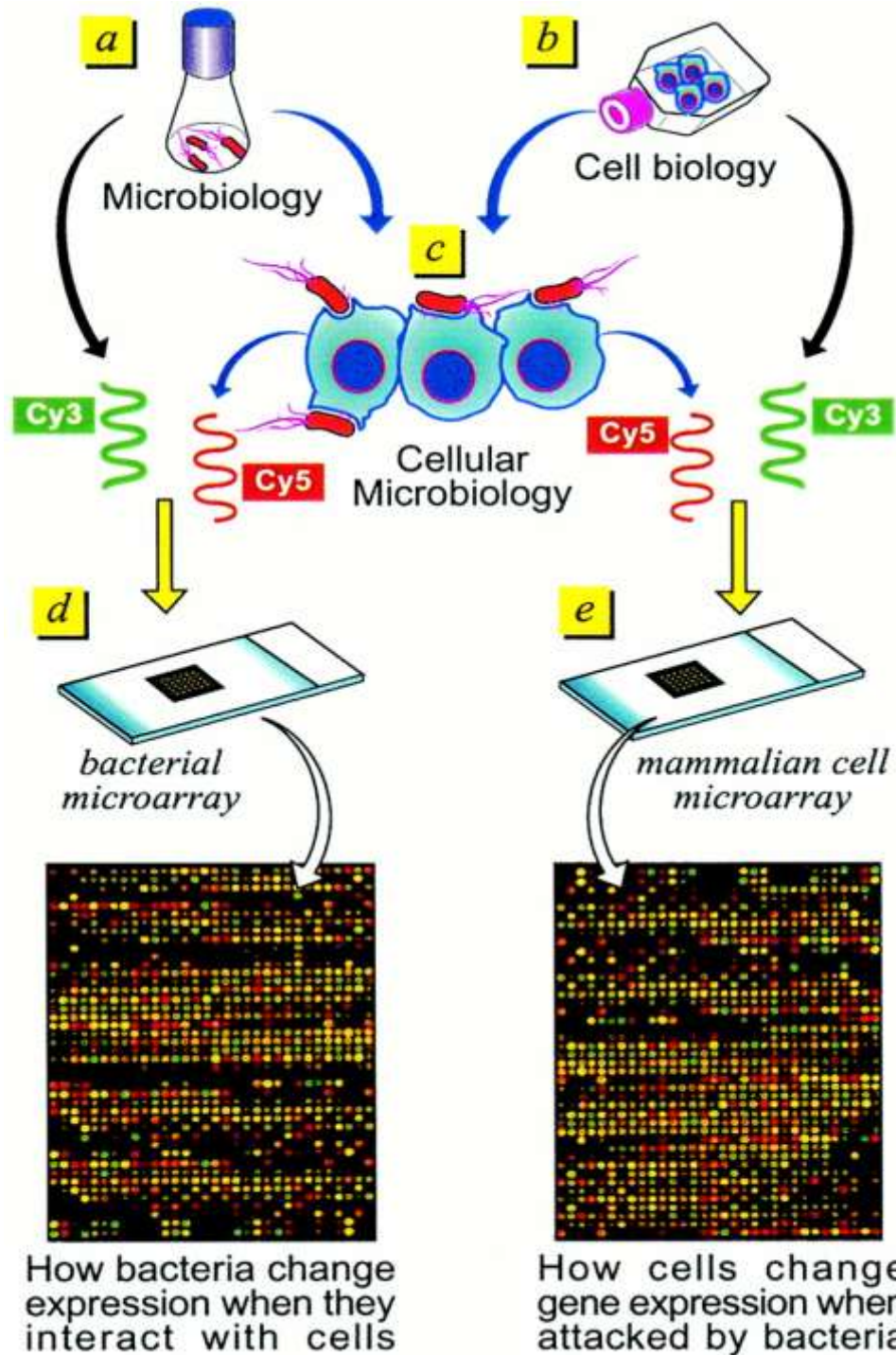
Not
expressed



Gene D
sovraespresso in
tessuti normali

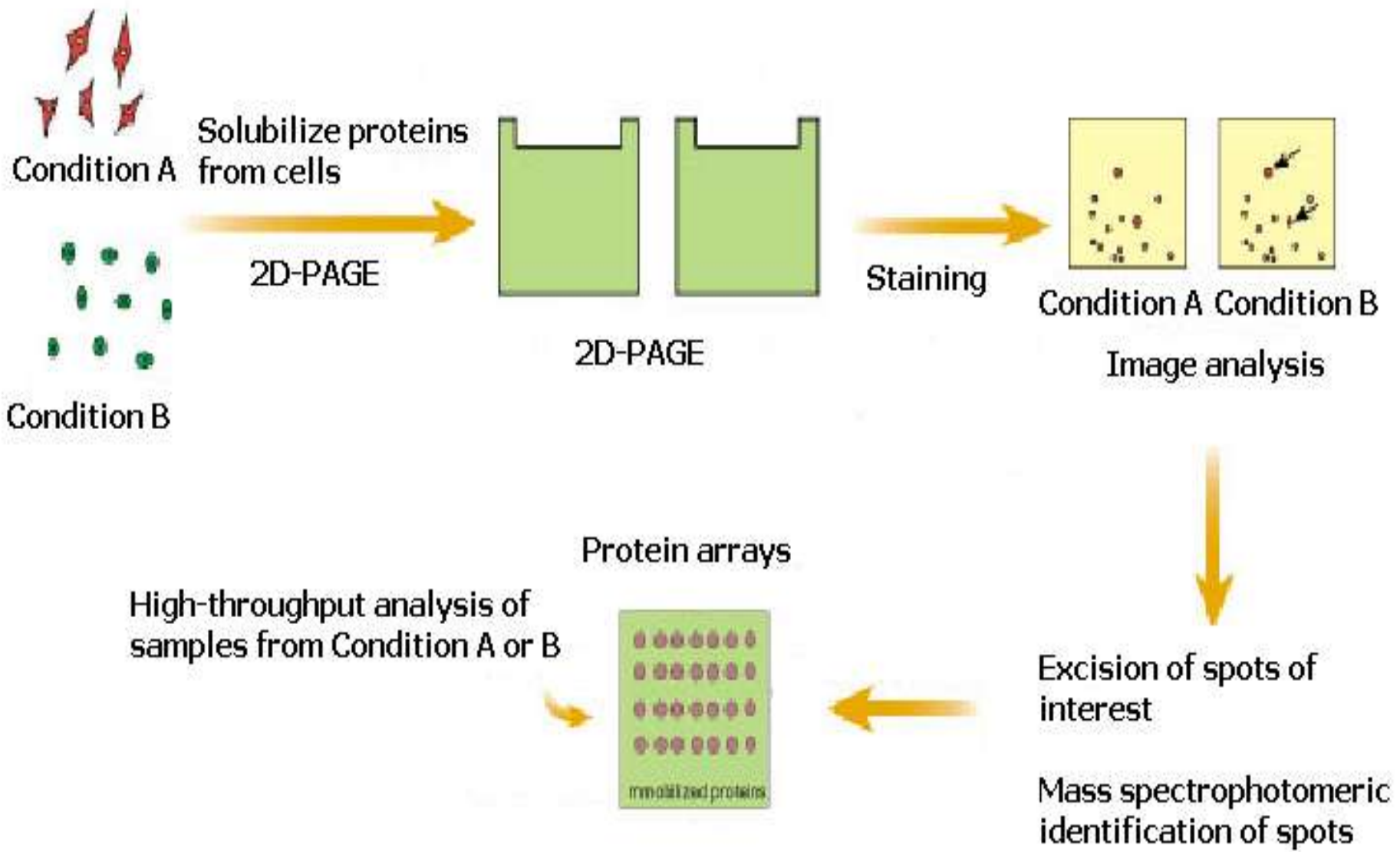
Gene E
sovraespresso
nei tumori

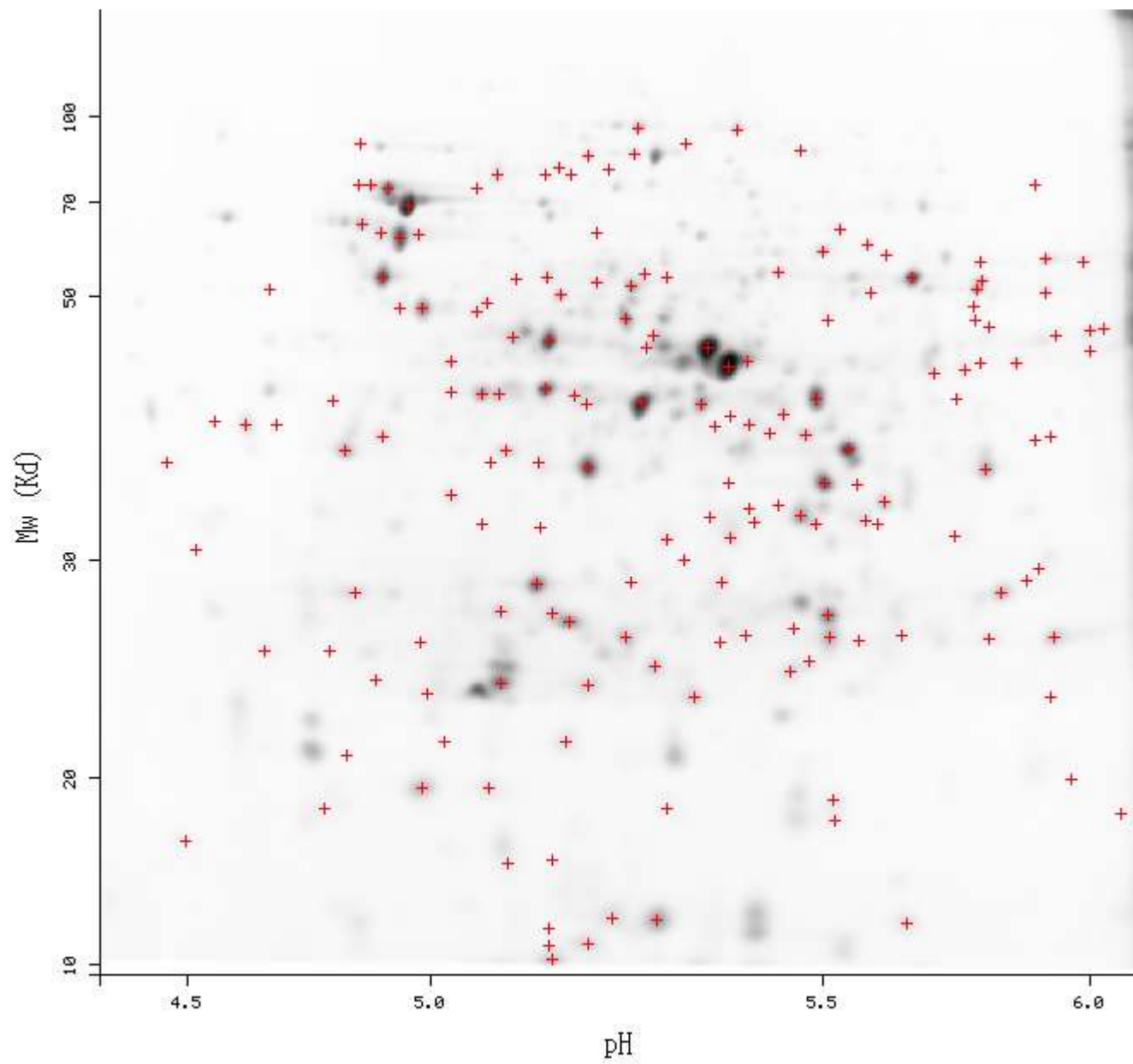
-



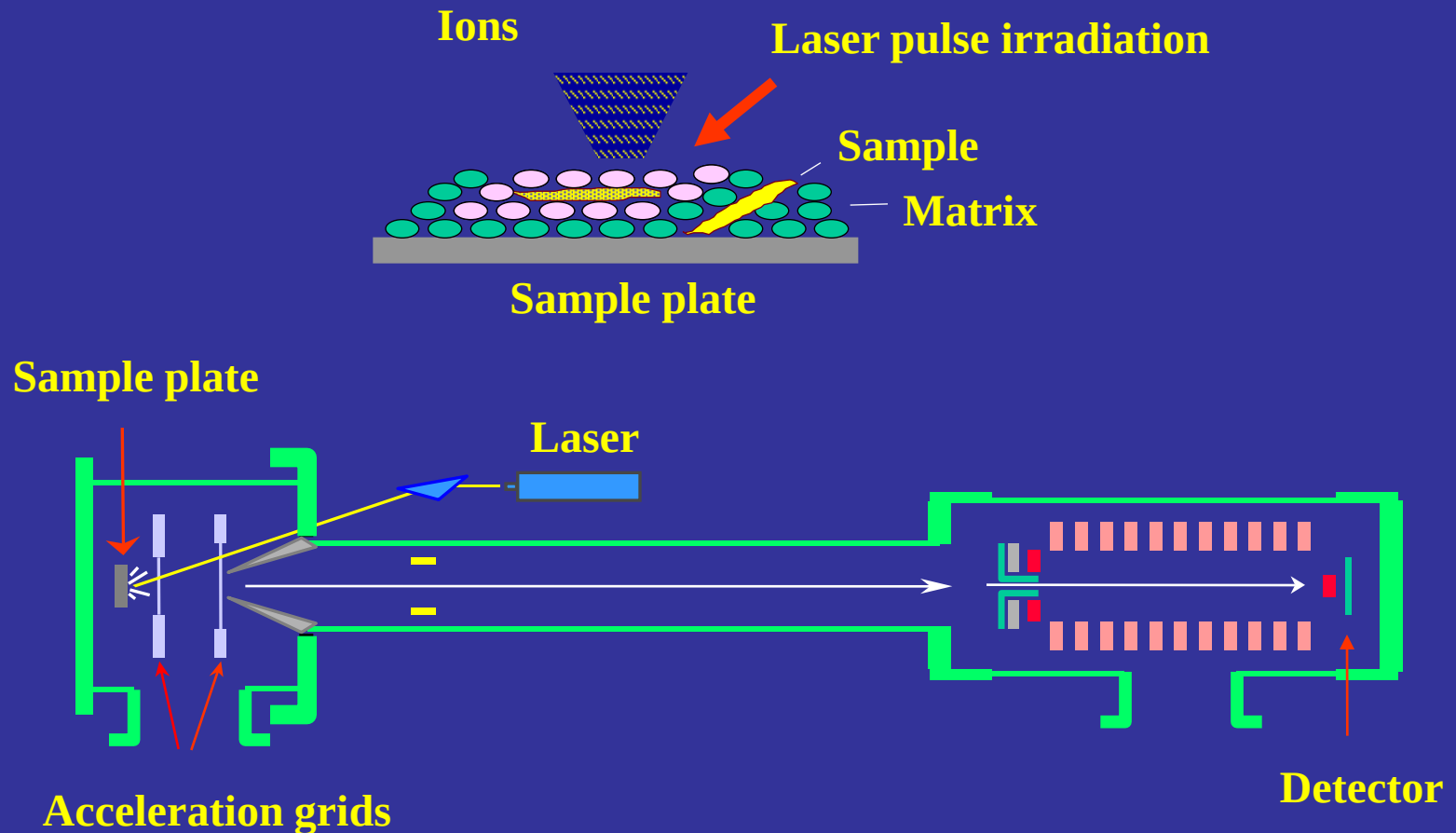
Il proteoma



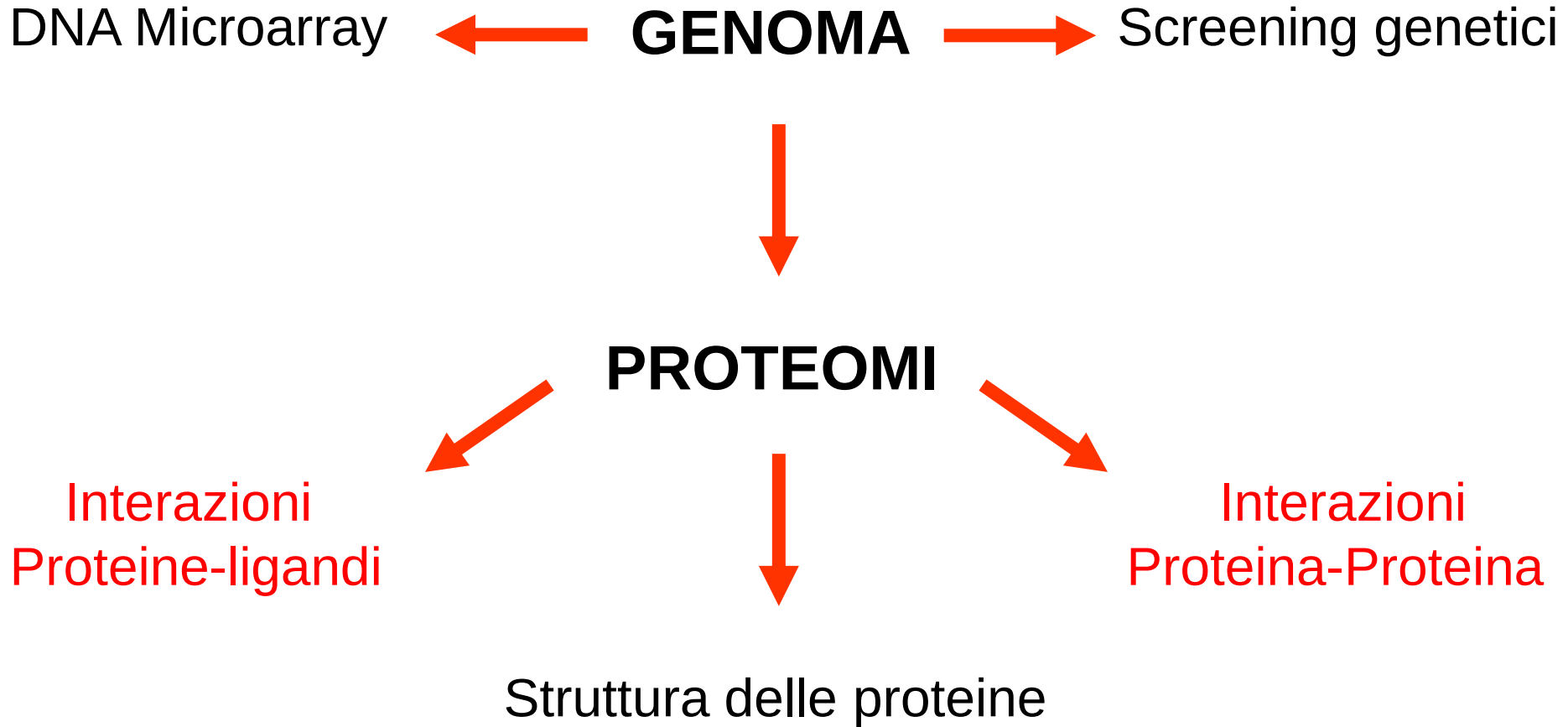




Matrix-Assisted Laser Desorption Time-of-Flight Mass Spectrometry (MALDI-TOF)



Al di là dei progetti di sequenziamento



L'ERA POST-GENOMICA

La **proteomica funzionale** utilizza diverse tecnologie complementari

- DNA Microarray

Utili per ottenere un profilo di trascrizione dell'intero genoma

- Interazioni Proteina-Ligando

- Per scoprire inibitori delle proteine
- Per scoprirne le funzioni

- Interazioni Proteina-Proteina

- Per identificare la rete di interazioni regolative
- Per scoprirne le funzioni

Gli array di proteine

Legame di piccole molecole

Modificazioni post-traduzionali

Interazioni proteina-proteina

Interazioni proteina DNA

Saggi enzimatici

Mappatura degli epitopi

