



Università degli Studi di Modena e Reggio Emilia
Dipartimento di Scienze Biomediche
Sezione di Chimica Biologica
Via Campi, 287 41100, Modena, Italy ¹

CULTURA E VITA – UNIMORE – A.A. 2010-2011

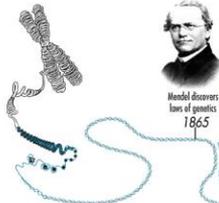
“Storia della Medicina e Antropologia Medica” VII
Direttore: Prof. Ugo Fabio

“BIOTECNOLOGIE MEDICHE”

Prof. Sergio Ferrari

Centro didattico di Ateneo – Facoltà di Medicina e Chirurgia

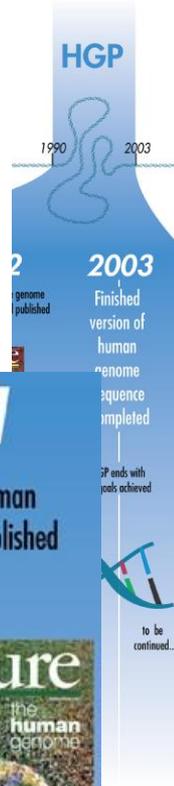
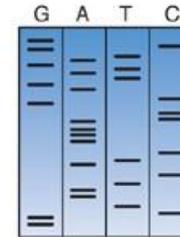
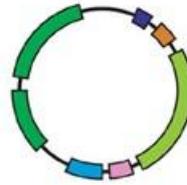
Aula T01, mercoledì 16 marzo 2011, ore 16.



Barrington Brown, Courtesy of Photo Researchers, Inc.



	U	C	A	G	
U	Pho Leu	Ser	Tyr stop stop	Cys stop Trp	U C A G
C	Leu	Pro	His Gln	Arg	U C A G
A	Ile Met	Thr	Asn Lys	Ser Arg	U C A G



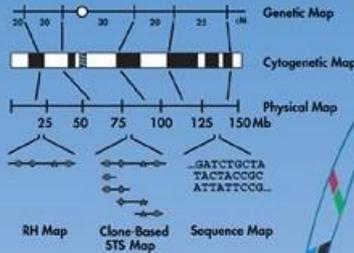
1990

Human Genome Project (HGP) launched in the U.S.



1991

First U.S. Genome Centers established



Cohen and Boyer develop recombinant DNA technology 1972

Issuing of Belmont Report on the use of human in re 1979

Sanger and Maxam & Gilbert

GenBank

2000

Draft version of human genome sequence completed

President Clinton and Prime Minister Blair support free access to genome information

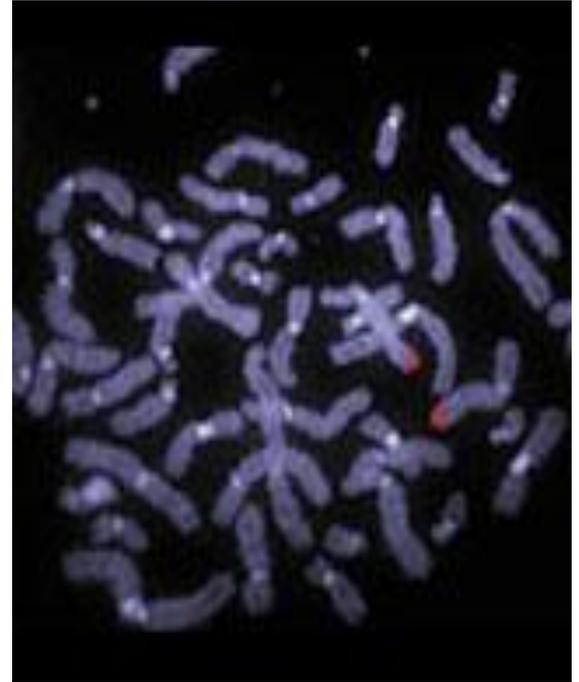
2001

Draft version of human genome sequence published

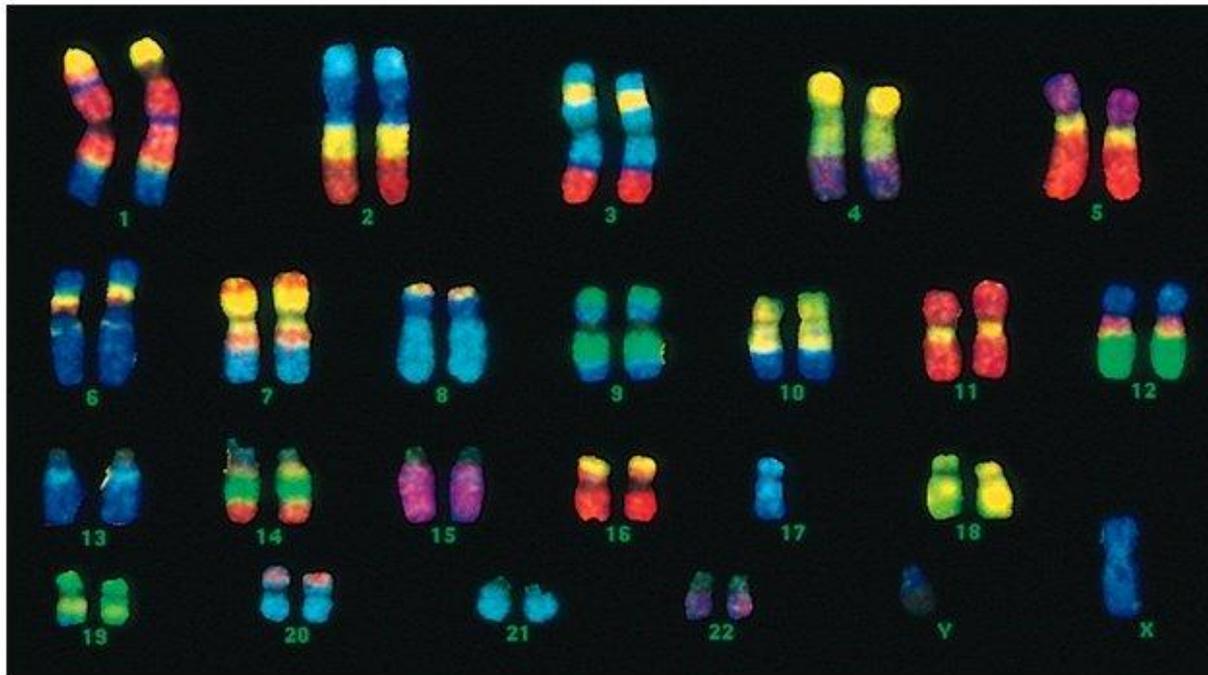
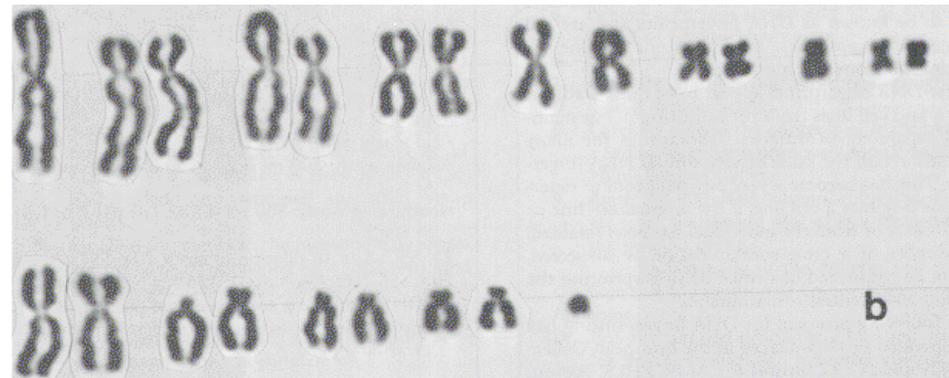
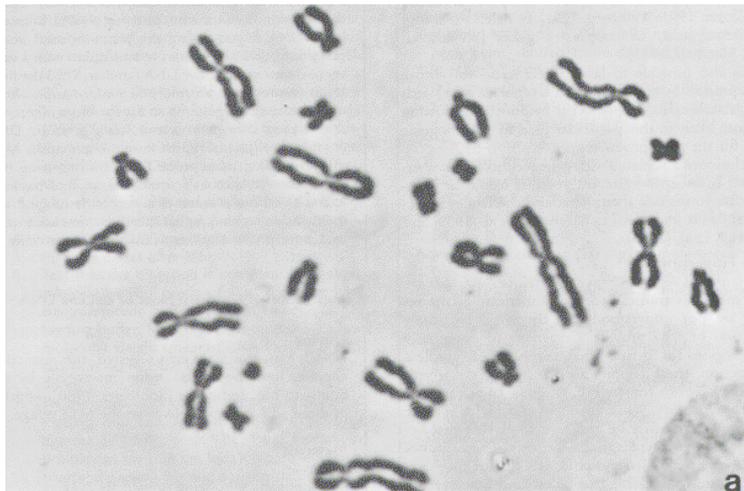




Il nucleo di ogni cellula contiene tutta l'informazione genetica necessaria per la costruzione dell'intero organismo. Questa informazione è contenuta nel DNA, ripartito nei cromosomi.



Tecniche di identificazione-Analisi del cariotipo

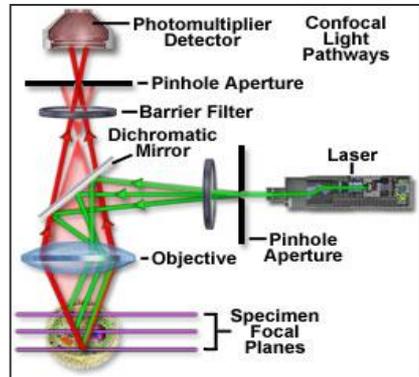
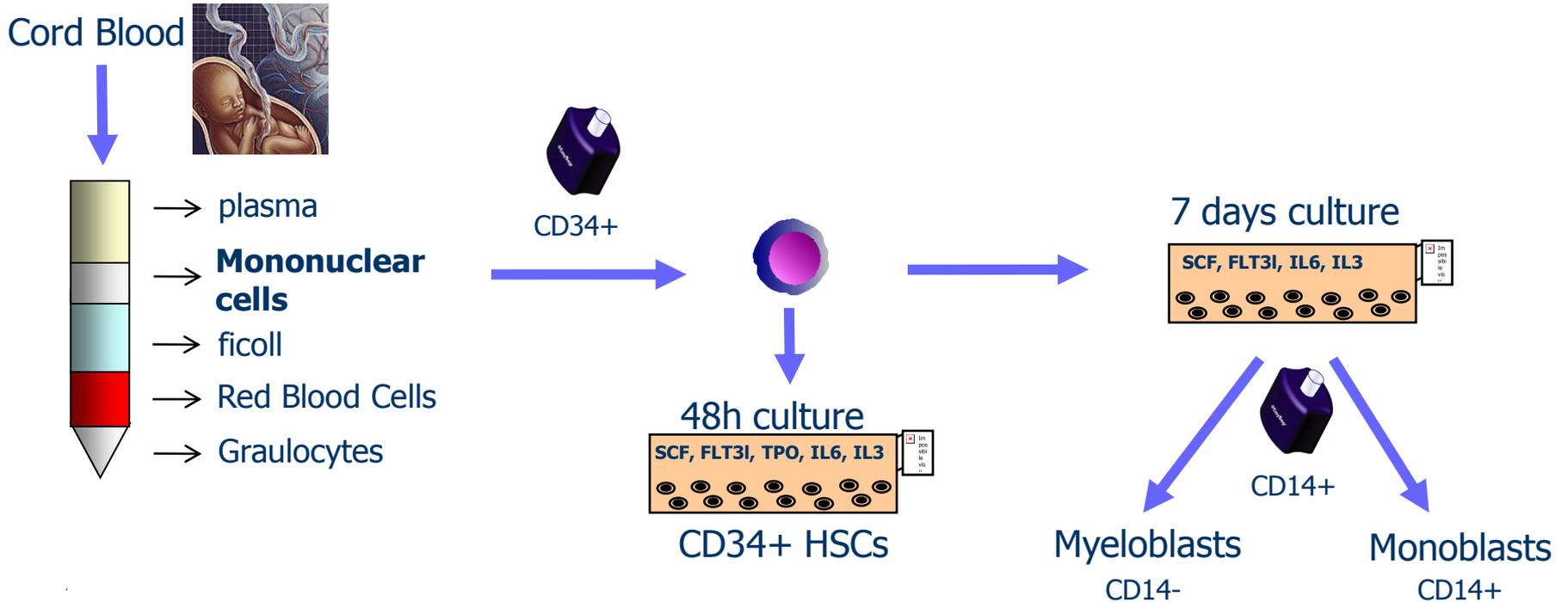


a, b:bandeggio G

c: chromosome painting

c

Experimental Design



3D FISH Analysis

CELL BIOLOGY

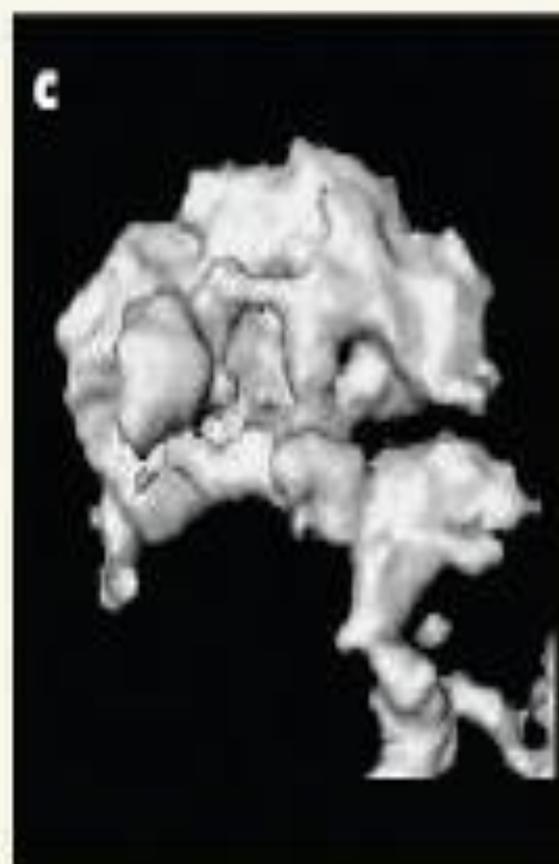
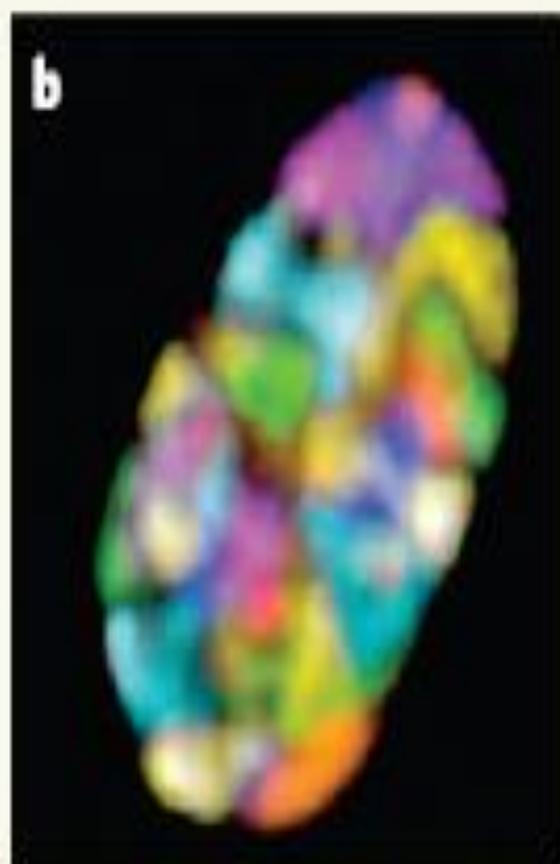
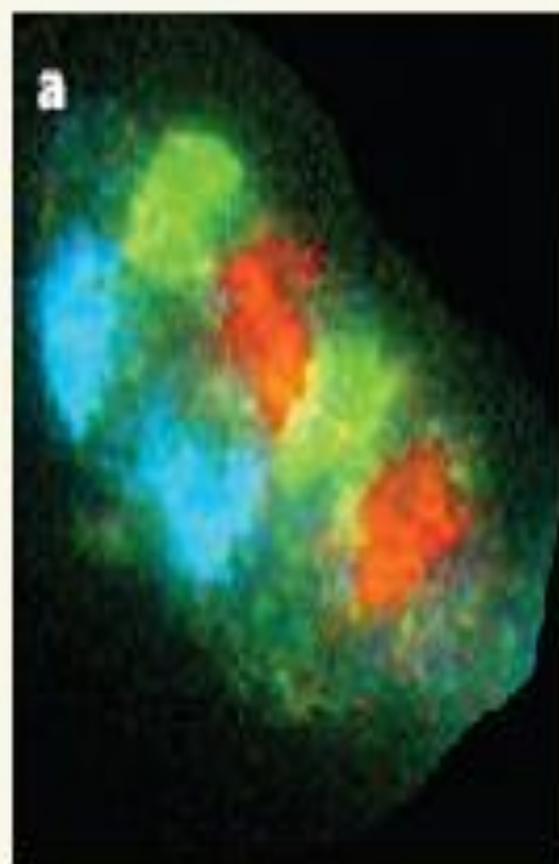
Chromosome territories

Karen J. Meaburn and Tom Misteli

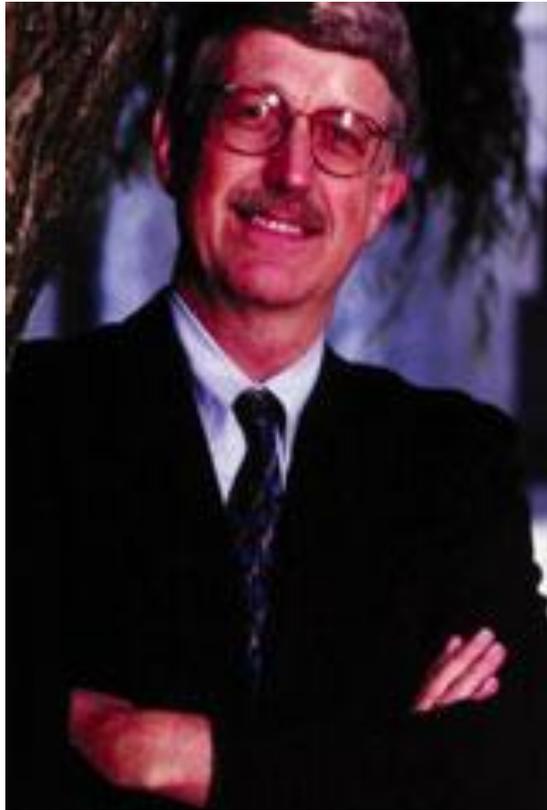
The natural habitat of eukaryotic genomes is the cell nucleus, where each chromosome is confined to a discrete region, referred to as a chromosome territory. This spatial organization is emerging as a crucial aspect of gene regulation and genome stability in health and disease.

Vol 445|25 January 2007

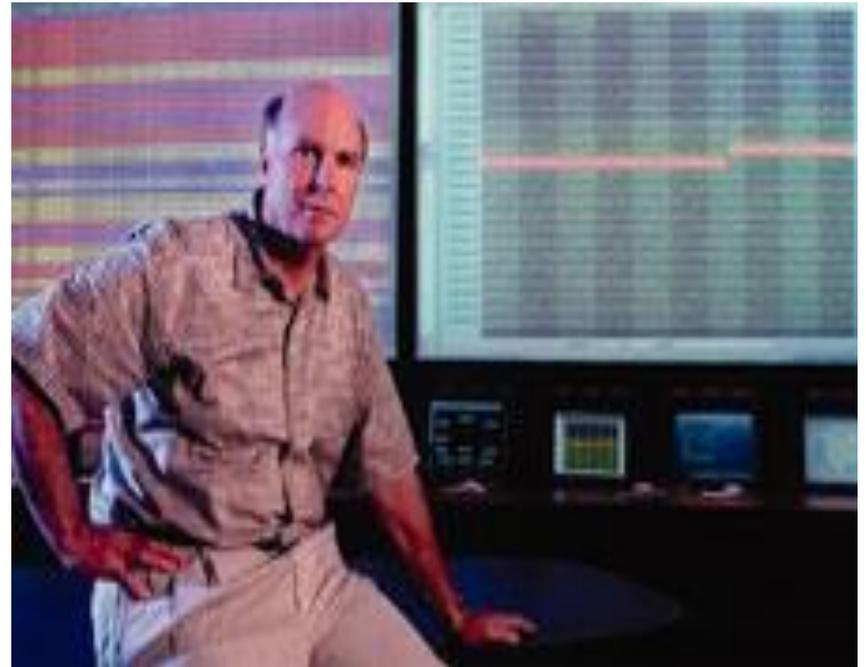
nature



The Human Genome Project



Francis Collins
National Institutes of Health (NIH)



Craig Venter
Celera Genomics

26 giugno 2000...

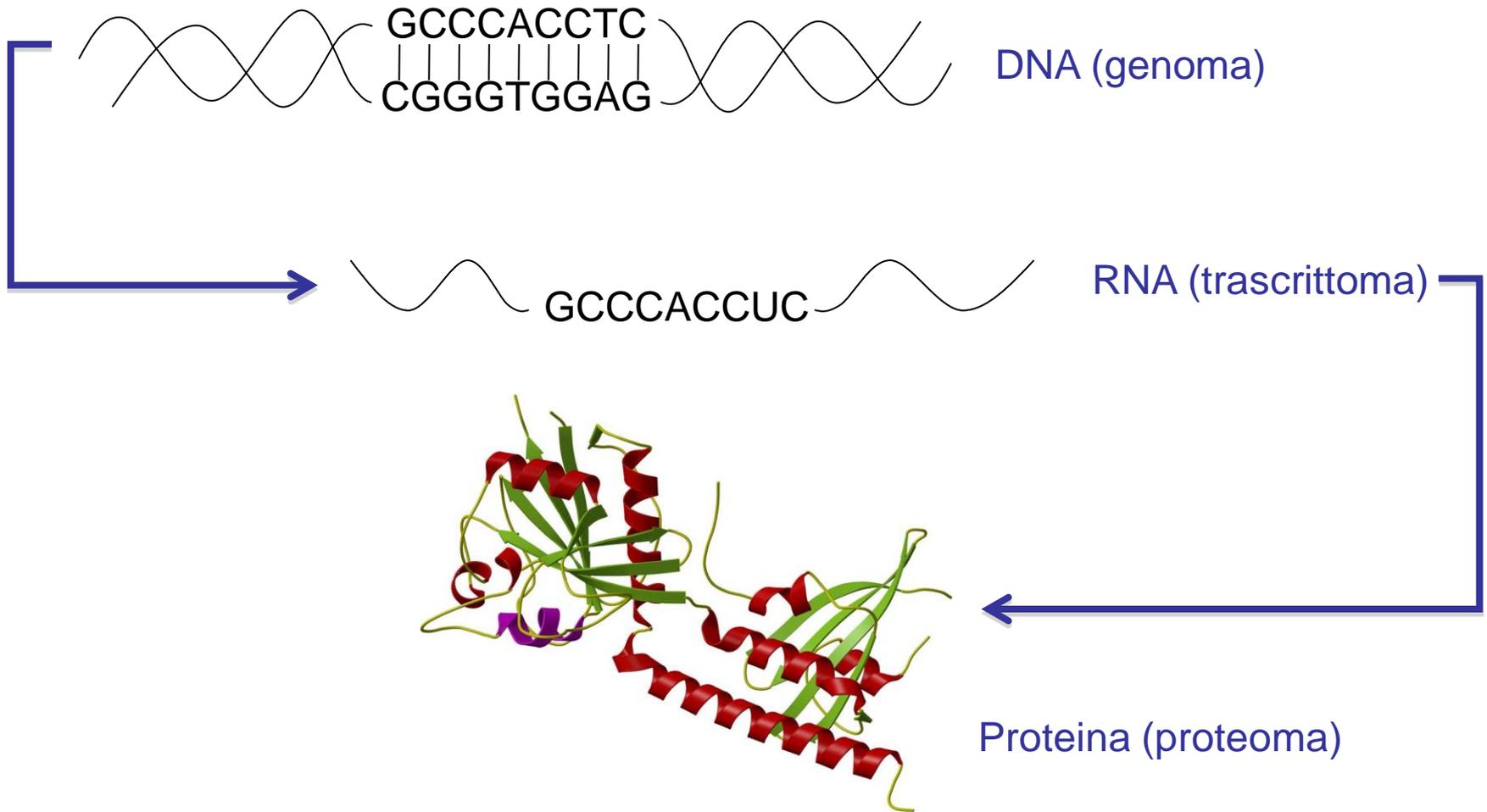
- ...una **data storica**
- ...uno **sforzo congiunto** tra Stati Uniti e Gran Bretagna, la conclusione di una gara **avvincente**
- ...si parla anche di **farmaci**, ma non perchè i concorrenti li hanno usati
- ... quanto perchè, forse, se ne potranno progettare di **migliori**



la sequenza del genoma umano



Le basi dell'espressione genica





SCIENCEPHOTOLIBRARY





Centro Interdipartale Di Ricerche Genomiche



◆ TABELLA 20.1

Obiettivi fondamentali della genomica

1. Stabilire banche dati e interfacce di ricerca integrate e disponibili in Internet.
2. Generare e ordinare sequenze genomiche e sequenze di geni espressi.
3. Identificare e annotare tutti i geni codificati da un determinato genoma.
4. Costruire atlanti di espressione genica.
5. Incrementare i dati funzionali, tra cui quelli sulle caratteristiche biochimiche e fenotipiche dei geni.
6. Caratterizzare la diversità di sequenza del DNA.
7. Fornire risorse per eseguire comparazioni fra i genomi.



FANTONI, BOZZARO, DEL SAL,
FERRARI, TRIPODI

**BIOLOGIA CELLULARE
E GENETICA**

Sequenziamento ABI 377/310

PE Navigator TM

HLA Nav. all.							
	210	220	230	240	250	260	
1 FHCRC#10F	GAG-AGTTTCG	TGCGCTTCGA	CAGCGACGTG	GGGGAGTTCC	GGGCGGTGAC	GGAGCTGGGG	
2 FHCRC#10R	GAGGAGTTTCG	TGCGCTTCGA	CAGCGACGTG	GGGGAGTTCC	GGGCGGTGAC	GGAGCTGGGG	
3 FHCRC#12F	GAGGAGTTTCG	CGCGCTTCGA	CAGCGACGTG	GGGGAGTTCC	GGGCGGTGAC	GGAGCTGGGG	
4 FHCRC#12R	GAGGAGTTTCG	CGCGCTTCGA	CAGCGACGTG	GGGGAGTTCC	GGGCGGTGAC	GGAGCTGGGG	
	270	280	290	300	310	320	
5 DPB1. CON. i	CGGCCTGMKG	NGGWSTACTG	GAACAGCCAG	AAGGACHTCC	TGGAGGAGRR	GCGGGCAKTG	
1 FHCRC#10F	CGGCCTGATG	AGGAGTACTG	GAACAGCCAG	AAGGACATCC	TGGAGGAGRA	GCGGGCAGTG	
2 FHCRC#10R	CGGCCTGATG	AGGAGTACTG	GAACAGCCAG	AAGGACATCC	TGGAGGAGRA	GCGGGCAGTG	
3 FHCRC#12F	CGGCCTGCTG	CGGAGTACTG	GAACAGCCAG	AAGGACATCC	TGGAGGAGRA	GCGGGCAGTG	
4 FHCRC#12R	CGGCCTGCTG	CGGAGTACTG	GAACAGCCAG	AAGGACATCC	TGGAGGAGRA	GCGGGCAGTG	

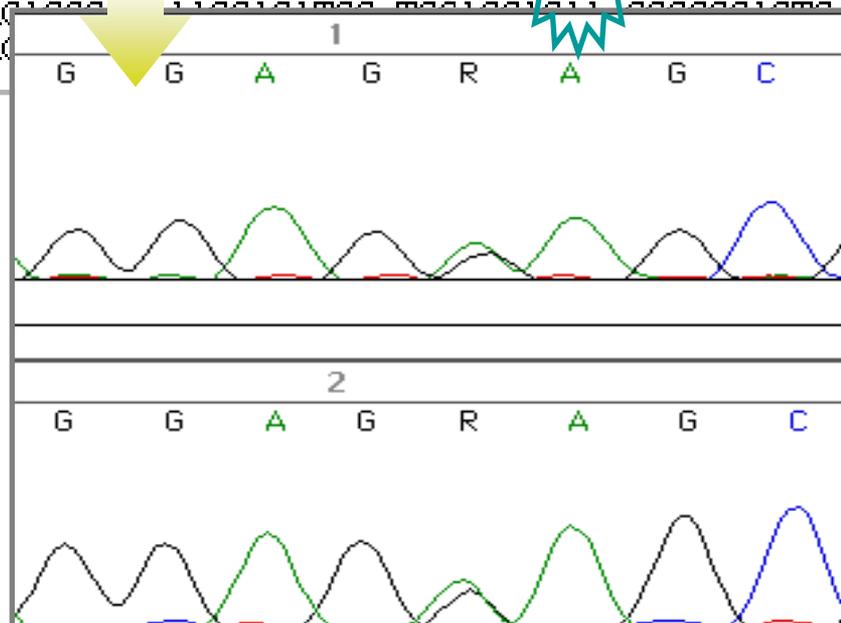
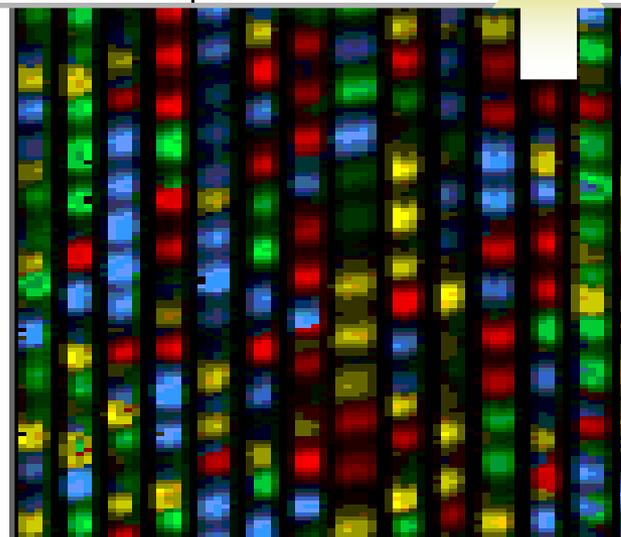
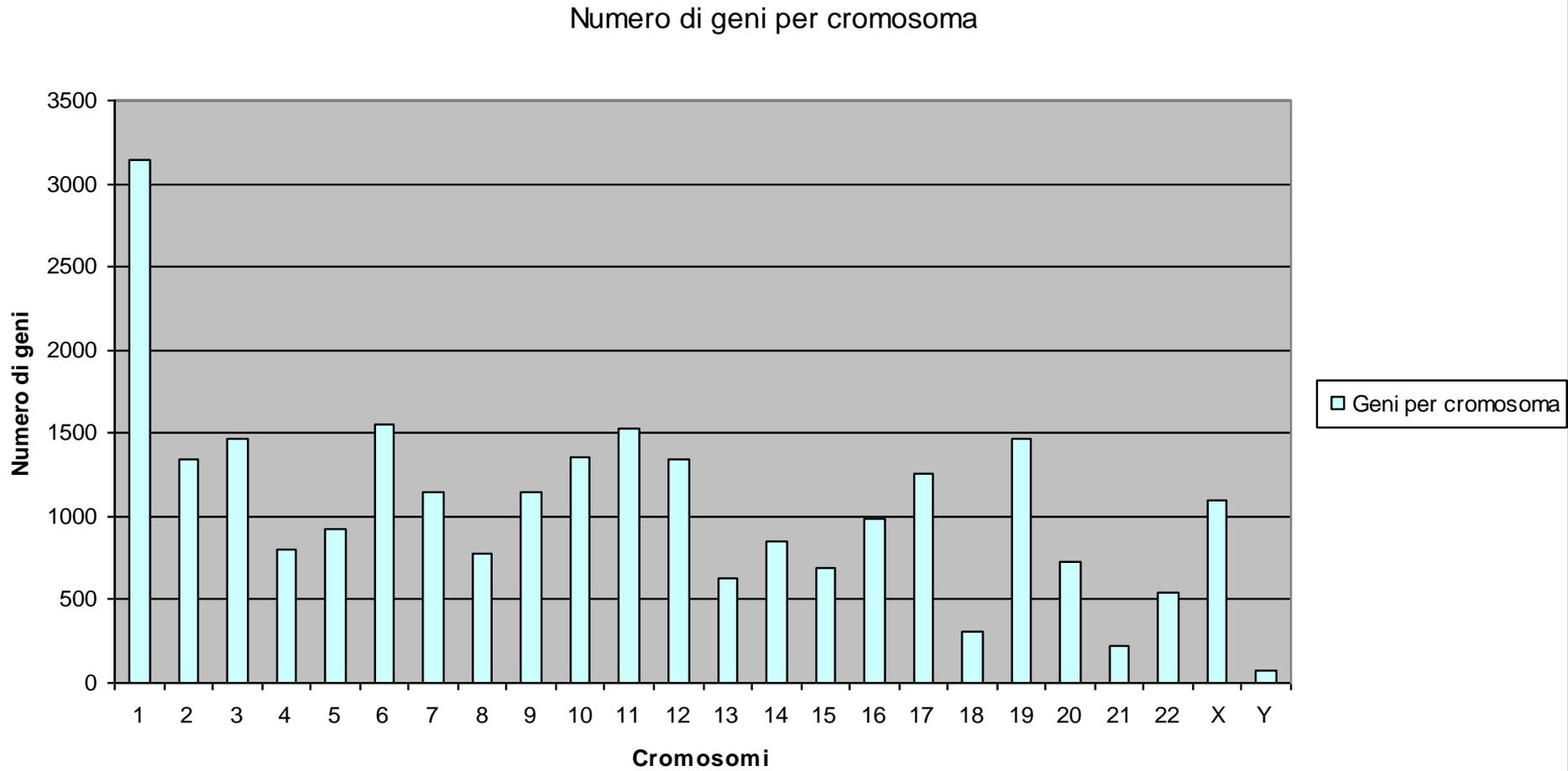


Figura 3



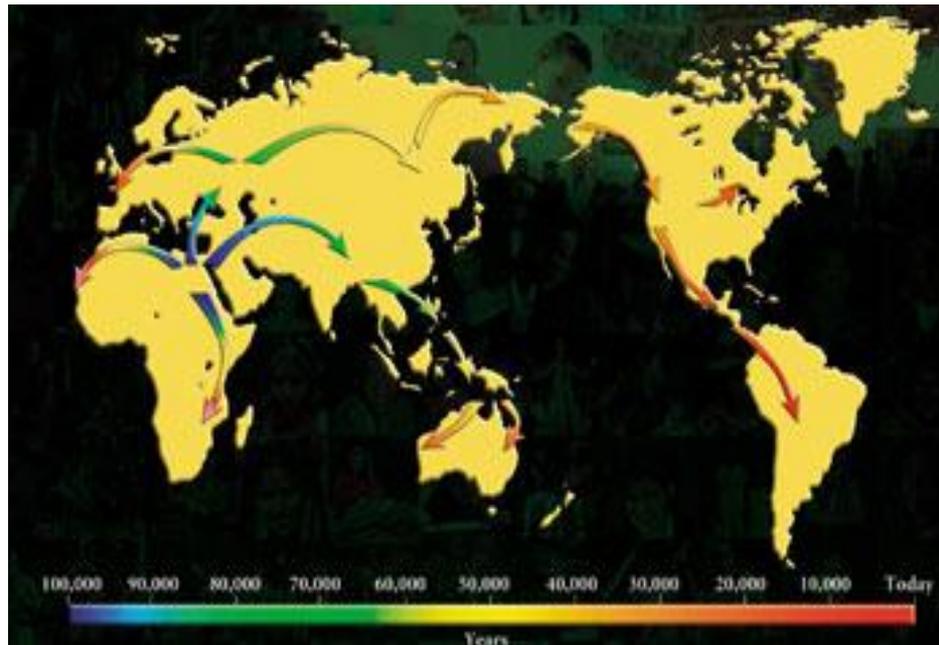
The Human Genome Project

- **Geni identificati: ~ 25.000**
- **SNPs: 9×10^6**

The Human Genome Project

Popolazioni Umane

Tutti gli individui condividono sequenze genomiche che sono al 99.9% identiche. Il rimanente 0.1% è responsabile di tutta la diversità genetica tra gli individui.



La tipizzazione delle diversità genetiche (SNPs) ci consente di tracciare l'evoluzione della razza umana e delle sue migrazioni attraverso la terra.



Human genome variation:

Molecular basis for phenotypic variation

Single nucleotide polymorphisms (SNPs)

GATTAGATC**G**CGATAG
 AG
 GATTAGATC**T**CGATAG
 AG

~0.1% of the genomes of any two individuals differ due to SNPs

Structural variants (SVs) [e.g. kb-Mb-sized deletions, insertions, inversions]

Chromosome

Genes from reference genome

Deletion: A B C

Insertion: A B D C

Inversion: C B A

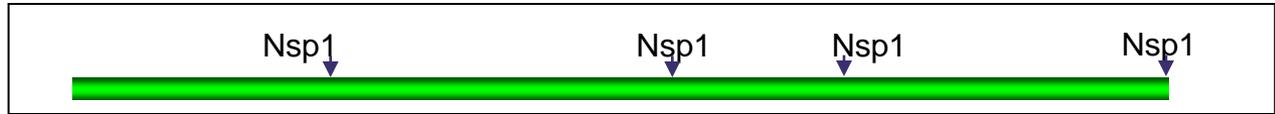
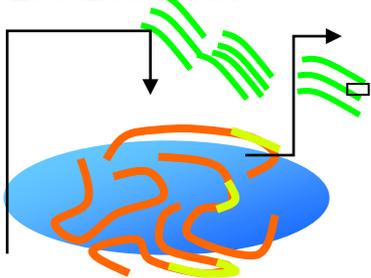
• Not well studied

• Short Interval technologies are not well suited to analyze these types of variations

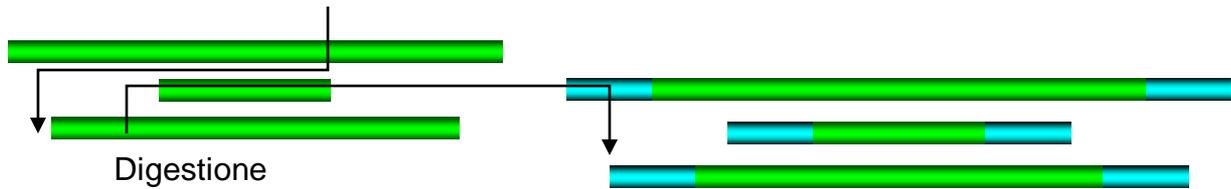
• Conventional technologies hold potential for significant discovery

presumably >0.1% of the genome [adapted from Feuk *et al.*, *Nat. Rev. Genet.* 2006; see also e.g. Freeman *et al.*, *Genome Res.* 2006, or Sharp *et al.* *Annu. Rev. Genomics Hum. Genet.* 2006.]

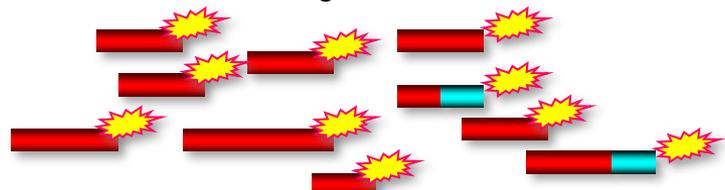
Estrazione DNA



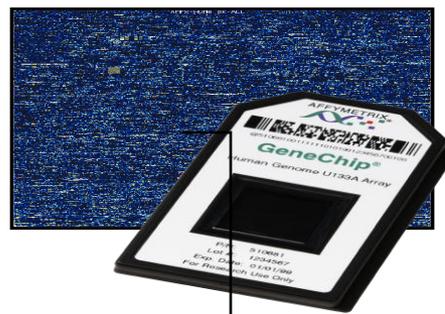
Digestione



Ligazione di adattatore

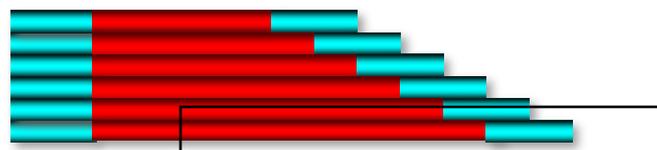


Frammentazione e marcatura

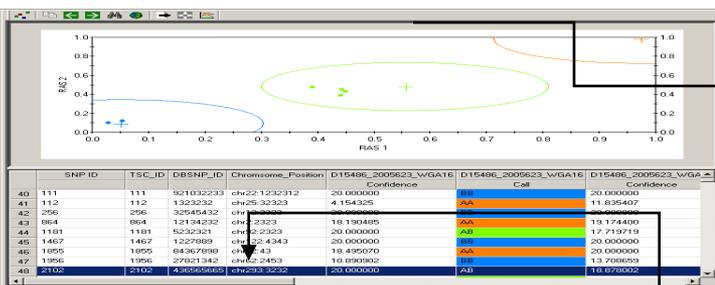


Ibridizzazione, lavaggi e scanning

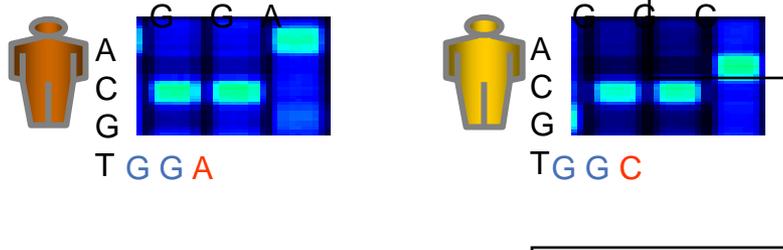
SUMMARY	
TSC ID	0001956
dbSNP RS ID	997238
Alleles	A/G
Flanking Sequence	atctcaggatcacaca[A/G]:attctagggttcattgg
PHYSICAL MAP	
Chromosome	10
Physical Position	60864097 (June 2002)
Cytoband	10q21.2



PCR



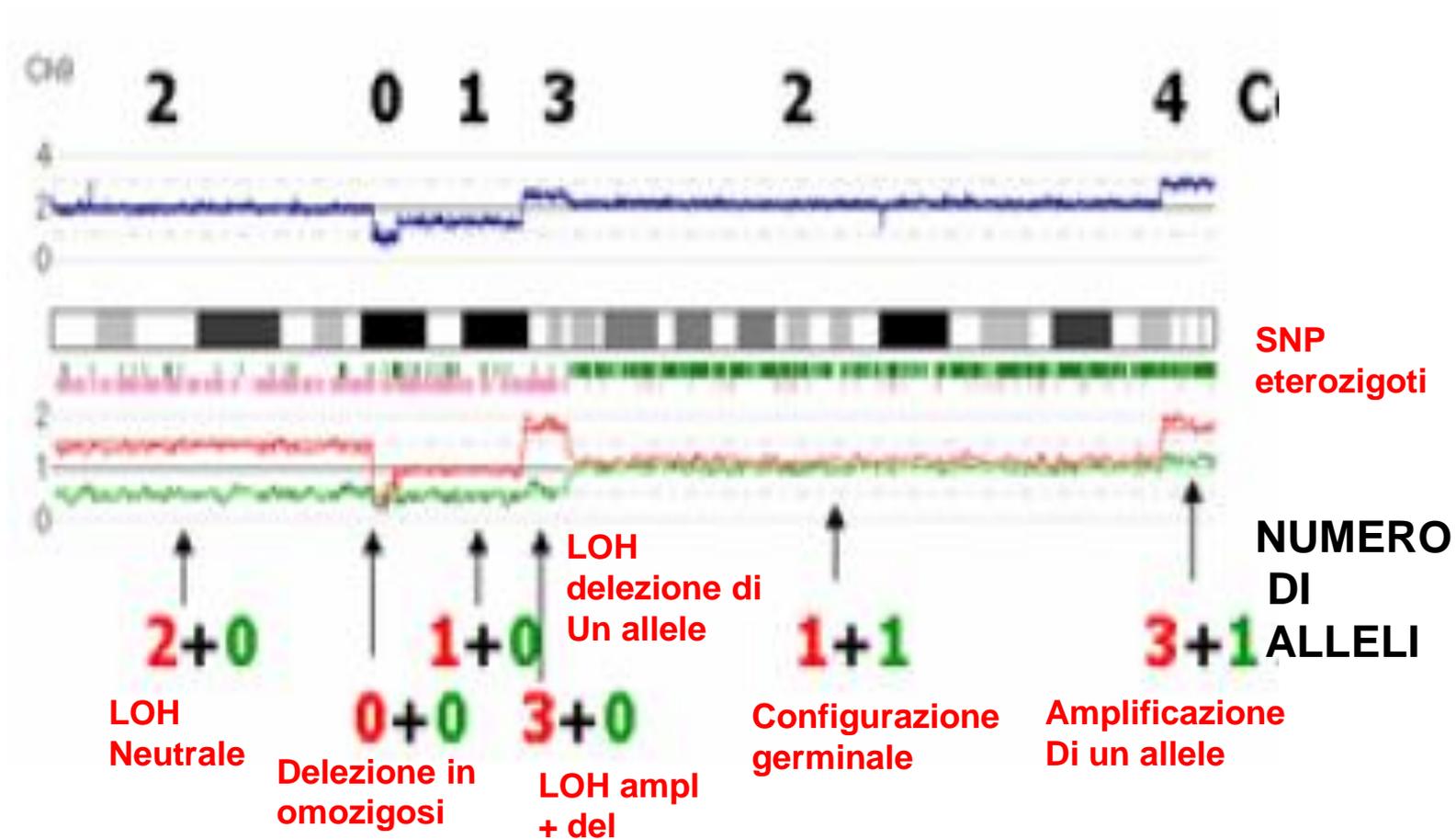
Analisi dei dati



Analisi della annotazioni dei genotipi

Cromosoma 9

NUMERO DI
COPIE DI DNA



Probabilità di omozigosi

SNP eterozigoti

NUMERO DI ALLELI

LOH Neutrale

Delezione in omozigosi

LOH ampl + del

Configurazione germinale

Amplificazione Di un allele



ChAS: Intuitive Graphical Interface

Upper pane: Karyoview, segments, cytoregions, overlap map, graphs

The screenshot displays the Chromosome Analysis Suite (ChAS) interface. The main window is titled "Chromosome Analysis Suite [For Research Use Only. Not for use in diagnostic procedures.]". The top menu bar includes File, View, Reports, Analysis, References, and Help. A toolbar contains various icons for file operations and analysis. The main workspace is divided into several panes:

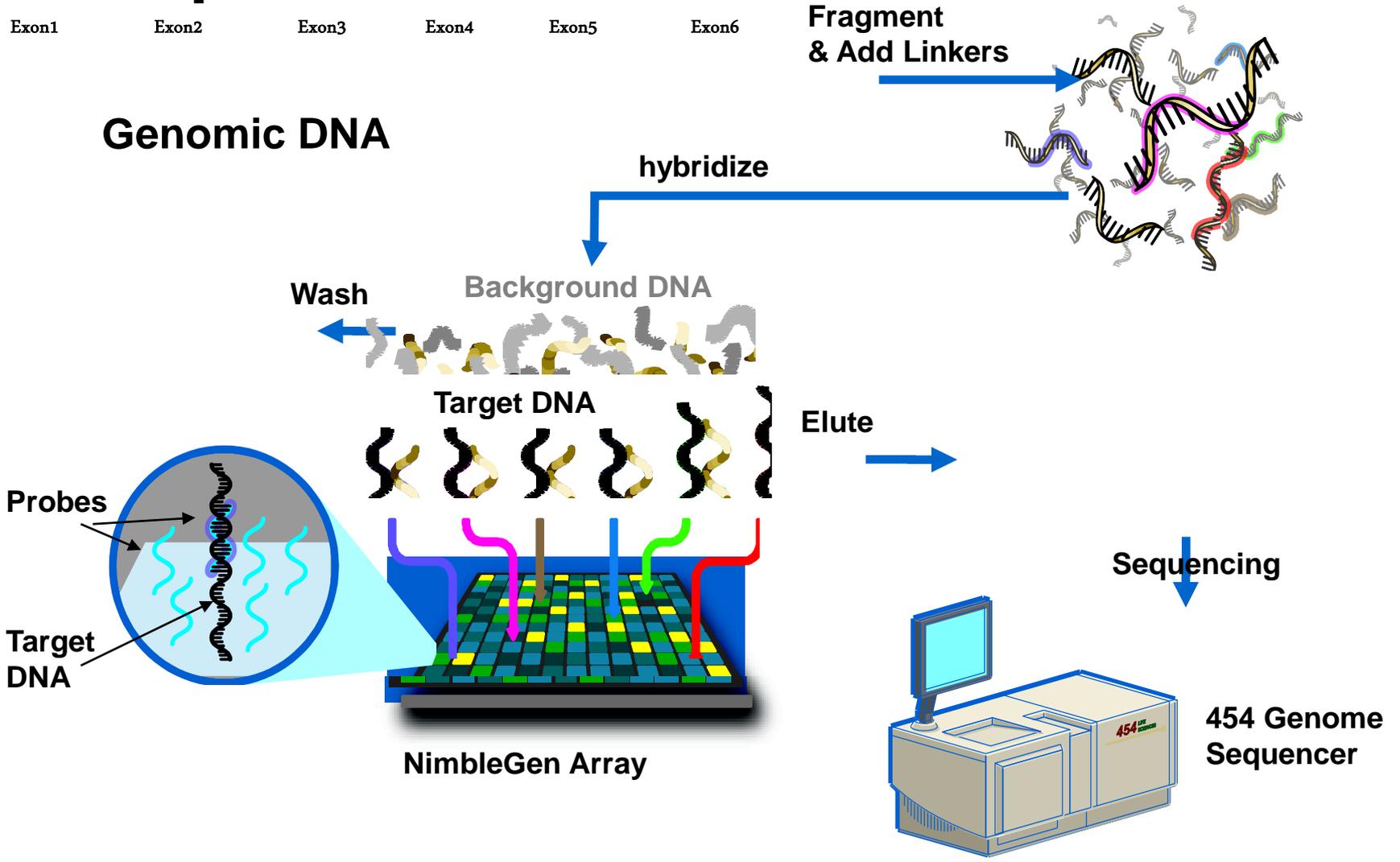
- Files list:** Located on the left, it contains a list of files and data types, including "711 Cytogenetics R", "BACs", "DGV", "FISH Clones", "Genes", "OMIM", "Segmental Duplicat", and "sno/miRNA".
- Data types:** Located below the Files list, it shows a list of data types with checkboxes for "Gain", "Loss", "Copy Number State", "Log2 Ratio", "Weighted Log2 Ratio", "Mosaicism", "LCSH", "SmoothSignal", and "Allele Peaks".
- Upper pane:** This pane is divided into five tabs: "Karyoview", "Segments", "CytoRegions", "Overlap Map", and "Graphs". The "Karyoview" tab is currently active, showing a karyotype of human chromosomes. A blue box highlights a specific region on chromosome 12.
- Detail view:** Located below the upper pane, it shows a detailed view of the selected region. It includes a list of data types for the selected region, such as "Beta1_M_01_Cyto_VH_20090202.Cytogenetics.cyto2.cychp: Copy Number State (segments)", "Beta1_M_01_Cyto_VH_20090202.Cytogenetics.cyto2.cychp: Copy Number State", "Beta1_M_01_Cyto_VH_20090202.Cytogenetics.cyto2.cychp: Weighted Log2 Ratio", "Beta1_M_01_Cyto_VH_20090202.Cytogenetics.cyto2.cychp: Allele Peaks", "711 Cytogenetics Relevant and Haploinsufficiency Genes_Custom Regions.bed", "DGV", "FISH Clones", "Genes", "OMIM", and "Markers Cytogenetics_Array dbShp: 128". Below this list is a genomic track showing the selected region from 70000kb to 85000kb, with labels for p31.3, p31.1, and p22.3.
- Selected chromosome view:** Located on the right, it shows a vertical view of the selected chromosome (12) with a green bar indicating the selected region. Labels for p31.3, p31.1, q12, q32.1, q41, and q43 are visible.

Annotations with brackets and labels are present:

- "Files list" points to the left-hand file browser.
- "Data types" points to the data type selection panel.
- "Detail view" points to the bottom pane showing genomic tracks.
- "Selected chromosome view" points to the right-hand vertical chromosome view.

At the bottom left, it says "Restricted Mode: Off". At the bottom right, it says "User: dflauc".

Microarray Sequence Capture



Unique combination of read length & reads per run ensure the broadest applications portfolio

De Novo Sequencing

- *Microorganisms* (genome plasticity)
- Plants, Animals
- BACs, YACs, Fosmids, Viruses etc.

Resequencing

- Whole Genomes
- Disease associated regions
- Somatic mutations (cancer research via amplicon sequencing)
- Structural variations of the human genome

Transcriptome Analysis

- Expression profiling (e.g. SAGE-like, CAGE-like, GIS-PET)
- EST-sequencing
- Full length cDNA sequencing

Gene Regulation Studies

- Identification of transcription factor binding sites (ChIP-Sequencing)
- Identification and quantification of sncRNAs sequences

Epigenetic Changes

- DNA-Methylation patterns

Metagenomes & Microbial Diversity

- Shotgun sequencing of the metagenome
- 16S amplicon sequencing

Ancient DNA

- Neanderthals, Mammoths and many more



Structural Variation of the Human Genome

Andrew J. Sharp, Ze Cheng, and Evan E. Eichler

Department of Genome Sciences, University of Washington, Howard Hughes Medical Institute, Seattle, Washington 98195; email: eee@gs.washington.edu

Annu. Rev. Genomics Hum. Genet. 2006.
7:407-42

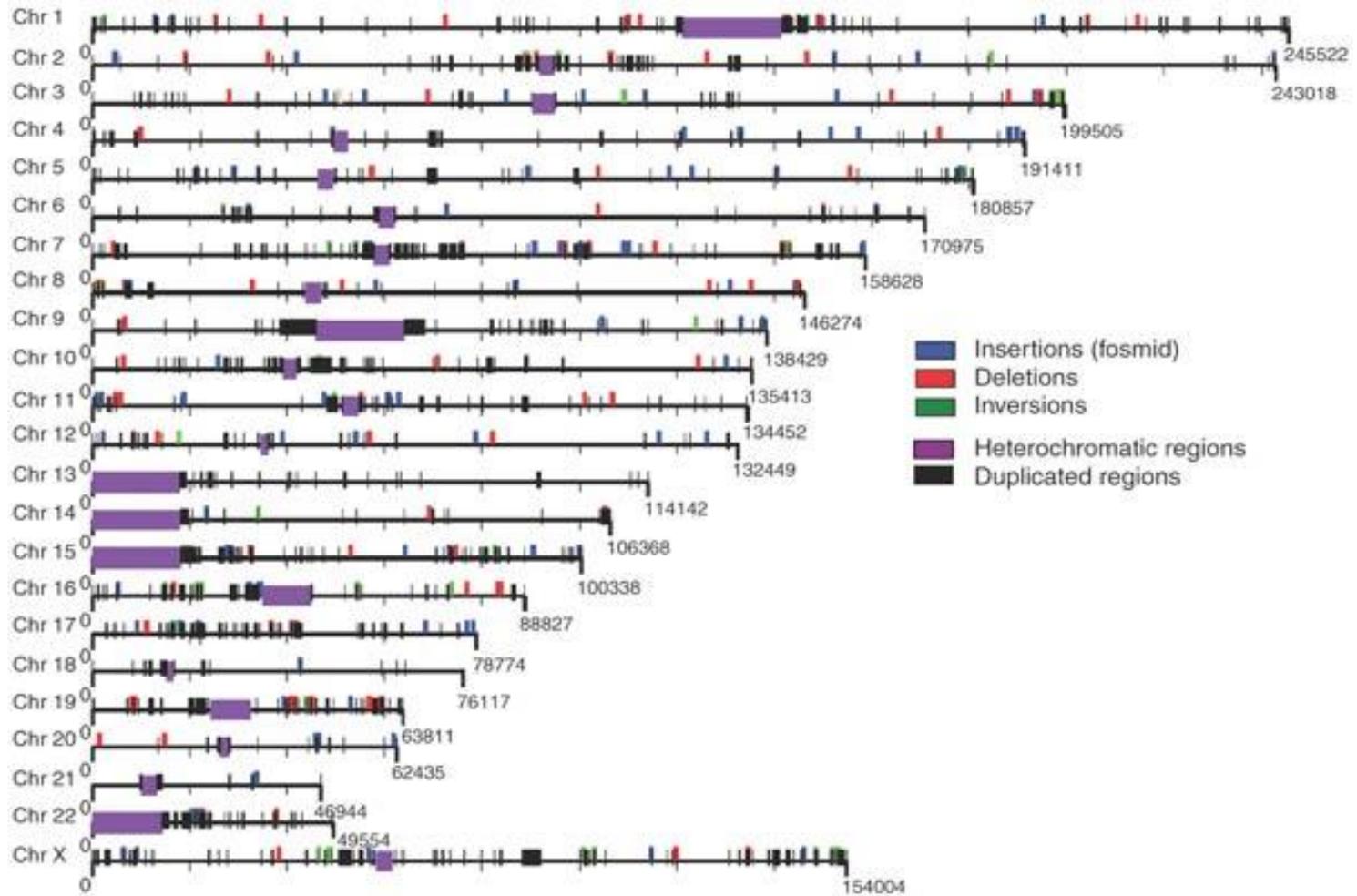


Figure 2. Structural variation map.

The schematic summarizes the distribution of insertions, deletions and inversions on each human chromosome. A total of 297 clusters were identified: 139 insertions, 102 deletions and 56 inversions breakpoints ([Supplementary Table 1](#) online). Across the genome, 163 of the structural variants map to regions of segmental duplication.

The Diploid Genome Sequence of an Individual Human

Samuel Levy^{1*}, Granger Sutton¹, Pauline C. Ng¹, Lars Feuk², Aaron L. Halpern¹, Brian P. Walenz¹, Nelson Axelrod¹, Jiaqi Huang¹, Ewen F. Kirkness¹, Gennady Denisov¹, Yuan Lin¹, Jeffrey R. MacDonald², Andy Wing Chun Pang², Mary Shago², Timothy B. Stockwell¹, Alexia Tsiamouri¹, Vineet Bafna³, Vikas Bansal³, Saul A. Kravitz¹, Dana A. Busam¹, Karen Y. Beeson¹, Tina C. McIntosh¹, Karin A. Remington¹, Josep F. Abril⁴, John Gill¹, Jon Borman¹, Yu-Hui Rogers¹, Marvin E. Frazier¹, Stephen W. Scherer², Robert L. Strausberg¹, J. Craig Venter¹

1 J. Craig Venter Institute, Rockville, Maryland, United States of America, 2 Program in Genetics and Genomic Biology, The Hospital for Sick Children, and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada, 3 Department of Computer Science and Engineering, University of California San Diego, La Jolla, California, United States of America, 4 Genetics Department, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

ARTICLES

Initial sequence of the chimpanzee genome and comparison with the human genome

The Chimpanzee Sequencing and Analysis Consortium*

Here we present a draft genome sequence of the common chimpanzee (*Pan troglodytes*). Through comparison with the human genome, we have generated a largely complete catalogue of the genetic differences that have accumulated since the human and chimpanzee species diverged from our common ancestor, constituting approximately thirty-five million single-nucleotide changes, five million insertion/deletion events, and various chromosomal rearrangements. We use this catalogue to explore the magnitude and regional variation of mutational forces shaping these two genomes, and the strength of positive and negative selection acting on their genes. In particular, we find that the patterns of evolution in human and chimpanzee protein-coding genes are highly correlated and dominated by the fixation of neutral and slightly deleterious alleles. We also use the chimpanzee genome as an outgroup to investigate human population genetics and identify signatures of selective sweeps in recent human evolution.

La rivoluzione “omica”

Human Genome Project (2003)

Viene pubblicata la sequenza eucromatica del genoma umano.

Nature 2004; 431 (7011): 931-945.

Phase I HapMap project (2005)

Viene prodotta la prima mappa degli aplotipi.

Nature 2005; 437(7063):1299-1320

Encyclopedia of DNA Elements (ENCODE) project (2007)

Identificazione e analisi degli elementi funzionali (in fase pilota) nell' 1% del genoma umano.

Nature 2007; 447(7146):799-816

1000 Genomes Project in 2008

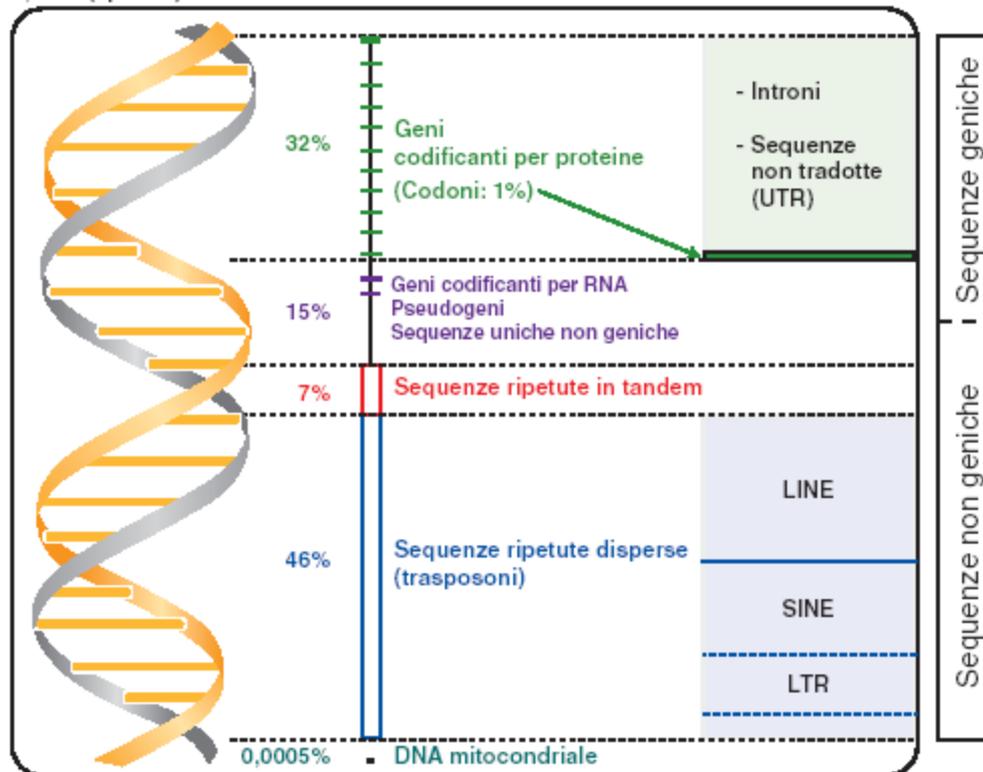
Un piano per catturare le diversità umane in 1000 genomi.

Science 2008; 319(5863):395

TABELLA 1.2 Il contenuto in DNA delle cellule

Organismo	Contenuto in DNA aploide (in milioni di bp)
Batteri	
<i>Mycoplasma</i>	0,6
<i>E. coli</i>	4,6
Eucarioti unicellulari	
<i>Saccharomyces cerevisiae</i> (lievito)	12
<i>Dictyostelium discoideum</i>	70
<i>Euglena</i>	3000
Piante	
<i>Arabidopsis thaliana</i>	125
<i>Zea mays</i> (granoturco)	5000
Animali	
<i>Caenorhabditis elegans</i> (nematode)	97
<i>Drosophila melanogaster</i> (mosca della frutta)	180
Pollo	1200
Pesce zebra	1700
Topo	3000
Uomo	3000

GENOMA UMANO
3,2 Gb (aploide)



◆ FIGURA 33.17

Organizzazione generale del genoma umano. Lo schema rappresenta l'estensione proporzionale complessiva delle diverse classi di sequenze nel genoma umano. L'area occupata sul genoma dai geni codificanti per proteine ed RNA è disegnata in modo da richiamare simbolicamente la particolare suddivisione in esoni ed introni tipica di queste sequenze (gli esoni sono simboleggiati da barre orizzontali). Le percentuali sono indicative, perché non tengono conto delle sovrapposizioni parziali di alcuni tipi di sequenza nella stessa regione. La percentuale complessiva di codoni codificanti è stimata intorno all'1,2%, tuttavia qui sono stati considerati solo i geni la cui espressione sia stata confermata sperimentalmente all'Ottobre 2006, secondo i dati della Tabelle 33.4 e 33.5. Altre spiegazioni sono date nel testo.

A. FANTONI - S. BOZZARO - G. DEL SAL
S. FERRARI - M. TRIPODI

BIOLOGIA CELLULARE E GENETICA

Parte Seconda
GENETICA

A cura di
A. FANTONI - M. TRIPODI



PICCIN

FANTONI, BOZZARO, DEL SAL,
FERRARI, TRIPODI

BIOLOGIA CELLULARE
E GENETICA (Parte 2°)

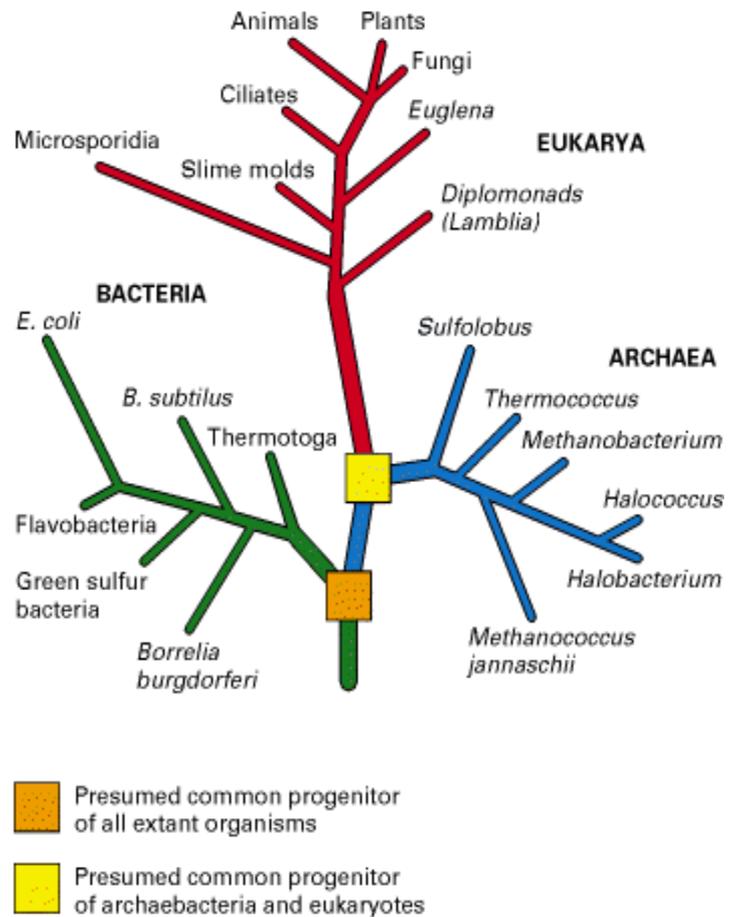


Figure 1-5. The three kingdoms of organisms are related through common sequences of their ribosomal RNAs. Their lineage depicts a view of how all life on earth, from simple bacteria to complex mammals, evolved from a common, single-celled progenitor.



Figura 1.18 *Drosophila melanogaster*
(Darwin Dale/Photo Researchers,
Inc.).

Figura 1.19 *Arabidopsis thaliana*
(Jeremy Burgess/Photo Researchers,
Inc.).



(A)



Figura 1.21 Pesce zebra

(A) Un embrione dopo 24 ore di crescita. (B) Un pesce adulto. (A, cortesemente fornita da Charles Kimmel, University of Oregon; B, cortesemente fornita da S. Kondo)

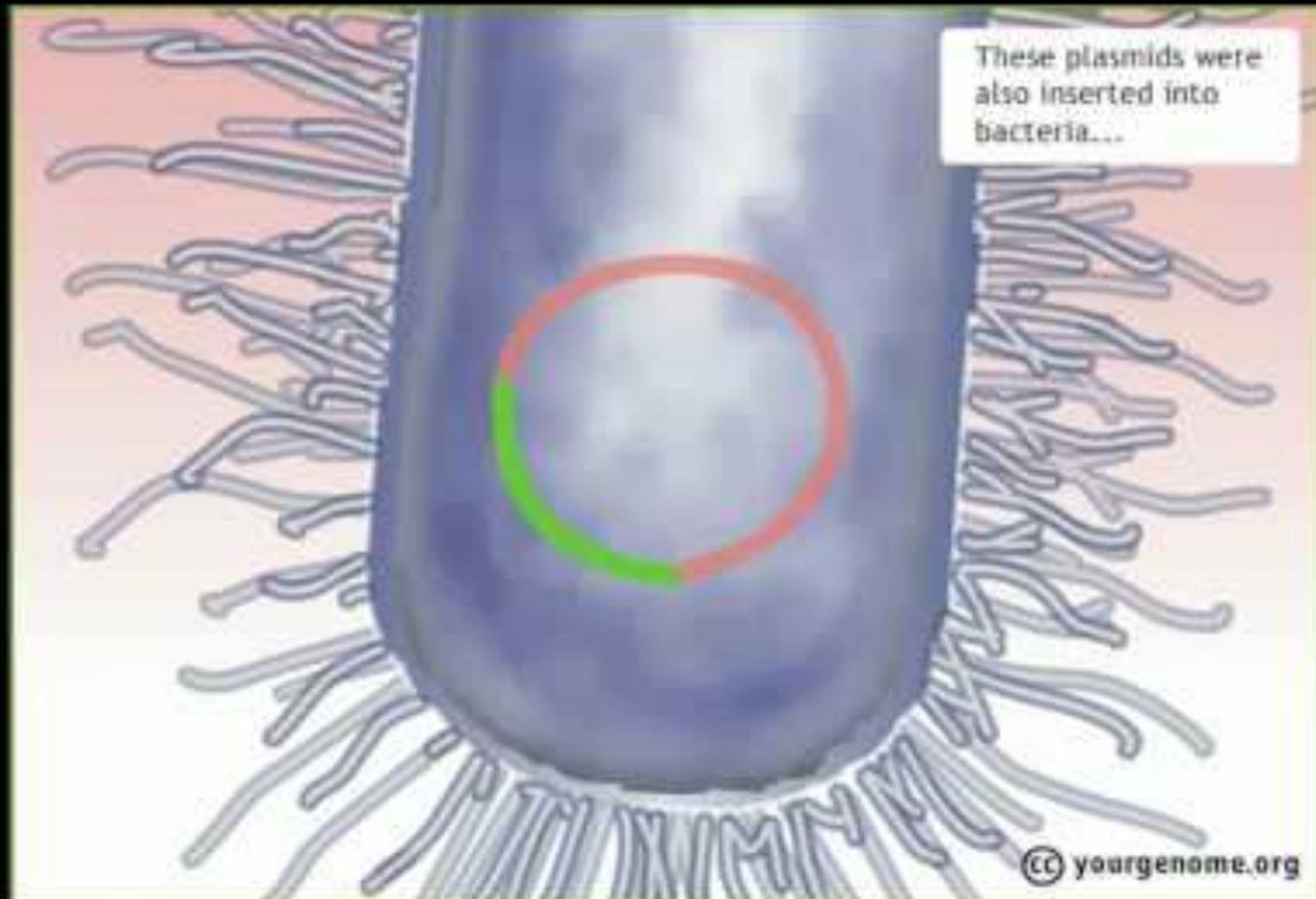
(B)



Figura 1.6 Immagine di *E. coli* al microscopio elettronico

La cellula è circondata dalla parete cellulare, all'interno della quale si trova la membrana plasmatica. Il DNA è compattato nel nucleotide. (Menge e Wurtz/Biozentrum, University of Basel/Science Photo Library/Photo Researchers, Inc.).



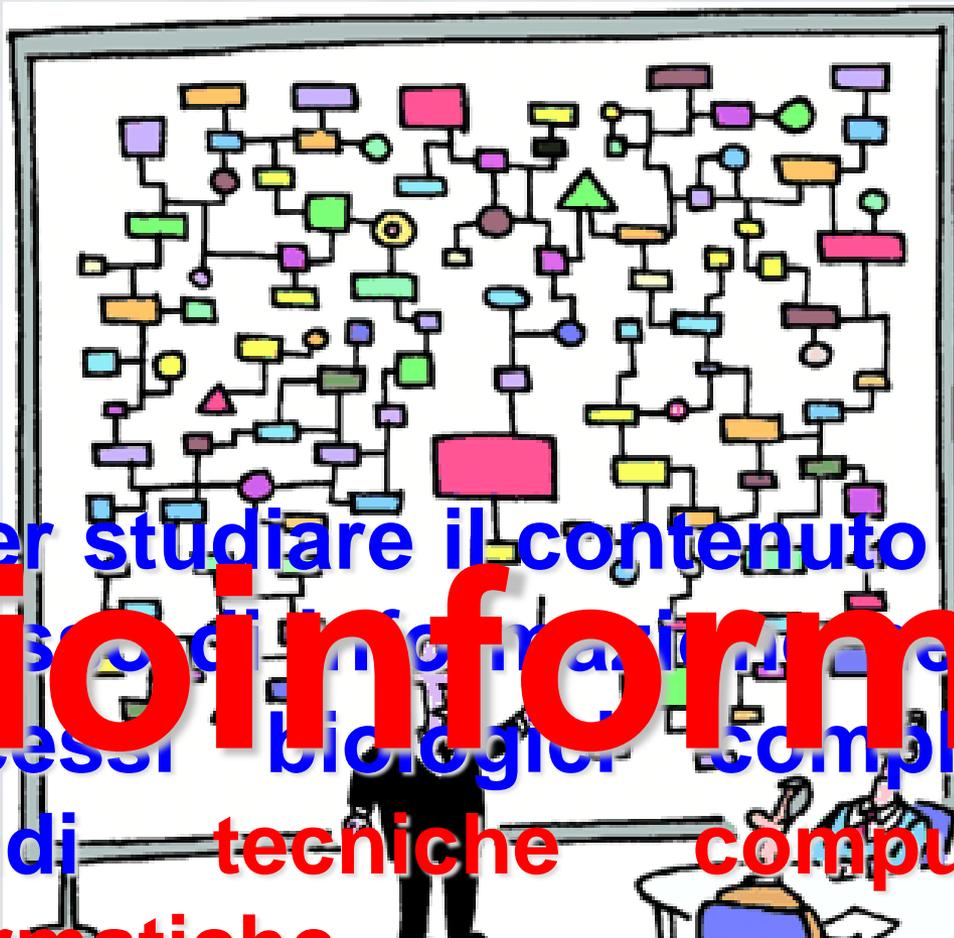
A diagram of a bacterium, shown as a blue, cylindrical cell with a textured surface of flagella. Inside the cell, a circular plasmid is depicted, divided into two segments: a red segment and a green segment. The background is a light pinkish-white gradient.

These plasmids were also inserted into bacteria...

Applicazioni delle Biotecnologie

Semplici batteri in grado di:

- ✓ Produrre carburanti puliti fino a sostituire il petrolio come fonte di energia
- ✓ Sintetizzare antibiotici per combattere le nuove infezioni del pianeta
- ✓ Eliminare la CO₂ in eccesso nell'atmosfera
- ✓ Frenare il riscaldamento globale



... per studiare il contenuto informativo ed
Bioinformatica
... di un sistema biologico e di un sistema
processi biologici complessi servono
quindi tecniche computazionali ed
informatiche



"And that's why we need a computer."



Where can I find annotations?



Genes
Proteins
Expression levels



Sequences
Genes
Proteins



Map Location



Expression levels



Microarray annotations



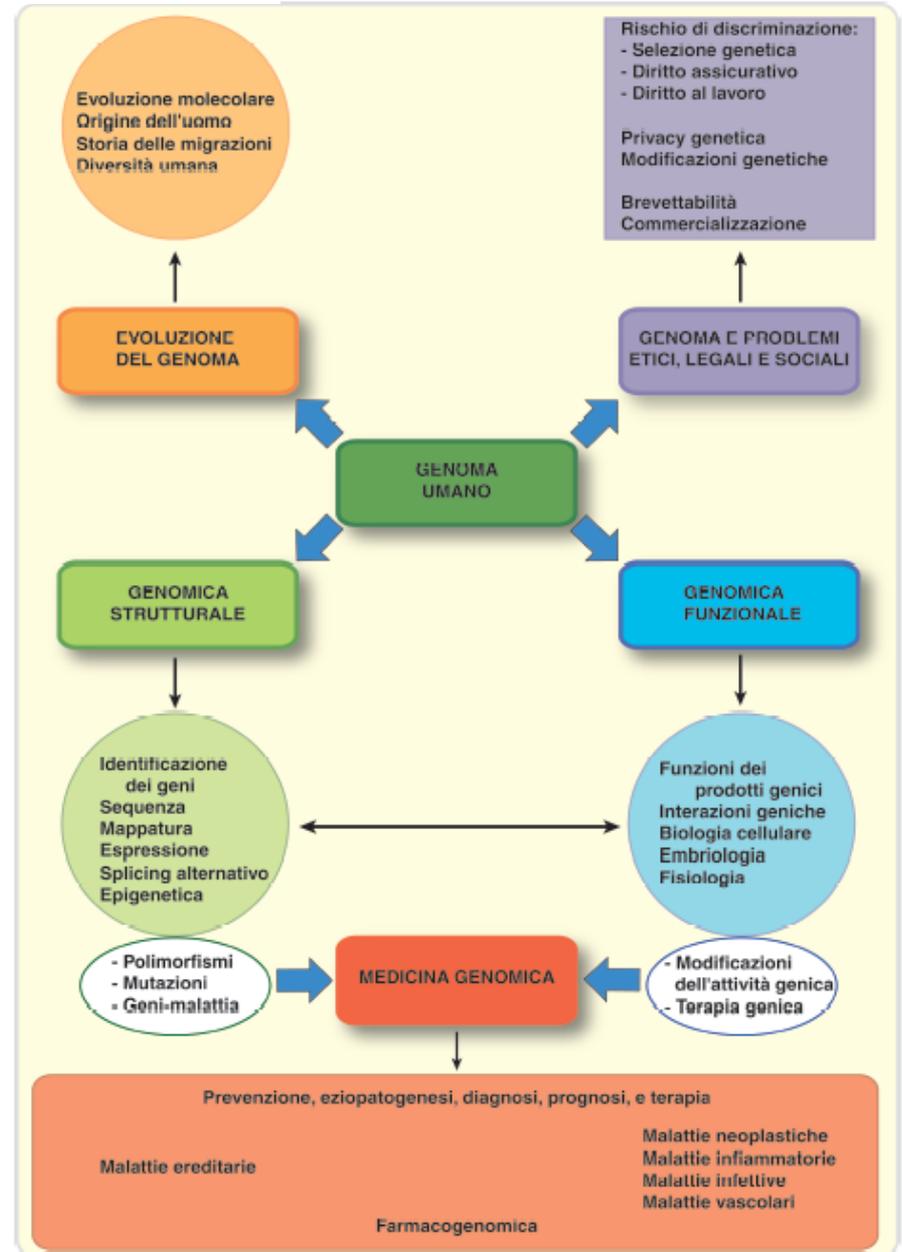
Genes & pathology
Allelic variants
Animal models



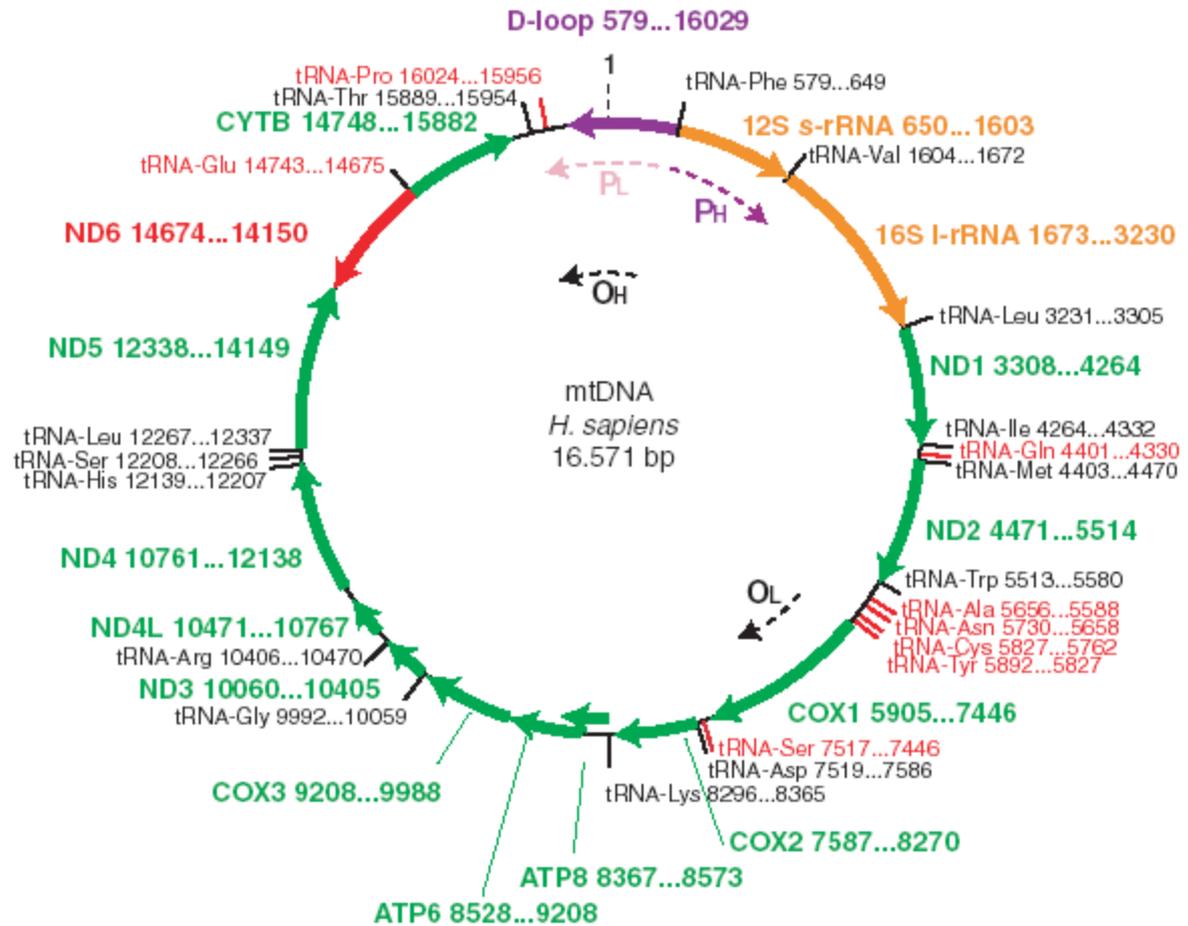
Literature!

♦ FIGURA 33.1

Le ricadute del sequenziamento del genoma umano in biologia e medicina. Spiegazione nel testo. Il concetto base dell'illustrazione è ripreso da Vezzoni et al., 1998.



FANTONI, BOZZARO, DEL SAL,
FERRARI, TRIPODI
**BIOLOGIA CELLULARE
E GENETICA (Parte 2°)**

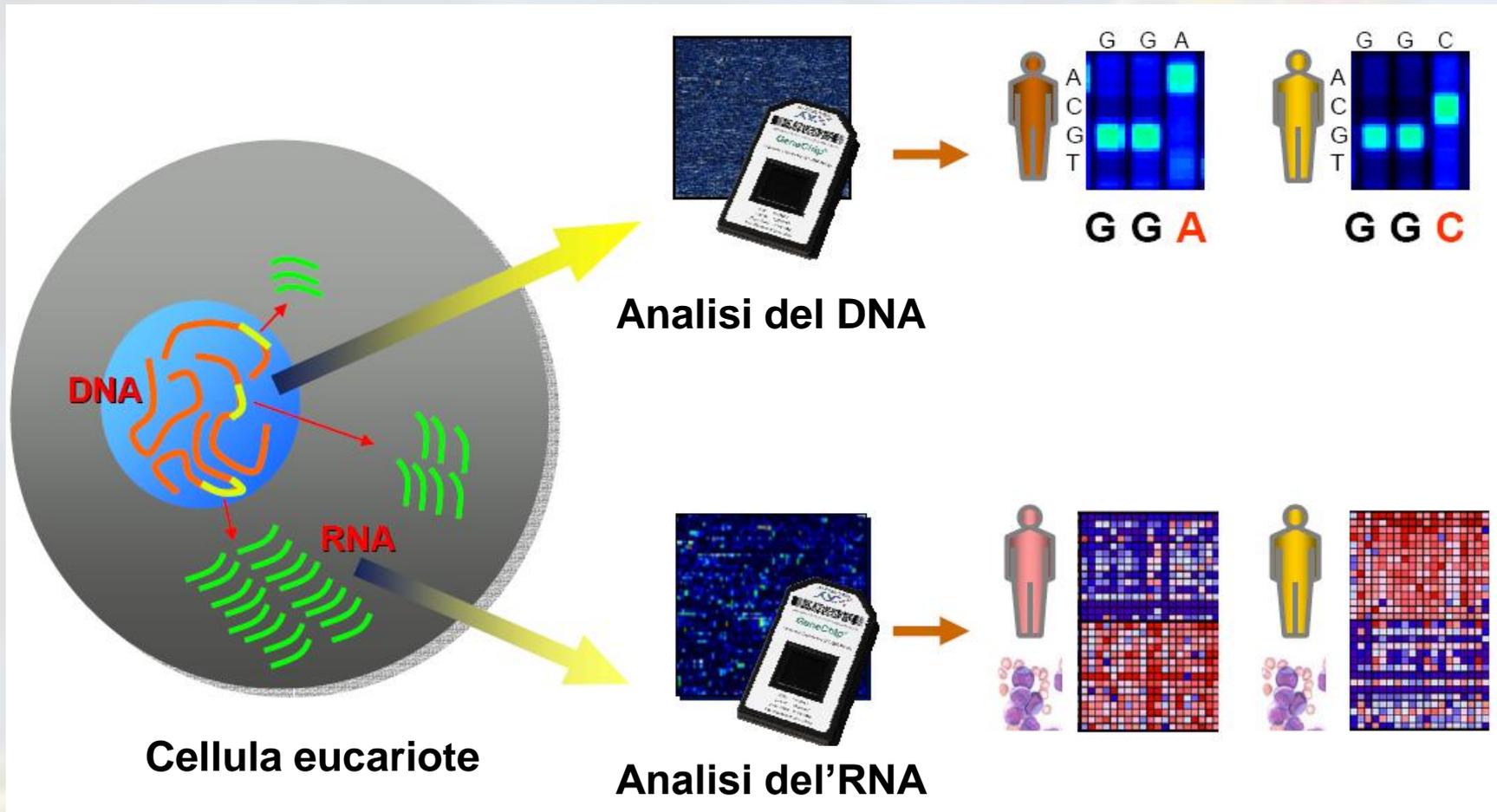


◆ **FIGURA 33.34**

Mapa del DNA mitocondriale umano. In verde i geni codificanti per proteine, in arancione i geni per gli rRNA, in viola la regione a tripla elica. È indicata la direzione di trascrizione per ciascun gene e sono evidenziati in colore rosso i geni la cui sequenza codificante si trova sul filamento "L" (in antisenso rispetto al filamento "H"). PH e PL: avvio e direzione della trascrizione per il filamento, rispettivamente, "H" e "L"; OH e OL: origine e direzione della replicazione del filamento, rispettivamente, "H" e "L". Sono indicate le coordinate esatte per ogni gene, riferite alla sequenza *GenBank* n. NC_001807.4 (filamento "H").



Analisi del trascrittoma





Affymetrix GeneChip platform 450: GeneChip Scanner 3000 and GeneChip Fluidics Station 450

X LAB



♦ TABELLA 33.3

Principali tipi di RNA non codificanti per proteine (ncRNA)

Le stime si riferiscono a quelli meglio caratterizzati, ma il numero effettivo totale può essere molto più grande (vedi testo). I geni mitocondriali non sono stati considerati. I simboli indicati per ciascun tipo possono servire a rintracciare informazioni maggiori per ciascun RNA o proteina nelle banche dati. nt: lunghezza in nucleotidi dell'RNA; Pol: RNA polimerasi implicata nella trascrizione dell'ncRNA; GenBank: numero di accesso alla sequenza nella banca dati GenBank (*Entrez Nucleotide*); RP: proteine ribosomiali; L e S: subunità ribosomiale maggiore e minore, rispettivamente; TF: fattori implicati nella traduzione; SRP: *signal recognition particle* (particella di riconoscimento del segnale).

Classe/Tipo	Processo biologico	Interazione molecolare (principali molecole riconosciute)	Localizzazione cellulare	Lunghezza (in nucleotidi, nt)	Numero di geni (per genoma aploide)	Pol	GenBank
<i>rRNA (RNA ribosomiali)</i>							
28S	Traduzione	rRNA 5,8S, RP, TF	Ribosoma (L)	5.025 nt	~400	I	M11167
18S	Traduzione	RP, TF	Ribosoma (S)	1.868 nt	~400	I	M10098
5,8S	Traduzione	rRNA 28S, RP	Ribosoma (L)	159 nt	~400	I	JO1866
5S	Traduzione	rRNA 28S, RP	Ribosoma (L)	121 nt	~1.000	III	V00589
<i>tRNA (RNA transfer)</i>							
tRNA	Traduzione	mRNA, rRNA, RP, TF	Citoplasma	69-99 nt	450 (48 diversi anticodori)	III	es. tRNA(Gln): X56184
<i>snRNA (piccoli RNA nucleari)</i>							
<i>U1 (RNU1)</i>	Splicing Inizio della trascrizione Ciclo cellulare?	pre-mRNA; U1A TFIIH CCNH (ciclina H)	Spinosoma Nucleo Nucleo	163-173 nt	30 (cluster in 1p36.3)	II	es. <i>RNU1A</i> : J00318
<i>U2 (RNU2)</i>	Splicing	pre-mRNA; <i>U6</i>	Spinosoma	186 nt	6-30 (cluster in 17q21-q22)	II	X59360
<i>U6 (RNU6)</i>	Splicing	pre-mRNA; <i>U2</i>	Spinosoma	106 nt	44	III	es. <i>RNU6A</i> : M14486
<i>7SK (RN7SK)</i>	Trascrizione	CDK9/CCNT1 (regola la Pol II)	Nucleo	331 nt	1 (6p12.1)	III	X05490
<i>snoRNA (piccoli RNA nucleolari) (I geni per questi RNA sono spesso localizzati negli introni di altri geni)</i>							
C/D box (famiglia)	Metilazione dell'RNA	rRNA (snRNA, mRNA)	Nucleolo	~70 nt	>150 loci	III	es. <i>SNORD2</i> : AJ609488
H/ACA box (famiglia)	Pseudouridilazione dell'RNA	rRNA (snRNA, mRNA)	Nucleolo; corpi di Cajal	~130 nt	>90 loci	III	es. <i>SNODA1</i> : AJ609425
<i>TERC</i>	Sintesi dei telomeri	DNA telomerico	Nucleo (telomerasi)	451 nt	1 (3q26)	II	NR_001566
<i>HBII-5</i>	Splicing	mRNA di HTR2C (recettore per la serotonina)	Spinosoma	76 nt	1 (11q11.2)	II	AC124303
<i>Altri ncRNA</i>							
<i>XIST</i>	Inattivazione dell'X	Cromosoma X	Nucleo	19.271 nt	1 (Xq13.2)	II	NR_001564
<i>RNA antisense</i>	Regolazione della espressione genica	RNA					
es. <i>TSIX</i> es. <i>NUDT6</i>	Inattivazione dell'X Regolazione di <i>FGF2</i>	RNA di <i>XIST</i> mRNA di <i>FGF2</i>	Nucleo Citoplasma	>30.000 nt 1.195 nt	1 (Xq13.2) 1 (4q26)	II II	AL353804 NM_007083
<i>7SL (RN7SL1)</i>	Trasporto delle proteine	Proteine della SRP	SRP	299 nt	1 (14q22.1)	III	X04248
<i>miRNA (microRNA)</i>	Regolazione della espressione genica	RNA	Citoplasma	22 nt	>450 loci dispersi (3 cluster paraloghi nei cromosomi: 13, X e 7)	III	es. <i>MIRN20A</i> : AJ421738

Regioni non codificanti di pre-mRNA e mRNA: 5' UTR; 3' UTR; Introni

A. FANTONI - S. BOZZARO - G. DEL SAL
S. FERRARI - M. TRIPODI

BIOLOGIA CELLULARE E GENETICA

Parte Seconda
GENETICA

A cura di
A. FANTONI - M. TRIPODI



PICCIN



MAS 5.0: Microarray Suite 5.0



- Controllo strumentazione
- Rilevazione intensita' di segnale
- Analisi assolute
- Analisi comparative

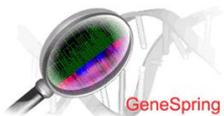
MicroDB: Microarray Data Base

- Database di gestione dei dati

DMT: Data Mining Tool

- Elaborazione dati contenuti nei database generati da MicroDB
- Opzioni di "filtering", "sorting", "querying", ecc...
- Analisi statistica (medie, deviazioni standard, t-test, ecc...)
- Clustering: SOM (Self Organizing Maps)
Pearson's Correlation Coefficient Clustering

Gene Spring™ 5.0



- Visualizzazione dei dati mediante "scatter plots", dendrogrammi, mappe cromosomiche, "pathways", ecc...
- Opzioni di "filtering", "sorting", "querying", ecc...
- Analisi statistica (ANOVA, normalizzazione, ecc..)
- Clustering: "gene tree", "experiment tree", SOM, "k-means", ecc...

GO Mining Tool

- Go mapping:
identificazione dei processi biologici principalmente attivati

Gene Ontology: tool for the unification of biology

The Gene Ontology Consortium*

Genomic sequencing has made it clear that a large fraction of the genes specifying the core biological functions are shared by all eukaryotes. Knowledge of the biological role of such shared proteins in one organism can often be transferred to other organisms. The goal of the Gene Ontology Consortium is to produce a dynamic, controlled vocabulary that can be applied to all eukaryotes even as knowledge of gene and protein roles in cells is accumulating and changing. To this end, three independent ontologies accessible on the World-Wide Web (<http://www.geneontology.org>) are being constructed: biological process, molecular function and cellular component.



The goal of the Gene Ontology™ Consortium is to produce a dynamic controlled vocabulary for specific biological domains that can be used to describe gene products in any organism, even as knowledge of gene and protein roles in cells is accumulating and changing.

The GO work includes 3 extensive ontologies describing:

➤ **Molecular Function**

the tasks performed by individual gene products; examples are *transcription factor* and *DNA helicase*

➤ **Biological Process**

broad biological goals, such as *mitosis* or *purine metabolism*, that are accomplished by ordered assemblies of molecular functions

➤ **Cellular Component**

subcellular structures, locations, and macromolecular complexes; examples include *nucleus*, *telomere*, and *origin recognition complex*



GO categories

Molecular function	Biological process	Cellular component
enzyme.F	metabolism.P	extracellular.C
transcription factor.F	cell growth and maintenance.P	plasma membrane.C
nucleic acid binding.F	cell cycle.P	integral plasma membrane protein.C
DNA binding.F	cell organization and biogenesis.P	membrane.C
RNA binding.F	transport.P	nucleus.C
ligand binding or carrier.F	cell motility.P	nuclear membrane.C
calcium binding.F	cell communication.P	nucleolus.C
transporter.F	signal transduction.P	chromosome.C
protein kinase.F	stress response.P	cytoplasm.C
protein phosphatase.F	developmental processes.P	cytosol.C
cell cycle regulator.F	death.P	cytoskeleton.C
cell adhesion molecule.F	apoptosis.P	mitochondrion.C
defense/immunity protein.F	biological_process unknown.P	mitochondrial membrane.C
enzyme inhibitor.F		Golgi apparatus.C
chaperone.F		cytoplasmic vesicle.C
structural protein.F		endoplasmic reticulum.C
cell surface antigen.F		lysosome.C
signal transducer.F		peroxisome.C
receptor.F		cellular_component unknown.C
ligand.F		
enzyme activator.F		



Method

MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data

Scott W Doniger*, Nathan Salomonis*, Kam D Dahlquist*[†], Karen Vranizan*[‡], Steven C Lawlor* and Bruce R Conklin*^{†§}

Addresses: *Gladstone Institute of Cardiovascular Disease, University of California, San Francisco, CA 94141-9100, USA. [†]Cardiovascular Research Institute, and [‡]Departments of Medicine and Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143, USA. [§]Functional Genomics Lab, University of California, Berkeley, CA 94720, USA.

Correspondence: Bruce R Conklin. E-mail: bconklin@gladstone.ucsf.edu

Published: 6 January 2003

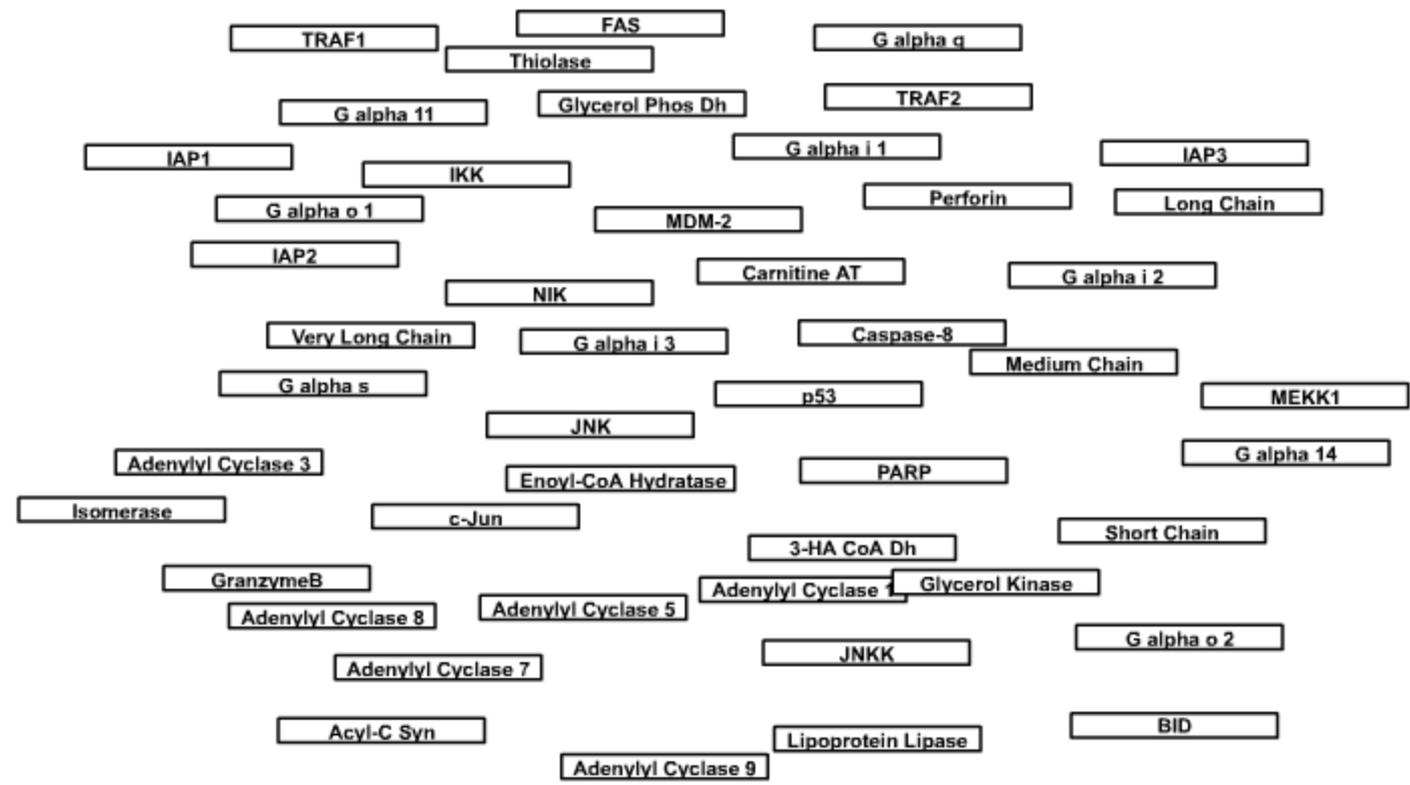
Genome Biology 2003, 4:R7

Received: 11 September 2002

Revised: 8 October 2002

Accepted: 8 November 2002

DNA-arrays randomly list thousands of genes



Genes can be grouped by biological function

- G alpha q
- G alpha 11
- G alpha 14
- G alpha o 1
- G alpha o 2
- G alpha i 1
- G alpha i 2
- G alpha i 3
- G alpha s
- Adenylyl Cyclase 1
- Adenylyl Cyclase 3
- Adenylyl Cyclase 5
- Adenylyl Cyclase 7
- Adenylyl Cyclase 8
- Adenylyl Cyclase 9

- Medium Chain
- Short Chain
- Glycerol Phos Dh
- Long Chain
- Carnitine AT
- Isomerase
- Glycerol Kinase
- Thiolase
- Enoyl-CoA Hydratase
- 3-HA CoA Dh
- Acyl-C Syn
- Very Long Chain
- Lipoprotein Lipase

- GranzymeB
- Perforin
- FAS
- TRAF1
- TRAF2
- IAP1
- IAP2
- IAP3
- NIK
- IKK
- MEKK1
- JNKK
- JNK
- c-Jun
- PARP
- MDM-2
- p53
- BID
- Caspase-8



GenMAPP is a tool that allows for organization of these groupings

G protein Signaling

- G alpha q
- G alpha 11
- G alpha 14
- G alpha o 1
- G alpha o 2
- G alpha i 1
- G alpha i 2
- G alpha i 3
- G alpha s
- Adenylyl Cyclase 1
- Adenylyl Cyclase 3
- Adenylyl Cyclase 5
- Adenylyl Cyclase 7
- Adenylyl Cyclase 8
- Adenylyl Cyclase 9

Fatty Acid Degradation

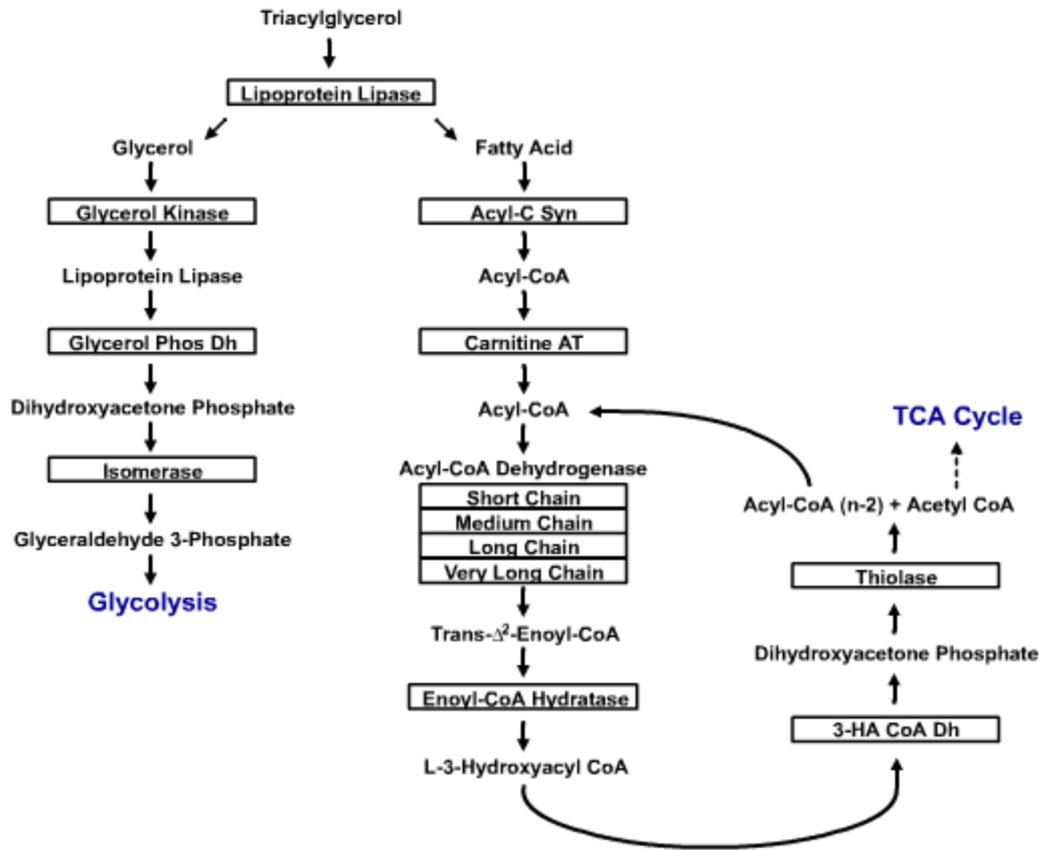
- Medium Chain
- Short Chain
- Glycerol Phos Dh
- Long Chain
- Carnitine AT
- Isomerase
- Glycerol Kinase
- Thiolase
- Enoyl-CoA Hydratase
- 3-HA CoA Dh
- Acyl-C Syn
- Very Long Chain
- Lipoprotein Lipase

Apoptosis

- GranzymeB
- Perforin
- FAS
- TRAF1
- TRAF2
- IAP1
- IAP2
- IAP3
- NIK
- IKK
- MEKK1
- JNKK
- JNK
- c-Jun
- PARP
- MDM-2
- p53
- BID
- Caspase-8

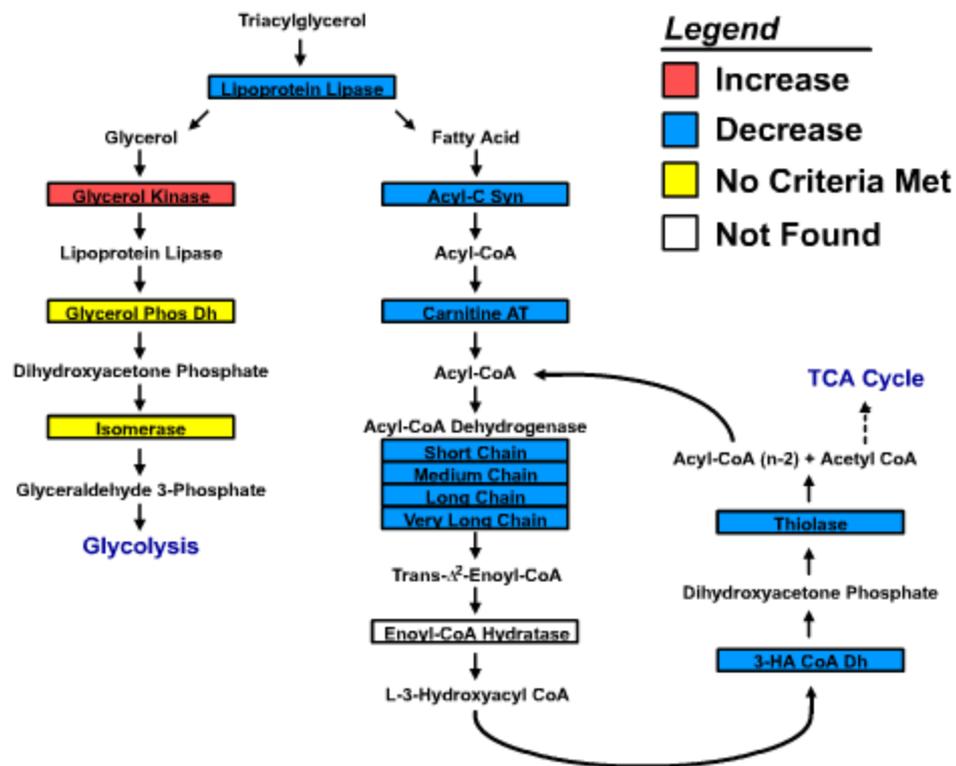


Fatty Acid Degradation



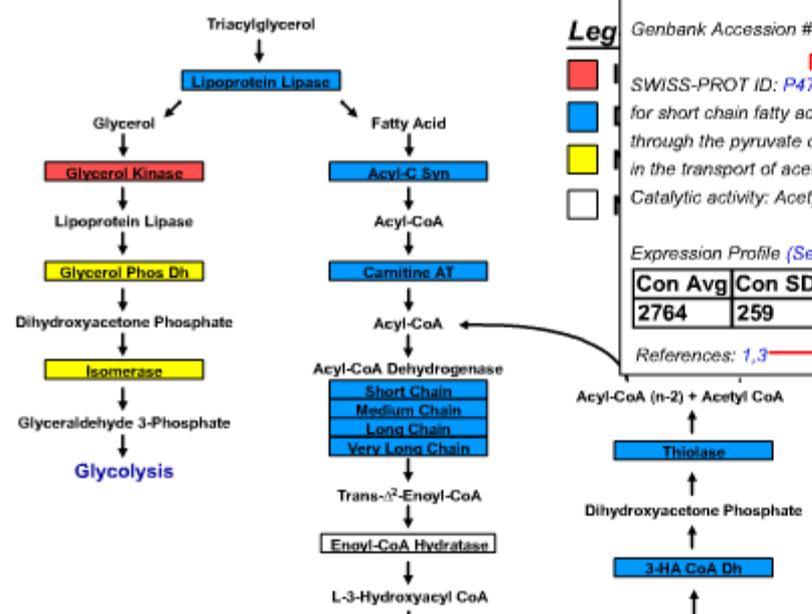
GenMAPP can be used to visualize DNA-array data

Fatty Acid Degradation



How GenMAPP works

Fatty Acid Degradation



Carnitine Acetyltransferase
 Blast Definition: *M.musculus* mRNA for carnitine acetyltransferase
 Genbank Accession #: [X85983](#)
 SWISS-PROT ID: [P47934](#) (note: function: carnitine acetylase is specific for short chain fatty acids. Carnitine acetylase seems to affect the flux through the pyruvate dehydrogenase complex. It may be involved as well in the transport of acetyl-coa into mitochondria.
 Catalytic activity: Acetyl-coA + carnitine = coA + O-acetylcarnitine.)
 Expression Profile (See above description):

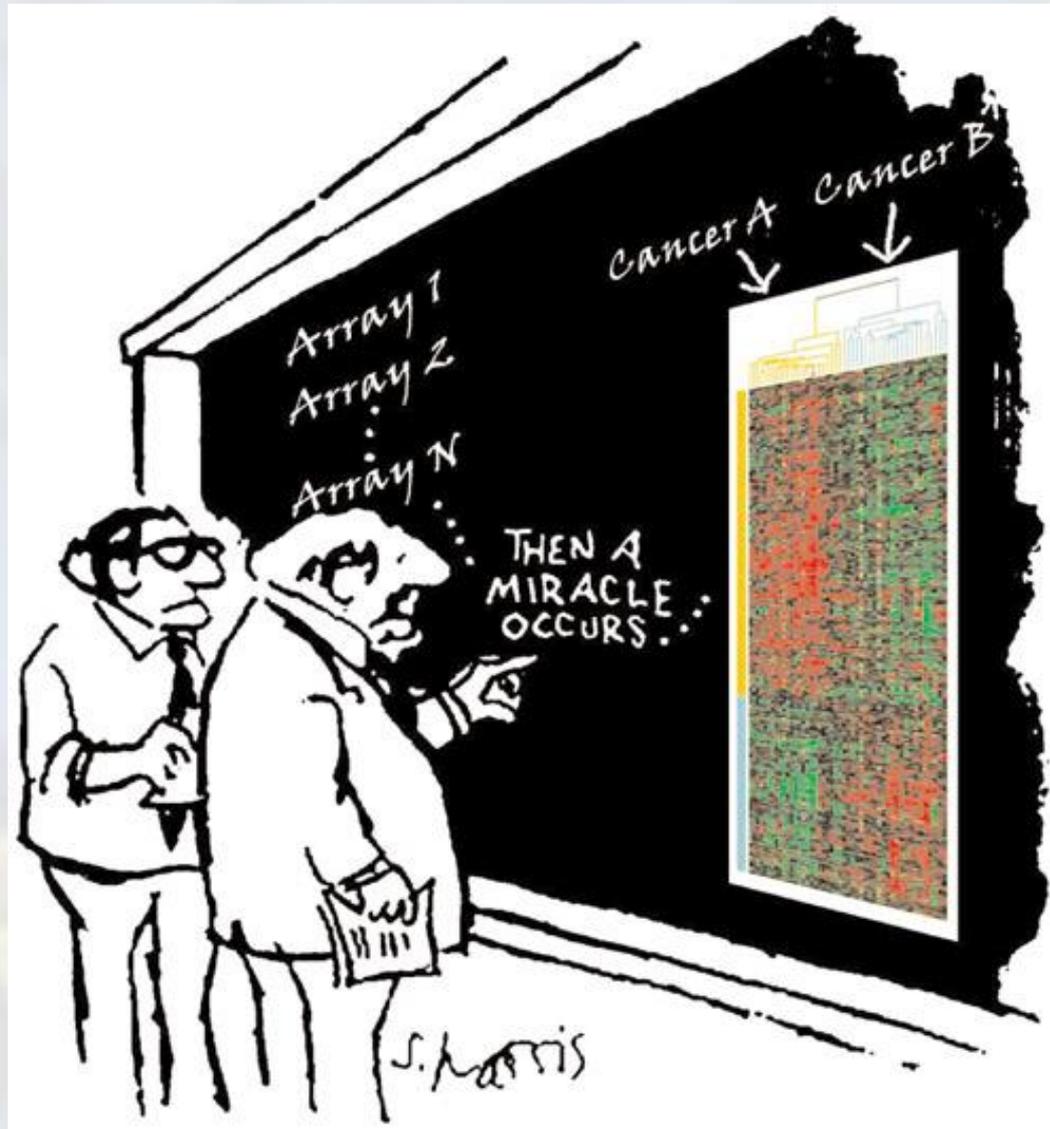
Con Avg	Con SD	8wks Avg	8wks SD	p-value	fold	lower ci	upper ci
2764	259	1611	288	0.0000	0.58	0.5050	0.6725

References: [1,3](#)

**Hyperlinks to
 Swiss-Prot,
 GenBANK ENTREZ,
 and PubMed references**

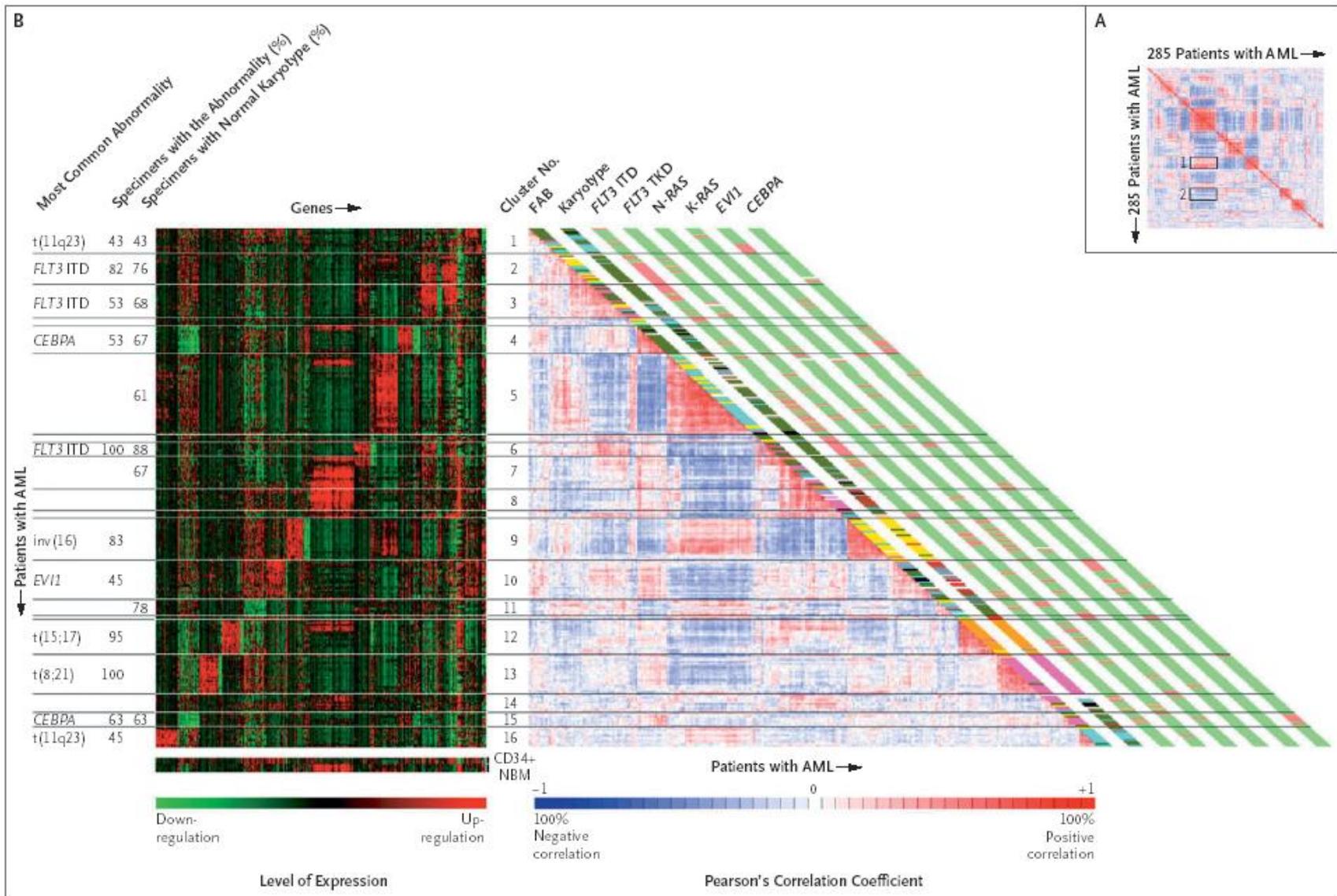


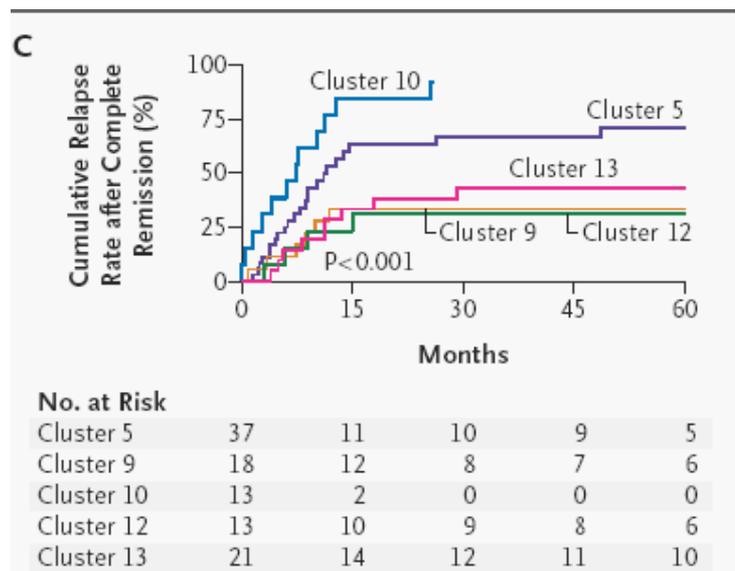
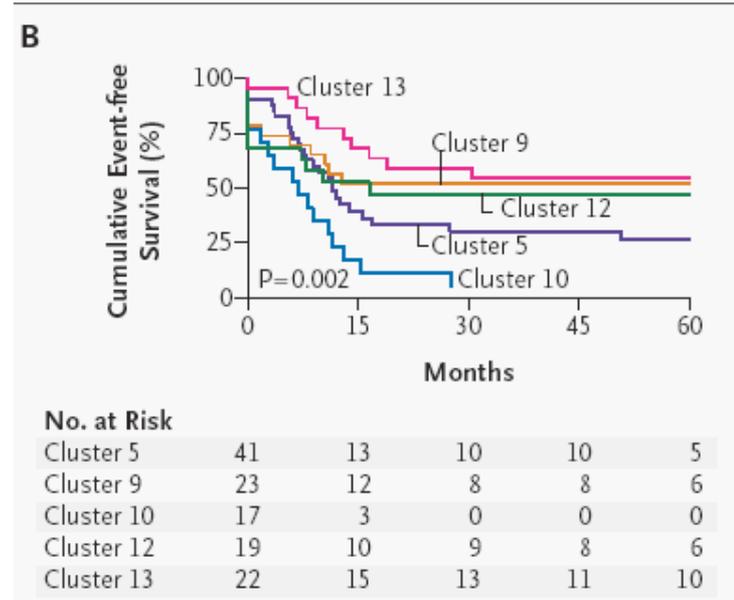
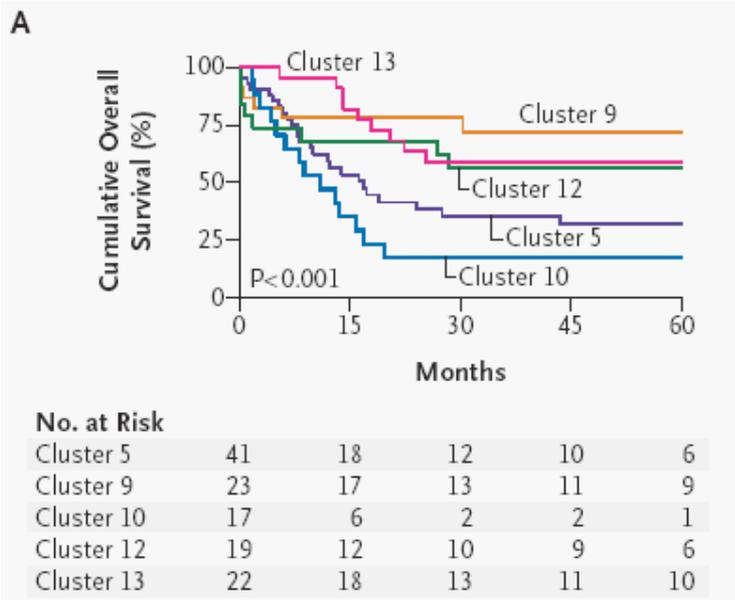
Analisi del trascrittoma



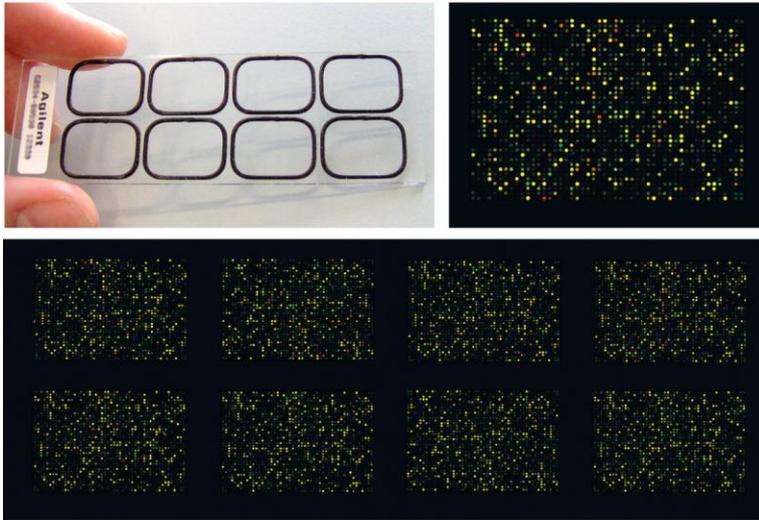
Prognostically useful gene-expression profiles in acute myeloid leukemia.

Valk P., Verhaak G.W., Beijen A., Erpenlinck C.A.J., et al.
New Engl. J. Med. 350: 1617-1628, 2004





Applicazioni “commerciali”



**APPROVATO
DALL'FDA 2007**
U.S. Food and Drug Administration
Rockville MD 20857-0001



ORIGINAL ARTICLE

Gene expression signature of primary imatinib-resistant chronic myeloid leukemia patients

O Frank¹, B Brors², A Fabarius¹, L Li³, M Haak³, S Merk⁴, U Schwindel¹, C Zheng¹, MC Müller¹, N Gretz³, R Hehlmann¹, A Hochhaus¹ and W Seifarth¹

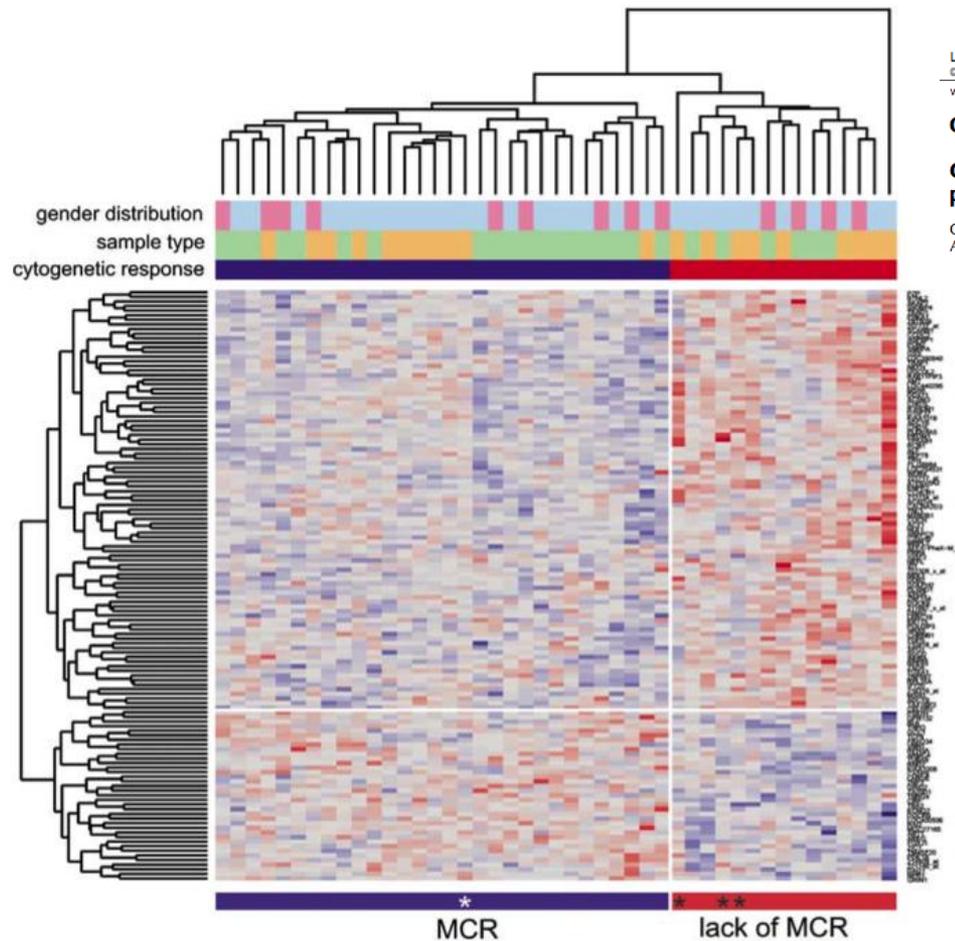
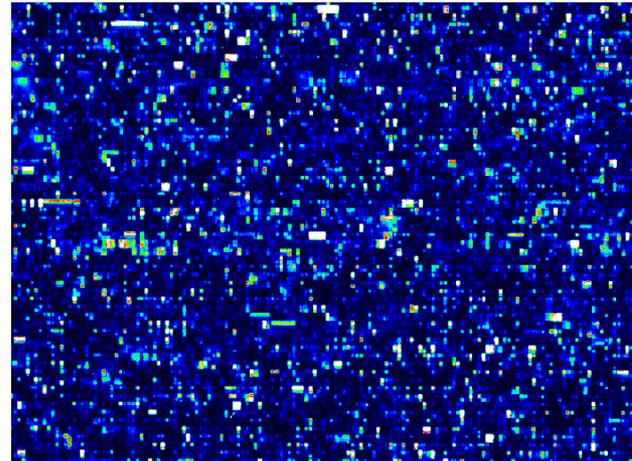
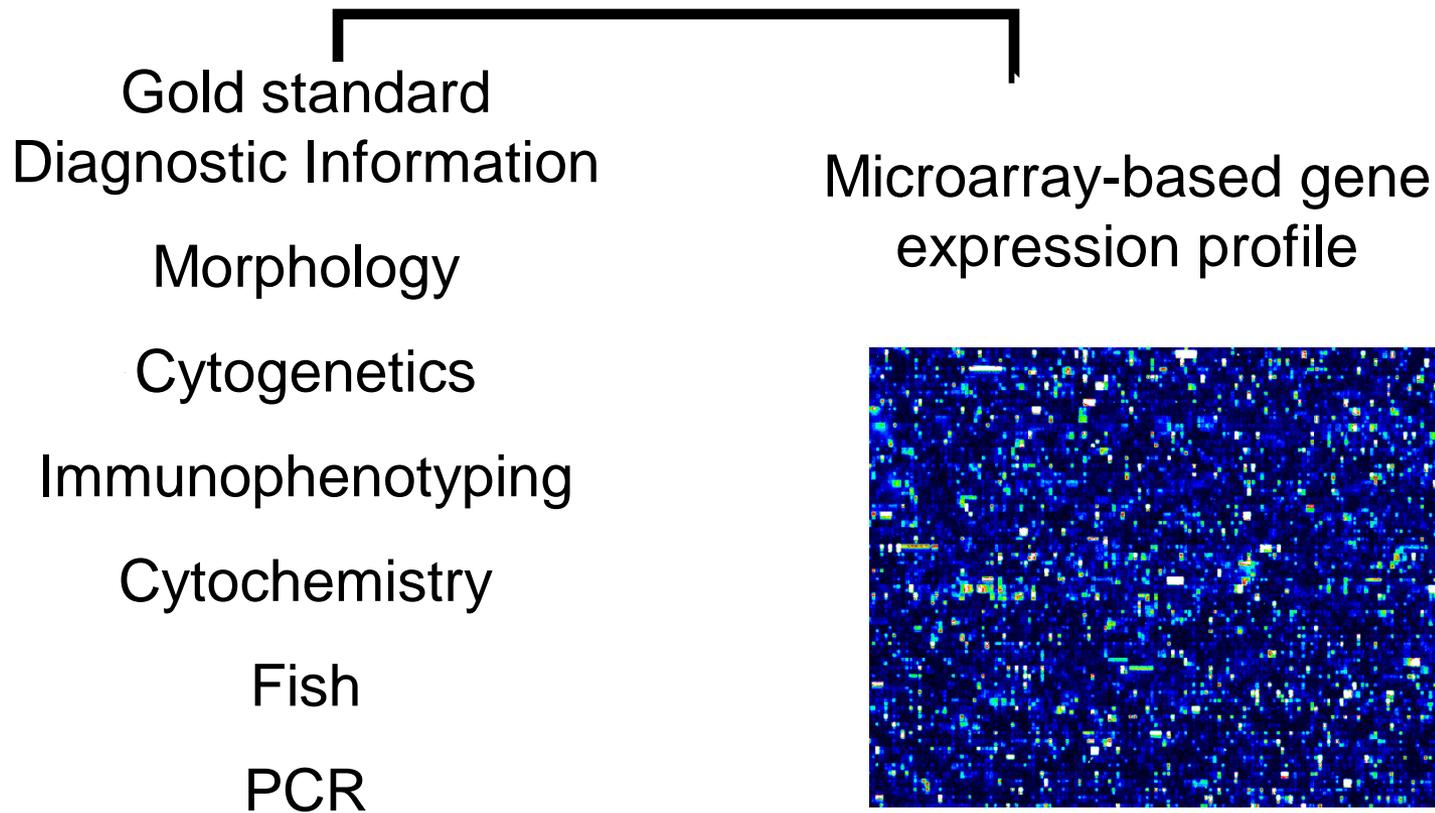


Figure 1 Outcome-specific gene expression signature in CML. Hierarchical cluster analysis of CML samples of 36 patients with different cytogenetic responses to imatinib treatment within the first 12 months of therapy. The heatmap combines training set ($n=41$) and test set samples ($n=4$), the latter being marked by asterisks. Using SVM-RFE, 128 genes that were required to discriminate between both diagnostic patient groups according to their cytogenetic response profile were identified. Heatmaps were constructed using z-transformed gene expression values, that is, measurements were scaled to have gene-wise zero mean and gene-wise unit variance. Columns (samples) and rows (probe sets) of the gene expression matrix were reordered by hierarchical clustering using an Euclidean distance metric and the complete linkage algorithm. Colored bars above the heatmap illustrate (i) gender distribution (pink=female, light blue=male), (ii) sample type (green=PB; orange=BM) and (iii) cytogenetic response (blue=MCR, red=lack of MCR).

Clinical accuracy of the microarray test as compared to standard leukemia laboratory methods (“gold standard”)

n = 4000 patients

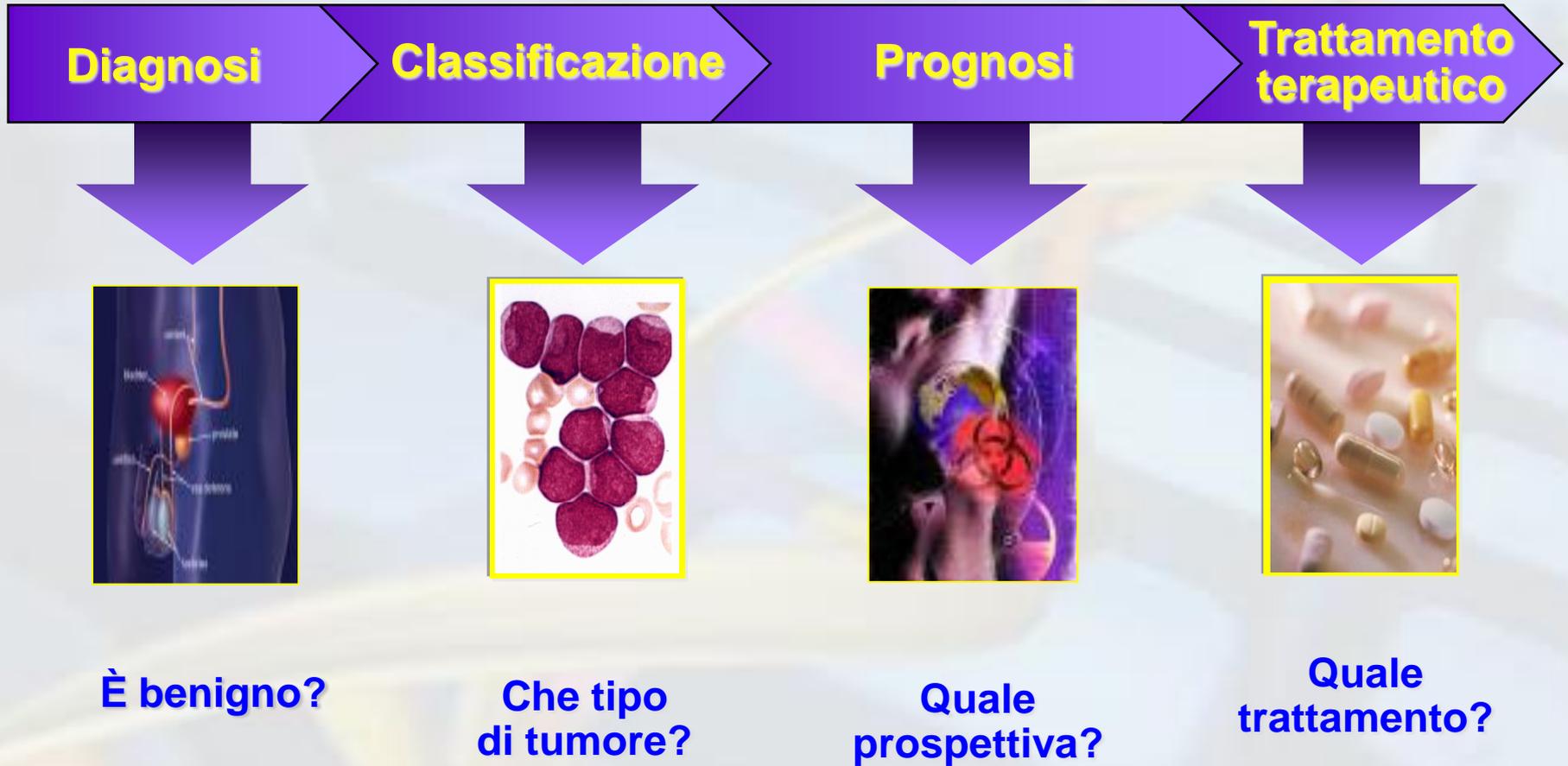


Classification Scheme (Proposal)

21-class scheme

- ✓ ALL with t(8;14)
- ✓ ALL with t(11q23)/MLL
- ✓ ALL with t(9;22)
- ✓ ALL with t(12;21)
- ✓ ALL with t(1;19)
- ✓ ALL hyperdiploid
- ✓ ALL other/normal
- ✓ T-ALL immature
- ✓ T-ALL cortical
- ✓ AML with t(8;21)
- ✓ AML with t(15;17)
- ✓ AML with inv(16)
- ✓ AML (11q23)/MLL
- ✓ AML with complex aberrant karyotype
- ✓ AML other/normal
- ✓ CML
- ✓ CLL with del(13q)
- ✓ CLL with del(17p)
- ✓ CLL other/normal
- ✓ MDS
- ✓ Non-leukemia

Analisi del trascrittoma



Into the Mines: extracting meaning from
microarrays

DATA are not
KNOWLEDGE, and
knowledge is not
UNDERSTANDING

John Quackenbush

Laboratorio di proteomica (DSB) Spettrometro di massa (CIGS)

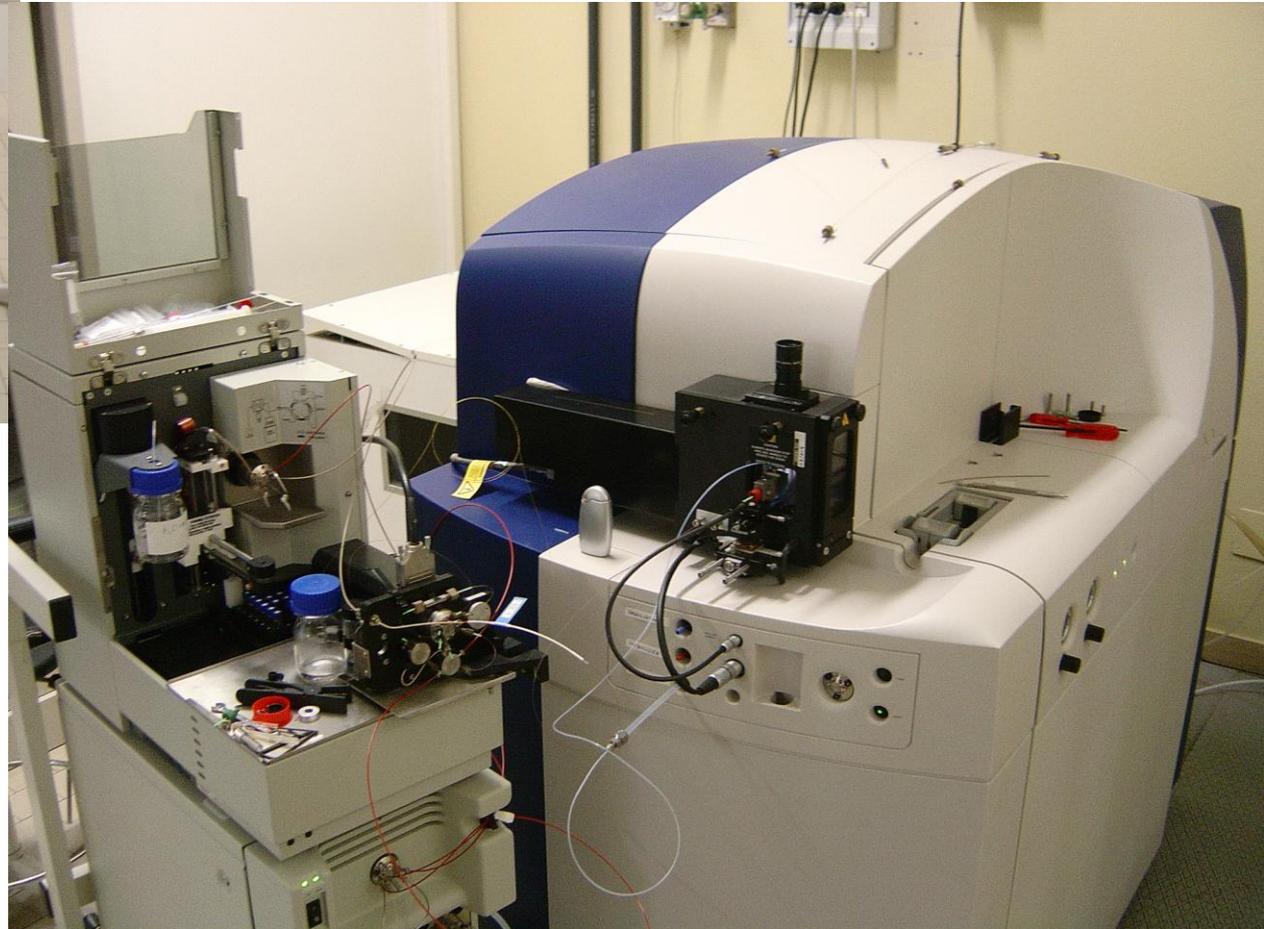
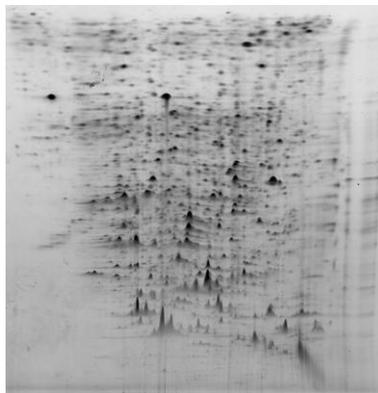


Figura 12

Proteomica



1. Estrazione delle proteine cellulari



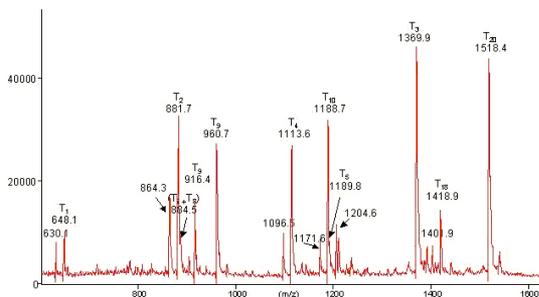
2. Separazione su gel bidimensionale



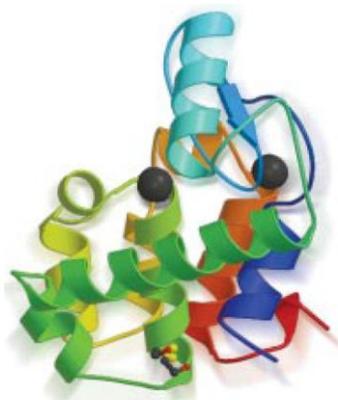
3. Prelievo degli spot e digestione enzimatica



4. Analisi con lo spettrometro di massa

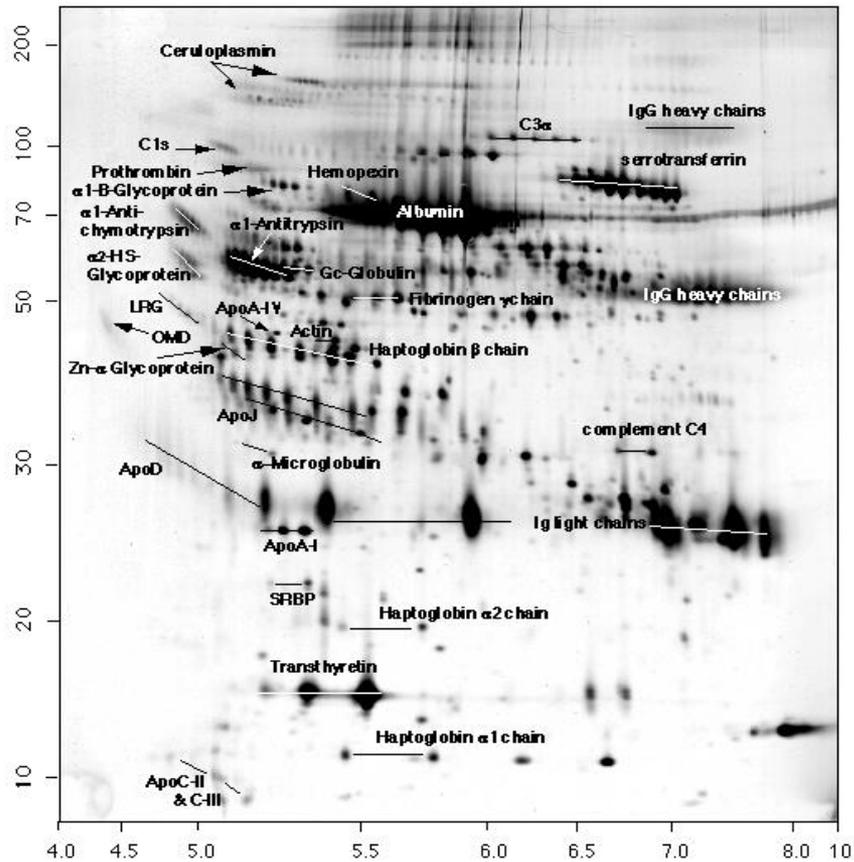


5. Confronto dello spettro con i database disponibili

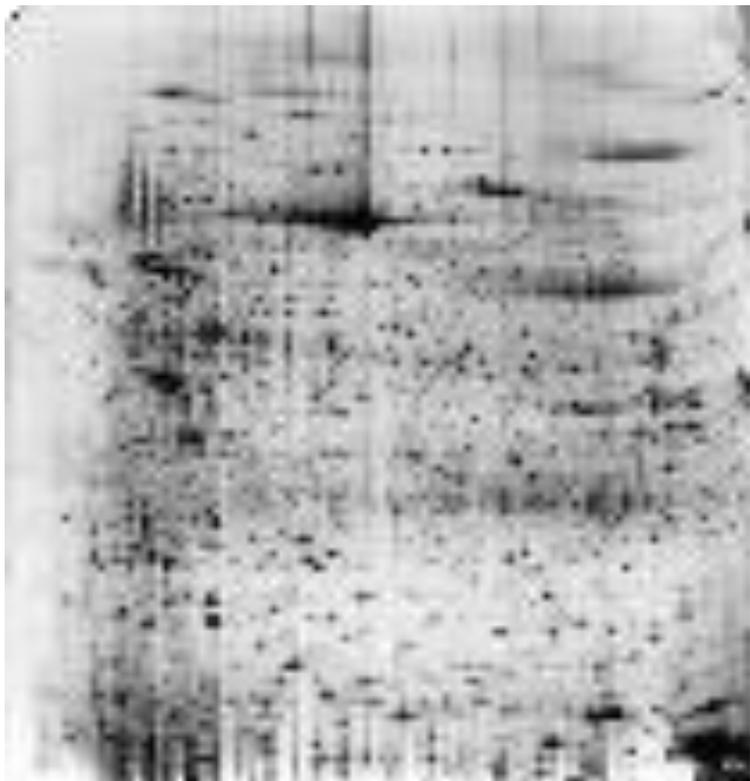


6. Identificazione della proteina

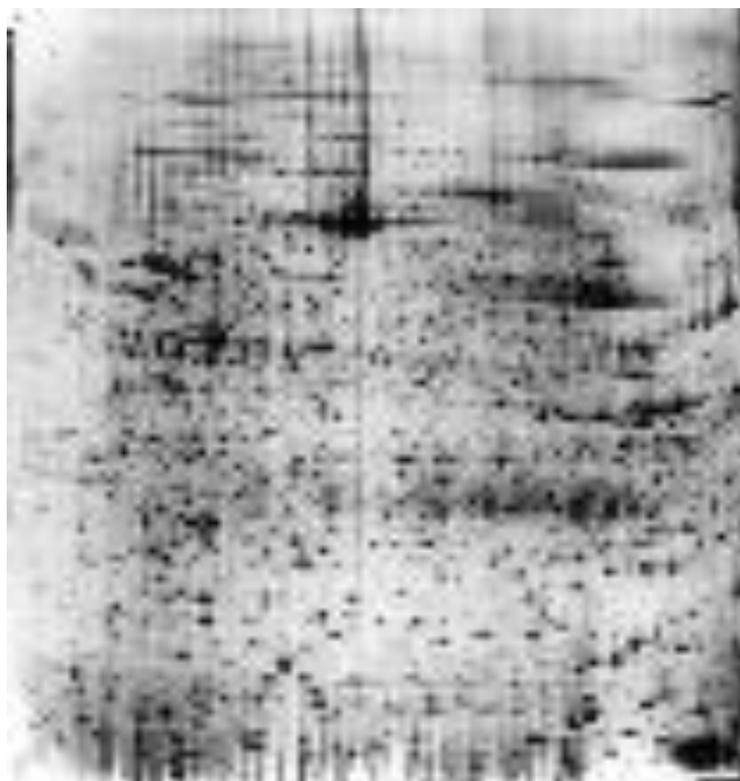
IL GEL BIDIMENSIONALE CI DA' L'IMPRONTA DIGITALE DI UNA CELLULA



CELLULA NORMALE

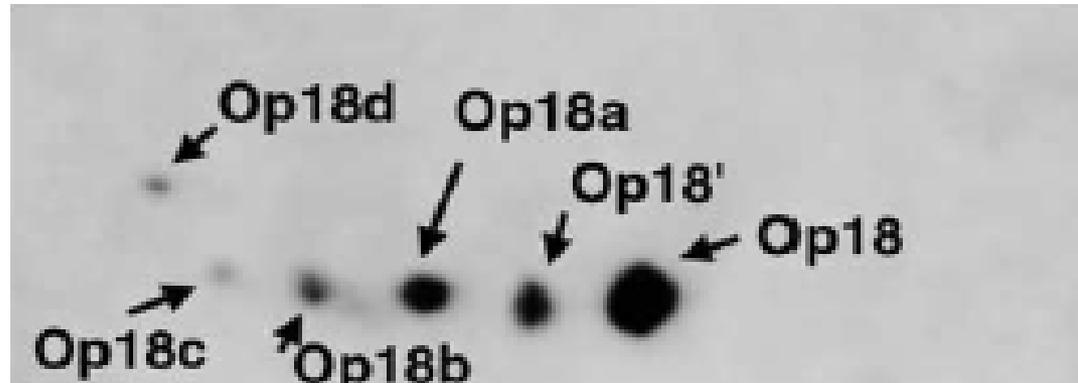


CELLULA TUMORALE

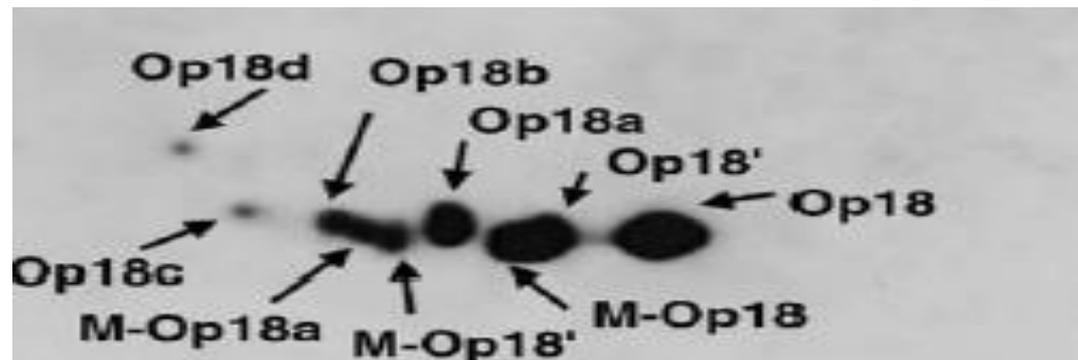


IDENTIFICAZIONE DI MARCATORI TUMORALI

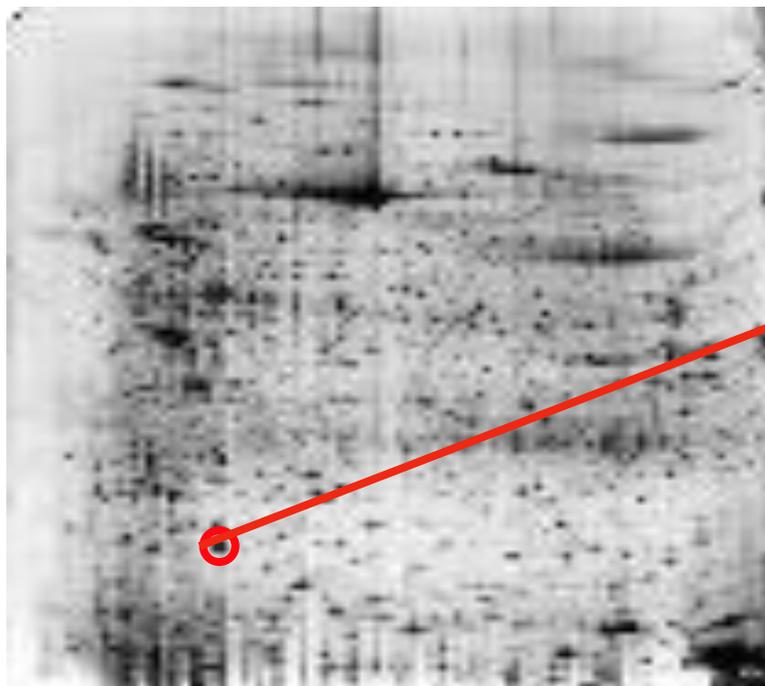
CELLULA NORMALE



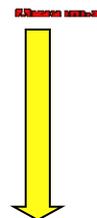
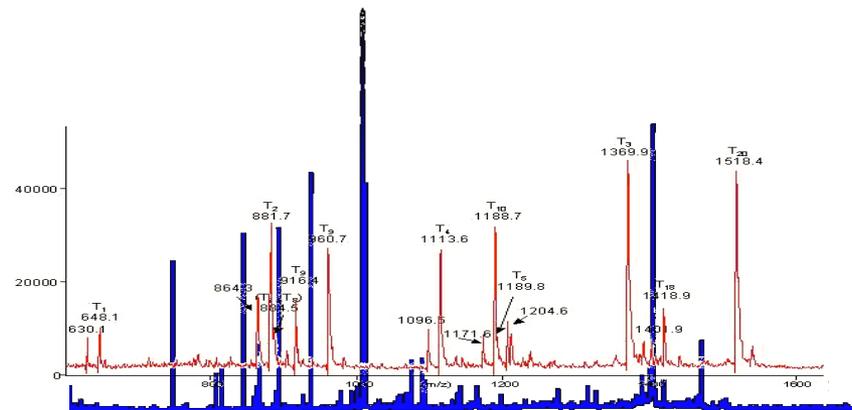
CELLULA TUMORALE



GEL BIDIMENSIONALE



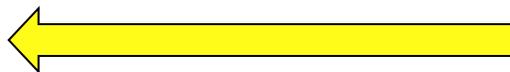
SPETTROMETRIA



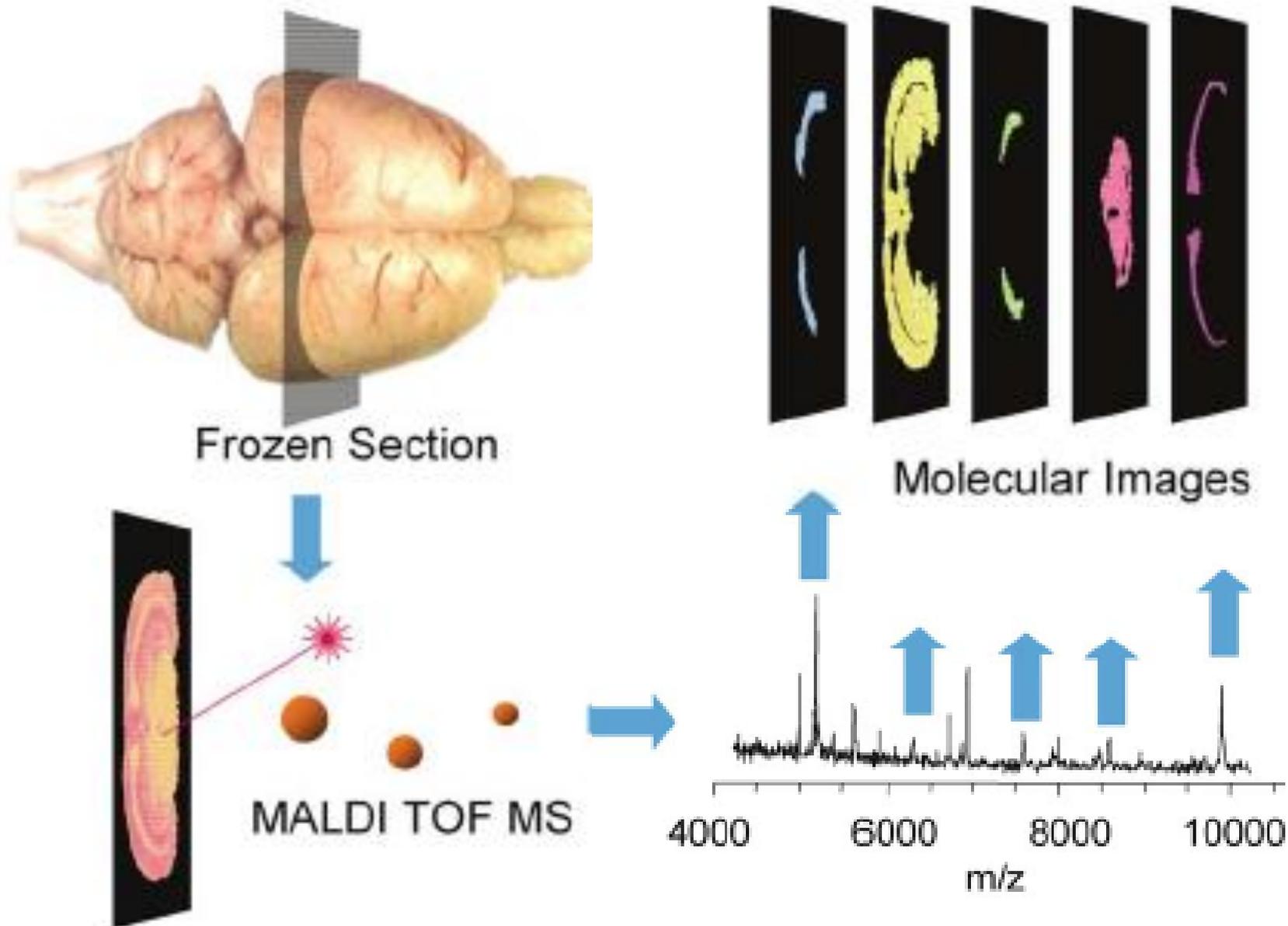
MAPPA PEPTIDICA



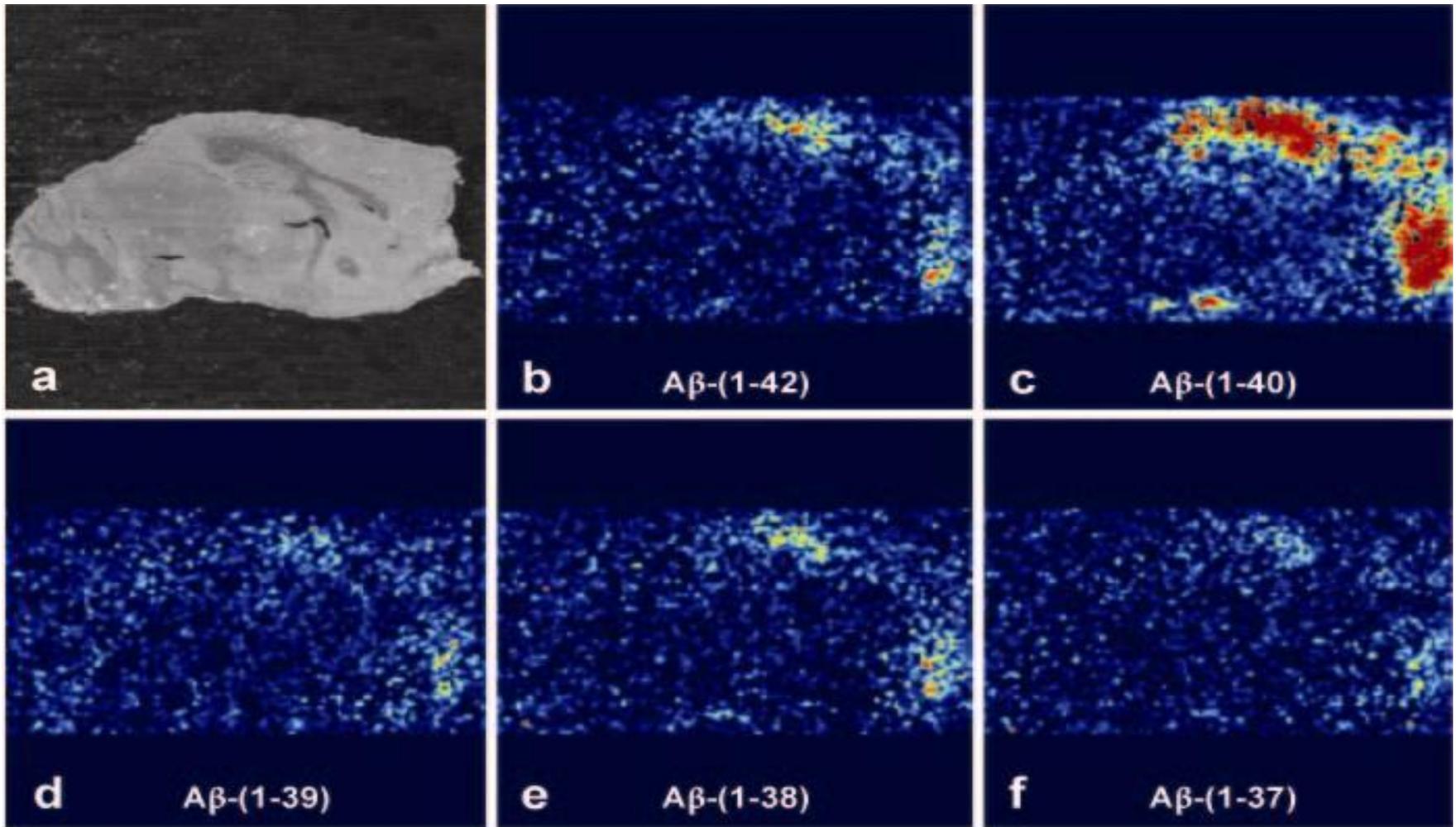
IDENTIFICAZIONE
DELLA PROTEINA



Mass spectrometer as a molecular microscope

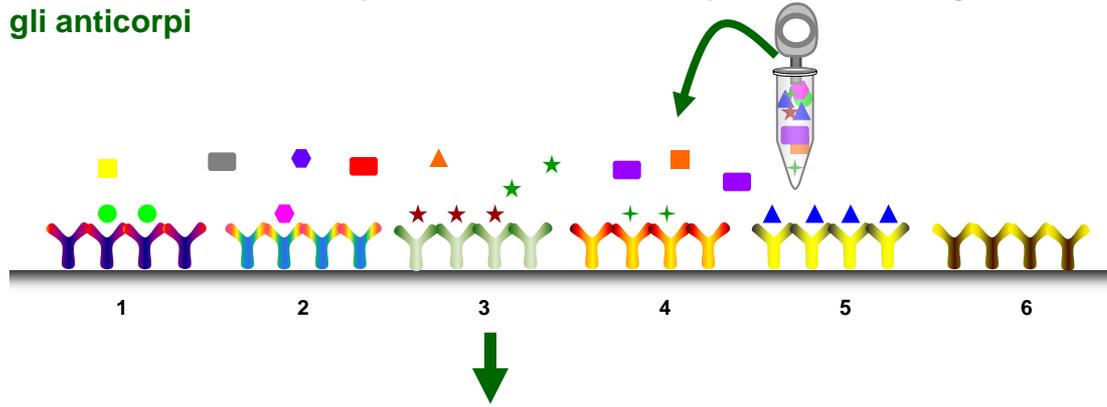


IMS images from a diseased brain section from an Alzheimer's patient.

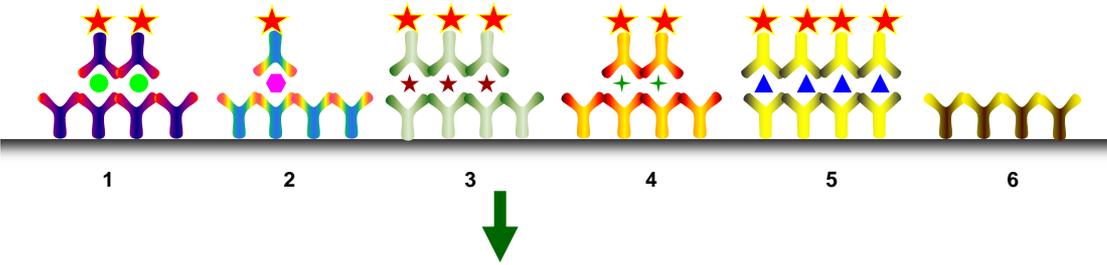


(a) Optical image of the sagittal diseased brain section. (b) $A\beta$ -(1-42) molecular image (m/z 4515.1). (c) $A\beta$ -(1-40) (m/z 4330.9). (d) $A\beta$ -(1-39) (m/z 4231.7). (e) $A\beta$ -(1-38) (m/z 4132.6). (f) $A\beta$ -(1-37) (m/z 4075.5).

1. Incubazione del campione solubile sulla superficie dell'array sulla cui superficie sono immobilizzati gli anticorpi



2. Addizione di una miscela solubile di anticorpi marcati diretti contro lo stesso pannello antigenico



3. Incubazione, sviluppo e quantificazione del segnale

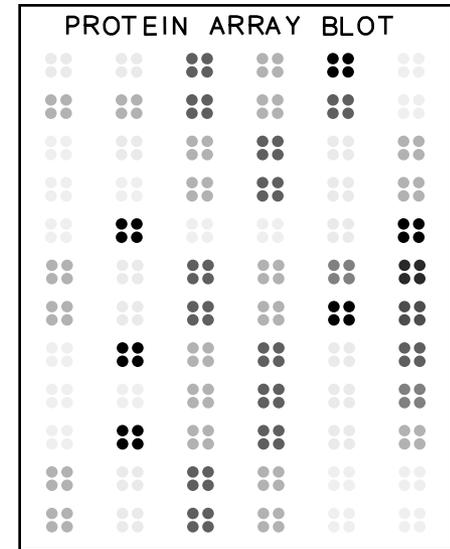


Figura 13



Map of Network Summary View

Node Name: MAF
([Neighborhood Explorer](#))

LL Description

v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)

Synonyms

2810401A20Rik, C-MAF, Maf2

Source Id

209347_s_at, 209348_s_at, 206363_at

Protein Family, Domain

basic domain, bZIP domain, DNA binding, leucine zipper domain, protein binding, RNA polymerase II transcription factor, transcription regulator

Subcellular Location

Edge Summary:

[MAF/EP300](#)

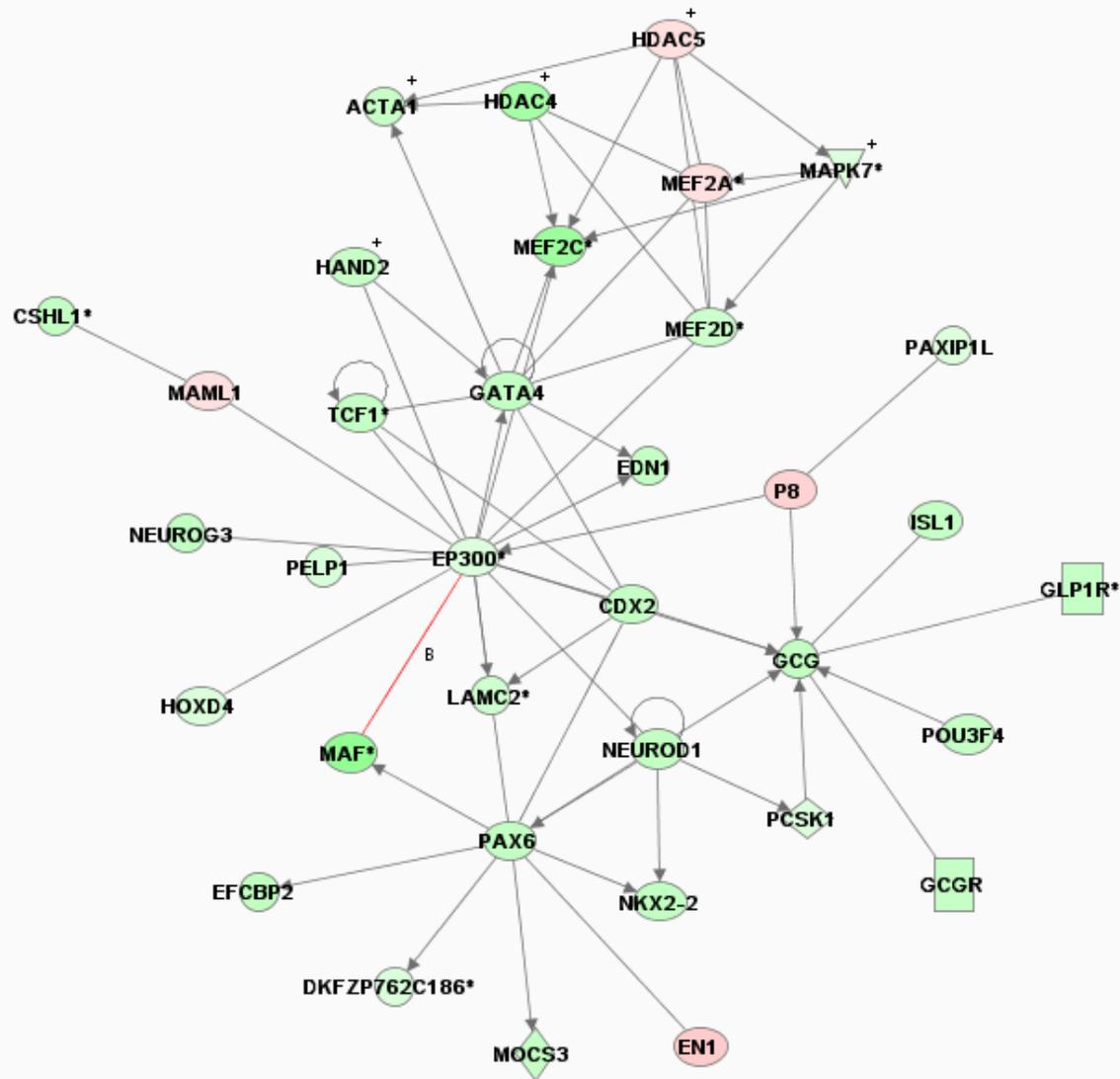
Physical Interactions [3]

Binding Events [3]

Binding of human [p300](#) [[EP300](#)] protein and [C-MAF](#) [[MAF](#)] protein occurs in nuclei of Cos-1 cells.

Binding of mouse [P300](#) [[Ep300](#)] protein and mouse [C-maf](#) [[Maf](#)] protein occurs in lysate of mouse lens.

Binding of human [p300](#) [[EP300](#)] protein and



CELL LAB



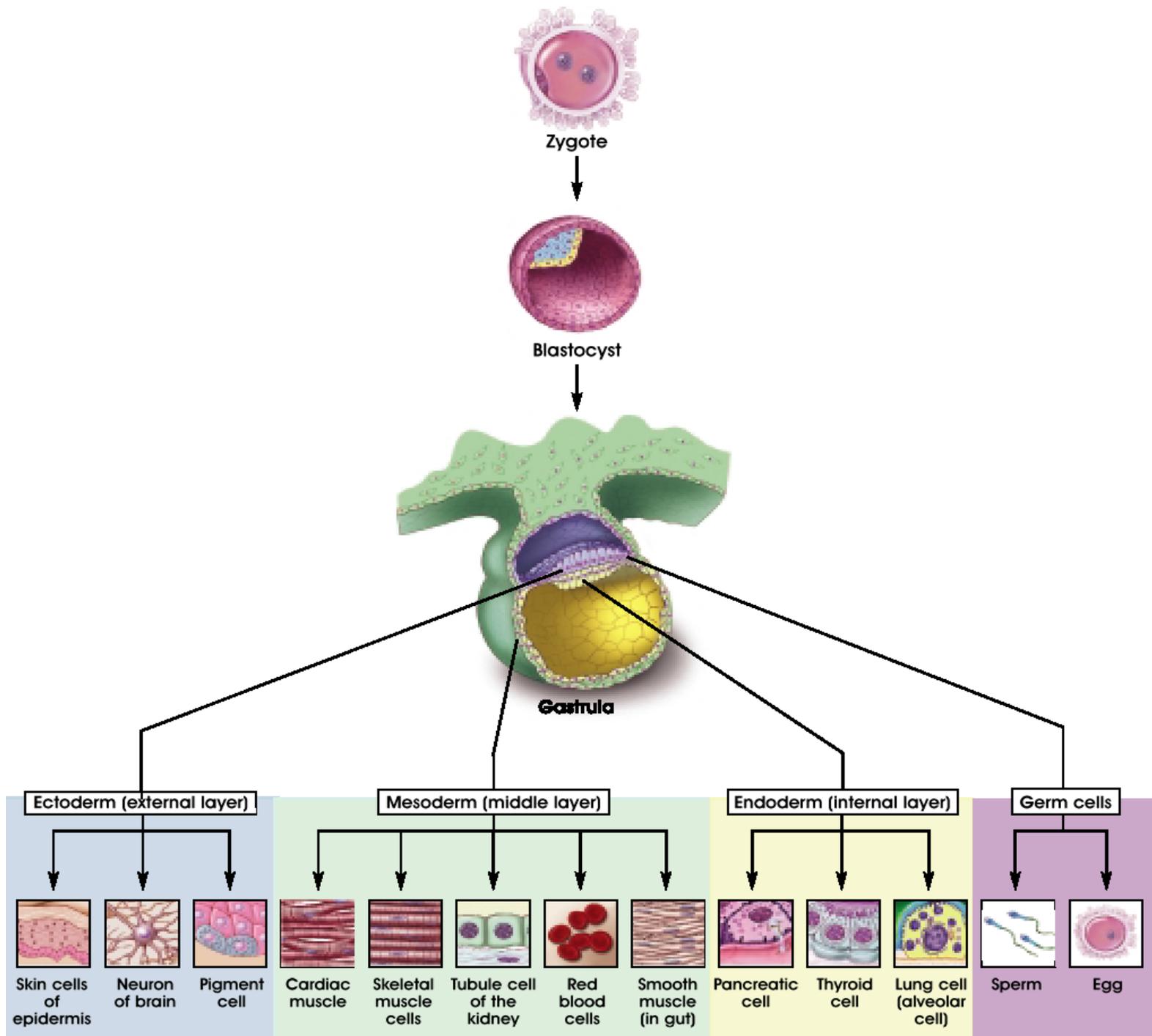




Figura 1.22 Il topo come modello dello sviluppo dell'uomo

Un bambino e un topo mostrano gli stessi difetti di pigmentazione (piebaldismo) come risultato delle mutazioni avvenute a carico di un gene necessario per la normale migrazione dei melanociti (le cellule responsabili della pigmentazione della pelle) durante lo sviluppo embrionale. (Cortesemente fornite da R.A. Fleischman, Markey Cancer Center, University of Kentucky).

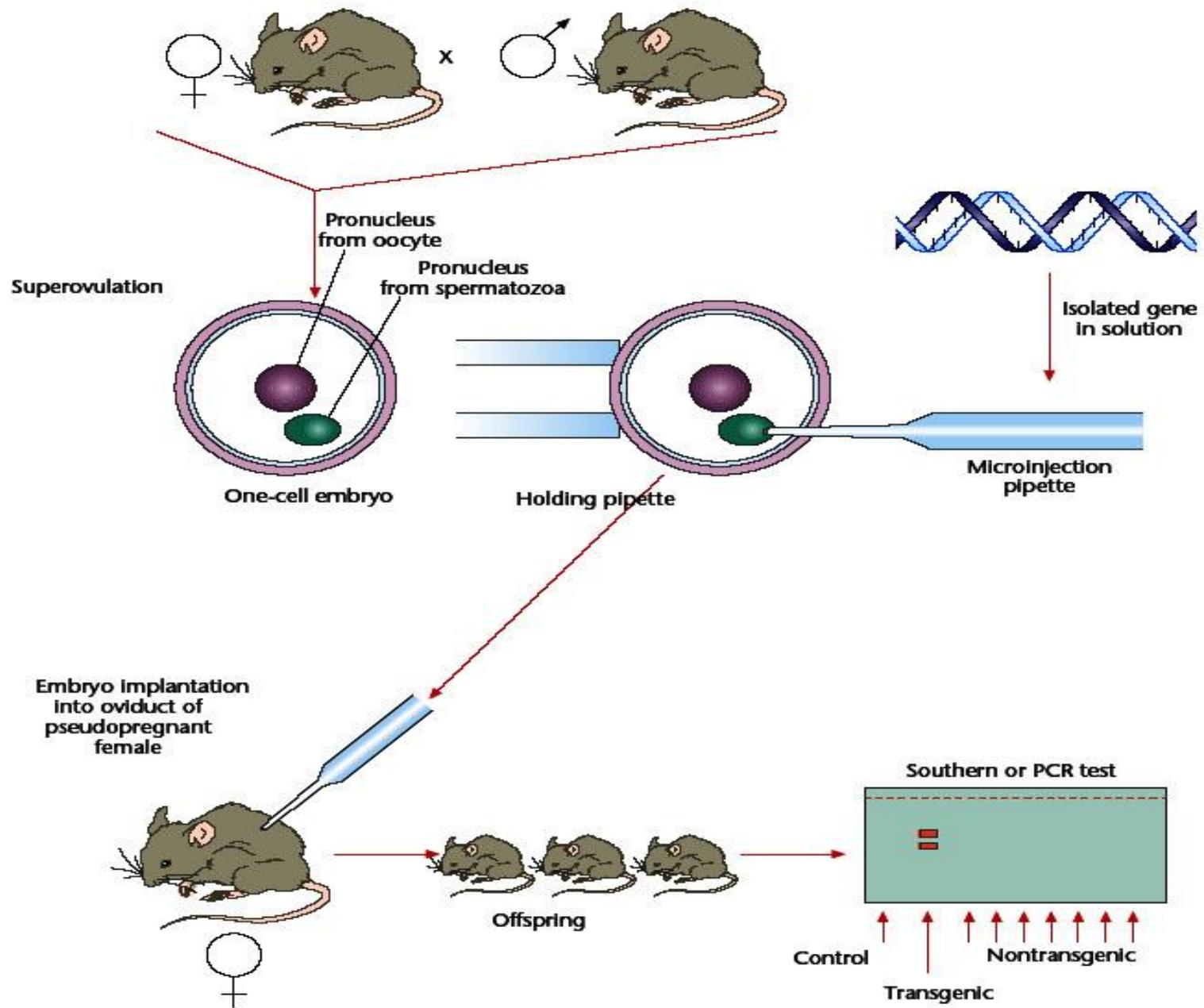
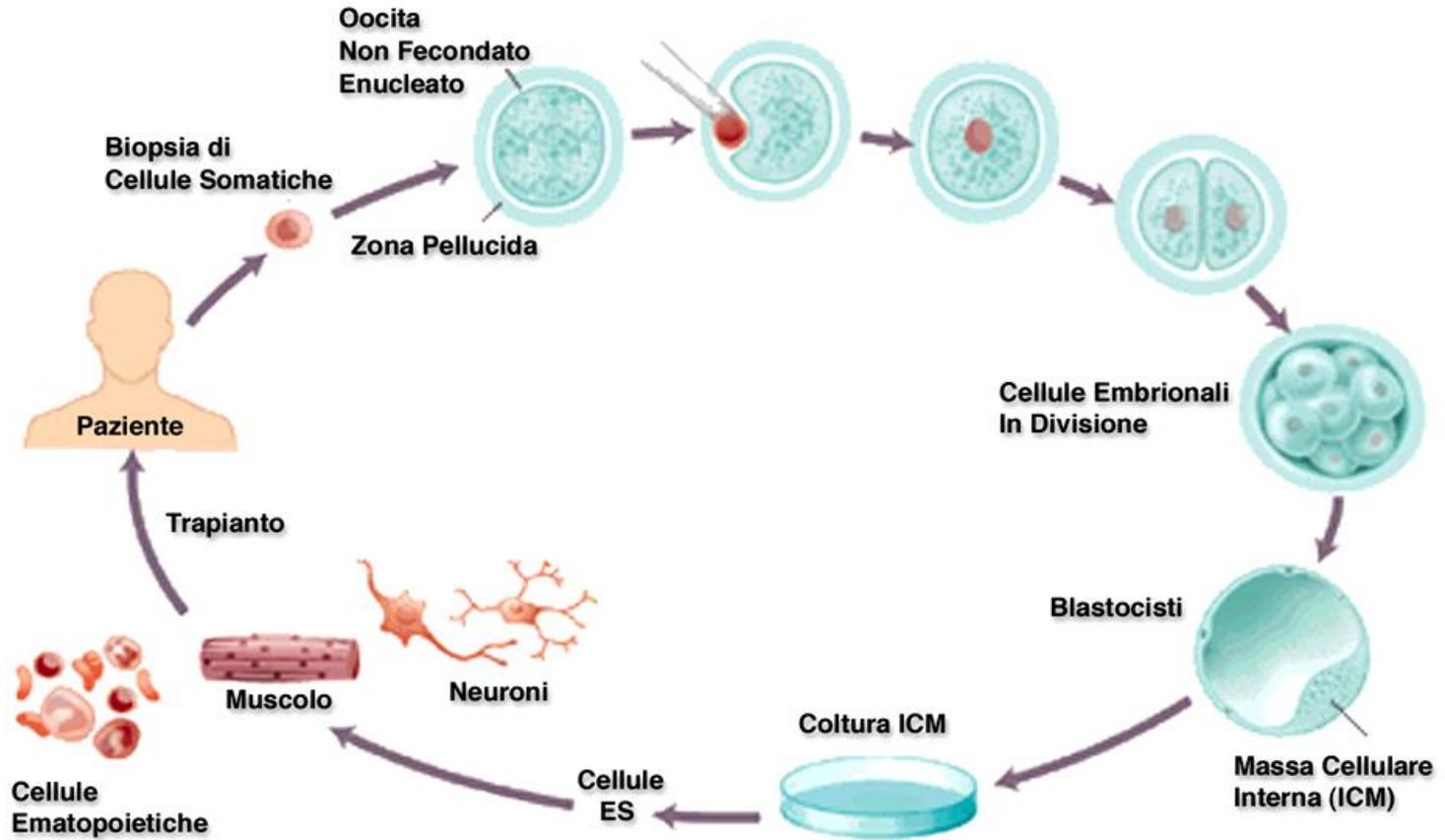


Figure 1 The generation of transgenic animals by gene microinjection. The embryos obtained by superovulation or by *in vitro* fertilization receive the foreign genes and are developed in foster mothers. Transgenes are detected and transmitted to progeny by normal reproduction. PCR, polymerase chain reaction.

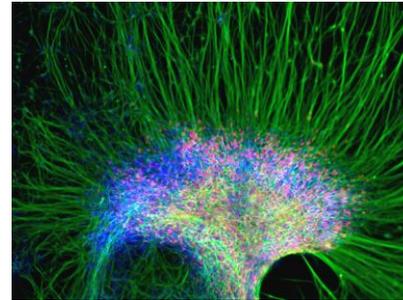
Trasferimento Nucleare da Cellule Somatiche



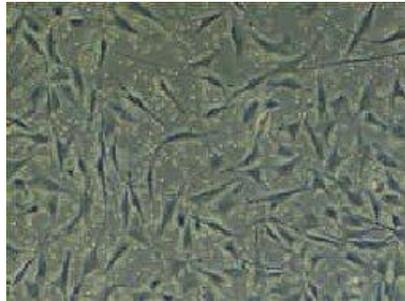
Fenotipizzazione di diverse cellule staminali



Hematopoietic
stem cells

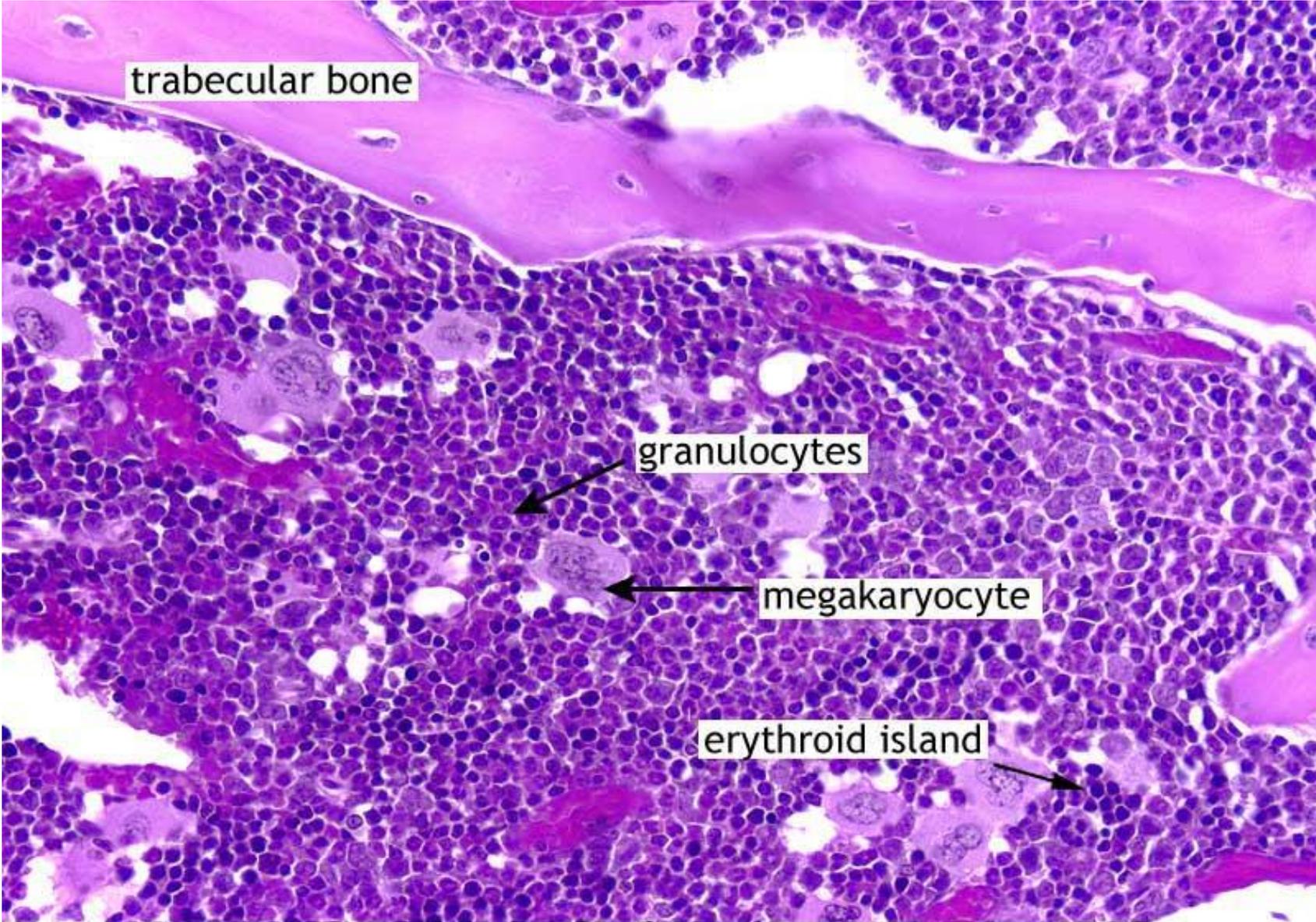


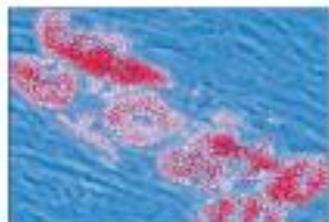
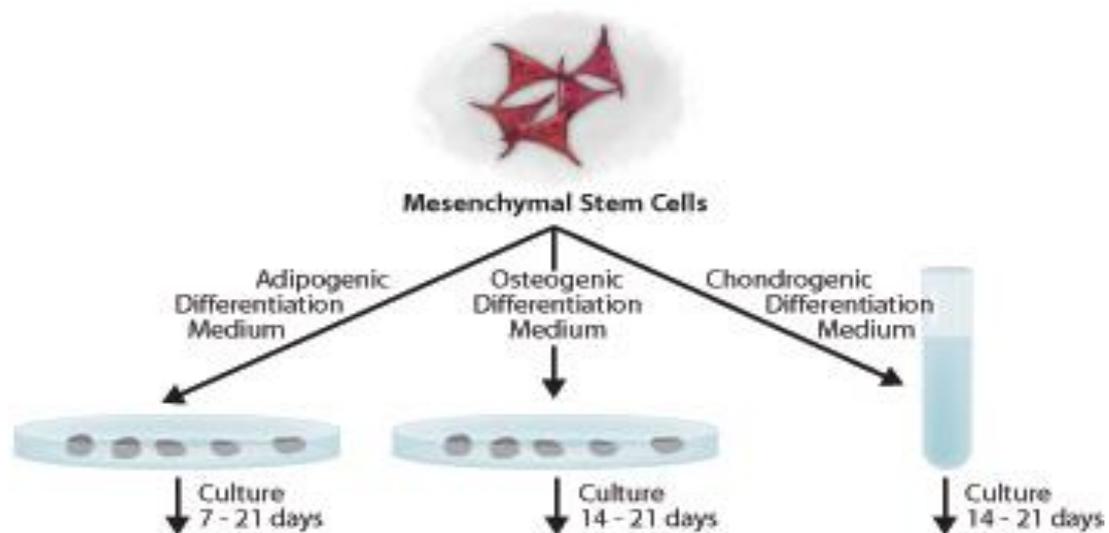
Neural stem cells



Mesenchymal stem cells

Midollo osseo



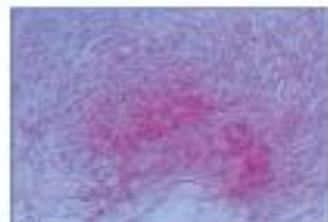


A

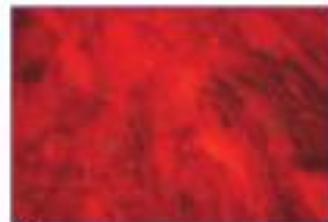


B

Adipocytes derived from MSCs differentiated *in vitro* with Adipogenic Supplement and identified by (A) Oil Red staining and (B) using R&D Systems' goat anti-mouse FABP-4 affinity-purified polyclonal antibody.



A

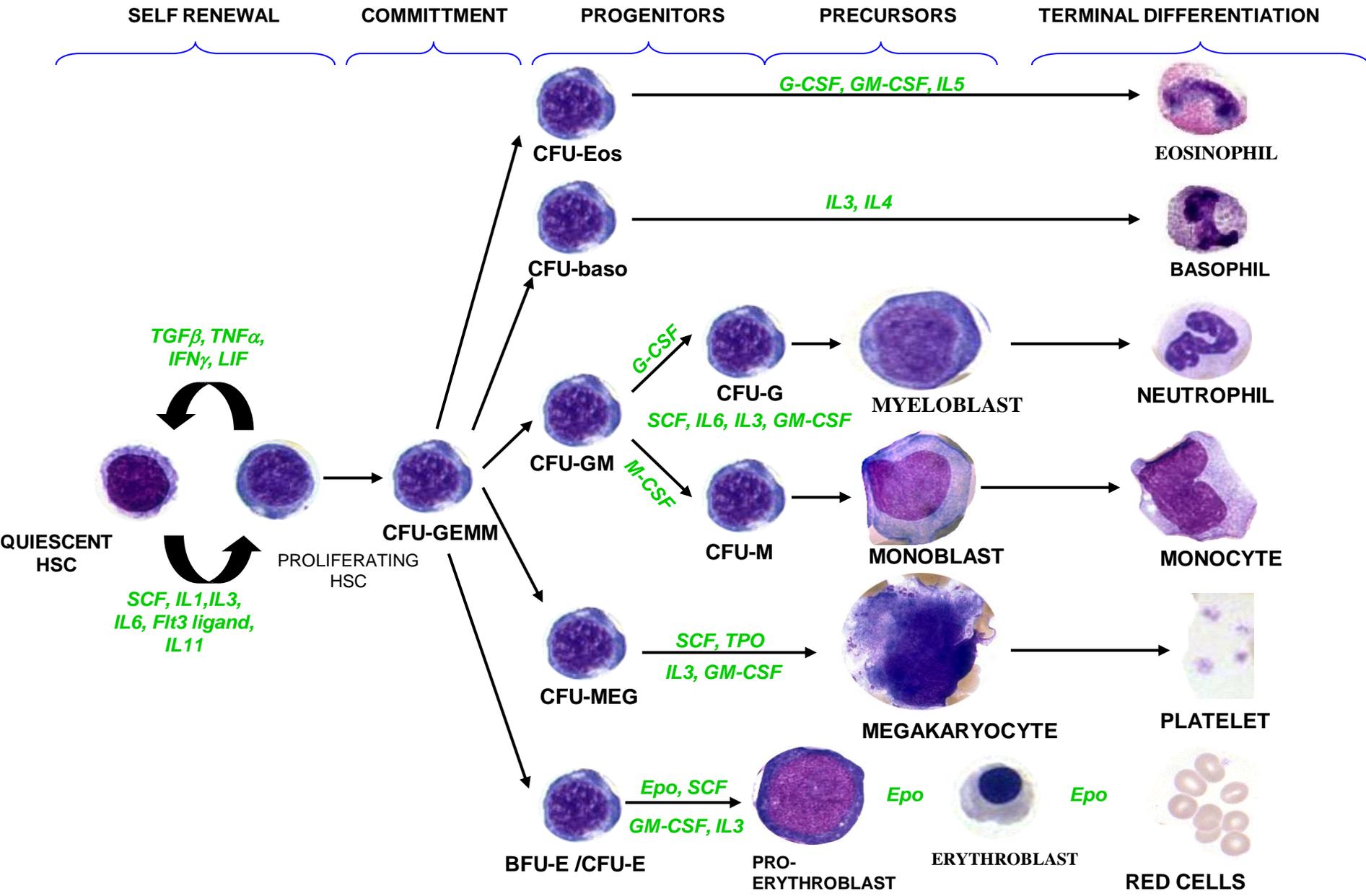


B

Osteocytes derived from MSCs differentiated *in vitro* with Osteogenic supplement and identified by (A) Alzarlan red staining and (B) using R&D Systems' mouse anti-human Osteocalcin monoclonal antibody.

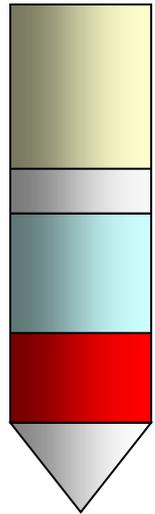


Chondrocytes derived from MSCs differentiated *in vitro* with chondrogenic Supplement and identified using R&D Systems' goat anti-human Aggrecan affinity-purified polyclonal antibody.



NORMAL MYELOPIESIS





→ plasma

→ mononucleate

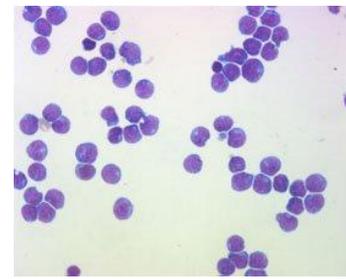
→ ficoll

→ globuli rossi

→ granulociti

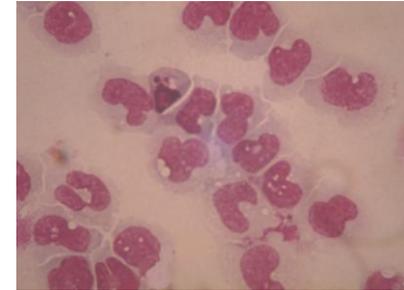
α CD34

HSC



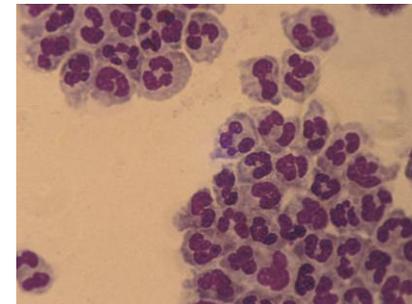
α CD14

monociti

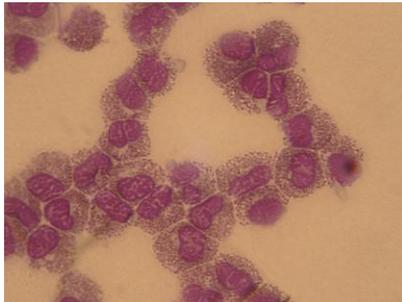


α CD16

neutrofili (CD16+)



Eosinofili (CD16-)



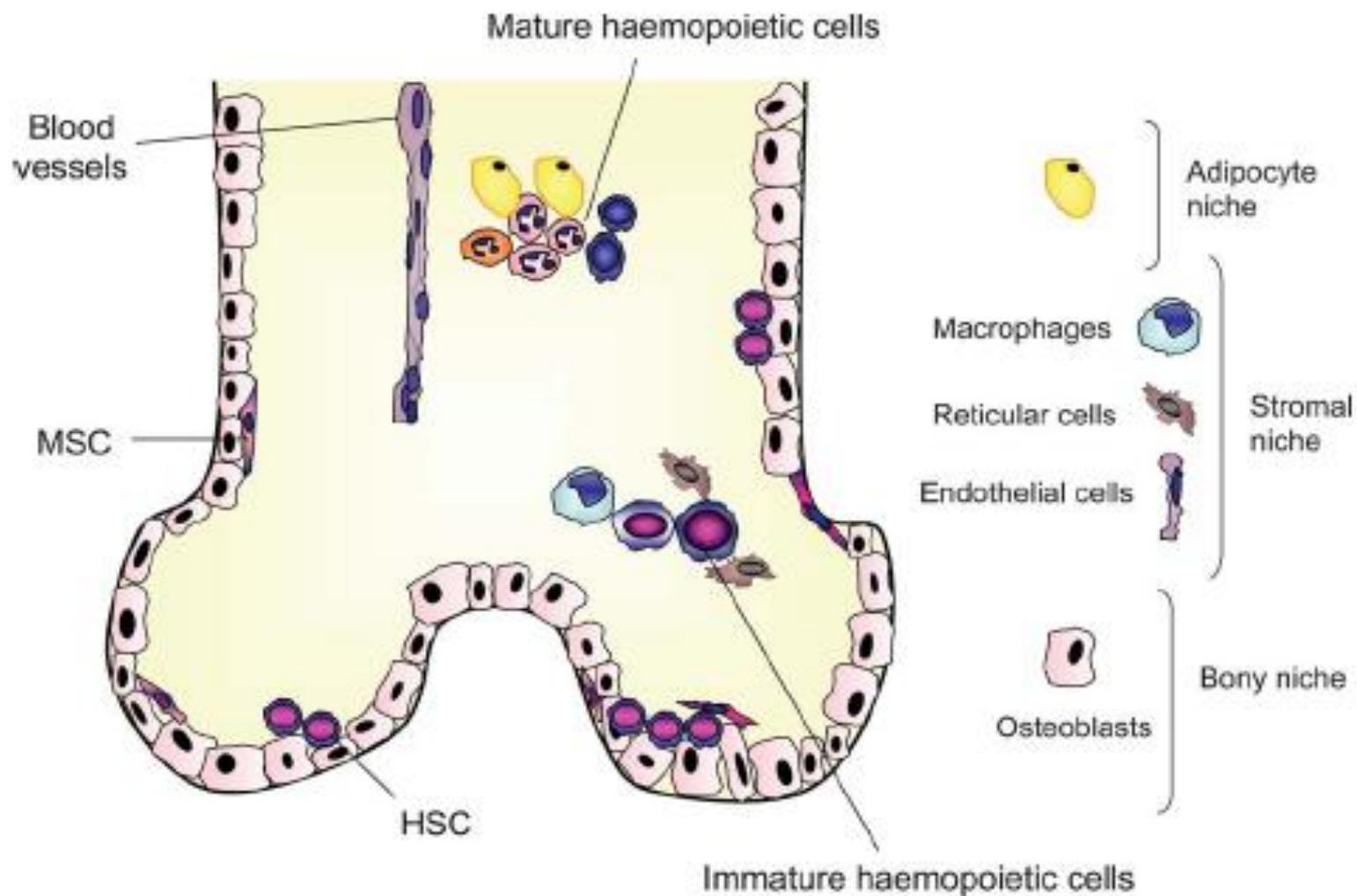
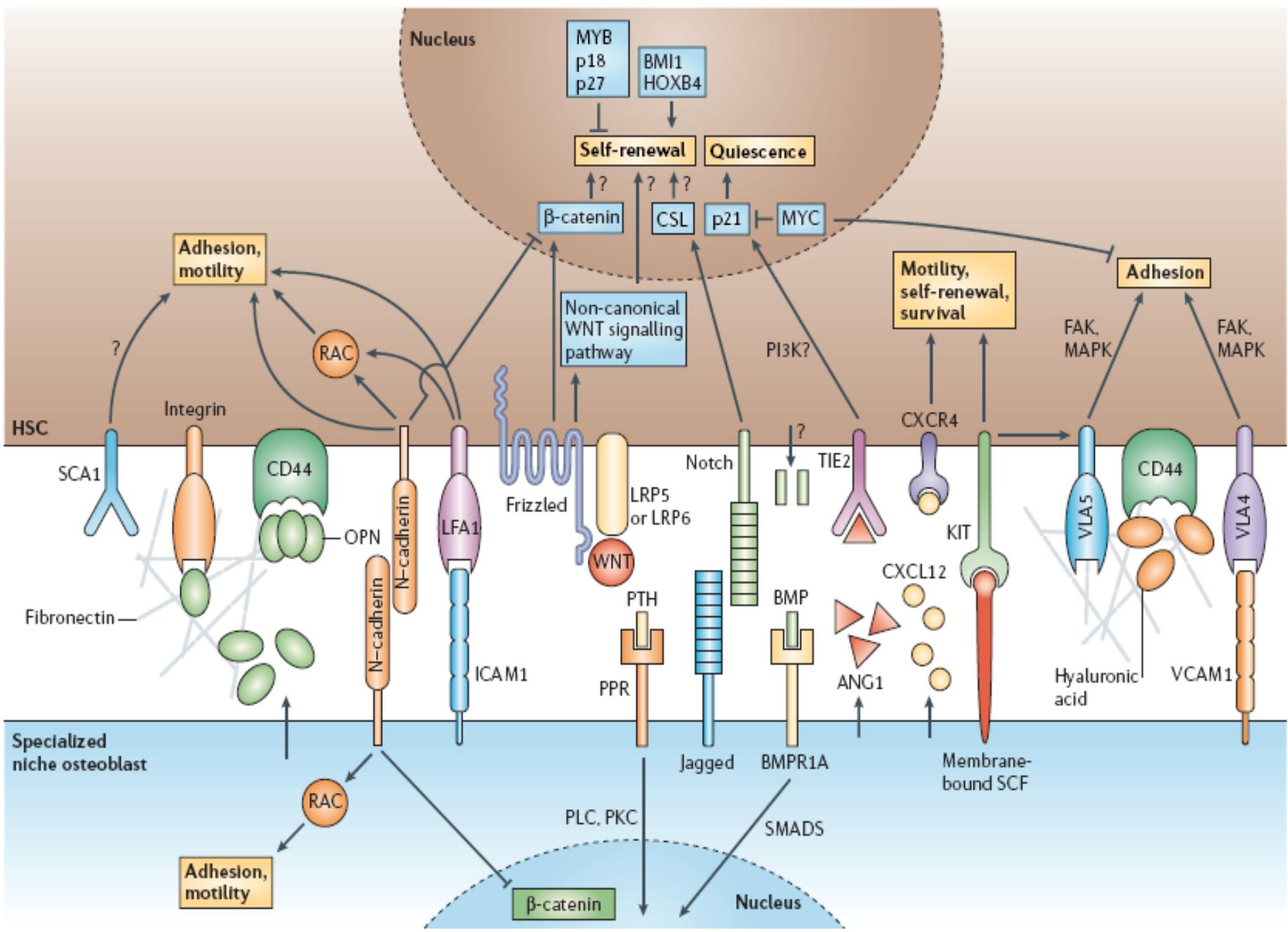
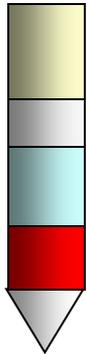


Figure 1 The cellular components of the haemopoietic niche.



Sangue di CB



- plasma
- mononucleate
- ficoll
- globuli rossi
- granulociti

Pcr quantitativa



Profilo di espressione DNA microarrays



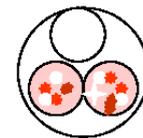
Analisi morfologica



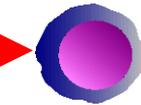
Immunofenotipo



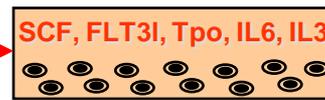
Saggio clonogenico



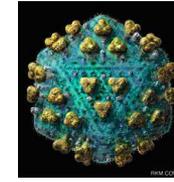
α CD34



HSC CD34+



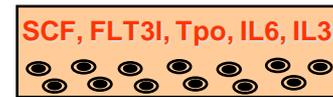
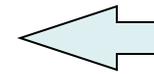
pRRL.PGK.FTX.IE.SIN-18



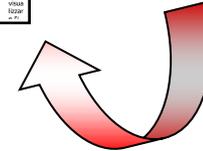
Non oltre 24h



Efficienza d'infezione:
WB & citofluorimetria
24h-48h



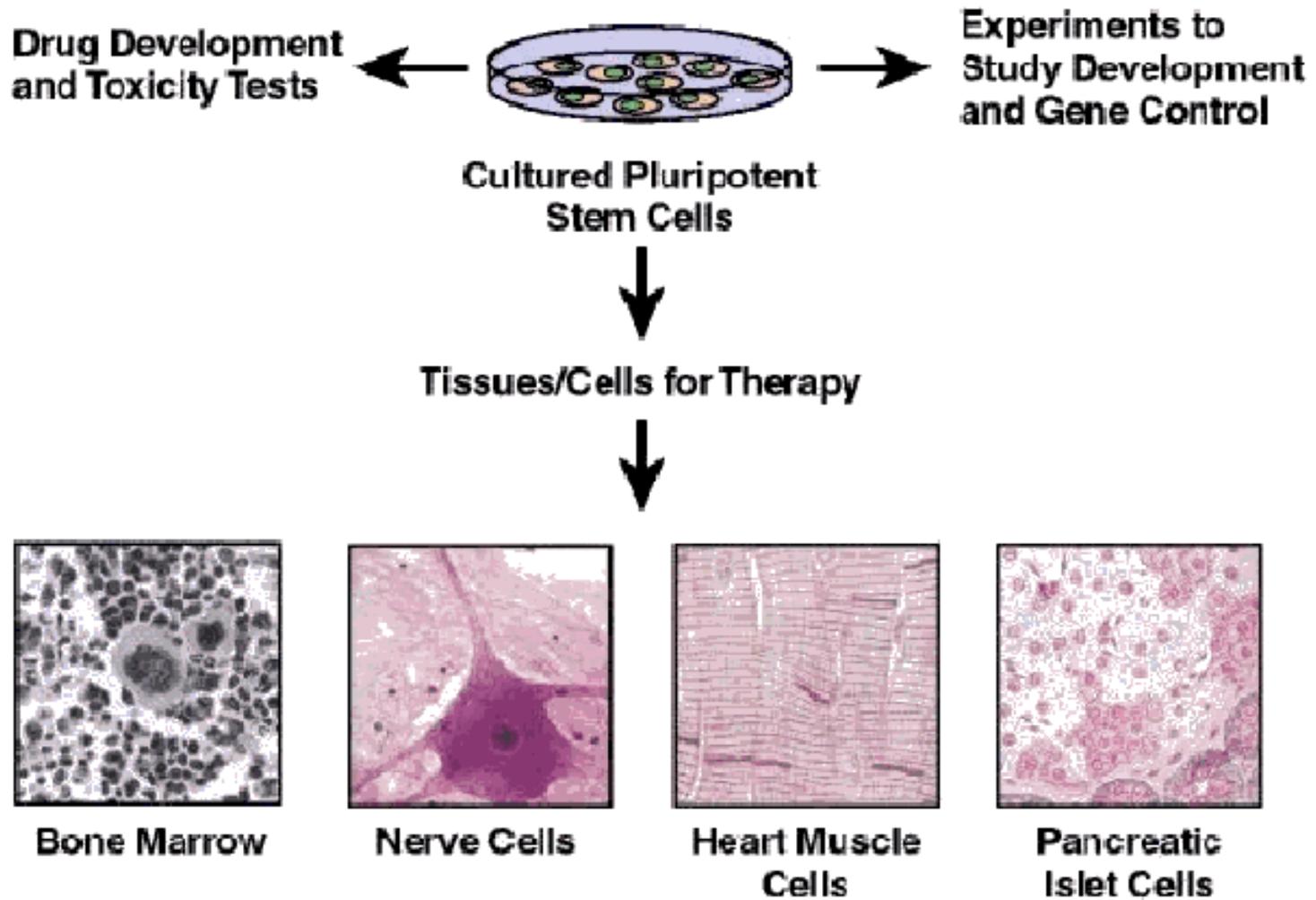
TAMOXIFEN
500nM-1µM 72h



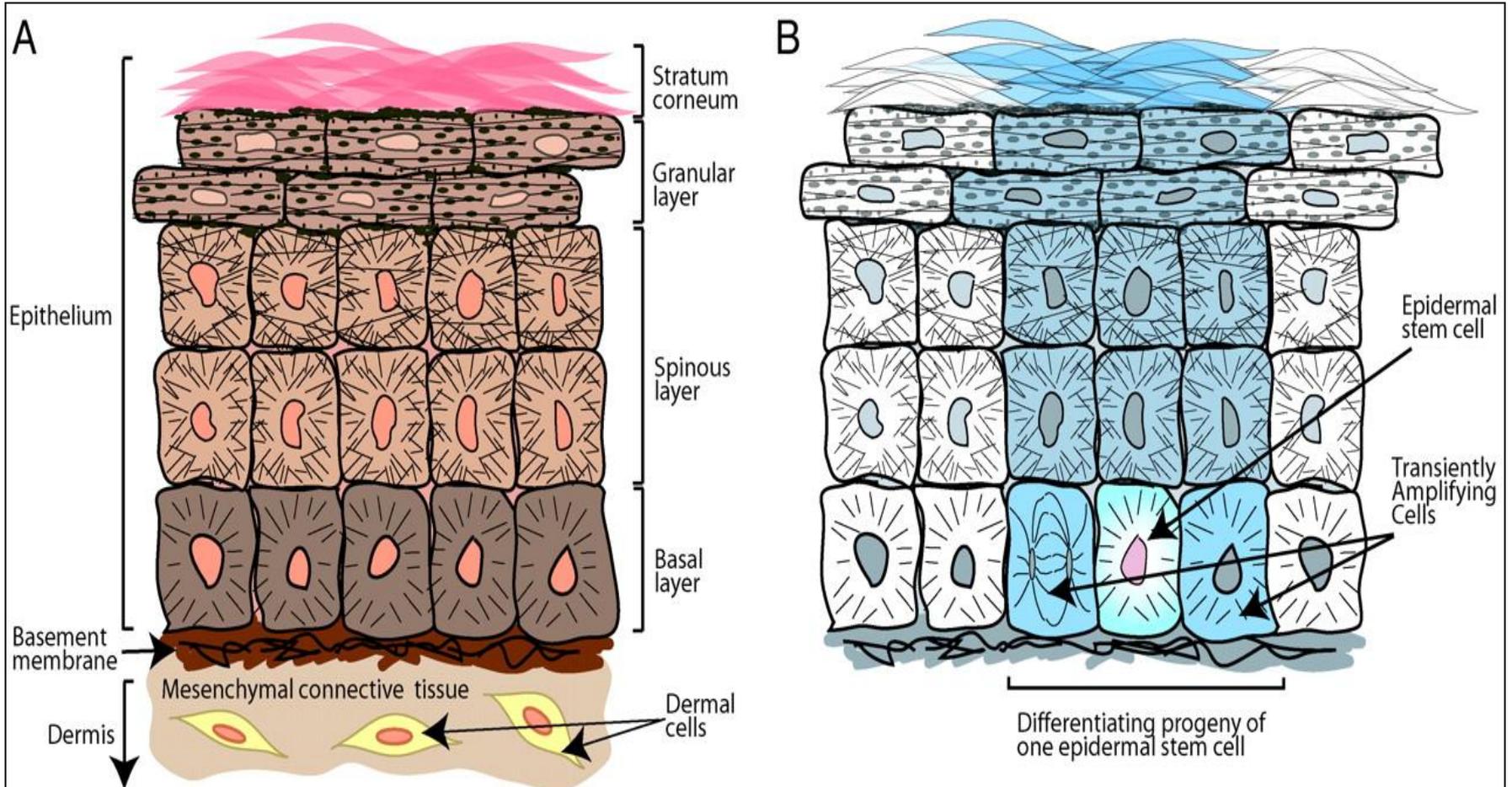
Traslocazione nucleare:
immunocitochimica



The Promises of Stem Cell Research

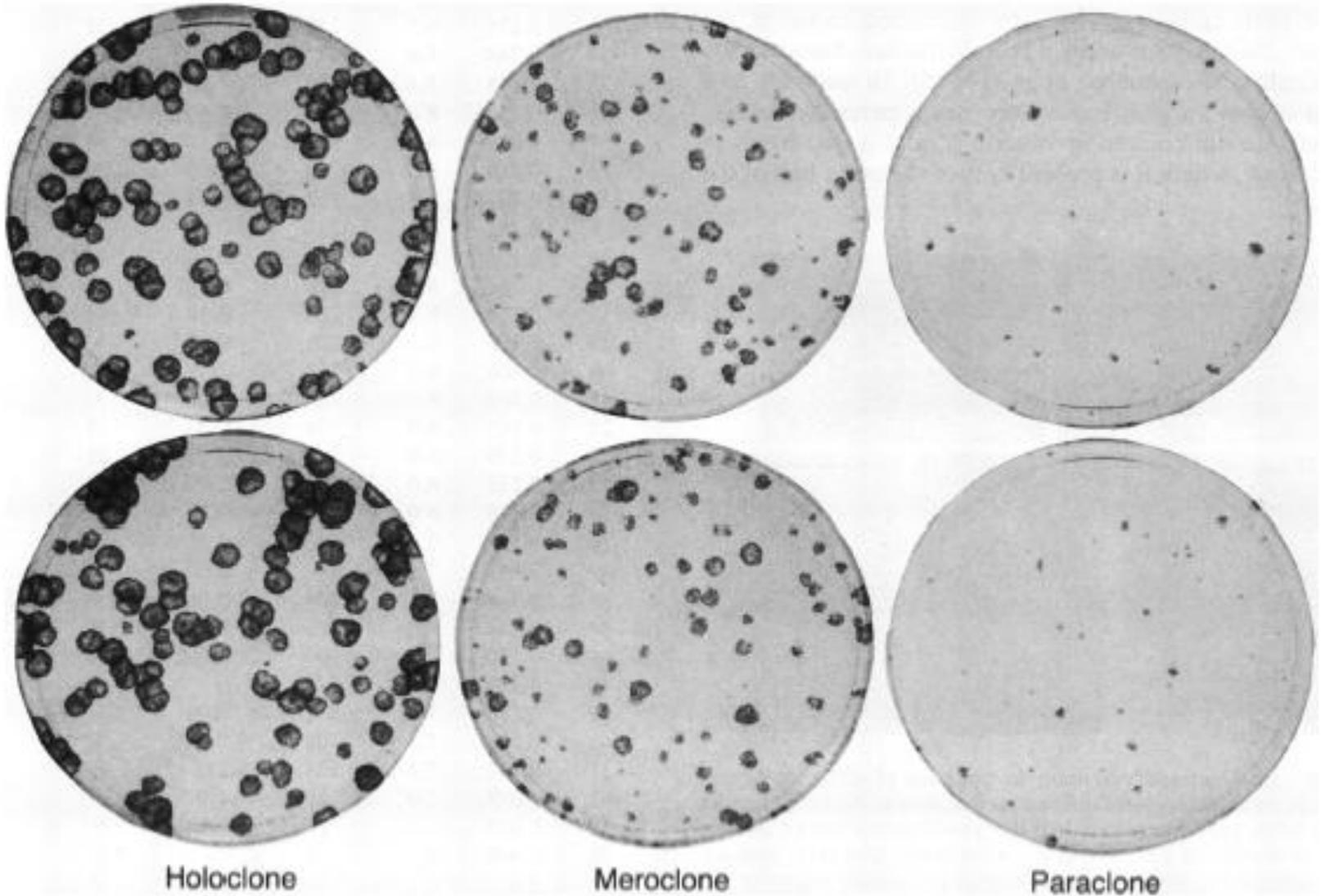


Epiteli stratificati

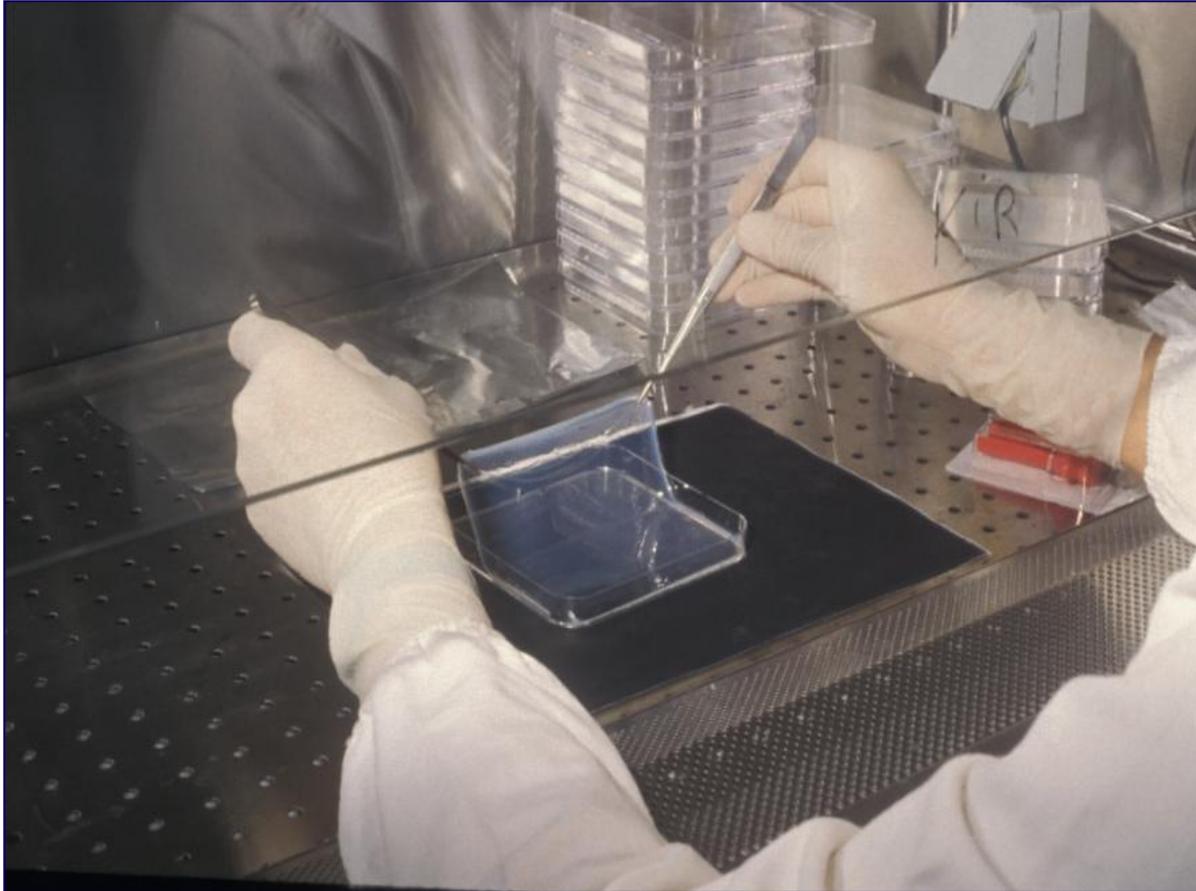


A) Stratificazione dell'epidermide **B)** Differenziamento delle cellule epidermiche.

Cheratinociti dello strato basale

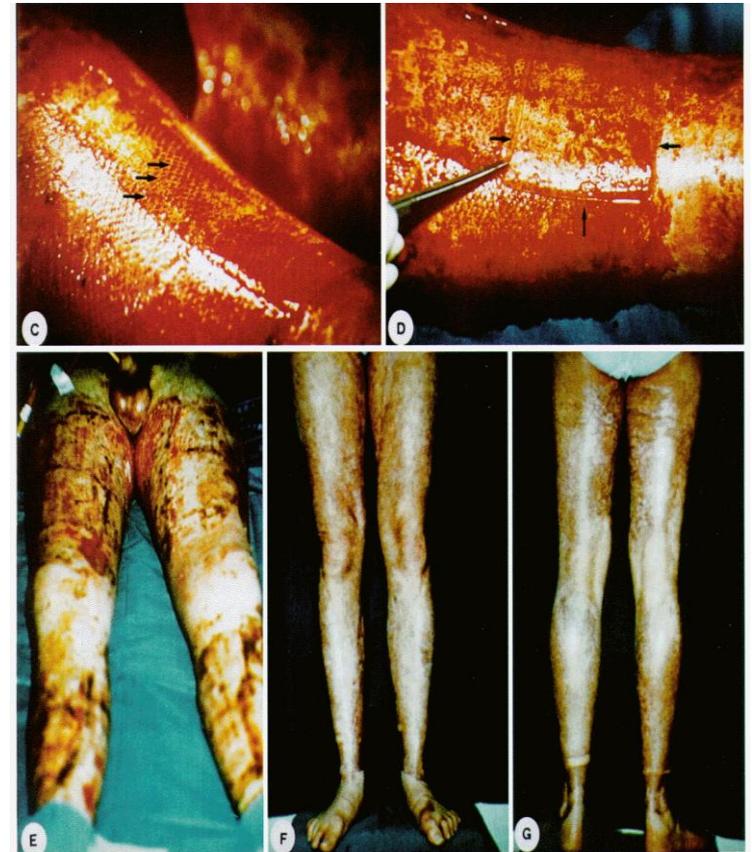
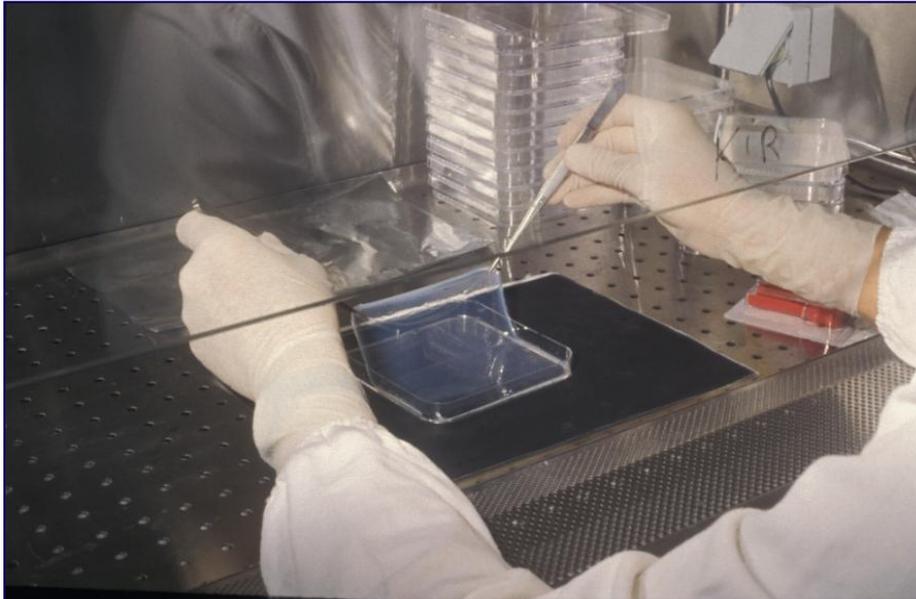


Colture di cheratinociti su colla di fibrina (2)



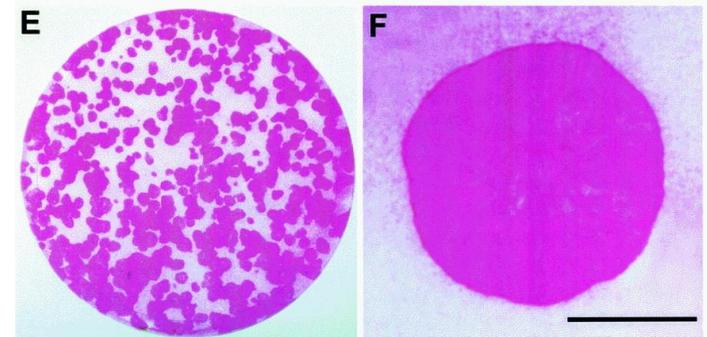
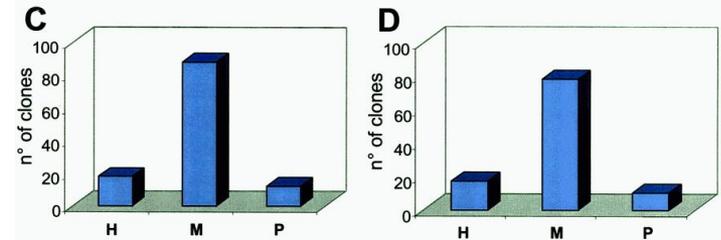
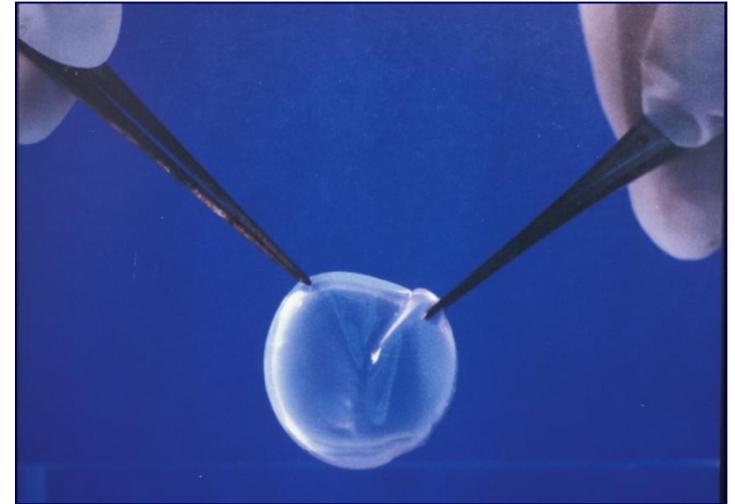
La manipolazione dei lembi di pelle, il loro trasporto e il loro impianto sono agevolati dalla colla di fibrina, utilizzata come substrato, che viene poi facilmente rimossa dai meccanismi fisiologici dell'organismo deputati a questa funzione.

Medicina rigenerativa (trapianto di cute)

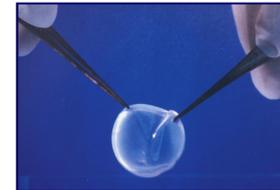
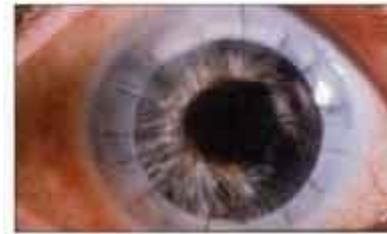
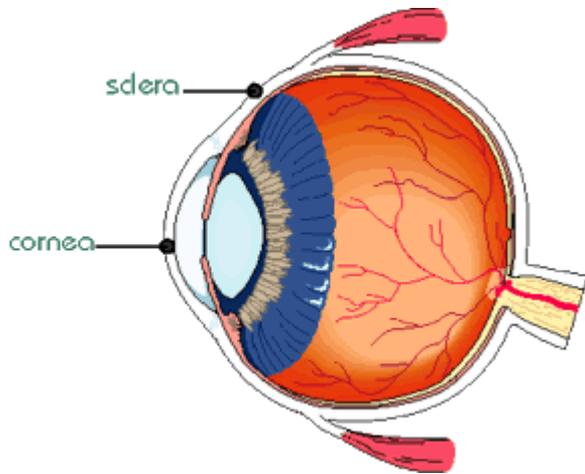


Produzione di cornee su colla di fibrina

- Biopsia di 1-2 mm² dal limbus dell'occhio controlaterale sano di ogni paziente.
- Cellule cresciute per 14-16 giorni in condizioni idonee su gel di fibrina fino ad ottenere un lembo di 3 cm di diametro.
- Trapianti effettuati con questa tecnica di coltura sono stati eseguiti su 116 pazienti affetti da assenza totale o molto severa di cellule del limbus.

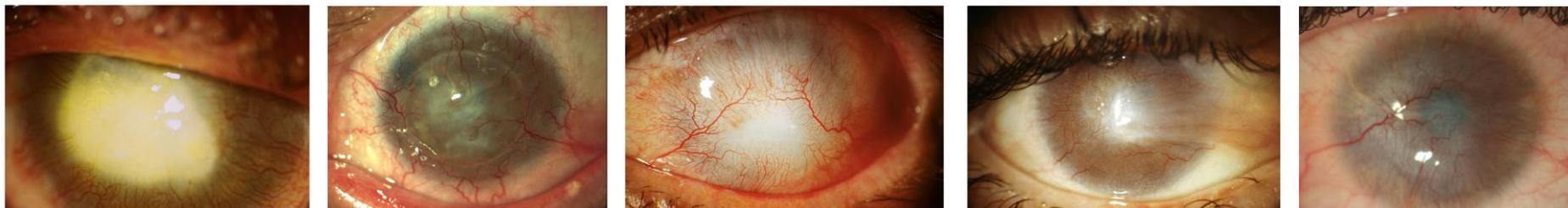


Medicina rigenerativa (trapianto corneale)



Corneal regeneration by cultures of limbal stem cells (up to 10 years follow-up)

BEFORE: SEVERE SYMPTOMS AND LOSS OF VISION



AFTER: TOTAL RESTORATION



2 yrs follow-up

2 yrs follow-up

2 yrs follow-up

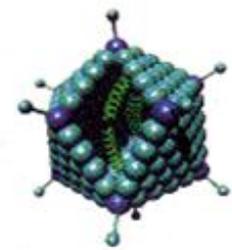
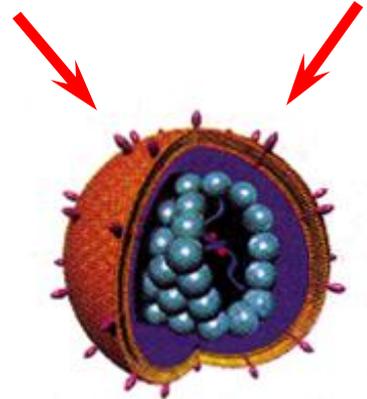
4 yrs follow-up

4 yrs follow-up

full recovery of visual acuity

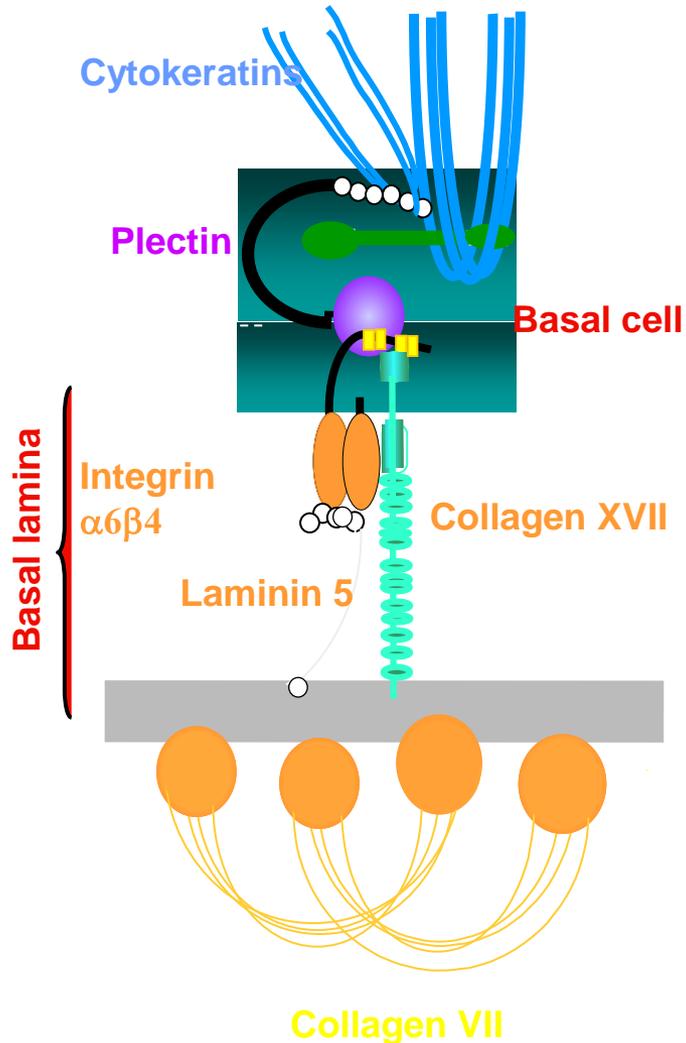
Onco-retro-virus

Lenti-virus



	Retrovirus	Adenovirus	Virus adeno-associati	Liposomi	DNA «nudo»
Alcuni potenziali vantaggi	Integrano i geni nei cromosomi dell'ospite, consentendo una stabilità a lungo termine	La maggior parte non causa gravi malattie; possono accogliere geni estranei di grandi dimensioni	Integrano i geni nei cromosomi dell'ospite; non causano malattie umane note	Non hanno geni virali e pertanto non causano malattie	Come i liposomi; si prevede che sia utile per le vaccinazioni
Alcuni difetti dei vettori esistenti	I geni si integrano a caso, pregiudicando a volte i geni dell'ospite; molti infettano solo cellule in divisione	I geni a volte funzionano transitoriamente, per la mancata integrazione o l'attacco del sistema immunitario	Non possono accogliere geni estranei di grandi dimensioni	Sono meno efficienti dei virus nel trasferire geni alle cellule	È inefficiente nel trasferimento genico e instabile in gran parte dei tessuti dell'organismo

Epidermolisi bollose



EBS

**KRT5
KRT14**

EBS-MD *PLEC1*

JEB

*LAMA 3
LAMB 3
LAMC 2
ITGA 6
ITGB 4*

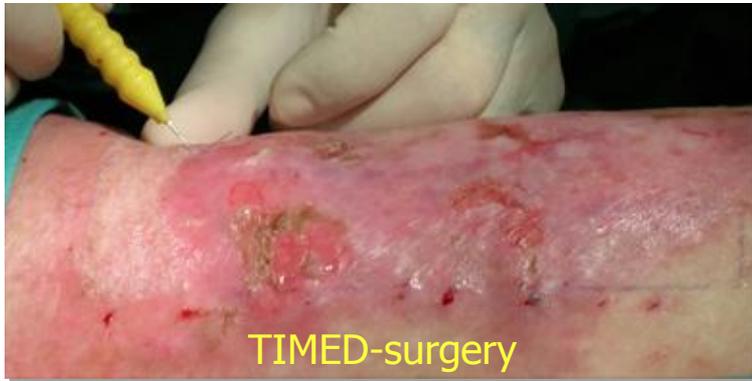
DEB

*COL17A1
COL7A1*



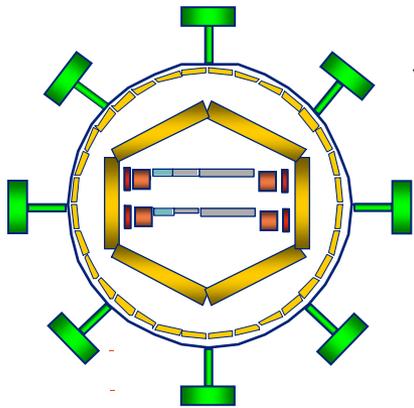
Graft of autologous transduced cultured epidermis

Mavilio et al., *Nature Medicine*, 2006, 12:1397-1402

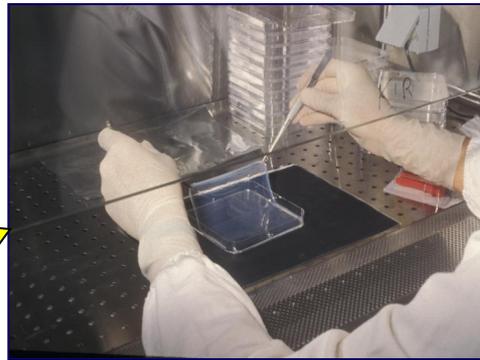


Terapia genica

Produzione di vettori
retro/lentivirali



Trasduzione *ex vivo*
di colture di cheratinociti



Trapianto su paziente delle
cellule trasdotte



EPITHELIAL STEM CELLS AND REGENERATIVE MEDICINE

Michele De Luca



Centre for Regenerative Medicine "*Stefano Ferrari*"
Department of Biomedical Sciences
University of Modena and Reggio Emilia, Modena, Italy

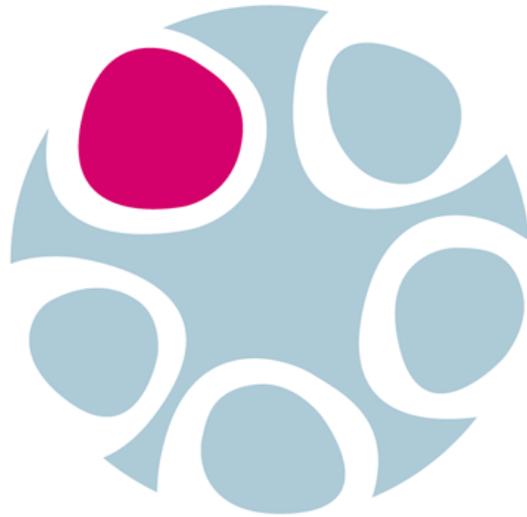
OBSTACLES TOWARDS STEM CELL-MEDIATED REGENERATIVE MEDICINE

1

CONSTRUCTION OF A GMP FACILITY (opening: October 2008): **13.5 Millions €**

CENTER FOR REGENERATIVE MEDICINE “Stefano Ferrari” CRM, UNIVERSITY OF MODENA AND REGGIO EMILIA (4,000 m²)





HOLOSTEM
T E R A P I E A V A N Z A T E

Gaining Insight into the Strict Requirements for the Market Approval of Cell-Based Therapies

Andrea Chiesi

CEO

London, December, 3rd 2009