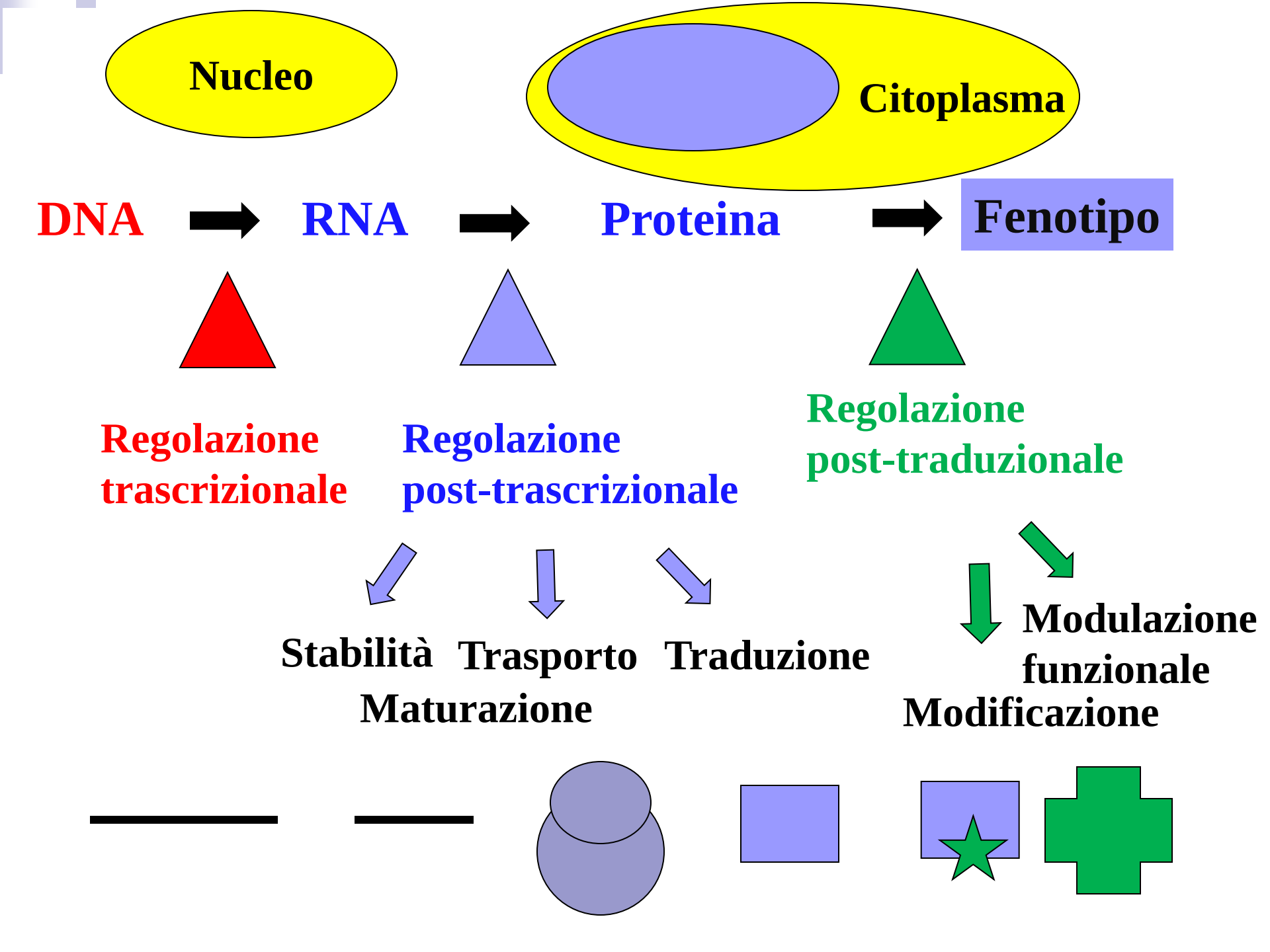
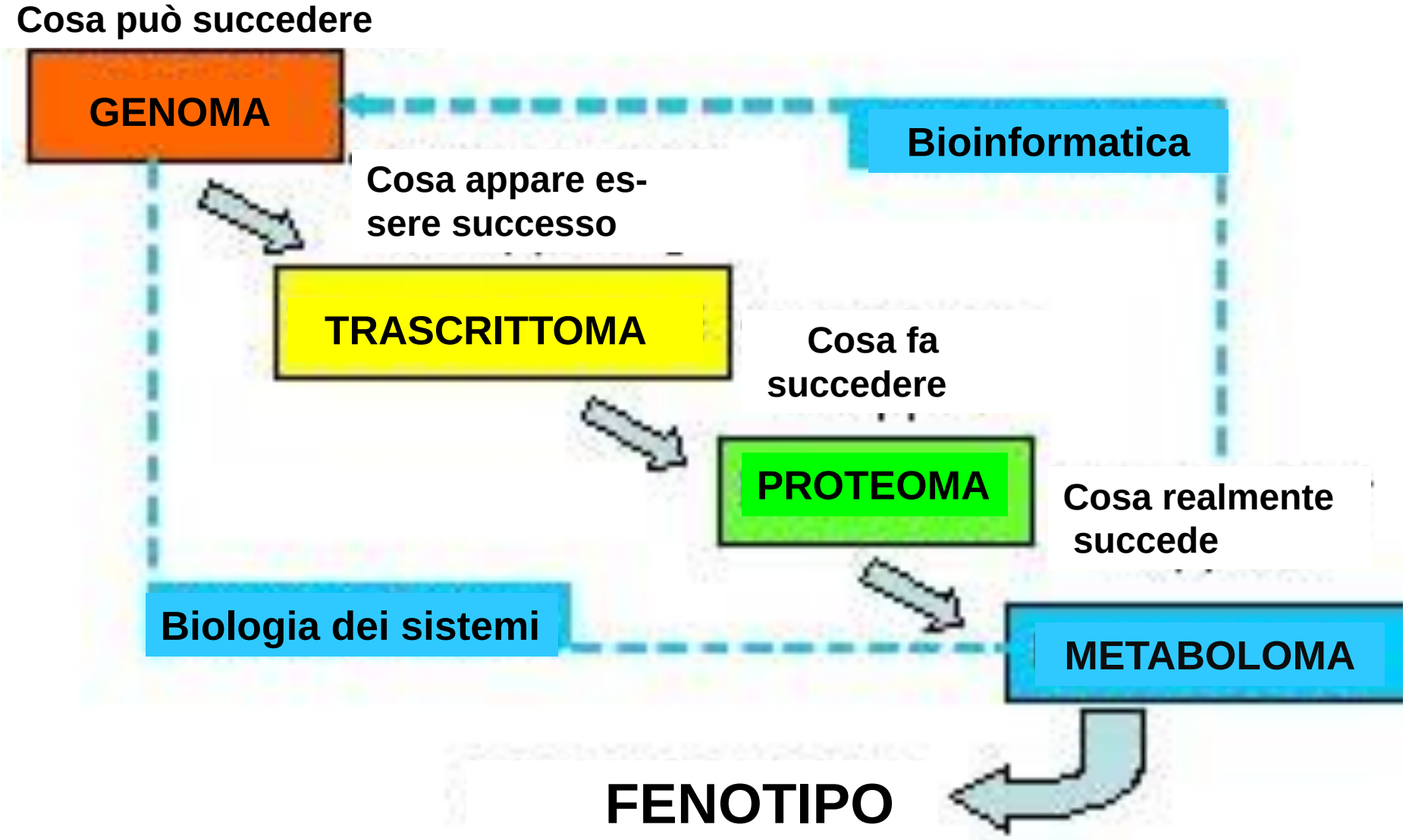
23. Genomica funzionale/ Postgenomica

- La **Genetica** studia il ruolo del **singolo gene** nel produrre un determinato fenotipo.

- La **Genomica strutturale** studia **l'organizzazione del Genoma**.

- La **Genomica funzionale** studia il **funzionamento** coordinato dei geni nel produrre un determinato fenotipo.


Nucleo
Citoplasma
DNA
RNA
Proteina
Fenotipo
Regolazione trascrizionale
Regolazione post-trascrizionale
Regolazione post-traduzionale
Stabilità
Trasporto
Traduzione
Maturazione
Modulazione funzionale
Modificazione


Cosa può succedere
GENOMA
Bioinformatica
Cosa appare es-
sere successo
TRASCRITTOMA
Cosa fa
succedere
PROTEOMA
Cosa realmente
succede
Biologia dei sistemi
METABOLOMA
FENOTIPO

I. La trascrittomica

- Il **genoma** può essere definito come **l'insieme di tutti i geni** che compongono un organismo e la genomica la disciplina che ne affronta lo studio

- Analogamente possiamo definire
  - **trascrittoma l'insieme di tutti i geni trascritti**
  - **trascrittomica la disciplina che se ne occupa.**

- E ancora possiamo parlare di proteomica, metabolomica, farmaco-genomica ecc.

Trascrittomica.

- Sappiamo che
- **l'espressione genica** a livello trascrizionale, **influisce** in modo importante **sull'espressione genica complessiva**
- **solo un sotto insieme di geni viene trascritto**, in un dato momento o in una particolare localizzazione,
- ciò in **risposta a complessi stimoli interni ed esterni** alla cellula,
- da qui l'interesse di conoscere le dinamiche globali dei profili di trascrizione.

“One-gene approach”

- Il gene di interesse e’ espresso in un tessuto o in un dato momento dello sviluppo ? Quanto e’ attivo dal punto di vista trascrizionale ?

- Real Time PCR
- PCR semiquantitativa
- Ibridazione DNA genico o cDNA con RNA totale o poly(A)+RNA (Northern blot)
- Ibridazione *in situ*

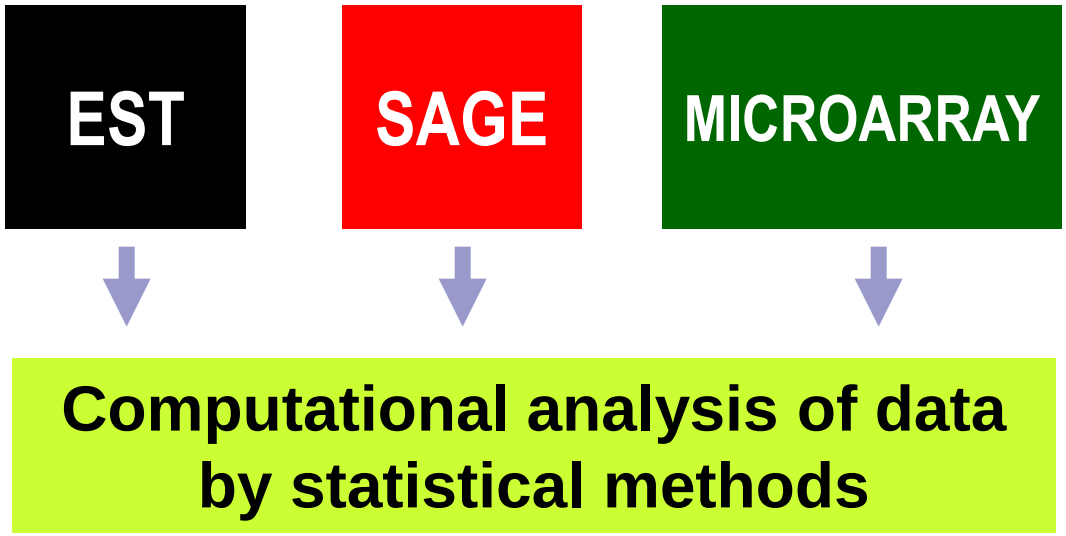
“Large-scale approach”

- Quali geni sono espressi in un tessuto ed in un dato momento dello sviluppo ?
- Quanto ciascuno di essi e’ attivo dal punto di vista trascrizionale ?

Profilo d’espressione del genoma (TRASCRITTOMA)

A. Metodi per lo studio dell'espressione genica su larga scala:

1. ESTs
2. SAGE
3. Microarray



1. Cosa sono le EST?

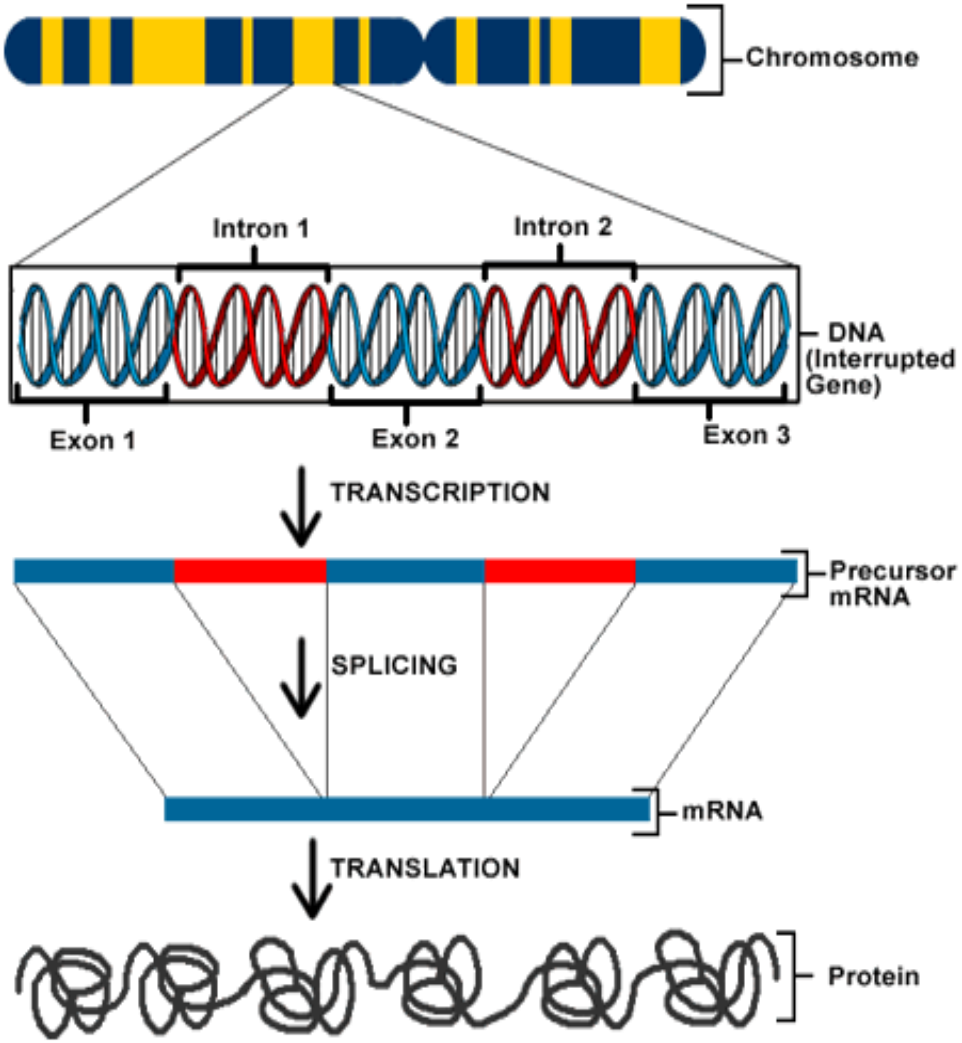
ESTs

❑ **Expressed Sequence Tags.**

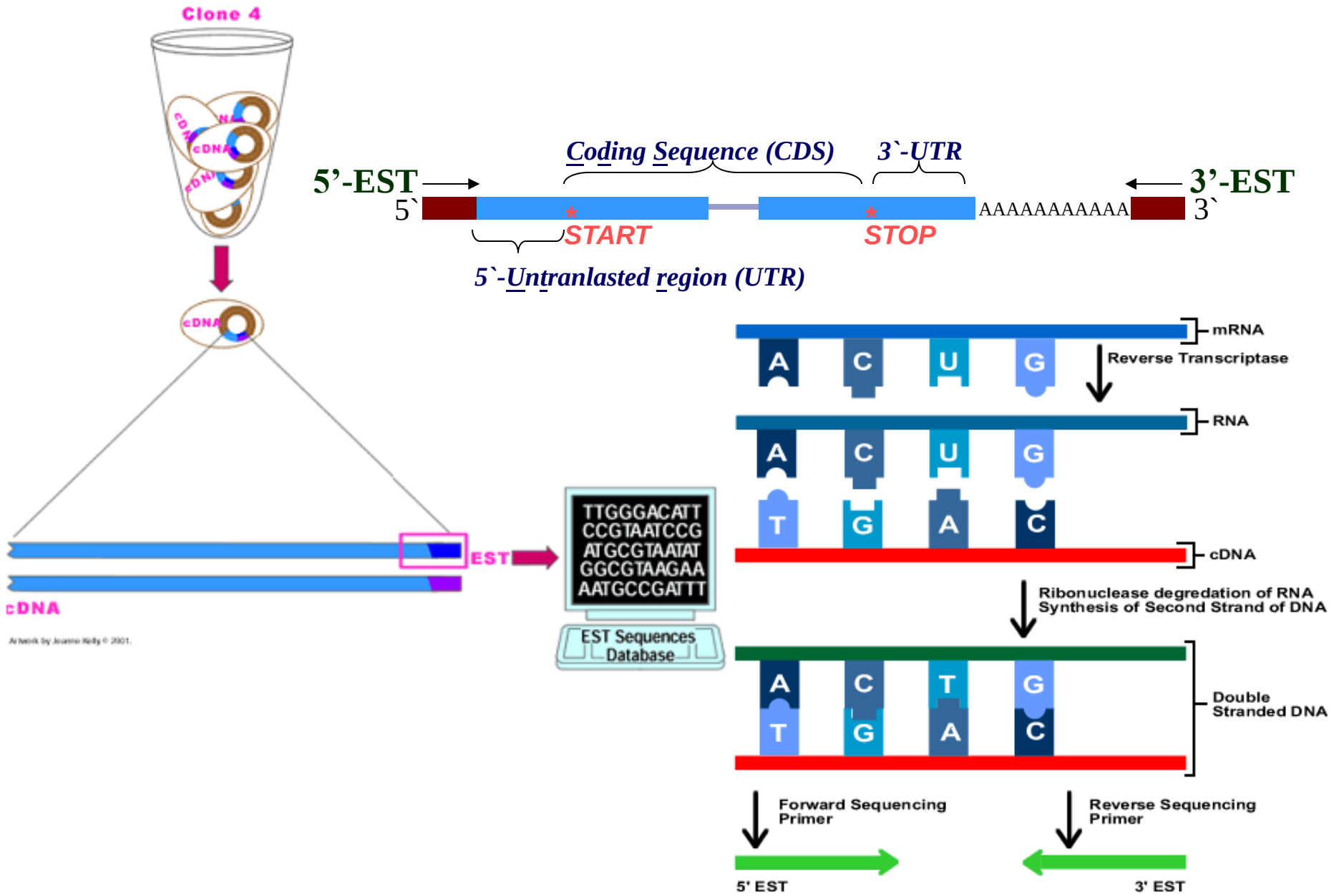
- **Brevi frammenti di sequenze di DNA** (generalmente lunghe **200- 500 nt**
- generate dal sequenziamento di uno o entrambe le estremità di un gene espresso.

▪ L'idea è di:

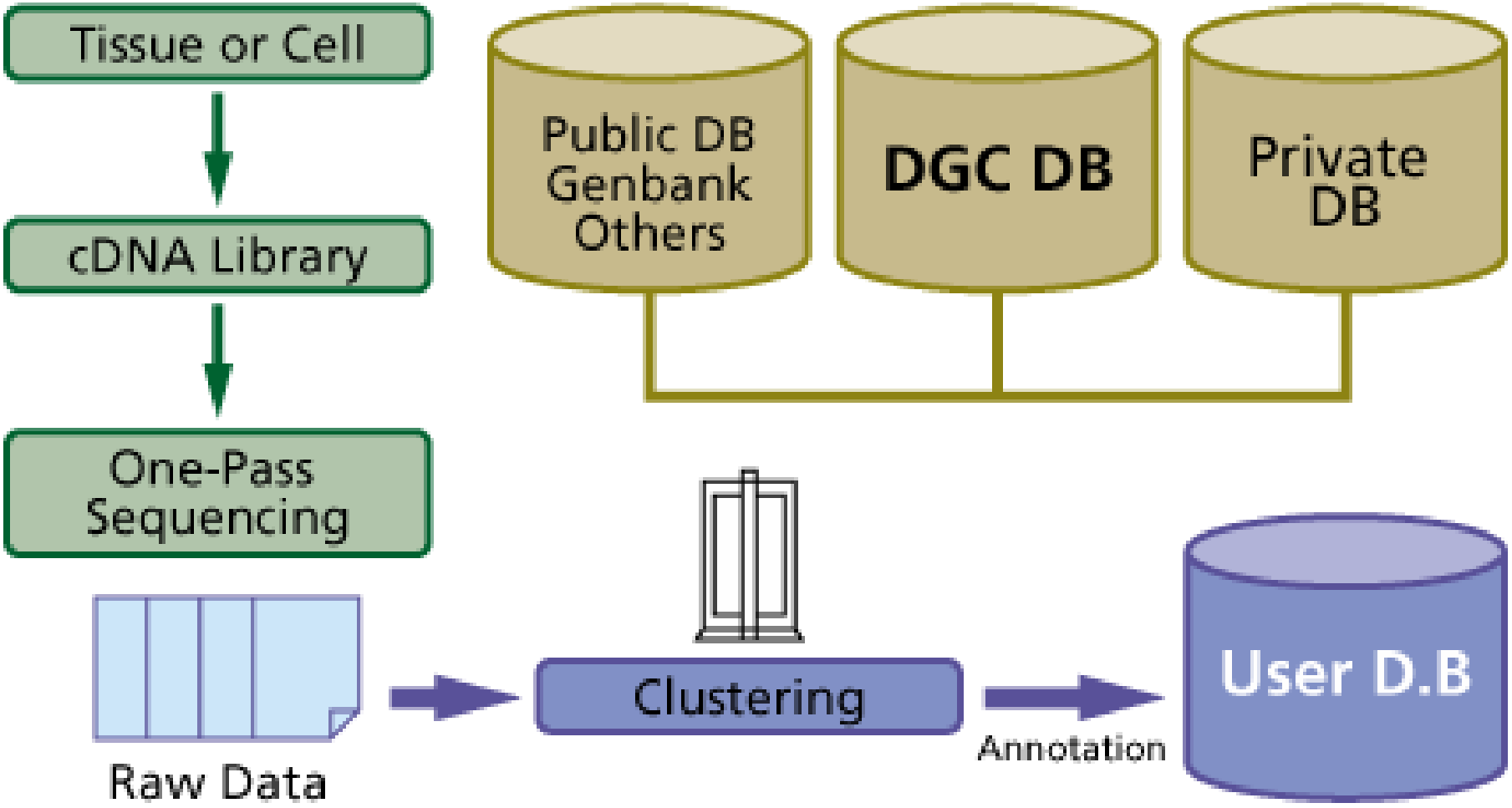
- sequenziare frammenti di **DNA che rappresentano geni espressi** in specifiche cellule, tessuti, o organi da differenti organismi-
- Usare queste **"tags"** per identificare un **gene da una porzione di DNA cromosomico DNA tramite appaiamento delle basi**
- La probabilità associata con l'identificazione dei geni varia tra le sequenze genomiche degli organismi
- dipende dalle dimensioni del genoma e dalla presenza o assenza di **introni (*i.e.* sequenze di DNA che interrompono la continuità della sequenza codificante di un gene)**



Expressed Sequence Tags (*EST*)



EST Analysis



Basic Features and Tools of an Automated EST Analysis Pipeline

- ▲ Relational database (Oracle 8i)
- ▲ Automatic data validation
- ▲ Quality score generation
- ▲ Automatic trimming of low-quality, vector, adaptor, poly-A tails, low-complexity and contaminant sequences
- ▲ Automatic running of selected blast algorithms, with user-defined parameters, user selected reference databases, and storage of top results (by user-defined cutoffs) in the database
- ▲ Includes a web interface for viewing the data in the database, according to the permissions allowed to the viewer (by individual, project, lab or institution)
- ▲ Includes a Java tool for dbEST submission of newly generated ESTs at intervals define by the users
- ▲ System can be readily and simply deployed at any of the partner's institutions
- ▲ Includes methods for defining a Unigene set for a library.

Additional functionalities are needed by the members of the current co-development group, including:

- Tissue or organism, integration of gene expression data.
- Annotations: Gene ontology annotations, functional motif annotation, metabolic pathways annotations, signal transduction pathways.

2.SAGE (Serial Analysis of Gene Expression)

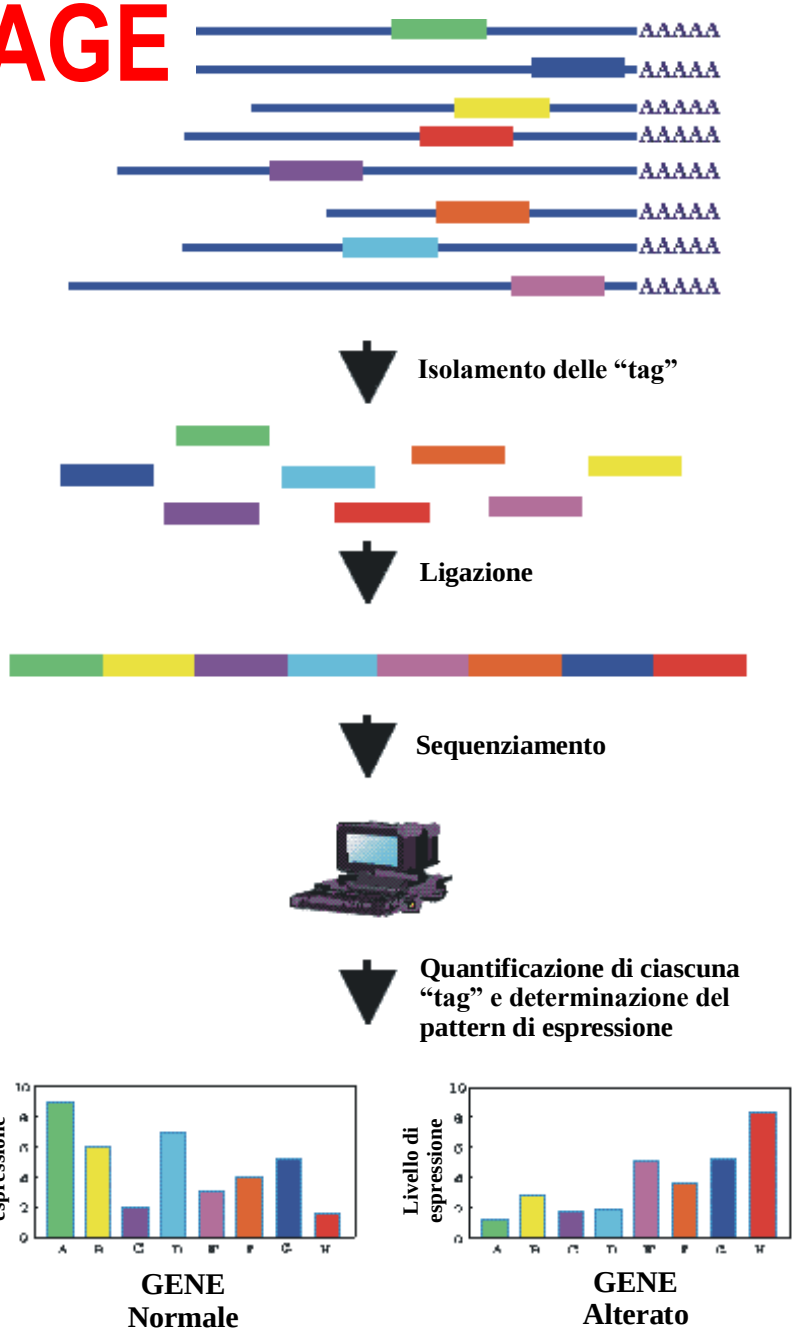
SAGE

- SAGE è un metodo sperimentale, ideato per utilizzare i vantaggi del sequenziamento su larga scala, al fine di **avere informazioni quantitative di espressione**
- Con questa tecnica e' **possibile stimare il livello d'espressione di ciascun gene,**
  - questo attraverso la **misura del numero di volte in cui la TAG** che lo rappresenta **compare in un campione abbastanza grande di TAGs** sequenziate a partire dal messaggero del tessuto in analisi

Tag to Gene mapping ←→ Gene to Tag mapping

- Consiste nel **sequenziamento da messaggeri cellulari di brevi oligo-nucleotidi,** che fungono da etichette di sequenza (TAG)

SAGE

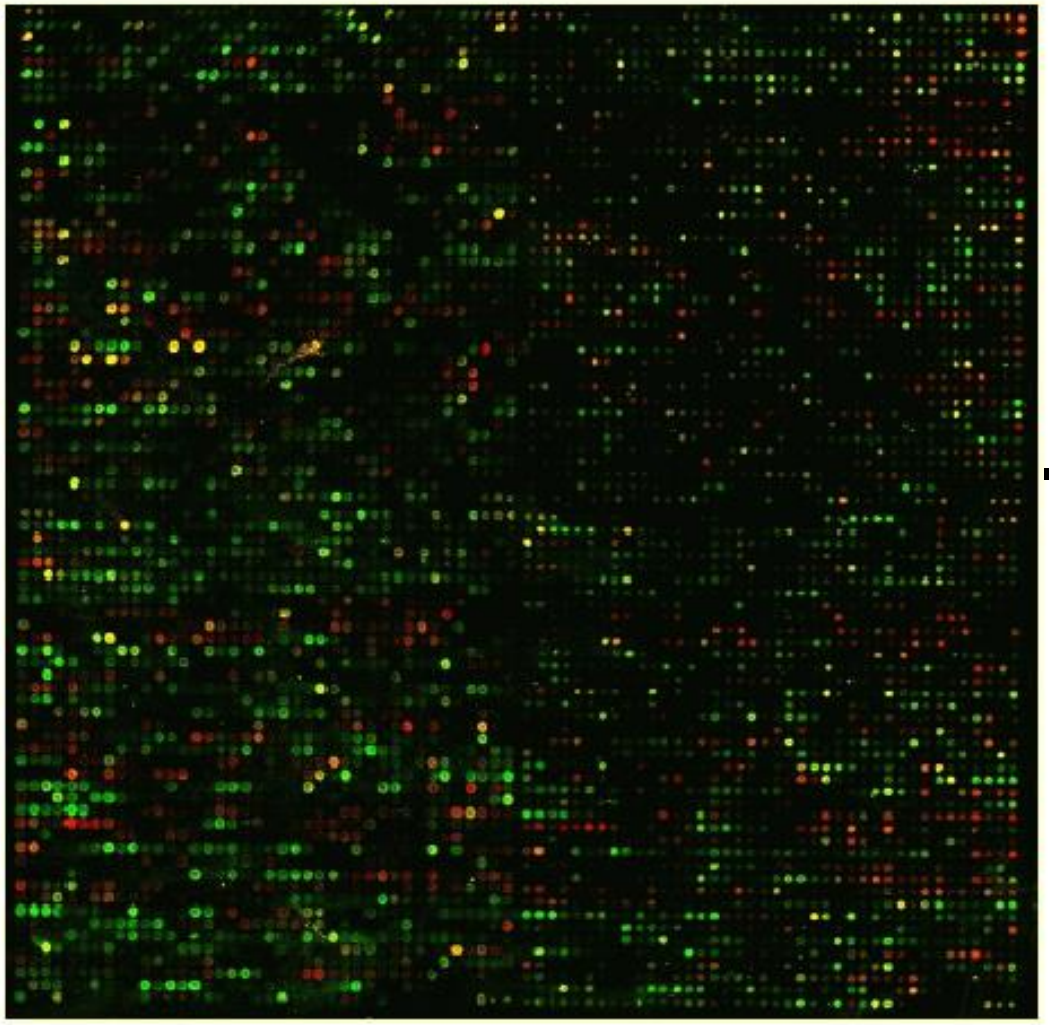


- una sequenza di **9 pb permette di identificare 4^9** (262.144) diversi trascritti (una "tag" viene ottenuta da una posizione specifica di ogni trascritto).

- le "tag" possono essere **unite insieme in serie, a costituire lunghe molecole di DNA, c**he vengono clonate e sequenziate.

- il **numero di volte in cui una singola "tag"** viene osservata permette di **quantificare l'abbondanza del messaggero** identificato nella popolazione dei messaggeri e, indirettamente, il **livello di espressione del gene corrispondente.**

3. Esperimenti di Microarray

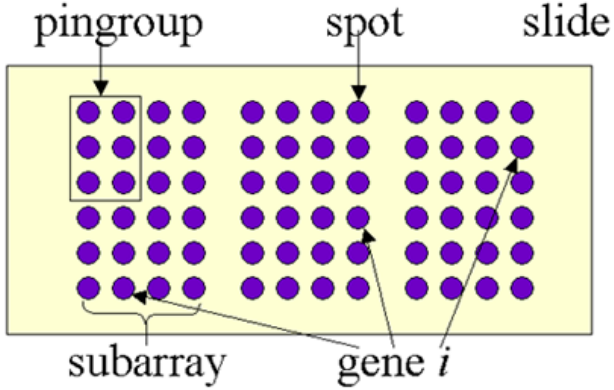


Permettono l'analisi dell'espressione di migliaia di geni simultaneamente

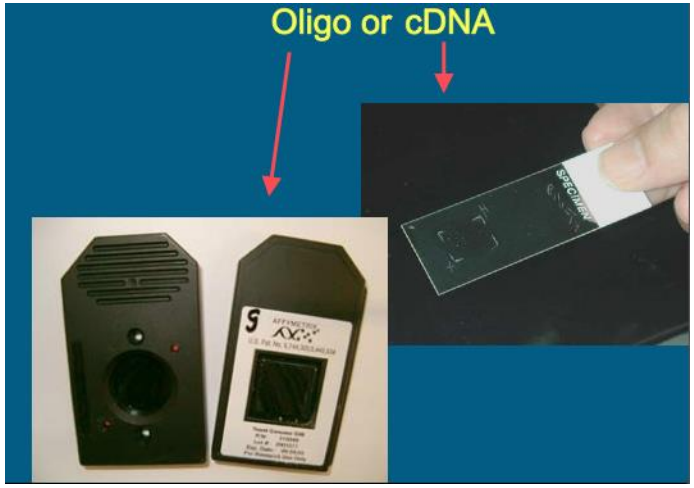
DNA microarray

- **La trascrittomica è resa possibile**, oltre che dalle accresciute conoscenze sulle sequenze geniche e sulla crescente disponibilità di cDNA, da una tecnica, il DNA **microarray, che permette di comparare tra loro un gran numero di geni.**





- Un microarray è **una matrice solida di piccole dimensioni** , di solito di nylon o, più di **recente di vetro**, su cui sono immobilizzate, **in posizioni definite** migliaia di sequenze note di:
  - **cDNA (Agilent)**
  - **oligonucleotidi sintetici (Affymetrix) .**



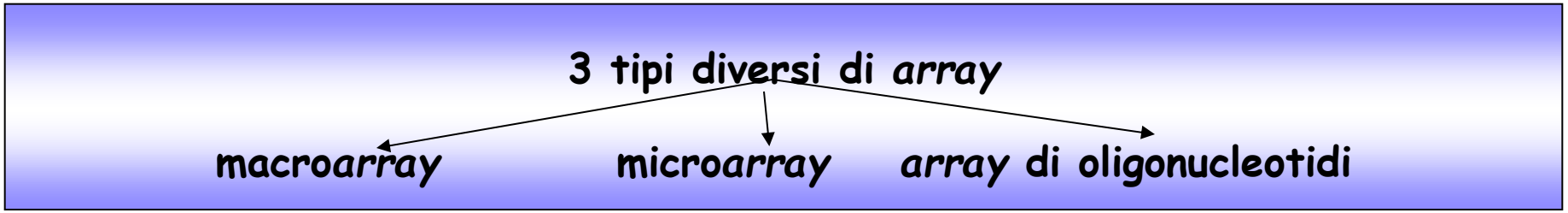
- Attualmente è possibile costruire dei chips ad alta densità (i vetrini su cui posizionano gli spots di DNA) contenenti fino a 10.000 cDNA/cm^2 e fino a 500.000 oligo/ cm^2.

Tecnologia degli "*array*" di DNA

- Un **array di DNA** è una grande **collezione di frammenti di DNA disposti** in file **ordinate su un supporto solido chiamato "*chip*"**.

- I *chip* sono formati da **moltissime molecole di DNA (detti *sonde*)** depositate in una posizione nota su un **supporto a formare una griglia** (da cui il nome *array*) che consente di identificarle in MODO UNIVOCO.
- Il **supporto, di** solito, è un **vetrino da microscopio** che ha le dimensioni, più o meno, di un francobollo.
- Ogni **sonda è costituita da DNA a singola elica** che può corrispondere
  - al **cDNA di un gene**,
  - ad **una sequenza genomica**
  - oppure ad **un oligonucleotide** (cDNA o genomico).

- **Per costruire un *array* si possono:**
  - **usare cloni di cDNA** prelevati direttamente da micropozzetti
  - **produrre frammenti di DNA per PCR** che saranno poi depositate in posizioni specifiche su un supporto solido usandone **microscopiche gocce**
  - **sintetizzare oligonucleoti**di direttamente su un **supporto solido** che corrispondono a porzioni di geni specifici



Costruzione dei microarrays

- Le sequenze di DNA sono legate ai microarrays essenzialmente con due metodi:
  - ***microdispenser robotizzati*** che utilizzano tecnologie note come ink jetting
  - ***mechanical microspotting*.**

- La prima tecnologia **utilizza una propulsione piezoelettrica** che ricorda quella utilizzata nelle stampanti a getto d'inchiostro.
- La seconda funziona essenzialmente con il principio della capillarità:
  - Un capillare viene intinto in una goccia di DNA.
  - La soluzione di DNA per capillarità viene attratta sul microspillo e da questo depositato sull'array per contatto fisico.
  - Dopo aver completato un ciclo la punta viene lavata e il ciclo viene ripetuto da un robot per n cicli.
- Alternativamente, al posto dei cDNA vengono utilizzate sequenze **oligo-nucleotidiche che vengono sintetizzate *in sit*u**, direttamente sul chip, con una tecnologia che ricorda la fotolitografia

a.DNA microarray a cDNA (Agilent)

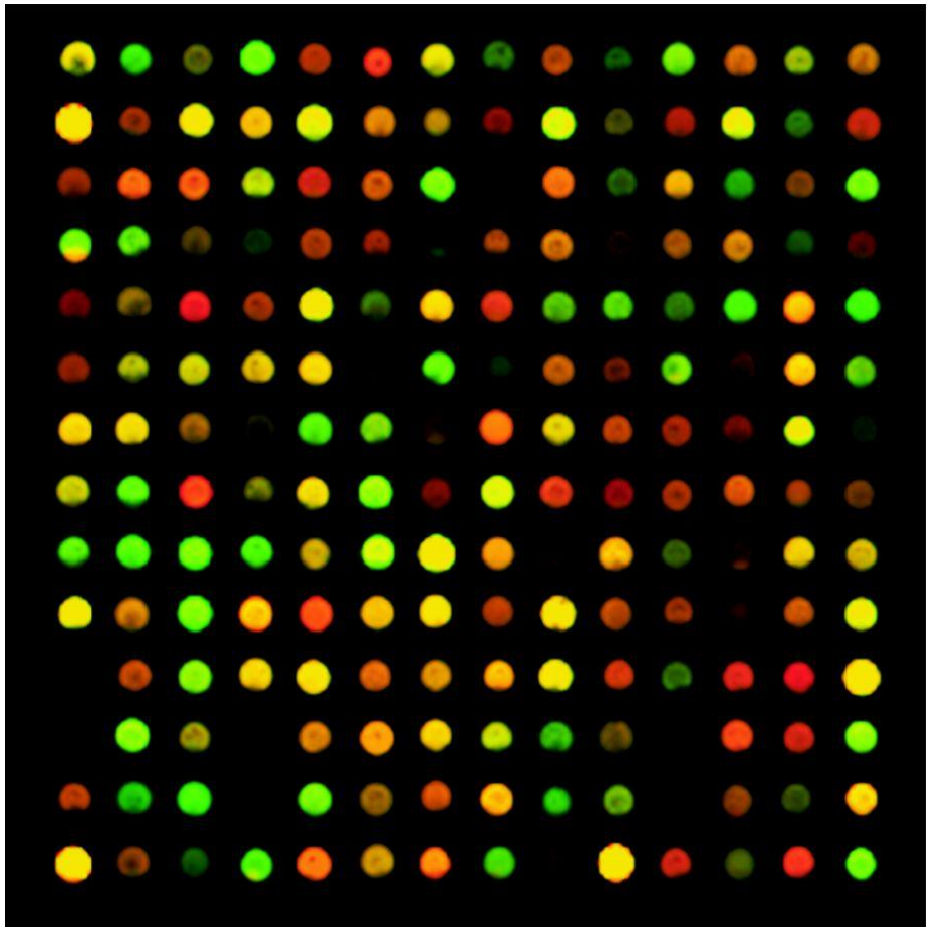
- Gli mRNA provenienti da due popolazioni o due condizioni sperimentali diverse (due canali), per esempio un mutante e un wild type, vengono retrotrascritti con marcatori fluorescente diversi, uno che emette nel **rosso (cyc5)** e uno che emette nel **verde (cyc3)**
- Le due popolazioni di cDNA marcate si riuniscono, si ibridizzano ai microarrays e vengono irraggiate con un raggio laser per leggere l'eventuale fluorescenza. Vengono rilevati separatamente i segnali emessi nel rosso e nel verde. Quindi i due segnali vengono sovrapposti (merge).
- Se un **cDNA** della popolazione **campione cyc5** è più abbondante rispetto alla popolazione di **controllo cyc3**, lo spot contenente la **sequenza complementare sul microarray apparirà rossa.**
- Se, al contrario, uno specifico **cDNA è** più rappresentato nel **controllo cyc3** che nel campione cyc5, la sequenza complementare sul microarray apparirà verde.
- Se le due quantità sono equilibrate apparirà **giallo,** e se non c'è ibridazione apparirà **nero.**
- Dal **colore e dall'intensità della fluorescenza possiamo** quindi trarre **informazioni su migliaia di geni** rispetto ad un controllo o rispetto a due situazioni diverse (per esempio due campioni estratti da individui trattati a concentrazioni diverse di un farmaco)

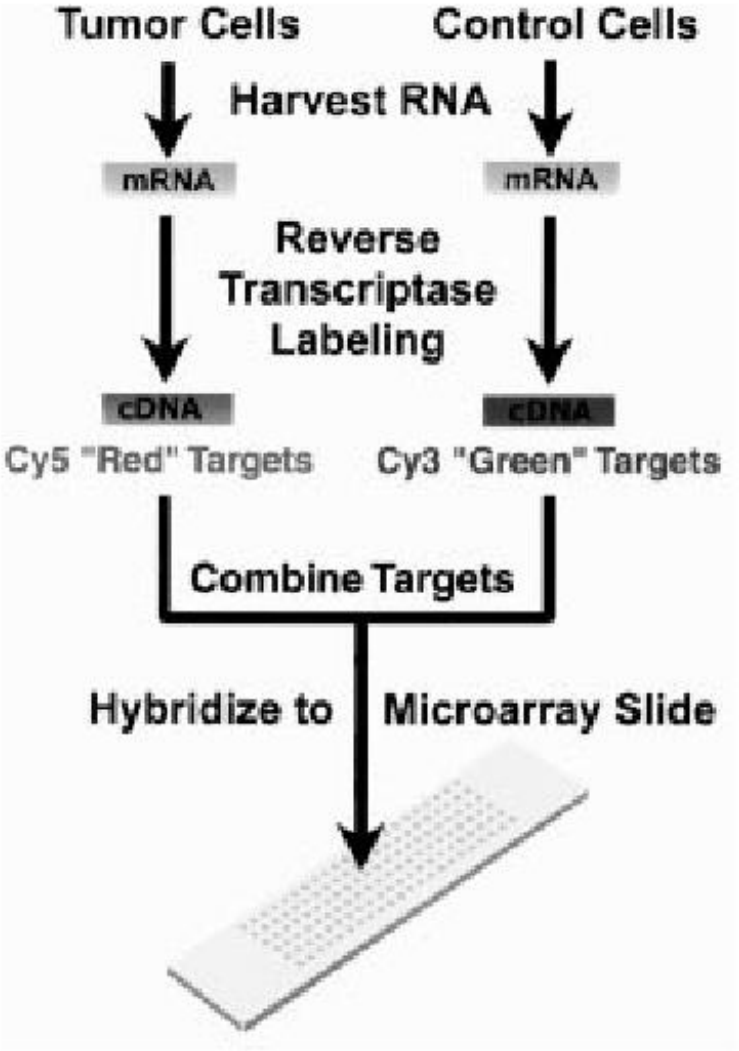
Microarray e analisi trascrizionale su scala genomica: trascrittoma

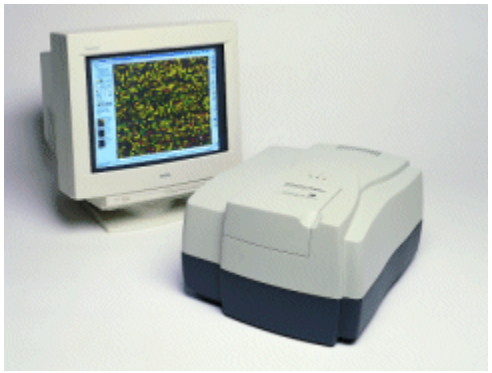
- I microarray: versioni modificate del Northern blotting , registrano simultaneamente il profilo di espressione di molti geni contemporaneamente

- **Microarray a due canali.**

- L'RNA è **isolato** da ogni campione, retrotrascritto in un cDNA a doppio filamento o in un cRNA **marcato** con un colorante fluorescente;
- è poi **ibridato** al vetrino che porta il microarray.
- I geni che sono più trascritti in un campione rispetto all'altro sono
  - o **rossi**
  - o **verdi,**
  - mentre quelli trascritti a livelli simili nei due campioni sono **gialli.**
- I geni che non sono trascritti appaiono **neri.**







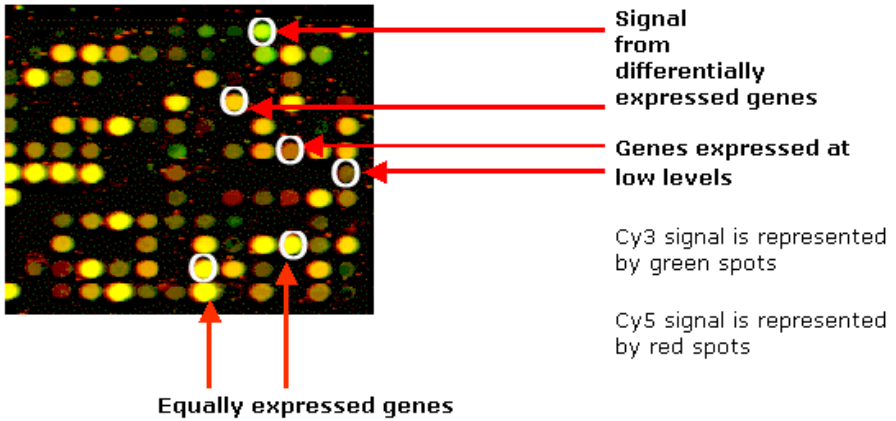


Figure 2. Schematic of target preparation and hybridization

Microarrays a cDNA (Agilent)

- cDNA microarrays
- Competitive hybridisation
- Relative expression levels
- Look at ratios: $\frac{Cy5}{Cy3}$ $\frac{red}{green}$

RNA sample A

AAAAA

Cy3

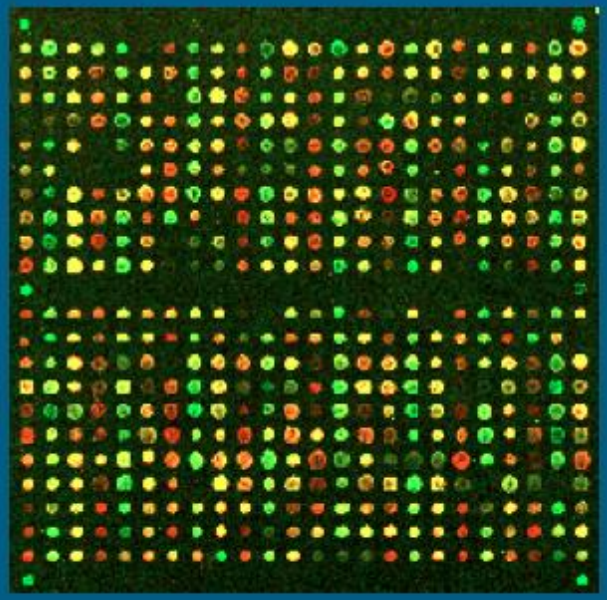
AAAAA

RNA sample B

AAAAA

Cy5

AAAAA

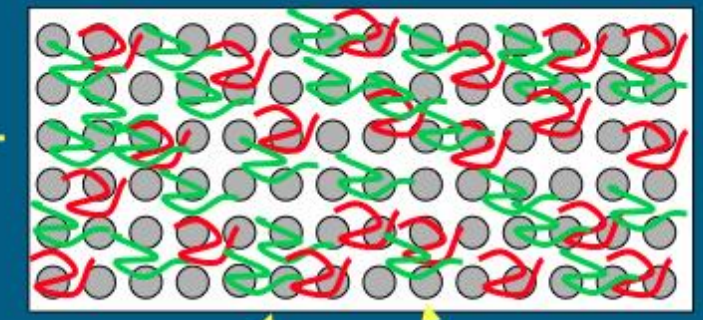


Cy5 < Cy3

Cy5 = Cy3

Cy5 > Cy3

No hybridis.



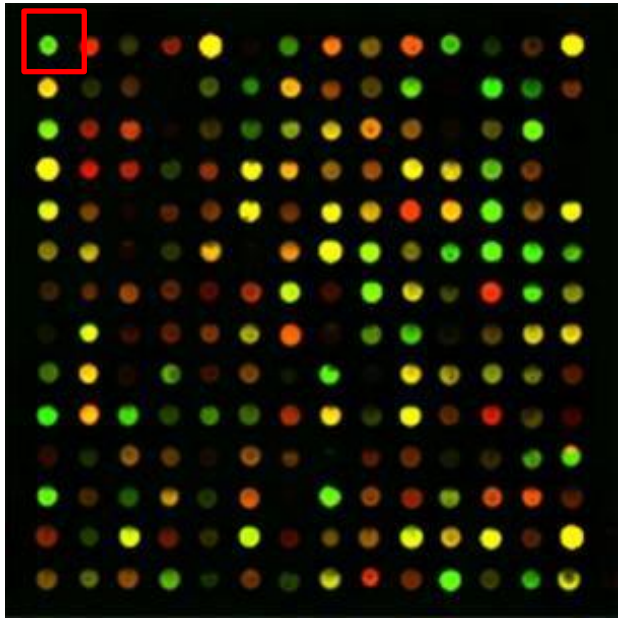
gene a

gene b

Ogni spot rappresenta un gene diverso.

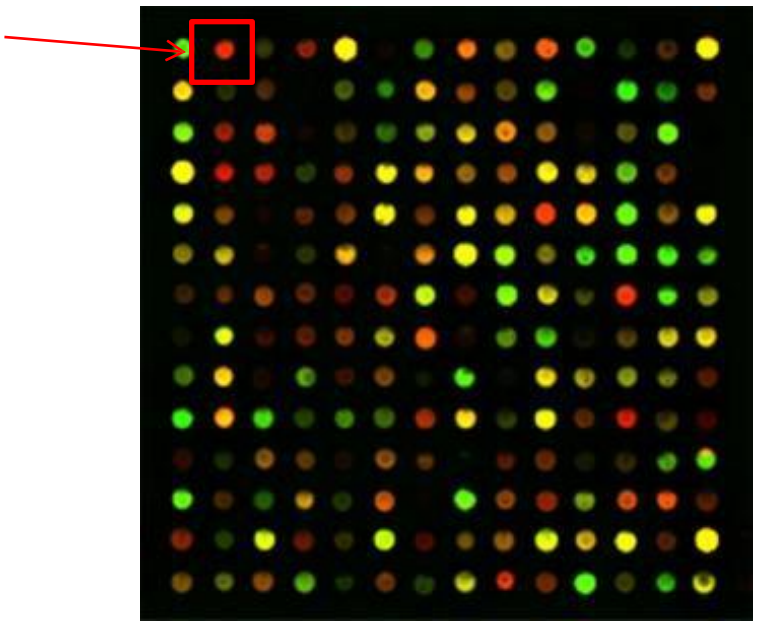
- Se lo spot corrispondente ad un determinato gene
- si illumina di **verde**

significa che quel gene è più espresso nel campione marcato con il verde



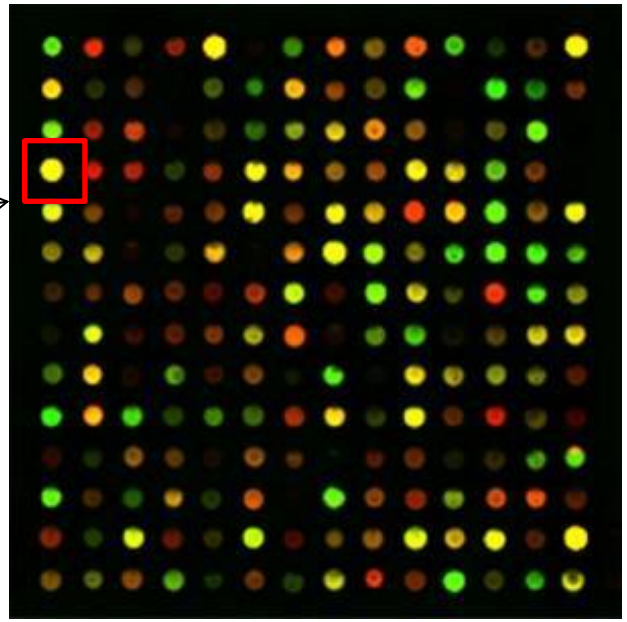
Ogni spot rappresenta un gene diverso.

- Se lo spot corrispondente ad un determinato gene si
- illumina di rosso
- significa invece che quel gene è più espresso nel campione marcato con il rosso

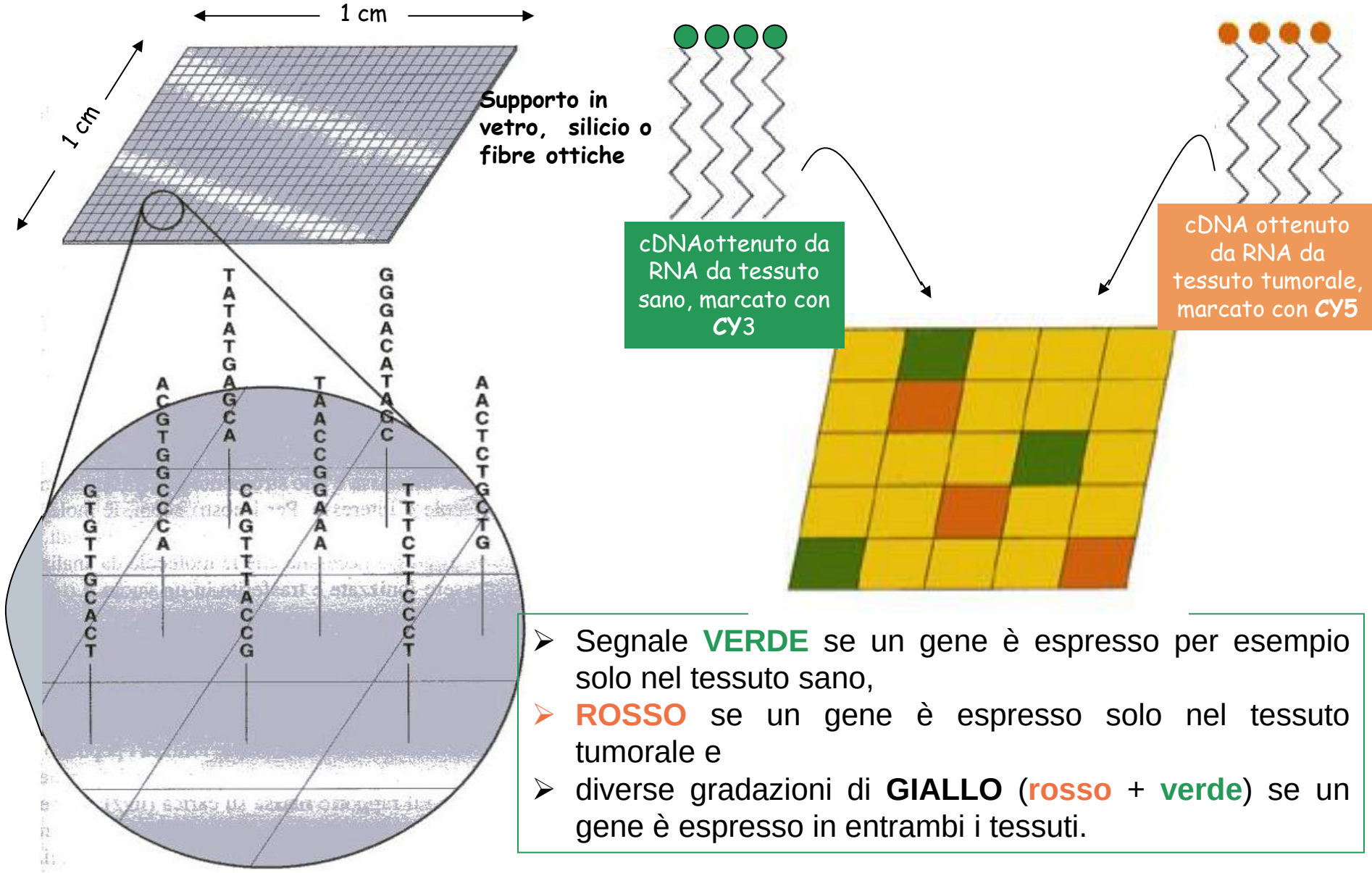


Ogni spot rappresenta un gene diverso.

- Se lo spot corrispondente ad un determinato gene si illumina di giallo significa invece che quel gene è espresso in ugual misura nei due campioni

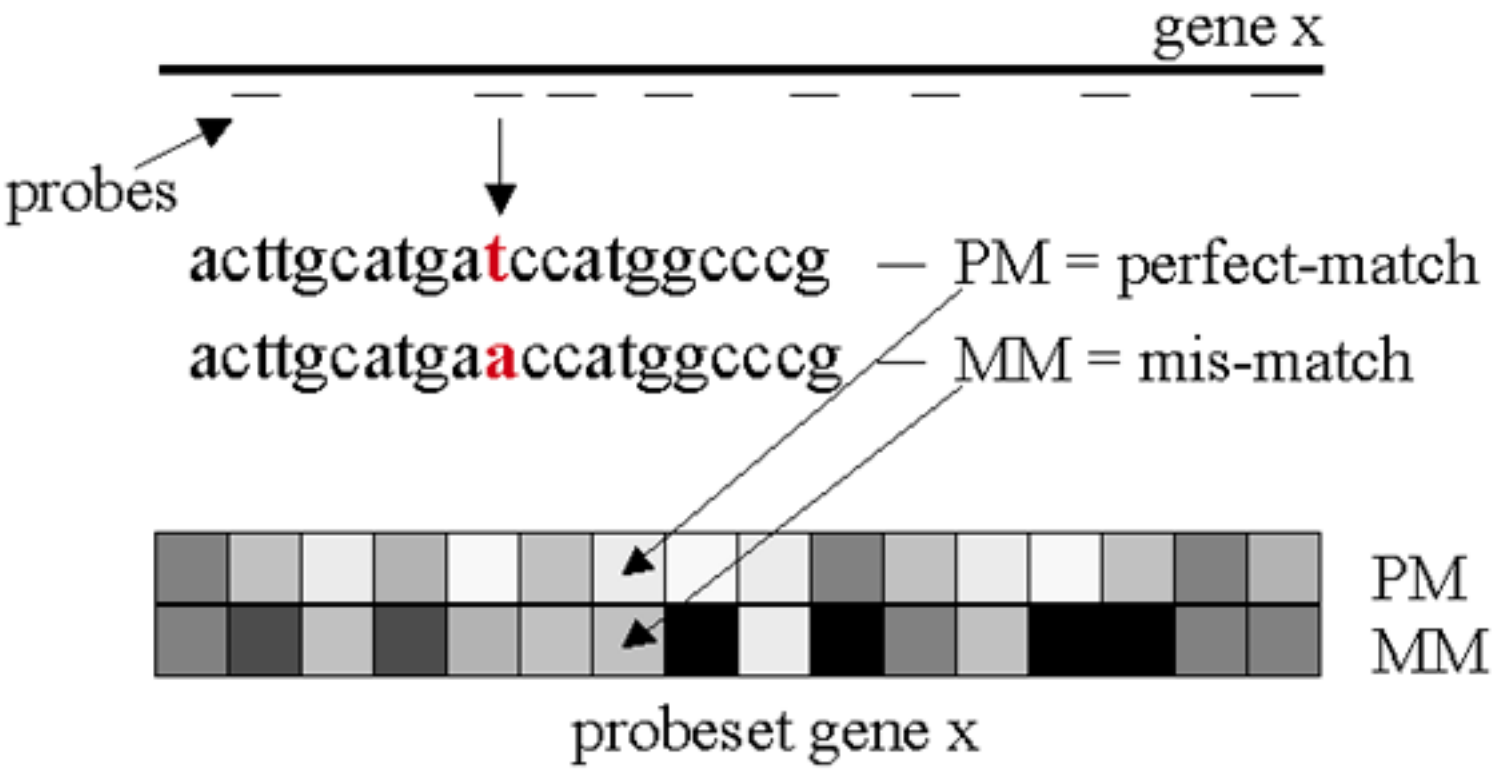


b) Esempio di microarray di oligonucleotidi per analisi differenziale dei trascritti tra popolazioni cellulari differenti



- Segnale **VERDE** se un gene è espresso per esempio solo nel tessuto sano,
- **ROSSO** se un gene è espresso solo nel tessuto tumorale e
- diverse gradazioni di **GIALLO** (**rosso** + **verde**) se un gene è espresso in entrambi i tessuti.

b) Microarrays a oligonucleotidi (Affymetrix)



- I microarrays della Affymetrix utilizzano un solo campione (1 canale).
- Utilizzano **corti oligonucleotidi di circa 20 mer.**
- Per ogni gene sono previsti coppie di oligo (probe-sets) di cui
  - uno uguale al 100% e
  - uno, che contiene una mutazione centrale, che si appaia male.
- In un **gene fortemente espresso** ci si aspetta che la **linea corrispondente al Perfect Match sia più forte.**

L'interpretazione dei dati

- Fare dei microarrays è una **tecnica costosa e delicata** ma il vero problema è quello di leggere l'informazione e darle un senso.

- Un tipico output di questa analisi è una lista lunghissima di dati comprensibili solo con l'ausilio di appositi software. Nel migliore dei casi possiamo avere, in formato testuale una lista di geni, spesso non caratterizzati.

- Il problema del recupero di informazione significativa da enormi masse di dati si è sviluppata come attività a sè stante detta **data mining** che interessa essenzialmente i bioinformatici.
- Sono stati sviluppati strumenti bioinformatici evoluti, come, per esempio GEO
- questi ci permettono di interpretare e consultare enormi set di dati con relativa facilità

- Download from the Gene Expression Omnibus at NCBI (http://www.ncbi.nlm.nih.gov/geo/)



DNA Microarrays: Applicazioni

- **I DNA microarray consentono lo screening simultaneo di migliaia di geni: screening high-throughput***
  - **Genotipizzazione genome-wide**
  - **Espressione genica tessuto-specifica**
  - **Analisi di mutazioni**

- ***Per high-throughput si intendono tutte quelle analisi scientifiche in grado di effettuare dei test**
  - su un numero molto grande di dati
  - in un tempo ristretto
  - ciò grazie a macchinari o strumentazioni automatizzate.

3. Analisi RNASeq

Campioni di interesse

Tessuto normale **Tessuto tumorale**

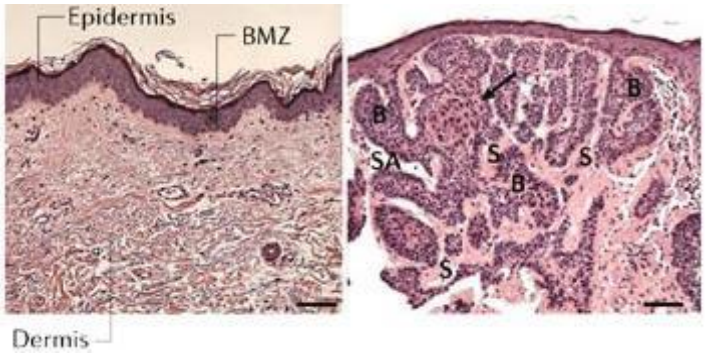
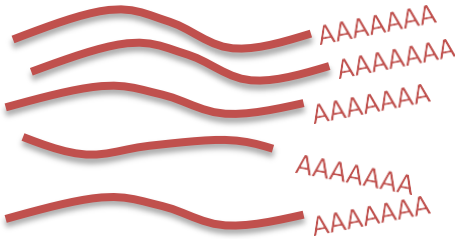
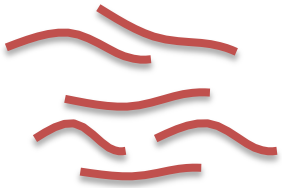


Immagine modificata da:
http://www.nature.com/nrc/journal/v6/n4/full/nrc1838.html

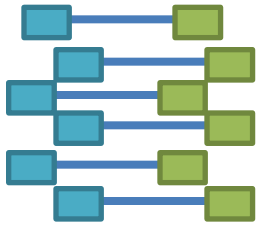
Isolamento dell'RNA/mRNA



Frammentazione chimica



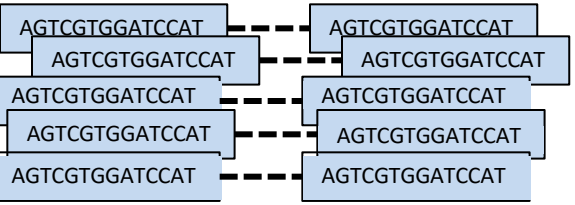
Conversione a cDNA e ligazione degli adattatori



Sequenziamento



Milioni di sequenze



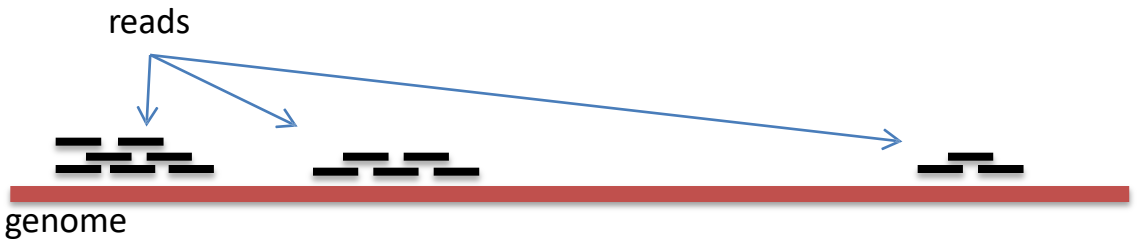
Principi dell'analisi RNASeq

- Quanto più un gene è espresso
- **tante più molecole di mRNA saranno presenti nel campione**
- **quindi tante più sequenze verranno generate.**

- E' quindi possibile utilizzare **il numero di sequenze corrispondenti a ciascun gene per ottenere una misura del livello di espressione del gene.**

- I valori di espressione ottenuti dall'RNA-Seq derivano dalla conta diretta delle read che mappano su un gene: **misura digitale**

Protocollo di analisi dati RNA-Seq



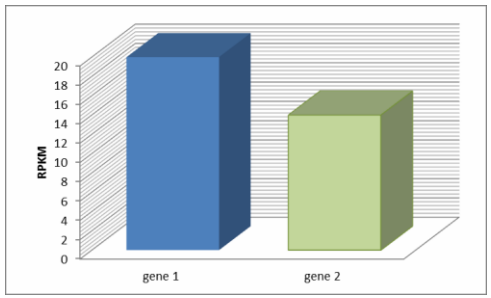
1. Allineamento su un genoma di riferimento



2. Assegnamento delle read ai geni annotati

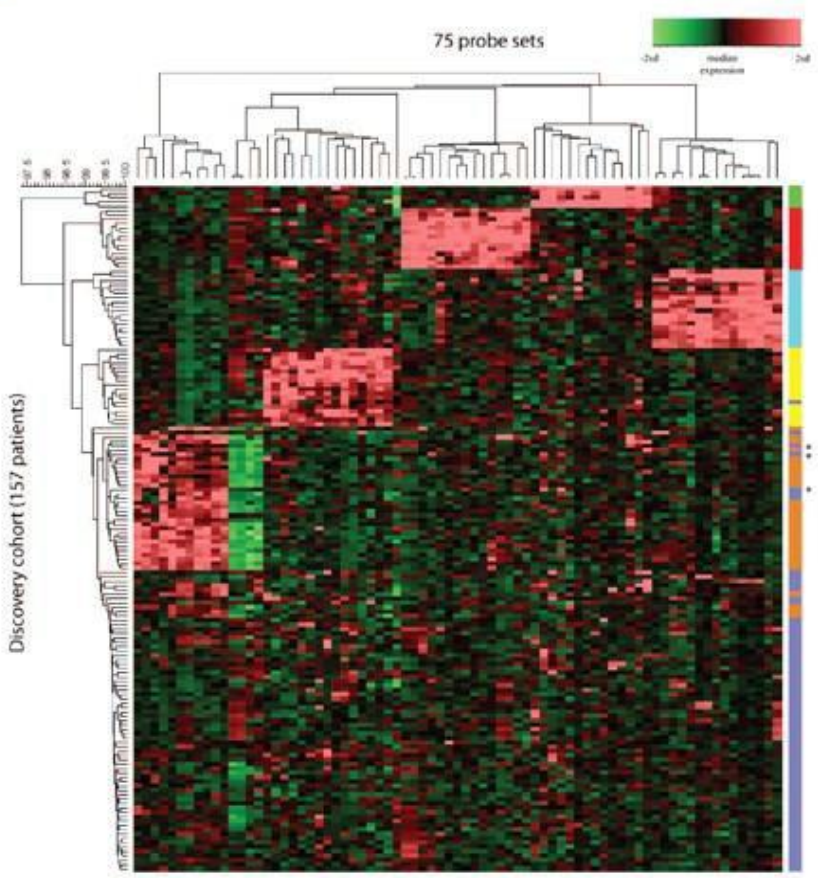


3.Rilevazione di eventuali geni "nuovi" non annotati



4. Quantificazione dell'espres-sione e analisi statistica

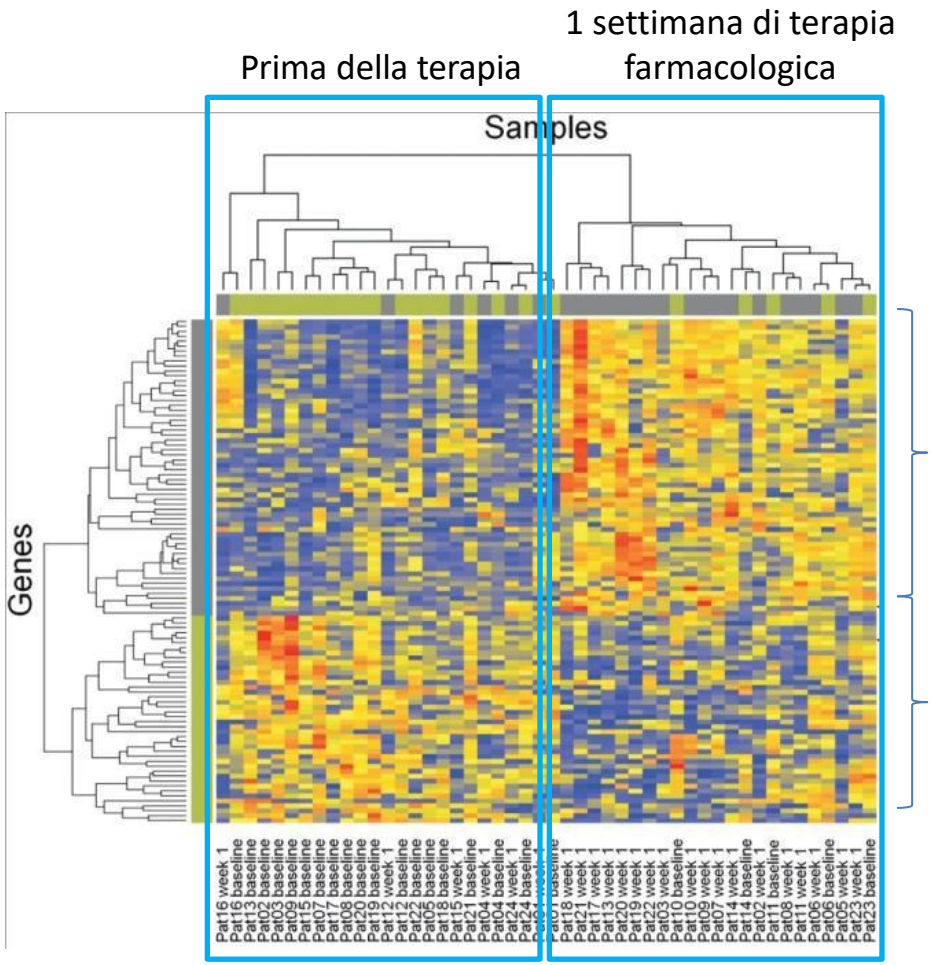
Esempio: classificazione delle malattie sulla base dei pattern di attività genica



- Clustering di dati di espressione di 75 geni (colonne) in 157 pazienti (righe).
- Il colore rosso indica un aumento del valore di espressione nel campione tumorale rispetto al normale per quel dato gene in quel paziente.
- Il colore verde indica una riduzione del valore di espressione

Brian V. Balgobind et al. Haematologica 2011;96:221-230

Esempio: caratterizzazione della risposta a terapie (farmacogenomica)



- L'heatmap mostra i livelli di espressione di 102 geni indotti o repressi in risposta al trattamento con interferone-B-1a in pazienti malati di sclerosi multipla

✓ **Geni indotti dal trattamento**

✓ **Geni repressi dal trattamento**

The Pharmacogenomics Journal (2012) **12,** 134-146; doi:10.1038/tpj.2010.77

II. Strategie molecolari per studiare la funzione di un gene

↑ Aumento espressione del gene endogeno:

- forte costitutiva
- ectopica
- inducibile

↓ Riduzione/silenziamento del gene endogeno:

- Anti-SENSO
- RNAi
- co-soppressione,

X Knock Out del gene.....

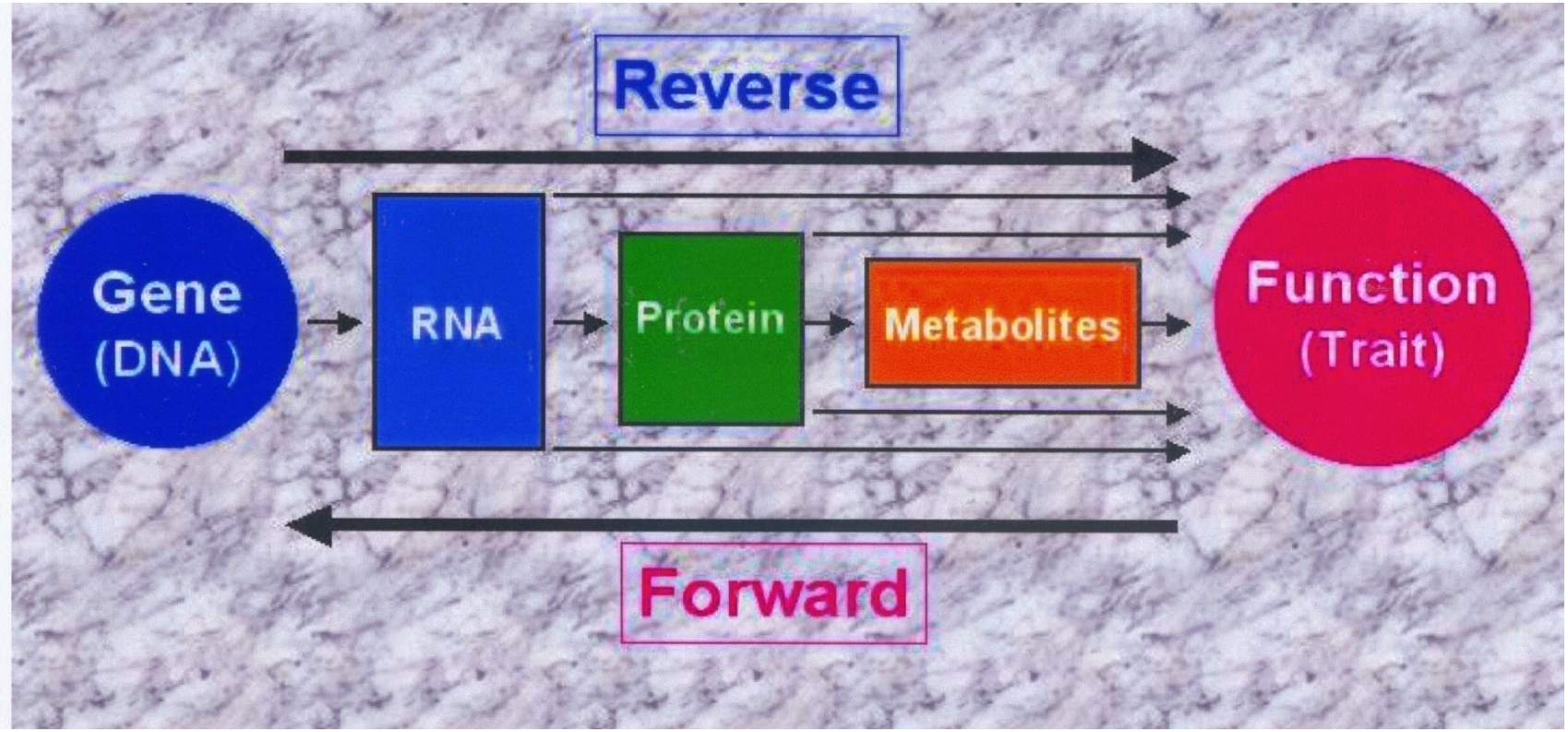
- genetica forward
- genetica reverse

Strategie genetiche per studiare la funzione genica

a. Inattivazione di un gene per identificare e caratterizzare un fenotipo

- Non tutti i geni sono “mutagenizzabili” perché:
  - **geni funzionalmente ridondanti**
    - geni a funzione simile ma non omologhi
    - ridondanza strutturale v/s ridondanza funzionale
  - **geni coivolti in diverse fasi dello sviluppo**,
    - mutazioni in tali geni possono causare fenotipi letali o pleiotropici

Functional Genomics approaches



Two different genetic approaches

Forward genetics

(from phenotype to gene)

- ID the mutant phenotype
- ID morphological, physiological, developmental differences
- ID the gene (DNA sequence)

Reverse genetics

(from gene to phenotype)

- ID a gene
- Mutate, knockout or over-express the gene
- Analyze morphological, physiological, or developmental effects (the phenotype)



Ways to inactivate genes

- Insertions
- TILLING
- RNAi

Strategie di Mutagenesi Inserzionale

a) T-DNA

- mutageno inserzionale che produce inserzioni stabili
- no “hot spots” di integrazione

b) Trasposoni

- il trasposone può essere exciso dal gene inattivato (reversione del fenotipo = complementazione)
- eventi di trasposizione avvengono in zone limitrofe al sito di inserzione

Mutagenesi da inserzione

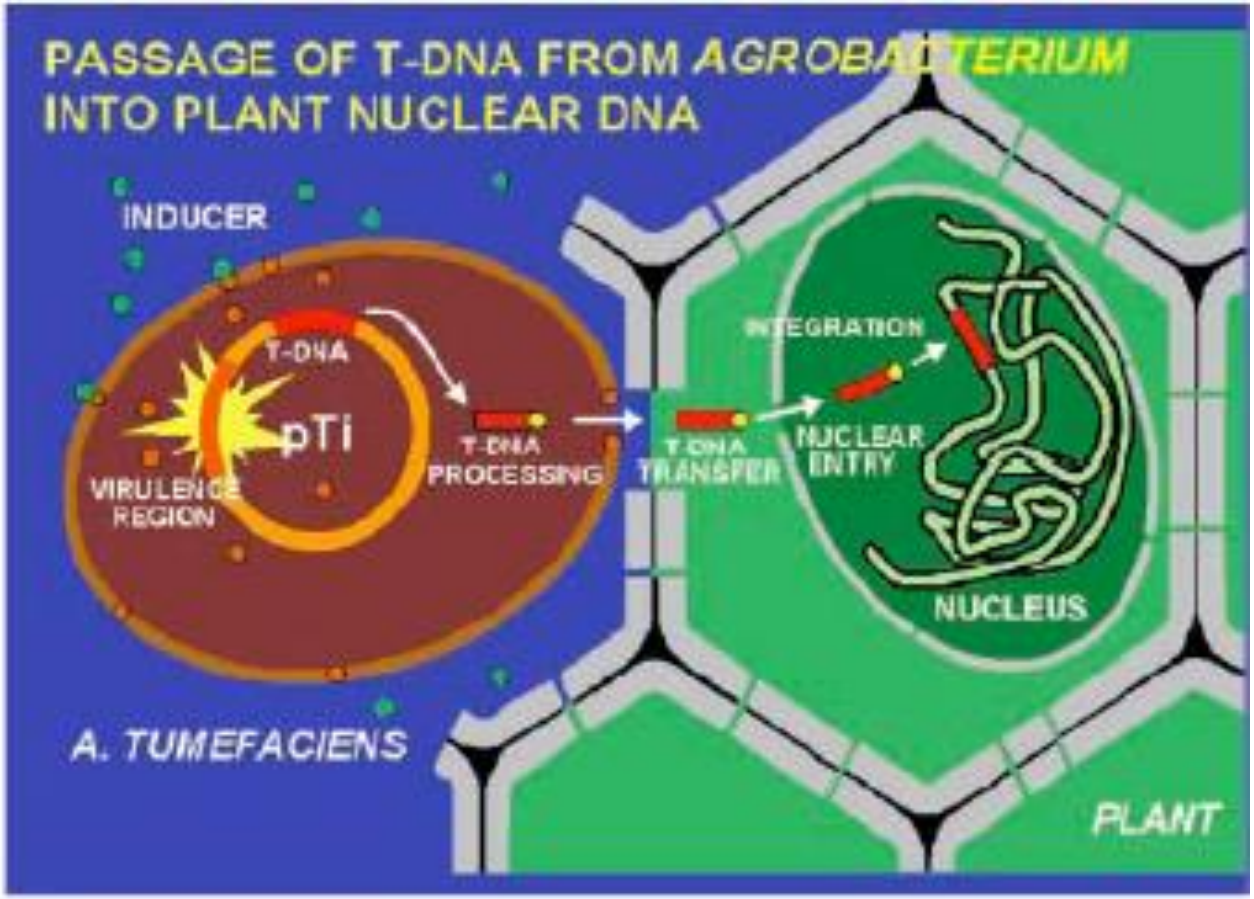
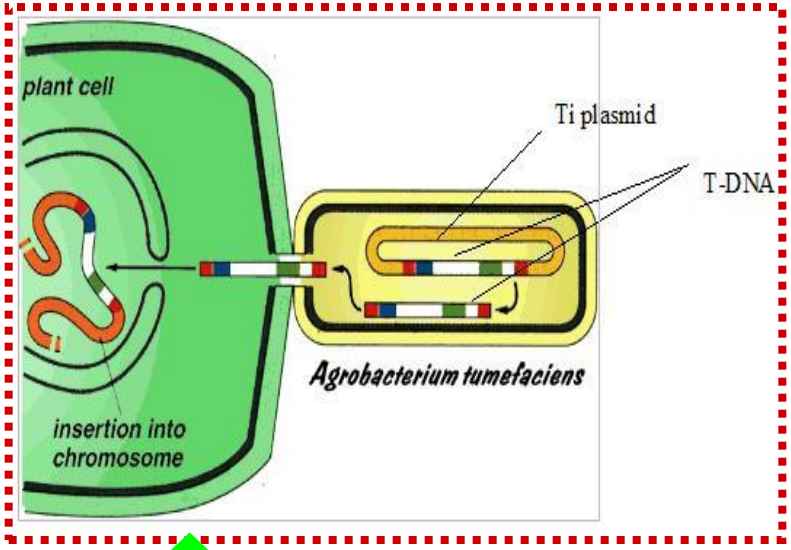


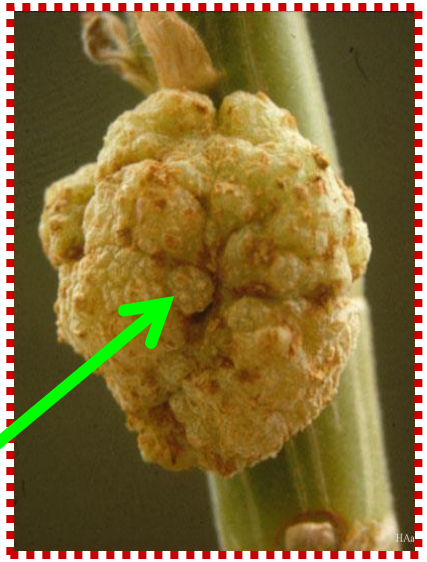
Figure 1. The path that the T-DNA travels from the stage of T-DNA processing from the Ti plasmid (pTi) to stable integration into the plant genome (Rossi et al., 1998).

Agrobacterium tumefaciens

- Soil bacterium.
- Causes crown gall tumors in dicotyledonous plants.
- T DNA (gall producing gene) occurs in Ti plasmid.
- Ti plasmid is used as vector for higher plants.
- Many genetically modified plants are produced using *A. tumifaciens.*



Mechanism



Tumor

Mutagenesi da inserzione

T –DNA

Agrobacterium tumefaciens

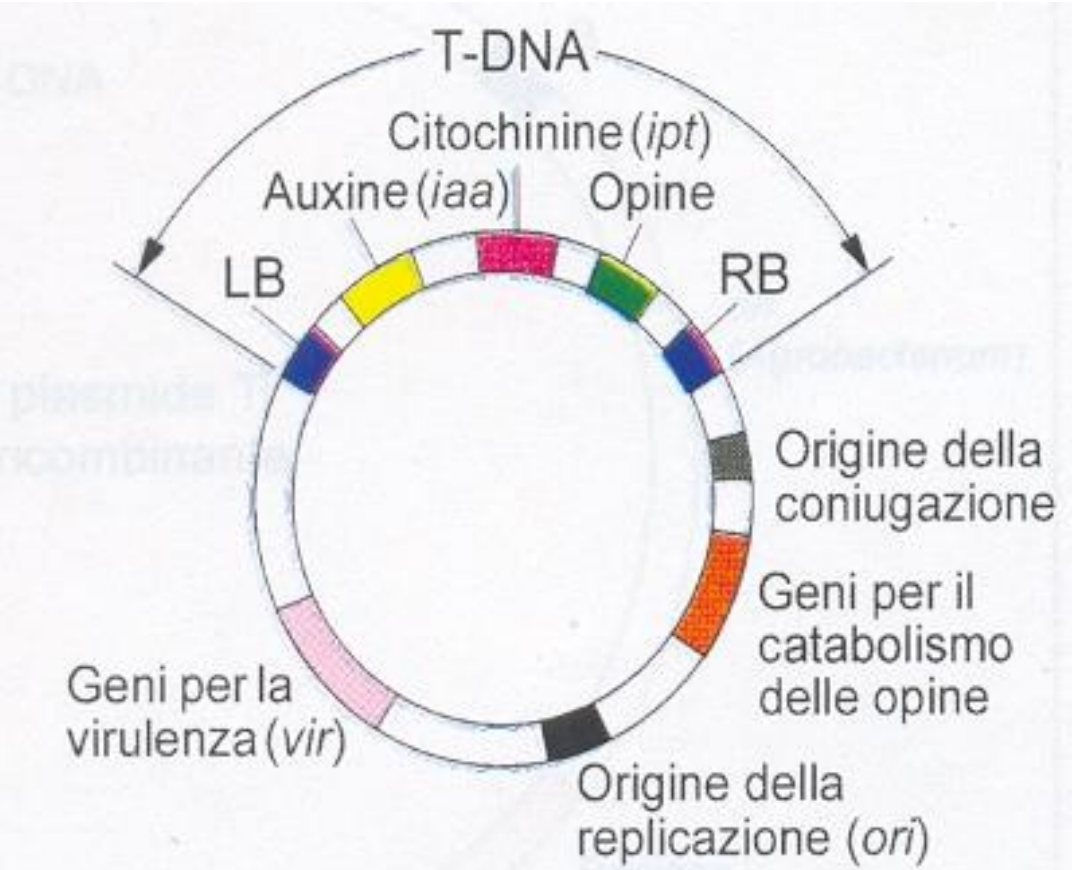


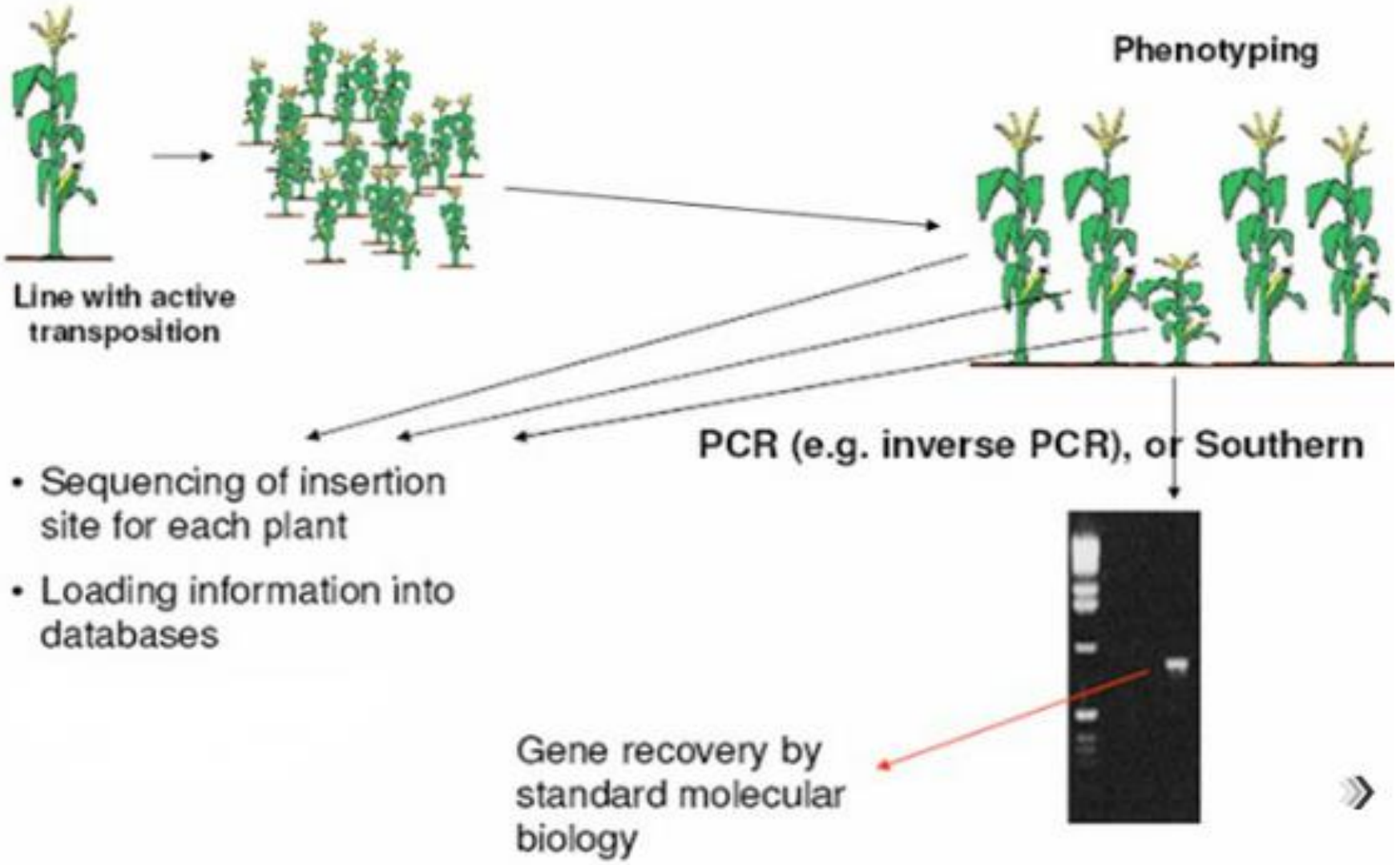
Fig. 19.29 – Rappresentazione schematica di un plasmide Ti con i principali elementi costitutivi.

Mutagenesi da inserzione: T –DNA tagging

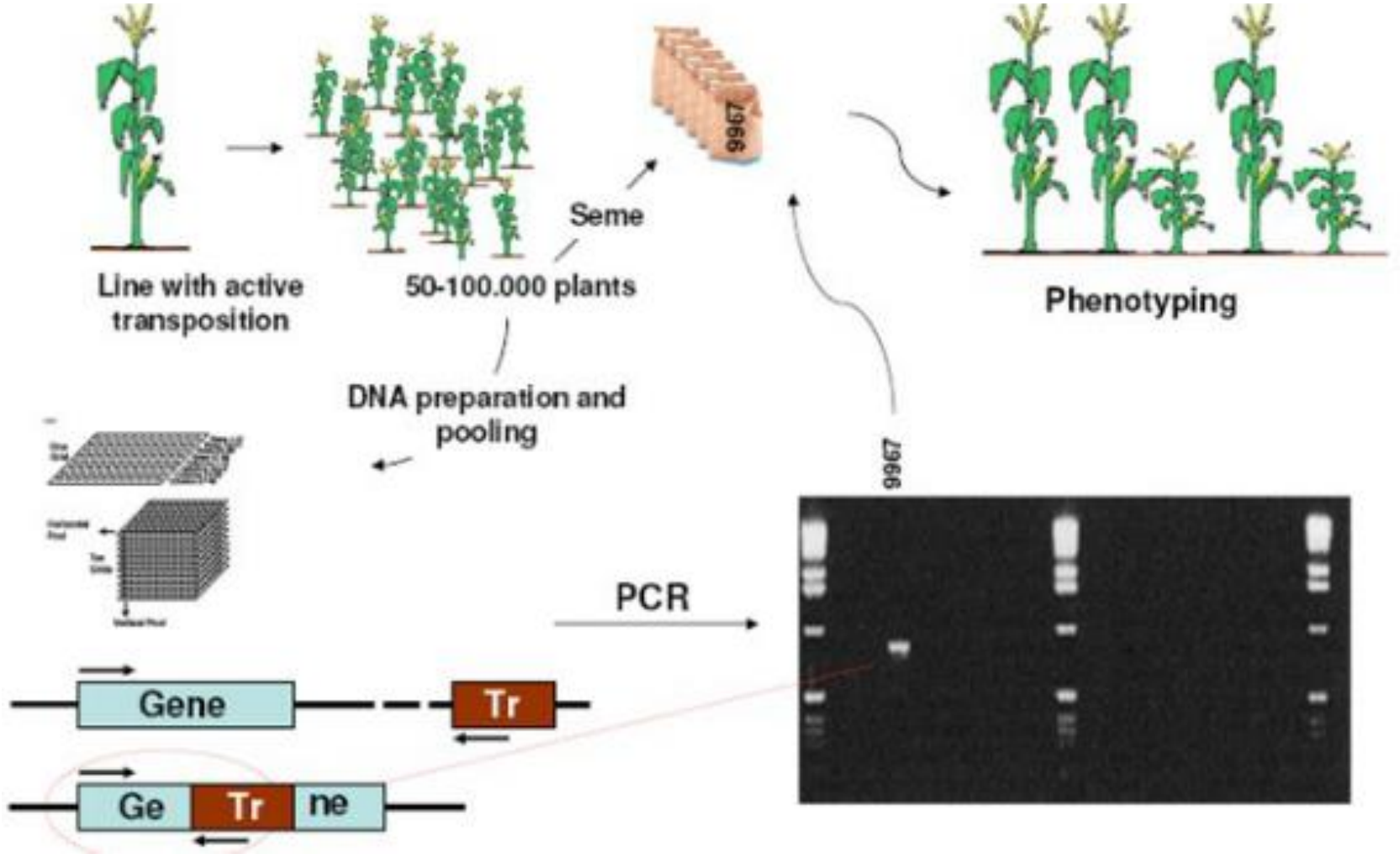
Arabidopsis: 225.000 linee indipendenti da inserzione con completa saturazione da mutagenesi del genoma

Insertion al element	Character	Number of lines	Web site	Reference
Ds	Gene trap	30 000	http://bk21.gsnu.ac.kr/cdhan	Chin *et al.* (1999)
Ds	Knock out	8 000	http://www.pi.csiro.au/fgrttpub/	Upadhyaya *et al.* (2002)
Ds	Knock out	18 000	http://www.tll.org.sg/res/sri.asp	Kolesnik *et al.* (2004)
Ds	Enhancer trap	10 000	http://genoplante-info.infobiogen.fr/oryzatagline/project.htm	Greco *et al.* (2003)
T-DNA	Gene trap	50 000	http://www.postech.ac.kr/life/pfg	Jeon *et al.* (2000)
T-DNA	Activation/ gene trap	50 000	http://www.postech.ac.kr/life/pfg	Jeong *et al.* (2002)
T-DNA	Enhancer trap	46 000	http://genoplante-info.infobiogen.fr/oryzatagline/project.htm	Sallaud *et al.* (2003)
T-DNA	Enhancer trap	42 000	http://www.genomics.zju.edu.cn/ricetdna	Chen *et al.* (2003)
Tos17	Knock out	50 000	http://tos.nais.affrc.go.jp	Miyao *et al.* (2003)

Insertional mutagenesis (forward genetics)



Insertional mutagenesis (reverse genetics)

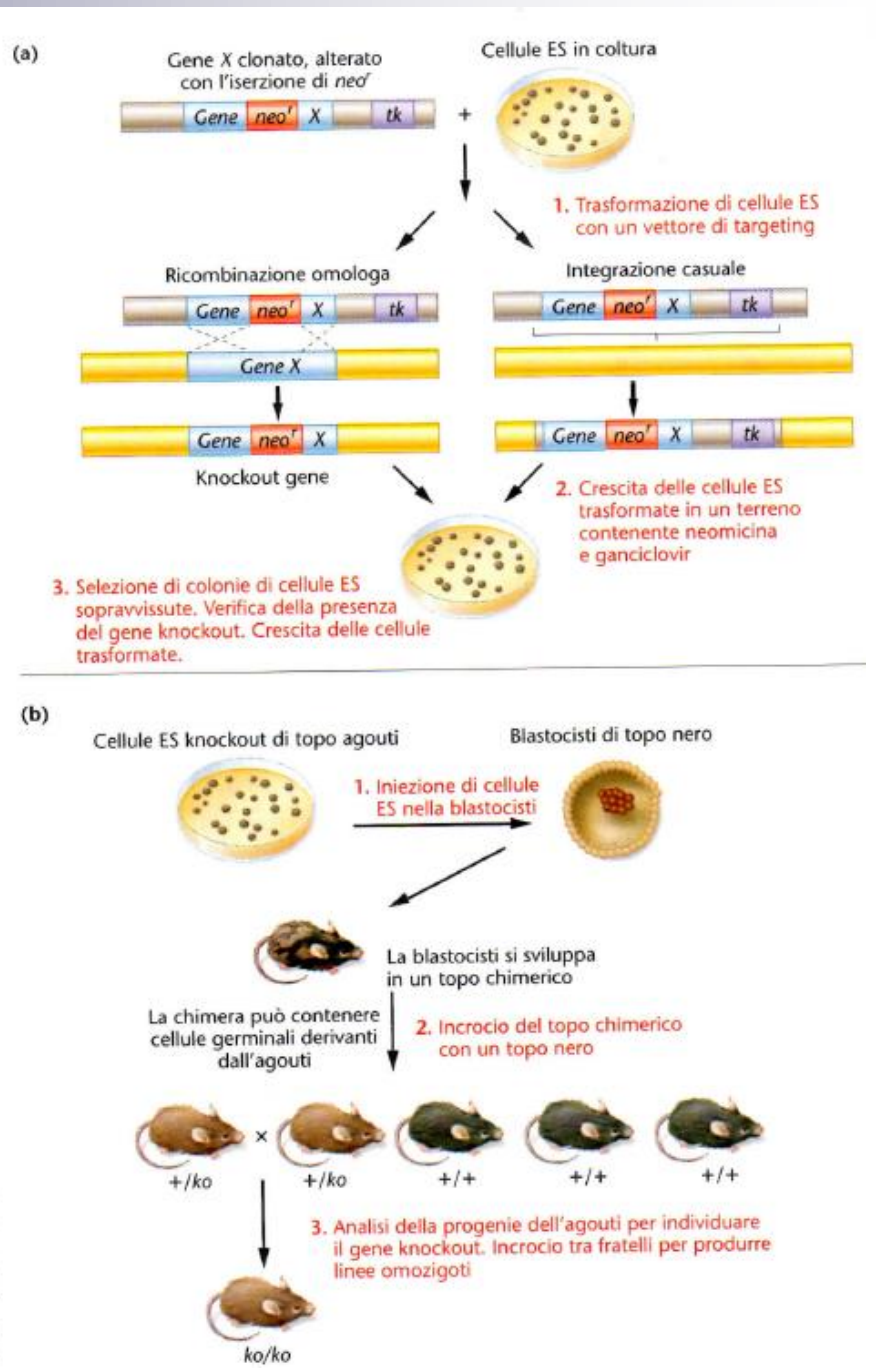


Dissezione della funzione genica: genomica e genetica

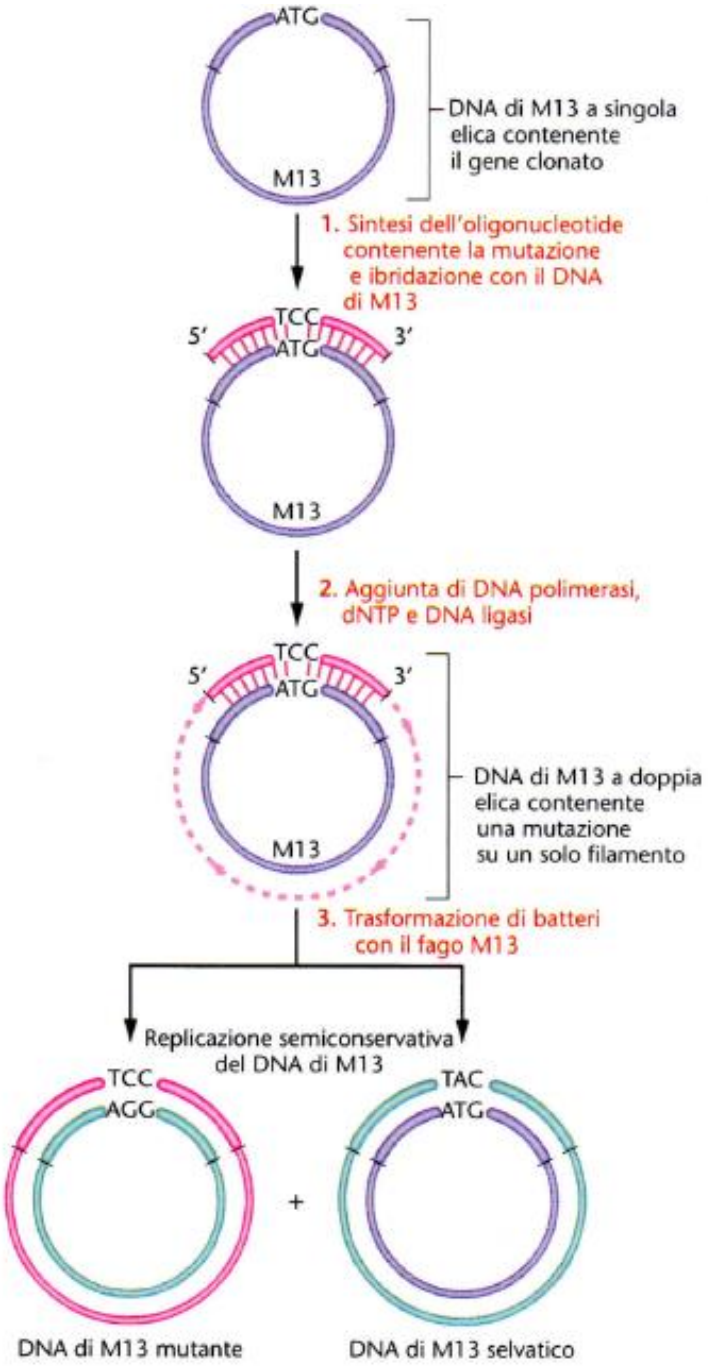
Creazione di un topo knockout.

a) Inserzione del gene knockout in cellule staminali embrionali (ES).
Le cellule ES provengono da un topo agouti (bruno).

b)
Creazione di un ceppo di topi knockout da cellule ES trasformate



Mutagenesi sito-specifica Mediata da oligonucleotide.

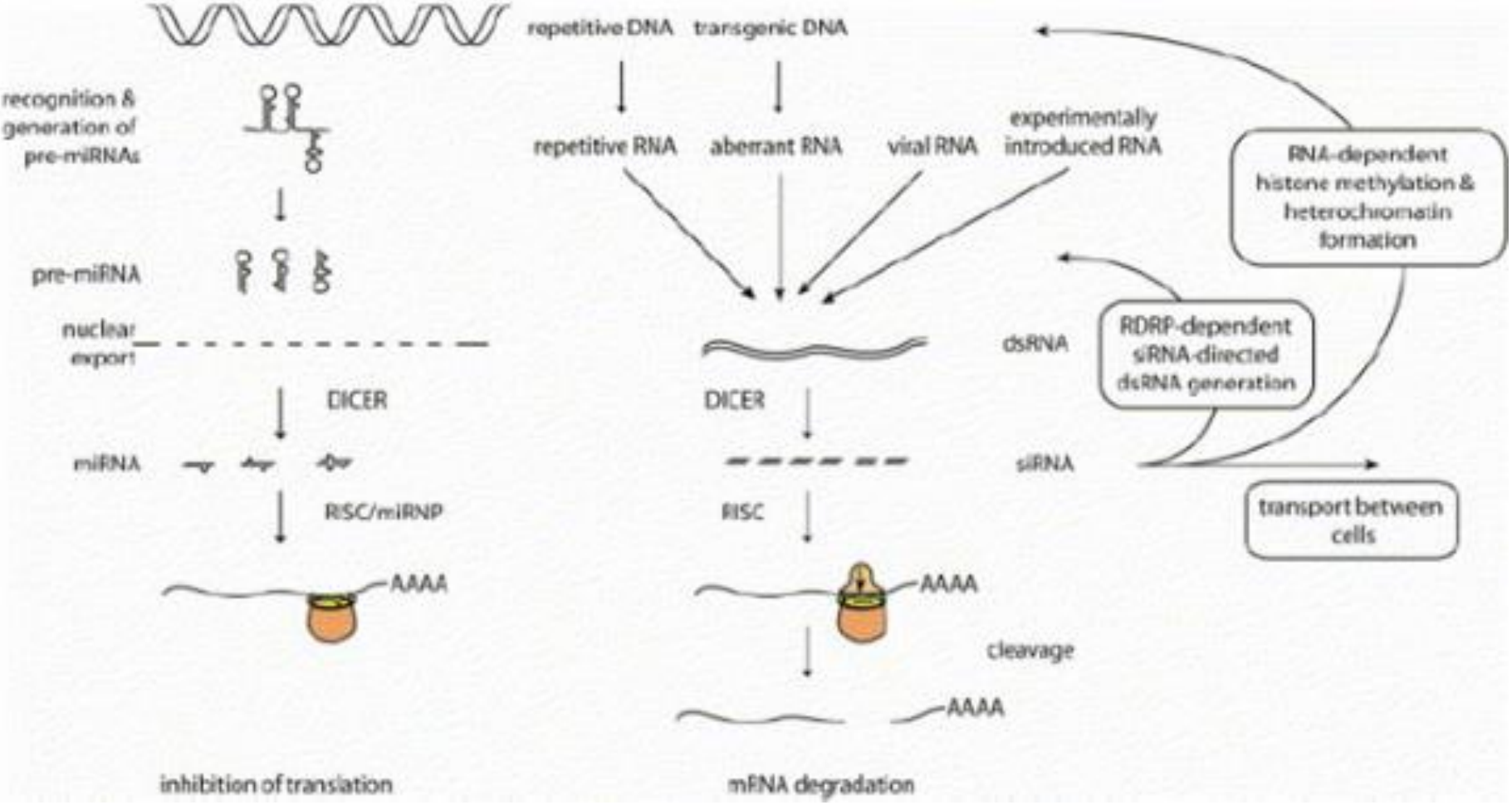


RNAi (gene knock-down)

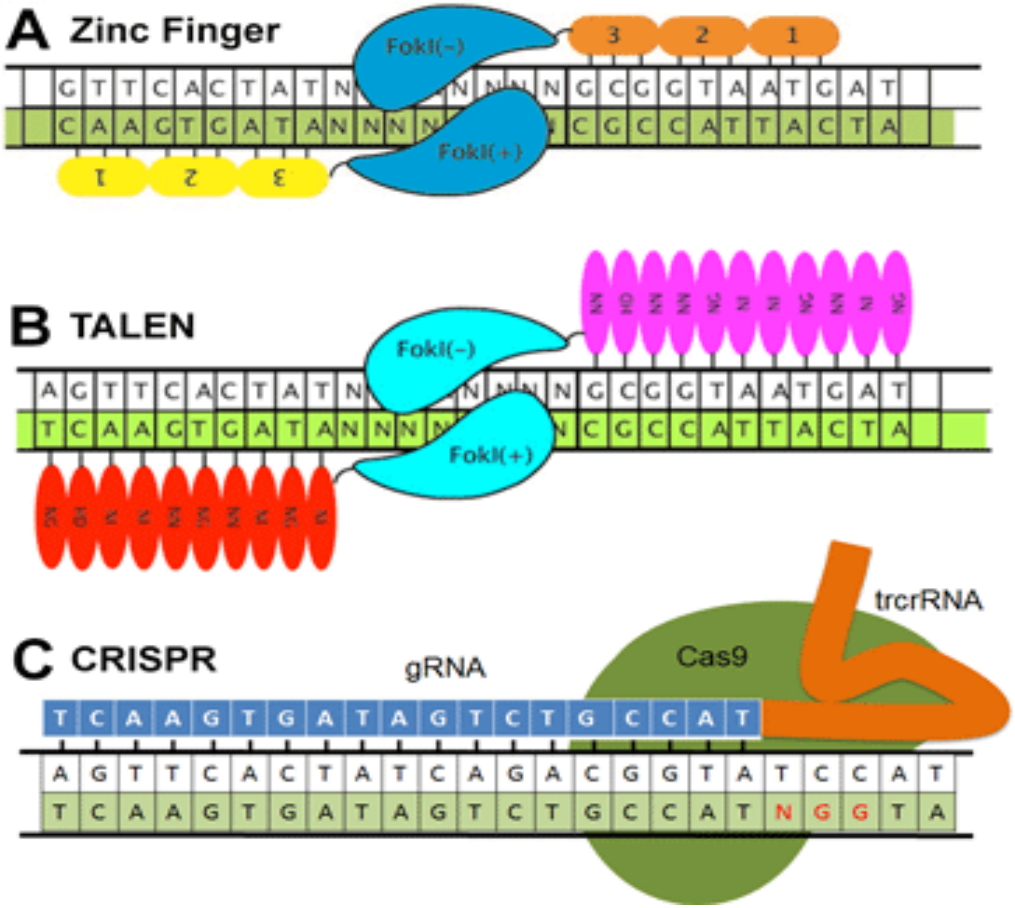
- **Mutagenesi da inserzione : tecnologia utile per genomica funzionale, ma ha limiti per :**

  - tempo richiesto per saturare un genoma.
    - knockout letali
    - mutanti non specifici
    - ridondanza genica causata da famiglie multigeniche e poliploidia
    - RNAi permette di silenziare uno, diversi., o tutti i menbri di una famiglia multigenica
    - è specifica

miRNA & RNAi (gene knock-down)

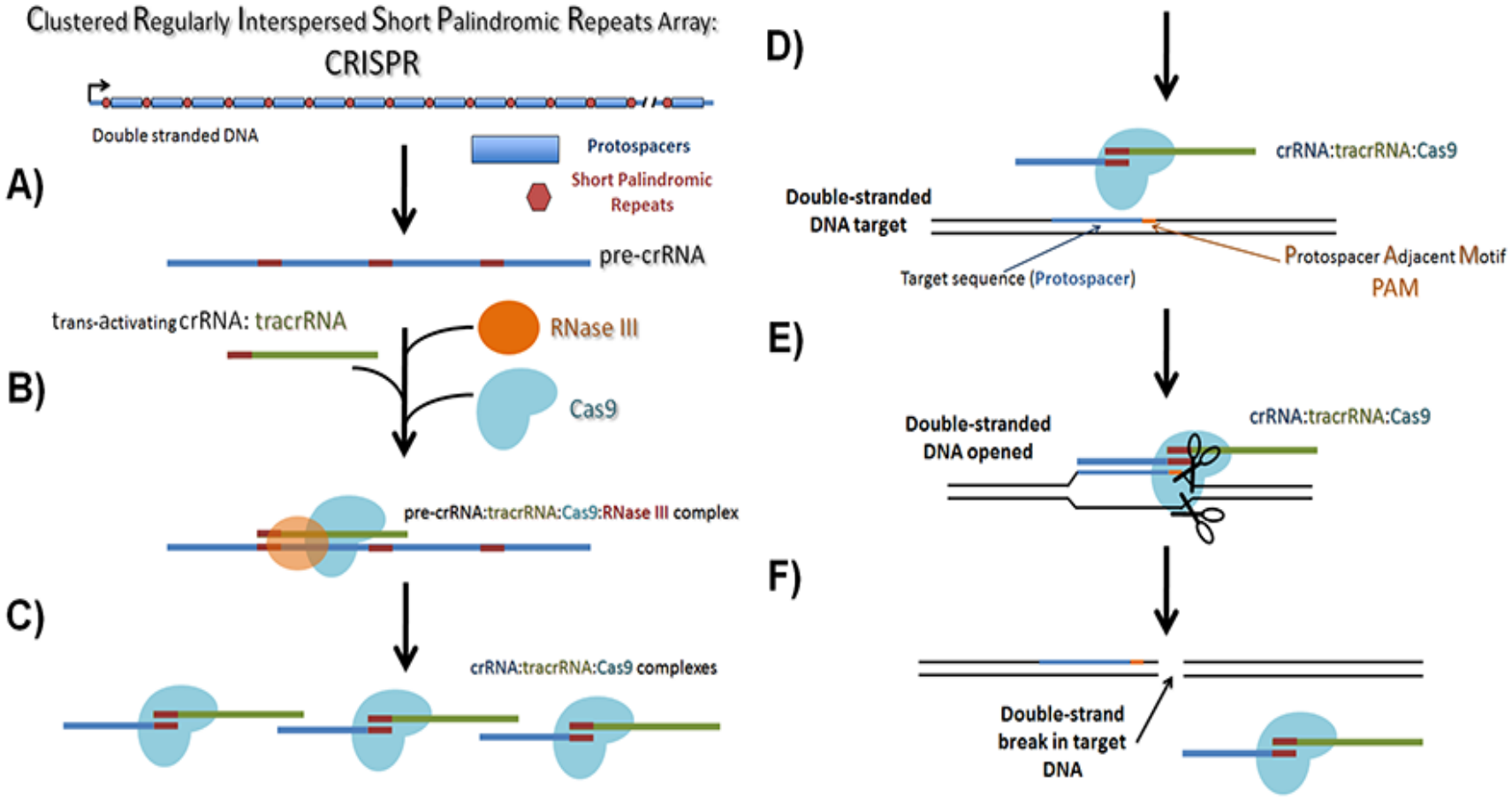


The current genome engineering technologies allow scientists to introduce double stranded breaks at specific sequences



Abbreviated list of examples of ZFN, TALEN, and CRISPR/Cas-mediated genome editing in human cells and model organisms

Type of modification	Organism	Genes	Nucleases	Refs
Gene disruption	Human	*CCR5*	ZFN	[65,91,92]
			TALEN	[25,52]
			CRISPR/Cas	[101]
	Human	TCR (T cell receptor)	ZFN	[94,95]
	Zebrafish	*gol* (Golden), *ntl* (No tail), *kra*	ZFN	[66,68]
	Pig	*GGTA1* (α1, 3-galactosyltransferase)	ZFN	[77]
		LDLR (LDL receptor)	TALEN	[76]
	Bovine	*ACAN12, p65*	TALEN	[76]
	Human	*EMX1, PVALB*	CRISPR/Cas	[102]
	Rat	*IgM, Rab38*	ZFN	[70]
	Arabidopsis	*ADH1, TT4*	ZFN	[81]
	C. elegans	*ben-1, rex-1, sdc-2*	ZFN/TALEN	[78]
	Hamster	*DHFR*	ZFN	[37]
	Drosophila	*yellow*	ZFN	[72]
	Rice	*OsSWEET14*	TALEN	[84]
Gene addition	Human	*OCT4, PITX3*	ZFN/TALEN	[45,46]
	Human	*CCR5*	ZFN	[97]
	Human	*F9* (Coagulation Factor IX)	ZFN	[86]
	Mouse	*Rosa26*	ZFN	[57]
	Human	AAVS1	ZFN	[45,96,97]
			TALEN	[46]
			CRISPR/Cas	[103]
	Human	*VEGF-A*	ZFN	[17]
	Zebrafish	*th* (tyrosine hydroxylase), *fam46c, smad5*	TALEN	[80]
	Maize	*IPK1*	ZFN	[82]
Gene correction	Human	*IL2RG*	ZFN	[44,85]
		A1AT (α_1-antitrypsin)	ZFN	[89]
		HBB (β-globin)	ZFN	[87,88]
		SNCA (α-synuclein)	ZFN	[90]
	Tobacco	*SuRA, SurRB* (acetolactate synthase)	ZFN	[83]
	Drosophila	*yellow*	ZFN	[71]



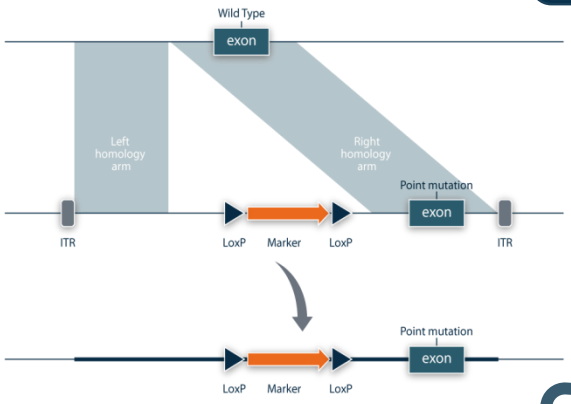
Within the bacterial genome, a CRISPR array contains many unique protospacer sequences that have homology to various foreign DNA (e.g. viral genome). Protospacers are separated by a short palindromic repeat sequence. **(A)** The CRISPR array is transcribed to make the pre-CRISPR RNA (pre-crRNA). **(B)** The pre-crRNA is processed into individual crRNAs by a special trans-activating crRNA (tracrRNA) with homology to the short palindromic repeat. The tracrRNA helps recruit the RNAse III and Cas9 enzymes, which together separate the individual crRNAs. **(C)** The tracrRNA and Cas9 nuclease form a complex with each individual, unique crRNA. **(D)** Each crRNA:tracrRNA:Cas9 complex seeks out the DNA sequence complimentary to the crRNA. In the Type II CRISPR system a potential target sequence is only valid if it contains a special Protospacer Adjacent Motif (PAM) directly after where the crRNA would bind. **(E)** After the complex binds, the Cas9 separates the double stranded DNA target and cleaves both strands after the PAM. **(F)** The crRNA:tracrRNA:Cas9 complex unbinds after the double strand break.

Genome editing

a technique where DNA is inserted, replaced or removed from a genome using artificially engineered nucleases

- **Target gene mutation**
- **Knockout gene**
- Study gene function
- Create transgenic organism
- Synthetic biology
- Gene therapy

Genome Editing Tools



Non-Nuclease

Nuclease

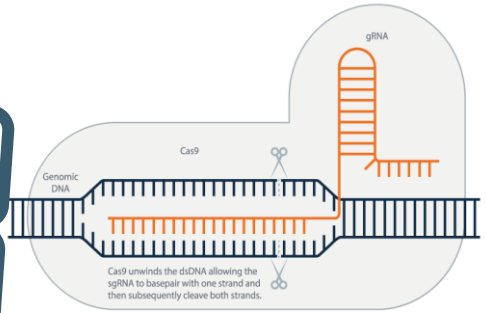
CRISPR-Cas9

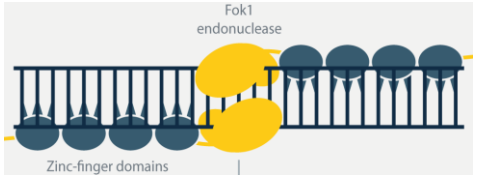
Meganucleases

TALENS

ZFNs

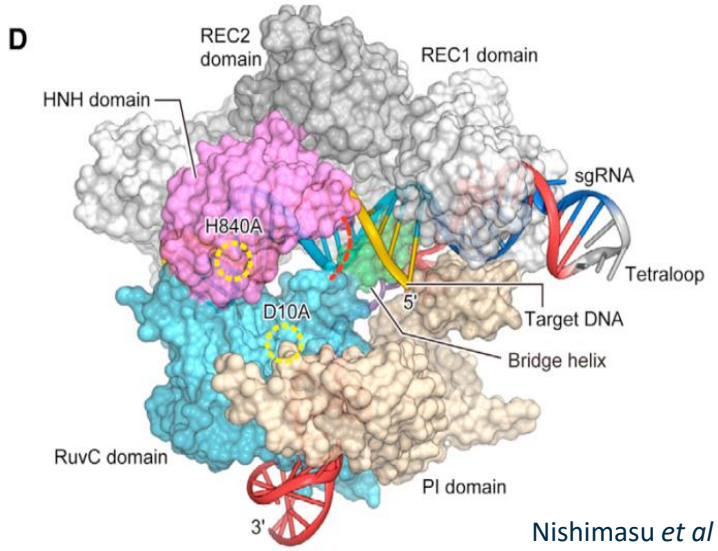
rAAV





rAAV:
adeno-associated virus -mediated gene editing

CRISPR Specificity

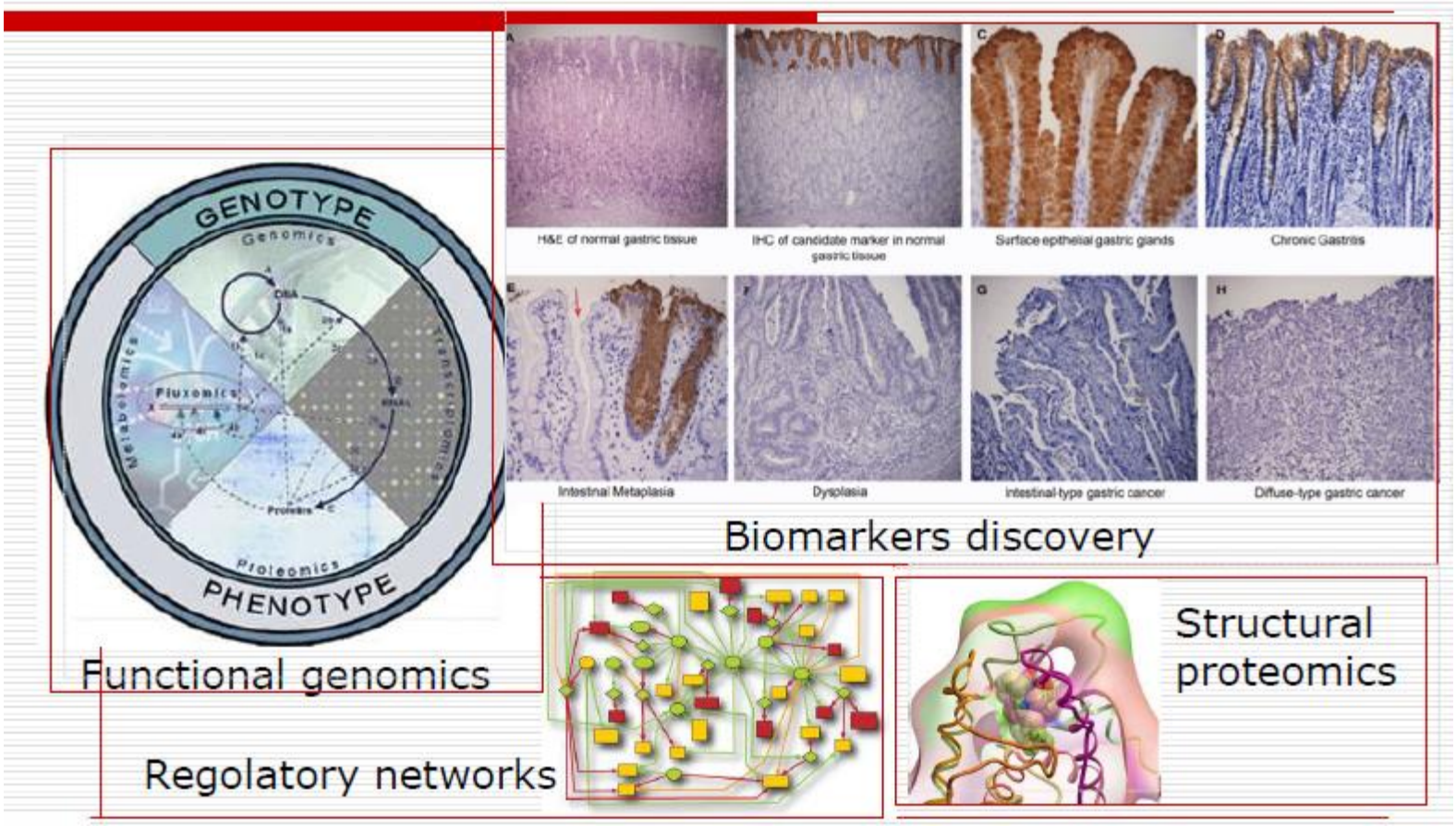


Nishimasu *et al* Cell

Cas9 Wild type	Cas9 Nickase (Cas9n)
Induces double strand break	Only “nicks” a single strand
Only requires single gRNA	Requires two guide RNAs for reasonable activity
Concerns about off-target specificity	Reduced likelihood of off-target events
High efficiency of cleavage Especially good for random indels (= KO)	Guide efficiency dictated by efficiency of the weakest gRNA

B.Proteomica:

Complemento proteico di un genoma



Proteomica:

studio delle proteine presenti in una cellula in un dato momento

Obiettivi

- **identificazione di tutte le proteine** espresse in una cellula in certe condizioni,
- **natura e l'ampiezza** di qualsiasi **modificazione post-traduzionale delle proteine,**
- natura e l'ampiezza delle **interazioni proteina-proteina nella cellula**
- **localizzazione delle proteine** nei diversi compartimenti cellulari.

Introduzione



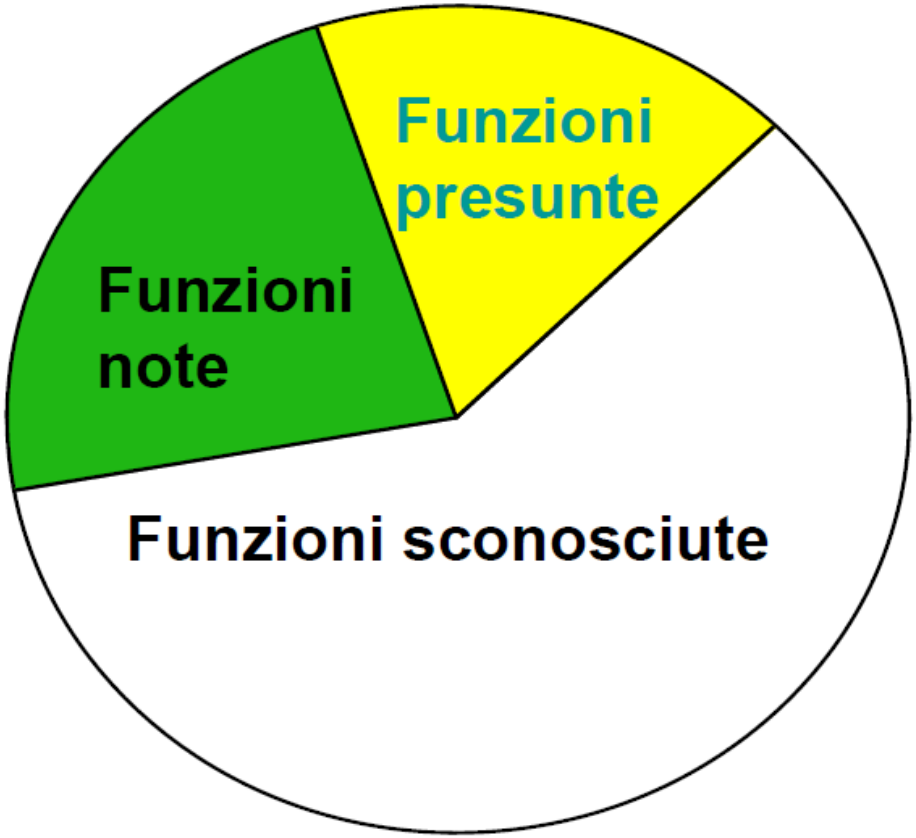
30.000 geni nel genoma umano
100.000 proteine nel proteoma umano



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

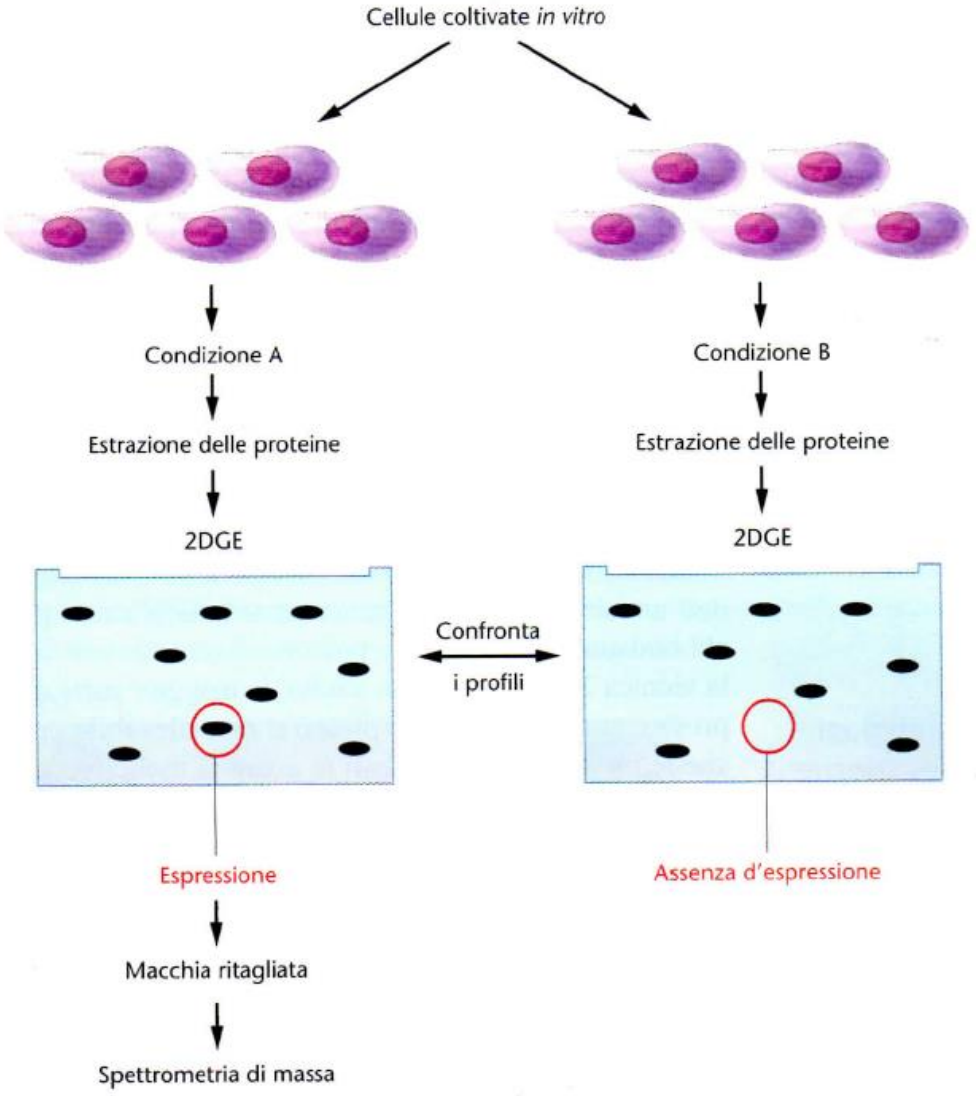
- Fino poco tempo fa i ricercatori studiavano l'espressione di un singolo gene
- Ora è possibile studiare simultaneamente l'espressione di tutti i geni di un organismo (questo può aiutare a capire meglio la funzione dei singoli geni nel contesto cellulare)

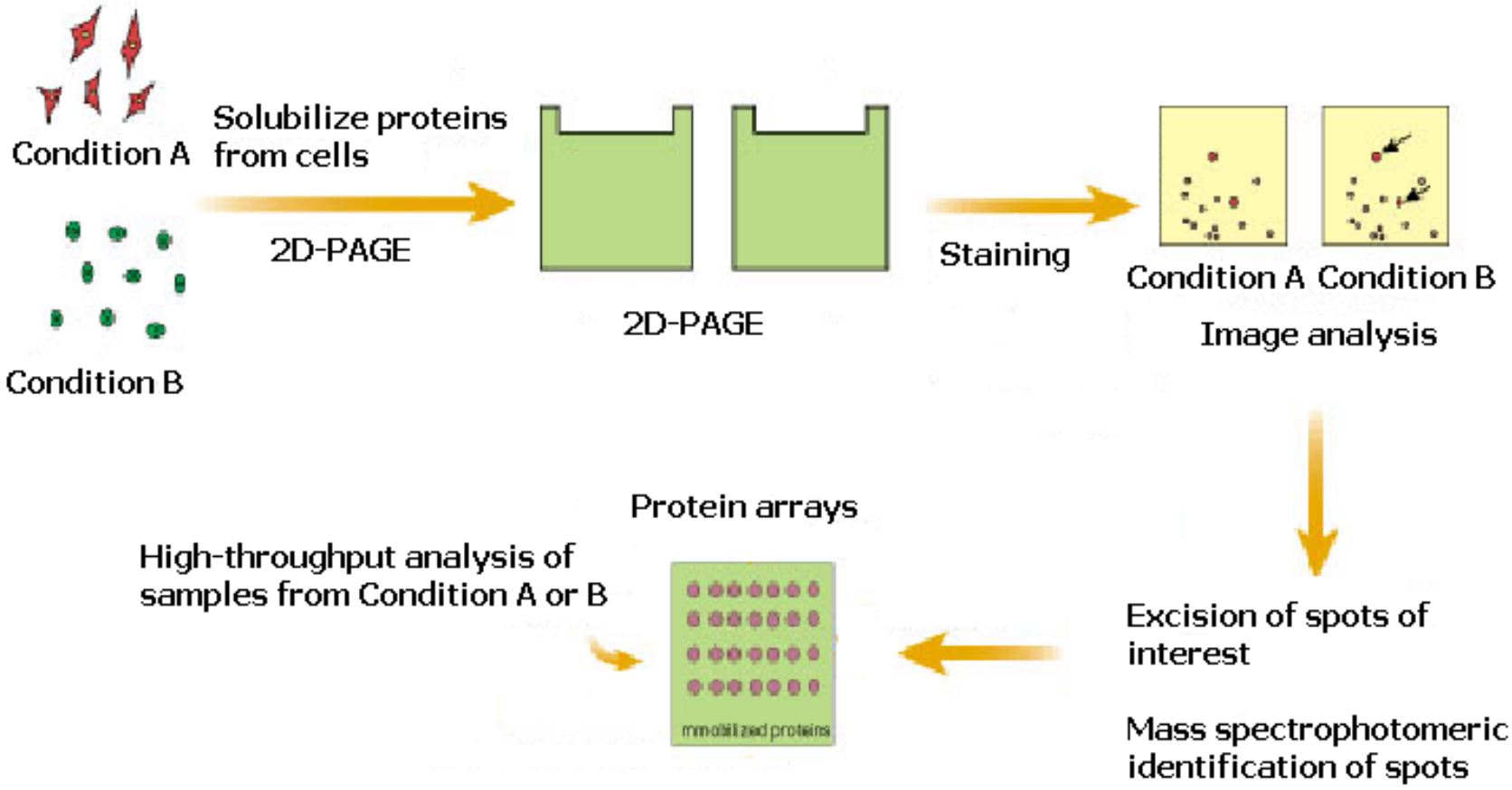
Perchè analizzare cosi tanti geni?

- <10% del geni umani sono stati studiati a livello della loro funzione individuale.
- Ma i geni totali sono circa 30,000
- I pannelli di espressione globale forniscono molte più informazioni
- soprattutto si ottengono informazioni non attese !

Proteomica: identificare e analizzare le proteine di una cellula

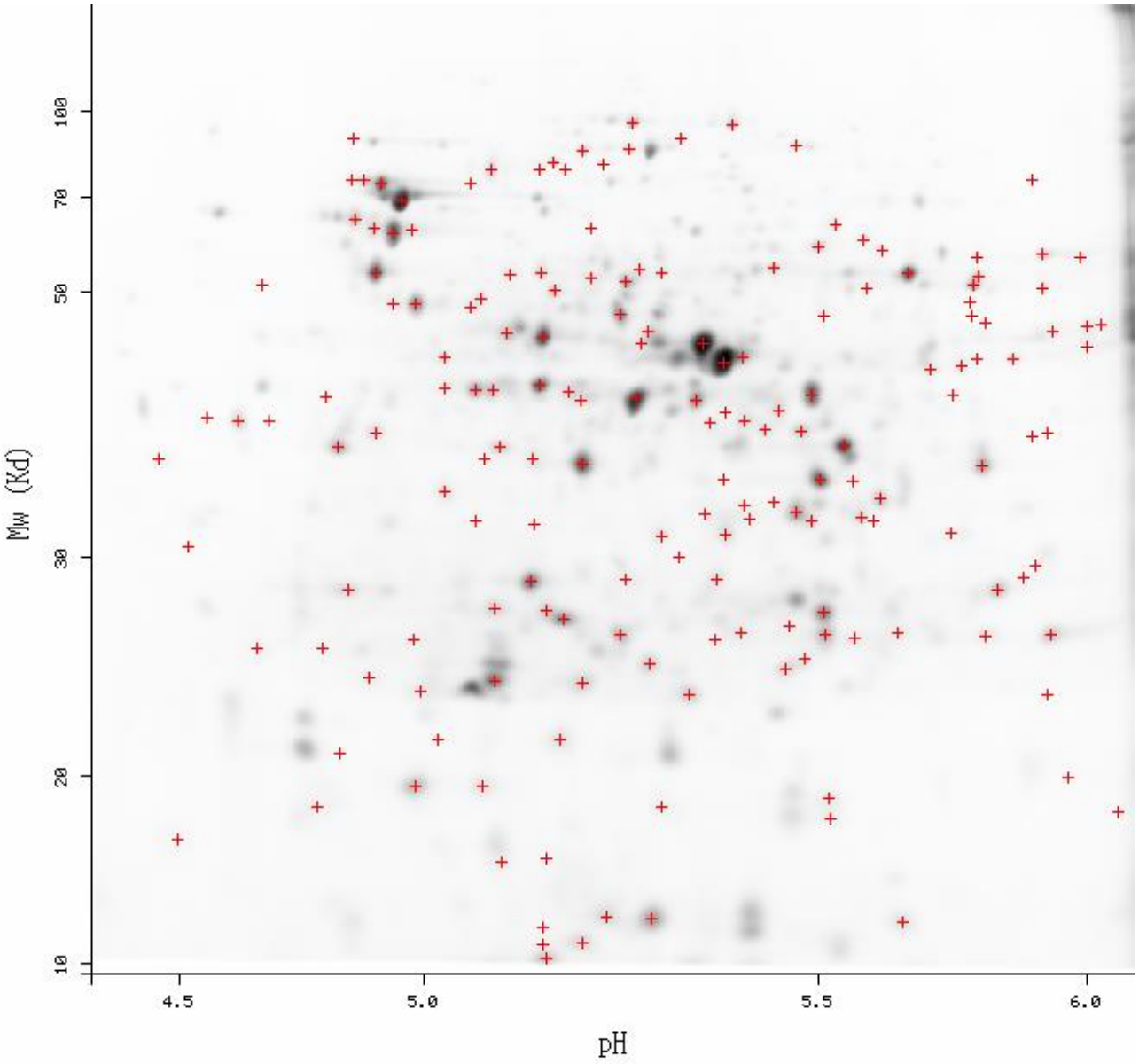
- In una tipica analisi proteomica le cellule sono esposte a due diverse condizioni (come condizioni di crescita, farmaci, ormoni).
- Dopo il trattamento,
  - le proteine sono estratte
  - separate mediante **2DGE.**
- È quindi confrontato l'insieme delle macchie ottenute per evidenziare l'espressione genica differenziale.
- Le macchie interessanti sono
  - tagliate dal gel,
  - digerite in frammenti peptidici
  - analizzate in spettrometria di massa per identificare la proteina presente nella macchia




Condition A
Solubilize proteins from cells
Condition B
2D-PAGE
2D-PAGE
Staining
Condition A
Condition B
Image analysis
Excision of spots of interest
Mass spectrophotomeric identification of spots
Protein arrays
High-throughput analysis of samples from Condition A or B
immobilized proteins

Il proteoma




Mw (Kd)
100
70
50
30
20
10
4.5
5.0
5.5
6.0
pH

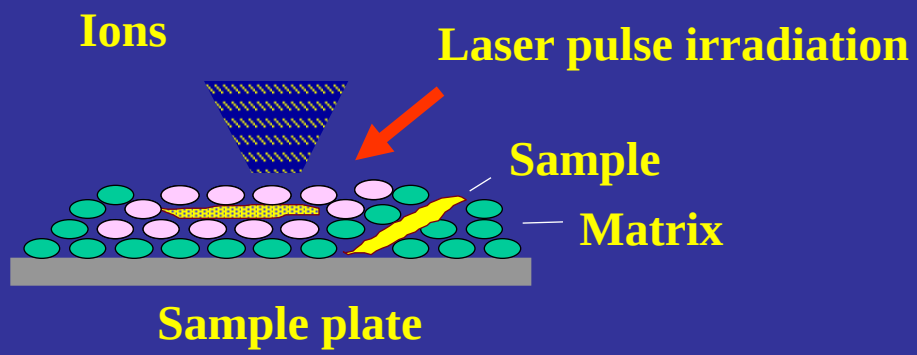
Spettrometria di massa (quantitativa -con applicazioni qualitative)

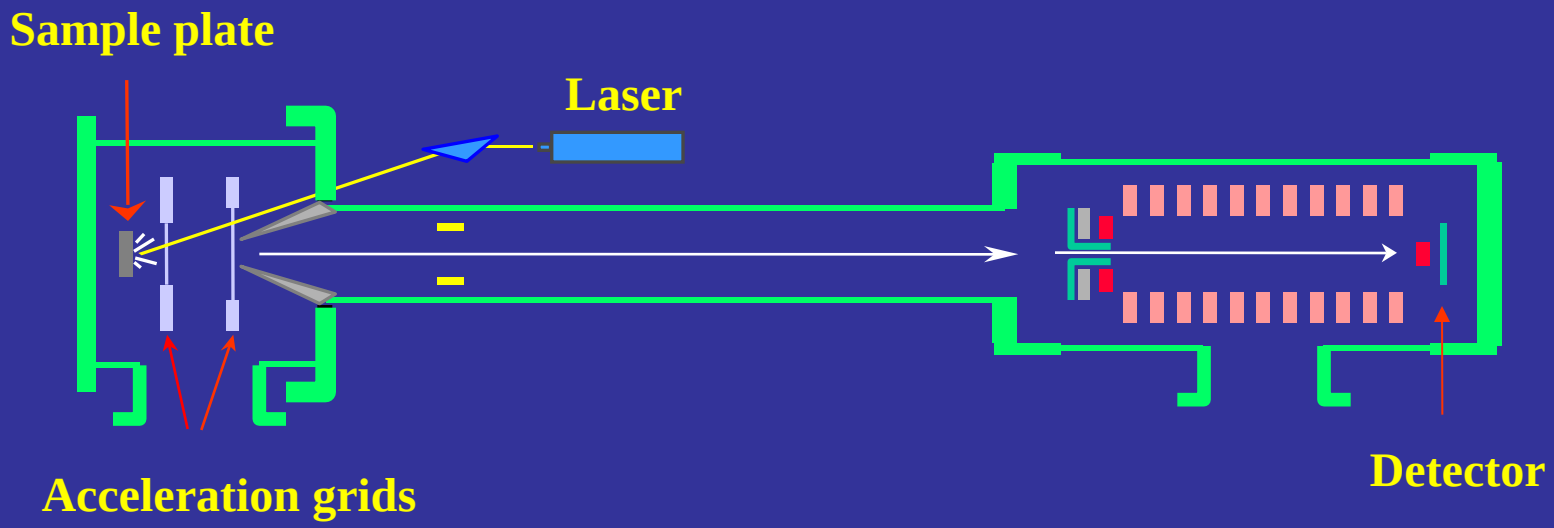
- Nella **SPETTROMETRIA DI MASSA** le molecole sono ionizzate e poi accelerate nel vuoto da un campo elettrico
- A secondo del tipo di ionizzazioni si ottengono spettri a picco singolo o multiplo. I primi sono utili per determinare le masse molecolari accurate, mentre i secondi servono a determinare altre proprietà molecolari

- Le particelle sono discriminate in vario modo sulla base del diverso rapporto tra massa e carica

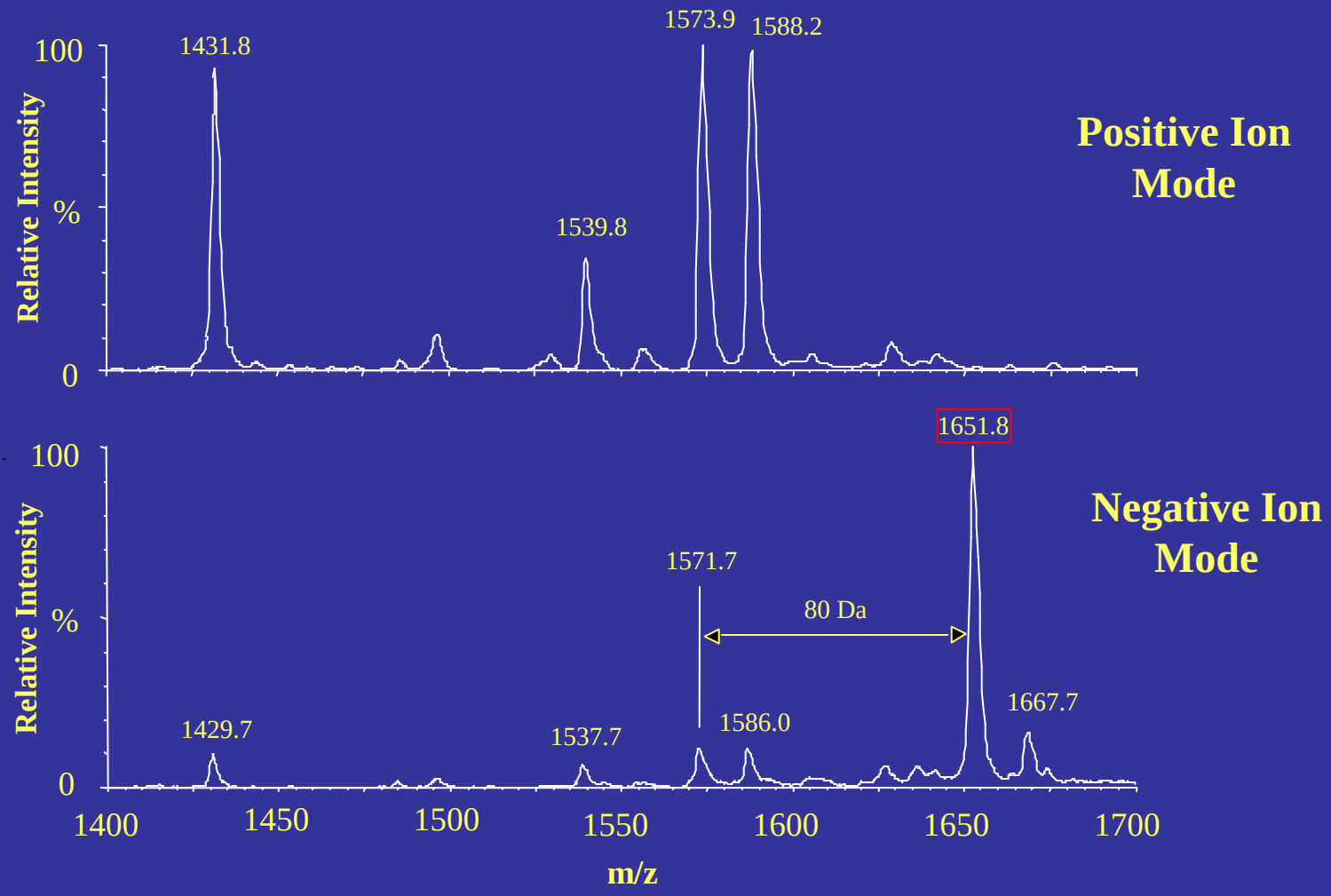
- I dati ottenuti possono essere utilizzati per:
  - calcolare il peso molecolare esatto di una molecola
  - ottenere informazioni sulla sua struttura ed eventuali modifiche post-traduzionali
  - determinare l'abbondanza di specie isotopiche

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

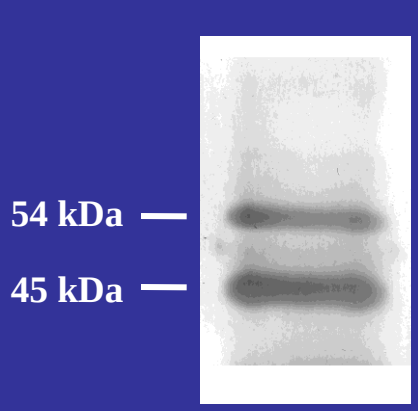




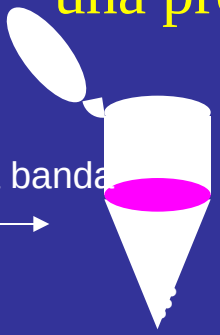
MALDI-TOF MS of Phosphopeptides



Caratterizzazione delle modificazioni post-traduzionali di una proteina



Preleva la banda

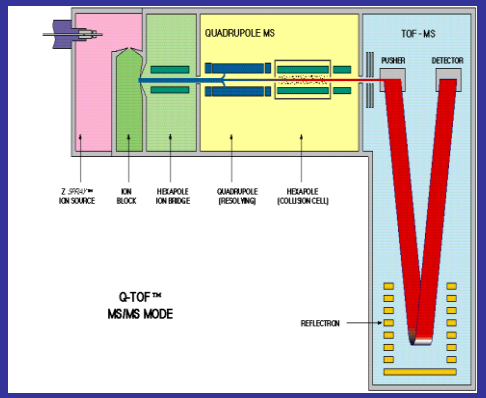


Incuba con tripsina, estrai I peptidi e de-salifica



MKKCTILVVASLLLVNSLLPGYGQNKIIQA
QRNLNELCYNEGNDNKLYHVLNSKNGKIYN
RNTVNRLLPMLRRKKNEKKNEKIERNNKLK
QPPPPPNPNDPPPPNPNDPPPPNPNDPPPP
NPNDPPPPNANDPPPPNANDPAPPNANDPA
PPNANDPAPPNANDPAPPNANDPAPPNAND
PAPPNANDPPPPNPNDPAPPQGNNNPQPQP
RPQPQPQPQPQPQPQPQPQPRPQPQPQPGG
NNNNKNNNNDDSYIPSAEKILEFVKQIRDS
ITEEWSQCNVTCGSGIRVRKRKGSNKKAED
LTLEDIDTEICKMDKCSSIFNIVSNSLGFV
ILLVLVFFN

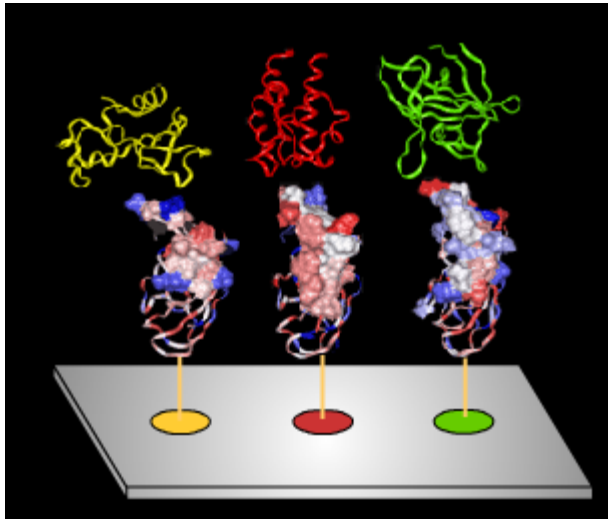
Determina la massa molecolare dei peptidi



Confronta I risultati ottenuti con quelli attesi

Gli array di proteine

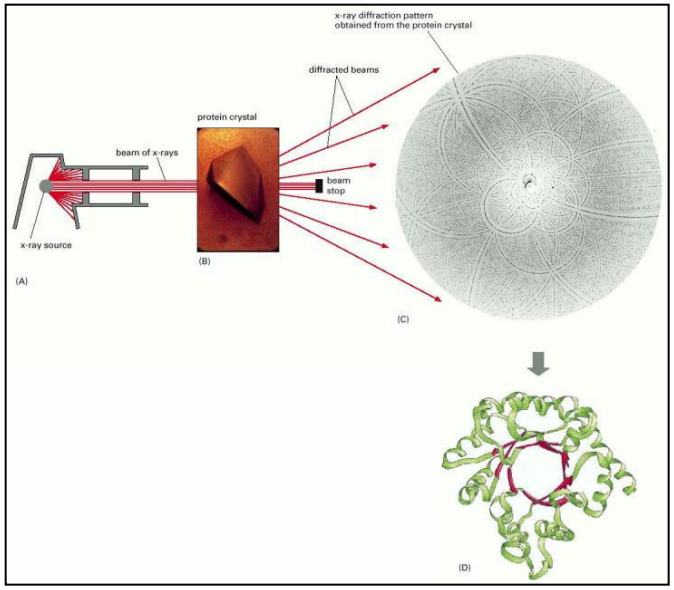
- Legame di piccole molecole
- Modificazioni post-traduzionali
- Interazioni proteina-proteina
- Interazioni proteina DNA
- Saggi enzimatici
- Mappatura degli epitopi



Genomica strutturale

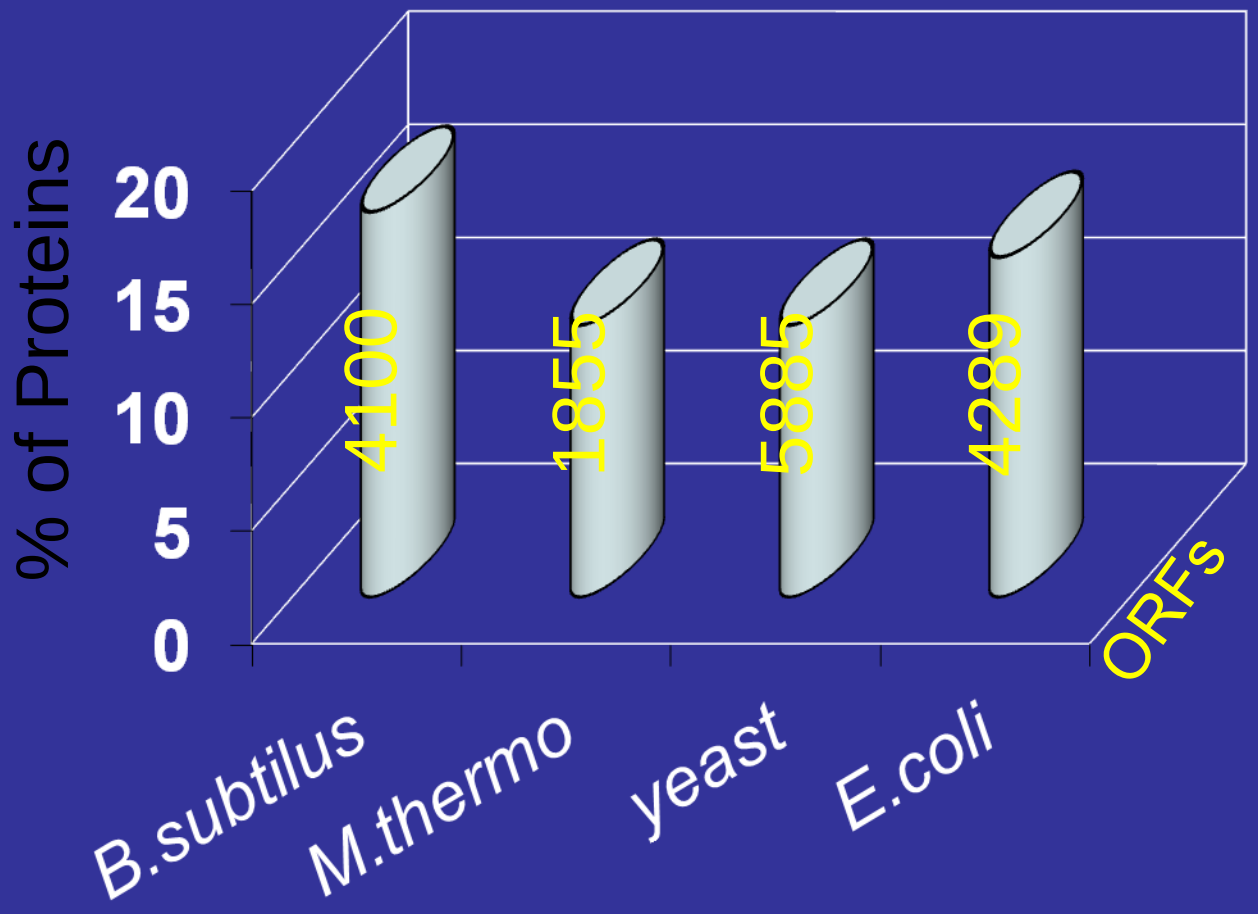
Obiettivi:

- **determinare la struttura 3D** delle proteine codificate da un intero genoma.



- Ovvero studio sistematico della **struttura 3D delle proteine mediante spettro-scopia con raggi-X**
- Studio sistematico delle interazioni tra proteine
- E' una naturale estensione dei proget-ti di sequenziamento genomico
- **L'obiettivo** è quello di **capire le funzioni molecolari e cellulari di tutti i prodotti genici.**
- Questo dovrebbe aiutare la progettazione razionale per es. di farmaci e sistemi diagnostici

PROTEINE CON OMOLOGHI 3D



Biologia dei sistemi

- Una delle conseguenze della massa di informazioni che i progetti genomici stanno producendo consiste in un nuovo modo di affrontare i problemi biologici in maniera integrata.
- L'approccio classico è sempre stato quello di studiare di affrontare i problemi biologici in maniera sostanzialmente indipendente; si studiavano singoli geni o singole proteine.
- In realtà il comportamento di queste molecole è probabilmente collegato ed integrato a quello di numerosi altri fattori nel contesto di reti biologiche ben integrate.

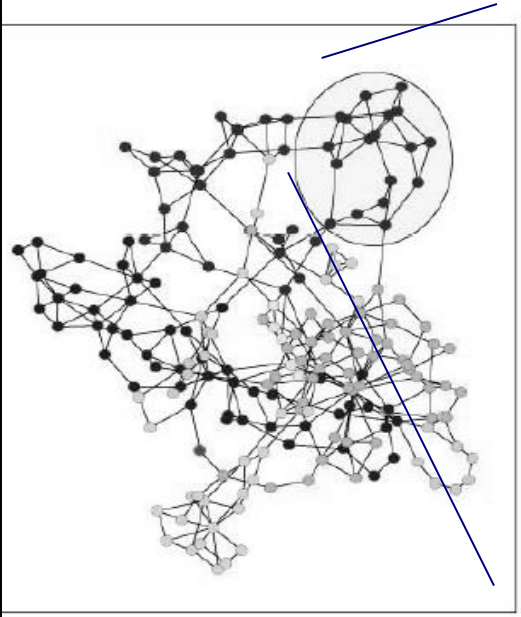
- Utilizzando la quantità di informazione che si sta accumulando a seguito dei programmi di sequenziamento genomico e gli strumenti bioinformatici che si stanno sviluppando per cercare di dare un senso a questi dati, si stanno imponendo approcci olistici allo studio dei genomi, la cosidetta
- **"Systems Biology" che cerca di affrontare i problemi biologici analizzando, in modo integrato, processi complessi.**

What is Systems Biology?

- identify *elements* (genes, molecules, cells, …)
- ascertain their *relationships* (co-expressed, interacting, …)
- *integrate* information to obtain **view of system as a *whole***

Large (genomic) systems

- many uncharacterized elements
- relationships unknown
- *computational analysis* should:
  - improve annotation
  - reveal relations
  - reduce complexity

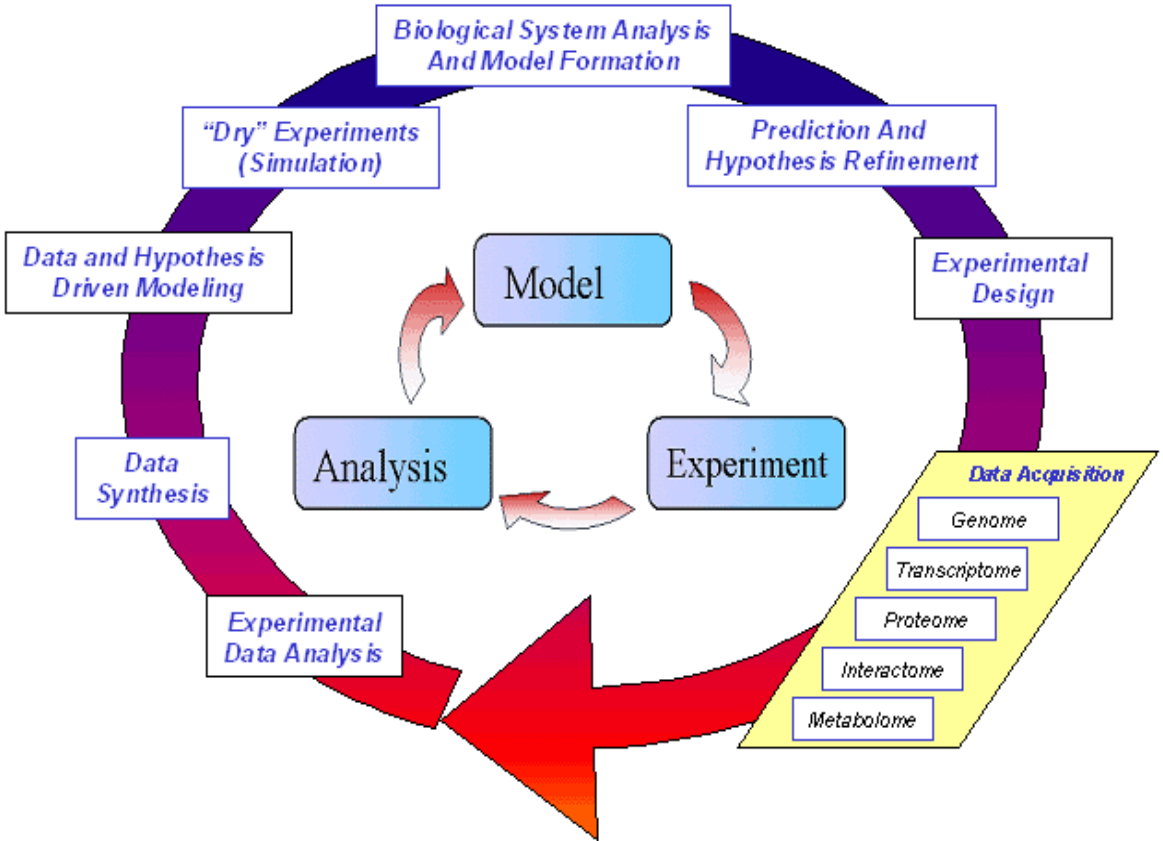


Small systems

- elements well-known
- many relationships established
- *quantitative modeling* of systems properties like:
  - Dynamics
  - Robustness
  - Logics

Biologia dei sistemi

- **Si basa nella computazione e nella modellizzazione dei diversi dati sperimentali.**
- **Interazioni dinamiche tra le varie molecole (proteine e acidi nucleici).**
- La crescita della bioinformatica ha facilitato la computazione dei diversi dati scientifici.



Obiettivi

- modelli per svelare i meccanismi che causano fenotipi alterate e elaborare nuove terapie e farmaci.

- Strumenti predittivi per la progettazione di cellule con proprietà desiderate economico ed affidabile.

- Medicina individualizzata e predittiva.

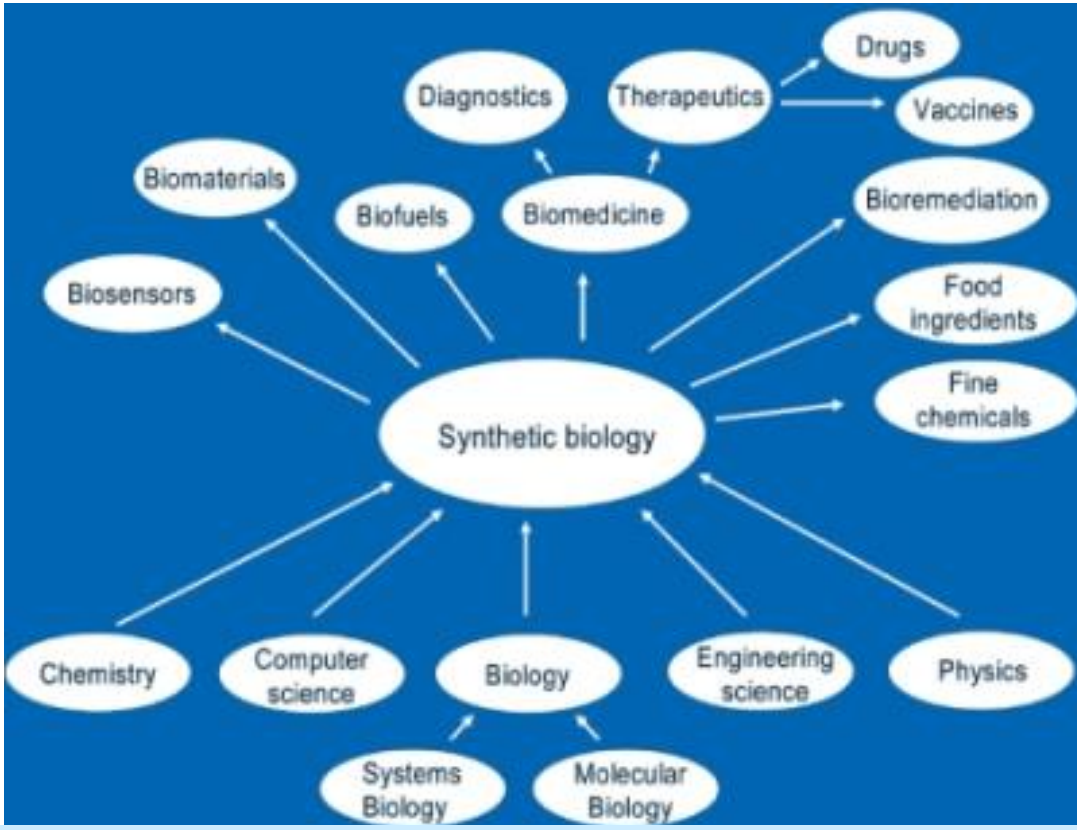
Synthetic biology

Industrial Synthetic Biology

- Progettazione e costruzione di strumenti e sistemi biologici.
- E' un settore di ricerche biologiche e tecnologiche che **combina biologia e ingegneria**,

pertanto spesso si sovrappone con la bioingegneria e ingegneria biomedica.

- Si avvale di una varietà di differenti strategie, metodologie, e discipline focalizzate sull'ingegneria biologica e biotecnologia.



- La biologia di sintesi mira a
  - **progettazione e sviluppo di nuovi componenti biologici,** quali ad esempio enzimi, circuiti genetici, e cellule
  - o la **riprogettazione di sistemi biologici** esistenti utilizzando DNA ing. e metodologie matematiche/logiche

Science, 25 MARCH 2016 • VOL 351 ISSUE 6280

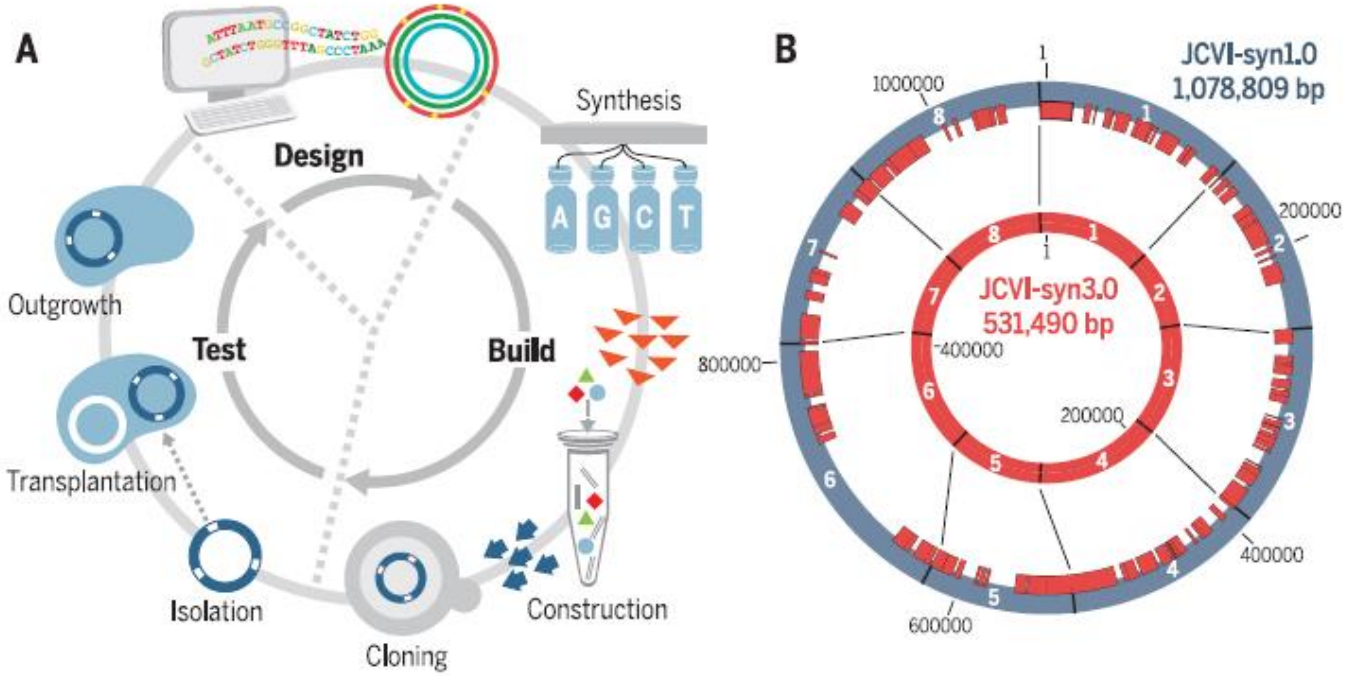
RESEARCH ARTICLE

SYNTHETIC BIOLOGY

Design and synthesis of a minimal bacterial genome

Clyde A. Hutchison III,[1*†] Ray-Yuan Chuang,[1†‡] Vladimir N. Noskov,[1] Nacyra Assad-Garcia,[1] Thomas J. Deerinck,[2] Mark H. Ellisman,[2] John Gill,[3] Krishna Kannan,[3] Bogumil J. Karas,[1] Li Ma,[1] James F. Pelletier,[4§] Zhi-Qing Qi,[3] R. Alexander Richter,[1] Elizabeth A. Strychalski,[4] Lijie Sun,[1||] Yo Suzuki,[1] Billyana Tsvetanova,[3] Kim S. Wise,[1] Hamilton O. Smith,[1,3] John I. Glass,[1] Chuck Merryman,[1] Daniel G. Gibson,[1,3] J. Craig Venter[1,3*]

‘**Whole genomes can now be built from chemically synthesized oligos and brought to life by installation into a receptive cellular environment.** We have applied whole-genome design and synthesis to the problem of minimizing a cellular genome’.

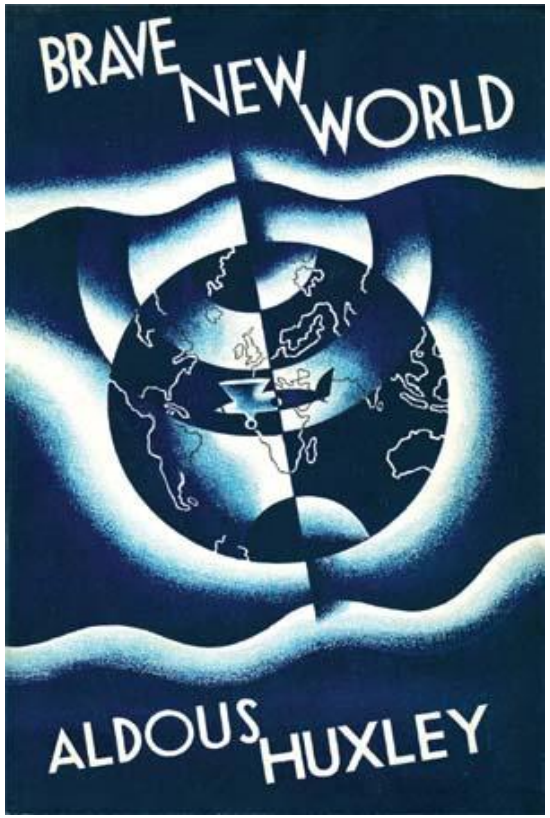


JCVI-syn3.0 (531 kbp, **473 genes).**

Four design-build-test cycles produced JCVI-syn3.0. (**A**) The cycle for genome design, building by means of synthesis and cloning in yeast, and testing for viability by means of genome transplantation. After each cycle, gene essentiality is reevaluated by global transposon mutagenesis. (**B**) Comparison of JCVI-syn1.0 (outer blue circle) with JCVI-syn3.0 (inner red circle), showing the division of each into eight segments. The red bars inside the outer circle indicate regions that are retained in JCVI-syn3.0. (**C**) A cluster of JCVI-syn3.0 cells, showing spherical structures of varying sizes (scale bar, 200 nm).

Raise ethical questions:

"How can we use this powerful tool in such a way as to ensure maximum benefit while minimizing risks?"



Eraclito



“nulla e’ permanente

tranne il cambiamento”

Nuova era:

quella della conoscenza e della competizione



Creatività
e Innovazione…

Si possono fare sempre più cose…che sono sempre migliorabili…
…..e non solo dal punto di vista tecnico……



‘’… E TU VUOI PRENDERE QUEST’UOMO, PARTI DEL QUALE SONO BREVETTATE DALL’IMPRESA GENESCOPE…”