

Analisi quantitativa in biologia molecolare

O. Valentini, B. Milanesi

Laboratori di Patologia Clinica
Azienda Ospedaliera di Desenzano del Garda (BS)

Riassunto

La quantificazione di specifiche sequenze genomiche, e più ancora quella dell'espressione genica, richiedeva in passato procedure molto laboriose ed era possibile solo presso strutture di ricerca. Oggi si dispone di strumenti e protocolli operativi che consentono di ottenere lo stesso risultato in maniera semplice e veloce, rendendo possibile l'effettuazione di questo tipo di indagini, almeno entro certi limiti, anche presso strutture di tipo clinico diagnostico. L'analisi di alcuni aspetti di questi procedimenti operativi, e in particolare delle condizioni necessarie per ottenere un dato quantitativo corretto, è utile per invogliare a ulteriori approfondimenti, sulla traccia anche di alcune indicazioni bibliografiche, finalizzati a stimolare la capacità critica dell'operatore.

Summary

Quantitative analysis in molecular biology.

Until not long ago the quantitative analysis of specific gene sequences and of gene expression was possible only in research centres and only by means of complex and laborious processes. Nowadays we have instruments and technology that enable us to do this kind of analysis in a simple and fast way even in clinical diagnostic centres. The study of some aspects of the processes by which it is possible to obtain the quantitative information is useful in order to lead to further mastering with suggested bibliography and to stimulate a critical vision in the operator.

Premessa

Grazie alle attuali tecniche di biologia molecolare è possibile ottenere in tempi rapidi e a costi contenuti informazioni di tipo quantitativo su sequenze specifiche di acido nucleico utili nella ricerca pura e applicata, nonché in svariati ambiti della diagnostica infettivologica, oncologica, delle malattie genetiche e in altri ambiti ancora, come quello medicolegale. Anche l'espressione genica è suscettibile di indagine quantitativa, che può essere condotta con le attuali tecnologie in tempi incomparabilmente più rapidi e in modo di gran lunga più semplice rispetto a un recente passato.

L'ottenimento di un dato quantitativo corretto presuppone il rispetto di precise condizioni operative ed è possibile solo mettendo in atto strategie per il controllo di ogni singola fase dell'iter analitico. Appresso verranno sommariamente descritti i principali strumenti analitici utilizzabili per la quantificazione di bersagli genomici e per una valutazione quantitativa dell'espressione genica. È importante conoscere nel dettaglio i meccanismi attraverso i quali si giunge al risultato quantitativo e le possibilità di controllo su tali meccanismi per poter valutare criticamente e correttamente il dato analitico stesso.

Richiami storici

In passato, fino alla fine degli anni Ottanta, le indagini di tipo quantitativo sugli acidi nucleici non potevano essere dirette su sequenze specifiche, con l'eccezione di quelle oggetto di clonazione, tecnica, questa, alquanto laboriosa e non utilizzabile al di fuori di strutture espressamente dedicate alla ricerca biologica. La quantificazione di sequenze specifiche richiedeva allora di solito il loro preliminare isolamento con metodiche di tipo cromatografico o elettroforetico e il successivo utilizzo di tecniche di chimica analitica o di marcatura con isotopi radioattivi. Per fare qualche esempio, oltre alle misure di assorbanza nell'ultravioletto, A_{260}/A_{280} , eventualmente previa digestione con ribonucleasi o desossiribonucleasi per eliminare rispettivamente RNA o DNA contaminante ($1 \text{ OD}_{260} = 50 \text{ microgr/ml}$ per DNA, $1 \text{ OD}_{260} = 40 \text{ microgr/ml}$ per RNA), si poteva ricorrere alla metodica di Burton¹ con difenilamina per il DNA o, sempre per il DNA, a quella con acido 3,5-diaminobenzoico in fluorescenza di Kissane e Robins^{2,3}. Per il dosaggio di mRNA eucariotico, poliadenilato, si potevano utilizzare tecniche come l'ibridazione con poli(U) marcato^{4,5}, in eccesso, seguita da rimozione del poli(U) non ibridato mediante ribonucleasi, precipitazione con acido tricloroacetico dell'ibrido e

quantificazione del marcatore, più spesso radioattivo che enzimatico. Si trattava di tecniche piuttosto laboriose, per le quali erano necessari ambienti attrezzati e una particolare preparazione degli operatori, che ne limitava l'utilizzo in ambiti circoscritti; il loro impiego in ambito diagnostico era affatto limitato. Questa situazione era destinata ad essere rivoluzionata con l'avvento, a partire dalla metà degli anni Ottanta, di nuove tecnologie, prima fra tutte la reazione a catena della DNA-polimerasi.

Il principio della metodica è frutto di un'idea espressa da Kary Mullis⁶⁻⁸ nel 1983 e fondamentale consiste nell'impiego di due oligonucleotidi complementari a sequenze presenti su eliche opposte del DNA utilizzati come innesco per la DNA-polimerasi, che viene a replicare in cicli successivi la sequenza di DNA delimitata dagli stessi inneschi, in modo tale che la quantità di sequenza può essere raddoppiata ad ogni ciclo replicativo. Il processo descritto da Mullis è molto simile a quello già descritto da un altro ricercatore, Gobind Khorana, tra la fine degli anni Sessanta e l'inizio dei Settanta⁹. Il processo descritto da Khorana e da lui indicato come "replicazione riparativa" comprendeva i seguenti passaggi: (1) estensione ad opera della DNA-polimerasi a partire da un innesco legato a una singola elica bersaglio, (2) separazione dell'elica neosintetizzata da quella bersaglio complementare e (3) reibridazione dell'innesco con ripetizione del ciclo. A quest'ultimo scienziato e ai suoi collaboratori va se non altro il merito di aver suggerito l'idea.

In origine la DNA-polimerasi utilizzata era quella di *Escherichia coli*, che doveva essere aggiunta ad ogni ciclo replicativo, poiché la temperatura necessaria per la denaturazione della doppia elica di DNA all'inizio di un ciclo provocava anche l'inattivazione irreversibile dell'enzima aggiunto nel ciclo precedente. Un notevole passo in avanti fu compiuto quando si arrivò a disporre di un enzima termostabile, ottenuto da un batterio, *Thermophilus aquaticus*, isolato dalle sorgenti termali calde e capace di mantenere gran parte della sua attività anche dopo numerosi cicli replicativi. Risolto questo problema, il processo si prestava a essere automatizzato grazie all'utilizzo di apparecchi termostatici in grado di compiere in successione tutti i cicli termici necessari per la reazione a catena. Ciò ha reso alquanto maneggevole l'impiego di questa metodica, contribuendo alla sua larga e rapida diffusione.

Le prime applicazioni della reazione DNA-polimerasica a catena (PCR, Polymerase Chain Reaction) erano essenzialmente di tipo qualitativo e solo con adattamenti laboriosi era possibile un suo utilizzo in senso quantitativo. Effettuando, per esempio, più reazioni in parallelo su una stessa quantità di campione e bloccando la reazione dopo differenti numeri di cicli si poteva determinare poi la quantità di prodotti, preferibilmente utilizzando marcatori radioattivi per aumentare la sensibilità del metodo; i dati ottenuti consentivano di estrapolare la quantità di sequenza specifica dell'acido nucleico bersaglio inizialmente presente nel campione biologico. Nel 1990 fu descritta da vari Autori^{10,11} una variante competitiva della PCR, consistente nell'amplificazione simultanea nella stessa miscela di reazione di due sequenze diverse, ma tuttavia amplificabili con identica efficienza, di cui una a concentrazione ignota (sequenza bersaglio del campione) e una a concentrazione nota

(calibratore o competitore interno). Quest'ultimo veniva ottenuto clonando una sequenza di DNA derivata da quella bersaglio per introduzione di un breve inserto, oppure per sostituzione di una o poche coppie di basi, in modo da ottenere un sito di restrizione artificiale (sequenza riconosciuta da una specifica endonucleasi di restrizione). La PCR competitiva veniva effettuata con quantità scalari di DNA competitore aggiunti nella miscela di amplificazione a una quantità fissa di sequenza bersaglio, cDNA o DNA genomico, del campione. La quantità di bersaglio inizialmente presente era ottenibile, nell'ipotesi di una identica efficienza di amplificazione, dal rapporto dei prodotti di amplificazione (rapporto tra bersaglio e competitore). All'iniziale quantificazione dei prodotti di amplificazione del bersaglio e del calibratore su gel si sostituì in seguito una più accurata quantificazione attraverso l'utilizzo di sonde specifiche per le due diverse sequenze con marcatura di tipo enzimatico. L'osservazione (Higuchi et al, 1992) che l'etidio bromuro, un composto fluorescente che si intercala nella doppia elica del DNA durante la sua sintesi, può essere usato come marcatore dei prodotti di amplificazione della PCR^{12,13} aprì la strada a nuove tecniche di rivelazione simultanea, e quindi di tipo cinetico, dei prodotti della PCR stessa. L'intensità della fluorescenza è direttamente proporzionale alla quantità di amplificato e può essere monitorata ciclo dopo ciclo. Stabilita una intensità soglia di fluorescenza, si può osservare che quest'ultima viene raggiunta dopo un numero di cicli amplificativi che è inversamente proporzionale alla quantità di sequenza bersaglio inizialmente presente nel campione: quanto maggiore è il numero di copie di DNA bersaglio presenti al ciclo zero, tanto minore è il numero di cicli necessari per evidenziare l'amplificato al livello soglia di fluorescenza. Nel 1993 i ricercatori della Perkin-Elmer combinarono il metodo cinetico di Higuchi con le sonde sensibili alle esonucleasi scoperte da Holland et al¹⁴, sviluppando così una variante della amplificazione in tempo reale, siglata come metodo TaqMan¹⁵. Altre varianti si aggiunsero in seguito, rendendo questa tecnica ancora più versatile. Attualmente esistono vari modi per ottenere informazioni di tipo quantitativo su sequenze specifiche di acido nucleico, che verranno appresso brevemente descritti.

Una valutazione quantitativa dell'espressione genica, in termini assoluti e comparativi, è possibile con i mezzi di indagine testè descritti, che consentono di quantificare l'RNA messaggero trascritto a partire da specifici geni, tuttavia è soltanto nell'ultimo decennio che si sono ideati e perfezionati mezzi di indagine di estrema potenza attraverso i quali è possibile la rapida analisi comparativa contemporanea di migliaia di geni: si tratta dei "microarray", cui è dedicata l'ultima parte di questa rassegna.

Quantificazione di sequenze specifiche di acido nucleico: principali tecniche

In prima approssimazione si possono distinguere: 1) tecniche di amplificazione, ciclotermica o isoterma, della sequenza genomica bersaglio e 2) tecniche di esclusiva amplificazione del segnale. Le tecniche di amplificazione ciclotermica comprendono la reazione DNA-polimerasica a catena nella sua forma classica, la reazione DNA-ligasica a catena e la PCR in tempo reale, nelle sue numerose va-

rianti; oltre a queste, meritano di essere sommariamente descritte, tra le metodiche di amplificazione genomica, quella isotermica, nelle versioni classica e in tempo reale, e, tra le tecniche di amplificazione del segnale, quella che fa uso del DNA ramificato (bDNA). Le tecniche amplificative basate sulla reazione a catena della DNA-polimerasi o della DNA-ligasi utilizzano come substrato il DNA. Per poter essere amplificato con tali tecniche, l'RNA deve essere prima retroscritto in cDNA, ad opera di una trascrittasi inversa. Nel caso della reazione a catena della polimerasi esiste anche la possibilità di utilizzare per entrambe le fasi di retroscrittura e di amplificazione un unico enzima, come la DNA-polimerasi di *Thermus thermophilus*, in grado di esplicare, oltre all'attività polimerasica, anche quella di retroscrittasi.

Amplificazione del segnale con "DNA ramificato" (bDNA)

È una tecnica di dosaggio¹⁶⁻²¹ che non altera la quantità di molecola bersaglio, ma la rivela direttamente attraverso una forte amplificazione del segnale. Prevede l'uso di oligodesossiribonucleotidi ramificati (bDNA, per "branched DNA"), che consentono di incorporare fino a circa 3000 molecole di enzima su un'unica molecola bersaglio di acido nucleico, che può essere, indifferentemente, DNA o RNA. Non richiede la purificazione dell'acido nucleico, né le particolari precauzioni previste con le tecniche di amplificazione genomica.

Vari oligonucleotidi sintetici specifici per la sequenza bersaglio sono utilizzati per la cattura di quest'ultima sulla superficie di una micropiastra (sonde di cattura). Un secondo gruppo di sonde (sonde di preamplificazione) media l'ibridazione delle sonde di DNA ramificato (sonde di amplificazione) all'acido nucleico bersaglio. Ogni sonda di DNA ramificato (bDNA) ha una quindicina di ramificazioni, ognuna delle quali può legare varie sonde marcate con enzima (fosfatasi alcalina), dette sonde di marcatura. Da ultimo viene aggiunto un substrato chemiluminescente (diossietano) e viene misurata l'emissione luminosa, la cui intensità è direttamente proporzionale alla quantità di acido nucleico bersaglio presente nel campione. La quantificazione è possibile utilizzando una curva di calibrazione, o di riferimento, ottenuta misurando la fluorescenza sviluppata con quantità note di DNA o RNA.

Amplificazione isotermica convenzionale e "in tempo reale"

Il metodo NASBA (Nucleic Acid Sequence-Based Amplification) e la sua variante TMA (Transcription Mediated Amplification) sono tecniche di amplificazione isotermica inizialmente descritte per l'RNA, ma adattabili anche al DNA, in cui si fa uso di un innesco iniziale che porta legata una sequenza promotore per l'RNA-polimerasi²²⁻³².

L'innesco-promotore si lega al filamento di RNA bersaglio, consentendo a una trascrittasi inversa di sintetizzare il singolo filamento complementare di DNA. Successivamente il filamento originale di RNA viene degradato ad opera di una ribonucleasi (Rnasi H), mentre il cDNA a singolo filamento rimane intatto. A questo punto si aggiunge un secondo innesco, che si lega al cDNA, consentendo la sin-

tesi dell'elica complementare. Il DNA così ottenuto conterrà la sequenza promotore per la RNA-polimerasi, che potrà trascriverlo in migliaia di copie di RNA complementare.

Nella metodica classica l'amplificato viene rivelato al termine del processo amplificativo, mentre nella variante "in tempo reale" la rivelazione simultanea all'amplificazione viene ottenuta utilizzando sonde complementari all'amplificato aventi struttura ad ansa, o a forcina, (per parziale complementarità intramolecolare) marcate con un composto fluorescente (fluoroforo) all'estremità 5' e con uno smorzatore (quencher) universale all'estremità 3'. Poiché le due estremità 5' e 3' della sonda sono tra loro complementari e si associano a formare un tratto a doppia elica, mentre nel suo tratto centrale la molecola è costituita da un filamento a singola elica ripiegato ad ansa, il fluoroforo legato all'estremità 5' e lo smorzatore legato all'estremità 3' sono tra loro vicini, e ciò causa uno smorzamento della fluorescenza^{33,34}. Quando la sonda trova la sua sequenza complementare sull'amplificato, il numero di legami a idrogeno che con essa può formare supera il numero dei legami a idrogeno intramolecolari; per questo motivo l'interazione tra sonda e amplificato è termodinamicamente favorita e, comportando la perdita della struttura secondaria ad ansa con conseguente allontanamento delle estremità 3' e 5', determina la rimozione dell'effetto di smorzamento, e quindi la comparsa di fluorescenza. La fluorescenza che si sviluppa è così direttamente proporzionale al numero di molecole di amplificato.

Nella configurazione descritta, l'amplificazione isotermica, sia del tipo a rivelazione postamplificativa (NASBA o TMA convenzionali), sia del tipo a rivelazione simultanea (in tempo reale) dà risultati di tipo prettamente qualitativo. È tuttavia possibile rendere quantitative queste metodiche attraverso l'introduzione nella miscela di reazione di quantità note di un competitore interno, o standard interno, che venga coamplificato con pari efficienza rispetto alla sequenza bersaglio. A tale scopo il competitore dovrebbe avere le stesse sequenze di innesco del bersaglio, differendo da quest'ultimo solo per poche basi nucleotidiche nella regione compresa tra gli inneschi, in modo da poter legare sonde di rivelazione discriminabili da quelle specifiche per il bersaglio e in grado di fornire un segnale distinto. Ciò si può ottenere legando marcatori diversi alle sonde specifiche per bersaglio e per competitore. Nella rivelazione postamplificativa una stessa quantità di amplificato, in cui sono presenti sia il bersaglio che il competitore, può essere cimentata in pozzetti diversi, rispettivamente con la sonda specifica per il bersaglio e con la sonda specifica per il competitore, ottenendo così due segnali distinti; dal rapporto tra l'intensità dei due segnali sarà possibile ottenere una valutazione quantitativa del bersaglio. Nel caso dell'amplificazione in tempo reale le due sonde saranno legate a fluorocromi diversi che emettono a lunghezze d'onda differenti e ben discriminabili e dal rapporto tra le intensità dei due segnali sarà possibile calcolare la quantità iniziale di bersaglio.

Amplificazione ciclotermica: reazione DNA-polimerasica a catena

Questo metodo^{8,35} utilizza un paio di oligonucleotidi sinte-

tici (iniziatori, o inneschi), complementari ciascuno a uno solo dei due filamenti del dsDNA bersaglio, che, con la loro estremità 3'-OH libera, fungono da innesco per la DNA-polimerasi. La regione delimitata dai due inneschi viene replicata ciclicamente in modo esponenziale. Ogni ciclo termico replicativo consta di tre fasi:

- 1) separazione dei due filamenti complementari del DNA a doppia elica a temperature superiori a 90 °C;
- 2) associazione ibridativa dei due inneschi oligonucleotidici, ciascuno con il rispettivo filamento ("annealing") a 50-75 °C;
- 3) sintesi, a partire da ciascun innesco, su ciascuno dei due filamenti, della catena complementare con conseguente duplicazione della regione genomica bersaglio, a 72-78 °C.

Come è stato ricordato, in origine era necessario aggiungere la DNA-polimerasi ad ogni ciclo, a causa dell'inattivazione termica dell'enzima. La tecnica ha subito un miglioramento formidabile con l'utilizzo di una DNA-polimerasi termostabile, ottenuta dal batterio termofilo *Thermus aquaticus* e perciò indicata come *Taq*, che rimane attiva per molti cicli replicativi.

Riportando in ordinate il logaritmo della quantità di amplificato e in ascisse il numero di cicli termici si ottiene una curva assimilabile a un ramo di sigmoide (Fig. 1). Durante la fase logaritmica lineare della curva, in condizioni ideali, la quantità di amplificato aumenta di circa 3 logaritmi decimali (ossia di un fattore 1000) ogni 10 cicli. Superato però un certo numero di cicli, la reazione si fa sempre meno efficiente, fino a raggiungere una fase di plateau, in corrispondenza della quale non si ha più alcun incremento dell'amplificato.

Il raggiungimento della fase di plateau è imputabile ai seguenti fattori:

- 1) graduale inattivazione termica della DNA-polimerasi;
- 2) diminuzione esponenziale nel corso del processo amplificativo del rapporto molare tra la DNA-polimerasi e il suo substrato (amplificato), fino al punto che la quantità di enzima può diventare limitante;
- 3) riduzione progressiva dell'efficienza di ibridazione ("annealing") per diminuzione esponenziale del rapporto tra le concentrazioni di innesco e di amplificato; in tali condizioni l'interazione ibridativa tra filamenti complementari di amplificato diventa sempre più frequente rispetto all'analogia interazione tra innesco e amplificato;
- 4) degradazione del DNA ad opera dell'attività 5'-3' esonucleasica associata alla *Taq* DNA-polimerasi.

Con le miscele di reazione normalmente utilizzate, la fase di plateau viene di norma raggiunta a causa dei fattori sopra elencati molto prima che vengano esauriti gli inneschi oligonucleotidici o i trifosfonucleotidi nella miscela stessa³⁵.

Reazione a catena della DNA-ligasi

La reazione ligasica a catena³⁶⁻⁴⁰ prevede l'utilizzo di 4 sonde, che a due a due delimitano sui due filamenti complementari del DNA bersaglio gli estremi della regione da amplificare. Dopo denaturazione termica della sequenza bersaglio e appaiamento delle quattro sonde a due a due sui due filamenti complementari, le sonde di ciascuna coppia, se tra loro adiacenti, vengono unite covalentemente ad

opera della DNA-ligasi. Quando le sonde di ogni coppia siano invece separate tra loro, è necessaria la presenza di una DNA-polimerasi che sintetizzi la regione mancante, interposta tra le sonde, dopo di che il legame covalente può essere ripristinato dalla DNA-ligasi. Ad ogni ciclo si avrà la duplicazione della sequenza bersaglio di acido nucleico, attraverso le fasi di denaturazione termica con separazione dei filamenti complementari, associazione delle sonde ai filamenti bersaglio, sintesi dei tratti mancanti e saldatura covalente finale.

Amplificazione ciclotermica in tempo reale (PCR cinetica)

La PCR "in tempo reale", o PCR "a rivelazione simultanea", o PCR cinetica⁴¹⁻⁵⁴, utilizza una strategia che consente di rivelare l'amplificato nel momento stesso in cui si forma. Questa tecnica, che si è molto evoluta rispetto alla sua primitiva versione in cui veniva impiegato l'etidio bromuro^{12,13} si basa sull'impiego di leganti fluorescenti, o di sonde marcate con composti fluorescenti, oppure di inneschi a marcatura fluorescente, che forniscono un segnale la cui intensità, misurabile in ogni singolo istante, è direttamente correlata con la quantità di amplificato in quello stesso istante.

Esistono varie configurazioni del sistema di rivelazione: si possono utilizzare composti fluorescenti che si legano alla doppia elica del DNA, come il SYBR Green I, oppure sonde "a idrolisi", sonde a forcina, sonde "adiacenti", od "osculanti", inneschi "a sole nascente", inneschi "scorpione"^{14,42,55-62}. Nella configurazione più nota, quella della sonda "a idrolisi"^{14,42}, la sonda, complementare a uno solo dei due filamenti di DNA della regione bersaglio in un tratto situato tra i due inneschi oligonucleotidici, è marcata alla sua estremità 5' con una molecola fluorescente (definita "reporter", o "fluoroforo"), come ad es. la 6-carbossifluoresceina, oppure il suo tetracloro o esacloroderivato, e all'estremità 3' con uno smorzatore di fluorescenza (definito "quencher", o "smorzatore" appunto), come ad es. la 6-carbossitetrametilrodamina⁵².

Quando la sonda è integra, l'emissione di fluorescenza da parte del fluoroforo è soppressa dallo smorzatore, a causa di un effetto di risonanza che interessa l'intera catena covalente del DNA interposto, noto come effetto Foerster^{63,64}, o F.R.E.T. (Fluorescence Resonance Energy Transfer, trasferimento di energia per risonanza in fluorescenza, effetto la cui intensità è inversamente proporzionale alla sesta potenza della distanza tra fluoroforo e smorzatore)⁶⁵⁻⁶⁸.

Durante la fase di ibridazione degli inneschi ai due filamenti della regione bersaglio, a uno dei due filamenti si lega anche la sonda. Nella fase di polimerizzazione, o fase di estensione, della PCR, la DNA-polimerasi, incontrando la sonda, la degrada per effetto della sua attività 5'-3' esonucleolitica¹⁴, rompendo la catena covalente interposta tra fluoroforo e smorzatore. Cessando così l'effetto di soppressione, si ha emissione di fluorescenza, che aumenta in modo esponenziale ad ogni ciclo amplificativo e presenta intensità direttamente proporzionale alla quantità di amplificato prodotto.

Poiché dopo un sufficiente numero di cicli amplificativi tutte le curve di amplificazione tendono a plateau, dove si perde la linearità nella lettura di tipo quantitativo, la lettura

sarà tanto più lineare quanto più precocemente nel corso della reazione amplificativa sarà possibile evidenziare il prodotto amplificato. Il momento migliore per la lettura quantitativa di tipo fluorimetrico è quindi quello in cui l'intensità di fluorescenza è minima, ma discriminabile dal rumore di fondo. Quanto maggiore è la quantità iniziale di sequenza bersaglio, tanto più precocemente si raggiungerà il livello soglia per la rilevazione, ossia tanto minore sarà il numero di cicli amplificativi necessario per raggiungere il livello di determinabilità.

Il numero di cicli amplificativi necessario per raggiungere l'intensità soglia rivelabile di fluorescenza e discriminabile dal rumore di fondo viene indicato come "ciclo soglia" (C_t , dove t sta per threshold = soglia) e corrisponde a un numero intero o frazionario, tanto maggiore quanto minore è la quantità di sequenza bersaglio presente nel campione in analisi. L'intensità soglia di fluorescenza è pari a 10 volte la deviazione standard del segnale corrispondente alla linea di base.

Riportando in assi cartesiani ortogonali i valori di ciclo soglia (C_t) contro il logaritmo del numero iniziale di copie si ottiene una linea retta.

Poiché nella PCR in tempo reale la rivelazione dell'amplificato avviene nella fase di maggiore linearità della curva amplificativa (quando l'intensità del segnale fluorescente e la quantità di amplificato prodotto sono perfettamente proporzionali e quando l'efficienza di amplificazione è costante e massimale), anziché in una fase vicina al plateau, l'ambito di linearità nel dosaggio risulta molto ampio, dell'ordine di almeno 7 o 8 logaritmi decimali, e la variabilità analitica è molto bassa^{53,54}. La PCR in tempo reale nasce con caratteristiche di tecnica fondamentalmente di tipo quantitativo, dal momento che, perfino in assenza di una curva di calibrazione esterna o di un calibratore interno, il raggiungimento della soglia di fluorescenza dopo un determinato numero di cicli amplificativi (ciclo soglia) già da solo fornisce una indicazione sulla quantità di acido nucleico presente nel campione al ciclo zero. La soglia di fluorescenza corrisponde infatti in linea teorica a una quantità fissa di acido nucleico per una determinata configurazione del sistema di rivelazione e dipende in buona misura dal tipo di fluoroforo impiegato. Una volta che sia stato rilevato il numero di cicli necessari per arrivare a quella precisa quantità di acido nucleico, ossia una volta che sia noto il numero corrispondente al ciclo soglia, è possibile calcolare dall'equazione (vedasi oltre) che descrive la cinetica di amplificazione la quantità molare di acido nucleico iniziale dividendo la quantità molare necessaria a fornire la fluorescenza soglia per il fattore di amplificazione $(1+E)$ elevato alla potenza n , dove n = ciclo soglia e dove E (efficienza di amplificazione) è approssimabile all'unità, poiché fino al ciclo soglia la cinetica è di tipo logaritmico lineare. Ovviamente, per le ragioni che verranno esposte, l'utilizzo di una curva di calibrazione esterna e, più ancora, di un calibratore interno, offre maggiori garanzie di sicurezza e di precisione nella misura.

Un vantaggio non trascurabile dell'amplificazione in tempo reale è la sua rapidità, dal momento che non è prevista una fase di rivelazione successiva a quella di amplificazione, essendo i due processi simultanei. Ciò si traduce in un risparmio di tempo e nella riduzione delle probabilità di

errori operativi che qualsiasi procedura manipolativa, nella fattispecie finalizzata alla rivelazione del prodotto amplificato, necessariamente comporta.

Un altro importante elemento a favore è rappresentato dalla altissima sensibilità di questa tecnologia amplificativa, non inferiore a quella della PCR convenzionale di tipo qualitativo e paragonabile a quella di una PCR "a innesto" ("nested PCR", dove l'amplificato di una prima reazione DNA-polimerasica a catena viene utilizzato come substrato per una seconda reazione amplificativa che utilizza innesci spostati più internamente rispetto a quelli della prima reazione). Ciò consente di superare, cancellandola, la dicotomia tra saggio qualitativo, in cui si ricerca una sensibilità elevata, e saggio quantitativo, di norma meno sensibile del precedente.

Un limite della metodica è rappresentato dal fatto che in una stessa miscela di reazione è possibile amplificare solo poche (fino a 6) sequenze bersaglio⁶⁹, poiché ogni sequenza bersaglio deve essere marcata con un fluorocromo diverso, che emetta a una lunghezza d'onda chiaramente discriminabile, affinché i segnali corrispondenti ai diversi amplificati non si sovrappongano tra loro.

Cinetica di formazione del prodotto nelle reazioni amplificative ciclotermiche^{13,70-73}

Per una valutazione critica del risultato quantitativo, in particolare con i metodi di amplificazione che sfruttano la reazione DNA-polimerasica a catena, è utile soffermarsi su alcuni aspetti di carattere cinetico, la conoscenza dei quali è di particolare importanza per definire le condizioni operative e le strategie analitiche finalizzate all'ottenimento di un risultato che sia la rappresentazione più fedele possibile del dato biologico reale.

Indicando con $M(0)$ la quantità iniziale di molecola bersaglio, la quantità di amplificato ottenuta dopo n cicli amplificativi è pari a:

$$M(n) = M(0)(1 + E)^n$$

che diventa: $M(n) = 2^n M(0)$ per $E = 1$, dove E indica l'efficienza di amplificazione, che varia in funzione di n , assumendo, dopo una prima fase di relativa stabilità che si osserva nei primi cicli, valori progressivamente decrescenti, dal valore massimo iniziale (che in condizioni ottimali è prossimo a 1, ossia al 100%) fino a un valore minimo finale pari a zero.

La graduale progressiva diminuzione del valore di E fino a 0 nel corso del processo amplificativo è dovuta ai fattori già ricordati (inattivazione termica dell'enzima, variazione dei rapporti molar, attività nucleasica, ecc.).

Considerando che se $a = b$, anche $\log a = \log b$, e ricordando che $\log(ab) = \log(a) + \log(b)$, e che $\log(a^n) = n \log(a)$, dall'equazione cinetica sopra riportata ricaviamo la sua equivalente forma logaritmica:

$$\log M(n) = n \log (1 + E) + \log M(0)$$

Ricordando che l'equazione canonica di una retta in assi cartesiani x e y è: $y = mx + q$, dove m è il coefficiente angolare e q l'intercetta sull'asse y delle ordinate, si può rappresentare graficamente la cinetica di amplificazione riportando in ordinate $\log M(n)$ e in ascisse n : si ottiene così una retta con coefficiente angolare pari a $\log (1 + E)$ e con intercetta $\log M(0)$. Questo tratto rettilineo (fase di linearità

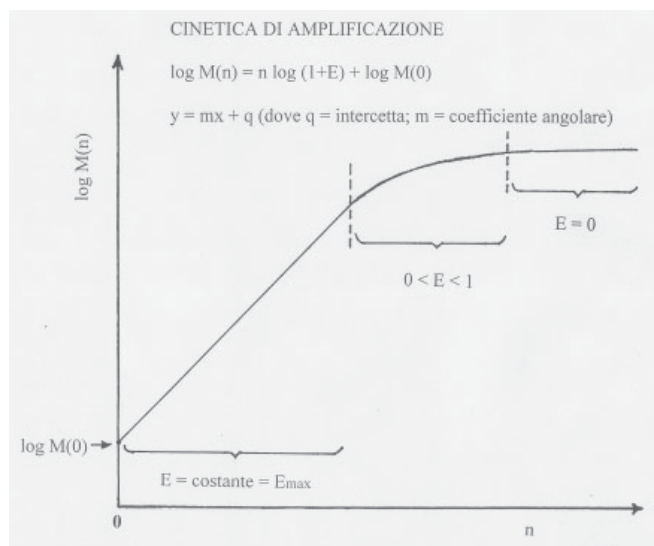


Figura 1. Cinetica di amplificazione della sequenza bersaglio, da una quantità molare iniziale $M(0)$ a una quantità molare finale $M(n)$. In ascisse è riportato il numero n di cicli termici, in ordinate il logaritmo della molarità. Il tratto rettilineo iniziale corrisponde a un valore costante e massimale di E (efficienza di amplificazione), che successivamente decresce gradualmente fino a raggiungere lo zero (tratto orizzontale della curva).

logaritmica), che si osserva nella prima parte del processo amplificativo, corrisponde a un valore di E pressochè costante e prossimo al valore massimo iniziale, idealmente pari a 1 in assenza di inibitori. All'aumentare di n , dopo un certo numero di cicli, l'efficienza di amplificazione E da costante diventa variabile (fase di transizione), diminuendo gradualmente fino a zero al crescere di n , e la retta diventa una curva, la cui tangente in ogni punto presenta coefficiente angolare pari a $\log(1 + E)$, gradualmente decrescente al diminuire di E ; quando $E = 0$ la tangente risulta parallela all'asse delle ascisse, avendo coefficiente angolare $\log(1+0) = \log(1) = 0$. A questo punto (fase di plateau) la curva ha assunto l'andamento di una retta parallela all'asse delle ascisse e il processo amplificativo è terminato (Fig. 1).

L'andamento testè descritto per la curva di amplificazione risulta diverso da quello tipicamente sigmoide, o a S, osservato sperimentalmente con tecniche di rivelazione simultanea quali la PCR in tempo reale (Fig. 2).

L'apparente incongruenza nella forma della curva di amplificazione quale risulta dall'espressione matematica e quale invece risulta dalla rivelazione strumentale "in tempo reale" è semplicemente dovuta al fatto che in quest'ultimo caso il sistema di rivelazione del segnale luminoso non è in grado, per il suo limite di sensibilità, di discriminare tra $\log M(0)$ e lo zero, come pure non è in grado di discriminare tra $\log M(n)$ e lo zero, fino a che $\log M(n)$ non superi un valore soglia rilevabile. Tutto il tratto di curva che precede tale valore soglia appare quindi, anziché una retta con intercetta $\log M(0)$ e coefficiente angolare $\log(1 + E)$, una retta pressochè coincidente con l'asse delle ascisse. La curva reale e quella ottenuta con la rilevazione strumentale coincidono solo dopo che il valore soglia di rilevazione è stato raggiunto e superato.

Dopo uno stesso numero di cicli amplificativi, in particolare partendo da differenti quantità iniziali di sequenza ber-

saglio, le curve di amplificazione di campioni diversi si potranno trovare in fasi differenti (in fase di linearità logaritmica, piuttosto che in fase di transizione o in fase di plateau). Ciò significa che il valore medio per ciclo dell'efficienza di amplificazione sarà diverso da campione a campione, e in particolare nei campioni giunti a plateau (dove $E = 0$) il valore medio di E sarà più basso che nei campioni che dopo lo stesso numero di cicli si trovano ancora nella fase di linearità logaritmica, caratterizzata da valori di E prossimi a 1. In altri termini, si avrà una sottostima dei primi, rispetto a una stima più corretta dei secondi. Se per tutti i campioni la lettura venisse fatta durante la fase di linearità logaritmica, la stima sarebbe corretta e nessuno verrebbe sottostimato, purchè però si potesse escludere la presenza di inibitori che possano diminuire il valore di E fin dal primo ciclo amplificativo, e quindi indipendentemente dalle condizioni che portano a plateau.

Nell'amplificazione ciclotermica in tempo reale la lettura, effettuata nel momento in cui il segnale fluorescente raggiunge un'intensità soglia discriminabile dallo zero, è quella teoricamente più precoce possibile e avviene, come si è detto, nella fase di linearità logaritmica, durante la quale il valore di E si mantiene costante e prossimo al suo valore massimo iniziale. Nell'istante in cui viene raggiunta la fluorescenza soglia il termine n dell'equazione cinetica coincide con C_t , dove C_t è il ciclo soglia, e l'equazione può essere riscritta nella forma:

$$M(C_t) = M(0)(1 + E)^{C_t}$$

La fluorescenza soglia è una costante per un dato tipo di

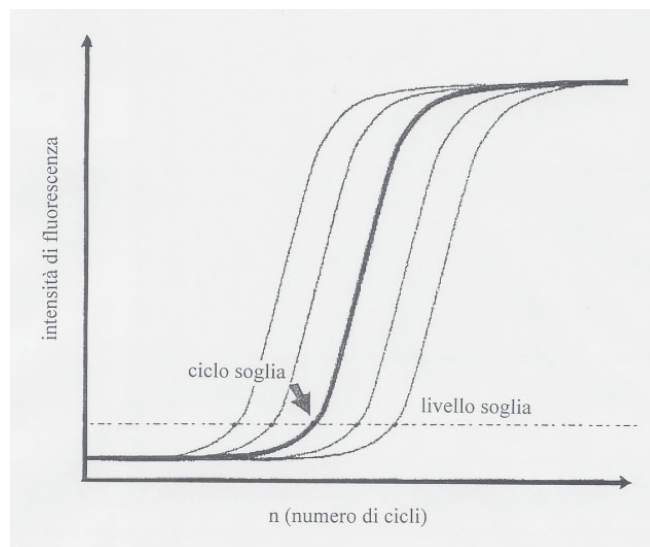


Figura 2. Cinetica di amplificazione quale si ottiene sperimentalmente con tecniche di rivelazione simultanea in fluorescenza. Il tratto orizzontale iniziale delle curve corrisponde a un valore zero strumentale. Lo strumento nelle fasi iniziali dell'amplificazione non è in grado di rivelare l'andamento esponenziale della cinetica amplificativa, poiché i segnali corrispondenti ai primi cicli sono di intensità inferiore a quella strumentalmente rilevabile. Allorchè la soglia di sensibilità strumentale viene raggiunta (ciclo soglia) e superata, lo strumento misura l'andamento rettilineo corrispondente a un valore costante e massimale di E (efficienza di amplificazione), che poi decresce fino a zero, del tutto come in Figura 1.

fluoroforo e corrisponde, come si è detto, a una precisa quantità molare di acido nucleico bersaglio. Per uno specifico tipo di fluoroforo vale la relazione:

$$M(Ct) = k \text{ (dove } k \text{ è una costante);}$$

ciò permette di ricavare direttamente il valore della quantità iniziale di bersaglio $M(0)$. Da:

$$M(Ct) = k = M(0)(1 + E)^{Ct}$$

si ottiene infatti: $M(0) = k / (1 + E)^{Ct} = k(1 + E)^{-Ct}$

Nel caso ideale in cui $E = 1$, il valore di $M(0)$ è facilmente ottenibile anche in assenza di una curva di calibrazione, essendo noto k e potendosi misurare Ct . Al di fuori di questo caso ideale, è comunque teoricamente possibile ottenere un valore medio di E utilizzando dei calibratori esterni in concentrazioni scalari: il valore di E può essere desunto dalla pendenza, o coefficiente angolare, che potrà variare tra una seduta analitica e l'altra, della retta di calibrazione. E' però una forzatura ipotizzare che il valore di E sia identico per tutti i campioni di una stessa seduta analitica; in linea generale bisogna invece ammettere che l'efficienza di reazione sia variabile da campione a campione: per ottenere $M(0)$ in queste condizioni si rende allora necessaria una strategia che consenta di eliminare la variabile E dall'equazione. Ciò è possibile attraverso l'introduzione in ogni singola miscela di reazione di un competitore, o calibratore, o controllo interno in quantità nota, che venga amplificato con la stessa cinetica del bersaglio. Uno standard amplificato a parte, ossia una curva di calibrazione esterna, può essere solo rappresentativa di un valore medio di E all'interno di una singola seduta analitica.

Utilità e caratteristiche del calibratore o riferimento interno⁷⁴⁻⁸²

Come si è ribadito, il risultato analitico può differire notevolmente in relazione al momento, durante il processo amplificativo, in cui esso viene rilevato (fase di linearità logaritmica, di transizione o di plateau), nonché all'efficienza di amplificazione iniziale, che può essere diminuita in presenza di inibitori. Per questi motivi è opportuno utilizzare un controllo o competitore interno. Per potersi definire tale, il controllo interno deve essere amplificato con la stessa efficienza della sequenza bersaglio.

Se $M(0)$ è la quantità iniziale di sequenza bersaglio presente nel campione, $C(0)$ la quantità iniziale di competitore (nota), $M(n)$ la quantità di sequenza bersaglio sintetizzata dopo n cicli di amplificazione e $C(n)$ la quantità di competitore sintetizzato dopo n cicli di amplificazione, si ha:

$$M(n) = M(0) (1 + E)^n \quad \text{e} \quad C(n) = C(0) (1 + E)^n$$

$$\text{Perciò } M(n)/C(n) = M(0) (1 + E)^n / C(0) (1 + E)^n$$

$$\text{e quindi } M(n) / C(n) = M(0) / C(0)$$

purchè E abbia lo stesso valore sia per C (calibratore) che per M (bersaglio): in queste condizioni dopo n cicli di amplificazione il rapporto quantitativo tra M e C rimane uguale a quello iniziale. Il risultato è indipendente dal numero dei cicli e dalla quantità di amplificato prodotto dalla reazione. Vengono così eliminate tutte le variabili che si riflettono sull'efficienza di reazione (composizione molare della miscela di reazione, inattivazione termica della DNA-

polimerasi o DNA-ligasi, presenza di inibitori, attività nucleasiche, ecc.). L'equazione finale non vale, ovviamente, se i valori di efficienza di amplificazione per standard e bersaglio sono diversi.

Per avere la stessa efficienza di amplificazione della sequenza bersaglio, il competitore deve avere caratteristiche molto simili a quest'ultima. La sua lunghezza deve differire il meno possibile da quella della sequenza bersaglio, poiché il tempo impiegato per la reazione enzimatica di polimerizzazione varia in funzione della lunghezza e ciò può tradursi in una diversa efficienza di amplificazione.

Anche la composizione in basi del competitore è importante, poiché dal rapporto tra basi puriniche e pirimidiniche dipende la temperatura di denaturazione-rinaturazione della doppia elica del competitore.

La sequenza del competitore deve essere ovviamente diversa, anche se per poche basi, da quella bersaglio, in quanto deve essere distinguibile nella fase di rivelazione. E' bene tuttavia che le differenze di sequenza non riguardino le regioni complementari agli inneschi, in modo che sia possibile utilizzare gli stessi inneschi sia per il bersaglio che per il competitore: ciò rende più probabile una uguale efficienza di amplificazione. Nel caso della PCR in tempo reale il competitore può essere rivelato con sonde fluorescenti che emettono a una lunghezza d'onda diversa da quella utilizzata per rivelare il DNA bersaglio⁷⁹, in modo che i due segnali non si sovrappongano tra loro.

Se il processo di amplificazione è preceduto da una fase di retrotrascrizione (come nel caso di HCV-RNA), il competitore interno dovrà essere utilizzato già a partire da questa fase. L'utilizzo di un competitore interno di DNA dopo la retrotrascrizione è metodologicamente scorretto e può portare a risultati altamente imprecisi, poiché anche l'efficienza della reazione di retrotrascrizione può discostarsi dai valori massimi teorici: questa variabile può essere eliminata solo retrotrascrivendo nella stessa miscela di reazione la sequenza bersaglio e un competitore interno. Affinchè sia assicurata una uguale efficienza di retrotrascrizione, la sequenza bersaglio e il competitore dovranno utilizzare lo stesso iniziatore (primer) per la trascrittasi inversa.

Nel processo analitico di determinazione quantitativa una variabile minore, ma non trascurabile, è rappresentata anche dall'efficienza di estrazione dell'acido nucleico. L'aggiunta del calibratore interno a quantità nota fin dalla fase di estrazione consente di monitorare l'intero procedimento⁷⁸, eliminando anche questa variabile. L'efficienza di estrazione, alta o bassa che sia, sarà la stessa per la sequenza bersaglio e per il competitore, perciò, al termine del processo estrattivo, il rapporto finale tra quantità di sequenza bersaglio e quantità di competitore sarà identico a quello iniziale, poiché, per le caratteristiche chimicofisiche del competitore, il processo di estrazione non potrà discriminare tra le due specie di acido nucleico.

E' da ultimo opportuno sottolineare la necessità di effettuare tutti i controlli intesi ad accertare la corretta concentrazione, nonché la perfetta integrità del competitore.

In appendice si potrebbero fare alcune considerazioni relative all'errore intrinseco, o incertezza teorica, nella misura.

A livello molecolare la quantità di prodotto di una reazione

ne chimica è legata alla probabilità con cui ogni singola molecola tende a reagire. Quando la popolazione molecolare è ampia la legge dei grandi numeri genera un comportamento di tipo deterministico. Una frazione delle molecole reagisce e le fluttuazioni di questa frazione si fanno trascurabili fino a ridursi a zero man mano che la popolazione molecolare aumenta e tende a infinito. Quando però la popolazione molecolare è molto piccola l'aspetto probabilistico si fa più evidente, ossia è possibile osservare grossolane fluttuazioni nella frazione di molecole che reagiscono (comportamento di tipo stocastico)⁷¹.

Nella reazione polimerasica a catena ciascuna molecola di DNA può essere duplicata non più di una volta con una probabilità p durante un singolo ciclo di amplificazione. Il fattore di amplificazione m , corrispondente al coefficiente $1+E$ dell'equazione cinetica, è legato alla probabilità p dalla semplice relazione $m = 1+p$. In realtà E e p sono la stessa cosa, poiché l'efficienza di amplificazione E corrisponde a livello molecolare alla probabilità di reazione, ovvero di duplicazione, che varia tra 0 e 1.

Quando $p = 1$ (probabilità massima, pari al 100%) si ha $m = 2$, ossia il raddoppio delle molecole di DNA dopo il ciclo amplificativo. Per $p = 0,5$ (probabilità pari al 50%) si ha $m = 1,5$ (incremento della massa di prodotto pari al 50% dopo il ciclo amplificativo). Si può dimostrare che il livello di incertezza, ossia il grado di errore nella misura, a parità di popolazione di partenza, decresce al crescere di m ; per un valore di m costante il grado di errore aumenta al diminuire della popolazione molecolare, raggiungendo un valore massimo per una popolazione costituita da un'unica molecola di DNA.

Per fare un esempio numerico, se indichiamo con $N(0)$ la popolazione iniziale, per $N(0) = 1$ e $m = 1,5$ si ottiene una incertezza relativa teorica pari al 255%, che si riduce al 99,1% per $N(0) = 1$ e $m = 1,9$. Per $N(0) = 100$ e $m = 1,5$ si ottiene una incertezza teorica pari al 25%, che si riduce al 10% per $N(0) = 100$ e $m = 1,9$ ⁷¹.

Analisi quantitativa dell'espressione genica: le "sonde a schiera" o "microarray"

Il termine "microschiera" (come anche "sonde a schiera") è una possibile traduzione dell'inglese "microarray"⁸³⁻¹⁰⁶, che sta per "microscopic glass array", ossia disposizione ordinata, dispiegamento, ovvero schieramento (atto ed effetto dello schierare, ordinamento o disposizione in schiera) su un vetrino da microscopio di elementi che consentono il legame specifico di geni o prodotti genici. Esso rende conto dell'architettura di uno strumento analitico di estrema sofisticazione, dove le molecole legate al supporto solido sono disposte esattamente come un esercito schierato pronto alla battaglia. Secondo il vocabolario italiano, schiera è una moltitudine ordinata in lunghe file una accanto all'altra o comunque raggruppata secondo determinati o prestabiliti criteri. Un termine alternativo potrebbe forse essere quello di "griglia". Ogni posizione dello schieramento, diviso in righe e colonne, è occupata da una singola specie molecolare di identità nota (acido nucleico o proteina o tessuto, a seconda del tipo di "microschiera") che, legandosi a molecole (acidi nucleici o proteine) presenti nel campione saggiato fornirà informazioni di tipo qualitativo e, in determinate condizioni, anche di tipo quantitativo, nel-

l'ambito della ricerca biologica, come nell'analisi di vie metaboliche, dell'espressione genica, dell'interazione tra proteine, dei polimorfismi, nonché di mutazioni, riarrangiamenti e delezioni a livello genomico, per citarne solo alcune. Le applicazioni sono sempre più numerose e spaziano dalla ricerca e diagnostica in ambito oncologico, alla microbiologia, alla fisiologia, alla farmacogenomica, a svariate branche della medicina interna.

La tecnologia della "microschiera" è stata sviluppata a partire dalla metà degli anni Novanta all'Università di Stanford ad opera di Schena e coll. ed è frutto dell'interazione tra varie discipline, come biologia, chimica, fisica, ingegneria, matematica e informatica. Lo sviluppo dei supporti a microschiera è andato di pari passo con quello dei microprocessori a semiconduttore, attualmente capaci di operare miliardi di operazioni al secondo (capacità operative dell'ordine dei gigahertz) su circuiti miniaturizzati delle dimensioni di frazioni di micrometro. I primi supporti a microschiera, nel 1995, contenevano 96 geni, legati ognuno a una superficie del diametro di 200 micrometri; nel 2001 contenevano già 30.000 geni, ognuno occupante una superficie del diametro di 16 micrometri. Attualmente sono disponibili "microarray" per l'intero genoma umano, che rappresentano circa 41.000 geni e trascritti.

Esistono diversi tipi di "microarray" ("microschiere"):

- 1) "schiere" a cloni di DNA micropipettati, dove in ogni posizione del supporto (un vetrino da microscopio) viene impresso e legato uno specifico clone di cDNA (DNA retrotrascritto su uno specifico mRNA). La lunghezza dei cDNA utilizzati è tipicamente di 500-2500 coppie di basi. I primi esperimenti furono fatti con questo tipo di microschiere, tuttora ampiamente usate nei saggi di espressione genica intesi a valutare la quantità di specifici mRNA trascritti a partire da un complesso di geni. Circa il 65% di tutte le pubblicazioni riguardanti lavori in cui si sono usate le schiere riguardano la tipologia a cDNA.
- 2) "schiere" di oligonucleotidi sintetizzati *in situ*. La sintesi degli oligonucleotidi, a singola elica e di lunghezza pari a 15-80 (più spesso intorno a 25) nucleotidi è ottenuta associando un procedimento di sintesi chimica con tecniche fotolitografiche che ricalcano quelle utilizzate nella fabbricazione miniaturizzata dei semiconduttori (microprocessori). La tecnologia fotolitografica è assimilabile a quella di una macchina stampatrice. Vengono prefabbricate fino a 100 maschere fotolitografiche e le quattro basi del DNA, ossia A, T, G, C sono poi stampate in "fogli" sovrapposti strato su strato fino a formare oligonucleotidi della lunghezza voluta. Il processo di fabbricazione del tipo a getto di inchiostro "stampa" le quattro basi nucleotidiche in maniera non molto diversa da quella di una stampante a quattro colori. Le testine di stampa rilasciano una quantità controllata di A, T, G, C. Una goccia si sovrappone all'altra, sintetizzando chimicamente un filamento di DNA. Per ogni specifico tipo di schiera si dovrà caricare il corrispondente "file" (insieme dei dati operativi) nell'elaboratore che controlla il sistema di stampa. Le schiere a oligonucleotidi sono più costose, ma possono raggiungere una densità di schieramento molto elevata e presentano notevole versatilità, in particolare nello studio dei profili di espres-

- sione genica e per la genotipizzazione. Circa un quarto del totale delle pubblicazioni riguardanti studi con microschiere concerne questa tipologia^{103,104}.
- 3) “schiere” per proteine. Sono più recenti, e ciò vale anche per quelle di tessuti, rispetto alle schiere con acido nucleico. Entrambe queste più nuove tipologie vengono utilizzate con sempre maggior frequenza e a tutt’oggi sono riportate, considerandole insieme, in quasi il 10% delle pubblicazioni che riguardano studi su microschiara. Al supporto solido possono essere legate specifiche proteine oppure specifiche molecole leganti proteine, come DNA, RNA, anticorpi o altri ligandi. Le schiere a proteine^{105,106} sono utilizzate per studiare l’attività biochimica di una proteina o l’interazione tra proteine, mentre le schiere a molecole leganti proteine per analizzare il profilo di espressione genica, intesa come sintesi proteica, cellulare o tissutale.
 - 4) “schiere” di tessuti. Sono costituite da fettine sottili (5 micrometri) di tessuto normale o patologico, in particolare neoplastico, fissate in formalina e fatte aderire su un vetrino, che di solito ne contiene un numero compreso tra 500 e 1000. Sono utilizzate per evidenziare in specifici tessuti normali o patologici la presenza di determinati DNA, RNA o proteine mediante ibridazione *in situ* o immunohistochimica.

Come funzionano le “schiere” molecolari?

Le sonde molecolari (DNA, proteine o altri leganti) schierate secondo un ordine noto sul supporto solido vengono cementate con acido nucleico o proteine del campione oggetto di studio, previamente marcati con un colorante fluorescente. Con uno scansore sarà possibile misurare in ogni punto della schiera l’intensità luminosa, ottenendo un valore numerico che esprime la quantità di molecola marcata legata ad ogni specifica sonda. E’ così possibile, per esempio, ibridando le sonde a DNA della schiera con mRNA estratto dalle cellule e marcato con fluorocromo, ottenere un valore numerico di espressione genica per ciascun gene rappresentato sulla schiera.

Le schiere ad alta densità (10.000-30.000 geni) trovano principalmente applicazione per vedere differenze qualitative nell’espressione genica e per scoprire quali geni sono coinvolti in un determinato processo. Le schiere a bassa densità (100-500 geni) sono invece più adatte per studiare l’espressione genica in termini propriamente quantitativi.

In qualsiasi tipo cellulare solo il 10-30% dei geni è espresso o attivo, cioè viene trascritto in RNA messaggero. Lo studio di quali geni sono iperespressi o ipoattivi nelle cellule malate aiuta a chiarire il meccanismo patogenetico e a indirizzare verso la terapia più efficace. Il quadro di espressione genica può peraltro in determinati casi essere patognomonico e come tale trovare specifico utilizzo nella medicina diagnostica.

Una corretta quantificazione di una specie molecolare bersaglio presente nel campione analizzato su microschiara richiede che vi sia una relazione di perfetta proporzionalità tra la quantità di molecola bersaglio presente nel campione e la quantità di molecola bersaglio legata dalla specifica sonda sulla schiera: ciò si traduce in un segnale luminoso di intensità direttamente proporzionale alla quantità di bersa-

glia. Questa condizione si verifica solo quando la quantità di sonda schierata specifica per il bersaglio è in eccesso rispetto alla quantità di bersaglio. Allorché invece la quantità di bersaglio supera la capacità legante della sonda, cosicché quest’ultima viene saturata, vale a dire allorché il bersaglio è in eccesso rispetto alla sonda, si ottiene un segnale luminoso corrispondente a una risposta massimale che non aumenta più di intensità all’aumentare del numero di molecole bersaglio nel campione: in questo caso l’analisi quantitativa non è possibile, poiché la risposta non è più proporzionale, ovvero, come si suol dire, non è più di tipo lineare. La condizione di eccesso di molecola bersaglio va sempre tenuta in considerazione, in particolare quando si analizzino miscele complesse, come quelle di mRNA totale: in queste ultime le differenze quantitative tra le singole specie di mRNA, ossia di mRNA trascritti a partire da geni diversi, possono essere enormi. Gli mRNA più abbondanti possono rappresentare anche l’1% del totale, mentre trascritti più rari possono non raggiungere neppure lo 0,001% del totale. Ciò porta a una saturazione selettiva delle sonde schierate specifiche per le specie abbondanti, che così non possono essere correttamente quantificate, mentre le sonde specifiche per i trascritti meno abbondanti si possono mantenere in eccesso, il che permette di conservare le condizioni di linearità nella misura. In condizioni di saturazione selettiva delle sonde i dati ottenibili risultano essere quantitativi per le specie poco rappresentate e selettivamente non quantitativi per le specie più abbondanti. Questo effetto, noto come compressione del segnale, può portare a una lettura erronea sulle fasce alte di concentrazione, ciò che si cerca di evitare in tutte escluse poche situazioni sperimentali. In alcune situazioni l’eccesso di bersaglio viene esplicitamente ricercato, per esempio per misurare la densità assoluta di sonda sulla schiera. In questo tipo di analisi si cerca di ottenere la saturazione completa della sonda, corrispondente alla concentrazione di bersaglio specifico oltre la quale non si verifica alcun aumento nell’intensità del segnale luminoso; confrontando quest’ultimo con quello di una quantità nota di marcatore fluorescente è possibile calcolare il numero assoluto di molecole sonda per unità di area sulla schiera, ossia la densità di sonda. Una condizione di saturazione selettiva delle sonde può rendersi necessaria per misurare la concentrazione di specie rare in miscele complesse, come i trascritti di geni poco espressi. In questi casi la schiera verrà cementata con forti quantità di mRNA totale, tali da saturare le sonde per i trascritti più abbondanti, fornendo al contempo quantità rivelabili di trascritti rari, che trovando un eccesso di sonda specifica possono venire correttamente quantificati. Nelle condizioni in cui si ha una lettura lineare per i trascritti più rappresentati i trascritti più rari potrebbero infatti generare un segnale eccessivamente debole o non rilevabile. Per poter confrontare, dopo averli memorizzati in forma digitale, i dati ottenuti da campioni biologici diversi e da esperimenti differenti, occorre disporre di segnali la cui intensità possa essere presa come termine di riferimento alla quale rapportare tutte le altre. Come riferimento si può utilizzare il segnale corrispondente ai cosiddetti “geni costitutivi” (housekeeping genes)¹⁰⁷, ossia a quei geni che, esplicando un ruolo centrale nel metabolismo cellulare, vengono espressi all’incirca in egual modo in tutte le cellule

e tessuti. In alternativa si possono costruire delle curve di calibrazione aggiungendo al materiale oggetto di analisi delle quantità note di specifici mRNA utilizzati come controllo. Le sonde specifiche sulla schiera, che dovranno ovviamente essere presenti in eccesso rispetto al bersaglio, legheranno quest'ultimo, in quantità nota, e le corrispondenti fluorescenze potranno essere utilizzate come termini di paragone: si otterrà una serie di segnali corrispondenti a quantità note di mRNA. Ciò permette una quantificazione sia assoluta che relativa. La curva standard, o di calibrazione, può servire anche a correggere l'effetto di compressione del segnale che si verifica in caso di eccesso di molecola bersaglio rispetto alla sonda.

I "microarray" costituiscono lo strumento più sofisticato e versatile per l'analisi quantitativa di geni e prodotti genici. Con una sola analisi è possibile raccogliere una quantità impressionante di dati, che può essere archiviata in forma digitale e confrontata con dati precedenti o successivi per poter dedurre, per esempio, quali geni sono coinvolti, e in quale misura, in determinati processi fisiologici, patologici o indotti da farmaci.

Queste "microschiere" molecolari possono essere utilizzate per analizzare l'intero genoma umano in un solo passaggio e per ottenere una informazione quantitativa sull'espressione genica per 30.000 geni umani in meno di 10 minuti. Citando le parole di Mark Schena, l'inventore delle "schiere" molecolari, "l'analisi con "microarray" è unica nella storia della biologia, perché nessun'altra tecnologia ha utilizzato un tale grado di sofisticazione e attento conoscenze da così tante discipline diverse, fornendo una visione altrettanto dettagliata della cellula."

Conclusioni

L'evoluzione tecnologica ha messo al servizio del ricercatore e del patologo strumenti di assai grande versatilità, in grado di dare risposte sempre più rapide e precise. L'analisi della "filosofia" del procedimento analitico, intesa come il percorso logico ad esso sotteso, e quella dei suoi punti critici non può che aiutare a meglio interpretare il risultato, nonché a realizzare e mettere in atto le migliori strategie operative possibili. Nessuna tecnologia va presa a scatola chiusa o peggio ancora subita: l'utilizzatore avrà tutti i vantaggi ad acquisire la massima padronanza, soprattutto teorica, del sistema in uso, sforzandosi di conoscerne e valutarne nella maniera più analitica i vari aspetti. Sarà così più agevole risolvere eventuali problemi operativi e sicuramente più soddisfacente lo stesso lavoro. La bibliografia che segue può servire da traccia per i necessari approfondimenti.

Bibliografia

- Burton K. A study of the conditions and mechanism of the diphenylamine reaction for the colorimetric estimation of DNA. *Biochem J* 1956; 62: 315-23.
- Kissane JM, Robins E. The fluorometric measurement of deoxyribonucleic acid in animal tissues with special reference to the central nervous system. *J Biol Chem* 1958; 233: 184-8.
- Hinegardner RT. An improved fluorometric assay for DNA. *Anal Biochem* 1971; 39: 197-201.
- Covey SN, Grierson D. The measurement of plant polyadenylic acid by hybridisation with radioactive polyuridylic acid. *Planta* 1976; 131: 75-9.
- Bishop JO, Rosbash M, Evans D. Polynucleotide sequences in eukariotic DNA and RNA that form ribonuclease-resistant complexes with polyuridylic acid. *J Mol Biol* 1974; 85: 75-86.
- Rabinow P. Making PCR: a story of biotechnology. University of Chicago Press, Chicago, 1996.
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol* 1986; 51: 263-73.
- Mullis K, Faloona F. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* 1987; 155: 335-50.
- Kleppe K, Ohtsuka E, Kleppe R, Molineux I, Khorana HG. Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalysed by DNA polymerases. *J Mol Biol* 1971; 56:341-61.
- Gilliland G, Perrin S, Blanchard K, Bunn HF. Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. *Proc Natl Acad Sci* 1990; 87: 2725-9.
- Kellogg DE, Sninsky JJ, Kwok S. Quantitation of HIV-1 proviral DNA by the polymerase chain reaction. *Anal Biochem* 1990; 189: 202-8.
- Higuchi R, Dollinger G, Walsh PS, Griffith R. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 1992; 10: 413-7.
- Higuchi R, Fockler C, Dollinger G, Watson R. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology* 1993; 11:1026-30.
- Holland PM, Abramson RD, Watson R, Gelfand DH. Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci* 1991; 88: 7276-80.
- Lee LG, Connell CR, Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids* 1993; 21: 3761.
- Wilber JC, Urdea MS. Quantification of viral nucleic acids using branched DNA signal amplification. In: *Molecular Methods for Virus Detection*. New York: Academic Press; 1996.p. 131-44.
- Detmer J, Lagier R, Flynn J, Zayat C, Kolberg J, Collins M, et al. Accurate quantification of hepatitis C virus RNA from all genotypes by using branched-DNA technology. *J Clin Microbiol* 1996; 34:901-7.
- Kern D, Collins M, Fultz T, Detmer J, Hamren S, Peterkin JJ, et al. An enhanced-sensitivity branched-DNA assay for quantification of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol* 1996; 34: 3196-202.
- Ross RS, Viazov S, Sarr S, Hoffmann S, Kramer A, Roggendorf M. Quantitation of hepatitis C virus RNA by third generation branched DNA-based signal amplification assay. *J Virol Methods* 2002; 101:159-68.
- Elbeik T, Surtihadi J, Destree M, Gorlin J, Holodniy M, Jortani SA, et al. Multicenter evaluation of the per-

- formance characteristics of the Bayer VERSANT HCV RNA 3.0 assay (bDNA). *J Clin Microbiol* 2004; 42: 563-9.
21. Yao JD, Beld MG, Oon LL, Sherlock CH, Germer J, Menting S, et al. Multicenter evaluation of the VERSANT hepatitis B virus DNA 3.0 assay. *J Clin Microbiol* 2004; 42: 800-6.
 22. Langabeer SE, Gale RE, Harvey RC, Cook RW, Mackinnon S, Linch DC. Transcription-mediated amplification and hybridisation protection assay to determine BCR-ABL transcript levels in patients with chronic myeloid leukaemia. *Leukemia* 2002; 16: 393-9.
 23. de Baar MP, van Dooren MW, de Rooij E, Bakker M, van Gemen B, Goudsmit J, et al. Single rapid real-time monitored isothermal RNA amplification assay for quantification of human immunodeficiency virus type 1 isolates from groups M, N, and O. *J Clin Microbiol* 2001; 39: 1378-84.
 24. Deiman B, van Aarle P, Sillekens P. Characteristics and applications of nucleic acid sequence-based amplification (NASBA). *Mol Biotechnol* 2002; 20: 163-79.
 25. Yates S, Penning M, Goudsmit J, Frantzen I, van de Weijer B, van Strijp D, et al. Quantitative detection of hepatitis B virus DNA by real-time nucleic acid sequence-based amplification with molecular beacon detection. *J Clin Microbiol* 2001; 39: 3656-65.
 26. Weusten JJAM, Carpay WM, Oosterlaken TAM, van Zuijlen MCA, van de Wiel PA. Principles of quantitation of viral loads using nucleic acid sequence-based amplification in combination with homogeneous detection using molecular beacons. *Nucleic Acids Res* 2002; 30:No.6e26.
 27. Leone G, van Schijndel H, van Gemen B, Kramer FR, Schoen CD. Molecular beacon probes combined with amplification by NASBA enable homogeneous, real-time detection of RNA. *Nucleic Acids Res* 1998; 26:2150-5.
 28. Greijer AE, Adriaanse HM, Kahl M, Tacken NM, Oldenburg N, Sijlmans A, et al. Quantitative competitive NASBA for measuring mRNA expression levels of the immediate early 1, late pp67, and immune evasion genes US3, US6 and US11 in cells infected with human cytomegalovirus. *J Virol Methods* 2001; 96: 133-47.
 29. Damen M, Sillekens P, Cuypers HT, Frantzen I, Melsert R. Characterization of the quantitative HCV NASBA assay. *J Virol Methods* 1999; 82:45-54.
 30. van Deursen PBH, Gunther AW, Spaarganen-van Riel CC, van den Eijnden MMED, Vos HL, van Gemen B, et al. A novel quantitative multiplex NASBA method: application to measuring tissue factor and CD14 mRNA levels in human monocytes. *Nucleic Acids Research* 1999; 27: e15 1-6.
 31. Weusten JJAM, Wouters PAWM, van Zuijlen MCA, van de Wiel PA. Stochastic processes defining sensitivity and variability of internally calibrated quantitative NASBA-based viral load assays. *Nucleic Acids Res* 2002; 30:No.24e137.
 32. Verjat T, Cerrato E, Jacobs M, Leissner P, Mougin B. Multiparametric duplex real-time nucleic acid sequence-based amplification assay for mRNA profiling. *Biotechniques* 2004; 37: 476-81.
 33. Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, et al. Molecular beacons as diagnostic tools: technology and applications. *Clin Chem Lab Med* 2003; 41:468-74.
 34. Tyagi S, Kramer FR. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 1996; 14:303-8.
 35. Marin MG. Diagnostica di laboratorio: tecniche di amplificazione genica. Edizioni Sorbona Milano 1999, pp XVI, 1-224.
 36. Barany F. Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc Natl Acad Sci USA* 1991; 88:189-93.
 37. Birkenmeyer LG, Mushahwar I. DNA probe amplification methods. *J Virol Methods* 1991; 35:117-26.
 38. Lee H. Infectious disease testing by ligase chain reaction. *Clin Chemistry* 1993; 39:729-30.
 39. Marshall RL, Laffler TG, Cerney MB, Sustachek JC, Kratochvil JD, Morgan RL. Detection of HCV RNA by the asymmetric gap ligase chain reaction. *PCR Methods Appl* 1994; 4:80-4.
 40. Hsuih TCH, Park YN, Zaretsky C, Wu F, Tyagi S, Kramer FR, et al. Novel, ligation-dependent PCR assay for detection of hepatitis C virus in serum. *J Clin Microbiol* 1996; 34:501-7.
 41. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res* 1996; 6:986-94.
 42. Mackay IM, Arden KE, Nitsche A. Real-time PCR in virology. *Nucleic Acids Research* 2002; 30:1292-305.
 43. Niesters HG. Clinical virology in real time. *J Clin Virol* 2002; 25(suppl 3):3-12.
 44. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000; 25: 169-93.
 45. Morcellin S, Rossi CR, Pilati P, Nitti D, Marincola FM. Quantitative real-time PCR: a powerful ally in cancer research. *Trends Mol Med* 2003; 9: 189-95.
 46. Payungporn S, Tangkijvanich P, Jantaradsamee P, Theamboonlers A, Poovorawan Y. Simultaneous Quantitation and genotyping of hepatitis B virus by real-time PCR and melting curve analysis. *J Virol Methods* 2004; 120: 131-40.
 47. Candotti D, Temple J, Owusu-Ofori S, Allain JP. Multiplex real-time quantitative RT-PCR assay for hepatitis B virus, hepatitis C virus, and human immunodeficiency virus type 1. *J Virol Methods* 2004; 118: 39-47.
 48. Improved quantitative real-time RT-PCR for expression profiling of individual cells. *Nucleic Acids Res* 2002; 30: e89-89.
 49. Komurian-Pradel F, Paranhos-Baccalà G, Sodoyer M, Chevallier P, Mandrand B, Lotteau V, et al. Quantitation of HCV RNA using real-time PCR and fluorimetry. *J Virol Methods* 2001; 95:111-9.
 50. Niesters HG. Quantitation of viral load using real-time amplification technology. *Methods* 2001; 25:419-29.
 51. Yang JH, Lai JP, Douglas SD, Metzger D, Zhu XH,

- Ho WZ. Real-time RT-PCR for quantitation of hepatitis C virus RNA. *J Virol Methods* 2002; 102:119-28.
52. Livak KJ, Flood SJA, Marmaro J, Giusti W, Deetz K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl* 1995; 4:357-62.
53. Enomoto M, Nishiguchi S, Shiomi S, Tanaka M, Fukuda K, Ueda T, et al. Comparison of real-time quantitative polymerase chain reaction with three other assays for quantitation of hepatitis C virus. *J Gastroenterol Hepatol* 2001; 16:904-9.
54. Kleiber J, Walter T, Haberhausen G, Tsang S, Babel R, Rosenstraus M. Performance characteristics of a quantitative homogeneous TaqMan RT-PCR test for HCV RNA. *J Mol Diagn* 2000; 2:158-66.
55. Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, et al. Molecular beacons as diagnostic tools: technology and applications. *Clin Chem Lab Med* 2003; 41:468-74.
56. Tyagi S, Marras SAE, Kramer FR. Wavelength-shifting molecular beacons. *Nature Biotechnology* 2000; 18: 1191-6.
57. Sum SS, Wong DK, Yuen MF, Yuan HJ, Yu J, Lai CL, et al. Real-time PCR assay using molecular beacon for quantitation of hepatitis B virus DNA. *J Clin Microbiol* 2004; 42(8): 3438-40.
58. Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, et al. Molecular beacons as diagnostic tools: technology and applications. *Clin Chem Lab Med* 2003; 41:468-74.
59. Bonnet G, Tyagi S, Libchaber A, Kramer FR. Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc Natl Acad Sci USA* 1999; 96: 6171-6.
60. Whitcombe D, Theaker J, Guy SP, Brown T, Little S. Detection of PCR products using self-probing amplicons and fluorescence. *Nat Biotechnol* 1999; 17:804-7.
61. Nazarenko IA, Bhatnager SK, Hohman RJ. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res* 1997; 25: 2516-21.
62. Nazarenko I, Lowe B, Darfler M, Ikononi P, Schuster D, Rashtchian A. Multiplex quantitative PCR using self-quenched primers labeled with a single fluorophore. *Nucleic Acids Res* 2002; 30: n. 9e37.
63. Foerster T. Zwischenmolekulare energiewanderung und fluoreszenz. *Ann phys* 1948; 6: 55-75.
64. Foerster T. In: *Modern Quantum Chemistry*. New York: Sinanoglu Academic. 1965; pp. 93-137.
65. Ha T, Enderle T, Ogletree DF, Chemla DS, Selvin PR, Weiss S. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl Acad Sci USA* 1996; 93: 6264-8.
66. Clegg RM. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol* 1992; 211:353-88.
67. Selvin P. Fluorescence resonance energy transfer. *Methods Enzymol* 1995; 246:300-34.
68. Didenko V V. DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *Biotechniques* 2001; 31:1106-21.
69. Lee LG, Livak KJ, Mullah B, Graham RJ, Vinayak RS, Woudenberg TM. Seven-color, homogeneous detection of six PCR products. *Biotechniques* 1999; 27:342-49.
70. Peccoud J, Jacob C. Statistical estimations of PCR amplification rates. *C.R. Acad. Sci. Paris Serie I*, 1996; 322: 736-68.
71. Peccoud J, Jacob C. Theoretical uncertainty of measurements using quantitative polymerase chain reaction. *Biophys J* 1996; 71:101-8.
72. Freeman WM, Walker SJ, Vrana KE. Quantitative RT-PCR: pitfalls and potential. *Biotechniques* 1999; 26:112-25.
73. Orlando C, Pinzani P, Pazzagli M. Developments in quantitative PCR. *Clin Chem Lab Med* 1998; 36:255-69.
74. Meijerink J, Mandigers C, van de Locht L, Toennissen E, Goodsaid F, Raemaekers J. A novel method to compensate for different amplification efficiencies between patient DNA samples in quantitative real-time PCR. *J Mol Diagn* 2001; 3: n.2.
75. Clementi M, Menzo S, Manzin A, Bagnarelli P. Quantitative molecular methods in virology. *Arch Virol* 1995; 140:1523-39.
76. Kaneko S, Murakami S, Onora M, Kobayashi K. Quantitation of hepatitis C virus RNA by competitive polymerase chain reaction. *J Med Virol* 1992; 37:278-82.
77. Clementi M, Bagnarelli P, Manzin A, Menzo S. Competitive polymerase chain reaction and analysis of viral activity at the molecular level. *Genet Anal Tech Appl* 1994; 11:1-6.
78. Miyachi H, Masukawa A, Ohshima T, Fusegawa H, Hirose T, Impraim C, et al. Monitoring of inhibitors of enzymatic amplification in polymerase chain reaction and evaluation of efficacy of RNA extraction for the detection of hepatitis C virus using the internal control. *Clin Chem Lab Med* 1998; 36:571-5.
79. Gruber F, Falkner FG, Dorner F, Haemmerle T. Quantitation of viral DNA by real-time PCR applying duplex amplification, internal standardization and two-colour fluorescence detection. *Appl Environ Microbiol* 2001; 67:2837-9.
80. Ovstebo R, Haug KBF, Lande K, Kierulf P. PCR-based calibration curves for studies of quantitative gene expression in human monocytes: development and evaluation. *Clin Chem* 2003; 49: 425-32.
81. Castelain S, Descamps V, Thibault V, Francois C, Bonte D, Morel V, et al. TaqMan amplification system with an internal positive control for HCV RNA quantitation. *J Clin Virol* 2004; 31: 227-34.
82. Vandesompele J, De Preter K, Pattyn F, Poppe B, van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 2002; 3: 1-11.
83. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a

- complementary DNA microarray. *Science* 1995; 270: 467-70.
84. Schena M. *Microarray analysis*. Wiley-Liss ed. 2003; pp XIV 1-630.
 85. Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet* 2001; 2: 418-27.
 86. Draghici S. *Data analysis tools for DNA microarrays*. Chapman and Hall/CRC Press 2003.
 87. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet* 1999; 21: 10-4.
 88. De Risi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 1997; 278: 680-6.
 89. Yang YH, Speed T. Design issues for cDNA microarray experiments. *Nat Rev Genet* 2002; 3(8): 579-88.
 90. Churchill GA. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 2002; 32(Suppl): 490-5.
 91. Hartman JL IV, Tippery NP. Systematic quantification of gene interactions by phenotypic array analysis. *Genome Biology* 2004; 5: R49(1-18).
 92. Steinmetz LM, Davis RW. Maximizing the potential of functional genomics. *Nat Rev Genet* 2004; 5: 190-201.
 93. Csete ME, Doyle JC. Reverse engineering of biological complexity. *Science* 2002; 295: 1664-9.
 94. Affymetrix. *Gene chip expression analysis: technical manual*. <http://www.affymetrix.com>
 95. Affymetrix. *Microarray suite 5.0 user's guide 2002 edition*. Edited by: Affymetrix, Santa Clara, CA, USA, Affymetrix Inc, 2001.
 96. Quackenbush J. Microarray data normalization and transformation. *Nat Genet* 2002; 32(Suppl): 496-501.
 97. Gracey AY, Cossins AR. Application of microarray technology in environmental and comparative physiology. *Annu Rev Physiol* 2003; 65: 231-59.
 98. Eisen M, Brown P. DNA arrays for analysis of gene expression. *Methods in Enzymol* 1999; 303: 179-205.
 99. Hessner MJ, Wang X, Khan S, Meyer L, Schlicht M, Tackes J, et al. Three color cDNA microarrays: quantitative assessment through the use of fluorescein-labeled probes. *Nucleic Acids Res* 2003; 31: e14.
 100. Wang X, Ghosh S, Guo S-W. Quantitative quality control in microarray image processing and data acquisition. *Nucleic Acids Res* 2001; 29: E75-82.
 101. Wang X, Hessner MJ, Wu Y, Pati N, Ghosh S. Quantitative quality control in microarray experiments and the application in data filtering, normalization and false positive rate prediction. *Bioinformatics* 2003; 19: 1341-7.
 102. Park T, Yi SG, Kang SH, Lee S, Lee YS, Simon R. Evaluation of normalization methods for microarray data. *BMC Bioinformatics* 2003; 4: 33.
 103. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet* 1999; 21: 20-4.
 104. Draghici S. Statistical intelligence: effective analysis of high-density microarray data. *Drug Discov Today* 2002; 7: S55-63.
 105. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 2000; 289: 1760-3.
 106. Zhu H. Global analysis of protein activities using proteome chips. *Science* 2001; 293: 2101-5.
 107. Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, et al. Housekeeping genes as internal standards: use and limits. *J Biotechnol* 1999; 75: 291-5.