

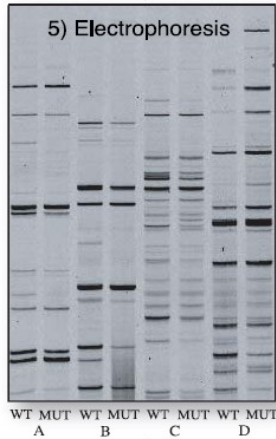
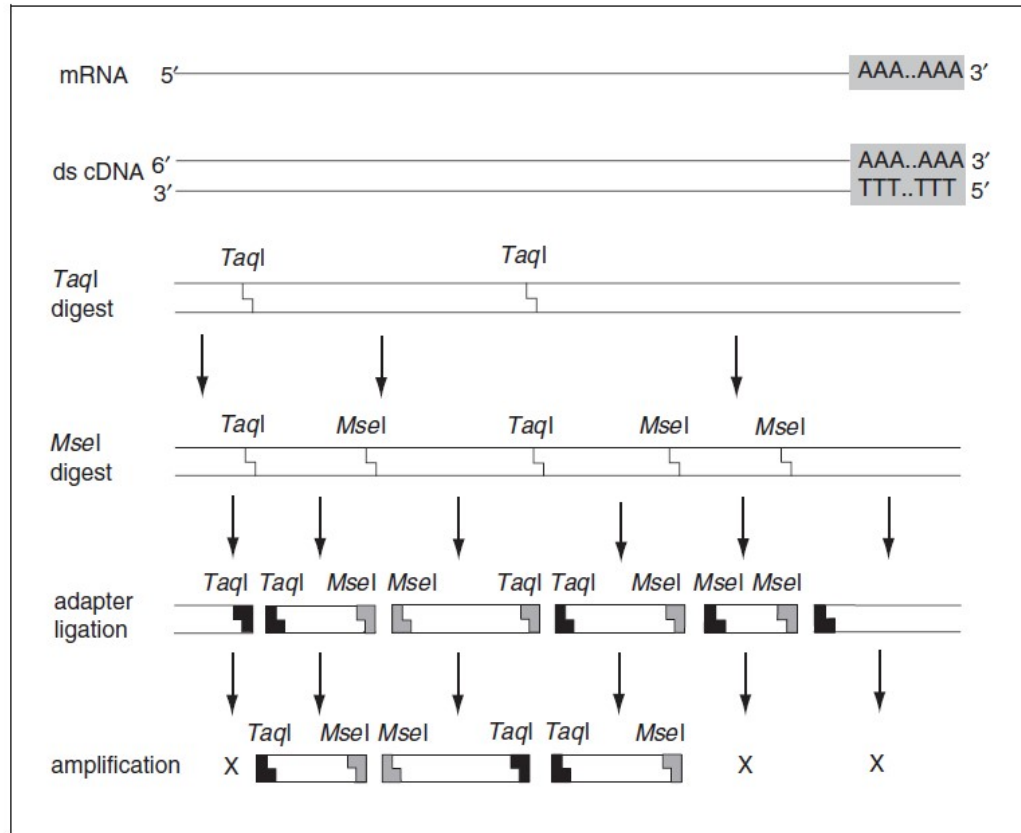
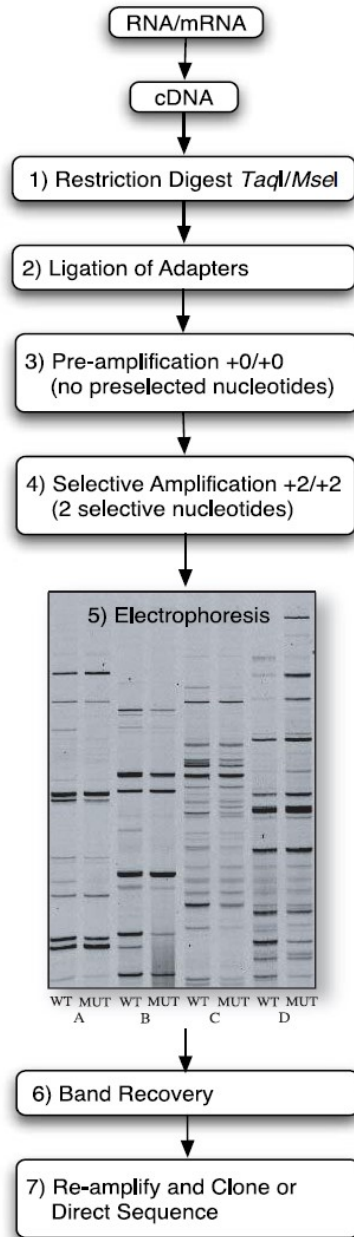
# ANALISI DI ESPRESSIONE SU LARGA SCALA

# ANALISI DI ESPRESSIONE SU LARGA SCALA

Studio dell'espressione di un grande numero di geni (migliaia) allo stesso tempo

Non richiede necessariamente il sequenziamento del genoma dell'organismo di interesse

# cDNA-AFLP



*EcoRI*-primer +0: 5'-GACTGCGTACCAATTC-3'

*EcoRI*-primer +1: 5'-GACTGCGTACCAATTCA-3'

*EcoRI*-primers +2: 5'-GACTGCGTACCAATTCAN-3'

*EcoRI*-primers +3: 5'-GACTGCGTACCAATTCANN-3'

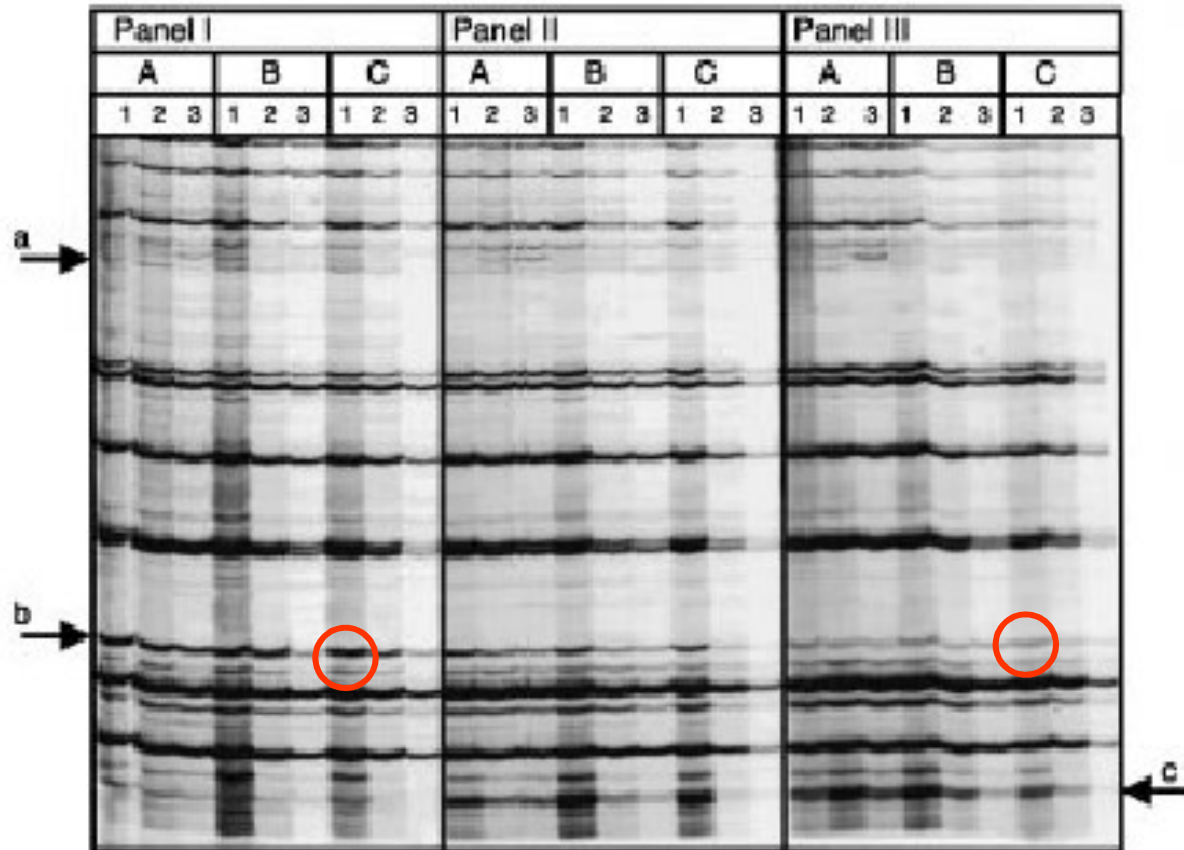
*MseI*-primer +1: 5'-GATGAGTCCTGAGTAAC-3'

*MseI*-primers +2: 5'-GATGAGTCCTGAGTAACN-3'

*MseI*-primers +3: 5'-GATGAGTCCTGAGTAACNN-3'

# cDNA-AFLP

## Fingerprinting su gel di poliacrilammide



A = 10 cicli; B = 15 cicli; C = 20 cicli

1 = diluizione 1:10 del template, 2= dil. 1:50; 3= dil. 1:100

## cDNA-AFLP

- Permette il riconoscimento di mRNA differenzialmente espressi, senza forti restrizioni riguardo il sistema sperimentale
- Essendo basata sulla PCR, mostra una elevata sensibilità.
- Rispetto a tecniche precedenti come il "differential display" mostra una maggiore specificità nei trascritti amplificati e riproducibilità dei profili di espressione (condizioni di PCR più stringenti, permesse dall'impiego di "primer" specifici)
- Costa poco!

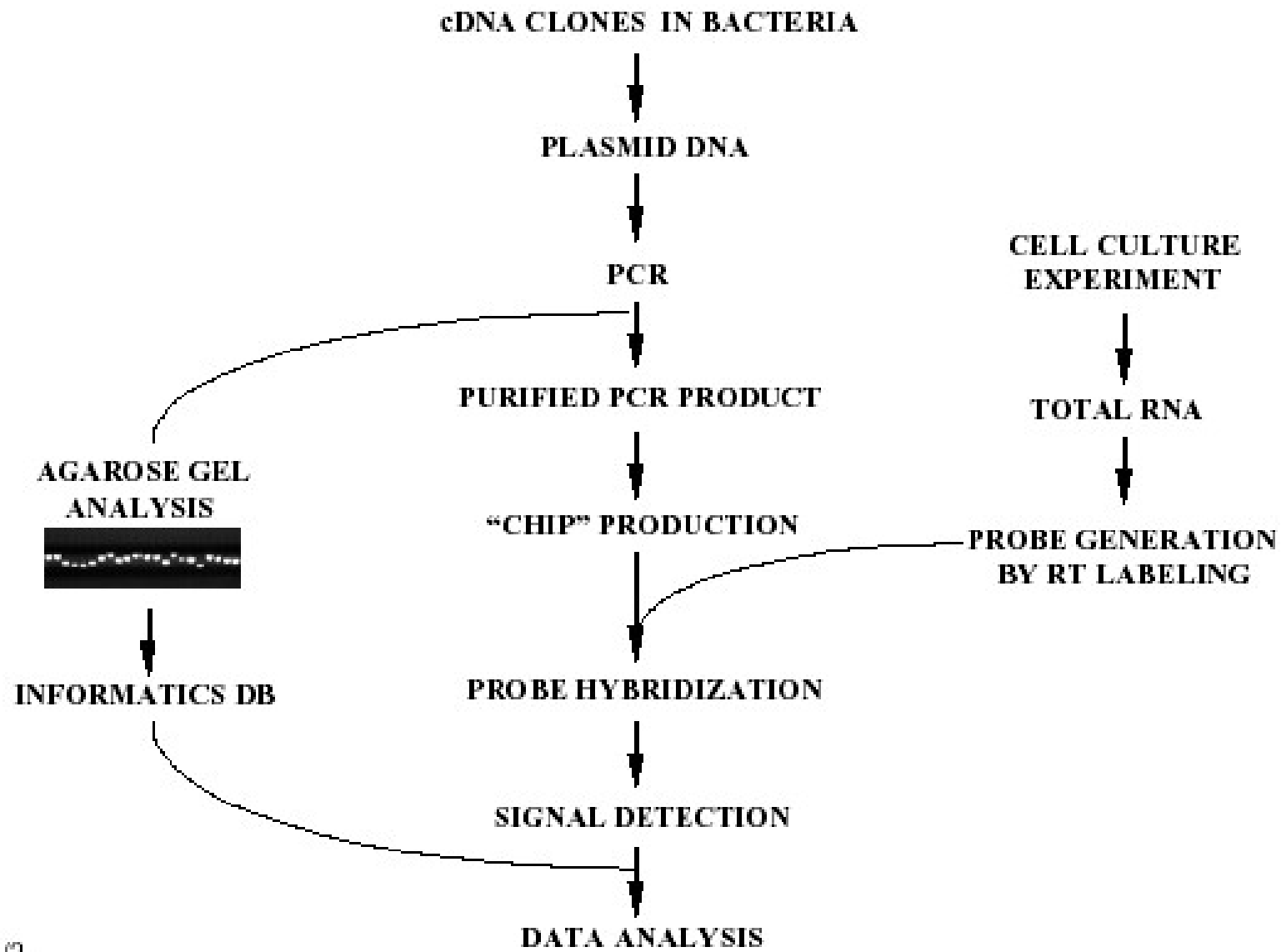
# ANALISI DI ESPRESSIONE CON MICROARRAYS

- Sequenze di DNA (**sonde**) specifiche per un alto numero di geni immobilizzate in una superficie molto piccola
- Ibridazione di RNA marcati (**target**) sull'array rilevata tramite fluorescenza
- La scansione dell'array con un laser fornisce valori di **fluorescenza** proporzionali alla quantità di RNA marcato corrispondente a ciascuna sonda

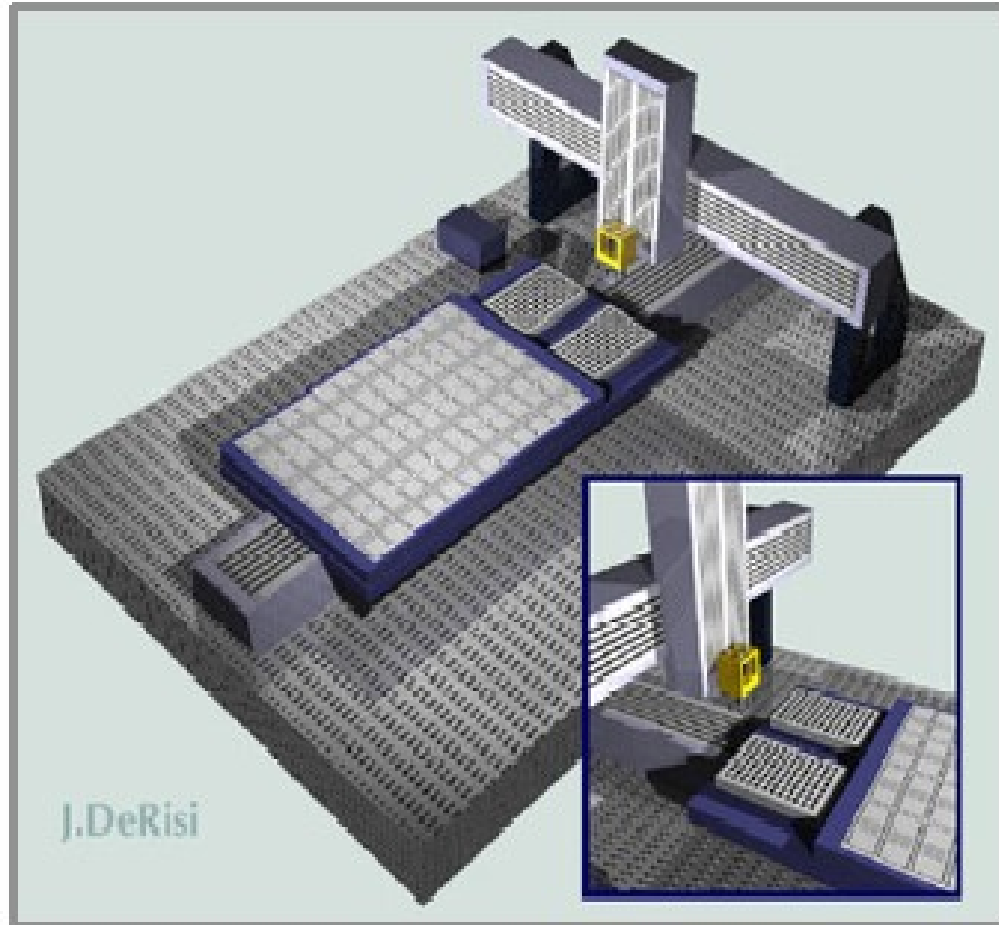
# Microarrays a cDNA

- Analisi comparativa (due campioni allo stesso tempo)
- Problemi di cross-ibridazione
- Costo più contenuto
- Si può fare "in casa"

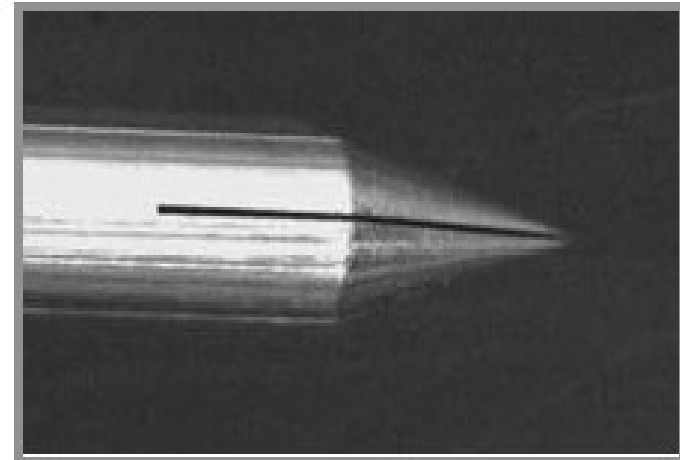




## X-Y-Z Robot

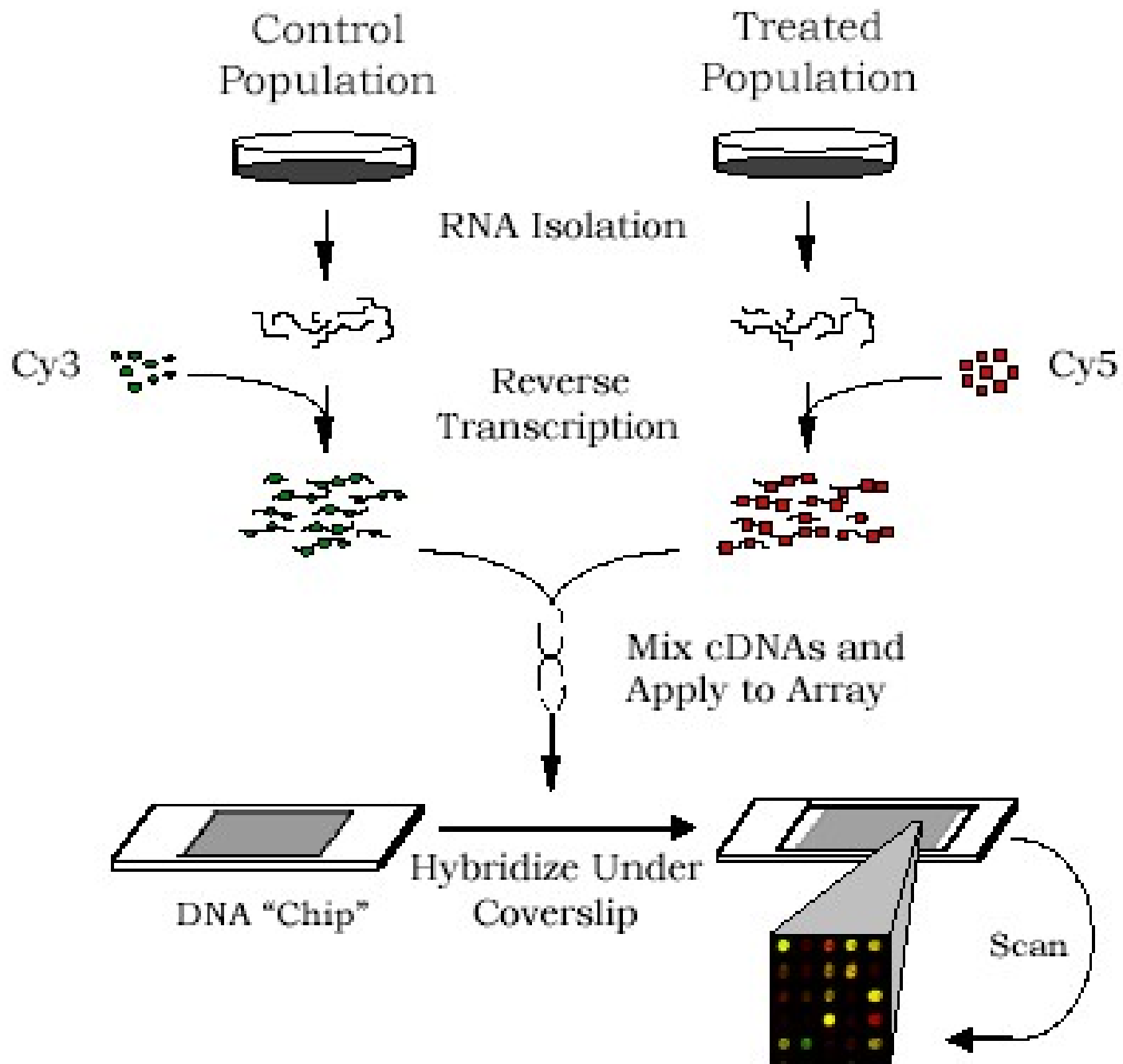


## Quill-pen Spotting Tip ("fountain-pen" design)

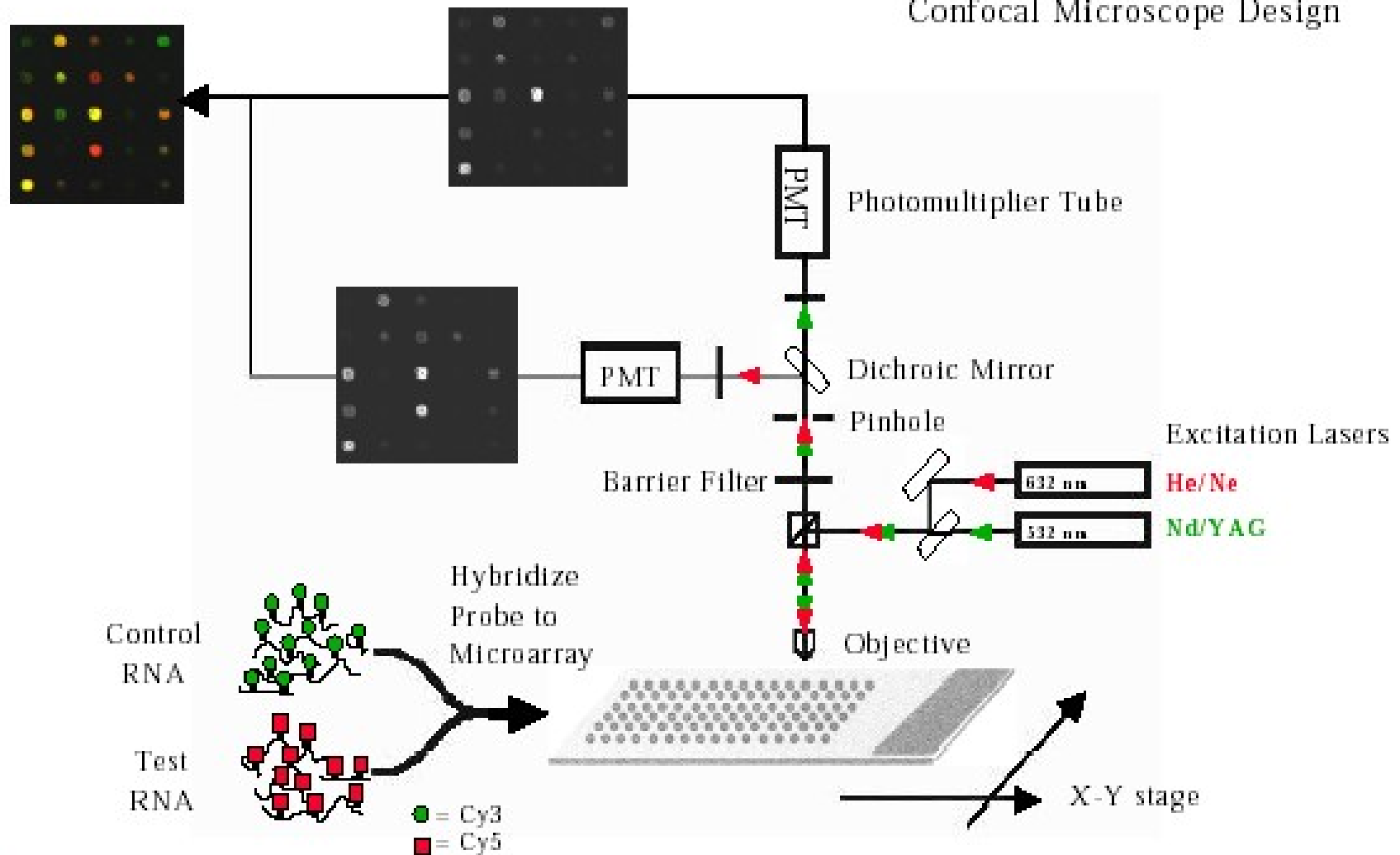


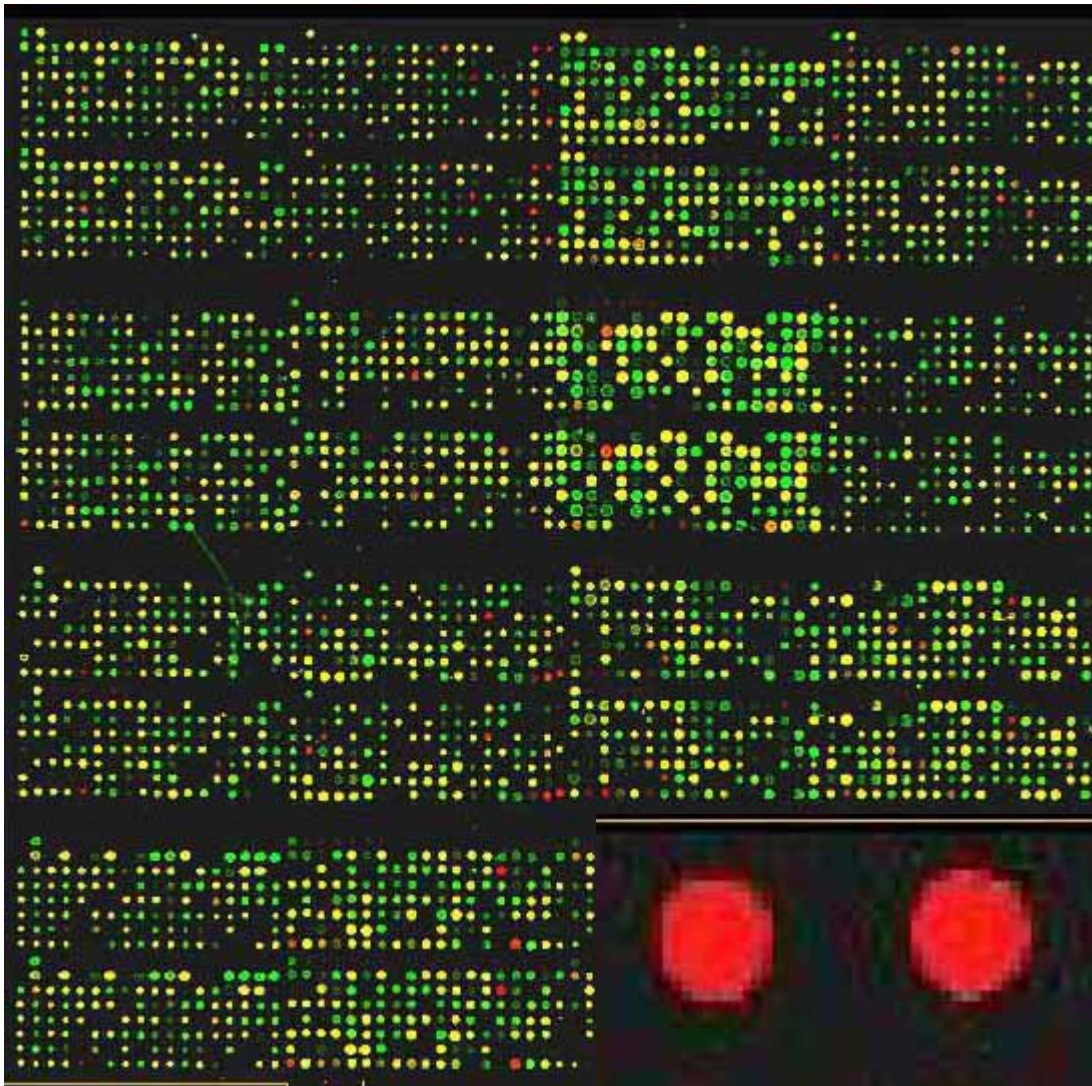
1/8

Image source: <http://cmgm.Stanford.edu/pbrown/>



# Confocal Microscope Design



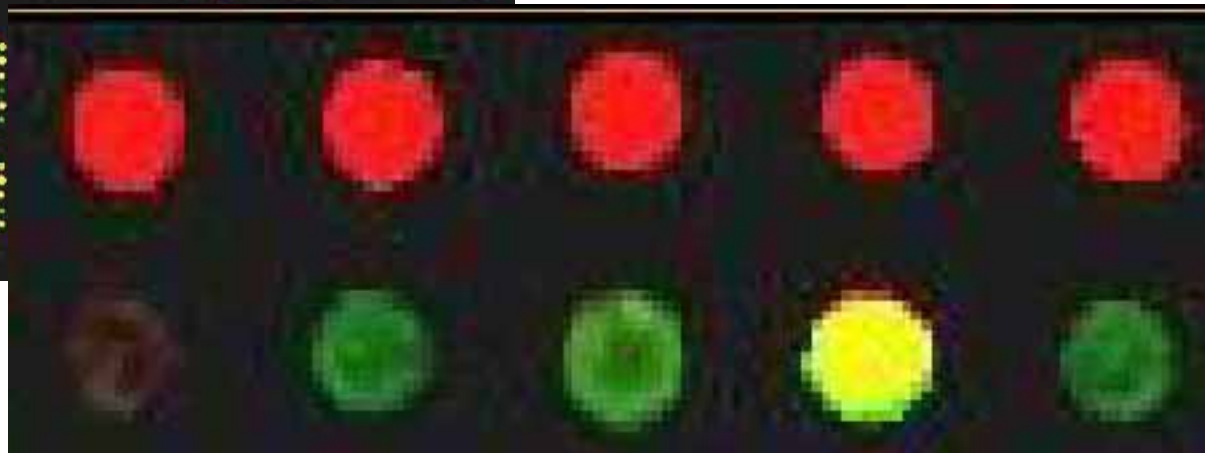


Spot rosso:  $Cy5 > Cy3$

Spot verde:  $Cy3 > Cy5$

Spot giallo:  $Cy3 = Cy5$

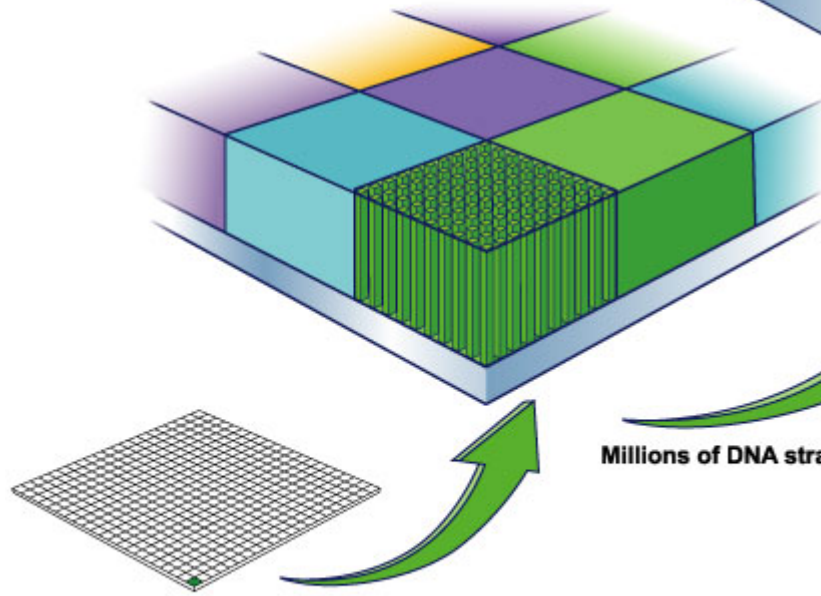
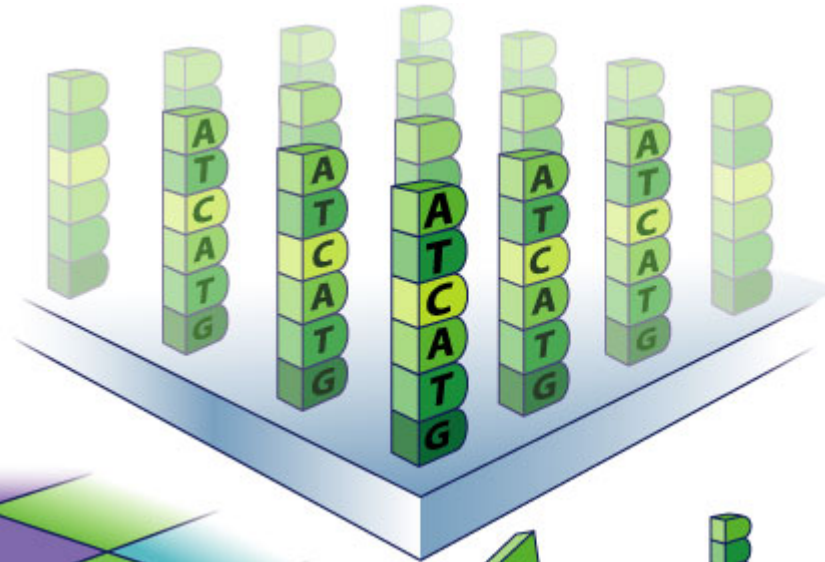
Spot nero: no ibridazione



## Microarrays ad **oligonucleotidi** (GeneChips)

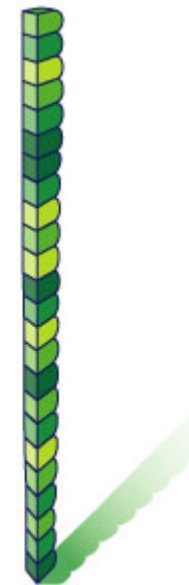
- Analisi di un campione alla volta
- Forniscono valori assoluti di espressione
- Alta specificità
- Costo elevato

# Microarray ad oligonucleotidi (GeneChip)



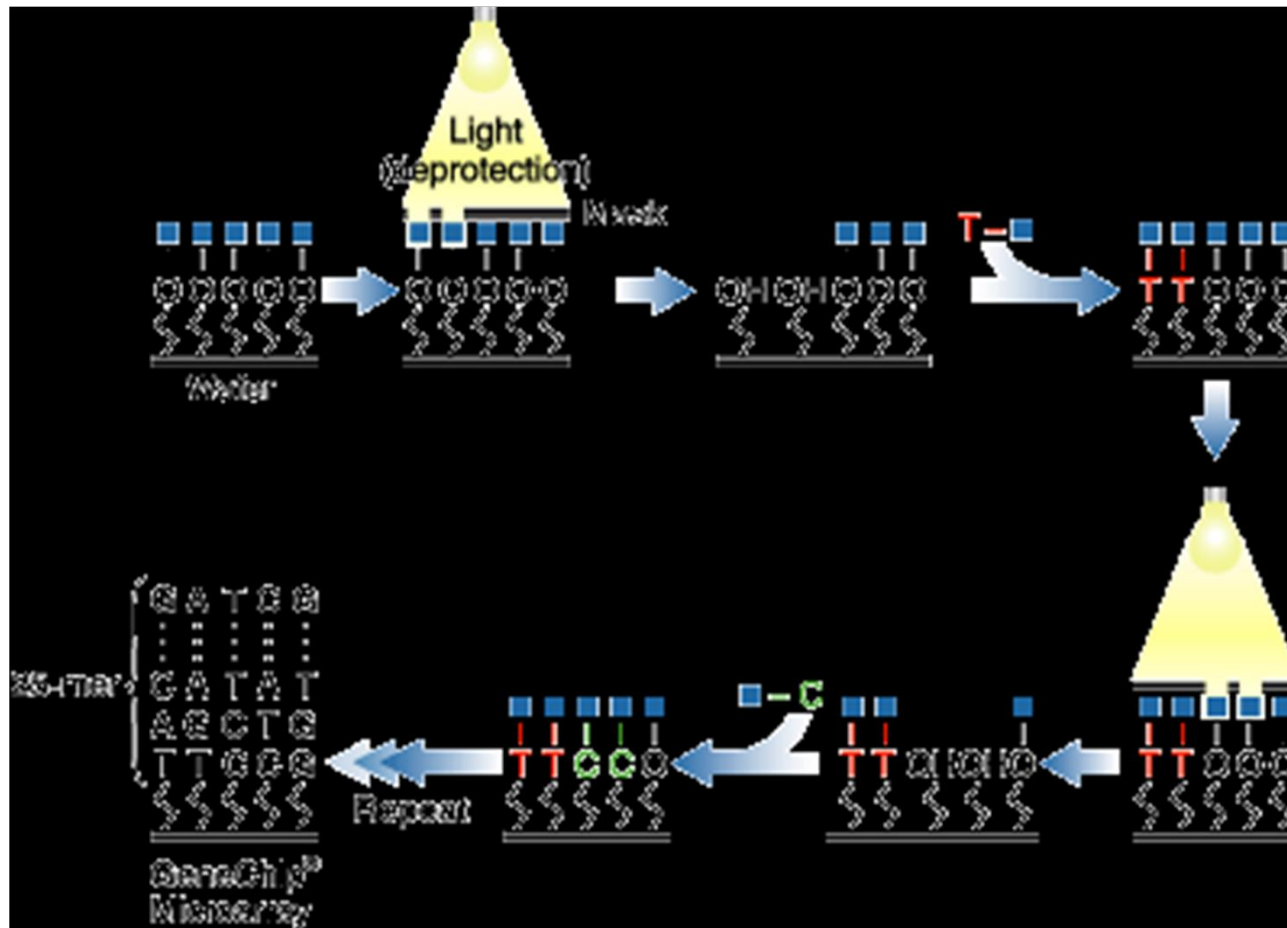
500,000 cells on each GeneChip™ array

Millions of DNA strands built up in each cell



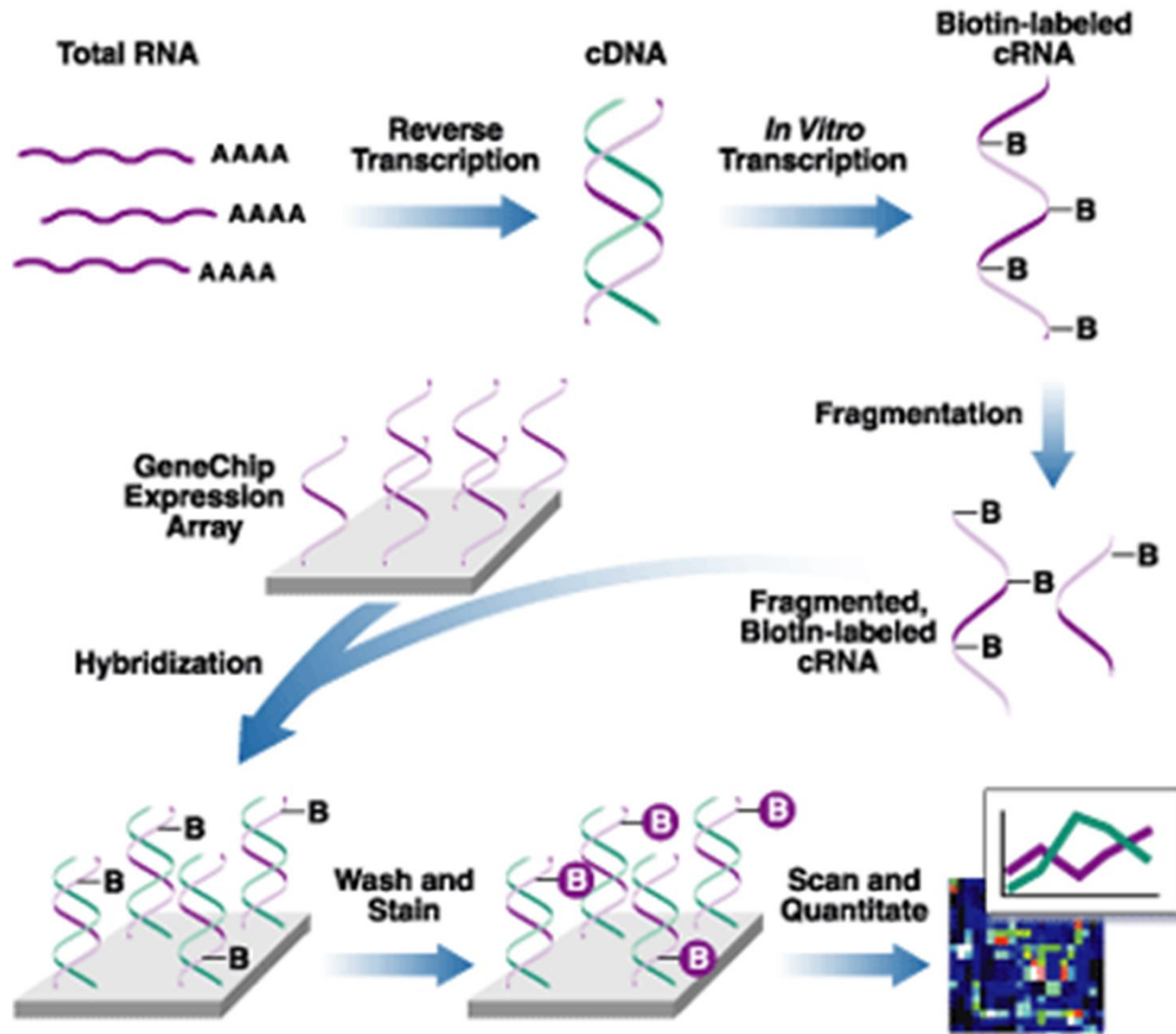
Actual strand = 25 base pairs

# Processo litografico per la fabbricazione di GeneChips della Affymetrix





# Saggio di espressione usando GeneChips

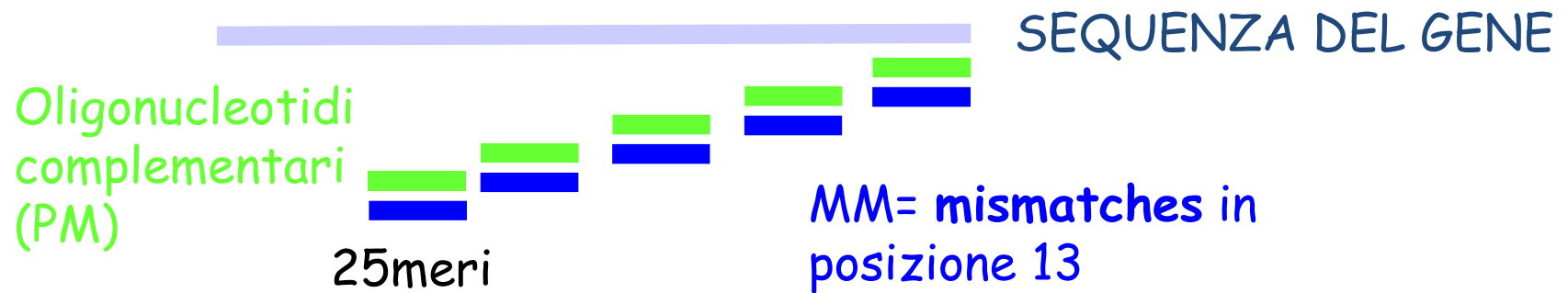


- Ogni gene sul chip è rappresentato da 25-meri
- 16-20 sequenze selezionate per rappresentare ciascun gene
- Ogni 25-mero è chiamato **perfect match (PM)** (appaiamento perfetto)
- Ogni PM è accompagnato da un **mismatch probe (MM)** (sonda ad appaiamento imperfetto)

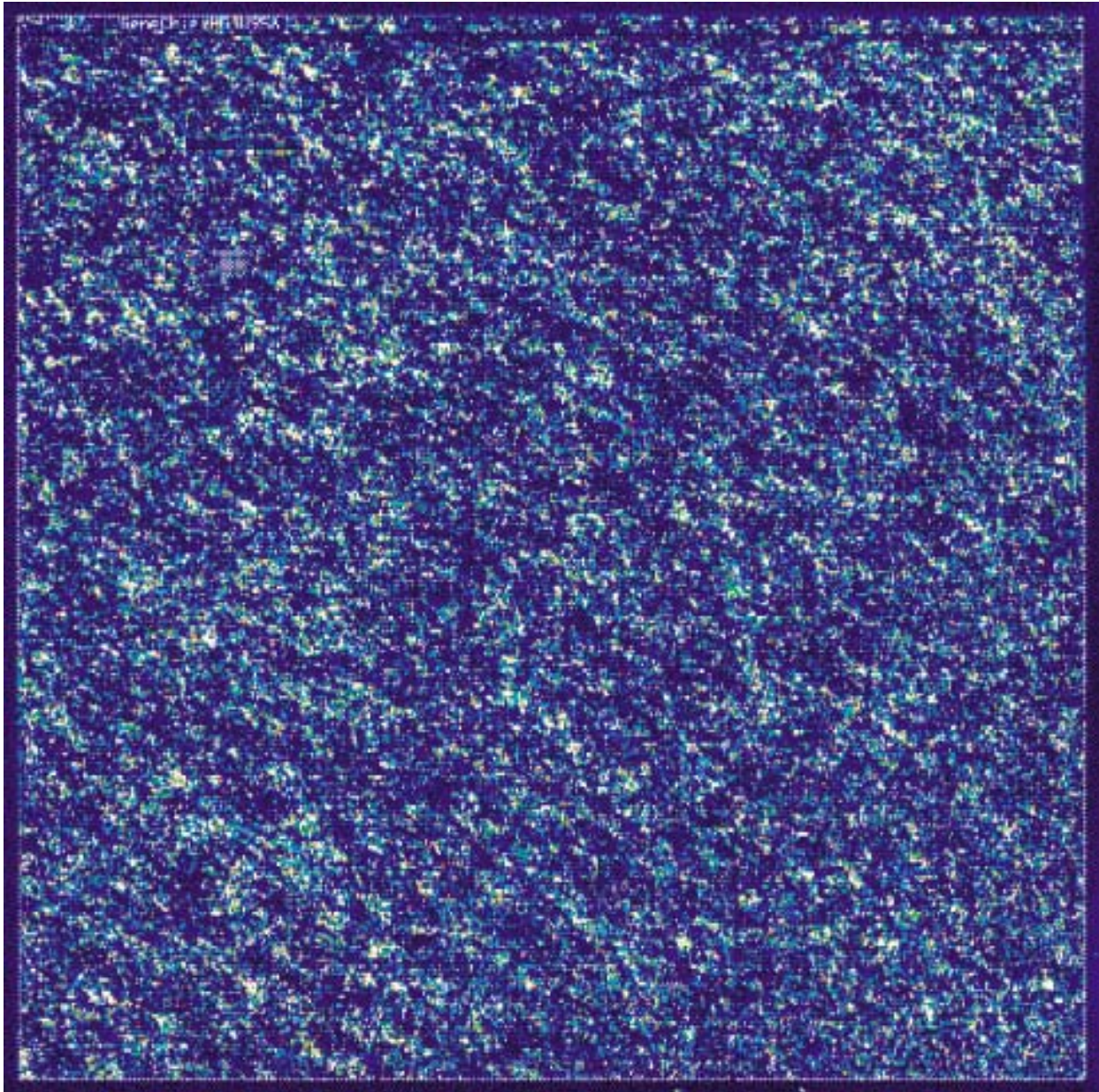
Probe on chip { *PM* : *ATCGACTGATGCATGCATCCATCAT*  
*MM* : *ATCGACTGATGC***C***TGCATCCATCAT*

Sample Probe

*TAGCTGACTACGTACGTAGGTAGTA*



1. MM è più sensibile a cambiamenti nella concentrazione del target
2. MM serve come controllo per ibridazione aspecifica



# SOFTWARE PER ANALISI DI DATI DA MICROARRAYS

- 1) Per disegnare sonde (oligos, primers per PCR...)
- 2) Per gestire immagini e assegnare ogni segnale ad un gene, valutare background, ecc...
- 3) Per analisi statistica (es. dopo analisi repliche esperimenti)
- 4) Per filtrare, raggruppare e dividere i dati ottenuti in modo da ottenere informazioni utili
- 5) Per estrarre informazioni su particolari geni da database pubblici e non

Sort

R/G Norm

Descending

Submit

Display

- % > B635+2SD
- B532 SD
- B635 SD
- Blast match
- Category

Show Controls

Hide Rank

Line

1 to 50

Filters

Spot Flag = 0

Ch1 Net > 150

Ch2 Net > 150

Name cci 0

Gene cci 0

% > B532+1SD = 0

Rank	NAME	GENE	LINK	CH1D	CH2DN	RAT1N	RAT2N	MRAT	REGR	CORR	FLAG	CAT	DESC
1	YBR072W	HSP26	<a href="#">SGD YPD</a>	540	1594	0.34	2.95	3.03	6.03	0.94		hsp	heat shock protein
2	YER106W		<a href="#">SGD YPD</a>	16	47	0.34	2.94	2.01	1.43	0.30			hypothetical protein
3	YHL028W		<a href="#">SGD YPD</a>	855	2477	0.35	2.90	3.10	5.90	0.89			similarity to mucin and other ser-thr rich protein
4	YGR139W		<a href="#">SGD YPD</a>	65	174	0.37	2.68	2.92	3.65	0.64			questionable ORF
5	YIL119C	RPI1	<a href="#">SGD YPD</a>	385	975	0.39	2.53	2.69	5.16	0.90			negative regulator of ras-cAMP pathway
6	YOR186W		<a href="#">SGD YPD</a>	49	119	0.41	2.43	2.26	2.16	0.33			hypothetical protein
7	YDL181W	INH1	<a href="#">SGD YPD</a>	342	756	0.45	2.21	2.42	3.24	0.70			inhibitor of mitochondrial ATPase
8	YDL021W	GPM2	<a href="#">SGD YPD</a>	560	1216	0.46	2.17	2.34	3.98	0.89			homology to phosphoglycerate mutase
9	YLR312C		<a href="#">SGD YPD</a>	64	139	0.46	2.17	3.39	1.75	0.36			hypothetical protein
10	YFR054C		<a href="#">SGD YPD</a>	76	164	0.46	2.16	3.53	0.94	0.21			hypothetical protein
11	YOR049C		<a href="#">SGD YPD</a>	1718	3613	0.48	2.10	2.14	4.44	0.93			hypothetical protein
12	YOL104C		<a href="#">SGD YPD</a>	218	456	0.48	2.09	2.16	3.99	0.90			hypothetical protein
13	YBR297W	MAL33	<a href="#">SGD YPD</a>	649	1356	0.48	2.09	2.19	3.87	0.86		hwt	maltose fermentation regulatory protein
14	YGL140C		<a href="#">SGD YPD</a>	48	99	0.48	2.06	2.27	2.17	0.73			hypothetical protein
15	YHR046C		<a href="#">SGD YPD</a>	61	125	0.49	2.05	1.98	2.83	0.69			similarity to myo-inositol-1(or 4)-monophosphatase
16	YIL111W	COX5B	<a href="#">SGD YPD</a>	764	1569	0.49	2.05	2.12	4.40	0.96		ccc	cytochrome-c oxidase chain Vb
17	YKL097W-A	CWP2	<a href="#">SGD YPD</a>	1392	2856	0.49	2.05	2.27	4.16	0.81		srp	cell wall mannoprotein
18	YER188W		<a href="#">SGD YPD</a>	304	600	0.51	1.97	1.96	3.90	0.93			hypothetical protein
19	YPL280W		<a href="#">SGD YPD</a>	33	65	0.51	1.97	2.55	1.59	0.43			"strong homology to hypothetical proteins YOR391c, YMR322c and YDR533c"
20	YLR012C		<a href="#">SGD YPD</a>	36	71	0.51	1.97	2.51	1.85	0.64			hypothetical protein
21	YDR278C		<a href="#">SGD YPD</a>	157	305	0.51	1.94	2.32	2.48	0.48			hypothetical protein
22	YIR042C		<a href="#">SGD YPD</a>	425	823	0.52	1.94	1.97	4.00	0.94			hypothetical protein
23	YCR102C		<a href="#">SGD YPD</a>	217	418	0.52	1.93	1.95	3.65	0.84			similarity to C.caerbonum toxD gene
24	YIL015C-A		<a href="#">SGD YPD</a>	903	1720	0.53	1.90	1.93	3.65	0.86			
25	YER135C		<a href="#">SGD YPD</a>	52	98	0.53	1.88	2.24	1.60	0.35			hypothetical protein
26	YPL166W		<a href="#">SGD YPD</a>	301	563	0.53	1.87	1.87	2.83	0.73			hypothetical protein
27	YBR054W	YRO2	<a href="#">SGD YPD</a>	847	1569	0.54	1.85	1.78	3.91	0.96		hsp	similarity to HSP30 heat shock protein Yro1p
28	YDR134C		<a href="#">SGD YPD</a>	10814	19924	0.54	1.84	1.85	3.92	0.97			similarity to FLC1 protein

## ANALISI DI CLUSTERS

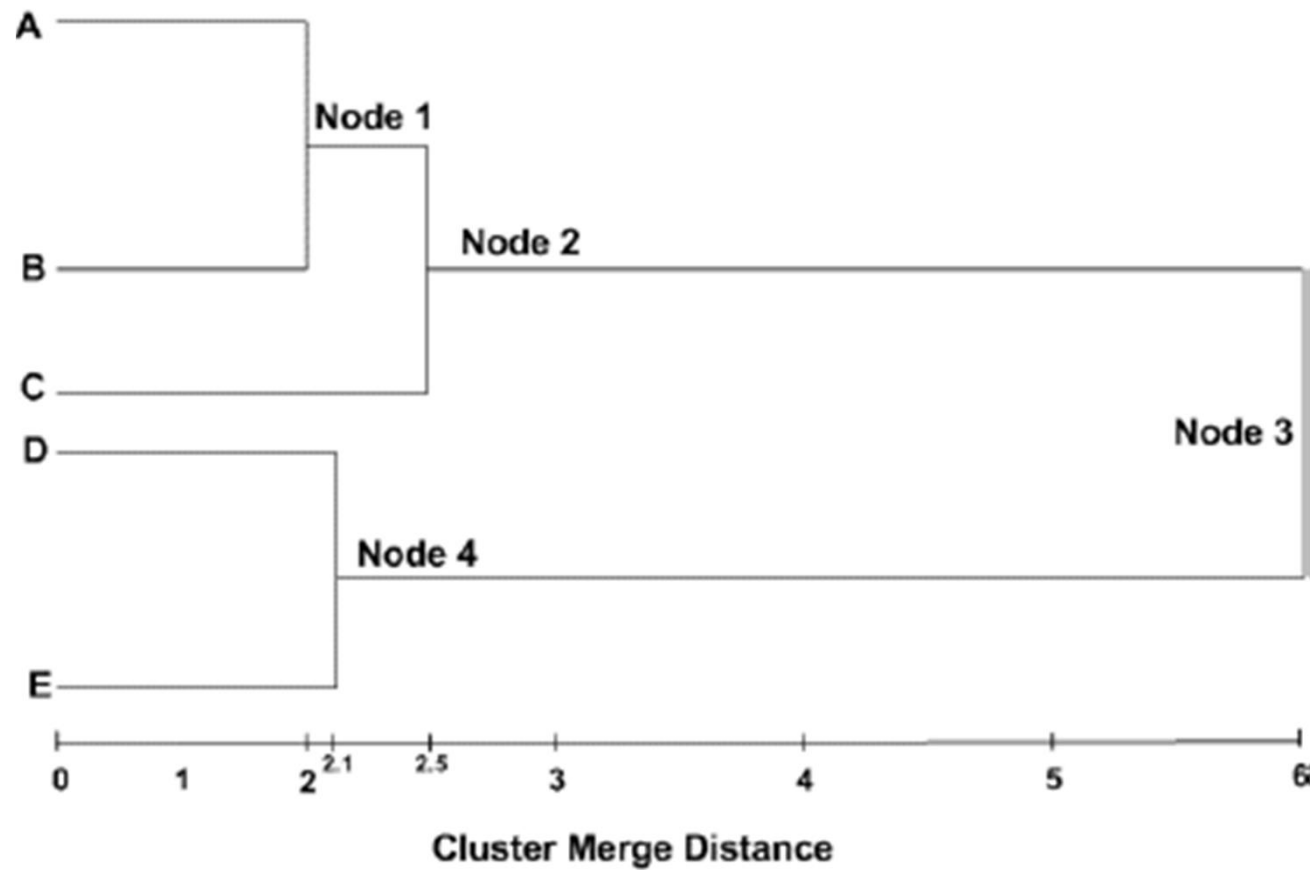
- **Clustering** = analisi statistica multivariata usata per raggruppare geni o esperimenti interi in gruppi (**clusters**) separati basati sul loro comportamento statistico.
- Serve a trovare somiglianza tra **esperimenti** (in base ai dati di espressione in tutti i geni) o tra **geni** (in base all'espressione in tutti i campioni) e **raggruppare simili geni o esperimenti**.
- Si basa su metodi matematici per trovare somiglianze in gruppi di dati.

# METODI PER RAGGRUPPARE DATI DI MICROARRAYS

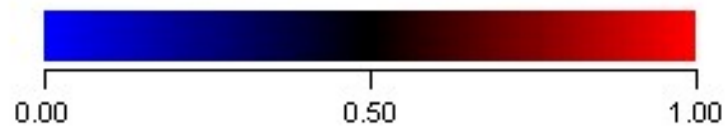
- 1) **Clustering gerarchico agglomerativo**: costruzione di un albero filogenetico, partendo da un gene e andando a distribuire gli altri in base alla loro somiglianza
- 2) **SOM** (= **Self-Organizing Maps**): basato su reti neuronali
- 3) **K-means**: raggruppa i geni in un numero predeterminato di clusters (non gerarchico)



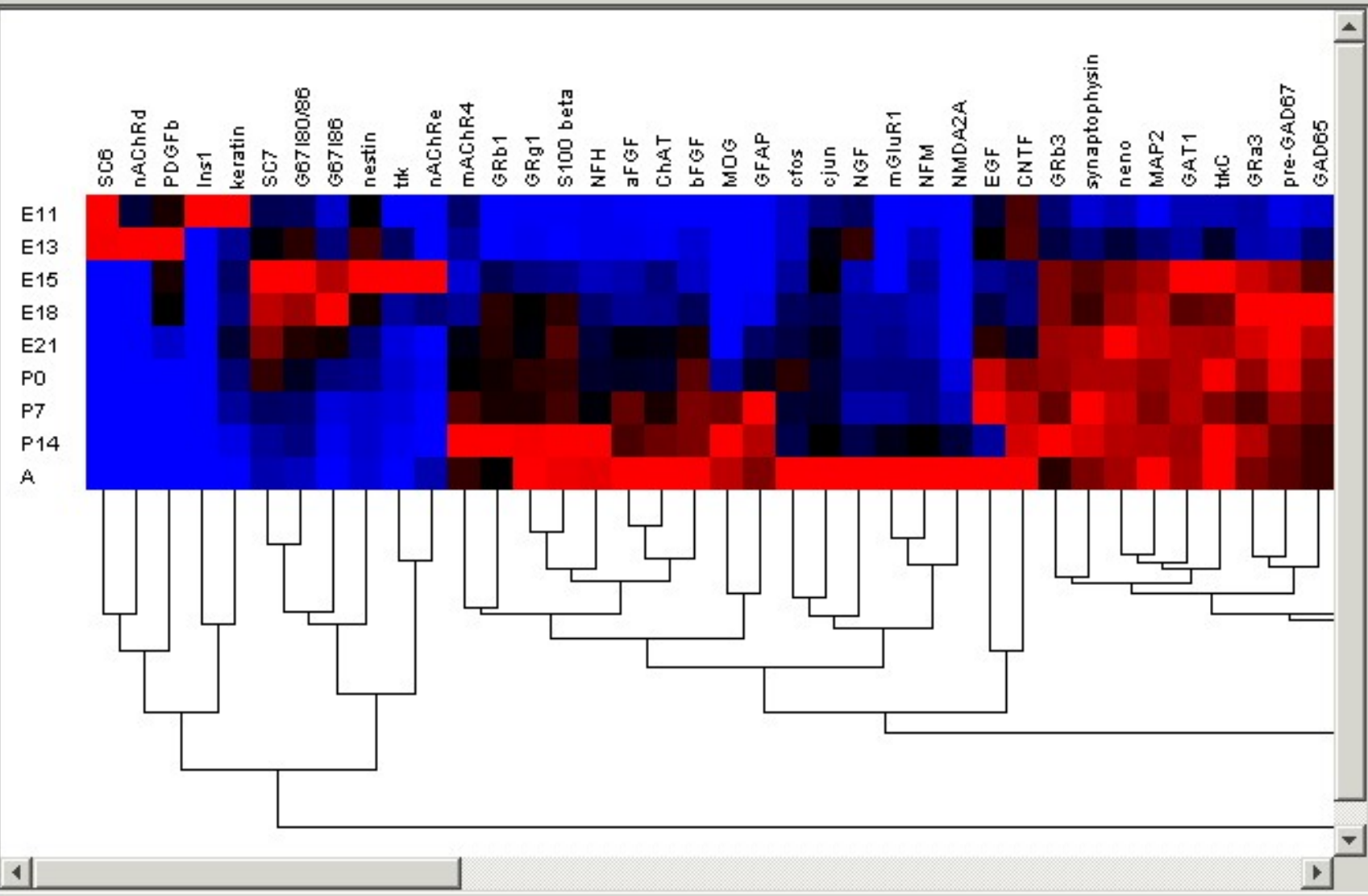
# Clustering gerarchico agglomerativo



Dendrogram Plot: Hier: genes | Euclid | average

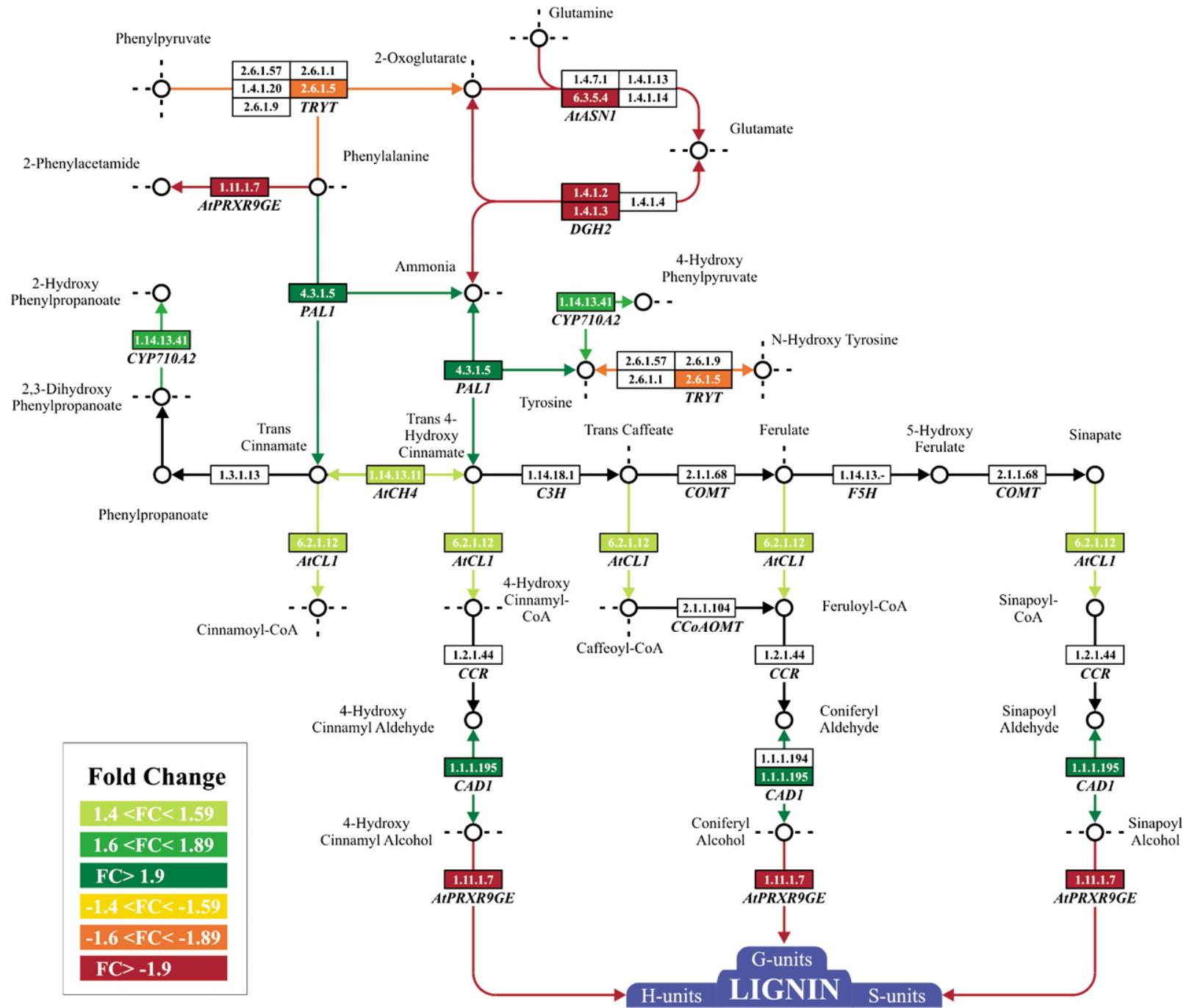


Color by: <no variables defined>



# CHE TIPO DI INFORMAZIONI DA' UN'ANALISI DEL PROFILO TRASCRIZIONALE DI UNA PIANTA?

- 1) Caratterizzazione degli effetti globali di mutazioni o espressione alterata di geni
- 2) Identificazione di geni attivati o repressi in risposta a determinati trattamenti (es. diversi patogeni)
- 3) Identificazione di sequenze regolatrici nei promotori di geni con simile pattern di espressione
- 4) Identificazione di vie metaboliche regolate da determinati trattamenti



**ANALISI DEL PROFILO TRASCRIZIONALE DI  
ARABIDOPSIS DOPO INOCULO CON PATOGENI  
VIRULENTI E AVIRULENTI**

*Pseudomonas syringae* = batterio patogeno ad ampio spettro d'ospite.

Ceppi specifici sono virulenti in specie vegetali diverse

Es.:

- *P. syringae* pv *phaseolicola* (Psp): patogeno del fagiolo
- *Pseudomonas syringae* pv. *tomato* (Pst): patogeno del pomodoro
- *P. syringae* pv *maculicola* (Psm): patogeno delle crucifere

*Ps. syringae* pv. *tomato* (**Pst**) e pv. *maculicola* (**Psm**) sono **virulenti** su *A. thaliana*



*Ps. syringae* pv. *phaseolicola* (**Psp**) non è virulento su *Arabidopsis* (**resistenza non-ospite**)

Esistono geni di resistenza di Arabidopsis che conferiscono resistenza "gene-per-gene" contro ceppi di Psm o Pst che portano specifici geni di avirulenza:

Gene R

RPS2

RPM1

Gene Avr

AvrRpt2

AvrB e AvrRpm1



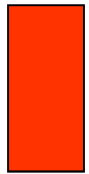
Come differiscono tra loro le risposte di difesa indotte durante interazioni compatibili ed incompatibili?

→analisi del profilo trascrizionale con microarray in piante inoculate con Psm e Pst con o senza geni avr e con Psp (Tao et al., 2003)

- Usata Affymetrix GeneChip AtGenome1 Array (circa 8,000 geni di Arabidopsis)

- Analisi effettuata a tempi diversi

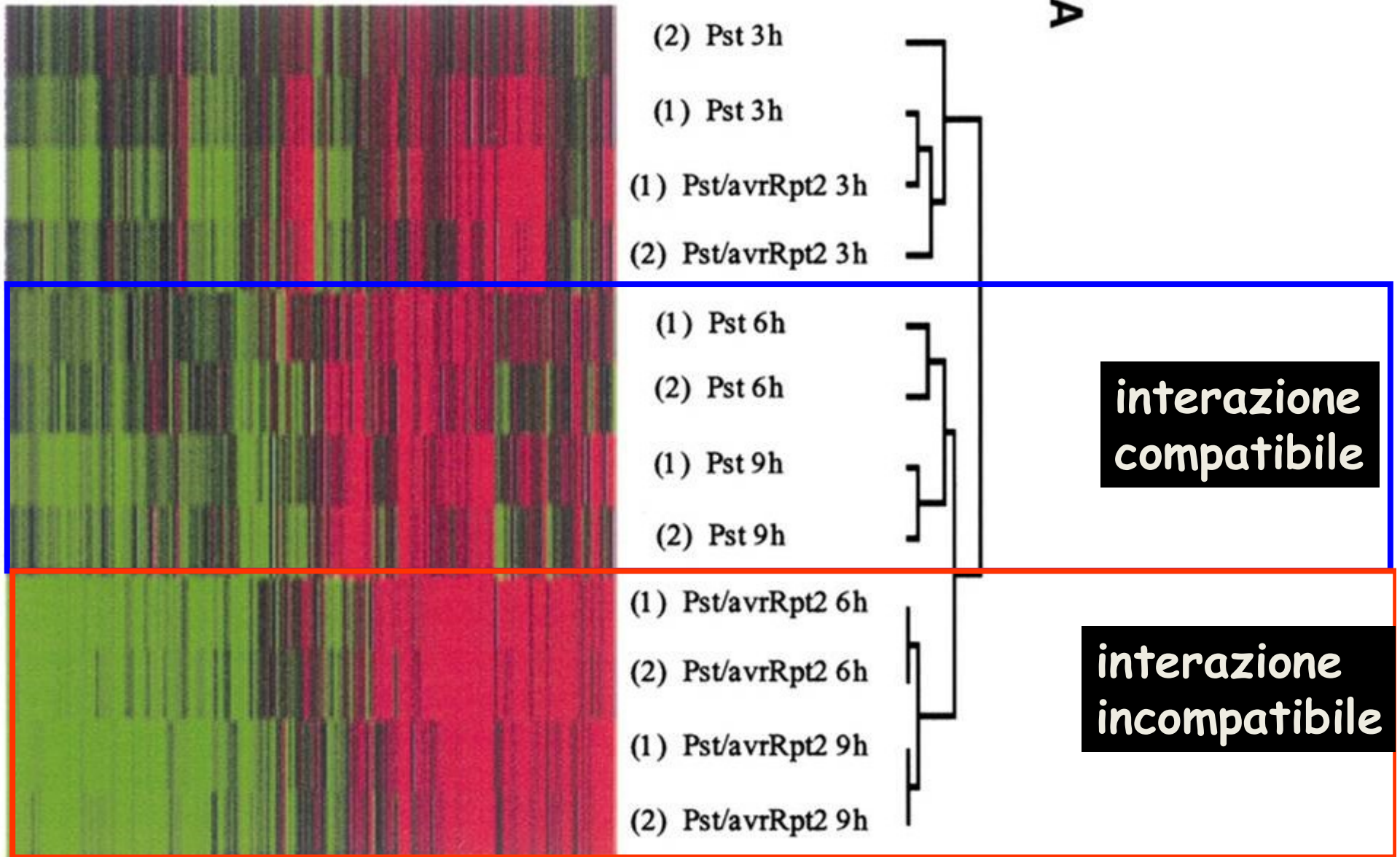
- Sviluppati algoritmi matematici per valutare somiglianze e differenze nel pattern di espressione a prescindere dal livello assoluto di espressione dei singoli geni



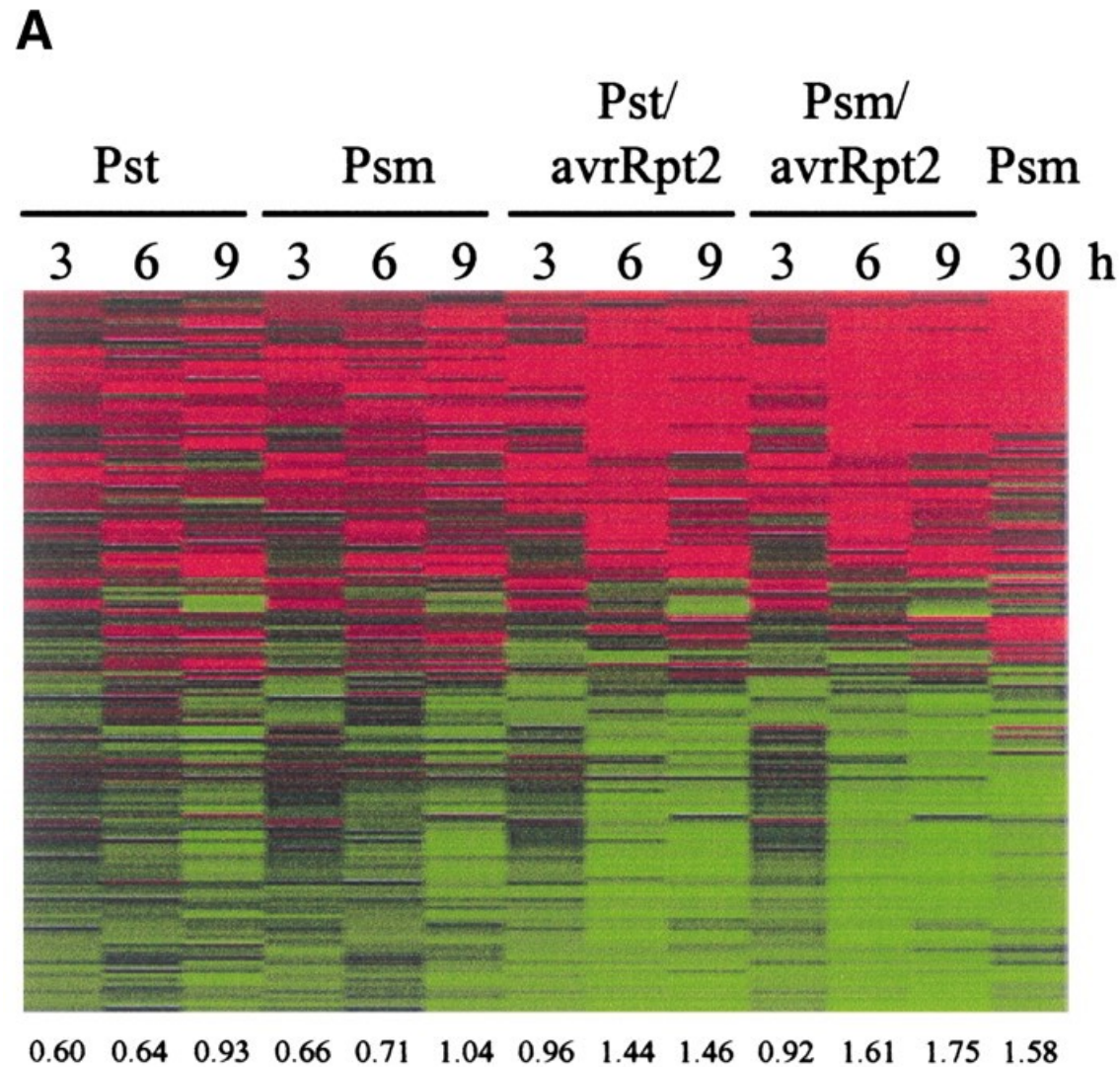
espressione > controllo



espressione < controllo



Le risposte indotte tardi durante la reazione compatibile sono simili a quelle indotte precocemente nell'interazione incompatibile



Analisi del profilo trascrizionale indotto da Psp (resistenza non ospite) ha dato risultati simili a quelli osservati durante la reazione incompatibile dovuta a resistenza gene-per-gene.

Il livello medio di induzione o repressione genica è apparso però più basso.

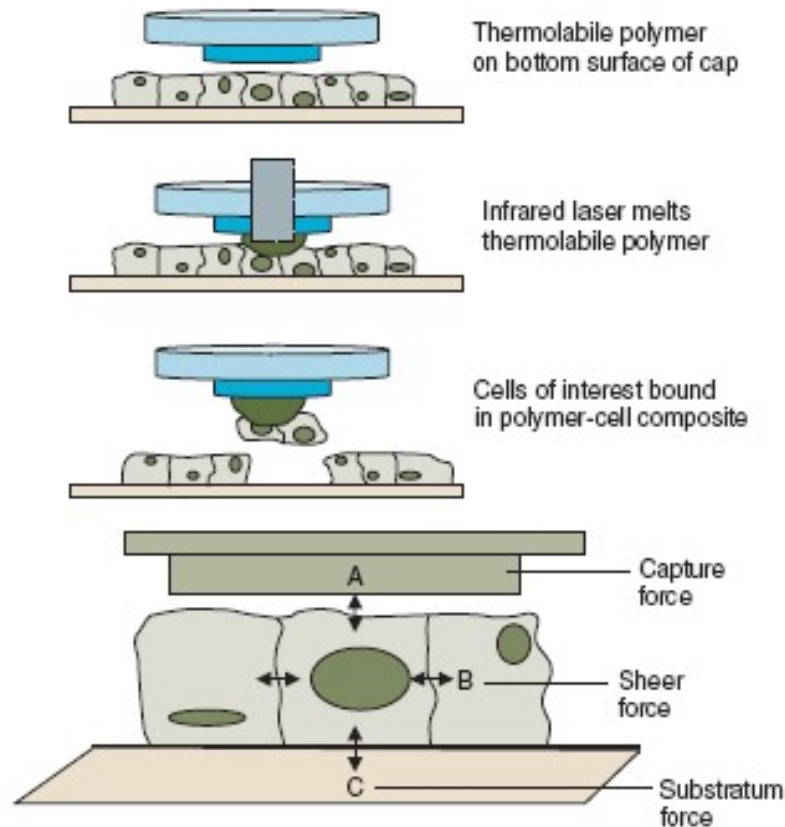
## CONCLUSIONI

- Le differenze a livello di espressione genica tra risposta compatibile ad incompatibile sono quantitative, ma non qualitative
- Simili risposte si osservano più velocemente nella risposta incompatibile che in quella compatibile
- Risposte indotte durante la risposta incompatibile mediata da RPS2 ed RPM1 o durante l'interazione non-ospite sono simili

# Gene expression atlas

- **Transcriptome map or gene expression atlas:** a global estimation of gene expression in all possible cells, tissues, organs, or parts of an organism during the life cycle of the organism from embryogenesis to senescence.
- **Whole-transcriptome technologies** in the past two decades have allowed the construction of transcriptome maps with various resolutions.
- Detailed transcriptome maps can also be applied to study the genes of **non-model organisms**.

# LCM: LASER CAPTURE MICRODISSECTION



- Visualizzazione delle cellule di interesse mediante un microscopio a luce invertita
- Trasferimento dell' energia del laser ad un polimero termolabile con formazione di un complesso polimero-cellula (sistema IR) o fotovolatilizzazione delle cellule che circondano l'area selezionata (sistema UV)
- Rimozione delle cellule di interesse dalla sezione di tessuto eterogeneo rimuovendo la pellicola
- Lo strumento è interfacciato con tubi da centrifuga in cui le cellule vengono recuperate per ulteriori analisi



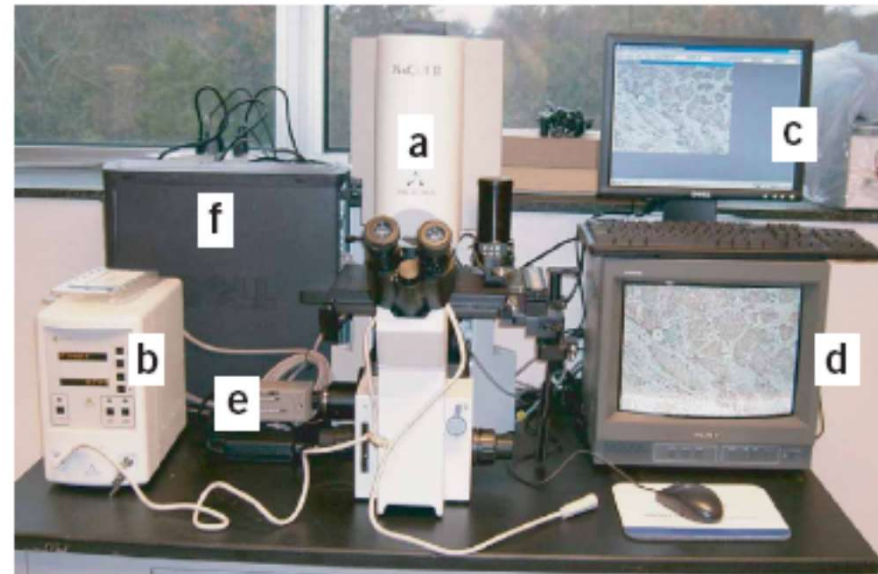
## DIVERSI TIPI DI STRUMENTI LCM

SISTEMI CON RAGGIO IR

SISTEMI CON RAGGIO UV

SISTEMI COMBINATI IR/UV

Esistono sistemi manuali (es. PixCell) o automatizzati (es. AutoPix) che utilizzano software che forniscono supporto nel riconoscimento morfologico dei diversi tipi cellulari



## LCM in pianta

- Se i tessuti vegetali sono congelati prima del fissaggio i grandi cristalli di ghiaccio tra i tessuti e le cellule danneggiano la morfologia delle cellule
- le cellule vegetali sono connesse da una rigida parete vegetale che le costringe a rimanere attaccate alle altre cellule durante l' LCM
- per risolvere questi problemi trattamento delle cellule vegetali con agente crioprotettore sotto vuoto prima del congelamento
- Plasmolisi per evitare contaminazione con cellule non di interesse

## LCM e amplificazione dell'RNA

Utilizzo di due tipi di fissanti chimici per preservare i tessuti:

•fissativi coagulanti (alcol, acetone...)



Migliore resa di RNA

75% etOH 25% acido acetico

•fissanti cross-linking

37-40% formaldeide, 95% alcol etilico  
(10%:5%:50%)



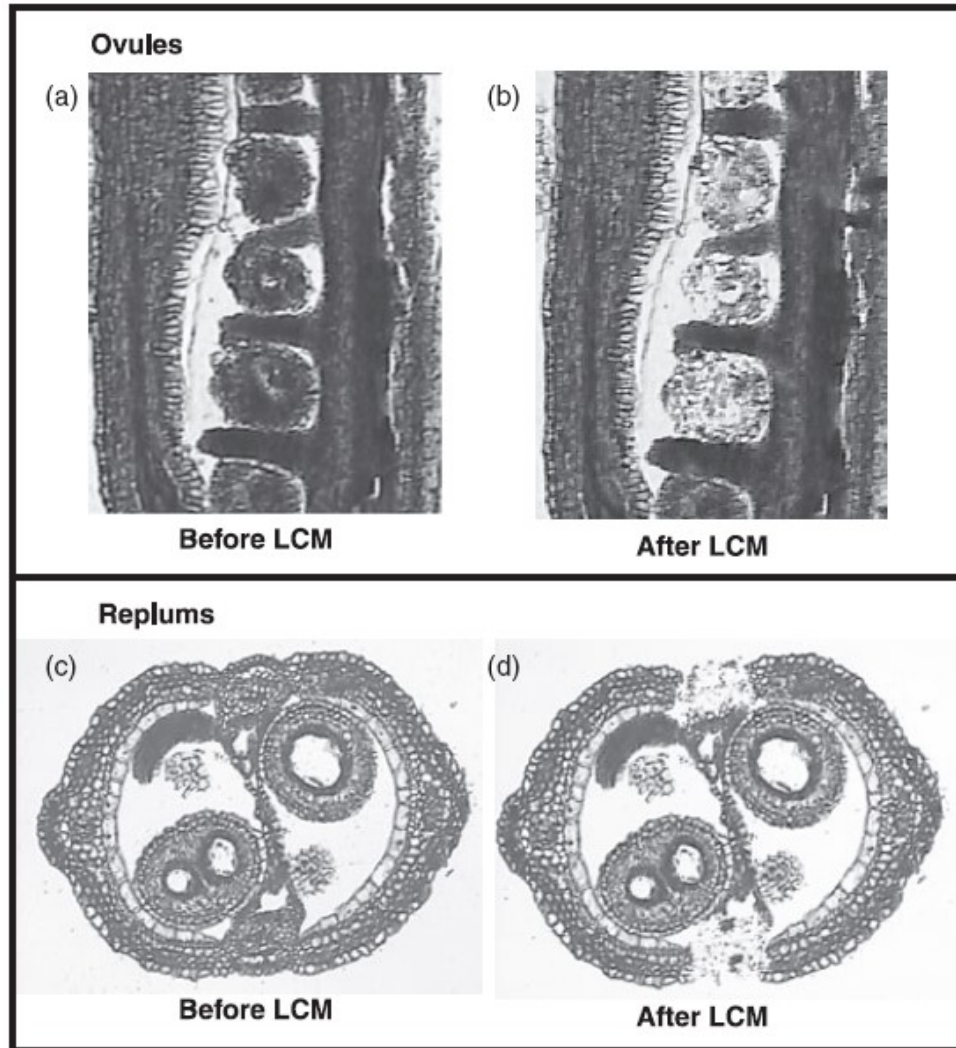
Migliore morfologia dei tessuti

**Laser capture microdissection of plant cells from tape-transferred paraffin sections promotes recovery of structurally intact RNA for global gene profiling**

Plant Journal 48, 628, 2006

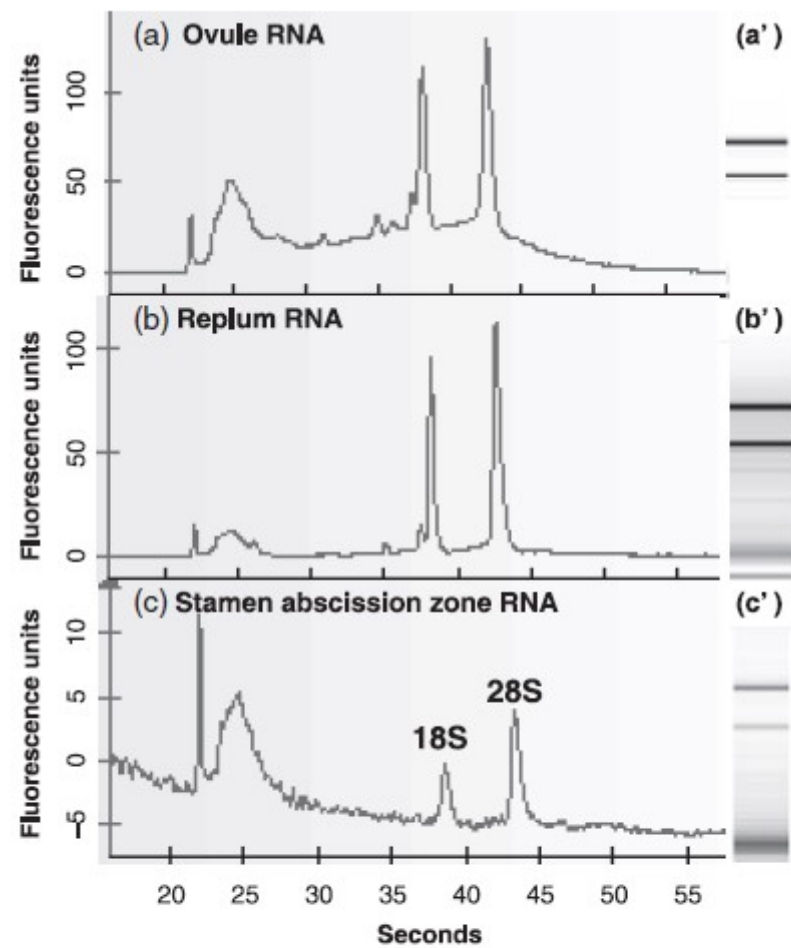
*Protocollo modificato di **LCM** basato su fissazione con paraffina per ottenere una maggiore resa e qualità dell'mRNA cellulare*

Tessuto in paraffina fissato con acido acetico ed etanolo



LCM ovuli

LCM cellule del replo



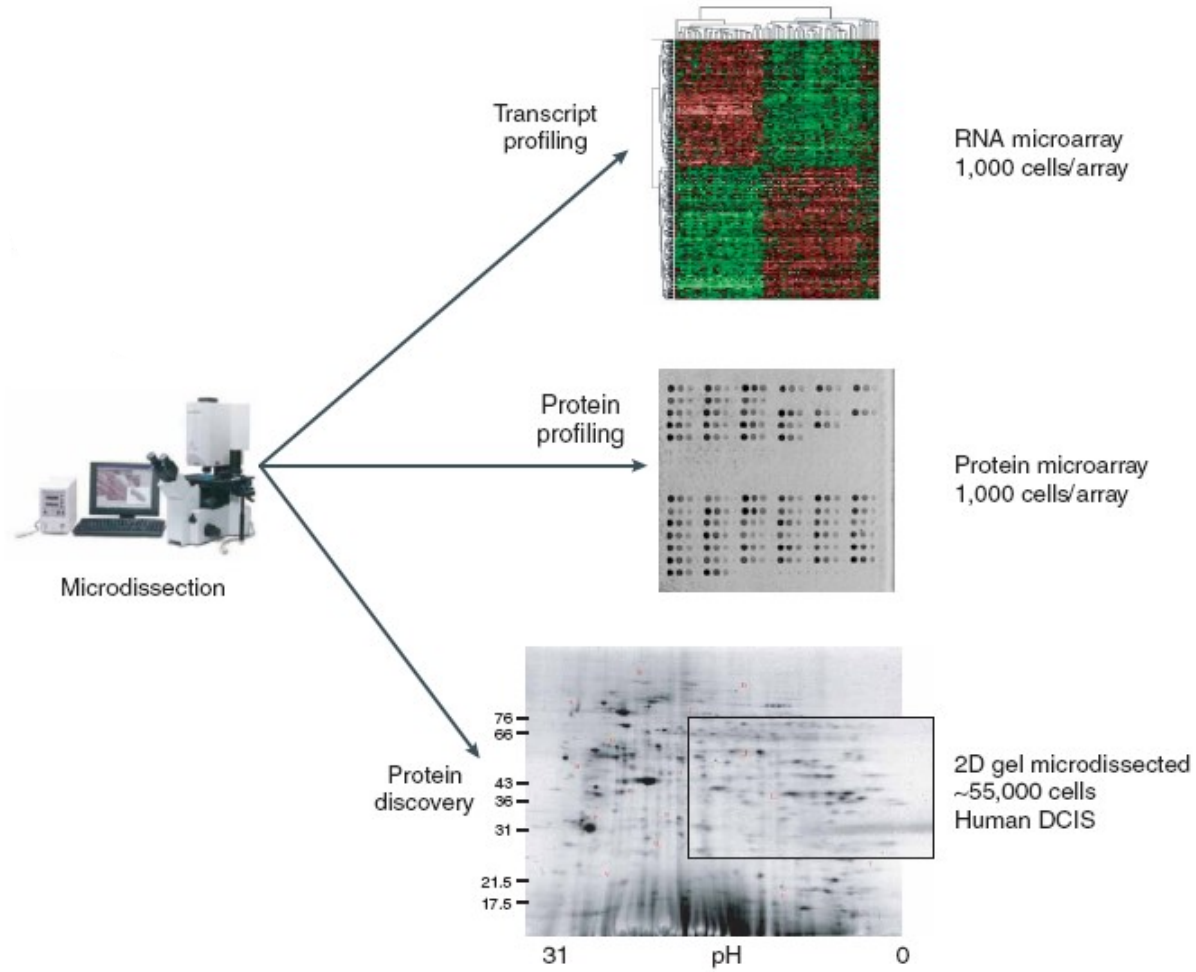
**Figure 3.** Qualitative assessment of RNA integrity from cells laser-captured from paraffin sections transferred to slides with tape.

(a) Fluorescence profile and (a') gel-like image of total RNA from laser-captured ovules.

(b, b') Total RNA from replums.

(c, c') Total RNA from stamen abscission zones.

# ANALISI SUCCESSIVE ALL' LCM



GENOMICS ARTICLE

## Laser-Capture Microdissection, a Tool for the Global Analysis of Gene Expression in Specific Plant Cell Types: Identification of Genes Expressed Differentially in Epidermal Cells or Vascular Tissues of Maize<sup>(W)</sup>

Mikio Nakazono,<sup>a,b,1</sup> Fang Qiu,<sup>a,c,1</sup> Lisa A. Borsuk,<sup>d</sup> and Patrick S. Schnable<sup>a,c,d,2</sup>

<sup>a</sup>Department of Agronomy, Iowa State University, Ames, Iowa 50011

<sup>b</sup>Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

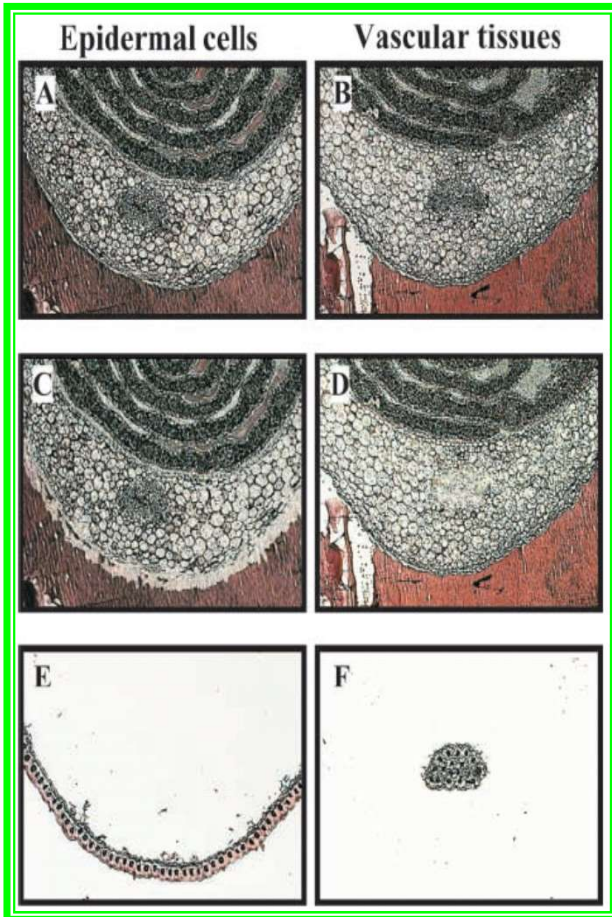
<sup>c</sup>Center for Plant Genomics, Iowa State University, Ames, Iowa 50011

<sup>d</sup>Department of Zoology and Genetics, Iowa State University, Ames, Iowa 50011





# LCM e amplificazione dell'RNA



Ottenuti 2 set di cellule epidermiche e vascolari contenenti ciascuno più di mille cellule portando a isolare circa 40 ng di RNA

**Table 1.** Yields of Amplified RNA from Laser-Captured Maize Coleoptile Cells

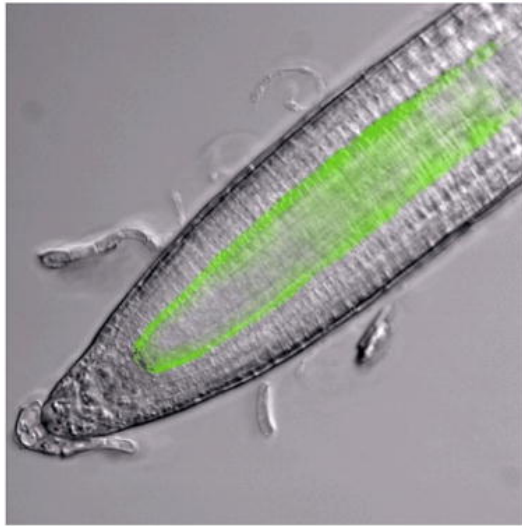
Experiment	Cell Source	Estimated No. of Captured Cells	RNA Yield (ng)	RNA per Cell (pg)	Second-Round Amplification	
					Yield ( $\mu$ g)	Amplification <sup>a</sup>
1	Epidermis	11,100	39.5	3.56	24.5	$\times 62,000$
	Vascular tissues	11,700	43.0	3.67	46.9	$\times 109,000$
2	Epidermis	16,100	35.2	2.19	28.6	$\times 81,300$
	Vascular tissues	18,700	39.0	2.08	25.4	$\times 65,100$

<sup>a</sup> Calculated based on the assumption that poly (A<sup>+</sup>) RNA is 1% of total RNA (Sugiura and Takeda, 2000).

# QUANTE CELLULE SONO NECESSARIE?

Molecule	Methodology/assay	Cellular yield/area of microdissection
DNA	Loss of heterozygosity	100–1,000 cells
DNA	Imprinting/DNA methylation	200 cells
gDNA	Genetic mosaic analysis	2,000 cells
RNA	cDNA library construction	25,000 cells (93 ng total RNA)
		5,000 cells (14.7–18.6 ng total RNA)
RNA	Gene-expression arrays	100 cells from FFPE
RNA	Real-time RT-PCR	1,400 cells
		0.8–1.0 × 10 <sup>6</sup> μm <sup>2</sup>
		200 and 1,000 cells
		22,000 cells/37.5 ng RNA
		10,000 cells/40 ng RNA from maize
		single cell
RNA	QRT-PCR	100 cells/1 reaction or 2,000 cells/200 μl
		4,000–5,000 cells
Protein	Western blot	500 cells (optimized blotting procedure)
		2,500 cells
		8,000–10,000 cells
Protein	2D gel electrophoresis	50,000–100,000 cells
		3.7 mm <sup>2</sup> area
		10,000 cells (100–200 μg in 350 μl)
		20,000–25,000 cells
		50,000 cells
Protein	2D-DIGE	30,000 cells/40 μl
Protein	Molecular profiling: reverse-phase protein microarray	5,000–30,000 cells
Protein	Mass spectrometry: MALDI or LC/MS-MS	50,000–100,000 cells (ICAT and LC/MS)
		10,000–15,000 cells
		25,000 μm <sup>2</sup>
		300 microvessels
Protein	Mass spectrometry: SELDI	1,500 cells
		3,000–5,000 cells

1a

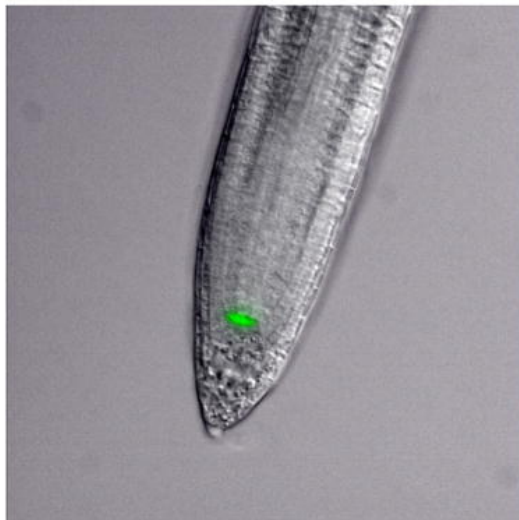


Cell type-specific GFP marker lines in the *Arabidopsis* root

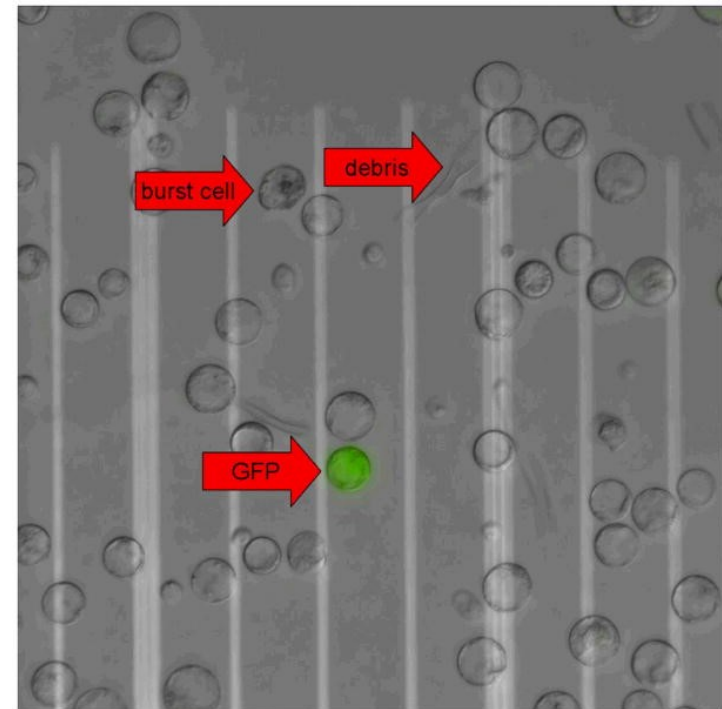
a)  $P_{SCR}::GFP$

b)  $P_{WOX5}::GFP$

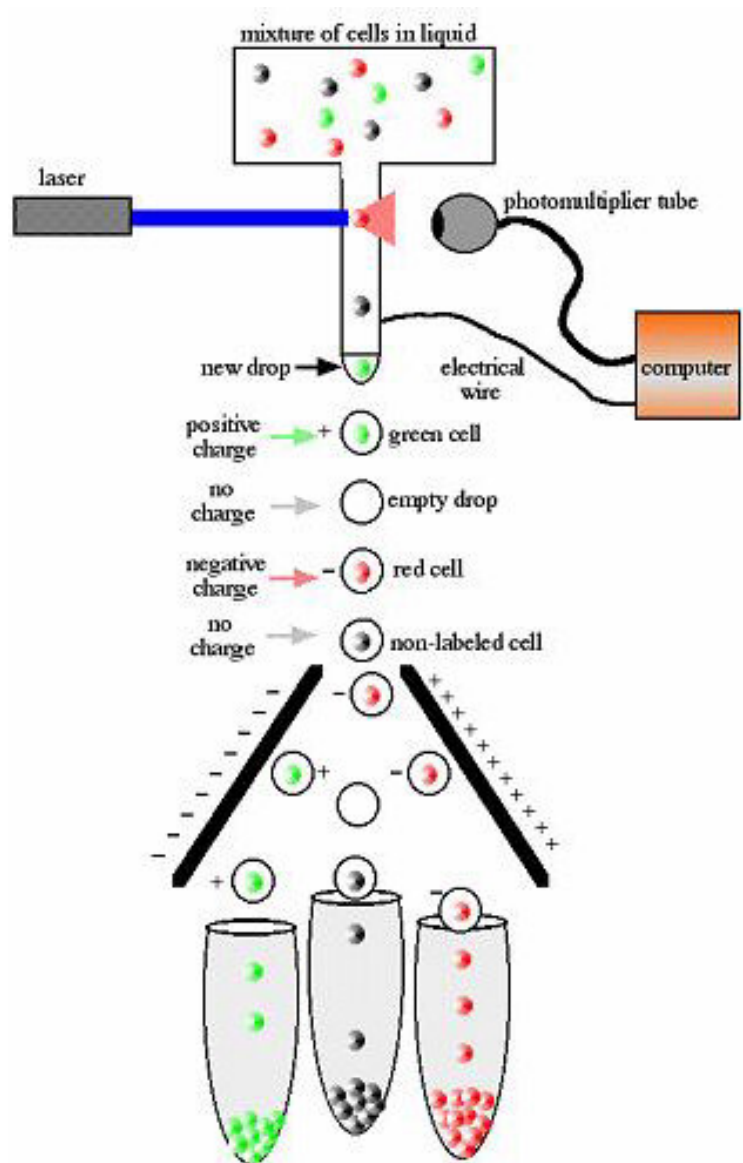
1b



Protoplasting



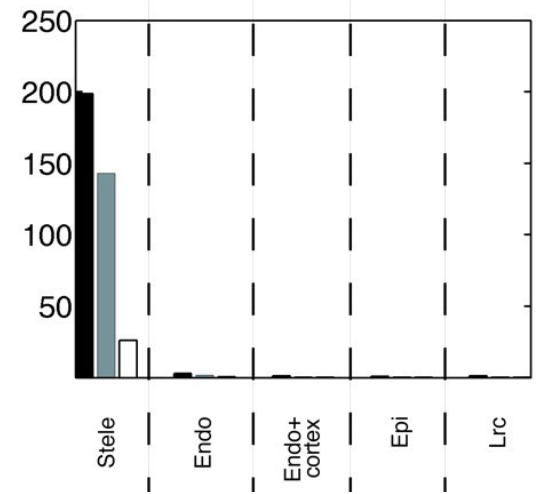
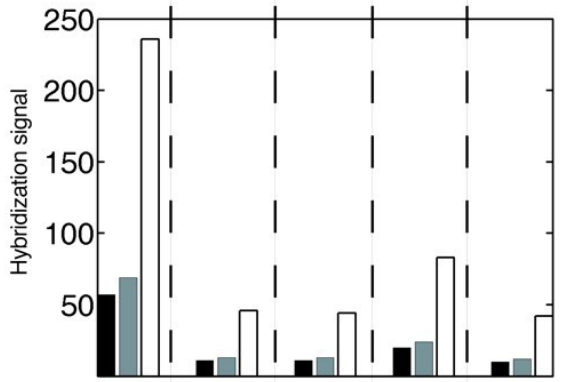
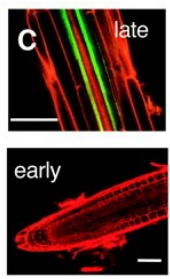
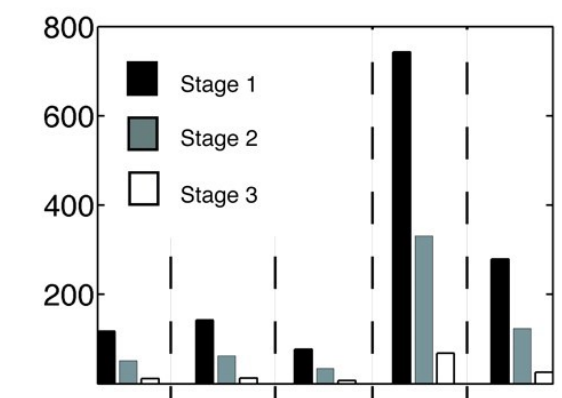
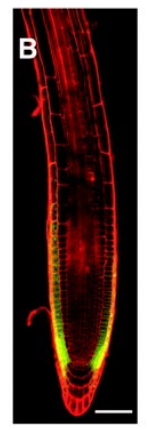
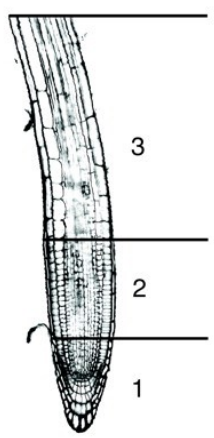
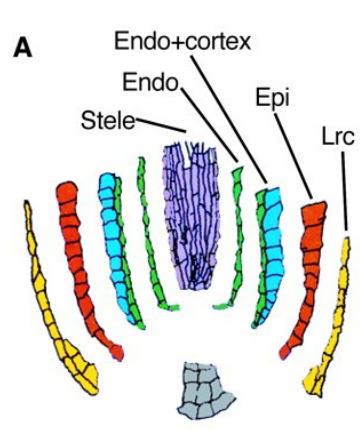
# FLUORESCENCE ACTIVATED CELL SORTING (FACS)



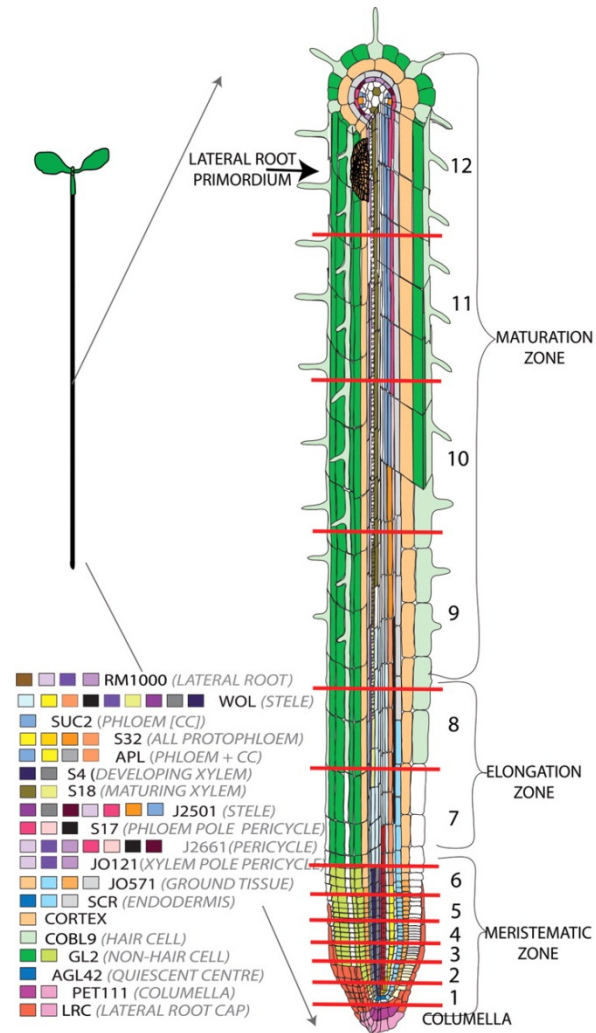
# A Gene Expression Map of the *Arabidopsis* Root

Kenneth Birnbaum,<sup>1</sup> Dennis E. Shasha,<sup>2</sup> Jean Y. Wang,<sup>3</sup>  
Jee W. Jung,<sup>1</sup> Georgina M. Lambert,<sup>4</sup> David W. Galbraith,<sup>4</sup>  
Philip N. Benfey<sup>3\*</sup>

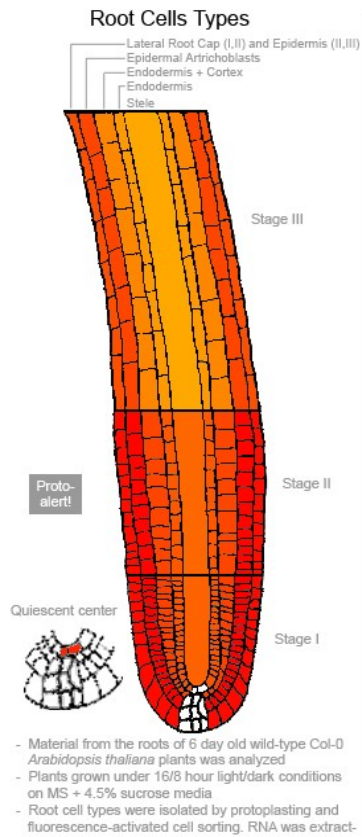
12 DECEMBER 2003 VOL 302 SCIENCE [www.sciencemag.org](http://www.sciencemag.org)



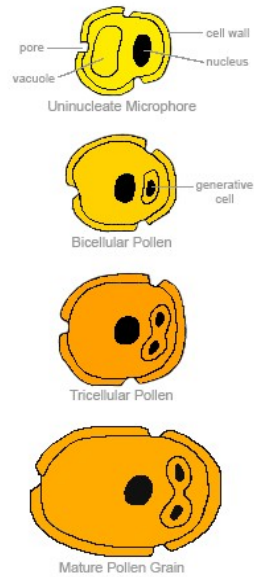
# Microarray expression profiles of 19 fluorescently sorted GFP-marked lines were analyzed



S M Brady et al. Science 2007;318:801-806

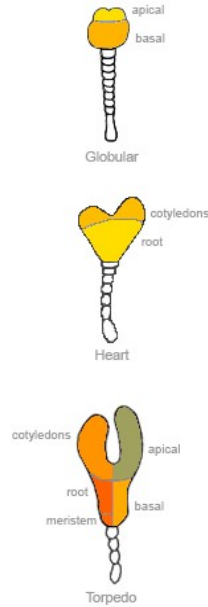


### Microgametogenesis (Pollen Development)



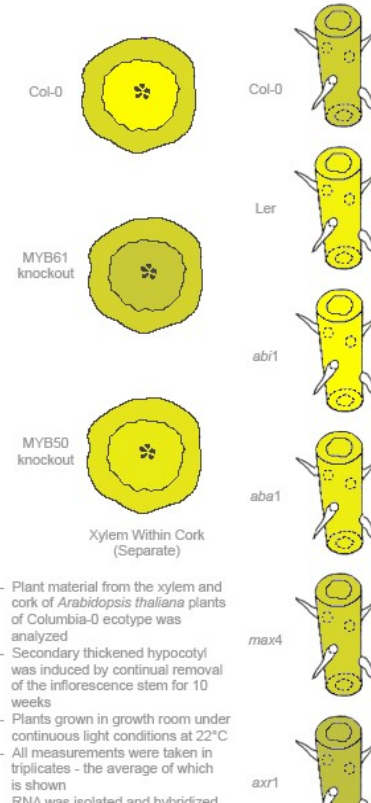
- Plant material from the pollen of 5.10 growth stage wild-type *Arabidopsis thaliana* plants of Ler-0 ecotype was analyzed
- Plants grown under 16/8 hour light/dark conditions at 21°C
- All measurements were taken in duplicates - the average of which is shown

### Embryo Development



- Plant material from embryos of wild-type Col-0 *Arabidopsis thaliana* plants was isolated by laser-capture microdissection
- Plants grown under 16/8 hour light/dark conditions
- RNA was amplified and hybridized to the ATH1 GeneChip. Note: 3' bias!
- All measurements were taken in triplicates - the average is shown

### Xylem and Cork



- Plant material from the xylem and cork of *Arabidopsis thaliana* plants of Columbia-0 ecotype was analyzed
- Secondary thickened hypocotyl was induced by continual removal of the inflorescence stem for 10 weeks
- Plants grown in growth room under continuous light conditions at 22°C
- All measurements were taken in triplicates - the average of which is shown



# LIMITI DELLA TRASCRITTOMICA

- Un gene può codificare per più proteine (modificazioni post-trascrizionali)
- I livelli di trascritto non sempre corrispondono ai livelli di proteina
- L'attività delle proteine è regolata a diversi livelli (es. fosforilazione, ubiquitinazione, legame di cofattori...)

Altri approcci: proteomica, metabolomica...