

Ibridazione di acidi nucleici: principi e applicazioni

CONCETTI CHIAVE

Nei saggi di ibridazione degli acidi nucleici, popolazioni ben caratterizzate di acidi nucleici o oligonucleotidi (sonde) sono utilizzate per identificare sequenze a esse correlate in campioni sperimentali contenenti popolazioni complesse e spesso scarsamente definite di acidi nucleici.

L'ibridazione degli acidi nucleici si basa sulla specificità di appaiamento tra le basi. Le molecole di acido nucleico della sonda e del campione sperimentale sono portate allo stato di singolo filamento e mescolate fra loro per consentire la formazione di molecole eteroduplex tra la sequenza della sonda e qualunque sequenza totalmente o parzialmente complementare a essa (sequenza bersaglio) contenuta nella popolazione del campione sperimentale.

Le molecole eteroduplex sono rilevate mediante la marcatura di una popolazione di acidi nucleici in soluzione acquosa e la sua successiva ibridazione con una seconda popolazione di acidi nucleici non marcati e fissati su un supporto solido. In seguito al lavaggio, che rimuove le molecole di acido nucleico marcato non ibridate sulla sequenza bersaglio, la marcatura rimanente sul supporto solido dovrebbe rilevare le molecole di eteroduplex formate dall'appaiamento della sonda con la sequenza bersaglio.

La stabilità dell'eteroduplex dipende dal grado di complementarietà tra le basi della sonda e quelle della sequenza bersaglio, ed è influenzata da parametri quali la lunghezza del segmento di acido nucleico appaiato e la temperatura e la composizione ionica della soluzione in cui avviene la reazione di ibridazione.

Le sonde a DNA o a RNA comunemente utilizzate si estendono per qualche centinaio di nucleotidi e vengono utilizzate per identificare sequenze bersaglio caratterizzate da un alto grado di similarità di sequenza con la sonda.

Le sonde costituite da corti oligonucleotidi (meno di 20 basi) possono essere utilizzate per discriminare sequenze bersaglio che differiscano tra loro per un singolo nucleotide.

Gli acidi nucleici vengono marcati mediante l'incorporazione di nucleotidi contenenti radioisotopi o gruppi chimici opportunamente modificati per essere facilmente rilevati.

Numerosi saggi di ibridazione prevedono l'immobilizzazione di una popolazione di acidi nucleici contenenti la sequenza bersaglio a un supporto solido e la successiva esposizione a una soluzione contenente la sonda marcata. Il DNA immobilizzato può essere stato precedentemente purificato o può essere presente all'interno di cellule o cromosomi immobilizzati.

L'ibridazione su microarray consente di effettuare diversi saggi di ibridazione simultaneamente. Migliaia di sonde a DNA o oligonucleotidiche non marcate sono fissate a un supporto solido sotto forma di matrici ad alta densità e sono successivamente utilizzate per effettuare lo *screening* di una popolazione complessa di molecole marcate di DNA o RNA in soluzione.

La principale applicazione dell'ibridazione su microarray consiste nell'analisi dei profili di espressione genica e delle variazioni a livello del DNA.

Le molecole di DNA sono molto lunghe e si frammentano facilmente quando vengono isolate dalle cellule, rendendo difficile la loro successiva analisi. Nel Capitolo 6 è stato descritto il modo in cui il clonaggio del DNA, effettuato in cellule viventi o *in vitro*, permette la replicazione selettiva di singole molecole di DNA per ottenerne un alto numero di copie. Se viene amplificato con tali modalità, il DNA risulta a tutti gli effetti purificato: possono essere così studiate specifiche sequenze di DNA genomico o di cDNA. In questo capitolo prenderemo in considerazione un approccio del tutto diverso: invece di cercare di purificare singole sequenze di acidi nucleici, l'obiettivo è di rintracciarle con alta specificità all'interno di una popolazione complessa che rappresenta un campione di interesse medico o biologico.

Nell'ambito della ricerca di base, l'ibridazione di acidi nucleici è spesso utilizzata per rintracciare molecole di RNA e ottenere in questo modo informazioni sull'espressione dei geni di interesse. Questa tecnica può tuttavia essere utilizzata anche per definire la relazione tra sequenze di DNA provenienti da diverse fonti. Per esempio, una particolare sequenza di DNA può essere utilizzata per identificare altre sequenze ad essa strettamente correlate sia in specie diverse, sia in individui diversi appartenenti alla stessa specie, o ancora per identificare sequenze correlate contenute nello stesso genoma cui appartiene la sequenza di DNA di partenza. L'ibridazione degli acidi nucleici è anche utilizzata come metodo per identificare forme alleliche o trascritti anomali associati a malattie genetiche.

Principi dell'ibridazione con acidi nucleici

Nei saggi di ibridazione, una popolazione di molecole note di acido nucleico è utilizzata per vagliare una popolazione di acidi nucleici poco caratterizzati

L'ibridazione degli acidi nucleici rappresenta uno strumento fondamentale per la genetica molecolare. Tale tecnica sfrutta la capacità di acidi nucleici a singolo filamento che presentino complementarità parziale o totale tra loro di formare molecole a doppio filamento mediante l'appaiamento delle basi (**ibridazione**). Per aiutare i lettori privi di particolare dimestichezza con la terminologia sull'argomento, nel Box 7.1 è riportato un elenco dei termini più usati.

L'ibridazione degli acidi nucleici si può effettuare con svariate modalità, ma presenta un principio unificante comune: una popolazione *nota* e ben caratterizzata di molecole di acido nucleico o di oligonucleotidi sintetici viene utilizzata per vagliare una popolazione complessa di acidi nucleici in un campione di interesse medico o biologico scarsamente caratterizzato.

Il campione da analizzare può contenere molecole di DNA, per esempio DNA genomico totale estratto da cellule del sangue di un singolo individuo o da un tipo particolare di cellula tumorale, oppure può contenere RNA, l'RNA totale oppure gli RNA messaggeri espressi da una specifica linea cellulare o tessuto. In ogni caso, per poter prendere parte alla reazione di ibridazione, le molecole di acido nucleico del campione di interesse devono essere convertite allo stato di singolo filamento (**denaturazione**). Il DNA cellulare si trova in natura sotto forma di molecole a doppio filamento, ma anche l'RNA possiede in quantità significative regioni interne a doppio filamento dovute alla formazione di legami idrogeno intrafilamento. Tali legami possono essere dissociati con diversi metodi. La denaturazione iniziale prevede di solito il riscaldamento o il trattamento con soluzioni alcaline, ma se è necessario mantenere lo stato a singolo filamento per lungo tempo, gli acidi nucleici vengono trattati con molecole fortemente polari quali formamide o urea.

La popolazione nota di acidi nucleici da utilizzare per la scansione del campione di interesse è costituita da sequenze di acidi nucleici definite, le qua-

BOX 7.1 VOCABOLARIO PER L'IBRIDAZIONE DI ACIDI NUCLEICI

Appaiamento. Formazione di legami a idrogeno tra due molecole di acido nucleico a singolo filamento. Se due molecole a singolo filamento mostrano una *complementarietà di basi* sufficiente, esse formeranno una molecola di DNA a doppio filamento. Il termine appaiamento ha un significato opposto al termine *denaturazione*.

Complementarietà delle basi. Il grado col quale le sequenze di due molecole di acido nucleico a singolo filamento sono in grado di formare molecole a doppio filamento mediante appaiamento delle basi secondo le regole di Watson-Crick (A appaiata a T o U; C appaiata a G).

Denaturazione. Separazione dei due filamenti di una molecola di DNA mediante la rottura dei legami a idrogeno che li uniscono. Può essere ottenuta mediante riscaldamento o esposizione a soluzioni alcaline o solventi altamente polari, quali urea o formamide. Il termine denaturazione ha pertanto un significato opposto al termine *appaiamento*.

DNA chip. Una qualunque matrice ad alta densità (*microarray*) contenente cloni di DNA o oligonucleotidi e utilizzabile per un saggio di ibridazione.

Eteroduplex. Acido nucleico a doppio filamento formato per appaiamento di basi tra due acidi nucleici a singolo filamento che non derivano dallo stesso allele. Si può avere il 100% di appaiamento delle basi tra le due sequenze di un eteroduplex, in particolare quando una delle due sequenze di ibridazione è una sonda oligonucleotidica oppure quando le due sequenze derivano da alleli dello stesso gene o sono strettamente correlate dal punto di vista evolutivo.

Feature (riferito a un microarray a DNA o a oligonucleotidi). Il numero elevato di molecole di DNA o oligonucleotidi identici che si trovano in una qualunque posizione all'interno del microarray.

Ibridazione in situ. Reazione di ibridazione in cui una sonda marcata viene ibridata con molecole di acido nucleico bersaglio in maniera tale da localizzare queste ultime su un preparato di cellule o cromosomi.

Microarray. Superficie solida sulla quale possono essere immobilizzate, secondo specifiche coordinate, molecole di interesse, in un formato a griglia ad alta densità da utilizzare per particolari analisi. Un microarray a oligonucleotidi o a DNA è costituito da numerose molecole di DNA o di oligonucleotidi non marcate, localizzate in posizioni predefinite sull'*array* per fungere da sonde in un saggio di ibridazione. Ogni specifica

posizione contiene diverse migliaia di copie identiche di un particolare tipo di oligonucleotide o molecola di DNA, e rappresenta una *feature*.

Omoduplex. Molecola di DNA a doppio filamento formata in seguito al riappaiamento tra due sequenze a singolo filamento originate dallo stesso allele. Un omoduplex mostra sempre il 100% di appaiamento delle basi tra le due sequenze ed è pertanto più stabile di un eteroduplex.

Ribosonda. Una sonda a RNA.

RNA antisenso. Una sequenza di RNA complementare a un trascritto a RNA, tale da consentire l'appaiamento delle basi tra le due sequenze.

Saggio di ibridazione. Esperimento in cui una popolazione ben caratterizzata di acidi nucleici o oligonucleotidi (denominata *sonda*) è convertita alla forma di singola elica e successivamente utilizzata per cercare sequenze bersaglio complementari in una popolazione scarsamente caratterizzata di sequenze di acidi nucleici con le quali formare eteroduplex.

Sequenza bersaglio. Sequenza di acido nucleico che mostra una similarità di sequenza con una sonda sufficiente per appaiarsi con essa in un saggio di ibridazione e formare un eteroduplex stabile.

Similarità di sequenza. Il grado col quale due sequenze risultano identiche nella sequenza o nella complementarietà tra le basi.

Sonda (probe). Una popolazione nota di acido nucleico o oligonucleotidica, utilizzata in un saggio di ibridazione per vagliare una popolazione complessa di acidi nucleici allo scopo di identificare, in base alla formazione di eteroduplex, sequenze *bersaglio* correlate.

Stringenza di ibridazione. Il grado col quale le condizioni sperimentali di ibridazione tollerano la presenza di basi non appaiate nella molecola eteroduplex. In condizioni di alta stringenza, solamente le sequenze perfettamente complementari possono appaiarsi, tuttavia anche eteroduplex con un numero significativo di basi non appaiate possono risultare stabili se la stringenza viene attenuata mediante abbassamento della temperatura di appaiamento o aumento della concentrazione salina.

Temperatura di fusione (T_m). Temperatura corrispondente al punto di mezzo nella transizione di una molecola di acido nucleico dalla forma a doppio filamento a quella a singolo filamento.

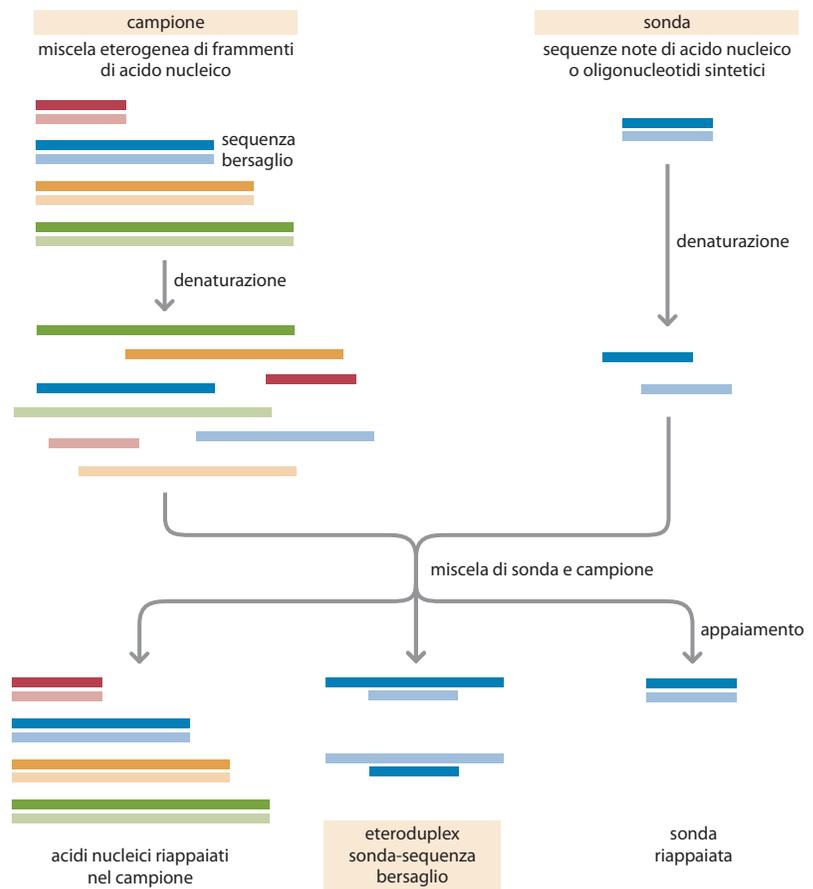
li devono anch'esse essere denaturate, oppure da oligonucleotidi sintetici a singolo filamento. Entrambi i tipi di molecole di questa popolazione fungeranno da **sonde** per localizzare nel campione di interesse molecole di **acido nucleico bersaglio** totalmente o parzialmente complementari a esse. Per fare ciò, il campione di interesse a singolo filamento e la popolazione di molecole che rappresenta la sonda vengono mescolati per consentire l'appaiamento tra i singoli filamenti complementari della sonda e del bersaglio (*annealing*), con l'obiettivo di produrre molecole **eteroduplex** sonda-bersaglio (Figura 7.1); la specificità dell'interazione tra la sonda e le sequenze bersaglio dipende dal grado di complementarietà tra le basi nei due filamenti interagenti.

Gli eteroduplex tra sonda e bersaglio sono rilevabili più agevolmente dopo immobilizzazione su un supporto solido

L'efficienza nella rilevazione degli eteroduplex sonda-bersaglio in soluzione è piuttosto bassa. Per facilitarne l'identificazione, il campione di acidi nucleici in esame oppure la popolazione di molecole della sonda vengono immobilizzati su una qualche forma di supporto solido, molto spesso una membrana di plastica, un vetrino o un chip di quarzo. L'altra popolazione di acido nu-

Figura 7.1 Formazione di eteroduplex sonda-bersaglio in un saggio di ibridazione di acidi nucleici.

Un campione di interesse consistente in una miscela complessa di acidi nucleici e una popolazione predefinita di sonde costituite da sequenze note di acido nucleico o oligonucleotidiche allo stato di singolo filamento vengono miscelati e lasciati appaiare. Le sequenze che in precedenza erano appaiate nel campione di interesse e nella sonda si riappaieranno tra loro per formare omoduplex (in basso a sinistra e a destra). Tuttavia, si formeranno anche nuove molecole di eteroduplex tra la sonda e le sequenze bersaglio che possiedono sequenze totalmente o parzialmente complementari a essa (in basso al centro). Le condizioni di ibridazione possono essere modificate per facilitare la formazione di eteroduplex. In questo modo, le sonde riconoscono e identificano selettivamente acidi nucleici correlati presenti in una popolazione complessa.



leico – la sonda o il campione bersaglio di interesse, rispettivamente – viene invece fornita in soluzione acquosa ed è specificamente marcata mediante il legame con una molecola contenente un particolare radioisotopo oppure un gruppo chimico che può essere facilmente rilevato in qualche modo. La popolazione di acido nucleico marcato viene quindi distribuita sopra il supporto solido in modo che le due popolazioni di acidi nucleici possano interagire e formare eteroduplex.

La Figura 7.2 mostra una di queste possibilità, in cui la sonda è marcata e si trova in soluzione, mentre il campione da saggiare è immobilizzato su un supporto solido. Le molecole di DNA bersaglio nel campione catturano le molecole di sonda marcata che presentano una sequenza totalmente o parzialmente complementare a esse, formando eteroduplex sonda-bersaglio. Alcune molecole della sonda marcata possono anche legare in maniera non specifica il supporto solido o altre molecole di acido nucleico che non rappresentano il bersaglio specifico, ma, dopo l'ibridazione, il supporto solido viene lavato ripetutamente, cosicché la sola marcatura che viene rilevata sul supporto proviene dalle molecole di eteroduplex sonda-bersaglio. Per ottimizzare la formazione di questi eteroduplex, l'acido nucleico bersaglio è solitamente presente in largo eccesso rispetto alla sonda. L'alternativa possibile – utilizzare sonde immobilizzate e marcare gli acidi nucleici nel campione di interesse – sarà considerata più avanti in questo capitolo, quando verrà trattata l'ibridazione su microarray.

La denaturazione e l'appaiamento sono condizionati dalla temperatura, dall'ambiente chimico e dall'estensione dei ponti idrogeno

Quando sequenze totalmente o parzialmente complementari si associano per formare molecole a doppio filamento, il numero di ponti idrogeno che si for-

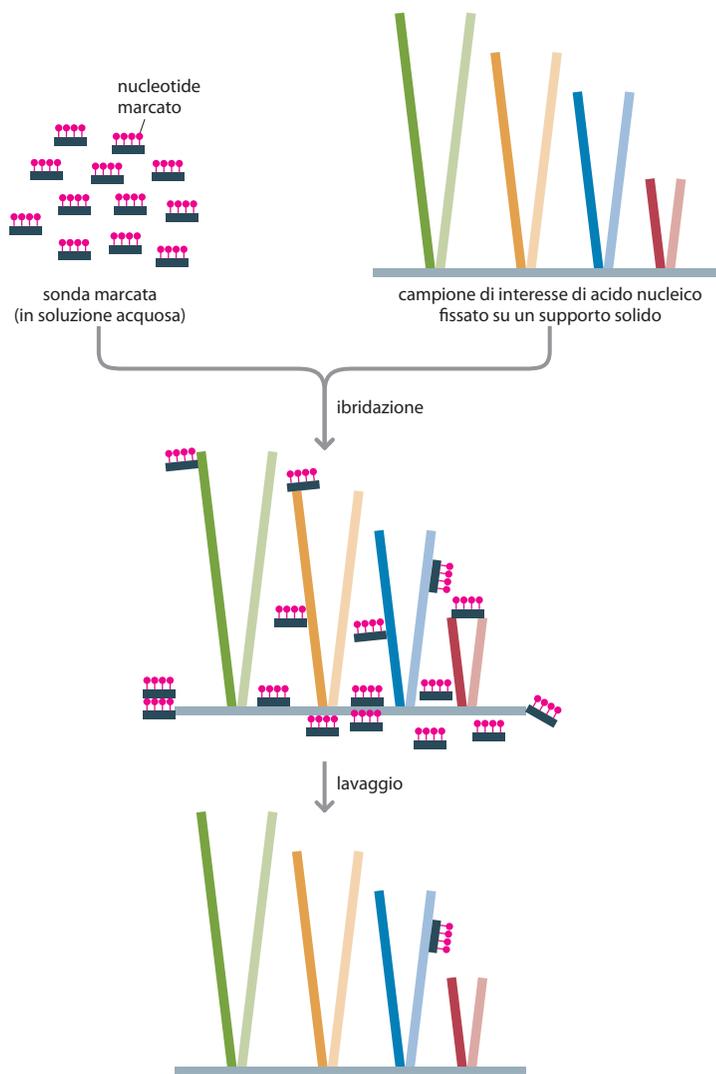


Figura 7.2 Gli eteroduplex sonda-bersaglio sono rilevati con maggior efficienza su un supporto solido.

Di norma, i saggi di ibridazione implicano il legame della popolazione del campione di interesse sulla superficie di un supporto solido (come mostrato in figura) oppure la fissazione sul supporto della popolazione di sonde. La popolazione immobilizzata è legata in modo che sequenze differenti siano depositate in posizioni predefinite sulla superficie solida, ma non necessariamente ancorate a una estremità come mostrato in figura. L'altra popolazione è marcata e successivamente esposta alla popolazione immobilizzata. Qui utilizziamo l'esempio mostrato in Figura 7.1, in cui si presume che la sonda sia omogenea e il campione di interesse sia una miscela eterogenea di molecole di DNA differenti. Per semplicità, è mostrato solo uno dei filamenti marcati della sonda. Le molecole di sonda marcata a singolo filamento sono distribuite sopra il supporto solido e possono essere catturate in seguito all'ibridazione con le molecole bersaglio complementari legate al supporto. Dopo l'ibridazione, il lavaggio rimuove ogni eccesso di sonda rimasta in soluzione o legata non specificamente alla superficie del supporto solido. La marcatura rimanente su quest'ultimo rappresenta quindi gli eteroduplex sonda-bersaglio e permette l'identificazione della sequenza bersaglio.

mano dipende dalla lunghezza della regione eteroduplex. Le molecole di acido nucleico lunghe possiedono più ponti idrogeno e di conseguenza è necessaria più energia per spezzare questi legami. L'aumento della stabilità della molecola a doppio filamento non è tuttavia proporzionale alla sua lunghezza: gli effetti dovuti a cambiamenti della lunghezza sono particolarmente evidenti solo per lunghezze limitate. Come descritto nella prossima sezione, il *mismatch* (appaiamento errato) delle basi riduce sensibilmente la stabilità della molecola a doppio filamento.

La composizione in basi e l'ambiente chimico rappresentano ulteriori parametri importanti nel determinare la stabilità del doppio filamento. Un'alta percentuale di coppie di basi GC comporta una maggiore difficoltà nella separazione dei due filamenti, in quanto la coppia GC è legata da tre ponti idrogeno, mentre la coppia AT solamente da due. La presenza di cationi monovalenti (per esempio, Na^+) stabilizza i ponti idrogeno nelle molecole a doppio filamento, mentre molecole fortemente polari (quali formamide e urea) rompono i ponti idrogeno e agiscono pertanto da denaturanti chimici. Anche il progressivo aumento di temperatura rende i ponti idrogeno instabili e alla lunga ne provoca la dissociazione. La temperatura corrispondente al punto intermedio nella transizione dallo stato a doppio filamento a quello a singolo filamento prende il nome di **temperatura di fusione** (T_m). Per i genomi dei mammiferi, che presentano una composizione pari a circa 40% GC,

TABELLA 7.1 EQUAZIONI PER IL CALCOLO DELLA T_m

Ibrido	T_m (°C)
DNA-DNA	$81,5 + 16,6(\log_{10}[\text{Na}^+]^a) + 0,41(\%GC^b) - 500/L^c$
DNA-RNA o RNA-RNA	$79,8 + 18,5(\log_{10}[\text{Na}^+]^a) + 0,58(\%GC^b) + 11,8 (\%GC^b)^2 - 820/L^c$
Oligo-DNA o oligo-RNA ^d	
< 20 nucleotidi	$2 (l_n)$
20-35 nucleotidi	$22 + 1,46(l_n)$

^aAnche per altri cationi monovalenti, ma accurato solo nel range 0,01-0,4 M. ^bAccurato solo per %GC tra 30 e 70%. ^cL, lunghezza del duplex in paia di basi. ^dOligo, oligonucleotide; l_n , lunghezza effettiva del primer = $2 \times$ (numero di basi G+C) + (numero di basi A+T). Per ogni incremento dell'1% di formamide, la T_m diminuisce di circa 0,6 °C e la presenza di urea 6 M diminuisce la T_m di circa 30 °C.

il DNA, in soluzioni tampone il cui pH e concentrazione salina riproducono condizioni fisiologiche, si denatura a una T_m di circa 87 °C. La T_m di un ibrido perfetto formato da sonde a DNA, RNA o oligonucleotidiche può essere determinata utilizzando formule standard (Tabella 7.1).

Condizioni stringenti di ibridazione aumentano la specificità della formazione di molecole a doppio filamento

Quando si effettua una reazione di ibridazione di acidi nucleici, le condizioni di ibridazione sono deliberatamente scelte in maniera da ottimizzare la formazione di eteroduplex, anche se ciò può comportare un certo grado di ibridazione non specifica. Per esempio, la temperatura di ibridazione utilizzata risulta spesso fino a 25 °C inferiore rispetto alla T_m , e quindi le molecole della sonda possono appaiarsi anche a molecole di acido nucleico con una sequenza vagamente correlata, oltre alle molecole bersaglio strettamente complementari. In questi casi, si dice che la **stringenza dell'ibridazione** è bassa.

Dopo aver indotto la formazione di legami forti tra sonda e bersaglio, vengono sovente effettuati ripetuti lavaggi in condizioni progressivamente meno permissive nei confronti delle molecole di eteroduplex con *mismatch* tra le basi. Questo risultato può essere facilmente perseguito aumentando progressivamente la temperatura, oppure diminuendo gradualmente la concentrazione di NaCl nel tampone di lavaggio. Il progressivo aumento della stringenza di ibridazione può rivelare le diverse sequenze bersaglio man mano più correlate alla sequenza della sonda. L'ultimo lavaggio corrisponde a un'alta stringenza di ibridazione, e garantisce la specificità degli eteroduplex.

Gli eteroduplex sonda-bersaglio sono termodinamicamente più stabili quando la regione a doppio filamento contiene un appaiamento perfetto delle basi. La presenza di *mismatch* diminuisce infatti la T_m , e per normali sonde a DNA ogni punto percentuale (1%) di *mismatch* determina un decremento della T_m pari a 1 °C. Questo effetto si attenua tuttavia col progressivo aumento della dimensione della regione di appaiamento, e quindi, se l'estensione totale della regione di complementarietà è relativamente elevata (più di 100 paia di basi) può ancora essere tollerato un considerevole grado di *mismatch*. Al contrario, se la regione di complementarietà è limitata, come è il caso delle sonde costituite da oligonucleotidi (di norma 15-20 nucleotidi), le condizioni di ibridazione possono essere scelte in modo che la presenza di un singolo *mismatch* renda l'eteroduplex instabile (Figura 7.3).

La cinetica della riassociazione del DNA dipende anche dalla sua concentrazione

La velocità con la quale singoli filamenti di DNA complementari si riassociano per formare molecole a doppia elica dipende dalla concentrazione inizia-

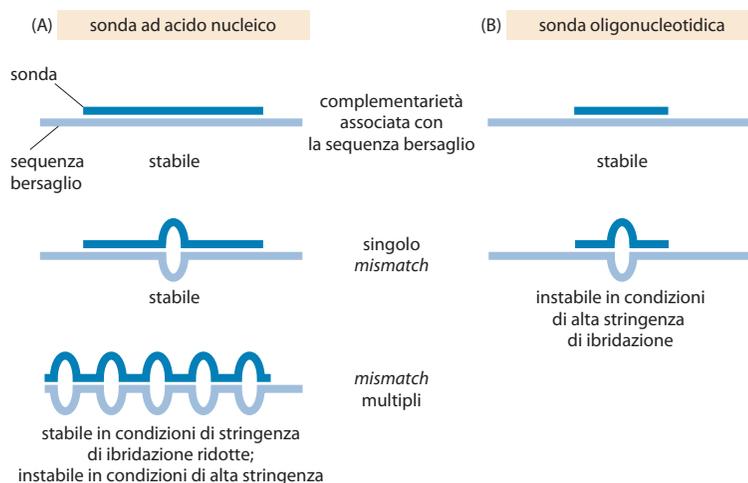


Figura 7.3 Lo sfruttamento dei **mismatch** tra sonda e sequenza bersaglio durante l'ibridazione.

(A) Se l'ibridazione viene effettuata in condizioni di bassa stringenza, sonde abbastanza lunghe (più di 100 nucleotidi) potranno formare eteroduplex stabili con sequenze bersaglio simili, anche se non identiche a esse. Si possono quindi effettuare analisi comparative interspecie e identificare membri di una famiglia genica lontanamente correlati. (B) Le sonde oligonucleotidiche corte sono invece meno tolleranti nei confronti di un appaiamento errato. Effettuando l'ibridazione in condizioni di alta stringenza, è quindi possibile selezionare soltanto i duplex caratterizzati dall'appaiamento completo tra le basi, e in questo modo distinguere anche sequenze alleliche che differiscono tra loro per un singolo nucleotide.

le del campione di DNA. In presenza di un'alta concentrazione di sequenze di DNA complementare, il tempo impiegato da una data molecola di DNA a singolo filamento per localizzare un filamento complementare e formare una regione duplex risulta più basso. La cinetica di riassociazione può essere misurata utilizzando la concentrazione iniziale (C_0) della specifica sequenza di DNA, espressa in moli per litro, e il tempo di reazione (t), espresso in secondi. Tuttavia, il valore C_0t (generalmente noto come "valore cot ") cambia anche in funzione della temperatura di riassociazione e della concentrazione di cationi monovalenti. Si usa pertanto definire dei valori fissi di riferimento, di norma una temperatura di riassociazione pari a 65°C e una concentrazione di NaCl pari a $0,3\text{ M}$.

La frequenza delle sequenze bersaglio in un campione di interesse può variare enormemente. Quando una sonda omogenea viene ibridata con una popolazione eterogenea di acidi nucleici presenti nel campione di interesse, la concentrazione di una qualunque sequenza bersaglio può essere molto bassa, determinando così una bassa velocità di riassociazione. Per esempio, se una sonda per il gene della β -globina viene ibridata con un campione di DNA genomico umano totale, le sequenze bersaglio saranno presenti a una concentrazione molto bassa (il gene della β -globina rappresenta lo $0,00005\%$ del DNA genomico umano). In questi casi, per facilitare la reazione di ibridazione, la quantità di DNA bersaglio deve essere aumentata, e quindi sono richiesti parecchi microgrammi di DNA genomico umano. Viceversa, sequenze che risultano altamente ripetute rappresenteranno una quota cospicua del DNA bersaglio e mostreranno quindi un velocità di riassociazione più elevata.

Anche l'intensità del segnale di ibridazione è proporzionale al numero di copie della sequenza bersaglio presenti nel campione: geni in singola copia produrranno segnali di ibridazione relativamente deboli, mentre sequenze altamente ripetute produrranno segnali molto intensi. Se una particolare sonda è eterogenea e contiene una sequenza a copia singola, per esempio un gene particolare, ma anche numerose sequenze di DNA altamente ripetute, il debole segnale di ibridazione prodotto dalle sequenze a basso numero di copie verrà completamente mascherato dal forte segnale di ibridazione generato dal DNA ripetuto. Questo effetto può essere annullato mediante un passaggio di pre-ibridazione conosciuto come *ibridazione competitiva*. Questo accorgimento consiste nell'aggiungere alla sonda marcata una popolazione di DNA non marcato e arricchito di DNA ripetuto. Questa miscela di acidi nucleici viene denaturata e successivamente lasciata riassociare; in tal modo, gli elementi ripetuti presenti nella sonda marcata vengono efficacemente neutralizzati: gli elementi ripetuti marcati si appaiano infatti alle sequenze ripetute com-

plementari non marcate, e quindi le uniche sequenze della sonda che risulteranno disponibili per la successiva ibridazione sono le sequenze non ripetute.

7.2 Marcatura di acidi nucleici e oligonucleotidi

In un esperimento di ibridazione con acidi nucleici, una determinata popolazione di acidi nucleici oppure di oligonucleotidi viene marcata e successivamente fatta ibridare alle sequenze a essa complementari presenti in una seconda popolazione di acidi nucleici. In questa sezione, ci occuperemo dei principi e dei metodi relativi alla marcatura delle sonde costituite da popolazioni omogenee di sequenze di DNA, RNA o oligonucleotidi. Le reazioni di ibridazione in cui è il campione di acidi nucleici di interesse a essere marcato implicano normalmente la marcatura di popolazioni eterogenee e generalmente complesse di DNA o RNA, e saranno trattate nella Sezione 7.4.

A partire da substrati a DNA, a RNA o oligonucleotidici, possono essere preparate differenti classi di sonde

Quando un saggio di ibridazione prevede l'uso di sonde marcate, il substrato per la reazione di marcatura può essere costituito da DNA, RNA o da un oligonucleotide sintetico, ottenuti mediante diversi metodi. Come già accennato, la sonda marcata deve essere preventivamente denaturata mediante riscaldamento, per separare i filamenti associati formando la doppia elica o per interrompere legami idrogeno intrafilamento.

Le sonde a DNA convenzionali vengono solitamente isolate mediante clonaggio del DNA in sistemi cellulari oppure in seguito ad amplificazione del DNA mediante PCR. Le dimensioni del DNA clonato in cellule viventi possono variare in un ampio intervallo, da 0,1 kb fino a centinaia di kilobasi, mentre il DNA clonato mediante PCR presenta una dimensione spesso inferiore a poche kilobasi. In entrambi i casi, di solito le sonde si trovano inizialmente nello stato a doppio filamento. Esse vengono generalmente marcate mediante incorporazione di dNTP (deossinucleotidi trifosfato) marcati nel corso di una reazione di sintesi di DNA *in vitro*. Poiché traggono generalmente origine da molecole di DNA a doppio filamento, le sonde a DNA a singolo filamento rappresentano in realtà una miscela di sequenze provenienti da entrambi i filamenti e pertanto non sono particolarmente indicate per rilevare specifici trascritti a RNA (in quanto possono identificare simultaneamente trascritti derivati sia dal filamento senso che dal filamento antisenso).

Le sonde a RNA sono prodotte utilizzando molecole di RNA a singolo filamento di dimensioni variabili, da poche centinaia di basi a parecchie kilobasi. L'RNA viene preparato a partire da DNA precedentemente clonato in particolari vettori di espressione plasmidici (vedi Capitolo 6) dotati di una sequenza promotore derivata da un fago. A partire da questo specifico promotore, la corrispondente RNA polimerasi fagica è in grado di trascrivere l'inserito di DNA clonato e produrne delle copie a RNA. Alla fine dunque, poiché la reazione di sintesi di RNA è effettuata utilizzando i quattro rNTP, di cui almeno uno marcato, vengono generate specifiche molecole di trascritto a RNA marcato a partire dall'inserito di DNA clonato. Sonde a RNA *antisense* a singolo filamento sono particolarmente indicate per l'identificazione di specifici trascritti *sense* complementari.

Le sonde oligonucleotidiche sono sintetizzate direttamente allo stato di singolo filamento, presentano una lunghezza molto limitata (di norma 15-50 basi) e sono prodotte per sintesi chimica, non mediante clonaggio. La sintesi comporta l'aggiunta di mononucleotidi, uno alla volta, a un mononucleotide di partenza immobilizzato su un supporto solido. Le sonde oligonucleotidiche vengono di norma marcate mediante incorporazione di un gruppo

marcato all'estremità 5', e possono distinguere anche alleli che differiscono tra loro per un singolo nucleotide (vedi Figura 7.3).

Le sonde di acido nucleico lunghe sono generalmente prodotte mediante incorporazione di nucleotidi marcati nel corso della sintesi di un filamento di DNA

Per particolari applicazioni, gli acidi nucleici vengono marcati mediante l'incorporazione di gruppi marcati all'estremità delle molecole. Gli oligonucleotidi a singolo filamento sono generalmente marcati utilizzando una proteina-chinasi che aggiunge un gruppo fosfato marcato all'estremità 5', mentre frammenti di DNA di lunghezza superiore vengono marcati mediante metodiche alternative, quali l'utilizzo di primer modificati dotati di un gruppo marcato all'estremità 5': nel corso della reazione di PCR, il primer con l'estremità 5' marcata viene incorporato nel prodotto di PCR.

La marcatura all'estremità comporta tuttavia che soltanto uno o pochi gruppi marcati siano inseriti in ogni singola molecola, e quindi essi rappresentano soltanto una piccola frazione della massa complessiva della sonda, specialmente nel caso di lunghe sequenze nucleotidiche. Di conseguenza, il metodo più popolare per la marcatura di DNA o RNA sfrutta la marcatura durante la sintesi di un filamento: si sintetizza un nuovo filamento di DNA o RNA *in vitro* e durante la sintesi vengono incorporati nucleotidi marcati in diverse posizioni lungo l'intera estensione del filamento neo-sintetizzato. Di norma, una DNA o RNA polimerasi utilizza un opportuno substrato a DNA per produrre copie a DNA o RNA, che vengono marcate utilizzando una miscela di marcatura in cui almeno uno dei quattro dNTP o rNTP è marcato.

Marcatura del DNA mediante *nick translation*

Nella *nick translation*, si introducono rotture a singolo filamento (*nicks*) nel DNA per generare estremità esposte 3'-idrossiliche e 5'-fosfato. Un *nick* può essere prodotto utilizzando un'opportuna endonucleasi, quale la deossiribonucleasi pancreatica (DNasi I), e il *nick* così esposto funge successivamente da innesco per l'introduzione di nuovi nucleotidi da parte della DNA polimerasi I di *E. coli*, un enzima a più subunità dotato dell'attività sia di DNA polimerasi sia di esonucleasi 5' → 3'. Contemporaneamente all'aggiunta di nuovi nucleotidi in corrispondenza del lato 3'-idrossilico del *nick* a opera dell'attività DNA-polimerasica, l'attività 5' → 3' esonucleasica rimuove i nucleotidi preesistenti dall'altro lato del *nick*. Di conseguenza, il *nick* sarà progressivamente spostato lungo il DNA in direzione 5' → 3' (Figura 7.4). Se la reazione è effet-

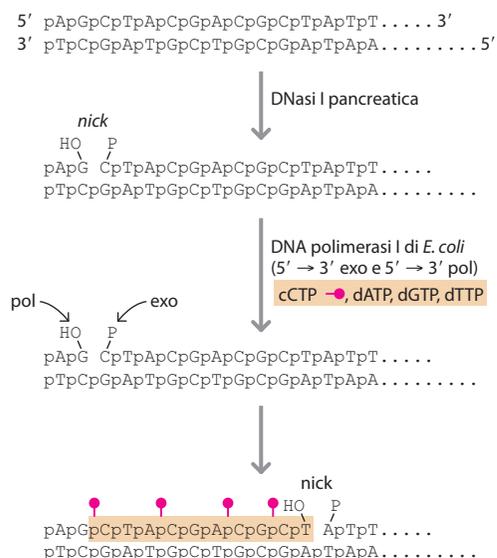


Figura 7.4 Marcatura del DNA per *nick translation*. La DNasi I pancreatica introduce interruzioni a singolo filamento (*nicks*) mediante il taglio di legami fosfodiesterici interni (p), producendo un'estremità 3'-idrossilica e un gruppo 5'-fosfato. Questi *nicks* forniscono il substrato per la DNA polimerasi I di *E. coli*, un enzima a più subunità. Questa proteina possiede due attività enzimatiche: una esonucleasi 5' → 3' (exo) attacca l'estremità 5' esposta dal *nick* e rimuove in sequenza i nucleotidi in direzione 5' → 3'; una DNA polimerasi (pol) aggiunge invece nuovi nucleotidi al gruppo 3' OH esposto, proseguendo in direzione 5' → 3', sostituendo quindi i nucleotidi rimossi dalla esonucleasi e spostando lateralmente il *nick*. Almeno uno dei deossinucleotidi trifosfato (dNTP) è marcato (indicato in figura dal pallino rosso sul CTP), pertanto in questo caso la marcatura viene incorporata ogni volta che una citosina viene incorporata nel filamento di nuova sintesi.

tuata a una temperatura relativamente bassa (circa 15 °C), essa non procede oltre lo stadio di una completa sostituzione della sequenza nucleotidica preesistente. Sebbene non avvenga alcuna sintesi netta di nuovo DNA a queste temperature, la reazione determina l'incorporazione di nucleotidi marcati in sostituzione dei nucleotidi non marcati preesistenti.

Marcatura del DNA mediante *random priming*

Il metodo di marcatura del DNA mediante *random priming* utilizza una miscela complessa di numerosi esanucleotidi differenti, in grado di appaiarsi in maniera casuale alle sequenze esanucleotidiche complementari presenti in uno stampo di DNA e di innescare pertanto la sintesi di un nuovo filamento di DNA (Figura 7.5). La sintesi dei nuovi filamenti di DNA complementare è catalizzata dalla subunità Klenow della DNA polimerasi I di *E. coli* (dotata dell'attività polimerasica, ma priva dell'attività 5' → 3' esonucleasica normalmente associata).

Marcatura per sintesi di un filamento mediante PCR

La classica reazione di PCR può essere modificata in modo da includere uno o più precursori nucleotidici marcati, che vengono pertanto incorporati nel prodotto di PCR in tutta la sua estensione.

Marcatura di RNA

Le sonde a RNA (*riboprobe*s) possono essere prodotte mediante trascrizione *in vitro* di molecole di DNA clonato in un vettore di espressione plasmidico opportuno. I vettori di espressione contengono una sequenza promotore immediatamente adiacente al sito di inserzione del DNA, consentendo a una qualunque sequenza di DNA inserita nel vettore di essere trascritta. Vengono generalmente utilizzati dei forti promotori fagici, come quelli derivati dai fagi SP6, T3 e T7, e vengono successivamente fornite le RNA polimerasi corrispondenti per garantire la trascrizione specifica del DNA clonato. L'uso di ribonucleotidi marcati consente quindi di utilizzare le molecole di RNA sintetizzate come sonde (Figura 7.6).

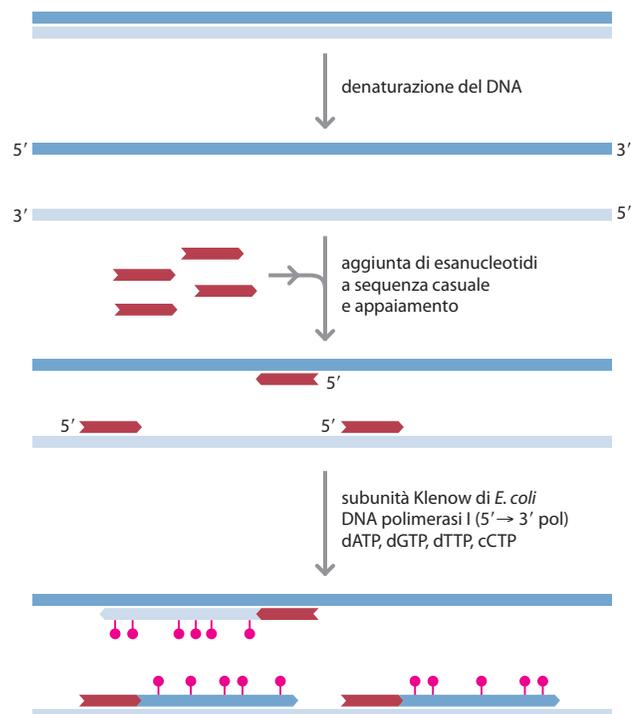


Figura 7.5 Marcatura del DNA per *random priming*. Il DNA a doppio filamento viene denaturato e successivamente viene aggiunta una miscela di esanucleotidi a sequenza casuale. Questi si legheranno al DNA a singolo filamento nei siti a essi complementari, fungendo da innesco per la sintesi di nuovi filamenti di DNA marcati. In questi casi si usa la subunità Klenow della DNA polimerasi I di *E. coli*, che possiede l'attività 5' → 3' polimerasica ma non l'attività esonucleasica.

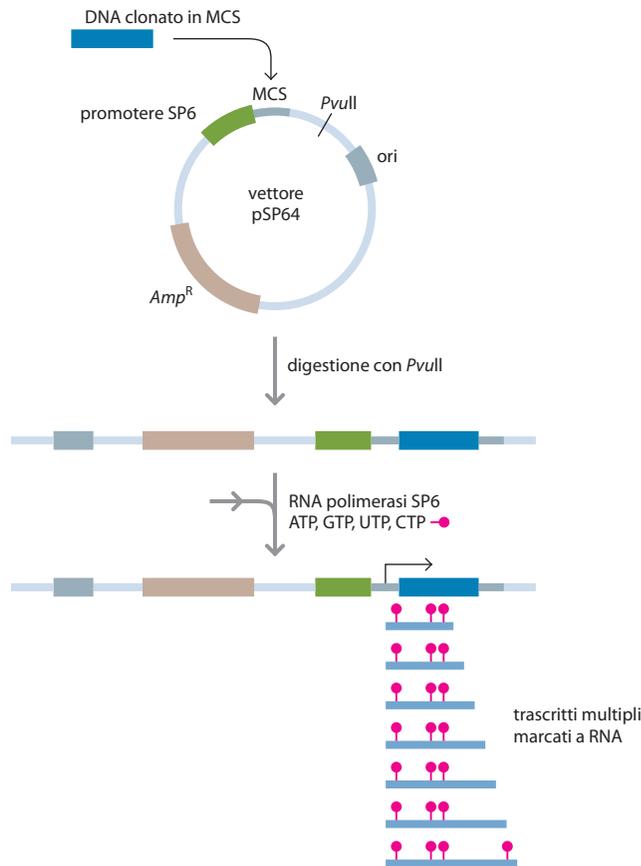


Figura 7.6 Le sonda a RNA sono solitamente prodotte trascrivendo inserti di DNA clonati con una RNA polimerasi fagica. Il vettore plasmidico pSP64 contiene una sequenza promotore per la RNA polimerasi del fago SP6 associata al sito multiplo di clonaggio (MSP), oltre a un'origine di replicazione (ori) e al gene per la resistenza all'ampicillina (Amp^R). Un opportuno frammento di DNA adatto viene clonato nel MCS e successivamente il DNA ricombinante purificato viene linearizzato mediante digestione con l'enzima di restrizione *PvuII*. Si aggiungono poi la RNA polimerasi SP6 e una miscela di NTP, di cui almeno uno marcato, per avviare la trascrizione a livello di uno specifico sito nel promotore SP6 ed estenderla lungo l'inserto di DNA clonato. Si producono così trascritti multipli di RNA marcato corrispondenti all'inserto di DNA. Numerosi vettori di espressione simili sono dotati di due diversi promotori fagici, quali i promotori dei fagi T3 e T7, situati ai due lati del sistema di clonaggio multiplo, e le rispettive RNA polimerasi possono quindi essere utilizzate per generare ribosonde a partire da ciascuno dei due filamenti di DNA clonato.

I radioisotopi possono essere utilizzati per marcare gli acidi nucleici, ma sono pericolosi e hanno una emivita breve

Tradizionalmente, si usava marcare gli acidi nucleici mediante incorporazione di nucleotidi contenenti un radioisotopo che potesse essere rilevato in soluzione o, più comunemente, in un campione solido (**autoradiografia**, Box 7.2). Un autoradiogramma fornisce una rappresentazione bidimensionale della distribuzione della marcatura radioattiva presente nel campione originale. L'intensità del segnale autoradiografico dipende sia dall'energia della radiazione emessa che dalla durata dell'esposizione.

Il radioisotopo ^{32}P è stato largamente utilizzato nei saggi di ibridazione degli acidi nucleici, in quanto emette particelle β ad alta energia facilmente rilevabili. Tuttavia, le stesse particelle si diffondono con alta efficienza, per cui anche il segnale diffonde, e risultano quindi non appropriate per ottenere una risoluzione fine. Radioisotopi alternativi quali il ^{35}S sono maggiormente indicati per tecniche che prevedono lo studio dell'espressione genica in cellule e tessuti, per i quali è richiesta una buona risoluzione a livello morfologico.

I radioisotopi sono facilmente rilevabili, ma rappresentano un rischio per la salute. Inoltre, la radioattività decade col tempo, rendendo necessaria la preparazione di nuove sonde prima di ogni esperimento. Pertanto, attualmente si fa un largo utilizzo di marcature non radioisotopiche basate su specifici gruppi chimici stabili, oltre che facilmente rilevabili.

Nella marcatura non isotopica degli acidi nucleici sono comunemente utilizzati i fluorofori

La marcatura non isotopica degli acidi nucleici prevede l'incorporazione nella sonda di nucleotidi contenenti un gruppo chimico o una molecola che possano essere velocemente e specificamente rilevati mediante un saggio diret-

BOX 7.2 AUTORADIOGRAFIA

L'autoradiografia registra la posizione di un composto marcato radioattivamente presente in un campione solido mediante la produzione di un'immagine su un'emulsione fotografica. Nelle applicazioni in genetica molecolare, i composti radioattivi sono spesso rappresentati da molecole di DNA o proteine, e il campione solido può consistere in cromatina fissata o campioni di tessuto montati su un vetrino. In alternativa, campioni di DNA o proteine sono fissati in un gel da elettroforesi opportunamente disidratato oppure sulla superficie di una membrana di nylon o un filtro di nitrocellulosa. Il campione solido è posto in stretto contatto con una lastra radiografica, un sottile foglio di plastica ricoperto da un'emulsione fotografica. Quest'ultima consiste in una sospensione di cristalli di alogenuro d'argento in una fase gelatinosa trasparente. Nell'attraversare la fase gelatinosa, l'emissione radioattiva del campione converte gli ioni Ag^+ in atomi di argento. La posizione dei cristalli di alogenuro di argento così modificati può essere rilevata nella fase di sviluppo, un processo di amplificazione del segnale in cui i rimanenti ioni Ag^+ inalterati presenti nel cristallo vengono ridotti per produrre argento metallico. I cristalli di alogenuro

di argento inalterati vengono poi rimossi nel processo di fissazione. Le aree scure prodotte sulla lastra fotografica forniscono quindi una rappresentazione bidimensionale della distribuzione del composto radioattivo nel campione originale. L'autoradiografia diretta è più indicata per la rilevazione di radionuclidi che emettono radiazione β a bassa o media intensità (quali ^3H o ^{35}S), mentre particelle ad alta energia (per esempio, quelle emesse dal ^{32}P) attraversano la pellicola disperdendo la maggior parte della loro energia (Tabella 1). Per campioni che emettono radiazione ad alta energia, è quindi necessario introdurre una variazione tecnica, grazie alla quale l'energia emessa viene convertita in luce mediante un'opportuna sostanza chimica (scintillatore o fluoroforo). L'autoradiografia indiretta fa quindi uso di lastre ricoperte di uno scintillatore solido inorganico, da utilizzare come schermo di intensificazione, che vengono posizionate dietro la lastra autoradiografica. Le emissioni che attraversano l'emulsione fotografica vengono quindi assorbite dallo schermo intensificatore e convertite in luce, la quale è anch'essa in grado di ridurre gli ioni Ag^+ e intensificare l'immagine autoradiografica diretta.

TABELLA 1 CARATTERISTICHE DEI RADIOISOTOPI DI USO COMUNE PER LA MARCATURA

Isotopo	Emivita	Tipo di decadimento	Energia di emissione (MeV)	Tempo di esposizione	Adattabilità a studi ad alta risoluzione
^3H	12,4 anni	β^-	0,019	molto lungo	eccellente
^{32}P	14,3 giorni	β^-	1,710	breve	scarsa
^{33}P	25,5 giorni	β^-	0,248	intermedio	intermedia
^{35}S	87,4 giorni	β^-	0,167	intermedio	intermedia

to, nel quale il gruppo stesso fornisce il segnale di marcatura. Spesso si utilizza un **fluoroforo**, un gruppo chimico facilmente rilevabile in quanto capace di assorbire energia a una specifica lunghezza d'onda (**lunghezza d'onda di eccitazione**) e successivamente riemettere l'energia assorbita a una lunghezza d'onda superiore ma ugualmente specifica (**lunghezza d'onda di emissione**) (Box 7.3).

In alternativa, si può utilizzare un saggio indiretto in cui il gruppo chimico incorporato nella sonda funge da **reporter** e viene successivamente riconosciuto e legato con alta specificità da una particolare molecola di affinità, quale per esempio un anticorpo specifico. La molecola di affinità utilizzata è legata a un **marcatore**, costituito da un gruppo chimico o una molecola che possano essere in qualche modo evidenziati (Figura 7.7).

Esistono due principali sistemi per la rilevazione indiretta della marcatura. Il sistema biotina-streptavidina si basa sull'altissima affinità esistente tra due ligandi naturali. La biotina (una vitamina) funge da reporter e viene legata specificamente dalla proteina batterica streptavidina, con una costante di affinità (definita anche costante di dissociazione) pari a 10^{-14} , una delle più alte conosciute in biologia. Le sonde biotinilate possono essere prodotte includendo semplicemente un nucleotide biotinilato nella reazione di marcatura (Figura 7.8). La streptavidina funge successivamente da molecola di affinità. Un altro gruppo reporter largamente utilizzato è rappresentato dalla digossigenina, uno steroide estratto da piante di *Digitalis* (vedi Figura 7.8). In questo caso, come molecola di affinità viene utilizzato uno specifico anticorpo diretto contro la digossigenina.

BOX 7.3 MARCATURA CON ELEMENTI FLUORESCENTI DI ACIDI NUCLEICI

La marcatura con elementi fluorescenti degli acidi nucleici fu sviluppata negli anni 80 e si è rivelata uno strumento estremamente valido in numerose applicazioni, compresa l'ibridazione *in situ* su preparati cromosomici o su tessuti, e il sequenziamento automatico del DNA.

Un **fluoroforo** è un gruppo chimico che se esposto alla luce a una specifica lunghezza d'onda (eccitazione) assorbe energia, e successivamente la riemette a una lunghezza d'onda specifica ma maggiore (Tabella 1 e Figura 1). La marcatura diretta degli acidi nucleici con fluorofori viene effettuata mediante l'incorporazione di un nucleotide modificato (spesso la 2' deossiridina-5' trifosfato) contenente un fluoroforo appropriato.

Vengono sovente utilizzati anche sistemi di marcatura indiretta. In questo caso, il fluoroforo da utilizzare come marcatore viene legato a una molecola di affinità (quale streptavidina oppure un anticorpo anti-digossigenina) in grado di legare specificamente nucleotidi modificati contenenti un gruppo *reporter* (quale biotina o digossigenina) (vedi Figura 7.7).

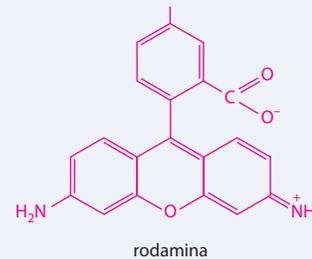
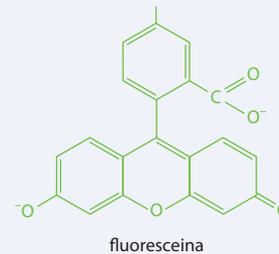


TABELLA 1 FLUOROFORI UTILIZZATI PER LA MARCATURA DEGLI ACIDI NUCLEICI

Fluoroforo	Lunghezza d'onda massima (nm)	
	Eccitazione	Emissione
Blu		
AMCA	350	450
DAPI	358	461
Verde		
FITC	492	520
Fluoresceina (vedi Figura 1)	494	523
Rosso		
CY3	550	570
TRITC	554	575
Rodamina (vedi Figura 1)	570	590
Rosso Texas	596	620
CY5	650	670

AMCA, aminometilcumarina; DAPI, 4',6'-diamino-2-fenilindolo; FITC, fluoresceina isotiocianato; CY3, indocarbocianina; TRITC, tetrametilrodamina isotiocianato; CY5, indocarbocianina.

Rilevamento di acidi nucleici marcati con fluorofori

Gli acidi nucleici marcati con fluorofori possono essere rilevati con appositi scanner a luce laser o mediante microscopia a fluorescenza. I fluorofori sono rilevati dirigendo un fascio di luce prodotto da un'apposita sorgente (un laser ad argon per il sequenziamento automatico del DNA, oppure una lampada a vapori di mercurio nella microscopia in fluorescenza) attraverso un opportuno filtro. Il filtro è strutturato in modo da trasmettere solo la luce alla *lunghezza d'onda di eccitazione* desiderata. Nella microscopia a fluorescenza, questa luce viene riflessa direttamente sul campione marcato e posizionato su un vetrino da microscopio mediante l'uso di uno specchio dicroico, in grado di riflettere la luce di una particolare lunghezza d'onda, ma nello stesso tempo permettere alla luce di altre lunghezze d'onda di attraversarlo (Figura 2). La

Figura 1 Struttura di due fluorofori comuni. Il TRITC e numerosi altri fluorofori sono derivati dalla rodamina.

luce eccita quindi il fluoroforo per generare la fluorescenza; nel fare ciò, il fluoroforo emette luce a una lunghezza d'onda leggermente superiore (*lunghezza d'onda di emissione*). La luce emessa dal fluoroforo torna indietro e attraverso nuovamente lo specchio dicroico, e mediante un apposito filtro raggiunge l'obiettivo del microscopio. Un secondo dispositivo di separazione della luce può inoltre consentire la registrazione della luce emessa mediante una telecamera CCD (*charge-coupled device*).

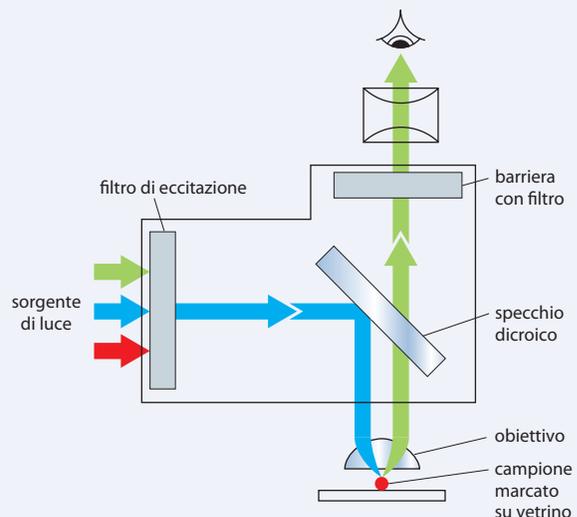
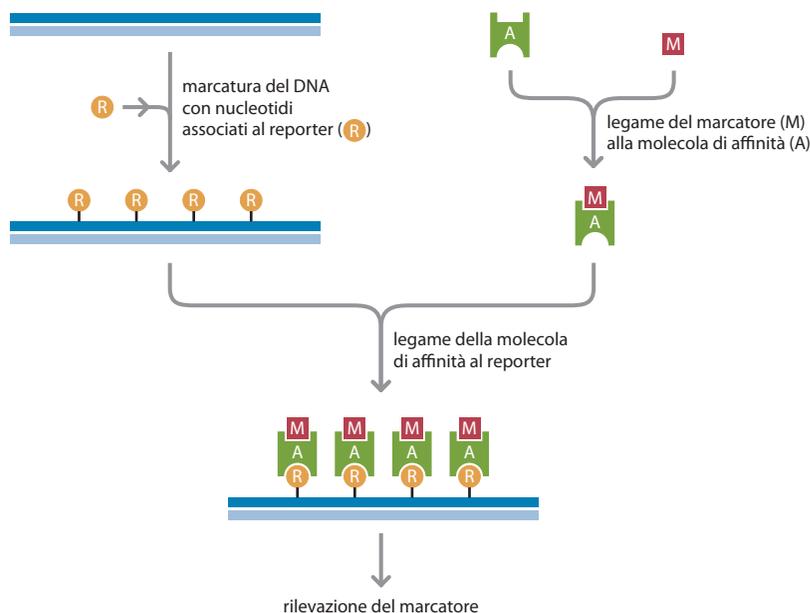


Figura 2 Microscopia in fluorescenza. I filtri di eccitazione consentono il passaggio solo alla luce con una lunghezza d'onda appropriata (in questo esempio, blu). La luce blu trasmessa viene quindi riflessa dallo specchio dicroico (che scompone il fascio luminoso) in direzione del campione di interesse marcato, il quale emette fluorescenza sotto forma di luce ad una lunghezza d'onda superiore (in questo caso, luce verde). Questa luce verde attraversa direttamente sia lo specchio dicroico sia una seconda barriera con filtro, che blocca eventuali segnali fluorescenti non specifici, permettendo di passare attraverso l'obiettivo del microscopio soltanto alla luce verde emessa dal campione.

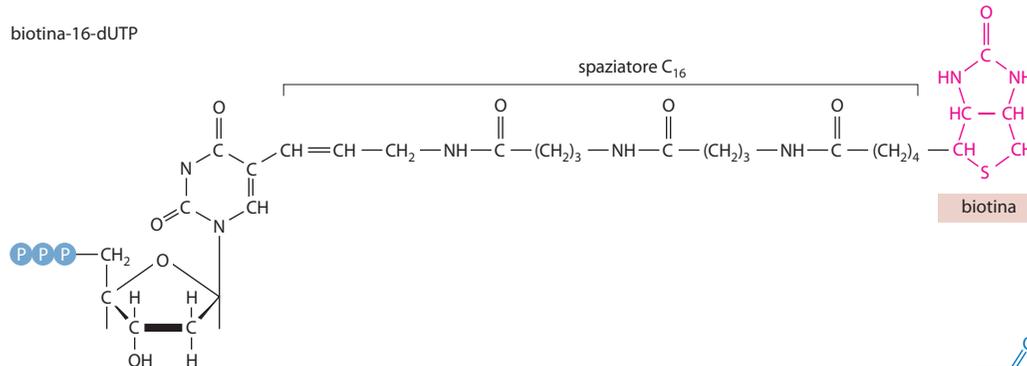
Figura 7.7 Rilevamento indiretto di gruppi marcati negli acidi nucleici.

Gli acidi nucleici possono essere marcati con gruppi chimici che non vengono rilevati direttamente. I gruppi chimici incorporati fungono invece da gruppi *reporter*, che vengono legati con elevata specificità da una molecola di affinità dotata di un marcatore facilmente rilevabile. Il marcatore può essere evidenziato in vari modi. Se è dotato di uno specifico gruppo fluorescente, può essere rilevato mediante microscopia in fluorescenza. Un'alternativa comune prevede l'utilizzo di un enzima quale la fosfatasi alcalina per convertire un substrato in un prodotto colorato che può essere analizzato con tecniche colorimetriche.



Una vasta gamma di molecole o a gruppi marcatori può essere coniugata a molecole di affinità quali streptavidina e anticorpi anti-digossigenina. La lista comprende diversi fluorofori rilevabili tramite microscopia in fluorescenza, oppure enzimi quali la fosfatasi alcalina e la perossidasi, che permettono di rilevare il segnale grazie a saggi colorimetrici o di chemiluminescenza.

biotina-16-dUTP



digossigenina-11-dUTP

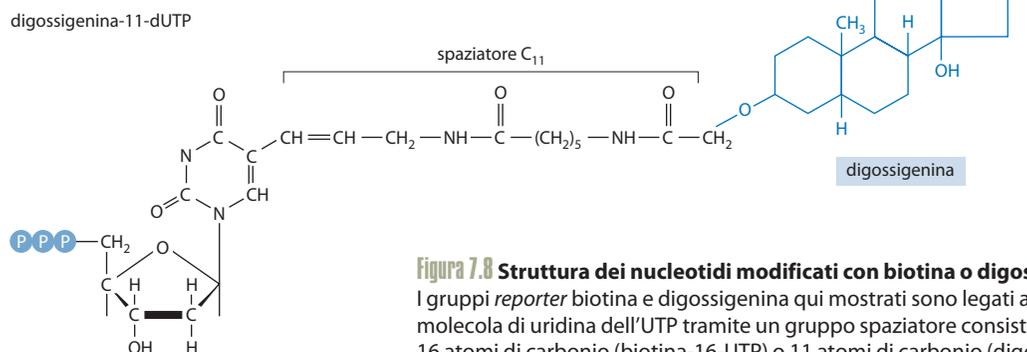


Figura 7.8 Struttura dei nucleotidi modificati con biotina o digossigenina.

I gruppi *reporter* biotina e digossigenina qui mostrati sono legati al carbonio in 5' della molecola di uridina dell'UTP tramite un gruppo spaziatore consistente costituito da 16 atomi di carbonio (biotina-16-UTP) o 11 atomi di carbonio (digossigenina-11-UTP). I gruppi spaziatori sono necessari per garantire la separazione fisica del gruppo *reporter* dall'ossatura dell'acido nucleico, in modo che il gruppo *reporter* possa sporgere a sufficienza da consentire alla molecola di affinità di legarlo efficacemente.

7.3 ibridazione su acidi nucleici bersaglio immobilizzati

Numerose strategie di ibridazione degli acidi nucleici prevedono l'immobilizzazione della popolazione di acidi nucleici bersaglio su un supporto solido e la successiva esposizione di quest'ultimo a una soluzione contenente una sonda marcata. Di norma, la sonda è omogenea e lo scopo consiste nell'identificare sequenze bersaglio all'interno della popolazione complessa di acidi nucleici contenuta in un campione di interesse. La popolazione di acidi nucleici nel campione, purificata in precedenza, può essere depositata direttamente sul supporto solido, oppure frazionata preventivamente sulla base del peso molecolare. In alternativa, possono venire fissate su un supporto solido preparazioni di cellule o di cromosomi, che successivamente sono trattate in modo da rendere il loro DNA o RNA disponibile per l'ibridazione.

In seguito alla formazione di eteroduplex tra la sonda e la molecola bersaglio, l'eccesso di sonda marcata non ibridata viene eliminato, e il supporto viene ulteriormente lavato e asciugato. Dopo questo trattamento, si può ritenere che la marcatura rilevata sul supporto derivi dalle molecole di eteroduplex costituite da un filamento bersaglio fissato sul supporto solido e legato tramite ponti idrogeno al filamento complementare della sonda marcata.

L'ibridazione *dot-blot* permette screening rapidi e utilizza spesso sonde oligonucleotidiche allele-specifiche

Nell'ibridazione *dot-blot*, il campione in esame consiste in una soluzione acquosa di DNA o RNA purificati, per esempio DNA genomico totale umano, che viene semplicemente depositata sopra una membrana di nitrocellulosa o nylon e lasciata asciugare. La variante dello *slot-blot* prevede di depositare il DNA attraverso una singola apertura operata in un apposito stampo. In entrambi i metodi, le sequenze di DNA in esame vengono denaturate prima dell'ibridazione con la sonda marcata.

Un'utile applicazione del *dot-blotting* consente di discriminare alleli che differiscono anche per una singola sostituzione nucleotidica. A tale scopo, si allestiscono sonde oligonucleotidiche allele-specifiche (ASO) a cavallo della sequenza che contiene in sito nucleotidico variante. Le sonde ASO si estendono solitamente per 15-20 nucleotidi e sono normalmente utilizzate in condizioni di ibridazione tali per cui il duplex di DNA tra la sonda e il filamento bersaglio risulti stabile solamente in presenza di una complementarità di basi *assoluta*: un singolo *mismatch* tra la sonda e il filamento bersaglio è sufficiente a rendere instabile la corta regione di eteroduplex. Generalmente si allestiscono due sonde ASO distinte rappresentanti i due alleli in esame, di modo che la differenza di un singolo nucleotide tra i diversi alleli sia localizzata nel segmento centrale della sequenza oligonucleotidica (per rendere massima l'instabilità termodinamica del duplex con un *mismatch*, una regione non correttamente appaiata). Tale distinzione può essere impiegata in una vasta gamma di obiettivi diagnostici e di ricerca, ad esempio per distinguere l'allele dell'anemia falciforme dall'allele normale nel gene della β -globina (Figura 7.9).

Figura 7.9 Il test per la mutazione dell'anemia falciforme mediante ibridazione *dot-blot*. La mutazione che provoca l'anemia falciforme consiste nella sostituzione di un singolo nucleotide ($A \rightarrow T$) nel codone 6 del gene della β -globina, che porta alla sostituzione GAG (Glu) \rightarrow GTG (Val). Sono stati disegnati oligonucleotidi allele-specifici (ASO) in grado di riconoscere la sequenza nucleotidica fiancheggiante la sostituzione: β^A -ASO è specifico per l'allele normale della β -globina e β^S -ASO per l'allele mutante. Lo schema del *dot-blot* illustrato mostra il risultato dell'ibridazione delle due sonde ASO in condizioni di alta stringenza. Con la sonda per l'allele normale (β^A -ASO) il risultato è positivo (cerchio pieno) per individui normali e per gli eterozigoti, ma negativo per individui omozigoti per la mutazione dell'anemia falciforme (cerchio vuoto). Con la sonda per l'allele mutato (β^S -ASO) il risultato è positivo per individui omozigoti per la mutazione e per gli eterozigoti, ma negativo per gli individui normali.

sequenze geniche	allele β^A allele β^S	5' GTG CAC CTG ACT CCT GAG GAG AAG TCT GCC 3' 5' GTG CAC CTG ACT CCT GTG G TG AAG TCT GCC 3'
oligonucleotidi allele-specifici (ASO)		β^A -ASO TG ACT CCT GAG GAG AAG TC β^S -ASO TG ACT CCT GTG G TG AAG TC
<i>dot-blot</i>	sonda β^A -ASO	
	sonda β^S -ASO	
genotipo		$\beta^A\beta^A$ $\beta^A\beta^S$ $\beta^S\beta^S$

Le ibridazioni mediante *Southern* e *northern blot* rilevano DNA e RNA frazionati in base al peso molecolare

Ibridazione mediante *Southern blot*

Un campione di DNA purificato viene digerito con uno o più endonucleasi di restrizione, producendo frammenti di dimensione pari a centinaia o migliaia di paia di basi. I frammenti di restrizione sono successivamente separati in base alla loro dimensione mediante elettroforesi su gel di agarosio (Box 7.4), denaturati e trasferiti su una membrana di nitrocellulosa o di nylon. Una sonda marcata viene quindi ibridata al DNA bersaglio legato alla membrana, e successivamente rilevata per autoradiografia (Figura 7.10).

Un'importante applicazione delle ibridazioni mediante *Southern blot* consente di identificare sequenze bersaglio simili ma non identiche al gene utilizzato come sonda. La sequenza bersaglio può consistere in membri di una famiglia di geni (o più in generale sequenze di DNA) evolutivamente correlati contenuti nello stesso genoma, oppure in un diretto equivalente della sonda contenuto in un altro genoma. Nel secondo caso, si possono analizzare genomi di diversi individui della stessa specie oppure di specie differenti (uno *zooblot*). Nel momento in cui si dimostra che una sonda è correlata ad altre sequenze non ancora caratterizzate, si può successivamente provare a isolare gli altri membri della famiglia mediante lo *screening* di opportune banche a DNA.

Ibridazione per *northern blot*

L'ibridazione *northern blot* rappresenta una variante dell'ibridazione *Southern* in cui il campione di interesse contiene RNA non digerito invece del DNA. Lo scopo principale di questo metodo consiste nell'ottenere informazioni sul profilo di espressione di specifici geni. Una volta clonato un gene

BOX 7.4. ELETTROFORESI DEGLI ACIDI NUCLEICI

Lo scheletro dei legami fosfodiesterici negli acidi nucleici comporta che essi siano dotati di numerosi gruppi fosfato carichi negativamente; pertanto, in presenza di un campo elettrico, gli acidi nucleici migrano in direzione dell'elettrodo carico positivamente. Se la migrazione viene effettuata in un gel poroso, le molecole di acido nucleico possono essere separate in base alla loro dimensione. Ciò avviene in quanto il gel poroso agisce come un setaccio, per cui le molecole di acido nucleico più piccole attraversano agevolmente i pori del gel, mentre quelle di dimensione maggiore vengono rallentate da forze frizionali.

In precedenza, molecole di acido nucleico di dimensione ridotte o intermedie venivano separate utilizzando gel di agarosio, nei quali i singoli campioni vengono caricati in appositi pozzetti in modo che essi possano utilizzare corsie distinte per la migrazione nel gel (Figura 1). Alla fine della corsa elettroforetica, il gel viene colorato con apposite molecole, quali bromuro di etidio o verde SYBR, in grado di legare gli acidi nucleici ed emettere fluorescenza se esposte a radiazione ultravioletta. Per frammenti compresi tra 0,1 e 30 kb, la velocità

di migrazione dipende prevalentemente dalla lunghezza dei frammenti e poco o per nulla dalla composizione in basi. Tuttavia, l'elettroforesi convenzionale su gel di agarosio non è indicata per la separazione di molecole di acido nucleico molto piccole o molto grandi. È possibile modificare la concentrazione di agarosio nel gel per incrementare la risoluzione: diminuendo la concentrazione si può facilitare la separazione di frammenti più grandi, mentre aumentandola si facilita la separazione dei frammenti più piccoli. Per ottenere una risoluzione decisamente migliore per le molecole piccole di acido nucleico si preferisce utilizzare gel di poliaccrilamide, che vengono usati per il sequenziamento del DNA al fine di separare frammenti che differiscono in lunghezza anche per un singolo nucleotide.

Anche frammenti di DNA molto grandi sono separati con scarsa efficienza nell'elettroforesi convenzionale su gel di agarosio. In questi frangenti, si utilizza una speciale apparecchiatura per elettroforesi nella quale viene applicato un campo elettrico discontinuo. Nel corso dell'elettroforesi, la direzione o la polarità del campo elettrico vengono cambiate a intervalli regolari. Il campo elettrico può essere modulato in modo da stabilire un particolare orientamento tra i poli positivo e negativo per un tempo breve, che viene successivamente invertito per un intervallo di tempo altrettanto breve, in un regime di alternanza ricorrente tra i due orientamenti del campo elettrico. Nella *elettroforesi su gel a campo modulato* (*pulsed-field gel electrophoresis*), l'inversione intermittente del campo elettrico forza il DNA a riorientarsi in continuazione per poter migrare nella direzione del nuovo campo elettrico; il tempo impiegato per effettuare il riorientamento dipende dalla lunghezza della molecola di acido nucleico. Pertanto, diventa possibile separare frammenti di DNA di dimensioni fino a parecchie megabasi.

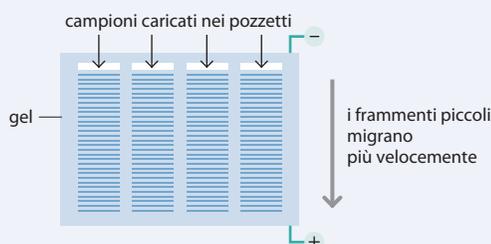


Figura 1 Elettroforesi in gel.

Figura 7.10 Ibridazione mediante Southern blot. Un campione complesso di DNA viene digerito con enzimi di restrizione. I frammenti ottenuti sono successivamente caricati nei pozzetti di un gel di agarosio e separati in base alle loro dimensioni mediante elettroforesi (vedi Box 7.4). Il gel è trattato con soluzioni alcaline per denaturare i frammenti di DNA e successivamente posizionato a contatto di una membrana di nitrocellulosa o nylon. Il DNA sarà trasferito dal gel alla membrana, la quale viene poi imbevuta in una soluzione contenente una sonda radioattiva di DNA a singolo filamento. Dopo l'ibridazione, la membrana viene lavata per rimuovere la sonda in eccesso, asciugata e poi posizionata a contatto con una lastra radiografica: in base alla posizione della sonda marcata si produrrà sulla lastra un'immagine latente, che sviluppando la lastra verrà successivamente rivelata come una banda di ibridazione (vedi Box 7.2).

di interesse, lo si può utilizzare come sonda e ibridare su un northern blot contenente campioni di RNA isolati da tessuti diversi (Figura 7.11). I dati così ottenuti possono fornire informazioni sulla gamma di tipi cellulari nei quali il gene in questione è espresso e sulla abbondanza relativa del trascritto, valutata dall'intensità relativa del segnale autoradiografico prodotto dalla banda di ibridazione. Inoltre, rivelando trascritti di dimensioni diverse, questa tecnica può fornire indicazioni circa l'esistenza di isoforme differenti, dovute per esempio all'uso di promotori, siti di *splicing* o siti di poliadenilazione alternativi.

In un saggio di ibridazione *in situ*, un campione di DNA o RNA è immobilizzato all'interno di preparati cromosomici o cellulari

Alcuni saggi di ibridazione si basano sull'ibridazione diretta con acidi nucleici bersaglio in cromosomi o cellule fissate su vetrini da microscopio. Poiché l'acido nucleico è immobilizzato all'interno di strutture native, si dice che l'ibridazione viene condotta *in situ*.

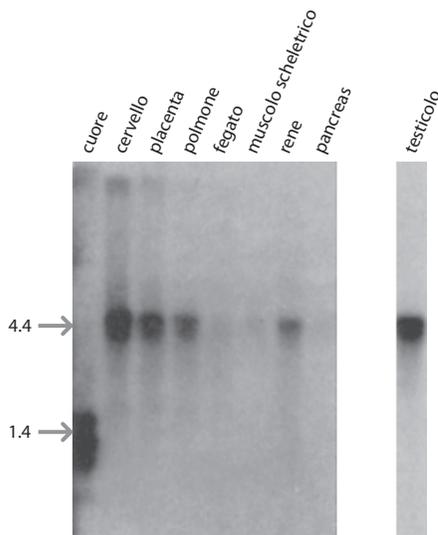
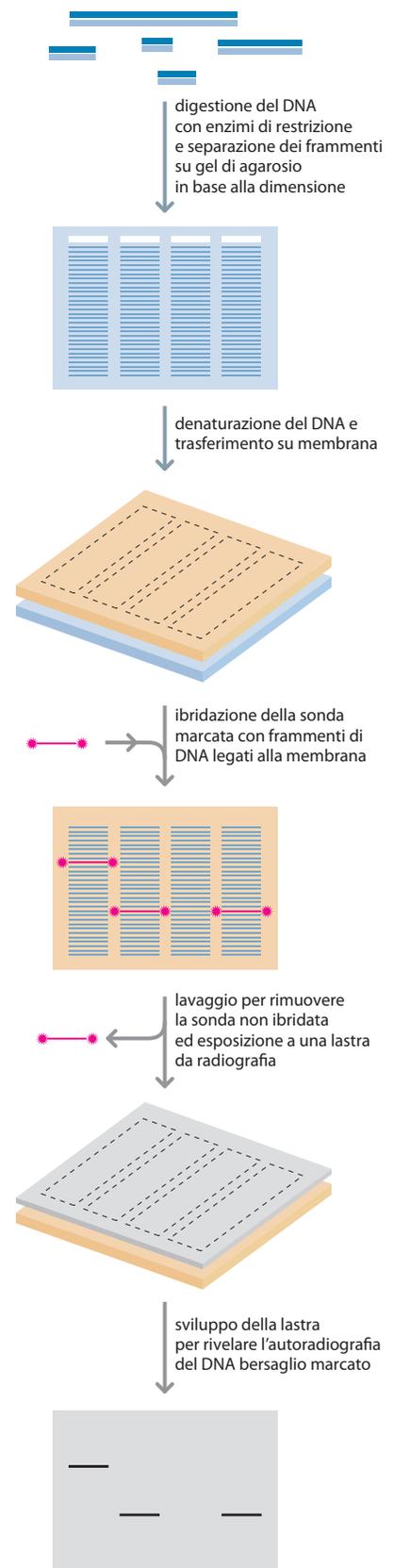


Figura 7.11 Ibridazione per northern blot. Nel northern blot si utilizzano campioni di RNA totale o una frazione purificata di mRNA poly(A)⁺ preparati da tessuti o cellule di interesse. L'RNA viene frazionato in base alla dimensione mediante elettroforesi, trasferito su una membrana e ibridato con una sonda ad acido nucleico opportunamente marcata. Nell'esempio qui mostrato, la sonda è il cDNA marcato del gene *FMR1* (*fragile X mental retardation syndrome*). Il risultato mostra l'espressione comparata di *FMR1* in diversi tessuti: l'espressione maggiore viene osservata nel cervello e nel testicolo (4.4 kb), con livelli di espressione decrescenti nella placenta, nel polmone e nei reni, rispettivamente, ed espressione appena rilevabile nel fegato, nel muscolo scheletrico e nel pancreas. Sebbene il trascritto da 4,4 kb non sia rilevabile nel cuore, si evidenziano in questo tessuto trascritti a basso peso molecolare (circa 1,4 kb) che potrebbero derivare da eventi di *splicing* alternativo o di trascrizione anomala. [Da Hinds HL, Ashley CT, Sutcliffe JS et al. (1993) *Nat. Genet.* 3,36-43. Con autorizzazione di Macmillan Publisher Ltd.]



Ibridazione *in situ* su cromosomi

Una semplice procedura per la mappatura di geni e di altre sequenze di DNA consiste nell'ibridare un'appropriata sonda di DNA marcato su un campione di DNA cromosomico denaturato *in situ*. A tale scopo, si allestisce un preparato cromosomico su un vetrino, utilizzando per lo più cromosomi metafasici o prometafasici da linfociti del sangue periferico o da linee cellulari linfoblastoidi. L'RNA e le proteine vengono rimossi dal campione mediante trattamento con RNasi e proteinasi K e il rimanente DNA cromosomico viene denaturato tramite trattamento con formamide. Il DNA denaturato è quindi disponibile per l'ibridazione *in situ*: una soluzione contenente una sonda marcata viene distribuita sopra il preparato, che è poi coperto con un vetrino coprioggetto.

A seconda della particolare tecnica utilizzata, il bandeggio cromosomico (vedi Capitolo 2) può essere effettuato prima o dopo l'ibridazione. Come risultato, il segnale ottenuto, dopo che l'eccesso di sonda non ibridata è stato rimosso, può essere correlato con il profilo delle bande cromosomiche per identificare la posizione di mappa delle sequenze di DNA riconosciute dalla sonda. L'ibridazione *in situ* su cromosomi è stata rivoluzionata dall'utilizzo di sonde marcate con elementi fluorescenti nella cosiddetta tecnica di ibridazione fluorescente *in situ* (FISH, vedi Capitolo 2).

Ibridazione *in situ* su tessuto

Una sonda marcata può essere anche ibridata sull'RNA presente in sezioni di tessuto. Sezioni di tessuto molto sottili vengono tagliate con un criostato a partire da blocchi di tessuto in paraffina oppure da tessuto congelato, e successivamente montate su un vetrino sul quale si applica una miscela di ibridazione, coprendo poi il tutto con un vetrino coprioggetto.

Sonde complementari di RNA a singolo filamento (cRNA) rappresentano il reagente preferito per questo tipo di indagine. Poiché esse devono risultare complementari all'RNA messaggero di un gene, queste ribosonde antisenso sono ottenute clonando un gene in orientamento inverso all'interno di un vettore appropriato, come per esempio pSP64 (vedi Figura 7.6). In questi casi, per sintetizzare trascritti marcati a partire dal filamento di DNA opposto a quello normalmente trascritto *in vivo*, viene utilizzata la polimerasi del fago.

Le ribosonde possono essere marcate con un radioisotopo, quale ^{33}P o ^{35}S , e successivamente rilevate mediante autoradiografia; per localizzare dei grani d'argento ci si avvale della *microscopia in campo oscuro*. In questo caso non si consente alla luce diretta di raggiungere l'obiettivo, ma i raggi luminosi sono invece diretti su un lato del preparato, in modo che solamente la luce riflessa possa attraversare le lenti del microscopio e il segnale appaia come un oggetto illuminato su uno sfondo scuro. È tuttavia possibile ottenere segnali migliori con la *microscopia in campo chiaro*, in cui l'immagine è ottenuta in seguito alla trasmissione diretta della luce attraverso

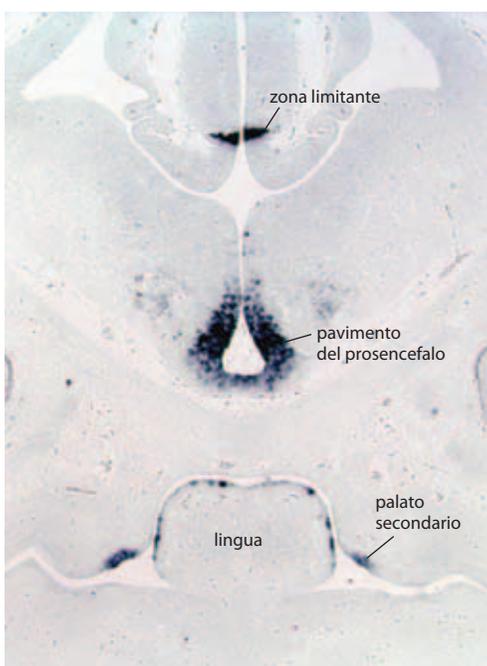


Figura 7.12 Ibridazione *in situ* su tessuto. L'immagine mostra il profilo di espressione prodotto in seguito all'ibridazione di una ribosonda antisenso per il gene *Sonic hedgehog*, marcata con digossigenina, su una sezione di tessuto da un embrione di topo al giorno 13,5 dello sviluppo. La sezione è stata effettuata a livello di una parte del prosencefalo (in alto e in mezzo) e della mascella superiore e inferiore (in basso). La rilevazione della sonda è stata effettuata con un saggio per la fosfatasi alcalina accoppiato allo sviluppo di colore blu. Si osserva una forte espressione in regioni molto specifiche del sistema nervoso centrale in via di sviluppo, come indicato, ma anche nel palato secondario. (Per gentile concessione di Mitsuhiro Nakatomi e Heiko Peters, Università di Newcastle.)

so il campione (Figura 7.12). In alternativa, le sonde sono sottoposte a marcatura in fluorescenza e la rilevazione è effettuata mediante microscopia in fluorescenza (vedi Box 7.3).

I saggi di ibridazione possono essere utilizzati per saggiare colonie batteriche contenenti DNA ricombinante

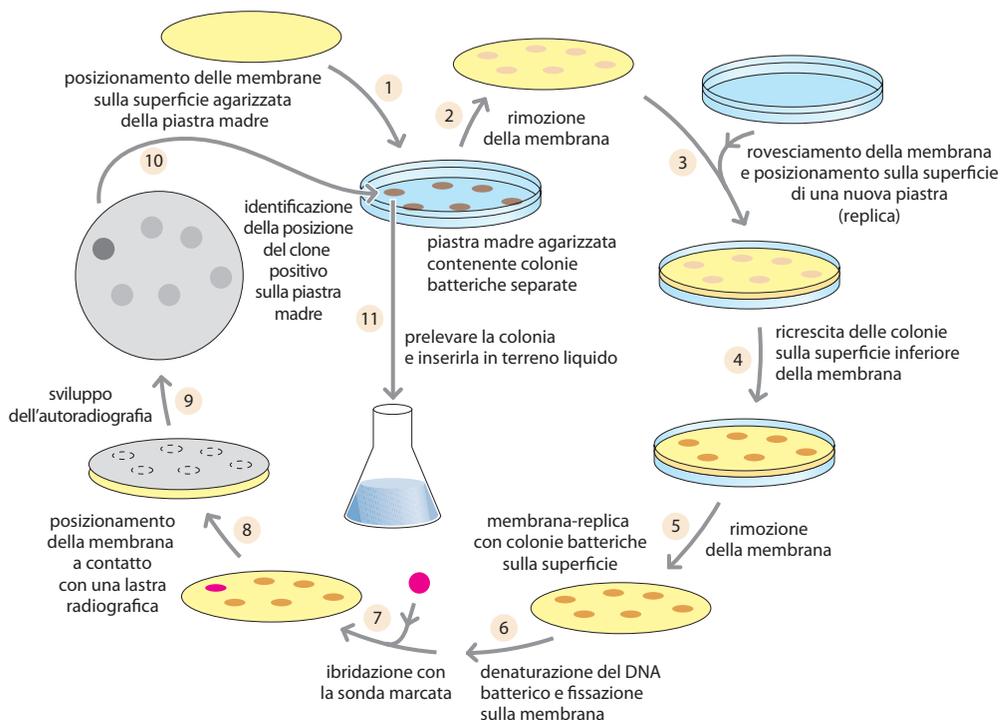
Le colonie batteriche contenenti DNA ricombinante sono spesso sottoposte a *screening* per identificare quelle che contengono sequenze correlate a una sonda marcata di interesse. A tale scopo, si fanno crescere le colonie su un terreno agarizzato e successivamente si trasferiscono per contatto diretto su una membrana di nitrocellulosa o di nylon, in un procedimento noto come *ibridazione su colonia* (Figura 7.13). In alternativa, la miscela di cellule viene distribuita su una membrana di nitrocellulosa o nylon posizionata sulla superficie del terreno contenuto in una piastra Petri, in modo che le colonie batteriche si formino direttamente sulla superficie della membrana. In entrambi i metodi, la membrana viene successivamente esposta a una soluzione alcalina per denaturare il DNA prima di procedere con l'ibridazione con una sonda di acido nucleico marcato.

Dopo l'ibridazione, la soluzione contenente la sonda viene rimossa e il filtro viene accuratamente lavato, asciugato e processato per rilevare la sonda marcata a esso legata. La posizione dei segnali generati dalla sonda viene successivamente confrontata con la posizione originale delle colonie batteriche nella piastra madre, in modo da identificare quelle contenenti inserti di DNA correlati alla sonda utilizzata. Queste colonie possono quindi essere prelevate singolarmente e fatte crescere in un terreno di coltura liquido per ottenere una quantità di cellule sufficiente per l'estrazione e la purificazione del DNA ricombinante di interesse.

Quando divenne possibile allestire genoteche complesse di DNA, si rese necessario mettere a punto metodi più efficienti e automatizzati per effettuare lo *screening* delle colonie: oggi, sono disponibili speciali apparecchiature robotizzate che prelevano le colonie distribuite in piastre a micropozzetti e le trasferiscono sulla membrana da ibridazione secondo coordinate linea-

Figura 7.13 Ibridazione mediante blot su colonia.

Quando un filtro di nylon o nitrocellulosa viene posizionato sulla superficie di una piastra di agar contenente singole colonie batteriche, alcune cellule derivanti dalle colonie aderiscono alla membrana. La membrana viene poi rimossa e trasferita su una nuova piastra, per consentire alle cellule trasferite di crescere e formare nuove colonie. La replica sulla membrana così ottenuta conterrà la stessa rappresentazione fisica delle colonie batteriche della piastra originale (piastra madre). La membrana-replica viene trattata per denaturare il DNA a essa legato e fissarlo sulla sua superficie, viene poi asciugata, e successivamente ibridata con una sonda di acidi nucleici marcata. Dopo il lavaggio, effettuato per rimuovere la sonda marcata non legata in eccesso, si utilizza una lastra radiografica per produrre un autoradiogramma che riveli la posizione delle colonie marcate sulla membrana-replica, in modo da identificare la colonia corrispondente nella piastra madre originale. Le colonie positive possono quindi essere prelevate e fatte crescere singolarmente in coltura liquida.



ri prestabilite. I risultanti filtri di cloni ad alta densità possono consentire lo *screening* rapido ed efficiente di una genoteca, come descritto nel Capitolo 8.

7.4 Saggi di ibridazione basati su microarray

L'ibridazione su microarray consente di effettuare saggi di ibridazione in parallelo utilizzando migliaia di sonde differenti immobilizzate

Le potenti e innovative tecnologie di ibridazione su microarray furono sviluppate nei primi anni '90 per consentire saggi massivi in parallelo, nei quali si effettuano simultaneamente nelle stesse condizioni sperimentali migliaia di singole reazioni di ibridazione. In un primo momento, l'ibridazione su microarray fu concepita come uno strumento per la mappatura su larga scala e il sequenziamento del DNA, ma la principale forza trainante per lo sviluppo di questa tecnologia fu la possibilità di analizzare simultaneamente l'espressione di un numero molto elevato di geni. Nel momento in cui furono allestite le banche a cDNA e divenne disponibile un numero sempre maggiore di cloni di cDNA ben caratterizzati, i ricercatori cominciarono a pianificare metodi per l'analisi globale dell'espressione genica su scala genomica. A tale scopo era necessario effettuare ibridazioni "multiplex", utilizzando contemporaneamente un gran numero di sonde a DNA gene-specifiche. In seguito, l'ibridazione su microarray è stata utilizzata per saggiare sia le varianti del DNA sia i profili di espressione genica.

Se utilizzato in un saggio di ibridazione, un **microarray** è costituito da diverse migliaia di sonde oligonucleotidiche o a DNA non marcate, fissate su un supporto di vetro o su una opportuna superficie secondo coordinate spaziali ben definite, in un formato rappresentato da una griglia ad alta densità. Un campione di interesse, che consiste in una soluzione acquosa contenente una popolazione complessa di DNA o RNA denaturato e marcato in fluorescenza, viene ibridato con le molecole sonda posizionate sul microarray. Dopo i passaggi di lavaggio e asciugatura, la marcatura fluorescente legata all'array viene rilevata grazie a uno *scanner* a luce laser ad alta risoluzione, e il segnale emesso da ogni macchiolina presente sull'array è analizzato utilizzando appositi programmi per l'analisi delle immagini che convertono il segnale emesso in una gamma di diversi colori a seconda dell'intensità del segnale stesso.

Prenderemo in considerazione esempi pratici dell'utilizzo di microarray nei capitoli successivi, come nel Capitolo 8, dove illustreremo come viene analizzata l'espressione genica. Qui ci occupiamo invece dei principi di base di questa metodologia e dei relativi aspetti tecnologici. Un esperimento di ibridazione su microarray consiste nel favorire, per ognuna delle migliaia di sonde differenti presenti nell'array, la formazione simultanea di eteroduplex con le rispettive sequenze bersaglio da una popolazione di acidi nucleici marcati presenti in un campione di interesse. In corrispondenza di *ogni singola* posizione sulla superficie del vetrino sono disponibili per l'ibridazione numerose copie identiche di *un singolo* tipo di oligonucleotide o acido nucleico (Figura 7.14).

Le sequenze bersaglio possono essere presenti in quantità diverse nel campione in esame, e la quantità di uno specifico DNA bersaglio legato da ogni singola sonda dipenderà dalla frequenza di quello specifico DNA bersaglio nel campione di interesse. Una sonda che legni un acido nucleico relativamente poco abbondante possiederà quindi una minore marcatura fluorescente associata, producendo un segnale di ibridazione debole, mentre le posizioni sull'array che risulteranno fortemente fluorescenti dopo il lavaggio corrispondono alle sonde che avranno legato un acido nucleico bersaglio presente in quantità elevate nel campione in esame. La quantificazione del marcatore fluorescente legato a ogni posizione nel microarray fornirà quindi una

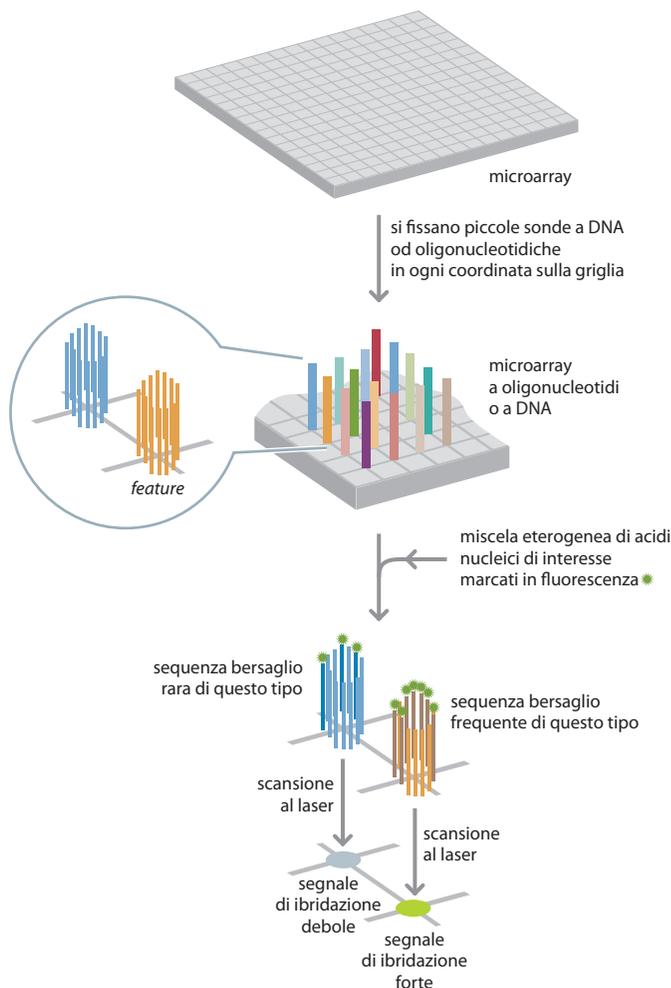


Figura 17.14 Principi dell'ibridazione su microarray. Un microarray consiste in una superficie solida sulla quale è possibile fissare molecole secondo coordinate specifiche in una matrice ad alta densità. L'ibridazione su microarray utilizza array a oligonucleotidi o a DNA che vengono allestiti fissando migliaia di sonde sintetiche di DNA o oligonucleotidi a singolo filamento in specifiche posizioni predeterminate. Come illustrato nella sezione ingrandita, in corrispondenza di ogni specifica posizione sono legate diverse migliaia di copie identiche di un singolo tipo di sonda a oligonucleotide o a DNA (una cosiddetta *feature*). Viene poi aggiunto il campione di interesse, costituito da una serie eterogenea di frammenti marcati di DNA o trascritti a RNA, per consentire l'ibridazione con le sonde presenti sull'array. Le sequenze bersaglio legate da alcune sonde possono essere abbondanti e produrre un forte segnale di ibridazione; per altre sonde possono invece esserci poche sequenze bersaglio presenti nel campione di interesse, e il segnale di ibridazione prodotto sarà quindi più debole. Dopo i passaggi di lavaggio e asciugatura dell'array, i segnali di ibridazione per le diverse migliaia di sonde vengono rilevati mediante una scansione al laser, ottenendo una massa di dati enorme da un singolo esperimento.

quantità enorme di dati circa le frequenze relative di numerose sequenze bersaglio presenti nel campione esaminato.

I microarray basati su oligonucleotidi ad alta densità forniscono strumenti estremamente potenti per l'analisi di campioni complessi di RNA o DNA

La costruzione dei microarray a DNA di prima generazione prevedeva la deposizione robotizzata di cloni di DNA o di oligonucleotidi sintetici sulla superficie di vetrini da microscopia trattati chimicamente. Tuttavia, attualmente si preferiscono microarray a DNA che usano sonde oligodeossinucleotidiche, grazie alla loro particolare versatilità. Gli oligonucleotidi possono essere disegnati in modo che siano specifici per una singola sequenza bersaglio, e la sintesi ordinata di oligonucleotidi implica la possibilità di produrre rapidamente ed efficacemente microarray ad alta densità. In molti casi, l'affidabilità dei dati può inoltre essere definita più efficacemente; per esempio, l'espressione di un singolo gene può essere monitorata simultaneamente con molte sonde oligonucleotidiche differenti, ognuna specifica per una regione particolare del gene di interesse.

La maggior parte delle moderne applicazioni di questa tecnologia fa uso di microarray prodotti grazie a ingegnose tecnologie alternative sviluppate da due aziende. L'approccio seguito da Affymetrix è basato sulla produzione di chip al silicio, e ha introdotto nell'uso comune il termine **chip a DNA** (sebbene attualmente questo termine possa in realtà indicare qualunque microarray di acidi nucleici o oligonucleotidi ad alta densità). La tecnologia sviluppata da Illumina è invece basata sull'assemblaggio casuale di array di biglie in micropozzetti.

I microarray a oligonucleotidi di Affymetrix

La tecnologia dei chip Affymetrix implica la sintesi fotomediata di sonde oligonucleotidiche sul microarray *in situ*, ed è in grado di produrre più di 6 milioni di popolazioni diverse di oligonucleotidi su una superficie di soli 1,7 cm². Ogni tipo di oligonucleotide viene assemblato aggiungendo sequenzialmente mononucleotidi a una molecola *linker* che termina con un gruppo protettivo fotolabile. Si utilizza quindi una “fotomaschera” (fotolitografo) per determinare quale posizione sul microarray esporre a una fonte esterna di luce per provocare la distruzione del gruppo fotolabile. Successivamente, si utilizza una reazione di accoppiamento chimico per aggiungere un particolare tipo di nucleotide ai nuovi siti privi della protezione, e il processo viene ripetuto in sequenza utilizzando maschere diverse per ognuno dei tre nucleotidi rimanenti, per ottenere l’aggiunta di un singolo nucleotide – A, C, G o T – a ogni molecola *linker*. L’intero processo è poi ripetuto per estendere la sintesi degli oligonucleotidi per altre 24 basi circa. A seconda della particolare disposizione di fori ricavati nelle maschere utilizzate in ogni singolo passaggio, è quindi possibile assemblare sulla superficie del microarray specifiche sequenze oligonucleotidiche in posizioni predeterminate (Figura 7.15).

I microarray a oligonucleotidi di Illumina

La tecnologia BeadArray™ di Illumina si basa sull’assemblaggio casuale di array di microsferi in specifici pozzetti. Le biglie consistono in microsferi in silicio di 3 μm di diametro e costituiscono gli elementi dell’array. Gli oligonucleotidi vengono presintetizzati in modo da raggiungere lunghezze superiori ai 70 nucleotidi (suddivisi in un codice di riferimento di 25 nucleotidi e in una sequenza gene-specifica di 50 nucleotidi). Ogni oligonucleotide viene accoppiato a una particolare partita di biglie e ogni singola biglia può portare più di 100 000 oligonucleotidi identici. Su questo stesso principio si basa una tecnologia innovativa di sequenziamento messa in commercio da Illumina, come descritto nel Capitolo 8.

Le biglie contenenti diversi oligonucleotidi sono raggruppate e successivamente distribuite su microarray prefabbricati dotati di micropozzetti a spaziatura regolare, concepiti per accogliere singole biglie. Queste ultime vengono pertanto immobilizzate nella cavità dei micropozzetti e viene decodificato il codice di riferimento a 25 nucleotidi, permettendo l’identificazione di ogni singola biglia. La distribuzione delle sonde sull’array risulta pertanto casuale e le loro posizioni relative sono determinate in un secondo momento. L’ibridazione di acidi nucleici bersaglio marcati con gli oligonucleotidi complementari presenti nell’array di biglie può quindi essere monitorata effettuando la scansione dell’intero array per rilevare la presenza di marcatura associata a specifici pozzetti.

L’ibridazione su microarray è utilizzata principalmente per definire profili trascrizionali o analizzare variazioni a carico del DNA

Sebbene la tecnologia per produrre microarray a DNA sia stata introdotta solo di recente, sono già state sviluppate numerose applicazioni importanti, e il loro impatto sulla futura ricerca biomedica e sugli approcci diagnostici è destinato a essere significativo. La maggior parte delle applicazioni implica la rilevazione di trascritti, oppure lo studio delle variazioni a carico del DNA.

La prima delle due applicazioni – tuttora molto importante – consiste nella misurazione simultanea della quantità di trascritti per molti geni. Originariamente, venivano saggiati sottogruppi limitati rispetto al numero totale di geni presenti in genomi complessi, quale il genoma umano, ma attualmente si possono indagare contemporaneamente i geni contenuti nell’intero genoma. Questa applicazione sarà ulteriormente descritta nel Capitolo 8. La definizione dei profili trascrizionali ha prodotto svariate applicazioni, in

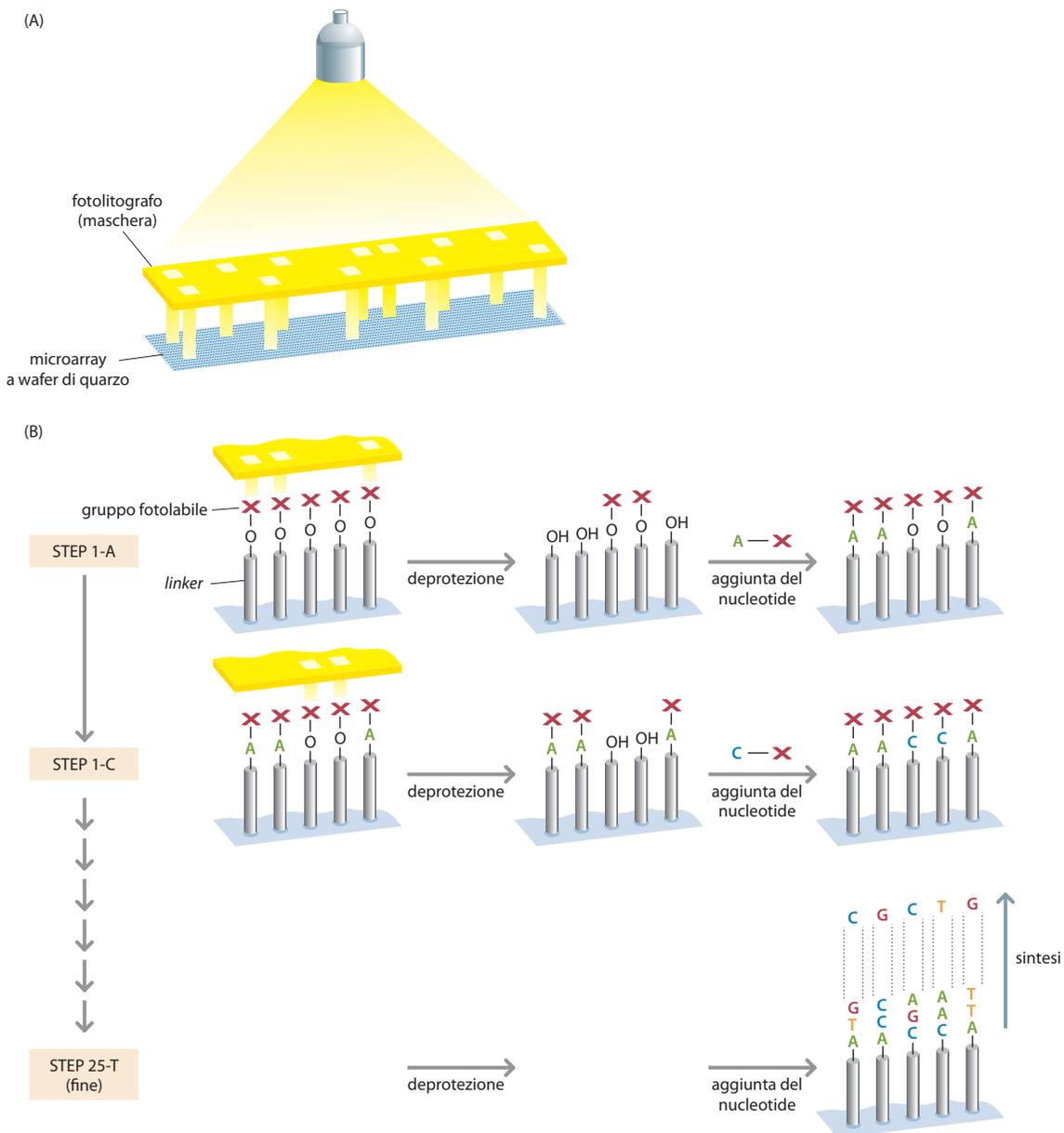


Figura 7.15 Costruzione di un microarray a oligonucleotidi mediante la sintesi *in situ* fotodiretta di oligonucleotidi.

(A) I microarray a oligonucleotidi di Affymetrix sono costruiti sintetizzando migliaia di oligonucleotidi in posizioni predeterminate su un microarray costituito da uno strato di quarzo, utilizzando un fotolitografo (maschera) per determinare la posizione di ogni nucleotide che viene aggiunto. (B) Ogni singolo oligonucleotide viene assemblato a partire da una molecola *linker* che è fissata all'array e la cui estremità libera porta un gruppo protettivo fotolabile. Una maschera dotata di fori in posizioni specifiche viene posizionata sopra l'array e successivamente viene fatta passare attraverso le aperture nella maschera luce proveniente da una sorgente esterna. Nelle singole posizioni dell'array esposte alla luce, i gruppi fotolabili vengono distrutti

(deprotezione). Un mononucleotide specifico accoppiato al gruppo fotolabile viene quindi aggiunto all'array, dove reagirà con i filamenti deprotetti. Il processo viene ripetuto tre volte utilizzando altre tre maschere dotate di fori in posizioni differenti, per introdurre in sequenza uno degli altri tre nucleotidi. Il risultato netto è che ogni molecola *linker* sarà dotata di un singolo nucleotide, A, C, G o T, attaccato specificamente con modalità pre-programmata. L'uso in sequenza di quattro maschere diverse per consentire l'accoppiamento dei quattro nucleotidi differenti viene ulteriormente reiterato per altri 24 cicli. Il risultato finale è l'assemblaggio di oligonucleotidi lunghi 25 basi, ognuno dotato di una sequenza predeterminata, in corrispondenza di ogni specifica coordinata spaziale del microarray.

particolare viene usata per confrontare i profili di espressione di tipi cellulari strettamente correlati, che rappresentano cellule altrimenti identiche tranne che per l'eventuale presenza di una particolare mutazione patogena, o ancora cellule identiche esposte in maniera differenziale a particolari condizioni ambientali predefinite, quali la presenza o meno di un farmaco. È possibile

valutare anche altri aspetti dell'espressione genica, quali la presenza di varianti di *splicing*, rilevata grazie all'utilizzo di specifiche sonde oligonucleotidiche.

Stanno inoltre prendendo progressivamente piede microarray specifici per particolari *pathway*, allo scopo di monitorare singoli profili di espressione per una serie di geni funzionalmente correlati che prendono parte a una via biologica comune. Anche le applicazioni diagnostiche stanno assumendo una crescente importanza. Per esempio, i profili trascrizionali per geni cruciali possono fornire tracciati molecolari in grado di definire i diversi tipi o stadi tumorali.

Numerose applicazioni dei microarray sono state inoltre concepite per indagare la variazione a livello del DNA. Oltre a saggiare mutazioni a carico di geni-malattia noti, ingenti sforzi sono stati profusi per indagare e catalogare **polimorfismi a singolo nucleotide (SNP)** nell'uomo. Come illustreremo più avanti, l'analisi della variazione del DNA su larga scala, per identificare grosse delezioni e duplicazioni subcromosomiche, utilizza spesso microarray basati su sonde derivate da cloni di DNA di cromosomi batterici artificiali.

Inoltre, va menzionato il fatto che i microarray possono avere numerose altre applicazioni, oltre all'ibridazione, quali gli array di proteine, gli array in grado di rilevare il legame tra acidi nucleici e proteine, e altre ancora.

LETTURE DI APPROFONDIMENTO

Generali

Sambrook J e Russell DW (2001) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Marcatura in fluorescenza

Invitrogen Molecular Probes—The Handbook: A Guide To Fluorescent Probes And Labeling Technologies. <http://www.invitrogen.com/site/us/en/home/References/Molecular-Probes-The-Handbook.html>

Ibridazione *in situ*

Wilkinson D (1998) *In Situ Hybridization: A Practical Approach*, 2nd ed. IRL Press.

Principi e tecnologia di base dell'ibridazione su microarray

Autori vari (1999) The Chipping Forecast. *Nat. Genet.* 21 (Suppl.), 1-60.

Autori vari (2002) The Chipping Forecast II. *Nat. Genet.* 32 (Suppl.), 465-551.

Tecnologie dei microarray a oligonucleotidi di affymetrix e illumina

Affymetrix GeneChip™. <http://www.affymetrix.com/technology/index.affx>

Gunderson KL, Kruglyak S, Graige MS et al. (2004) Decoding randomly ordered DNA arrays. *Genome Res.* 14, 870-877.

Illumina BeadArray™. <http://www.illumina.com/pages.ilmn?ID=5>

Applicazioni dell'ibridazione su microarray

Falciani F (ed.) (2007) *Microarray Technology Through Applications*. Taylor and Francis.

Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. *Nat. Rev. Genet.* 7, 200-210.