

27. GENETICA

V. Nigro, S. Aurino

L'ORGANIZZAZIONE DEL GENOMA UMANO

Il progetto genoma umano si è formalmente concluso nell'aprile 2003, quando sono state rese disponibili, ad accesso gratuito, le sequenze di DNA che corrispondono alla porzione eucromatinica dei 22 autosomi, del cromosoma X e del cromosoma Y.

Il totale delle basi sequenziate ammonta ad oltre 2.850.000.000 nucleotidi per genoma aploide maschile a cui devono essere sommate oltre 200 milioni di nucleotidi che rappresentano la porzione eterocromatinica del genoma per un totale di oltre tre miliardi e 100 milioni di nucleotidi. Ogni cellula contiene due copie di tale corredo cromosomico, un corredo di origine materna ed uno di origine paterna. Queste due copie non sono equivalenti, ma restano distinte da un differente pattern di metilazione che inibisce selettivamente l'espressione genica: in alcuni loci è inibito l'allele paterno, in altri l'allele materno. Tutte le cellule di un essere umano hanno nel loro nucleo un identico patrimonio genetico, con una sola piccola eccezione che riguarda i loci delle immunoglobuline nei linfociti B e del TCR nei linfociti T. Quello che differenzia le varie linee di cellule è l'uso di tali sequenze di DNA, che fa sì che una cellula epatica utilizzi parti del genoma differenti da quelle che sono utilizzate, ad esempio, da un neurone.

Il numero di basi di un genoma determina anche la quantità di DNA presente nel nucleo di una cellula. Nelle cellule umane tale quantità ammonta a circa 8 pg di DNA, molto simile alla quantità di DNA ritrovata nelle cellule di altri mammiferi, come, ad esempio, il topo che ha un contenuto di DNA molto simile all'uomo con un genoma aploide composto da 2.9 miliardi di basi. La quantità di DNA è indicata con la lettera c. Si potrebbe ipotizzare che quanto più è complesso un essere vivente tanto più lunga deve essere la sequenza di DNA presente nel nucleo delle cellule che lo compongono. Tale affermazione è in parte vera: un nematode (*C. elegans*) ha un genoma di circa 150 milioni di basi, mentre un insetto (*Drosophila melanogaster*) ha un genoma di 300 milioni di basi e un rospo di circa 1 miliardo di basi. Tuttavia esistono vistosissime eccezioni, tanto che il mais ha stranamente un genoma molto più grande di quello umano.

Allora la differenza potrebbe risiedere nel numero di

geni presenti. Ad esempio, il mais ha nel proprio genoma elementi di DNA di giunzione o riempitivo che alcuni autori hanno definito DNA spazzatura (junk DNA). Il genoma di *C. elegans* (nematode) ha circa 19.000 geni, mentre l'uomo sembra avere 25.000 geni, vale a dire poco di più. Ma la complessità è infinitamente superiore a quella di un nematode se solo si pensa al fatto che il cervello di un uomo è formato da centinaia di miliardi di cellule mentre quello di un verme da poche centinaia.

Contare i geni è un'operazione molto più difficile di quanto potrebbe sembrare a prima vista. Per stabilire dal DNA cos'è un gene e cosa non lo è si deve conoscere quale parte del genoma viene trascritta dalla RNA polimerasi DNA dipendente da tutte le cellule di un uomo in tutte le fasi del suo sviluppo. Per fare ciò sono state generate genoteche di copie dell'RNA presenti in ciascun tessuto dette DNA copia o cDNA, utilizzando un enzima denominato trascrittasi inversa. Tale enzima prodotto soprattutto dai retrovirus e necessario per la loro replicazione catalizza la trasformazione dell'RNA in cDNA, l'inverso di quello che fa la RNA polimerasi. Il cDNA prodotto da ciascuna cellula rappresenta quindi la porzione di genoma utilizzata; ogni cDNA diverso è prodotto da un gene diverso e contare i cDNA equivale a contare i geni. Per arrivare a tale scopo è stato intrapreso il progetto EST (expressed sequence tags). Il progetto EST consiste nel sequenziare una porzione di ciascun cDNA utile per poter identificare il gene di origine. La scelta dei cDNA da sequenziare è assolutamente casuale, con una procedura molto simile a quella della collezione delle figurine autadesive dei calciatori da attaccare sull'album. All'inizio ogni cDNA preso a caso è nuovo, ma per completare la collezione ci impareremo sempre più spesso in cDNA già noti. E poi compariranno molto più frequentemente i cDNA di geni con abbondante produzione di RNA, mentre sarà molto più difficile incontrare i geni debolmente espressi o trascritti magari in un solo citotipo in una brevissima fase dello sviluppo embrionale. Ogni laboratorio che inizia a sequenziare cDNA trova esclusivamente doppietti di cDNA già identificati. Ma la maggiore complessità del genoma umano non va cercata solo nel numero dei geni ma anche nella loro possibilità di dar vita a prodotti proteici alternativi. Il meccanismo principale è lo splicing alternativo. Lo splicing è un'operazione che vie-

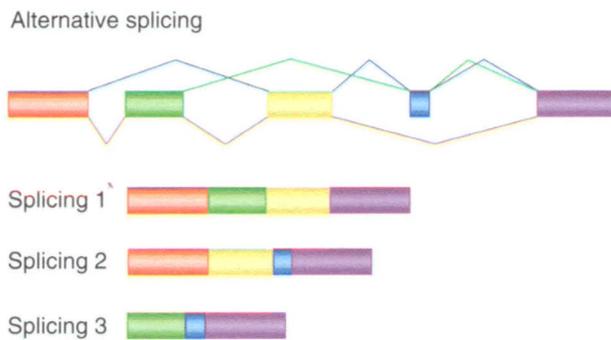


Fig. 1. Splicing alternativo. La figura mostra tre differenti varianti di splicing che daranno origine a tre differenti proteine da un solo gene.

ne effettuata su un film e che viene detta montaggio: si tagliano le parti di pellicola con scene considerate inutili e si saldano tra loro nel modo più appropriato le parti con un senso logico e sequenziale per lo spettatore. La cellula dal trascritto primario di RNA toglie via delle sequenze "inutili" e salda i pezzi in modo sequenziale.

Si calcola che oltre il 60% dei geni umani abbia più di una variante di splicing e alcuni geni meglio studiati come la tropomiosina arrivano a produrre oltre 800 varianti diverse a partire da un solo gene.

Altri meccanismi sono legati alla possibilità di tradurre o meno gli mRNA in proteine. Quest'aspetto è tutt'altro che scontato. Le cellule si riservano sempre la possibilità di non tradurre un RNA in proteina o tradurre una parte più piccola o di frammentare la proteina prodotta in pezzi più piccoli con funzione differente. Esiste poi una lunghissima serie di modificazioni post-traduzionali delle proteine, quali fosforilazione e defosforilazione, miristilazione, glicosilazione, etc. che ne cambiano radicalmente le funzioni

TEST GENETICI

Un test genetico è l'analisi di DNA, RNA o cromosomi per individuare alterazioni dovute ad una patologia genetica, ereditabile o meno. Occorre distinguere tra *test clinici* e *test di ricerca*.

I **test clinici** hanno lo scopo di fornire al paziente o alla famiglia una risposta utile per l'inquadramento diagnostico, la prevenzione o un trattamento terapeutico. Ci sono costi e tempistiche dei test clinici molto variabili in funzione della complessità diagnostica. I risultati sono riportati per iscritto con un commento ed un'illustrazione dettagliata delle tecniche impiegate e del significato del referto. I laboratori che forniscono test clinici devono seguire delle procedure approvate e standardizzate, anche

se per la maggior parte delle patologie genetiche, in quanto rare, non esistono riferimenti utili. Le linee guida si riferiscono pertanto ai caratteri generali delle indagini più che ai contenuti tecnici della diagnosi della singola patologia genetica. Un'organizzazione utile è la ripartizione dei test clinici relativi alle patologie più rare tra centri di riferimento in rete tra loro, in modo da assicurare che ciascun centro maturi la necessaria competenza in relazione al test effettuato.

I test clinici possono essere di cinque categorie:

- 1) **test diagnostici**, per confermare o escludere una malattia genetica nota o sospetta in un individuo sintomatico; tali test sono approvati se utili alla diagnosi per individui di qualsiasi età. Spesso sono necessari più test basati su tecniche indipendenti e i risultati possono avere implicazioni riproduttive o psicologiche anche per altri membri della stessa famiglia;
- 2) **test predittivi**, offerti ad individui asintomatici in una famiglia con una storia di una patologia genetica. Possono essere di tipo "presintomatico", come ad esempio nella corea di Huntington o di "predisposizione" come nelle mutazioni di BRCA1. I test predittivi sono indicati solo se la diagnosi precoce consente un qualche intervento atto a ridurre la morbilità o la mortalità. È importante valutare le implicazioni psicologiche del test e far precedere l'identificazione della mutazione specifica della famiglia;
- 3) **test di portatore**, per identificare individui che hanno una mutazione genica a trasmissione recessiva. In genere tale test è offerto a individui con una storia familiare o anche a gruppi etnici con un elevato numero di portatori di uno specifico allele patologico;
- 4) **test prenatali**, effettuati in gravidanza per determinare lo stato di salute del feto. I test di routine sono basati sui prodotti dell'amniocentesi e dei villi coriali (CVS). Tecniche più particolari includono la biopsia placentare, la biopsia cutanea fetale o la cordocentesi. I test prenatali specifici per singole patologie possono essere effettuati solo se si conoscono in anticipo le mutazioni legate alla patologia nella famiglia;
- 5) **test genetico preimpianto**, non consentito in Italia, è effettuato dopo una procedura di fecondazione in vitro. Si basa sull'analisi di una singola cellula da un embrione a 8 cellule. È effettuato in pochi centri specializzati e solo per alcune patologie. È in genere offerto a coppie con alta probabilità di un figlio affetto da una malattia genetica grave.

I **test di ricerca** hanno invece come scopo primario la migliore comprensione di un disordine genetico o lo sviluppo di una nuova tecnologia diagnostica. Per le malattie genetiche rare o rarissime, i test di ricerca sono spesso gli unici strumenti disponibili per la diagnosi e la prevenzione. A volte pochissimi gruppi di ricerca nel

mondo eseguono quel test che necessariamente non può essere standardizzato o proposto come test clinico. In genere, il centro copre parte dei costi con i fondi di ricerca. I test di ricerca non sono soggetti ad una regolamentazione e il ricercatore può rifiutare l'esecuzione del test in relazione agli obiettivi del progetto di ricerca.

LABORATORIO DI GENETICA MEDICA

Il laboratorio di genetica medica ha il compito di:

- valutare la legittimità delle richieste di test genetici diagnostici, con particolare riguardo ai casi di test preclinici o di suscettibilità, anche in riferimento alle linee guida e ai principi bioetici;
- informare esaurientemente gli utenti ed i medici richiedenti, ove non filtrati dalla struttura clinica, sul significato, i limiti l'attendibilità e la specificità del test genetico e le sue implicazioni;
- eseguire i test genetici utilizzando criteri di qualità, riconosciuti a livello nazionale e internazionale, per garantire l'appropriatezza e l'uniformità dei risultati;
- fornire nel referto esaurienti e dettagliate spiegazioni

sull'interpretazione dei risultati che siano comprensibili, anche ai non addetti ai lavori, e uniformati alle raccomandazioni delle Società scientifiche nazionali e internazionali.

Il laboratorio di genetica medica esegue tre differenti tipologie di test genetici:

- 1) analisi genomica generale;
- 2) test diretto di mutazione;
- 3) test di variazioni alleliche.

Analisi genomica generale

L'analisi genomica generale è effettuata nella diagnostica citogenetica. L'analisi del **cariotipo** consente la visualizzazione d'insieme dell'assetto cromosomico (Fig. 2). Tale test è effettuato in diagnostica prenatale o postnatale anche senza una previsione di mutazione.

In diagnostica prenatale le cellule sono prelevate mediante amniocentesi: questa consiste nel prelevare del liquido amniotico dall'utero, dove si trovano cellule fetali (derivanti dalle vie respiratorie e urinarie del feto) che

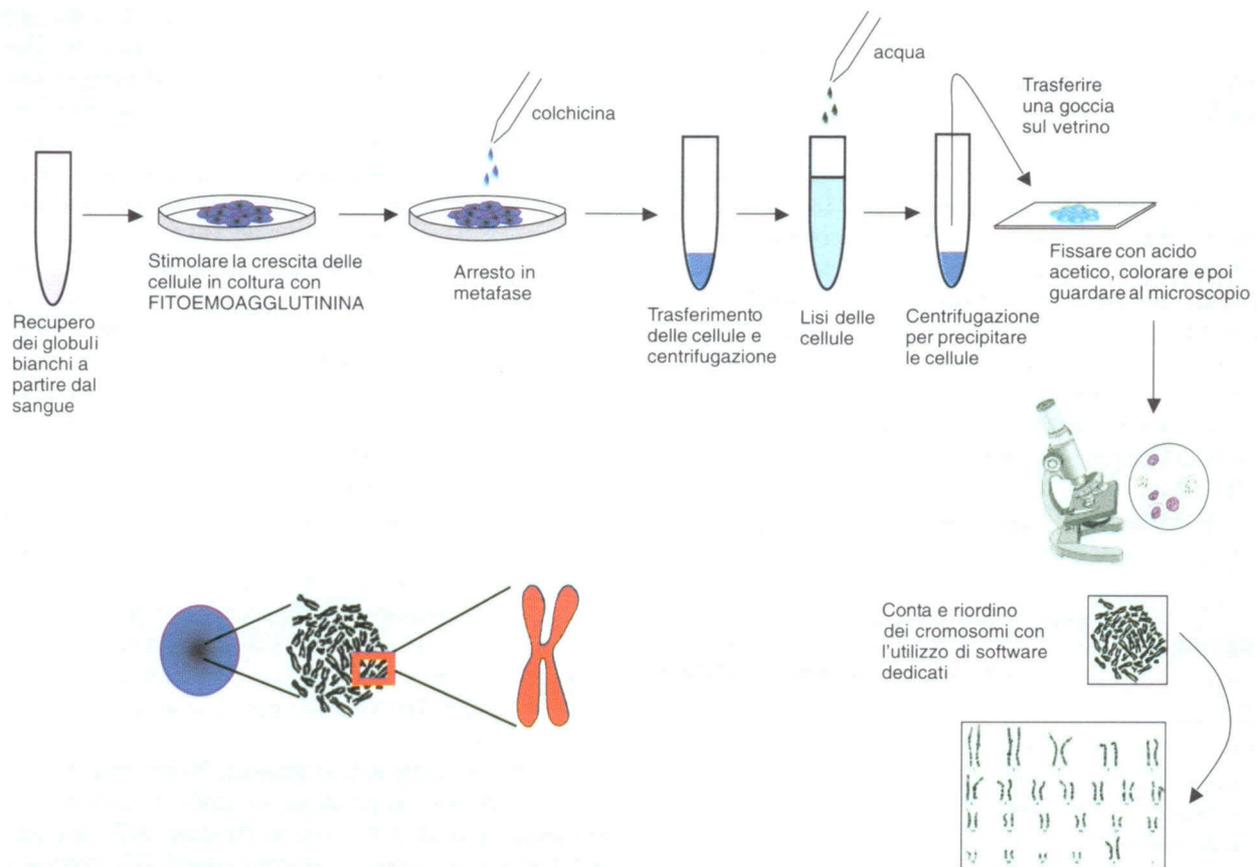


Fig. 2. Analisi del cariotipo.

possono essere anche utilizzate per analisi molecolari. Il prelievo si esegue tra la 15ma e la 18ma settimana di gestazione sotto controllo ecografico per accertare la posizione della placenta, la quantità di liquido amniotico, le dimensioni e la vitalità del feto e la presenza di eventuali malformazioni fetali. Tra le indicazioni principali per l'amniocentesi ci sono: età materna superiore ai 35 anni, un precedente figlio affetto da anomalia cromosomica, una patologia genetica in famiglia, test effettuati dal siero della gestante che indichino un rischio aumentato di un'anomalia cromosomica fetale. Questi ultimi sono **test di screening** (ovvero di valutazione della probabilità di malattia) e non esami diagnostici. Essi forniscono solo indicazioni circa l'opportunità di approfondire, o meno, in donne di età inferiore a 35 anni le indagini su possibili anomalie cromosomiche del feto. Tra i test più noti ci sono la determinazione dell'AFP (alfa-fetoproteina) nel siero materno, il duotest, il tritest e il quadritest. L'AFP è una proteina simile all'albumina prodotta dal fegato fetale. Valori bassi indicano un rischio aumentato di trisomia 21, mentre valori aumentati un rischio di difetti del tubo neurale.

Il **duotest** (double screen) include la valutazione del PAPP-A (Pregnancy Associated Plasma Protein A) e la frazione libera della gonadotropina corionica (free-βHCG). Viene effettuato tra la 10ma e la 13ma settimana di gravidanza dal siero della gestante. Valori di PAPP-A inferiori a 0,5 MoM (multipli della mediana) se combinati a free-βHCG maggiore di 2 MoM, indicano un rischio di trisomia 21. I valori bassi di PAPP-A si associano invece a free-βHCG normale nella sindrome di Turner e minore di 0,5 MoM nelle trisomie 13 e 18 e nella triploidia materna. La free-βHCG invece aumenta considerevolmente nella triploidia paterna. Tale analisi si integra con la valutazione della translucenza nucale (NT = nuchal translucency) fetale a 12-14 settimane che, per via ecografica, rileva l'aumento di liquido nucale nei feti affetti da trisomia 13, 18, 21, da sindrome di Turner e da triploidia paterna. Anche altre cromosomopatie con difetti cardiaci congeniti possono comportare un aumento di NT.

Il **tritest** (triple screen) che si esegue tra la 12ma e la

15ma settimana di gestazione determina l'AFP nel siero materno, l'estriolo non coniugato (uE3) e la βHCG. La valutazione combinata con altri parametri (età materna, peso etc.) consente di individuare le donne con rischio statistico aumentato di partorire un feto affetto da difetti di chiusura del tubo neurale (TND, spina bifida), da trisomia 21 o da trisomia 18. Con valori bassi di AFP e di uE3 associati a valori molto alti di βHCG è molto probabile che il bambino sia affetto da sindrome di Down e, conseguentemente, può essere opportuno eseguire l'amniocentesi. Se, invece, ci sono valori di AFP, di free-βHCG e di uE3 inferiori almeno alla metà di quelli normali, il feto potrebbe essere portatore della sindrome di Edwards, malattia ad esito infausto caratterizzata dalla trisomia del cromosoma 18.

Il **quadritest** (quad screen) è effettuabile più tardivamente tra la 15ma e la 18ma settimana. Aggiunge al tritest la misurazione dei livelli di inibina.

L'abortività ripetuta non è indicazione alla amniocentesi, a meno che uno dei due genitori non sia portatore di una traslocazione bilanciata, possibile causa degli aborti. È universalmente attribuito un rischio di abortività legata alla sola procedura pari allo 0,5% (1:200), rischio che aumenterebbe se il prelievo è effettuato prima della 15ma settimana. La negatività al test, tuttavia, non significa normalità cromosomica o assenza di difetti genetici. Il limite è rappresentato dalla risoluzione del bandeggio che dipende dalla tecnica di preparazione e colorazione dei cromosomi. Con un bandeggio Giemsa o bandeggio G, che costituisce la colorazione di routine, si possono distintamente identificare i cromosomi e alla risoluzione di 400-550 bande e si possono rilevare riarrangiamenti strutturali maggiori di 5 Mbp (Fig. 3).

La **FISH** (fluorescence *in situ* hybridization) rappresenta una tecnica per l'approfondimento e la verifica diagnostica del cariotipo. Ai cromosomi parzialmente denaturati è fatta ibridare una sonda fluorescente. Tale sonda è costituita da una specifica sequenza di DNA realizzata *in vitro* con l'incorporazione di nucleotidi fluorescenti. Dopo la separazione delle eliche, la sonda si appaierà solo sul tratto di cromosoma contenente la sequenza di DNA complementare ad essa. Solo quel tratto di cromosoma emetterà un segnale fluorescente. Alterazioni nella posizione del segnale rispetto a quella attesa in un corredo cromosomico normale indicheranno un possibile riarrangiamento. Esistono sonde FISH che ibridano con le regioni centromeriche o con le regioni subtelomeriche o che sono specifiche di un solo gene o di un intero cromosoma (Fig. 4).

La **CGH** (comparative genomic hybridization) è un altro esempio di test genomico generale in cui è possibile valutare quantitativamente la presenza delle sequenze di DNA. Le variazioni di quantità valutabili includono le delezioni (in meno) o le duplicazioni e le amplificazioni

TABELLA I. Duotest (double screen)

| | Translucenza nucale | free-βHCG | PAPP-A |
|--------------------|---------------------|-----------|--------|
| Trisomia 21 | ++ | ++ | - |
| Trisomia 13, 18 | +++ | -- | -- |
| S. di Turner | ++++ | +/- | - |
| Triploidia materna | +/- | ---- | ---- |
| Triploidia paterna | +++ | ++++ | +/- |

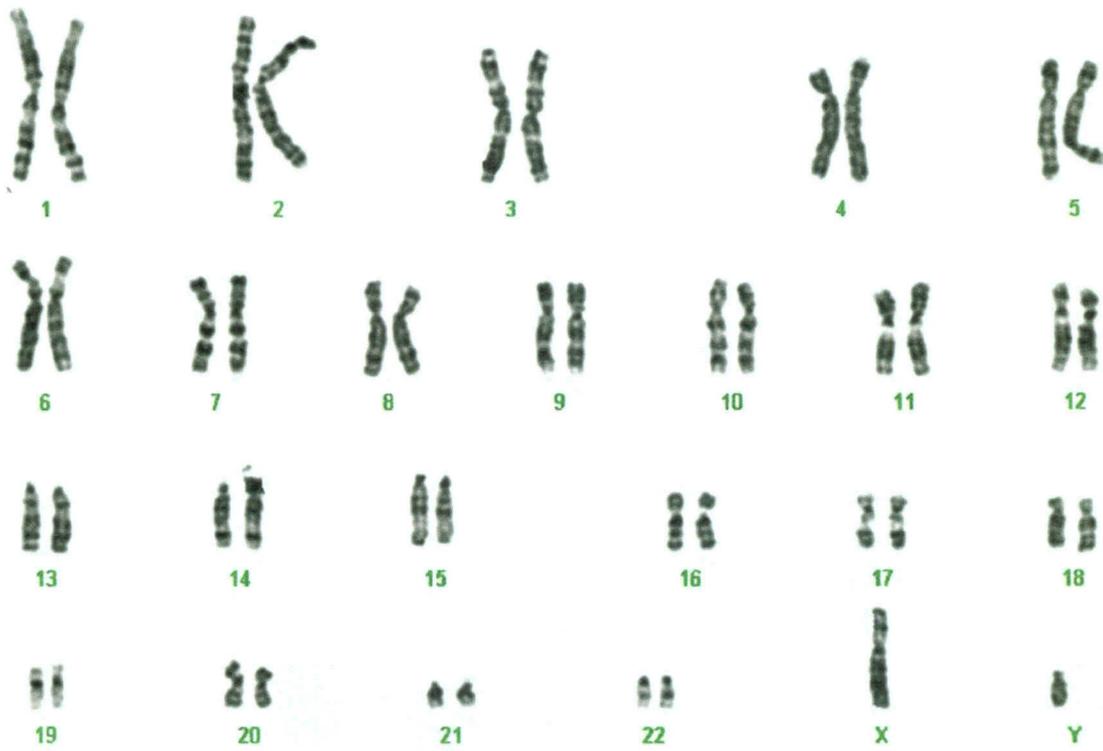


Fig. 3. Bandeggio G.

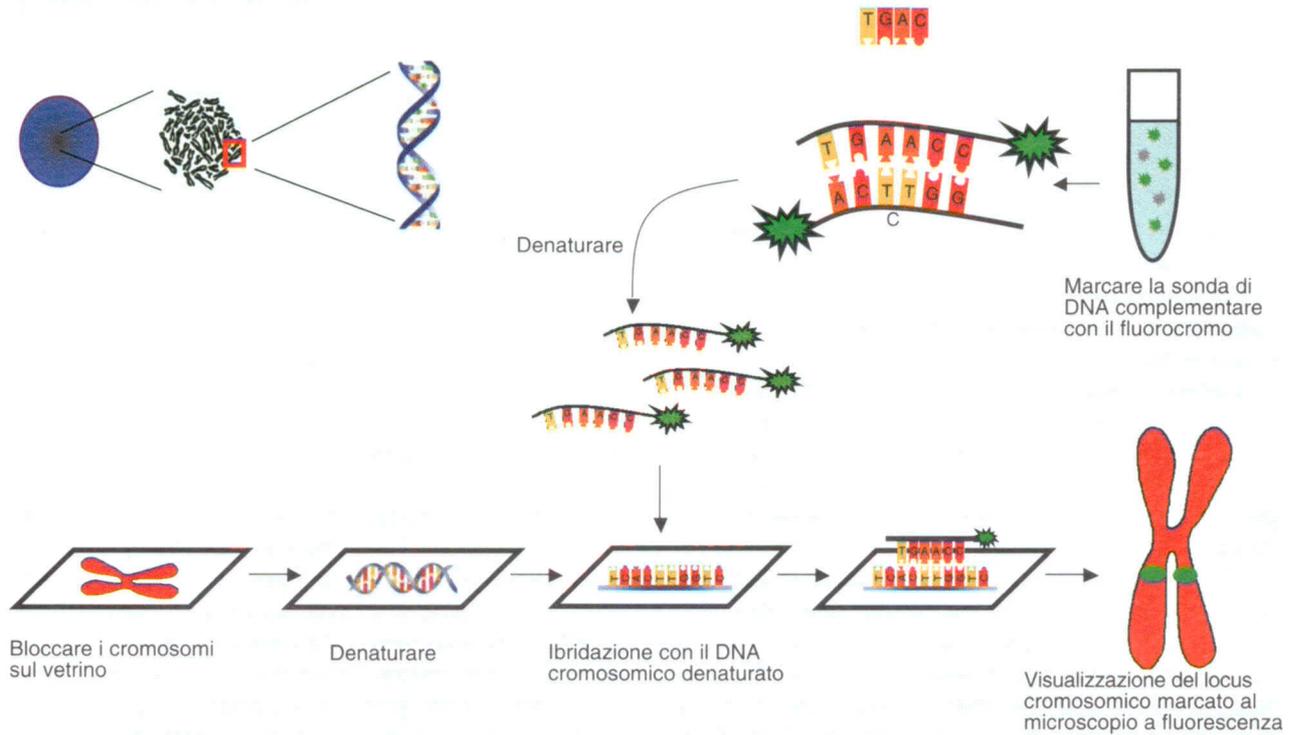


Fig. 4. FISH.

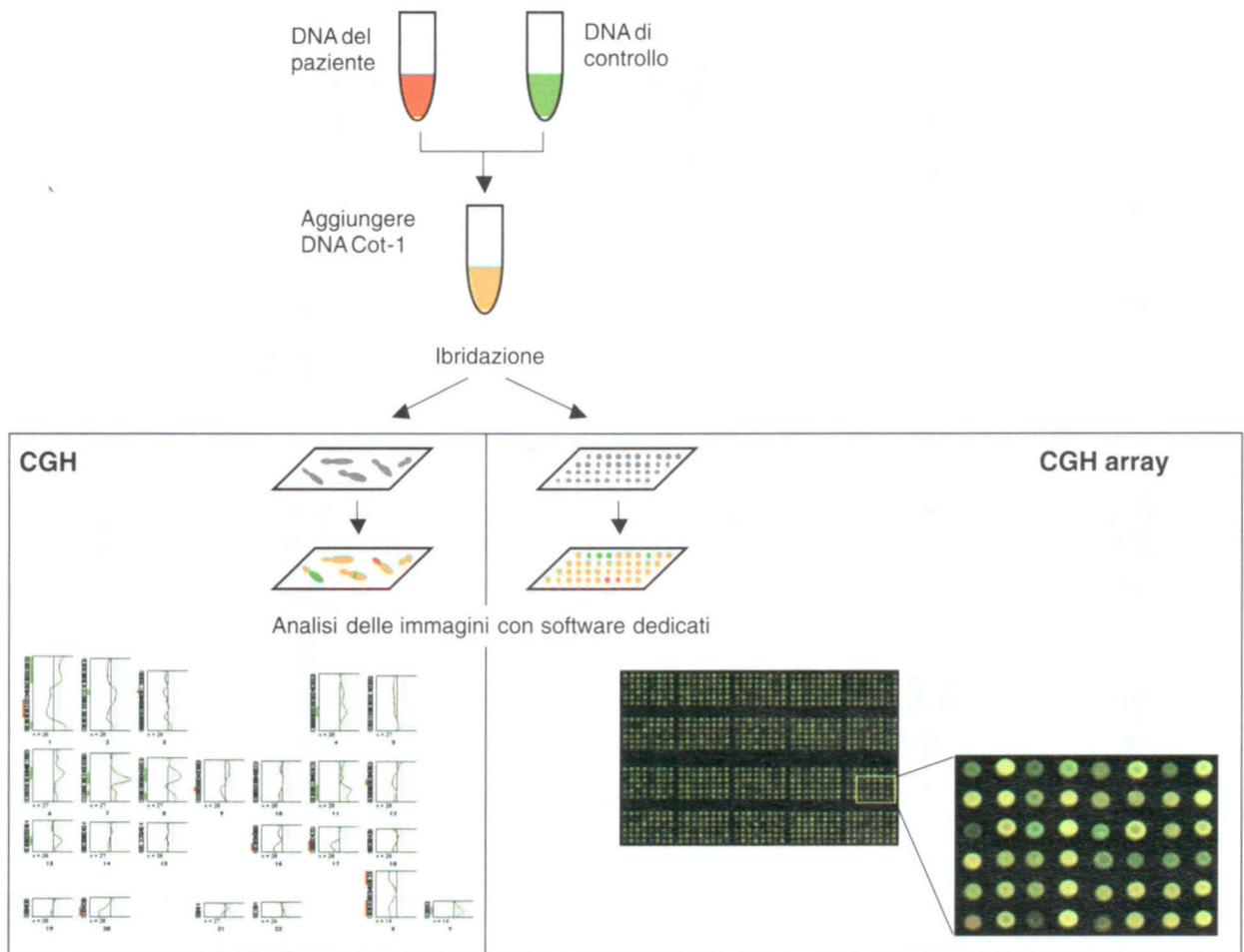


Fig. 5. CGH e CGH array.

geniche (variazioni in più). Il DNA da testare è reso fluorescente con un fluorocromo rosso mentre un campione di DNA di controllo è marcato con un fluorocromo verde. Entrambe le sonde dopo opportuno mascheramento con sequenze umane ripetute (Cot-1) sono fatte ibridare con i cromosomi metafasici, come illustrato per la tecnica FISH. La CGH è utilizzata preferenzialmente per studiare l'assetto cromosomico generale delle cellule cancerose isolate dai pazienti. L'aumentata espressione di oncogeni o la diminuita espressione di oncosoppressori può essere immediatamente individuata (Fig. 5).

La **CGH array** è una tecnica di recente introduzione ed in rapida evoluzione. È una tecnica simile alla CGH, ma priva di cromosomi, sostituiti invece da una moltitudine di sequenze di DNA corrispondenti alle differenti regioni cromosomiche. Il vantaggio rispetto alla CGH è dato dalle enormi potenzialità offerte dalle tecnologie di sintesi del DNA direttamente su chip. Questo consente di rea-

lizzare chip contenenti centinaia di migliaia di differenti sequenze di DNA, ciascuna corrispondente ad una precisa posizione cromosomica. Il vantaggio è immediatamente percepibile: una risoluzione del segnale quantitativo che arriverà a tratti di DNA di pochi nucleotidi (Fig. 5).

Test diretti di mutazioni

La **PCR** (polymerase chain reaction) o reazione a catena della polimerasi è alla base di quasi tutti i test diretti. Ideata nel 1985 dal premio Nobel per la chimica, Kary Mullis, si è rapidamente affermata per la sua apparente semplicità ed economicità. Si basa su un processo di copiatura di un frammento di DNA a partire da due sequenze di innesco (dette primers) ed operata da una DNA polimerasi. Sono necessari cicli di temperatura in cui il DNA si denatura, l'innesco si lega alla sequenza bersa-

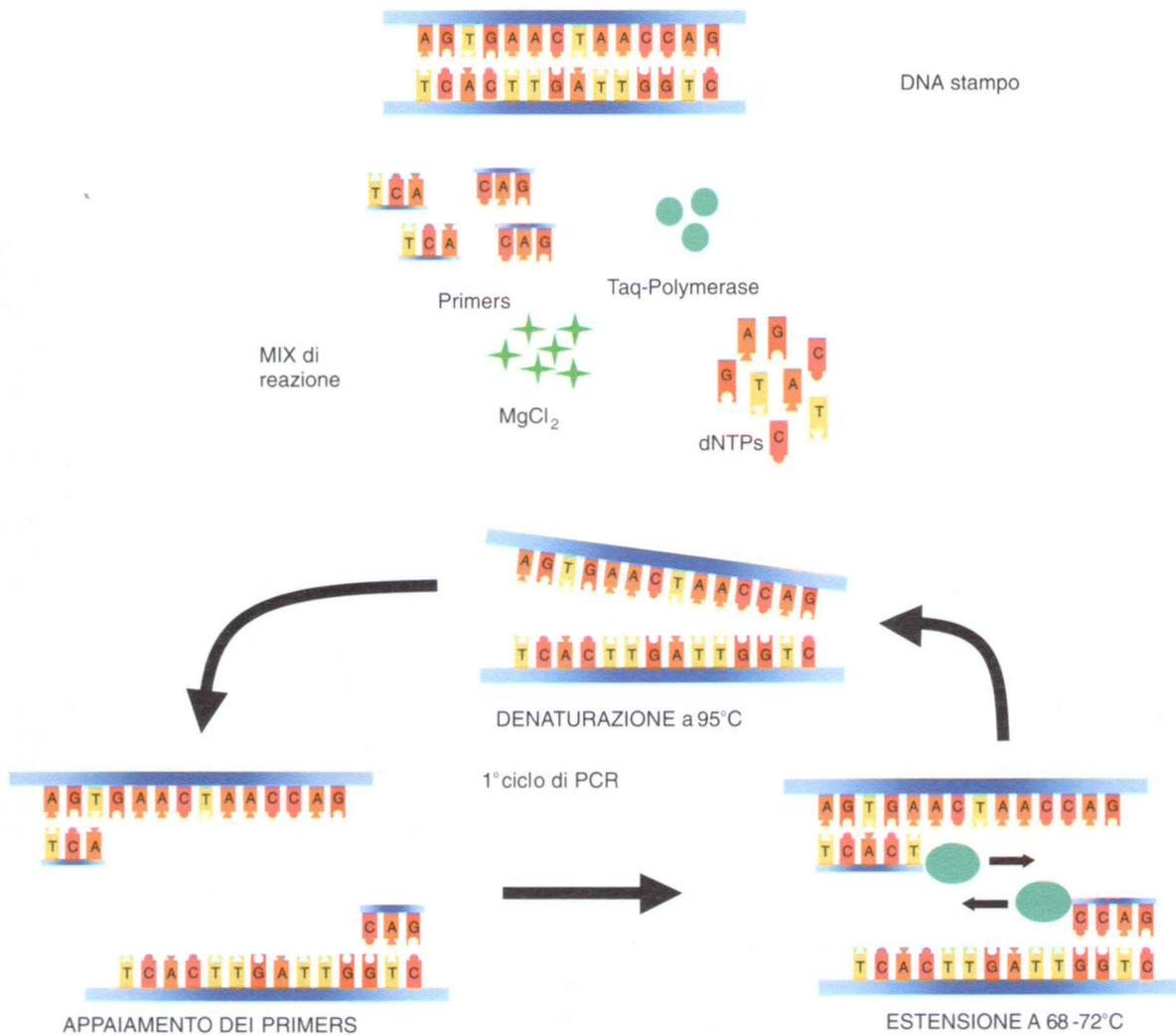


Fig. 6. PCR.

glio ed infine la polimerasi ricopia il DNA a partire dall'innesco (Fig. 6).

Dopo ogni ciclo la quantità di DNA copiata raddoppia, perché ogni copia è stampo di una nuova copia e alla fine del processo da ogni molecola di DNA possono essere generate centinaia di milioni di copie. Tale quantitativo è visualizzabile direttamente mediante fluorescenza su gel di agarosio. Il DNA prodotto può essere studiato mediante numerose tecniche (vedi sotto).

Per i test diretti occorre distinguere tre differenti tipologie:

- test di **mutazioni attese** in un solo gene in alcune specifiche posizioni;
- test di **mutazioni imprevedibili** nella posizione, ma attese in un dato gene (eterogeneità allelica);
- test di **mutazioni con molti geni** potenzialmente responsabili (eterogeneità genica).

Mutazioni attese

Le mutazioni sono attese nel gene in una o poche altre posizioni specifiche. Questo si verifica o in patologie autosomiche dominanti monomorfe, dove c'è una specifica mutazione a determinare il fenotipo clinico, o nelle patologie a trasmissione autosomica recessiva con pochi alleli relativamente frequenti. In tal caso le mutazioni sono presenti in eterozigosi nei portatori sani che sono due/quattro ordini di grandezza più numerosi dei soggetti affetti. Ad esempio, se gli affetti fossero 1:2.500 (q^2) i portatori sani del tratto recessivo sarebbero 1:25 ($2pq$), cioè 100 portatori per ogni affetto, ma se gli affetti fossero 1:1.000.000, cioè 57 in Italia, i portatori sarebbero 1:500, cioè oltre 100.000 in Italia. La tipologia di tali mutazioni è fortemente condizionata dalla provenienza etnica, come ad esempio la mutazione della beta-globina che

causa l'anemia falciforme nella popolazione di origine africana occidentale.

I test di mutazione già nota o attesa condividono alcune tecnologie utilizzate per la genotipizzazione delle varianti alleliche a singola base. Il test **ASO** (allele-specific oligonucleotide) si basa sull'impiego di un oligo sintetico non più lungo di 20nt che ibrida con una sequenza di DNA, ma si dissocia, in particolari condizioni di stringenza, se la stessa sequenza ha una variazione puntiforme. Gli ASO sono utilizzati come sonde marcate per riconoscere le mutazioni più comuni.

Il test **ARMS** (amplification refractory mutation sensitive) è il più semplice e diretto. La PCR determina la produzione di un amplificato solo se i prodotti contengono una sequenza specifica di DNA riconosciuta dall'estremità 3' di una delle sequenze d'innescio. In genere si preferisce introdurre un altro errore di appaiamento (mismatch) alla terzultima base in modo da aumentare la specificità del test. Infatti se il primer ha il mismatch sia in terzultima sia in ultima base non potrà fungere da innescio, mentre con il solo mismatch in terzultima base la PCR partirà ugualmente (Fig. 7).

Se il test di mutazioni attese è negativo in genere si procede con un'analisi più completa (vedi b). Un caso a parte si verifica quando gli individui affetti hanno lo stesso tipo di mutazione (malattia da espansione di triplette), ma non la stessa mutazione.

Mutazioni imprevedibili

Le mutazioni previste nel gene, ma imprevedibili nella posizione costituiscono la maggioranza dei casi di malattie genetiche. I disordini geneticamente eterogenei possono essere causati da molte differenti mutazioni in un singolo gene (eterogeneità allelica). Esempi di eterogeneità allelica sono dati dalla fibrosi cistica e dalle emoglobinopatie. Ad esempio, HbVar (<http://globin.bx.psu.edu/hbvar>) è un database di varianti umane dell'emoglobina e a tutt'oggi include 1.323 differenti mutazioni di cui l'81% sono sostituzioni nucleotidiche e il 13% piccole inserzioni o delezioni. Sono stati sviluppati **chips** che forniscono il rilevamento in parallelo di centinaia di mutazioni note in poche ore con un buon rapporto tra segnale e rumore anche in eterozigoti. Il caso più estremo di eterogeneità allelica è dato da patologie monoalleliche, a trasmissione recessiva X-linked o dominante causate da nuove mutazioni. L'elevata frequenza di nuove mutazioni si associa in genere a una importante presenza di riarrangiamenti maggiori quali delezioni e duplicazioni.

Il **sequenziamento** dei prodotti di PCR è ritenuto il metodo più diretto e lo standard aureo per determinare le mutazioni in caso di eterogeneità allelica. Oggi si basa sempre più sull'impiego di sequenziatori capillari che hanno sostituito il sequenziamento manuale. La tecnica si basa sull'idea di Sanger di ricopiare un'elica di DNA a

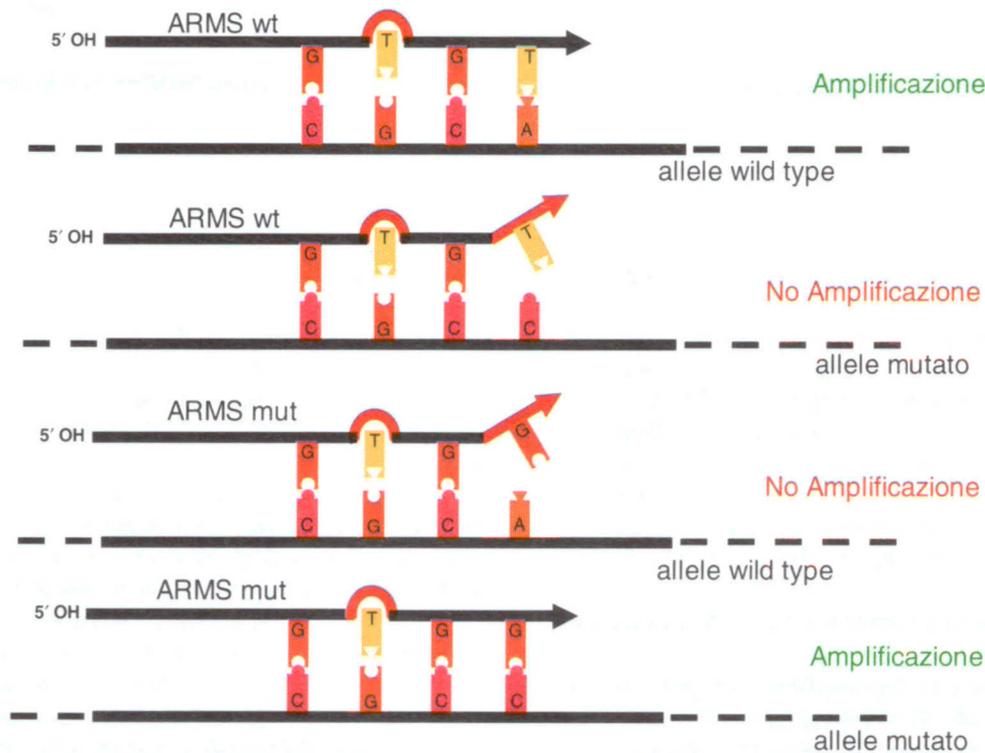


Fig. 7. ARMS.

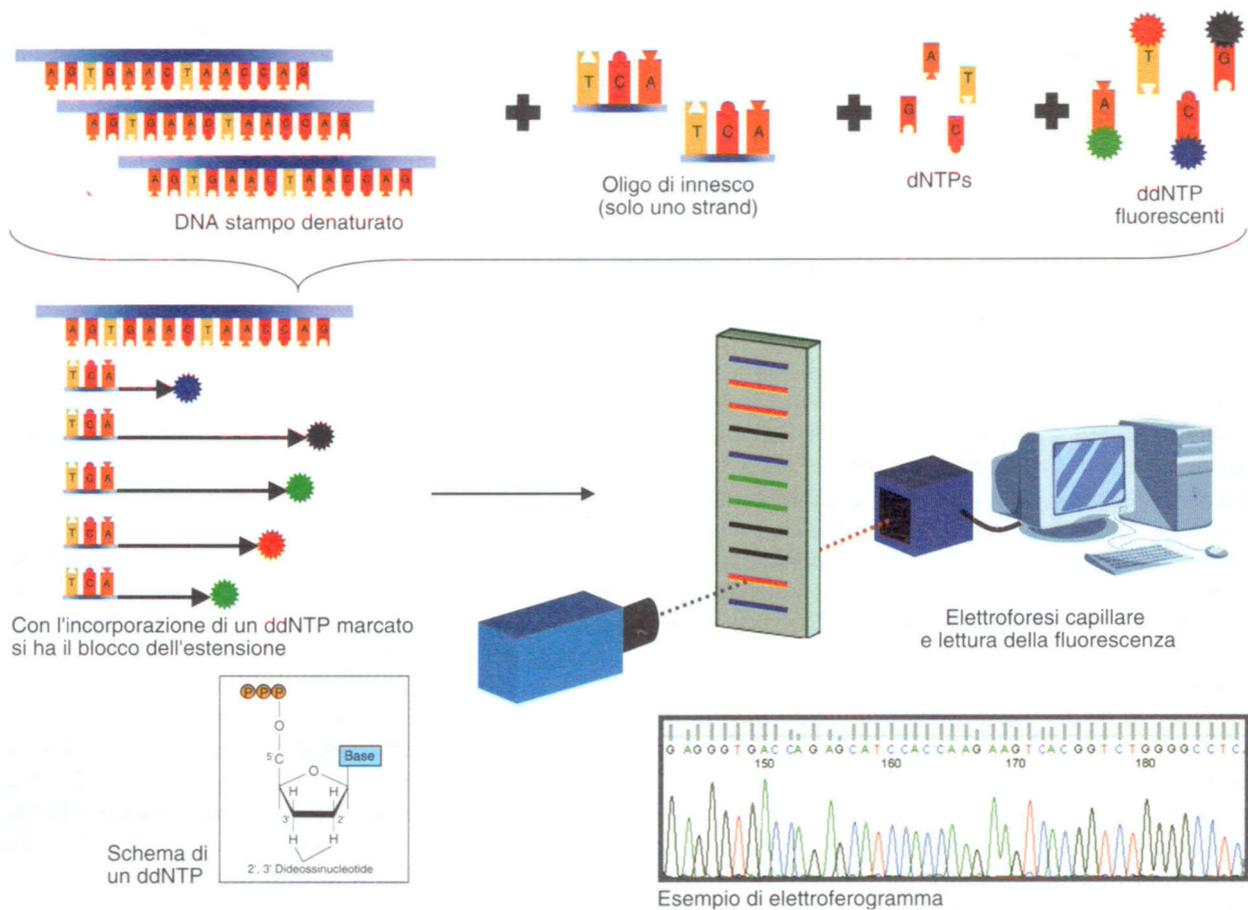


Fig. 8. Sequenziamento del DNA.

partire da uno stampo a singola elica ed un primer d'innesco. A differenza della PCR, il primer è singolo e il DNA viene ricopiato solo a partire dallo stampo originale e non da una copia. L'elegante stratagemma utilizzato da Sanger fu di introdurre, in aggiunta ai deossinucleotidi trifosfato naturali (dNTP), una quantità 10-20 volte inferiore di di-deossinucleotidi (ddNTP). Questi non avendo l'ossigeno in posizione 3' fungono da terminatori. La molecola in estensione si arresta quindi nel momento in cui incorpora un ddNTP e la polimerasi lascia il frammento incompleto. Nel sequenziamento automatico ciascun ddNTP (ddATP, ddTTP, ddCTP, ddGTP) emette una fluorescenza di colore differente: questo permette di stabilire quale base ha interrotto la catena ad ogni data posizione e perciò la successione delle basi in un frammento di DNA (Fig. 8).

Ogni reazione consente la determinazione ottimale di circa 400-600 nucleotidi a partire da 40 nucleotidi a valle dell'innesco. Ci sono in commercio sequenziatori da 1, 4, 16, 96, 192 o 384 capillari per esigenze differenti. La

tecnica è la prima scelta a patto che il gene non sia di dimensioni eccessivamente grandi. I problemi del sequenziamento sono il costo umano e materiale, gli alleli minoritari; i falsi positivi e negativi. Per i costi occorre tenere in considerazione il fatto che essi si sono notevolmente ridotti negli ultimi anni, restando tuttavia in Italia al di sopra di 3,00€ a reazione. Occorre ricordare che è sempre necessario confrontare una reazione di sequenziamento con quella di un controllo normale perché spesso si determinano false basi in corrispondenza di specifiche sequenze ricche di guanine e citosine. È inoltre utile una reazione di sequenza con un innesco di direzione antiparallela, in modo che eventuali errori possano essere corretti. Tutto questo fa lievitare i costi ed i tempi.

Gli alleli minoritari sono quelli presenti solo in una parte delle molecole sequenziate. Il caso più semplice è dato da mutazioni in eterozigosi studiate a partire dall'mRNA retrotrascritto a cDNA. La riduzione quantitativa dell'allele mutato può far visualizzare il solo allele normale, generando un falso negativo. Altri casi sono quello

relativi a mutazioni mosaico o mutazioni presenti nel DNA di un tessuto neoplastico infiltrante un tessuto normale non mutato.

Il sequenziamento può riguardare un prodotto di PCR derivato da una sequenza di DNA genomico oppure da una copia di un mRNA (cDNA) ottenuto mediante retrotrascrizione. Nel primo caso si procede in genere all'analisi dei soli esoni e delle sequenze immediatamente fiancheggianti. Non si visualizzano però le mutazioni introniche profonde che sono situate al di fuori delle sequenze amplificate dalla PCR. L'mRNA può essere di difficile reperimento, se il gene da studiare è espresso in un tessuto non accessibile (occhio, cervello) o durante una limitata finestra temporale durante lo sviluppo in uno specifico citotipo. L'alternativa è studiare l'mRNA da un citotipo differente e derivante da trascrizione ectopica o illegittima. Ad esempio quasi tutti gli RNA dai leucociti possono essere retrotrascritti in cDNA ed amplificati mediante doppia PCR con primers nidificati (nested PCR). I leucociti sono sempre accessibili e l'mRNA può essere isolato mediante semplici procedure o kits in commercio. Per evitare la degradazione dell'RNA il kit stesso contiene inibitori di RNasi. Tuttavia, utilizzando il trascritto ectopico (dai leucociti) invece del trascritto specifico, c'è il rischio di falsi positivi e falsi negativi. Si hanno falsi positivi allorché la tecnica mette in evidenza una mutazione che fa modificare l'mRNA dei soli leucociti, ma l'mRNA del tessuto non è compromesso. Si hanno falsi negativi quando la mutazione lascia intatto l'mRNA dei leucociti ma ha un effetto sensibile nel tessuto coinvolto nella patologia.

Mutazioni con molti geni

Quando un gene ha dimensioni molto grandi ed un organizzazione genomica complessa (>>10 esoni) o una patologia genetica può essere dovuta a mutazioni in geni diversi (eterogeneità genica), il sequenziamento diventa troppo oneroso. In tal caso si utilizzano tecniche di "mutation scanning" in grado di stabilire, con un costo ed un carico di lavoro sostenibili, se il frammento di DNA è identico ad un frammento standard o contiene uno o più nucleotidi diversi. A questo punto saranno sottoposti ad analisi di sequenza non tutti i frammenti, ma solo quelli "sospetti", con un risparmio di sequenze inutili. La prima e più rudimentale tecnica di "mutation scanning" si chiama **SSCP** (single strand conformation polymorphism); essa si basa sulla discriminazione della conformazione tridimensionale che può assumere un frammento di DNA a singola elica, ottenuto dopo una denaturazione a cui segue una corsa elettroforetica non denaturante. La singola elica si ripiega su se stessa in modo sequenza-specifico. Un gel di acrilammide a bassissimo cross-linking (1:200)

consente di effettuare un'analisi comparativa tra le conformazioni e il frammento con variazione nucleotidica. Il frammento che mostrerà un profilo elettroforetico anormale sarà sottoposto a sequenziamento (Fig. 9).

Il limite dell'SSCP è la bassa sensibilità che si aggira intorno al 70% per frammenti fino a 400 bp. Con più corse elettroforetiche in condizioni non denaturanti differenti (aggiunta di 5-10% saccarosio o glicerolo nel gel, pH o varie temperature di corsa) si possono ottenere sensibilità fino al 100%. La tecnica di chiama **DOVAM-S**, ma evidentemente aumenta anche molto il carico di lavoro. Una sensibilità molto maggiore con un tempo di lavoro inferiore è offerta dal **dHPLC**, tecnica utilizzata fin dal 1995 dal gruppo di Cavalli Sforza per la determinazione di polimorfismi del cromosoma Y. **DHPLC** è l'acronimo di denaturing high-performance liquid chromatography. È una tecnica in HPLC a fase inversa in cui una fase stazionaria costituita da microsferette omogenee di polistirene divinilbenzene è in grado di trattenere una molecola polare come il DNA grazie all'azione di un controione, rappresentato dal Trietil Ammonio Acetato (TEAA) (Fig. 10B). La parte polare del TEAA interagisce con il DNA e la parte apolare con la fase stazionaria. Essa si basa sui differenti tempi di eluizione di eteroduplex e omoduplex DNA applicando un gradiente di acetonitrile. Il DNA omoduplex è quello nativo, mentre l'eteroduplex si forma quando sono denaturati e rinaturati nella stessa provetta due frammenti di DNA che hanno una piccola differenza nella sequenza di basi (Fig. 10A). Sono così simili da poter ibridare tra loro, ma in modo imperfetto. Eteroduplex si formano spontaneamente durante la PCR tra i due alleli di un DNA eterozigote, oppure l'operatore fa ibridare nella stessa provetta un DNA controllo con un DNA da testare. La cromatografia deve essere "denaturing", cioè condotta ad una temperatura in genere tra 54 e 68°C, in funzione della sequenza del frammento da analizzare. A predire la temperatura ottimale provvede un opportuno software, ma il metodo empirico a volte è più efficace. L'eteroduplex a questa temperatura si comporta cromatograficamente in modo differente dall'omoduplex, eluisce ad una concentrazione di acetonitrile più bassa indicando così la presenza di una mutazione in un campione. Il vantaggio della procedura è la sua automazione. Ogni campione di DNA è prelevato da piastre con 96 o 384 posizioni ed analizzato in un tempo di 3-5 minuti senza intervento umano.

Test di variazioni alleliche

I test di variazioni alleliche sono identificati come test di genotipizzazione di marcatori genetici. Essi non sono legati alla patologia genetica mendeliana in modo diretto, ma rivestono una grande importanza in due cam-

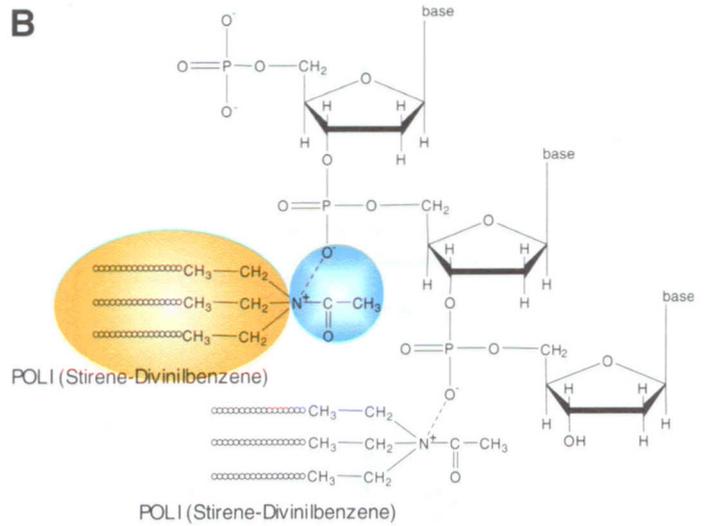
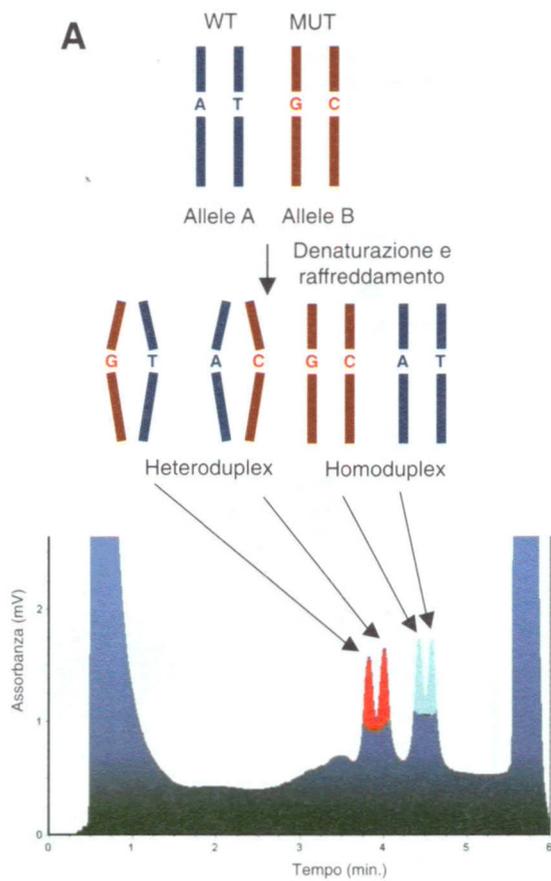


Fig. 10. DHPLC.

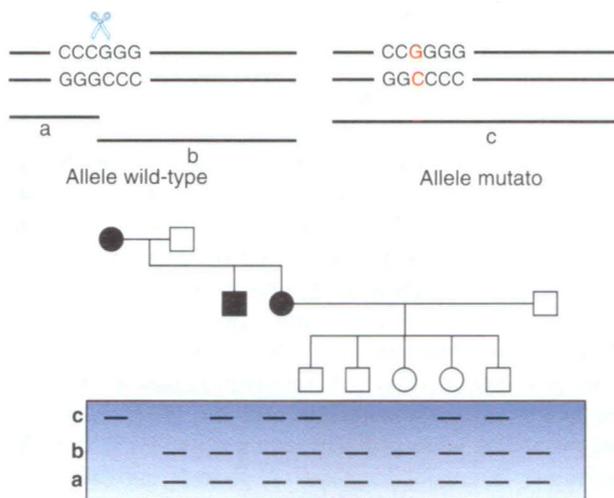


Fig. 11. RFLP.

invece marcatori biallelici dovuti a variazioni di sequenza del DNA. Il singolo SNP è una differenza tra i cromosomi che si osserva mediante il sequenziamento del DNA (Fig. 13).

Esistono oltre 10 milioni di SNPs identificati e inseriti in un database online. Ciascun SNP ha un limitato contenuto informativo e l'allele più frequente è spesso in omozigosi, ma l'utilità nasce dal fatto che si considerano combinazioni di molti SNPs (aploblocchi) che segregano insieme perché adiacenti e mai separati da fenomeni di ricombinazione meiotica. Oggi possono essere determinati quasi tutti gli aploblocchi di un individuo mediante analisi su chip.

Le *CNV* (copy number variations) sono variazioni nel numero di copie di segmenti di DNA identici, più lunghi dei microsatelliti, dovuti a fenomeni di delezione o duplicazione. La tecnica più semplice impiegata per identificare le *CNV* si chiama *MLPA* (multiple ligation probe assay) in cui due oligonucleotidi sono costruiti in modo da ibridare sul DNA l'uno adiacente all'altro. Quando l'i-

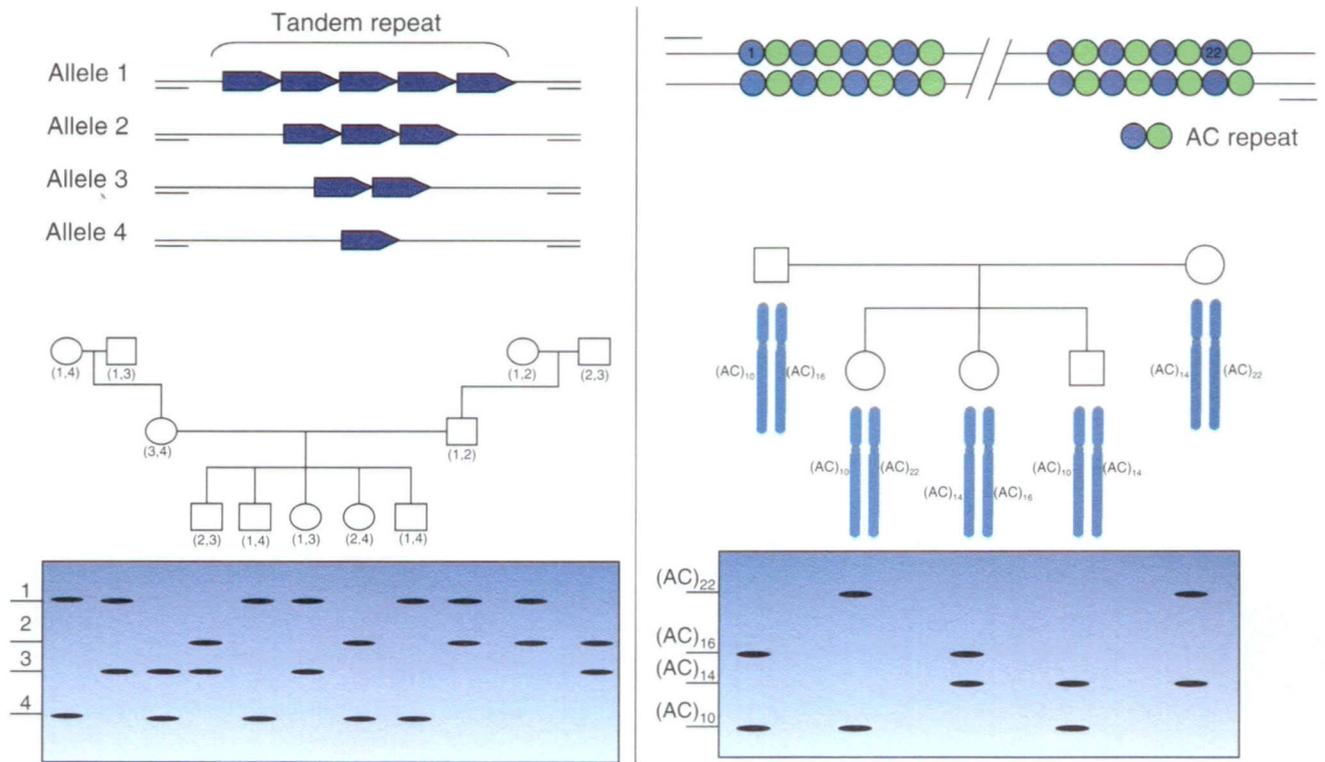


Fig. 12. Microsatelliti.

bridazione è avvenuta una ligasi provvede a saldare gli oligonucleotidi in una sequenza più lunga rilevabile mediante PCR: il prodotto della PCR sarà sempre proporzionale allo stampo di DNA di partenza e quindi al numero di CNV. Un'evoluzione della tecnica si chiama MLPA array (Fig. 14).

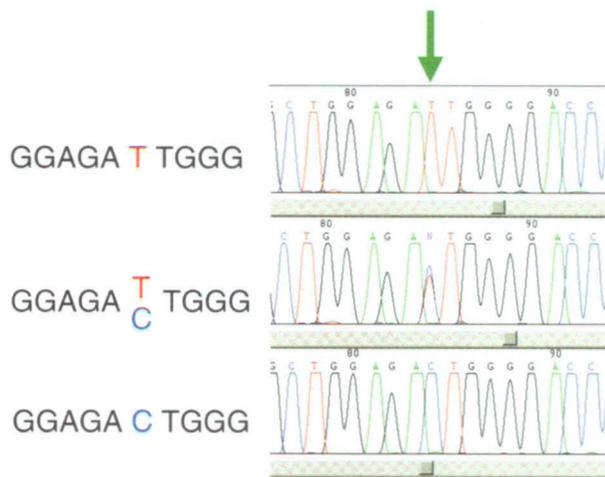


Fig. 13. Single Nucleotide Polymorphism (SNP).

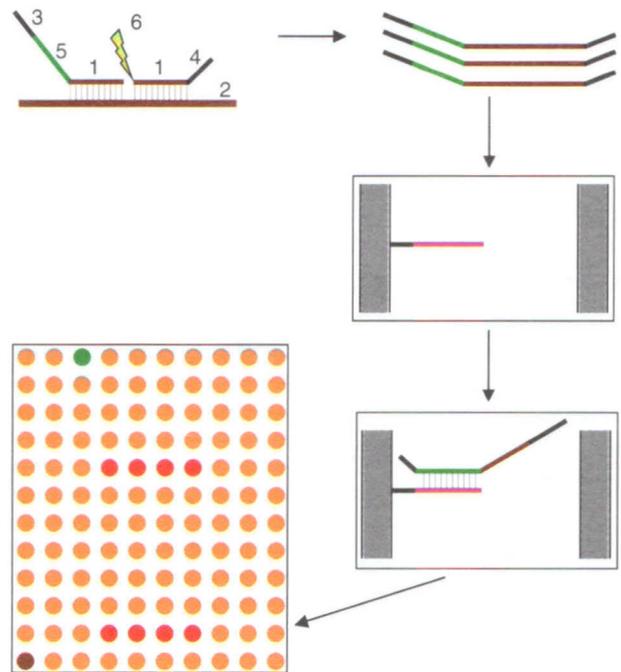


Fig. 14. MLPA-array. Legenda: 1: Oligonucleotide specifico per la sequenza target; 2: DNA genomico denaturato 5 min a 98°C; 3: Oligo Forward per PCR coniugato al 5' con Cy5 (verde); 4: Oligo Reverse per PCR; 5: Sequenza complementare alla sonda dell'array (20 nt); 6: sito di legame.

BIBLIOGRAFIA ESSENZIALE

- Chromosome Abnormalities and Genetic Counseling (Oxford Monographs on Medical Genetics, No. 46) by RJ McKinlay Gardner and Grant R. Sutherland 2003.
- Oxford Desk Reference Clinical Genetics (Oxford Desk Reference Series) by Firth HV, Hurst JA, 2005.
- Smith's Recognizable Patterns of Human Malformation (Hardcover) by Kenneth Lyons Jones and David W. Smith, 6th Edition, Saunders Aug 2005.
- Strachan T, Read AP, Human Molecular Genetics, 3rd Edition, Garland Science 2005.

Thompson & Thompson genetics in medicine, 7th Edition, Nussbaum RL, McInnes RR, Willard HF. Elsevier Health Sciences 2007.

Indirizzi internet utili per la consultazione

www.geneclinics.org
www.orpha.net
www.rarediseases.org
www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM