



**Collegio Ghislieri**

Centro per la Comunicazione e la Ricerca

*Progetto "Progressi in Biologia e Medicina"*

---

*7° Corso di formazione avanzata*

# **Nanomedicina: dalle nanostrutture alla terapia personalizzata**

7-11 aprile 2008, Collegio Ghislieri, Pavia

*A cura di Carlo Bernasconi*

*7° Corso di formazione avanzata*

**Nanomedicina: dalle nanostrutture  
alla terapia personalizzata**





**Collegio Ghislieri**

Centro per la Comunicazione e la Ricerca

*Progetto “Progressi in Biologia e Medicina”*

---

*7° Corso di formazione avanzata*

**Nanomedicina: dalle nanostrutture  
alla terapia personalizzata**

7-11 aprile 2008, Collegio Ghislieri, Pavia

*A cura di Carlo Bernasconi*

© Copyright 2008  Edizioni Medico-Scientifiche - Pavia

*Edizioni Internazionali srl*  
Divisione EDIMES - Edizioni Medico-Scientifiche - Pavia

Via Riviera, 39 - 27100 Pavia  
Tel. 0382526253 - Fax 0382423120  
E-mail: edint.edimes@tin.it

Tutti i diritti sono riservati.  
Nessuna parte può essere riprodotta in alcun modo  
(compresi i microfilm e le copie fotostatiche)  
senza il permesso scritto dell'editore.

# Indice

Prefazione ..... pag. IX

## Genomica

1. Le nanostrutture della vita: il mondo dei microRNA ..... » 3  
*CarloAlberto Redi*

## Le tecnologie della proteomica

2. Spettrometria di massa biomolecolare ..... » 15  
*Piero Pucci*

3. Elettroforesi bidimensionale ..... » 20  
*Luca Bini, L. Bianchi, R. Cianti, C. Landi, M. Puglia,  
C. Trombetta, D. Perini, A. Armini, V. Pallini*

4. Tecnologie proteomiche multidimensionali in fase liquida ..... » 29  
*Pierluigi Mauri*

## Genomica medica

5. L'avvento della genomica medica: dalle malattie monogeniche rare ai fattori di rischio genetici di malattie comuni ..... » 37  
*Carlo Bernasconi*

6. Mitochondrial DNA: a female perspective in recent human evolution ..... » 45  
*Antonio Torroni*

7. Mitochondrial disorders ..... » 50  
*Massimo Zeviani*

### **Biomarcatori in oncologia**

8. Profilo di espressione genica del tessuto polmonare normale  
 e rischio di cancro polmonare ..... » 63  
*Tommaso A. Dragani*
9. Approccio proteomico per la scoperta di biomarcatori di tumori:  
 un futuro per la diagnosi precoce e la terapia ..... » 68  
*Francesca Lavatelli, Giampaolo Merlini*
10. Marcatori genetici e risposta alla terapia di neoplasie  
 ematologiche ..... » 74  
*Mario Petrini*
11. Patient-specific tumor idiotype protein or idiotype (VDJ)-derived  
 class i-restricted peptides as biomarkers for the development  
 of anti-tumor vaccines in multiple myeloma ..... » 82  
*Roberto M. Lemoli*
12. Lo studio del profilo di espressione genica:  
 applicazioni cliniche ..... » 90  
*Sergio Ferrari*
13. Capire il cancro ..... » 100  
*Lucio Luzzatto*

### **Diagnosi molecolare delle neoplasie ematologiche**

14. Sviluppo di signatures molecolari per la diagnosi e la prognosi  
 delle leucemie mieloidi acute ..... » 107  
*Enrico Tagliafico*
15. Diagnosi molecolare dei linfomi: implicazioni cliniche ..... » 114  
*Silvia Franceschetti, Clara Deambrogi, Michaela Cerri,  
 Paola Riccomagno, Stefania Cresta, Lorenzo De Paoli,  
 Antonio Ramponi, Davide Rossi, Annarita Conconi,  
 Daniela Capello, Gianluca Gaidano*
16. Approccio genomico integrato per la classificazione molecolare  
 del mieloma multiplo ..... » 121  
*Antonino Neri, Sonia Fabris, Luca Agnelli*

17. Meccanismi molecolari e cellulari della trasformazione leucemica nelle SMD ..... » 128  
*Paolo Bernasconi, Marina Boni, Paola Maria Cavigliano, Rita Zappatore, Irene Dambruoso, Ilaria Giardini, Barbara Rocca, Silvia Calatroni, Paolo Tarantino, Marilena Caresana*

### **Genomica e nanotecnologie nelle malattie cardiovascolari**

18. Nanotecnologie nella diagnostica delle cardiomiopatie ..... » 145  
*Fabiana Gambarin, Nicola Marziliano, Maurizia Grasso, Eloisa Arbustini*
19. Sindromi aritmogeniche ereditarie ..... » 155  
*Silvia G. Priori*
20. Genomica delle piastrine e trombofilia ..... » 163  
*Ilaria Canobbio, Cesare Balduini*
21. Dalla genetica ai trials clinici in cardiologia ..... » 173  
*Stefano Ghio*

### **Genomica e proteomica nelle malattie neurodegenerative**

22. Indagini di proteomica clinica per l'identificazione di marcatori molecolari delle malattie neurodegenerative ..... » 181  
*Andrea Urbani*
23. Proteomica della malattia di Alzheimer: comprensione dei meccanismi di malattia e identificazione di biomarcatori ..... » 185  
*Stefano Govoni*
24. Definizione del profilo genomico della sclerosi multipla ..... » 190  
*Roberto Furlan*
25. Biomarcatori per la sclerosi laterale amiotrofica ..... » 195  
*Gabriele Mora, Caterina Bendotti*

### **Medicina predittiva e farmacogenomica**

26. Medicina predittiva: il profilo genetico di predisposizione alla malattia e il ruolo dell'ambiente ..... » 207  
*Paola Mandich*



27. Farmacogenomica: scopi ed applicazioni ..... » 213  
*Angelo Nicolin*
28. Le basi genetiche della variabilità individuale della risposta  
ai farmaci ..... » 217  
*Diego Fornasari*
29. Implicazioni della farmacogenomica nella pratica clinica  
in oncologia ..... » 221  
*Silvia Marsoni*

### **Farmacogenomica e nuovi farmaci**

30. Struttura delle proteine e scoperta di nuovi farmaci ..... » 229  
*Lucio Toma*
31. La struttura delle monoammina ossidasi umane A e B ..... » 235  
*Claudia Binda, Daniele Bonivento, Andrea Mattevi*
32. I passaggi successivi per lo sviluppo di un nuovo farmaco.  
Identificazione e sviluppo di PHA-739358, un potente inibitore  
delle Aurora chinasi con profilo di selettività chinasica compatibile  
con attività antitumorale ..... » 241  
*Maurizio Pulici*
33. Computer chemistry and molecular modelling ..... » 247  
*Christian Orrenius*

## Prefazione

Il termine “nanomedicina” è stato recentemente introdotto dal *National Institutes of Health* per indicare l’applicazione di nanotecnologie allo studio e monitoraggio di sistemi biologici, e alla diagnosi e terapia di malattie. Quindi, gli obiettivi primari della nanomedicina sono il progresso della conoscenza delle nanostrutture (intese come minuscoli costituenti delle cellule dell’organismo, con dimensioni attorno al milionesimo di millimetro) e il loro rapido trasferimento in clinica per scopi diagnostici e terapeutici.

Con tali obiettivi negli ultimi anni si sono sviluppati parecchi progetti di ricerca con l’utilizzo di nanotecnologie come supporto alla genomica e alla più moderna proteomica, mentre in clinica i primi progetti di nanomedicina potenzialmente più interessanti hanno riguardato le malattie tumorali, cardiovascolari e neurodegenerative. Particolarmente promettente è poi apparso l’impiego di nanotecnologie in farmacogenomica, sia per gli aspetti di medicina predittiva, sia per lo sviluppo di nuovi farmaci.

Nell’organizzazione di questo nostro Corso di formazione avanzata per Medici, Biologi e Biotecnologi è stata quindi programmata la trattazione dei vari argomenti sopra riportati, partendo dalle conoscenze scientifiche di ordine biologico per passare poi alle implicazioni pratiche di ordine clinico e farmacologico. In particolare, sessioni successive sono dedicate alla genomica e alle tecnologie della proteomica, alla genomica medica, all’identificazione e impiego di biomarcatori in oncologia, alla diagnosi molecolare delle neoplasie ematologiche (non mi è stato possibile tradire la mia origine!), alla genomica e alle nanotecnologie nelle malattie cardiovascolari e nelle malattie neurodegenerative, ai riferimenti di farmacogenomica in medicina predittiva e nello sviluppo di nuovi farmaci.

In accordo con l’orientamento generale dei nostri Corsi, maggior rilievo è stato riservato agli aspetti biomedici rispetto a quelli più strettamente tecnologici. I Docenti delle singole lezioni sono Ricercatori particolarmente esperti nell’ambito dei vari argomenti trattati, e a nome di tutti i partecipanti al Corso li ringrazio per aver accettato l’invito a svolgere il difficile compito di integrare nei singoli settori le conoscenze ritenute tradizionali con le più significative recenti acquisizioni. Tale compito è reso ancor più impegnativo dalla prospettiva delle profonde modificazioni che le nuove conoscenze potranno indurre nell’esercizio della medicina.

*Carlo Bernasconi*

Pavia, 7 aprile 2008



# **GENOMICA**



# Le nanostrutture della vita: il mondo dei microRNA

*CarloAlberto Redi*

Laboratorio di Biologia dello Sviluppo, Università di Pavia. Direttore Scientifico Fondazione IRCCS Policlinico San Matteo, Pavia

## Nanomedicina

Il termine nanomedicina fa la sua comparsa verso la fine degli anni '90 dell'ultimo secolo mentre la prima pubblicazione che esplicitamente impiega questo termine è dell'anno 2000 (Institute for Scientific Information, Thompson, Philadelphia, PA, USA). Non è semplice trovare una definizione precisa di cosa si debba intendere per nanomedicina. Ed ancor più difficile è trovare un consenso generale tra gli addetti delle varie discipline scientifiche che concorrono nella pratica alle realizzazioni dei programmi di ricerca, ed a quelle terapeutiche, di nanomedicina. Ciò è dovuto al fatto che ancora oggi vi sono almeno due principali accezioni semantiche per il termine stesso di nanomedicina. Il che, a sua volta, dipende dalla sfumata ed imprecisa capacità di riconoscere quale disciplina fondante gioca un ruolo chiave nel determinare i campi di interesse. A volte è la biotecnologia, a volte la tecnologia di microsistemi, più spesso è la sovrapposizione delle due.

Al punto che forse sarebbe meglio sostituire il termine nanotecnologia con quello di "tecnologie convergenti". Certamente però impiegare il termine "tecnologie convergenti in medicina" non migliora la necessità di porre dei confini precisi. Per questo motivo è bene riportare lo stato dell'arte e segnalare che alcuni ricercatori definiscono nanomedicina in termini molto vaghi come la tecnologia che usa strumenti molecolari a fini diagnostici e terapeutici. Altri fanno riferimento all'originale significato di nanotecnologia e quindi si riferiscono a tutte quelle tecnologie che impiegano effetti fisici che accadono a livello di nanoscala (i.e., sotto i 100 nm).

Secondo questa ultima versione, la nanomedicina è definita come quella disciplina che impiega oggetti a livello di nanoscala (materiali nanostrutturati) per ottenere "effetti medici" (a fini diagnostici o di terapia). È chiaro poi che gli effetti medici non sono limitati al livello di nanoscala. Questi includono *drug delivery* e terapia, in vivo imaging, in vitro diagnostics, biomateriali e *active implants*, per

ricordarne alcuni. Si deve poi ricordare che il mondo della nanomedicina legato alle terapie di “medicina personalizzata” non include strettamente anche altri effetti medici (terapeutici) di pratiche di particolare rilievo e che si basano su fenomeni che accadono comunque a livello di nanoscala. Un esempio per tutti, sviluppatosi in anni recenti e di estrema attualità a livello internazionale, è quello della adroterapia, il trattamento dei tumori grazie al loro irraggiamento con protoni e ioni (che fanno parte della famiglia degli adroni) del carbonio. Pavia è sede del Centro Nazionale di Adroterapia Oncologica e vede in questi giorni l’inizio sperimentale delle proprie attività. Queste particelle (prodotte da sincrotoni e ciclotroni) hanno la particolarità di poter essere indirizzate verso la massa neoplastica con elevatissima precisione. Gli ioni, inoltre, grazie al loro elevato effetto radiobiologico, risultano particolarmente utili per trattare tumori radiore-sistenti come quelli che colpiscono il distretto testa-collo, il polmone e il pancreas. Grazie alla possibilità di collimare con grande precisione il fascio di protoni, questi ultimi hanno come utilizzo di elezione organi e zone anatomiche a rischio: occhi, base del cranio o colonna vertebrale.

Come può risultare intuitivo, al di là degli aspetti più generali delle nanotecnologie e della nanomedicina, nel mondo delle nanostrutture biologiche (e quindi del loro impiego in nanomedicina) un ruolo particolare è giocato dagli acidi nucleici e certamente dalla attuale capacità tecnologica della loro manipolazione.

### **Il mondo dei micro RNA**

A questo riguardo negli ultimi 4-5 anni si è verificata una vera rivoluzione che ha comportato la nascita di nuovi paradigmi concettuali in Biologia e aperto nuove possibilità di applicazioni terapeutiche in medicina. Questo tema sarà trattato in modo particolare quale esempio di ricerca e applicazione possibile oggi nella cosiddetta medicina molecolare. Come testimoniato dai premi Nobel assegnati nell’anno 2006 ai giovanissimi statunitensi Andrew Fire (Stanford University School of Medicine Stanford, CA) e Craig Mello (University of Massachusetts Medical School Worcester, MA) “for their discovery of RNA interference - gene silencing by double-stranded RNA”, il mondo degli RNA è oggi uno dei temi, se non il tema, centrale della ricerca biomedica. E questo per il semplice motivo che queste ricerche stanno svelando le basi molecolari dei fenomeni fisiologici dello sviluppo e quindi le basi molecolari delle patologie e, di conseguenza, permettono la dissezione a livello molecolare della complessità biologica. Stanno inoltre rivoluzionando il nostro approccio alle terapie e ponendo le basi delle cosiddette terapie farmacogenomiche personalizzate.

Di rilievo è notare i tempi in cui tutto ciò sta accadendo, stiamo parlando degli ultimi anni della ricerca biologica! A questo riguardo invito caldamente gli studenti del corso a visitare il sito della Fondazione Nobel (<http://www.nobelprize.org>) ed a leggere attentamente le prolusioni di Craig Mello e di Andrew Fire per la pregevole esposizione dei contenuti delle ricerche da loro compiute e per la chiarezza della visione del profilo storico che ha portato a sviluppare tali ricerche.

Negli ultimi cinque anni i biologi, lavorando con una panoplia di organismi ed

impiegando svariate tecniche molecolari (fondamentalmente per studiare la espressione genica), si sono trovati dinnanzi la stessa sorpresa: la trascrizione avviene su una larga parte del genoma, oltre il 50% del genoma, e produce una elevata quantità di trascritti la cui funzione è sconosciuta (TUF, Transcripts of Unknown Function). Esiste in altre parole una complessa architettura trascrizionale che include:

- 1) isoforme alternative di RNA per ciascuna proteina;
- 2) trascrizione di RNA antisenso;
- 3) una grande quantità di RNA non codificante (ncRNA) con rilevanti funzioni biologiche.

Questa diversità trascrizionale può finalmente lasciare intuire, se non spiegare, come sia possibile che un relativamente simile numero di geni codificanti proteine sia stato trovato in *Drosophila*, nei Nematodi (*Caenorabditis elegans*) e nell'uomo, a fronte della impressionante differenza fenotipica. Ora due semplici calcoli possono dare l'idea del problema sul quale la ricerca è indirizzata. È noto infatti che solo circa 1.5% del genoma umano è costituito da regioni codificanti proteine ma, in base alla lunghezza media dei geni codificanti proteine, le unità di trascrizione codificanti proteine dovrebbero comprendere circa il 30-40% del genoma. I TUF costituiscono almeno il 50% del nostro genoma e circa un terzo di questi trascritti non codifica per proteine! Se da un lato è evidente che solo una completa visione del trascrittoma umano potrà spiegarci queste discrepanze, dall'altro questi numeri ci indicano che il ncRNA è un componente essenziale del trascrittoma umano ed ha una funzione biologica, svolge un ruolo, che non è ancora chiarito sebbene sia stato individuato. Il termine ncRNA veniva generalmente riferito a RNA che non codifica per proteine, come l'rRNA, il tRNA, piccole molecole di RNA coinvolte nella traduzione del segnale (snRNA) e piccole molecole di RNA nucleolare (snoRNA) coinvolte nelle modificazioni degli rRNA. Tra questi ncRNA, la gran parte degli snoRNA deve ancora essere chiaramente identificata ed analizzata.

È di rilievo concettuale ricordare che il paradigma centrale della Biologia molecolare, sviluppato in base agli studi su organismi come *Escherichia coli*, vedeva l'RNA svolgere la funzione di intermediario informazionale tra una sequenza di DNA (un gene) e la proteina da esso codificata. L'idea sottesa era che la gran parte della informazione genetica che specifica i fenotipi è espressa come proteine le quali non solo svolgono diverse funzioni strutturali e catalitiche ma svolgono anche funzioni regolatorie in diversi sistemi fisiologici. Tutto ciò è ampiamente vero nei procarioti e si riteneva altrettanto vero negli eucarioti. Dal che discendeva l'idea che la gran parte del genoma era costituito da DNA junk, DNA spazzatura (Redi et al., 2002), nonostante già Jacob e Monod ben 46 anni orsono avessero predetto che l'RNA doveva essere ampiamente coinvolto come mezzo di regolazione della espressione genica. In *Escherichia coli* sono stati identificati più di 60 ncRNA e più di 200 sono stati predetti per via bioinformatica. Di grande interesse il fatto che piccoli ncRNA sono stati identificati anche in altri batteri, ed anche in rappresentanti degli Archaea, e risultano omologhi di Argonata. Quest'ultima è una famiglia di endonucleasi che si lega specificatamente agli



RNA e che svolge un ruolo centrale nel meccanismo d'azione dei micro RNA (miRNA) e nei piccoli RNA ad interferenza (RNAi) negli eucarioti. I miRNA costituiscono una ampia famiglia di ncRNA di 21-22 nucleotidi che interagisce con specifici mRNA, in siti precisi, inducendone il taglio e la inibizione della traduzione. I miRNA si trovano nelle piante, negli animali ed in alcuni virus. Sono stati scoperti in *Caenorhabditis elegans* dove è stato chiarito che il meccanismo d'azione grazie al quale regolano l'espressione di vari mRNA è basato sul riconoscimento di sequenze parzialmente complementari a quelle degli mRNA. La gran parte dei miRNA è conservata evolutivamente tra specie anche a livello intergenerico; alcuni addirittura tra invertebrati e vertebrati. Molti miRNA hanno specifiche attività in alcune fasi dello sviluppo e presentano quadri di espressione che sono specifici dei vari tipi cellulari. Comunque, per la gran parte dei miRNA dei Mammiferi, la specificità di espressione deve ancora essere chiarita.

### **L'interferenza da RNA**

Ciò non toglie che il mondo dei miRNA stia già producendo applicazioni nell'ambito delle tecnologie convergenti, sia a livello della bioinformatica che della diagnostica e terapia. Risulta immediato infatti pensare di poter interferire con il processamento della informazione genetica attraverso la manipolazione di alcuni suoi intermediari, gli RNA delle varie famiglie e classi, al fine di ottenere degli effetti biologici specifici a valle del processamento stesso: in altre parole, in medicina, ciò significa tentare di ottenere effetti terapeutici interferendo con le vie di processamento della informazione genetica basata sugli RNA. A questo riguardo già si stanno sviluppando diverse sperimentazioni (Hannon e Rossi, 2004; Rinaudo et al., 2007), basate sulla interferenza da RNA, per il trattamento di alcune patologie che brevemente saranno qui considerate. Il grande interesse per il potenziale uso della metodica di RNAi discende come ovvio dalla altissima specificità con la quale si può inibire l'espressione di uno specifico gene grazie alla potenza mirata della metodica stessa. Prima di considerare alcuni esempi delle attuali sperimentazioni è bene ricordare, almeno in via schematica, su quali basi molecolari si attua la interferenza da RNA. Questa è fondamentalmente dovuta al riconoscimento base specifico, alla ibridazione, tra catene di acidi nucleici (mRNA e miRNA) ed alla azione di alcuni specifici enzimi del metabolismo degli acidi nucleici. Schematicamente, le cellule trascrivono brevi segmenti di DNA in miRNA che regolano la espressione genica grazie all'appaiamento con mRNA e producendo così una interferenza che abbassa (spegne) il livello della proteina prodotta. Il miRNA viene trascritto come un lungo precursore che avvolgendosi a doppia elica forma una porzione a forcina di capello di circa 80 bp. Un enzima taglia la porzione a forcina di capello ed una proteina la esporta dal nucleo al citoplasma. Qui, un altro enzima taglia l'occhiello al termine della forcina mentre una delle due eliche, lunga 21-22 bp, viene caricata su un complesso proteico di "silenzamento". Il miRNA aiuta, grazie all'appaiamento con una regione non tradotta dell'mRNA di una proteina, il complesso proteico di silenzamento a riconoscere in modo specifico il proprio bersaglio ed a determi-

nare il taglio del mRNA che così non può più essere tradotto. In tal modo la produzione della proteina codificata dal mRNA si spegne. Sulla base di questa schematica visione del meccanismo molecolare su cui si basa l'RNAi diviene intuitivo capire che grazie alla modulazione della regolazione della espressione dell'informazione genetica veicolata dalle varie classi di RNA è possibile manipolare RNA specifici al fine di trattare varie patologie. Queste ultime spaziano dagli oncogeni (cancro) ai fattori di crescita sino a bersagli così specifici come i polimorfismi di singoli nucleotidi (malattie genetiche). Inoltre, la metodica della interferenza da RNA può essere potenzialmente studiata per il trattamento di infezioni virali come quelle causate dal virus della epatite C (HCV) e da quello della immunodeficienza (HIV). I maggiori dubbi sull'impiego terapeutico in medicina umana riguardano le difficoltà legate alla somministrazione farmacologica degli RNA capaci di produrre l'effetto della interferenza e la loro potenziale tossicità. Sebbene nessuna applicazione terapeutica sia stata approvata da alcuna agenzia regolativa, alcuni trials clinici di RNAi che impiegano sia RNA sintetici sia RNA virali sono in corso di attuazione.

### **RNAi e cancro**

Sono molti gli studi che hanno impiegato la tecnica della RNAi per controllare la proliferazione cellulare e dunque non stupisce che sebbene non vi siano attualmente trials clinici in corso per il controllo della crescita neoplastica, l'RNAi sia stata già impiegata a tal fine. Mediante l'uso di oligonucleotidi antisense disegnati per colpire selettivamente il gene anti-apoptotico *BCL2* (noto come *gene-sense*, della Genta Inc.) sono stati ottenuti risultati molto soddisfacenti nel trattamento del melanoma metastatico (seppure ancora in combinazione con la chemio convenzionale).

Fattore chiave per il successo dei trattamenti con tecnologia RNAi è la scelta della via di somministrazione che permetta di raggiungere solo le cellule bersaglio e di ottenere al loro livello la concentrazione utile per gli effetti terapeutici. Altri aspetti di rilievo sono quelli legati alla stabilità delle piccole molecole di RNA disegnate ed al tempo di permanenza dell'effetto provocato: in genere si tratta di sette-dieci giorni. Il tempo di dimezzamento (half-life) delle molecole di RNA impiegate per l'RNAi è ancora insufficiente, troppo rapidamente le nucleasi seriche degradano l'RNA impiegato abbassandone in modo critico il livello di molecole circolanti e quindi abbassando il numero delle molecole giunte a bersaglio e potenzialmente utili alla interferenza. Certamente la via sistemica di somministrazione deve essere migliorata.

Come deve esserlo la capacità di penetrazione dell'RNA nudo a doppia elica nei tessuti e nelle cellule, così da raggiungere concentrazioni efficaci terapeuticamente. L'RNAi è stata anche impiegata per colpire selettivamente la espressione di oncogeni che codificano per proteine di fusione come la Bcr-Abl p210, caratteristica della leucemia mieloide cronica (CML), ottenendo la prova che la RNAi è una ottima terapia anticancro. Le principali opzioni di trattamento per la CML sono la chemio, il trapianto allogenico di midollo e di recente l'impiego

degli inibitori delle tirosinchinasi (il famoso “imatinib” commercializzato come *gleevec/glivic*, della Novartis). Purtroppo un numero sempre crescente di pazienti ha sviluppato resistenza contro l’imatinib, riducendone così il potenziale terapeutico. La buona notizia è che grazie alla RNAi, con molecole di RNA sia a elica semplice che doppia, si è ottenuto un significativo spegnimento della espressione di Bcr-Abl p210. E ciò che è ben più importante è che lo spegnimento è selettivo per la sola oncoproteina p210 ed il suo mRNA. Ora, poichè i tumori ematologici sono spesso trattati con il trapianto di midollo, una possibile applicazione terapeutica potrebbe essere quella di trasfettare le cellule staminali ematopoietiche con un vettore che porta un costrutto avente come bersaglio l’mRNA che codifica per la oncoproteina p210. In tal modo si dovrebbe proteggere il paziente dalla ripresa della malattia dovuto alla proliferazione di poche latenti cellule staminali leucemiche. Chiaramente, per rendere efficiente una simile terapia è necessario migliorare lo spettro degli attuali vettori virali in uso e l’efficienza della trasfezione (100% è il titolo richiesto!) delle cellule staminali da reinfondere.

### **RNAi per il trattamento della infezione da HIV**

Sebbene negli ultimi anni la infezione da HIV risulti ben più controllata grazie alla azione di nuove molecole ad azione farmacologica di grande efficienza, si è però verificata l’insorgenza di varianti virali, e di tossicità, legata proprio all’uso dei nuovi farmaci. Di conseguenza, vi è ancora un grande interesse (bisogno) per lo sviluppo di nuovi approcci terapeutici antivirali. In effetti, grazie al fatto che la storia del ciclo vitale ed il pattern di espressione genica dell’HIV sono ben conosciuti, l’HIV è stato il primo agente infettivo che si è tentato di colpire con la metodica dell’RNAi mediante l’impiego di RNA sintetici aventi come bersaglio gli RNA di alcuni geni espressi precocemente o tardivamente nel corso del ciclo vitale (rev, gag, env e trascrittasi inversa) su linee cellulari o su colture primarie di cellule ematopoietiche infettate. Nonostante il successo ottenuto in colture di linee cellulari con la inibizione da RNAi della espressione di alcuni geni, gli stessi risultati non sono stati ottenuti direttamente sulle cellule da infezione primaria poichè l’alto tasso di mutazione che il virus presenta porta alla selezione di mutanti capaci di evitare la interferenza.

Una strategia alternativa è la repressione della espressione dei cofattori cellulari necessari per la infezione. Recettori (i.e., CD454) e corecettori (i.e., CXCR4 e CCR557) sono stati spenti con successo grazie alla RNAi, ottenendo così la inibizione della replicazione dell’HIV, sia in varie linee cellulari umane che in colture primarie di T linfociti, cellule staminali ematopoietiche e macrofagi. La scelta di quali cofattori colpire specificatamente è di rilievo nello sviluppare strategie di controllo della infezione con la metodica dell’RNAi. Di particolare interesse è il co-recettore CCR5: questo co-recettore non è essenziale per il normale funzionamento del sistema immunologico e individui omozigoti per una delezione di 32 bp risultano resistenti alla infezione da HIV (gli eterozigoti mostrano una ritardata progressione all’AIDS). Grazie all’uso di un vettore lentivirale per trasdurre

del RNA anti-CCR5 si è ottenuta una selettiva inibizione della espressione di CCR5 in linfociti umani capace di produrre una infettività sino a sette volte minore rispetto ai controlli. Purtroppo questi linfociti risultano ancora sensibili alla variante del virus che usa un altro co-recettore, CXCR4; poichè CXCR4 è essenziale per il fisiologico funzionamento delle cellule staminali ematopoietiche non si può pensare di colpire con l'RNAi questo co-recettore, e neppure il recettore CD4, per sviluppare una terapia anti HIV. In altre parole, vi sono grandi difficoltà nello stabilire quali cofattori colpire poichè anche le cellule non infette risultano bersaglio (in modo analogo alla tossicità riscontrata con l'impiego di molecole ad azione antiretrovirale nelle attuali terapie). Inoltre vi è da considerare la immunogenicità dei vettori virali impiegati per la somministrazione degli RNA stessi, il che preclude l'impiego di somministrazioni multiple. Il sistema di elezione potrebbe essere quello di isolare cellule T dal paziente e poi trasdurre, espandere e reinfondere queste cellule nello stesso paziente. Ci si attende che nei prossimi tre - quattro anni la tecnologia a RNAi per il trattamento della infezione da HIV si sposti dal laboratorio al letto del paziente.

### **RNAi per il trattamento dell'epatite virale**

Uno dei principali problemi sanitari a livello internazionale è certamente quello delle epatiti causate dai virus B e C della epatite (HBV e HCV); attualmente vi sono diverse centinaia di milioni di individui infetti ed esiste un vaccino per la prevenzione della sola epatite da virus B. Le HBV e HCV sono dunque due importanti bersagli di una potenziale terapia basata su RNAi. Ad oggi non vi è comunque un risultato capace di portare allo svolgimento di trials clinici in uomo; vi è però una prova del principio che la tecnologia a RNAi può essere valida. Infatti, McCaffrey e collaboratori (McCaffrey et al., 2003) impiegando un anti-RNA per HBV hanno ottenuto un significativo (99%) spegnimento del genoma virale in epatociti di topo.

### **L'interferenza da RNA per le malattie genetiche**

Alcuni risultati promettenti al fine del trattamento delle malattie genetiche sono venuti dalla dimostrazione che polimorfismi di singoli nucleotidi (SNP) sono bersaglio selettivo per la RNAi nei trascritti di alleli mutanti. È così possibile disegnare molecole di RNA capaci di interagire (ibridare) con la sola catena dell'allele mutato e non con la forma *wild*, nonostante una sola coppia di basi sia mutata, e con ciò determinare la degradazione del solo mRNA dell'allele mutato lasciando inalterato e funzionante l'mRNA della forma *wild*. A questo riguardo, un esempio di grande interesse è quello legato agli studi sulla sclerosi laterale amiotrofica (ALS) causata da mutazioni nel gene della superossidodismutasi Cu e Zn dipendente (*SOD1*).

Data il ruolo cruciale giocato dal gene SOD1 è essenziale determinare la selettiva eliminazione dei soli trascritti dell'allele mutato. Molte delle mutazioni di SOD1 sono SNP. Ding e collaboratori (Ding et al., 2003) hanno dimostrato la

potenziale azione terapeutica del RNAi nel trattamento della ALS silenziando selettivamente l'allele mutato. Anche il trattamento di alcune patologie neurologiche degenerative quali la malattia di Huntington sono potenzialmente aggredibili con la tecnologia dell'RNAi. In questo caso giocano un ruolo chiave nello stabilirsi e nella progressione della malattia la formazione di poliglutammine codificate da ripetizioni CAG. Stante la frequenza delle ripetizioni CAG in trascritti normali diviene impossibile la somministrazione sistemica di molecole di RNA disegnate per colpire selettivamente i trascritti mutati. La recente possibilità tecnica di somministrare in loco, e solo in alcune regioni cerebrali, molecole di RNA permette di evitare la somministrazione sistemica e di colpire selettivamente i trascritti mutati giusto là dove necessario. Questo avanzamento tecnico unito al disegno di molecole capaci di riconoscere SNP nei trascritti mutati apre nuovi scenari terapeutici per le malattie neurologiche degenerative che nei prossimi quattro/sei anni dovrebbero trovare piena attuazione.

### **Le sfide della tecnologia a RNAi nella pratica terapeutica**

La principale preoccupazione nell'uso della tecnologia del RNAi è certamente quella di disegnare molecole capaci di agire in maniera selettiva solo su alcuni intermediari del processamento della informazione genetica basata sugli RNA. In altre parole, la prima preoccupazione nella pianificazione di un intervento basato su RNAi è quella che la molecola in uso non interagisca con le vie della RNAi endogena che normalmente si esplicano nelle cellule. Non va dimenticato infatti che questa è presente in tutte le cellule ed è, come già detto nella introduzione, una parte essenziale del processamento della informazione genetica nel corso dello sviluppo e della differenziazione cellulare. Il secondo aspetto di maggior rilievo nello sviluppo delle strategie terapeutiche con RNAi è la via di somministrazione. L'infusione a livello del midollo osseo è fattibile ma non assicura il raggiungimento della necessaria concentrazione, per la efficacia terapeutica, a livello di singola cellula bersaglio. L'approccio alternativo, la somministrazione grazie a vettori virali, incontra tutti i problemi e gli ostacoli legati alla terapia genica, poichè si presenta come una forma variante di questa solo per il costrutto impiegato; i vettori virali e retrovirali sono infatti gli stessi.

Resta comunque riconosciuto a livello della comunità scientifica internazionale la grande validità concettuale dell'approccio alla pratica terapeutica basata sulla tecnologia da RNAi e molti sforzi si stanno compiendo in questa direzione. Per lo sviluppo di strategie per lo studio della espressione genica ed il disegno di molecole utili alla sua manipolazione con la tecnica dell'RNAi esistono già alcuni siti di istituzioni accademiche di eccellenza che offrono *free* la possibilità di pianificare i propri esperimenti. La domanda oggi è infatti come e quando la RNAi rivoluzionerà il trattamento delle malattie, la terapia medica, come è già riuscita a rivoluzionare la ricerca biologica sulla funzione dei geni.

Per chi si vuol cimentare e ritiene giunto il momento di espandere il proprio armamentario di tecniche nella propria pratica di ricerca, sono disponibili diversi siti che ben illustrano i diversi passaggi del disegno sperimentale ed accompa-

gnano i nuovi entrati nel mondo dei micro RNA:  
[http://web.mit.edu/icbp/resources/rna\\_%20interference.html](http://web.mit.edu/icbp/resources/rna_%20interference.html)  
[https://bioinfo.wistar.upenn.edu/siRNA\\_tutorial.htm](https://bioinfo.wistar.upenn.edu/siRNA_tutorial.htm)  
<https://bioinfo.wistar.upenn.edu/siRNAdescrip.htm>  
<http://www.cshl.edu/public/SCIENCE/hannon.html>  
<http://www.rockefeller.edu/labheads/tuschl/sirna.html>

### **Bibliografia essenziale**

1. Ding, H. et al. Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. *Aging Cell* 2: 209-217, 2003.
2. McCaffrey AP. et al. Inhibition of hepatitis B virus in mice by RNA interference. *Nature Biotechnol.* 6: 639-644, 2003.
3. Mattick JS., Makunin IV. Non-coding RNA. *Human Molecular Genetics*, 15: R17-R29, 2006.
4. G.J. Hannon e J. J. Rossi. Unlocking the potential of the human genome with RNA interference. *Nature*, 431: 371-378, 2004.
5. Redi CA., Zuccotti M., Garagna S. L'altro genoma. *Le Scienze*, 409: 36-42, 2002.
6. Rinaudo K., Bleris L., Maddamsetti R., Subramanian S., Weiss R., Benenson Y. A universal RNAi-based logic evaluator that operates in mammalian cells. *Nature Biotechnology*, 25: 795-801, 2007.
7. Willingham AT., Gingeras TR. TUF Love for "Junk". *Cell*, 125: 1-6, 2006.



# **LE TECNOLOGIE DELLA PROTEOMICA**





## Spettrometria di massa biomolecolare

*Piero Pucci*

CEINGE Biotecnologie Avanzate s.c.a r.l., Dipartimento di Chimica Organica e Biochimica,  
Università di Napoli Federico II

La moderna ricerca in biologia é attualmente indirizzata verso lo studio della struttura delle macromolecole biologiche ed, in particolare, delle loro interazioni con altre macromolecole o con piccoli metaboliti, il fenomeno del riconoscimento molecolare grazie al quale i sistemi viventi funzionano. Una profonda conoscenza delle funzioni di queste macromolecole dipende dalla comprensione a livello molecolare di tali interazioni, che a sua volta dipende dalla definizione della loro struttura.

Tra le varie metodologie analitiche utilizzate per la caratterizzazione strutturale delle proteine, la spettrometria di massa occupa un posto speciale dal momento che essa è in grado di misurare una proprietà intrinseca delle molecole, il loro peso molecolare, con una sensibilità elevatissima. Per queste sue caratteristiche la spettrometria di massa ha svolto un ruolo preminente negli ultimi dieci anni nella biologia strutturale ed è stata utilizzata in una sorprendente varietà di applicazioni, specialmente per quanto riguarda la soluzione di problemi in cui le metodologie biochimiche classiche spesso non producono dati significativi.

Eppure, le macromolecole biologiche essendo di grandi dimensioni e di natura polare non sono facilmente trasferite in fase gassosa e trasformate in specie cariche (ioni) in modo da poterle analizzare. Infatti, solo durante gli anni novanta sono state sviluppate una serie di metodologie che hanno reso possibile la ionizzazione di molecole polari, termolabili ed ad alto peso molecolare. La spettrometria di massa ad Electrospray (ESMS) si basa sulla capacità di molecole proteiche di diventare poliprotonate quando una loro soluzione diluita viene spruzzata da un capillare metallico attraverso un campo elettrico. Durante questo processo si formano specie proteiche multiprotonate  $(M+nH)^{n+}$  che danno origine nello spettro di massa ad una serie di picchi consecutivi; ogni segnale corrisponde ad una popolazione di molecole proteiche contenente un particolare numero  $n$  di protoni. L'analisi ESMS è in grado di determinare il peso molecolare di proteine fino a oltre 80kDa con un'accuratezza straordinaria, impensabile con altre metodologie analitiche, limitando l'errore della misura a circa 1 unità di massa ogni 20.000 Da.

La spettrometria di massa MALDI (Matrix Assisted Laser Desorption Ionization) si basa invece sull'irradiazione di un campione proteico immerso in una opportu-

na matrice utilizzando un raggio laser. Durante questo processo, le molecole di analita vengono ionizzate sotto forma di specie monoprotionate,  $MH^+$ , originando spettri di massa molto semplici in cui ogni segnale corrisponde ad un singolo analita. Questa metodologia quindi offre la possibilità di analizzare miscele di analiti, ad esempio le miscele di peptidi provenienti dalla digestione di proteine, ed ottenere informazioni sui componenti di tali miscele senza ricorrere a preliminari passaggi di separazioni cromatografiche.

Infine, i moderni spettrometri di massa possono combinare due fasi dell'analisi in tandem dando luogo alla metodologia denominata spettrometria di massa tandem (tandem MS o MS/MS). In questa metodologia, un singolo peptide presente in una miscela può essere selezionato nel primo spettrometro di massa e quindi introdotto in una zona denominata cella di collisione e frammentato per collisione con particelle di gas inerte (Argon, Elio o Azoto). I frammenti risultanti sono quindi separati nel secondo spettrometro di massa dando luogo al cosiddetto spettro di frammentazione o spettro MS/MS. Dal momento che i peptidi frammentano preferenzialmente sul legame peptidico, l'interpretazione dello spettro di frammentazione produce informazioni sulla sequenza dei peptidi stessi.

Le moderne applicazioni della spettrometria di massa nel campo della Biologia Strutturale e della Proteomica sono indirizzate verso tre obiettivi principali:

- 1) la caratterizzazione strutturale di proteine ricombinanti, un aspetto estremamente importante nell'industria biotecnologica;
- 2) la definizione e l'analisi delle modifiche post-traduzionali, un campo in cui le informazioni ottenibili dalle metodologie di manipolazione del DNA sono di uso estremamente limitato;
- 3) l'identificazione di proteine incognite in diversi aspetti della ricerca in Proteomica.

In questa presentazione saranno discusse alcune applicazioni della spettrometria di massa ai diversi aspetti della ricerca sottolineati in precedenza, descrivendo una strategia generale di analisi strutturale che si esplica attraverso tre fasi distinte ma fortemente integrate tra di loro, la determinazione del peso molecolare accurato di proteine intatte, la determinazione della mappa peptidica di proteine e l'identificazione di proteine frazionate su gel mono- e/o bi-dimensionale.

### **Determinazione del peso molecolare accurato di proteine intatte**

In molti casi la determinazione diretta del peso molecolare accurato di una proteina può fornire utili indicazioni sulla struttura della proteina stessa. Infatti il confronto tra il valore di massa determinato sperimentalmente e quello atteso consente di confermare la correttezza della sequenza amminoacidica della proteina o di identificare immediatamente la presenza di una anomalia strutturale. Data l'accuratezza necessaria nella maggioranza delle applicazioni, la spettrometria di massa ad electrospray (ESMS) costituisce il metodo ideale per la misura del peso molecolare delle proteine.

L'analisi ESMS costituisce un modo semplice, rapido ed efficace per verificare la correttezza della sequenza proteica dedotta dalla sequenza nucleotidica del gene

corrispondente. Allo stesso modo è possibile controllare la struttura primaria di proteine ricombinanti; infatti un requisito assolutamente indispensabile di tali proteine é la loro assoluta identità con il prodotto naturale. Il controllo di questa identità é ancora più importante quando i prodotti dell'ingegneria genetica sono destinati a scopo terapeutico. Talvolta la semplice differenza di massa tra il peso molecolare sperimentale e quello teorico fornisce indicazioni sulla natura stessa dell'anomalia presente.

Inoltre, la verifica della struttura primaria di proteine ricombinanti mediante analisi ESMS costituisce anche un sofisticato criterio di purezza, in grado di evidenziare la presenza di altri componenti proteici presenti in tracce e difficilmente separabili ed evidenziabili con altri metodi analitici. Infine, nel caso dell'analisi di alcune emoglobine umane varianti, denominate emoglobine Lepore, la determinazione del peso molecolare accurato della globina variante è addirittura diagnostico per identificare il particolare tipo molecolare di Hb Lepore.

In altri casi, la misura del peso molecolare di una proteina definisce soltanto la presenza di una variante genetica o di una modifica post-traduzionale o di una anomalia strutturale, ma non è in grado di definire il sito e la natura di tale variazione. È necessario in questi casi ricorrere ad un'indagine più dettagliata della struttura proteica attraverso la determinazione della mappa peptidica della proteina.

### **Determinazione della mappa peptidica di proteine**

La localizzazione esatta del sito e la definizione della natura di eventuali anomalie strutturali presenti in una proteina vengono determinate mediante la costruzione della mappa peptidica della proteina stessa. In questo modo è possibile caratterizzare accuratamente sostituzioni amminoacidiche, modificazioni post-traduzionali inclusa la glicosilazione, o presenza di eventuali ponti disolfuro.

Il concetto di mappa peptidica è molto familiare ai chimici delle proteine sin dalla metà degli anni cinquanta e la filosofia di questo approccio è ancora valida ai giorni nostri mentre le tecniche utilizzate sono totalmente diverse. La costruzione della mappa di una proteina avviene attraverso varie fasi; in primo luogo la proteina in oggetto è idrolizzata in modo chimico e/o enzimatico producendo una miscela di peptidi che deve essere successivamente analizzata. Ed è qui che la caratteristica peculiare della spettrometria di massa, costituita dalla possibilità di analizzare miscele di peptidi, rende questa metodologia analitica ideale per la determinazione della mappa peptidica di proteine. Infatti è possibile analizzare direttamente la miscela di peptidi proveniente dall'idrolisi della proteina in esame mediante MALDI/MS senza ricorrere a preliminari separazioni cromatografiche. I segnali di massa osservati negli spettri sono quindi correlati ai peptidi presenti nella sequenza della proteina. Ogni segnale dello spettro è associato al corrispondente peptide sulla base del peso molecolare e, quando possibile, tenendo conto della specificità dell'enzima proteolitico utilizzato, venendo così a determinare una mappa peptidica della proteina stessa. La presenza di eventuali segnali anomali non riconducibili a peptidi presenti nella sequenza amminoacidica

della proteina identifica errori di traduzione, delezione, inserzione, sostituzioni amminoacidiche, modifiche post-traduzionali o presenza di ponti disolfuro che possono successivamente essere assegnati combinando tecniche di chimica delle proteine con analisi di spettrometria di massa.

### **Identificazione di proteine frazionate su gel mono- e/o bi-dimensionale**

Il termine proteoma è utilizzato oggi per indicare l'analisi globale di proteine in miscele molto complesse provenienti da organelli cellulari, intere cellule o addirittura organismi. I moderni studi sul proteoma possono essere divisi essenzialmente in due settori principali, il proteoma di espressione che tende alla definizione qualitativa e quantitativa dell'aumento e/o diminuzione dei livelli di proteine, e il proteoma funzionale che tenta di identificare i componenti di complessi multiproteici e vie di trasduzione del segnale. Entrambi questi approcci si basano sul frazionamento delle miscele di proteine mediante elettroforesi mono- e/o bidimensionale (1D o 2D-gel) e sulla successiva identificazione delle bande proteiche mediante tecniche di spettrometria di massa.

L'identificazione di bande proteiche da 2D-gel si ottiene essenzialmente mediante la procedura definita Peptide Mass Fingerprinting, la cui origine deriva dalle mappe peptidiche di massa descritte in precedenza, utilizzando la spettrometria di massa MALDI-MS. Lo spot proteico è scisso dal gel, digerito *in situ* con opportuni enzimi proteolitici e la miscela peptidica risultante è direttamente analizzata mediante MALDI-MS. La filosofia di questo approccio è che ogni proteina idrolizzata con un agente proteolitico produce una miscela di peptidi che è caratteristica di quella proteina. L'identificazione delle proteine quindi dipende dalla capacità di associare il set di peptidi prodotti alla proteina da cui esso proviene. Nella strategia del Peptide Mass Fingerprinting, i valori di massa accurati dei peptidi determinati mediante MALDI-MS vengono utilizzati, insieme ad altri parametri quali il tipo di proteasi usata per l'idrolisi o il peso molecolare presunto della proteina, come determinanti specifici di un particolare idrolizzato peptidico. I valori di massa registrati negli spettri vengono introdotti in alcuni programmi di ricerca disponibili in rete (ProFound, Mascot, MS-Fit, ecc.) e paragonati con quelli provenienti dalla digestione virtuale di tutte le proteine presenti nella banca dati. La coincidenza di un set di valori di massa sperimentali con quelli teorici consente l'identificazione delle proteine.

Nel caso invece dell'analisi di bande da gel monodimensionale che generalmente contengono più di una specie proteica, o quando il Peptide Mass Fingerprinting non fornisce dati definitivi, vengono utilizzate metodologie di spettrometria di massa tandem basate sulla ionizzazione ad electrospray (LC-ESMS/MS). La miscela di peptidi proveniente dall'idrolisi *in situ* della banda da gel viene analizzata mediante tecniche LC-MS/MS. I peptidi sono separati mediante HPLC capillare ed introdotti direttamente nella sorgente electrospray dello spettrometro di massa man mano che eluiscono dalla colonna cromatografica. Nella sorgente i peptidi sono ionizzati e ed il loro peso molecolare determinato.

Contemporaneamente gli ioni peptidici prodotti sono isolati uno alla volta e fram-

mentati nella cella di collisione producendo uno spettro di frammentazione da cui è possibile ottenere informazioni anche parziali di sequenza. Tutte le informazioni ottenute dall'analisi (peso molecolare e sequenza anche parziale dei frammenti proteolitici) sono utilizzati per interrogare le banche dati di proteine mediante i programmi di ricerca citati in precedenza. È stato dimostrato che le informazioni ottenute dalla sequenza parziale anche di due soli peptidi sono spesso sufficienti ad identificare la proteina in banca dati.

### **Bibliografia essenziale**

1. Matthias Mann, Ronald C. Hendrickson and Akhilesh Pandey Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* 70: 437-73, 2001.
2. Juri Rappsilber and Matthias Mann What does it mean to identify a protein in proteomics? *Trends Biochem Sci.* 27: 74-78, 2002.
3. Bernd Thiede, Wolfgang Höhenwarter, Alexander Kraha, Jens Mattow, Monika, Schmid, Frank Schmidt, Peter R. Jungblutb Peptide mass Fingerprinting. *Methods* 35: 237-247, 2005.

## **Elettroforesi bidimensionale**

*Luca Bini, L. Bianchi, R. Cianti, C. Landi, M. Puglia, C. Trombetta, D. Perini, A. Armini, V. Pallini*

Laboratorio di Proteomica Funzionale, Dipartimento Biologia Molecolare,  
Università degli Studi di Siena

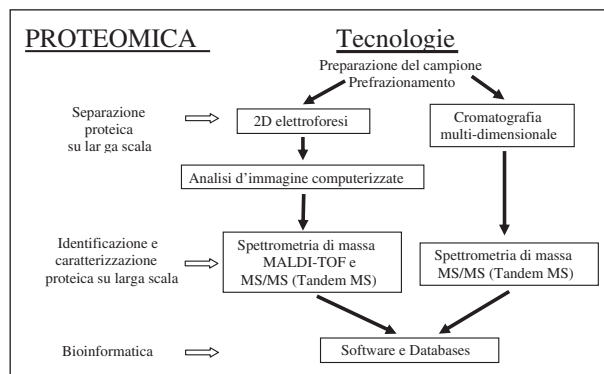
Dal 1995 ad oggi sono stati completati oltre 200 progetti di sequenziamento globale di genoma di vari Procarioti (Fleischmann et al., 1995; Fraser et al., 1995) ed organismi eucariotici complessi uni e pluricellulari tra cui l'uomo (Venter et al., 2001). Molti studi sono stati svolti per individuare sequenze di cDNA, corrispondenti a porzioni di genoma sicuramente espresse ed immagazzinate in specifiche banche dati chiamate EST (Expressed Sequence Tags). I dati attualmente disponibili per tali sequenze non sono sempre sufficienti per stabilire la funzione biologica di una proteina, ma rappresentano il punto di partenza per lo sviluppo di ulteriori esperimenti con l'obiettivo di ottenere informazioni sui patterns di espressione genica in condizioni normali e patologiche, compresa la differenziazione e la morte cellulare. Lo screening differenziale di librerie di cDNA (Adams et al., 1991) ed il cosiddetto "differential display" (Liang and Pardee 1992) di messaggeri eucariotici e la rispettiva versione quantitativa, forniscono un valido metodo per l'identificazione dei geni espressi in diversi tipi cellulari e per la misura delle differenze quantitative dell'espressione genica (Skena et al., 1995).

L'analisi globale degli mRNA presenti in un dato organismo, in un dato momento (trascrittoma), ci fornisce importanti informazioni sull'attività trascrizionale dei geni, ma non riesce a fornire elementi significativi sulle caratteristiche della proteina codificata, come la sua reale abbondanza nel contesto cellulare, dato che la quantità di messaggero non è necessariamente proporzionale alla proteina presente (Anderson and Seilhamer, 1997). Inoltre le cellule eucariotiche esprimono diverse migliaia di proteine e ciascun prodotto genico primario può subire oltre a fenomeni di splicing alternativo, una vasta gamma di modificazioni co e post-traduzionali (es. fosforilazioni, glicosilazioni, ossidazioni ecc.), fondamentali per lo svolgimento della funzione proteica, che aumentano enormemente il numero di elementi proteici presenti in una cellula o in un tessuto. Per indicare questo fenomeno è stato introdotto il neologismo PROTEOMA ad indicare tutte le PROTEINE espresse da un genOMA e successivamente modificate, durante l'intero arco della vita di una cellula o di un tessuto (Kahn 1995, Wilkins et al.,

1996, 1996a). Contrariamente al genoma, un'entità sostanzialmente statica e che, a meno di specifiche mutazioni, rimane inalterata nel corso dell'intera vita di un organismo, il proteoma rappresenta un'entità estremamente dinamica e contestodipendente, la cui composizione varia al variare dello stato di espressione del genoma e quindi in funzione del tempo, dello stato cellulare, delle condizioni ambientali e in generale dello stato fisio-patologico dell'intero organismo (Oliver, 2000). Gli approcci sperimentali legati all'analisi proteomica di campioni biologici complessi si pongono come obiettivo la visualizzazione e la caratterizzazione su larga scala dell'espressione proteica globale caratteristica del campione in analisi. Possono essere annotate in questo modo le proprietà temporali, spaziali e funzionali delle singole proteine (inclusi sia i prodotti genici primari sia i prodotti di modificazioni co- e post-traduzionali) che un dato genoma è in grado di esprimere in un certo momento, in un dato compartimento cellulare, nel contesto di uno specifico processo biologico.

Le idee e le procedure tecniche per lo studio del proteoma hanno preso forma nei primi anni novanta con lo scopo di descrivere con la maggiore ampiezza possibile i componenti proteici di cellule e tessuti, derivanti dalla espressione genica e dalle successive modificazioni. Le potenzialità e i limiti della tecnica separativa basilare, l'elettroforesi bidimensionale e lo spettro delle applicazioni della proteomica, in campo sia biomedico sia biologico sono stati descritti da Herbert et al., 1997, Hochstrasser D., 1997 e Williams and Pallini, 1997.

La proteomica rappresenta lo studio del proteoma di una cellule o di un organismo, utilizzando tecnologie su larga scala per la separazione, l'identificazione e la quantificazione delle proteine presenti. Oggi giorno esistono principalmente due tipi di approcci proteomici (Fig. 1), il primo definito, "classico", prevede l'uso dell'elettroforesi bidimensionale ad alta risoluzione su gradiente di pH immobilizzato (Bjellqvist et al., 1982 e 1993) per la separazione contemporanea di migliaia di componenti proteiche, dell'analisi computerizzata, per individuare le variazioni qualitative e quantitative, della spettrometria di massa per l'identificazione proteica su larga scala (Shevchenko et al., 1996, Aebersold and Mann, 2003) e della bioinformatica che contribuisce con specifici programmi e banche dati (Appel et al., 1994). Il secondo è basato essenzialmente sulla separa-



**Fig. 1** - Rappresentazione schematica delle principali tecnologie legate alla proteomica.

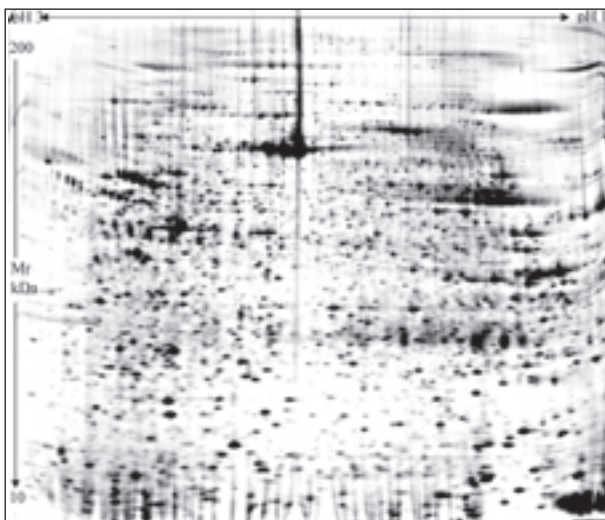


zione proteica attraverso tecniche cromatografiche spesso multidimensionali e sull'uso della spettrometria di massa per l'identificazione e la caratterizzazione dei peptidi separati (Washburn et al., 2001). Quest'ultimo approccio ha il vantaggio di permettere l'identificazione anche di proteine presenti all'interno della cellula in basse concentrazioni o proteine molto idrofobiche che difficilmente vengono viste con l'elettroforesi bidimensionale, ma, in verità, non riescono a fornire molte informazioni sulle modificazioni post-traduzionali e/o la quantità di proteina presente (Rabilloud, 2002, Schratzenholz, 2004), se non applicando specifiche marcature pre-analisi.

### Elettroforesi bidimensionale ad alta risoluzione

L'elettroforesi bidimensionale (2-D PAGE) rimane il mezzo più potente per studi di proteomica comparativa, grazie al suo elevato potere risolutivo che permette la visualizzazione e l'analisi di oltre 3000 specie proteiche distinte in un singolo gel elettroforetico (Fig. 2).

Le procedure standard di separazione tramite 2-D PAGE prevedono la completa denaturazione e riduzione del campione proteico, che viene poi sottoposto a due successivi processi di separazione elettroforetica sulla base di due distinti parametri chimico-fisici delle proteine, la carica e la massa. La prima separazione avviene per focalizzazione isoelettrica su gel di poliacrilammide in gradienti di pH immobilizzati (IPG), dove le varie proteine presenti nel campione vengono separate in funzione dei rispettivi valori di punto isoelettrico (pI), cioè in funzione della propria carica netta. Le proteine già separate secondo il proprio pI vengono poi sottoposte ad una seconda separazione elettroforetica, ortogonale alla prima, su gel di poliacrilammide in presenza di SDS e quindi separate in funzione della propria massa ( $M_r$ ). Questa doppia separazione condotta su due assi orto-



**Fig. 2** - Elettroferogramma bidimensionale colorato con nitrato di argento di una biopsia di carcinoma mammario.

gonali produce un'immagine che può essere letta come un piano cartesiano, dove ciascun prodotto proteico è individuabile sulla base di due specifiche coordinate elettroforetiche relative ai valori di pI e Mr. Il pattern proteico sul gel di poliacrilammide viene generalmente rivelato attraverso procedure di colorazione chimica post-elettroforesi, es. nitrato di argento o blue Coomassie, e poi letto con un densitometro e trasformato in immagine digitale. Tali immagini vengono poi analizzate con particolari programmi computerizzati che consentono di valutare tutte le differenze qualitative e quantitative che possono essere presenti tra campioni differenti (cellule normali e patologiche, cresciute in condizioni diverse, ecc.).

Attraverso i gel bidimensionali, in cui ogni macchia proteica ha una posizione univoca determinata dal peso molecolare e dal punto isoelettrico, si può studiare visivamente come al variare delle condizioni ambientali, di crescita, di differenziamento o patologiche, cambi l'espressione di gruppi o di singole proteine, mettendo in evidenza "sets" di polipeptidi accomunati da relazioni e/o legami metabolici, così come possono essere visualizzate classi di proteine accomunate da caratteristiche modificazioni post-traduzionali (fosforilazioni, glicosilazioni, ossidazioni, ecc.) (O'Farrell 1975; Bjellqvist et al., 1993; Herbert et al., 1997, Modesti et al., 2001, Rottoli et al., 2005). Potendo inoltre disporre delle sequenze espresse o genomiche di un organismo, tramite tecniche di spettrometria di massa e di sequenziamento proteico parziale, si può risalire dalla macchia proteica al gene codificante, attraverso ricerche di omologia con sequenze note depositate in banche dati. Una volta verificata la fattibilità di questo approccio, si può passare ad analizzare e monitorare sottosistemi (o subsets) specifici di proteine. Per questo ci possiamo avvalere dell'immunoblotting con anticorpi specifici utilizzando come sistema di rivelazione la chemiluminescenza (Modesti et al., 2001, Magi et al., 2004, Rottoli et al., 2005). Inoltre l'approccio combinato dell'elettroforesi bidimensionale con tecniche di pre-frazionamento, in grado di ridurre la complessità del campione di partenza, e tecniche di "contigazione" dei gel elettroforetici, in grado di ampliare specifiche finestre sperimentali di separazione elettroforetica, ne fanno attualmente la procedura di separazione proteica di elezione (Poland et al., 2003, Ravichandran et al., 2004).

Oggi giorno l'identificazione e la caratterizzazione proteica è ottenuta con la spettrometria di massa che tramite il calcolo della massa relativa, identifica le proteine. Questa tecnica, applicata negli ultimi anni nel campo proteomico, ha consentito di avere una quantità di identificazioni sufficienti per parlare di "high-throughput approach" anche per lo studio del proteoma (Schevchenko et al., 1996; Mann and Talbo, 1996, Roepstorff, 1997; Gygi et al., 1999; James et al., 1993). Le proteine caratteristiche di ogni campione studiato ed evidenziate dalla 2D e dalla rispettiva analisi computerizzata vengono identificate soprattutto con la spettrometria di massa MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization-Time of Flight) e ESI (ElectroSpray Ionization)-Ion Trap. I dati provenienti dall'analisi spettrometrica dei peptidi ottenuti per digestione con triplicina di spot proteici separati per elettroforesi bidimensionale vengono utilizzati, sfruttando specifici software disponibili on-line, per identificare la proteina di partenza. L'identificazione con la spettrometria di massa MALDI TOF si basa su

di un approccio di “*peptide mass fingerprinting*“, dove la lista di masse determinate sperimentalmente (corrispondenti al peso dei peptidi di tripsinizzazione) rappresenta una “impronta digitale” della proteina di partenza e può essere confrontata con le masse teoriche dei frammenti ottenuti per digestione virtuale “*in silico*“ di tutte le proteine presenti nei *database*. L’identificazione è ottenuta quando le masse dei peptici sperimentali sono uguali a quelle dei peptidi teorici. Questo tipo di approccio fornisce un metodo di identificazione rapida ed applicabile su larga scala. La spettrometria di massa ESI-IonTrap sfrutta, invece, la frammentazione dei singoli peptidi derivanti dalla miscela dei peptidi triptici per dare informazioni sulla loro sequenza amminoacidica. In questo strumento, i peptidi sono ionizzati con una ionizzazione “*electrospray*” direttamente in fase liquida. I peptidi sono poi spruzzati in un “*tandem mass spectrometer*” che ha l’abilità di separare peptidi da una miscela complessa, isolandone uno per volta e dissociandolo nei rispettivi amminoacidi. I dati di frammentazione possono essere utilizzati oltre che per cercare nelle banche dati di sequenze proteiche, anche per identificare amminoacidi modificati post-traduzionalmente, permettendo l’identificazione della posizione modificata all’interno della sequenza.

Infine, la scienza del proteoma si avvale dell’ausilio indispensabile della bioinformatica che ha il compito fondamentale di fornire programmi specifici e maneggevoli per organizzare e decifrare le migliaia di informazioni che vengono prodotte e permettere l’immagazzinamento di tutti questi dati in banche dati. L’estrema riproducibilità dell’elettroforesi bidimensionale ha permesso, anche, la costruzione di banche dati specifiche per proteine separate tramite elettroforesi bidimensionale dove, oltre alla parte testuale che riguarda l’identificazione e la caratterizzazione dei polipeptidi, sono presenti immagini di gel di riferimento adoperabili direttamente per paragonare i propri gel (SWISS-2DPAGE, <http://www.expasy.ch/ch2d/ch2d-top.html>; SIENA-2DPAGE, (Appel 1997).

Pur avendo raggiunto elevati standard di riproducibilità questa procedura risente tuttora di una variabilità sperimentale che spesso si riflette nella necessità di almeno triplicare il numero di gel elettroforetici effettuati su ciascuno specifico campione, al fine di ridurre in parte tale variabilità. Una procedura innovativa di marcatura fluorescente pre-elettroforetica, denominata DIGE (Differential In Gel Electrophoresis), rappresenta attualmente il sistema di rivelazione dei pattern proteici più affidabile per le analisi delle variazioni quantitative su campioni diversi (Tonge et al., 2001). Il sistema DIGE prevede la marcatura fluorescente differenziale di due campioni diversi prima della separazione elettroforetica che possono essere poi caricati nello stesso gel elettroforetico, che sottoposto a scansione differenziale produrrà le immagini digitali distinte dei due campioni. In questa procedura un campione A viene marcato con il fluoroforo Cy3 e un campione B viene marcato con il fluoroforo Cy5. Un terzo fluoroforo (Cy2) viene utilizzato per la marcatura della miscela dei due campioni, per essere poi utilizzato come standard interno per l’analisi delle differenze quantitative tra le immagini ottenute dal campione A colorato con Cy3 e dal campione B colorato con Cy5. La procedura di analisi quantitativa dei due campioni richiede un programma specifico disponibile commercialmente, il DeCyder Differential Analysis software, che

rispetto ai tradizionali software di analisi di immagini bidimensionali contiene algoritmi per la co-rivelazione delle immagini fluorescenti e per l'analisi quantitativa specificamente studiati per il sistema DIGE. Tali algoritmi utilizzano la miscela dei due campioni, marcata separatamente con il terzo fluoroforo, come standard interno in cui tutte le proteine dei singoli campioni sperimentali sono rappresentate. Il sistema DIGE/DeCyder rappresenta attualmente il metodo più affidabile per l'analisi quantitativa differenziale di sistemi proteici complessi, poiché permette un'analisi differenziale in-gel, eliminando così la variabilità sperimentale normalmente rilevabile nell'analisi differenziale di campioni diversi corsi su gel diversi.

## Applicazioni

Oggi giorno, in molti campi della biologia, viene seguito un approccio di studio globale come quello proteomico per capire quali siano i meccanismi di regolazione alla base dei vari fenomeni molecolari. Possiamo ricordare l'applicazione in biomedicina e in clinica, soprattutto nello studio dei tumori, delle malattie infettive e di quelle polmonari, dove lo studio del proteoma viene usato come mezzo diagnostico e prognostico oltreché per definire l'eventuale patogenesi della malattia (Bini et al., 1997, Rottoli et al., 2005, Celis e Gromov, 2003, Celis et al., 2004, Dwek and Alaiya, 2003, Zang et al., 2004, Okuzawa, 1994, Chen et al., 2003, Yanagisawa et al., 2003, Alaiya et al., 1997 e 2002, Friedman et al., 2004). Tale approccio è di grande utilità in tutte le patologie che scaturiscono da numerose e complesse interazioni geniche, oltreché da vari eventi cellulari influenzati da differenti fattori ambientali e sotto regolazioni epigenetiche. Per questo motivo un approccio globale, soprattutto basato sullo studio dei prodotti genici, come è quello proteomico, può essere fondamentale per ottenere una visione dei cambiamenti proteici implicati nel processo patogenetico. Inoltre nel campo clinico la proteomica offre il vantaggio di potere identificare e caratterizzare marcatori proteici multivariati che possono caratterizzare lo stato patologico o un trattamento farmacologico molto meglio dello studio di marcatori individuali. Infatti un insieme di marcatori proteici indipendenti l'uno dall'altro ed associati ad una patologia, considerati nel loro insieme come un unico marcatore multivariato, saranno meno soggetti alle variazioni genetiche o ambientali che non il livello di una singola proteina (Anderson, 2005).

Anche le grandi industrie farmaceutiche stanno ora usando quest'approccio per valutare l'efficacia e le eventuali conseguenze che i farmaci e le sostanze tossiche in genere possono avere sull'individuo. Altre applicazioni, non meno importanti, riguardano lo studio della complessità e delle proteine immunogeniche, oltre che il controllo di qualità dei cibi.

In definitiva l'approccio proteomico può fornire informazioni complementari e supplementari a quelle ottenute dall'analisi degli acidi nucleici, fornendo una fotografia istantanea dell'espressione proteica globale di un determinato tessuto o organismo e permettendo l'analisi delle variazioni dinamiche che in esso si verificano in seguito ad eventi fisio-patologici.

## Bibliografia

1. Adams MD., Kelley JM., Gocayne JD., Dubnick M., Polymeropoulos MH., Xiao H., Merril CR., Wu A., Olde B., Moreno RF., et al. "Complementary DNA sequencing: expressed sequence tags and human genome project". *Science*, 1991; 252 (5013): 1651-1656.
2. Aebersold R., Mann M. Mass spectrometry-based proteomics. *Nature* 2003, 422: 198-207.
3. Alaiya AA., Franzen B., Fujioka K., Moberger B. Schedvins, et al. Phenotypic analysis of ovarian carcinoma: polypeptide expression in benign, borderline and malignant tumors. *Int. J. Cancer* 1997; 73: 678-683.
4. Alaiya AA. Franzen B., Hagman A., Dysvik B., Roblick, UJ., et al. Molecular classification of borderline ovarian tumors using hierarchical cluster analysis of protein expression profiles. *Int. J. Cancer* 2002; 98: 895-899.
5. Anderson L., Seilhamer J. A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 1997; 18: 533-537.
6. Anderson L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J. Physiol* 2005; 563: 23-60.
7. Appel RD, Bairoch A, Hochstrasser DF. A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem Sci* 1994; 19: 258-260.
8. Bini L., Magi B., Marzocchi B., Arcuri F. et al. Protein expression profiles in human breast ductal carcinoma and histologically normal tissue. *Electrophoresis* 1997; 18: 2832-2841.
9. Bjellqvist B., Ek K., Righetti PG., Gianazza E., Gorg A., Westermeier R., Postel W. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* 1982; 6: 317-39.
10. Bjellqvist B., Pasquali C., Ravier F., Sanchez JC., Hochstrasser D. A nonlinear wide-range immobilized pH gradient for two-dimensional electrophoresis and its definition in a relevant pH scale. *Electrophoresis* 1993; 14: 1357-1365.
11. Celis JE., Gromov P. Proteomics in translational cancer research: toward an integrated approach. *Cancer Cell*. 2003; 3: 9-15.
12. Celis JE., Gromova I., Moreira JM., Cabezon T., Gromov P. Impact of proteomics on bladder cancer research. *Pharmacogenomics* 2004; 5: 381-394.
13. Chen G., Gharib TG., Wang H. Huang CC., Kuick R. et al. Protein profiles associated with survival in lung adenocarcinoma. *Proc. Natl. Acad. Sci. USA* 2003, 100: 13537-13542.
14. Dwek MV., Alaiya AA. Proteome analysis enables separate clustering of normal breast, benign breast and breast cancer tissues. *Br. J. Cancer* 2003, 89: 305-307.
15. Fleischmann RD., Adams MD., White O., Clayton RA. et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 1995; 269: 496-512.
16. Fraser CM., Gocayne JD., White O., Adams MD., Clayton RA., Fleischmann RD., Bult C., Kerlavage AR., Sutton G., Kelley JM., et al. "The minimal gene complement of *Mycoplasma genitalium*". *Science*, 1995; 270 (5235): 397-403.

17. Friedman DB., Hill S., Keller JW., Merchant NB., Levy SE. et al. Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics* 2004; 4: 793-811.
18. Gygi SP., Han DK., Gingras AC., Sonenberg N., Aebersold R. Protein analysis by mass spectrometry and sequence database searching: tools for cancer research in the post-genomic era. *Electrophoresis* 1999; 20: 310-319.
19. Herbert B., Sanchez J-C., Bini L. Two-dimensional electrophoresis: The state of the art and future directions. In Wilkins M., Williams K., Appel R., Hochstrasser D., editors. *Proteome research: new frontiers in functional genomics*. Springer, Berlin Heidelberg New York 1997; 13-33.
20. Hochstrasser DF (1997). "Clinical and biomedical applications of proteomics". In *Proteome research: new frontiers in functional genomics*, Wilkins M., Williams K., Appel R., Hochstrasser D. (Eds), Springer-Verlag Berlin, Heidelberg, New York, 1997; 187-219.
21. James P., Quadroni M., Carafoli E., Gonnet G. Protein identification by mass profile fingerprinting. *Biochem. Biophys. Res. Commun.* 1993; 195: 58-64.
22. Kahn P. "From genome to proteoma: looking at a cell's proteins". *Science*, 1995; 270 (369): 369-70.
23. Liang P. and Pardee AB. "Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction". *Science*, 1992; 257 (5072): 967-71.
24. Magi B., Ettore A., Liberatori S., Bini L., et al. Selectivity of protein carbonylation in the apoptotic response to oxidative stress associated with photodynamic therapy: a cell biochemical and proteomic investigation. *Cell Death Differ.* 2004; 11: 842-852.
25. Mann M., Talbo G. Developments in matrix-assisted laser desorption/ionization peptide mass spectrometry. *Curr. Opin. Biotechnol.* 1996; 7: 11-19.
26. Modesti A., Bini L., Carraresi L., Magherini F., Liberatori S., Pallini V., Manao G., Pinna LA., Raugei G., Ramponi G. Expression of the small tyrosine phosphatase (Stp1) in *Saccharomyces cerevisiae*: a study on protein tyrosine phosphorylation. *Electrophoresis* 2001; 22: 576-585.
27. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 1975; 250: 4007-4021.
28. Okuzawa K., Franzen B., Lindholm J., Linder S., Hirano T., Bergman T., Ebihara Y., Kato H., Auer G. Characterization of gene expression in clinical lung cancer materials by twodimensional polyacrylamide gel electrophoresis. *Electrophoresis* 1994; 15: 382-390.
29. Oliver S. "Proteomics: Guilt-by-association goes global". *Nature* 2000; 403: 601-603.
30. Poland J, Cahill MA, Sinha P. Isoelectric focusing in long immobilized pH gradient gels to improve protein separation in proteomic analysis. *Electrophoresis* 2003; 24: 1271-1275.
31. Rabilloud T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* 2002; 2: 3-10.
32. Ravichandran V., Lubell J., Vasquez GB., Lemkin P., Sriram RD., Gilliland



- GL. Ongoing development of two-dimensional polyacrylamide gel electrophoresis data standards. *Electrophoresis* 2004; 25: 297-308.
33. Roepstorff P. Mass spectrometry in protein studies from genome to function. *Curr. Opin. Biotechnol.* 1997; 8: 6-13.
  34. Rottoli P., Magi B., Cianti R., Bargagli E., Vagaggini C., Nikiforakis N., Pallini V., Bini L.. Carbonylated proteins in bronchoalveolar lavage of patients with sarcoidosis, pulmonary fibrosis associated with systemic sclerosis and idiopathic pulmonary fibrosis. *Proteomics* 2005; 5: 2612-2618.
  35. Schena M., Shalon D., Davis RW., Brown PO. "Quantitative monitoring of gene expression patterns with a complementary DNA microarray". *Science*, 1995; 270 (5235): 467-470.
  36. Schratzenholz A. Proteomics: how to control highly dynamic patterns of millions of molecules and interpret changes correctly? *Drug Discovery Today: Technologies* 2004; 1: 1-8.
  37. Shevchenko A., Jensen ON., Podtelejnikov AV., Sagliocco F. et al. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. USA* 1996; 93: 14440-14445.
  38. Tonge R., Shaw J., Middleton B., Rowlinson R., Rayner S., Young J., et al. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 2001; 1: 377-396.
  39. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM et al. "The sequence of the human genome". *Science*, 2001; 291 (5507): 1304-51.
  40. Washburn MP., Wolters D., Yates JR., 3<sup>rd</sup>. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 2001; 19: 242-247.
  41. Wilkins MR., Pasquali C., Appel RD., Ou K., Golaz O., Sanchez JC., Yan JX., Gooley AA., Hughes G., Humphery-Smith I., Williams KL., Hochstrasser DF. "From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis". *Biotechnology (N Y)*, 1996; 14 (1): 61-65.
  42. Wilkins MR., Sanchez JC., Williams KL., Hochstrasser DF. "Current challenges and future applications for protein maps and post-translational vector maps in proteome projects". *Electrophoresis*, 1996a; 17 (5): 830-838.
  43. Williams KL, Pallini V (1997). "Biological applications of proteomics". In *Proteome research: new frontiers in functional genomics*, Wilkins M., Williams K., Appel R., Hochstrasser D. (Eds). Springer-Verlag Berlin, Heidelberg, New York, 1997; 221-237.
  44. Yanagisawa K., Shyr Y., Xu BJ., Massion PP., Larsen PH., White BC. et al. Proteomic patterns of tumour subsets in nonsmall-cell lung cancer. *Lancet* 2003; 362: 433-439.
  45. Zang L., Toy DP., Hancock WS., Sgroi DC., Karger BL. Proteomic analysis of ductal carcinoma of the breast using laser capture microdissection, LC-MS, and 16O/18O isotopic labeling. *J. Proteome Res.* 2004; 3: 604-612.

## Tecnologie proteomiche multidimensionali in fase liquida

*Pierluigi Mauri*

Laboratorio di Proteomica e Metabolomica, Istituto di Tecnologie Biomediche (ITB-CNR), Segrate (MI)

Negli ultimi anni lo studio del proteoma ha assunto sempre maggiore importanza, suscitando un interesse crescente nella comunità scientifica. La proteomica trova applicazione in numerosi settori della ricerca, primo fra tutti il campo medico, al fine di individuare marcatori di patologie e/o bersagli terapeutici.

Tradizionalmente, lo studio dei profili proteici viene effettuato mediante la separazione su gel elettroforetico bidimensionale (2Dgel) delle proteine e la loro identificazione attraverso l'analisi MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) dei digeriti triptici (1, 2). Un miglioramento nella identificazione si è avuto con l'utilizzo della electrospray tandem mass spectrometry (ESI-MS/MS), che permette il sequenziamento dei peptici (3) ed un più facile riconoscimento delle modificazioni post-traduzionali, quali fosforilazioni e glicazioni.

Negli ultimi anni sono stati proposti metodi alternativi alla separazione in gel bidimensionali. Tra questi ricordiamo la tecnologia dei microarray di proteine (4), che utilizza supporti solidi, chip o microsferette, che permettono di legare le diverse proteine in modo più o meno specifico a seconda del legante utilizzato. Una applicazione di tale tecnologia è rappresentata dal SELDI (Surface Enhanced Laser Desorption/Ionization), in cui le diverse proteine selezionate dalle diverse superficie del chip vengono analizzate per spettrometria di massa a tempo di volo (TOF, Time of Flight) (5).

Sono state proposte altre tecniche per la separazione delle proteine, che vanno dall'elettroforesi capillare all'accoppiamento di tecniche cromatografiche come il chromatofocusing e la fase inversa (6).

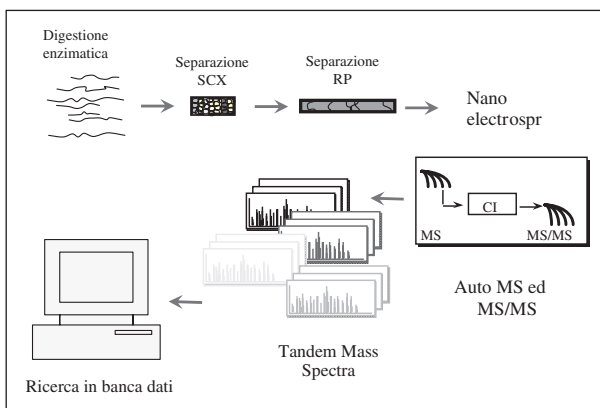
Recentemente si è fatta strada una nuova strategia proteomica, definita "topdown sequencing", che prevede la frammentazione diretta in spettrometria di massa di proteine intere senza digestione enzimatica. La misurazione del peso molecolare della proteina intatta e dei suoi frammenti, attraverso il confronto con database di sequenze, permette l'identificazione della proteina stessa. Il maggiore vantaggio di questo approccio è la potenziale copertura di sequenza al 100% e l'identificazione immediata di modifiche post-traduzionali (7). Tuttavia, tale approccio è ancora poco utilizzato perchè necessita metodi di ionizzazione poco diffusi come l'ECD (Electron Capture Dissociation) e l'IRMPD (Infrared



Multiphoton Dissociation), di spettrometri ad altissima risoluzione, e quindi, molto costosi come i FTICR (Fourier Transform Ion Cyclotron Resonance) (8) e sistemi di calcolo complessi e da validare.

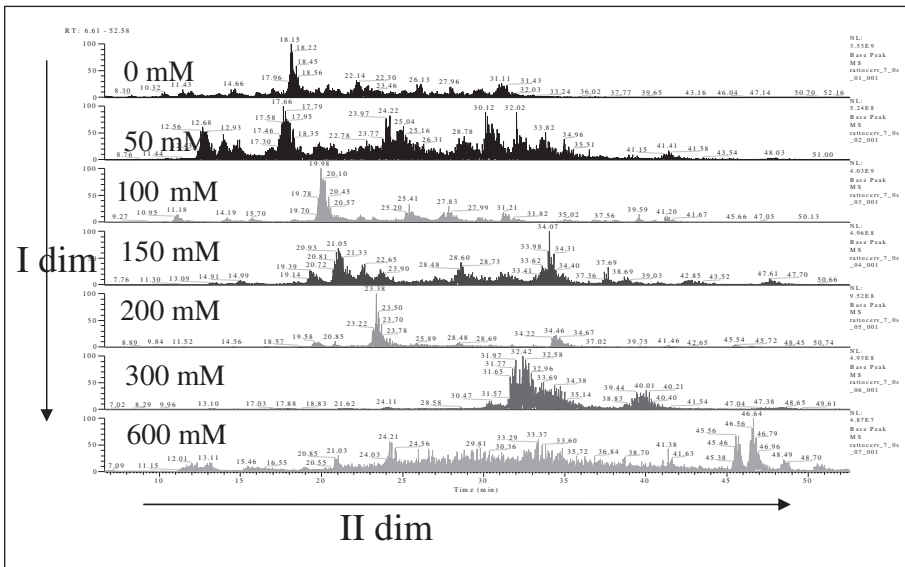
Tra le innovazioni metodologiche proposte nell'ambito dell'analisi proteomica quella che sta suscitando maggiore interesse è sicuramente l'approccio basato sulla microcromatografia bidimensionale accoppiata alla tandem mass spectrometry (2DC-MS/MS) (9). Tale approccio, chiamato anche MudPIT (Multidimensional Protein Identification Technology), o "shotgun proteomics" (proteomica su larga scala), permette di identificare simultaneamente, in modo rapido ed automatizzato, un gran numero di proteine a partire dal campione biologico, senza la necessità di frazionamenti e purificazioni (10). Il campione viene digerito enzimaticamente ed i peptidi che ne derivano sono separati per cromatografia bidimensionale (solitamente per scambio ionico SCX, seguita da una fase inversa, RP), seguita dalla analisi diretta in spettrometria MS/MS (11). Attraverso software di elaborazione (ad esempio SEQUEST o MASCOT) è poi possibile caratterizzare le sequenze dei peptidi ottenuti dalla digestione, e quindi, di identificare le proteine correlate. Tale approccio consente di identificare molte proteine (da alcune centinaia a migliaia) in un'unica analisi ed in tempi relativamente brevi (ore) senza dover ricorrere necessariamente alla separazione preventiva delle proteine. La metodologia è completamente automatizzata e permette di identificare agevolmente anche proteine difficilmente rilevabili con i gel 2D, quali ad esempio quelle basiche ( $pI > 9.5$ ) o con pesi molecolari estremi ( $< 10$  kDa,  $> 250$  kDa) o proteine molto idrofobiche (ad esempio quelle presenti nelle membrane) (12). Inoltre, la metodologia MudPIT permette di utilizzare una più ampia varietà di enzimi proteolitici, tra cui quelli con una bassa specificità di taglio, quali pepsina e proteinasi K (13, 14). La figura 1 riassume i principali step della metodologia MudPIT.

La gran quantità di spettri di massa (diverse migliaia) che si ottengono da un singolo esperimento MudPIT ha determinato la necessità di appositi strumenti informatici per l'analisi quali-quantitativa ed il confronto dei dati. I tempi di elaborazione sono stati ridotti con lo sviluppo di sistemi basati sul calcolo parallelo (PVM, Parallel Virtual Machine), utilizzando reti convenzionali di computer od

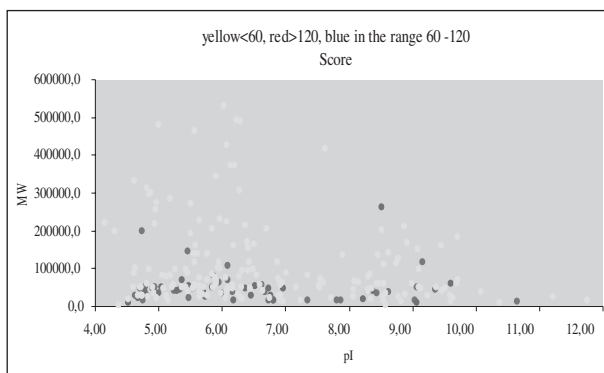


**Fig. 1** - Principali stadi dell'analisi proteomica MudPIT.

apposite architetture a “nodi” multipli (cluster). Per quanto riguarda la validazione dei dati sono stati sviluppati software specifici, permettono di eliminare le proteine la cui sequenza è interamente contenuta nella sequenza di un’altra proteina. Tali programmi permettono anche di filtrare le assegnazioni di SEQUEST in base ad alcuni parametri di analisi (DtaSelect (15)) o di validare le assegnazioni su base probabilistica (PeptideProphet-ProteinProphet (16)). Il software Contrast permette il confronto dei risultati ottenuti da DtaSelect in diverse analisi. Come esempio si riporta qui l’analisi della parte idrofobia (pellet) di un omogeneizzato di cervello. Dopo la digestione con tripsina la miscela peptidica risultante è stata caricata nel sistema MudPIT. In questo modo si è ottenuto in modo automatico sia la separazione dei peptidi sia la loro rilevazione (Fig. 2): la colonna a scambio ionico (I dimensione) separa i peptidi secondo la loro carica e li eluisce a “pacchetti” mediante l’aumento a “gradini” della concentrazione salina (da 0 a 600 mM di cloruro d’ammonio); ogni “pacchetto” di peptidi eluiti dalla colonna SCX vengono caricati direttamente sulla colonna a fase inversa ed ulteriormente separati mediante un gradiente di acetonitrile (II dimensione). Dalla colonna RP i peptidi vengono eluiti direttamente nello spettrometro di massa (di solito una trappola ionica) che ne rileva gli ioni molecolari (attraverso lo spettro “full scan”) e gli ioni di frammentazione (tandem mass spectrometry, MS/MS). Dagli spettri MS ed MS/MS vengono individuate le sequenze aminoacidiche di molti peptidi e da questi le proteine di partenza ad essi correlate. In questo modo si sono identificate centinaia di proteine presenti nel pellet di partenza. Uno strumento informatico utile per rappresentare le lunghe liste di proteine identificate attraverso l’approccio MudPIT è quello denominato MProMa



**Fig. 2** - Esempio di analisi 2DC-MS/MS di un digerito proteico della parte idrofobia di un omogeneizzato di cervello di ratto.



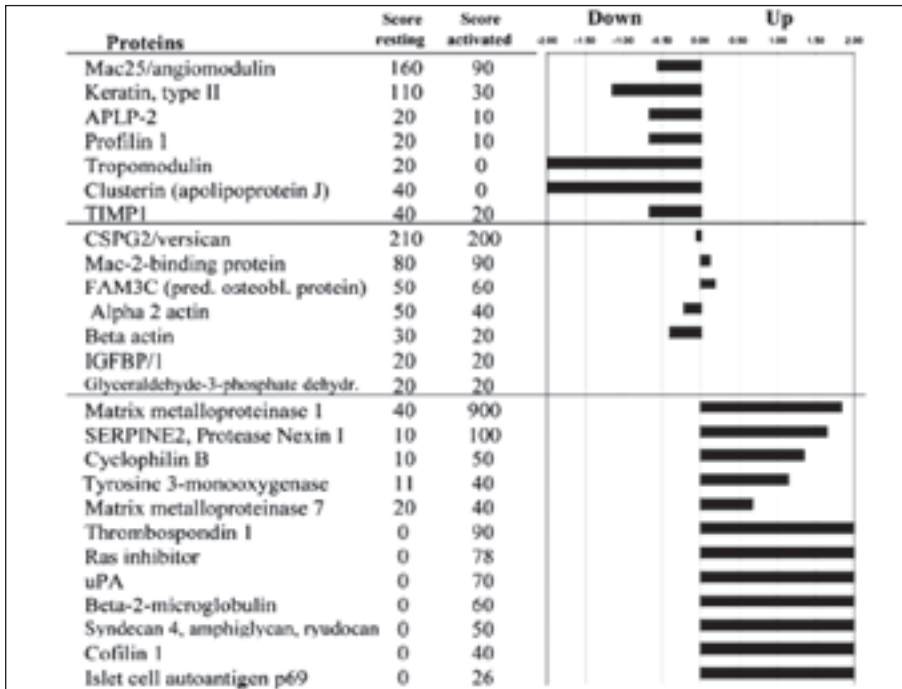
**Fig. 3** - Mappa 2D ottenuta dalle proteine identificate dall'analisi sperimentale di Figura 2.

(*MultidimensionalAlgorithmProteinMap*) (13, 17). Esso permette sia di costruire mappe bidimensionali a posteriori utilizzando pI e peso molecolare virtuali delle proteine identificate sia di ottenere il confronto tra le liste proteiche da analisi diverse. Ad esempio, dalla lista delle proteine individuate con l'analisi MudPIT riportata in Figura 2 si è ottenuta, attraverso il software MAPROMA, una mappa 2D che fornisce sia la distribuzione delle proteine sia il grado di confidenza nella identificazione di ogni proteina. (Fig. 3). Ogni "spot riportato in Figura 3 corrisponde ad una proteina identificata e si può facilmente osservare che sono presenti sia proteine ad alto PM (>300 kDa) sia proteine con pI>10. Inoltre, basandosi su molte osservazioni sperimentali si è potuto evidenziare che lo "score value" ed altri indici forniti dal software di indentificazione SEQUEST possono essere correlati alla quantità relativa di proteina presente in campioni diversi (13, 17-20). Ciò è stato tradotto in un algoritmo, inserito in MAPProMA, che permette di effettuare una prima rapida valutazione della quantità relativa delle stesse proteine identificate in campioni diversi e di costruire degli istogrammi differenziali di espressione.

Ovviamente, per ottenere le concentrazioni assolute ed in modo più preciso è necessario utilizzare altre strategie (ad esempio con isotopi stabili, tipo ICAT, (21)) che però richiedono ulteriori indagini, reattivi e costi.

Una applicazione dell'analisi proteomica MudPIT nell'ambito della ricerca di biomarcatori di malattie è relativa alle proteine rilasciate (secrete) da linee cellulari ottenute da tumore pancreatico. L'analisi 2DC-MS/MS dei digeriti triptici dei surnatanti ha permesso l'identificazione di circa 50 proteine, di cui 9 non erano mai state correlate al tumore pancreatico (17). Utilizzando l'algoritmo sviluppato in MAPProMA sono state valutate le differenze relative di espressione, confermate con altri metodi (ad esempio, immunostochimica). Inoltre, successivamente altri autori hanno confermato (22) i potenziali biomarcatori identificati mediante MudPIT (fattori angiogenici, migrazione cellulare ed al.).

In conclusione, l'approccio proteomico MudPIT, pur richiedendo maggiori investimenti iniziali rispetto al sistema 2Dgel, sta stimolando sempre più l'interesse dei ricercatori, che lo stanno utilizzando sempre più per studiare rapidamente i profili di espressione al fine di individuare biomarcatori di malattie e/o target proteici per terapie.



Relative abundance of selected proteins in supernatant from resting and activated Suit-2 cells. Da Mauri et al., FASEB J, 2005.

*Ringraziamenti - Si ringrazia la Fondazione CARIPO per i finanziamenti nell'ambito della piattaforma Trascrittomica e Proteomica (Guard WP5).*

## Bibliografia

1. Corthals GL., Wasinger VC., Hochstrasser DF., Sanchez JC. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 2000; 21: 1104-15.
2. Gygi SP., Corthals GL., Zhang Y., Rochon Y., Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A* 2000; 97: 9390-5.
3. Wysocki VH., Resing KA., Zhang Q., Cheng G. Mass spectrometry of peptides and proteins, *Methods* 2005; (35): 211-222.
4. Poetz O., Schwenk JM., Kramer S., Stoll D., Templin MF., Joos TO., Protein microarrays: catching the proteome, *Mech. ageing Devel* 2005, (126): 161-170.
5. Tang N, Tornatore P, Weinberger SR, Current developments in SELDI affinity technology, *Mass spectrometry reviews*, 2004; (23): 34-44.
6. Cu M., Siminian MH., Betgovarguez E., Roby KW., Threadgrill GJ., MALDI spotting in the biomek 3000 enables multidimensional proteome profiling, *The application notebook* 2004; 6-7.
7. Kelleher NL., Top-down proteomics, *Anal. Chem.* 2004; 17: 196-203.

8. Wysocki VH., Resing KA., Zhang Q., Cheng G., Mass spectrometry of peptides and proteins, *Methods*, 2005; (35): 211-222.
9. Link AJ., Eng J., Schieltz DM., Carmack E., Miz GJ., Direct analysis of protein complexes using mass spectrometry, *Nat Biotechnol*, 1999; (17): 676-682.
10. Washburn MP., Wolters D., Yates JR III., Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat Biotechnol*, 2001; (19): 242-247.
11. Swanson SK., Washburn MP., The continuing evolution of shotgun proteomics, *Drug Discovery Today*, 2005; 10 (10): 719-725.
12. Wu CC., and Yates JR. 3rd. The application of mass spectrometry to membrane proteomics. *Nat Biotechnol* 2003; 21: 262-267.
13. Mauri PL. Proteomics studies based on microchromatography coupled to tandem mass spectrometry and new algorithm (MAProMA) for data handling of 2DC-MS/MS results. Presentation at 16<sup>th</sup> International Mass Spectrometry Conference, Edinburgh, Scotland 2003.
14. Wu CC., MacCoss MJ., Howell KE. and Yates JR., 3rd. A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol* 2003; 21: 532-538.
15. Tabb DL., McDonald WH., Yates JR. III, DTaselect and Contrast: tools for assembling and comparing protein identifications, *J. Proteome Res.*, 2002; 21 (1): 21-26.
16. Nesvizhskii AI., Keller A., Kolker E., Aebershold R., A statistical Model for identifying Proteins by tandem Mass Spectrometry, *Anal. Chem.*, 2003; 17 (75): 4646-6458.
17. Mauri PL., Scarpa A., Nascimbeni AC., Benazzi L., Parmagnani E., Mafficini A., Della Peruta M., Bassi C., Miyazaki K., Sorio C. Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers. *FASEB Journal* 2005; 19 (9): 1125-7.
18. Regonesi ME., Del Favero M., Basilico F., Briani F., Benazzi L., Tortora P., Mauri PL., Analysis of the Escherichia coli RNA degradosome composition by a proteomic approach, *Biochimie*, 2006; 88 151-161.
19. Liu H., Sadygov RG., Yates JR., III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 2004; 76: 4193-4201.
20. Gao J., Opiteck GJ., Friedrichs MS., Dongre AR., and Hefta SA. Changes in the protein expression of yeast as a function of carbon source. *J. Proteome Res.* 2, 2003; 643-649.
21. Gygi SP., Rist B., Gerber SA., Turecek F., Gelb MH., and Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 1999; 17: 994-999.
22. Chen R., Yi EC., Donohoe S., Pan S., Eng J., Cooke K., Crispin DA., Lane Z., Goodlett DR., Bronner MP., Aebersold R., Brentnall TA. Pancreatic cancer proteome: the proteins that underlie invasion, metastasis, and immunologic escape. *Gastroenterology* 2005; 129 (4): 1187-97.

# **GENOMICA MEDICA**



## **L'avvento della genomica medica: dalle malattie monogeniche rare ai fattori di rischio genetici di malattie comuni**

*Carlo Bernasconi*

Già Professore Ordinario di Ematologia, Università di Pavia.  
Consulente ematologo e coordinatore della ricerca sulle cellule staminali,  
IRCCS Fondazione "Salvatore Maugeri", Pavia

È conoscenza senz'altro antica che l'ereditarietà influenzi la salute dell'uomo. Tuttavia, solo verso la metà del 19° secolo Mendel ha fatto le prime osservazioni fondamentali sui meccanismi di trasmissione ereditaria di caratteri costituzionali condizionanti il fenotipo, e tali conoscenze sono state applicate all'uomo solo a partire dall'inizio del '900. Durante tutto il secolo scorso sono stati gradualmente precisati e descritti i processi di trasmissione ereditaria di malattie definite come "genetiche", ma si può senz'altro affermare che tali conoscenze sono state molto importanti per la cura di pochi ammalati e poco importanti per la cura di molti ammalati.

Il recente sequenziamento del genoma umano e il rapido sviluppo delle conoscenze in questo settore di studi hanno portato negli ultimi anni ad un rinnovato interesse per la genetica. Tuttavia, rimane ancora da precisare il reale ruolo che l'informazione genetica può svolgere nella pratica medica, e anche la terminologia merita di essere chiarita. Ad esempio, quale è esattamente la differenza fra genetica e genomica? Genetica è lo studio di singoli geni e dei loro effetti. Genomica, termine coniato circa venti anni fa (McKusick e Ruddle, 1987), è lo studio non dei singoli geni ma delle funzioni e delle interazioni di tutti i geni nel genoma.

Gli obiettivi della genomica sono più ampi e ambiziosi di quelli della genetica, e sono indirizzati a chiarire i meccanismi di malattie comuni, quali il carcinoma della mammella e il carcinoma del colon-retto, l'infezione da HIV, la tubercolosi, il lupus eritematoso sistemico, la malattia di Parkinson, la malattia di Alzheimer e molte altre. Infatti, queste malattie comuni sono multifattoriali e derivano da interazioni di parecchi geni e di fattori ambientali. Inoltre, è da sottolineare che variazioni genetiche possono svolgere un ruolo non solo patogenetico, ma anche protettivo nell'insorgenza e nell'espressione di alcune malattie. Ecco quindi che la genomica sta rapidamente assumendo un'importanza sempre maggiore nella pratica medica (Gutmacher e Collins, 2002).



## Il genoma umano

Noi sappiamo che meno del 2% del genoma umano codifica per le proteine, mentre più del 50% rappresenta sequenze ripetitive di vario tipo, la cui funzione è molto poco conosciuta. Questi tratti di sequenze ripetitive, talvolta erroneamente indicati come “DNA di scarto”, costituiscono una registrazione storica della biologia evolutiva, forniscono una ricca sorgente di informazioni per la genetica di popolazioni e la genetica medica, e introducendo cambiamenti nelle regioni codificanti sono agenti attivi per mutazioni all’interno del genoma (Lander et al., 2001).

Ora l’intera sequenza del DNA di circa 3,1 gigabasi è essenzialmente completata; tuttavia, non sappiamo ancora con precisione quanti geni contiene il genoma umano. I dati attuali indicano un numero approssimativo di 30.000-35.000 geni, parecchio inferiore di quello che in un primo tempo si era ritenuto. Solo di una parte di questi geni è stata riconosciuta o prospettata la funzione, e la ricerca in questo settore di studi prosegue molto attivamente.

In circa 1.000 geni sono state identificate mutazioni cause di malattie. Tuttavia, è probabile che praticamente tutti i geni siano capaci di causare malattie se alterati in modo sostanziale. In passato si riteneva che un gene fabbricasse una proteina; oggi però sappiamo che, attraverso il meccanismo dello *splicing* alternativo (un meccanismo regolatore mediante il quale variazioni nella incorporazione di esoni, o regioni codificanti, nel RNA messaggero porta alla produzione di più di una proteina correlata, o isoforme; Graveley, 2001), dai 30.000-35.000 geni possono derivare più di 100.000 proteine. Oltre allo *splicing* alternativo altri fenomeni epigenetici, quali la metilazione e le modificazioni degli istoni, possono alterare gli effetti di un gene. Infine, un complesso insieme di segnali molecolari consente a geni specifici di essere accesi (espressi) o spenti in specifici tessuti e a tempi determinati.

I geni sono distribuiti in modo non uniforme lungo il genoma umano. Alcuni cromosomi, particolarmente il 17,19 e 21, sono relativamente densi di geni se paragonati ad altri come il 4,8, 13, 18 e Y (Venter et al., 2001). Inoltre, la densità di geni varia all’interno di ciascun cromosoma, essendo più elevata nelle aree ricche in citosina e guanina, piuttosto che in quelle ricche in adenina e timina.

Non tutti i geni risiedono nei cromosomi del nucleo; parecchie decine, coinvolte nel metabolismo energetico, sono sui cromosomi mitocondriali (Wallace, 1999). Siccome le uova sono ricche in mitocondri e gli spermatozoi no, il DNA mitocondriale è usualmente ereditato dalla madre. Quindi i geni mitocondriali (e le malattie dovute a variazioni nella loro sequenza di DNA) sono trasmessi lungo la linea materna, con modalità completamente distinte da quelle dei geni nucleari.

## Malattie monogeniche

Nel corso del 20° secolo l’associazione di osservazioni cliniche e ricerche genetiche ha chiarito l’ereditarietà di condizioni patologiche derivate da un singolo gene, o monogeniche, o definite anche “malattie mendeliane” poiché trasmesse in modo conforme con le leggi di Mendel. Migliaia di condizioni causate da muta-

zioni in singoli geni sono state identificate e catalogate in un volume (McKusick, 1998), e successivamente in un compendio informatico noto come OMIM (*Online Mendelian Inheritance in Man*).

Delle singole malattie sono state precisate le modalità di trasmissione ereditaria: autosomica dominante, autosomica recessiva, legata al cromosoma X. Negli ultimi anni sono stati poi descritti altri meccanismi per l'eredità monogenica, quali l'ereditarietà mitocondriale, l'*imprinting* e la disomia uniparentale.

Le malattie monogeniche sono quasi sempre malattie rare. Anche le più frequenti, come l'emocromatosi ereditaria (incidenza approssimativa 1:300 persone), la fibrosi cistica (incidenza approssimativa 1:3000), la deficienza di alfa<sub>1</sub>-antitripsina (incidenza approssimativa 1:1700) o la neurofibromatosi (incidenza approssimativa 1:3000) colpiscono non più di 1 su parecchie centinaia di individui in Europa e nel Nord America. Tuttavia, l'effetto totale delle malattie monogeniche è importante, riguardando non solo il singolo paziente ma anche prospettive di salute pubblica. Infatti, la comprensione dei meccanismi attraverso i quali fattori genetici possono causare malattie monogeniche ha fornito importanti informazioni su processi fisiopatologici basali, che sottostanno a disordini correlati e che si osservano con maggior frequenza rispetto alla malattia genetica. Per esempio, le conoscenze riguardanti l'ipercolesterolemia familiare, una malattia genetica che nella nostra popolazione colpisce all'incirca solo 1 su 500 individui, sono state fondamentali per capire la patogenesi dell'arteriosclerosi, che colpisce un'elevata percentuale di persone; inoltre, le stesse conoscenze hanno portato allo sviluppo di nuovi farmaci (le statine), che sono fra i medicinali più ampiamente prescritti (Goldstein et al., 2001).

### Tipi di mutazione

Ci sono diversi modi per classificare le mutazioni. Un modo è secondo i meccanismi causali; un altro modo è secondo i loro effetti funzionali.

Se si segue il criterio dei meccanismi, le mutazioni puntiformi (cioè il cambio di una singola base nella sequenza di DNA) sono le più frequenti. Un tipo sono le mutazioni di senso sbagliato (*missense mutations*), dovute a una sostituzione nella sequenza delle tre basi, o codone, che codifica per un aminoacido alternativo; queste mutazioni possono essere *conservative* o *non-conservative*, distinzione riguardante la struttura dell'aminoacido alternativo, simile o molto differente rispetto all'aminoacido originario. Un tipo di mutazioni puntiformi molto più dannoso è rappresentato dalle mutazioni nonsense (*nonsense mutations*), che generano un codone "stop", che invece di codificare per un aminoacido causa l'interruzione della proteina. Un altro tipo di mutazioni sono le mutazioni di cambio strutturale (*frame-shift mutations*), nelle quali la sostituzione nel codone di una singola base modifica la successiva struttura di lettura del gene, portando frequentemente a una prematura interruzione della sintesi proteica.

In termini di effetti funzionali, invece che di meccanismi, molte variazioni nella sequenza del genoma umano non hanno effetto fenotipico. Queste sono mutazioni silenti, con sostituzione di una base con un'altra, ma il codone risultante codi-

fica per lo stesso aminoacido. Mutazioni che possono non cambiare il fenotipo sono anche quelle in cui il codone alterato sostituisce un aminoacido con un altro simile, che produce solo piccole modifiche nella funzione della proteina che il gene codifica; questo è il caso delle mutazioni conservative. Nelle mutazioni non-conservative invece l'aminoacido sostituito è molto differente, e con ogni probabilità influenza il fenotipo.

Le mutazioni possono causare malattia attraverso varie modalità; fra queste la modalità più frequente è la perdita di funzione, che altera il fenotipo diminuendo la quantità dell'attività funzionale di una proteina. Per esempio, mutazioni nel gene della glucosio-6-fosfato deidrogenasi (*G6PD*) sul cromosoma X diminuiscono l'attività funzionale di questo enzima, causando un'anemia emolitica acuta se un maschio (che ovviamente ha solo una copia di cromosoma X) con la mutazione è esposto a certi farmaci, come sulfamidici, primachina e nitrofurantoina. Poiché non esistono geni che dipendono da agenti farmacologici, in condizioni naturali la deficienza di *G6PD* causa una crisi emolitica acuta nei maschi che ne sono affetti quando questi ingeriscono le fave (favismo), poiché l'enzima è importante nella degradazione di una componente delle fave (Luzzatto et al., 2001).

Alcune mutazioni causano malattia attraverso l'acquisizione di una funzione, nei casi in cui la proteina sintetizzata ha la capacità di esercitare una nuova funzione, che si rivela tossica. L'espansione delle triplette esoniche CAG, che si associa alla malattia di Huntington e all'atassia spinocellulare, appare generare disordini neurologici attraverso la produzione di proteine che funzionano in modo anormale a causa dell'espansione di tratti poliglutammici (CAG codifica infatti per l'acido glutammico; Goldberg et al., 1996). Queste mutazioni che fanno aumentare una funzione sono frequentemente ereditate in modo dominante, poiché una singola copia del gene mutato può alterare la funzione.

Si potrebbe ritenere che le mutazioni che interessano il 98,5% circa del genoma che non codifica per le proteine non influenzino il fenotipo. In realtà non è così, poiché anche le mutazioni regolatorie agiscono alterando l'espressione di un gene. Infatti, una mutazione regolatoria può portare ad una caduta di espressione di un gene, ad una sua inaspettata espressione in un tessuto in cui è usualmente silente, oppure a una modifica del tempo in cui esso viene espresso. Esempi di mutazioni regolatorie associate a malattie sono quelli nella regione vicina al gene *FMRI* (che causa la sindrome dell'X fragile; Jin e Warren, 2000), nella regione vicina al gene per l'insulina (con aumento del rischio per il diabete mellito di tipo 1; Todd, 1999) e in un sito regolatorio del gene del collagene di tipo I (con aumento del rischio per l'osteoporosi; Grant et al., 1996).

Le mutazioni possono anche diminuire il rischio di una malattia. Un esempio di questa possibilità è una delezione 32-bp (una *frame-shift mutation*) in un gene recettore per una chemochina, il *CCR5*. Le persone che sono omozigote per questa delezione sono quasi completamente resistenti all'infezione con HIV di tipo 1, e quelle che sono eterozigote per la delezione presentano una più lenta progressione dall'infezione all'AIDS. Questi effetti derivano dal fatto che il gene *CCR5* svolge un ruolo importante nei meccanismi di entrata del HIV nelle cellule (Dean et al., 1996).

## Geni nelle malattie comuni

Gli studi di genomica medica stanno dando un grande contributo alla conoscenza dei meccanismi di insorgenza di malattie comuni e complesse, come l'ipertensione, il diabete mellito e l'asma. Sino ad ora la maggior parte dei geni coinvolti in malattie comuni sono stati identificati grazie alla loro alta penetranza. Alcuni esempi sono le mutazioni in *BRCA1* e *BCRA2* (aumentato rischio di carcinoma della mammella e dell'ovaio; Nathanson et al., 2001), *HNPCC* (aumentato rischio di carcinoma del colon-retto senza poliposi ereditaria; Lynch, 1999), *MODY 1*, *MODY 2* e *MODY3* (aumentato rischio di diabete mellito; Froguel e Velho, 1999) e nel gene per l' $\alpha$ -sinucleina (correlata alla malattia di Parkinson; Mouradian, 2002). Se un individuo possiede una di queste mutazioni, la probabilità di sviluppare la malattia corrispondente è elevata. Tuttavia, si deve osservare che ciascuna di queste mutazioni altamente penetranti associate a malattie comuni ha una prevalenza nella popolazione generale solo di uno su parecchie centinaia a parecchie migliaia di individui.

Dal punto di vista della salute pubblica, geni con mutazioni meno altamente penetranti ma molto più prevalenti hanno un maggior effetto sulla popolazione rispetto a geni che sono altamente penetranti ma poco frequenti. Mutazioni di questo tipo sono state segnalate in geni quali l'*APC* (che aumenta il rischio del carcinoma del colon-retto; Fearnhead et al., 2001) e il fattore V Leiden (che aumenta il rischio di trombosi; Major et al., 2000).

Un esempio del raffronto fra gli effetti di mutazioni rare altamente penetranti verso mutazioni frequenti meno penetranti è fornito dalla malattia di Alzheimer. Mutazioni rare in geni di presenilina 1, presenilina 2 o precursore proteico di  $\beta$ -amiloide sono cause altamente penetranti di malattia di Alzheimer ad esordio precoce; infatti, la malattia di Alzheimer si sviluppa circa all'età di 60 anni nella maggior parte delle persone eterozigote per una mutazione in uno di questi geni (Campion et al., 1999); tuttavia, siccome molto poche sono gli individui che mostrano mutazioni di questi geni, esse svolgono un ruolo solo in meno dell'1% di tutti i casi di Alzheimer. Al contrario, l'allele  $\epsilon 4$  dell'apolipoproteina E, che aumenta il rischio della malattia di Alzheimer ad esordio tardivo, agisce in modo più subdolo. Uno studio finlandese ha trovato che la malattia di Alzheimer si sviluppa fra i 70 e gli 80 anni in circa il 21% delle persone che sono omozigote per l'allele  $\epsilon 4$ , nell'8% di quelle che sono eterozigote, e solo nel 3% di quelle senza l'allele  $\epsilon 4$  (Kuusisto et al., 1994). Siccome il 26% circa della popolazione è eterozigote e il 2% è omozigote per l'allele  $\epsilon 4$  dell'apolipoproteina (Roses, 1998), questo fattore genetico svolge un ruolo causale in molti più ammalati di Alzheimer rispetto alle mutazioni dei geni per la presenilina 1, presenilina 2 e precursore proteico di  $\beta$ -amiloide, considerate complessivamente (Guttmacher e Collins, 2002).

## Variazioni nel genoma umano

Una caratteristica del genoma umano rilevante sul piano medico e sociale è che, in media, due persone non consanguinee condividono più del 99,9% delle loro

sequenze di DNA (Venter et al., 2001). Ciò significa che, siccome il genoma umano è costituito da più di tre bilioni di paia di basi, le sequenze di DNA di due individui non consanguinei variano per milioni di basi. Poiché il genotipo di una persona rappresenta una mescolanza dei genotipi dei genitori, ognuno di noi è eterozigote per almeno tre milioni di basi. Attualmente è in atto un grande sforzo di ricerca per catalogare queste varianti, comunemente indicate come “polimorfismi di nucleotidi singoli” (*single-nucleotide polymorphisms, SNPs*) e per correlare queste specifiche variazioni del genotipo con specifiche variazioni del fenotipo importanti per la salute.

Alcune correlazioni SNP-fenotipo esprimono un risultato diretto dell'influenza dell'SNP sulla salute. Tuttavia, più frequentemente l'SNP è semplicemente un marcatore di diversità biologica, che correla con la salute per la sua vicinanza al fattore genetico che è la causa reale di una malattia. Il termine di vicinanza è solo una grossolana misura di vicinanza fisica; ma dimostra che, poiché il materiale genetico è passato attraverso 5.000 generazioni dal nostro comune pool ancestrale africano, la ricombinazione fra SNP e fattore genetico è avvenuta solo raramente. In termini genetici il fenomeno dell'associazione nonrandom di una popolazione di alleli in loci vicini viene indicato come “disequilibrio di fase dei gameti”, o più comunemente “*linkage disequilibrium*”.

Oltre agli sforzi fatti per catalogare gli SNP correlandoli al fenotipo, si debbono sottolineare quelli fatti per mappare gli aptotipi. Mentre gli SNP rappresentano una variante di un singolo nucleotide, gli aptotipi consistono in sequenze di nucleotidi notevolmente lunghe (all'incirca 25.000), che, come qualsiasi variante, tendono ad essere ereditate intere. SNP e aptotipi saranno la chiave per gli studi di associazione (cioè gli studi di persone affette verso soggetti controllo) necessari per identificare i fattori genetici in malattie complesse e comuni, come gli studi di familiarità sono stati importanti per l'identificazione dei geni coinvolti nelle malattie monogeniche (Reich et al., 2001; Goldstein et al., 2001). Almeno sino a quando non sarà possibile eseguire in clinica il sequenziamento dell'intero genoma di un singolo paziente, l'identificazione di SNP e aptotipi rappresenta lo sforzo attuale di utilizzare la genomica medica per una terapia personalizzata.

Dall'insieme delle recenti conoscenze di genomica, che sono state qui brevemente riportate, risulta evidente la necessità che il medico acquisisca rapidamente il concetto della variabilità genetica, delle sue interazioni con l'ambiente, delle sue implicazioni per la cura dei pazienti. Infatti, da disciplina di base per la ricerca biomedica, la genomica sta rapidamente guadagnando una posizione centrale nella medicina clinica.

## **Bibliografia**

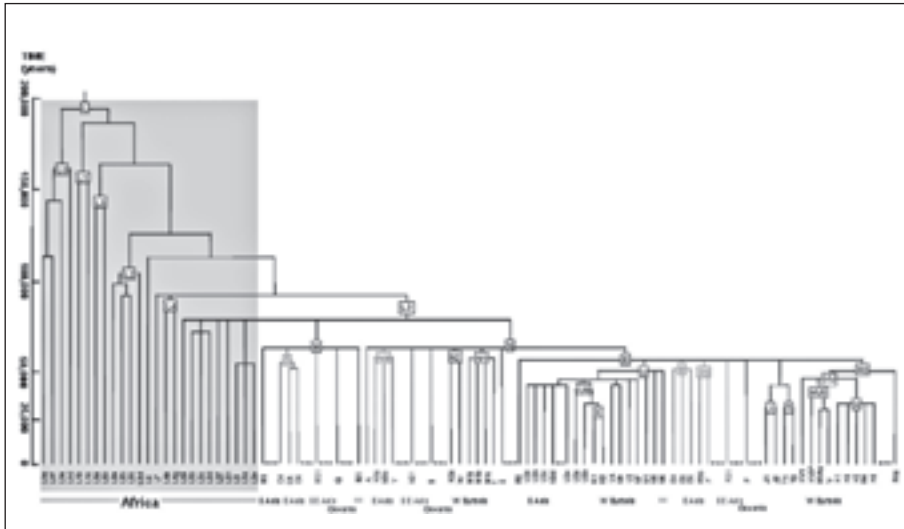
1. Champion D., Dumanchin C., Mannequin D. et al., Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am. J. Hum. Genet.* 65, 664, 1999.
2. Dean M., Carrington M., Winkler C. et al., Genetic restriction of HIV-1

- infection and progression to AIDS by a deletion allele of the *CKR5* structural gene. *Science*, 273, 1856, 1996.
3. Fearnhead N.S., Britton M.P., Bodmer W.F. The ABC of APC. *Hum. Mol. Genet.*, 2001, 721, 2001.
  4. Froguel P., Velho G. Molecular genetics of maturity-onset diabetes of the young. *Trends Endocrinol. Metab.*, 10, 142, 1999.
  5. Goldberg Y.P., Nicholson D.W., Rasper D.M. et al. Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamic tract. *Nat.Genet.*, 13, 442,1996.
  6. Goldstein DB., Weale ME. Population genetics: linkage disequilibrium holds the key. *Curr. Biol.*, 11, R576, 2001.
  7. Goldstein JL., Hobbs HH., Brown MS. et al. Familial hypercholesterolemia. In: Scriver CR., Beauder AL., Sly WS., Valle D., eds. *The metabolic and molecular bases of inherited diseases*. 8th ed. Vol 2. McGraw-Hill, New York, 2001.
  8. Grant S.F.A., Reid D.M., Blake G. et al. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I  $\alpha 1$  gene. *Nat Genet.*, 14, 203, 1996. Graveley B.R. Alternative splicing: increased diversità in the proteomic world. *Trends Genet.*, 17, 100, 2001.
  9. Guttmacher AE., Collins FS. *Genomic medicine - A primer*. N. Engl. J. Med., 347, 1512, 2002.
  10. Jin P., Warren ST. Understanding the molecular basis of fragile X syndrome. *Hum. Mol. Genet.*, 9, 901, 2000.
  11. Kuusisto J., Koivisto K., Kervinen K. et al. Association of apolipoproteine E phenotypes with late onset Alzheimer's disease: population base study. *Brit. Med. J.*, 309, 636, 1994.
  12. Lander ES., Linton LM., Birren B. et al. Initial sequencing and analysis of the human genome. *Nature*, 409, 2001.
  13. Luzzatto L., Mehta A., Vulliamy T. Glucose 6-phosphate dehydrogenase deficiency. In: Scriver CR., Beauder AL., Sly WS., Valle D., eds. *The metabolic and molecular bases of inherited diseases*. 8th ed. Vol 3. McGraw-Hill, New York, 2001.
  14. Lynch HT. Hereditary nonpolyposis colorectal cancer (HNPCC). *Cytogenet. Cell Genet.*, 86, 130, 1999.
  15. Major DA., Sane DC., Harrington DM. Cardiovascular implications of the factor V Leiden mutation. *Am. Heart. J.* 140, 189, 2000.
  16. McKusick VA., Ruddle FH. A new discipline, a new name, a new journal. *Genomics*, 1, 1, 1987.
  17. Mouradian MM. Recent advances in the genetics and pathogenesis of Parkinson disease. *Neurology*, 58, 179, 2002.
  18. Nathanson K.N., Wooster R., Weber B.L. Breast cancer genetics: what we know and what we need. *Nat. Med.*, 7, 552, 2001.
  19. *Online Mendelian Inheritance in Man, OMIM™*. Baltimore: McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University, 2000. (Accessed October 15, 2002, at <http://www.ncbi.nlm.nih.gov/omim/>).

20. Reich DE., Cargill M., Bolk S. et al. Linkage disequilibrium in the human genome. *Nature*, 411, 199, 2001.
21. Roses AD. Alzheimer diseases: a model of gene mutations and susceptibility polymorphisms for complex psychiatric diseases. *Am. J. Med. Genet.*, 81, 49, 1998.
22. Todd J.A. From genome to aetiology in a multifactorial disease, type 1 diabetes. *Bioessays*, 21, 164, 1999.
23. Venter JC., Adams MD., Myers EW. et al. The sequence of the human genome. *Science*, 291, 1304, 2001.
24. Wallace D.C. Mitochondrial diseases in man and mouse. *Science*, 283, 1482, 1999.





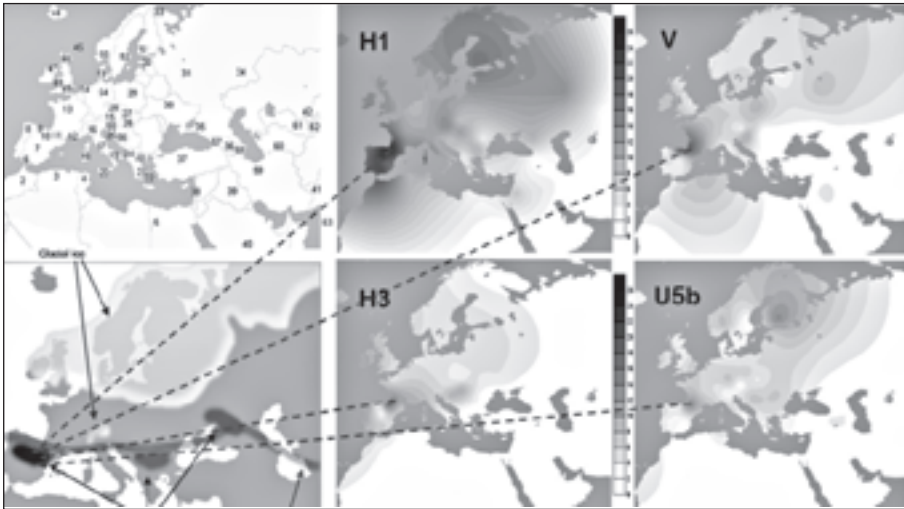


**Fig. 2** - Schematic representation of the worldwide phylogeny of human mtDNA. African haplogroups are in green and those of other geographical regions are in other colors (from Olivieri et al., 2006).

maternal lineages. The time frame of these mutations is such that human mtDNA contains a molecular record not only of the genealogical history but also of the migrations of women who transmitted it through the generations. Because the process of molecular differentiation is relatively fast and occurred mainly during and after the recent process of dispersal into different parts of the world, subsets of mtDNA variation (haplogroups) generally tend to be restricted to particular geographic areas and populations (Fig. 2).

Recently, mtDNA studies have entered a new phase: the blossoming of complete genome analyses. Of course, sequencing complete mtDNAs is much more expensive and requires much more labour than restriction analysis or simply sequencing the control region of the molecule. However, the efforts are paying off, as the phylogenetic resolution of the mtDNA tree has been greatly improved, meaning that, in turn, phylogeographic interpretations can be given correspondingly greater precision in terms of the timing and direction of human dispersals. Therefore, despite the fact that mtDNA makes up only a tiny fraction of our total genome, the deciphering of its evolution is profoundly changing our perception about how modern humans spread across our planet. I will illustrate the phylogeographic approach with four case studies: the initial dispersal out of Africa (1), an ancient back-migration to Africa (2), the human recolonization of Europe after the Last Glacial Maximum (3-5) (Fig. 3), and the origin of the Etruscans (6).

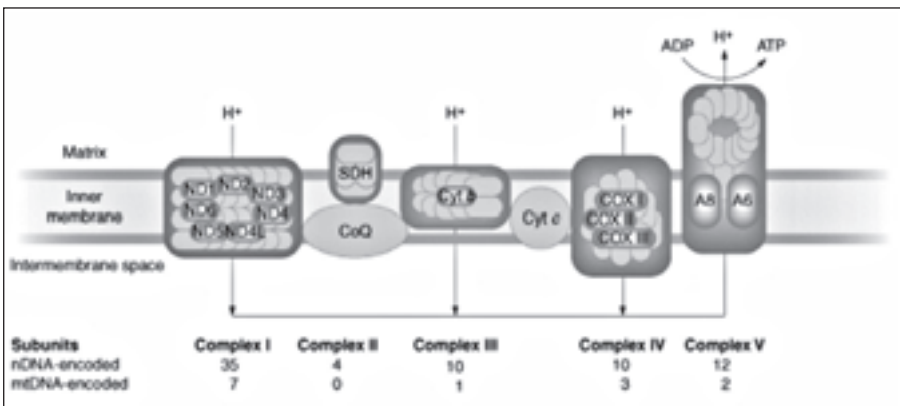
However, mtDNA is not only a string of DNA with peculiar features useful for evolutionary studies; it contains genes involved in the generation of cellular energy through the process of oxidative phosphorylation (Fig. 4). Thus, the high mutation rate of the mitochondrial genome has resulted not only in the accumulation of the broad spectrum of sequence polymorphisms that distinguish hap-



**Fig. 3** - Refuge areas during the Last Glacial Maximum (LGM) and frequency maps of mtDNA haplogroups H1, H3, V and U5b. The geographic location of the populations surveyed is shown in the upper left-hand side of the panel (modified from Torroni et al., 2006).

logroups and sub-haplogroups, but also in a large number of disease-causing mutations. These diseases generally manifest in organs and tissues that rely heavily on mitochondrial energy production (7), and Leber hereditary optic neuropathy (LHON) (MIM 535000) was the first disease to be linked with an mtDNA point mutation (8).

In addition to the search for disease-causing mutations, the sequence variation of the human mitochondrial genome has been the focus of numerous studies investigating its role as a contributing factor in the pathogenesis of neurodegenerative



**Fig. 4** - Schematic representation of the mitochondrial respiratory chain. The thirteen mtDNA-encoded subunits are shown in red and nuclear DNA (nDNA)-encoded subunits are in violet. In the lower part the number of nDNA and mtDNA-encoded subunits for each complex are reported.

diseases. Indeed, it has been proposed that certain mtDNA polymorphisms, or combinations thereof, may confer a higher risk for certain complex diseases - in other words that the supposedly “neutral” mutations that define haplogroups and sub-haplogroups play a role in disease expression. The first strong evidence for such a haplogroup effect was reported in 1997, and this came again from studying LHON. Four concomitant investigations carried out in patients of diverse European origins revealed that the LHON mutations 11778/ND4 and 14484/ND6 were preferentially associated with the Western Eurasian haplogroup J (9-12). This observation led to the idea that the mutational motif that defines J might increase the penetrance of the LHON mutations and the risk of disease expression (11).

After these findings about LHON, association studies with haplogroups became a common approach to unravel the role of mtDNA sequence variation in various disorders and phenotypes, and have grown exponentially. For example, haplogroup J has also been associated with successful ageing, and associations have been reported between haplogroup U and occipital stroke in migraine, haplogroup J and optic neuritis, haplogroup T and asthenozoospermia, haplogroup H and migraine without aura and cyclic vomiting syndrome, haplogroup D4h and bipolar disorder, haplogroups J2 and K and sprint athletic performances. However, except for LHON, virtually all reported associations between haplogroups and diseases have remained provisional. The findings of one study have rarely been replicated by studies in other populations, and conflicting associations have been often reported. One striking example of inconsistency is highlighted by the recent studies on Parkinson Disease (PD). It has been proposed that a reduced risk of PD for Europeans is associated with haplogroups J and K (13), or the supercluster UKJT (14), or haplogroup K alone (15), while other studies suggested an increased risk with either the super-haplogroup JT (16), or the supercluster JTIWX (17).

A very recent study carried on patients with LHON provides important clues about haplogroups effects (18). It revealed that mtDNA association studies are probably affected by a major specific problem: the resolution of the mtDNA haplogroup structure is generally too low, and the definition of sub-haplogroups is necessary. An update on this issue will be provided.

In conclusion, for the small mtDNA molecule the final level of molecular resolution - that of complete sequences - has been reached, but the task of determining and interpreting the worldwide variation and its potential role in disease expression has only just begun, with tens of thousands of mtDNAs to be selected worldwide for complete sequencing in the next few years.

## References

1. Macaulay V. et al. (2005). Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308: 1034-1036.
2. Olivieri A. et al. (2006). The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science* 314: 1767-1770.
3. Achilli A. et al. (2004). The molecular dissection of mtDNA haplogroup H

- confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* 75: 910-918.
4. Achilli A. et al. (2005). Saami and Berbers-an unexpected mitochondrial DNA link. *Am J Hum Genet* 76: 883-886.
  5. Torroni A. et al. (2006). Harvesting the fruit of the human mtDNA tree. *Trends Genet* 22: 339-345.
  6. Achilli A. et al. (2007). Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. *Am J Hum Genet* 80: 759-768.
  7. Wallace DC. (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annu Rev Genet* 39: 359-407.
  8. Wallace DC. et al. (1988). Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* 242: 1427-1430.
  9. Brown MD. et al. (1997). Clustering of Caucasian Leber hereditary optic neuropathy patients containing the 11778 or 14484 mutations on an mtDNA lineage. *Am J Hum Genet* 60: 381-387.
  10. Hofmann S. et al. (1997). Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. *Hum Mol Genet* 6: 1835-1846.
  11. Lamminen T. et al. (1997). mtDNA haplotype analysis in Finnish families with Leber hereditary optic neuroretinopathy. *Eur J Hum Genet* 5: 271-279.
  12. Torroni A. et al. (1997). Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *Am J Hum Genet* 60: 1107-1121.
  13. van der Walt JM. et al. (2003). Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease. *Am J Hum Genet* 72: 804-811.
  14. Pyle A. et al. (2005). Mitochondrial DNA haplogroup cluster UKJT reduces the risk of PD. *Ann Neurol* 57: 564-567.
  15. Ghezzi D. et al. (2005). Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur J Hum Genet* 13: 748-752.
  16. Ross OA. et al. (2003). mt4216C variant in linkage with the mtDNA TJ cluster may confer a susceptibility to mitochondrial dysfunction resulting in an increased risk of Parkinson's disease in the Irish. *Exp Gerontol* 38: 397-405.
  17. Autere J. et al. (2004). Mitochondrial DNA polymorphisms as risk factors for Parkinson's disease and Parkinson's disease dementia. *Hum Genet* 115: 29-35.
  18. Carelli V. et al. (2006). Haplogroup effects and recombination of mitochondrial DNA: novel clues from the analysis of Leber hereditary optic neuropathy pedigrees. *Am J Hum Genet* 78: 564-574.

## Mitochondrial disorders

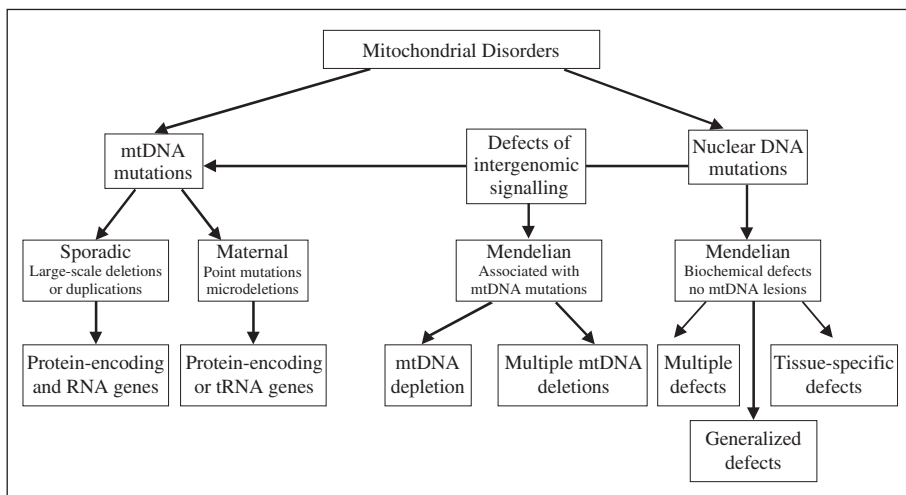
*Massimo Zeviani*

U.O. Neurogenetica Molecolare, Istituto Nazionale Neurologico "Carlo Besta", Milano

Mitochondrial diseases comprise a wide range of clinical phenotypes associated with failure of mitochondrial oxidative phosphorylation (OXPHOS). The best-known morphological finding of these disorders is the transformation of scattered muscle fibres into "ragged red fibres" (RRF), an accumulation of abnormal mitochondria under the sarcolemmal membrane. An important biochemical hallmark is the presence of high levels of lactic acid, in blood as well as in cerebro-spinal fluid.

### Classification

OXPHOS is at the cross-point of the two separated and relatively autonomous genetic systems, the nuclear and the mitochondrial genomes (nDNA, mtDNA). The identification of mutations of mtDNA has provided the basis for the current classification of mitochondrial disorders (Fig. 1). A first group of syndromes is characterised by the presence of sporadic or maternally transmitted mtDNA muta-



**Fig. 1** - Classification of mitochondrial diseases.

tions. A second group is characterised by the association of mtDNA abnormalities with Mendelian transmission of the trait (defects of nucleus-mitochondrial signalling). Finally, mutations in nuclear genes coding for mitochondrially located peptides may lead to mitochondrial disease in the presence of normal mtDNA.

### Mutations of mtDNA

Under normal conditions, the mitochondrial genotype of an individual is composed of a single mtDNA species (homoplasmy). Mutations of mtDNA can lead to *heteroplasmy*, where the wild-type and the mutant genomes co-exist intracellularly. Only when mutated copies of mitochondrial genomes accumulate over a certain threshold, the deleterious effects of the mutation will be expressed phenotypically. In addition, phenotypic expression will depend upon the nature of the mutation, its tissue distribution, and the relative reliance of each organ system on the mitochondrial energy supply.

According to the molecular and genetic features of the mutation of mtDNA, clinical syndromes can be due to:

- a) Large-scale rearrangements of mtDNA.
- b) Point mutations and micro-rearrangements of mtDNA.

### Rearrangements of mtDNA

Large-scale rearrangements of mtDNA can be either single partial mtDNA deletions or, more rarely, partial duplications. Both types of mutation are heteroplasmic (i.e. they coexist with wild-type mtDNA). Three main clinical phenotypes are associated with these mutations: *Kearns-Sayre syndrome* (KSS), *sporadic progressive external ophthalmoplegia* (PEO), and *Pearson's syndrome*.

*Kearns-Sayre syndrome* is a sporadic, severe disorder characterised by the invariant triad of PEO, pigmentary retinopathy, and onset before age 20. Frequent additional symptoms are a progressive cerebellar syndrome, cardiac conduction blocks and increased protein content in the CSF.

*Sporadic PEO* is characterised by bilateral ptosis and ophthalmoplegia, frequently associated with proximal muscle weakness and exercise intolerance.

*Pearson's bone marrow-pancreas syndrome* is a rare sporadic disorder of early infancy characterised by sideroblastic anemia with pancytopenia and exocrine pancreatic insufficiency. Interestingly, infants surviving into childhood may develop the clinical features of KSS.

In most cases, the disorders caused by mtDNA rearrangements occur sporadically. Most likely, they originate during oogenesis in the patient's mother. Transmission of single duplications or deletions of mtDNA from an affected mother to a child has occasionally been reported, which prompted to attribute a 5% risk of recurrency for these conditions. For diagnostic purposes, muscle biopsy is recommended because in both KSS and PEO the presence of a mtDNA rearrangement can often be detected only in muscle DNA. In Pearson's syndrome, mtDNA deletions are detected in blood as well as in other tissues.

## Point mutations of mtDNA

Point mutations of mtDNA are usually maternally inherited, and affect either protein-synthesis genes (tRNA or rRNA) or protein-encoding genes. The former may impair the overall mitochondrial protein synthesis, while the latter impair the activity of the specific respiratory chain complex to which the mutant protein belongs. In spite of the very many mtDNA mutations identified as pathogenic in humans (Mitomap 2006), only four of them recur with relatively high frequency in different human populations: the A3243G, G8344A, T8993G and A11778G, associated with MELAS, MERRF, NARP and LHON syndromes, respectively.

*Mitochondrial Encephalomyopathy, Lactic Acidosis and Strokelike episodes (MELAS)* is defined by the presence of (i) stroke-like episodes due to focal brain lesions often localised in the parieto-occipital lobes, (ii) lactic acidosis and/or RRF. Other signs of central nervous system involvement include dementia, recurrent headache and vomiting, focal or generalized seizures, and deafness. Diabetes, intestinal pseudo-obstructions and cardiomyopathy may also complicate the clinical course

MELAS was first associated with a heteroplasmic point mutation in the tRNA<sup>Leu(UUR)</sup>, an A->G transition at position 3243. Other MELAS-associated point mutations were later reported, although the A3243G remains by far the most frequent. The genotype-phenotype correlation of the A3243G mutation is rather loose since the A3243G mutation has been detected in several patients with maternally-inherited PEO, isolated myopathy, cardiomyopathy, or in pedigrees with maternally inherited diabetes mellitus and deafness.

*Myoclonus Epilepsy with Ragged-Red Fibres (MERRF)* is a maternally inherited neuromuscular disorder characterised by myoclonus, epilepsy, muscle weakness and wasting, cerebellar ataxia, deafness and dementia.

The most commonly observed mutation of mtDNA associated with MERRF is an A->G transition at np 8344 in the tRNA<sup>Lys</sup> gene. Even though the genotype-phenotype correlation between MERRF syndrome and the A8344G mutation is tighter than that of other mutations, the A8344G transition has also been reported in phenotypes as different as Leigh syndrome, isolated myoclonus, familial lipomatosis, and isolated myopathy.

*Neurogenic weakness, Ataxia and Retinitis Pigmentosa (NARP)* is a maternally-inherited syndrome associated with a heteroplasmic T->G transversion at position 8993 in the ATPase 6 subunit gene. RRF fibres are absent in the muscle biopsy.

When the percentage of mutant mtDNA is more than 95% patients show the clinical, neuroradiologic and neuropathologic findings of Leigh syndrome (hence called MILS, Maternally Inherited Leigh Syndrome). The two phenotypes of NARP and MILS may coexist in the same family.



*Leber's Hereditary Optic Neuropathy (LHON)* is characterised by bilateral, acute or subacute loss of central vision due to optic atrophy. The visual defect can occasionally be associated with cardiac conduction abnormalities (pre-excitation syndrome). Penetrance of LHON is much higher in males than in females, and onset is usually in the second and third decade.

A total of 17 mtDNA mutations are now reported to be associated with LHON. However, based on numerous genetic, clinical and biochemical parameters, mutations at three nucleotide positions, np 11778 (subunit 1 of complex I, ND1), np 3460 (subunit 4 of complex I, ND4), and np 14484 (subunit 6 of complex I, ND6) encompass most of LHON families in all human populations, and are considered as high-risk and primary LHON mutations. A combination of LHON and dystonia has for long been recognized as a distinct entity. A pathogenic G->A transition at np 14459 in the ND6 gene (complex I) has been reported in three independent LHON-dystonia pedigrees.

### **Other mtDNA point mutations**

Several other point mutations of mtDNA have been detected in single patients or in pedigrees affected with different phenotypes. For instance, *Myopathy and Cardiomyopathy (MMC)* is associated with a number of different tRNA gene mutations, and aminoglycoside-induced or spontaneous, *non-syndromic progressive deafness* has been associated with a homoplasmic A->G transition at np 1555 in the 12S rRNA gene.

### **Disorders due to nDNA mutations**

A clinical-genetic classification of mitochondrial disorders due to mutations in nuclear genes includes:

- 1) disorders due to mutations in respiratory chain subunits;
- 2) disorders due to mutations in ancillary proteins;
- 3) disorders due to faulty intergenomic communication affecting the maintenance and expression of mtDNA;
- 4) disorders due to defects in biosynthetic enzymes for lipids or cofactors;
- 5) disorders due to gene defects encoding proteins involved in mitochondrial biogenesis.

### **Mutations in respiratory chain subunits**

Although 72 of the 85 subunits of the OXPHOS system are encoded by nuclear DNA, mutations in nuclear genes have so far been identified in a minority of patients.

Isolated Complex I deficiency is the most common biochemical finding in OXPHOS disease being present in about one-third of the cases. It is composed of 45 proteins, seven of which are encoded by mtDNA, while the remaining 38 are encoded by the nuclear genome. Mutations in mtDNA ND genes are detected in



only 20% of the children with complex I deficiency, suggesting that the vast majority of patients may have mutations in nuclear genes encoding structural Complex I subunits or assembly factors. Extensive sequence analysis revealed mutations in several structural-subunit genes (*NDUFS1*, *NDUFS2*, *NDUFS3*, *NDUFS4*, *NDUFS6*, *NDUFS7*, *NDUFS 8*, *NDUFV1* and *NDUFV2*). The clinical picture is variable but the majority of patients show early onset neurological symptoms including hypotonia, ataxia, psychomotor retardation or Leigh syndrome (LS, see below). Cardiomyopathy has also been reported in patients with mutations in the genes encoding *NDUFS2* and *NDUFV2*.

Complex II is composed of four nuclear encoded subunits. Mutations in *SDHA*, the gene encoding the 75 kDa flavoprotein subunit of complex II, have occasionally been found in children with Leigh syndrome (see below) but also in two sisters with a late onset neurodegenerative disease. Mutations in *SDHB*, *SDHC* and *SDHD* have been associated with pheochromocytoma and with hereditary paraganglioma.

### Mutations in ancillary proteins

A second group includes nuclear disease genes encoding factors, which are not part of the protein backbone, however are involved in the formation, turnover, and function of the respiratory chain complexes. The most important gene of this group is *SURF-1*, which is responsible for Leigh Syndrome associated with severe deficiency of cytochrome c oxidase (LSCOX). *SURF-1* encodes a COX assembly factor; the SURF-1 protein is absent in all tissues of affected LSCOX individuals (Tiranti et al., 1999).

*Leigh syndrome* (LS, MIM 256000), or *Subacute Necrotizing Encephalomyelopathy*, is an early-onset progressive neurodegenerative disorder. Affected infants show severe psychomotor delay, cerebellar and pyramidal signs, dystonia, respiratory abnormalities, incoordination of ocular movements, and recurrent vomiting. MRI shows focal symmetric lesions in the brainstem, thalamus and posterior columns of the spinal cord. RRF are consistently absent in skeletal muscle. LS is a genetically heterogeneous entity. In some cases it is attributable to mtDNA mutations (e.g. the T8993G "NARP/MILS" mutation), in others to an autosomal recessive defect of a nuclear gene. In still other cases, the defect is X-linked or sporadic, as in the case of the defect of the E1-alpha subunit of PDH. In any case, all defects described to date in patients with LS affect the terminal oxidative metabolism and are likely to impair energy production.

Other COX assembly factors include *SCO2*, *SCO1*, *COX10*, and *COX15*, associated with COX-deficient Leigh syndrome or other multisystemic fatal infantile disorders, in which encephalopathy is accompanied by cardiomyopathy (*SCO2*, *COX15*), nephropathy (*COX10*), or hepatopathy (*SCO1*).

Mutations in a complex III assembly protein, *BCS1L*, have also been associated with Leigh-like syndromes and with a lethal infantile disorder dubbed GRACILE (growth retardation, aminoaciduria, cholestasis, iron overload, lactic acidosis, and early death).

The only pathogenic mutation in a complex V assembly gene so far identified, *ATP12*, was associated with congenital lactic acidosis and a fatal infantile multi-systemic disease involving brain, liver, heart, and muscle.

The newest addition to this group of disorders are due to mutations in three complex I assembly factors, *NDUFA12L*, *NDUFAF1* and *C6ORF66*. The first gene was associated with a clinical presentation resembling a leukoencephalopathy with vanishing white matter; the additional two genes have been linked to a severe cardioencephalomyopathy.

### Defects of intergenomic signaling

Autosomal disorders classified as defects of nuclear-mitochondrial intergenomic communication can be associated with the accumulation of mtDNA large-scale rearrangements (mtDNA breakage syndromes) or characterized by severe reduction of the mtDNA copy number. The most frequent clinical presentations are:

- 1) an adult-onset encephalomyopathy, defined clinically by PEO, genetically by the autosomal dominant (ad) or, more rarely, recessive (ar) transmission of the trait, and molecularly by the presence of multiple deletions of mtDNA;
- 2) an autosomal recessive multisystem disorder known as mitochondrial neurogastrointestinal encephalomyopathy, characterized by combined accumulation of multiple deletions and partial depletion of mtDNA;
- 3) a spectrum of, recessive neurological syndromes ranging from typical infantile hepatopathic poliodystrophy (Alpers-Huttenlocher syndrome) to juvenile onset sensory-ataxia neuropathy, dysarthria and ophthalmoplegia to a combination of spinocerebellar ataxia and epilepsy with or without external ophthalmoplegia;
- 4) early-onset, organ-specific autosomal recessive syndromes associated with profound mtDNA depletion.

Maintenance of mtDNA requires the concerted activity of several nuclear-encoded factors that participate in its replication, as part of the mitochondrial replisome or by supplying dNTPs to mitochondria.

#### *mtDNA breakage syndromes*

As already mentioned, autosomal dominant mtDNA breakage syndrome are characterized by the accumulation of multiple mtDNA rearrangements in critical tissues, such as skeletal muscle and brain. Most of the adPEO families carry heterozygous mutations in one of three genes: *ANT1*, encoding the muscle-heart-specific mitochondrial adenine nucleotide translocator, *Twinkle*, encoding a mtDNA helicase, and *POLG1*, encoding the catalytic subunit of the mtDNA-specific polymerase gamma. Of these, *POLG1* is particularly relevant for several reasons. First, *POLG1* mutations are the most common cause of mendelian PEO (see: <http://dir-apps.niehs.nih.gov/polg/index.cfm>).

Second, specific mutations or a combination of mutations can be associated with a wide spectrum of clinical presentations, including autosomal dominant (ad) or recessive (ar) PEO, SANDO (sensory ataxic neuropathy, dysarthria, ophthalmo-

plegia), or its variants (e.g. spino-cerebellar ataxia and epilepsy, SCAE), and also infantile Alpers-Huttenlocher syndrome, characterized by liver insufficiency and gray matter disorder (poliodystrophy). While *POLG1*-associated PEO is due to the accumulation of multiple mtDNA deletions, the latter syndromes are rather associated with tissue-specific mtDNA depletion. This clinical heterogeneity can be explained by the complexity of the enzyme, which comprises an exonuclease domain with predominantly proofreading function and a polymerase domain, which is mostly in charge of mtDNA replication. Only one, heterozygous dominant mutation has been identified in *POLG2*, which encodes for the accessory subunit of pol gamma.

Multiple mtDNA deletions and mtDNA depletion coexist in the skeletal muscle of patients with MNGIE (mitochondrial neurogastrointestinal encephalomyopathy), an autosomal recessive disorder of young adults characterized by PEO, peripheral neuropathy, leukoencephalopathy, and severe gastrointestinal dysmotility leading to cachexia and early death. The gene responsible for MNGIE encodes the enzyme thymidine phosphorylase (TP), which promotes the phosphorolysis of thymidine into thymine and deoxyribose-phosphate. Defects of thymidine phosphorylase result in systemic accumulation of thymidine and deoxyuridine, which leads to deoxynucleotide pool imbalance and mtDNA instability.

#### *MtDNA depletion syndromes*

MtDNA depletion syndrome (MDS) is transmitted as an autosomal recessive trait, and is phenotypically heterogeneous. Southern blot analysis is diagnostic, demonstrating the severe reduction of mtDNA in affected tissues (up to 98% in the most severe forms). Some children present with myopathy, others with liver failure in infancy, and some with multisystem involvement. Consistent with the different phenotypes, mtDNA depletion may affect either a specific tissue (most commonly muscle or liver and brain) or multiple organs, including heart, brain, and kidney. MDS has been linked to mutations in eight genes.

Mutations in the gene encoding mitochondrial Thymidine kinase (*TK2*) are responsible of approximately 20% of myopathic MDS cases.

Encephalomyopathic MDS has been associated with mutations in *SUCLA2*, the gene encoding the  $\beta$ -subunit of the adenosine diphosphate-forming succinyl coenzyme A synthetase ligase, while mutations in the  $\beta$ -subunit of GDP-forming succinate-CoA ligase causes a fatal infantile acidosis with combined muscle and liver mtDNA depletion. Recently, mutations in p53-dependent ribonucleotide reductase (*RRM2B*) have been reported in children with mtDNA depletion in muscle and proximal tubulopathy.

As already mentioned, *POLG1* mutations are a common cause of mtDNA depletion especially of the hepatocerebral variants known as Alpers-Huttenlocher syndrome. However, an Alpers-like phenotype has also been associated with recessive mutations in *PEO1*, the gene encoding mitochondrial helicase *Twinkle*. Finally, a fraction of cases with hepato-cerebral MDS has been associated with defects in two additional genes, namely *DGUOK* and *MPV17*. *DGUOK* is the

gene encoding the mitochondrial deoxyguanosine kinase (dGK), while *MPV17* encodes a small mitochondrial membrane protein of unknown function.

#### *mtDNA translation defects*

The autochthonous translation of mtDNA proteins is carried out by an RNA apparatus composed of tRNAs and rRNAs synthesized *in situ* from the corresponding mitochondrial genes and numerous protein factors encoded by nuclear DNA genes, including 77 mitoribosomal proteins, several tRNA maturation enzymes (e.g. pseudouridylate synthases), the aminoacyl-tRNA synthetases, the translation initiation, elongation and termination factors, etc. Mutations in several of these genes have recently been found, in most cases in single patients or families. An exception is *DARS2*, the gene encoding mitochondrial aspartyl-tRNA synthetase, mutations of which have been identified in 38 patients belonging to 30 families. In contrast with other translation defects, *DARS2* mutations are not associated with clearcut defects of the respiratory chain activities, and the clinical manifestations are those of progressive motor impairment caused by leukodystrophic changes in the brainstem and upper spinal cord.

### **Defects of the mitochondrial lipid milieu**

Barth syndrome (BTHS) is an X-linked recessive disorder characterized by mitochondrial myopathy, cardiopathy, growth retardation, and leucopenia. Methylglutaconic aciduria is documented in BTHS patients.

The tafazzin (*TAZ*) gene responsible for this disorder encodes a family of proteins (“tafazzins”) that are homologous to phospholipid acyltransferases involved in the biosynthesis of cardiolipin, a phospholipid species present in the inner membrane of mitochondria.

Cardiolipin plays a role in the structural stabilization and functional modulation of the respiratory chain complexes.

### **CoQ defects**

Coenzyme Q<sub>10</sub> (CoQ<sub>10</sub>), a lipid-soluble component of virtually all cell membranes, transports electrons from complexes I and II to complex III in the mitochondrial respiratory chain and is essential for the stability of complex III.

Primary CoQ<sub>10</sub> deficiency causes clinically heterogeneous diseases:

- 1) encephalomyopathy characterized by the triad of recurrent myoglobinuria, brain involvement and ragged-red fibers;
- 2) severe infantile multi-systemic disease;
- 3) cerebellar ataxia;
- 4) Leigh syndrome with growth retardation, ataxia and deafness.

These disorders are transmitted as autosomal recessive traits and in most cases respond to CoQ<sub>10</sub> supplementation. The molecular defect has been identified in only a few infantile forms due to mutations in *COQ2*, *PDSS1*, *PDSS2*, three enzymes of the CoQ biosynthetic pathway.

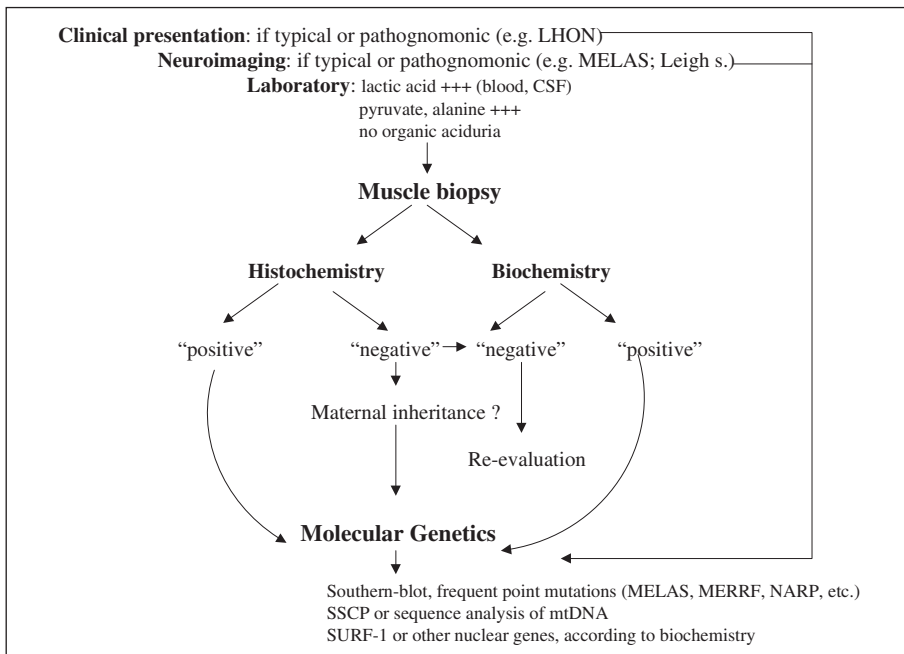
## Defects in mitochondrial biogenesis

The mitochondria are dynamic organelles that constantly fuse and divide. An equilibrium between fusion and fission controls the morphology of the mitochondria, which appear as dots or elongated tubules depending on the prevailing process.

Mitochondrial fusion requires the concerted action of three GTPases, mitofusin 1 and 2 and OPA1. For mitochondrial fission, another GTPase, called dynamin-like protein 1 (DLP1) is required. Mutations in genes required for mitochondrial fusion result in the fragmentation of mitochondria, due to ongoing fission in the absence of fusion.

The disruption of proteins required for mitochondrial fission results in tubulized, “spaghetti-like” mitochondria that are excessively long and interconnected. Heterozygous mutations in *MFN2* cause Charcot-Marie-Tooth type 2A, while mutations in *OPA1* cause dominant optic atrophy, the most common inherited optic neuropathy.

In some cases, optic atrophy may be associated with a more complex syndrome including myopathy, PEO, hearing loss and ataxia, and caused by the accumulation of multiple mtDNA deletions in skeletal muscle. Very recently, a mutation affecting *DLP1* has been described in a severe infantile encephalopathy. Since *DLP1* is also present in peroxisomes, both elongated mitochondria and peroxisomes were found in this patient.



**Fig. 2** - Diagnostic algorithm in mitochondrial disorders.

## **Diagnostic considerations**

The diagnosis of mitochondrial disorders relies upon clinical, morphological, biochemical and molecular genetic studies (Fig. 2). A careful clinical examination is necessary to correctly address the subsequent diagnostic protocol.

Neuroimaging investigations have identified several patterns of involvement of the CNS. The observation of abnormalities within the brainstem, basal ganglia, white matter, or the presence of multiple cortico-subcortical areas of abnormal signal intensity may suggest the correct diagnosis, particularly in patients with atypical presentations. MRI spectroscopy is a non-invasive technique that can detect increased lactate levels in the brain, and low ATP production in both brain and muscle. Useful laboratory findings, especially in infants and children, include the presence of elevated levels of lactic acid and pyruvic acid in blood and CSF, and the absence of organic acidurias. The absence of RRF, or other morphological signs of mitochondrial pathology, does not necessarily exclude the diagnosis. The availability of muscle tissue is mandatory in patients with phenotypes associated with large scale rearrangements of mtDNA, that can be easily missed in leukocytes. By contrast, point mutations of mtDNA can be detected also in blood samples. However, in maternally inherited phenotypes associated with heteroplasmic point mutations the percentage of mutant mtDNA is often significantly higher in muscle than in other tissues.



# **BIOMARCATORI IN ONCOLOGIA**





## **Profilo di espressione genica del tessuto polmonare normale e rischio di cancro polmonare**

*Tommaso A. Dragani*

Dipartimento di Oncologia Sperimentale e Laboratori, Fondazione IRCCS Istituto Nazionale Tumori, Milano

Facendo seguito ad una nostra osservazione in un modello sperimentale, abbiamo testato l'ipotesi che il profilo di espressione genica del tessuto polmonare normale potesse essere associato al rischio di sviluppare un tumore polmonare. Allo scopo, abbiamo confrontato pool di RNA preparati da tessuto polmonare normale di pazienti con adenocarcinoma polmonare (casi) o di pazienti con metastasi polmonari di tumori non polmonari (controlli). È stata condotta l'ibridizzazione degli RNA su microarrays contenenti tutti i geni noti di uomo. Il profilo di espressione genica di 85 geni è risultato in grado di distinguere i casi dai controlli, e i fumatori dai non fumatori. Elevati livelli di mRNA di uno di tali geni, AZGP1, sono risultati associati significativamente con lo status di malattia. I nostri risultati sostengono l'ipotesi che differenze nei livelli di espressione genica nel tessuto polmonare normale possano essere predittive della predisposizione genetica al cancro polmonare nell'uomo.

### **Introduzione**

In ceppi di animali di laboratorio sono state descritte associazioni tra il profilo di espressione genica del tessuto normale e la suscettibilità genetica allo sviluppo tumorale nello stesso organo; tali osservazioni hanno riguardato sia il polmone (Gariboldi et al., 2003) che altri organi (Riggs et al., 2005; Fujiwara et al., 2003; Falvella et al., 2002). Nell'uomo, fattori di predisposizione genetica al cancro polmonare sono stati implicati sia nei fumatori (Sellers et al., 1994) che nei non-fumatori (Wu et al., 2004).

Abbiamo voluto studiare se il profilo di espressione genica del tessuto polmonare normale poteva distinguere i pazienti con adenocarcinoma (ADCA) polmonare dai pazienti di controllo, che non avevano sviluppato un tumore polmonare. Scopo ultimo di questo tipo di studi è l'identificazione di un profilo di espressione genica che possa permettere, attraverso l'uso di metodi non invasivi, quali ad esempio l'analisi delle proteine dello sputo, di fornire una stima del rischio genetico individuale di sviluppare un tumore polmonare.

## Metodi e Risultati

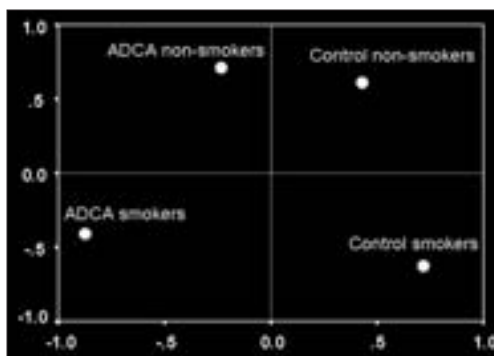
Un piccolo frammento di tessuto polmonare normale è stato prelevato, a distanza dal tumore, da lobi polmonari rimossi chirurgicamente da pazienti con ADCA polmonare (casi) o da pazienti con metastasi polmonari di tumori non polmonari (controlli). È stata condotta l'analisi del profilo di espressione genica su pool di RNA preparati dai casi o dai controlli, che sono stati ibridizzati su piattaforma Affymetrix (Human U133 2.0 plus GeneChip). L'analisi quantitativa dell'espressione genica è stata condotta con il metodo SYBRGreen su apparecchio ABI GeneAmp 7900 (Applied Biosystems). L'analisi immunostochimica della proteina AZGP1 (zinc-alpha2-glycoprotein, ZAG) è stata condotta su tissue arrays commerciali, consistenti di 59 casi di tumore polmonare o del loro corrispondente tessuto normale.

L'analisi di ibridizzazione su microarray ha permesso di individuare 85 trascritti indipendenti, caratterizzati da elevato coefficiente di variazione tra i vari pool di RNA e da differenze di almeno 4 volte tra due o più campioni. L'analisi statistica delle componenti principali ha dimostrato che era possibile distinguere i tessuti polmonari provenienti dai casi (pazienti con ADCA) da quelli provenienti dai controlli. Inoltre, era possibile anche distinguere i pazienti fumatori dai non fumatori (Fig. 1).

I risultati dell'analisi con i microarray sono stati validati con la RT-PCR quantitativa, che ha dimostrato una correlazione statisticamente significativa con i dati ottenuti dai microarray, indicando che quest'ultima analisi ha individuato variazioni reali nei livelli dei trascritti.

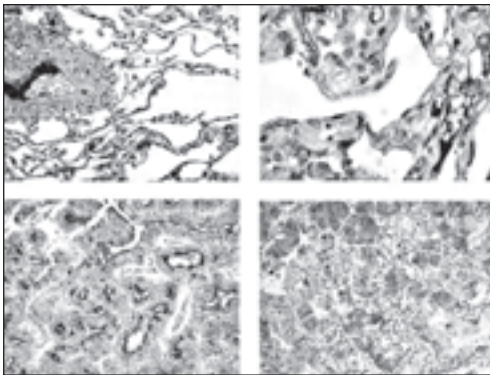
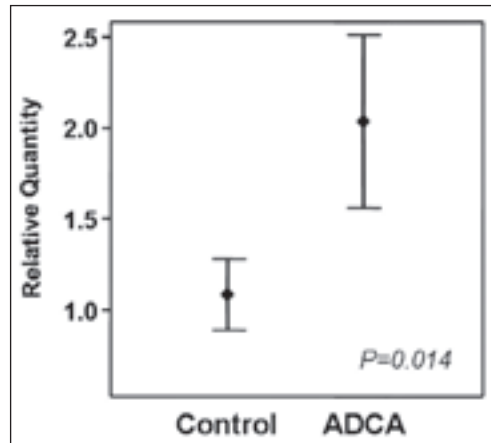
Nel gruppo degli 85 trascritti, è stata condotta un'analisi statistica preliminare ed uno studio di validazione in real-time RT-PCR, che hanno portato a focalizzare la nostra attenzione sul gene AZGP1. I livelli di tale gene sono stati analizzati in tutti i campioni disponibili, rappresentati da un totale di 108 campioni tra casi e controlli. I livelli di espressione di AZGP1 sono risultati altamente variabili, con valori di RQ da 0.02 a 24.1 (media  $\pm$  S.E. =  $1.6 \pm 0.3$ ).

Non è stata osservata alcuna associazione statistica significativa tra i livelli di mRNA di AZGP1 e il sesso, lo status di fumatore, l'età alla diagnosi della malattia, o lo stadio clinico. Invece, un'associazione statisticamente significativa è stata osservata tra lo status di malattia, con i pazienti con ADCA che hanno mostrato



**Fig. 1** - Analisi delle componenti principali della matrice di correlazione dei livelli di espressione genica nel tessuto polmonare normale. Sono stati analizzati i livelli di 85 geni che hanno mostrato i più elevati coefficienti di variazione tra i campioni analizzati. Il primo asse fattoriale, che spiega il 38% della varianza totale, contrasta i casi (pazienti con ADCA polmonare; a sinistra) dai controlli (pazienti con metastasi polmonare di tumore non-polmonare; a destra). Il secondo asse fattoriale, che spiega il 36% della varianza, contrasta i soggetti fumatori (in basso) da quelli non fumatori (in alto).

**Fig. 2** - Espressione di AZGP1 nel tessuto polmonare normale analizzata mediante RT-PCR. I livelli medi dei trascritti di AZGP1 sono risultati circa 2 volte più alti nei pazienti con ADCA polmonare (n=55) che nei pazienti con tumore non-polmonare (n=53) ( $P=0.014$ , Kruskal-Wallis's test).



**Fig. 3** - Immunohistochemica della proteina ZAG protein. Nel tessuto polmonare normale (in alto), la maggior parte degli alveoli non è positiva (sinistra), sebbene alcune cellule alveolari mostrino sporadicamente una debole colorazione citoplasmatica (destra). Nei tessuti tumorali (in basso), una debole colorazione granulare, concentrata lungo il bordo interno delle ghiandole tumorali è stata osservata in un campione di ADCA polmonare (sinistra); un carcinoma bronchiolo-alveolare ha mostrato una debole ed uniforme colorazione citoplasmatica (destra).

livelli medi di mRNA di AZGP1 circa 2 volte più elevati rispetto a quelli dei pazienti non-ADCA (controlli) ( $P=0.01$ ) (Fig. 2).

L'analisi immunohistochemica ha rivelato assenza di proteina AZGP1 nel tessuto polmonare normale di tutti i casi analizzati, tranne che per una debole colorazione citoplasmatica di poche cellule alveolari (Fig. 3, in alto). Anche la maggior parte dei campioni di cancro polmonare è risultata negativa per la proteina AZGP1, mentre l'espressione di tale proteina è stata chiaramente individuata, a livelli elevati, in solo 4 casi, costituiti da un carcinoma a cellule squamose, un ADCA (Fig. 3, in basso a sinistra), un carcinoma bronchiolo-alveolare (Fig. 3, in basso a destra), ed un carcinoma mucoepidermoide. Un'altro carcinoma spinocellulare ed un altro ADCA hanno mostrato una debole colorazione positiva per AZGP1 in alcune cellule.

## Discussione

In questo studio abbiamo confrontato il tessuto polmonare normale di pazienti con ADCA polmonare con lo stesso tipo di tessuto di pazienti con metastasi pol-

monari di tumori non-polmonari. Il profilo di espressione genica di 85 trascritti è risultato in grado di distinguere i pazienti con ADCA polmonare dai pazienti con tumore non-polmonare (Fig. 1).

L'analisi quantitativa in real-time RT-PCR non ha però individuato differenze statisticamente significative, tra casi e controlli, per la maggior parte dei trascritti, eccetto per quelli del gene AZGP1. Tale risultato può essere spiegato dal fatto che, per i singoli trascritti, le differenze di espressione sono relativamente modeste, tuttavia, il profilo di espressione globale di tutti e 85 i trascritti, potrebbe essere in grado di fornire una stima del rischio individuale di cancro polmonare.

Il trascritto del gene AZGP1, che codifica per una proteina solubile di 41-kDa di peso molecolare (ZAG), ha mostrato livelli più elevati nel tessuto polmonare normale dei pazienti con ADCA rispetto ai pazienti con tumore non-polmonare (Fig. 2). L'espressione della proteina codificata dal gene AZGP1 era stata riportata da altri lavori in associazione con la prognosi e la sopravvivenza di pazienti con cancro della prostata (Lapointe et al., 2004; Henshall et al., 2006).

La nostra analisi preliminare ha confermato che l'abitudine al fumo rappresenta una fonte notevole di variabilità per profilo di espressione genica del tessuto polmonare normale. Tra i geni identificati dall'analisi di microarray e che abbiamo confermato con la RT-PCR quantitativa, possiamo segnalare HAS1 (hyaluronan synthase), CYP1A1 (cytochrome P450, family 1, subfamily A, polypeptide 1), e AHRR (aryl-hydrocarbon receptor repressor).

In conclusione, sebbene i nostri risultati sono da intendersi come risultati preliminari che necessitano di conferma in studi di più grandi dimensioni, essi suggeriscono che uno specifico profilo di espressione genica del tessuto polmonare normale può distinguere i pazienti con ADCA polmonare (casi) dai pazienti con tumore non-polmonare (controlli). Queste osservazioni sono quindi a supporto dell'ipotesi che la predisposizione genetica abbia un ruolo nel rischio individuale di cancro polmonare e che essa sia accompagnata da alterazioni nel profilo di espressione genica che sono identificabili già nel tessuto polmonare normale. Se tali risultati saranno confermati, essi potranno quindi fornire nuove opportunità per migliorare la prevenzione e la diagnosi precoce di tale tipo di tumore.

### **Bibliografia essenziale**

1. Falvella FS., Pascale RM., Gariboldi M., Manenti G., De Miglio MR., Simile MM., et al. (2002). Stearoyl-CoA desaturase 1 (Scd1) gene overexpression is associated with genetic predisposition to hepatocarcinogenesis in mice and rats. *Carcinogenesis*, 23: 1933-1936.
2. Fujiwara K., Ochiai M., Ubagai T., Ohki M., Ohta T., Nagao M., et al. (2003). Differential gene expression profiles in colon epithelium of two rat strains with distinct susceptibility to colon carcinogenesis after exposure to PhIP in combination with dietary high fat. *Cancer Sci.*, 94: 672-678.
3. Gariboldi M., Spinola M., Milani S., Pignatiello C., Kadota K., Bono H., et al. (2003). Gene expression profile of normal lungs predicts genetic predisposition to lung cancer in mice. *Carcinogenesis*, 24: 1819-1826.

4. Henshall SM., Horvath LG., Quinn DI., Eggleton SA., Grygiel JJ., Stricker PD., et al. (2006). Zinc-alpha2-glycoprotein expression as a predictor of metastatic prostate cancer following radical prostatectomy. *J. Natl. Cancer Inst.*, 98: 1420-1424.
5. Lapointe J., Li C., Higgins JP., Van de RM., Bair E, Montgomery K., et al. (2004). Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. Natl. Acad. Sci. USA*, 101: 811-816.
6. Riggs PK., Angel JM., Abel EL. and DiGiovanni J. (2005). Differential gene expression in epidermis of mice sensitive and resistant to phorbol ester skin tumor promotion. *Mol. Carcinogen.*, 44: 122-136.
7. Sellers TA., Chen PL., Potter JD. Bailey-Wilson JE., Rothschild H. and Elston RC. (1994). Segregation analysis of smoking-associated malignancies: evidence for Mendelian inheritance. *Am. J. Med. Genet.*, 52: 308-314.
8. Wu PF., Lee CH., Wang MJ., Goggins WB., Chiang TA., Huang MS., et al. (2004). Cancer aggregation and complex segregation analysis of families with female non-smoking lung cancer probands in Taiwan. *Eur. J. Cancer*, 40: 260-266.

## **Approccio proteomico per la scoperta di biomarcatori di tumori: un futuro per la diagnosi precoce e la terapia**

*Francesca Lavatelli, Giampaolo Merlini*

Centro per lo Studio delle Amiloidosi, Laboratorio di Biotecnologie,  
Fondazione IRCCS Policlinico San Matteo, Università di Pavia

L'applicazione della proteomica allo studio della patologia umana è un campo in rapida evoluzione e la cui traslazione da tecnica di ricerca a strumento a diretta disposizione del medico sembra sempre più a portata di mano. Scopo principale della proteomica clinica è fornire nuovi mezzi per una diagnosi precoce e per una identificazione del profilo proteico di malattia, che apra la strada ad una terapia modellata sull'individuo.

L'analisi di cambiamenti nel proteoma - inteso come il contenuto proteico globale - di tessuti o fluidi biologici, permette infatti di identificare proteine associate alla presenza di una condizione morbosa, che possano fungere da marcatori diagnostici in fase precoce, indicatori di prognosi o predittori di risposta ad un trattamento. La principale differenza dell'approccio proteomico rispetto a quello tradizionale, basato sullo studio di marcatori singoli, consiste nella possibilità di individuare anomalie nell'espressione proteica anche senza ipotesi a priori sulle possibili alterazioni in causa, e nel rilievo di un profilo di multiple proteine la cui espressione risulta globalmente modificata in relazione alla presenza di malattia. La proteomica clinica ha trovato applicazione soprattutto nello studio di malattie croniche ed in campo oncologico. In particolare, l'individuazione di biomarcatori di tumore può avere un impatto clinico enorme, nel selezionare i soggetti a rischio di sviluppo di patologia neoplastica, nella diagnosi precoce e nel monitoraggio in tempo reale della risposta alla terapia.

### **Tecniche analitiche in proteomica**

L'analisi proteomica combina tecniche separative ad alta risoluzione con metodi basati sulla spettrometria di massa, per l'identificazione, la caratterizzazione e la quantificazione di proteine in miscele proteiche complesse.

Le caratteristiche analitiche della spettrometria di massa, in particolare elevata sensibilità, alta risoluzione e rapidità, rendono possibile lo studio e l'identificazione sistematici e riproducibili di migliaia di proteine in un singolo campione. L'avvento relativamente recente di tecniche che consentono la ionizzazione e

l'analisi di macromolecole biologiche, quali la ionizzazione tramite MALDI (Matrix-Assisted Laser Desorption/Ionization) ed ESI (ElectroSpray Ionization) ha permesso, di fatto, lo sviluppo della moderna proteomica.

A monte dell'analisi spettrometrica, tuttavia, la complessità e l'elevato *range* dinamico delle proteine in tessuti o fluidi biologici rendono necessario l'uso di tecniche separative per pre-frazionarne i componenti ed arricchire la miscela delle proteine espresse a bassa concentrazione. Elettroforesi mono- e bidimensionale (2D-PAGE) e separazioni mediante cromatografia multidimensionale (2D Protein LC/LC, MudPIT) sono i metodi più comunemente usati a tale scopo. La tecnica SELDI (Surface-Enhanced Laser Desorption/Ionization) si basa invece sull'immobilizzazione delle proteine su chip la cui superficie è stata modificata chimicamente, consentendo il legame selettivo e l'arricchimento di particolari classi di proteine, che possono essere direttamente analizzate dallo spettrometro, solitamente con analizzatore TOF (Time of Flight). Il profilo di picchi presenti nello spettro di massa, la cui intensità differisce tra pazienti e controlli, viene quindi rilevato mediante algoritmi matematici ed è esso stesso usato come marcatore, anche senza la necessità di identificare la proteina che costituisce ogni picco. La minima processazione del campione richiesta da questo approccio ne ha fatto uno strumento di particolare interesse nel campo della proteomica clinica.

L'uso di sofisticate risorse bioinformatiche è un altro cardine su cui si basa ogni analisi proteomica ed interviene ad ogni livello di analisi. La creazione di banche dati proteiche accessibili in rete e lo sviluppo di algoritmi per interrogarle è il presupposto per ogni identificazione di proteine basata sulla spettrometria di massa; la comparazione di mappe proteomiche ottenute tramite elettroforesi o cromatografia multidimensionale non sarebbe possibile senza l'utilizzo di programmi dedicati, e sofisticati algoritmi matematici sono alla base della possibilità di analizzare ed individuare profili di espressione proteica associati alla presenza di malattia.

### **Ricerca proteomica di biomarcatori tumorali in fluidi e tessuti biologici**

L'assunto alla base della ricerca di biomarcatori in oncoproteomica è che la presenza del tumore si associ ad una alterazione dell'espressione proteica tissutale ed, in ultima analisi, si traduca in modificazioni rilevabili nel sangue.

Ad oggi, la maggior parte dei marcatori tumorali utilizzati in clinica soffrono di limitate sensibilità e specificità, che ne riducono le potenzialità diagnostiche, specialmente in fase precoce. Una delle sfide fondamentali dell'oncoproteomica è il disegno di un metodo che permetta di rilevare lesioni cancerose o pre-cancerose ad uno stadio localizzato e pre-invasivo, mediante analisi di fluidi biologici. Nell'ipotesi che la presenza del tumore e la sua interazione con il microambiente producano una cascata di cambiamenti nelle proteine espresse, è stato ipotizzato che l'uso di un pannello di marcatori multipli (che, utilizzando la tecnica SELDI, si traduce nel rilievo di un sottogruppo di picchi discriminatori nello spettro di massa) possa avere sensibilità e specificità maggiori che test basati su marcatori singoli.



La ricerca di marcatori tumorali nel siero e nel plasma è resa problematica dall'elevato *range* dinamico delle proteine che vi si trovano. Gli strumenti analitici hanno limiti di sensibilità finiti, che precludono il rilievo di proteine presenti in quantità troppo bassa in una miscela complessa. Nel plasma, inoltre, una ventina di proteine costituiscono da sole il 99% del totale (1), mascherando il cosiddetto "proteoma a basso peso molecolare" (2), nel quale sono compresi, spesso in concentrazione pico o femtomolare, citochine, chemochine, ormoni e frammenti proteolitici di proteine più grosse, e di cui fanno parte la maggior parte dei marcatori tumorali correntemente utilizzati in clinica. Per arricchire il campione delle proteine espresse in minore quantità, l'analisi di questi fluidi biologici è solitamente preceduta dalla deplezione di albumina e delle principali proteine plasmatiche, mediante tecniche cromatografiche o per immunoaffinità.

L'analisi proteomica di plasma e siero è stata applicata allo studio di una ampia gamma di tumori, tra cui neoplasie di ovaio, prostata, mammella, apparato genito-urinario, gastrointestinale, fegato, polmone, e sistema nervoso (3-8). La maggior parte dei lavori sono stati condotti mediante tecnica SELDI-TOF. Per ognuno di questi tipi tumorali sono stati identificati potenziali pattern diagnostici;<sup>3-8</sup> molti di essi hanno dimostrato preliminarmente elevate sensibilità e specificità, superiori a quelle di marcatori tradizionali (Tab. 1) (9).

Nel complesso, enzimi, frammenti proteolitici inattivi di enzimi o inibitori delle proteasi fanno spesso parte dei pannelli discriminatori rilevati. In uno studio,<sup>10</sup> condotto da due centri indipendenti su pazienti con tumore dell'ovaio, è stato rilevato un pannello di tre marcatori la cui concentrazione variava in relazione alla presenza della malattia, costituiti da apolipoproteina A-I, un frammento di trans-tiretina ed un frammento proteolitico della catena H4 dell'inibitore dell'inter-tripsina. La sensibilità di un modello multivariato basato sull'uso combinato dei tre marcatori e del CA-125 è risultata nettamente superiore a quella del CA-125 da solo nello screening del tumore ovarico infiltrante in stadio precoce (Tab. 1) (10).

In un secondo studio, mirato alla ricerca di un profilo sierico di marcatori di carcinoma prostatico(7), previo arricchimento cromatografico del proteoma a basso

**Tab. 1** - Comparazione di pattern diagnostici proteomici e attuali marcatori tumorali.

Tumore	Marcatori proteomici		Attuali marcatori tumorali		
	Sensibilità (%)	Specificità (%)	Marcatore	Sensibilità (%)	Specificità (%)
Vescica (11)	80	90-97	NMP22	31	95
Mammella (12)	93	91	CA 15-3	63	80-88
Colon-retto (13)	91	93	CEA	43	-
Stomaco (14)	83	95	CEA	49	-
Fegato (15)	94	86	AFP	50	90
Polmone (16)	87	80	Cyfra21-1	63	94
Ovaio (10)	83	94	CA-125	57	-
Pancreas (17)	78	97	CA 19-9	72	-
Prostata (18)	83	97	PSA	86	20-34

(Modificata da Cho ECS, Mol Cancer, 2007).

peso molecolare su resina lipofilica, un set di 36 picchi discriminatori negli spettri di massa ha mostrato una sensibilità diagnostica del 94.1% ed una specificità del 99%, quando testato in cieco su un ampio gruppo di pazienti ed individui non affetti.

Nessuno di questi test, è, tuttavia, ancora pronto per un uso clinico diretto. Per la maggior parte di essi, è infatti ancora necessaria una fase di validazione su coorti di pazienti ampie ed indipendenti da quelle utilizzate nella fase di ricerca, con analisi quantitativa delle alterazioni osservate.

Anche le urine, che presentano una complessità minore rispetto al plasma, sono un materiale promettente per la ricerca di marcatori in vari tipi di tumore, soprattutto quelli dell'apparato genito-urinario (5). Particolarmente interessante è lo studio dei cosiddetti esosomi, strutture vescicolari a basso peso molecolare ricche di proteine di membrana e citosoliche.

L'approccio proteomico è correntemente utilizzato anche nella analisi diretta del tessuto tumorale, alla ricerca di proteine che consentano di tracciarne una "impronta digitale", migliorando l'accuratezza diagnostica e permettendo di ottenere una classificazione molecolare delle neoplasie. Diverse metodiche possono essere usate, che spaziano dallo studio dell'intero tessuto, allo specifico arricchimento di cellule coinvolte nel processo morboso (mediante, ad esempio, *laser capture microdissection*), al frazionamento di strutture subcellulari (1). Oltre all'isolamento del proteoma complessivo, è inoltre possibile l'arricchimento selettivo dei cosiddetti sub-proteomi, quali fosfo- e glicoproteoma, cioè sottoinsiemi di proteine accomunate da particolari caratteristiche chimico-fisiche, che si traducono, da un punto di vista patogenetico, in caratteri funzionali (19). Il rilievo di tirosin-fosfoproteomi tumore-specifici (19) è stato proposto come un nuovo potenziale metodo per la classificazione di neoplasie umane.

Un approccio di sviluppo recente per lo studio del tessuto è il cosiddetto "MALDI Imaging" (20), tecnica che consente di generare una lista di masse di proteine e peptidi da sezioni tissutali direttamente montate sulla piastra target dello spettrometro. Spettri di massa vengono acquisiti a livello di tutta la sezione di campione analizzata, in punti molto vicini tra loro. A seconda della distribuzione di particolari ioni scelti come marcatori di stati patologici (per esempio, di trasformazione maligna di una cellula), possono essere create delle immagini virtuali del tessuto in esame, che ne rappresentino le caratteristiche organiche. Questo approccio è in studio per consentire una miglior definizione dei margini tumorali ed è stato utilizzato con successo nella classificazione molecolare di neoplasie polmonari e cerebrali (21).

### **Problemi aperti e prospettive future**

Mentre a livello sperimentale l'analisi proteomica è uno strumento consolidato e potente, la traslazione dal laboratorio alla routine clinica sembra ancora essere gravata da alcune limitazioni. Dopo un iniziale entusiasmo, è emersa l'evidenza che, anche a causa della mole e della complessità delle informazioni generate, fattori confondenti possono intervenire in ciascuno dei passaggi che portano dalla

raccolta del campione all'interpretazione dei dati, e possono così influenzare la riproducibilità e significatività biologica dei risultati (22).

A questo va aggiunta la scarsità, ancora oggi, di informazioni conclusive su quale debba essere considerata l'esatta natura del proteoma normale di fluidi quali plasma ed urine, che rende difficile definire accuratamente i limiti della variabilità individuale.

Poco dopo la pubblicazione dei primi studi sul tumore dell'ovaio (4), condotti mediante analisi SELDI-TOF e che mostravano altissime sensibilità e specificità dei pattern discriminatori rilevati, la re-analisi dei dati da parte di gruppi indipendenti ha messo in dubbio la validità dei risultati raggiunti, dimostrando come fattori esterni dipendenti dalla raccolta e processazione dei campioni, più che parametri legati alla presenza di malattia, avessero dato origine a pattern proteici "pseudodiagnostici", in grado di distinguere fra pazienti e controlli (22-24). Il clamore sollevato da questo episodio ha gettato dubbi sulla applicabilità clinica della proteomica in tempi brevi ed ha messo in evidenza i punti ancora deboli del sistema. In seguito, sono stati sollevati altri dubbi sulla rilevanza biologica di alcuni studi pubblicati (22-24). È infatti risultato evidente come molto spesso proteine di fase acuta, marcatori non specifici della risposta dell'ospite ad una patologia, fossero tra i principali componenti dei pannelli di marcatori rilevati (6, 22). La maggior parte dei lavori sono stati, inoltre, condotti su tumori in fase clinica avanzata, nell'ipotesi, probabilmente semplificativa e non dimostrata, che le caratteristiche biochimiche di neoplasie avanzate e precoci siano simili, e che quindi i risultati potessero essere estesi allo screening precoce.

Sforzi collettivi della comunità scientifica sono oggi in corso per la definizione di criteri da adottare nel campo della proteomica clinica. La necessità di ottenere dati riproducibili e biologicamente plausibili sta portando alla definizione di criteri di standardizzazione degli esperimenti, soprattutto per rilevare ed eliminare quanto più possibile l'insorgenza di bias nella fase pre-analitica, analitica e bioinformatica e per definire criteri di validazione dei dati raccolti, che possano rendere concreto l'utilizzo della proteomica clinica, che rimane una metodica di enorme potenzialità, nel laboratorio di chimica clinica.

## **Bibliografia**

1. Zolg JW, Langen H. How industry is approaching the search for new diagnostic markers and biomarkers. *Mol Cell proteomics* 2004; 3: 354-354.
2. Tirumalai RS, Chan KC, Prieto DA, et al. Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics* 2003; 2: 1096-1103.
3. Simpkins F, Czecheowicz JA, Liotta L, et al. SELDI-TOF mass spectrometry for cancer biomarker discovery and serum proteomic diagnostics. *Pharmacogenomics* 2005; 6: 647-53.
4. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002; 359: 572-77.
5. Hu S., Loo J. A., Wong D.T. Human body fluid proteome analysis. *Proteomics* 2006; 6: 6326-53.

6. Petricoin E, Wulfkuhle J, Espina V, et al. Clinical proteomics: revolutionizing disease detection and patient tailoring therapy. *J Proteome Res* 2004; 3: 209-17.
7. Mobley JA, Lam YW, Lau KM, et al. Monitoring the serological proteome: the latest modality in prostate cancer detection. *J Urol* 2004; 172: 331-337.
8. Bonney GK, Crafen RA, Prasad R, et al. Circulating markers of biliary malignancy: opportunities in proteomics? *Lancet Oncol* 2008; 9: 149-158.
9. Cho GCS. Contribution of oncoproteomics to cancer biomarker discovery. *Mol Cancer* 2007; 6: 25-37.
10. Zhang Z, Bast RC, Yu Y, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 2004; 64: 5882-5890.
11. Mueller J, VonEggingel F, Driesch D, et al. ProteinChip technology reveals distinctive protein expression profiles in the urine of bladder cancer patients. *Eur Urol* 2005; 47: 885-894.
12. Li J, Zhang Z, rosenzweig J, et al. Proteomic and bioinformatic approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem* 2002; 48: 1296-1304.
13. Chen YD, Zheng S, Yu JK, et al. Artificial neural network analysis of SELDI mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population. *Clin Cancer Res* 2004; 10: 8380-8385.
14. Poon TC, Sung JJ, Chow SM, et al. Diagnosis of gastric cancer by serum proteomic fingerprinting. *Gastroenterology* 2006; 130: 1858-1864.
15. Ward DJ, Cheng Y, N'Kontchou G, et al. Changes in the serum proteome associated with the development of hepatocellular carcinoma in hepatitis C-related cirrhosis. *Br J Cancer* 2006; 94: 287-292.
16. Yang SY, Xiao XY, Zhang WG, et al. Application of serum SELDI proteomic patterns in diagnosis of lung cancer. *BMC Cancer* 2005; 5: 83.
17. Koopman J, Zhang Z, White N, et al. Serum diagnosis of pancreatic adenocarcinoma using SELDI mass spectrometry. *Clin Cancer Res* 2004; 10: 860-868.
18. Adam BL, Qu Y, Davis JW, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res* 2002, 62: 3609-3614.
19. Lim YP. Mining the tumor phosphoproteome for cancer markers. *Clin Cancer Res* 2005; 11: 3163-3169.
20. Caldwell LR, Caprioli RM. Tissue profiling by mass spectrometry. *Mol Cell Proteomics* 2005; 4: 394-401.
21. Yanagisawa K, Shyr Y, Xu BJ, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 2003; 362: 433-439.
22. Beretta L. Proteomics from the clinical perspective: many hopes and much debate. *Nature Methods* 2007; 4: 785-6.
23. Check E. Proteomics and cancer: running before we can walk? *Nature*, 2004; 429: 496-7.
24. Diamandis E. Serum proteomic profiling by Matrix-Assisted Laser Desorption-Ionization Time of Flight mass spectrometry for cancer diagnosis: next steps. *Cancer Res* 2006; 66: 5540-41.

## Marcatore genetici e risposta alla terapia di neoplasie ematologiche

*Mario Petrini*

Dipartimento di Oncologia, dei Trapianti e delle Nuove Tecnologie in Medicina, Università di Pisa

La terapia delle neoplasie maligne si basa ampiamente sui modelli elaborati a partire dalle curve Gompertz che forniscono un razionale matematico agli schemi di chemioterapia ed al numero di cicli da utilizzare. L'assunto che la crescita della massa è più lenta nelle fasi iniziali e avanzate che non in un momento intermedio permette di calcolare, sulla base della cinetica, l'ipotetico azzeramento di una malattia chemiosensibile. Poiché le neoplasie solide hanno caratteristiche intrinseche di resistenza legate alla massa e alla vascolarizzazione e, in definitiva, alla permeabilità ai farmaci, le malattie maligne ematologiche apparivano come le più adatte ad essere affrontate secondo questi modelli.

Più recentemente è apparso evidente che le cellule staminali neoplastiche possono avere intrinsechi meccanismi di resistenza e, in ultima analisi, non eradicabili. L'esempio della leucemia mieloide cronica è illuminante, da questo punto di vista, e l'enorme difficoltà, se non l'impossibilità, ad eradicare questa malattia con glivec è documentata anche da modelli teorici.

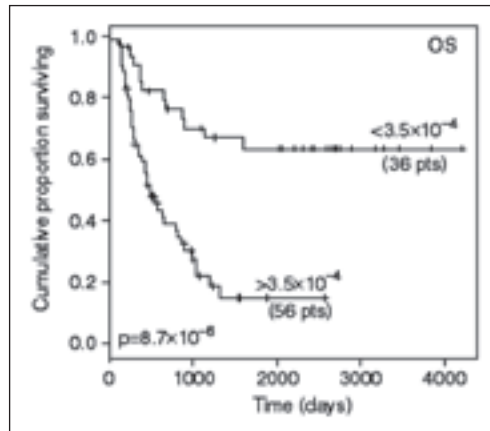
La necessità di utilizzare tecniche che possano rilevare la malattia con grande specificità ed una sensibilità superiore a quella dei metodi morfologici/biochimici è, oggi ampiamente riconosciuta. Lo studio molecolare permette, frequentemente, di individuare mutazioni specifiche per una determinata malattia e, grazie all'elevata sensibilità, di documentare la presenza/persistenza di cellule neoplastiche.

In termini generali un marcatore genetico può rilevare la lesione patogenetica responsabile di una determinata patologia e, quindi, contribuire a chiarire la patogenesi ma, in questa sede, interessa principalmente la possibilità di utilizzazione come indicatore di presenza di malattia.

Un secondo tipo di marcatori è quello che, pur non essendo direttamente implicato nella patogenesi, permette, tuttavia, di identificare la presenza di cellule neoplastiche. Infine, occorre considerare la possibilità di identificare l'espressione di geni associati alla prognosi e che possono condizionare le scelte terapeutiche.

Tra i geni del primo gruppo *pml/rar $\alpha$*  e *bcr/abl* sono paradigmatici.

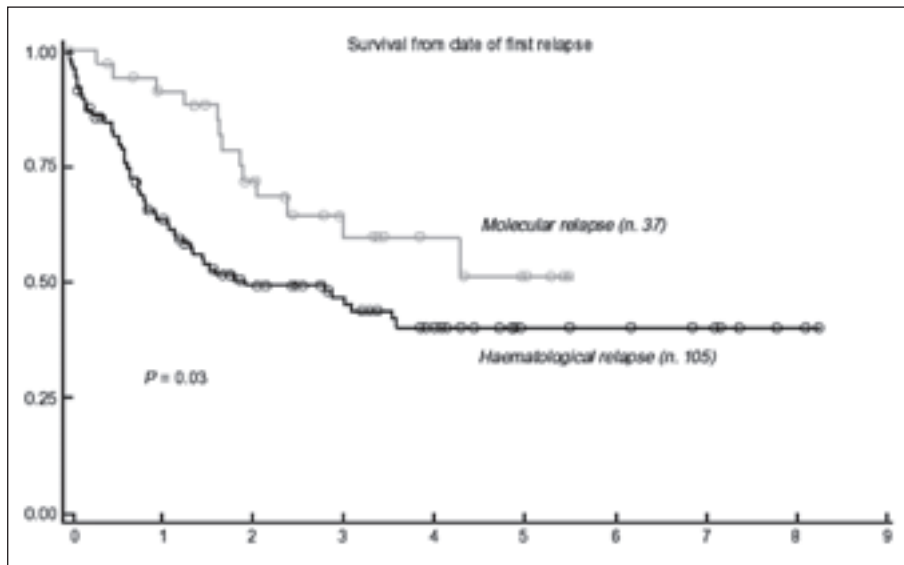
*pml/rar $\alpha$*  deriva dalla traslocazione *t(15;17)* che inducendo la formazione di un gene ibrido determina la produzione di una proteina modificata che non permet-



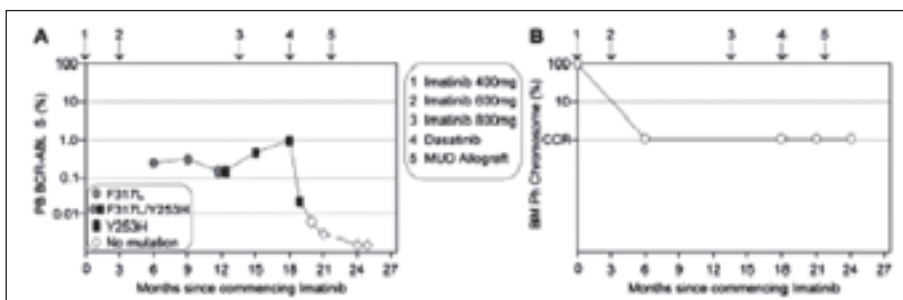
**Fig. 1** - Ruolo della remissione molecolare dopo consolidamento (1).

te la normale interazione del complesso recettore-ligando con il sistema che controlla l'acetilazione degli istoni. La presenza del gene ibrido è patognomonica e diagnostica della leucemia acuta promielocitica. La sua negativizzazione è indice della remissione molecolare e correla strettamente con la clinica: la ricomparsa di *pml/rarα* precede la recidiva clinica e permette di impostare un trattamento in fase estremamente precoce.

Lo studio di un altro gene di fusione, *bcr/abl*, ha dimostrato una importanza rilevante nel management della leucemia mielode cronica. Mentre, da un lato, il monitoraggio programmato della malattia minima residua ha permesso di docu-



**Fig. 2** - Migliore prognosi dei pazienti affetti da leucemia promielocitica trattati alla recidiva molecolare rispetto a quelli trattati alla ripreca ematologica (2).



**Fig. 3** - L'incremento di bcr/abl remissione ematologica e citogenetica permette, in questo paziente di sospettare l'insorgenza al glivec e di passare efficacemente ad altra terapia mentre lo studio citogenetico non dimostra modificazioni (3).

**Tab. 1** - Geni di fusione nelle Leucemie Mieloblastiche Acute (4).

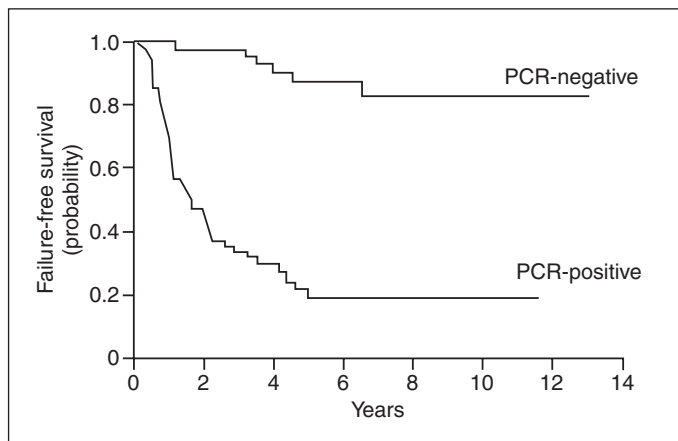
Translocation	FAB	Fusion Gene	Frequency	System	Outcome
t(8;21)(q22;q22)	Specially, M2 Rarely, M1 or M4	AML1-ETO	12	FISH, RT-PCR	Favourable
Inv(16)(p13;q22)	Specially, M4Eo Rarely, M2, M5 M1, M6, M7	CBFb-MYH11	10	FISH, RT-PCR	Favourable
t(15;17)(q22;q21)	M3	PML-RAR $\alpha$	8	FISH, RT-PCR anti-PML antibody	Favourable
t(11;17)(q23;q21)	M3	PLZF-RAR $\alpha$	<1	RT-PCR	Unfavourable
t(11;17)(q13;q21)	M3	NUMA1-RAR $\alpha$	<1	RT-PCR	Unfavourable
Der17	M3	STAT5b-RAR $\alpha$	<1	RT-PCR	?
t(5;17)(q35;q21)	M3	NPM-RAR $\alpha$	<1	RT-PCR	?
t(9;22)(q34;q11)	Specially, M1 Rarely, M2b or M4 More rarely, M0		1-2	FISH, RT-PCR	Unfavourable
t(9;1)(p21;q23)	Specially, M5a Rarely, M5b or M4 More rarely, M0 M1 or M2	MLL-AF9	2	FISH, RT-PCR	Unfavourable intermediate
t(10;11)(p21;q23)	Specially, M5a Rarely, M5b or M4 More rarely, M0 M1 or M2	MLL-AF10	<1	FISH, RT-PCR	Unfavourable intermediate
t(7;11)(q15;p15)	Specially, M2 or M4	NUP98-HOXA9	<1	FISH, RT-PCR	Unfavourable
t(11;19)(q23;p13.1)	Specially, M5a Rarely, M5b or M4 More rarely, M0 M1 or M2	MLL-ELL	<1	FISH, RT-PCR	Unfavourable intermediate
t(6;9)(p23;q34)		DEK-CAN	<1	FISH, RT-PCR	Unfavourable

mentare come la LMC non sia eradicabile nella maggior parte dei casi con il trattamento a base di specifici inibitori delle tirosin-kinasi, appare evidente come i pazienti che ottengano una significativa riduzione (almeno tre logaritmi) del gene ibrido rilevato mediante PCR quantitativa abbiano un miglior andamento. In particolare, l'incremento dell'espressione del trascritto in successivi controlli permette di modificare il trattamento prima della recidiva citogenetica.

**Tab. 2** - Geni di fusione nelle Leucemie Linfoblastiche acute (4).

Chromosomal translocation	Fusion gene	Line-age	Frequency (%)	Detection system	Prognostic outcome
t(9;2)(q34;q11)	BCR-ABL	B	5 (pediatric) 25 (adult)	FISH, RT-PCR	Unfavourable
t(12;21)(p13;q22)	TEL-ALM1	B	20 (pediatric) <1 (adult)	FISH, RT-PCR	Favourable
t(9;11)(q21;q23)	MLL-AF9	B	<1	FISH, RT-PCR	Unfavourable
t(11;19)(q23;p13.3)	MLL-ENL	B	<1	FISH, RT-PCR	Unfavourable
t(1;19)(q23;p13)	E2A-PBX1	B	5 (pediatric) 3 (adult)	FISH, RT-PCR	Unfavourable
t(17;19)(q23;p13)	E2A-HLF	B	<1	FISH, RT-PCR	?
t(4;11)(q21;q23)	MLL-AF4	B	2 (pediatric) <1 (adult)	FISH, RT-PCR	Unfavourable
t(8;14)(p24;q32) <sup>a</sup>	MYC-IgH	B	<1	FISH, RT-PCR	Unfavourable
del6	?	B	10 (pediatric) <1 (adult)	FISH, RT-PCR	Unfavourable
t(11;14)(p13;q11) <sup>a</sup>	LMO2-TCR8	T	<1	FISH, RT-PCR	Unfavourable
t(10;14) <sup>a</sup>	HOX11-TCR8	T	<1	FISH, RT-PCR	Unfavourable
t(1;14) <sup>a</sup>	TALI-TCR8	T	<1		Unfavourable
t(11;19)(q23;p13.3)	MLL-ENL	T	<1	FISH, RT-PCR	Unfavourable
t/de(9)(p21)	INK4A, p19 <sup>ARFy</sup>	T	10 (pediatric) <1 (adult)	FISH, RT-PCR	Unfavourable
Microdeletion Ip32	SIL-TAL1	T	26 (pediatric) <16 (adult)	FISH, RT-PCR	Unfavourable

<sup>a</sup>These translocations result in gene deregulation rather than in a protein.



**Fig. 4** - Ruolo della MRD nei pazienti trapiantati con Prodotti aferetici manipolati (8).

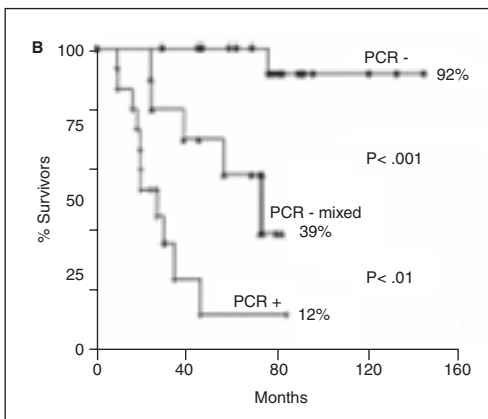




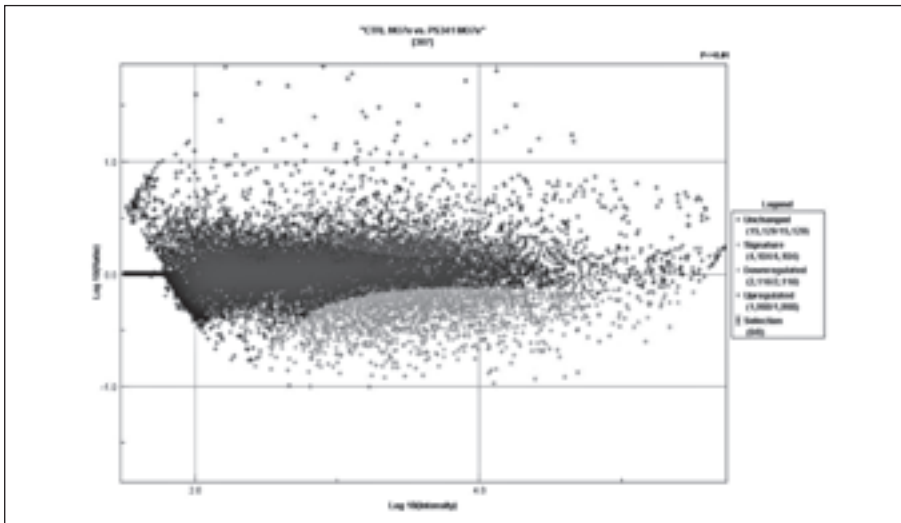
**Fig. 5** - Andamento dell'autotrapianto in relazione al livello di contaminazione delle sacche aferetiche. Dati dalla casistica dell'Ematologia di Pisa.

Lo studio molecolare ha permesso di documentare numerosi geni di fusione correlati con leucemie acute: alla maggior parte di questi geni viene attribuito un valore prognostico. Occorre, tuttavia, sottolineare che l'eradicazione dell'espressione di questi trascritti non ha sempre un valore predittivo dell'evoluzione e, ad esempio, la persistenza del trascritto  $CBF\beta /MYH11$  non impone un approccio terapeutico aggressivo.

Per contro, numerosi studi confermano il ruolo prognostico di mutazioni all'interno di geni cruciali nella trasmissione del segnale intracellulare. Mutazioni e duplicazioni all'interno del gene  $FLT3$  hanno, nelle leucemie acute, un ruolo prognostico rilevante tanto da determinare il grado di aggressività della malattia e suggerire un approccio terapeutico trapiantologico in prima remissione. Il valore delle mutazioni di  $FLT3$  ha un valore prognostico indipendente anche dall'espressione della multidrug resistance 1 (5). In maniera opposta, mutazioni di  $NPM1$ , associandosi ad un cariotipo normale nelle LMA (6), ha un valore prognostico favorevole. Numerosi dati suggeriscono che la valutazione di alcuni geni particolarmente



**Fig. 6** - DFS in relazione allo stato del midollo post trapianto (9).



**Fig. 7** - Modulazione dell'espressione di oltre 4.000 geni in una linea cellulare (Mo7) verso controllo dopo esposizione a PS 341. Dal database dell'Ematologia di Pisa e di Modena.

te espressi in corso di leucemie acute possa contribuire alla determinazione della malattia minima residua pur se la loro espressione non è, di per sé, diagnostica. Un esempio interessante è rappresentato da WT-1 che è frequentemente molto espresso sia nelle leucemie acute che nelle sindromi mielodisplastiche (7).

Lo studio dell'espressione di geni di fusione si è esteso a numerose altre patologie ematologiche quali, ad esempio, le sindromi ipereosinofile e nelle mieloproliferative croniche Ph-negative, ma una disamina analitica esula dalle finalità di questa relazione. Anche nelle malattie linfoproliferative la formazione di geni ibridi ha un ruolo diagnostico e forse patogenetico. Nelle forme B linfocitarie numerosi esempi dimostrano come traslocazioni portino geni rilevanti sotto il dominio delle regioni regolatrici dei geni delle immunoglobuline determinando, ad esempio, l'iper espressione di BCL1 nel linfoma mantellare, di BCL6 in numerosi casi di linfoma diffuso a grandi cellule ed alla alterata espressione di BCL2 nel linfoma follicolare. Lo studio di questi geni può essere fondamentale nella valutazione della malattia minima residua: le implicazioni cliniche della persistenza di malattia minima non sono uniformemente condivise ma, in generale, la negativizzazione della malattia a livello molecolare sembra accompagnarsi ad una più lunga DFS, almeno nei linfomi mantellari e nei follicolari, specialmente dopo trattamenti immuno-chemioterapici e/o trapiantologici. In quest'ultimo setting, numerosi lavori dimostrano come la negativizzazione molecolare post-trapianto si accompagni ad una migliore prognosi. Anche l'impiego di sacche aferetiche da utilizzare per autotrapianto prive di contaminazione ha un valore prognostico.

Non solo la completa negatività dei prodotti aferetici trapiantati modifica la DFS ma il livello di contaminazione ha un ruolo rilevante nel determinare la prognosi del paziente trapiantato.

Tuttavia l'impatto più rilevante lo ha la positività o negatività del midollo osseo dopo autotrapianto.

In assenza di specifici marcatori neoplastici, può, frequentemente, essere seguita la clonalità valutando il riarrangiamento dei geni per le immunoglobuline o per il T cell receptor in modo da seguire la massa della malattia. In questo senso sono state proposte numerose metodiche che si basano sull'identificazione del clone realizzando primers paziente-specifici o seguendo l'amplificazione del riarrangiamento clonale identificato con metodi di sequenziamento (10).

È, ormai, ampiamente dimostrato che una percentuale di pazienti che abbiano ottenuto una remissione molecolare, ricadranno. Benchè sia sempre possibile ipotizzare che la recidiva rappresenti l'insorgenza di una nuova patologia verso la quale il paziente è "predisposto", la maggior parte delle recidive è legata alla mancata eradicazione della patologia di base. Le differenti percentuali di ricadute molecolari dopo la remissione a seconda dell'aggressività della iniziale terapia fortemente indicano che tra la negativizzazione molecolare diagnosticata con i mezzi attuali e l'eradicazione esistono numerosi logaritmi di possibile malattia che dovranno essere esplorati con metodiche maggiormente sensibili.

Più recentemente, l'impiego dei microarrays ha permesso di identificare patterns di attivazione/inibizione caratteristici di forme neoplastiche specifiche. Fino ad oggi, l'impiego più vasto di questi studi consisteva nell'identificare un numero ristretto di marcatori capace di identificare uno specifico tipo di malattia. Un elegante esempio di questa applicazione è dato dal lavoro di Basso et al., nella leucemia a cellule capellute (11).

L'insieme delle modificazioni di espressione genica appare, tuttavia, estremamente complesso e la possibilità di utilizzare il profilo di espressione in modo da caratterizzare la singola patologia valutando l'intero genoma o ampie selezioni di geni coinvolti appare estremamente promettente, anche se non immediatamente realizzabile su larga scala. Le modificazioni di espressione potrebbero anche predire la sensibilità a diversi trattamenti in modo da disegnare le cure sul singolo paziente. In quest'ottica la possibilità di aumentare la sensibilità delle tecniche molecolari e di implementare i modelli matematici di analisi dei dati contribuirà a rendere ancora più attraente questo tipo di indagini.

## Bibliografia

1. Buccisano F., Maurillo L., Gattei V., Del Poeta G., Del Principe ML., Cox MC., Panetta P., Irno Consalvo M., Mazzone C., Neri B., Ottavini L., Fraboni D., Tamburini A., Lo-Coco F., Amadori S. and Venditti A. The kinetics of reduction of minimal residual disease impacts on duration of response and survival of patients with acute myeloid leukemia. *Leukemia* 2006; 20: 1783-1789.
2. Lo-Coco F., Breccia M., Diverio D. The importance of molecular monitoring in acute promyelocytic leukaemia. *Best Pract Res Clin Haematol.* 2003; 16: 503-520.
3. Branford S. Chronic Myeloid Leukemia: Molecular Monitoring in Clinical Practice. *American Society of Hematology* 2007: 376-383.

4. Lo-Coco F., Breccia M., Noguera N., Miller WH. Diagnostic value of detecting fusion proteins derived from chromosome translocations in acute leukemia *Best Practice & Research Clinical Haematology* 2003; 16: 653-670.
5. Galimberti S., Rossi A., Palumbo GA., Morabito F., Guerrini F., Vincelli I. Fazzi R., Petrini M. FLT3 mutations do not influence MDR-1 gene expression in acute leukemia. *Anticancer Res.* 2003; 23: 3419-3426.
6. Falini B., Mecucci C., Tiacci E., Alcalay M., Rosati R., Pasqualucci L., La Starza R., Diverio D., Colombo E., Cartucci A., Bigerna B., Pacini R., Pucciarini A., Liso A., Vignetti M., Fazi P., Meani N., Pettrossi V., Saglio G., Mandelli F., Lo-Coco F., Pelicci PG., Martelli MF. Cytoplasmic Nucleophosmin in Acute Myelogenous Leukemia with a Normal Karyotype *N Engl J Med* 2005; 352: 254-266.
7. Significant co-expression of WT1 and MDR1 genes in acute myeloid leukemia patients at diagnosis. *Eur. J. Haematol* 2004; 72: 45-51.
8. Freedman AS., Neuberg D., Mauch P., Soiffer RJ., Anderson KC., Fisher DC., Schlossman R., Alyea EP., Takvorian T., Jallow H., Kuhlman C., Ritz J. Nadler LM., Gribben JG. Long-term follow-up of autologous bone marrow transplantation in patients with relapsed follicular lymphoma. *Blood* 1999; 94: 3325-3333.
9. Corradini P., Tarella C., Zallio F., Doderò A., Zanni M., Valagussa P., Gianni AM., Rambaldi A., Barbui T., Cortelazzo S. Long-term follow-up of patients with peripheral T-cell lymphomas treated up-front with high-dose chemotherapy followed by autologous stem cell transplantation. *Leukemia.* 2006; 20: 1533-1538.
10. Galimberti S., Brizzi F., Mameli M., Petrini M. An advantageous method to evaluate IgH rearrangement and its role in minimal residual disease detection. *Leukemia Res* 1999; 23: 921-929.
11. Basso K., Liso A., Tiacci E., Benedetti R., Pulsoni A., Foa R., Di Raimondo F., Ambrosetti A., Califano A., Klein U., Dalla Favera R., Falini B. Gene expression profiling of hairy cell leukemia reveals a phenotype related to memory B cells with altered expression of chemokine and adhesion receptors. *J Exp Med.* 2004; 199: 59-68.

## **Patient-specific tumor idiotype protein or idiotype (VDJ)-derived class I-restricted peptides as biomarkers for the development of anti-tumor vaccines in multiple myeloma**

*Roberto M. Lemoli*

Istituto di Ematologia e Oncologia Medica, Policlinico Sant'Orsola, Bologna

Monocyte-derived dendritic cells (DCs) were infused in multiple myeloma (MM) patients after autologous stem cell transplantation. DCs were loaded with the tumor idiotype (whole protein (=6) or idiotype (VDJ)-derived class I-restricted peptides (=9)) and keyhole limpet hemocyanin (KLH). Patients safely received three subcutaneous and two intravenous injections of increasing DC doses. Seven of 14 evaluable patients developed an idiotype-specific T-cell proliferative response and increased IFN- $\gamma$ -secreting T cells, 4/14 showed an idiotype-positive skin test and 2/2 increased anti-idiotype CTL-p. T-cell responses were observed only after subcutaneous DC injections and idiotype-derived peptides were equally effective as the whole protein to stimulate immunity.

### **Introduction**

Dendritic cells (DCs) are professional antigen-presenting cells showing the most potent stimulatory activity on T-cell immunity (1). Murine DCs loaded with tumor-associated antigens, induce specific anti-tumor responses *in vitro* and *in vivo* (1, 2). In humans, several pilot clinical trials support the role of DCs in anti-tumor immunity (2). Early data on DC vaccination of multiple myeloma (MM) and lymphoma patients indicate that stimulation of an anti-idiotype (Id) T-cell response may sometimes be feasible (2). However in MM, the heterogeneity of the trial designs (e.g source of DCs, route of administration, number of injections, patient selection and antigen formulation) as well as the functional deficiencies of circulating DCs (3) and Id-specific T cells, which often become tolerant by peripheral (4) or central (5) deletion as disease progresses, and the limited number of patients enrolled in each trial may account for the modest number of immune responses observed and the poor clinical results.

CD14<sup>+</sup> monocytes from MM patients can be reproducibly induced to differenti-

ate into fully functional, mature, CD83<sup>+</sup> DCs according to Good Manufacturing Practice (GMP) guidelines (6, 7). Here, we report the safety profile, the immunologic and clinical evaluation of DC therapy in 15 MM patients with low tumor burden after tandem autologous stem cell transplantation (ASCT) who did not tolerate maintenance therapy or showed initial signs of disease progression under therapy. The trial was designed to assess the role of:

- 1) two different routes of DC vaccine administration (subcutaneous, sc; or intravenous, iv);
- 2) increasing doses of Id-pulsed DCs;
- 3) two types of Id formulation (i.e. whole protein or Id-derived peptides).

## Patients and methods

### *Patients*

Fifteen patients with symptomatic MM who had received double ASCT were enrolled in this phase I/II study. The clinical characteristics of study patients are reported in Table 1. After transplantation, MM patients were initially treated with interferon- $\alpha$  (IFN) (3 MU 3 times/week) as maintenance therapy. Monthly dexamethasone (Dex) was added in case of disease progression. DCs were administered to 4 patients, who did not tolerate maintenance treatment, before any serologic MM progression, and to 11 relapsing patients. This trial was approved by the Italian Ministry of Health (N.8136-Di21-783) and the local ethical committee (28/2002/U). Informed consent was obtained from study patients prior to enrollment according to the Declaration of Helsinki.

### *Tumor antigen*

Peptides derived from the VDJ IgH gene rearrangements were produced for the nine patients, for whom monoclonal Ig was not available. The VDJ gene rearrangement was identified for each patient as previously reported (8, 9). The putative amino-acidic sequence corresponding to each patient-specific VDJ IgH region was obtained from the ExpASY Proteomic Server ([www.expasy.org](http://www.expasy.org)). By applying structural prediction models derived either from the Bioinformatics & Molecular Analysis Section of The National Institutes of Health web site or from the International Immunogenetics Information System<sup>®</sup> web site (<http://bimas.cit.nih.gov> and <http://imgt.cines.fr>, respectively) all the Ig-deduced protein sequences were evaluated for peptides of nine and ten aminoacids in length, which could potentially bind to MHC class I molecules. For each patient, two peptides were selected: one deriving from the three complementary determining regions (CDR) IgH rearranged regions and one from the framework (FW) IgH rearranged region. The peptides were selected according the highest prediction score. In two cases, only peptides from FW3 and FW2 were chosen due to the exceedingly low score of CDR-derived peptides (Table 1). Monoclonal IgG or IgA were purified from serum using protein-G column (Hi Trap protein G Sepharose, Pharmacia Biotech, Uppsala Sweden) as reported elsewhere (10).

**Tab. 1** - Clinical characteristics of study patients and vaccine formulation.

Pts	Sex/ age	Stage	Ig isotype	Pre-Vax status	Auto-TX (N=2) cond.	Time from TX (months)	Id formulation Whole protein	Peptides
1	F/61	II A	IgG/k	VGPR	MEL BU-MEL	60	IgG+KLH	-
2	M/55	I A	IgG/k	Progression	MEL BU-MEL	30	IgG+KLH	-
3	M/50	III A	IgA/k	VGPR	MEL BU-MEL	60		ARVGAWAEAFD (10) CDR3 DNAKNSLYL (9) FW3 +KLH
4	M/63	III A	IgA/k	VGPR	MEL BU-MEL	60		SLIVADTAL (9) FW3 AIVALPDDL (9) CDR3 +KLH
5	F/59	I A	IgG/k	VGPR	MEL BU-MEL	30	IgG+KLH	-
6	M/55	II A	IgG/k	Progression	MEL BU-MEL	61		RLSPGKALE (9) FW2 RVVIAASSP (9) CDR3 +KLH
7	M/54	I A	IgG/k	MCR	MEL BU-MEL	19		FTNYDINWV (9) CDR1 SLTSDDTAV (9) FW3 +KLH
8	M/37	III A	IgG/k	Progression	MEL BU-MEL	18	IgG+KLH	-
9	M/59	III A	IgG/k	Progression	MEL BU-MEL	73		YDSGERWRL (9)CDR3 YIISDADSV (9) CDR2/FW3 +KLH
10	M/62	I A	IgG/k	CR	Mel Bu-Mel	9		SLRAEDTAV (9) FW3 QAPGKLEWV (10) FW2 +KLH
11	M/59	I A	IgA/k	PR	MEL BU-MEL	108	IgA+KLH	-
12	M/68	I A	IgA/λ	Progression	MEL BU-MEL	48		SLRVEDTALY (10)FW3 ALYYCAKEDY (10)FW2 +KLH
13	F/58	I A	IgG/k	Progression	MEL BU-MEL	25		FCARVLRDV (9) FW3 SLRLTSVTA (9) CDR3 +KLH
14	F/54	I A	IgG/λ	VGPR	MEL BU-MEL	88		SLESRLAISV (10) FW2 SPLDSAGYSK (10) CDR3 +KLH
15	M/66	III A	IgA/λ	PR	MEL MEL	24	IgA+KLH	-

Legend: VGPR, very good partial remission; MCR, molecular complete remission; CR, complete remission; PR, partial remission; MEL, melphalan; BU, busulphan; Tx, transplantation.

### Production of DC-based vaccine

Monocyte-derived DCs were generated and loaded with the tumor antigen (50  $\mu\text{g}/\text{mL}$ ) and keyhole limpet hemocyanin (KLH, Byosin, Germany; 50  $\mu\text{g}/\text{mL}$ ) as earlier reported (6, 7).

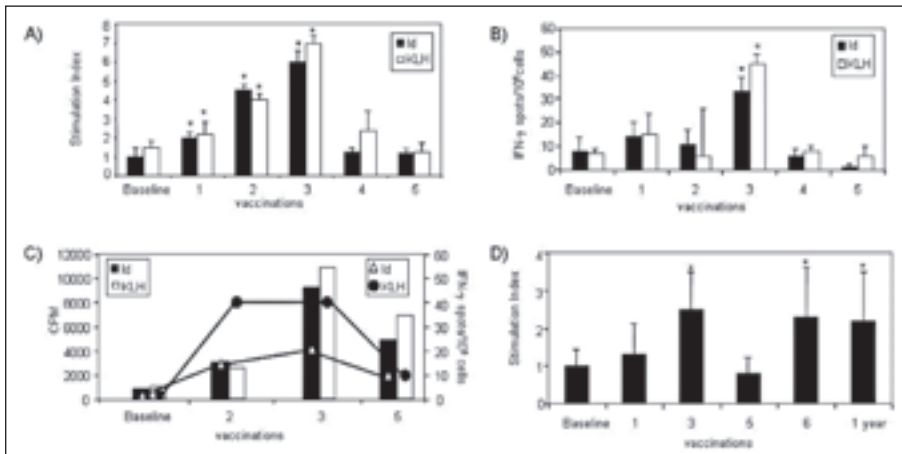
### Trial design and immunoresponse to DC vaccination

DC-based immunotherapy consisted of 3 sc and 2 iv administrations, 2 weeks apart, at increasing doses of 5x-, 10x-, 50x10<sup>6</sup> cells and 10x-, 50x10<sup>6</sup> cells, respectively. Additional sc monthly injections of 50x10<sup>6</sup> DCs were scheduled in case of stable disease and DC availability. A careful analysis was performed to test whether the two different routes of administration (i.e. sc and iv), increasing DC doses, as well as different tumor antigen formulations, resulted in efficient T-cell responses or toxicity. Humoral and T-cell responses were evaluated according to previously published conventional assays (11-14).

## Results and discussion

### KLH-specific immune response

All evaluable patients (n =14) showed a KLH-specific antibody response. Nine out of 14 patients showed a KLH-specific T-cell proliferative response, 8/14



**Fig. 1** - Kinetics of the immune response to DC-based anti Id vaccination. T-cell proliferative responses (i.e. >50% increase compared to baseline) and the concentration of PB IFN- $\gamma$  producing T cells are reported in panels A and B, respectively, for all (=14) treated patients. The correlation between these two T-cell activation parameters is clearly demonstrated in panel C depicting a representative clinically responding patient (N. 2). Overall, T-cell responses become measurable after the first sc DC injection, the peak values are observed after the third sc vaccination with the highest number of DCs (i.e. 5x 10<sup>7</sup>) and decrease after iv injections regardless of the number of DCs infused (i.e. 1 or 5 x 10<sup>7</sup>). Anti-Id proliferation, which was found to be decreased after iv infusion, was restored in 3 patients (N. 2, 7, 10) by sc injection of 5 x 10<sup>7</sup> DCs and the proliferative response lasted up to one year following therapy (D). The data are presented as mean  $\pm$  SD and were compared using Wilcoxon's test. P values for T-cell proliferation to Id and KLH were statistically significant (\*, <0.05) after the first, second and third vaccination. ELISPOT assay showed statistically significance (\*, <0.05) after the third s.c. vaccination.

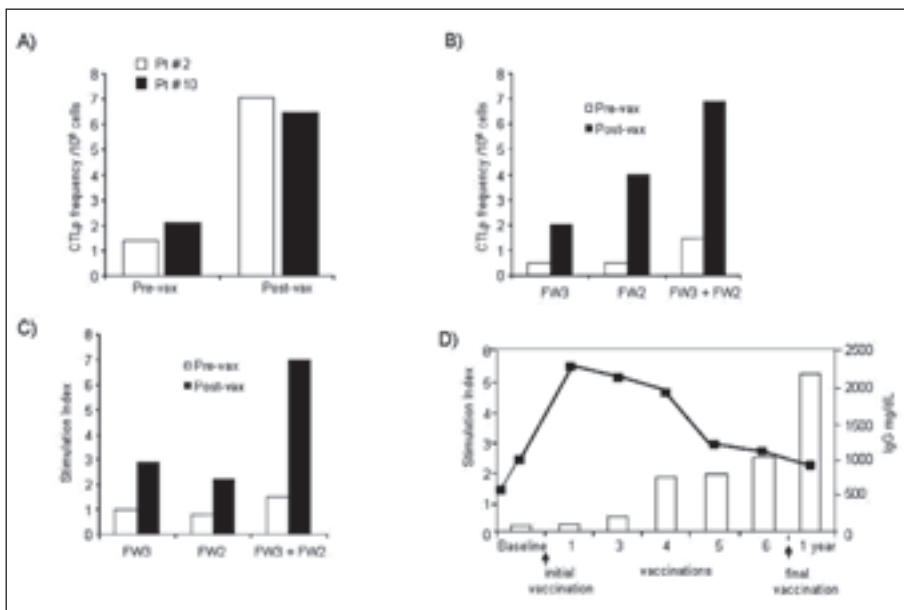


showed a significant increase in the concentration of circulating KLH-specific IFN- $\gamma$  secreting T cells after the third vaccination, and 9/14 had a positive delayed-type hypersensitivity reaction (DTH) test. T-cell responses became measurable after the first DC injection, the peak values were observed after the third vaccination and decreased thereafter as shown in Fig. 1.

Thus, despite heavy prior treatment, MM patients retained a certain degree of immunocompetence and mounted B- and T-cell primary immunoresponses to a strong xenoantigen.

#### *Id-specific immune response*

None of the patients mounted a detectable Id-specific antibody response. Conversely, 7/14 patients showed an Id-specific T-cell proliferative response that was not present before vaccination. Moreover, 7/14 significantly increased IFN- $\gamma$ -secreting T-cell levels measured by ELISPOT analysis after the third vaccination, and 4/14 demonstrated an Id-positive DTH test. T-cell response peaked after the third sc DC injection, and declined thereafter (Fig. 1). Anti-Id proliferation was restored up to 1 year by sc DC injection in 3/3 patients (N.2, 7, 10) who received 1, 3 and 4 additional monthly sc DC injections, respectively (Fig. 1d).



**Fig. 2** - Increased frequency of anti-Id CTLp following DC vaccination. In 2/2 evaluable patients CTLp frequency was found to be increased one year after DC vaccination (A). One of the two patients showing increased CTLp frequency after treatment was vaccinated with FW-derived peptides. In this case, we demonstrated the development of proliferative (C) and cytotoxic T-cell response (B) to both peptides and their combination was synergistic (B, C).

One of the two MM patients (N. 2) achieved PR which is maintained after 40 months. The decrease of serum M-component (indicated by the solid line) was strictly associated with the development of an anti-Id T-cell response (empty bars) (panel D).

Anti-Id T-cell immune responses were similar in patients receiving the Id as whole protein or VDJ-derived peptides (data not shown).

The frequency of anti-Id CTLp was assessed at one year from DC vaccination in 2 patients (N. 2 and 10) with stable disease. As shown in Fig. 2a, infusion of Id-pulsed DCs increased the frequency of cytotoxic T cells. In one patient (N. 10), vaccinated with FW-derived peptides, we demonstrated the development of anti-tumor T-cell proliferative (Fig. 2c) and cytotoxic responses (Fig. 2b) to both peptides and their combination was synergistic.

#### *Safety and clinical outcome*

There were no serious adverse events during DC therapy regardless of the number of DCs infused or the route of administration. No significant changes in routine blood values nor evidence of any autoimmune reaction were recorded.

Overall, 6/14 evaluable patients have stable serologic disease, one patient achieved PR and seven patients progressed. One other patient (N. 15) is presently evaluable only for safety and not for DC therapy efficacy due to the limited follow up time (1 month). Specifically, 6/8 patients in CR or VGPR showed stable disease after a median of 26 months from vaccination. Conversely, only 1/6 patients vaccinated in progression achieved PR which is maintained at 40 months from treatment. In this last case, the decrease of serum M-component was strictly associated with the development of an anti-Id proliferative T-cell response (Fig. 2d).

Our data demonstrate the efficacy of *ex-vivo*-generated DCs to stimulate T-cell responses in MM patients. Half of the study patients showed an anti-Id T-cell proliferative response, increased numbers of Id-specific IFN- $\gamma$ -secreting T cells and, in some cases, enhanced CTL-p frequency and positive DTH reaction despite the immunogenic weakness of the tumor Id. Interestingly, we found peak values after the third sc injection followed by a rapid decline concomitant with iv infusion. Anti-Id T-cell responses were then restored in 3/3 patients who received additional sc injections of  $50 \times 10^6$  DCs. Thus, we provide evidence that sc administration is more effective than iv infusion to induce immunological responses as shown in melanoma patients (15). The immunological responses were not different between patients immunized with the whole Id protein or Id (VDJ)-derived patient-specific peptides. Thus, our epitope-prediction analysis was effective in identifying peptides that could potentially bind to one or more HLA molecules expressed by the patients and these epitope represent a target for T-cell reactivity. Contrasting results are available on the capacity of FW-derived peptides to stimulate T cells (16, 17). In our study, we showed, in one evaluable patient, the post-vaccination increase of CTLp for neoplastic plasma cells following the administration of FW-derived peptides-pulsed DCs. We chose to utilize only HLA class I-restricted peptides because of limitations of the structural prediction models employed in this study and the effectiveness of such peptides in eliciting T-cell responses (15). KLH was then used as a “helper” epitope to potentiate the vaccine formulation (18). However, given the importance of CD4<sup>+</sup> cells in maximizing anti-tumor cytotoxicity, future trials should likely include HLA class II peptides.

Experimental and preliminary clinical data have emphasized the role of anti-tumor vaccination in the elimination of minimal residual disease (2). Our results support this notion, since 6 of 8 patients vaccinated in CR or VGPR showed stable disease without further treatment after a median follow up of 26 months whereas only 1/6 patients vaccinated with overt disease achieved PR.

In summary, whereas iv administrations of DCs were ineffective to stimulate T-cell responses, sc injections of cryopreserved idio-type-pulsed DCs were safe and induced anti-MM immunological responses regardless of the loading antigen formulation.

## References

1. Bancherou J., Steinman RM. Dendritic cells and the control of immunity. *Nature* 1998; 392: 245-52.
2. Bocchia M., Bronte V., Colombo MP., et al. Antitumor vaccination: where we stand. *Haematologica* 2000; 85: 1172-1206.
3. Ratta M., Fagnoni F., Curti A., et al. Dendritic cells are functionally defective in multiple myeloma: The role of interleukin-6. *Blood* 2002; 100: 230-7.
4. Bogen B. Peripheral T-cell tolerance as a tumor escape mechanism: deletion of CD4+ T cells specific for a monoclonal immunoglobulin iditype secreted by a plasmacytoma. *Eur J Immunol* 1996; 26: 2671-9.
5. Lauritzsen GF., Hoofgard PO., Schenk K., Bogen B. Clonal deletion of thymocytes as a tumor escape mechanism. *Int J Cancer* 1998; 78: 216-22.
6. Ratta M., Curti A., Fogli M., et al. Efficient presentation of tumor iditype to autologous T-cells by CD83+ dendritic cells derived from highly purified circulating CD14+ monocytes in multiple myeloma patients. *Experimental Hematology* 2000; 28: 931-40.
7. Motta MR., Castellani S., Rizzi S., et al. Generation of dendritic cells from CD14+ monocytes positively selected by immunomagnetic adsorption for multiple myeloma patients enrolled in a clinical trial of anti-iditype vaccination. *Br J Haematol* 2003; 121: 240-50.
8. Lemoli RM., Martinelli G., Zamagni E., et al. Engraftment, clinical and molecular follow-up of patients with multiple myeloma who were reinfused with highly purified CD34+ cells to support single or tandem high-dose chemotherapy. *Blood* 2000; 95: 2234-9.
9. Martinelli G., Terragna C., Lemoli RM., et al. Molecular remission after allogeneic or autologous transplantation of hematopoietic stem cells for multiple myeloma. *J Clin Oncol* 2000; 18: 2273-81.
10. Curti A., Isidori A., Ferri E., et al. Generation of dendritic cells from positively selected CD14+ monocytes for anti-tumor immunotherapy. *Leukemia Lymphoma* 2004; 45: 1419-28.
11. Massaia M., Borrione P., Battaglio S., et al. Iditype vaccination in human myeloma: generation of tumor-specific immune responses after high-dose chemotherapy. *Blood* 1999; 94: 673-83.

12. Coscia M., Mariani S., Battaglio S., et al. Long-term follow-up of idiotype vaccination in human myeloma as a maintenance therapy after high-dose chemotherapy. *Leukemia* 2004; 18: 139-45.
13. Comoli P., Montagna D., Moretta A., et al. Alloantigen-induced human lymphocytes rendered non responsive by a combination of anti-CD80 monoclonal antibodies and cyclosporin-A suppress mixed lymphocyte reaction in vitro. *J Immunol* 1995; 155: 5506-11.
14. Comoli P., Pedrazzoli P., Maccario R., et al. Cell therapy of stage IV nasopharyngeal carcinoma with autologous Epstein-Barr virus-targeted cytotoxic T lymphocytes. *J Clin Oncol* 2005; 23: 8942-9.
15. Thurner B., Haendle I., Roder C., et al. Vaccination with Mage-3A1 peptide-pulsed mature, monocyte-derived dendritic cells expands specific cytotoxic T cells and induce regression of some metastases in advanced stage IV melanoma. *J Exp Med* 1999; 190: 1669-78.
16. Trojan A., Schultze JL., Witzens M., et al. Immunoglobulin framework-derived peptides function as cytotoxic T-cell epitopes commonly expressed in B-cell malignancies. *Nat Med* 2000; 6: 667-72.
17. Baskar S., Kobrin CB., Kwak LW. Autologous lymphoma vaccines induce human T-cell responses against multiple unique epitopes. *J Clin Invest* 2004; 113: 1498-1510.
18. Timmerman JM, Levy R. Linkage of foreign carrier protein to a self-tumor antigen enhances the immunogenicity of a pulsed dendritic cell vaccine. *J Immunol* 2000; 164: 4797-4803.

## **Lo studio del profilo di espressione genica: applicazioni cliniche**

*Sergio Ferrari*

Dipartimento di Scienze Biomediche, Sezione di Chimica Biologica,  
Università di Modena e Reggio Emilia, Modena

La ricerca biologica e biomedica è in un momento di transizione significativo, determinato da due principali fattori: la conoscenza della sequenza del DNA genomico umano e lo sviluppo di nuove tecnologie per utilizzarla. Sfortunatamente, la conoscenza della sequenza del DNA non fornisce informazioni su come i singoli geni agiscano, su come la cellula operi, su come le cellule si organizzino a formare organismi, su cosa non “funzioni” in caso di malattia, o come poter dirigere lo sviluppo di un farmaco. Entrano quindi in gioco gli studi funzionali, il cui scopo non è semplicemente quello di fornire un elenco di tutti i geni e le informazioni sulle loro funzioni, ma di capire come le componenti lavorino insieme a formare cellule ed organismi. Tra gli strumenti più efficaci e versatili per questo genere di studi vi sono gli arrays di oligonucleotidi o di DNA complementare, chiamati DNA microarrays, microarrays di oligonucleotidi o chips. L'insieme dei geni trascritti dal DNA genomico è indicato come profilo d'espressione o “trascrittoma”, e caratterizza il fenotipo e la funzione della cellula. Al contrario del genoma, il “trascrittoma” è molto dinamico e cambia rapidamente e drammaticamente in risposta a qualsiasi cambiamento od anche durante eventi cellulari, come la replicazione del DNA, la divisione cellulare, il differenziamento, l'apoptosi, la fagocitosi, la mobilità o la trasformazione neoplastica. Inoltre, cambiamenti nel quadro multi genico dell'espressione possono fornire indizi circa i meccanismi regolatori, le funzioni e le vie biochimiche cellulari. Ancora, studi sul profilo di espressione genica sono utili ad individuare cause e conseguenze delle malattie, come agiscono i farmaci sulle cellule e quali prodotti genici possono essere bersagli molecolari adatti per la terapia o la scoperta di biomarcatori utili per la diagnosi. Va notato tuttavia che l'RNA messaggero è soltanto un prodotto intermedio della via di produzione di una proteina, ma gli studi delle proteine si mostrano a tutt'oggi molto difficili e meno sensibili nel predire variazioni funzionali della cellula. I livelli di RNA messaggero offrono importanti informazioni sullo stato funzionale della cellula e, per la maggior parte dei geni, un aumento del livello trascrizionale si traduce in un aumento della quantità di RNA messaggero e di proteina tradotta. Il sequenziamento e l'analisi del genoma

umano indicano che circa 25.000-30.000 geni codogenici vengono attivamente trascritti, potendo però generare con lo splicing alternativo un numero maggiore di prodotti proteici. Analisi comparative in diversi contesti cellulari, utilizzando la tecnologia degli arrays, hanno contribuito ad approfondire le conoscenze su processi biologici fondamentali, come la proliferazione, il differenziamento, il processo apoptotico e la trasformazione neoplastica. Storicamente, studi tesi in questa direzione hanno sempre proceduto un gene per volta, ma l'avvento di questa nuova tecnologia rende possibile raccogliere informazioni da migliaia di geni, disegnando un quadro globale dell'espressione genica e suggerendo nuovi meccanismi di regolazione dei processi biologici. Ad esempio, è stato possibile indagare interessanti contesti biologici:

- a) espressione genica differenziale nelle fasi successive del ciclo cellulare o nel differenziamento cellulare;
- b) meccanismi molecolari alla base della trasformazione e della progressione neoplastica;
- c) identificazione dei geni espressi durante il programma genetico pro-apoptotico;
- d) modificazioni dell'espressione genica durante le infezioni virali; e individuazione dei geni interessati nel metabolismo di farmaci.

Le analisi di espressione e di clustering nei diversi contesti cellulari forniranno utili informazioni sulla complessità di molti programmi genetici ed aiuteranno l'individuazione delle loro alterazioni in caso di malattia. Esistono più di 200 tipi diversi di cancro e ciascuno può essere considerato un contesto cellulare unico, ciò che rende la classificazione tramite analisi con DNA microarrays un valido strumento nella lotta contro i tumori. Un recente studio ha preso in considerazione i profili di espressione mediante cDNA microarrays in 60 linee cellulari neoplastiche di diversa origine, anche emopoietica, con la prospettiva di comparare tali profili con i tumori originari. Inoltre negli ultimi dieci anni in quasi tutti i tipi di tumori sono stati eseguiti esperimenti di DNA microarrays con risultati contrastanti e non sempre clinicamente utili.

### **Basi tecnologiche dei DNA microarrays**

La tecnologia dei DNA microarrays si è sviluppata a partire dalla fine degli anni novanta e ha consentito lo studio dell'espressione di migliaia di geni contemporaneamente. La comunità scientifica ha poi cominciato ad applicarla ad un ampio spettro di processi biologici sia normali che patologici. Tale metodologia permette infatti la fenotipizzazione molecolare di un contesto cellulare. Due tipi di DNA microarrays sono attualmente in uso, gli oligonucleotidi arrays e gli spotted arrays. Numerose sono inoltre le piattaforme tecnologiche che permettono di eseguire questo tipo di indagine con procedura semiautomatizzata. Una delle applicazioni più importanti degli arrays è, come già accennato, il monitoraggio dell'espressione genica globale di una cellula a livello degli RNA messaggeri. I risultati che si ottengono dalle diverse piattaforme non sempre sono riproducibili e quindi confrontabili fra loro.

## **Il profilo di espressione genica globale o trascrittoma**

L'insieme delle migliaia di geni espressi o trascritti dal DNA genomico in un determinato contesto cellulare prende il nome di profilo di espressione genica (PEG) o trascrittoma, e permette di determinare il fenotipo molecolare della cellula (fenotipizzazione molecolare). Come primo passo nel processo di sintesi proteica, il DNA genomico viene trascritto in RNA messaggero, e si assume che ad un RNA messaggero corrisponda una proteina, anche se in realtà ciò non sempre è vero, in quanto esiste un controllo traduzionale dell'espressione genica.

Possiamo comunque assumere che differenze nell'espressione genica sono responsabili delle differenze morfologiche e fenotipiche della cellula in risposta a stimoli microambientali o a perturbazioni di vario tipo (nutrizionali, condizionamento genico, temperatura, interazioni patogeni-ospite, trattamento con composti biologicamente attivi, ecc.). A differenza del genoma, il trascrittoma è altamente dinamico e si modifica rapidamente e drammaticamente in risposta a perturbazioni o anche durante i normali eventi cellulari quali proliferazione, differenziamento, apoptosi, attivazione, fagocitosi ecc. Per la comprensione della funzione dei geni, sapere quando, dove e come un gene è espresso (in termini di abbondanza di RNA messaggero) è fondamentale per capire l'attività e il ruolo biologico della proteina da esso codificata. Inoltre, e più rilevante, le variazioni dei pattern di espressione permettono di valutare i meccanismi di regolazione a livello trascrizionale, di sviluppare mappe trascrizionali e quindi di capire un ampio spettro di funzioni biologiche, quali vie metaboliche sono attivate o quali vie di trasduzione del segnale sono coinvolte nel processo biologico oggetto di studio. Lo studio del trascrittoma è stato negli ultimi 8 anni ampiamente utilizzato non solo per lo studio dei processi biologici normali ma anche per studiare le cause e le conseguenze di malattie, le interazioni tra farmaci e funzioni cellulari, lo sviluppo di nuovi farmaci mirati su bersagli genici. L'uso di questa tecnologia ha richiesto la risoluzione di problemi legati al disegno sperimentale, all'uso di piccole quantità di RNA messaggero ottenibile da campioni biologici (ad es. biopsie), allo sviluppo di software per la corretta analisi dei dati e alla valutazione della significatività statistica dei risultati ottenuti, al fine di estrarre dati biologicamente rilevanti utilizzabili nella pratica clinica. Sulle riviste internazionali sono stati pubblicati migliaia di articoli scientifici nei quali è stato studiato il trascrittoma. I campi applicativi di tale tecnologia in biologia e medicina sono molteplici, e fra questi troviamo lo studio dei programmi genetici di processi biologici attivabili dalle cellule (quali quello proliferativo, differenziativo, apoptotico), lo studio dei meccanismi molecolari alla base della trasformazione e progressione delle cellule tumorali, la classificazione molecolare dei diversi tipi di tumori, lo studio della interferenza dei farmaci con la trascrizione genica, lo studio della farmacoresistenza, lo studio dei meccanismi patogenetici di malattie complesse e l'individuazione di nuovi geni bersaglio per lo sviluppo di strategie terapeutiche mirate. L'insieme dei geni che sono coespressi in un processo biologico normale o patologico viene definito "signature" e può avere una valenza diagnostica, prognostica o patogenetica se individuato in una malattia. Segnaliamo inoltre la possibilità di consultare in rete anche dati relative

agli esperimenti di DNA microarrays, in particolare Microarrays and gene expression (<http://www.mged.org/mage>) e Stanford microarrays Database (<http://genome-www5.stanford.edu>).

### **Analisi dei dati di microarrays**

Cinque sono le fasi fondamentali dell'analisi dei dati di microarrays:

- 1) disegno sperimentale;
- 2) processamento;
- 3) inferenza;
- 4) classificazione;
- 5) validazione.

#### *1) Il disegno sperimentale*

Il disegno sperimentale interferisce con l'efficienza e la validità degli esperimenti. C'è un consenso generale della comunità scientifica sul fatto che la replica biologica degli esperimenti è essenziale. Possiamo programmare una replica tecnica, cioè usare l'RNA estratto da un singolo campione biologico per l'ibridazione a diversi DNA arrays. Questa replica permette di valutare la variabilità tecnica dei diversi DNA chip e quindi costituisce un controllo di qualità. La replica biologica viene fatta partendo invece da diversi esperimenti biologici e quindi diversi campioni ibridati e permette di valutare la variabilità e quindi la riproducibilità del sistema biologico oggetto dello studio. La domanda che sorge spontanea è: quanti replicati biologici sono necessari per poter avere una significatività statistica dei risultati? Se quello che si deve valutare è l'espressione differenziale fra due contesti cellulari sono necessarie dalle 3 alle 5 repliche. Riunire in pool gli RNA derivati da esperimenti o campioni omogenei è possibile e in alcuni casi necessario quando l'RNA che si può estrarre dai singoli campioni è piccolo e non sufficiente per programmare un esperimento di DNA microarray. Questa risoluzione elimina le possibili varianti biologiche individuali e dà risultati in parte diversi dalla ibridazione dei singoli campioni. Infine gli esperimenti biologici vanno rigorosamente programmati usando le stesse condizioni sperimentali e possibilmente gli stessi reagenti.

#### *2) Processamento*

La scansione delle immagini dei dati di ibridazione su DNA microarrays è una procedura di acquisizione necessaria sia per la normalizzazione, cioè il processo per cui l'intensità dei vari spot sono aggiustati tenendo conto della variabilità fra i diversi esperimenti e piattaforme, che per rimuovere le variazioni sistematiche. Inoltre permette la trasformazione di dati, cioè l'applicazione di specifiche funzioni matematiche cosicché i dati sono cambiati in una forma differente, ad esempio log<sub>2</sub> (fold change: comparazione fra i livelli di espressione fra due distinti campioni). Il filtraggio dei dati, cioè la selezione dei dati importanti per l'interpretazione dell'esperimento. Ad esempio se sono interessato ai geni differenzialmente espressi è evidente che filtro, cioè escludo dall'analisi, i geni espres-



si costitutivamente o non espressi. Un esempio di filtraggio dei dati che riguarda gli arrays a due colori è la sottrazione del fondo.

### 3) *Inferenza*

Si tratta di trarre conclusioni sulla veridicità dell'ipotesi formulata. Ad es. l'ipotesi da verificare è: esistono differenze nell'espressione genica quando coltivo le stesse cellule nella condizione A o nella condizione B? Nel trarre le conclusioni si possono commettere molti errori quali la presenza di falsi positivi o falsi negativi. Esistono molti metodi di analisi dei dati che permettono di evitare questi errori che non abbiamo la possibilità di trattare dettagliatamente. Fra questi possiamo annoverare il t-test, ANOVA, la tecnica di regressione logistica ecc. Nella comunità scientifica vi è unanime accordo che non è più sufficiente prender in considerazione il solo "fold change" come parametro dell'espressione differenziale.

### 4) *Classificazione*

Il processo di classificazione consiste nell'ordinare i geni in categorie preesistenti, in questo caso si parla di "supervised classification" anche denominata assegnazione o predizione di classe o discriminazione, o di sviluppare un set di categorie in cui i geni vengono raggruppati "unsupervised classification", detta anche scoperta di classe. Diversi algoritmi sono stati sviluppati per questo tipo di analisi di "clustering".

### 5) *Validazione*

Nella maggior parte dei casi si preferisce validare i dati di espressione mediante altre tecniche ed in particolare mediante RT-PCR quantitativa. È evidente che se limito il mio studio all'espressione genica valuto di base l'attività trascrizionale della cellula in quanto pochi sono gli esempi di stabilizzazione e compartimentalizzazione degli RNA messaggeri. Posso validare i dati microarrays anche studiando le proteine codificate dagli RNA messaggeri mediante le diverse metodologie della proteomica, ma in realtà sono numerosi ormai gli esempi in letteratura che non tutti gli RNA messaggeri vengono decodificati a proteine.

## **Rappresentazione grafica dei dati derivati dall'analisi del trascrittoma**

Il percorso informatico per analizzare i dati di espressione ottenuti dagli esperimenti di DNA microarrays può essere lungo e laborioso ma fondamentalmente viene adattato alla domanda che il ricercatore si pone in relazione al sistema biologico oggetto di studio. Una visualizzazione relativamente semplice che mi permette di valutare fra due o più contesti cellulari la statistica dei geni espressi nella sua globalità è il diagramma di Venn. Questa rappresentazione grafica è utile nelle analisi esplorative basilari sui profili di espressione genica, per confrontare geni espressi o liste di geni differenzialmente espressi.

Anche lo scatter plot consente di visualizzare graficamente differenze nel profilo di espressione genica tra due contesti cellulari. Il clustering gerarchico permette

di visualizzare graficamente il grado di similitudine tra i profili di espressione di più campioni o di più geni. Il grado di similitudine reciproca tra i differenti profili di espressione è indicato dalla lunghezza dei rami del dendrogramma. Con lo stesso data set utilizzato per il clustering gerarchico posso, mediante l'analisi delle componenti principali (PCA) stabilire la distanza Euclidea fra i diversi insiemi di geni.

Se l'esperimento che ho programmato prevede diversi campioni che fanno parte ad esempio di una cinetica di induzione del differenziamento cellulare posso valutare mediante l'uso di programmi analitici specifici, come Gene Spring, l'andamento coerente di insiemi di geni che vengono riuniti per tipologia di espressione (indotti, silenziati o costitutivamente espressi).

### **Annotazioni funzionali e classificazione dei geni in famiglie**

L'annotazione funzionale dei geni è un passaggio obbligato verso la comprensione delle interazioni tra i geni e i prodotti genici, come realizzato attraverso gli studi di espressione genica e di proteomica, che è fondamentale per la comprensione dell'attivazione o inattivazione di funzioni cellulari specifiche ed in particolare di moduli funzionali (insieme di geni che permettono l'attivazione di un processo biologico). Attualmente il consorzio Gene Ontology (GO) ha in atto un progetto per annotare tutte le caratteristiche funzionali del gruppo completo di geni identificati dai diversi progetti genoma e quindi anche quello umano (HGP); il sito di GO è accessibile all'indirizzo <http://www.gene-ontology.org>. Tale sito classifica le funzioni dei geni in tre categorie principali, che sono:

- 1) processo biologico, cioè il processo in cui rientra la funzione della proteina, ad esempio proliferazione, differenziamento, apoptosi, immunità naturale ecc.;
- 2) funzione molecolare, che definisce l'attività della proteina, ad esempio fattore trascrizionale, molecola d'adesione, recettore, enzima, proteina legante il calcio ecc.;
- 3) componente cellulare, che si riferisce alla localizzazione della proteina nei diversi compartimenti cellulari, ad esempio nucleo, citoplasma, mitocondri, nucleolo, lisosomi, reticolo, membrane ecc.

Ciascuna di queste categorie è organizzata in modo gerarchico dalla funzione generale a quella più specifica. L'organizzazione del sito è dinamica e prevede un continuo aggiornamento delle categorie, che va di pari passo con la disponibilità di nuove acquisizioni. Le annotazioni funzionali dei geni sono riportate anche in altri siti quali ad esempio Genecards <http://www.genecards.org/index.shtml> e OMIM (On line Mendelian Inheritance in Men) <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>.

Sono disponibili programmi di analisi dei dati, disponibili in rete, utili allo scopo di classificare funzionalmente i geni come GenMAPP o ricostruire vie di trasduzione del segnale, MAPPFinder (<http://www.GenMAPP.org>), o NetAffx, software di analisi funzionale dei geni a disposizione degli utenti della piattaforma Affymetrix, che sono fondamentali per l'interpretazione dei dati di espressione genica.

## **Le mappe trascrizionali**

Un ulteriore sviluppo dell'analisi dei dati ottenuti dai profili di espressione genica, reso possibile dallo sviluppo di nuovi softwares, è l'integrazione dei dati relativi al trascrittoma con le informazioni della localizzazione dei geni sul DNA genomico. Tale tipo di analisi permette di stabilire interessanti correlazioni tra posizione genomica, funzione biologica e pattern di espressione genica. Con questa metodica di analisi applicata a diversi sistemi biologici si sono individuati domini cromatinici iperespressi o silenti.

## **Analisi di DNA microarray e classificazione dei tumori**

Il potenziale di indagine dei DNA microarrays nello studio globale dell'espressione genica è stato recepito fin da subito ed è stato largamente usato per definire nuove sottoclassi di malattie, per identificare nuovi marcatori biologici associati a malattie o per distinguere sottoclassi all'interno della stessa patologia. In particolare, lo studio del PEG è stato utilizzato per classificare diversi tipi di tumore ("signature diagnostica") o per predire la risposta alla chemioterapia e quindi la sopravvivenza dei pazienti trattati ("signature prognostica"). Un esempio è lo studio di DNA microarrays condotto in Olanda su pazienti affette da carcinoma mammario.

Tale studio ha permesso l'individuazione di una signature prognostica di 70 geni predittivi dello sviluppo di metastasi e quindi di un aggravamento del tumore. Tale evidenza ha portato a modificare gli schemi terapeutici ed ha una valenza predittiva molto superiore rispetto alla diagnostica convenzionale basata sull'esame istologico del tumore e sull'esame clinico. Un altro esempio, fra i molti che si potrebbero riportare, è la sottoclassificazione molecolare delle leucemie mieloidi acute.

Tali leucemie sono tradizionalmente divise in 7 sottoclassi diverse in base ai criteri diagnostici convenzionali, che riguardano lo studio della morfologia, l'espressione di enzimi cellulari evidenziata con coloranti cromogenici (citochimica), l'espressione di proteine di superficie riconosciute da anticorpi specifici (immunofenotipo), la citogenetica per lo studio delle alterazioni cromosomiche, e le metodiche molecolari per lo studio di alterazioni di geni specifici. Uno studio cooperativo fra diversi gruppi di ricerca, pubblicato recentemente, ha sottoclassificato tali leucemie in base al PEG in 16 sottogruppi che hanno una diversa risposta alla terapia, che va da una buona responsività alla refrattarietà al trattamento.

Uno studio analogo, sempre condotto su pazienti affetti da leucemie mieloidi acute, ha invece individuato una "signature" prognostica di 133 geni predittivi che correla con la sensibilità dei pazienti alla terapia. Tuttavia deve essere ricordato che questo tipo di studi richiede una stretta collaborazione fra clinici, laboratoristi e bioinformatici, e che spesso più istituzioni devono collaborare a distanza per la raccolta di un numero sufficiente di pazienti da studiare. Un'altra applicazione importante dei DNA arrays riguarda la farmacogenomica.

## DNA microarrays e farmacogenomica

La farmacogenomica è lo studio dell'impatto della variabilità genomica individuale sulla risposta ai farmaci, sulla loro efficacia terapeutica e sul loro metabolismo, ed è forse da considerare una delle prime applicazioni del HGP nella pratica clinica in medicina. Negli anni cinquanta questo campo di studio era chiamato farmacogenetica, ma con l'aumento delle conoscenze sul genoma umano è stato adottato il termine di farmacogenomica. Con i DNA microarrays è virtualmente possibile esaminare la risposta dell'intero genoma della cellula quando esposta a farmaci e quindi stabilire come il farmaco è in grado di modificare l'espressione genica.

Nel paragrafo precedente sono stati riportati alcuni esempi di come si possa arrivare ad una "signature" prognostica, cioè a definire un insieme di geni associati alla farmacoresistenza. Il relativamente nuovo settore della farmacogenomica contribuirà notevolmente a capire il meccanismo d'azione dei farmaci e il meccanismo di risposta o di adattamento delle cellule in risposta all'esposizione a farmaci. Infatti, le differenze nel profilo di espressione genica indotte da farmaci o da composti biologicamente attivi possono essere analizzate per individuare nuove attività di questi prodotti. Un simile approccio sperimentale, chiamato "forward pharmacology", può portare all'individuazione di nuovi effetti biologici di un farmaco già conosciuto e possibilmente a nuove applicazioni terapeutiche in considerazione dei geni coinvolti nell'attività del farmaco stesso. Inoltre possono essere individuati e studiati nuovi geni bersaglio. È altrettanto evidente che lo sviluppo di nuovi farmaci prevede necessariamente lo studio della loro interferenza con il genoma, e quindi lo studio del PEG diventerà indispensabile per individuare possibili effetti tossici e stabilire l'indice terapeutico. Quando questa metodologia è applicata allo studio dei tumori, ed in particolare allo sviluppo di nuove strategie terapeutiche, il termine più comunemente usato è Oncogenomica (Strausberg RL. et al., Nature 2004).

## Bibliografia

1. International Human genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature 431: 931-945, 2004.
2. Lockhart DJ. and Winzeler EA. Genomics, gene expression and DNA arrays. Nature 405: 827, 2000.
3. Lipshutz RJ., Fodor SPA., Gingeras TR., Lockhart DJ. High density synthetic oligonucleotide arrays. Nature Genetics (Suppl.) 21: 20, 1999.
4. Epstein CB., Butow RA. Microarray technology - enhanced versatility, persistent challenge. Curr. Opinion Biotech. 11: 36, 2000.
5. Pandey A., Mann M. Proteomics to study genes and genomes. Nature 405: 837, 2000.
6. Young RA. Biomedical discovery with DNA arrays. Cell 102: 9, 2000.
7. Cho RJ, Campbell MJ., Winzeler EA., Steinmetz L., Conway A., Wodicka L., Wolfsberg TG., Gabrielian AE., Landsman D., Lockhart DJ., Davis RW.

- A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular Cell* 2: 65, 1998.
8. Zhu H, Cong JP, Mamtora G, Gingeras T, Shenk T. Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *PNAS* 95:14470, 1998.
  9. Der SD, Zhou A, Bryan R, Williams G, Silverman RH. Identification of genes differentially regulated by interferon  $\alpha$ ,  $\beta$ , or  $\gamma$  using oligonucleotide arrays. *PNAS* 95: 15623, 1998.
  10. Tagliafico E, Tenedini E, Bergamaschi A, Manfredini R, Percudani R, Siena M, Zanocco-Marani T, Grande A, Montanari M, Gemelli C, Torelli U, Ferrari S. Gene expression profile of Vitamin D3 treated HL-60 cells shows a phenotypic but not a complete functional conversion to monocytes. Submitted.
  11. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P. Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genetics* 24: 227, 2000
  12. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531, 1999.
  13. Quackenbush J. Microarray analysis and tumor classification. *New England Journal of Medicine* 354: 2463-2472, 2006.
  14. Quackenbush J. Computational analysis of microarray data. *Nature Reviews in Genetics* 2: 418-427, 2001.
  15. Mark Schena *Microarray Analysis*, Wiley-Liss 2003.
  16. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: disarray to consolidation and consensus. *Nature Reviews in Genetics* 7: 55-65, 2006.
  17. Doniger SW, Salomonis N, Dahlquist KD, Vranizan K, Lawlor SC, Conklin BR. MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biology* 4: R7, 2003.
  18. Ashurst JL, Collins JE. Gene annotation: prediction and testing. *Annu. Rev. Genomics Hum. Genet.* 4: 69-88, 2003.
  19. Ashburner M, Ball CA, Blake JA, et al Gene Ontology: tool for unification of biology. *Nature Genetics* 25: 25-29, 2000.
  20. Yamishita T, Honda M, Takatori H, Nishino R, Hoshino N, Kaneko S. Genome wide transcriptome mapping analysis identifies organ specific gene expression pattern along human chromosomes. *Genomics* 84: 867-875, 2004.
  21. Liang P, Pardee AB. Analysing differential gene expression in cancer. *Nature Rev. Cancer* 3: 869-876, 2003.
  22. Van de Vijver MJ, He YD, van't Veer LJ, et al. A gene expression signature as a predictor of survival in breast cancer. *New England Journal of Medicine* 347: 1999-2009, 2002.

23. Valk PJ., Verhaak RGW., Beijen MA. et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *New England Journal of Medicine* 350: 1617-1628, 2004.
24. Bullinger L., Dohner K., Bair E. et al. Use of Gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *New England Journal of Medicine* 350:1605-1616, 2004.
25. Tagliafico E., Tenedini E., Manfredini R., Grande A., Ferrari F., Roncaglia E., Bicciato S., Zini R., Salati S., Bianchi E., Gemelli C., Montanari M., Vignudelli T., Zanocco-Marani T., Parenti S., Paolucci P., Martinelli G., Piccaluga PP., Baccarani M., Specchia G., Torelli U., Ferrari Ser. "Identification of a molecular signature predictive of sensitivity to differentiation induction in Acute Myeloid Leukemia. *Leukemia* 20: 1751-1758, 2006
26. Ginsburg GS., Konstance RP., Allsbrook JS., Schulman KA. Implications of pharmacogenomics for drug development and clinical practice. *Archives of Internal Medicine* 165: 2331-2336, 2005.
27. Strausberg RL., Simpson AJG., Old LJ., Riggins GJ. Oncogenomics and the development of new cancer therapies. *Nature* 429: 469-474, 2004.

## Capire il cancro

*Lucio Luzzatto*

Direttore Scientifico Istituto Toscano dei Tumori, Firenze  
Cattedra di Ematologia, Università di Firenze

Da tempo immemorabile i tumori devono essere stati riconosciuti come gravi malattie in molte culture: ma probabilmente non di rado venivano confusi con altre malattie che danno tumefazioni, come processi infettivi di vario genere. Forse proprio per evitare equivoci, e in virtù dell'aspetto di tumori avanzati che per fortuna oggi raramente vediamo, è entrata in uso da secoli la parola «cancro» (greco *καρκινος* latino, francese, inglese *cancer*; tedesco *Krebs*), il nome di un crostaceo e della costellazione in cui il sole entra alla fine di giugno. Così un segno dello Zodiaco venne ad evocare, volutamente o no, un aggressore o un parassita che si abbarbica a un organo del corpo e ne succhia la sostanza e la forza.

Fu Rudolf Virchow, il patriarca della patologia dell'Ottocento, a formulare la «teoria cellulare» del cancro. In sostanza, l'ipotesi di Virchow era che la cellula cambiasse in qualche modo la propria natura, diventando neoplastica. Per designare questo processo, Virchow usò la parola «degenerazione»; oggi preferiamo dire che la cellula normale *si trasforma* in cellula neoplastica. Non era facile capire quale fosse l'essenza della patologia tumorale, poiché i singoli tumori sono diversissimi: occorre perciò uno sforzo notevole per astrarre, da questa eterogeneità, quali fossero gli elementi più caratteristici. Apparve chiaro che, in generale, un tessuto tumorale differisce da un tessuto normale per almeno due aspetti. Il ritmo di divisione cellulare era elevato: questo poteva spiegare la crescita del tumore. La proliferazione cellulare, anziché essere ordinata e controllata come nei tessuti normali dell'embrione e dell'adulto, appare fuori controllo, e spesso in modo drammatico.

### **Il ciclo cellulare**

Per comprendere come la crescita cellulare divenga sregolata, occorre far riferimento al processo della divisione cellulare, la cui regolazione si è raffinata nel corso dell'evoluzione, ma le cui caratteristiche fondamentali sono già presenti in organismi unicellulari come i lieviti: in questi organismi molto più semplici degli uomini, è stato più facile chiarire i meccanismi di controllo, che sono di due tipi. Da un lato, la cellula deve decidere se e quando dividersi; dall'altro, una volta

attivato il programma di divisione, la cellula vuole essere certa che esso sarà eseguito correttamente. Il controllo è esercitato per mezzo di «posti di blocco». Per esempio, se una cellula allo stadio cosiddetto G1 non è pronta per S (sintesi del DNA), il processo abortisce. Lo stesso succede se i cromosomi non si separano correttamente nella fase M (mitosi). L'importanza di questi controlli è evidente: essi mirano a prevenire in tutti i modi la formazione di cellule anomale. È facile immaginare che nelle cellule umane i controlli siano ancora più rigorosi che nei lieviti.

Il processo di replicazione è basato sulla ben nota struttura a doppia elica della molecola del DNA. Ogni filamento dell'elica funge da stampo per la realizzazione di un nuovo filamento. Per avere la certezza che ciascun duplicato sia fedele, il processo di replicazione è sottoposto a una continua sorveglianza, in modo che eventuali errori siano subito scoperti. In altre parole, ha luogo una continua «revisione delle bozze», che mira a correggere immediatamente eventuali errori. Ciò nonostante, nessun processo fisico è perfetto, e gli errori, sebbene raramente, hanno luogo. Molti errori non avranno conseguenze, perché gran parte del DNA non è costituito da geni. Tuttavia, di tanto in tanto un singolo errore, che in genetica chiamiamo *mutazione puntiforme*, può avere gravi conseguenze. Se un dattilografo, copiando un manoscritto, commettesse un errore ogni milione di battute, nessuno potrebbe accusarlo di negligenza; ma poiché parliamo di miliardi di caratteri, ciò corrisponderebbe a qualche migliaio di errori.

### **Le mutazioni somatiche sono eventi stocastici**

In un organismo normale la proliferazione cellulare è strettamente controllata. Il giusto incremento di una popolazione di cellule richiede una quantità appropriata di fattori di crescita, il normale svolgimento del ciclo cellulare, l'entrata e uscita dal ciclo al momento opportuno. Qualche volta, se un errore durante il ciclo cellulare non è altrimenti correggibile, la cellula attiva un meccanismo di autodistruzione programmata chiamato apoptosi. Ognuno di questi fenomeni dipende dall'intervento di varie proteine, codificate da altrettanti geni. Non conosciamo il numero totale dei geni coinvolti, ma una stima approssimativa permette di stimare che il totale sia di qualche centinaio. Qualunque di questi geni può mutare: se ciò avviene, siccome un gene mutante è sempre un gene, quando duplica se stesso durante il ciclo cellulare - verosimilmente, questa volta, senza errori - le cellule figlie avranno il gene mutante; e così via. Dopo un certo tempo, ci sarà una popolazione di cellule tutte con il gene mutante; e dal momento che tutte le cellule in questione derivano dalla prima cellula in cui si è verificata la mutazione, possiamo parlare di «clone mutante». In tutte le cellule del clone, la proteina codificata dal gene mutante sarà anomala. Spesso l'anomalia della proteina ha conseguenze trascurabili e non nuoce alla salute. Se però una mutazione riguarda, ad esempio, un gene che controlla l'uscita dal ciclo cellulare, le cellule del clone mutante avranno difficoltà a passare dal ciclo riproduttivo alla fase di riposo. Una simile mutazione produrrebbe quindi un clone con la tendenza a crescere eccessivamente.



Mentre la replicazione del DNA è di natura decisamente deterministica, il verificarsi di una mutazione è un evento stocastico imprevedibile. In altre parole: in ogni mutazione c'è un elemento di casualità. Siccome il nostro corpo è formato da circa  $10^{18}$  (o un miliardo di miliardi) di cellule somatiche; è quindi inevitabile che, col passare degli anni, le mutazioni diventino sempre più numerose.

### **La trasformazione di una cellula da normale in cancerosa**

Se il clone che abbiamo appena descritto potesse essere già considerato un tumore maligno, probabilmente noi tutti saremmo vittime di questa patologia prima di avere superato i cinque anni di età. Ma il controllo della proliferazione cellulare è così importante che si sono evoluti numerosi meccanismi di controllo e creazione di copie di riserva. Perciò, per eludere la normale supervisione della proliferazione la cellula deve infrangere le regole non una volta sola, ma ripetutamente. In altre parole, il processo di trasformazione della cellula da normale a cancerosa non è affatto banale: richiede senza dubbio più di una mutazione, e va quindi considerato un *processo a tappe*.

Il controllo finemente regolato della riproduzione delle cellule dell'organismo è stato paragonato, in modo forse un po' rozzo, alla guida efficiente di un'automobile, nel senso che implica l'uso sia dell'acceleratore che dei freni. Visto che il tumore è una proliferazione patologica, è logico immaginare che una mutazione favorirà la sua comparsa in uno di due modi: o facendo funzionare troppo un gene promotore della crescita (come quando si preme troppo l'acceleratore); o danneggiando un gene inibitore della crescita (come quando si guastano i freni). Nel primo caso il gene mutato si chiama *oncogene*, perché lo consideriamo un gene che favorisce la comparsa del cancro (il gene normale, non mutato, sarebbe quindi un *proto-oncogene*). Nel secondo caso, il gene coinvolto si chiama *gene oncosoppressore*. Dal momento che i geni esistono sempre in paia, è importante precisare che nel caso dell'oncogene una mutazione è sufficiente a causare problemi, perché il gene mutante è libero di premere l'acceleratore anche se il suo partner non lo fa. Al contrario, se un oncosoppressore è mutato, l'altro gene agirà da freno; perciò entrambi i geni della coppia devono diventare mutanti prima che la possibilità di frenare vada perduta. Non c'è dubbio che ci rimangono non pochi oncogeni e non pochi oncosoppressori da identificare; ma è interessante che finora l'elenco dei primi sia molto più corto di quello dei secondi.

Se il cancro sia una sola malattia o molte malattie è stato oggetto di vivaci discussioni: era una questione di sostanza, o una mera disputa terminologica? Oggi l'analisi molecolare rende forse superata la controversia. In tutti i tumori troviamo che una serie di mutazioni somatiche ha trasformato una cellula normale in una cellula anormale: le anomalie sono svariate, ma il denominatore comune è che in quella cellula si è perso il controllo della crescita. In questo senso tutti i tumori insieme sono un gruppo unitario. D'altra parte, in ogni tipo di tumore i geni coinvolti sono diversi: in questo senso possiamo considerare il cancro come una collezione di patologie distinte. In altre parole, entrambe le impostazioni

sono corrette ed hanno ripercussioni pratiche. Infatti, se il cancro è in sostanza una sola patologia, ciò che abbiamo appreso su un tumore maligno può aiutarci a capire le proprietà di tutti gli altri. Al tempo stesso, siccome vi sono tante vie diverse che portano alla trasformazione neoplastica, è necessario escogitare misure preventive e terapeutiche su misura per ogni singolo tipo di cancro.

### **Perché può svilupparsi il cancro**

Per qualunque malattia è naturale chiedersi quali siano le cause: e se la malattia è frequente non saranno solo gli addetti ai lavori a porsi questa domanda. Anche se a tutt'oggi non possiamo dare una risposta completa, abbiamo dei dati di fatto piuttosto chiari.

1) La frequenza del cancro aumenta di molto con l'età. In effetti, anche se il cancro sfortunatamente può insorgere nei giovani e perfino nei bambini, in generale è una malattia degli anziani. Questa regola generale è più o meno valida a seconda del tipo di neoplasia; nell'insieme, statistiche accurate hanno mostrato che la frequenza del cancro aumenta in modo più che proporzionale col crescere dell'età. Evidentemente, siccome la trasformazione della cellula normale in cellula maligna è un processo a tappe che richiede diverse mutazioni somatiche, nelle persone anziane queste mutazioni hanno avuto più tempo per accumularsi.

2) Il ruolo dell'eredità. Il cancro non è mai ereditario, ma un'augmentata probabilità di ammalarsi può essere ereditaria. Alcuni geni mutanti, se ereditati, aumentano in generale il rischio di sviluppare un tumore; altri geni mutanti aumentano il rischio di ammalarsi di un particolare tipo di cancro - per esempio, di un cancro della retina, del seno o del colon. Anche nelle persone che hanno ereditato una maggiore probabilità di ammalarsi, una o più mutazioni somatiche sono necessarie all'insorgere della neoplasia. Perciò, un rischio più elevato non implica la certezza di sviluppare il cancro. Ci sono almeno due modi in cui le mutazioni ereditarie possono agire. Alcune aumentano la probabilità che abbiano luogo certe mutazioni somatiche; in altri casi, una mutazione che avvicina una cellula alla condizione della malignità può essere direttamente trasmessa al figlio da uno dei genitori invece che essere acquisita. In simili casi, il maggior rischio di ammalarsi dipende dal fatto che tutte le cellule di una persona sono una tappa più avanti lungo il percorso che porta al cancro. Oltre a quei geni mutanti ereditari per i quali è stato già dimostrato che aumentano significativamente il rischio di sviluppare un tumore, ce ne possono essere altri dotati di un effetto meno evidente ma pur sempre rilevante. Nelle famiglie con un'alta percentuale di tumori maligni può essere giustificato offrire test genetici per stabilire se un membro della famiglia ha o no il gene mutante. Il test deve essere effettuato solo dopo consulenza appropriata, in un centro con personale specializzato nella genetica dei tumori.

3) Il ruolo dell'ambiente. Sappiamo che i tumori hanno origine dalle cellule dei nostri tessuti, ma fattori ambientali possono accelerare il processo di trasformazione da cellula normale a cellula tumorale, o aumentando la frequenza delle mutazioni somatiche o attraverso altre modalità. Possiamo dire con certezza che

gli agenti ambientali implicati sono molteplici: comprendono agenti fisici, come le radiazioni, compresi i raggi ultravioletti; agenti chimici come gli idrocarburi aromatici e l'asbesto; agenti infettivi come il virus del papilloma e il batterio *Helicobacter pylori*. Da recenti analisi vari fattori, pertinenti all'ambiente e ai comportamenti delle persone, sono emersi come tali da dare un contributo significativo allo sviluppo dei tumori in molte popolazioni: Soprappeso/obesità; Scarsità di frutta e verdura nella dieta; Scarsità di esercizio fisico; Fumo di sigaretta; Uso eccessivo di alcool; Attività sessuale non protetta; Inquinamento atmosferico urbano; Fumo da combustibili solidi in ambienti confinati; Uso di aghi da iniezione contaminati. È importante sottolineare però che il contributo di questi nove fattori non è paritetico: complessivamente, il fumo è il colpevole massimo.

### **Tradizioni e innovazione nel trattamento dei tumori**

Dal punto di vista medico, capire le cause del cancro è un mezzo: il fine è di controllare la malattia. Non possiamo qui addentrarci nella questione complessa del trattamento dei tumori, i cui cardini sono la chirurgia, le radiazioni e la chemioterapia. Ovviamente, l'intervento chirurgico mira alla rimozione completa primaria è in rapporto troppo stretto con strutture anatomiche vitali, o perché si è diffuso in altre parti del corpo. Lo scopo della radioterapia e della chemioterapia è distruggere le cellule neoplastiche ovunque siano, nuocendo il meno possibile al resto dell'organismo. Poiché la caratteristica principale delle neoplasie è la crescita abnorme, gli agenti antitumorali [che comprendono sia (b) che (c)] sono scelti o perché inibiscono la crescita delle cellule maligne, o della neoplasia, la quale, però, non è sempre possibile, o perché il cancro perché ne favoriscono la morte, o per entrambe le ragioni. Inevitabilmente, questi agenti influiscono anche sulle cellule normali: per questo motivo sono detti citotossici, e causano effetti collaterali più o meno seri: il motivo per cui vengono utilizzati è che in genere le cellule neoplastiche proliferano più in fretta di quelle normali e quindi sono colpite in misura maggiore dalle sostanze citotossiche. La maggior parte degli agenti antitumorali prende di mira processi cellulari che sono direttamente o indirettamente legati alla divisione cellulare o al suo controllo. Per individuare sostanze utili alla terapia è necessario effettuare molte ricerche, prima nelle colture cellulari e poi sugli animali da esperimento. Con il progredire delle nostre conoscenze circa le lesioni molecolari prodotte nei tumori dalle mutazioni somatiche, sempre più si potranno elaborare a scopo terapeutico sostanze che prendano di mira differenze qualitative, e non solo quantitative, tra cellule tumorali e cellule normali. Simili farmaci, essendo più specifici, dovrebbero dimostrarsi molto meno tossici. È per questo che gli agenti che cercano di raggiungere il tallone d'Achille della cellula neoplastica vengono chiamati, nel gergo corrente, 'farmaci intelligenti'.

**DIAGNOSI MOLECOLARE  
DELLE NEOPLASIE EMATOLOGICHE**



## **Sviluppo di signatures molecolari per la diagnosi e la prognosi delle leucemie mieloidi acute**

*Enrico Tagliafico*

Dipartimento di Scienze Biomediche, Sezione di Chimica Biologica,  
Università di Modena e Reggio Emilia, Modena

Alcuni eventi, eccezionali nella loro ricorrenza o per le novità che annunciano, segnano delle tappe fondamentali, alle quali l'umanità finisce per fare riferimento, per un tempo più o meno lungo. Gli uomini del nostro tempo hanno avuto il privilegio di assistere da spettatori ad almeno due di questi appuntamenti con la storia. Il primo è stato il passaggio del millennio, che ha anche concluso un secolo, rivoluzionario sulla qualità della vita e per la forte accelerazione che ha impresso al progresso scientifico. Il secondo è stato la decodificazione del genoma umano, che ha generato una serie di aspettative, sulla qualità della vita e sul benessere individuale. All'indomani della conclusione del progetto genoma umano, uno dei più noti quotidiani italiani metteva a confronto, a tutta pagina, le reazioni di due genetisti. Al primo, che etichettava l'evento come "la più grande scoperta dell'umanità", faceva eco il secondo che ammoniva "non cambierà la vita di nessuno, quantomeno nei prossimi dieci anni". Queste due posizioni sintetizzano una parte significativa dei pensieri sull'argomento. Da un lato l'entusiasmo per un traguardo annunciato da tempo, che di fatto ha creato uno spartiacque nel mondo scientifico, separando simbolicamente ciò che è stata, da ciò che è destinata a diventare la ricerca biologica. Questa transizione tra il genoma e il postgenoma si prepara a trasformare le tendenze e gli orientamenti della biomedicina, che sposta l'attenzione verso nuovi paradigmi: dall'eziologia, alla patogenesi delle malattie; dallo studio dell'azione, alla regolazione del gene; dalle malattie semplici, a quelle complesse; dall'analisi di un singolo gene, a più geni, famiglie geniche, vie metaboliche e sistemi; dallo studio di un singolo gene-malattia, alla suscettibilità alle malattie complesse; dall'analisi genomica (il patrimonio ereditario di una persona), all'assetto degli RNA cellulari (trascrittoma) e delle proteine (proteoma), alla distribuzione qualitativa e quantitativa degli intermedi metabolici (metaboloma). La tecnologia dei DNA microarrays è uno tra gli strumenti più potenti e versatili dell'era post-genomica e la possibilità di studiare l'espressione genica è uno degli aspetti più interessanti di questa tecnologia. Un solo esperimento permette di studiare l'espressione di 22.000 sequenze geniche. L'insieme dei geni trascritti od espressi

dal DNA genomico viene chiamato “profilo di espressione” o “trascrittoma”; esso definisce il fenotipo di una cellula e di conseguenza le sue funzioni. L’analisi dell’espressione genica ha permesso, infatti, di valutare l’espressione genica differenziale della proliferazione e del differenziamento cellulare; i meccanismi molecolari che caratterizzano la trasformazione e la progressione tumorale, il programma pro-apoptotico e i geni coinvolti nella risposta a molti farmaci. Negli ultimi tre anni il nostro laboratorio si è concentrato sullo studio di analisi del “trascrittoma” utilizzando la piattaforma tecnologica Affymetrix GeneChip.

Il nostro laboratorio ha una lunga tradizione nello studio della regolazione genica dell’emopoiesi normale e delle alterazioni di tale processo in corso di neoplasie ematologiche. Il cancro è una malattia che presenta una spiccata individualità: pazienti con lo stesso tipo di tumore possono presentare differenti alterazioni genetiche e diversi comportamenti clinici. Per quanto concerne le neoplasie ematologiche (che sono il nostro diretto campo di interesse), anche se ancora molta strada si deve ancora percorrere molto si è compreso, sui meccanismi patogenetici che portano allo sviluppo della Leucemia. Molto meno è conosciuto sui meccanismi che determinano l’evoluzione della malattia e, soprattutto, che cosa determinare, in una particolare leucemia, la risposta ai trattamenti. In sintesi molto deve ancora essere conosciuto sui meccanismi molecolari che influiscano nella determinazione della prognosi clinica. In particolare

La prognosi nella leucemia mieloide acuta (LMA) (patologia della quale ci siamo particolarmente interessati) è stata fino ad ora stabilita in base alla citochimica e soprattutto alla morfologia, la quale però, pur rimanendo tuttora il primo strumento diagnostico di tale patologia, non possiede alcun significato prognostico indipendente (1-3). In seguito, nella scelta del protocollo terapeutico e nella formulazione di ipotesi prognostiche sono subentrati anche altri controlli quali l’immunofenotipizzazione, gli esami citogenetici e la ricerca di marcatori molecolari. Vari studi riferiscono che le caratteristiche cliniche e i parametri di laboratorio riscontrabili in un paziente affetto da LMA possono avere un significato prognostico, tuttavia sono davvero pochi quei fattori prognostici che con analisi multivariate si dimostrino realmente indipendenti. È noto che molte aberrazioni cromosomiche influenzano la prognosi, per quanto l’impatto relativo di ciascuna di queste mutazioni non sia completamente compreso. È più probabile che la biologia della malattia e la risposta al trattamento siano determinate da una combinazione di geni piuttosto che da una singola mutazione. I test citogenetici sono considerati il singolo fattore prognostico migliore nei pazienti di età inferiore ai 60 anni.

Per i pazienti che presentano la traslocazione t(15;17), l’inversione (16) o anomalie cromosomiche che coinvolgono il core binding factor (t(8;21)), la prognosi sembra essere favorevole, con un rischio approssimativo di ricaduta del 30%, mentre i pazienti con anomalie del cromosoma 5 o 7 o con anomalie cromosomiche multiple hanno all’incirca un 75% di probabilità di ricaduta. La maggior parte dei pazienti tuttavia non appartiene a queste due categorie e ad essa viene associato un rischio standard. Inoltre sembra che tale suddivisione sia insufficiente, specialmente per quei pazienti che presentano una citogenetica a rischio standard e hanno una probabilità di ricaduta altamente variabile. Per questo grup-

po eterogeneo di pazienti sono stati definiti nuovi marcatori molecolari che possiedono un significato prognostico e che soprattutto possono aprire la strada a terapie biologicamente mirate.

Sono stati proposti molti marcatori molecolari in grado di predire la prognosi nella LMA, ma solo alcuni di questi sembrano avere un significativo impatto prognostico indipendente. Molti studi hanno mostrato che mutazioni nel recettore tirosin chinasi FLT3 in grado di determinare una sua sovraespressione sono fortemente associate ad un aumento del rischio di ricaduta e ad una ridotta sopravvivenza generale nei pazienti di età inferiore ai 60 anni, indipendentemente dal loro assetto citogenetico (4-8). La presenza di FLT3/ITD (internal tandem duplication) può essere persino più importante della citogenetica nel predire il rischio di ricaduta e la sopravvivenza senza malattia (7). Mutazioni di tipo FLT3/ITD sono più frequenti in pazienti con leucemia promielocitica acuta (LPA) (37%) e in quelli con cariotipo normale (34%); inoltre i pazienti a rischio standard che presentano tale mutazione hanno un tasso di ricaduta significativamente maggiore se comparato a quello dei pazienti in cui essa è assente (74 vs 48%). Rimane però ancora da chiarire l'impatto prognostico della positività per FLT/ITD nei pazienti LPA come per quelli delle altre LMA (9-11). Il riarrangiamento del gene MLL (mixed lineage leukemia), identificato in pazienti con citogenetica normale, è una duplicazione in tandem della porzione interna del gene MLL, che si estende tra gli esoni 2 e 6 o gli esoni 2 e 8. Questa parziale duplicazione in tandem (PTD) è presente in circa il 10% di tutti i pazienti LMA con assetto citogenetico normale e nel 90% di quei pazienti che presentano come unica anomalia la trisomia del cromosoma 11. In alcuni studi è stato proposto che lo status MLL PTD sia uno dei più significativi fattori prognostici indipendenti in grado di predire la sopravvivenza nei pazienti con cariotipo normale (12), ma altri studi non sono riusciti a confermare questo risultato, pertanto l'importanza di tale mutazione come fattore prognostico indipendente rimane incerta (13).

CEBPA è un fattore di trascrizione coinvolto nella regolazione della mielopoiesi. Di recente è stato studiato un mutante di CEBPA che sembra in grado di predire una prognosi favorevole nei pazienti con cariotipo normale, indipendentemente da mutazioni di FLT3 e da parziali duplicazioni in tandem di MLL (14, 15).

Il gene del tumore di Wilms (WT1) è sovraespresso nei pazienti affetti da LMA, ed è stato riportato di recente che un alto livello di espressione di questo gene è il fattore prognostico sfavorevole indipendente più significativo, in grado di influenzare la sopravvivenza senza malattia e il rischio di ricaduta nei pazienti di età inferiore ai 60 anni, ancor più della conta dei leucociti all'esordio, delle mutazioni di FLT3 e della citogenetica (16-18).

Uno dei più importanti meccanismi di resistenza alla chemioterapia è l'efflusso rapido del farmaco, mediato da vari membri della superfamiglia dei trasportatori ATP-binding cassette, quali ad esempio MDR1 (multidrug resistance gene 1) che codifica per la glicoproteina P, MRP1 (multidrug resistance-associated protein), LRP (lung resistance protein) e "sorcin", la proteina solubile legante il calcio correlata alla resistenza ai farmaci. Tali geni sono espressi soprattutto nei pazienti più anziani, il che può in parte spiegare la minor risposta di questi pazienti alla



chemioterapia. Non è stato però tuttora chiarito se la positività per MDR abbia un simile impatto prognostico sfavorevole indipendente anche sulla resistenza alla chemioterapia (CR) e sui risultati a lungo termine nei pazienti più giovani (19-21). Inoltre il fatto che un numero significativo di pazienti possa presentare una ricaduta nonostante una bassa espressione di MDR1/MRP1 suggerisce il coinvolgimento di altri meccanismi di resistenza intracellulare.

Infine, anche i profili di espressione genica studiati in microarrays acquisiranno probabilmente maggiore significato prognostico (22, 23). Gli ultimi anni hanno visto l'esplosione di una grande mole di dati e di ricerche che hanno visto nello studio dei profili di espressione genica un mezzo potentissimo nello studio delle AML. I profili di espressione genica hanno permesso di evidenziare i sottotipi già conosciuti di AML, sottolineando che essi esistono come entità separate anche a livello di profili di espressione genica, ma hanno soprattutto permesso di evidenziarne dei nuovi. Bullinger et al., (22) hanno postulato l'esistenza di due nuove sottoclassi di AML con cariotipo normale ed associate a significative differenze nella sopravvivenza. Valk et al., (23) hanno identificato 16 clusters di signatures omogenee alcune delle quali associate a caratteristiche citogenetiche e molecolari note, altre correlati ad anomalie non prima associate a sottotipi omogenei di AML, alcune delle quali (monosomie 7 e 5 e t(9:22)) associate ad un outcome particolarmente sfavorevole.

Per quanto riguarda la prognosi nelle leucemie pediatriche, che è l'ambito in cui si sviluppa il nostro progetto di ricerca, Yagi et al. (24) hanno utilizzato l'analisi dei dati di microarray per predire gli esiti clinici delle LMA e per trovare geni la cui espressione sia correlata a prognosi sfavorevole. Sebbene la LMA possa insorgere indistintamente sia in età infantile che in età adulta, sono state descritte alcune caratteristiche ben precise che distinguono le LMA pediatriche da quelle adulte (25). Yagi et al., hanno identificato 35 geni associati alla prognosi della LMA pediatrica, selezionando i geni differenzialmente espressi tra i pazienti con esito sfavorevole e quelli con esito positivo. In seguito, analizzando i restanti pazienti che presentavano fenotipi intermedi, hanno confermato che l'espressione di tali geni è effettivamente associata alla prognosi, e hanno pertanto concluso che possono essere utili per predire al momento della diagnosi l'esito del trattamento sul paziente pediatrico affetto da LMA.

La logica conseguenza di questi studi è stata quella di incentrare notevoli sforzi nella definizione di pannelli con il numero minimo di geni capaci di predire classi di AML prognosticamente separate. Numeri differenti di geni sono stati quindi associati a varie classi (vecchie e nuove) di AML. Purtroppo, almeno fino ad ora, nonostante molti studi abbiano definito delle signature collegate all'esito, rimane purtroppo un'ampia porzione di dati in qualche modo divergente, quindi resta un problema critico chiarire le caratteristiche trascrizionali essenziali della progressione neoplastica. Un discorso diverso può invece essere fatto riguardo alla classificazione generale delle leucemia acuta. Haferlach et al., hanno chiaramente dimostrato come il gene expression profile consente di classificare in modo accurato e riproducibile tutte le classi conosciute di LA (26, 27).

Chiaramente, nonostante che questi studi abbiano mostrato potenzialità incredi-

bilmente ampie, la principale sfida immediata è la definizione di profili associati all'outcome e soprattutto alla predizione della risposta ai farmaci.

Alcuni studi sono stati pubblicati a questo riguardo, Heuser et al., hanno descritto un pannello genico che discrimina le AML resistenti alla chemioterapia convenzionale da quelle sensibili (28). In questo campo di interesse il nostro gruppo di ricerca ha apportato un contributo.

In particolare abbiamo pubblicato un lavoro (28) nel quale descriviamo lo sviluppo di una "signature" molecolare, ossia un gruppo di geni differenzialmente espressi, predittiva della risposta agli agenti differenzianti Acido All Trans Retinoico e Vitamina D3 in sei linee cellulari stabilizzate di LMA, eterogenee, come i blasti dei pazienti, per la risposta agli agenti differenzianti. I dati di espressione genica sono stati ottenuti utilizzando una piattaforma per l'analisi di espressione Affymetrix. La stessa signature è stata poi validata con successo in cellule di 28 pazienti affetti da LMA, che sono state trattate in vitro con ATRA e VD3; tale analisi è stata condotta mediante Taqman Real Time PCR con tecnologia Applied Biosystems. Inoltre, è stata eseguita una meta-analisi di dati di espressione presenti in un database pubblico su 86 pazienti affetti da LMA. La signature individuata, si è dimostrata efficace nel distinguere i pazienti responsivi dai non responsivi agli agenti inducenti.

## Bibliografia

1. Tallman MS. Relevance of pathologic classifications and diagnosis of acute myeloid leukemia to clinical trials and clinical practice. *Cancer Treat. Res.* 2004; 121: 45-67.
2. Tallman MS, Kim HT, Paietta E, Bennett JM, Dewald G, Cassileth PA, Wiernik PH, Rowe JM. Acute monocytic leukemia (French-American-British classification M5) does not have a worse prognosis than other subtypes of acute myeloid leukemia: a report from the Eastern Cooperative Oncology Group. *J. Clin. Oncol.* 2004; 22: 1276-1286.
3. Haferlach T., Kern W., Schoch C., Schnittger S., Sauerland MC., Heinecke A., Buchner T., Hiddemann W. A new prognostic score for patients with acute myeloid leukemia based on cytogenetics and early blast clearance in trials of the German AML Cooperative Group. *Haematologica* 2004; 89: 408-418.
4. Kiyoi H., Naoe T., Nakano Y., Yokota S., Minami S., Miyawaki S., Asou N., Kuriyama K., Jinnai I., Shimazaki C., Akiyama H., Saito K., Oh H., Motoji T., Omoto E., Saito H., Ohno R., Ueda R. Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood* 1999; 93: 3074-3080.
5. Yamamoto Y., Kiyoi H., Nakano Y., Suzuki R., Kodera Y. et al. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood* 2001; 97: 2434-2439.
6. Kottaridis PD., Gale RE., Linch DC. Prognostic implications of the presence of FLT3 mutations in patients with acute myeloid leukemia. *Leuk. Lymphoma* 2003; 44: 905-913.

7. Kottaridis PD., Gale RE., Frew ME., Harrison G. Langabeer SE., Belton AA., Walker H., Wheatley K., Bowen DT., Burnett AK., Goldstone AH., Linch DC. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood* 2001; 98: 1752-1759.
8. Schnittger S., Schoch C., Dugas M., Kern W., Staib P., Wuchter C., Loffler H., Sauerland CM., Serve H., Buchner T., Haferlach T., Hiddemann W. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood* 2002; 100: 59-66.
9. Shih LY., Kuo MC., Liang DC., Huang CF., Lin TL., Wu JH., Wang PN., Dunn P., Lai CL. Internal tandem duplication and Asp835 mutations of the FMS-like tyrosine kinase 3 (FLT3) gene in acute promyelocytic leukemia. *Cancer* 2003; 98: 1206-1216.
10. Au WY., Fung A., Chim CS., Lie AK., Liang R., Ma ES., Chan CH., Wong KF., Kwong YL.. FLT-3 aberrations in acute promyelocytic leukaemia: clinicopathological associations and prognostic impact. *British Journal of Haematology* 2004; 125: 463-469.
11. Ozeki K., Kiyoi H., Hirose Y., Iwai M., Ninomiya M., Kodera Y., Miyawaki S., Kuriyama K., Shimazaki C., Akiyama H., Nishimura M., Motoji T., Shinagawa K., Takeshita A., Ueda R., Ohno R., Emi N., Naoe T. Biologic and clinical significance of the FLT3 transcript level in acute myeloid leukemia. *Blood* 2004; 103: 1901-1908.
12. Dohner K., Tobis K., Ulrich R., Frohling S., Benner A., Schlenk RF, Dohner H. Prognostic significance of partial tandem duplications of the MLL gene in adult patients 16 to 60 years old with acute myeloid leukemia and normal cytogenetics: a study of the Acute Myeloid Leukemia Study Group Ulm. *J. Clin. Oncol.* 2002; 20: 3254-3261.
13. Steudel C., Wermke M., Schaich M., Schakel U., Illmer T., Ehninger G., Thiede C. Comparative analysis of MLL partial tandem duplication and FLT3 internal tandem duplication mutations in 956 adult patients with acute myeloid leukemia. *Genes Chromosomes. Cancer* 2003; 37: 237-251.
14. Frohling S., Schlenk RF, Krauter J., Thiede C., Ehninger G., Haase D., Harder L., Kreitmeier S., Scholl C., Caligiuri MA., Bloomfield CD., Dohner H., Dohner K. Acute myeloid leukemia with deletion 9q within a noncomplex karyotype is associated with CEBPA loss-of-function mutations. *Genes Chromosomes. Cancer* 2005; 42: 427-432.
15. Frohling S., Schlenk RF., Stolze I., Bihlmayr J., Benner A., Kreitmeier S., Tobis K., Dohner H., Dohner K. CEBPA mutations in younger adults with acute myeloid leukemia and normal cytogenetics: prognostic relevance and analysis of cooperating mutations. *J. Clin. Oncol.* 2004; 22: 624-633.
16. Barragan E., Cervera J., Bolufer P., Ballester S., Martin G., Fernandez P.,

- Collado R., Sayas MJ., Sanz MA. Prognostic implications of Wilms' tumor gene (WT1) expression in patients with de novo acute myeloid leukemia. *Haematologica* 2004; 89: 926-933.
17. Ogawa H., Ikegame K., Kawakami M., Tamaki H. WT1 gene transcript assay for relapse in acute leukemia after transplantation. *Leuk. Lymphoma* 2004; 45: 1747-1753.
  18. Ogawa H., Tamaki H., Ikegame K., Soma T., Kawakami M., Tsuboi A., Kim EH., Hosen N., Murakami M., Fujioka T., Masuda T., Taniguchi Y., Nishida S., Oji Y., Oka Y., Sugiyama H. The usefulness of monitoring WT1 gene transcripts for the prediction and management of relapse following allogeneic stem cell transplantation in acute type leukemia. *Blood* 2003; 101: 1698-1704.
  19. Tan Y., Li G., Zhao C., Wang J., Zhao H., Xue Y., Han M., Yang C. Expression of sorcin predicts poor outcome in acute myeloid leukemia. *Leukemia Research* 2003; 27: 125-131.
  20. Galimberti S., Testi R., Guerrini F., Fazzi R., Petrini M. The clinical relevance of the expression of several multidrug-resistant-related genes in patients with primary acute myeloid leukemia. *J. Chemother.* 2003; 15: 374-379.
  21. Schaich M., Koch R., Soucek S., Repp R., Ehninger G., Illmer T. A sensitive model for prediction of relapse in adult acute myeloid leukaemia with t(8;21) using white blood cell count, CD56 and MDR1 gene expression at diagnosis. *British Journal of Haematology* 2004; 125: 477-479.
  22. Valk PJ., Verhaak RG., Beijen MA., Erpelinck CA., Barjesteh van Waalwijk van Doorn-Khosrovani, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, Delwel R. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* 2004; 350: 1617-1628.
  23. Bullinger L., Dohner K., Bair E., Frohling S., Schlenk RF, Tibshirani R., Dohner H., Pollack JR. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N. Engl. J. Med.* 2004; 350: 1605-1616.
  24. Yagi T., Morimoto A., Eguchi M., Hibi S., Sako M., Ishii E., Mizutani S, Imashuku S., Ohki M., Ichikawa H. Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* 2003; 102: 1849-1856.
  25. Downing JR., Shannon KM. Acute leukemia: a pediatric perspective. *Cancer Cell* 2002; 2: 437-445.
  26. Haferlach T., Kohlmann A., Bacher U., Schnittger S., Haferlach C., Kern W. Gene expression profiling for the diagnosis of acute leukaemia. *Br J Cancer.* 2007; 96: 535-540.
  27. Haferlach T., Kohlmann A., Schnittger S., et al. Global approach to the diagnosis of leukemia using gene expression profiling. *Blood.* 2005; 106: 1189-1198.
  28. Tagliafico E., Tenedini E., Manfredini R., Grande A., Ferrari F., Roncaglia E., Bicciato S., Zini R., Salati S., Bianchi E., Gemelli C., Montanari M., Vignudelli T., Zanocco-Marani T., Parenti S., Paolucci P., Martinelli G., Piccaluga PP., Baccarani M., Specchia G., Torelli U., Ferrari S. (2006). Identification of a molecular signature predictive of sensitivity to differentiation induction in acute myeloid leukemia. *Leukemia* 20 (10) 1751-8.

## **Diagnosi molecolare dei linfomi: implicazioni cliniche**

*Silvia Franceschetti, Clara Deambrogi, Michaela Cerri, Paola Riccomagno, Stefania Cresta, Lorenzo De Paoli, Antonio Ramponi, Davide Rossi, Annarita Conconi, Daniela Capello, Gianluca Gaidano*

Dipartimento di Medicina Clinica e Sperimentale & BRMA, Divisione di Ematologia, Università del Piemonte Orientale "Amedeo Avogadro", Novara

La classificazione WHO identifica entità cliniche definite in base ad un approccio multidisciplinare, in grado di combinare dati clinici, morfologici, immunofenotipici e genetici. Tuttavia, l'approccio ai linfomi non-Hodgkin è complicato dalla estrema eterogeneità ancora esistente nell'ambito delle categorie nosologiche identificate dalla classificazione WHO. Pazienti con diagnosi apparentemente affini possono avere presentazioni cliniche, profili molecolari e prognosi significativamente diverse. Ne deriva la necessità di perseguire la ricerca di nuovi marcatori molecolari che consentano l'identificazione di subset di pazienti che possano beneficiare di approcci differenziati. Ad oggi, sono stati caratterizzati un grande numero di marcatori molecolari che hanno consentito di chiarire e definire patogenesi, criteri diagnostici e fattori prognostici nell'ambito dei linfomi non-Hodgkin.

Classicamente, la patogenesi dei linfomi non-Hodgkin viene ricondotta ad un processo multifasico che implica lesioni molecolari di proto-oncogeni e geni onco-soppressori. Nell'ambito delle lesioni molecolari dei proto-oncogeni, il principale meccanismo di deregolazione è rappresentato dalla traslocazione cromosomica. Mediante questo meccanismo, il proto-oncogene può essere allontanato dalle proprie strutture regolatorie ed essere posto sotto nuovi elementi di controllo, che tipicamente ne determinano la deregolazione della espressione. Alternativamente, la traslocazione cromosomica può determinare la formazione di un "trascritto di fusione", derivante dalla fusione dei due geni coinvolti nella rottura cromosomica. I proto-oncogeni attivati dalla formazione di proteine di fusione generano proteine chimeriche che mostrano nuove proprietà biologiche in grado di sostenere il processo di linfomagenesi.

L'inattivazione bi-allelica di geni onco-soppressori avviene classicamente per mutazione inattivante di un allele e delezione dell'altro allele. Un tipico meccanismo di acquisizione di mutazioni caratteristico dei linfomi è caratterizzato dalla "ipermutazione somatica", che fisiologicamente riguarda i geni delle immuno-

globuline, ma può estendersi in maniera aberrante (ipermutazione somatica “aberrante”) anche ad altri geni. In casi più rari, l’inattivazione bi-allelica è conseguita attraverso una doppia mutazione su entrambe gli alleli o per delezione omozigote del gene. Un ulteriore meccanismo di inattivazione dei geni onco-soppressori è rappresentato dalla metilazione delle regioni regolatorie del gene, che ne determinano una ridotta espressione.

Meccanismi addizionali coinvolti nella patogenesi dei linfomi sono rappresentati dall’infezione dell’ospite da parte di virus o microorganismi, che attraverso meccanismi diretti o indiretti creano una condizione necessaria, ma in genere non sufficiente, allo sviluppo di cloni linfoidi. Tra i virus oncogeni associati a linfoma che agiscono con meccanismi diretti, citiamo EBV, HHV-8, HTLV-1: essi presentano nel loro genoma oncogeni con capacità trasformanti che hanno un’azione iniziante o promuovente la linfomagenesi. Altri virus, quali HIV o HCV, e microorganismi come *H.pylori*, promuovono invece indirettamente la trasformazione neoplastica, causando immunodepressione, inducendo una stimolazione antigenica cronica del sistema immune, o determinando modificazioni del microambiente cellulare.

Nell’ambito dei linfomi non-Hodgkin a cellule B o T periferiche, nuove implicazioni clinico-prognostiche dei marcatori molecolari sono state segnalate principalmente nel linfoma diffuso a grandi cellule B, nel linfoma follicolare, e nel linfoma mantellare, mentre sono ancora in fase preliminare nell’ambito dei linfomi a cellule T.

### **Linfoma diffuso a grandi cellule B**

Le traslocazioni cromosomiche coinvolgenti la banda 3q27 e svariati siti cromosomici, inclusi il locus delle immunoglobuline, rappresentano una alterazione frequente nell’ambito dei linfomi diffusi a grandi cellule B (Diffuse Large B-cell Lymphoma, DLBCL). Il proto-oncogene coinvolto, definito *BCL-6*, appartiene alla famiglia dei fattori trascrizionali contenenti domini “zinc finger” ed è espresso selettivamente dalle cellule B del centro germinativo. Altre traslocazioni descritte riguardano il gene *BCL-2* e *cMYC*, anch’essi posti sotto il controllo trascrizionale dei geni regolatori del locus delle immunoglobuline. Ulteriori meccanismi patogenetici descritti nel DLBCL includono la ipermutazione somatica aberrante di proto-oncogeni, incluso *BCL-6*, la delezione del “FAS death-domain” e mutazioni inattivanti di *p53*.

Ad oggi l’analisi dell’espressione genica mediante la metodica del “gene expression profiling” (GEP), ha consentito di postulare l’esistenza di sottotipi distinti nell’ambito dei DLBCL, caratterizzati da specifiche “molecular signatures”. In particolare, i primi studi di GEP applicati al DLBCL identificarono due gruppi prognostici distinti caratterizzati da una istogenesi diversa: il primo gruppo, definito “germinal-center B-cell like” (GCB-DLBCL), riflette il GEP di cellule B del centro germinativo e presenta prognosi favorevole; il secondo gruppo, definito “activated B-cell like” (ABC-DLBCL), riflette il GEP di cellule B attivate “in vitro” e di cellule B post-centro germinativo e presenta prognosi sfavorevole.

Questi dati fanno tuttavia riferimento a casistiche trattate in era pre-rituximab; trial clinici più recenti effettuati su casistiche di pazienti trattati con regimi polichemioterapici in combinazione con rituximab sembrano non confermare la rilevanza prognostica della distinzione tra i due gruppi.

Lavori più recenti hanno analizzato casistiche di DLBCL secondo altri pannelli di espressione genica delineando tre gruppi distinti:

- 1) “oxidative phosphorylation (Ox Phos)”;
- 2) “B-cell receptor/proliferation (BCR)”;
- 3) “Host Response (HR)”.

Il primo gruppo presenta una aumentata espressione di geni che regolano la funzione mitocondriale, il trasporto degli elettroni, l'apoptosi e i proteosomi. Tipicamente, si tratta di DLBCL caratterizzati da lesioni genetiche coinvolgenti i pathways apoptotici, incluse le traslocazioni di *BCL-2* e la delezione del “FAS death domain”. Nell'ambito del secondo gruppo, viene osservata la iperespressione di geni coinvolti nella regolazione del ciclo cellulare, nella riparazione del DNA, nel “signaling” a partire dal B-cell receptor, insieme a numerosi fattori trascrizionali specifici della cellula B, incluso *BCL-6*. In particolare, traslocazioni coinvolgenti il locus di *BCL-6* sono particolarmente frequenti nel sottogruppo BCR, offrendo pertanto un potenziale target terapeutico per i nuovi peptidi inibitori di *BCL-6*. Infine, il sottogruppo HR esprime una spiccata risposta immune/infiammatoria coerente con la tipica infiltrazione da parte di cellule T e cellule dendritiche proprie dei T-cell/histiocyte-rich B-cell lymphoma.

Sottogruppo a sè stante sono i linfomi primitivi del mediastino caratterizzati da un GEP affine a quello del linfoma di Hodgkin classico, con la tipica “cytokine signature” e l'attivazione costitutiva della via di NF- $\kappa$ B, insieme alla presenza di specifiche lesioni genetiche, quali la duplicazione dei cromosomi 9p e 2p che includono il locus di *JAK2* (9p24) e di *REL* (2p16).

Poiché l'analisi GEP dei DLBCL è verosimilmente importante ai fini prognostici, ma non è praticabile nella gestione clinica di routine, studi di immunoistochimica hanno tentato di riprodurre l'impatto prognostico del GEP utilizzando “marcatori surrogati”. Questi studi hanno prevalentemente utilizzato CD10, *bcl6* e MUM1/IRF4. La loro utilità nella stratificazione prognostica dei DLBCL è comunque ancora dibattuta.

Come già rimarcato, tali studi fanno riferimento a casistiche retrospettive rituximab-naive. Recenti osservazioni cliniche in pazienti trattati con rituximab-CHOP dimostrerebbero il superamento della rilevanza prognostica di alcune precedenti osservazioni, quali l'outcome sfavorevole associato alla iperespressione di *BCL-2* o l'outcome favorevole associato alla iperespressione di *BCL-6*.

### **Linfoma follicolare**

I linfomi follicolari sono caratterizzati dalla attivazione di *BCL-2* nel 90% dei casi. A differenza della maggior parte dei proto-oncogeni, *BCL-2* non regola direttamente la proliferazione cellulare ma è preposto al controllo della morte cellulare programmata (apoptosi). L'attivazione di *BCL-2* nei linfomi follicolari



avviene tramite la traslocazione  $t(14;18)(q32, q21)$  coinvolgente il locus di *BCL-2* ed il locus delle immunoglobuline. La traslocazione deregola l'espressione di *BCL-2*, che viene pertanto espresso ad alti livelli proteggendo la cellula dall'apoptosi. La traslocazione di *BCL-2* fornisce alla cellula un vantaggio di sopravvivenza ma non è di per sé sufficiente ad indurre la linfomagenesi. I linfomi follicolari grado 3 possono più comunemente non portare la  $t(14;18)$ , ma presentare iperespressione di *BCL-2* grazie a meccanismi alternativi. Ad oggi, la maggiore rilevanza assunta dal marcatore molecolare *BCL-2* riguarda la valutazione della malattia minima residua nell'ambito di coorti di pazienti trattati con immuno-chemioterapia convenzionale o terapia ad alte dosi. Pazienti che negativizzano il marcatore dopo il trattamento presentano una prognosi favorevole in termini di sopravvivenza libera da progressione.

Nell'ambito dei linfomi follicolari, anche l'analisi di GEP ha condotto alla identificazione di profili distinti associati a diversa prognosi. Le due "signatures" GEP identificate nei linfomi follicolari sono definite "risposta immune tipo 1 (IR-1)" e "risposta immune tipo 2 (IR-2)", poichè i geni coinvolti includono loci espressi differenzialmente dalle cellule del sistema immunitario normale. Entrambe le "signatures" vengono caratterizzate nell'ambito della popolazione CD19-negativa delle biopsie di linfoma follicolare, ovvero nella frazione composta da cellule T, monocito/macrofagi ed altre cellule del microambiente. La "signature" IR-1 comprende geni ristretti alle cellule T tra cui *CD7*, *CD8B*, *ITK* e *LEF1*, e si associa a prognosi favorevole. La "signature" IR-2 comprende invece geni iperespressi da macrofagi e cellule dendritiche, tra cui *CD64*, *TLR5* e varie componenti del complemento, e si associa a prognosi sfavorevole.

Una possibile implicazione terapeutica di questa osservazione riguarda la disponibilità di anticorpi monoclonali potenzialmente attivi anche sul microambiente, tra cui l'anti-*CD80*, dotati di una potenziale azione sinergica con la chemioterapia convenzionale.

Una frazione di linfomi follicolari progredisce a DLBCL: questa trasformazione clinico-istologica si accompagna a mutazione di *p53* nel 90% dei casi, ma può anche presentare inattivazione dei geni oncosoppressori *p15* e *p16*, ipermutazione somatica aberrante di molteplici proto-oncogeni e, più raramente, attivazione di *c-MYC*.

## **Linfoma mantellare**

La traslocazione  $t(11;14)$  associata al linfoma mantellare attiva il proto-oncogene ciclina D1 (anche noto come *BCL-1*), un fattore di controllo del ciclo cellulare. Poichè la traslocazione si associa selettivamente al linfoma mantellare nello spettro dei linfomi non-Hodgkin, essa rappresenta un utile marcatore diagnostico per la diagnosi differenziale di questo linfoma.

L'analisi GEP ha rivelato più di 1.000 geni espressi differenzialmente fra i linfomi mantellari ed altri tipi di linfoma. Questi studi hanno quindi permesso di identificare una "signature" propria del linfoma mantellare, indipendentemente dalla iperespressione di ciclina D1. Ne deriva che alcuni linfomi, pur ciclina D1 negativi (<10% dei casi), esprimono la "signature" propria dei linfomi mantellari; in



alcuni di questi casi, sono coinvolte la ciclina D2 o la ciclina D3. Inoltre, l'analisi GEP ha permesso di identificare categorie prognostiche distinte di linfoma mantellare. Infine, l'analisi GEP permette il riconoscimento molecolare della variante blastoide del linfoma mantellare, caratterizzata da iperespressione di cyclin-dependent kinase (CDK) 4 e di CDC28 protein kinase 1. CDK4 si associa con ciclina D1 e favorisce la progressione del ciclo cellulare attraverso il checkpoint G1/S. L'iperespressione di CDC28 protein kinase 1 abroga l'inibizione del complesso ciclina D1/CDK4 da parte dell'inibitore CDK p27/Kip1. Più recente è l'acquisizione di un ruolo biologico dei meccanismi anti-apoptotici nella farmaco-resistenza propria del linfoma mantellare.

La comprensione della patogenesi del linfoma mantellare consente la definizione di strategie terapeutiche mirate, tra cui l'utilizzo degli oligonucleotidi antisenso o gli inibitori della istone deacetilasi in grado di ridurre i livelli di ciclina D1, gli inibitori o modulatori delle cyclin-dependent kinase (cdk), o ancora l'utilizzo di farmaci che interferiscono con i meccanismi anti-apoptotici che potrebbero essere integrati nella terapia citoriduttiva convenzionale.

### **Linfomi a cellule T-periferiche**

Nell'ambito dei linfomi T, la t(2;5) si associa elettivamente ai linfomi non-Hodgkin T a grandi cellule anaplastiche (Anaplastic Large Cell Lymphoma, ALCL). La t(2;5) costituisce uno dei pochi esempi di proteina di fusione nei linfomi non-Hodgkin, ed è costituita dai geni *NPM* e *ALK*. *NPM* codifica per una proteina nucleolare, *ALK* codifica per una tirosina chinasi normalmente espressa nelle cellule T. Altre anomalie genetiche, quali la trisomia del cromosoma 3, 5, 8 e X, le delezioni del 6q, i riarrangiamenti del 7q, la monosomia 13 o la delezione di 13q14 sono alterazioni citogenetiche ricorrenti, ma non hanno ancora condotto alla caratterizzazione dei geni implicati.

Ad oggi, l'analisi del GEP nei linfomi T ha fornito risultati preliminari che ancora necessitano di ulteriore validazione. Un primo studio di genomica ha differenziato due sottogruppi di linfomi a cellule T periferiche (Peripheral T-cell Lymphoma, PTCL): uno a prognosi favorevole, associato alla espressione dei geni della via di NF- $\kappa$ B, e un secondo gruppo a prognosi sfavorevole, associato alla scarsa espressione dei geni correlati ad NF $\kappa$ B. Proprio grazie agli studi di GEP è stato recentemente caratterizzato il ruolo di *PDGFRA* come gene potenzialmente coinvolto nella patogenesi dei PTCL e quindi come potenziale bersaglio terapeutico. Infine, un ultimo studio ha caratterizzato tre sottogruppi prognostici di PTCL in base al profilo di espressione citochinica: cattiva prognosi per il tipo associato ad espressione di CCR4, intermedia per l'espressione di CXCR3 e favorevole per il tipo associato ad espressione di CCR3.

Riguardo l'espressione di citochine tra gli istotipi meglio caratterizzati, negli studi di GEP il linfoma T angio-immunoblastico (Angioimmunoblastic T-cell Lymphoma, AILT) si associa tipicamente ad un "fenotipo Th1" caratterizzato dalla espressione di citochine quali CXCR3, TNF receptor OX40, e CXCL13. In particolare, quest'ultimo marcatore è uno tra i geni maggiormente espressi da

parte delle cellule T regolatorie del centro germinativo (CG-Th), da qui l'ipotesi che l'istogenesi del AILT sia riconducibile a questo tipo cellulare. Gli ALCL sono invece associati ad un "fenotipo Th2" caratterizzato dall'espressione delle citochine CCR3 e CCR4, e dei geni IL13R, FOS e JUNB. Si tratta comunque di modelli che necessitano di ulteriore validazione, laddove gli studi di GEP sono limitati principalmente dalla significativa proporzione di cellule "non clonali" infiltranti il clone tumorale.

### **Implicazioni cliniche e prospettive future**

Numerosi esempi dimostrano come i marcatori molecolari siano ormai strumenti validati e indispensabili nella pratica diagnostica laddove ci consentono l'identificazione di specifici tipi nosologici. Come già da tempo accade nelle leucemie acute e nella patologia mieloide, anche tra i linfomi non-Hodgkin i marcatori molecolari stanno inoltre accrescendo la loro importanza nel monitoraggio della malattia minima residua, oggi raffinata grazie alle metodiche di PCR quantitativa. Infine, i marcatori hanno un ruolo indispensabile nella costruzione di modelli prognostici che consentano di adattare la terapia al paziente. La definizione prognostica è tuttavia un "target in movimento", e l'introduzione di nuovi farmaci può significativamente alterare la rilevanza di fattori prognostici precedentemente acquisiti. Quindi, l'emergere di nuove strategie terapeutiche, che agiscono in sinergia con la chemioterapia convenzionale, necessita di sempre nuove conferme che accertino la applicabilità dei marcatori molecolari.

### **Bibliografia**

1. Alizadeh AA., Eisen MB., Davis RE., Ma C., Lossos IS., Rosenwald A., et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; 403: 503-511.
2. Buckstein R., Pennel N., Berinstein NL., et al. What is remission in follicular lymphoma, and what is its relevance? *Best Pract Res Clin Haematol.* 2005; 18: 27-56.
3. Colomo L., Lopez-Guillermo A., Perales M., et al. Clinical impact of the differentiation profile assessed by immunophenotyping in patients with diffuse large B-cell lymphoma. *Blood* 2003; 101: 78-84.
4. Dave SS., Wright G., Tan B., et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. *N Engl J Med* 2004; 351: 2159-2169.
5. Gaidano G., Dalla-Favera R. Pathobiology of non-Hodgkin's lymphomas. In *Hematology. Basic Principles and Practice*, 4th edition. Hoffman R., Benz EJ., jr., Shattil SJ., Furie B., Cohen HJ., Silberstein L., McGlave P. (eds.). Philadelphia: Elsevier, 2005; 1307-1324.
6. Hans CP., Weisenburger DD., Greiner TC., et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* 2004; 103: 275-282.

7. Hofmann WK., de Vos S., Tsukasaki K., et al. Altered apoptosis pathways in mantle cell lymphoma detected by oligonucleotide microarray. *Blood*. 2001; 98: 787-794.
8. Lossos IS., Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J Clin Oncol* 2006; 24: 995-1007.
9. Martinez-Delgado B. Peripheral T-cell lymphoma gene expression profiles. *Hematol Oncol* 2006; 24: 113-119.
10. Monti S., Savage KJ., Kutok JL., et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood*. 2005; 105: 1851-1861.
11. Mounier N., Briere J., Gisselbrecht C., et al. Rituximab plus CHOP (R-CHOP) overcomes bcl-2-associated resistance to chemotherapy in elderly patients with diffuse large B-cell lymphoma (DLBCL). *Blood*. 2003; 101: 4279-4284.
12. Nyman H., Adde M., Karjalainen-Lindsberg ML., et al. Prognostic impact of immunohistochemically defined germinal-center phenotype in diffuse large B-cell lymphoma patients treated with immunochemotherapy. *Blood*. 2007; 109: 4930-4935.
13. Piccaluga PP., Agostinelli C., Zinzani PL., Baccarani M., Dalla-Favera R., Pileri SA. Expression of platelet-derived growth factor receptor alpha in peripheral T-cell lymphoma not otherwise specified. *Lancet Oncol* 2005; 6: 440.
14. Rosenwald A., Wright G., Chan WC., Connors JM., Campo E., Fisher RI., et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large B-cell lymphoma. *N Engl J Med* 2002; 346: 1937-1947.
15. Winter JN., Weller EA., Horning SJ., et al. Prognostic significance of bcl-6 expression in DLBCL treated with CHOP or R-CHOP: a prospective correlative study. *Blood*. 2006; 107: 4207-4213.

## **Approccio genomico integrato per la classificazione molecolare del mieloma multiplo**

*Antonino Neri, Sonia Fabris, Luca Agnelli*

Dipartimento di Scienze Mediche, Università di Milano.  
Fondazione IRCCS Ospedale Maggiore Policlinico, U.O. di Ematologia 1  
e Centro Trapianti di Midollo, Milano

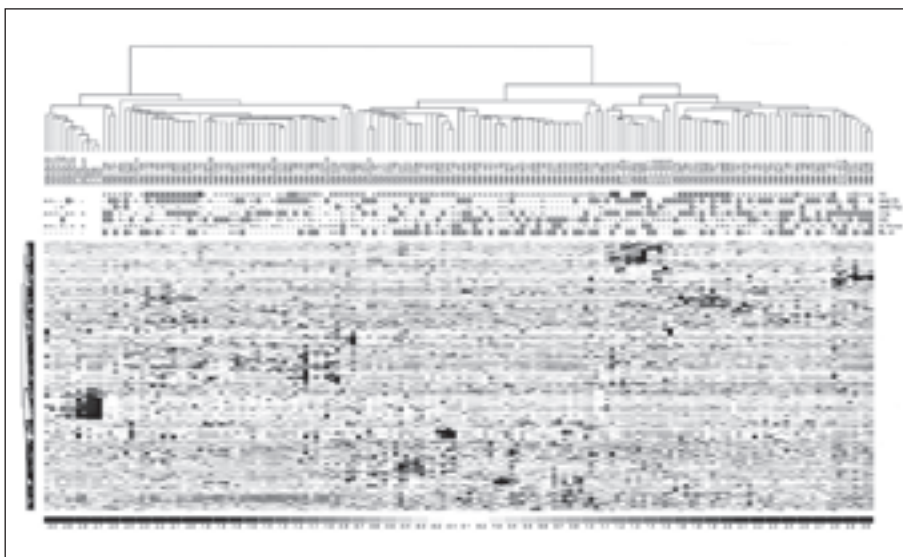
Il Mieloma Multiplo (MM) è una neoplasia incurabile dovuta alla proliferazione clonale della plasmacellula (PC) nel midollo osseo. Il MM rappresenta il 10% dei tumori ematologici ed è caratterizzato da una marcata eterogeneità clinica: può infatti presentarsi come una gammopatia monoclonale di origine incerta (MGUS, una condizione pre-maligna che può evolvere in MM conclamato nel 1% dei casi), mieloma asintomatico (SMM), mieloma conclamato e sintomatico, o forma extra-midollare (leucemia plasmacellulare, PCL).

Il MM si associa ad una marcata instabilità cariotipica con aberrazioni cromosomiche sia strutturali che numeriche (1). Circa il 50% dei casi presenta iperdiploidia, con un numero complessivo compreso tra 48 e 74 cromosomi e caratterizzata da una trisomia non-random di un gruppo di cromosomi, quali il 3, 5, 7, 9, 11, 15, 19, 21; i rimanenti casi non-iperdiploidi comprendono i casi ipodiploidi (fino a 44-45 cromosomi), pseudodiploidi (44-47 cromosomi) e ipotetraploidi/tetraploidi (75 o più cromosomi). I MM non-iperdiploidi presentano una maggiore frequenza di delezione del cromosoma 13, di extracopie o amplificazioni del cromosoma 1q nonché di traslocazioni cromosomiche coinvolgenti il locus delle catene pesanti delle immunoglobuline (IgH), sul cromosoma 14q32, che determinano la deregolazione dei geni CCND1 (11q13), CCND3 (6p21.1), MAF (16q24), MAFB (20q12) o FGFR3/MMSET (4p16.3) (2). Dai dati disponibili in letteratura, la presenza di specifiche alterazioni strutturali o numeriche sembra correlare con la prognosi; in particolare, lo stato di iperdiploidia sembra associarsi ad una migliore sopravvivenza, mentre la presenza della delezione del 13, extracopie di 1q e la traslocazione t(4;14) o quelle coinvolgenti i fattori trascrizionali Maf identificherebbero gruppi di pazienti con prognosi sfavorevole (2, 3). I recenti sviluppi tecnologici hanno consentito di affiancare alla citogenetica convenzionale altre metodologie di indagine, permettendo una migliore definizione delle alterazioni molecolari presenti nella malattia. In particolare, oltre alle ormai acquisite analisi mediante Ibridazione Fluorescente In-situ (FISH), Ibridazione

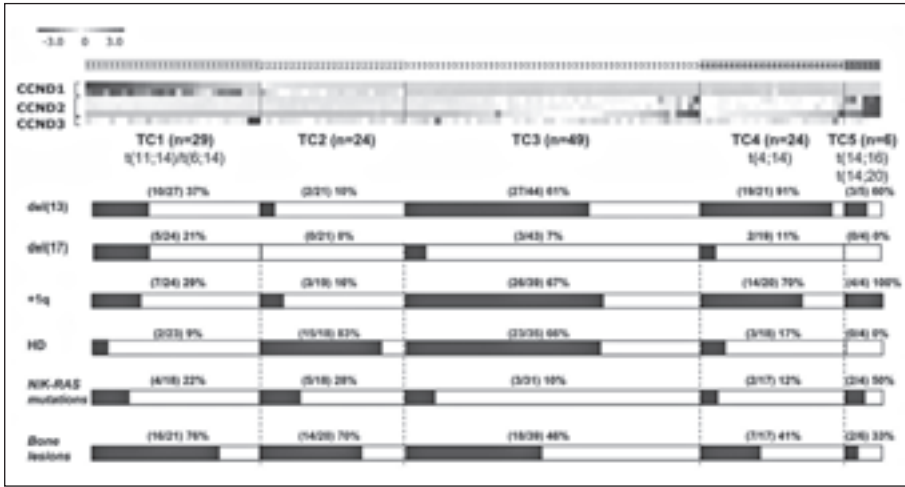
Genomica Comparativa (CGH) e PCR quantitativa, gli studi mediante nanotecnologie hanno incrementato la possibilità di classificare i MM su base molecolare, per una migliore definizione della prognosi della malattia. Lo studio dei profili di espressione genica globale (GEP) su microarray ad oligonucleotidi, condotti anche nel nostro ad altri laboratori hanno identificato patterns trascrizionali caratteristici, quali quelli associati alla deregolazione dei geni target delle traslocazioni cromosomiche, all'espressione delle Cicline D, alla delezione del cromosoma 13, all'iperdiploidia, alla presenza di extracopie di 1q, così come a differenti fattori prognostici (4-10).

In Figura 1 è riportata una analisi non supervisionata di espressione genica di PC CD138+ purificate da controlli normali, MGUS, MM e PCL che dimostra come il profilo di espressione genica delle plasmacellule mielomatose è diverso da quello delle plasmacellule normali ma non da quello delle plasmacellule purificate da pazienti con MGUS (Fig. 1) (10).

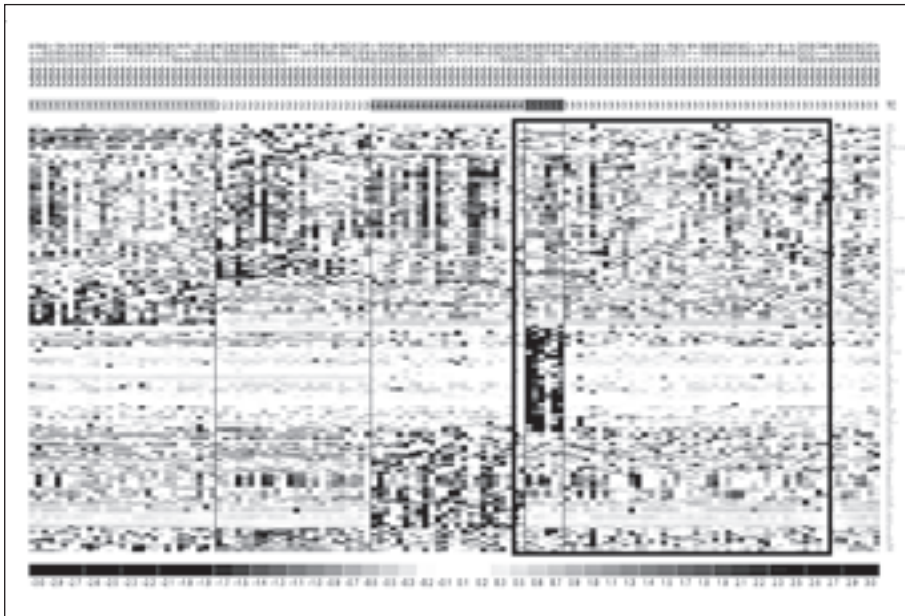
Inoltre, nell'ambito dei MM si distinguono differenti sottogruppi, alcuni dei quali non identificabili con le più comuni traslocazioni cromosomiche. È stata recentemente proposta una classificazione molecolare del MM basata sulla presenza delle traslocazioni e dell'espressione delle cicline (TC classification). In Figura 2 sono riportate le caratteristiche molecolari ed il profilo di espressione genica dei 5 gruppi TC su un pannello di 132 casi di MM del nostro laboratorio.



**Fig. 1** - Analisi non supervisionata dei profili di espressione genica di PC CD138+ purificate da 4 controlli normali, 12 MGUS, 132 MM e 9 casi di PCL. Il dendrogramma è stato generato usando un algoritmo di "clustering" gerarchico basato sul metodo di "linkage" centroide. Nella matrice ogni colonna rappresenta un campione e ogni riga un gene. I campioni sono raggruppati in accordo coi livelli di espressione dei geni maggiormente variabili. La barra con la scala di colori rappresenta i cambiamenti relativi di espressione genica normalizzati dalla deviazione standard. Sono indicate le caratteristiche cliniche e molecolari dei pazienti e i gruppi TC (vedi Figura 2). I pazienti positivi per le delezioni 13q e 17p, amplificazione 1q, stato iperdiploide, mutazioni dei geni RAS e lesioni ossee (BL, bone lesions) sono identificati con il simbolo +.

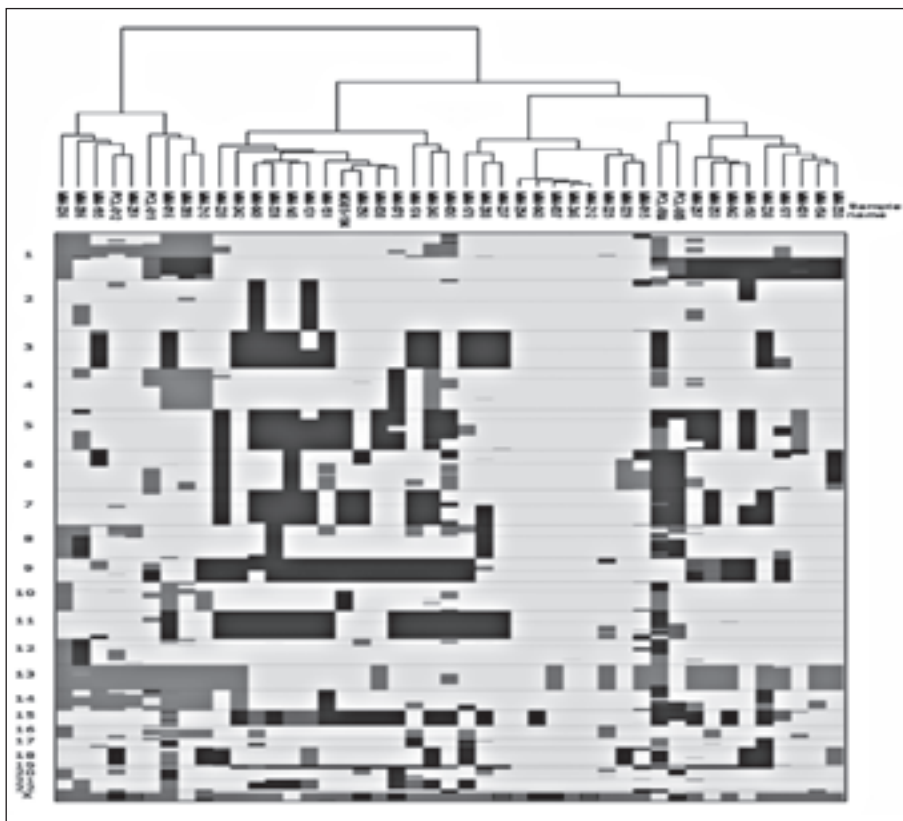


**Fig. 2** - Alterazioni genetiche e caratteristiche patologiche nei gruppi TC stratificati secondo i livelli di ciclina D1,2 e 3 e traslocazioni del locus IGH. Il gruppo TC1 è caratterizzato dalla t(11;14) o (6;14) ed over-espressione rispettivamente della ciclina D1 o D2; il gruppo TC2 da assenza di traslocazioni note e moderata espressione di ciclina D1; il gruppo TC3 da espressione della ciclina D2, i gruppi TC4 e TC5 dall'espressione della ciclina D2 e rispettivamente dalla traslocazione t(4;14) e da quelle coinvolgenti i geni *MAF*. I pazienti sono stati caratterizzati per le principali anomalie genetiche, incluse le delezioni 13q e 17p e l'amplificazione 1q, lo stato iperdiploide (HD), le mutazioni dei geni *RAS*.



**Fig. 3** - Analisi multi-classe dei cinque gruppi di MM stratificati in accordo alla Classificazione TC. I gruppi 1-4 presentano uno specifico profilo di espressione genica che non è evidenziabile nei pazienti TC3 (riquadro nero).





**Fig. 4** - Analisi non supervisionata dei profili genomici di PC CD138+ di 45 pazienti con MM. La mappa genomica dei pazienti, ordinati in base al dendrogramma, descrive le alterazioni individuate mediante GeneChip<sup>®</sup> Affymetrix Human Mapping 50k XbaI. In asse Y è indicata la localizzazione cromosomica; le linee tratteggiate rappresentano la localizzazione del centromero. La scala colorimetrica del numero di copie inferite rappresenta delezioni bialleliche (verde chiaro), perdite (verde scuro), normalità (grigio), guadagni (rosso scuro), amplificazioni (rosso chiaro).

In particolare abbiamo dimostrato che il gruppo TC2 (che presenta moderati livelli di ciclina D1 in assenza di traslocazioni cromosomiche note) è caratterizzato dall'overespressione di geni coinvolti nella biosintesi delle proteine a livello traduzionale e si associa alla iperdiploidia (Fig. 3) (6, 7).

Parallelamente, l'analisi mediante microarray ha consentito lo studio del profilo genomico. Al riguardo, nel MM sono stati compiuti studi utilizzando array-CGH (11), consentendo una prima definizione ad alta risoluzione delle caratteristiche genomiche della neoplasia. L'utilizzo di SNP (Single Nucleotide Polymorphism) microarray con oligonucleotidi ad alta densità, introdotti più recentemente, consente lo studio di anomalie cromosomiche numeriche (guadagni o perdite) ad una risoluzione inferiore a 100 Kb, superiore rispetto alla citogenetica convenzionale e a quella degli array-CGH. Recenti studi su linee cellulari e tumori primari di MM condotti nel nostro ed altri laboratori hanno dimostrato che gli SNP micror-

ray possono essere utilizzati per individuare sbilanciamenti allelici e perdita di eterozigosi e la loro associazione con le caratteristiche molecolari della neoplasia precedentemente descritte (12, 13). Inoltre, questi studi hanno evidenziato la possibilità di combinare l'analisi genomica con i profili trascrizionali, per identificare geni la cui espressione risulta alterata in associazione a specifiche lesioni genomiche.

I nostri studi hanno riguardato l'analisi dei profili genomici di campioni di MM mediante ibridazione su GeneChip® Mapping 50k Array, basati sulla tecnologia degli SNP microarray. I dati ottenuti al momento mostrano una elevata eterogeneità nella distribuzione delle alterazioni nei diversi pazienti, e consentono di definire alterazioni ricorsive e caratterizzanti i vari sottogruppi, presenti ad alte frequenze (guadagno dei cromosomi coinvolti nell'iperdiploidia; monosomia del cromosoma 13, delezione del braccio 1p, guadagno del braccio 1q) nonchè caratterizzanti in maniera contestuale i singoli pazienti (quali ad esempio la presenza contemporanea di del (4p) e del (1p)). Un'analisi non supervisionata rivela la consistenza dei sottogruppi caratterizzati da iperdiploidia (gruppo 1); guadagni ricorrenti nei cromosomi 3, 5, 7, 9, 11, 15, 19, 21) ed alta frequenza di del (13) e +1q (gruppo 2); ed alta frequenza di del (1p), del (4p), del (13), del (14), del (16), del (22) (gruppo 3).

L'identificazione dei microRNA (miRNA), brevi sequenze di RNA non-codificanti e regolatorie dell'attività genica, ha aggiunto un ulteriore livello di complessità alla biologia del cancro. I miRNA sono infatti spesso localizzati su regioni frequentemente coinvolte in alterazioni cromosomiche, ed è stato inoltre dimostrato che queste, così come eventi epigenetici, possono contribuire all'alterata espressione di miRNA, con la conseguente deregolazione di geni sottoposti al loro controllo. Queste osservazioni suggeriscono che i miRNA possano avere un ruolo importante nella patogenesi tumorale (14, 15), ed una loro espressione sbilanciata è stata dimostrata in diversi tumori inclusi quelli ematologici (16). Al momento, non sono ancora presenti dati al riguardo sul MM; tuttavia il MM, per le sue peculiari caratteristiche genotipiche (la cui complessità presenta analogie con i tumori solidi), rappresenta un buon modello per lo studio della loro possibile deregolazione. Le potenzialità di tale studio sono incrementate dall'integrazione dei dati derivati dai profili trascrittomico e genomico ottenuti con microarrays.

Il primo livello di indagine, nel nostro laboratorio, ha riguardato l'analisi di espressione dei miRNA in linee di mieloma. Analogamente a quanto da noi descritto per l'espressione genica (13), a fronte dell'elevata eterogeneità ciascuna linea presenta un profilo di espressione di miRNA peculiare; l'espressione dei microRNA non sembra essere associata alla presenza nelle linee di alcuna delle traslocazioni note, mentre è in corso di indagine la possibile deregolazione conseguente ai frequenti sbilanciamenti allelici osservati nelle linee di mieloma.

In conclusione, le informazioni derivanti da un'analisi integrativa compiuta ai diversi livelli sopra esposti possono contribuire a meglio definire il network patogenetico del MM e conseguentemente identificare nuovi potenziali targets terapeutici. Un aspetto di fondamentale importanza in questo contesto è legato all'ap-



plicazione di approcci statistici e matematici che siano in grado di integrare le informazioni provenienti dalle diverse analisi, e ancor più in grado di gestire la complessità di tali informazioni per riportare il dato matematico al contesto biologico. A questo scopo, sono stati recentemente sviluppati nuovi algoritmi specifici sia per l'analisi integrata dei dati generati da microarrays, sia per investigare networks di interazione genica con putativa funzione di regolazione della fisiologia cellulare (17-20).

## **Bibliografia**

1. Seidl S., Kaufmann H., Drach J. New insights into the pathophysiology of multiple myeloma. *Lancet Oncol.* 2003; 4: 557-564.
2. Fonseca R., Barlogie B., Bataille R. et al. Genetics and cytogenetics of multiple myeloma: a workshop report. *Cancer Res.* 2004; 64: 1546-1558.
3. Bergsagel PL., Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *J Clin Oncol.* 2005; 23: 6333-6338.
4. Shaughnessy JD., Jr., Zhan F., Burington BE. et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood.* 2007; 109: 2276-2284.
5. Zhan F., Huang Y., Colla S. et al. The molecular classification of multiple myeloma. *Blood.* 2006; 108: 2020-2028.
6. Agnelli L., Bricciato S., Mattioli M. et al. Molecular classification of multiple myeloma: a distinct transcriptional profile characterizes patients expressing CCND1 and negative for 14q32 translocations. *J Clin Oncol.* 2005; 23: 7296-7306.
7. Agnelli L., Fabris S., Bricciato S. et al. Upregulation of translational machinery and distinct genetic subgroups characterise hyperdiploidy in multiple myeloma. *Br J Haematol.* 2007; 136: 565-573.
8. Agnelli L., Bricciato S., Fabris S. et al. Integrative genomic analysis reveals distinct transcriptional and genetic features associated with chromosome 13 deletion in multiple myeloma. *Haematologica.* 2007; 92: 56-65.
9. Fabris S., Ronchetti D., Agnelli L. et al. Transcriptional features of multiple myeloma patients with chromosome 1q gain. *Leukemia.* 2007; 21: 1113-1116.
10. Mattioli M., Agnelli L., Fabris S. et al. Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. *Oncogene.* 2005; 24: 2461-2473.
11. Carrasco DR., Tonon G., Huang Y. et al. High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. *Cancer Cell.* 2006; 9: 313-325.
12. Walker BA., Leone PE., Jenner MW. et al. Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. *Blood.* 2006; 108: 1733-1743.
13. Lombardi L., Poretti G., Mattioli M. et al. Molecular characterization of

- human multiple myeloma cell lines by integrative genomics: insights into the biology of the disease. *Genes Chromosomes Cancer*. 2007; 46: 226-238.
14. Calin GA., Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer*. 2006; 6: 857-866.
  15. Calin GA., Croce CM. MicroRNAs and chromosomal abnormalities in cancer cells. *Oncogene*. 2006; 25: 6202-6210.
  16. Calin GA., Ferracin M., Cimmino A. et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med*. 2005; 353: 1793-1801.
  17. Friedman N. Inferring cellular networks using probabilistic graphical models. *Science*. 2004; 303: 799-805.
  18. Basso K., Margolin AA., Stolovitzky G. et al. Reverse engineering of regulatory networks in human B cells. *Nat Genet*. 2005; 37: 382-390.
  19. Callegaro A., Basso D., Bicciato S. A locally adaptive statistical procedure (LAP) to identify differentially expressed chromosomal regions. *Bioinformatics*. 2006; 22: 2658-2666.
  20. Margolin AA., Nemenman I., Basso K. et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. 2006; 7 Suppl 1: S7.
  21. Olshen AB., Venkatraman ES., Lucito R., Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004; 5: 557-572.
  22. Chen C., Ridzon DA., Broomer AJ. et al. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*. 2005; 33: 179-185.

## **Meccanismi molecolari e cellulari della trasformazione leucemica nelle SMD**

*Paolo Bernasconi, Marina Boni, Paola Maria Cavigliano, Rita Zappatore,  
Irene Dambroso, Ilaria Giardini, Barbara Rocca, Silvia Calatroni,  
Paolo Tarantino, Marilena Caresana*

Divisione di Ematologia, Fondazione IRCCS Policlinico San Matteo, Università di Pavia

Le sindromi mielodisplastiche (SMD) sono disordini clonali della cellula staminale ematopoietica (CSE) caratterizzati da emopoiesi inefficace e citopenia di grado variabile nel sangue periferico (Shadduck et al., 2007). L'incidenza delle SMD è in continuo aumento per il progressivo invecchiamento della popolazione. Negli Stati Uniti si stimano 15,000-20,000 nuovi casi all'anno con un'incidenza complessiva di circa 55,000 nuovi casi anno. Corey et al. (2007) hanno riportato una frequenza annuale di 2-12/100.000 nuovi casi, ma di 50/100.000 nuovi casi quando vengano analizzati soggetti di età  $\geq 70$  anni. Le SMD possono insorgere improvvisamente (forme *de novo*) o dopo esposizione a carcinogeni ambientali o a chemio- radioterapia eseguita per una precedente neoplasia (forme secondarie, t-MDS).

Con le leucemie acute mieloide (LAM) le SMD non solo presentano importanti somiglianze ma anche rilevanti differenze biologiche. Le SMD non possono essere semplicemente considerate preleucemie. Infatti, solo in una minoranza di pazienti con SMD, circa un terzo, la storia naturale della malattia consiste nell'evoluzione da un quadro di citopenia periferica rappresentato da anemia, piastrinopenia, causa di facili e frequenti sanguinamenti, e leucopenia, causa di frequenti infezioni, variamente associate ad un quadro di leucemia acuta, più spesso mieloide (LAM). Man mano che la malattia progredisce aumenta la percentuale di cellule indifferenziate (blasti) nel midollo osseo come pure l'incidenza di anomalie cromosomiche clonali ed il quadro clinico diventa del tutto sovrapponibile a quello di una LAM. La maggioranza dei pazienti con SMD però non evolve in LAM, ma muore di complicanze legate alla progressione della citopenia periferica conseguente all'insufficienza midollare.

Sia nelle SMD che nelle LAM il difetto genetico responsabile della malattia probabilmente si sviluppa in una cellula dotata della capacità di auto-mantenersi, proprietà ad esclusivo appannaggio della cellula staminale ematopoietica (CSE). Tale capacità è assolutamente necessaria per garantire la persistenza nel tempo della popolazione displastica/leucemica. È stato quindi suggerito che la stamina-

lità della cellula bersaglio dell'evento neoplastico ed il difetto genico siano ugualmente importanti per la patogenesi delle SMD e delle LAM. Tuttavia dati recenti sembrano indicare che alcune lesioni molecolari fanno sì che la cellula colpita mantenga il proprio immunofenotipo ed il proprio profilo genico, ma attivi una "leukemia self-renewal-associated signature" senza riprogrammare totalmente la propria espressione genica. Le lesioni geniche presenti nelle SMD sono diverse da quelle presenti nelle LAM. Infatti, la sovra-espressione o il riarrangiamento di EVI1, le mutazioni di NPM1, la fusione genica NUP98-HOXD13 e l'"haplo-insufficienza" del gene RPS14 sono lesioni capaci di ricreare il fenotipo mielodisplastico, mentre la cooperazione tra mutazioni di classe I e mutazioni di classe II sono capaci di ricreare il fenotipo leucemico. Le mutazioni osservate nelle SMD aumenteranno sicuramente le nostre conoscenze circa la patogenesi della malattia perché permetteranno di superare le difficoltà sinora incontrate dal trapianto di cellule midollari mielodisplastiche in modelli murini (xenotrapianti). Una delle principali differenze tra SMD e LAM è rappresentata dal ruolo svolto dal microambiente nella patogenesi delle due malattie. Nelle SMD si tratta di un ruolo rilevante come dimostrano la difficoltà di ottenere sostegni stromali a partire da cellule mononucleate di midollo osseo, l'aumentata produzione di interleuchina 6 e di TNF- $\alpha$  da parte dei fibroblasti e macrofagi midollari, la variabile espressione di metalloproteinasi 9 (MMP-9) da parte dei monociti. Viceversa, nelle LAM il ruolo del microambiente midollare è del tutto marginale e si ritiene che la malattia sia determinata da una competizione tra cellula normale e cellula leucemica per la stessa nicchia ematopoietica.

Dati sperimentali suggeriscono che nelle SMD, specialmente nelle forme a basso rischio, prevalgono le vie di segnale che inducono un aumento dei livelli di apoptosi, mentre nelle LAM prevalgono le vie di segnale che causano un aumento della sopravvivenza cellulare. Man mano che le SMD a basso rischio evolvono in quelle ad alto rischio ed in LAM si osserva una progressiva riduzione dell'apoptosi ed una progressiva attivazione delle vie di segnale che favoriscono la sopravvivenza cellulare. Le SMD avanzate e le LAM non solo condividono varie vie di segnale, la più importante delle quali è quella presieduta dal "nuclear factor kappa B" (NF- $\kappa$ B), ma differiscono anche per varie vie di segnale. La conoscenza di tutte queste vie potrebbe tradursi in un miglior uso dei farmaci oggi disponibili con l'identificazione di quelli più specifici per quel particolare sottotipo di SMD o citotipo di LAM. Nella presente trattazione analizzeremo dapprima le difficoltà incontrate nella dimostrazione della cellula staminale mielodisplastica, le lesioni molecolari delle SMD e delle LAM, i meccanismi apoptotici delle SMD e successivamente le vie di trasformazione delle SMD in LAM.

### **Cellula staminale mielodisplastica**

Negli anni 50-60 Till e McCulloch suggerirono che l'ematopoiesi normale fosse organizzata in modo gerarchico e fosse quindi sostenuta da un ristretto numero di CSE. Successivamente, studi sperimentali condotti in modelli murini indicarono che l'esistenza della CSE poteva essere confermata solo dai risultati ottenuti *in*

*vivo* ed a lungo termine dei trapianti di sospensioni cellulari midollari. Negli anni successivi l'infusione nel topo "non-obese diabetic severe combined immunodeficient" (NOD/SCID) di sospensioni cellulari midollari ottenute da soggetti sani e leucemici (xenotrapianti) permise di dimostrare l'effettiva esistenza non solo della CSE normale ma anche di quella leucemica (CSL). Infatti, per quanto concerne quest'ultima esperimenti di purificazione cellulare e di marcatura clonale hanno permesso di stabilire che l'ematopoiesi leucemica, similmente a quella normale, è sostenuta da classi eterogenee di cellule dotate di distinti profili funzionali e che solo una piccola frazione di tali cellule è in grado di generare leucemia se trapiantata nel topo (Lapidot et al., 1994; Bonnet & Dick, 1997). Viceversa, mancano a tutt'oggi prove a favore dell'esistenza della cellula staminale mielodisplastica a causa dei costanti insuccessi degli xenotrapianti con progenitori emopoietici ottenuti da pazienti affetti da SMD. Il mancato attecchimento dello xenotrapianto dipende non solo dagli elevati livelli di apoptosi dei progenitori emopoietici mielodisplastici, dimostrati dai risultati delle colture *in vitro* e dall'impossibilità di creare linee cellulari mielodisplastiche, ma probabilmente anche da un difettoso funzionamento del microambiente midollare, dimostrato dall'impossibilità di generare supporti stromali a partire dalle cellule mononucleate di pazienti affetti da SMD (Manakova et al., 2000). Vari studi hanno poi sottolineato la difficoltà di espandere *in vitro* cellule prelevate a pazienti con SMD, vuoi perché i progenitori granulocito-macrofagici, eritroidi e megacariocitopoietici danno frequentemente origine a colonie anormali, predittive di cattiva prognosi, vuoi perché nelle colture a lungo termine di midollo osseo si osserva una significativa riduzione dei progenitori più immaturi misurata come formazione di colonie a 5 settimane. Inoltre, il fatto che la sospensione cellulare midollare impiegata contenga progenitori mielodisplastici accanto a progenitori normali rende ancora più problematica l'interpretazione dei risultati forniti dalle colture *in vitro* (Corey et al., 2007).

Recentemente, l'impiego di topi NOD/SCID maggiormente immunodeficienti a causa di un deficit di beta2microglobulina ha parzialmente migliorato l'attecchimento degli xenotrapianti condotti con progenitori mielodisplastici (Thanopoulou et al., 2004). Tuttavia, anche in questi topi il tessuto ematopoietico midollare, inizialmente ricco di cellule granulopoietiche e con persistenza sino a 16 settimane di cellule marcate dall'anomalia cromosomica clonale, viene con il passare del tempo ripopolato dalle cellule normali residue che si espandono e hanno progressivamente il sopravvento sulle cellule displastiche. I due soli xenotrapianti che hanno mostrato una discreta persistenza nel tempo sono stati quelli che avevano utilizzato una sospensione cellulare ottenuta da pazienti affetti anemia refrattaria con eccesso di blasti in trasformazione (RAEB-t), sottotipo FAB di SMD più strettamente imparentato sul piano biologico con le LAM.

### **Lesioni geniche nelle SMD e nelle LAM**

Anomalie cromosomiche clonali si osservano nel 40-60% delle SMD *de novo*, nel 50-60% delle LAM *de novo* e nel 70-90% delle SMD/LAM secondarie (Olney &

LeBeau, 2001; Byrd et al., 2001; Bernasconi et al., 2006). Bisogna però sottolineare che le anomalie citogenetiche delle SMD de novo sono diverse da quelle delle LAM de novo e questo fa supporre che i meccanismi patogenetici responsabili delle due malattie siano diversi. Nelle SMD de novo prevalgono le delezioni cromosomiche ed i difetti cromosomici numerici che si osservano rispettivamente nel 50% e nel 10% dei pazienti, mentre nelle LAM de novo prevalgono le traslocazioni cromosomiche che si osservano nel 40% dei pazienti. L'ipotesi che SMD e LAM de novo abbiano una diversa patogenesi è ulteriormente rafforzata dal fatto che l'instabilità cromosomica, responsabile dello sviluppo di anomalie non clonali nel 10% dei pazienti e di cloni non relati nel 5% dei pazienti, è un evento relativamente frequente nelle SMD de novo mentre raro nella LAM de novo. L'importante ruolo dell'instabilità cromosomica nella patogenesi delle SMD è ulteriormente avvalorato dalla comparsa di nuove anomalie cromosomiche nei pazienti con 5q- responsivi alla terapia con lenalidomide.

Anche i dati molecolari sembrano sottolineare una diversa patogenesi delle due malattie. Infatti, è stato proposto che per lo sviluppo di una LAM siano necessari due tipi di mutazioni, indicate come mutazioni di classe I e di classe II (Deguchi & Gilliland, 2002). Le prime hanno come bersaglio geni responsabili della trasduzione del segnale (FLT3, RAS, KIT, ecc), mentre le mutazioni di classe II hanno come bersaglio geni che codificano per fattori trascrizionali (AML1, RARA, EVI1, WT1, ecc.). La cooperazione tra questi due tipi di mutazione, inizialmente descritta nel 34% dei pazienti con leucemia acuta promielocitica (cooperazione tra FLT3 e PML-RARA), è stata poi confermata anche in pazienti con altri citotipi FAB di LAM. Infatti, mutazioni di classe I sono state descritte nella maggior parte dei pazienti che all'analisi cromosomica presentavano traslocazioni cromosomiche, mentre mutazioni di classe I e di classe II sono state identificate nei pazienti che all'analisi cromosomica mostravano un cariotipo normale (Baldus et al., 2007). Viceversa, non vi sono dati al riguardo per quanto concerne le SMD. Infatti, si è sempre ritenuto che la lesione genica rilevante per la patogenesi di questo disordine onco-ematologico fosse costituita dalla perdita di un gene onco-soppressore seguita dalla perdita o dalla mutazione dell'altro allele situato sul cromosoma omologo apparentemente normale all'analisi cromosomica convenzionale. Tuttavia, a tutt'oggi nessun gene oncosoppressore è stato individuato all'interno delle regioni di minima comune delezione definite grazie all'impiego di metodiche di citogenetica molecolare (FISH) e di biologia molecolare. La sola delezione per cui è valida quest'ipotesi è quella della banda p13 del cromosoma 17 che determina la perdita di un allele del gene p53 seguita nel 70% dei pazienti dalla perdita o dalla mutazione dell'altro allele situato sul cromosoma 17 omologo apparentemente normale all'analisi cromosomica convenzionale. Dati molto recenti indicano invece che uno dei possibili meccanismi patogenetici della SMD possa essere l'"haploinsufficienza" di geni che codificano per proteine ribosomiali. Infatti è stato dimostrato da studi di "RNA interference" che l'"haploinsufficienza" per il gene RPS14 determina un blocco nel "processing" del RNA pre-ribosomiale e nella formazione della sub-unità ribosomiale 40S ricreando le caratteristiche biologiche e cliniche della sindrome del 5q-

(Ebert et al., 2007). Alternativamente, l'instabilità genica e l'elevata incidenza di delezioni cromosomiche nelle SMD potrebbero suggerire un difettoso funzionamento dei geni che presiedono al riparo del DNA. Tra questi il gene della nucleofosmina 1 (NPM1) sembra svolgere un ruolo importante dal momento che controlla la progressione nel ciclo cellulare e la risposta allo stress della cellula staminale e dei progenitori emopoietici. Nel 2005 Grisendi e collaboratori hanno creato topi che presentavano un numero diverso di alleli NPM1 funzionanti e topi ipomorfici (hy). I topi NPM1<sup>-/-</sup> e i topi NPM1<sup>hy/hy</sup> morivano di anemia e di difetti dell'organogenesi tra le giornate E11.5 ed E16.5 di vita embrionale, mentre i topi NPM1<sup>+/-</sup> sviluppavano un disordine ematologico che somigliava ad una sindrome mielodisplastica. Li collaboratori (2006) hanno dimostrato che il gene NPM1 induce la proliferazione della CSE favorendo la transizione dalla fase G1 alla fase S del ciclo cellulare, blocca i regolatori negativi del ciclo cellulare riduce l'espressione di p53 e di p19<sup>ARF</sup>. Quest'ultima azione probabilmente protegge la CSE dalla morte per apoptosi ed è quindi necessaria per la sua integrità genomica e per la sua sopravvivenza. La bassa espressione di p53 potrebbe favorire la progressione nel ciclo cellulare impedendo il riparo di un eventuale danno genomico o, in alternativa, potrebbe favorire la sopravvivenza della cellula permettendo che essa ripari un possibile danno genomico. Tuttavia, i dati di Li collaboratori (2006) sembrano contraddire quest'ultima possibilità. Comunque sia, anche nelle SMD non è possibile escludere l'esistenza di una cooperazione tra mutazioni di classe I e di classe II dal momento che mutazioni nella via di segnale presieduta da RAS, nei geni che regolano il ciclo cellulare e nei geni che codificano per fattori trascrizionali si osservano frequentemente nelle SMD a cariotipo normale/anormale al momento della progressione in LAM.

Una cooperazione tra mutazioni di classe I e di classe II è stata recentemente confermata da Pedersen-Bjergaard et al. (2006) anche nelle SMD/LAM secondarie. È stato anzi dimostrato che si osserva sempre un solo tipo di anomalia cromosomica, le mutazioni di AML1 non si associano mai ad altre traslocazioni cromosomiche, le mutazioni di FLT3 non si associano mai a mutazioni di RAS, BRAF, cKIT e PTPN11 e, fatto ancora più importante, le mutazioni di AML1 e di RAS si sviluppano al momento della trasformazione della SMD in LAM.

Tuttavia, nonostante tutte queste evidenze a sostegno di una cooperazione tra più mutazioni, vi sono dati che indicano che alcune lesioni geniche sono da sole in grado di causare il fenotipo leucemico o mielodisplastico. Nella LAM i difetti genici dotati di questa caratteristica sono i geni di fusione TLS-ERG, MOZ-TIF MLL-ENL e MLL-AF9. Un recente studio ha dimostrato che il gene di fusione MLL-AF9 introdotto in progenitori emopoietici già commissionati in senso mieloidale è in grado di generare leucemia nel topo. Inoltre, le CSL di questi topi sono in grado di trasferire leucemia in topi riceventi. Anche il fenotipo mielodisplastico sembra poter essere riprodotto da alcuni difetti genici specifici. Tra questi bisogna ricordare la forzata espressione di EVI1 e l'espressione del gene di fusione NUP98-HOXD13. La sovra-espressione di EVI1 determina in modelli murini un'eccessiva proliferazione delle cellule del midollo osseo con ridotta espressione dei geni che controllano la differenziazione eritroide e la formazione delle pia-



strine e quindi una profonda e fatale citopenia (Buonamici et al., 2004). Perciò, in questi topi la sovra-espressione del gene *EVII* causa una malattia del tutto sovrapponibile alla SMD anche se inaspettatamente non si ha la progressione in LAM. Viceversa, topi transgenici per il gene di fusione *NUP98-HOXD13* sviluppano una malattia che riproduce fedelmente tutte le caratteristiche più rilevanti della SMD, cioè la citopenia periferica, la displasia midollare, l'apoptosi e la trasformazione in LAM (Lin et al., 2005). Questi topi sviluppano quindi una SMD sempre fatale poiché muoiono di complicanze determinate o dalla citopenia o dall'evoluzione in LAM.

### **Apoptosi e difetti citogenetici/molecolari nelle SMD**

È stato osservato che l'apoptosi non solo compromette la produzione di cellule emopoietiche mature, ma induce anche lesioni geniche nelle cellule che riescono a superare il processo di morte cellulare programmata. Infatti, è stato osservato che le nucleasi effettrici dell'apoptosi possono causare traslocazioni cromosomiche favorenti lo sviluppo di leucemia (pro-leucemogeniche) e che le cellule con questi difetti cromosomici possono sfuggire all'apoptosi e sopravvivere (Betti et al., 2003). Il più studiato bersaglio di clivaggio apoptotico sito-specifico è contenuto all'interno del gene *MLL*, corrisponde alla regione telomerica del *BCR*, la regione che contiene tutti i punti di rottura delle traslocazioni osservate nelle MDS/AML secondarie, e co-localizza con un sito di ipersensibilità alla DNAasi. Altri siti sensibili alla nucleasi sono contenuti nei geni *AF9*, *TEL* e *AML1*. Betti et al., (2003) hanno suggerito tre fasi nello sviluppo delle traslocazioni pro-leucemogeniche indotte dall'apoptosi. In una prima fase il programma apoptotico è da solo sufficiente a creare rotture nella doppia catena all'interno dei geni *MLL* e *AF9*, successivamente le rotture scompaiono per creazione di un gene di fusione *MLL-AF9*, quindi le cellule con la traslocazione *MLL-AF9* prodotta dall'apoptosi possono dividersi. Sarebbe questa la ragione per cui la traslocazione *MLL-AF9* non è presente unicamente in cellule neoplastiche ma anche in cellule normali che verrebbero eliminate da meccanismi fisiologici e che comunque diverrebbero leucemiche solo dopo aver acquisito ulteriori lesioni geniche (Higuchi et al., 2002).

Altri studi hanno riportato che sarebbe il difetto citogenetico stesso a determinare la suscettibilità della cellula displastica all'apoptosi (Sloand et al., 2007). Questo evento si verificherebbe soprattutto nei pazienti con trisomia 8 (+8) isolata. Rispetto alle cellule con monosomia 7, le cellule con +8 esprimono Fas e annessina V a livelli più alti delle cellule con monosomia 7. Inoltre, la popolazione trisomica contiene un numero significativamente maggiore di cellule CD34+ in apoptosi rispetto alla popolazione monosomica. Siccome le cellule con la trisomia sono circondate da linfociti T attivati ed i pazienti con tale trisomia rispondono bene alla terapia immuno-soppressiva è stato proposto che le cellule con +8 presentino un neoantigene capace di evocare una reazione immunologica. Questa possibilità è stata confermata da uno studio di citometria a flusso che ha individuato nei pazienti con +8 una significativa espansione di cellule T CD8+ e



ha ipotizzato che il bersaglio della reazione immunologica sia la proteina codificata dal gene WT1. Infatti, le cellule CD34+ trisomiche mostrano livelli di mRNA codificante per WT1 e livelli di proteina significativamente più alti rispetto a quelli delle cellule CD34+ normali. WT1 evoca una risposta immunologica specificamente mediata dai linfociti CD8+ e tale risposta viene abrogata dalla terapia immuno-soppressiva. Sloan e collaboratori (2007) hanno anche dimostrato che la popolazione trisomica esprime alti livelli di c-myc, ciclina D1 e survivina, è resistente all'apoptosi indotta dalle radiazioni ionizzanti e meno suscettibile alla sospensione dei fattori di crescita delle cellule diploidi normali. Questo pattern anti-apoptotico potrebbe essere determinato da un mancato accoppiamento dell'attivazione della caspasi 3 con la morte cellule, da un mancato accoppiamento dell'attivazione della caspasi 8 con cambiamenti dei potenziali della membrana mitocondriale, e dalla sovra-espressione di survivina e ciclina D1 indotta da c-myc.

Anche la monosomia 7 (-7) ha un'importante influenza sul microambiente midollare. Si tratta dell'anomalia cromosomica che più spesso si osserva nei pazienti affetti da anemia aplastica severa (AAS) al momento dell'evoluzione in SMD o LAM. Alcune casistiche hanno riportato una maggiore incidenza di monosomia 7 e di evoluzione in SMD/LAM nei pazienti con AAS a lungo trattati con G-CSF. Quest'associazione non è stata però confermata da altre casistiche. Pertanto, bisogna ancora stabilire se la popolazione con monosomia 7 è indotta dal G-CSF o è già presente alla diagnosi ed indotta a proliferare dal G-CSF stesso. Siccome è stato osservato che le cellule con -7 sono più sensibili al G-CSF delle cellule diploidi (Sloan et al., 2006), è stato ricercato un difetto nella proteina recettoriale G-CSFR. Gli studi molecolari hanno però escluso che le cellule CD34+ monosomiche presentino mutazioni nel gene del G-CSFR e hanno invece dimostrato una sovra-espressione del RNA messaggero che codifica per la IV isoforma del G-CSFR. Tale isoforma determina una resistenza all'apoptosi inducendo una sovra-regolazione della via di segnale presieduta da Akt (Aarts et al., 2004), con effetti del tutto sovrapponibili a quelli prodotti dalla mutazione dominante del G-CSFR nei pazienti pediatrici con neutropenia congenita severa (Zhu et al., 2004). La via di segnale presieduta da G-CSFR è anche bersaglio del gene che codifica per il fattore trascrizionale CCAAT. Questo gene, mappato sulle braccia lunghe del cromosoma 7 in una regione frequentemente deleta nelle SMD, induce una ridotta espressione del G-CSFR modulando l'espressione di C/EBP $\alpha$  (Khanna-gupta et al., 2001). I pazienti con -7/del(7q) mostrano una ridotta espressione di CCAAT ed una sovra-espressione di C/EBP $\alpha$  e di G-CSFR. Perciò, in questi pazienti la sovra-espressione della proteina C/EBP $\alpha$ , che normalmente induce la differenziazione mieloide, determina una prematura apoptosi mieloide intra-midollare del tutto sovrapponibile a quella che si osserva anche nei pazienti con neutropenia congenita severa (Zhuang et al., 2006).

Anche la delezione 5q ha probabilmente un'importante influenza sull'apoptosi intra-midollare. Vari studi pre-clinici e clinici hanno dimostrato che la lenalidomide è estremamente efficace nel sopprimere la popolazione clonale con 5q-, sia che l'anomalia sia presente come singolo difetto o associata ad ulteriori difetti del

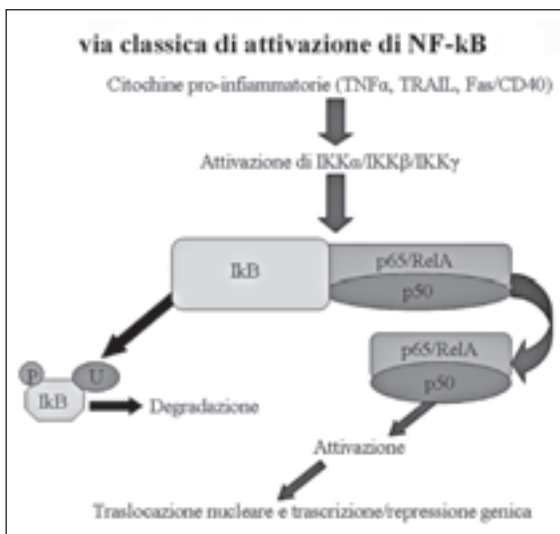
cariotipo (List et al., 2006). La lenalidomide, molecola più potente della talidomide, riduce l'espressione di mRNA codificante per TNF- $\alpha$  e per citochine pro-infiammatorie ed induce la secrezione di IL-2 e  $\gamma$ -interferone. In base a questi dati ed all'inibizione di IL-12, è stato proposto che la lenalidomide induca l'attivazione e l'espansione delle cellule T e natural killer, abbia un effetto anti-proliferativo agendo sulla fase G0-G1 del ciclo cellulare, attivi la via di segnale presieduta dal recettore dell'eritropoietina, riduca l'angiogenesi indotta *in vivo* e la migrazione endoteliale indotta *in vitro* dal fattore di crescita. Queste due ultime azioni sono raggiunte con un'inibizione della fosforilazione di Akt. Quanto sopra riportato suggerisce che la lenalidomide sia in grado di riassetare i segnali di crescita e di differenziazione del micro-ambiente midollare.

#### *Vie di trasformazione delle SMD in LAM*

Il passaggio da una SMD a basso rischio ad una ad alto rischio e l'evoluzione in LAM sono contrassegnati da una progressiva riduzione dei livelli di apoptosi e da un aumento della proliferazione cellulare (Kerbay et al., 2005). Questi cambiamenti sono strettamente correlati a meccanismi molecolari e cellulari. I primi sono stati in parte già descritti, mentre per quanto concerne i secondi bisogna ricordare che l'espressione del recettore R1 del TNF, aumentata nei pazienti con anemia refrattaria (AR) si riduce al momento della progressione in RAEB, RAEB-t e LAM; per contro i recettori R2 presentano un opposto pattern di espressione (Swanobori et al., 2003). I recettori R1 trasmettono segnali protettivi per la cellula attraverso la via del "nuclear factor kappaB" (NF- $\kappa$ B) e segnali pro-apoptotici attraverso la via del "TNF-receptor associated death domain" (TRADD), di FADD e della caspasi 8, mentre i recettori R2 trasmettono solo segnali protettivi per la cellula attraverso la via di segnale presieduta da NF- $\kappa$ B. Un'altra citochina che gioca un ruolo fondamentale nella progressione della SMD a LAM è FLIP [FLICE ("Fas-associated death domain like IL $\beta$ -converting enzyme) inhibitory protein"] (Benesch et al., 2003). Si tratta di un inibitore di Fas e di TRAIL che può essere presente in due diverse varianti di "splicing", FLIP<sub>Long</sub> (FLIP<sub>L</sub>) e FLIP<sub>Short</sub> (FLIP<sub>S</sub>) rispettivamente. È stato recentemente osservato che i livelli di RNA messaggero codificante per le due varianti ed i livelli dei due tipi di proteina sono strettamente correlati ai tassi di apoptosi. In particolare, i livelli di espressione di FLIP<sub>L</sub> RNA delle cellule CD34+ e delle cellule mononucleate di midollo osseo di soggetti sani sono superiori a quelli dei soggetti affetti da SMD ed i livelli di FLIP<sub>L</sub> RNA delle cellule CD34+ si correlano negativamente con i tassi di apoptosi sia negli individui sani che in quelli con SMD. Al contrario, i livelli di espressione di FLIP<sub>S</sub> RNA delle cellule CD34+ e delle cellule mononucleate sono più alti nei soggetti con SMD che in quelli sani e sono direttamente correlati ai tassi di apoptosi. Inoltre, i livelli di proteina FLIP<sub>L</sub>, essendo strettamente correlati con la percentuale di blasti midollari, sono facilmente determinabili nelle SMD in fase avanzata, mentre sono di difficile determinazione o assolutamente trascurabili nelle SMD a basso rischio e nei controlli sani. Pertanto, questo studio ha sottolineato che entrambe le varianti di "splicing" di FLIP si correlano significativamente ai tassi di apoptosi

e che la sovra-espressione di FLIP<sub>L</sub> si accompagna ad una progressione della malattia. Altri studi hanno dimostrato che FLIP<sub>L</sub> e FLIP<sub>S</sub> attraverso il reclutamento di proteine adattatorie possono indurre l'attivazione delle vie di segnale presiedute da NF- $\kappa$ B e da Erk trasformando segnali di morte cellulare in segnali di sopravvivenza cellulare (Kerbaux et al., 2005; Takada et al., 2005; Braun et al., 2006). È stato osservato che la famiglia dei fattori trascrizionali NF- $\kappa$ B rappresenta la più importante via di segnale per la trasformazione delle SMD a basso rischio in SMD ad alto rischio ed in LAM. L'attivazione di NF- $\kappa$ B viene indotta dall'interleukina 1, dal TNF- $\alpha$ , dalle radiazioni ionizzanti e da vari agenti chemioterapici e coinvolge vari recettori di superficie che utilizzano diverse proteine intracellulari producendo segnali diversi che convergono verso una stessa via di segnale ed inducono l'attivazione del complesso I $\kappa$ B (IKK) con conseguente inibizione di NF- $\kappa$ B (Fig. 1).

Al momento dell'attivazione cellulare, I $\kappa$ B viene fosforilato, poli-ubiquitinato e quindi degradato dal proteosoma. In questo modo NF- $\kappa$ B viene liberato dal complesso inibitorio e trasferito nel nucleo dove regola l'espressione di geni che controllano l'apoptosi, la crescita cellulare, la differenziazione cellulare e l'infiammazione. La persistente attivazione di NF- $\kappa$ B è il principale meccanismo di protezione della cellula dall'apoptosi; nella cellula neoplastica è la più rilevante via di induzione di resistenza ai chemioterapici (Birkenkamp et al., 2004). Recentemente, Kerbaux e collaboratori (2005) hanno dimostrato che l'attivazione di NF- $\kappa$ B si correla con lo stadio di SMD. Nelle SMD in fase iniziale le cellule CD34+ presentavano livelli di NF- $\kappa$ B più bassi delle cellule CD34-, facendo ipotizzare che i progenitori ematopoietici più immaturi possano presentare alti livelli di apoptosi, mentre nelle SMD in fase avanzata si verificava l'opposto e veniva quindi ipotizzata l'attivazione di segnali anti-apoptotici. Inoltre, l'azione anti-apoptotica di FLIP<sub>L</sub> veniva ulteriormente confermata dal fatto che il rapporto



**Fig. 1** - Via classica di attivazione del NF- $\kappa$ B. Diverse citochine pro-infiammatorie causano l'attivazione del complesso IKK $\alpha$ /IKK $\beta$ /IKK $\gamma$ . Con l'attivazione cellulare I $\kappa$ B è fosforilato (P) dalla kinasi I $\kappa$ B (IKK), poli-ubiquitinato (U) e degradato dal proteosoma. Dopodiché, NF- $\kappa$ B [p65 (RelA) e p50] è liberato dal complesso inibitorio e trasferito nel nucleo dove regola l'espressione di diversi geni coinvolti nella crescita cellulare, differenziazione, infiammazione e apoptosi. TNF = "tumor necrosis factor"; TRAIL = "TNF-related apoptosis inducine ligand".

FLIP<sub>L</sub>/FLIP<sub>S</sub> era basso nelle SMD in fase iniziale ed alto nelle SMD in fase avanzata. Braun e collaboratori (2005), in contrasto con altri studi, hanno poi rivelato che nella cellula staminale emopoietica e nei progenitori immaturi dei pazienti con SMD in fase avanzata si verificava la traslocazione nucleare della sub-unità

**Tab. 1** - Difetti cromosomici più comuni nelle SMD.

Cariotipo	Frequenza (%) - SMD de novo/SMD secondarie
Anormale	40-70/80-95
Difetti bilanciati	10/4
Difetti sbilanciati	30/12
Anomalie cromosomi 5/7	20/75-90
Anomalie specifiche	
Del(5q)	15/40
-7,del(7q)	5/55
+8	10/10
Del(20q)	5/7
-Y	10/-
Anomalie 17p	7/10
Traslocazioni 11q23	5/5
Cariotipi complessi (≥3 anomalie)	10-20/90

**Tab. 2** - Anomalie cromosomiche nelle LAM.

Cariotipo	Citotipo più comune	Incidenza (%)
Anormale		50-60
Traslocazioni bilanciate		40
Aneuploidie		10
Cariotipi complessi (≥3 difetti)		10
Delezioni		4
<b>Traslocazioni coinvolgenti il “core binding factor”</b>		
t(8;21)(q22;q22) → gene di fusione: AML1/ETO	M2	8
t(3;21)(q26;q22) → gene di fusione: AML1/EVI1 or AML1/MDS1	M0,M1,M2	1-2
inv(16)(p13q22)/t(16;16)(p13;q22) → gene di fusione: CBFβ/SMMHC	M4	10-12
<b>Traslocazioni coinvolgenti il recettore α dell'acido retinoico</b>		
t(15;17)(q22;q12) → gene di fusione: PML/RARA	M3	15
t(11;17)(q23;q12) → gene di fusione: PLZF/RARA	M2-M3	2
<b>Translocations involving the Mixed Lineage Leukemia (MLL) gene</b>		
t(9;11)(p22;q23) → gene di fusione: MLL/AF9	M4	4
t(11;19)(q23;p13.3) → gene di fusione: 1MLL/ENL	M4	Raro
<b>Translocations involving transcription co-activators</b>		
t(8;16)(p11;p13) → gene di fusione: MOZ/CBP	M4	Raro
inv(8)(p11q13) → gene di fusione: MOZ/TIF	M4	Raro
t(11;16)(q23;p13) → gene di fusione: MLL/CBP	M4	Raro
<b>Other translocations</b>		
t(6;9)(p23;q34) gene di fusione: DEK/CAN	M2	2
t(3;3)(q21;q26) gene di fusione: EVI1/Ribophorin 1	Nessuno	1
inv(3)(q21q26) gene di fusione: EVI1/Ribophorin 1	Nessuno	1

**Tab. 3** - Lesioni molecolari nelle LAM a cariotipo normale.

Gene	Difetto molecolare	Localizzazione cromosomica	Incidenza (%)
FLT3	Internal tandem duplication Mutazione puntiforme del dominio tirosin kinasico	13q12	15-30 5-15
NPM1	4-base pair duplication	5q35	45-60
ABCG2	Sovra-espressione	4q22	42
MLL	Partial tandem duplication	11q23	5-10
BAALC	Sovra-espressione	8q22.3	25
CEBPA	Mutazioni C- e N-terminal	19q13.1	15-20
KIT	Mutazioni puntiformi	4q11	<10
WT1	Sovra-espressione	11p13	70
N-Ras	Mutazioni puntiformi	1p13.2	20-30
ERG	Sovra-espressione	21q22.3	16

Legenda: FLT3 = "Fms-related tyrosine kinase 3"; NPM1 = "Nucleophosmin 1"; BRCP = "Breast cancer resistance protein"; MLL = "Mixed-lineage leukaemia"; BAALC = "Brain and acute leukaemia"; CEBPA = "CCAAT/enhancer binding protein (C/EBP) alpha"; KIT = "v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog"; WT1 = "Wilm's Tumour"; N-Ras = "Neuroblastoma RAS viral oncogene homolog"; ERG = "v-ets erythroblastosis virus E26 oncogene-like".

p65 del NF- $\kappa$ B con conseguente attivazione della stessa via di segnale. L'attivazione di NF- $\kappa$ B sembrava ristretta alle sole cellule dotate di anomalie cromosomiche almeno per quanto concerneva la monosomia 7 e la trisomia 8 e la sua inibizione induceva un aumento dell'apoptosi nelle SMD ad alto rischio, mentre non aveva alcun effetto nelle SMD a basso rischio. Questo dato, oltre a confermare l'esistenza di una correlazione inversa tra attivazione di NF- $\kappa$ B e livelli di apoptosi, dimostrava che l'attivazione di NF- $\kappa$ B non poteva essere indotta né da TNF, né da TRAIL. Infatti, l'attivazione di Fas/CD95 si verifica unicamente nelle cellule con trisomia 8 (Sloand et al., 2007) e sia TNF che TRAIL, essendo entrambi agenti diffusibili, potrebbero determinare l'attivazione di NF- $\kappa$ B anche nelle cellule normali. È stato quindi proposto che l'attivazione di NF- $\kappa$ B possa avvenire con un meccanismo autonomo e possa essere causata da mutazioni secondarie. Questa ipotesi sarebbe anche confermata dall'attivazione della via di NF- $\kappa$ B in cellule blastiche prive di anomalie citogenetiche. Inoltre, l'attivazione di NF- $\kappa$ B potrebbe essere indotta da altre vie di segnale oltre a quella presieduta da NIK/IKK/I $\kappa$ B dal momento che effettori a valle della via di segnale presieduta Ras-GTPasi, ad esempio il sistema Raf/MEK/ERK, potrebbero interagire con la via di NF- $\kappa$ B (Birkenkamp et al., 2004).

Recentemente, Birkenkamp e collaboratori (2004) hanno osservato che nei blasti di LAM l'attivazione di Ras è significativamente correlata con la capacità di legarsi al DNA di NF- $\kappa$ B. Il fatto che nei pazienti con attivazione costitutiva della capacità DNA-legante di NF- $\kappa$ B sia Ras che PKB siano costitutivamente attivati e che l'inibitore della farnesil-transferasi L-744832 non inibisca solo la fosforilazione di PKB ma anche la capacità di legarsi al DNA di NF- $\kappa$ B dimostra che Ras determina l'attivazione di NF- $\kappa$ B attraverso la via di segnale PI3-K/PKB (Birkenkamp et al., 2004). Lo stesso studio ha indicato che nelle LAM l'attivazione di Ras è molto più frequente della mutazione e che né la mutazione né l'attivazione costitutiva di FLT3 sono responsabili dell'attivazione di NF- $\kappa$ B.

Takahashi e collaboratori (2005) hanno invece dimostrato che la sovra-espressione di FLT3 determina una sovra-regolazione di una molecola reporter ed induce un aumento dell'espressione di IL-6. Gli Autori hanno però ipotizzato che la sovra-espressione di FLT3 possa attivare la via di segnale presieduta da NF- $\kappa$ B non solo grazie ad un effetto diretto ma anche grazie ad altri fattori che agiscono in sinergia con NF- $\kappa$ B nell'indurre la produzione di IL-6. Questi risultati tra loro in contraddizione sono in accordo con il modello già proposto della cooperazione tra più mutazioni per la patogenesi delle LAM *de novo* e delle t-SMD/t-LAM. Infatti, secondo questo modello il fenotipo neoplastico sarebbe prodotto da mutazioni diverse che attivano vie di segnale diverse dotate di meccanismi d'azione talora diversi ma talaltra assolutamente sovrapponibili.

Altri studi hanno sottolineato il ruolo della via di segnale PI3-K/Akt nelle SMD ad alto rischio (Nyåkern et al., 2006; Follo et al., 2007). È stato dimostrato che, similmente ai blasti leucemici delle LAM, le cellule mononucleate della SMD ad alto rischio presentano un'attivazione per fosforilazione della kinasi Akt, un'alta espressione della sub-unità catalitica p110 $\delta$  di classe I del PI3-K ed una ridotta espressione di PTEN, eventi tutti responsabili della riduzione dei livelli di apoptosi nelle SMD in fase avanzata. Ancor più recentemente, Follo e collaboratori (2007) hanno esaminato lo stato funzionale di alcuni bersagli posti a valle di Akt e hanno dimostrato che i livelli di espressione del "mammalian target of rapamycin" (mTOR) e di alcuni suoi bersagli posti più a valle ("4-E-binding protein 1" e "p70 ribosomal S6 kinase") non sono determinabili nei soggetti sani, sono invece bassi nei pazienti con SMD a basso rischio ed alti nelle SMD ad alto rischio. Inoltre, le cellule CD33+ dei pazienti con SMD ad alto rischio non esposte a rapamicina presentano una percentuale di cellule apoptotiche inferiore a quella dei controlli sani e dei soggetti con SMD a basso rischio. Questo dato dimostra che in tali cellule si verifica l'attivazione di meccanismi anti-apoptotici. L'esposizione a rapamicina è in grado di inibire almeno alcune delle vie che attivano tali meccanismi ed è quindi in grado d'indurre la morte cellulare programmata. Inoltre, la via di segnale presieduta da Erk1/2 era attivata nei controlli sani e nelle SMD a basso rischio, normale nelle SMD ad alto rischio ed attivata assieme a PKB nelle LAM (Follo et al., 2007). Però, nelle SMD e nelle AML l'attivazione di Erk1/2 era probabilmente ottenuta attraverso vie di segnale diverse. Infatti, nelle SMD ad alto rischio l'attivazione di Akt avveniva indipendentemente da PI3K poiché inibitori di PI3k e di PI3K $\delta$ , molto efficaci nell'indurre apoptosi nelle LAM, erano invece incapaci di impedire la proliferazione cellulare nelle SMD ad alto rischio (Billottet et al., 2006; Follo et al., 2007).

## Bibliografia

1. Shadduck RK., Latsko JM., Rossetti JM., Haq B., Abdulhaq H. Recent advances in myelodysplastic syndromes. *Experimental Hematology*, 2007; 35: 137-143.
2. Corey SJ., Minden MD., Barber DL., et al. Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nature Reviews Cancer*, 2007; 7: 118-129.

3. Lapidot T, Sirard C, Vormoor J, et al A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, 1994; 367: 645-648.
4. Bonnet D., Dick JE. Human acute myeloid leukaemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine*, 1997; 3: 730-737.
5. Manakova T., Gerasimova I., Tsvetaeva N., Momotyuk K. Functional disturbances in hematopoietic microenvironment in various forms of myelodysplastic syndromes. *Bull. Exp. Biol. Med*, 2000; 130: 832-834.
6. Thanopoulou E., Cashman J., Kakagianne T., et al. Engraftment of NOD/SCID beta2 microglobulin null mice with multilineage neoplastic cells from patients with myelodysplastic syndromes. *Blood*, 2004; 103: 4285-4293.
7. Olney JH., Le Beau MM. The cytogenetics of myelodysplastic syndromes. *Best Practice & Res. Clin. Haematol.*, 2001; 14: 479-495.
8. Byrd JC., Mrózek K., Dodge RK., et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukaemia: results from Cancer and Leukemia Study Group B (CALGB 8461). *Blood*, 2002; 100: 4325-4336.
9. Bernasconi P., Klersy C., Boni M., et al. World Health Organization classification in combination with cytogenetic markers improves the prognostic stratification of patients with de novo primary myelodysplastic syndromes. *Brit. J. Haematol*, 2006; 137: 193-205.
10. Deguchi K., Gilliland DG. Cooperativity between mutations in tyrosine kinases and in hematopoietic transcription factors in AML. *Leukemia*, 2002; 16: 740-744.
11. Baldus C., Mrózek K., Marcucci G., et al. Clinical outcome of *de novo* acute myeloid leukaemia patients with normal cytogenetics is affected by molecular genetic alterations: a concise review. *Brit. J Haematol*, 2007; 137: 387-400.
12. Ebert BL., Pretz J., Bosco J., et al Identification of RPS14 as the 5q-Syndrome gene by RNA interference screen. *Blood*, 2007; 110: 8a (Abstract 1).
13. Grisendi S., Bernardi R., Rossi M., et al. Role of nucleophosmin in embryonic development and tumorigenesis. *Nature*, 2005; 437: 147-153.
14. Li J., Sejas DP., Rani R., et al. Nucleophosmin regulates cell cycle progression and stress response in hematopoietic stem/progenitor cells. *J. Biol. Chem.*, 2006; 281: 16536-16545.
15. Pedersen-Bjergaard J., Christiansen DH., Desta F., et al. Alternative genetic pathways and cooperating genetic abnormalities in the pathogenesis of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia*, 2006; 20: 1943-1949.
16. Buonamici S., Li D., Chi Y., et al. EVI1 induces myelodysplastic syndrome in mice. *J. Clin. Invest*, 2004; 114: 713-719.
17. Lin YW., Slape C., Zhang Z., et al. NUP98-HOXD13 transgenic mice devel-



- op a highly penetrant, severe myelodysplastic syndrome that progresses to acute leukaemia. *Blood*, 2005; 106: 287-295.
18. Betti CJ., Villalobos MJ., Diaz MO., et al. Apoptotic stimuli initiate MLL-AF9 translocations that are transcribed in cells capable of division. *Cancer Res*, 2003; 63: 1377-1381.
  19. Higuchi M., O'Brien D., Kumaravelu P., et al. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukaemia. *Cancer Cell*, 2002; 1: 63-74.
  20. Sloand EM., Pfannes L., Chen G., et al. CD34 cells from patients with trisomy 8 myelodysplastic syndrome (MDS) express early apoptotic markers but avoid programmed cell death by up-regulation of antiapoptotic proteins. *Blood*, 2007; 109: 2399-2405.
  21. Sloand EM., Yong ASM., Ramkissoon S., et al. Granulocyte colony-stimulating factor preferentially stimulates proliferation of monosomy 7 cells bearing the isoform IV receptor. *Proc. Natl. Acad. Sci. USA*, 2006; 103: 14483-14488.
  22. Aarts LH., Roovers O., Ward AC., et al. Receptor activation and two distinct COOH-terminal motifs control G-CSF receptor distribution and internalisation kinetics. *Blood*, 2004; 103:571-579.
  23. Zhu QS., Robinson LJ., Roginskaya V., et al. G-CSF-induced tyrosine phosphorylation of Gab2 is Lyn kinase dependent and associated with Akt and differentiative, not proliferative, responses. *Blood*, 2004; 103: 3305-3312.
  24. Khanna-Gupta A., Zibello T., Sun H., et al. (2001) C/EBP mediates myeloid differentiation and is regulated by the CCAAT displacement protein (CDP/cut). *Proc. Natl. Acad. Sci. USA*, 2001; 98: 8000-8005.
  25. Zhuang D., Qiu Y., Kogan SC., et al. Increased CCAAT enhancer-binding protein  $\epsilon$  (C/EBP $\epsilon$ ) expression and premature apoptosis in myeloid cells expressing Gfi-1 N382S mutant associated with severe congenital neutropenia. *J. Biol. Chem.*, 2006; 281: 10745-10751.
  26. List A., Dewald G., Bennet J., et al. Lenalidomide in the Myelodysplastic Syndrome with chromosome 5q deletion. *New Engl. J Med*, 2006; 355: 1456-1465.
  27. Kerbauy DMB., Lesnikov V., Abbasi N., et al. NF- $\kappa$ B and FLIP in arsenic trioxide (ATO)-induced apoptosis in myelodysplastic syndromes (MDSs). *Blood*, 2005; 106: 3917-3925.
  28. Benesch M., Platzbecker U., Ward J., et al. Expression of FLIP<sub>Long</sub> and FLIP<sub>Short</sub> in bone marrow mononuclear and CD34+ cells in patients with myelodysplastic syndrome: correlation with apoptosis. *Leukemia*, 2003; 17: 2460-2466.
  29. Takada Y., Andreeff M., Aggarwal BB. Indole-3-carbinol suppresses NF- $\kappa$ B and I $\kappa$ B kinase activation, causing inhibition of expression of NF- $\kappa$ B-regulated antiapoptotic and metastatic gene products and enhancement of apoptosis in myeloid and leukaemia cells. *Blood*, 2005; 106: 641-649.
  30. Braun B., Carvalho G., Coquelle A., et al. NF- $\kappa$ B constitutes a potential ther-



- apeutic target in high-risk myelodysplastic syndrome. *Blood*, 2006; 107:1156-1165.
31. Birkenkamp KU., Geugien M., Schepers H., et al. Constitutive NF- $\kappa$ B DNA-binding activity in AML is frequently mediated by a RAS/PI3-K/PKB-dependent pathway. *Leukemia*, 2004; 18: 103-112.
  32. Follo MY., Mongiorgi S., Bosi C., et al. The Akt/mammalian target of rapamycin signal transduction pathway is activated in high-risk myelodysplastic syndromes and influences cell survival and proliferation. *Cancer Res*, 2007; 67: 4287-4294.
  33. Nyåkern M., Tazzari PL., Finelli C., et al. Frequent elevation of Akt kinase phosphorylation in blood marrow and peripheral blood mononuclear cells from high-risk myelodysplastic syndrome patients. *Leukemia*, 2006; 20: 230-238.
  34. Billotet C., Grandage VL., Gale RE, et al. A selective inhibitor of the p110 isoform of PI 3-kinase inhibits AML cell proliferation and survival and increases the cytotoxic effects of VP16. *Oncogene*, 2006; 25: 6648-6659.

**GENOMICA E NANOTECNOLOGIE  
NELLE MALATTIE CARDIOVASCOLARI**



## **Nanotecnologie nella diagnostica delle cardiomiopatie**

*Fabiana Gambarin, Nicola Marziliano, Maurizia Grasso, Eloisa Arbustini*

Centro per le Malattie Genetiche Cardiovascolari, Fondazione IRCCS Policlinico San Matteo, Pavia

Le *nanoscienze* costituiscono il punto di incontro di discipline diverse che vanno dalla fisica quantistica, alla chimica supramolecolare, dalla scienza dei materiali, alla biologia molecolare e rappresentano una realtà ormai affermata nel mondo della ricerca. Le *nanotecnologie*, che sono invece ancora nella fase iniziale del loro sviluppo, puntano a sfruttare e ad applicare i metodi e le conoscenze derivanti dalle nanoscienze. Esse fanno riferimento ad un insieme di tecnologie, tecniche e processi che richiedono un approccio multidisciplinare e consentono la creazione e utilizzazione di materiali, dispositivi e sistemi con dimensioni a livello nanometrico. In sintesi, con nanotecnologie si intende la capacità di *osservare, misurare e manipolare la materia su scala atomica e molecolare*. Le possibilità offerte dalle nanotecnologie di analizzare campioni biologici in maniera rapida e *high-throughput* (i.e. molti geni/loci contemporaneamente) ben si adatta alla necessità di una diagnostica molecolare avanzata in un contesto di malattia “mendeliana” ma dalla forte eterogeneità genetica come accade nelle cardiomiopatie; inoltre le nanotecnologie permettono l’identificazione - ad esempio tramite lo studio di associazione - di loci non ancora descritti come causativi di questa patologia.

### **Lo scenario clinico delle cardiomiopatie**

Le cardiomiopatie “idiopatiche” sono malattie primitive del muscolo cardiaco, non causate quindi da malattia coronarica, valvolare, congenita, o sistemica. La classificazione clinica si basa sul fenotipo e comprende quattro gruppi maggiori: le cardiomiopatie dilatative (CMD), ipertrofiche (CMI), restrittive (CMR) ed aritmogene del ventricolo destro (ARVD); esistono forme non ancora completamente caratterizzate come il ventricolo sinistro non compatto (LVNC) e la sindrome del *left ventricular apical ballooning* (LVABS), la cui appartenenza al gruppo delle cardiomiopatie è dibattuta (1).

La CMD è caratterizzata da dilatazione e disfunzione ventricolare sinistra e si manifesta tipicamente con scompenso cardiaco. È la prima causa di scompenso cardiaco terminale che porta alla necessità di trapianto di cuore.

La CMI è caratterizzata da ipertrofia ventricolare sinistra asimmetrica nella mag-

gioranza dei casi localizzata a livello settale (prevalente), apicale e della parete libera (meno frequente). Può essere clinicamente silente oppure può manifestarsi con aritmie e morte improvvisa, addirittura come prima manifestazione clinica; può esordire con sintomi da ostruzione all'efflusso ventricolare sinistro nei casi di CMI ostruttiva; può evolvere verso la dilatazione ventricolare sinistra e lo scompenso cardiaco ed essere, specie nelle fasi evolutive terminali, clinicamente simile alla cardiomiopatia dilatativa. Per i problemi aritmici ventricolari è una patologia nella quale è spesso necessario l'impianto di un defibrillatore.

La CMR è caratterizzata da disfunzione diastolica, spessori ventricolari sinistri normali e importante dilatazione biatriale. Esistono CMI che possono comportarsi clinicamente come CMR, ma sono comunque morfologicamente caratterizzate da ispessimento/ipertrofia ventricolare sinistra segmentaria o diffusa.

La ARVC è definita dalla combinazione di criteri morfologici macro- e microscopici, (sostituzione fibroadiposa della parete ventricolare destra, più o meno estesa), anomalie elettriche ed aritmie potenzialmente fatali. Il ventricolo sinistro non compatto definisce una condizione caratterizzata da prevalenza della porzione trabecolata non compatta su quella compatta del ventricolo sinistro. Non è ancora unanime il consenso sul fatto che si tratti di una vera cardiomiopatia, in quanto la variabilità anatomica del rapporto porzione compatta/non compatta è molto ampia e la funzione ventricolare sinistra nonché le dimensioni possono essere normali anche in ventricoli sinistri con elevato grado di trabecolatura (2).

Il LVABS è una patologia acuta, transitoria caratterizzata da "paralisi" della porzione anteroapicale del ventricolo sinistro che simula clinicamente l'infarto acuto del miocardio, ma che non è causata da patologia coronarica. È generalmente innescata da un trigger emotivo molto consistente e colpisce solo la donna solitamente post-menopausa. La sua appartenenza alle cardiomiopatie non è unanimemente condivisa; la patogenesi risiede in un possibile *imbalance* di espressione e distribuzione dei recettori adrenergici del ventricolo sinistro.

Lo scompenso cardiaco evolvente fino alla necessità di trapianto cardiaco e le aritmie potenzialmente letali nell'età giovane adulta sono le due grandi conseguenze delle cardiomiopatie.

### **Prevalenza delle diverse forme di cardiomiopatia**

La prevalenza delle CMD è stimata essere di 1:10.000, quella delle CMI di 1:500 in popolazioni di giovani adulti, quella delle CMR è ignota (è di fatto molto rara) e quella delle ARVC di 1: 5000 (1, 3). Per ciascun fenotipo, esistono numerose fenocopie solo apparentemente simili.

Quindi nel nostro Paese il numero complessivo degli individui affetti da cardiomiopatie è atteso essere di almeno 150.000 individui. La stragrande maggioranza dei casi di cardiomiopatia riguarda una popolazione di giovani ed adulti, di entrambi i sessi, anche se all'interno di ciascun gruppo vi sono patologie che possono essere causate da difetti di geni che mappano sul cromosoma X ed esprimersi quindi tipicamente nei pazienti di genere maschile e da difetti del DNA mitocondriale a trasmissione matrilineare, ma espressi sia in maschi che femmine.

## Cardiomiopatie sporadiche e familiari

Una recente revisione della classificazione esistente ha introdotto per ciascun gruppo di cardiomiopatie la suddivisione in forme familiari e non familiari (1). Sono definite familiari quelle cardiomiopatie che sono dimostrate ricorrere in almeno due individui della stessa famiglia (genitori-figli, germani) (4, 5) o che, in seguito a dimostrazione della causa genetica alla base del disordine, possono essere clinicamente presenti anche in un solo membro della famiglia ove vi sia una madre portatrice (cardiomiopatia *X-linked*) (6) ovvero, in caso di difetto del DNA mitocondriale (7), quando la malattia si manifesti in un solo membro della famiglia, caratterizzato da livelli di eteroplasmia della mutazione causale nel miocardio tali da far esprimere il difetto sul piano clinico. In presenza di un unico affetto accertato nella famiglia si parla di cardiomiopatie sporadiche; questa definizione può essere ribaltata quando nel corso del follow-up emerge un secondo membro affetto.

A partire dagli anni '80 ad oggi la percentuale di CMD familiari è salita dal 7% al 65% grazie a studi di *screening* clinico (visita, ECG ed ecocardiogramma) che hanno incluso i parenti del probando e che hanno documentato che almeno il 10% dei probandi ha almeno un parente affetto in fase precoce ancora asintomatica (4); a tre anni di *follow-up* una percentuale di almeno il 6% di famiglie vede emergere un ulteriore membro affetto in fase asintomatica precoce (8). A contribuire e confermare l'elevata percentuale di forme familiari è subentrato il grosso contributo della ricerca dedicata alle basi genetiche delle CMD (v. paragrafo successivo) che ha spostato le vecchie ipotesi di malattia acquisita su base prevalentemente post-miocarditica a malattia prevalentemente ereditaria e con basi genetiche identificabili. La CMI è invece familiare in almeno il 70% dei casi, e nei rimamenti è spesso possibile identificare una mutazione genica "*de novo*" (3). Queste cardiomiopatie sono quindi malattie familiari quasi per definizione; lo screening clinico delle famiglie evidenzia spesso membri affetti asintomatici, non consapevoli della loro condizione e malattia. La CMR pura è familiare almeno nel 50% dei casi; il fenotipo restrittivo coincide con la cardiomiopatia restrittiva quando non siano presenti patologie infiltrative (per esempio l'amiloidosi cardiaca) o sistemiche (per es. la sclerodermia) che simulino la cardiomiopatia restrittiva idiopatica. La stessa proporzione (circa 50% di forme familiari) si registra nel gruppo delle ARVC (1).

## La genetica clinica nelle cardiomiopatie ereditarie

La stragrande maggioranza delle CMD familiari (circa l'85%) si trasmette come carattere autosomico dominante, una minoranza come carattere X-linked recessivo (7%), autosomico recessivo (5%) e matrilineare (3%) (4, 6, 7). L'esperienza clinica acquisita negli ultimi decenni ha documentato queste modalità di trasmissione in modo incontrovertibile. Sono rimasti aperti alcuni problemi oggetto di attuale ricerca:

- 1) la penetranza incompleta: in una certa percentuale delle cardiomiopatie autosomiche dominanti è possibile osservare per esempio un genitore ed un figlio affetti, con cugini di primo grado della stessa via genitoriale affetti ma aventi il genitore correlato sano;

- 2) la penetranza tardiva: si parla di età-dipendenza della comparsa del fenotipo clinico, per cui solo regolari programmi di monitoraggio clinico rivelano nel tempo la malattia. Forme sicuramente familiari emergono clinicamente come tali solo con l'età;
- 3) l'espressione clinica variabile: all'interno di stesse famiglie vi possono essere membri affetti da fenotipo CMD e altri con fenotipo CMI. Alcuni membri possono esordire con disturbi della conduzione o del ritmo, altri con scompenso cardiaco acuto;
- 4) l'espressione clinica incompleta: alcuni membri di famiglie con cardiomiopatie clinicamente ben caratterizzate possono presentare per decenni forme lievi di malattia, mentre altri membri presentano, già all'esordio, forme gravissime rapidamente fatali;
- 5) l'anticipazione clinica: a tutt'oggi non vi corrisponde una base genetica nota (com'è invece per altre patologie note con anticipazione; la malattia tende a comparire sempre più precocemente nelle generazioni). Nelle famiglie di pazienti con cardiomiopatia si osservano genitori e figli affetti da identica patologia che giungono al trapianto gli uni verso la quinta-sesta decade di vita e gli altri nella seconda o terza decade di vita. Non essendo ancora completamente nota la storia naturale di molte cardiomiopatie a partire dall'esordio presintomatico (la malattia è infatti riconoscibile in fase precoce solo mediante esami strumentali che non necessariamente vengono espletati in assenza di sintomi), è possibile che non si tratti di una vera anticipazione ma che geni modificatori o fattori ambientali si sovrappongano a quelli genetici nell'influenzare la storia naturale della malattia.

A spiegare parte di queste difficili situazioni cliniche entra la non ancora completa conoscenza delle basi genetiche delle cardiomiopatie (v. paragrafo successivo).

### **Basi di genetica molecolare delle cardiomiopatie**

Le cardiomiopatie ereditarie sono caratterizzate da eterogeneità genetica: difetti presenti in geni diversi possono essere associati allo stesso fenotipo clinico. Passando dalla genetica clinica a quella molecolare emerge infatti come nei grandi raccoglitori clinici delle CMD, CMI etc. si collochino di fatto più malattie causate da geni diversi e da difetti diversi. Il percorso dell'identificazione delle basi di genetica molecolare delle cardiomiopatie è stato rallentato dalla difficile caratterizzazione clinica: i segni e sintomi delle cardiomiopatie apparivano a volte così simili da non sollevare il sospetto che fossero malattie diverse tra loro. Il primo grande aiuto è venuto dalla verifica della familiarità clinica (4, 5, 8): anche se difficili da "vedere", le cardiomiopatie emergevano in più membri di stesse famiglie. Quindi si è cominciato a ricercare l'associazione tra tratti clinici e marcatori genici per cercare una corrispondenza (un legame) tra determinati marcatori genetici e fenotipo (analisi di *linkage*). Da queste prime fasi della ricerca sono emersi i primi loci candidati a contenere i geni che potevano essere alla base delle malattie nelle famiglie studiate. Negli ultimi 20 anni sono stati così identificati decine e decine di geni-malattia. Contemporaneamente si sviluppavano le metodiche di analisi del

**Tab. 1**

CMD	CMI	CMR	ARVC
LMNA	CMI sarcomerica	Desmin	PKP2
Dystrophin (DYS)	MYBPC3	TNNI3	DSG1
Tafazzin	MYH7		DSC2
LDB3 (CYPHER-ZASP)	MYH6	Fenocopie, es.	DSG3
Epicardin	TNNT2	Amiloidosi ereditarie	DSP
CRYAB	TNNC1	TTR	JUP
T-CAP	TPM1	ApoA1	TGFB3
DYS Associated Glycoproteins' genes (n = 10)	TNNI3	Etc.	Mutazioni di gesti geni si possono associare anche a CMD
ACTC	MYO6		
Desmin	MYL2	Sclerodermia,	
Emerin	MYLK2	Sarcoidosi	
Sarcomeric HCM:3/15 geni	MYL3	Emocromatosi	
ARVC: 7 geni	TTN		
MtDNA	ANT1	Miopatie	
(tutti i geni >16.000 bp)	ACTC	Con basi genetiche note (>20)	
Fenocopie con evoluzione end-stage dilatativa	Desmin		
	Fenocopie es.		
	LAMP2		
	PRKAG2		
	GLA		
	MtDNA		
	PTPN11, Etc.		

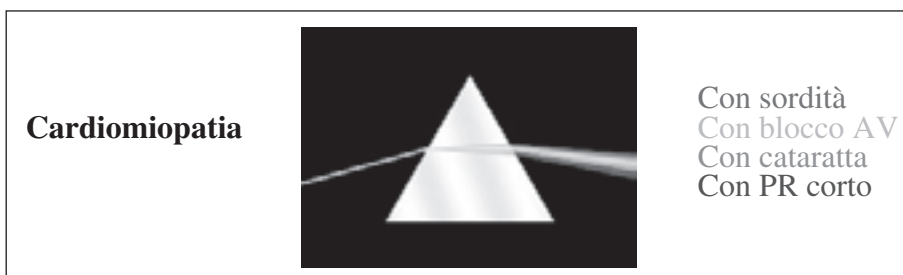
DNA e da queste a loro volta nascevano sistemi *high throughput* che potevano in tempi rapidi analizzare i geni candidati nei probandi e nei parenti affetti all'interno delle famiglie studiate. Di pari passo si sviluppavano anche tecniche sempre più avanzate di indagini tissutali e cellulari, che consentivano di avviare verifiche degli effetti dei difetti genici, e quindi la possibilità di studiare quantitativamente il trascrittoma, il proteoma, il metaboloma. Una ulteriore difficoltà incontrata dalla traslazione della genetica molecolare delle cardiomiopatie è stata l'eterogeneità genetica. Sono oggi noti oltre 20 geni associati a CMD, oltre 15 geni associati a CMI, almeno 7 geni associati a ARVC e alcuni geni associati a CMR. Stessi geni possono inoltre causare fenotipi clinici diversi. Oltre ad essere numerosi, i geni-malattia identificati sono anche complessi e di ampie dimensioni, tali cioè da rendere difficile l'analisi di tutti i geni senza disporre di una guida clinica che indirizzi un dato fenotipo verso un gene più probabile. La tabella 1 ha il solo scopo di illustrare parzialmente l'eterogeneità genetica dei principali gruppi di cardiomiopatie (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM>) e l'esistenza di fenocopie che devono essere riconosciute perchè esistono terapie malattia-specifiche.

Ciononostante, l'analisi di tutti i geni malattia noti in serie consecutive di pazienti clinicamente ben caratterizzati ha portato ad un volume di conoscenze tali e ad un numero di famiglie/pazienti che condividevano mutazioni dello stesso gene tali da favorire l'identificazione, all'interno di queste famiglie e gruppi di pazienti, di marcatori clinici apparentemente poco rilevanti (per es. un blocco atrioventricolare, oppure una ipoacusia associata, un intervallo PR breve, un WPW, una cataratta precoce, una trabecolatura ventricolare sinistra più marcata, etc.) più specifici rispetto a quello principale CMD, o CMI etc., ma ad esso associato. Al



fenotipo clinico principale cioè si va oggi affiancando la ricerca delle cosiddette *red flags* cliniche che caratterizzano quelle cardiomiopatie che originano da difetti dello stesso gene o difetti di gruppi di geni attivi nell'ambito dello stesso *pathway* biomolecolare.

La strategia oggi è quella di “partire da un fascio di luce bianca per ottenere la rifrazione dello spettro di colori che lo compongono”: per esempio, a partire dalla diagnosi principale di CMD si identificano diverse forme (quelle causate da difetti del gene *LMNA*, quelle da difetti di *ACTC*, *epicardina*, etc.). Mentre non sono facilmente scomponibili i grandi gruppi di cardiomiopatie (CMD, CMI, CMR, ARVC), sono invece “scomponibili” i fenotipi specifici quando si riesca, sul piano clinico, a riconoscere marcatori specifici associati al fenotipo principale (la figura sottostante illustra il concetto) utili per indirizzare la ricerca del giusto gene per quella particolare cardiomiopatia.



Quindi la strategia per un uso ottimizzato sulle risorse delle tecnologie disponibili oggi ai fini della diagnostica delle cardiomiopatie è la combinazione di marcatori di genetica clinica (modalità di trasmissione), di cardiologia clinica (tratti cardiaci fenotipici associati), di medicina clinica (tratti extracardiaci associati), di biochimica (per es. aumento del CKP, o disionie); successivamente se ne verifica la segregazione con il fenotipo cardiomiopatico e si orienta l'analisi genetica molecolare al gene candidato più appropriato (6, 7, 9, 13).

### Quali risultati ha portato questa ricerca

Nel 1987 il nostro gruppo di ricerca ha acquisito il primo termociclatore (PCR), nel 1992 il primo sequenziatore automatico, quindi il primo sintetizzatore di oligonucleotidi (*service* per più di 10 anni per Policlinico ed Università), quindi via via negli anni altri termociclatori (oggi uno per operatore), altri sequenziatori, piattaforme per ricerca *high throughput* di varianti geniche, piattaforme per trascrittomiche e sistemi robotizzati per l'estrazione e processazione rapida degli acidi nucleici. Dalla ricerca si è passati alla traslazione assistenziale. Oggi si può parlare di diagnostica molecolare delle cardiomiopatie ereditarie.

Il tasso di insuccessi all'inizio è stato impressionantemente elevato, quasi scagghiante. In seguito, l'integrazione clinica e genetica continua, l'implementazione ed i percorsi di qualità in laboratorio (certificazione ISO 9001 dal 2004) hanno portato all'impiego clinico delle informazioni ottenute in ambito di genetica molecolare. Ad oggi sono 494 i pazienti con cardiomiopatia nei quali è stata

identificata la mutazione (o le mutazioni, in quanto sono emerse eterozigotà doppie e composte nel 10% circa dei pazienti) presso il nostro Centro.

### **A cosa serve oggi conoscere la base genetica della malattia?**

#### *Diagnosi*

All'interno delle CMD parliamo per esempio di:

- *cardiodistrofinopatie*: spesso misconosciute perchè la CMD può esserene la manifestazione clinica più rilevante, in assenza di miopatie conclamate o comunque clinicamente evidenti (6);
- *tafazzinopatie*, nelle quali il problema clinico neonatale è di estrema gravità e che solo se superato grazie ad un'ottima gestione dopo la nascita può consentire non solo la sopravvivenza ma un miglioramento clinico radicale, quasi una guarigione: il ruolo che le tafazine svolgono nella vita fetale viene infatti vicariato da altri geni (13);
- *cardiolaminopatie*, ovvero CMD con fenotipo clinico caratterizzato dall'associazione largamente prevalente di BAV e di elevato rischio di eventi, sia morte improvvisa che evoluzione verso la fase *end-stage* della malattia con necessità di trapianto cardiaco; oggi sappiamo che queste malattie hanno penetranza completa, età dipendente, e che entro l'età di 60 anni tutti i portatori di difetti del gene *LMNA* esprimono il fenotipo (9);
- *epicardinopatie*, malattie autosomiche dominanti associate a sordità (14, 15);
- *cardiozaspopatie*, associate a non compattazione del ventricolo sinistro (2).

Nell'ambito delle CMI, è ormai chiaro che l'identificazione di una mutazione in uno dei geni-malattia non deve coincidere con la conclusione dell'analisi genetica per l'esistenza di almeno il 10% di eterozigotà doppie o composte (16-18). Il numero di controlli da valutare per ciascuna mutazione identificata deve essere di almeno 200, data la non rara possibilità di subpolimorfismi missenso che rischiano di essere interpretati come mutazioni causali, specie in famiglie con pochi membri viventi nelle quali la verifica della segregazione del difetto con il fenotipo è difficile, oppure in presenza di mutazioni *de novo*.

#### **Prognosi**

Il paradigma emerso negli anni '90: "CMI con ipertrofia severa = mutazioni di MYH7; CMI con ipertrofia di grado lieve = mutazione dei gene MYBPC3" (19) non è stato confermato. Sicuramente i fenotipi più gravi sono quelli che si associano a eterozigotà composte (16-18). L'elevata frequenza di CMI (1:500) (3) comporta anche un'elevata probabilità di incroci tra individui portatori di varianti geniche che isolatamente non genererebbero malattia o causerebbero malattie non severe. I geni codificanti per le troponine sono di fatto alla base di CMI con comportamenti funzionali restrittivi pur in presenza di ipertrofia anche solo lieve; questo non sembra valere per il gene della troponina I, i cui difetti possono causare un fenotipo CMR puro. I tratti clinici associati al fenotipo principale si sono rivelati utili per indagare le fenocopie della CMI: la malattia di Anderson-Fabry,

la Malattia di Danon, la sindrome di Noonan, e molte altre malattie rare che possono esprimersi anche con CMI.

Nell'ambito delle CMR, nel cui contesto deve essere sempre esclusa l'amilodiosi (prima causa di cardiomiopatia restrittiva in generale, ma sempre caratterizzata da bassi voltaggi ECG periferici ed aumentati spessori ecocardiografici, mai presente in età giovanile), sono ormai chiari i fenotipi associati a difetti del gene della desmina [CMR e blocco AVB con accumuli intramiocitari di desmina (11,20)] e quelli associati a difetti del gene *TNNI3* senza blocco AVB, con scompaginamento (*disarray*) dei miociti in assenza di accumuli intrasarcoplasmatici.

### **Decisioni cliniche basate sui risultati della genotipizzazione**

Nell'ambito di *CMD da difetti del gene della LMNA* deve essere rivalutato l'uso dei betabloccanti nello scompenso, data la presenza di difetto della conduzione atrio-ventricolare; deve essere inoltre rivista l'indicazione all'impianto di defibrillatori, in quanto il rischio di aritmie potenzialmente fatali è maggiore in questo tipo di CMD rispetto ad altre; i portatori ancora sani di mutazioni devono essere incoraggiati ad attività sportive ludiche-ricreative non agonistiche, in quanto l'agonismo costituisce un fattore prognostico negativo (9, 21).

Gli atleti agonisti con ipertrofie *borderline* possono beneficiare della genotipizzazione ove non vengano raggiunti i criteri diagnostici per l'idoneità o l'esclusione su base soltanto clinica e di imaging. Lo sport diventa in questi pazienti uno dei determinanti del peggioramento dell'ipertrofia cardiaca, nonché momento di rischio per aritmie ipercinetiche potenzialmente fatali (22).

Il *timing* per l'inserimento in lista trapianto dei pazienti con malattia di Danon deve essere accuratamente vagliato ed anticipato in quanto, una volta raggiunta la fase di accumulo avanzato intramiocitario, l'evoluzione verso la fase terminale della malattia è rapida. Le famiglie con tafazzinopatie (sindrome di Barth) chiedono oggi la diagnosi prenatale: l'elevato tasso di aborti e morti neonatali impone un'accurata diagnosi ed una gestione intensiva del neonato cardiomiopatico che, superati i primi anni di vita, può raggiungere una quasi guarigione (13).

Le fenocopie di CMI associate a malattie da accumulo oggi possono essere trattate con terapia enzimatica sostitutiva (malattia di Anderson Fabry ed altre malattie da accumulo per le quali esistano terapie specifiche (per es. glicogenosi).

**Grants.** *Le ricerche svolte dal 1984 ad oggi sono state sostenute da grants del Ministero della Salute, Istituto Superiore di Sanità, Regione Lombardia, INSERM, Telethon e Fondazioni Cariplo e BRE.*

### **Bibliografia**

1. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, Dubourg O, Kühn U, Maisch B, McKenna WJ, Monserrat L, Pankuweit S, Rapezzi C, Seferovic P, Tavazzi L, Keren A. Classification of the

- Cardiomyopathies: A Position Statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 2008; 29: 270-276.
2. Zaragoza MV, Arbustini E, Narula J. Noncompaction of the left ventricle: primary cardiomyopathy with an elusive genetic etiology. *Curr Opin Pediatr*. 2007; 19: 619-627.
  3. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT and Bild DE. Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. *Circulation* 1995; 92: 785-789.
  4. Gavazzi A, Repetto A, Scelsi L, Inserra C, Laudisa ML, Campana C, Specchia C, Dal Bello B, Diegoli M, Tavazzi L, Arbustini E. Prevalence, inheritance and characteristics of familial non-X-linked dilated cardiomyopathy. *Eur Heart J* 2001; 22: 73-81.
  5. Grünig E, Tasman JA, Kucherer H, Franz W, Kubler W and Katus HA. Frequency and phenotypes of familial dilated cardiomyopathy. *J Am Coll Cardiol* 1998; 31: 186-194.
  6. Arbustini E, Diegoli M, Morbini P, Dal Bello B, Banchieri N, Pilotto A, Magani F, Grasso M, Narula J, Campana C, Gavazzi A. prevalence and characteristics of dystrophin defects in adult male patients with dilated cardiomyopathy. *J Am Coll Cardiol* 2000; 35: 1760-1768.
  7. Arbustini E, Diegoli M, Fasani R, Grasso M, Morbini P, Banchieri N, Bellini O, Dal Bello B, Porcu E, Pilotto A, Magrini G, Campana C, Pellegrini C, Fortina P, Gavazzi A, Viganò M. Mitochondrial DNA mutations and mitochondrial abnormalities in dilated cardiomyopathy. *Am J Pathol* 1998; 153: 1501-1510.
  8. Repetto A, Serio A, Pasotti M, Fontana, A, Bertoletti, A, Scelsi L, Magrini G, Monti L, Campana C, Ghio S, Tavazzi L, Arbustini E. Rescreening of "healthy" relatives of patients with dilated cardiomyopathy identifies subgroups at risk of developing the disease. *Eur Heart J*, 2004, S6: F54-F60.
  9. Arbustini E, Pilotto A, Repetto A, Grasso M, Negri A, Diegoli M, Campana C, Scelsi L, Baldini E, Gavazzi A, Tavazzi L. Autosomal dominant dilated cardiomyopathy with atrioventricular block: a lamin A/C defect-related disease. *J Am Coll Cardiol* 2002; 39: 981-990.
  10. Pilotto A, Marziliano N, Pasotti M, Grasso M, Costante AM, Arbustini E. AlphaB-crystallin mutation in dilated cardiomyopathies: low prevalence in a consecutive series of 200 unrelated probands. *Biochem Biophys Res Commun* 2006; 346: 1115-1117.
  11. Arbustini E, Pasotti M, Pilotto A, Pellegrini C, Grasso M, Previtali S, Repetto A, Bellini O, Azan G, Scaffino M, Campana C, Piccolo G, Viganò M, Tavazzi L. Desmin accumulation restrictive cardiomyopathy and atrioventricular block associated with desmin gene defects. *Eur J Heart Fail* 2006; 8: 477-483
  12. Marziliano N, Pilotto A, Grasso M, Pasotti M, Arbustini E. Deletion of Glu

- at codon 13 of the TCAP gene encoding the titin-cap-telethonin is a rare polymorphism in a large Italian population. *Mol Genet Metab* 2006; 89: 286-287.
13. Marziliano N, Mannarino S, Nespoli L, Diegoli M, Pasotti M, Malattia C, Grasso M, Pilotto A, Porcu E, Raisaro A, Raineri C, Dore R, Maggio PP, Brega A, Arbustini E. Barth Syndrome Associated with Compound Hemizyosity and Heterozygosity of the TAZ and LDB3 Genes. *Am J Med* 2007; 143: 907-915.
  14. Arbustini E, Diegoli M, Pasotti M, Grasso M, Marziliano N, De Donato G. Gene symbol: CMD1J. Disease: SensoriNeural Hearing Loss (SNHL). *Hum Genet* 2005; 117: 297.
  15. Arbustini E, Diegoli M, Pasotti M, Grasso M, Marziliano M, Delogu A, Carminati M. Gene symbol: CMD1J. Disease: Dilated cardiomyopathy. *Hum Genet* 2005; 117: 297.
  16. Van Driest SL, Vasile VC, Ommen SR, Will ML, Tajik AJ, Gersh BJ and Ackerman MJ. Myosin binding protein C mutations and compound heterozygosity in hypertrophic cardiomyopathy. *J Am Coll Cardiol* 2004; 44: 1903-1910.
  17. Ingles J, Doolan A, Chiu C, Seidman J, Seidman C and Semsarian C. Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. *J Med Genet* 2005; 42: e59.
  18. Arbustini E, Fasani R, Morbini P, Diegoli M, Grasso M, Dal Bello B, Marangoni E, Banfi P, Banchieri N, Bellini O, Comi G, Campana C, Gavazzi A, Danesino C, Viganò M. Coexistence of mitochondrial DNA and of b-Myosin heavy chain mutations in hypertrophic cardiomyopathies with late dilating congestive evolution. *Heart* 1998; 80: 548-558.
  19. Niimura H, Bachinski LL, Sangwatanaroj S, Watkins H, Chudley AE, McKenna W, Kristinsson A, Roberts R, Sole M, Maron BJ, Seidman JG and Seidman CE. Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *N Engl J Med* 1998; 338: 1248-1257.
  20. Arbustini E, Morbini P, Grasso M, Fasani R, Verga L, Bellini O, Dal Bello B, Campana C, Piccolo G, Febo O, Opasich C, Gavazzi A, Ferrans VJ. Restrictive cardiomyopathy, atrio-ventricular block and mild to subclinical myopathy in patients with desmin-immunoreactive material deposits. *J Am Coll Cardiol* 1998; 31: 645-653.
  21. Pasotti M, Klersy C, Pilotto A, Marziliano N, Rapezzi C, Serio A, Mannarino S, Gambarin F, Favalli V, Grasso M, Agozzino M, Campana C, Gavazzi A, Febo O, Marini M, Landolina M, Mortara A, Piccolo G, Viganò M, Tavazzi L, Arbustini E. Long-term outcome and risk stratification in dilated cardiomyopathies. *J Am Coll Cardiol* 2008, in press.
  22. Maron BJ and Pelliccia A. The heart of trained athletes: cardiac remodeling and the risks of sports, including sudden death. *Circulation* 2006; 114: 1633-44.

## Sindromi aritmogeniche ereditarie

Silvia G. Priori

IRCCS Fondazione Salvatore Maugeri, U.O. di Cardiologia Molecolare, Pavia

A partire dai primi anni Novanta, l'applicazione delle tecniche di biologia molecolare nel campo della cardiologia clinica ha permesso di identificare l'origine genetica di alcune forme di morte cardiaca improvvisa (MCI) in giovani con cuore strutturalmente sano. Grazie al contributo della genetica è stato scoperto che mutazioni su geni che codificano per canali ionici cardiaci possono provocare un'alterazione nell'equilibrio delle correnti elettriche responsabili del controllo del potenziale d'azione cardiaco e facilitare l'insorgenza di tachicardia e fibrillazione ventricolare in assenza di anomalie strutturali del cuore. Queste scoperte hanno permesso di identificare diverse entità cliniche associate a rischio di MCI su base genetica, quali la sindrome del QT lungo (LQTS), la sindrome di Brugada (SB), la Tachicardia Ventricolare Polimorfica Catecolaminergica (CPVT), la sindrome del QT Corto (SQTS). Diversi geni sono stati finora associati a queste patologie ed è attualmente evidente che mutazioni diverse su uno stesso gene possono provocare sindromi differenti (Tab. 1). In questo articolo verrà discusso il

**Tab. 1** - I geni implicati nell'eziopatogenesi delle patologie aritmogene ereditarie. Si noti come mutazioni diverse su uno stesso gene possono causare fenotipi diversi e come generalmente più di un gene sia implicato nell'eziologia di uno stesso fenotipo clinico.

Patologia	Locus cromosomico	Ereditarietà	Gene	Proteina
LQT1	11p15.5	AD	KCNQ1	subunità $\alpha$ canale ionico per IKs
LQT2	7q53-q36	AD	KNCH2	subunità $\alpha$ canale ionico per IKr
LQT3	3p21	AD	SCN5A	subunità $\alpha$ canale ionico per corrente INa
LQT4	4q25-q27	AD	ANK2	ankirina B
LQT5	21q22.1-q22.2	AD	KNCE1	subunità $\beta$ canale ionico per Iks (minK)
LQT6	21q22.1-q22.2	AD	KNCE2	subunità $\beta$ canale ionico per IKr
LQT7	17q23.1-q24.2	AD	KNCJ2	canale ionico per IK1
LQT8	12p13.3	AD	CACNA1C	canale tipo L per il $Ca^{2+}$
JLN1	11p15.5	AR	KCNQ1	subunità $\alpha$ canale ionico per IKs
JLN2	21q22.1-q22.2	AR	KNCE1	subunità $\beta$ canale ionico per IKs
BRS1	3p21	AD	SCN5A	subunità $\alpha$ canale ionico per corrente INa
SQTS1	7q53-q36	AD	KNCH2	subunità $\alpha$ canale ionico per IKr
SQTS2	11p15.5	AD	KCNQ1	subunità $\alpha$ canale ionico per IKs
SQTS3	17q23.1-q24.2	AD	KNCJ2	canale ionico per IK1
CPVT1	1q42.1-q42.3	AD	hRyR2	recettore cardiaco per la rianodina
CPVT2	1p11-p13.3	AR	CASQ2	calsequestrina cardiaca

ruolo assunto dalla genetica molecolare in cardiologia, non soltanto nel permettere la diagnosi e il riconoscimento dei pazienti affetti da tali patologie, ma anche nell'identificare nuovi parametri prognostici e nuove strategie terapeutiche basati sulla presenza di singoli difetti genetici.

### **La diagnosi genetica e l'importanza di identificare i portatori silenti di mutazione**

Le malattie aritmogene ereditarie sono, generalmente, associate a caratteristiche cliniche ed ECGrafiche che permettono di sospettare la diagnosi negli individui affetti. Per esempio, la presenza di un intervallo QTc prolungato e associato ad alterazioni della ripolarizzazione è la caratteristica tipica della LQTS; la presenza di soprasslivellamento del tratto ST >2 mm nelle derivazioni precordiali destre e talora di un BBD completo od incompleto è patognomonico per la SB; un intervallo QT di durata breve e la presenza di onde T alte e appuntite si associa alla SQTS. Tuttavia, fin dalle prime descrizioni di queste malattie, è apparso evidente che non tutti i pazienti affetti da difetto genetico presentano il fenotipo clinico. Nel caso della LQTS, circa il 30% dei pazienti portatori di mutazione genetica ha un intervallo QT nei limiti della norma (1). La penetranza dell'intervallo QTc come caratteristica fenotipica della malattia, ovvero la percentuale di pazienti con intervallo QTc  $\geq 440$  msec se uomini e  $\geq 460$  msec se donne, risulta pari soltanto al 70% (1). Tuttavia, studi su un vasto gruppo di individui con LQTS hanno dimostrato che, anche se il loro ECG risulta normale, i portatori silenti di mutazione genetica per LQTS hanno un rischio del 15-20% di diventare sintomatici per sincope o arresto cardiaco prima dei 40 anni di età, se lasciati privi di terapi (2). Inoltre, è noto che l'esercizio fisico agonistico, l'ipokaliemia o l'assunzione di farmaci non cardiovascolari associati ad un rischio di prolungamento dell'intervallo QTc vanno evitati dagli individui affetti da LQTS. La diagnosi genetica rappresenta l'unico strumento per riconoscere i portatori silenti di LQTS. In questo modo, tali individui possono proteggersi da situazioni per loro pericolose e sono resi consapevoli del rischio di trasmettere la mutazione alla propria progenie.

La LQTS non è la sola malattia aritmogena ereditaria in cui la diagnosi genetica gioca un ruolo fondamentale per la gestione del paziente. Anche nella Sindrome di Brugada ad esempio, l'analisi genetica ha un ruolo primario per confermare la diagnosi in situazioni cliniche dubbie quali ad esempio quello rappresentate da un profilo ECGrafiche simili a quadri di ripolarizzazione precoce frequenti nei giovani. Nella valutazione dei familiari di soggetti con diagnosi di Sindrome di Brugada l'analisi genetica aiuta il clinico ad identificare i soggetti a rischio in quanto, come per la Sindrome del QT Lungo, un ECG normale non permette di escludere la presenza del difetto genetico e quindi di un rischio aritmico. Come è noto la diagnosi di SB può essere smascherata in soggetti con ECG basale nella norma da un test provocativo farmacologico con somministrazione di flecainide (3). Tuttavia, il nostro gruppo ha dimostrato che il potere predittivo positivo del test con flecainide non è così alto come considerato in passato e comunque non è



certamente prossimo al 100%, pertanto il test genetico ha un importante ruolo nella identificazione di questi soggetti (4).

Un altro esempio dell'utilità della diagnosi genetica nella gestione clinica del paziente e dei suoi famigliari è riscontrabile nella CPVT.

La CPVT è un'altra patologia aritmogena ereditaria associata a rischio di sincope e MCI durante esercizio fisico o stress emotivo (5). Il quadro aritmico tipico e pressoché diagnostico è rappresentato da una tipica tachicardia ventricolare con pattern bidirezionale, che può facilmente degenerare in tachicardia ventricolare polimorfica o fibrillazione ventricolare. La tachicardia ventricolare bidirezionale è facilmente riproducibile durante test da sforzo (5, 6). Tuttavia l'ECG basale dei pazienti affetti da CPVT non presenta alterazioni patognomiche (5). Prima dell'identificazione dei geni responsabili per la malattia, tali pazienti sono stati spesso considerati come affetti da LQTS con intervallo QTc normale. Tuttavia, a differenza dei portatori silenti di LQTS, la prognosi dei pazienti con CPVT è particolarmente maligna, in quanto il rischio di MCI in assenza di terapia è pari al 30% prima dei 40 anni d'età (6). Anche in questo caso, la diagnosi genetica si è rivelata fondamentale per identificare questi pazienti, che sono considerati ad alto rischio di MCI e in cui un approccio terapeutico aggressivo è consigliato anche negli individui ancora asintomatici al momento della diagnosi.

### **Le correlazioni genotipo-fenotipo e il loro ruolo nella stratificazione prognostica e nell'approccio terapeutico**

Diversi geni possono causare la stessa patologia aritmogena ereditaria. Sono noti finora 8 diversi geni responsabili di LQTS (7-14), 3 geni responsabili di SQTs (14-17), 2 geni associati a CPV (5, 18) (Tab 1). La LQTS è stata la prima delle patologie aritmogene ereditarie, le cui basi genetiche sono state identificate. Grazie alla collaborazione di diversi gruppi internazionali è stato possibile creare ampi registri di pazienti in Europa e negli Stati Uniti. La disponibilità di una popolazione relativamente numerosa di soggetti affetti da LQTS ha permesso di applicare i risultati dell'analisi genetica nell'identificazione di correlazioni genotipo-fenotipo (19), utilizzabili per la stratificazione del rischio (2) e l'approccio terapeutico (20). Questi studi si sono principalmente focalizzati sulle tre più frequenti varianti di LQTS. La LQTS di tipo 1, o LQT1, è causata da mutazioni sul gene *KNCQ1*, che codifica per il canale ionico per la corrente cardiaca di potassio IKs (8). Mutazioni sul gene *KNCH2*, che codifica per il canale ionico per la corrente di potassio IKr, causano la variante LQT2 (9). Invece, la forma LQT3 è associata a mutazioni sul gene *SCN5A*, che codifica per il canale ionico per la corrente depolarizzante di sodio, I<sub>Na</sub> (7). Un aumento del tono adrenergico è considerato il principale fattore scatenante dei sintomi in questa patologia. Tuttavia, uno studio (19) su un'ampia popolazione di pazienti affetti da queste tre varianti ha rivelato che pazienti con diverso difetto genetico sono più sensibili a fattori scatenanti differenti. I pazienti con LQT1 sono a maggior rischio di sintomi durante esercizio fisico e, principalmente, durante il nuoto. Lo stress emotivo e i rumori improvvisi rappresentano il maggior fattore scatenante i sintomi nei



pazienti affetti da LQT2. Invece, i pazienti affetti da LQT3 rappresentano un gruppo particolare nell'ambito della malattia, in quanto è risultato evidente che sono a maggior rischio di sintomi a riposo o durante il sonno (19). Queste informazioni hanno importanti implicazioni cliniche, in quanto possono orientare il cardiologo ad essere più risoluto nell'interruzione dell'attività sportiva agonistica in un giovane con LQT1 rispetto ad un paziente affetto da LQT3, così come a suggerire di evitare la presenza di sorgenti di rumore improvviso, quali il telefono o la sveglia, nella stanza di un paziente con LQT2. Un recente studio su un gruppo di 647 soggetti geneticamente affetti da LQTS ha dimostrato che il genotipo rappresenta anche uno dei principali strumenti per la stratificazione prognostica (2). I dati pubblicati dal nostro gruppo hanno evidenziato che i pazienti affetti da LQT2 o LQT3 hanno un rischio maggiore di essere sintomatici rispetto ai pazienti affetti da LQT1 (2). Lo stesso studio ha rivelato che il genotipo, insieme alla durata dell'intervallo QTc e al sesso rappresenta, in analisi multivariata, un fattore indipendente di rischio di sintomi (2).

Recentemente, il nostro gruppo ha investigato la possibilità che il genotipo possa assumere un ruolo fondamentale anche in termini di risposta alla terapia (20). Nella LQTS, i betabloccanti rappresentano la terapia di prima scelta in quanto si sono rivelati molto efficaci. Tuttavia, è noto (21) che l'incidenza di recidive in terapia, in particolare nei pazienti già sintomatici per arresto cardiaco, è considerevole. Lo studio effettuato dal nostro gruppo, ha rivelato che l'efficacia anti-aritmica dei betabloccanti per la LQTS è influenzata dal genotipo (20). Il genotipo LQT2 e LQT3, infatti, rappresentato un fattore di rischio di recidive nonostante i betabloccanti, insieme ad altri parametri clinici, quali un intervallo QTc >500 msec o l'insorgenza di sintomi prima dei 7 anni di età (20).

I risultati della diagnosi genetica, pertanto, si sono rivelati fondamentali del predire la storia naturale della malattia in un soggetto portatore, rappresentando, così, un nuovo strumento nelle mani del cardiologo per una migliore gestione del paziente. Almeno nel caso della LQTS, il genotipo può essere considerato dai cardiologi come uno dei parametri utili nella definizione del profilo clinico dei pazienti, come qualunque altro dato strumentale, dall'ecocardiografia alla stimolazione elettrica programmata.

Gli studi genotipo-fenotipo e l'uso delle informazioni genetiche nella stratificazione prognostica non sono ancora applicabili ad altre patologie aritmogene ereditarie. In alcuni casi, come per la SB, un solo gene-malattia è tuttora noto (22) e in grado di identificare solo il 20-25% dei soggetti con diagnosi clinica. In altri casi, come nel caso della SQTS, anche se tre geni-malattia sono stati associati alla sindrome, il numero esiguo di pazienti finora giunti all'attenzione clinica non permette di effettuare studi di popolazione.

Un ulteriore tentativo di correlazioni genotipo-fenotipo è stato svolto dal nostro gruppo nei pazienti affetti da CPVT (6). Un solo gene-malattia è noto per la forma autosomica dominante della malattia: il gene *hRyR2* che codifica per il recettore per la rianodina RyR2. Tuttavia, a differenza di quanto si è verificato per la SB, la ricerca di mutazioni sul gene *hRyR2* permette di identificare circa il 70% dei pazienti con sospetto clinico di malattia. Il confronto tra un gruppo di pazien-

ti affetti da CPVT e geneticamente affetti ed un gruppo con diagnosi clinica ma con genotipo negativo ha dimostrato che la presenza di una mutazione sul gene *hRyR2* correla con un'insorgenza di sintomi precoce (6). Inoltre lo stesso studio ha rivelato che i maschi geneticamente affetti hanno un rischio aumentato di divenire sintomatici rispetto alle femmine<sup>6</sup>. Pertanto, considerando l'alta malignità della malattia, la consapevolezza che un paziente di sesso maschile è portatore di mutazione sul gene *hRyR2* può rivelarsi di aiuto al cardiologo nell'orientarsi verso una scelta terapeutica più aggressiva.

### Direzioni future

L'identificazione dei geni responsabili di diverse patologie aritmogene ha permesso di utilizzare le informazioni derivate dalla genetica per delineare nuove strategie nell'approccio clinico e terapeutici nei pazienti affetti. Un passo futuro nell'interazione tra cardiologia clinica e biologia molecolare è rappresentato dalla possibilità di rendere le tecniche molecolari più accessibili per un uso clinico e non esclusivamente limitate alla ricerca di base. Uno dei principali ostacoli all'uso su larga scala della analisi genetica è rappresentato dal costo di queste tecniche applicate alle malattie aritmogene in cui molto spesso i pazienti risultano portatori di mutazioni "private" e pertanto l'analisi genetica richiede lo screening di tutta la sequenza codificante di ciascun gene. Recentemente uno studio del nostro gruppo<sup>4</sup> ha, tuttavia, dimostrato che almeno nel caso della LQTS, circa il 50% dei pazienti risulta portatore di mutazioni su un gruppo di 64 codoni genici sui tre principali geni noti. Questa scoperta ha permesso di sviluppare un approccio diagnostico in tre fasi, che permette di restringere l'analisi genetica, almeno in fase iniziale, su questo ristretto gruppo di codoni genici, riducendo tempi e costi dell'analisi molecolare. La possibilità di effettuare diagnostica genetica in tempi brevi e costi limitati è il primo passo per rendere più accessibile in tutti i centri clinici l'identificazione dei pazienti affetti.

Questo nuovo approccio è, al momento, proponibile soltanto per la LQTS, in quanto questa è l'unica patologia aritmogena ereditaria in cui una vasta popolazione di pazienti è stata identificata. La possibilità di estendere gli studi di correlazioni genotipo-fenotipo e di stratificazione prognostica alle altre patologie aritmogene si basa sulla consapevolezza che un'ampia collaborazione internazionale tra diversi gruppi di ricerca può ridurre i tempi per raccogliere dati su ampie popolazioni di soggetti affetti da SB, SQTS o CPVT e la scoperta di nuovi genimaleattie. Inoltre, il futuro dell'interazione tra biologia molecolare e cardiologia clinica è rappresentato dalla possibilità di individuare strategie terapeutiche di tipo gene-specifiche o mutazione-specifiche. Studi preliminari hanno già permesso di introdurre, in casi selezionati, tentativi di terapie basate sul genotipo, quali l'uso di mexiletina in alcuni pazienti affetti da LQT3 (23-26) o di risparmiatori di potassio per alcuni pazienti con LQT2 (2). Una crescente collaborazione tra cardiologi clinici, che diagnosticano e curano tali pazienti, e ricercatori di base, che investigano a livello molecolare le possibilità di terapia genica, rappresenta il più valido strumento per raggiungere una migliore strategia di gestio-

ne nell'ambito di queste patologie, il cui costo in termini di vite umane è ancora troppo alto.

## Bibliografia

1. Napolitano C., Priori SG., Schwartz PJ., Bloise R., Ronchetti E., Nastoli J., Bottelli G., Cerrone M., Leonardi S. Genetic testing in the long QT syndrome: development and validation of an efficient approach to genotyping in clinical practice. *JAMA* 2005; 294: 2975-80.
2. Priori SG., Schwartz PJ., Napolitano C., Bloise R., Ronchetti E., Grillo M., Vicentini A., Spazzolini C., Nastoli J., Bottelli G., Folli R., Cappelletti D. Risk stratification in the long-QT syndrome. *N Engl J Med* 2003; 348: 1866-74.
3. Brugada R., Brugada J., Antzelevitch C., Kirsch GE., Potenza D., Towbin JA., Brugada P. Sodium channel blockers identify risk for sudden death in patients with ST-segment elevation and right bundle branch block but structurally normal hearts. *Circulation* 2000; 101: 510-5.
4. Priori SG., Napolitano C., Gasparini M., Pappone C., Della Bella P., Brignole M., Giordano U., Giovannini T., Menozzi C., Bloise R., Crotti L., Terreni L., Schwartz PJ. Clinical and genetic heterogeneity of right bundle branch block and ST-segment elevation syndrome: A prospective evaluation of 52 families. *Circulation* 2000; 102: 2509-15.
5. Priori SG., Napolitano C., Tiso N., Memmi M., Vignati G., Bloise R., Sorrentino V., Danielli GA. Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation* 2001; 103: 196-200.
6. Priori SG., Napolitano C., Memmi M., Colombi B., Drago F., Gasparini M., DeSimone L., Coltorti F., Bloise R., Keegan R., Cruz Filho FE., Vignati G., Benatar A., DeLogu A. Clinical and molecular characterization of patients with catecholaminergic polymorphic ventricular tachycardia. *Circulation* 2002; 106: 69-74.
7. Wang Q., Shen J., Splawski I., Atkinson D., Li Z., Robinson JL., Moss AJ., Towbin JA., Keating MT. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell* 1995; 80: 805-11.
8. Wang Q., Curran ME., Splawski I., Burn TC., Millholland JM., VanRaay TJ., Shen J., Timothy KW., Vincent GM., de Jager T., Schwartz PJ., Towbin JA., Moss AJ., Atkinson DL., Landes GM., Connors TD, Keating MT: Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat Genet* 1996; 12: 17-23.
9. Curran ME., Splawski I., Timothy KW., Vincent GM., Green ED., Keating MT. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* 1995; 80: 795-803.
10. Mohler PJ., Schott JJ., Gramolini AO., Dilly KW., Guatimosim S., duBell WH., Song LS., Haurogne K., Kyndt F., Ali ME., Rogers TB., Lederer WJ., Escande D., Le Marec H., Bennett V. Ankyrin-B mutation causes type 4

- long-QT cardiac arrhythmia and sudden cardiac death. *Nature* 2003; 421: 634-9.
11. Splawski I., Tristani-Firouzi M., Lehmann MH., Sanguinetti MC., Keating MT. Mutations in the hminK gene cause long QT syndrome and suppress IKs function. *Nat Genet* 1997; 17: 338-40.
  12. Abbott GW., Sesti F., Splawski I., Buck ME., Lehmann MH., Timothy KW. Keating MT., Goldstein SA. MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* 1999; 97: 175-87.
  13. Tristani-Firouzi M., Jensen JL., Donaldson MR., Sansone V., Meola G., Hahn A., Bendahhou S., Kwiecinski H., Fidzianska A., Plaster N., Fu YH., Ptacek LJ., Tawil R. Functional and clinical characterization of KCNJ2 mutations associated with LQT7 (Andersen syndrome). *J Clin Invest* 2002; 110: 381-8.
  14. Splawski I., Timothy KW., Sharpe LM., Decher N., Kumar P., Bloise R., Napolitano C., Schwartz PJ., Joseph RM. Condouris K., Tager-Flusberg H., Priori SG., Sanguinetti MC., Keating MT. Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* 2004; 119: 19-31.
  15. Brugada R., Hong K., Dumaine R., Cordeiro J., Gaita F., Borggrefe M., Menendez TM., Brugada J., Pollevick GD., Wolpert C., Burashnikov E., Matsuo K., Wu YS., Guerschicoff A., Bianchi F., Giustetto C., Schimpf R., Brugada P., Antzelevitch C. Sudden death associated with short-QT syndrome linked to mutations in HERG. *Circulation* 2004; 109: 30-5.
  16. Bellocq C., van Ginneken AC., Bezzina CR., Alders M., Escande D., Mannens MM., Baro I., Wilde AA. Mutation in the KCNQ1 gene leading to the short QT-interval syndrome. *Circulation* 2004; 109: 2394-7.
  17. Priori SG., Pandit SV., Rivolta I., Berenfeld O., Ronchetti E., Dharamoon A., Napolitano C., Anumonwo J., di Barletta MR., Gudapakkam S., Bosi G., Stramba-Badiale M., Jalife J. A novel form of short QT syndrome (SQT3) is caused by a mutation in the KCNJ2 gene. *Circ Res* 2005; 96: 800-7.
  18. Lahat H., Pras E., Olender T., Avidan N., Ben-Asher E., Man O., Levy-Nissenbaum E., Khoury A., Lorber A., Goldman B., Lancet D., Eldar M. A missense mutation in a highly conserved region of CASQ2 is associated with autosomal recessive catecholamine-induced polymorphic ventricular tachycardia in Bedouin families from Israel. *Am J Hum Genet* 2001; 69: 1378-84.
  19. Schwartz PJ., Priori SG., Spazzolini C., Moss AJ., Vincent GM., Napolitano C., Denjoy I., Guicheney P., Breithardt G., Keating MT., Towbin JA., Beggs AH., Brink P., Wilde AA., Toivonen L., Zareba W., Robinson JL., Timothy KW., Corfield V., Wattanasirichaigoon D., Corbett C., Haverkamp W., Schulze-Bahr E., Lehmann MH., Schwartz K., Coumel P., Bloise R. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. *Circulation* 2001; 103: 89-95.
  20. Priori SG., Napolitano C., Schwartz PJ., Grillo M., Bloise R. Ronchetti E., Moncalvo C., Tulipani C., Veia A., Bottelli G., Nastoli J. Association of long

- QT syndrome loci and cardiac events among patients treated with beta-blockers. *JAMA* 2004; 292: 1341-4.
21. Moss AJ., Zareba W., Hall WJ., Schwartz PJ., Crampton RS., Benhorin J., Vincent GM., Locati EH., Priori SG., Napolitano C., Medina A., Zhang L., Robinson JL., Timothy K., Towbin JA., Andrews ML. Effectiveness and limitations of beta-blocker therapy in congenital long-QT syndrome. *Circulation* 2000; 101: 616-23.
  22. Chen Q., Kirsch GE., Zhang D., Brugada R., Brugada J., Brugada P., Potenza D., Moya A., Borggrefe M., Breithardt G., Ortiz-Lopez R., Wang Z., Antzelevitch C., O'Brien RE., Schulze-Bahr E., Keating MT., Towbin JA., Wang Q. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* 1998; 392: 293-6.
  23. Schwartz PJ., Priori SG., Locati EH., Napolitano C., Cantu F., Towbin JA., Keating MT., Hammoude H., Brown AM., Chen LS. Long QT syndrome patients with mutations of the SCN5A and HERG genes have differential responses to Na<sup>+</sup> channel blockade and to increases in heart rate. Implications for gene-specific therapy. *Circulation* 1995; 92: 3381-6.
  24. Schwartz PJ., Priori SG., Dumaine R., Napolitano C., Antzelevitch C., Stramba-Badiale M., Richard TA., Berti MR., Bloise R. A molecular link between the sudden infant death syndrome and the long-QT syndrome. *N Engl J Med* 2000; 343: 262-7.
  25. Schulze-Bahr E., Fenge H., Etzrodt D., Haverkamp W., Monnig G., Wedekind H., Breithardt G., Kehl HG. Long QT syndrome and life threatening arrhythmia in a newborn: molecular diagnosis and treatment response. *Heart* 2004; 90: 13-6.
  26. Kehl HG., Haverkamp W., Rellensmann G., Yelbuz TM., Krasemann T., Vogt J., Schulze-Bahr E. Images in cardiovascular medicine. Life-threatening neonatal arrhythmia: successful treatment and confirmation of clinically suspected extreme long QT-syndrome-3. *Circulation* 2004; 109: e205-6.
  27. Compton SJ., Lux RL., Ramsey MR., Strellich KR., Sanguinetti MC., Green LS. Keating MT., Mason JW. Genetically defined therapy of inherited long-QT syndrome. Correction of abnormal repolarization by potassium. *Circulation* 1996; 94: 1018-22.

## Genomica delle piastrine e trombofilia

*Ilaria Canobbio, Cesare Balduini*

Dipartimento di Biochimica "A. Castellani", Sezione Scienze MM.FF.NN., Università di Pavia

Le piastrine sono piccoli frammenti cellulari anucleati (2-3  $\mu\text{m}$ ), derivati dai megacariociti, e circolanti nel sangue umano ad una concentrazione di 250.000-350.000 cellule per  $\mu\text{l}$  di sangue. La principale funzione delle piastrine è quella di mantenere l'integrità vascolare e partecipare ai processi della coagulazione e della emostasi primaria. Le piastrine sono infatti in grado di riconoscere una lesione sulla superficie di un vaso sanguigno, di aderire alla matrice sottoendoteliale esposta attraverso specifici recettori di membrana e di attivare una complessa cascata di eventi che culmina con l'attivazione e l'aggregazione piastrinica e la formazione del tappo emostatico (1). Le piastrine sono inoltre coinvolte nei processi infiammatori e nella risposta immunitaria (2, 3), nelle infezioni batteriche e virali (4), nei fenomeni di aterosclerosi (5), e nella progressione dei tumori (6). Un'alterata funzionalità conseguente ad un anomalo numero di piastrine, ad irregolarità morfologiche, ad alterata espressione o malfunzionamento dei recettori di membrana o delle proteine coinvolte nella trasduzione del segnale, può essere alla base di gravi difetti dell'emostasi quali le patologie emorragiche da un lato e le patologie tromboemboliche dall'altro. Per questi motivi appare evidente che le piastrine rivestono un'importanza fondamentale in numerosi processi fisiologici e giustificano il crescente interesse per questi elementi del sangue nella fisiopatologia di molteplici malattie.

In questo breve studio verranno prima di tutto presi in esame i principali recettori piastrinici e i loro rispettivi ligandi per comprendere i fondamentali meccanismi di attivazione piastrinica in condizioni fisiologiche. In seguito verrà analizzato in dettaglio il ruolo delle piastrine nelle malattie trombotiche, verranno discusse le principali osservazioni cliniche in proposito, ed in particolare verrà studiata la correlazione tra mutazioni a livello del genoma piastrinico e predisposizione all'insorgenza di trombosi.

### **Principali recettori piastrinici e loro ligandi**

Le piastrine esprimono diversi recettori specifici di membrana. Essi svolgono un ruolo molto importante nella funzionalità piastrinica in quanto sono in grado di interagire sia con proteine adesive del sottoendotelio esposte in seguito ad un danno nella parete vasale, sia con agonisti extracellulari in grado di promuovere

l'attivazione e la conseguente aggregazione piastrinica (7). Molti fenomeni di adesione cellulare sono mediati dalle *integrine* (8). Le integrine sono eterodimeri transmembrana costituiti da una subunità  $\alpha$  ed una subunità  $\beta$ .

associate non covalentemente a formare dimeri con attività recettoriale per svariate proteine adesive. Il dominio extracellulare contiene la regione di legame con il ligando. Il dominio citoplasmatico di entrambe le subunità è in grado di interagire con le proteine del membranoscheletro e del citoscheletro e di mediare i segnali provenienti dal versante extracellulare con la riorganizzazione interna della cellula (9). La più abbondante integrina espressa sulla membrana delle piastrine è l'integrina  $\alpha_{IIb}\beta_3$  (10). In seguito all'attivazione questo recettore media l'interazione delle piastrine con il fibrinogeno ed è essenziale nel promuovere il processo di aggregazione tra piastrine adiacenti. L'integrina  $\alpha_{IIb}\beta_3$  è inoltre in grado di legare il fattore von Willebrand (VWF), la fibronectina e la trombospodina. L'integrina  $\alpha_2\beta_1$  media l'adesione delle piastrine al collagene e ne promuove l'attivazione (11). L'integrina  $\alpha_v\beta_3$  rappresenta il recettore piastrinico per la vitronectina, mentre l'integrina  $\alpha_2\beta_5$  media il legame alla fibronectina. La laminina viene riconosciuta specificatamente dall'integrina  $\alpha_2\beta_6$ .

Accanto ai recettori per l'adesione delle piastrine alle componenti del sottoendotelio, la membrana plasmatica di queste cellule presenta numerose altre glicoproteine che agiscono da recettori per ligandi in grado di indurre l'attivazione (agonisti) o l'inibizione (antagonisti) piastrinica. Questi ligandi possono essere molecole solubili oppure strutture multimeriche. I recettori per gli agonisti solubili tra cui la trombina (PAR1, PAR4), il trombossano  $A_2$  ( $TxA_2R\alpha/\beta$ ), l'ADP (P2Y1, P2Y12), la serotonina (5-HT<sub>2</sub>) e altri, sono recettori a sette domini transmembrana accoppiati a proteine G (GPCR), e mediano l'attivazione attraverso un complesso sistema di effettori e secondi messaggeri (calcio, cAMP, fosfoinositidi) (12, 13).

Le strutture multimeriche interagiscono con specifiche glicoproteine e complessi glicoproteici (GP), fisicamente eterogenei che promuovono l'attivazione attraverso l'attivazione di recettori per il frammento cristallizzabile delle immunoglobuline G (Fc $\gamma$ R) contenenti sequenze ITAM di attivazione a cui sono fisicamente associati e la fosforilazione di proteine segnale a valle. Il collagene, oltre che con l'integrina  $\alpha_2\beta_1$ , interagisce con il complesso GPVI-Fc  $\gamma$ R promuovendo l'attivazione cellulare (14). Il VWF interagisce con il complesso GPIb-IX-V a livello della subunità  $\alpha$  della GPIb. Il legame del VWF al GPIb-IX-V media l'adesione e l'attivazione piastrinica in condizioni di elevato flusso ematico attraverso l'attivazione del recettore a bassa affinità per le IgG, Fc $\gamma$ RIIA (15).

Le principali caratteristiche dei recettori piastrinici e dei rispettivi ligandi sono riportati in tabella 1.

L'importanza dei recettori nella funzionalità piastrinica è confermata dal fatto che numerose patologie emorragiche sono causate da una diminuita espressione o dal malfunzionamento degli stessi. Tra queste la trombostenia di Glanzmann e la Sindrome di Bernard-Soulier sono malattie rare dovute alla mancanza rispettivamente del recettore per il fibrinogeno, l'integrina  $\alpha_{IIb}\beta_3$ , e del recettore per il VWF, complesso GP Ib-IX-V.



**Tab. 1 -** Recettori piastrinici e loro principali caratteristiche.

<b>Proteine di membrana</b>	<b>Note caratteristiche</b>	<b>Ligandi</b>	<b>Funzione</b>
<b>Integrine</b>			
Integrina $\alpha_{IIb}\beta_3$	GPIIb-IIIa	Fibrinogeno (VWF, Fibronectina, Vitronectina, Trombospondina)	Adesione stabile, Aggregazione, Attivazione
Integrina $\alpha_2\beta_1$	GPIa-IIa	Collagene	Adesione
Integrina $\alpha_3\beta_1$	GPIc-IIa	Fibronectina	Adesione
Integrina $\alpha_6\beta_1$	GPIc'-IIa	Laminina	Adesione
Integrina $\alpha_v\beta_3$	GPV-IIIa	Vitronectina	Adesione?
<b>Glicoproteine recettoriali di membrana</b>			
GP Ib-IX-V	Complesso glicoproteico (famiglia Leucine-rich-repeat)	VWF (Trombina)	Adesione (alto flusso) Attivazione
GP VI	Glicoproteina (famiglia Immunoglobuline)	Collagene	Attivazione
GP IV	Glicoproteina (CD 36)	Collagene (Trombospondina)	Adesione
<b>Immunorecettori</b>			
FcR $\gamma$ -chain	Subunità di 14 kDa associata a immunorecettori e alla GPVI	-	Attivazione
Fc $\gamma$ RIIA	Immunorecettore	Immunocomplessi	Attivazione
<b>Recettori accoppiati a proteine G (GPCR)</b>			
PAR-1	Rec. accoppiato a $G_q$ e a $G_{12}$	Trombina	Attivazione
PAR-4	Rec. accoppiato a $G_q$ e a $G_{12}$	Trombina	Attivazione
TxR $\alpha$	Rec. accoppiato a $G_q$	TxA <sub>2</sub> (PGH <sub>2</sub> )	Attivazione
TxR $\beta$	Rec. accoppiato a $G_q$	TxA <sub>2</sub> (PGH <sub>2</sub> )	Attivazione
$\alpha_2$ -adrenergico	Rec. accoppiato a $G_i$	Epinefrina	Attivazione
P2Y <sub>1</sub>	Rec. accoppiato a $G_q$	ADP	Attivazione
P2Y <sub>12</sub>	Rec. accoppiato a $G_{12}$ e a $G_i$	ADP	Attivazione
5-HT <sub>2</sub>	Rec. accoppiato a $G_q$	Serotonina	Attivazione

### Adesione, attivazione e aggregazione piastrinica

Le piastrine esercitano la loro funzione principale nei processi dell'emostasi e della coagulazione. L'esposizione delle fibrille di collagene della parete vasale danneggiata rappresenta lo stimolo iniziale per l'attivazione delle piastrine *in vivo*. Il collagene interagisce con i propri recettori specifici presenti sulla superficie piastrinica (GPVI, integrina  $\alpha_2\beta_1$ ) e questo evento ha la duplice funzione di iniziare l'adesione delle cellule al sottoendotelio e di innescare i processi di attivazione delle piastrine (16). In questa fase il VWF immobilizzato nella matrice sottoendoteliale riveste un ruolo molto importante in quanto è in grado di mediare l'adesione *in vivo* in condizioni di elevato flusso ematico. In seguito ad un danno nella parete vasale, le piastrine aderiscono rapidamente al VWF presente nel sottoendotelio attraverso il riconoscimento ed il legame con il complesso GP Ib-IX-V piastrinico. L'interazione tra la GP Ib $\alpha$  e il VWF porta alla formazione di una serie di legami crociati tra piastrine e matrice ed alla associazione del com-



plesso GP Ib-IX-V al membranoscheletro (17). Le cellule adese reversibilmente successivamente “rotolano” sulla superficie del vaso danneggiato formando contatti stabili con le integrine piastriniche ( $\alpha_{\text{Ib}}\beta_3$ ,  $\alpha_2\beta_1$ ,  $\alpha_5\beta_1$ ) e l’adesione viene rafforzata dall’interazione con le altre componenti della matrice, laminina e fibronectina (18).

L’adesione delle piastrine al collagene e al VWF promuove anche l’attivazione delle cellule con conseguente cambio di forma, secrezione dei granuli e aggregazione. Il cambio di forma consiste nell’emissione di pseudopodi e nella centralizzazione dei granuli piastrinici. Gli pseudopodi sono lunghe estensioni cellulari che contengono elementi citoscheletrici come l’actina, l’ABP, l’ $\alpha$ -actinina e i microtubuli. L’estensione degli pseudopodi aumenta l’area di contatto delle piastrine con il sottoendotelio e con altre piastrine e stabilizza i legami che si vanno formando. Contemporaneamente all’emissione degli pseudopodi si assiste alla migrazione dei granuli verso il centro della cellula. Questo evento è dipendente dagli ioni calcio e dalla forza contrattile promossa dal legame tra i filamenti di actina e miosina ed è associato alla perdita da parte della piastrina della forma discoidale tipica dello stato a riposo e al rimodellamento del corpo cellulare in una struttura sferica.

Un secondo evento importante che segue l’attivazione delle piastrine è l’espressione di funzioni recettoriali adesive sulla superficie delle cellule. Molti recettori adesivi sono espressi in maniera inattiva oppure in concentrazione molto bassa sulle piastrine a riposo in modo da evitare la formazione di aggregati circolanti in condizioni normali. In seguito all’attivazione essi manifestano la loro attività. Tra questi il più importante nella funzionalità piastrinica è l’integrina  $\alpha_{\text{Ib}}\beta_3$ . Come già ricordato, l’integrina  $\alpha_{\text{Ib}}\beta_3$  è in grado di legare fibrinogeno, VWF, vitronectina, fibronectina e altre proteine adesive. Il complesso è espresso nelle piastrine a riposo in forma inattiva, e l’attivazione piastrinica è richiesta per convertire il recettore in uno stato conformazionale competente per l’interazione con il fibrinogeno, attraverso un meccanismo conosciuto come attivazione come “inside-out” (19). La stimolazione piastrinica ad opera di diversi agonisti, inoltre, porta ad un aumento del numero di complessi presenti sulla superficie piastrinica, principalmente dovuto ad una redistribuzione dalle membrane del sistema di membrane interno e dei granuli intracellulari (20). Il fibrinogeno è una molecola dimerica e simmetrica e possiede due siti di legame a ciascuna estremità; questo permette il legame del fibrinogeno contemporaneamente a due recettori posti su piastrine adiacenti fungendo da ponte intracellulare e mediando l’aggregazione. In seguito all’attivazione l’integrina  $\alpha_{\text{Ib}}\beta_3$  promuove essa stessa attivazione cellulare mediante un meccanismo di segnalazione “outside-in”. L’aggregazione quindi viene rafforzata dalla partecipazione di altre molecole adesive, tra cui la trombospondina. La trombospondina è presente nel sottoendotelio e nei granuli piastrinici. Essa si lega al suo recettore specifico, la GP IV, e contemporaneamente lega il fibrinogeno associato l’integrina  $\alpha_{\text{Ib}}\beta_3$ , stabilizzando in questo modo l’interazione piastrina-piastrina (21).

Contemporaneamente al progredire dell’aggregazione le piastrine secernono il contenuto dei granuli intracellulari, rilasciando sostanze ad attività protromboti-

ca (ADP) e proteine adesive (VWF, fibrinogeno) in grado di rinforzare l'attivazione o cooperare al reclutamento di nuove cellule al trombo nascente attraverso un meccanismo di regolazione positiva. Inoltre la fusione delle membrane dei granuli con la membrana piastrinica aumenta il numero di recettori esposti e permette l'esposizione di nuove molecole di adesione, quali ad esempio la P-selectina, coinvolte nell'interazione con altri tipi cellulari tra cui neutrofilo. Sinergicamente alla secrezione dei granuli un importante evento che amplifica la risposta cellulare ed aumenta il reclutamento delle piastrine nell'aggregato è la formazione di eicosanoidi prodotti dal metabolismo dell'acido arachidonico. L'attivazione comporta la liberazione di acido arachidonico dai fosfolipidi di membrana ad opera della fosfolipasi  $A_2$ . L'acido arachidonico viene successivamente metabolizzato attraverso due vie: la prima attuata dall'enzima lipossigenasi con formazione di acido 12-idrossiperossido-5,8,10,14-eicosatetraenoico (12-HPETE) e acido 12-idrossi-5-8-10-14-eicosatetraenoico (12-HETE) che agiscono principalmente come fattori chemiotattici per i neutrofilo e partecipano ai processi di metastasi tumorale stimolando la retrazione delle cellule endoteliali e favorendo la migrazione extravasale delle cellule tumorali; e la seconda attraverso l'enzima ciclossigenasi con formazione finale di trombossano  $A_2$  ( $TxA_2$ ), un potente attivatore piastrinico.

### **Genomica piastrinica e trombofilia**

La trombofilia è la tendenza alla coagulazione eccessiva del sangue, con predisposizione a sviluppare trombosi, sia di tipo venoso che arterioso. I fenomeni trombotici rappresentano oggi la principale causa di morte nei paesi occidentali e per questo motivo emerge un crescente interesse nel definirne le cause e prevenirne l'insorgenza. La trombosi è il risultato di una serie complessa di numerosi fattori di tipo molecolare, genetico, cellulare e ambientale che interagiscono tra di loro in tempi e modi differenti non ancora del tutto compresi. I principali fattori di rischio per l'insorgenza di patologie trombotiche sono l'assunzione di contraccettivi orali, l'obesità, la gravidanza, il fumo, il diabete e gli interventi chirurgici. Tra i fattori di rischio di tipo genetico rivestono grande importanza mutazioni in specifici fattori coinvolti nella coagulazione, in particolare il fattore II (protrombina) e il fattore V (fattore V di Leiden) per i quali la correlazione genotipo-predisposizione all'insorgenza di trombosi è stata ampiamente studiata e dimostrata (per una review completa e dettagliata si veda 22, 23). In questo contesto il coinvolgimento diretto delle piastrine, principali responsabili dei fenomeni della coagulazione, nella predisposizione e nell'insorgenza di trombosi di tipo venoso e arterioso, non è ancora stato chiaramente stabilito (24). In questo paragrafo cercheremo di prendere in esame i maggiori studi clinici in cui è stata suggerita una possibile correlazione tra genomica piastrinica e trombofilia. Questi studi si sono occupati in particolare della presenza di specifici polimorfismi a livello dei geni che codificano per i principali recettori di membrana piastrinici, integrina  $\alpha_{IIb}\beta_3$ , complesso GP Ib-IX-V, e integrina  $\alpha_2\beta_1$  e della loro eventuale associazione alla predisposizione all'insorgenza di trombosi sia di tipo venoso

**Tab. 2** - Riassunto delle evidenze correlazione polimorfismo piastrinico-trombofilia.

Polimorfismo	Fenotipo	Correlazioni fenotipo malattia	Associazioni genotipo malattia
$\beta_3$ Pro33Leu	Nessun effetto sul legame al fibrinogeno, iper reattività?	Sconosciute	Inconsistenti
GPIb $\alpha$ VNTR	Sconosciuto	Sconosciute	Studi preliminari mostrano associazione
GPIb $\alpha$ C3550T	Sconosciuto	Sconosciute	Inconsistenti
Thr145Met	Sconosciuto	Sconosciute	Inconsistenti
GPIb $\alpha$ allele Kozak C	Sconosciuto	Sconosciute	Inconsistenti
$\alpha_2$ C807T	Associato con alterata espressione del recettore, alterato legame al collagene	Sconosciute	Inconsistenti

che arterioso. Nella tabella 2 sono riassunti le principali evidenze di correlazione tra polimorfismi in recettori piastrinici e trombofilia.

### Integrina $\alpha_{IIb}\beta_3$

L'integrina  $\alpha_{IIb}\beta_3$  è il principale recettore per il fibrinogeno, ed è un eterodimero  $\alpha_{IIb}$  (150kDa)  $\beta_3$  (90kDa). Una comune mutazione puntiforme nella  $\beta_3$  porta alla sostituzione di una Prolina in posizione 33 con una Leucina (allele PLA2), con un'incidenza nella popolazione europea di circa il 15%. Uno studio preliminare condotto nel 1996 dal gruppo di Weiss (25) ha dimostrato per la prima volta una correlazione tra la presenza dell'allele PLA2 e un'aumentata incidenza di rischio di infarto del miocardio (di circa 2.8 volte). Alcuni studi successivi hanno confermato l'associazione tra allele PLA2 e aumentato rischio trombotico nell'infarto del miocardio e nelle malattie coronariche (26), tuttavia un altrettanto consistente numero di studi ha confutato queste osservazioni (27, 28). Un'analisi di correlazione tra l'allele PLA2 e malattie cerebrovascolari non ha mostrato alcuna associazione. Analogamente nessuna correlazione è stata finora dimostrata fra polimorfismi in  $\beta_3$  e trombosi di tipo venoso (27).

### Complesso GP Ib-IX-V

Il complesso GP Ib-IX-V è il principale recettore per il VWF, ma è in grado di legare e mediare anche l'adesione piastrinica indotta da trombina e collagene. Esso è costituito dal prodotto di quattro distinti geni, la subunità  $\alpha$  della GPIb (140kDa) legata covalentemente alla subunità  $\beta$  (25kDa), la GPIIX (22kDa) e la GPV (82kDa), in rapporto tra loro di 2:2:2:1. La corretta espressione di GPIb $\alpha$ , GPIb $\beta$  e GPIIX, ma non di GPV, è necessaria all'espressione del complesso integro e funzionale sulla superficie piastrinica (29). Nel complesso GP Ib-IX-V sono stati identificati tre polimorfismi nel gene che codifica per la subunità  $\alpha$  della GPIb, potenzialmente coinvolti nella predisposizione alla trombosi.

Il primo è un polimorfismo di lunghezza per un numero variabile di sequenze

ripetute in tandem (VNTR) di 39 paia di basi che codifica per 13 aminoacidi nella regione extracellulare glicosilata (macroglucopetide) e porta alla formazione di 4 varianti definite D, C, B, A in base al numero crescente di ripetizioni (da 1 a 4). Uno studio condotto su 101 pazienti con malattie coronariche acute, 104 pazienti con malattie cardiovascolari, e 95 pazienti con trombosi venosa ha riportato un'associazione tra genotipo e trombosi arteriosa, ma non venosa (30).

Il secondo polimorfismo è una sostituzione puntiforme nel gene per la GPIIb $\alpha$  (C3550T) che porta alla sostituzione di un residuo di treonina in posizione 145 con una metionina (Thr145Met). Uno studio preliminare ha dimostrato un'associazione tra malattie coronariche ma non tra infarto del miocardio e questo polimorfismo (31).

Il terzo polimorfismo riguarda il sito di inizio della trascrizione (allele Kozak C). È stato dimostrato che l'allele Kozak influenza il livello di espressione di GPIIb $\alpha$  sulla superficie piastrinica (32), e che per questo motivo correla con un aumentato rischio all'insorgenza di malattie coronariche (33); tuttavia queste osservazioni non sono state confermate da ulteriori e più approfonditi studi.

### **Integrina $\alpha_2\beta_1$**

L'integrina  $\alpha_2\beta_1$  è un eterodimero  $\alpha_2$  (167 kDa)  $\beta_1$  (130kDa) e media l'adesione e l'attivazione piastrinica alla matrice collagenica (34). È stato osservato un dimorfismo nel gene che codifica per  $\alpha_2$  in posizione 807 C/T che correla con una diversa espressione del recettore sulla superficie piastrinica (C bassa espressione del recettore, T alta espressione). È stato dimostrato che il tasso di adesione delle piastrine al collagene di tipo I dipende dal numero di recettori  $\alpha_2\beta_1$  esposti sulla superficie piastrinica (35). In particolare è stata osservata un'associazione tra l'allele 807T e l'incidenza di infarto del miocardio in soggetti inferiori ai 62 anni di età (36). Studi su un'ampia casistica non hanno tuttavia confermato la correlazione tra allele 807T e un'aumentata incidenza di infarto del miocardio (37).

### **Conclusioni**

Le malattie tromboemboliche comprendono diverse tipologie di disordini, tra cui infarto del miocardio, trombosi venosa e arteriosa, malattie coronariche, malattie cerebrovascolari ed embolia polmonare e rappresentano in generale una delle principali cause di morte nei paesi occidentali. I fattori che predispongono all'insorgenza di malattie tromboemboliche sono diversi e riguardano fattori ambientali e genetici tra loro interconnessi. Tra questi è stata dimostrata chiaramente una correlazione tra polimorfismi in specifici fattori della coagulazione e trombofilia. Recentemente è stato suggerito un possibile ruolo delle piastrine nella predisposizione all'insorgenza delle malattie tromboemboliche, in quanto le piastrine rappresentano i principali elementi figurati coinvolti nei processi di coagulazione ed emostasi. Sono stati identificati polimorfismi a livello dei geni che codificano per i principali recettori piastrinici,  $\alpha_{IIb}\beta_3$ , complesso GP Ib-IX-V,  $\alpha_2\beta_1$ . Studi di correlazione genotipo-malattia sono stati compiuti su un'ampia casistica di pazienti

affetti da malattie tromboemboliche, utilizzando diversi parametri di valutazione (età, sesso, la presenza di episodi ricorrenti). Come riassunto in tabella 2, questi studi clinici hanno dato risultati contrastanti, indicando che una diretta correlazione tra genomica piastrinica e trombofilia non può a tutt'oggi considerarsi dimostrata. Si può peraltro prevedere che il rapido progredire delle conoscenze sui meccanismi della funzionalità piastrinica permetterà in un prossimo futuro di proporre nuove associazioni tra corredo genetico e predisposizione alle patologie trombotiche.

## Bibliografia

1. Gresele P., Page C., Fuster V., Vermylen J. Platelets in thrombotic and non-thrombotic disorders. Ed. Cambridge University Press, 2002.
2. Kekomäki R. Platelet function and immune response. *J Pediatr Hematol Oncol.* 2003; 25: S19-23.
3. Klinger M.H., Jelkmann W. Role of blood platelets in infection and inflammation. *J Interferon Cytokine Res.* 2002; 22: 913-22.
4. Fitzgerald J.R., Foster T.J., Cox D. The interaction of bacterial pathogens with platelets. *Nat Rev Microbiol.* 2006; 4: 445-57.
5. Davi G., Patrono C. Platelet activation and atherothrombosis. *N Engl J Med.* 2007; 357: 2482-94.
6. Rickles F.R. Mechanisms of cancer-induced thrombosis in cancer. *Pathophysiol Haemost Thromb.* 2006; 35: 103-10.
7. Ruggeri Z.M., Mendolicchio G.L. Adhesion mechanisms in platelet function. *Circ Res.* 2007; 100: 1673-85.
8. Shattil S.J., Newman P.J. Integrins: dynamic scaffolds for adhesion and signaling in platelets. *Blood.* 2004; 104: 1606-15.
9. Prévost N., Kato H., Bodin L., Shattil S.J. Platelet integrin adhesive functions and signaling. *Methods Enzymol.* 2007; 426: 103-15.
10. Ma Y.Q., Qin J., Plow E.F. Platelet integrin alpha(IIb)beta(3): activation mechanisms. *J Thromb Haemost.* 2007; 5: 1345-52.
11. Clemetson K.J., Clemetson J.M. Platelet collagen receptors. *Thromb Haemost.* 2001; 86: 189-97.
12. Traynelis S.F., Trejo J. Protease-activated receptor signaling: new roles and regulatory mechanisms. *Curr Opin Hematol.* 2007; 14: 230-5.
13. Murugappan S., Shankar H., Kunapuli S.P. Platelet receptors for adenine nucleotides and thromboxane A2. *Semin Thromb Hemost.* 2004; 30: 411-8.
14. Moroi M., Jung S.M. Platelet glycoprotein VI: its structure and function. *Thromb Res.* 2004; 114: 221-33.
15. Canobbio I., Balduini C., Torti M. Signalling through the platelet glycoprotein Ib-V-IX complex. *Cell Signal.* 2004; 16: 1329-44.
16. Andrews R.K., Berndt M.C. Platelet physiology and thrombosis. *Thromb Res.* 2004; 114: 447-53.
17. Ruggeri Z.M. Von Willebrand factor: looking back and looking forward. *Thromb Haemost.* 2007; 98: 55-62.

18. Jackson S.P. The growing complexity of platelet aggregation. *Blood*. 2007 Jun 15; 109 (12): 5087-95.
19. Shattil S.J. Signaling through platelet integrin alpha IIb beta 3: inside-out, outside-in, and sideways. *Thromb Haemost*. 1999; 82: 318-25.
20. Wencel-Drake J.D., Boudignon-Proudhon C., Dieter M.G., Criss A.B., Parise L.V. Internalization of bound fibrinogen modulates platelet aggregation. *Blood*. 1996; 87: 602-12.
21. Furie B., Furie B.C. In vivo thrombus formation. *J Thromb Haemost*. 2007; 5: 12-7.
22. Lane D.A., Grant P.J. Role of hemostatic gene polymorphisms in venous and arterial thrombotic disease. *Blood*. 2000; 95: 1517-32.
23. Feinbloom D., Bauer K.A. The role of other hemostatic factors in predicting arterial thrombotic events. *Arterioscler Thromb Vasc Biol*. 2006; 26 e 29.
24. Bray P.F. Integrin polymorphisms as risk factors for thrombosis. *Thromb Haemost*. 1999; 82: 337-44.
25. Weiss E.J., Bray P.F., Tayback M., Schulman S.P., Kickler T.S., Becker L.C., Weiss J.L., Gerstenblith G., Goldschmidt-Clermont P.J. A polymorphism of a platelet glycoprotein receptor as an inherited risk factor for coronary thrombosis. *N Engl J Med*. 1996; 334: 1090-4.
26. Carter A.M., Ossei-Gerning N., Grant P.J. Platelet glycoprotein IIIa PIA polymorphism in young men with myocardial infarction. *Lancet*. 1996; 348: 485-6.
27. Ridker P.M., Hennekens C.H., Schmitz C., Stampfer M.J., Lindpaintner K. PIA1/A2 polymorphism of platelet glycoprotein IIIa and risks of myocardial infarction, stroke, and venous thrombosis. *Lancet*. 1997; 349: 385-8.
28. Herrmann S.M., Poirier O., Marques-Vidal P., Evans A., Arveiler D., Luc G., Emmerich J, Cambien F. The Leu33/Pro polymorphism (PIA1/PIA2) of the glycoprotein IIIa (GPIIIa) receptor is not related to myocardial infarction in the ECTIM Study. *Etude Cas-Temoins de l'Infarctus du Myocarde*. *Thromb Haemost*. 1997; 77: 1179-81.
29. López J.A., Andrews R.K., Afshar-Kharghan V., Berndt M.C. Bernard-Soulier syndrome. *Blood*. 1998; 91: 4397-418.
30. Gonzalez-Conejero R., Lozano M.L., Rivera J., Corral J., Iniesta J.A., Moraleda J.M., Vicente V. Polymorphisms of platelet membrane glycoprotein Ib associated with arterial thrombotic disease. *Blood*. 1998; 92: 2771-6.
31. Ardissino D., Mannucci P.M., Merlini P.A., Duca F., Fetiveau R., Tagliabue L., Tubaro M., Galvani M., Ottani F., Ferrario M., Corral J., Margaglione M. Prothrombotic genetic risk factors in young survivors of myocardial infarction. *Blood*. 1999; 94: 46-51.
32. Kaski S., Kekomäki R., Partanen J. Systematic screening for genetic polymorphism in human platelet glycoprotein Ibalpha. *Immunogenetics*. 1996; 44: 170-6.
33. Santoso S., Zimmermann P., Sachs U.J., Gardemann A. The impact of the Kozak sequence polymorphism of the glycoprotein Ib alpha gene on the risk and extent of coronary heart disease. *Thromb Haemost*. 2002; 87: 345-6.

34. Moroi M., Jung S.M. Integrin-mediated platelet adhesion. *Front Biosci.* 1998; 3: d719-28.
35. Kunicki T.J. The role of platelet collagen receptor (glycoprotein Ia/IIa; integrin alpha2 beta1) polymorphisms in thrombotic disease. *Curr Opin Hematol.* 2001; 8: 277-85.
36. Santoso S., Kunicki T.J., Kroll H., Haberbosch W., Gardemann A. Association of the platelet glycoprotein Ia C807T gene polymorphism with nonfatal myocardial infarction in younger patients. *Blood.* 1999; 93: 2449-53.
37. Croft S.A., Hampton K.K., Sorrell J.A., Steeds R.P., Channer K.S., Samani N.J., Daly M.E. The GPIa C807T dimorphism associated with platelet collagen receptor density is not a risk factor for myocardial infarction. *Br J Haematol.* 1999; 106: 771-6.

## **Dalla genetica ai trials clinici in cardiologia**

*Stefano Ghio*

Divisione di Cardiologia, Fondazione IRCCS Policlinico San Matteo, Pavia

Negli ultimi anni lo sviluppo della Medicina ha proceduto con un passo estremamente veloce, utilizzando le nuove scoperte della scienza e della tecnologia per migliorare i processi diagnostici e la terapia. La Cardiologia, disciplina ad elevata componente tecnologica, si può dire che abbia fatto passi da gigante. Le opportunità diagnostiche offerte dalle nuove metodiche di immagine, quali la tomografia computerizzata multidetettore, la risonanza magnetica, le più recenti applicazioni ecocardiografiche, sono sotto gli occhi di tutti. In tema di terapia, non soltanto il cardiologo dispone ora di farmaci in grado di contrastare efficacemente le diverse patologie cardiovascolari, incluse malattie rare che una volta erano considerate incurabili quali l'ipertensione arteriosa polmonare, ma è in grado di intervenire per dilatare stenosi coronariche, ablare aritmie, impiantare dispositivi con funzione antiaritmica e di resincronizzazione meccanica nei pazienti affetti da scompenso cardiaco. Tutto ciò ha reso la professione del cardiologo da un lato sempre più stimolante, dall'altro anche decisamente impegnativa. Il cardiologo di oggi deve aggiornarsi continuamente, rivedendo criticamente ed adeguando rapidamente la propria attività clinica ai suggerimenti della letteratura internazionale prima ancora che le linee guida indichino gli standard di riferimento; il che vuol dire che il cardiologo di oggi deve essere in grado di comprendere e di interpretare i risultati dei trials clinici e degli studi genetici pubblicati in letteratura. Conoscenze genetiche e metodologia di ricerca clinica basata sui trial controllati-randomizzati sono infatti i due settori che negli ultimi 20 anni hanno contribuito in maniera maggiore all'avanzamento delle conoscenze in ambito cardiovascolare.

L'era dei trial nasce, poco più di 20 anni fa, dalla convinzione che il mondo della terapia fosse troppo empirico e vago e che fosse invece necessario e possibile raggiungere un grado maggiore di certezza riguardo all'efficacia e sicurezza dei trattamenti terapeutici (1). I trial sono stati lo strumento applicativo di questo nuovo approccio. Oggi gli enti regolatori e la comunità scientifica ritengono che niente meglio dei trial randomizzati-controllati possa dare solidità di prova ai risultati di uno studio. La randomizzazione infatti minimizza i bias noti ed ignoti di impostazione della ricerca e bilancia le caratteristiche note ed ignote del gruppo testato e di quello di controllo. Ovvio che la randomizzazione da sola non basta a fare un buon trial (2). Gli obiettivi devono essere chiari e basati su ipotesi corrette dal punto di vista fisiopatologico e plausibili dal punto di vista clinico; il disegno



dello studio deve per quanto possibile prevedere analisi per sottogruppi perché si possa poi cercare di comprendere chi sono i pazienti che beneficeranno maggiormente della nuova terapia e chi invece no; la gestione dello studio e l'interpretazione dei risultati devono essere indipendenti da condizionamenti di ogni genere. Né ci si può aspettare che la dimostrazione di efficacia nel trial di un nuovo trattamento possa da solo modificare l'epidemiologia o la storia naturale di una malattia: c'è da verificare "nel mondo reale" la sicurezza ed efficacia del trattamento testato nel trial, creando le premesse perché la terapia venga implementata in modo corretto, evitando da un lato usi inappropriati e dall'altro il sottoutilizzo del farmaco stesso (3).

Il futuro della ricerca clinica però non si basa solo sui trials. Il trial, pur dimostrando l'efficacia della una terapia nella popolazione di pazienti affetti, non ha la possibilità di predire i risultati della terapia nel singolo paziente. L'identificazione a priori dei potenziali responders ad una terapia è un punto cruciale, non solo per quanti ritengono che ci sia bisogno in qualche modo di ri-equilibrare l'entusiasmo per la medicina basata sulle evidenze, dove i risultati dei grandi trials dominano il processo di decisione clinica, ma anche dal punto di vista economico: la moltiplicazione delle terapie, soprattutto di quelle invasive e costose, non è un processo che può andare avanti senza fine. La farmacogenetica, individualizzare cioè la terapia basandosi su una conoscenza del patrimonio genetico del singolo paziente che ci permetta di comprendere la variabilità individuale nella risposta al farmaco, è una speranza per il futuro: i polimorfismi dei target recettoriali e dei componenti delle vie metaboliche attivate dai farmaci possono infatti differenziare, anche in modo rilevante, le risposte biologiche individuali ai farmaci. Ma già oggi studiare le caratteristiche genetiche del singolo paziente vuol dire spostare l'attenzione dalla malattia al paziente come entità unica. Diverse malattie cardiovascolari sono monogeniche, ad esempio le cardiomiopatie, ed i progressi nell'ambito della loro caratterizzazione genetica sono stati notevoli (4). Nuove conoscenze sul genoma umano, nuovi e più potenti approcci di analisi di linkage applicabili su larghe popolazioni potranno poi svelare geni che contribuiscono in maniera sostanziale al rischio anche nell'ambito di malattie poligeniche come l'infarto del miocardio, in cui tra l'altro l'ereditarietà si combina in maniera complessa con diverse componenti ambientali (5).

Le considerazioni svolte sopra possono sembrare molto generiche e forse risulta difficile coglierne il senso pratico. Affrontiamo quindi una singola patologia, lo scompenso cardiaco, per renderci conto nel particolare di:

- 1) come i trials clinici non solo hanno migliorato il modo di trattare la malattia, ma hanno addirittura radicalmente cambiato il nostro modo di concepirla;
- 2) come ai trials clinici siano seguiti studi di outcome per monitorare il processo di implementazione della terapia nel mondo reale, massimizzando i benefici e minimizzando i rischi per i pazienti;
- 3) come venga affrontato il problema dei "responders" alla terapia;
- 4) come infine sia cambiato l'approccio clinico al paziente (ed alla famiglia) con il riconoscimento di difetti genetici responsabili dell'insorgenza di cardiomiopatie.

## Il caso dello scompenso cardiaco

Fino a 20 anni fa, l'insufficienza cardiaca era vista come una condizione acuta, caratterizzata da episodi di edema polmonare, oliguria ed edemi periferici; il modello fisiopatologico interpretativo era quello secondo il quale le alterazioni emodinamiche determinavano in maniera sostanziale gli aspetti clinici e la progressione della malattia (6). Lo scopo della terapia era quindi risolvere rapidamente i sintomi migliorando l'emodinamica: utilizzando cioè farmaci come diuretici, vasodilatatori ed inotropi positivi. L'ipotesi che i meccanismi emodinamici siano i principali responsabili della prognosi infausta dei pazienti con insufficienza cardiaca è stata ampiamente smentita dai trials clinici. Ad eccezione della digitale i cui risultati sulla mortalità sembrano essere neutri, i trial clinici hanno concordemente dimostrato un effetto sfavorevole sulla mortalità di tutti gli inotropi positivi sperimentati nei pazienti con insufficienza cardiaca cronica (prenalterolo, xamoterolo, amrinone, milrinone, enoximone, ibopamina, pimobendan, vesnarnone), nonostante la loro somministrazione abbia generalmente determinato un miglioramento emodinamico significativo e talora anche un miglioramento della capacità funzionale. L'aumentata mortalità sembra essere principalmente dovuta alla maggiore incidenza di morte improvvisa. Anche due potenti vasodilatatori quali il flosequinan e l'epoprostenolo, hanno determinato un peggioramento della sopravvivenza a tal punto da causare, in entrambi i casi, l'interruzione prematura dei trial in cui erano testati. Altri vasodilatatori come l'amlodipina e la felodipina hanno dato risultati neutri sulla mortalità ed a tutt'oggi solo i nitrati possono essere utilizzati con sicurezza nello scompenso cronico.

Alla fine degli anni '80 ci si è resi conto che l'attivazione del sistema renina-angiotensina-aldosterone e quella del sistema nervoso simpatico, entrambe causate dalla disfunzione cardiaca ed inizialmente considerate un meccanismo di compenso alla disfunzione meccanica del cuore, potevano avere un ruolo determinante nel peggioramento della malattia stessa; lo scompenso cardiaco è stato quindi considerato una sindrome "neuro-ormonale" ed il vero bersaglio della terapia farmacologica è diventato antagonizzare questa attivazione. I trial con gli ACE-inibitori ed i betabloccanti hanno non solo permesso di confermare queste ipotesi ma hanno anche direttamente contribuito a generare la nuova ipotesi interpretativa dello scompenso. Il primo trial è stato il CONSENSUS, in cui pazienti con insufficienza cardiaca in IV classe funzionale NYHA nonostante una terapia con dosi ottimali di diuretici, digitale ed altri vasodilatatori sono stati randomizzati a enalapril o placebo. Come è noto, lo studio è stato interrotto precocemente dopo l'inclusione di soli 253 pazienti per una riduzione altamente significativa della mortalità con enalapril: (50% a 6 mesi e 46% a 12 mesi). A seguire innumerevoli studi con diversi ACE-inibitori, in pazienti con scompenso cardiaco cronico o post-infartuati, con sintomatologia severa o con disfunzione ventricolare sinistra asintomatica, hanno tutti coerentemente dimostrato l'efficacia di questa categoria di farmaci. I meccanismi della riduzione di mortalità con ACE inibitori sono verosimilmente multipli. Tra questi, è certamente fondamentale il rallentamento del rimodellamento ventricolare dovuto, a sua volta, alla riduzione dello

stress miocardico e del consumo miocardico di ossigeno, all'inibizione dell'ipertrofia, della fibrosi miocardica e dell'apoptosi, nonché, indirettamente, alla minore stimolazione di altri meccanismi neuromorali, quali il sistema simpatico e l'endotelina. In accordo con questa ipotesi, è l'osservazione che gli ACE inibitori hanno ridotto soprattutto l'incidenza delle morti per insufficienza cardiaca ingravescente. Oggi l'antagonismo del sistema renina-angiotensina-aldosterone viene perseguito in maniera più consistente associando altri farmaci che interferiscono a livelli diversi su questo sistema, come i sartanici e gli antialdosteronici, con risultati clinicamente più che favorevoli.

Il passo successivo sono stati i betabloccanti. Negli anni '70 ed '80 questi farmaci erano formalmente contro-indicati per la loro azione di depressione della contrattilità miocardica. I trials multicentrici hanno invece dimostrato una significativa riduzione dell'incidenza di ospedalizzazioni, eventi cardiovascolari e mortalità globale nei pazienti trattati. Il rationale al loro impiego è che il sistema simpatico, soprattutto a livello cardiaco, è attivato molto precocemente, già nella fase in cui la disfunzione ventricolare è asintomatica e soprattutto non è solo la conseguenza della disfunzione cardiaca ma può essere causa del suo progressivo peggioramento. Una prolungata esposizione a livelli elevati di catecolamine plasmatiche interferisce con il corretto funzionamento dei recettori adrenergici e può addirittura avere effetti tossici determinando danno miocardico ed induzione di apoptosi. La somministrazione di betabloccanti provoca una upregulation dei recettori, l'espressione di isoforme diverse di proteine contrattili e, paradossalmente, i betabloccanti possono oggi essere considerati i soli farmaci ad azione inotropica positiva nello scompenso. Purtroppo i betabloccanti sono farmaci difficili da utilizzare: prima che si manifestino i loro effetti positivi prevalgono gli effetti emodinamici di depressione della contrattilità ed il rischio che il paziente possa inizialmente peggiorare è reale; è necessario iniziare la terapia con dosi molto basse e titolare il dosaggio progressivamente nel corso di qualche mese controllando strettamente il paziente. L'implementazione nella pratica clinica di questi farmaci è quindi stata piuttosto lenta e per favorirla è stato necessario perseguire a livello nazionale una strategia di "partecipazione ed istruzione", fatta di incontri con i medici, distribuzione di documenti e studi multicentrici osservazionali. In quest'ottica si è svolto lo studio BRING-UP I, che ha arruolato quasi 3.000 pazienti: all'inizio meno del 25% dei pazienti assumeva betabloccanti ma al termine la percentuale era salita a più del 50%.

Negli anni l'interpretazione dello scompenso come sindrome neuroormonale ha subito una sua evoluzione: innanzitutto si è ipotizzato che nuovi peptidi (ad es. le citochine) che non possono essere considerati ormoni in senso stretto abbiano un ruolo nel determinare l'evoluzione della malattia (da sindrome "neuro-ormonale" a neuro-umorale). L'attenzione verso aspetti clinici come l'anemia e l'insufficienza renale ha poi messo in evidenza la loro rilevanza prognostica ed ha indotto una maggiore attenzione nei confronti delle alterazioni che lo scompenso cardiaco induce in organi e tessuti diversi dal cuore. Infine il cuore è ritornato prepotentemente al centro dell'attenzione dei clinici, con le nuove possibilità di migliorare la meccanica di contrazione nei pazienti affetti da scompenso e distu-

ri della conduzione intraventricolare. La terapia di resincronizzazione mediante stimolazione biventricolare migliora i sintomi, la capacità funzionale, riduce le ospedalizzazioni per scompenso e riduce la mortalità nei pazienti con scompenso cardiaco e blocco di branca sinistra, anche se sono già in terapia farmacologica massimale. Il miglioramento clinico si associa ad un rimodellamento inverso del ventricolo sinistro (riduzione dei volumi ed aumento della frazione di eiezione) talora così marcato da far parlare, per la prima volta in questa patologia, di vera remissione della malattia. Tuttavia una percentuale stimata intorno al 30% dei pazienti o non migliora o peggiora nonostante la stimolazione biventricolare. Ovvio che nessuna terapia in medicina funziona sempre nel 100% dei pazienti, ma in considerazione del fatto che la resincronizzazione è una terapia definitiva (i farmaci possono essere interrotti in qualsiasi momento mentre un pacemaker può essere spento e se lo si vuole espantare è necessario un altro intervento chirurgico), è una terapia ad alto costo, è una terapia che comporta comunque dei rischi per il paziente soprattutto se le condizioni cliniche sono avanzate, l'attenzione dei ricercatori in questi ultimi anni si è focalizzata sulla possibilità di identificare *a priori* i potenziali responders. Molti parametri di dissincronia meccanica, generalmente ecocardiografici, sono stati proposti in letteratura; un ampio dibattito si è aperto sulla necessità o meno di indicare nelle linee guida la necessità di selezionare con l'ecocardiografia i pazienti prima dell'impianto. Ad oggi, non siamo ancora in grado di proporre un parametro di facile utilizzo in tutti i laboratori di ecocardiografia.

Il riconoscimento che molte cardiomiopatie sono potenzialmente ereditarie e che spesso l'ereditarietà è riconducibile ad una singola mutazione genetica causale (malattie monogeniche) significa per il cardiologo clinico innanzitutto avere una particolare attenzione nello svolgere l'anamnesi del paziente. Non diagnosticare una cardiomiopatia familiare comporta conseguenze non tanto sulla strategia e sulle opzioni terapeutiche del paziente in studio, quanto sulla possibilità di informare il resto della famiglia, in cui potrebbero esserci degli individui già affetti anche se asintomatici che potrebbero beneficiare di una diagnosi e terapia precoce e sulla possibilità di spiegare al malato i rischi per la sua eventuale prole. Il lavoro di counseling genetico è estremamente delicato e complesso e solleva molti problemi etici e medico-legali; esso è strettamente legato a programmi di screening e di re-screening dei familiari dei pazienti che giungono all'attenzione del cardiologo per una cardiomiopatia potenzialmente ereditaria. I costi, anche in termini di risorse umane, necessari per portare avanti su larga scala questi programmi sono però così alti che la caratterizzazione genetica dei pazienti e la diagnosi pre-clinica sono tuttora appannaggio di pochi centri dedicati allo studio di queste malattie, in genere centri in cui sono già attivi più ampi programmi di trapianto di cuore e di polmone (7).

## Conclusioni

Nell'era della medicina basata sull'evidenza, dell'ottimizzazione del rapporto costo beneficio delle prestazioni sanitarie, dell'appropriatezza delle metodiche

invasive e non invasive, al cardiologo si richiedono non solo competenze cliniche ma anche la capacità:

- 1) di implementare nel mondo reale i risultati dei trials clinici controllati;
- 2) di tradurre in pratica gli insegnamenti derivanti dalla caratterizzazione genetica delle malattie ereditarie.

Ovviamente non è possibile essere contemporaneamente un clinico, un trialista ed un genetista, ma è possibile e necessario che il paziente riceva tutti i benefici derivanti da un approccio multidisciplinare alla malattia (8).

### **Bibliografia**

1. Tavazzi L., Maggioni AP., Tognoni G. Participation versus education: the GISSI story and beyond. *Am Heart J* 2004; 148: 222-9.
2. Tavazzi L. evidence-based therapeutic strategies. There is the need to bridge the gap between simplified megatrials and individual prescriptions. *Ital Heart J* 2001; 2 (2): 73-77.
3. Tavazzi L., Garattini S., Mannuccio PM., Remuzzi G. Note sulla ricerca clinica sul farmaco in Italia. *GIC* 2006; 7 (1): 62-67.
4. Weatherall DJ. Genetics and molecular biology: from a monastery garden to rebuilding the heart. *Dialogues in Cardiovascular Medicine* 2006; 11: 148-155.
5. Damani SB, Topol EJ. Future use of genomics in coronary artery disease. *J Am Coll Cardiol* 2007; 50: 1933-40.
6. Tavazzi L. Terapia dell'insufficienza cardiaca dopo vent'anni di trial. *Cardiologia* 2004; 342-355.
7. Pasotti M, Agozzino M, Repetto A, Arbustini E. un approccio operativo allo screening familiare nelle cardiomiopatie (dall'anamnesi all'analisi genetica): quando, come, dove, con quali costi. *Cardiologia* 2004; 314-328.
8. Kats AM. Blossoms on the tree of cardiology: some predictions for the coming decade. *Dialogues in Cardiovascular Medicine* 2006; 11: 157-167.

**GENOMICA E PROTEOMICA  
NELLE MALATTIE  
NEURODEGENERATIVE**



## **Indagini di proteomica clinica per l'identificazione di marcatori molecolari delle malattie neurodegenerative**

*Andrea Urbani*

U.O. Biochimica Analitica, Ce.S.I., Fondazione Università "Gabriele D'Annunzio", Chieti

The complete human genome sequence opened the route to a new holistic scientific movements, characterised by the investigation of the functional elements of our genetic heredity mainly in respect to multifactorial diseases such as cancer, diabetes, dementias, etc. The Clinical Proteomics tries to highlight the relations between the main molecular physiological effectors, proteins, with defined physio-pathological states. The development of a new generations of Mass Spetrometers is deeply changing the technologies employed in these investigations, enabling direct quantitative analysis of clinical samples such as, blood, urine, saliva, biopsy samples. Our Proteomics Group is focused in the investigation of differential protein expression profiles and metabolites pattern in neurodegenerative diseases. We have actively established a set of novel mass spectrometry investigations in order develop potential assays for molecular biomarkers identification and characterisation. The achieved results will be discussed during the lecture.

### **Introduction**

Following the great promise of the human genome project, a plethora of novel studies has been being carried out to investigate the molecular mechanisms of the encoded genes. The main interest of clinical research has consequently shifted toward an improved understanding of the inter-relations of functional elements and transcribed genome regions with specific physiopathologic conditions. This enthusiasm brought novel technological breakthroughs as much as intellectual developments in the experimental design. These were particularly driven by an holistic view of biological phenomena, which are considered the consequence of organic wholes that are greater than the simple sum of their components. In fact, the human genome project provided us with the first holistic view of our genetic repository, shading new light in the inductive investigation of other molecular and functional wholes such as transcripts, proteins, metabolites, molecular interactions, etc.



Proteome investigations, called proteomics, are the natural development of this approach, now encompassing not only all the proteins encoded in any given cell, but also the set of all proteins isoforms and modifications, the interactions between them, the structural description of proteins and their higher-order complex. As proteins are the main effectors of physiological functions proteomics spans over a wide area of the present 'post genomic' investigations. This high-throughput biochemistry will be crucial for the emerging field of system biology, contributing at a direct level to a full description of cellular function and biological phenomena (1, 2). However, the sequencing of the human genome and that of numerous pathogens has opened the door for proteomics by providing a sequence-based framework for mining proteomes, most difficult problems ahead is to find out how genes contribute to diseases that have complex pattern of inheritance, such as in the cases of diabetes, asthma, cancer and mental illness. In all these cases, no single gene has implicated in the disease process obtaining a dramatic complex situation.

There is high interest in applying proteomics to foster a better understanding of disease dynamics. In particular clinical biochemistry has found a new era taking advantage from this holistic point of view. The interest in clinical proteomics is due in part to the prospects that a proteomics approach to disease investigations will overcome some of the limitations of other approaches. Because most drug targets are proteins, it is inescapable that proteomics will promote drugs discovery by the identification of new targets for therapeutics with more effective strategies for clinical practice (3) (Fig. 1).

We have actively established a set of novel mass spectrometry investigations in order develop potential assays for molecular biomarkers identification and characterisation. 2D Electrophoresis approach has been complemented by MALDI-TOF mass spectrometry protein profiles either in the screening of complete protein extracts and in direct analysis of tissue samples (4). Novel bioinformatics tools have been developed to specifically analysed MALDI-TOF-MS molecular profiles and possibly outline spectra differences (5, 6).

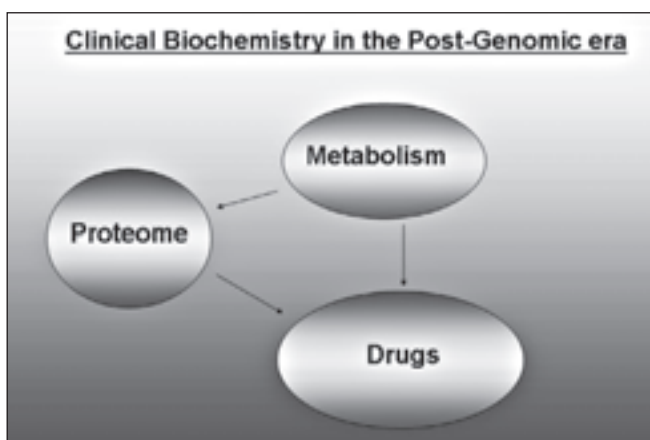


Fig. 1

We have investigated the cerebral spinal fluid (CSF) of Alzheimer Disease affected patients versus subjects without cognitive impairment (7). Our results indicated a differential distribution of TTR-cysteine and TTR-cysteinglycine adducts. In fact, both oxidised forms of TTR are significantly less abundant in the Alzheimer Disease group ( $P=0.0001$ ). The investigated population has been diagnosed using the relative ratio of conjugated forms over the free TTR. A sensitivity  $>90\%$  and a specificity  $>70\%$  was outlined in ROC curve analysis when the overall cohort is discriminated by the TTR-Cysteine signals. Transthyretin is the only known protein in the CSF synthesised uniquely by the choroid plexus. It posses binding capacity of the beta-amyloid peptide and a number of amyloidosis diseases have been associated with its allele variants. Transthyretin has a free reactive sulphidryl moiety located on the Cys10 residue which might undergo to different oxidation reactions.

We have pursued an independent clinical proteomics investigation on sera of Multiple Sclerosis affected patients employing linear MALDI-TOF mass spectrometry (8). Our investigation have been mainly focused on the disease profile course characterised by the relapsing-remitting pattern (RR). Spectra have been acquired in Linear positive ion mode on a Reflex IV and Ultraflex III (Bruker-Daltonics) MALDI-TOF-MS after solid phase extraction on a reverse-phase material. Our results highlight a differential distribution of the proteins in the range of 6,000-10,000 AMU in the sera of patients affected by MS in respect to both healthy subjects and patients affected by other neurological diseases. These data allow to build a mathematical model which enables multiple sclerosis diagnosis from 10 ul of sera samples with 93% specificity and 91% sensitivity. Moreover, recent studies have suggested the possible involvement of additional non proteic mediators of autoimmune brain inflammation such as lipids. In order to investigate characteristic serum metabolite profiling as the result of the onset of disease we have developed a liquid chromatography/mass spectrometry method followed by statistical multivariate analysis investigation. Once improved the method was applied in a preliminary clinical cross-sectional metabonomics investigation of multiple sclerosis affected patients vs healthy subjects. Our results outline the possibility to discriminate on principal component analysis (PCA), based on 2,251 different metabolites, the disease affected subjects from healthy controls.

## References

1. Tyres M., Mann M. *Nature* 422, 13 March 2003; 193-197.
2. Csete ME., Doyle JC. *Science* 295, 2002; 1664-1669.
3. Hanash S. *Nature* 422, 13 March 2003; 226-232.
4. D'Aguanno S., Del Boccio P., Bernardini S., Ballone E., Di Ilio C., Federici G., Urbani A. *Clin Chem Lab Med.* 2007; 45 (4): 437-49. Review.
5. Mantini D., Petrucci F., Del Boccio P., Pieragostino D., Di Nicola M., Lugaresi A., Federici G., Sacchetta P., Di Ilio C., Urbani A. *Bioinformatics.* 2008 Jan 1; 24 (1): 63-70. Epub 2007 Nov 14.

6. Mantini D., Petrucci F., Pieragostino D., Del Boccio P., Di Nicola M., Di Ilio C., Federici G., Sacchetta P., Comani S., Urbani A. *BMC Bioinformatics*. 2007 Mar 26; 8: 101.
7. Biroccio A., Del Boccio P., Panella M., Bernardini S., Di Ilio C., Gambi D., Stanzione P., Sacchetta P, Bernardi G, Martorana A, Federici G, Stefani A, Urbani A. *Proteomics*. 2006 Apr; 6 (7): 2305-13.
8. Del Boccio P., Pieragostino D., Lugaesi A., Di Ioia M., Pavone B., Travaglini D., D'Aguanno S., Bernardini S., Sacchetta P., Federici G., Di Ilio C., Gambi D., Urbani A. *Ann Neurol*. 2007 Aug; 62 (2): 201-4; discussion 205.

## **Proteomica della malattia di Alzheimer: comprensione dei meccanismi di malattia e identificazione di biomarcatori**

*Stefano Govoni*

Dipartimento di Farmacologia Sperimentale ed Applicata, Università di Pavia

Le ricerche sulle basi neurobiologiche della malattia hanno guidato nel passato il disegno dei farmaci attualmente utilizzati nel trattamento della malattia di Alzheimer e continuano a fornire le basi teoriche per lo sviluppo di quelli futuri secondo una scansione temporale che è quella propria dello sviluppo delle nuove terapie e che, purtroppo, è piuttosto lenta. Circa trent'anni fa gli studi di neurochimica dimostranti un deficit della trasmissione colinergica nella malattia di Alzheimer fornirono il razionale per lo sviluppo dei farmaci entrati in terapia alla fine degli anni '80, gli inibitori dell'acetilcolinesterasi. Prima dello sviluppo degli inibitori d'ultima generazione, l'interesse sui farmaci colinergici era stato limitato dai numerosi effetti collaterali e dalla scarsa efficacia clinica dei composti sino ad allora sviluppati. Nei primi anni '90 l'approvazione da parte di molti Enti Regolatori, di vari inibitori dell'acetilcolinesterasi con l'indicazione specifica per la malattia di Alzheimer ha rilanciato sia il campo di indagine sul sistema colinergico sia, in generale, il settore delle ricerche sulla malattia di Alzheimer.

Nel complesso da queste ricerche sono emerse nuove informazioni sul processo patogenetico della malattia. In particolare, attorno alla metà degli anni '80, con la scoperta della natura proteica dei costituenti delle placche senili e dei gomitoli neurofibrillari, si è assistito ad una rivoluzione epocale nel modo di intendere le basi biologiche della malattia. In breve tempo si è transitati da una concezione classica di tipo neurotrasmettitoriale, basata sui deficit colinergici osservati nell'AD, ad una concezione incentrata su disfunzioni di alcune proteine cerebrali: la demenza di Alzheimer come proteinopatia. In questo ambito si collocano le informazioni sugli effetti neurotossici di beta amiloide e delle sue forme aggregate e sulla possibilità che questi composti oltre ad esercitare effetti diretti rendano il tessuto nervoso più suscettibile ad una serie di danni che comprendono il mancato afflusso di sangue o di nutrienti, l'esposizione a radicali liberi e a neurotossine. L'attività neurotossica di beta amiloide è a tutt'oggi esplorata sotto vari aspetti ed ha dato avvio allo sviluppo di farmaci anti-amiloide compresa la così detta "vaccinazione". Le esperienze con il vaccino hanno fornito risultati ambi-

gui che non permettono di trarre conclusioni definitive sull'efficacia del trattamento, senza considerare i seri effetti collaterali, tra cui l'encefalite, evidenziati nella principale sperimentazione clinica effettuata che ne hanno determinato la precoce interruzione.

### **Dalla neurobiologia allo sviluppo razionale dei farmaci per la malattia di Alzheimer**

Il peptide beta amiloide o A $\beta$  è generato dal metabolismo proteolitico di un precursore denominato APP (amyloid precursor protein) che per semplicità è distinto in amiloidogenico o non amiloidogenico, rispettivamente risultante nella genesi di peptide amiloide o nella genesi di prodotti metabolici non patogenetici. Nel corso dell'ultimo decennio sono state identificate e caratterizzate le attività proteolitiche responsabili di questo complesso metabolismo. La produzione di beta amiloide è opera dell'azione sequenziale di due proteasi note come beta e gamma secretasi la cui inibizione farmacologica potrebbe in linea di principio ridurre il carico di amiloide nel cervello dei pazienti AD. Parallelamente è anche noto che il metabolismo non amiloidogenico è condotto grazie ad una proteasi denominata alfa secretasi, la cui attivazione è regolata da numerosi meccanismi intracellulari. Tra questi i sistemi di trasduzione del segnale a valle dei recettori colinergici che dimostrano come gli inibitori della colinesterasi possono agire anche sui meccanismi di produzione di amiloide. Nel complesso esistono diverse classi di farmaci che agiscono secondo il principio dell'interruzione dell'evento patologico (per ora limitato all'ipotesi amiloide) o eventi a valle di esso, ed alcuni farmaci hanno già raggiunto le fasi II o III della sperimentazione clinica.

È possibile che ci si trovi di fronte ad una nuova era di trattamenti per la malattia di Alzheimer. Le nuove classi di farmaci in fase di sperimentazione includono l'immunoterapia, con una evoluzione rispetto alla sperimentazione originale del "vaccino", insieme a molecole in grado di inibire la produzione o l'aggregazione di beta amiloide. Questo ultimo fenomeno, l'aggregazione, sembra avere particolare rilievo e numerose osservazioni concordano sul fatto che solo particolari forme di aggregazione (oligomeri a peso molecolare elevato, ma non in forma fibrillare) siano neurotossiche, ponendo anche il problema della attività biologica (nulla? diversa?) delle forme monomeriche solubili e degli intermedi. A livello di strategie anti-amiloide i problemi di sicurezza e tollerabilità che hanno accompagnato la sperimentazione del vaccino e degli inibitori di gamma secretasi sono stati affrontati e le sperimentazioni cliniche continuano. Per esempio nel maggio 2007 la Elan/Wyeth annuncia l'inizio della sperimentazione in fase III di un anticorpo monoclonale umanizzato contro beta amiloide (AAB-001 - bapineuzumab) per il trattamento delle forme lievi-moderate di AD. Questa strategia è definita come immunoterapia passiva e comporta la somministrazione di anticorpo al paziente in dosi ripetute. Tra le nuove molecole che hanno recentemente dimostrato una promettente attività citiamo il Tarenflurbil, una molecola derivata dal flurbiprofene con attività contro la produzione di beta amiloide attraverso l'inibizione di gamma secretasi.

Di rilievo, infine, dal punto di vista concettuale una recente sperimentazione su un numero limitatissimo di pazienti nei quali sono state chirurgicamente impiantate nel cervello cellule (fibroblasti) geneticamente modificate in laboratorio a produrre NGF (il fattore di crescita neuronale scoperto da Rita Levi Montalcini). I risultati clinici ottenuti sono stati modesti, per lo meno a fronte della tecnologia avanzata richiesta per realizzare l'intervento e dell'invasività dell'intervento, e quindi non è possibile prevedere in questo momento se si tratta di una tecnica che avrà sviluppi clinici significativi e diffusi, ma certamente costituisce la prima prova diretta nell'uomo che un fattore di crescita nervoso, somministrato in modo efficace e in sede adeguata, può avere un effetto neurotrofico e rallentare il processo degenerativo dei neuroni colinergici.

### **Le difficoltà del tradurre le nuove conoscenze sulla biologia della malattia in strategie dirette a correggere un quadro proteico alterato e del condurre una corretta sperimentazione clinica**

Il quadro generale lascia prevedere la possibilità di individuare nuovi diversi bersagli biologici e quindi nuovi farmaci per la malattia di Alzheimer, potenzialmente diretti alle cause o ai meccanismi molecolari da essa avviati a livello di proteine anormalmente prodotte, configurate o fosforilate. Tutte le classi di farmaci che si stanno studiando in questo momento, basano la loro ipotesi di efficacia sulla modificazione del corso della malattia, ovvero si configurano come "*disease modifying drugs*". Il concetto della modificazione del decorso della malattia è controverso ed è stato oggetto di un recente sforzo comune tra i ricercatori clinici per definire alcuni importanti principi da seguire. Il disegno delle sperimentazioni cliniche con l'obiettivo di dimostrare proprietà *disease modifying* per un farmaco pone in primo piano alcune domande importanti tra cui: quale popolazione studiare? Per quanto tempo? Con quali endpoints primari e secondari? Esistono marcatori surrogati di efficacia? Queste domande, già sottese, almeno in parte, agli studi sui farmaci sintomatici agenti in modo classico, diventano particolarmente pressanti per i nuovi farmaci che agiscono sull'assetto di alcune proteine cerebrali o su meccanismi molto più incisivi che nel passato sugli assetti cellulari.

Secondo la task force europea che ha elaborato il documento il primo problema controverso è la definizione dell'intervento "modificatore". Infatti intuitivamente per una patologia neurodegenerativa l'intervento dovrebbe risultare in una riduzione dell'evoluzione della patologia. Tuttavia dal punto di vista del paziente questo dovrebbe comunque tradursi in una riduzione o miglioramento stabile delle disabilità conseguenti alla patologia.

Queste osservazioni in qualche modo contrastano nella visione delle agenzie regolatorie mondiali poiché per esempio per l'FDA l'effetto deve essere basato sulla modifica del processo fisiopatologico (per esempio i livelli di beta amiloide) mentre per l'agenzia europea EMEA e per quella canadese l'approvazione di un farmaco con proprietà "*disease modifying*" richiede la dimostrazione di outcomes clinici sostanziali.

## **Biomarcatori e malattia di Alzheimer, un problema non ancora risolto**

La ricerca sulla malattia è resa complessa dalle incertezze diagnostiche nei pazienti in vita e dalla mancanza di biomarcatori di tratto e di stato di facile determinazione e, soprattutto, di biomarcatori legati alla patogenesi della malattia universalmente validati.

In questo ultimo ambito vi sono dati importanti su beta amiloide e proteina tau nel liquor, mentre sono ancora pochi gli studi sui marcatori plasmatici che avrebbero il vantaggio di una più facile applicabilità a grandi numeri di pazienti. In questo ultimo settore si collocano studi recenti su nuovi marcatori legati a proteine, come p53, importanti nel regolare il ciclo cellulare, e che possono essere legate in certe circostanze ai processi di invecchiamento. Studi recenti indicano come nella malattia di Alzheimer si osservi un accumulo di una forma conformazionalmente mutata di p53 in diverse cellule periferiche dei pazienti tra cui fibroblasti, ma anche leucociti circolanti, un'informazione che offre la sponda al possibile uso di questa misura come marcatore plasmatico e allo stesso tempo permette di esplorare ulteriori e nuove ipotesi patogenetiche sulle alterazioni di proteine che controllano il ciclo cellulare (e potenzialmente la plasticità neuronale) nelle demenze.

## **Bibliografia essenziale**

1. Christensen DD. (2007) Alzheimer's disease: progress in the development of anti-amyloid disease-modifying therapies. *CNS Spectr.* 12 (2): 113-116.
2. Govoni S. Alzheimer's disease: facing the biological complexity to treat the disease. *Pharmacol Res.* 2004 Oct; 50 (4): 381-3 (Editorial to a monographic issue of the journal devoted to AD and organized by S. Govoni).
3. Govoni S. (2004) Eziologia e patofisiologia della Malattia di Alzheimer. in: "Invecchiamento Cerebrale e Demenze", Trattato Italiano di Psichiatria, 3<sup>a</sup> edizione MASSON, Milano, 102- 124.
4. Kennedy GJ., Golde TE., Tariot PN., Cummings JL. (2007) Amyloid-Based interventions in Alzheimer's disease. *CNS Spectr.* 12 (12 Suppl 1): 1-14.
5. Lanni C., Racchi M., Mazzini G., Ranzenigo A., Polotti R., Sinfioriani E., Olivari L., Barcikowska M., Styczynska M., Kuznicki J., Szybinska A., Govoni S., Memo M., Uberti D. Conformationally altered p53: a novel Alzheimer's disease marker? *Mol Psychiatry.* 2007 Aug 7; [Epub ahead of print].
6. Racchi M, Govoni S. The pharmacology of amyloid precursor protein processing. *Exp Gerontol.* 2003 Jan-Feb; 38 (1-2): 145-57.
7. Racchi M., Govoni S. Rationalizing a pharmacological intervention on the amyloid precursor protein metabolism. *Trends Pharmacol Sci.* 1999 Oct; 20 (10): 418-23.
8. Racchi M., Sironi M., Caprera A., König G., Govoni S. Short- and long-term effect of acetylcholinesterase inhibition on the expression and metabolism of the amyloid precursor protein. *Mol Psychiatry.* 2001 Sep; 6 (5): 520-8.

9. Sabella S., Quaglia M., Lanni C., Racchi M., Govoni S., Caccialanza G., Calligaro A., Bellotti V., De Lorenzi E. Capillary electrophoresis studies on the aggregation process of beta-amyloid 1-42 and 1-40 peptides. *Electrophoresis*. 2004 Oct; 25 (18-19): 3186-94.
10. Vellas B., Andrieu S., Sampaio C., Wilcock G. (2007) Disease-modifying trials in Alzheimer's disease: a European task force consensus. *Lancet Neurol*; 6: 56-62.



## **Definizione del profilo genomico della sclerosi multipla**

*Roberto Furlan*

Unità di Neuroimmunologia, Dipartimento di Neuroscienze, Università Vita-Salute San Raffaele, DIBIT, Milano

Multiple sclerosis (MS) is a demyelinating disease of the central nervous system (CNS) of unknown etiology (Martino 1999). Its pathological hallmark is the presence within the CNS of inflammatory infiltrates containing few autoreactive T cells and a multitude of pathogenic nonspecific lymphocytes (Steinman 1996). Among T cells present in the CNS of MS patients, CD4<sup>+</sup> predominate although CD8<sup>+</sup> and B cells are also present (Lassmann 1997). The pathological process underlying MS determines the typical patchy CNS demyelination, ranging from demyelination with preservation of oligodendrocytes to complete oligodendrocyte loss and severe glial scarring (Lucchinetti 1996). MS is therefore considered an autoimmune organ-specific inflammatory demyelinating disease of the CNS triggered by a still unknown stimulus (virus, immune-mediated event?). T cells committed against CNS components (e.g. myelin proteins) provide the organ specificity of the pathogenic process and regulate the re-circulation within the CNS of non antigen-specific lymphomononuclear cells acting as effector cells by directly destroying oligodendrocytes and/or by releasing myelinotoxic substances (Steinman 1996).

After an immune-reaction occurring in the CNS (Martino 1999), whose primary trigger as well as the primary target (i.e. myelin components) are still partially unknown, primary inflammatory cytokines (TNF $\alpha$ , IL-1 $\beta$ , and, to some extent, IL-6) are locally released thus triggering, as a cascade, the production of secondary inflammatory mediators such as chemokines (Esiri 1997; Karpus 1998), colony stimulating factors (Battistini 1997) and lipids (Newcombe 1994; Huterer 1995; Narayana 1998). Due to the combined action of cytokines and chemokines, adhesion molecules are then up-regulated on blood-brain barrier (BBB) endothelial cells thus inducing recruitment of leukocytes from the periphery to the CNS. The recruitment of leukocytes within the CNS is paralleled by the activation of local “immune” cells, such as microglia, which then actively participate in the ongoing inflammatory process (Cuzner 1996; Li 1996; Selmaj 1991). The interplay between local immune cells, acting as antigen presenting cells, and myelin-specific blood-borne T cells, belonging to the normal autoimmune repertoire and trafficking to the CNS as a part of their physiological immune-surveillance activ-

ity of the brain, is the crucial event leading to the perpetuation of the inflammatory reaction. Finally, primary inflammatory cytokines *in situ* produced by glial cells (in response to the local inflammatory process) and by myelin-reactive T cells, induce per se a myelin breakdown (Esiri 1997) thus further fostering the whole immune process.

### **Biomarkers in multiple sclerosis**

An ideal biomarker reflects the neuropathology of MS lesions and can be validated in patients with neuropathologically confirmed MS. In addition, such a marker has high sensitivity and specificity for prediction of neurological disability. Furthermore, a marker should be reliable, reproducible, non-invasive, simple to use, and inexpensive.

The involvement of inflammatory molecule in the immunological processes occurring both in the periphery or in the CNS during MS pathogenesis makes inflammatory molecules suitable for being used as peripheral markers of disease (disease-trait) and/or disease activity (steady-trait). Thus, many papers have been published on markers of inflammatory and immunological processes in MS, but few on markers of axonal damage. The few published results are non conclusive if not contradictory (Teunissen 2005). Unsatisfactory results obtained so far are due to a very high background noise for all of tested markers, resulting in an insufficient signal to noise ratio. The reasons for this can be resumed in two main issues:

- 1) the intrinsic variability of the candidate markers in healthy and diseased individuals;
- 2) pre-analytical errors.

The first point can be addressed by the determination of composite markers, made of patterns of several different targets, increasing their prediction power by providing complementary information, and decreasing detection of non-specific modulation by increasing the number of markers to be concomitantly modulated to be significant.

Particular attention has been focused also on the second issue, on pre-analytical aspects, since pre-analytical errors account for 60% of total laboratory errors (Plebani 2006), and pre-analytical errors can occur in both clinical and research settings. The possibility to collect blood directly in RNA stabilizing agents containing tubes, and the reliability of recent techniques of RNA quantifications, indicate gene expression studies, thus mRNA quantification, as a technique with increased chances to yield informative results.

### **Genomics in multiple sclerosis**

Investigators have used microarray approaches to analyze gene expression profile differences in affected tissues of individuals with autoimmune diseases, such as the white matter in the brain in MS or synovium in rheumatoid arthritis (RA). Specific gene expression profiles have also been found in peripheral blood mononuclear cells of individuals with different autoimmune diseases (Maas

2002; Bennett 2003; Bomprezzi 2003; Han 2003; Jarvis 2004; Crow 2003; Lock 2002; Baechler 2003; Crow 2003; Batliwalla 2005; Olsen 2004). For example, an interferon signature has been found in systemic lupus erythematosus that is a function of disease severity (Crow 2003). An early disease signature has been found in RA (Olsen 2004). These are unique to an individual autoimmune disease. In addition, we have described a gene expression signature that is shared among several autoimmune diseases, including RA, systemic lupus erythematosus, type 1 diabetes (insulin-dependent diabetes mellitus), and MS (Maas 2002). We applied a multivariate statistical analysis on results obtained by measuring - by real-time RT-PCR - mRNA levels of 25 immunological relevant molecules in PBMCs from 198 MS patients. The combined measurement of mRNA levels of IL-1 $\beta$ , TNF $\alpha$ , TGF $\beta$ , CCL20 and CCR3 was able to distinguish MS patients from healthy individuals. CXCR5, CCL5, and CCR3 combined mRNA levels identify primary progressive MS patients while TNF $\alpha$ , IL-10, CXCL10 and CCR3 differentiate relapsing MS patients. Our results indicate that multi-parametric analysis of mRNA levels of immunological relevant molecules in PBMCs may represent a successful strategy for the identification of putative peripheral markers of disease state and disease activity in MS patients (Furlan 2005).

Currently we are applying this kind of approach to unbiased technical approaches like microarrays and deep sequencing, using, in parallel, also the animal model of MS, autoimmune experimental encephalomyelitis (EAE), to further dissect the association of peripheral markers with MS pathogenesis. Preliminary, unpublished results will be discussed during the lecture.

## References

1. Baechler EC., Batliwalla FM., Karypis G., Gaffney PM., Ortmann WA., Espe KJ., Shark KB., Grande WJ., Hughes KM., Kapur V., Gregerson PK., Behrens TW. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci USA* 2003; 100: 2610-2615.
2. Batliwalla FM., Baechler EC., Xiao X., Li W., Balasubramanian S., Khalili H., Damle A., Ortmann WA., Perrone A., Kantor AB., Gulko PS., Kern M., Furie R., Behrens TW., Gregersen PK. Peripheral blood gene expression profiling in rheumatoid arthritis. *Genes Immun* 2005; 6: 388-397.
3. Battistini L., Borsellino G., Sawicki G., Poccia F., Salvetti M., Ristori G., Brosnan CF. Phenotypic and cytokine analysis of human peripheral blood gamma delta T cells expressing NK cell receptors. *J Immunol* 1997; 159: 3723-30.
4. Bennett L., Palucka AK., Arce E., Cantrell V., Borvak J., Banchereau J., Pascual V. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 2003; 197: 711-723.
5. Bomprezzi R., Ringner M., Kim S., Bittner ML., Khan J., Chen Y., Elkhoulou A., Yu A., Bielekova B., Meltzer PS., Trent JM. Gene expression profile in multiple sclerosis patients and healthy controls: identify-

- ing pathways relevant to disease. *Hum Mol Genet* 2003; 12: 2191-2199.
6. Crow MK., Kirou KA., Wohlgemuth J. Microarray analysis of interferon-regulated genes in SLE. *Autoimmunity* 2003; 36: 481-490.
  7. Crow MK., Wohlgemuth J. Microarray analysis of gene expression in lupus. *Arthritis Res Ther* 2003; 5: 279-287.
  8. Cuzner ML., Gveric D., Strand C., Loughlin AJ., Paemen L., Opdenakker G., Newcombe J. The expression of tissue-type plasminogen activator, matrix metalloproteases and endogenous inhibitors in the central nervous system in multiple sclerosis: comparison of stages in lesion evolution. *J Neuropathol Exp Neurol* 1996; 55: 1194-204.
  9. Esiri MM., Gay D. The immunocytochemistry of multiple sclerosis plaques. In: Raine CS, McFarland HF, Tourtellotte WW, eds. *Multiple sclerosis. Clinical and pathogenetic basis*. London: Chapman & Hall; 1997; 173-86.
  10. Furlan R., Rovaris M., Martinelli Boneschi F., Khademi M., Bergami A., Gironi M., Deleidi M., Agosta F., Franciotta D., Scarpini E., Uccelli A., Zaffaroni M., Kurne A., Comi G., Olsson T., Filippi M., Martino G. Immunological patterns identifying disease course and evolution in multiple sclerosis patients. *J Neuroimmunol.* 2005; 165: 192-200.
  11. Goverman J., Woods A., Larson L., Weiner LP., Hood L., Zaller DM. Transgenic mice that express a myelin basic protein-specific T cell receptor develop spontaneous autoimmunity. *Cell* 1993; 72: 551-60.
  12. Han GM., Chen SL., Shen N., Ye S., Bao CD., Gu YY. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. *Genes Immun* 2003; 4: 177-186.
  13. Huterer SJ., Tourtellotte WW., Wherrett JR. Alterations in the activity of phospholipases A2 in postmortem white matter from patients with multiple sclerosis. *Neurochem Res* 1995; 20: 1335-43.
  14. Jarvis JN., Dozmorov I., Jiang K., Frank MB., Szodoray P., Alex P., Centola M. Novel approaches to gene expression analysis of active polyarticular juvenile rheumatoid arthritis. *Arthritis Res Ther* 2004; 6: R15-R32.
  15. Karpus WJ., Ransohoff RM. Chemokine regulation of experimental autoimmune encephalomyelitis: temporal and spatial expression patterns govern disease pathogenesis. *J Immunol* 1998; 161: 2667-71.
  16. Lafaille JJ., Nagashima K., Katsuki M., Tonegawa S. High incidence of spontaneous autoimmune encephalomyelitis in immunodeficient anti-myelin basic protein T cell receptor transgenic mice. *Cell* 1994; 78: 399-408.
  17. Lassmann H., Bruck W., Lucchinetti C, Rodriguez M. Remyelination in multiple sclerosis. *Mult Scler* 1997; 3: 133-6.
  18. Li H., Cuzner ML., Newcombe J. Microglia-derived macrophages in early multiple sclerosis plaques. *Neuropathol App Neurobiol* 1996; 22: 207-15.
  19. Lock C., Hermans G., Pedotti R., Brendolan A., Schadt E., Garren H., Langer-Gould A., Strober S., Cannella B., Allard J., Klonowski P., Austin A., Lad N., Kaminski N., Galli SJ., Oksenberg JR., Raine CS., Heller R., Steinman L. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis.

- Nat Med 2002; 8: 500-508.
20. Lucchinetti CF., Bruck W., Rodriguez M., Lassmann H. Distinct patterns of multiple sclerosis pathology indicates heterogeneity on pathogenesis. *Brain Pathol* 1996; 6: 259-74.
  21. Maas K., Chan S., Parker J., Slater A., Moore J., Olsen N., Aune TM: Molecular portrait of human autoimmune disease. *J Immunol* 2002; 169: 5-9.
  22. Martino G, Hartung HP. Immunopathogenesis of multiple sclerosis: the role of T cells. *Curr Opin Neurol.* 1999; 12; 309-321.
  23. Narayana PA., Doyle TJ., Lai D., Wolinsky JS. Serial proton magnetic resonance spectroscopic imaging, contrast-enhanced magnetic resonance imaging, and quantitative lesion volumetry in multiple sclerosis. *Ann Neurol* 1998; 43: 56-71.
  24. Newcombe J., Li H., Cuzner ML. Low density lipoprotein uptake by macrophages in multiple sclerosis plaques: implications for pathogenesis. *Neuropathol App Neurobiol* 1994; 20: 152-62.
  25. Olsen N., Sokka T., Seehorn CL., Kraft B., Maas K., Moore J., Aune TM. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. *Ann Rheum Dis* 2004; 63: 1387-1392 .
  26. Plebani M. Errors in clinical laboratories or errors in laboratory medicine? *Clin.Chem Lab Med* 2006; 44: 750-759.
  27. Selmaj KW., Raine CS., Cannella B., Brosnan CF. Identification of lymphotoxin and tumor necrosis factor in multiple sclerosis lesions. *J Clin Inv* 1991; 87: 949-54.
  28. Steinman L. A few autoreactive cells in an autoimmune infiltrate control a vast population of non specific cells: a tale of smart bombs and the infantry. *Proc Natl Acad Sci USA* 1996; 93: 2253-6.
  29. Teunissen CE., Dijkstra C., Polman C. Biological markers in CSF and blood for axonal degeneration in multiple sclerosis. *Lancet Neurol.* 2005; 4: 32-4.

## **Biomarcatori per la sclerosi laterale amiotrofica**

*Gabriele Mora<sup>1</sup>, Caterina Bendotti<sup>2</sup>*

<sup>1</sup>Dipartimento di Neurologia, Fondazione Salvatore Maugeri, IRCCS, Pavia

<sup>2</sup>Istituto di Ricerche Farmacologiche "Mario Negri", Milano

La sclerosi laterale amiotrofica (SLA) è una malattia neurodegenerativa ad esito fatale che colpisce i motoneuroni. La conseguenza funzionale è una paralisi muscolare progressiva che colpisce anche i muscoli respiratori portando il paziente a morte prematura generalmente entro i 5 anni dall'esordio dei sintomi. La SLA si presenta in forma sporadica (oltre il 90% dei casi) e familiare. Solo nel 20% delle forme familiari l'eziologia è conosciuta ed è correlata alla mutazione del gene SOD1. In tutti gli altri casi, l'eziologia è sconosciuta benché sia sempre più probabile l'ipotesi di una predisposizione genetica che, in sinergia con determinati fattori di rischio ambientale, sia responsabile dello sviluppo della patologia e della variabilità del suo decorso.

Attualmente la diagnosi di SLA è ancora una diagnosi di esclusione che si basa su criteri clinici che nella maggior parte dei casi si rendono evidenti solo a malattia conclamata. La mancanza di marcatori diagnostici e prognostici della malattia ha indirizzato la ricerca verso l'identificazione appunto di marcatori precoci di malattia il più possibile specifici e facilmente misurabili. A questo scopo vengono impiegate diverse tecnologie (genomica, proteomica e metabolomica) per l'analisi di campioni biologici quali sangue, liquido cerebrospinale e muscolo provenienti da pazienti SLA in confronto con pazienti affetti da altre malattie neurologiche o con individui sani.

Studi recenti hanno identificato alcune proteine dal sangue e dal liquor che potrebbero rappresentare dei biomarcatori per la SLA, tuttavia è necessaria una validazione di queste osservazioni preliminari su un'ampia popolazione di pazienti a differenti stadi di malattia per definire il ruolo di queste proteine nelle diverse fasi della malattia. L'identificazione di biomarcatori proteici specifici per la SLA daranno la possibilità non solo di sviluppare saggi di laboratorio per una diagnosi precoce della malattia ma potrà servire anche per valutare l'efficacia di una terapia nei clinical trials e per identificare nuovi bersagli farmacologici. I risultati di studi recenti di genomica e proteomica saranno discussi durante la presentazione.

## Introduzione

La sclerosi laterale amiotrofica (SLA) è una malattia neurodegenerativa ad esito fatale che colpisce i motoneuroni della corteccia motoria (I° motoneurone) e del bulbo-midollo spinale (II motoneurone). La conseguenza funzionale è una paralisi muscolare progressiva che colpisce anche i muscoli respiratori portando il paziente a morte prematura. Questa malattia presenta un'incidenza di circa 2 casi su 100.000 abitanti per anno con espressione ubiquitaria (vi sono solo dati nei Paesi Occidentali); la prevalenza è di circa 6 casi su 100.000. Attualmente in Italia sono circa 4.000 gli individui affetti da questa malattia. La SLA si presenta in forma sporadica (oltre il 90% dei casi) e familiare. Solo nel 10-20% delle forme familiari l'eziologia è conosciuta ed è correlata alla mutazione del gene SOD1. In tutti gli altri casi, l'eziologia è sconosciuta benché sia sempre più probabile l'ipotesi di una predisposizione genetica che, in sinergia con determinati fattori di rischio ambientale, sia responsabile dello sviluppo della patologia e della variabilità del suo decorso. Attualmente il tempo tra l'esordio dei sintomi e la diagnosi è di circa un anno, un tempo molto lungo se si considera che la prognosi è in media di 3 anni. La diagnosi infatti è basata principalmente sull'esclusione di malattie che possono mimare la SLA e su criteri clinici che nella maggior parte dei casi sono evidenti solo a malattia conclamata. Ciò si riflette negativamente sulla possibilità di un tempestivo intervento terapeutico. Infatti nonostante nei modelli animali di SLA esistano diversi esempi di trattamento in grado di rallentare la patologia quando applicati prima o all'inizio dell'esordio dei sintomi (1) nelle sperimentazioni cliniche tali molecole sono risultate inefficaci, forse anche perché somministrate quando la malattia è ad uno stadio conclamato (2). Recentemente sono state sviluppate e applicate le tecnologie di genomica, trascrittomica e proteomica per l'identificazione di marcatori biologici specifici per la SLA, che potrebbero essere di aiuto nella diagnosi, nella valutazione del decorso, e nell'interpretazione degli effetti terapeutici di nuove molecole di rilevanza clinica.

L'analisi genomica si basa sull'indagine di mutazioni e/o polimorfismi a carico di singoli nucleotidi nell'intero genoma o in intere classi di categorie di geni in un'ampia popolazione di individui affetti da malattia sporadica in confronto a controlli sani. L'analisi trascrittomica, attraverso l'uso di microarray è in grado, in un unico chip, di rappresentare tutti i geni trascrizionalmente attivi nella cellula o tessuto in esame. La rappresentazione che se ne ottiene è un quadro dettagliato dello stato funzionale della cellula o del tessuto patologico rispetto a quello sano per individuare alterazioni geniche sia qualitative che quantitative quali possibili cause di alterazioni proteiche. La proteomica permette un'analisi differenziale dei pattern di espressione proteica fra un campione sano ed uno patologico. Essa si basa sull'elettroforesi bidimensionale di proteine estratte da cellule o tessuti seguita dall'identificazione delle singole proteine mediante spettrometria di massa. Inoltre con la proteomica è possibile visualizzare le differenti modificazioni post-traduzionali che la proteina può subire in relazione alle particolari condizioni in cui essa si trova, tra cui le modificazioni ossidative che modificano



l'attività fisiologica e/o la degradazione della proteina stessa. Un esempio è la nitrificazione proteica che potrebbe svolgere un ruolo importante nella patogenesi della malattia, come vedremo più avanti.

Importanti indicazioni per l'identificazione di potenziali biomarcatori sono derivate dallo studio di modelli sperimentali della malattia quali i topi transgenici esprimenti il gene umano della SOD1 mutata. Questi topi sono caratterizzati dalla comparsa di una progressiva paralisi muscolare, con quadro clinico e patologico sovrapponibile a quanto osservato nei pazienti, sino alla morte precoce dell'animale.

Un modello transgenico ampiamente caratterizzato e utilizzato dalla maggior parte dei laboratori che si occupano di SLA è quello che esprime ad alti livelli la mutazione umana SOD1 G93A. Questo topo mostra i primi segni clinici della malattia a 80-90 giorni e muore a 130-150 giorni di vita (3).

Il vantaggio del modello animale è quello di conoscere esattamente quando avviene l'esordio della malattia permettendo quindi di monitorare eventuali alterazioni biochimiche in tessuti periferici facilmente reperibili (e.g. sangue, biopsie muscolari e cutanee) e nel sistema nervoso colpito, in relazione al grado di neurodegenerazione motoneuronale ed alla progressione sintomatologica della malattia. In questo articolo prenderemo in considerazione alcuni dei risultati principali ottenuti nel nostro laboratorio e in altri gruppi sui modelli murini di SLA riguardo all'identificazione di potenziali biomarcatori. Inoltre analizzeremo alcuni studi sul profilo proteomico del liquido cerebrospinale (CSF) di pazienti affetti da SLA.

### **Analisi genomica della SLA**

A partire dal 1993 con la scoperta della prima mutazione del gene della SOD1, diverse altre mutazioni genetiche sono state riconosciute in alcune forme di SLA familiare (4). Tuttavia, per quanto riguarda le forme sporadiche della malattia solo il recente sviluppo di nuove tecnologie genetiche basate sull'indagine di polimorfismi a carico di singoli nucleotidi ha offerto la possibilità di eseguire un rapido screening di una vasta popolazione di pazienti affetti da SLA sporadica per identificare delle varianti genetiche associate ad un aumentato rischio della malattia. Recentemente sono stati infatti pubblicati diversi studi di "whole-genome association" (WGA) in diverse coorti di pazienti e alcuni di essi hanno identificato dei geni potenzialmente associati alla malattia. Ad esempio nello studio eseguito dal Translational Genome Research Institute di Phoenix, Arizona su 1.152 pazienti è stato identificato un gene, FLJ10986, dalla funzione sconosciuta presente sul cromosoma 1, la cui associazione con la malattia è risultata significativa (5) Successivamente un gruppo olandese del Medical Center di Utrecht, riportava che il recettore 2 dell'inositolo 1,4,5 trifosfato (ITPR2), era associato con la SLA in 3 popolazioni europee (6). L'ITPR2 è una proteina implicata nel rilascio intraneuronale di calcio il cui eccesso è considerato tra le potenziali cause della degenerazione motoneuronale nella SLA. Lo stesso gruppo olandese più recentemente ha identificato in 4 coorti di pazienti provenienti da 4 diversi paesi un altro gene di suscettibilità per la SLA sporadica che codifica per l'enzima



dipeptidil-petidasi 6 (DPP6), la cui funzione è quella di regolare i neuropeptidi (7). Sia ITPR2 che DPP6, tuttavia, non hanno dimostrato varianti significative associate alla malattia nello studio WGA di una popolazione americana di 276 pazienti (8). Questo dimostra che solamente dall'analisi di una popolazione molto vasta di pazienti sia possibile identificare dei potenziali geni candidati di suscettibilità ed è probabile che per molti di essi l'associazione con la malattia sia dipendente da un'interazione con l'ambiente.

### **Analisi dell'espressione genica e proteica in un modello di SLA familiare**

L'applicazione di un sistema integrato di analisi informatica dei profili di trascrittomico e proteomico ottenuti da estratti di midollo spinale di topi transgenici SOD1G93A in confronto a topi controllo e l'uso del programma di "gene ontology" ha permesso di individuare delle variazioni di famiglie geniche associate a specifici processi biologici, prima dell'esordio dei sintomi. Per esempio, tra le categorie di proteine che sono significativamente alterate emerge una ridotta espressione delle proteine mitocondriali mentre risultano sovra-rappresentate proteine associate al metabolismo cellulare, alla regolazione dei segnali di trasduzione (come alcune protein-chinasi), all'attività antiossidante e all'attività lisosomiale per la degradazione proteica (9).

In linea con questi dati, in uno studio recente di analisi proteomica differenziale su gel bidimensionale, seguita dall'analisi di spettrometria di massa, anche noi abbiamo osservato alterazioni nell'espressione di proteine specifiche associate alla disfunzione mitocondriale, alla formazione di aggregati intracellulari e allo stress ossidativo nel midollo spinale di topi SOD1G93A allo stadio presintomatico in confronto con topi controllo (10). In particolare abbiamo trovato un aumento dei livelli di piruvato deidrogenasi E1 che trasforma il piruvato in acetylCoA per il ciclo dell'acido citrico mentre la catena beta dell'ATP sintasi, la subunità catalitica del complesso proteico che sintetizza ATP da ADP era significativamente ridotta. Le alterazioni dei livelli di questi enzimi possono essere un indice di una severa disfunzione mitocondriale. Questo è compatibile con le alterazioni morfologiche dei mitocondri che sono tra i primi segnali neuropatologici osservabili nei motoneuroni dei topi SOD1G93A. Anche nei pazienti affetti da SLA sporadica sono state riportate alterazioni della morfologia e/o della funzionalità mitocondriale in vari distretti cellulari periferici inclusi i terminali degli assoni motori. La disfunzione di questi organelli che può derivare da un eccessivo stress ossidativo della cellula per la presenza di SOD1 mutata o da altri meccanismi tossici può a sua volta determinare una elevata produzione di radicali liberi (es. perossinitrito e radicali idrossilici) e può provocare un deficit energetico della cellula per la ridotta produzione di ATP. Altre proteine che sono risultate essere differenzialmente espresse nei topi SOD1G93A allo stadio presintomatico sono degli enzimi coinvolti nella degradazione proteica come la serpina, un inibitore della serina proteasi e un enzima di ubiquitinazione dei substrati proteici (NEDD8 conjugating enzyme Ubc12) che sono rispettivamente ridotto o aumentato nei topi SOD1G93A rispetto ai controlli. Entrambe queste alterazioni potreb-

bero essere un tentativo di risposta delle cellule nervose a limitare l'aggregazione proteica che fa parte del processo degenerativo dei motoneuroni, come risulta dall'accumulo di inclusioni proteiche intracellulari evidenziate dall'analisi neuropatologica del midollo spinale di pazienti SLA e dei modelli murini della malattia. Oltre ad un risultato differenziale quantitativo, l'analisi proteomica su gel bidimensionale permette inoltre di individuare alterazioni nel pattern di alcune isoforme proteiche così come le alterazioni post-trasduzionali. Ad esempio la ciclofillina A, una chaperonina che protegge le cellule dallo stress ossidativo è risultata alterata nella composizione delle sue isoforme.

Sebbene il coinvolgimento nella patologia di queste variazioni della ciclofillina A sia ancora da definire, è interessante il fatto che esse sono in parte riscontrabili anche in cellule periferiche di pazienti affetti da SLA sporadica, costituendo così dei potenziali biomarcatori della malattia.

### **Proteomica ossidativa nei topi modello di SLA familiare e nei pazienti con SLA sporadica**

Tra le modificazioni post-trasduzionali delle proteine, l'ossidazione costituisce uno dei meccanismi più probabili come risultato di un eccessivo e prolungato stato di stress ossidativo. Infatti sia in pazienti affetti da SLA (sporadica e familiare) che in modelli sperimentali è stato rilevato un aumento delle specie reattive dell'ossigeno (ROS) e dell'azoto (RNS) associato alla modificazione ossidativa di varie proteine e altri componenti cellulari (11). In particolare, si ipotizza che il danno ossidativo mediato dal perossinitrito, il prodotto di reazione dell'ossido nitrico (NO) con il superossido, attraverso la nitratura a carico della tirosina, possa giocare un ruolo importante nella patogenesi della SLA (12). La nitratura in tirosina è considerata una modifica irreversibile che nella maggior parte dei casi porta ad un danno permanente della proteina coinvolta, che può diventare più suscettibile alla degradazione attraverso il sistema proteasoma/ubiquitina.

Recentemente abbiamo dimostrato che nei topi transgenici che esprimono il gene umano della SOD1 mutata i livelli di proteine ntrate sono significativamente aumentati nel midollo spinale prima dell'esordio dei sintomi, quando ancora è presente la maggior parte dei motoneuroni (13). Dall'analisi proteomica abbiamo inoltre potuto identificare alcune delle proteine maggiormente ntrate nei topi con SOD1 mutata rispetto ai controlli. Si tratta di proteine implicate nel metabolismo energetico (es. l'ATP sintasi e le enolasi alfa e gamma), nella degradazione proteica (es. heat shock cognate protein 71 KDa) e nell'assetto citoscheletrico della cellula (es. actina). Il meccanismo attraverso cui avviene un'aumentata nitratura proteica - se a seguito di un'eccessiva attivazione delle diverse ossido nitrico sintasi a produrre NO e quindi perossinitrito, e/o come conseguenza di una ridotta degradazione delle proteine ntrate - ancora non è chiaro. Nel midollo spinale dei topi FALS ad uno stadio avanzato della malattia abbiamo osservato una significativa riduzione dei livelli delle subunità costitutive del proteasoma (14), il che farebbe supporre che l'aumento delle proteine ntrate durante la progressione della malattia sia espressione di una ridotta degradazione.

Recentemente abbiamo analizzato i livelli di nitro-proteine nelle cellule mononucleate del sangue periferico (PBMC) di pazienti con SLA sporadica a diversi stadi della malattia e abbiamo osservato che alcune proteine nitate nel gel bidimensionale risultavano significativamente aumentate rispetto agli individui sani. Tra esse sono presenti due delle proteine che abbiamo visto essere maggiormente nitate anche nel midollo spinale dei topi SOD1 mutati allo stadio presintomatico. Si tratta della actina e dell'ATP sintasi. L'aumento delle proteine nitate nei PBMC dei pazienti SLA non sembra essere correlato alla gravità della malattia in quanto si osserva sia nei pazienti che sono moderatamente o severamente affetti. Lo sviluppo di anticorpi specifici contro le forme nitate di alcune proteine, attualmente in corso, ci permetterà di confermare questi dati su una più vasta popolazione di pazienti SLA mediante immunoblot monodimensionale e poter quindi sviluppare dei test diagnostici facilmente applicabili nel laboratorio clinico come supporto biochimico alla diagnosi della malattia (15).

Livelli aumentati di NT sono stati riscontrati anche nel CSF di pazienti con SLA sporadica (16) anche se in uno studio più recente tale effetto è stato confermato solamente in una sottopopolazione di pazienti (17). Attualmente nel nostro laboratorio è in corso l'analisi proteomica per l'identificazione delle proteine nitate nel liquor dei pazienti SLA. Queste indagini permetteranno soprattutto di identificare dei bersagli farmacologici specifici per la malattia.

### **Profilo proteomico del CSF di pazienti SLA**

Uno dei problemi che rende difficile lo studio dei meccanismi patogenetici della SLA e più in generale delle malattie neurodegenerative, è l'impossibilità di utilizzare il tessuto nervoso come mezzo d'indagine se non come reperto autoptico. Recentemente è stato ipotizzato che le variazioni nella composizione e nell'espressione proteica del CSF possono rispecchiare le alterazioni che avvengono a livello del tessuto nervoso, come ad esempio il rilascio di proteine o di frammenti di proteine da neuroni o glia affetti dalla patologia. Per questo una dettagliata analisi del profilo proteomico del CSF può fornire importanti informazioni sulla natura del processo degenerativo che avviene a livello dei motoneuroni e permettere così di identificare potenziali biomarcatori per una diagnosi precoce della malattia. Attualmente sono pochi gli studi che hanno approntato una analisi di questo tipo presumibilmente per la difficoltà di reperire sufficienti quantità di liquor. Tra questi, quelli più significativi che hanno fatto uso di tecnologie avanzate e sofisticate della spettrometria di massa come ad esempio la "surface-enhanced laser desorption/ionization time of flight mass spectrometry (SELDI-TOF-MS) hanno dimostrato l'alterazione nel contenuto di alcune specie proteiche delle quali la diminuzione della cistatina C, una proteina di secrezione appartenente alla classe degli inibitori delle cisteine proteasi che svolge un ruolo importante nell'omeostasi proteica extracellulare, sembra essere la più consistente e riproducibile (18, 19). La cistatina C è una componente dei corpi di Bunina che sono piccole inclusioni eosinofile che si riscontrano tipicamente nei motoneuroni di reperti autoptici di pazienti SLA. Una sua riduzione nel CSF potrebbe dunque essere legata al seque-

stro della proteina in queste inclusioni, tuttavia non è da escludere che esista un legame funzionale tra la riduzione di questa proteina e l'aggregazione proteica come è stato dimostrato in una rara forma ereditaria di angiopatia cerebrale amiloidea associata alla mutazione del gene della cistatina C.

## Conclusioni

Dalle osservazioni sopra riportate risulta chiaro che una notevole mole di dati stanno emergendo dalle differenti analisi di espressione genomica e proteomica e probabilmente solo dal confronto tra di essi sarà possibile disegnare un quadro quanto più completo e dettagliato sui possibili biomarcatori diagnostici e prognostici della SLA oltre che individuare potenziali bersagli terapeutici. Un approccio particolarmente importante sarà quello di combinare le informazioni provenienti dall'analisi genomica della popolazione SLA con le alterazioni di espressione genica e proteica in tessuti facilmente reperibili come il sangue al fine di individuare i potenziali fattori di rischio sia genetici che ambientali, della malattia. Inoltre poichè la SLA sporadica è ritenuta una patologia eterogenea questi confronti ci permetteranno di identificare dei sottogruppi di pazienti affetti dalla malattia che potrebbero rispondere in maniera diversa a specifici approcci terapeutici favorendo quindi una terapia individuale più mirata.

## Bibliografia

1. Carri MT., Grignaschi G., Bendotti C. Targets in ALS: designing multidrug therapies. *Trends Pharmacol. Sci.* 2006; 27: 267-273
2. Traynor BJ., Buijn L., Conwit R., Beal F., O'Neill G., Fagan SC., Cudkowicz ME. Neuroprotective agents for clinical trials in ALS: a systematic assessment. *Neurology* 2006; 67: 207
3. Bendotti C., Carri MT. Lessons from models of SOD1-linked familial ALS. *Trends Mol Med* 2004; 10: 393-400.
4. Valdmans PN., Rouleau GA. Genetics of familial amyotrophic lateral sclerosis. *Neurology*. 2008; 70: 144-52.
5. Dunckley T., Huentelman MJ., Craig DW., Pearson JV., Szelinger S., Joshipura K., Halperin RF., Stamper C, Jensen KR., Letizia D., Hesterlee SE., Pestronk A., Levine T., Bertorini T., Graves MC., Mozaffar T., Jackson CE., Bosch P., McVey A., Dick A., Barohn R., Lomen-Hoerth C., Rosenfeld J., O'connor DT., Zhang K., Crook R., Ryberg H., Hutton M., Katz J., Simpson EP., Mitsumoto H., Bowser R., Miller RG., Appel SH., Stephan DA. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N Engl J Med.* 2007; 357: 775-788.
6. Van Es MA., Van Vught PW., Blauw HM., Franke L., Saris CG., Andersen PM., Van Den Bosch L., de Jong SW., van 't Slot R., Birve A., Lemmens R., de Jong V., Baas F., Schelhaas HJ., Sleegers K., Van Broeckhoven C., Wokke JH., Wijmenga C., Robberecht W., Veldink JH., Ophoff RA., van den Berg LH. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis:

- a genome-wide association study. *The Lancet Neurology* 2007; 6: 869-877
7. Van Es MA., van Vught PWJ., Blauw HM., Franke L., Saris GH., Van Den Bosch L., de Jong SW., de Jong V., Baas F., van 't Slot R., Lemmens R., Schelhaas HJ., Birve A., Slegers K., Van Broeckhoven C., Schymick JC., Traynor BJ., Wokke JHJ., Wijmenga C., Robberecht W., Andersen PM., Veldink JH., Ophoff RA., Leonard H. van den Berg. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. *Nature Genetics*, 2008; 40: 29-31.
  8. Schymick JC., Scholz SW., Fung HC., Britton A., Arepalli S., Gibbs JR., Lombardo F., Matarin M., Kasperaviciute D., Hernandez DG., Crews C., Bruijn L., Rothstein J., Mora G., Restagno G., Chiò A., Singleton A., Hardy J., Traynor BJ. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *The Lancet Neurology* 2007; 6: 322-328.
  9. Lukas TJ., Luo WW., Mao H., Cole N., Siddique T. Informatics-assisted protein profiling in a transgenic mouse model of amyotrophic lateral sclerosis. *Molec. Cell. Proteomics* 2006; 5, 7: 1233-1244.
  10. Massignan T., Casoni F., Basso M., Stefanazzi P., Biasimi E., Tortarolo M., Salmone M., Gianazza E., Bendotti C., Bonetto V. Proteomic analysis of spinal cord of presymptomatic amyotrophic lateral sclerosis G93A SOD1 mouse. *Bioch. Biophys. Res. Comm.* 2007; 353: 719-725.
  11. Beal MF. Oxidatively modified proteins in aging and disease. *Free Radic Biol Med* 2002; 32: 797-803.
  12. Estévez AG., Spear N., Manuel SM., Barbeito L., Radi R., Beckman JS. Role of endogenous nitric oxide and peroxynitrite formation in the survival and death of motor neurons in culture. *Prog Brain Res* 1998; 118: 269-80.
  13. Casoni F., Basso M., Massignan T., Gianazza E., Cheroni C., Salmons M., Bendotti C., Bonetto V. Protein nitration in a mouse model of familial amyotrophic lateral sclerosis: possible multifunctional role in the pathogenesis. *J Biol Chem* 2005; 280: 16295-304.
  14. Cheroni C., Peviani M., Cascio P., Debiassi S., Monti C., Bendotti C. Accumulation of human SOD1 and ubiquitinated deposits in the spinal cord of SOD1G93A mice during motor neuron disease progression correlates with a decrease of proteasome. *Neurobiol Dis* 2005; 18: 509-522.
  15. Alimonti D., Sonetto V., Nardo G., Mantovani S., Gabelli S., Bendotti C., Mora G. Proteomic analysis of nitrated proteins from peripheral blood mononuclear cells of patients with sporadic ALS. *Amyotroph. Lateral Scler.* 2007; 8 (Suppl1): 122.
  16. Tohgi H., Abe T., Yamazaki K., Murata T., Ishizaki E., Isobe C. Remarkable increase in cerebrospinal fluid 3-nitrotyrosine in patients with sporadic amyotrophic lateral sclerosis. *Ann Neurol* 1999; 46: 129-31.
  17. Ryberg H., Soderling AS., Davidsson P., Blennow K., Caidahl K., Persson LI. Cerebrospinal fluid levels of free 3-nitrotyrosine are not elevated in the majority of patients with amyotrophic lateral sclerosis or Alzheimer's disease. *Neurochem Int* 2004; 45: 57-62.

18. Ranganathan S., Williams E., Ganchev P., Gopalakrishnan V., Lacomis D., Urbinelli L., Newhall K., Cudkowicz ME. Brown RH. Jr., Bowser R. Proteomic profiling of cerebrospinal fluid identifies biomarkers for amyotrophic lateral sclerosis. *J. Neurochem.* 2005; 95: 1461-1471.
19. Pasinetti G.M., Ungar L.H., Lange D.J., Yemul S., Deng H., Yuan X., Brown R.H., Cudkowicz M.E., Newhall K., Peskind E., Marcus S., Ho L. Identification of potential CSF biomarkers in ALS. *Neurology* 2006; 66: 1218-1222.



**MEDICINA PREDITTIVA  
E FARMACOGENOMICA**





## **Medicina predittiva: il profilo genetico di predisposizione alla malattia e il ruolo dell'ambiente**

*Paola Mandich*

Dipartimento di Neuroscienze, Oftalmologia e Genetica (DINOG), Sezione di Genetica Medica, Università di Genova

È ormai superato il concetto che i geni contengano tutta l'informazione necessaria per un programma fisso e predeterminato che viene seguito fedelmente portando a compimento un individuo la cui definizione finale è già predestinata nei geni che riceve dai genitori. Questa analogia minimizza il ruolo dell'ambiente che, invece, agisce in complessa interazione con i geni nel determinare lo sviluppo di ciascun individuo. Il gene porta l'informazione necessaria alla formazione di strutture chimiche complesse, le proteine, che sono costituenti fisiche del corpo. Senza i geni non esisterebbe il corpo. Allo stesso modo, senza un ambiente, prima nell'utero e poi fuori da esso, un individuo non potrebbe esistere. Perciò lo sviluppo *normale* di ogni individuo dipende dall'interazione tra geni e ambiente. Negli anni '60 e '70 vi è stato un ampio dibattito sul contributo del patrimonio genetico nel determinare il livello di intelligenza e il comportamento negli individui. Allo stesso modo è possibile discutere quale sia il relativo apporto di geni e ambiente alla *salute* e viceversa all'assenza di salute - *malattia* - di ciascuno di noi. È ormai un dato acquisito il fatto che geni e ambiente concorrano entrambi all'equilibrio di un sistema nel quale lo spostamento in un senso o nell'altro può dipendere da entrambi nello stesso modo o da uno dei due fattori in maniera prevalente ma sempre e comunque da un'interazione tra essi. La scoperta di geni responsabili di malattie ereditarie e la mappatura del genoma umano hanno ampliato la conoscenza di queste interazioni che rimangono però ancora in larga parte da comprendere. Dall'anno della scoperta della struttura del DNA sono stati ottenuti notevoli avanzamenti nella comprensione delle basi genetiche della variabilità dell'uomo e delle malattie. È nota la base genetica di più di 2000 patologie geneticamente determinate di tipo mendeliano semplice (un gene - una malattia) e sono quindi disponibili test genetici per uso clinico. La componente ereditaria in molte malattie comuni (diabete, ipertensione, obesità, tumori, malattia coronarica, ictus) è invece complessa e coinvolge generalmente varianti genetiche multiple che interagiscono con l'ambiente. Ambiente inteso in senso lato che comprende non solo l'ambiente circostante all'individuo ma l'individuo stes-

so con le sue abitudini (dieta, fumo, attività fisica ecc.) e le sue malattie (comorbidità). Sono note al momento poche varianti che predispongono a queste patologie, ciascuna delle quali contribuisce in modo modesto al rischio finale di malattia e interagisce a sua volta con fattori ambientali diversi a seconda dell'individuo e della popolazione in esame.

È quindi immaginabile che, a seconda del diverso contributo di fattori genetici o di fattori ambientali alla patogenesi di una malattia, la sua comparsa possa essere prevista in base al profilo genetico di un individuo. Mentre per patologie mendeliane dominanti a penetranza completa la presenza della mutazione genica che causa la malattia è generalmente associata alla comparsa della patologia, per le patologie complesse è necessario che il fattore genetico interagisca con fattori ambientali o dell'individuo per determinarne la comparsa. È possibile immaginare un continuum con un estremo al quale si trovano le malattie causate da un singolo gene e l'altro che include le patologie comuni nelle quali la componente ambientale e di stile di vita abbiano il ruolo predominante. In questo scenario il ruolo del profilo genetico di un individuo assume diversa importanza a seconda della posizione nella quale si trova la patologia in considerazione.

La definizione di profilo genetico è quindi ampia e può includere: l'identificazione della presenza di una mutazione patogenetica, la combinazione di mutazioni/polimorfismi che concorrono alla predisposizione per una data malattia, l'insieme di prodotti genici diversamente espressi in individui diversi o in tessuti diversi (sani/patologici).

I tumori, le malattie psichiatriche e la malattia cardiovascolare possono essere presi in considerazione come esempi di interazione profilo di predisposizione/ambiente.

La ricerca su cancro e ambiente ha portato alla conoscenza del tumore come processo caratterizzato da passaggi successivi nei quali nella cellula avvengono cambiamenti che determinano la perdita dei meccanismi di controllo della crescita e della replicazione cellulare. Il modello convenzionale usato per descrivere la carcinogenesi (Fig. 1) suppone l'esistenza di numerose mutazioni in successione a carico di un singolo clone cellulare che inducono la graduale perdita dei meccanismi di controllo della divisione cellulare e del mantenimento dell'integrità del patrimonio genetico. In una catena di eventi i carcinogeni ambientali possono innescare il processo; centinaia di geni vengono successivamente coinvolti nel trasporto e eliminazione degli xenobiotici, nei meccanismi di riparo e nella risposta immune o infiammatoria. Ciascuno di questi geni può avere varianti con livelli di attività diversi che possono quindi modificare la suscettibilità individuale ai tumori. Le cellule circostanti il clone patologico esercitano anch'esse un effetto diretto o indiretto sul processo. Il tumore risulta perciò da un processo complesso nel quale geni e fattori ambientali intervengono con multiple interazioni.

È stato stimato che le donne hanno un rischio di 1:9 di sviluppare un tumore della mammella nel corso della vita (a 85 anni). Questo rischio non è distribuito omogeneamente nella popolazione poiché alcune donne non avranno mai questo tumore mentre altre sono a rischio aumentato e avranno la probabilità di sviluppare più tumori della mammella nel corso della vita. Per queste donne ad alto

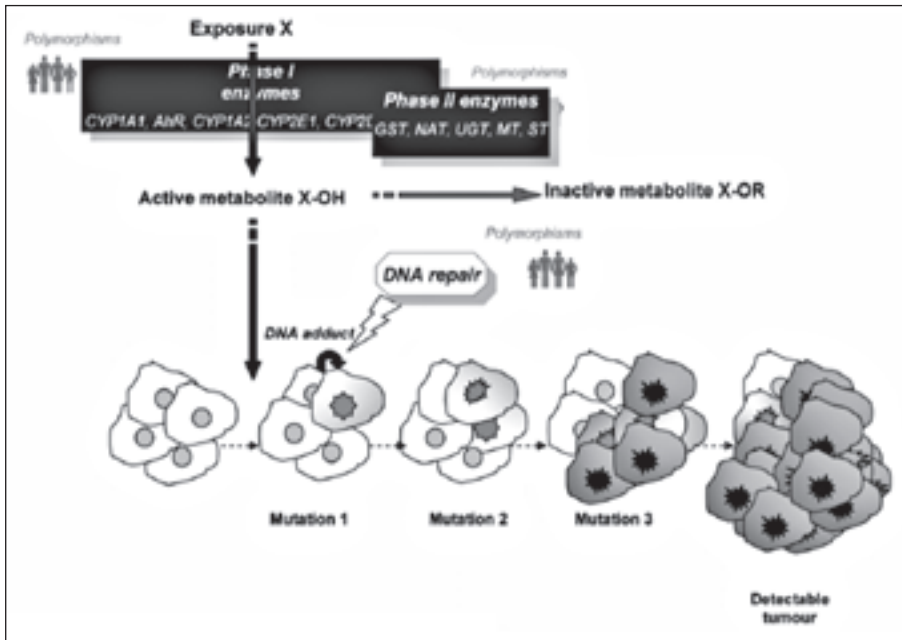


Fig. 1 - Da Clavel CR, Biologies 330:306-317, 2007.

rischio è necessario seguire un programma di sorveglianza/trattamento ad hoc. È quindi estremamente importante essere in grado di identificarle in base a un profilo genetico ad elevato rischio per garantire loro un trattamento adeguato. Sono noti oggi numerosi fattori di rischio per il tumore della mammella (Tab. 1) che includono sia abitudini di vita e caratteristiche della donna e di storia tumorale pregressa, sia fattori genetici predisponenti.

I fattori genetici possono modificare lievemente il rischio così come in modo drammatico. Infatti la presenza di storia familiare positiva aumenta il rischio di 2-3 volte mentre la presenza di mutazioni nei geni BRCA1 e BRCA2 aumenta il rischio fino al 60-70% (a 70 anni). Anche nel caso di mutazioni di questo tipo, con trasmissione autosomica dominante e elevata penetranza, è comunque necessario che altri fattori contribuiscano alla comparsa del tumore. Infatti la presenza della mutazione non è di per se necessaria né sufficiente a determinare la comparsa del tumore. È quindi l'interazione di un profilo genetico di predisposizione con altri fattori, legati all'ambiente o alle abitudini di vita della donna, che in conclusione determina il processo tumorale.

La psichiatria è un altro settore della medicina nel quale è ormai provata l'interazione tra fattori genetici predisponenti e fattori ambientali scatenanti. Un esempio di questa interazione che spazia dal comportamento *normale* di un individuo fino alla comparsa di patologia psichiatrica e che include anche la risposta alla terapia (farmacogenetica) è dato dal sistema di trasmissione serotoninergico e in particolare dal trasportatore della serotonina. La trasmissione serotoninergi-

ca ha influenza su un ampio range di comportamenti (assunzione di cibo, attività riproduttiva, attività motoria, controllo delle emozioni ecc.). Il trasportatore della serotonina (5-HTT) ha un ruolo regolatorio centrale rimuovendo la serotonina rilasciata nel vallo sinaptico. La proteina 5-HTT è codificata da un singolo gene, SLC6A4, la cui attività trascrizionale è modulata da 2 varianti polimorfiche

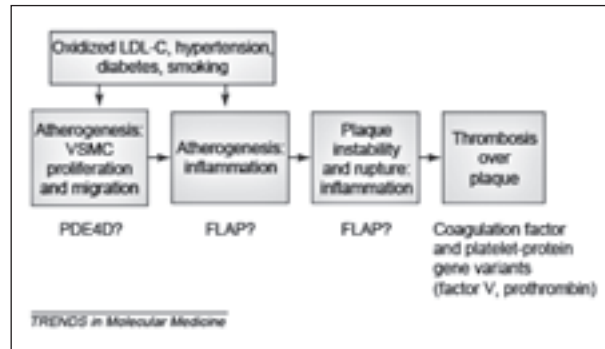
**Tab. 1** - Da Singletary ES, Ann Surg. 237: 474-82, 2003.

Risk factor	Category at risk	Comparison category	Relative risk
Alcohol intake (7)	2 drinks per day	Nondrinker	1.2
Body mass index (11)	80th percentile, age 55 or greater	20th percentile	1.2
Hormone replacement therapy with estrogen and progesterone (23)	Current user for at least 5 years	Never used	1.3
Radiation exposure (25, 26)	Repeated fluoroscopy	No exposure	1.6
	Radiation therapy for Hodgkin's disease	No exposure	5.2
Early menarche (27)	Younger than 12 years	Older than 15 years	1.3
Late menopause (27, 30)	Older than 55 years	Younger than 45	1.2-1.5
Age at first childbirth (25, 29, 31)	Nuliparous or 1st child after 30	1st child before 20	1.7-1.9
Current age (2)	66 or older	Less than 65	5.8
Past history of breast cancer (33, 34)	Invasive breast carcinoma	No history of invasive breast carcinoma	6.8
Other histologic findings (38, 39)	Lobular carcinoma in situ	No abnormality detected	16.4
	Ductal carcinoma in situ	No abnormality detected	17.3
Breast biopsy (40)	Hyperplasia without atypia*	No hyperplasia	1.9
	Hyperplasia with atypia	No hyperplasia	5.3
	Hyperplasia with atypia and positive family history	No hyperplasia, negative family history	11
Cytology (fine-needle aspiration, nipple aspiration fluid (42, 45)	Proliferation without atypia*	No abnormality detected	2.5
	Proliferation with atypia	No abnormality detected	4.9-5
	Proliferation with atypia and positive family history	No abnormality detected	18.1
Family history (479)	1st-degree relative 50 years or older with postmenopausal breast cancer	No 1st- or 2nd-degree with breast cancer	1.8
	1st-degree relative with postmenopausal breast cancer	No 1st or 2nd-degree with breast cancer	3.3
	2nd-degree relative with breast cancer	No 1st- or 2nd-degree relative with breast cancer	1.5
	Two 1st-degree relatives with breast cancer	No 1st- or 2nd-degree relative with breast cancer	3.6
Gemline mutation (46)	Heterozygous for BRCA1, age <40	Not heterozygous for BRCA1, age <40	200°
	Heterozygous for BRCA1, age 60-69	No heterozygous for BRCA1, age 60-69	15°

\*There is controversy over whether pathologic hyperplasia detected in breast biopsy samples is directly equivalent to cytologic hyperplasia detected in samples obtained through FNA or nipple aspiration.

°Begg (52) has suggested that these relative risks are subject to ascertainment bias and may overestimate the true risk associated with germline mutations in BRCA genes.

**Fig. 2** - Da Gulcher JR et al., *TRENDS in Molecular Medicine* 11, 2005.



(lunga e corta) presenti nel promotore del gene che ne determinano una differente espressione e funzione. È ormai dimostrato che la variante corta del promotore è associata a differenze individuali di personalità e a predisposizione alla depressione maggiore che viene però precipitata solo in associazione con eventi stressanti. È inoltre dimostrato che la variante lunga del promotore è associata a una migliore risposta alla terapia antidepressiva con farmaci inibitori del reuptake della serotonina e con una minore incidenza di effetti collaterali alla terapia. Lo studio di questo polimorfismo genetico può quindi dare importanti indicazioni pratiche al medico psichiatra. Sono oggi disponibili test genetici approvati per l'utilizzo in psichiatria.

Infine la malattia cardiovascolare (ictus e infarto del miocardio) rappresenta ancora una delle maggiori cause di morte nei paesi industrializzati nonostante il riconoscimento dei fattori di rischio (ipertensione, ipercolesterolemia, obesità, diabete, vita sedentaria e fumo). Numerosi studi hanno promosso screening di popolazione e il trattamento dei fattori di rischio in molti paesi. Ciò nonostante una discreta percentuale di pazienti non possiede nessuno di questi fattori di rischio acquisiti, suggerendo la presenza di altri fattori non identificati.

L'ictus ischemico rappresenta una sindrome complessa nella quale l'influenza di fattori di rischio modula processi patogenetici diversi con un ruolo differente di fattori genetici predisponenti. Studi genetico-epidemiologici e studi di genetica molecolare hanno dimostrato la presenza di una componente genetica nella predisposizione all'ictus ischemico.

Il ruolo dei fattori genetici va dall'identificazione di geni maggiori di suscettibilità responsabili di forme ereditarie a elevata penetranza fino a geni il cui ruolo concorre, insieme a fattori acquisiti, al processo che conduce alla patologia cerebrovascolare. L'evento cerebrovascolare è quindi il risultato dell'interazione tra il profilo genetico dell'individuo, i meccanismi di risposta infiammatoria e l'attività del sistema della coagulazione del sangue del paziente (Fig. 2). Anche in questo caso la presenza dei fattori genetici predisponenti non è condizione né necessaria né sufficiente a determinare la malattia cerebrovascolare. Il riconoscimento di un profilo genetico di rischio potrà in futuro consentire di limitare la comparsa della malattia e, possibilmente, di orientare il trattamento farmacologico in base a studi di farmacogenetica.

**Bibliografia essenziale**

1. Clavel CR. Progress in epidemiological understanding of gene-environment interactions in major diseases: cancer. *Biologies* 2007; 330: 306-317.
2. Singletary SE. Rating the risk factors for breast cancer. *Ann Surg.* 2003 Apr; 237 (4): 474-82.
3. Belmacher RH and Agam G. Major depressive disorder. *NEJM* 2008; 358: 55-68.
4. Canli T., Lesch KP.. Long story short: the serotonin transporter in emotion regulation and social cognition. *Nat Neurosci.* 2007 Sep; 10 (9): 1103-9.
5. Gulcher JR., Gretarsdottir S., Helgadottir A., Stefansson K. Genes contributing to risk for common forms of stroke. *TRENDS in Molecular Medicine* 2005 May; Vol. 11: 218 No. 5.
6. Worrall BB, Mychaleckyj. PDE4D and stroke. *Stroke*, 2006; 37: 1955.

## **Farmacogenomica: scopi ed applicazioni**

*Angelo Nicolin*

Dipartimento di Farmacologia, Chemioterapia e Tossicologia Medica "E. Trabucchi",  
Università di Milano

### **Farmacogenomica**

La Farmacogenomica è una branca della Farmacologia che si propone di utilizzare gli acidi nucleici quali bersagli molecolari per la prevenzione e la terapia delle malattie dell'uomo. Gli studi a scopo farmacologico iniziano alla fine degli anni 80 e la piena potenzialità applicativa si consolida con la decodificazione del genoma umano. La scoperta degli siRNA ha ulteriormente accresciuto l'attesa di rivoluzionarie terapie.

Gli obiettivi si discostano in parte da quelli della Farmacologia classica per riconoscere nelle alterazioni geniche e genomiche l'origine delle malattie e per la scelta degli acidi nucleici invece delle proteine quale bersaglio molecolare. Quindi essenziale è l'azione a livello del gene piuttosto che a livello della malattia. Si ritiene in questo modo di poter disporre di farmaci molto specifici perché basati sull'interazione complementare delle basi. Non secondaria è l'ambizione di preparare nuovi farmaci con ridotti effetti tossici rispetto agli attuali. L'obiettivo finale è poter guarire malattie come il cancro, infezioni virali e batteriche, riportare alla norma alterate espressioni geniche. La Farmacogenomica contiene inoltre la potenzialità di regolare l'espressione genica anche in senso positivo cioè inducendo aumentata espressione di geni oncosoppressori o potenziando l'espressione di alcuni caratteri fenotipici.

### **Caratteristiche**

Sono oggi allo studio numerose modalità sostanzialmente basate sull'utilizzo di brevi sequenze nucleotidiche preparate per sintesi chimica secondo strategie opportune. Questi approcci consentono di disegnare composti attivi qualora si conosca la sequenza del gene o del messaggero che si intende controllare. Mentre il bersaglio è quasi esclusivamente l'RNA messaggero del citoplasma, il farmaco è costituito da brevi (oligo) sequenze di DNA o RNA a singola o a doppia catena. La preparazione di questi composti richiede l'utilizzo di nucleotidi modificati che conferiscono alla sequenza nucleotidica resistenza alle nucleasi biologiche mantenendo le capacità di ibridizzazione delle sequenze naturali.



Una grave limitazione che ha sinora impedito una applicazione su larga scala è la difficoltà degli oligonucleotidi di superare le membrane cellulari. Il carattere idrofilico, le dimensioni molecolari, l'elevata caratterizzazione elettrica determina la effettiva incapacità di raggiungere i bersagli all'interno della cellula. Sebbene capaci di distribuirsi efficacemente nei diversi organi ad eccezione del SNC, il ridotto uptake cellulare ne impedisce l'efficacia terapeutica. Vari tipi di veicolazione consentono l'utilizzo degli oligonucleotidi in culture cellulari fornendo un contributo straordinario alla comprensione dei meccanismi molecolari delle funzioni cellulari accanto all'identificazione di pathway patologici. Uno sforzo tra i più impegnativi è in corso per consentire agli oligonucleotidi di superare il passaggio delle membrane cellulari per acquisire la piena efficacia terapeutica.

### **Modalità**

Varie modalità sono state sviluppate per ottenere il controllo dell'espressione genica (Fig. 1).

### **Regolazione di Bcl2**

Il nostro laboratorio è impegnato sin dagli anni 80 nel contribuire allo sviluppo della Farmacogenomica affrontando le potenzialità che venivano a prospettarsi con la strategia degli oligonucleotidi antisense. L'iniziale obiettivo era di ottenere evidenza sperimentale circa l'attività biologica e la specificità d'azione. Quale modello si è scelto la leucemia follicolare umana caratterizzata dalla traslocazione cromosomica t(14;18) che determina l'iper-espressione della proteina anti-apoptotica Bcl2. In questo tumore la sequenza nucleotidica nella regione di giunzione dei 2 cromosomi è specifica di ciascun tumore. Qualora efficace, l'ipotesi prevedeva un controllo clone-specifico dell'espressione della proteina Bcl2 ed una conseguente morte cellulare. I risultati hanno dimostrato che effettivamente ciascun oligonucleotide disegnato per interagire con una regione di giunzione era in grado di inibire la crescita di quello specifico clone risultando del tutto inefficace contro altre cellule follicolari o qualunque altra linea cellulare (Fig. 2).

Successivamente queste ricerche hanno portato alla scoperta di un RNA in orientamento antisense ectopicamente espresso nelle cellule di tumore follicolare. La funzione dell'RNA endogeno antisense è quella di inibire, mediante ibridizzazione, con la sequenza complementare del gene di bcl2, una struttura deputata a regolare la cinetica di degradazione dell'RNA messaggero. La regione Adenine-uridine Reach Element, ARE, è contenuta nella regione 3'UTR di molti messaggeri inducibili e ne regola di fatto i livelli di espressione (Fig. 3).

Lo studio di ARE espresso nel messaggero di bcl2 ci ha portato a scoprire l'azione di controllo di tipo feed-back da parte della proteina Bcl2. Infatti la velocità di degradazione del messaggero di bcl2 è regolata da una serie di Adenine-Uridine Binding Proteins (AUBPs) che agiscono in modo non sequenza specifica. La specificità di regolazione è dovuta all'azione della proteina sul proprio messaggero (Fig. 4).

Il controllo dell'espressione di bcl2 è stata ottenuta mediante oligonucleotidii antisenso, brevi sequenza di DNA a singola elica. In cellule normali, mimando l'azione dell'RNA antisenso endogeno di origine tumorale, brevi sequenze sintetiche di RNA hanno determinato l'inibizione della

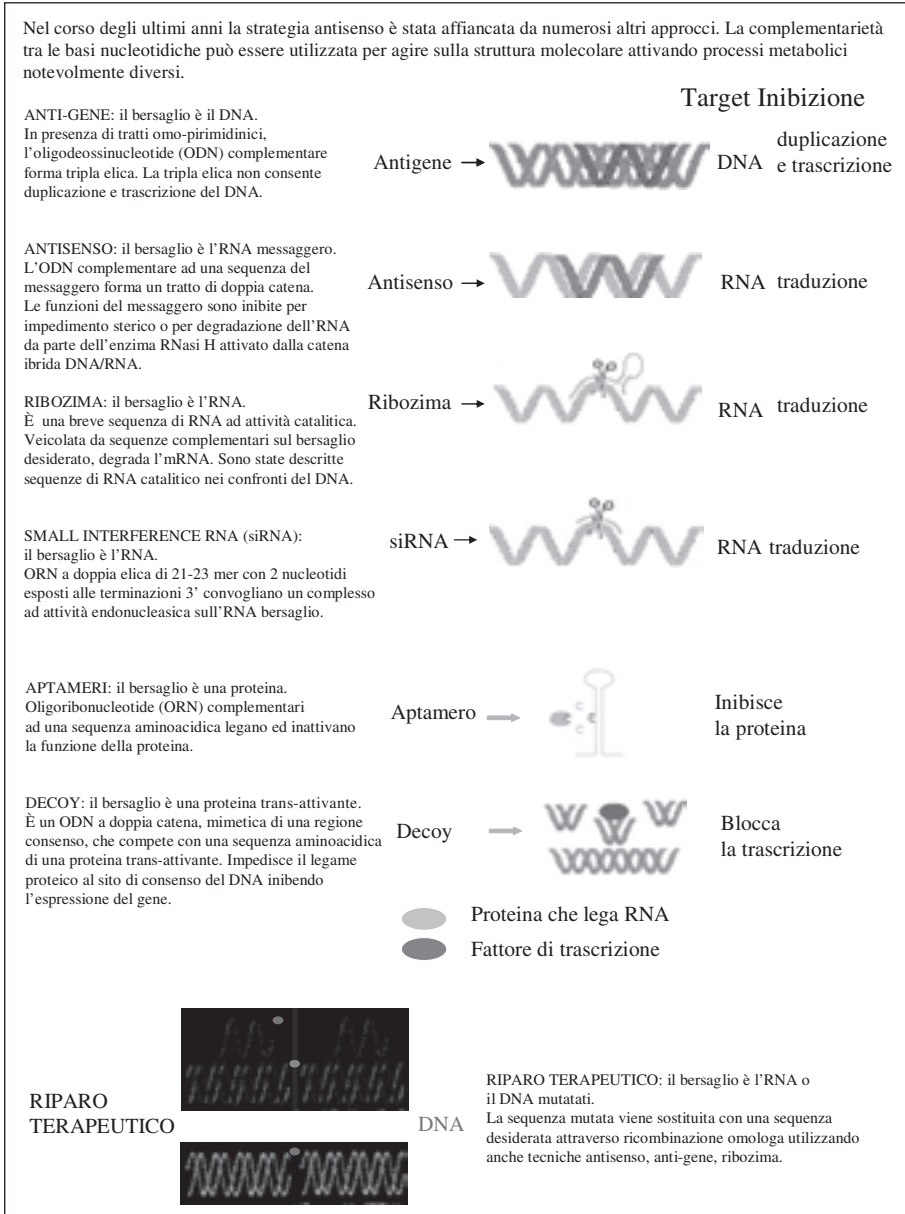


Fig. 1 - Strategie per regolare l'espressione genica.



## Le basi genetiche della variabilità individuale della risposta ai farmaci

*Diego Fornasari*

Dipartimento di Farmacologia, Chemioterapia e Tossicologia Medica "E. Trabucchi",  
Università di Milano

Le reazioni avverse ai farmaci costituiscono la quarta causa di morte nei paesi occidentali, dopo l'infarto del miocardio, i tumori e l'ictus. La maggior parte di queste imprevedibili morti è dovuta a variabilità inter-individuale nella risposta ai farmaci e l'aspetto speculare di questa variabilità è la mancata efficacia di un determinato trattamento farmacologico, con possibile esposizione del paziente ai soli eventi avversi del trattamento stesso senza alcun beneficio terapeutico (1). In passato questa variabilità inter-individuale veniva principalmente spiegata alla luce di fattori correlati alla fisio-patologia del paziente quali l'età, lo stato della funzionalità renale ed epatica o lo stile di vita, con particolare riferimento alla dieta e all'abuso di alcool e fumo. Sebbene tutti questi fattori siano certamente coinvolti e possano in specifiche circostanze giustificare da soli la diversa risposta clinica al medesimo trattamento farmacologico, attualmente si ritiene che la risposta individuale ai farmaci abbia sempre una componente genetica, la quale talvolta è decisamente preponderante sugli altri fattori.

La Farmacogenetica è la disciplina che studia come l'assetto genetico degli individui influenzi l'azione dei farmaci ad essi somministrati, favorendo o determinando lo sviluppo di reazioni avverse o fallimenti terapeutici. Il termine Farmacogenetica è stato coniato da Vogel nel 1959, ma il grande sviluppo di questa disciplina è coinciso con l'avvento delle tecniche del DNA ricombinante e più recentemente, con il completamento del Progetto Genoma Umano.

### Le basi genetiche della variabilità individuale nella risposta ai farmaci

Si ritiene che più del 90% dei geni nell'uomo contenga variazioni di sequenza tra individui, cioè che i geni siano presenti sotto forma di varianti alleliche diverse. Queste variazioni inter-individuali nella sequenza del DNA sono note come polimorfismi, i quali possono essere di diverso tipo, ma i più numerosi sono quelli a singolo nucleotide, che nella letteratura scientifica internazionale vengono identificati con la sigla SNPs (single nucleotide polymorphisms). Come detto, si tratta appunto di variazioni nella sequenza del DNA per cui due individui possono

presentare in una identica posizione dei rispettivi genomi due diversi nucleotidi. Si stima che esistano nel genoma umano circa dieci milioni di SNPs di cui circa 4 milioni sono già stati catalogati (2). Gli SNPs vengono classificati in base alla posizione che essi hanno nel contesto del gene. Si riconoscono pertanto: i) SNPs delle regioni codificanti (cSNP), che sono localizzati negli esoni e la loro presenza può modificare la sequenza aminoacidica di una proteina; ii) SNPs perigenici (pSNP), che interessano le regioni regolatorie (promotori, enhancer), le regioni specificanti le porzioni a monte e valle del mRNA, gli introni, con particolare riferimento alle giunzioni di splicing; la presenza di questi SNP può interferire con i livelli di espressione di una proteina (trascrizione, stabilità e traducibilità del messaggero), ma anche con la sua struttura nel caso in cui essi interferiscano con lo splicing; iii) SNPs random (rSNP) si trovano nelle regioni intergeniche, cioè in quelle regioni che non contengono geni, ma che costituiscono il 98,5% del genoma umano.

Le informazioni ottenute attraverso il Progetto Genoma Umano, estese grazie alla nascita di progetti *ad hoc* per lo studio dei polimorfismi, consentono di ritenere che vi siano in media quattro cSNP per gene; assumendo che il genoma umano contenga 30.000 geni, esisterebbero almeno 120.000 cSNP, i quali potrebbero determinare la presenza di 20-30.000 aminoacidi di differenza tra 2 individui-tipo (3). Non meno importante è poi il ruolo dei pSNP che si stima siano tra i 250.000 e i 500.000. Questi polimorfismi, interessando regioni di DNA che influenzano i livelli di trascrizione, traducibilità e stabilità del mRNA possono condizionare la quantità della relativa proteina presente, influenzando e diversificando profondamente il trascrittoma e quindi il proteoma di un determinato tessuto o organo tra individui diversi.

I cSNPs e pSNPs sono quindi tra i principali responsabili dell'esistenza di forme diverse di un medesimo gene, cioè dell'esistenza di varianti alleliche.

I geni che determinano la risposta ai farmaci possono essere distinti in due grandi classi: i geni codificanti per il bersaglio terapeutico primario, come per esempio recettori, canali ionici o enzimi e i geni codificanti per proteine coinvolte nell'assorbimento, metabolismo ed escrezione del farmaco. Polimorfismi a carico di geni appartenenti a queste due classi possono pertanto determinare modificazioni nell'azione di un farmaco sia di natura farmacocinetica che farmacodinamica provocando la comparsa di effetti avversi o l'assenza di risposta clinica ad un determinato trattamento.

### **Polimorfismi genetici e metabolismo ed escrezione dei farmaci**

Polimorfismi genetici a singolo nucleotide sono stati dimostrati in numerosi geni codificanti per enzimi operanti in ciascuna delle fasi di metabolizzazione dei farmaci con conseguenze funzionali diverse a carico del prodotto genico. In particolare, si è potuto per esempio dimostrare l'esistenza di varianti alleliche completamente inattive di membri della superfamiglia del citocromo P450, quali CYP2D6 CYP2C19 o CYP2A6, in cui i polimorfismi a singolo nucleotide causano l'introduzione di codoni di stop, con la conseguente produzione di proteine

troncate, o determinano sostituzioni aminoacidiche, con la conseguente produzione di proteine enzimaticamente inattive (4). In altri casi, come per gli enzimi di fase II N-acetyltransferasi di tipo 1 e 2 (NAT1 e NAT2), si sono identificate varianti alleliche con ridotta attività enzimatica o con ridotta emivita proteica: in entrambi i casi il fenotipo consisteva in un ridotto metabolismo di alcuni farmaci. È opportuno sottolineare che la gravità dei quadri clinici derivanti dalla somministrazione di un determinato farmaco dipende dall'assetto genotipico dell'individuo, relativamente agli alleli codificanti per lo specifico enzima coinvolto. Infatti, poiché nel caso degli enzimi preposti al metabolismo degli xenobiotici entrambi gli alleli contribuiscono alla produzione del prodotto proteico, la presenza di una variante allelica completamente inattiva in eterozigosi determinerà, in linea di principio, la produzione del 50% di enzima rispetto ad un soggetto con due alleli perfettamente attivi, mentre la presenza di una variante allelica completamente inattiva in omozigosi causerà l'assenza totale dell'enzima stesso. La presenza di varianti alleliche inattive è spesso studiata in relazione alla comparsa di reazioni avverse, poiché il mancato o ridotto metabolismo causano un'abnorme concentrazione di farmaco nel plasma per tempi prolungati, favorendone gli effetti tossici; tuttavia i farmaci non sempre vengono assunti come tali, ma talvolta come pro-farmaci, cioè come molecole inattive che necessitano di essere attivate proprio dagli stessi meccanismi di metabolizzazione degli xenobiotici sopra descritti; quindi un soggetto omozigote per una variante allelica inattiva di un enzima necessario alla conversione di un pro-farmaco in farmaco, andrà incontro ad un fallimento terapeutico, piuttosto che ad un evento avverso.

### **Polimorfismi genetici nei geni codificanti per i bersagli terapeutici primari dell'azione dei farmaci**

La presenza di polimorfismi nei geni codificanti per i bersagli molecolari dell'azione dei farmaci, come recettori, canali ionici, enzimi, pompe e trasportatori, ha un enorme impatto sulla risposta individuale al trattamento farmacologico. Se i polimorfismi a carico dei geni codificanti per gli enzimi del metabolismo dei farmaci sembrano essere principalmente responsabili della comparsa di eventi avversi (ma non sempre, si pensi ai casi descritti dei profarmaci), quelli nei geni codificanti per i target terapeutici sembrano influenzare principalmente, ma non esclusivamente, l'efficacia del trattamento. Per esempio, SNP delle regioni codificanti di recettori a sette domini transmembranari possono determinare, a seconda della localizzazione della sostituzione aminoacidica nel contesto della proteina, modificazioni nell'affinità di agonisti ed antagonisti, ridotto accoppiamento con le G-proteine, e quindi ridotta capacità di transdurre il segnale, o anche aumentata internalizzazione da agonisti e quindi ridotta espressione sulla superficie delle cellule. Esistono attualmente numerosi casi ben caratterizzati di polimorfismi nei bersagli molecolari dei farmaci.

Uno degli ambiti in cui la ricerca farmacogenetica ha prodotto i risultati più interessanti è in medicina cardiovascolare, in cui la relativa disponibilità di più farmaci per il trattamento della medesima patologia, si pensi all'ipertensione arte-

riosa essenziale, ha imposto l'esigenza di fondare la scelta del farmaco e più in generale l'intervento terapeutico su ulteriori elementi di razionalizzazione. Il paziente cardiovascolare è inoltre un paziente cronico in politerapia, in cui le probabilità di reazioni avverse sono aumentate dall'eventualità di interazioni tra farmaci. Per una classe di farmaci fondamentali in medicina cardiovascolare, gli anticoagulanti orali quali il warfarin, la Food and Drug Administration, in una nota del 2007, ha fortemente raccomandato la genotipizzazione dei pazienti, relativamente ai geni codificanti l'enzima metabolizzante e il bersaglio dell'azione del warfarin (VCORC), prima della somministrazione del farmaco, al fine di ottimizzare la terapia con una classe di molecole che è la seconda responsabile di reazioni avverse negli USA.

### **Il futuro della farmacogenetica**

Sebbene molto debba essere ancora fatto prima di un effettivo impiego nella pratica clinica quotidiana della genotipizzazione dei pazienti per individualizzare la terapia, è prevedibile che già nei prossimi 5-10 anni, in alcune aree terapeutiche, le modalità di prescrizione dei farmaci saranno irreversibilmente modificate dalla Farmacogenetica. Le conseguenze di questa nuova modalità di prescrizione sono evidenti: una migliore qualità dell'intervento terapeutico, con una più razionale scelta della dose, un uso più sicuro dei farmaci da parte del Medico, che potrà richiedere la genotipizzazione del paziente in relazione a determinati trattamenti terapeutici così come oggi richiede i comuni esami del sangue, una più razionale e motivata ricerca per sviluppare strategie terapeutiche alternative al fine di curare coloro i quali non possono assumere uno specifico farmaco, un risparmio di denaro dovuto all'eliminazione di terapie inefficaci e di ricoveri a causa degli effetti avversi dei farmaci.

### **Bibliografia**

1. Lazarou J., Pomeranz BH., Corey PN. Incidence of adverse drug reactions in hospitalized patient: a meta-analysis of prospective studies. *JAMA* (1998); 279: 1200-1205.
2. Gabriel SB., Schaffner SF., Nguyen H., et al. The structure of haplotype blocks in the human genome. *Science* 2002; 296: 2225-2229.
3. Cargill M., Altshuler D., Ireland J., et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Gen* 1999; 22: 231-238.
4. Ingelman-Sundberg M., Oskarson M., and McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *TiPS* 1999; 20: 342-349.

## Implicazioni della farmacogenomica nella pratica clinica in oncologia

*Silvia Marsoni*

SENDO Foundation, Milano

Uno dei problemi principali della pratica oncologica quotidiana è la variabilità individuale della risposta alla terapia dei farmaci antitumorali: lo stesso farmaco somministrato alla stessa dose spesso induce in pazienti risposte diverse. Questa variabilità si manifesta sia rispetto all'efficacia che alla tossicità dei farmaci ed oggi sappiamo che è dovuta alla presenza di fattori genetici nei singoli pazienti o a livello di popolazione. L'identità genetica nella specie *Homo sapiens sapiens* è in realtà molto alta - i vari genomi sono identici al 99.9% - ma ciononostante tra ogni coppia di individui genomi e proteomi possono differire rispettivamente per circa 3 milioni di basi nucleotidiche e per circa 100,000 aminoacidi il che lascia ampio spazio alla eterogeneità delle risposte ai farmaci. Due discipline studiano questo fenomeno da angoli di prospettiva diversa benché molti usino i due termini come sinonimi: la farmacogenetica e la farmacogenomica. In realtà la farmacogenetica è una disciplina nata formalmente nella metà del secolo scorso ( il termine è stato coniato da Friedrich Vogel, nel 1959) che studia la variabilità di risposta ad un farmaco dovuta a fattori genetici ereditari, negli individui o a livello di popolazione (è dunque in qualche modo una disciplina pertinente al campo dell'epidemiologia limitata allo studio di un numero definito di geni noti) mentre la farmacogenomica studia la correlazione tra il genoma in toto (ed i suoi prodotti) e la risposta al farmaco per individuare nuovi bersagli terapeutici ed è dunque una disciplina legata al vasto campo della scoperta e dello sviluppo di nuovi farmaci.

In altre parole la farmacogenomica estende la farmacologia alla genetica moderna il che schiude la porta a prospettive molto attraenti dal punto di vista clinico. Ed infatti benché la farmacogenomica sia una disciplina più giovane la sua curva di crescita è in ascesa verticale, soprattutto nell'ultima decade dopo il sequenziamento completo del genoma umano. Questo perché la farmacogenomica sostiene una delle più seducenti aspettative della medicina moderna: la medicina personalizzata. La "mission" della farmacogenomica è infatti quella di delucidare le basi genetiche delle differenze interindividuali di risposta ai farmaci e di usare queste conoscenze per predire la tossicità e/o l'efficacia dei farmaci nel singolo paziente.



La variabilità di risposta ai farmaci dipende da due fattori: cosa fa il nostro corpo al farmaco, (farmacocinetica) e cosa il farmaco fa al nostro corpo (farmacodinamica). Lo sviluppo dei farmaci antitumorali non può quindi prescindere dal porsi delle ben precise domande sin da ben prima la fase di sperimentazione clinica, indirizzando la farmacogenomica da un lato allo studio dei geni che codificano le proteine responsabili della biodisponibilità del farmaco, come gli enzimi del metabolismo ed i transporters, e dall'altro ai geni che codificano per il bersaglio terapeutico del farmaco, primi fra tutti per quanto riguarda l'oncologia gli oncogeni e gli oncosoppressori (ed i loro prodotti).

Questi studi sono fondamentali in campo oncologico perchè l'impatto dei fattori genetici, visto lo strettissimo indice terapeutico di questi farmaci, può accompagnarsi da un lato a tossicità letali, e dall'altro a tassi di risposta sorprendenti, rivelandosi sempre più un fattore fondamentale nel guidare le scelte terapeutiche. Importanti risultati sperimentali si stanno accumulando negli ultimi 10 anni, e sono oggi disponibili molti dati che spiegano come, almeno parzialmente, la variabilità interindividuale negli effetti collaterali di alcuni farmaci antitumorali sia imputabile a SNPs (single nucleotide polymorphism mutazioni in una sequenza genica con una prevalenza di almeno l'1%) degli enzimi del metabolismo.

Nel 2005 la Food and Drug Administration (FDA), l'agenzia regolatoria centrale americana ha mosso il primo passo in questa direzione riconoscendo che nei pazienti portatori di un particolare polimorfismo del gene che codifica per la diphosphate glucuronosyltransferase (UGT), vi era un rischio molto più elevato che nei pazienti non portatori di tossicità associata all'uso di irinotecan. Questo perché i pazienti portatori omozigoti del polimorfismo (UGT1A1) hanno una capacità più bassa della norma di metabolizzare SN38 il metabolita attivo di irinotecan e di conseguenza a parità di dosi sono esposti ad un rischio elevatissimo di grave mielosuppressione. Il "bugiardino" del farmaco venne quindi emendato a prevedere dosi iniziali ridotte del farmaco in simili pazienti.

Contemporaneamente la FDA approvò il test genetico per identificare i portatori del gene UGT1A1\*28 su DNA genomico isolato dal sangue periferico.

Un campo di applicazione ancora più importante, se possibile, della farmacogenomica è quello relativo alle implicazioni terapeutiche secondarie ai polimorfismi nei bersagli dei farmaci. Capire queste implicazioni significa infatti identificare quali pazienti (o sub popolazioni di pazienti) veramente risponderanno al farmaco, evitando quindi di trattare pazienti che non ne trarrebbero beneficio alcuno ma che sicuramente ne sperimenterebbero gli effetti collaterali. I due esempi più eclatanti in campo oncologico sono relativi a tumori ad alta incidenza come il ca del polmone ed il ca della mammella.

Dagli inizi del millennio si sa che alcune mutazioni germline nei geni oncosoppressori BRCA1 e BRCA2 (600185) conferiscono un rischio significativamente aumentato di cancro al seno ed all'ovaio in età precoce. Queste mutazioni sono responsabili rispettivamente del 40% e della quasi totalità dei cancri familiari al seno e dell'ovaio e dal 5 al 20% dei cancri ovarici e mammari in generale. Entrambe questi geni sono degli oncosoppressori e codificano per proteine con funzioni fisiologiche multiple tese a mantenere l'integrità genomica e a preveni-

re la proliferazione incontrollata delle cellule, ed in particolare sono coinvolte nelle funzioni di riparo ai danni del DNA. L'identificazione di queste mutazioni non solo può orientare la terapia chirurgica preventiva di queste pazienti ma le rende particolarmente sensibili ai trattamenti con farmaci inibitori del riparo del DNA come i PARP inibitori di recente introduzione per i quali i trials sono oggi in corso. Mutazioni attivanti Egfr, l'oncogene che codifica il recettore per l'Epithelial Growth Factor sono associate nei pazienti con carcinoma polmonare non a piccole cellule ad un cluster di genotipi particolarmente sensibili agli inibitori della tirosina chinasi recettoriale (erlotinib e gefitinb) e oggi la loro presenza orienta la scelta elettiva al trattamento con questi farmaci nel 10% della popolazione caucasica con probabilità di risposta intorno al 60-80% e durate del DFS intorno ai due anni, mentre altri polimorfismi nella regione del promoter sono stati recentemente associati alle manifestazioni di tossicità cutanea di questi farmaci.

Comprendere perchè la risposta ai farmaci sia così eterogenea nell'ambito dei diversi istotipi tumorali è sicuramente una delle priorità maggiori e oggi vi sono molti altri casi di polimorfismi genetici rilevanti per il trattamento dei tumori solidi (Tab. 1).

Tuttavia i geni "giusti" sono ancora da definire per la maggioranza dei farmaci anti-tumorali e quindi le implicazioni, nella pratica clinica quotidiana, delle conoscenze fin qui raggiunte vanno temperati rispetto sia alle attese dei pazienti, spesso troppo ottimisticamente influenzate dai media, sia alla oggettiva possibilità per il SSN, di offrire su base capillare e con qualità valida, un servizio di diagnostica costoso e ancora per larga parte sperimentale. La mia lettura cercherà di porre le più recenti e rilevante scoperte farmacogenomiche nel loro contesto clinico reale, discutendone i relativi punti di forza e di debolezza, e di illustrare la sfida che il concetto di medicina personalizzata pone allo sviluppo di nuovi farmaci in oncologia. Il caso che forse illustra meglio quale sia oggi il valore per l'oncologo praticante di queste informazioni è il caso della 5Fluorouracile (5FU) discusso di seguito.

Per la 5FU sono stati sviluppati una serie di markers predittivi ed è stata dimostrata, in particolare, un'associazione stretta tra tossicità/risposta ed alcune varianti di 3 geni che codificano rispettivamente per gli enzimi dihydropyrimidine dehydrogenase, methylene tetrahydrofolate reductase e thymidylate synthetase. L'identificazione di pazienti potenzialmente a rischio di sviluppare tossicità letali secondari a polimorfismi di questi geni permetterebbe ovviamente di esclu-

**Tab. 1**

Gene	Polimorfismo	Farmaco	Implicazione clinica
TS	5' regione (28 BP repeat G/C SNP)	5-Fu	Risposta, sopravvivenza, tossicità
DPD	IVS14 + 1G→a	5-FU	Tossicità
OPRT	213G→A	5-FU	Tossicità
UGT1A1	(TA) <sub>6/7</sub> TAAA	Irinotecan	Tossicità
XPD	Lys751Gln	Oxaliplatino	Risposta, Sopravvivenza
GSTP1	Ile105Val	Oxa- e cisplatino	DFS, sopravvivenza
EGFR	Mutazioni Ex 18,19,21	Inibitori tirosina chinasi	Risposta
CCND1	A870G	Cetuximab	Sopravvivenza

dere un trattamento con 5FU (o per lo meno ridurne il dosaggio), orientando il paziente verso forme meno tossiche e più attive di terapia. Tuttavia tutti i markers predittivi sperimentati fin ora, sia su base genotipica che fenotipica, quando valutati prospetticamente nell'ambito di studi controllati disegnati ad hoc hanno dimostrato ampi margini di incertezza circa la loro reale predittività, il che rendeva poco fattibile il loro uso su larga scala in condizioni di pratica clinica ospedaliera. Anche il più esaustivo degli studi recentissimamente pubblicato da Schwab e collaboratori che esamina tutti e tre questi markers (ed altri fattori non genetici) contemporaneamente in uno studio prospettico di larga scala con più di 600 casi trattati in monoterapia con 5FU finisce per concludere che questi geni giocano un ruolo limitato nello spiegare le tossicità relative al farmaco (ancorché sottolineano la presenza di una interazione fra genere maschile e DPYD non precedentemente definita), la cui eterogeneità di comparsa è ancora legata a componenti quali il sesso, le modalità di somministrazione del farmaco, la somministrazione concomitante di acido folinico ed altri fattori. Questo lavoro, per altro assai ben disegnato e condotto, nonostante i risultati apparentemente "negativi" illustra invece a mio parere proprio la necessità dell'approccio "olistico" farmacogenomico. Oggi infatti la possibilità di usare tecniche high-through-put per screenare numeri molto grandi di variabili schiude (almeno) la possibilità di studiare (e capire?) il complesso interplay molecolare di tutti gli elementi genetici, epigenetici e non-genetici che singolarmente sappiamo essere in qualche modo associati alla tossicità del farmaco incluso le disregolazioni nelle vie di cell signaling e/o di risposta al danno al DNA che potrebbero indirettamente orientare la risposta al farmaco.

In conclusione, fino ad oggi tutto il percorso dello sviluppo clinico di un farmaco dalla fase 1 alla sua eventuale commercializzazione è stato fondato su un approccio empirico che assicurava per quanto possibile la tollerabilità del dosaggio determinato durante il percorso di sviluppo nella stragrande maggioranza della popolazione, e determinava in quali istotipi tumorali e in che subpopolazioni di pazienti vi era una più alta probabilità di risposta. Tuttavia inevitabilmente una certa quota di pazienti dimostrava una più o meno grave intolleranza al farmaco e in una grande maggioranza di casi (per lo meno nei tumori solidi) la risposta non era raggiunta. Il potenziale pratico della farmacogenomica sta proprio nell'identificare pazienti apparentemente con la stessa diagnosi ma che in realtà sono geneticamente diversi e la cui diversità influenza sia la risposta al farmaco che la sua tollerabilità. Sia per quanto riguarda la tossicità che per quanto riguarda la sensibilità al farmaco cominciano oggi ad essere disponibili i primi dati di farmacogenomica che permettono al clinico di scegliere in maniera non più empirica, ma guidata, il tipo di terapia da scegliere o da escludere risparmiando tossicità inutili e magnificando le probabilità di un trattamento curativo.

### **Bibliografia essenziale**

1. Desai AA., Innocenti F., Ratain MJ. Pharmacogenomics: road to anticancer therapeutics nirvana? *Oncogene* 22: 6621-6628, 2003.

2. Relling MV., Dervieux T. Pharmacogenetics and cancer therapy. *Nat Rev Cancer* 1: 99-108, 2001.
3. Spitzer E., Abbaszadegan M., Schmidt F., et al. Detection of BRCA1 and BRCA2 mutations in breast cancer families by a comprehensive two-stage screening procedure. *Int J Cancer* 85: 474-481, 2000.
4. Schwab M., Zanger UM., Marx C., et al. Role of genetic and nongenetic factors for fluorouracil treatment-related severe toxicity: A prospective clinical trial. *J Clin Oncol* doi: 10.1200/JCO.2006.10.4182.
5. Maxwell PJ., Longley DB., Latif T., et al. Identification of 5-fluorouracil-inducible target genes using cDNA microarray profiling. *Cancer Res* 63: 4602-4606, 2003.
6. Lynch T., Bell DW., Sordella R, et al. Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib. *NEJM* Volume 350: 2129-2139.



**FARMACOGENOMICA  
E NUOVI FARMACI**



## Struttura delle proteine e scoperta di nuovi farmaci

*Lucio Toma*

Dipartimento di Chimica Organica, Università di Pavia

Storicamente il processo di scoperta di un farmaco è stato affrontato in modo empirico e poco organizzato. Solo a partire dagli anni '70, sulla spinta di discipline emergenti che si sono consolidate negli anni '80 e '90 (come la genomica, la strutturalistica o la chimica combinatoriale) si è arrivati a descrivere un modello più o meno uniforme che trova una generale applicazione.

Si deve però ricordare che tale modello non è assolutamente definitivo o valido in tutti i contesti: con la comparsa di nuove tecnologie e il ridimensionamento di tecnologie un tempo emergenti, il processo subisce dei mutamenti anche sostanziali. Detto questo, si può delineare uno schema generalmente applicato che descrive il moderno processo di Drug Discovery (Fig. 1) (1, 2). Questo schema prevede l'identificazione e la validazione in un contesto terapeutico di un target biologico (comunemente una proteina chiave nella genesi o nella progressione di una malattia). Una volta espresso e purificato, questo target può, da un lato essere sottoposto a screening per trovare una molecola (ad esempio un composto naturale o un prodotto di sintesi) che ne alteri l'attività, dall'altro può essere cristallizzato e la sua struttura tridimensionale determinata (ai raggi X o con l'NMR) in modo da iniziare la progettazione di inibitori basati sulla struttura (structure-based drug design). Identificato un modulatore dell'attività del target ancorché imperfetto (hit) si passa alla sua espansione chimica per generare una molecola intermedia (lead) che dimostri attività in un modello preclinico in vivo della malattia. Tale lead dovrà poi essere ottimizzato e fornire un candidato per la clinica (product candidate). Dopo parecchi anni e un percorso clinico che vede cadere in media oltre il 90% dei candidati, il prodotto può essere registrato e lanciato sul mercato.

Le basi concettuali per la scoperta di nuovi farmaci utilizzando l'approccio structure-based sono state formulate di fatto più di 100 anni fa da Emil Fischer. La sua ipotesi "lock and key", pur con i necessari aggiustamenti, è un motivo costantemente ricorrente nella moderna progettazione di farmaci. Tuttavia, solo negli ultimi decenni, è stato possibile entrare in possesso di una dettagliata descrizione delle "serrature" (locks), cioè dei target di natura proteica (enzimi e recettori), ai quali potenti e selettivi farmaci ("chiavi", keys) si possono adattare, fungendo da sub-



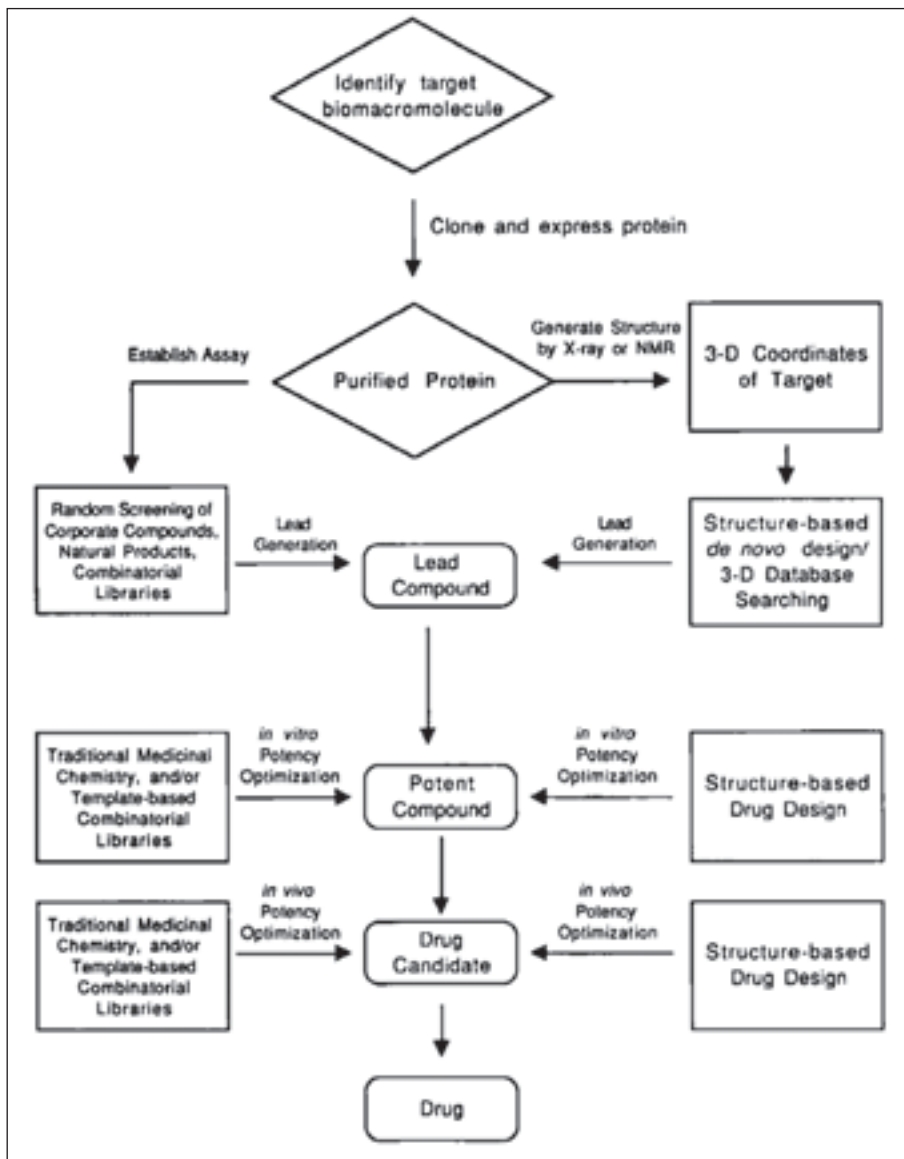


Fig. 1 - Schema del processo di drug discovery and design.

strati o ligandi. Notevoli successi sono stati e sono tuttora ottenuti mediante esplorazione della geometria e delle proprietà strutturali e conformazionali dei soli ligandi, cioè attraverso l'analisi struttura-attività di una serie di questi e la proposta di un modello farmacoforico. Comunque, questo approccio può fornire solo informazioni limitate riguardo ai partner proteici e richiede notevoli sforzi per il raggiungimento di risultati utili attraverso un continuo processo di prova ed errore.

Le strutture tridimensionali ottenute dalla cristallografia ai raggi X e dalla spettroscopia NMR contengono una quantità di informazioni che è di alcuni ordini di grandezza superiore rispetto a quella ottenuta attraverso un modello farmacoforico anche quando è disponibile una ben studiata serie di ligandi. La conoscenza della struttura tridimensionale della proteina rende possibile progettare farmaci basati su un dettagliato modello del sito di binding al quale i potenziali ligandi devono adattarsi.

Il principio fondamentale della progettazione basata sulla struttura del target è che dall'interazione tra proteina e ligando si forma un complesso la cui stabilità deriva dalla complementarità dei due partner interagenti; la formazione del complesso è quindi un processo reversibile la cui costante di equilibrio è intimamente connessa alle proprietà strutturali ed energetiche del complesso che si viene a formare. Le interazioni ligando-proteina possono essere di diversa natura e sono di norma interazioni di non legame quali ad esempio: interazioni elettrostatiche, dipolari, legami ad idrogeno, interazioni idrofobiche.

Le interazioni elettrostatiche sono dovute ai gruppi ionizzabili sulle catene laterali della proteina; infatti, a pH fisiologico i gruppi basici degli amminoacidi arginina, lisina e, in misura minore, istidina sono protonati e di conseguenza si comportano da cationi. Gruppi acidi, quali i gruppi carbossilici delle catene dell'acido aspartico e dell'acido glutammico sono deprotonati e si comportano da anioni. Ligandi e proteine si attraggono quando presentano gruppi con cariche di segno opposto. Le interazioni ioniche possono essere efficaci a distanze superiori di quelle richieste per altri tipi di interazione e possono persistere più a lungo. Una semplice interazione ionica può fornire un contributo energetico fino a 5 kcal/mol che diminuisce con il quadrato della distanza tra le cariche. Le interazioni dipolari (dipolo-dipolo e ione-dipolo) derivano da una non simmetrica distribuzione degli elettroni di un legame; come risultato di una maggiore elettronegatività di atomi quali l'ossigeno, l'azoto, lo zolfo e gli alogeni rispetto al carbonio, i legami tra carbonio ed uno di questi eteroatomi sono dei dipoli con carica parziale negativa sull'eteroatomo e positiva sull'atomo di carbonio. Questi dipoli del substrato o ligando possono essere attratti da altri dipoli (interazione dipolo-dipolo) o da ioni (interazione ione-dipolo) del sito attivo della proteina a condizione che cariche di segno opposto siano opportunamente allineate. Poiché la carica di un dipolo è più piccola di quella di uno ione, un'interazione dipolo-dipolo è più debole di un'interazione ione-dipolo.

I legami ad idrogeno sono un tipo di interazione dipolo-dipolo formata tra un protone di un gruppo X-H, dove X è un atomo elettronegativo, ed altri atomi elettronegativi (Y) contenenti una coppia di elettroni di non legame. I soli legami ad idrogeno significativi avvengono in molecole dove X è N, O o F. X rimuove densità elettronica dall'idrogeno così da renderlo parzialmente positivo e quindi fortemente attratto dagli elettroni di non legame di Y. I legami ad idrogeno sono di notevole importanza per l'attività biologica, sono essenziali per il mantenimento della struttura secondaria delle proteine ( $\alpha$ -eliche, foglietti- $\beta$ ), così come nella stabilizzazione di complessi ligando-proteina. Il contributo energetico di un legame ad idrogeno è nell'intervallo tra 1 e 3 kcal/mole.

Di fondamentale importanza nella stabilizzazione della conformazione nativa di una proteina, così come nella stabilizzazione di un complesso ligando-recettore (o substrato-enzima), sono le interazioni idrofobiche. In presenza di una molecola non-polare o che mostra una regione non-polare, le molecole d'acqua circostanti si riorientano, trovandosi di conseguenza in una situazione energetica meno stabile rispetto a quando sono presenti esse solamente. Quando due gruppi non polari, quali ad esempio un gruppo lipofilo di un substrato o di un ligando ed un gruppo non-polare del sito attivo di una proteina, ciascuno circondato da molecole di acqua, si avvicinano le molecole d'acqua vengono sottratte all'interazione con i gruppi apolari e diventano più disordinate portando ad un aumento di entropia e quindi ad una diminuzione dell'energia libera del complesso. Questa stabilizzazione è nota come interazione idrofobica. È stato stimato che una interazione metilene-metilene possa dare una stabilizzazione di circa 0.5 kcal/mole. Quale esempio di applicazione di queste problematiche si vuole portare lo sviluppo del processo di progettazione di inibitori della sialidasi che vengono usati come farmaci anti-influenza (3-5). L'influenza è una malattia respiratoria acuta che può essere il risultato di una infezione virale. Sono noti diversi tipi di virus influenzale: A, B e C, l'ultimo dei quali comunque non determina la comparsa di un quadro clinico. Sulla superficie del virus sono presenti due principali proteine, l'emoagglutinina e la sialidasi. L'emoagglutinina, una proteina di adesione, è omotrimerica e riconosce e lega i residui terminali di acido sialico (acido N-acetilneuramminico 1) che sono parte delle glicoproteine di superficie della cellula ospite mediando la fusione della capsula virale con la membrana della cellula ospite (Fig. 2). La sialidasi, una glicoidrolasi, è un omotetramero che taglia i residui di acido sialico dalle glicoproteine di superficie della cellula ospite e permette il rilascio della

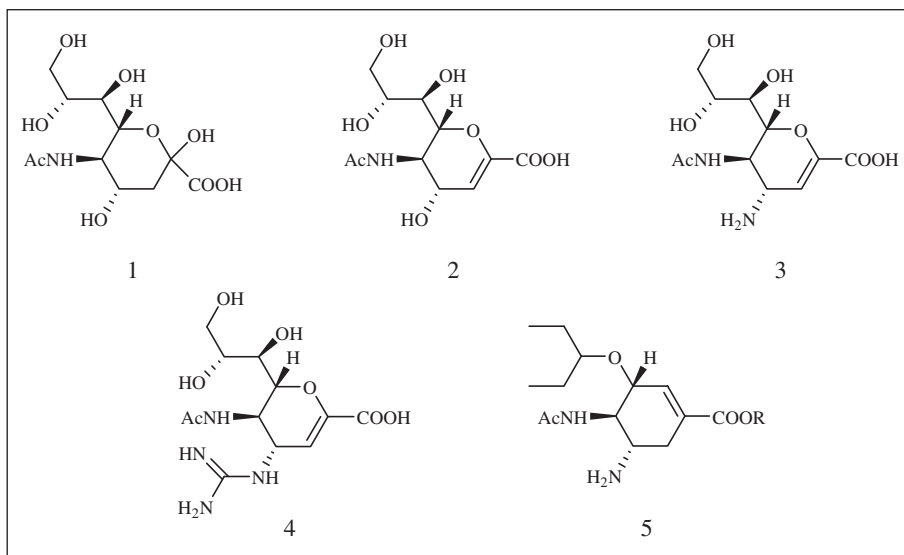
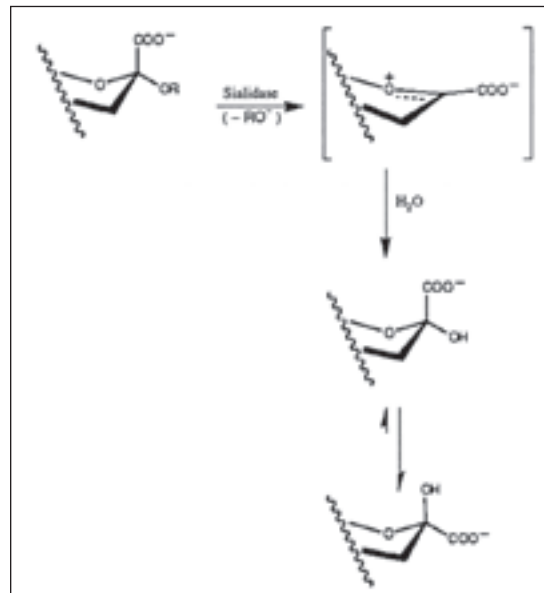


Fig. 2

progenie virale dalle cellule infettate. Molecole capaci di interferire con l'una o con l'altra delle suddette proteine dovrebbero rallentare la normale proliferazione del virus e permettere al sistema immunitario dell'ospite di colpirlo più efficacemente. La struttura cristallina della sialidasi del virus dell'influenza è stata riportata nel 1983. È stato evidenziato che gli amminoacidi che definiscono il sito attivo sono invarianti nelle sialidasi di tutti i ceppi influenzali di tipo A e B e che un certo numero di residui amminoacidici, i quali forniscono il supporto strutturale per i residui del sito attivo, sono conservati. Lo stato di transizione proposto per il meccanismo di azione della sialidasi è mostrato in figura 3. Uno degli inibitori della sialidasi identificati per primi è stato l'acido 2-desossi-2,3-dideidro-D-N-acetilneuramminico 2, che mostra una costante di inibizione verso la proteina a livello micromolare. È stato suggerito che il composto 2 sia un analogo dello stato di transizione, con il doppio legame in posizione 2(3) e l'atomo di O nell'anello che appiattiscono la geometria, così che 2 somigli al catione ossonio intermedio. Tuttavia 2 si è dimostrato non selettivo rispetto alle sialidasi virali, batteriche e di mammifero.

La struttura cristallina della sialidasi virale complessata sia con l'acido sialico 1 che con l'inibitore 2 si dimostrò molto utile in quello che può essere considerato uno dei primi esempi di structure-based rational drug design. Una delle più importanti conclusioni tratte dalla struttura cristallina della sialidasi prima e dopo il soaking con queste due molecole fu che non c'era virtualmente nessun cambiamento nella geometria del sito attivo per effetto della complessazione. Un esame delle interazioni più favorite energeticamente all'interno del sito attivo, effettuato con opportuni probes, mostrò la presenza di una tasca di dimensioni apprezzabili situata accanto al gruppo OH in posizione 4 dell'acido sialico. Il



**Fig. 3** - Meccanismo di idrolisi del residuo terminale di acido sialico catalizzata dalla sialidasi.

primo composto progettato e poi sintetizzato, il derivato amminico 3, mostrò un'inibizione due ordini di grandezza migliore di quella di 2 sulla sialidasi virale. Viste le dimensioni della tasca intorno al 4-OH, si pensò, con il supporto di studi di modellistica molecolare, che questa tasca potesse essere riempita da un gruppo carico positivamente di dimensioni ancora maggiori. Come risultato, fu preparato il derivato guanidinico 4 che mostrò una attività significativa nei confronti della sialidasi dei virus influenzali di tipo A e B. Questo composto è ora commercializzato e noto come zanamivir.

La scoperta dello zanamivir fornì la base per l'individuazione di molti altri farmaci antinfluenzali. Tra questi, viene qui descritto l'oseltamivir 5. La sostituzione della porzione glicerolica con gruppi alchilici fu esplorata sulla base del fatto che il suo scheletro idrofobico ha dei contatti con la proteina anche se i gruppi ossidrilici sul C8 e sul C9 interagiscono con Glu276. La sostituzione fu tentata nella speranza che l'ottimizzazione del carattere idrofobico portasse ad inibitori di maggior lipofilia mantenendo l'attività inibitoria. La non intuitiva sostituzione della porzione glicerolica con una catena alchilica portò al 3-pentil etere come catena laterale ottimizzata. Uno studio cristallografico del complesso con la sialidasi mostrò che Glu276 ruota e forma un ponte salino con Arg224 generando un'area idrofobica nel dominio di binding della porzione glicerolica. Sotto forma di estere etilico, l'oseltamivir fu approvato alla fine del 1999 ed è ora commercializzato come tamiflu.

## Bibliografia

1. Greer J., Erickson J., W., Baldwin JJ., Varney MD. Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* 1994; 37: 1035-1054.
2. Bohacek, RS., McMartin C., Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* 1996; 16: 3-50.
3. Dyason JC., von Itzstein M. Anti-influenza virus drug design: sialidase inhibitors. *Aust. J. Chem.* 2001; 54: 663-670.
4. De Clercq E. Antiviral agents active against influenza A virus. *Nature Rev.* 2006; 5, 1015-1025.
5. von Itzstein M. The war against influenza: discovery and development of sialidase inhibitors. *Nature Rev.* 2007; 6: 967-974.

## La struttura delle monoammina ossidasi umane A e B

*Claudia Binda, Daniele Bonivento, Andrea Mattevi*

Dipartimento di Genetica e Microbiologia, Università di Pavia

Le monoammina ossidasi (MAO A e MAO B) sono enzimi che rivestono un ruolo rilevante nel metabolismo cerebrale di neurotrasmettitori quali l'adrenalina, la dopamina e la serotonina e costituiscono quindi importanti bersagli farmacologici. La conoscenza della struttura tridimensionale delle monoammina ossidasi umane, è tappa promettente per lo sviluppo e per il miglioramento di nuovi farmaci da utilizzarsi nella terapia delle malattie neurologiche e psichiatriche. La struttura tridimensionale delle MAO A e B umane, recentemente ottenute con tecniche biocristallografiche, ha posto le basi per comprendere il meccanismo di riconoscimento e ossidazione dei substrati, aprendo la strada a studi mirati alla progettazione di nuovi inibitori di grande potenziale valore terapeutico.

### Introduzione

I flavoenzimi catalizzano un'enorme varietà di reazioni redox e partecipano a molte vie metaboliche fondamentali per la funzionalità della cellula (Fraaije e Mattevi 2000; Ghisla e Massey, 1989). Le monoammina ossidasi A e B (MAO A e MAO B) sono due flavoenzimi che negli ultimi 40 anni sono stati importanti "target" per le terapie farmacologiche (Weyler et al., 1990; Cesura e Pletsher, 1992). Questi enzimi catalizzano l'ossidazione FAD-dipendente di varie amine aromatiche. La MAO A e la MAO B sono prodotte da due geni distinti e sono formate rispettivamente da 527 e 520 aminoacidi con il 70% di identità di sequenza. Entrambi gli enzimi sono inseriti nella membrana mitocondriale esterna. Le MAO rivestono una grande importanza fisiologica in quanto la MAO A svolge un ruolo decisivo nella degradazione di neurotrasmettitori quali la serotonina e la dopamina mentre la MAO B partecipa al metabolismo di amine xenobiotiche e all'apoptosi delle cellule neuronali. Durante lo sviluppo i livelli di MAO A e MAO B sono regolati da ormoni, dalla dieta e dall'invecchiamento. Aumentati livelli di MAO A e MAO B nel cervello umano correlano con la diminuzione di dopamina e norepinefrina nel sistema nervoso centrale. L'aumento di MAO B nel cervello di pazienti affetti dal morbo di Alzheimer è un fenomeno ben documentato. L'analisi genetica di una famiglia olandese ha dimostrato che una mutazio-

ne nel gene della MAO A associato all'espressione di una forma inattiva dell'enzima è causa di un comportamento aggressivo da parte di individui maschi in condizioni di stress (i geni per le MAO A e MAO B sono localizzati sul cromosoma X; Cases et al., 1995). Risultati analoghi sono stati ottenuti con un topo MAO A-deficiente. L'inibizione delle MAO avviene anche nei fumatori di tabacco (Khalil et al., 2000) e almeno un composto presente nel tabacco (2,3,6-trimetil-1,4-naftochinone) è un potente inibitore reversibile.

I ruoli delle MAO A e MAO B nel metabolismo di neurotrasmettitori e amine biogeniche ha reso questi enzimi importanti target farmacologici. La lista dei farmaci di uso clinico che hanno come bersaglio le MAO è estesa. Attualmente, la principale indicazione degli inibitori della MAO A è la depressione mentale, mentre gli inibitori della MAO B sono principalmente utilizzati nel trattamento del morbo di Parkinson soprattutto come coadiuvanti della terapia basata sulla L-dopa (Cesura e Pletsher, 1992). Inoltre il lazabemide (un inibitore della MAO sviluppato dalla Roche) è nella fase III dei test clinici per il trattamento di pazienti affetti dal morbo di Alzheimer. Purtroppo questi composti hanno anche manifestato significativi effetti collaterali. Questi ultimi variano dal ben noto "cheese effect" (l'ipertensione scatenata dall'ingestione di cibi contenenti elevate concentrazioni di tiramina, quali il formaggio, il vino e la cioccolata) ad altri disordini quali danni epatici probabilmente causati da risposte autoimmuni. Questi effetti collaterali hanno causato un ridotto uso clinico degli inibitori della MAO e hanno stimolato un'ulteriore ricerca di inibitori specifici e reversibili. Una più dettagliata conoscenza delle proprietà strutturali e funzionali di questi due enzimi umani, come descritto in questa domanda di finanziamento, potrà dare un contributo sostanziale al miglioramento di questa classe di farmaci. In questo contesto va messo in evidenza che un altro aspetto delle MAO A e MAO B è costituito dalla loro capacità di metabolizzare xenobiotici aminici. Un esempio è il composto 1-metil-4-fenil-1,2,3,6-tetraidropiridina, un'impurezza trovata nell'eroina sintetica che viene ossidata dalla MAO in un metabolita altamente tossico che induce una sindrome simile al morbo di Parkinson.

### **La struttura tridimensionale della MAO B umana**

Abbiamo completato la determinazione della struttura tridimensionale della monoamina ossidasi B (MAO B) umana alla risoluzione di 1.6 Å (Fig. 1). Fondamentale per l'ottenimento di questo risultato è stata la possibilità di produrre grandi quantità di proteina usando un sistema di espressione con *Pichia pastoris* (Miller e Edmondson, 1999). La MAO B è una proteina di membrana e la scelta dei detergenti è stata cruciale per la cristallizzazione (Binda et al., 2001). Due forme cristalline sono state usate per la determinazione della struttura; una forma triclinica è stata ottenuta utilizzando il detergente laurildimetilaminaossido, mentre una seconda forma ortorombica è stata trovata impiegando il detergente zwittergent 3-12. La struttura cristallografica rivela che la MAO B comprende due domini con una topologia simile a quella già osservata in altri flavoenzimi quali la poliamina ossidasi. Il cofattore FAD è legato in una conformazione este-

**Fig. 1** - La struttura tridimensionale della maob umana è costituita da due catene proteiche identiche tra loro. Alla base di questa struttura dimerica emergono da ciascuna catena i segmenti  $\alpha$ -elicoidali che ancorano la proteina alla membrana. La struttura proteica è visualizzata come un nastro che segue il percorso della catena polipeptidica, evidenziando così i segmenti ad  $\alpha$ -elica e a  $\beta$ -foglietto che formano le cosiddette strutture secondarie della proteina. Il cofattore flavin-adenin-dinucleotide è disegnato in giallo.



sa nel cuore della struttura proteica. In entrambe le forme cristalline, l'enzima è dimerico suggerendo che questa forma oligomerica probabilmente è presente anche nell'ambiente fisiologico di membrana.

Entrambe le isoforme A e B della MAO sono legate alla membrana esterna dei mitocondri. Una caratteristica particolarmente evidente nella struttura è data dalla coda C-terminale (aminoacidi 460-520). Questa regione forma un segmento esteso che si estende sulla superficie proteica ed è poi seguito da un'alfa-elica. Quest'ultima protrude dal lato basale della struttura con il suo asse approssimativamente parallelo all'asse binario del dimero, in un'orientazione che pertanto è adatta ad ancorare la proteina alla membrana esterna mitocondriale. In aggiunta a quest'elica transmembrana, due segmenti apolari formano sulla superficie proteica due zone idrofobiche che sono probabilmente coinvolte nel legame alla membrana. Queste osservazioni sono in accordo con dati di mutagenesi che dimostrano che la troncatura della regione C-terminale non elimina la capacità della MAO di legarsi alla membrana (Mitoma e Ito, 1992). La struttura della MAO B è una delle pochissime strutture note (Bowie, 2000) di una proteina di membrana e presenta un meccanismo di ancoraggio diverso da quello osservato in altre proteine inserite nella membrana con un unico segmento polipeptidico, come la prostaglandina sintasi e la squalene ciclasi. Potrebbe essere simile alla modalità d'inserzione delle reduttasi P450 di mammifero, ma le strutture di queste ultime sono state ottenute dopo la rimozione della coda idrofobica mentre la struttura della MAO B comprende l'intera proteina.

La struttura della MAO B si caratterizza per la presenza di due cavità adiacenti nell'interno della proteina (Fig. 2). La cavità più ampia è posizionata direttamente di fronte alla flavina e forma il sito di legame del substrato occupato dall'ini-





**Fig. 2** - Le cavità d'ingresso e di legame del substrato all'interno di una molecola di MAO sono disegnate mediante superfici semi-trasparenti.

bitore pargilina. Per poter accedere a questa cavità, il substrato deve prima attraversare una cavità d'ingresso che è situata vicino al punto in cui la superficie proteica interseca la superficie della membrana. Questa osservazione solleva la possibilità che la superficie di membrana con il suo carattere anionico possa guidare il substrato aminico (carico positivamente) facilitandone l'ammissione nel sito di legame. Entrambe le cavità sono circondate da residui aromatici e alifatici che creano un ambiente fortemente apolare. In particolare nella cavità del substrato due catene laterali aromatiche (Tyr398 e Tyr435) formano una sorta di "sandwich" aromatico in quanto sono posizionate dirimpetto l'una all'altra con una orientazione perpendicolare alla flavina. La combinazione della flavina e delle catene laterali di Tyr398 and Tyr435 genera una gabbia aromatica che appare essere il sito di riconoscimento del substrato aminico.

### **La struttura tridimensionale della MAO A umana**

Nell'ambito del progetto abbiamo risolto la struttura tridimensionale della MAO A umana utilizzando due forme cristalline (De Colibus et al., 2005). L'analisi cristallografica ha riservato non poche sorprese poiché la MAO A umana differisce significativamente dalla MAO B umana, pur mostrando circa il 70% di identità di sequenza, e dalla MAO A di ratto con la quale ha il 92% di identità di sequenza. La prima caratteristica saliente è che nei cristalli la MAO A umana è presente

come monomero diversamente dalla MAO A di ratto e la MAO B umana che sono dimeriche. Questo diverso stato di oligomerizzazione è stato verificato non solo attraverso esperimenti di diffrazione a raggi X ma anche attraverso esperimenti di ultracentrifugazione. Un'altra differenza riscontrata è stata la dimensione della cavità del sito attivo che, nel caso della MAO A umana, risulta essere meno profonda ma più ampia ( $550 \text{ \AA}^3$ ) di quella della MAO B umana ( $700 \text{ \AA}^3$ ). Inoltre anche la cavità del sito attivo della MAO A di ratto si presenta differente da quella della MAO A umana poiché il loop che la delimita ha una conformazione molto più simile a quella della MAO B umana: questo si traduce in una riduzione del volume del sito attivo ( $450 \text{ \AA}^3$ ). In conclusione la sequenza amminoacidica della MAO A di ratto e della MAO A umana sono quasi identiche ma paradossalmente la conformazione del loop che delimita la cavità del sito attivo dell'enzima murino è molto più simile a quella della MAO B umana che non a quella della MAO A umana. Questa osservazione suggerisce che la conformazione del loop che delimita il sito attivo sia condizionata dalla formazione del dimero: quindi, lo stato di oligomerizzazione potrebbe influenzare la diversa capacità di legare sia il substrato sia l'inibitore.

## Conclusioni

Queste caratteristiche rivelate dalle strutture tridimensionali costituiscono il punto di partenza per studi volti alla caratterizzazione della specificità di substrato, del meccanismo di catalisi, del significato funzionale dell'ancoraggio alla membrana e della possibilità di sfruttare la cavità d'ingresso come sito di legame per farmaci.

## Bibliografia

1. Binda C., Newton-Vinson P., Hubalek F., Edmondson DE., Mattevi A. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat. Struct. Biol.* 2002; 9: 22-26.
2. Bowie JU. Membrane Proteins. Are we destined to repeat history?. *Curr. Opin. Struct. Biol.* 2000; 10: 435-437.
3. Cesura A., Pletsher A. The new generation of monoamine oxidase inhibitors. *Progress in Drug Research* 1992; 38: 171-297.
4. Cases O., Seif I., Grimsby J., Gasper P., Chen K., Muller, U., Aguet M., Babinet C., Shih JC., De Maeyer E. Aggressive Behavior and Altered Amounts of Brain Serotonin and Norepinephrine in Mice Lacking MAO A. *Science* 1995; 268: 1763-1766.
5. De Colibus L., Li M., Binda C., Lustig A., Edmondson DE., Mattevi A. Three-dimensional structure of human monoamine oxidase A (MAO A): relation to the structures of rat MAO A and human MAO B. *Proc. Natl. Acad. Sci.* 2005; 102: 12684-12689.
6. Fraaije MW., Mattevi A. Flavoenzymes: diverse catalysts with recurrent features. *Trends in Biochemical Sciences* 2000; 25: 126-132.

7. Ghisla S., Massey V. Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.* 1989; 181: 1-17.
8. Khalil AA., Steyn S., Catagnoli N. Isolation and Characterization of a Monoamine Oxidase Inhibitor from Tobacco Leaves. *Chem. Res. Toxicol.* 2000; 13: 31-35.
9. Miller JR., Edmondson DE. Structure-Activity Relationships in the Oxidation of para-Substituted Benzylamine Analogues by Recombinant Human Liver Monoamine Oxidase A. *Biochemistry* 1999; 38: 13670-13683.
10. Mitoma J., Ito A. Mitochondrial Targeting Signal of Rat Liver Monoamine Oxidase B is Located at its Carboxy Terminus. *J. Biochem.* 1992; 111: 20-24.
11. Weyler W., Hsu YP., Breakefield X. Biochemistry and Genetics of Monoamine Oxidase. *Pharma. Ther.* 1990; 47: 391-417.

## I passaggi successivi per lo sviluppo di un nuovo farmaco

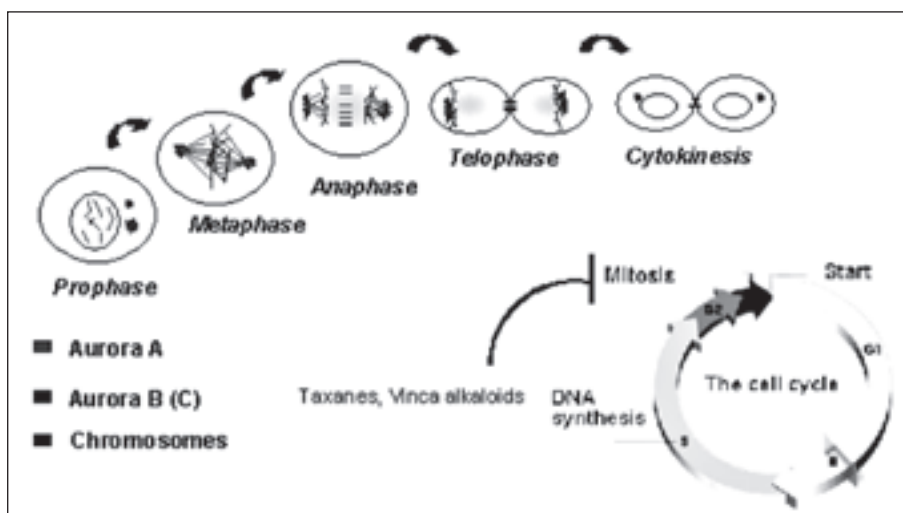
### Identificazione e sviluppo di PHA-739358, un potente inibitore delle Aurora chinasi con profilo di selettività chinastica compatibile con attività antitumorale

*Maurizio Pulici*

Nerviano Medical Sciences, Nerviano, Milano

Allo scopo di illustrare i passaggi che portano allo sviluppo di un nuovo farmaco, non sempre completamente standardizzabili, si è deciso di descrivere qui un “case-history”, quello che ha portato all’identificazione e sviluppo di PHA-739358, un potente inibitore delle Aurora chinasi con profilo di inibizione chinastica compatibile con attività antitumorale.

Le isoforme della chinasi Aurora (A, B, C) costituiscono una famiglia di serina-treonina chinasi coinvolte nel processo della mitosi. Aurora-A, che costituisce il target principale di PHA-739358, fu individuata nel 1995 e il suo legame con i tumori umani divenne evidente nel giro di pochi anni.



**Fig. 1** - Ruolo delle chinasi Aurora nella mitosi.

Aurora-A è coinvolta nella maturazione dei centrosomi, nell'organizzazione del fuso mitotico e nella separazione dei cromosomi. Essa viene espressa nelle prime fasi della mitosi (profase, metafase e anafase) e si trova localizzata oltre che sui centrosomi sulle porzioni polari del fuso mitotico. Aurora-B compare invece più tardivamente, ed è coinvolta nell'aggancio del cinetocoro alle fibre del fuso, nella condensazione dei cromosomi e nella citochinesi, ed infatti va incontro ad una drammatica rilocalizzazione durante la mitosi spostandosi dalla parte equatoriale verso quelle polari della cellula mano a mano che questa procede verso la divisione. Aurora-C è meno studiata ma si pensa replicare il ruolo di Aurora-B.

Le Aurora chinasi sono ampiamente validate come target antitumorali: esse sono infatti regolatori chiave del processo mitotico e sono indispensabili per la divisione cellulare; la loro inibizione si è dimostrata efficace in modelli tumorali animali, e in aggiunta, Aurora-A è un oncogene. Il loro legame con i tumori umani è stato pure ampiamente provato: ad esempio, sia Aurora-A che Aurora-B sono sovraespresse in un'ampia gamma di tumori umani, e spesso il gene di Aurora-A vi si trova amplificato.

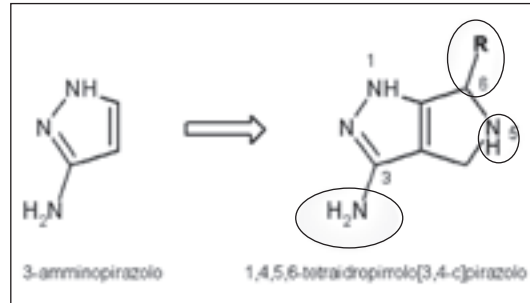
Il nostro progetto di ricerca di piccole molecole organiche come inibitori di Aurora-A venne varato nella prima parte del 2000 e ha condotto, già nel luglio 2004, ad iniziare degli studi clinici con un inibitore selettivo, PHA-739358, oggi in fase II di sviluppo clinico. Tale progetto va contestualizzato nel più ampio ambito della cosiddetta "piattaforma chinasica", sviluppata nella nostra azienda a partire dalla fine degli anni '90 con obiettivo generale l'individuazione e lo sviluppo di inibitori chinasici costituiti da piccole molecole organiche per il trattamento dei tumori umani. Sul versante chimico un tale approccio a piattaforma ha comportato in primo luogo l'arricchimento della collezione proprietaria di composti in potenziali inibitori chinasici. Saggiando questi composti con priorità massima nei progetti con un target chinasico si sarebbe potuto ottenere un incremento del "hit-rate", cioè del numero di composti attivi. Inoltre poiché le serie chimiche sviluppate erano originali cioè non note in letteratura, e i metodi per la loro espansione combinatoriale già validati, all'occorrenza si sarebbe potuto facilmente e rapidamente amplificare il numero di composti relati agli hits ed ottenere maggiori informazioni biologiche ed eventualmente una copertura brevettuale più ampia.

Le prime classi chimiche sviluppate erano basate sull'utilizzo dello scaffold 3-aminopirazolico, che era stato di recente identificato come templatato per inibitori chinasici ATP mimetici. Inserendo il farmacoforo del 3-aminopirazolo in un sistema biciclico come quello del 1,4,5,6-tetraidropirrolo[3,4-c]pirazolo sono stati ottenuti dei nuovi scaffold potenzialmente inibitori di chinasi.

Infatti il modo di legame verso l'enzima ipotizzato per questo tipo di scaffold prevede la formazione di un pattern di legami ad idrogeno con la hinge-region delle chinasi caratteristico di tutti gli inibitori chinasici ATP-mimetici. Inoltre, gli elementi di diversità portati dallo scaffold puntano verso diverse tasche della regione di binding dell'ATP permettendo idealmente una modulazione della potenza e della selettività delle molecole.

Lo scaffold pirrolopirazolico è stato convenientemente sintetizzato per mezzo di una procedura rapida ed efficiente, e la sua decorazione è stata effettuata in

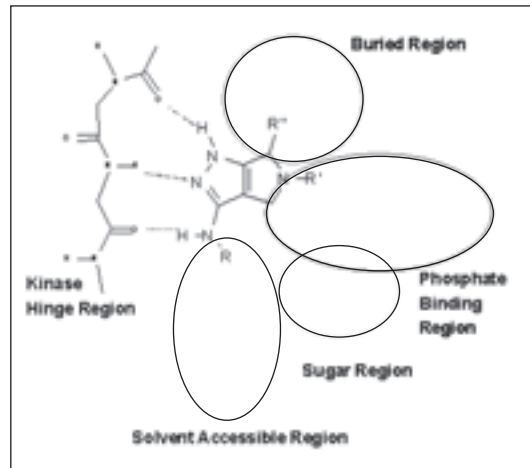
**Fig. 2** - Dal 3-amminopirazolo al 1,4,5,6-tetraidropirrolo[3,4-c]pirazolo.



maniera combinatoriale. Ciò ha comportato lo sviluppo di una metodologia di sintesi su fase solida che prevede il carico dello scaffold su di una resina isocianato e la successiva funzionalizzazione del gruppo amminico in posizione 3 prima e di quello in posizione 5 poi, seguita quindi dal distacco delle molecole dalla resina. Ciò ha permesso la preparazione di una “libreria” di quasi mille composti su 1.232 disegnati. Questi sono stati ottenuti con buona resa il che ne ha permesso, oltre alla totale caratterizzazione biologica *in vitro*, per un buon numero di composti la conservazione come polvere per successivi HTS (High Throughput Screening, Saggi ad alta portata) e la determinazione di un completo profilo di solubilità, permeabilità, stabilità chimica e metabolica.

Un primo set di 350 composti appartenenti a diverse classi chimiche note storicamente o disegnate come potenziali inibitori chinasi, venne testato contro Aurora-A ancor prima che l’intera libreria di pirrolopirazoli fosse completata. Infatti, solo 9 composti appartenenti a quest’ultima serie vi erano inseriti. Ponendo un cut-off di  $IC_{50}$  verso Aurora-A pari a  $0.5 \mu M$ , ben 27 composti (8% circa) vennero trovati attivi: questi appartenevano a 5 diverse classi chimiche, tra di esse i pirrolopirazoli erano rappresentati con un composto attivo (1/9!).

Diverse considerazioni relative alla potenza, solubilità ed espandibilità della clas-



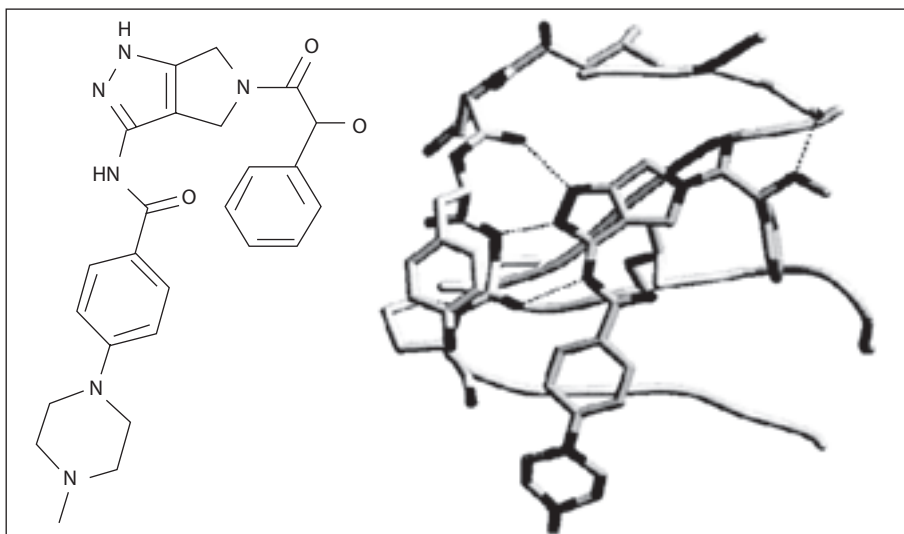
**Fig. 3** - Modo di legame dello scaffold 3-amminopirazolico.

se hanno poi condotto a puntare proprio sui pirrolopirazoli come classe da sviluppare per la ricerca di inibitori potenti e selettivi.

Determinante al fine dello sviluppo di inibitori di Aurora-A è stata la disponibilità della struttura cristallografica dell'enzima. Durante il corso della nostra ricerca sono stati risolti circa 30 complessi tra Aurora-A e rispettivi inibitori, il che ha fornito una valida guida al disegno di nuove molecole, a partire dal primo prototipo originato dalla libreria di pirrolopirazoli fino agli inibitori ottimizzati.

I primi pirrolopirazoli saggiati hanno mostrato una SAR (Structure-Activity Relationship) evidente: la presenza in posizione 3 dello scaffold di ammidi aromatiche, ed in particolare di benzammidi sostituite in posizione para, era associata con una buona potenza nell'inibizione di Aurora-A indipendentemente dal sostituente in posizione 5. Inoltre un trend di selettività nei confronti della chinasasi cdk2 emergeva chiaramente: ancora una volta, a parità di sostituente in posizione 5, le ammidi aromatiche erano più potenti nei confronti di Aurora-A mentre le corrispondenti fenilacetammidi erano più potenti come inibitori di cdk2.

Tra i membri della libreria vennero velocemente identificati dei composti potenti a livello enzimatico e dotati di un'attività antiproliferativa a concentrazioni nell'ordine del micromolare sulla linea tumorale HCT-116. Questi composti erano tutti poco solubili a causa della presenza di un ingombrante gruppo lipofilo. La sostituzione di quest'ultimo con uno più polare e ionizzabile ha portato le molecole ad essere non solo più idrosolubili, ma anche più potenti a livello antiproliferativo. In particolare nella mini-serie così preparata venne identificato PHA-680632, il primo composto-lead identificato tra i pirrolopirazoli inibitori di Aurora. Già ad una concentrazione di 1  $\mu$ M PHA-680632 mostra una buona modulazione dei biomarker di inibizione di Aurora quali il blocco in G2/M del ciclo cellulare, con accumulo di cellule poliploidi, e l'inibizione della fosforila-



**Fig. 4** - Formula di PHA-739358 e struttura del suo complesso con Aurora-A.

zione dell'istone H3 (entrambi connessi con l'inibizione di Aurora-B). La co-cristallizzazione di PHA-680632 con Aurora-A ne ha confermato il modo di legame, e una volta testato in vivo usando diversi modelli animali PHA-680632 ha mostrato una buona attività (misurata come TGI, Tumor Growth Inhibition) e tollerabilità.

L'analisi dei dati cristallografici relativi ai complessi tra Aurora-A e una serie di inibitori tra loro relati, tra cui PHA-680632, ha permesso di fare delle importanti considerazioni che sono culminate nel disegno di nuovi inibitori ancora più potenti dei precedenti. Si è arrivati così a PHA-739358 (Aur-A  $IC_{50}$  =13 nM, Aur-B  $IC_{50}$  =69 nM) il composto che è attualmente in fase di sviluppo clinico.

Anche PHA-739358 si lega ad Aurora-A come atteso, ed è ancora più potente di PHA-680632 sia in vitro che in vivo, mostrando una buona attività in numerosi modelli animali, diversi tra loro: xenotrapianti, tumori indotti da carcinogeni e tumori spontanei in animali transgenici. La sua farmacocinetica è lineare e mediante l'ausilio di un modello di PK/PD sviluppato in casa è possibile predire l'efficacia in dipendenza dallo schema di trattamento.

PHA-739358 è un inibitore di Aurora selettivo nei confronti di molte chinasi, incluse molte di quelle coinvolte nel ciclo cellulare. Tuttavia, come per ogni altro inibitore chinasi ad oggi sviluppato, la sua selettività non è assoluta. In effetti PHA-739358 colpisce altre chinasi con potenza nel basso nanomolare, tra queste ABL, RET TRK-A e FGFR-1. Ognuna di queste è implicata in determinati tumori umani, pertanto tali cross-reattività possono potenzialmente essere sfruttate terapeuticamente.

In particolare l'inibizione di ABL offre delle opportunità importanti: La traslocazione di BCR-ABL (Philadelphia Chromosome) è ritrovata nel 95% dei pazienti affetti da leucemia mieloide cronica (CML). Il primo inibitore chinasi mai entrato in uso clinico, il Gleevec, deve la sua efficacia terapeutica contro CML proprio grazie all'inibizione di ABL. Sfortunatamente però frequentemente i pazienti trattati con Gleevec sviluppano resistenza come conseguenza del selezionarsi di forme mutate di BCR-ABL. Tra queste quella T315I è una delle più comuni: PHA-739358 è in grado di inibirla dando la classica interazione con la hinge-region della chinasi, come dimostrato dai dati cristallografici del complesso tra T315IBCR-ABL e PHA-739358. I nostri studi clinici confermano questo comportamento, infatti pazienti affetti da CML con questa particolare mutazione hanno risposto completamente al trattamento con PHA-739358.

In conclusione, il disegno razionale di inibitori chinasi ATP-mimetici originali basati sullo scaffold 3-aminopirazolico e lo sviluppo di metodologie combinatoriali per la rapida espansione della classe ha consentito di individuare rapidamente dei composti lead come inibitori di Aurora chinasi. Le informazioni strutturali hanno poi consentito di ottimizzarli arrivando in conclusione a PHA-739358, il composto attualmente in fase II di sviluppo clinico.

Questo composto, oltre ad essere un potente inibitore di Aurora, ha delle cross-attività con altre chinasi il che offre l'opportunità di estenderne le applicazioni terapeutiche, come nel caso del trattamento di pazienti resistenti alle correnti terapie di trattamento con inibitori della chinasi ABL.



**Bibliografia essenziale**

1. Fancelli D., Berta D., Bindi S., Cameron A., Cappella P., Carpinelli P., Catana C., Forte B., Giordano P., Giorgini ML., Mantegani S., Marsiglio A., Meroni M., Moll J., Pittala V., Roletto F., Severino D., Soncini C., Storici P., Tonani R., Varasi M., Vulpetti A., Vianello P. Potent and selective Aurora inhibitors identified by the expansion of a Novel scaffold for protein kinase inhibition. *J. Med. Chem.* 2005; 48 (8): 3080-3084.
2. Fancelli D., Moll J., Varasi M., Bravo R., Artico R., Berta D., Bindi S., Cameron A., Candiani I., Cappella P., Carpinelli P., Croci W., Forte B., Giorgini ML., Klapwijk J., Marsiglio A., Pesenti E., Rocchetti M., Roletto F., Severino D., Soncini C., Storici P., Tonani R., Zugnoni P., Vianello P. 1,4,5,6-Tetrahydropyrrolo[3,4-c]pyrazoles: Identification of a Potent Aurora Kinase Inhibitor with a Favorable Antitumor Kinase Inhibition Profile. *J. Med. Chem.* 2006; 49 (24): 7247-7251.
3. Carpinelli P., Ceruti R., Giorgini ML., Cappella P., Gianellini L., Croci V., Degrassi A., Texido G., Rocchetti M., Vianello P., Rusconi L., Storici P., Zugnoni P., Arrigoni C., Soncini C., Alli C., Patton V., Marsiglio A., Ballinari D., Pesenti E., Fancelli D., Moll J. PHA-739358, a potent inhibitor of Aurora kinases with a selective target inhibition profile relevant to cancer. *Mol. Cancer Ther.* (2007), 6 (12, Pt. 1), 3158-3168.

## Computer chemistry and molecular modelling\*

Christian Orrenius

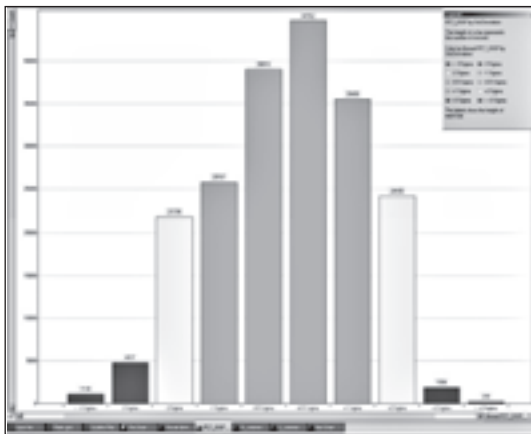
Nerviano Medical Sciences, Nerviano, Milano

Nella giungla dei concetti di Drug Discovery che coinvolgono l'utilizzo di un computer forse soltanto il "modelling molecolare" rimane ben definito con il suo significato classico ancora intatto. Le parole stesse (anche se si tratta di un campo di attività più vasto che mai) esprimono precisamente ed in modo apprezzabile quasi a tutti di che cosa si tratta senza lasciare molto spazio ad interpretazioni. Proprio per questo motivo, l'utilizzo del concetto "Computational Chemistry" (chimica computazionale) è stato preferito ed è venuto a comprendere una serie di tecnologie, operazioni ed attività di "Drug Design" molto diverse fra loro. L'industria farmaceutica è stata responsabile di questa trasformazione ed in parte anche l'università ha seguito questa tendenza. Chiaramente anche la chimica computazionale avrebbe il suo significato classico ben preciso nel campo dell'applicazione e dello sviluppo di algoritmi e di metodi di calcolo *ab initio*, semi-empirici e di meccanica molecolare (modelling molecolare). Nelle unità di chimica computazionale dell'industria farmaceutica di oggi invece si svolgono attività che vanno oltre al modelling molecolare come per esempio: l'analisi statistica di dati di screening ad alta resa ("HTS"); lo sviluppo di modelli statistici di tipo *intelligenza artificiale* per la predizione, in base alla struttura chimica, di proprietà chimico-fisiche e farmacocinetiche; l'enumerazione ed il design di librerie combinatoriali virtuali con la selezione di reagenti per la produzione di esse in base a proprietà cal-



**Fig. 1** - Esempi di concetti ed attività di Drug Design che possono coinvolgere chimici computazionali nell'industria farmaceutica.

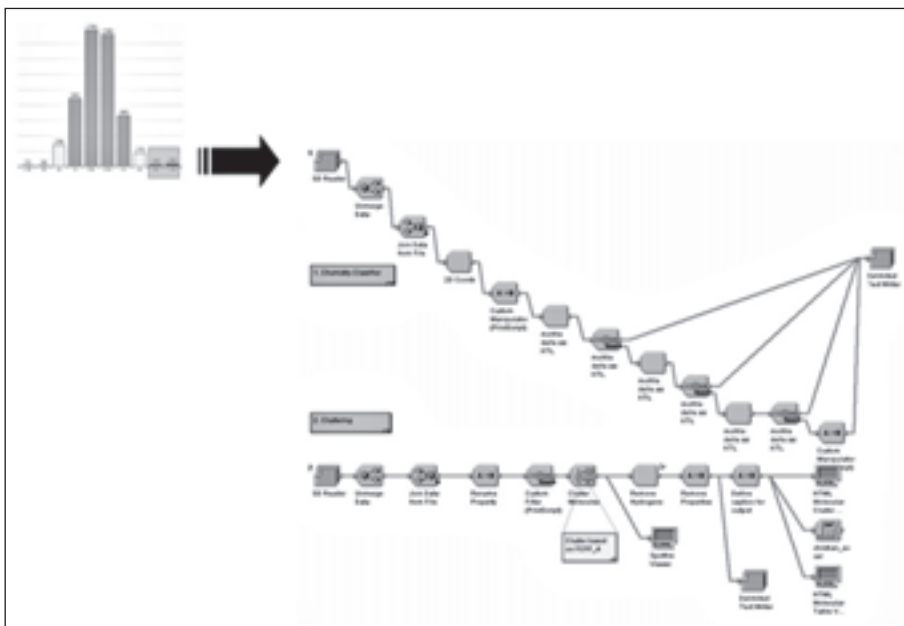
\*Leach, A. R., Molecular modelling. Pearson Education Limited: Harlow, 2001; Vol. 2.



**Fig. 2** - Esempio della distribuzione dei dati di attività biologica in uno screening “universale” (Spotfire, Tibotech USA).

colate sui prodotti finali; la selezione di insiemi di *screening compounds* commerciali da acquistare in base ai risultati ottenuti tramite simulazioni, predizioni e “analoging” (identificazione di molecole in base alla similarità strutturale calcolata); la preparazione di protocolli e metodi per la visualizzazione di dati complessi e numerosi quali per esempio i dati di attività biologica di insiemi di composti su pannelli di enzimi e linee cellulari.

Campagne di screening HTS producono da centinaia di migliaia a milioni di punti dati. Ad ogni dato di attività biologica è associato un composto chimico con la sua struttura. Dopo un primo controllo qualità in base a parametri standard per la valutazione di un saggio biochimico ( $Z'$ , high e low controls, effetti di piastra) i risultati vengono organizzati in categorie di deviazione standard (sigma) partendo dalla media del risultato glo-



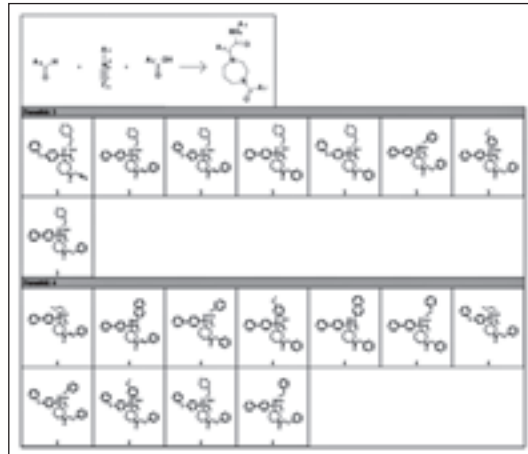
**Fig. 3** - Un tipico protocollo cheminformatico per la classificazione e la clusterizzazione di decine di migliaia di strutture chimiche (Pipeline Pilot, SciTegic, USA).

bale di % di inibizione. Tipicamente viene esaminata la numerosità dei composti risultanti nelle categorie 3 sigma (tre volte il valore di deviazione standard) e quella oltre 3 sigma per valutare il risultato del saggio. In un saggio di tipo “universale” in cui vengono testate librerie chimiche primarie o collezioni generiche di composti da screening, la distribuzione intorno allo zero risulta essere molto simile ad una gaussiana. Nel caso della distribuzione normale fino a 3 sigma si trova il 99.7% dei dati. In

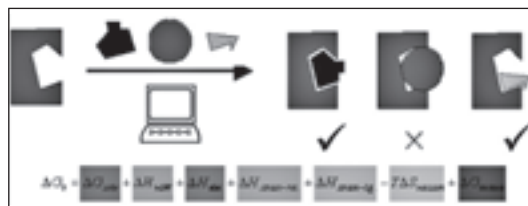
questo modo vengono identificate le molecole da classificare e clusterizzare in base alla loro struttura. Questa analisi permette di razionalizzare i dati di attività e rafforzarli in casi in cui risultino attivi più membri di una famiglia o di un cluster strutturale. La corrispondenza “activity singletons” - “structural singletons” viene accertata anche in categorie di attività più bassa. Là dove questa corrispondenza manca, i composti in questione potrebbero essere scartati in caso di abbondanza di potenziali hits. Alla fine di un saggio universale il numero di composti realmente attivi potrebbe risultare fra 100 e 1.000 composti per milione.

Il concetto introdotto qui, “Computer Chemistry”, si presta bene alle attività legate a “Library Design” ed enumerazione di librerie combinatoriali. In base alla struttura del prodotto virtuale generato vengono calcolati una serie di descrittori molecolari strutturali di cui la correlazione positiva o negativa con la probabilità di essere un prodotto oralmente biodisponibile è stata ipotizzata. Questa procedura assegna delle penalità ai prodotti virtuali che poi vengono visualizzate e quindi considerate nella scelta dei reagenti. Il protocollo NMS considera il peso molecolare, la Polar Surface Area, il numero di donatori e accettori di legami di idrogeno, l’AlogP (coefficiente di ripartizione ottanolo/acqua calcolato) ed il numero di legami rotabili come correlati negativamente e la solubilità in acqua come correlata positivamente alla probabilità che un composto sia oralmente biodisponibile.

Il “Virtual Screening” rimane come sfida principale per affermare la computer chemistry, il molecular modeling ed in parte anche la



**Fig. 4** - “Computer Chemistry” eseguita virtualmente con la generazione del prodotto finale e il suo “score”(Pipeline Pilot, SciTegic).



**Fig. 5** - Rappresentazione schematica di un esperimento di “Virtual Screening” (docking and scoring).

computational chemistry una volta per tutte. Il giorno in cui verrà regolarmente applicato con successo esso sarà il metodo di scelta di tutti i gruppi coinvolti nella ricerca di composti a basso peso molecolare con attività biologica ogni qual volta saranno disponibili dati strutturali o modelli strutturali affidabili dei recettori bersaglio. Molti passaggi rimangono da perfezionare ed alcuni proprio da ridisegnare quasi da capo. Abbiamo già dimostrato la qualità e la resa altissima nel trattare le piccole molecole con la Computer Chemistry disponibile oggi. Nella rappresentazione esaustiva della stereochimica, dei tautomeri e dello spazio conformationale è oggi possibile raggiungere un livello di perfezione che potrebbe garantire una performance più che soddisfacente. Frequentemente vengono anche identificate le orientazioni corrette rispetto a quelle osservate sperimentalmente nelle strutture cristallografiche disponibili. I problemi principali che rimangono sono:

- 1) la “premiatura” tramite scoring function del binding mode corretto una volta trovato;
- 2) una stima sensata del costo entropico speso dal sistema per “ordinarsi” in un determinato modo;
- 3) una stima del ruolo delle molecole d’acqua che in qualche modo partecipano;
- 4) la rappresentazione corretta del recettore in tutte le conformazioni incluse la flessibilità delle catene laterali ed i fenomeni di “induced fit”. Finché alcuni di questi problemi renderanno sempre necessaria la “visual inspection” da parte del ricercatore questo metodo non realizzerà il suo immenso potenziale.







Tipografia Viscontea  
Via Riviera, 39 - 27100 Pavia  
Tel. 0382/526253 r.a. - Fax 0382/423120



