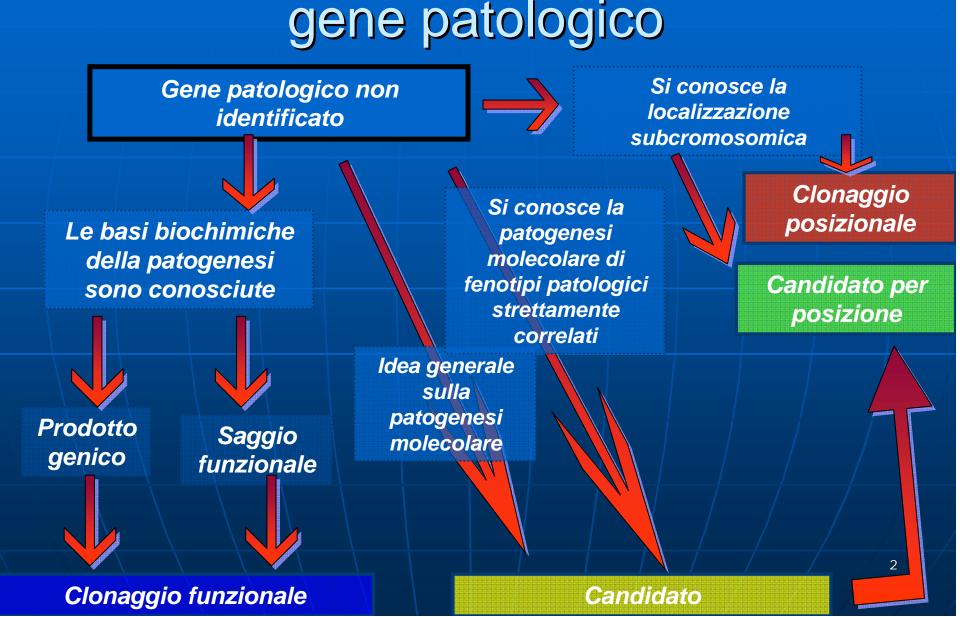
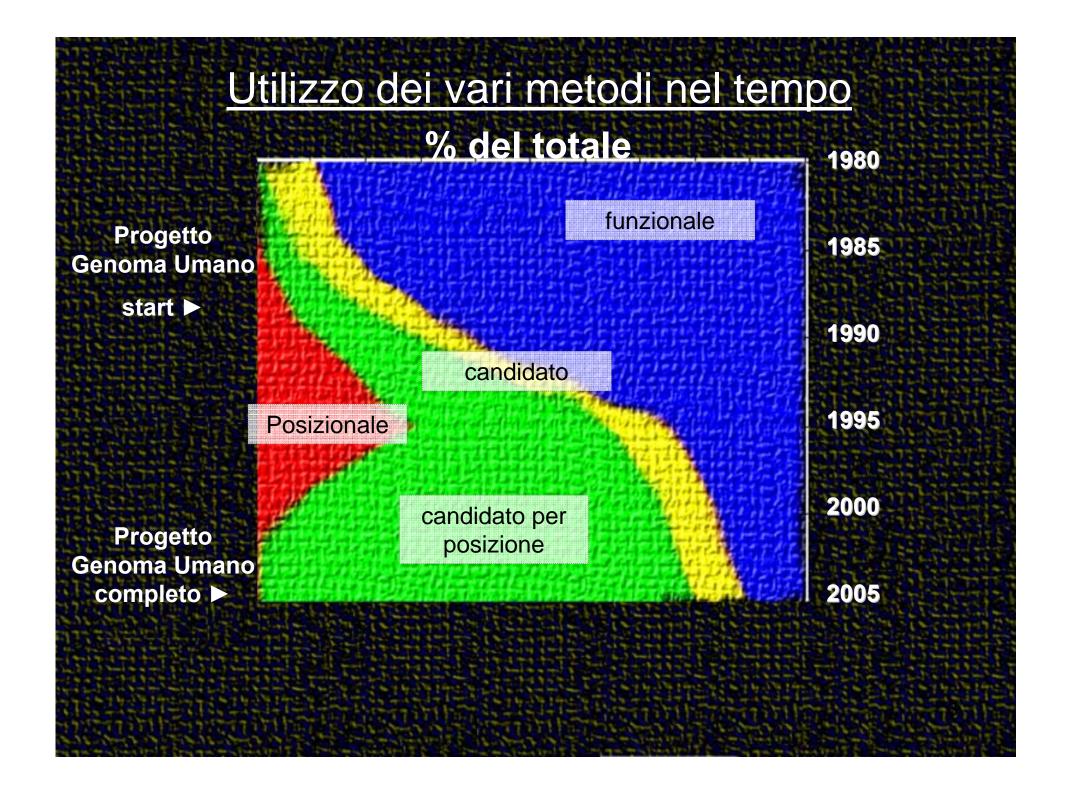
# Strategie per l'identificazione di geni malattia

WWW.FISIOKINESITERAPIA.BIZ

### Quattro metodi per identificare un gene patologico





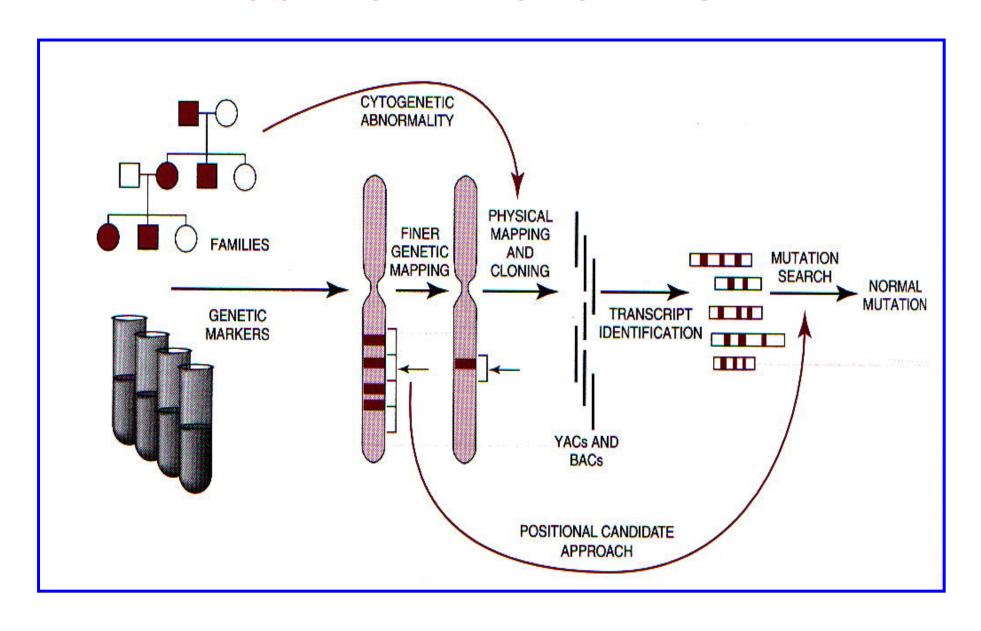
#### Clonaggio funzionale

 Alcune informazioni sulla funzione del gene vengono sfruttate per isolare un clone del gene. Se il prodotto genico è noto, la parziale purificazione del prodotto può permettere l'adozione di varie strategie per identificare il gene responsabile. Alternativamente, può essere usato un saggio funzionale per controllare la presenza del gene. Questo approccio è stato utile solo in pochi casi.

#### Clonaggio posizionale

 Significa isolare il gene conoscendo solo la sua localizzazione subcromosomica, senza utilizzare alcuna informazione riguardante la patogenesi o la funzione biochimica. La strategia è di cercare di costruire la mappa fisica e la mappa genetica della regione, precisare la localizzazione subcromosomica e poi identificare i geni presenti nella regione per valutarli come possibili geni patologici. Il clonaggio posizionale rimane arduo e sta diventando sempre meno necessario per l'accumularsi di informazioni che permettono un approccio posizionale al gene candidato

#### POSITIONAL CLONING



#### ANALISI DI LINKAGE (CONCATENAZIONE)

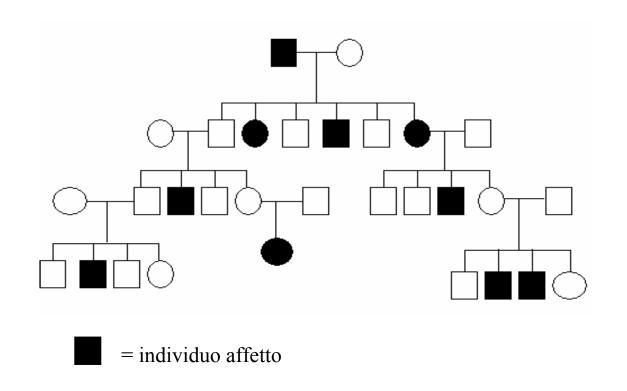
- •L'analisi di linkage permette determinare la posizione cromosomica di un locus responsabile di una determinata malattia/carattere genetico rispetto a marcatori polimorfici la cui localizzazione è nota
- •L'analisi di linkage è un approccio molto utile per il mappaggio e l'identificazione di geni responsabili di malattie genetiche Mendeliane

(oltre 1,400 geni identificati)

•Più difficoltosa è l'identificazione dei numerosi geni implicati in malattie più comuni, ad ereditarietà complessa (pochissimi geni identificati finora)

#### REQUISITI per l'ANALISI DI LINKAGE DI CARATTERI MENDELIANI

1) Una o più famiglie in cui segrega il carattere/malattia in questione



#### MARCATORI MOLECOLARI (DEL DNA)

#### RFLPs (Restriction Fragment Length Polymorphisms)

- •Presenza/assenza sito di taglio di enzima di restrizione
- •Biallelici
- •Southern blot / PCR

#### **MICROSATELLITI**

- ripetiz in tandem di 2-3-4 nucleotidi (CA)n
- •Molti alleli, molto informativi
- •Distribuiti in modo uniforme nel genoma. ≈ ogni 100 Kb
- •PCR / marcatura con fluorescenza

#### SNPs (Single Nucleotyde Polymorphisms)

- •Differenze di singola base, non necessariamente riconosciuti da enz. di restriz.
- •Biallelici
- •Molti frequenti,
- •Sviluppo di tecniche di genotyping automatizzate e in larga scala

#### ANALISI DI LINKAGE DI CARATTERI MENDELIANI

- $\triangleright$  Determinare la frequenza con cui due loci ricombinano fra loro alla meiosi = **TETHA** ( $\theta$ )
- se due loci sono su cromosomi diversi segregano indipendentemente. La probabilità che vengano ereditati insieme è del  $1/2 -> \theta = 50\%$
- Se due loci sono vicini fra loro sullo stesso cromosoma saranno ereditati insieme più frequentemente ->  $\theta$  < 50%

  Tanto più sono vicini, tanto più piccola è la probabilità che avvenga un crossing-over

LA FREQUENZA DI RICOMBINAZIONE (θ) E' UNA MISURA DELLA DISTANZA GENETICA

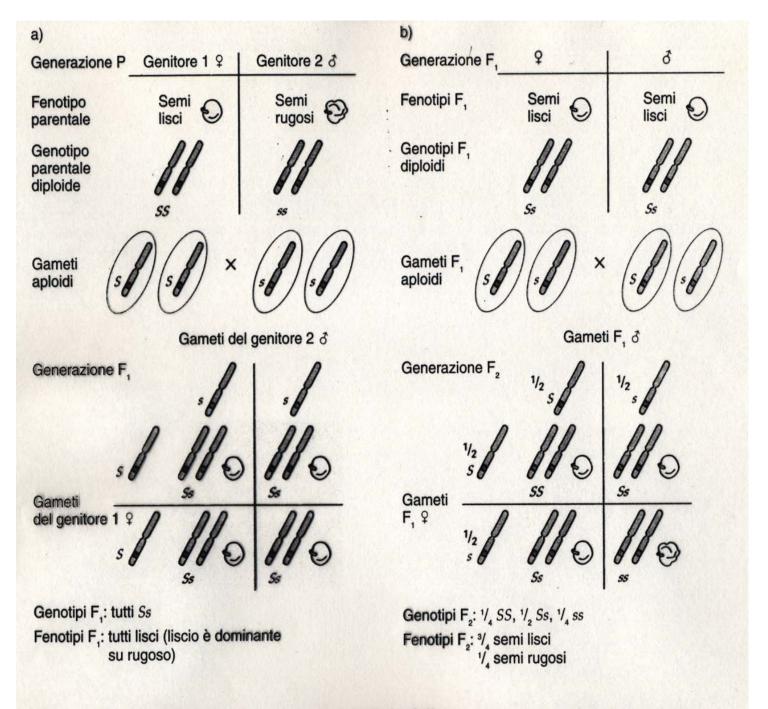


Figura # 2.11ab - Prima legge di Mendel: principio della segregazione dei fattori mendeliani

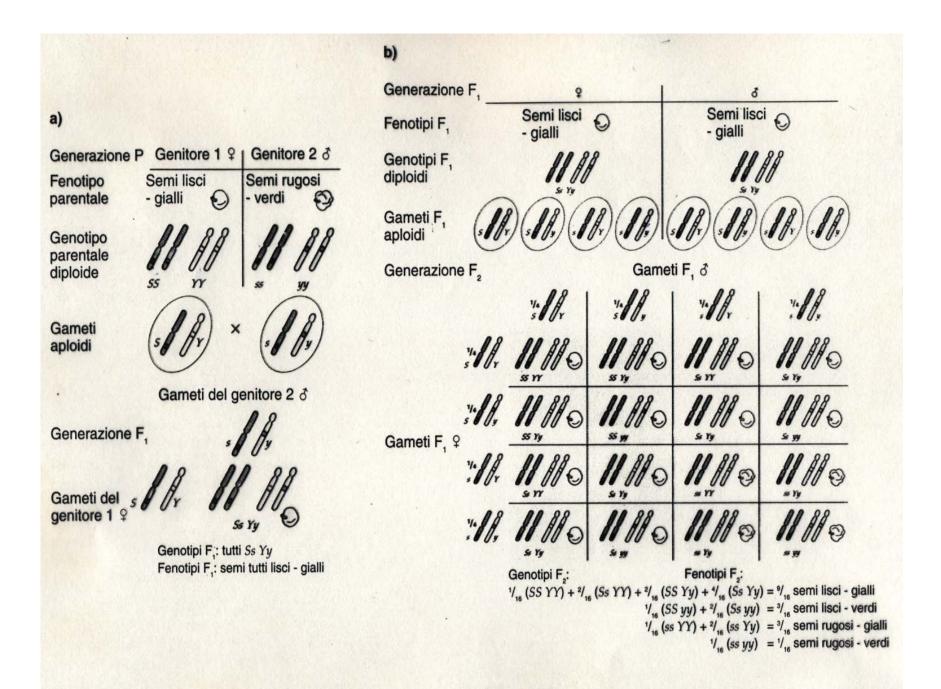


Figura # 2.15 - Principio dell'assortimento indipendente: (a) produzione della generazione F<sub>1</sub>; (b) genotipi della F<sub>2</sub> ed il rapporto fenotipico 9:3:3:1 usando il quadrato di Punnett

Topo albino c Topo shaker sh Normale C Normale Sh

c/sh x C/Sh

F1 normali CcShsh

#### $\mathbf{F1} \quad \mathbf{x} \quad \mathbf{F1} \longrightarrow \mathbf{F2}$

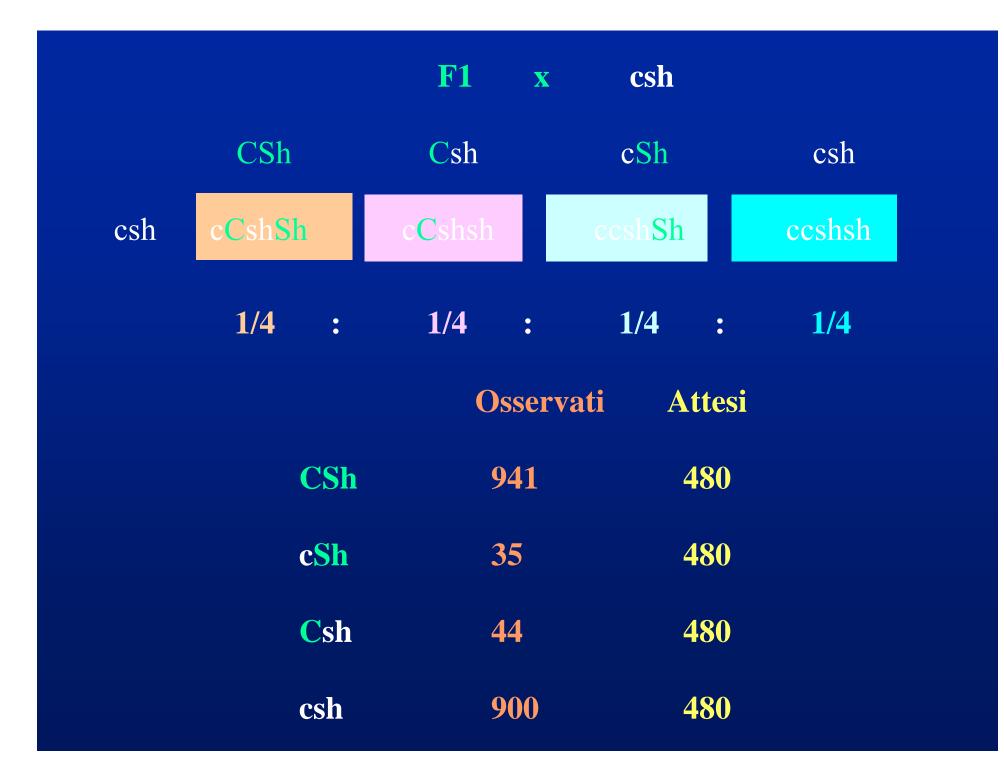
CSh Csh cSh csh **CCShSh CCS**hsh CSh CcShSh CcShsh **CC**shsh CcshSh **CCS**hsh Ccshsh Csh cSh cCShSh cCShsh ccShSh cc**Sh**sh cCshSh c**C**shsh ccshSh ccshsh csh

9/16 : 3/16 : 3/16 : 1/16

	Osservati	Attesi	
CSh	192	146.7	
cSh	9	48.9	
Csh	3	48.9	
csh	57	16.3	
Normali ———	• albini		
195	66	3 :	1
Normali	Shaker	•	
201	60		

CSh csh
CCShSh CcShsh
csh ccshsh

3 : 1



Normali  $\longrightarrow$  albini

985

935

1:1

Normali  $\longrightarrow$  Shaker

976

944

csh cCshSh csh

1 : 1

#### Il Linkage

La tipizzazione di famiglie ampie ed informative, dove segrega una specifica malattia genetica, consente di stabilire se esiste una "concatenazione" tra un marcatore specifico ed il locus malattia.

L' esistenza di un rapporto di concatenazione viene dimostrato tramite vari metodi tra cui quello del rapporto di massima verosimiglianza (o metodo dei lod scores).

La base logica di questo metodo è molto semplice ed è dovuta al fatto che in una progenie si possono osservare due situazioni, o quella del riassortimento indipendente di due geni (o di un marcatore con un gene) o quella della concatenazione degli stessi.

I lod scores vengono così calcolati prendendo in considerazione una meiosi alla volta e confrontando la probabilità dei genotipi osservati nelle ipotesi alternative di concatenazione (linkage) o di assortimento indipendente. Tuttavia anche quando due loci sono localizzati sullo stesso cromosoma bisogna tener conto che si possono osservare fenomeni di ricombinazione per il "crossing-over" che è dovuto alla distanza dei due loci.

La percentuale di ricombinazione è tanto maggiore quanto più distanti sono i loci.

La percentule di ricombinazione viene indicata con la lettera greca  $\theta$  e l'1% di ricombinazione corrisponde all'unità di mappa o centiMorgan (cM) dal nome di Morgan che scoprì la concatenazione genica nel 1911.

50% di ricombinazione Assenza di concatenzione  $\theta$ =0.5

#### Il Linkage

Si calcolano cioé due probabilità:

- 1) la probabilità L0 che la progenie di quella famiglia sia ottenuta in assenza di concatenazione ( $\theta$ =0.5).
- 2) la probabilità L1 che la stessa progenie si sia generata in presenza di concatenazione ( $\theta$ < 0.5).

Queste probabilità vengono più esattamente definite come verosimiglianza delle osservazioni nelle due ipotesi di indipendenza e di linkage.

> Si calcola quindi il rapporto di verosimiglianza L= L1/L0

e si cerca il valore di  $\theta$  per cui tale rapporto risulta massimo.

Per semplicità di calcolo si usa il logaritmo in base 10 del rapporto di verosimiglianza:

 $Z = \log (L1/L0)$ 

che in inglese viene denominato lod score.

La mappatura di una malattia si ottiene così quando il valore di Z è uguale o superiore a 3. In effetti questo valore sta ad indicare che l' ipotesi di concatenazione per un determinato valore di  $\theta$  è 1000 volte più probabile di quella di indipendenza

### Geni e patologie identificati con: Clonaggio posizionale

- **1986** 
  - Distrofia muscolare di Duchenne
  - Retinoblastoma
- 1989
  - Fibrosi cistica
- **1990** 
  - Neurofibromatosi di tipo 1
  - Tumore di Wilms
- 1991
  - Aniridia
  - Poliposi familiare del colon
  - Sindrome dell' X fragile
  - Distrofia miotonica

- **1993** 
  - Malattia di Huntingtion
  - Sclerosi tuberosa
  - Malattia di von Hippellindau
- **1994** 
  - Acondroplasia
  - Cancro mammario/ovarico a esordio precoce
  - Malattia del rene policistico
- **1995** 
  - Atrofia muscolo spinale

### Strategie con gene candidato indipendenti dalla localizzazioni

Si può sospettare che un gene sia responsabile di una patologia umana senza sapere niente della sua localizzazione cromosomica. Ciò può avvenire se un certo fenotipo assomiglia a un altro fenotipo, in animali o esseri umani, per il quale si conosca il gene responsabile o se la patogenesi molecolare suggerisce che il gene possa essere membro di una famiglia genica nota. Tali approcci hanno avuto successo solo raramente e sono stati superati dalle strategie con il gene candidato, di cui si conosce la localizzazione.

### Strategie con gene candidato conoscendo la localizzazione

 Una volta che una patologia sia stata mappata, sta diventando possibile utilizzare sempre più spesso la ricerca nelle banche dati per identificare i geni candidati. Con il continuo aumento del numero di geni umani mappati in specifiche regioni subcromosomiche questi approcci posizionali al gene candidato sono pronti a dominare il campo.

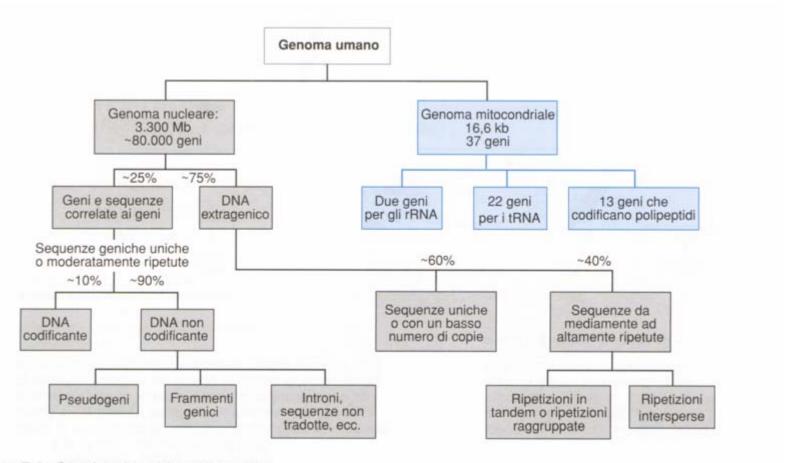


Figura 7.1: Organizzazione del genoma umano.

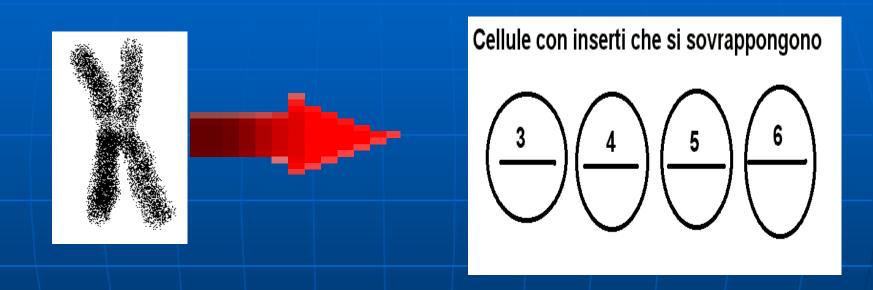
#### Progetto Genoma Umano

- Sequenziamento dell'intero Genoma Umano
- Identificazione dei circa 35000 geni che sono stimati essere presenti nel nostro Genoma

### Identificazione di un gene e correlazione con uno specifico fenotipo

- Riacquisizione del fenotipo normale in vitro
- Costruzione di modelli animali della malattia
- Ricerca di mutazioni causative all'interno di un gene candidato per una specifica malattia genetica in un gruppo di pazienti

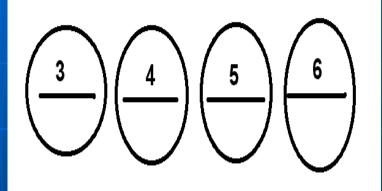
### Digestione parziale



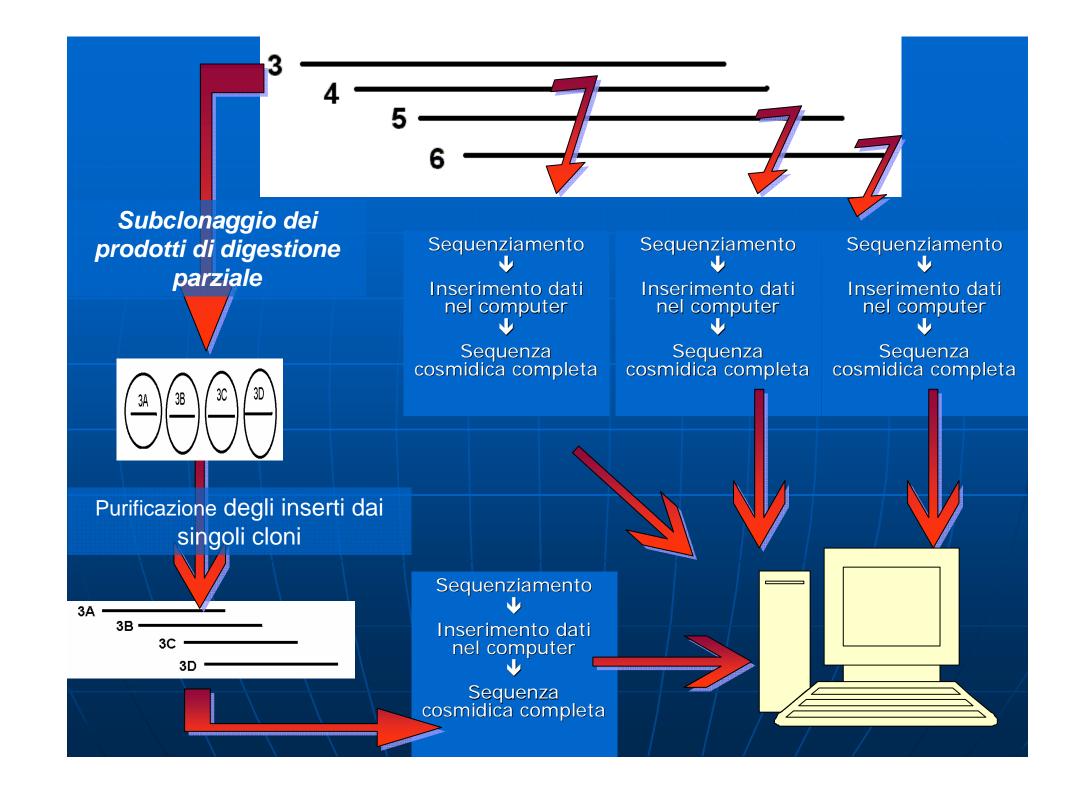
creazione di una libreria di DNA cosmidica

# Purificazione degli inserti dai singoli cloni

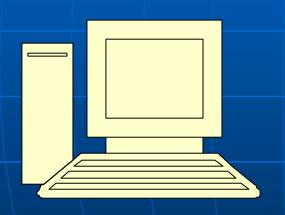
Cellule con inserti che si sovrappongono



3 4 5 6



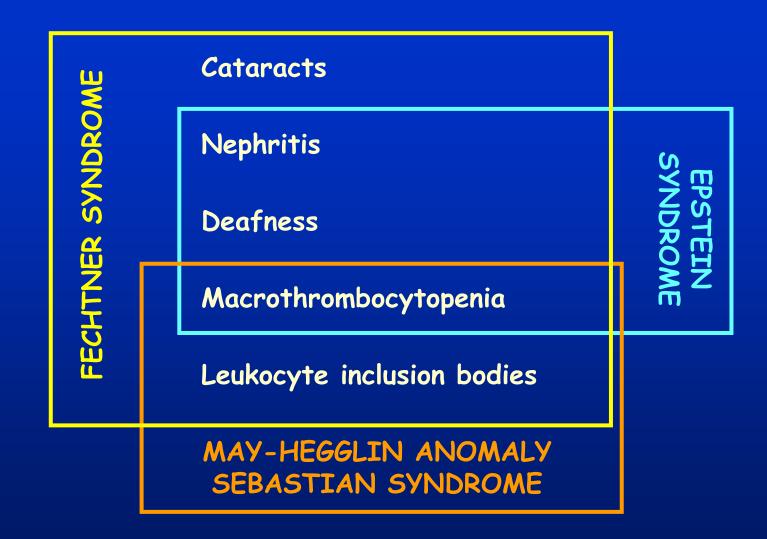


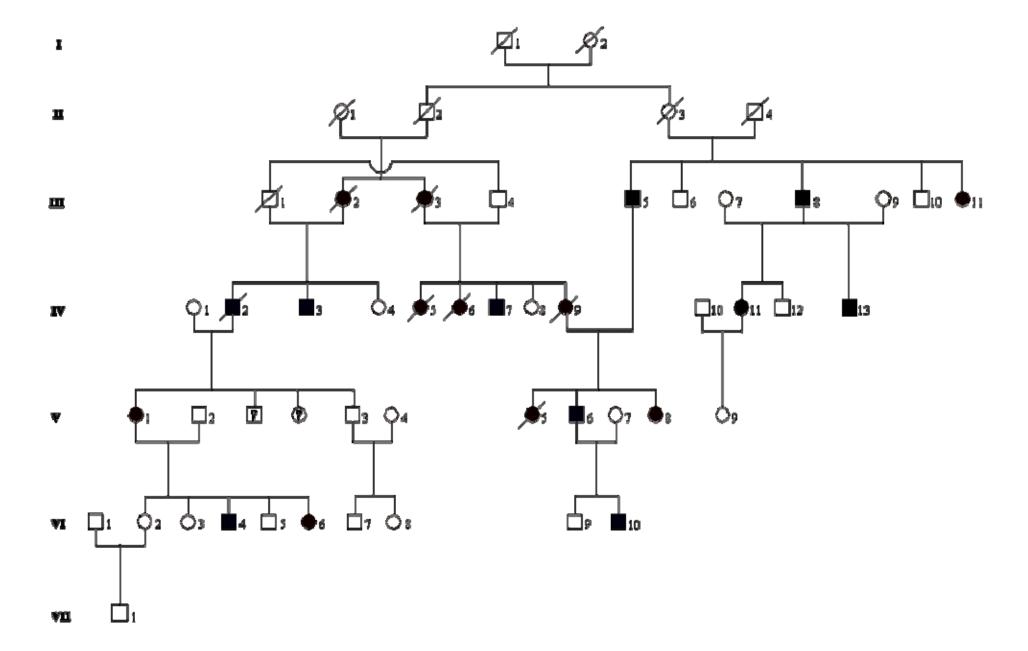


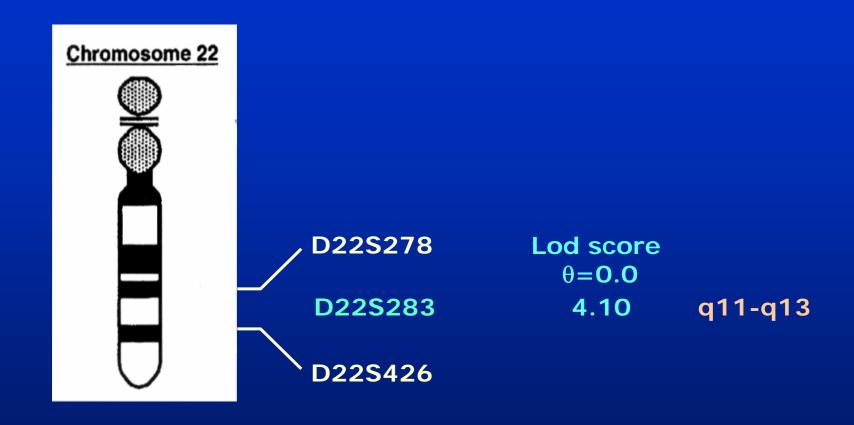
SEQUENZA
 GENOMICA COMPLETA
 O SEQUENZA DI UN
 INTERO CROMOSOMA

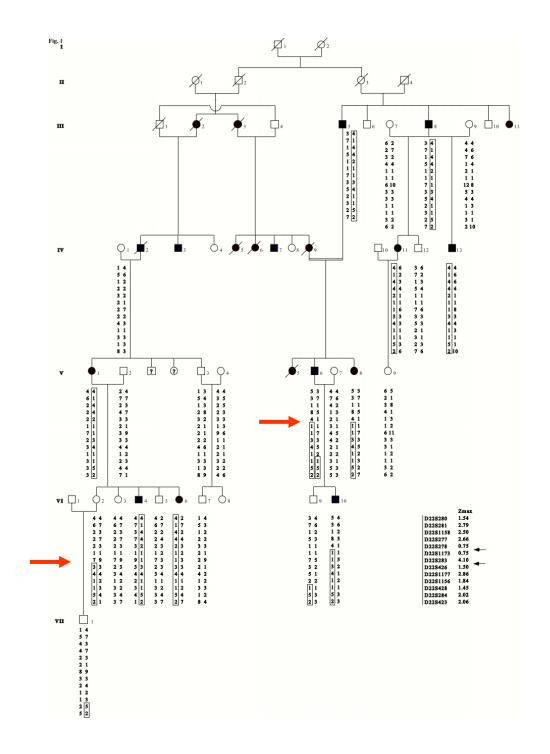
## Malattie (con proteine coinvolte) identificati con: candidati per posizione

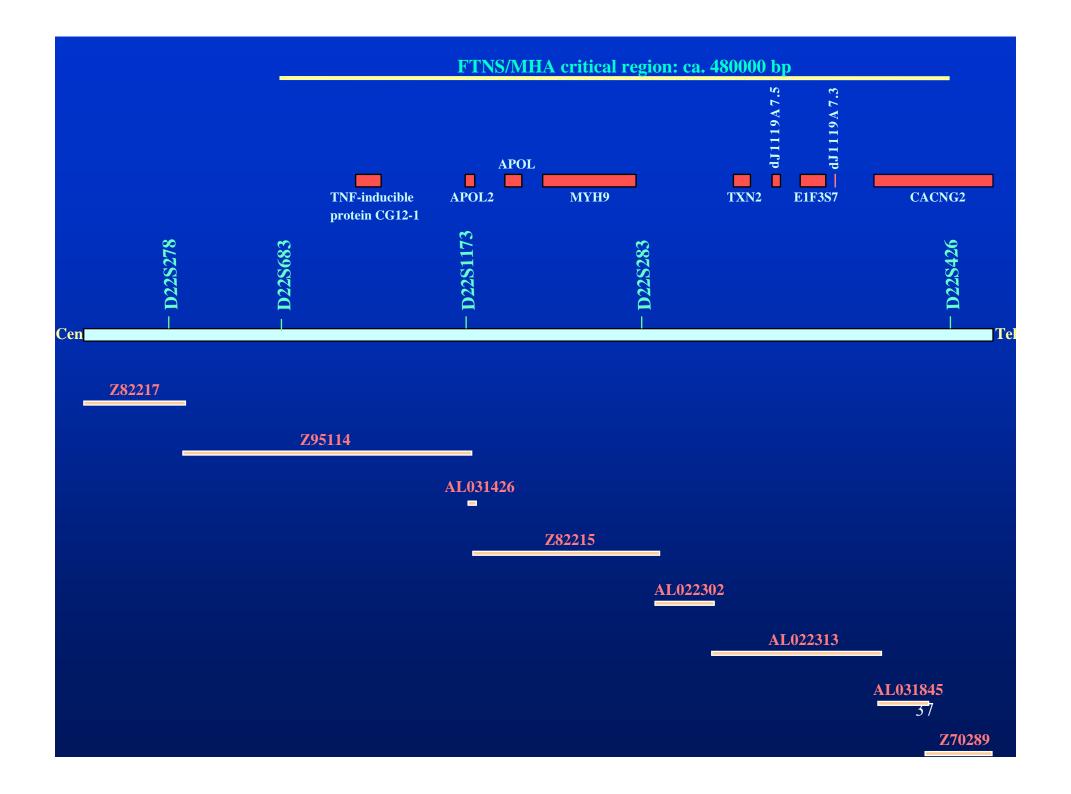
- Morbo di Alzheimer
  - precursore della proteina β-amiloide, apo-E
- Malattia di Charcot-Marie-Tooth di tipo 1A
  - Proteina mielina zero
- Malattia di Charcot-Marie-Tooth di tipo 1B
  - Proteina 22 mielina periferica
- Melanoma familiare
  - P16
- Cancro del colon non associato a poliposi ereditaria
  - hMSH 2, hMLH 1, hPMS 1, hPMS 2
- Ipotermia maligna
  - Recettore della rianodina
- Sindrome di Marfan
  - Fibrillina
- Neoplasia endocrina multipla di tipo 2A
  - Recettore RET della tirosina chinasi
- Retinite pigmentosa
  - Periferina e rodopsina
- Sindrome di Waardenburg di tipo 1
  - Paired box gene PAX 3





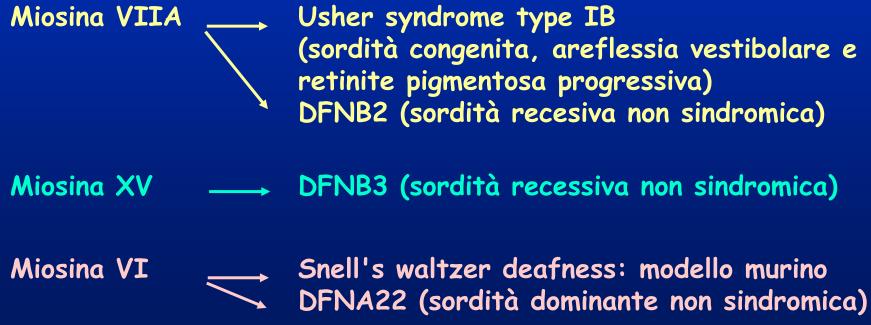


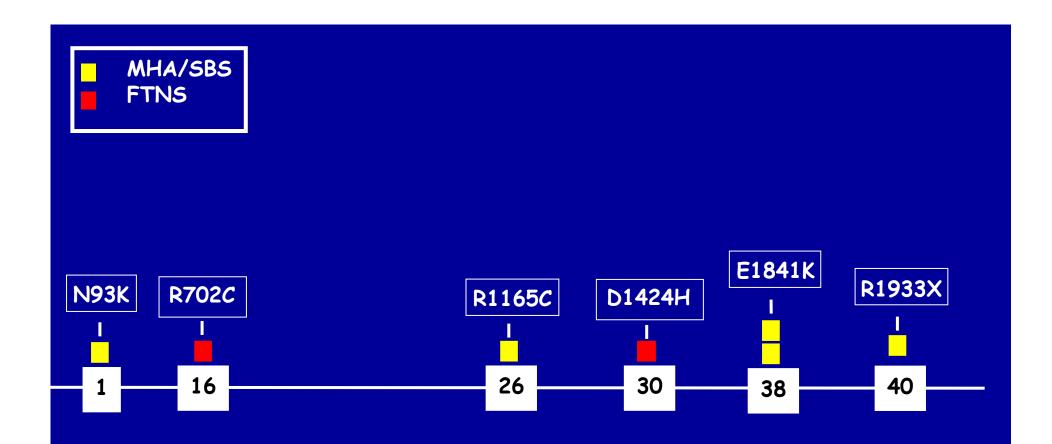




Il gene MYH9 codifica per la catena pesante della miosina non muscolare di tipo IIA (NMMHC-IIA) che è espressa in alcuni tessuti quali piastrine, rene, leucociti e coclea.

Altre miosine non convenzionali sono associate a malattie nell'uomo che mostrano sordità o difetti dell'occhio.





Motor domain

Coil coiled domain

ACD
Assembly competence domain