

Rapallo, 27 maggio 2005

Giochi tra geni.

Stefano Moretti

DIMA: Dipartimento di Matematica dell'Università di Genova

IST: Istituto per la Ricerca sul Cancro di Genova

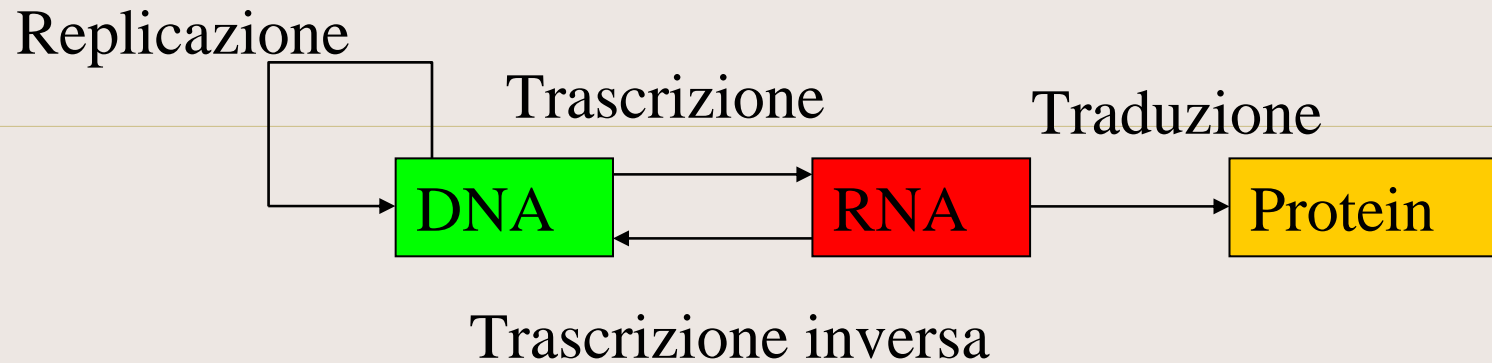
Fioravante Patrone

DIMA: Dipartimento di Matematica dell'Università di Genova

Punto della situazione:

- La teoria dei giochi e gli indici di potere:
 - *i geni al lavoro* (J.von Neumann (1944); J.Nash (1950); Shapley (1954))
- A cosa servono i 'microarray':
 - *il lavoro dei geni*: l'espressione genica;
- Un applicazione della TdG all'analisi di microarray:
 - giocare con i geni.

Principio fondamentale



DNA

ATATCGGCATCAGTCGATCGATCATCGATCGAT

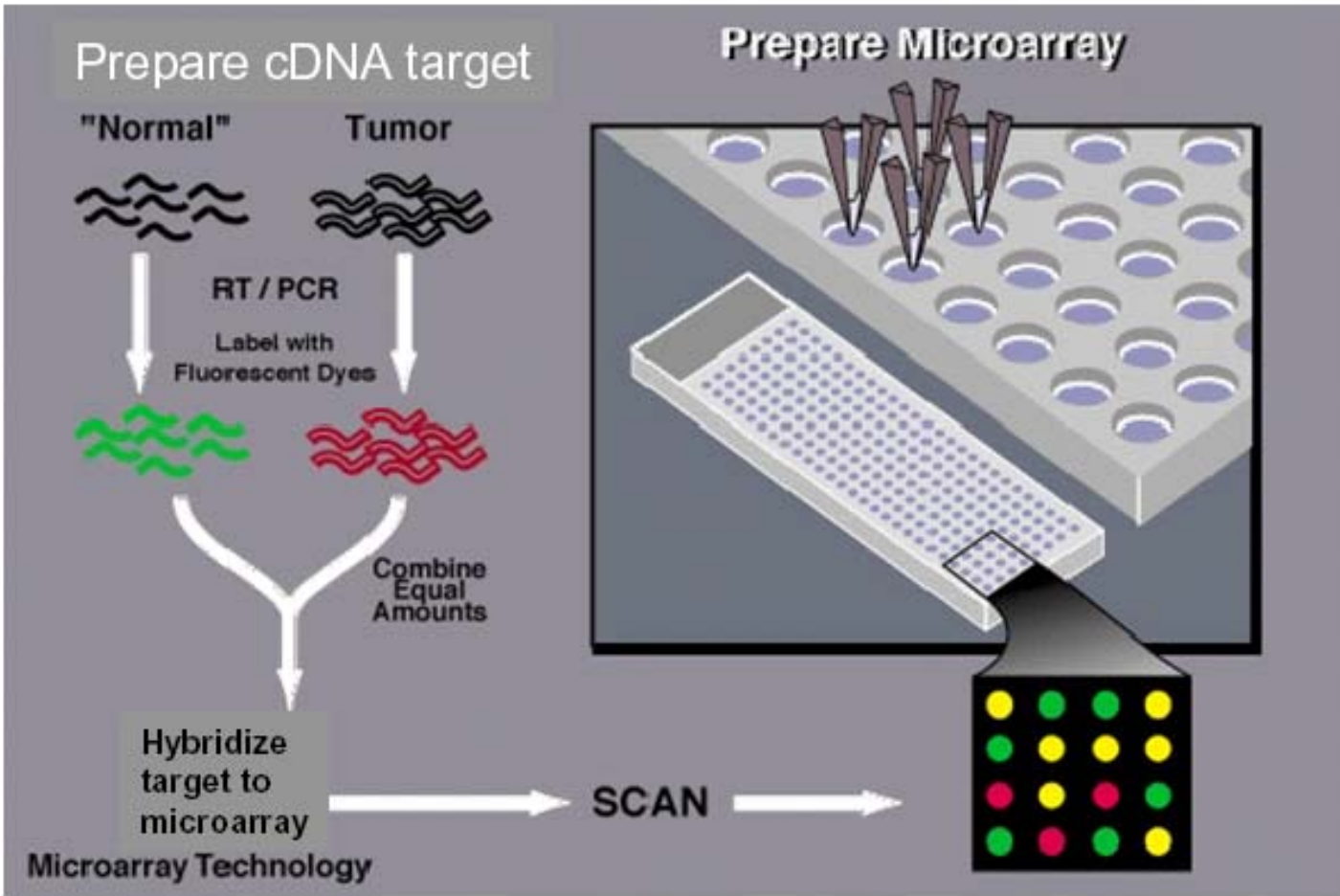
↓
mRNA

UAUAGCCGUAGUCAGCUAGCUAGUAGCUAGCUA

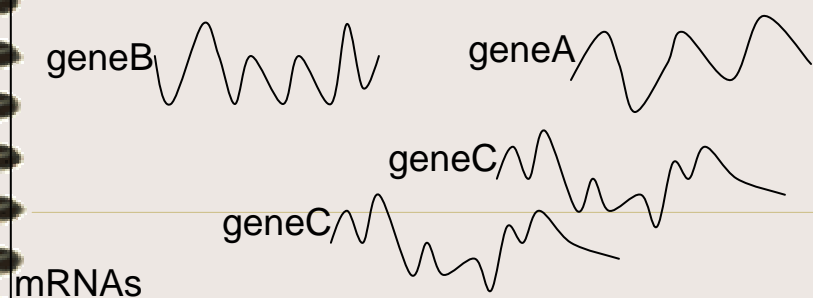
↓
cDNA

ATATCGGCATCAGTCGATCGATCATCGATCGAT

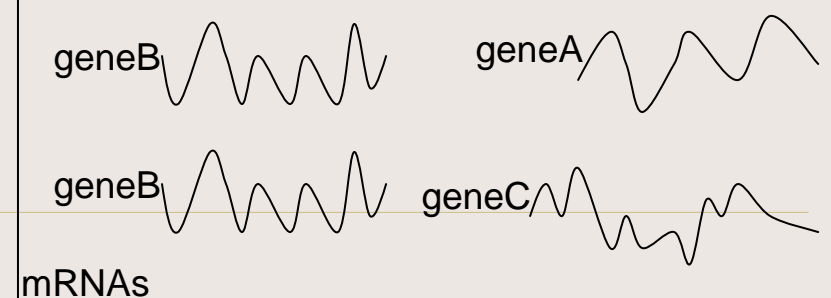
Espressione genica: quantità di proteine prodotte da un singolo gene in una cellula



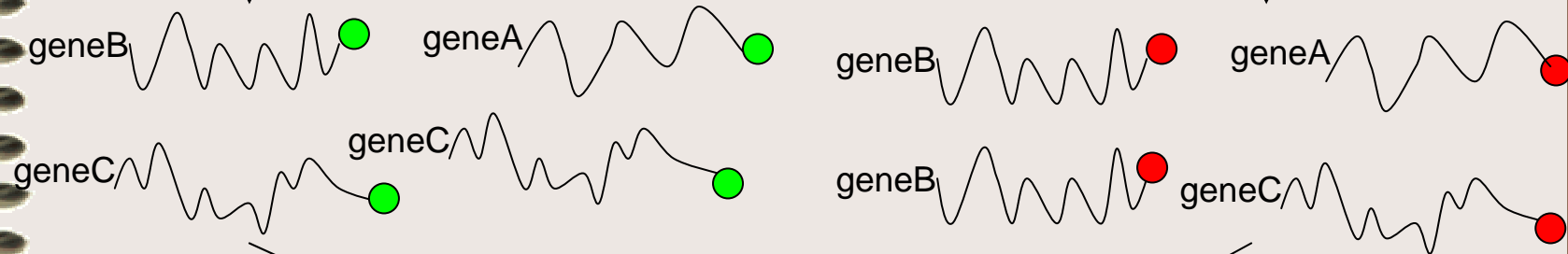
Cellula sana



Cellula tumorale



Reazione di marcatura fluorescente
con trascrizione inversa



Ibridizzazione

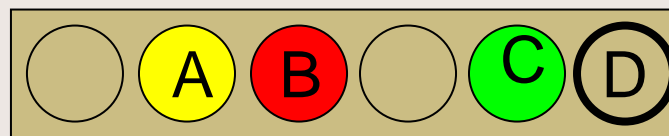
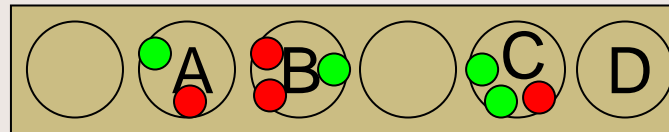


Immagine scannerizzata

	Tumore	Sana
Gene A	1	1
Gene B	2	1
Gene C	1	2
Gene D	0	0

Matrice di espressione

		Arrays					...
		Array1	Array2	Array3	Array4	Array5	
Genes	Gene1	0.46	0.30	0.80	1.51	0.90	...
	Gene2	-0.10	0.49	0.24	0.06	0.46	...
	Gene3	0.15	0.74	0.04	0.10	0.20	...
	Gene4	-0.45	-1.03	-0.79	-0.56	-0.32	...
	Gene5	-0.06	1.06	1.35	1.09	-1.09	...

$$M = \log_2(\text{Intensità del Rosso} / \text{Intensità del verde})$$

An image of a microarray

(from Van't Wout et al., 2003)



**Tabella dei dati di espressione
provenienti da un esperimento su
tre cellule malate di tumore**

	t1	t2	t3
g1	0.2	20	12
g2	11	9.8	8.6
g3	-7	2.4	-6.1

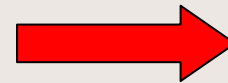


Tabella discretizzata

	t1	t2	t3
g1	0	1	1
g2	1	1	1
g3	1	0	1

- $M \approx 0$, il gene è espresso in maniera normale nella cellula malata \rightarrow assegno 0;
- $M \gg 0$ o $M \ll 0$, il gene è espresso in maniera anormale nella cellula malata \rightarrow assegno 1.

Dati di espressione di cellule tumorali

	t1	t2	t3
g1	0.2	20	12
g2	11	9.8	8.6
g3	7	2.4	6.1



Matrice discretizzata

	s1	s2	s3
g1	0	1	1
g2	1	1	1
g3	1	0	1

Dati di espressione di cellule normali

	s1	s2	s3
g1	4.1	6.3	2.7
g2	4.2	7.8	2.1
g3	5	3.5	0.5



cutoffs

<	>
2.7	6.3
2.1	7.8
0.5	5



Obiettivo

Matrice discretizzata

	s1	s2	s3
g1	0	1	1
g2	1	1	1
g3	1	0	1

Quantificare il potere dei geni nel determinare l'insorgenza della malattia, ovvero determinare in maniera oggettiva quanto ciascun gene sia rilevante per i meccanismi di insorgenza del tumore.

Costruiamo un gioco:

- I giocatori sono i geni;
- su ogni campione:
 - Se una coalizione assume valore **1** allora la coalizione fa insorgere la malattia
 - Se una coalizione assume valore **0** allora la coalizione non fa insorgere la malattia

	t1
g1	0
g2	1
g3	1

$$u_T(\{g_2, g_3\}) = u_T(\{g_2, g_3, g_1\}) = 1;$$

$$u_T(\{g_1\}) = u_T(\{g_2\}) = u_T(\{g_3\}) = \\ u_T(\{g_1, g_2\}) = u_T(\{g_1, g_3\}) = 0.$$

Gioco di unanimità:

$$N = \{g_1, g_2, g_3\}, T = \{g_2, g_3\}, S \subseteq N$$

$$u_T(S) = \begin{cases} 1 & \text{se } T \subseteq S; \\ 0 & \text{altrimenti.} \end{cases}$$

	t1
g1	0
g2	1
g3	1

Valore Shapley Sh per ogni $i \in N$:

$$Sh_i(u_T) = \begin{cases} 1/|T| & \text{se } i \in T; \\ 0 & \text{altrimenti.} \end{cases} \quad (\text{Eff, NP, Anon})$$

Gioco con più campioni

Matrice discretizzata

	s1	s2	s3
g1	0	1	1
g2	1	1	1
g3	1	0	1



Ad ogni coalizione S associo il numero medio di campioni in cui tale coalizione fa insorgere la malattia

Il microarray game corrispondente è $(\{g_1, g_2, g_3\}, v)$:

$$v(\{g_1, g_2\}) = v(\{g_3, g_2\}) = 1/3$$

$$v(\{g_1, g_2, g_3\}) = 1 \text{ e}$$

$v(S) = 0$ per ogni altra coalizione S .

Il valore Shapley è: $(5/18, 8/18, 5/18)$.

Come calcolare il valore Shapley di un microarray game?

Uso Add su giochi di unanimità!

Matrice discretizzata

	s1	s2	s3
g1	0	1	1
g2	1	1	1
g3	1	0	1

$$\begin{aligned} & (\text{Sh}(u_{\{g2,g3\}}) + \text{Sh}(u_{\{g1,g2\}}) + \text{Sh}(u_{\{g1,g2,g3\}})) / 3 = \\ & ((0, 1/2, 1/2) + (1/2, 1/2, 0) + (1/3, 1/3, 1/3)) / 3 = \\ & = (5/18, 8/18, 5/18) \end{aligned}$$

Dati sperimentali

- Selezione di dati ricavati mediante utilizzo di microarray da Alon et al. (1999) su cellule di tessuti tumorali e sane del colon.
- I microarray considerati sono **2000 geni** umani e 62 campioni cellulari in tutto (40 tumorali e 22 normali).

Expression values of one gene in 22 normal samples

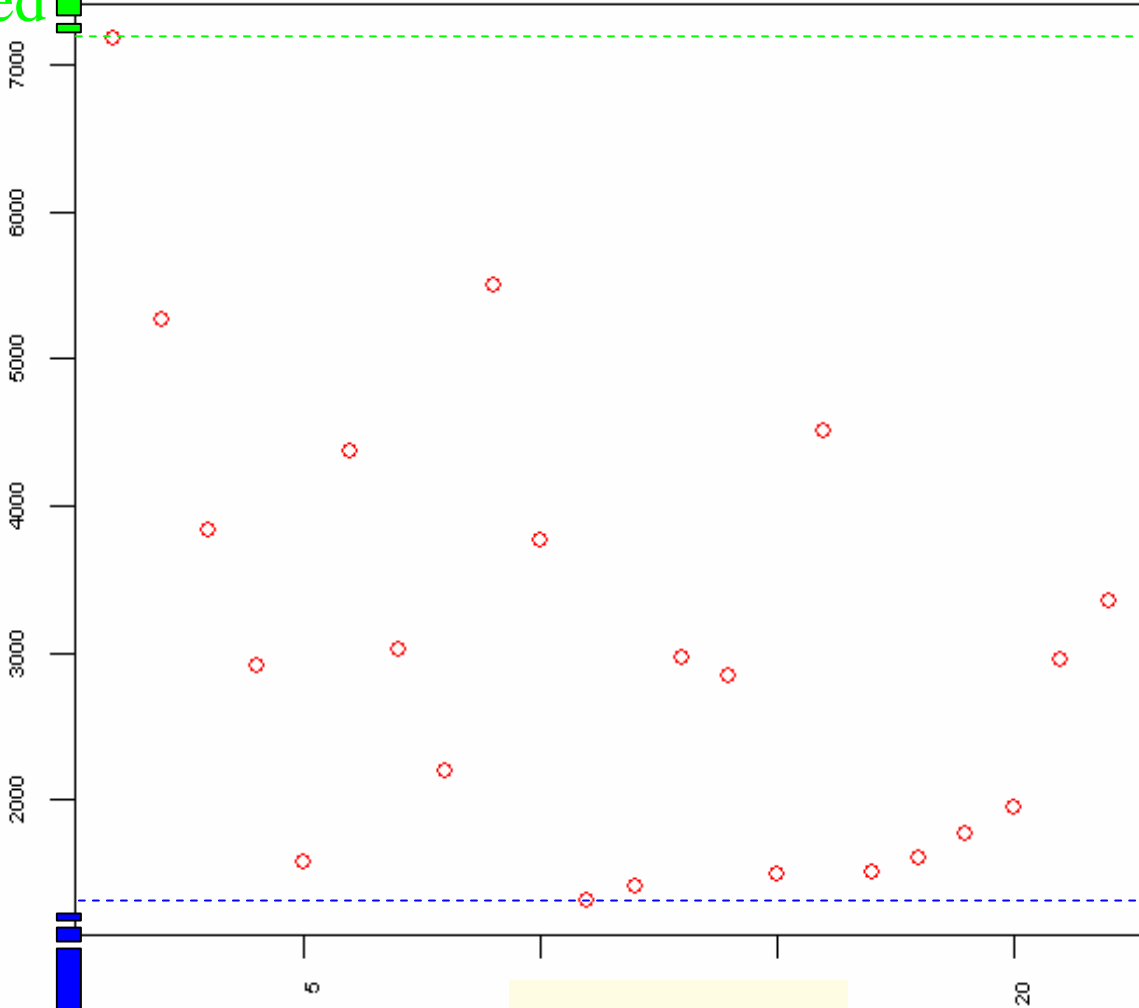
Expression values

Gene labels

gene labels

Values considered overexpressed

Values considered underexpressed



Alcuni risultati su dati sperimentali

- Abbiamo usato i dati ricavati mediante utilizzo di microarray da Alon et al. (1999) su cellule di tessuti tumorali e sane del colon.
- I microarray considerati hanno 6500 geni umani e 62 campioni cellulari in tutto (40 tumorali).

Nome gene	Indice di Shapley ($\times 10^{-3}$)
H.sapiens mRNA for GCAP-II/uroguanylin precursor	3.83
Nucleolin	3.56
Gelsolin precursor, Plasma	3.34
DNA-(Apurinic or apyrimidinic site) Lyase	3.23
Human vasoactive intestinal peptide (VIP)	3.21

!?

E' noto il coinvolgimento di questi geni nei processi tumorali

A spiral-bound notebook with a brown cover and a light beige page. The page is mostly blank, with a horizontal line near the top. The words "Fine ?" are written in the center of the page in a black, serif font. The spiral binding is on the left side.

Fine ?