

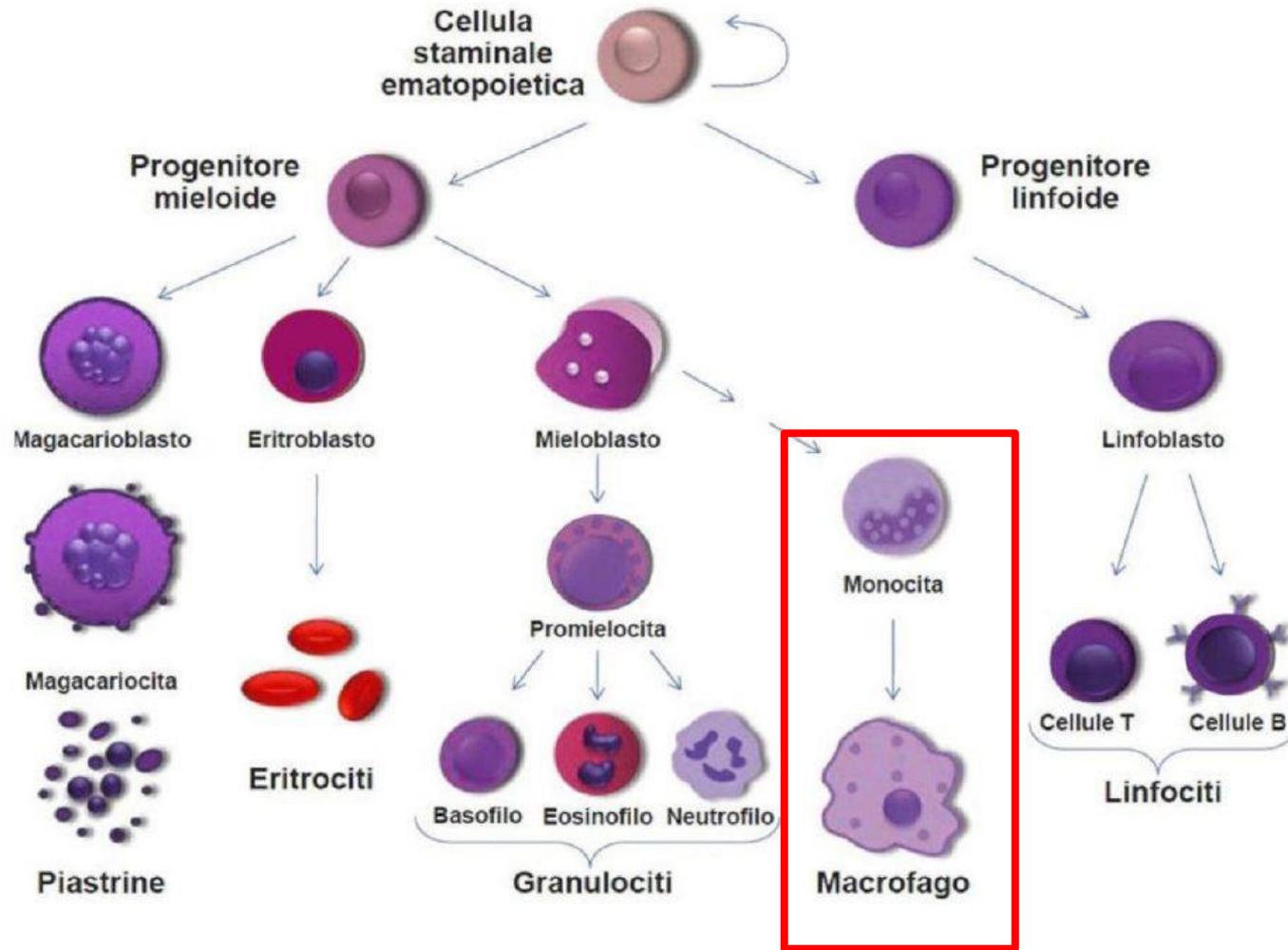


Istituto Zooprofilattico Sperimentale
del Lazio e della Toscana *M. Aleandri*

Studio dei marker di polarizzazione dei macrofagi

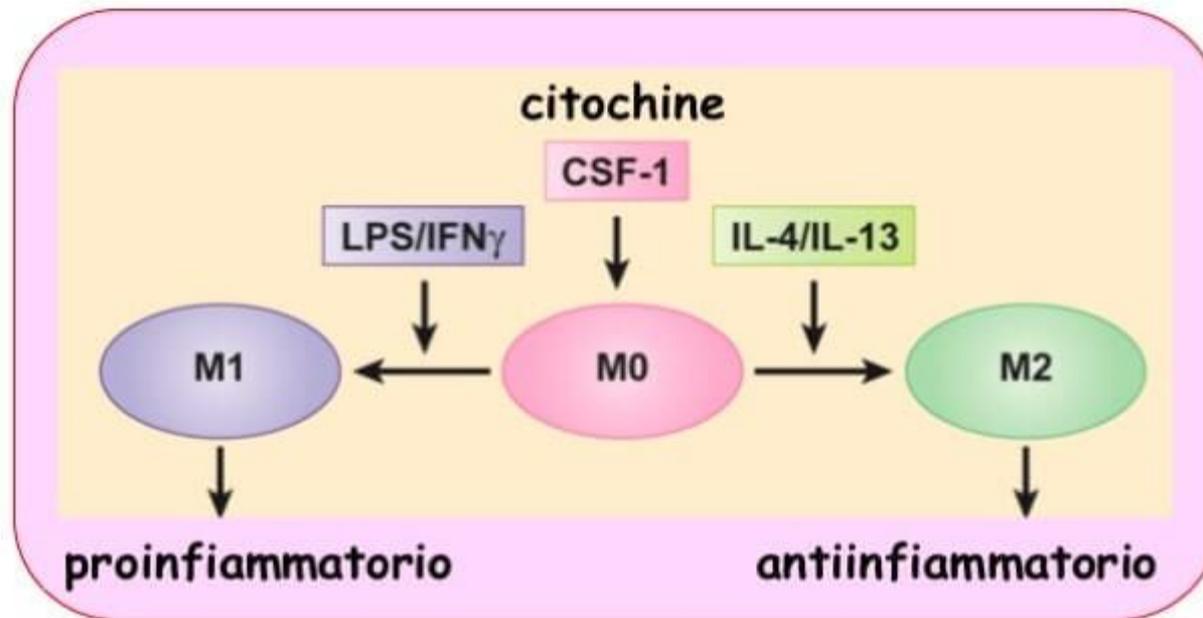
Giuseppe Manna



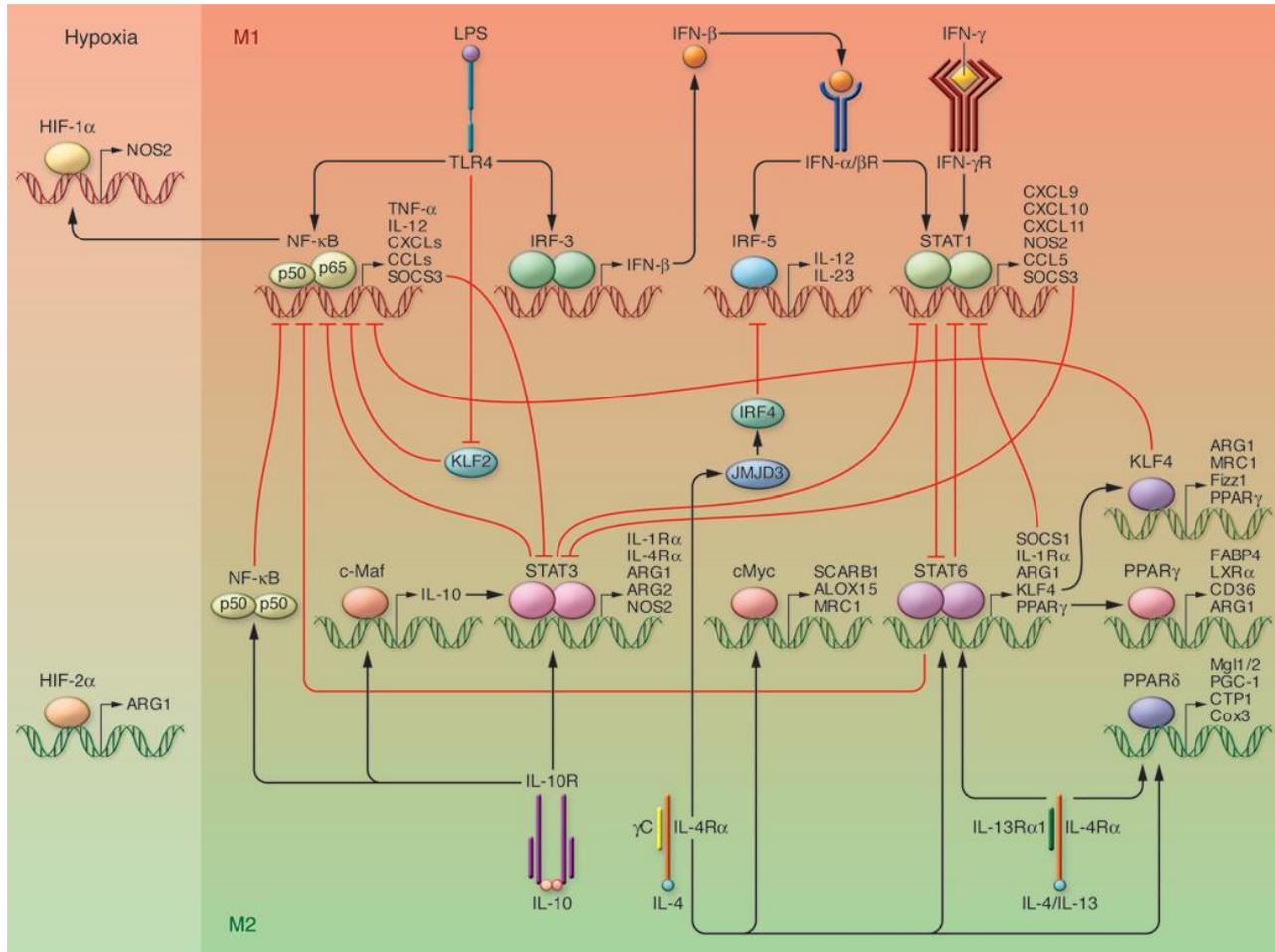


M1 stadio pro-infiammatorio

M2 stadio anti-infiammatorio



Profili di polarizzazione Macrofagi M1-M2





Il trascrittoma

Il trascrittoma è costituito dagli mRNA, che contengono in sé l'informazione genetica codificante le proteine, ma anche a tutte le altre forme di RNA intracellulari che svolgono, per lo più, funzioni regolatorie (dagli RNA di trasporto o tRNA, a quelli ribosomiali o rRNA, fino ai micro-RNA o ai long non-coding RNA).





Trascrittoma

L'espressione genica tessuto-specifica determina il fenotipo morfologico dei tipi cellulari e tissutali.

In ogni cellula differenziata e in ogni particolare momento dello sviluppo è attivo solo un sottoinsieme di geni. In tutti gli organismi viventi, infatti, le informazioni contenute nel genoma non si esprimono contemporaneamente, ma sono finemente regolate.





L'analisi dei trascrittomi può essere condotta attraverso l'utilizzo di differenti tecnologie:

Ibridazione su matrice solida (microarray) basati su spotting di molecole di cDNA.

Quantificazione mediante real time PCR dei trascritti di interesse

NGS (Next-Generation Sequencing) mediante la tecnica di **RNA-Seq** che permette il sequenziamento degli RNA messaggeri e degli altri RNA presenti.



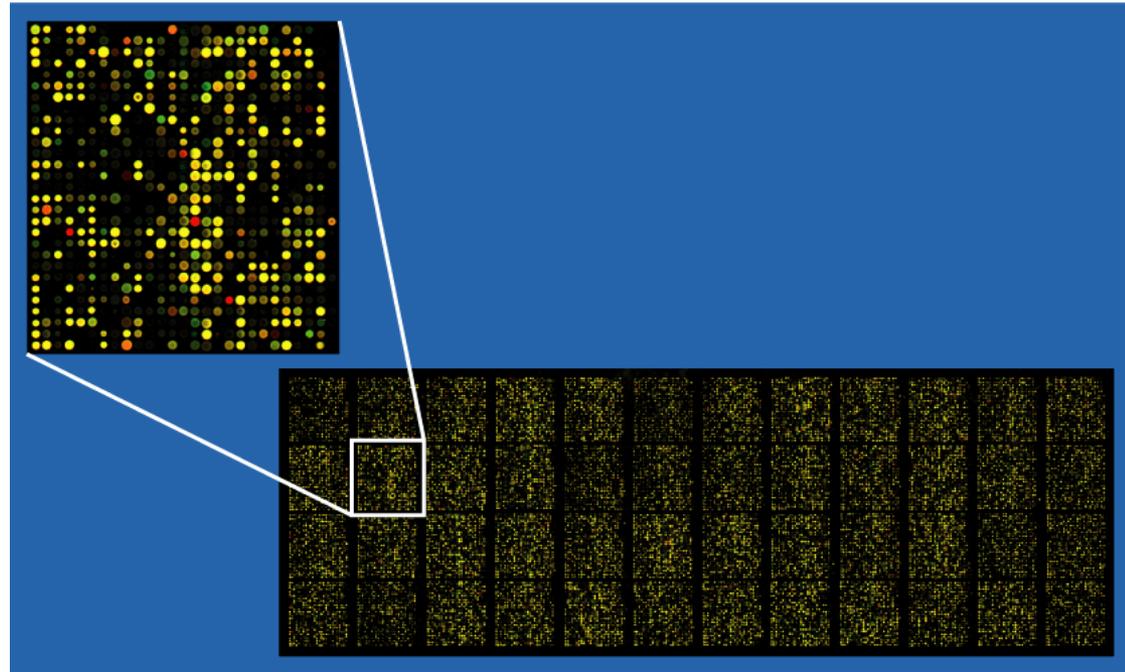


Istituto Zooprofilattico Sperimentale
del Lazio e della Toscana *M. Aleandri*

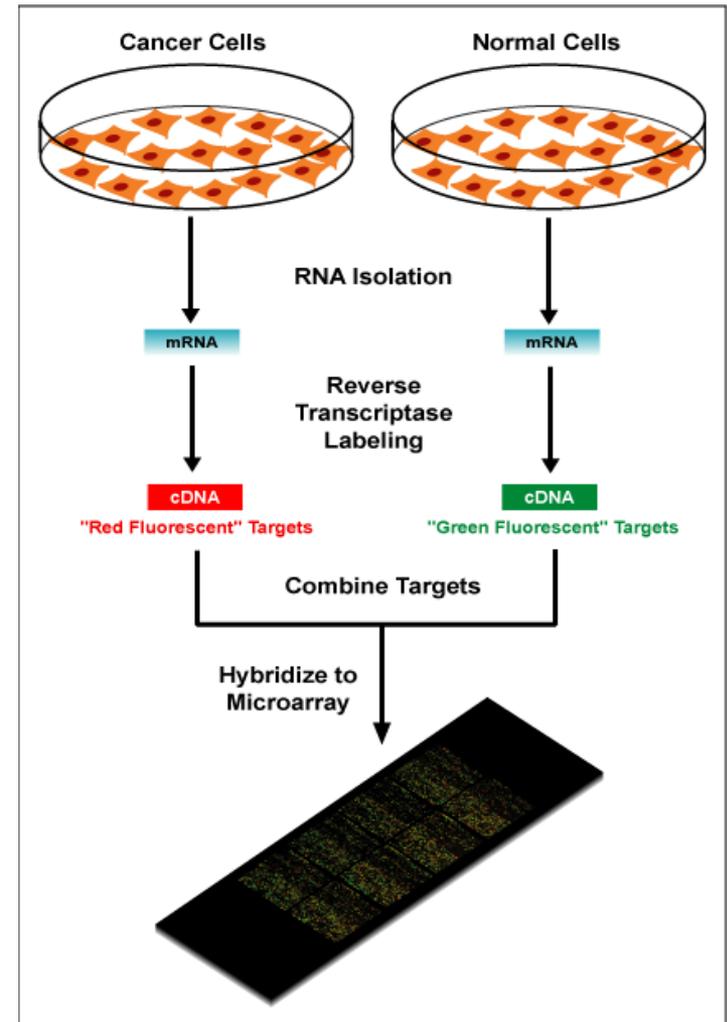
Un microarray di DNA è un insieme di microscopiche sonde di DNA attaccate ad una superficie solida come vetro, plastica, o chip di silicio formanti un array (matrice). Tali array permettono di esaminare simultaneamente la presenza di moltissimi geni all'interno di un campione di DNA

(che spesso può rappresentare anche tutto il genoma o il trascrittoma di un organismo).

Un utilizzo tipico è quello di confrontare il profilo di espressione genica di un individuo malato con quello di uno sano per individuare quali geni sono coinvolti nella malattia.



Le probe sono oligonucleotidi, cDNA o piccoli frammenti prodotti con la tecnologia PCR corrispondenti a mRNA. Questo tipo di microarray sfrutta l'ibridazione di DNA con cDNA da due campioni comparati, che sono marcati con due differenti fluorofori. I campioni possono essere miscelati e ibridizzati in un singolo microarray e quindi analizzati, permettendo la visualizzazione dei geni up-regolati e down-regolati contemporaneamente. In questo modo non può essere misurato il livello assoluto di espressione genica.





Microarray

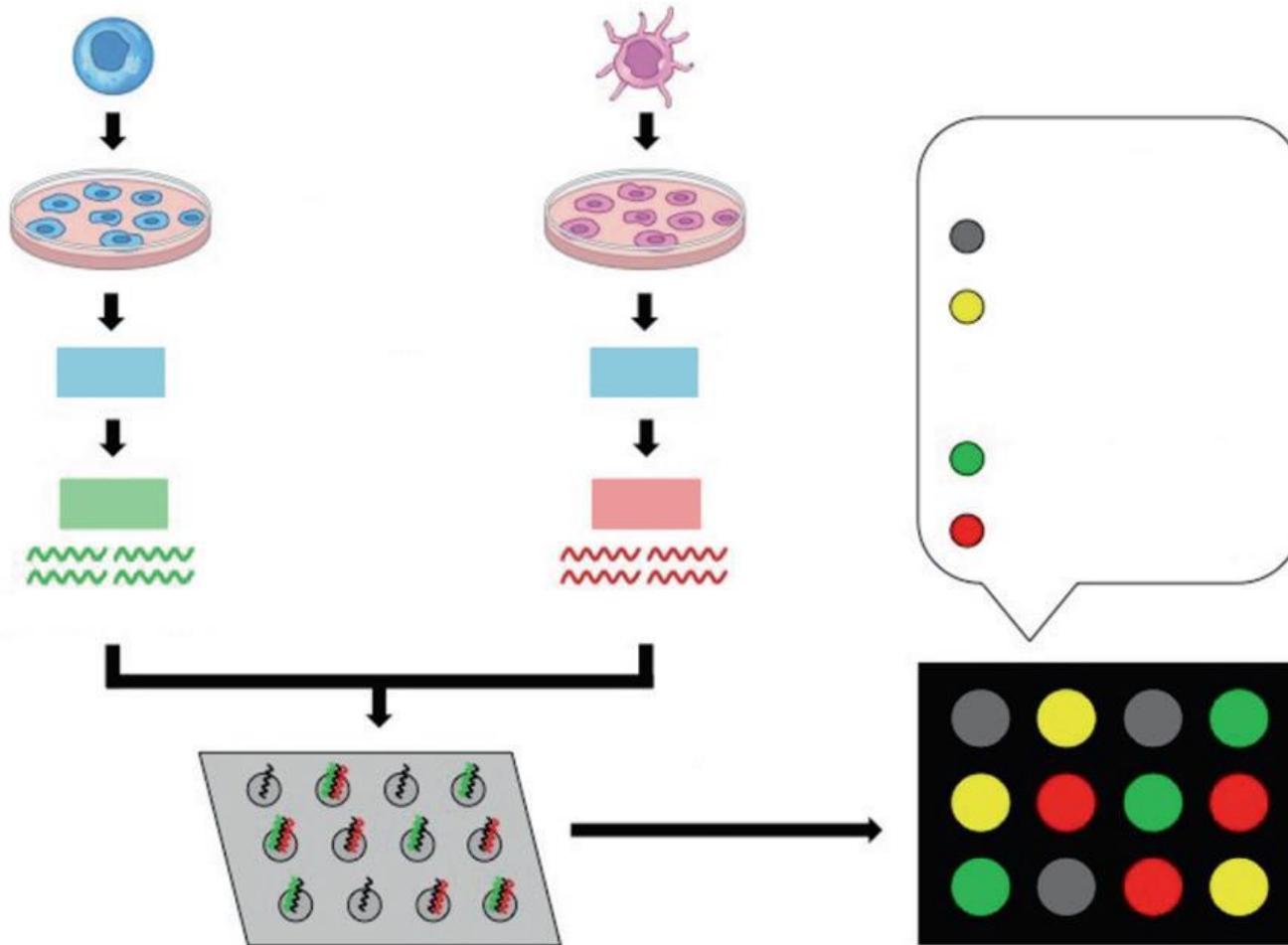
I **microarray** sfruttano la tecnica di ibridazione inversa, che consiste nel fissare tutti i segmenti di DNA (probe) su un supporto e nel marcare invece l'acido nucleico che vogliamo identificare (detto *target*). È una tecnica che è stata sviluppata negli anni '90 e oggi permette l'analisi dell'espressione genica monitorando in una sola volta gli RNA prodotti da migliaia di geni.

Per studiare gli mRNA, essi vengono prima estratti dalle cellule, convertiti in cDNA, e marcati con una sonda fluorescente. Quando si fa avvenire l'ibridazione fra la sonda presente sulla matrice e il cDNA target, quest'ultimo rimarrà legato alla sonda e può essere identificato semplicemente rilevando la posizione dove è rimasto legato.

La misura dell'espressione genica mediante microarray ha avuto un notevole sviluppo sia nel campo della ricerca di base che nella diagnostica medica, in particolare di malattie a base genetica, dove l'espressione genetica di cellule sane viene comparata con quella di cellule affette dalla malattia in esame.



Ibridazione su matrice solida (Microarray)



Microarray - svantaggi

Gli svantaggi nell'utilizzo dei microarray consistono nella presenza di artefatti di ibridazione incrociata, la quantificazione inesatta di geni poco o altamente espressi e la necessità di conoscere la sequenza da studiare a priori.

A causa di questi problemi tecnici, la trascrittomica è passata a metodi basati sul sequenziamento. Questi sono passati dal sequenziamento di Sanger delle librerie di tag di sequenze espresse, a metodi basati su tag chimici (ad es. Analisi seriale dell'espressione genica) e infine alla tecnologia attuale, il sequenziamento di nuova generazione del cDNA, in particolare RNA-Seq.





RNA-Seq

Con le tecniche NGS applicate all'RNA (RNA-Seq), è possibile analizzare il trascrittoma cellulare, utilizzando come input per la strumentazione, l'acido nucleico da sequenziare e quantificare.

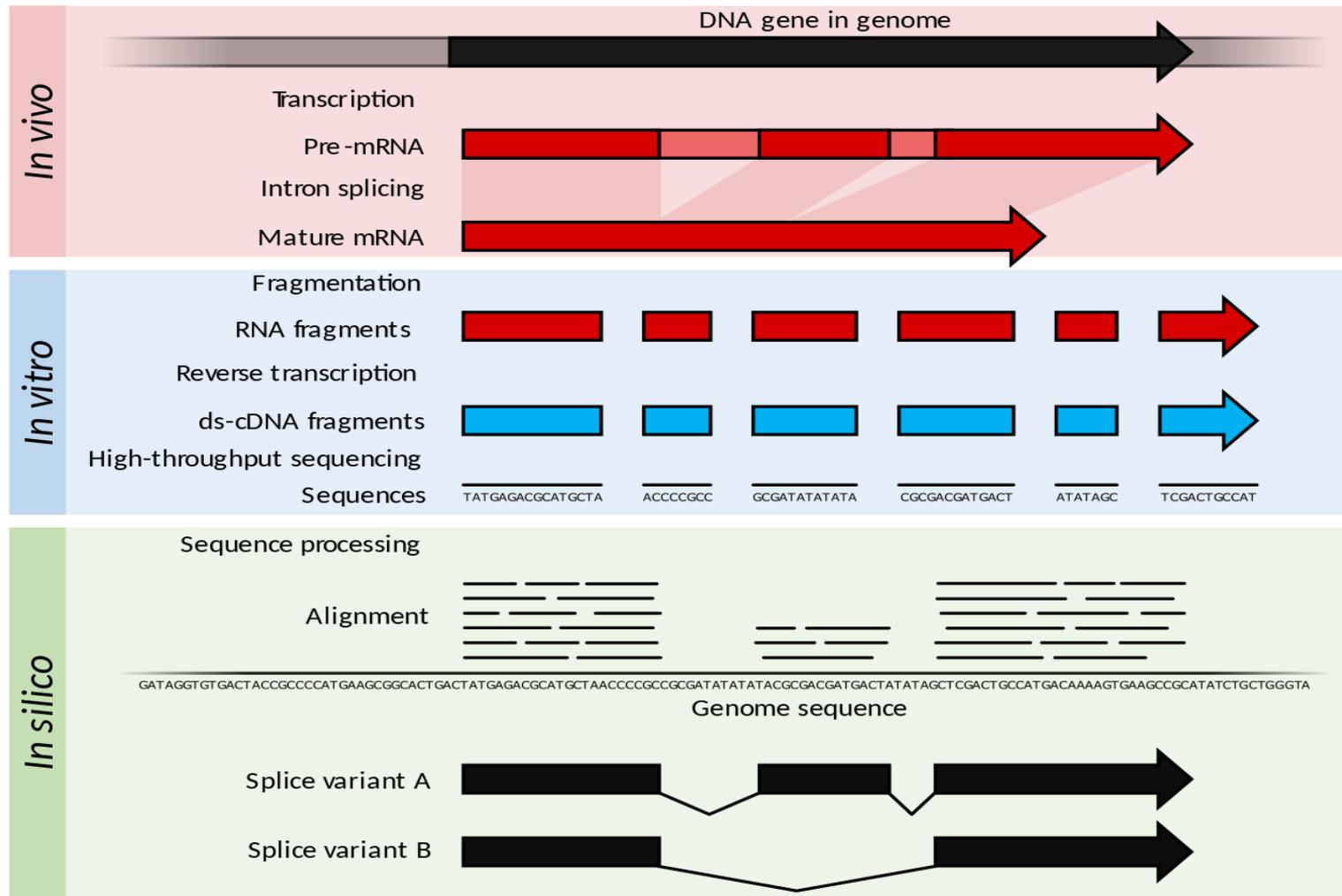
Gli RNA cellulari, dopo essere stati estratti dalla cellula, vengono retrotrascritti in DNA, sequenziato attraverso tecniche NGS.

Vengono così ottenute delle sequenze di DNA di lunghezza variabile in base alla tecnologia di sequenziamento utilizzata, definite reads. Tali sequenze (reads) vengono poi mappate su un genoma o un trascrittoma di riferimento per identificare i geni espressi nel campione in esame.

La tecnica è quantitativa: tanto maggiore è il numero di molecole di trascritto (RNA), tanto maggiore sarà il numero di read prodotte attraverso il sequenziamento.

Il totale delle read allineate su un gene (o un trascritto, o un esone), detto count, è una unità di misura dell'espressione del gene stesso, interpretabile attraverso modelli statistici







RNA-Seq

I vantaggi della tecnologia RNA-Seq rispetto a quella dei microarray, sempre meno utilizzati, sta essenzialmente nella possibilità di essere utilizzata anche quando non è nota la sequenza del gene in esame (quella che sarebbe necessaria per disegnare le probe) e nella maggiore precisione e più alta risoluzione della quantificazione delle misure.

Infatti, i dati di RNA-Seq, i count, sono una misura discreta, mentre nel caso dei microarray si tratta di dati analogici, ovvero del valore continuo dell'intensità luminosa.

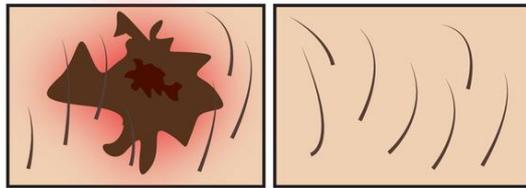
Si ottiene l'analisi dell'espressione differenziale, cioè l'identificazione di geni che presentano significative differenze del loro livello di espressione tra due o più condizioni sperimentali (interne o esterne alla cellula).

Si valuta cioè, se le differenze osservate tra i count delle diverse condizioni sperimentali siano o meno statisticamente significative.





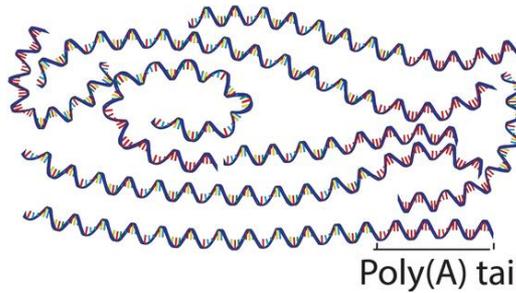
Samples of interest



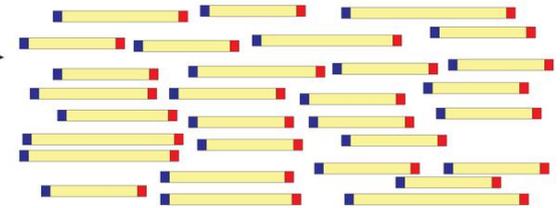
Condition 1
(e.g. tumor)

Condition 2
(e.g. normal)

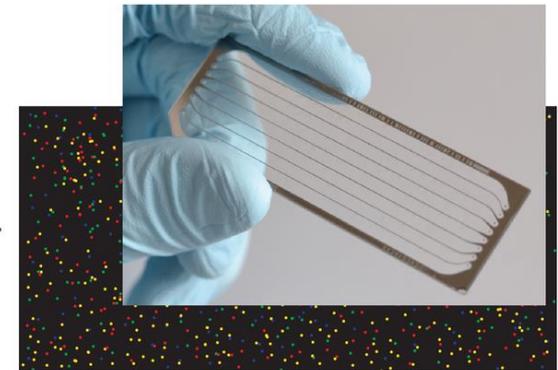
Isolate RNAs



Generate cDNA, fragment, size select, add linkers

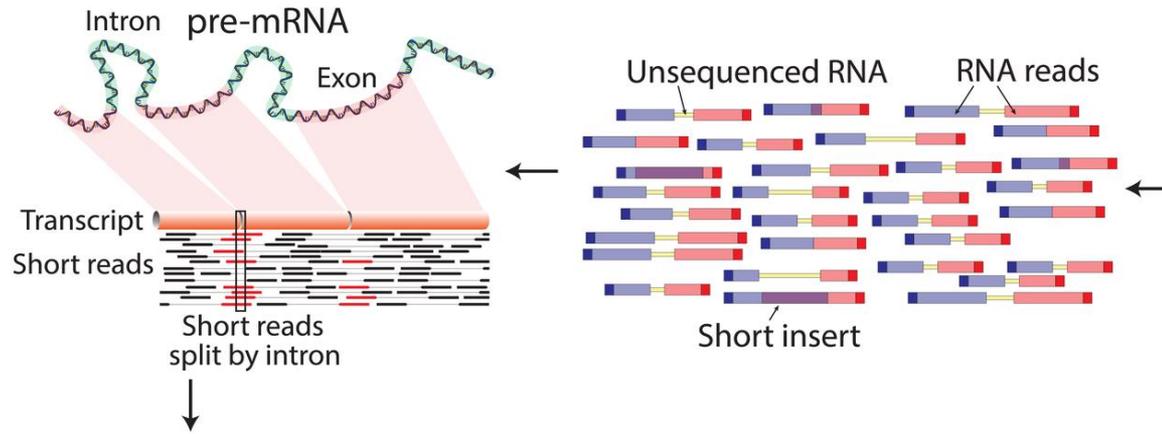


Sequence ends



100s of millions of paired reads
10s of billions bases of sequence

Map to genome, transcriptome, and predicted exon junctions



Downstream analysis



La quantità di dati generata è enorme, dunque richiede l'utilizzo di software di bioinformatica e statistica per elaborare le informazioni e creare modelli multidimensionali in grado di spiegare le interazioni all'interno del sistema o tra vari sistemi.

L'analisi dei dati di trascrittomici, avviene in generale attraverso l'utilizzo di metodi di clustering, metodi di statistica multivariata che raggruppano unità statistiche sulla base di misure di similarità o dissimilarità.

Simili rispetto a cosa? Si considerano i geni come punti nello spazio tra cui si misura la distanza: punti vicini sono raggruppati (clusterizzati) insieme.





Istituto Zooprofilattico Sperimentale
del Lazio e della Toscana *M. Aleandri*

Grazie

