Sinergie tra chimica, genomica e proteomica

di Pierfausto Seneci, Jan Miertus,
Antonio Amoroso e Stanislav Miertus

L'avvento di genomica e proteomica ha rivoluzionato il processo di identificazione e validazione di nuovi target nel processo di ricerca di nuovi farmaci. Sia la chimica di sintesi sia la chimica analitica o computazionale possono beneficiare di opportunità garantite da queste nuove discipline e dalle tecnologie ad esse legate. Questo contributo, attraverso esempi significativi, si propone di avvalorare questa tesi, purché si crei una sinergia fra chimici e biologi.

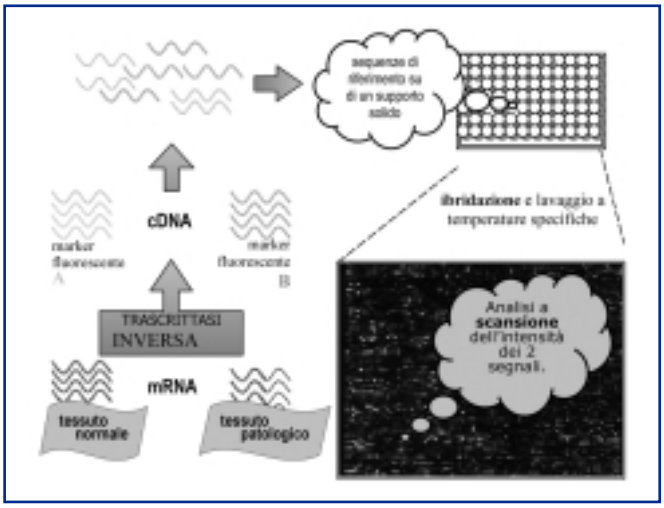

Figura 1 - Gene microarrays: differential gene expression, dal tessuto al segnale di fluorescenza

Molti fra i chimici che lavorano nel settore farmaceutico, soprattutto privato, hanno spesso vissuto con disagio il progressivo aumento d'importanza delle discipline biologiche nel definire le strategie di ricerca all'interno di aziende farmaceutiche. Fasi quali l'identificazione e la validazione di un target biologico, la messa a punto e l'effettuazione di uno screening biologico ad alta capacità (HTS), la caratterizzazione farmacologica e farmacocinetica/farmacodinamica di composti attivi sono ormai naturalmente considerati fra le variabili che, a seconda dei risultati ottenuti in un progetto specifico, ne determinano il successo o l'insuccesso. Inoltre, l'avvento in tempi recenti di genetica, bioinformatica, genomica, proteomica (e delle tecnologie ad esse strettamente legate) ha ulteriormente acuito questo senso di disagio fra i chimici, i quali credono di vedere un futuro sempre più cupo per il campo nel quale sono esperti e si sentono declassati a funzioni di supporto, quasi che la preparazione e l'ottimizzazione strutturale di composti attivi sia considerata una variabile di secondo piano rispetto a quelle sopra citate.

In questa breve prospettiva sul futuro (e sul presente) cercheremo proprio di provare il contrario, e cioè che l'avvento di tali discipline ha già procurato, e procurerà ancor più in futuro eccellenti opportunità scientifiche ed applicative per chimici sintetici ed analitici; esamineremo in maniera molto succinta (ma con dovizia di riferimenti bibliografici) alcuni dei settori in cui la natura e la rilevanza del contributo chimico è stata già dimostrata e sottolineeremo le aree che, a nostro avviso, riserveranno le maggiori occasioni per chimici motivati e dinamici (e per le loro aziende o istituzioni!).

P. Seneci, Nucleotide Analog Design AG, Monaco (Germania); J. Miertus, Sezione di Genetica Medica, Dipartimento di Scienze della Riproduzione e dello Sviluppo, Università di Trieste; A. Amoroso, Servizio di Genetica, IRCCS Burlo Garofolo, Trieste; S. Miertus, International Centre for Science and High Technology of United Nations Industrial Development Organization, Trieste. Stanislav.Miertus@ics.trieste.it

Genomica

Il sequenziamento del genoma umano, in dirittura d'arrivo sia da parte di un consorzio pubblico [1] sia di un'azienda privata [2], avrà ricadute di enorme impatto nel campo della medicina, sia a livello diagnostico sia terapeutico, e di conseguenza sulla qualità della vita dei pazienti che, al limite, potranno beneficiare di "trattamenti personalizzati" basati sull'analisi del loro proprio corredo genetico e sull'applicazione di terapie mirate ad eventuali difetti ereditari.

Dopo la comprensibile euforia legata al completamento del sequenziamento, ci si è resi conto che la sola sequenza nucleotidica del genoma umano non potrà permettere la caratterizzazione delle funzioni di ognuno dei geni in esso contenuti e la determinazione della loro rilevanza in processi patologici. Un ruolo essenziale verrà giocato in futuro quindi dalla cosiddetta "functional genomics" [3, 4], mirata appunto alla decifrazione della sequenza primaria del genoma per derivarne la struttura, la locazione fisica e soprattutto la funzione dei circa 30-35.000 geni in esso presenti. Tecniche bioinformatiche [5, 6], dove la sequenza del genoma umano sarà analizzata con software specifici allo scopo di determinare possibili sequenze geniche nuove e di caratterizzarle strutturalmente per la loro somiglianza con altre sequenze note, saranno fondamentali in questa fase. Lo saranno però anche nuove tecnologie affacciatesi sulla scena negli ultimi anni, oppure tecniche note che stanno sperimentando un ritorno d'interesse: fra le prime i cosiddetti "gene chips", o microarrays [7, 8] (Figura 1), che permettono fra le altre cose di evidenziare relazioni gene-malattia attraverso tecniche di espressione genica differenziale fra tessuti sani e tessuti malati [9, 10]; fra le seconde, l'utilizzo di organismi modello il cui genoma è noto [11, 12] per lo studio comparativo del loro genoma e di quello umano allo scopo di determinare nuovi bersagli terapeutici, od addirittura di elucidare nuovi percorsi biologici.

Se, o meglio, quando le discipline e le tecnologie sopra citate, insieme ad altri approcci che sicuramente ad esse si aggiungeranno, riusciranno a caratterizzare in maniera funzionale il genoma umano saremo in grado di vedere effetti enormi sull'intero processo di drug discovery.
Ciò sarà vero a partire dall'identificazione e validazione di nuovi bersagli per approcci terapeutici, passando da una migliore caratterizzazione tossicologica e metabolica delle molecole da sviluppare (toxicogenomics [13, 14], metabolomics [15] ecc.), per finire alla diagnosi e alla cura "personalizzata" di un vasto spettro di malattie in dipendenza dal corredo genico del singolo individuo (pharmacogenomics [16, 17]).
Fra i vari metodi alternativi a disposizione per una decodificazione funzionale del genoma umano, ci piace citare qui la cosiddetta "chemical genetics" [18-20], i cui principi guida sono riportati in Figura 2.
Supponiamo che una specifica alterazione fenotipica, osservabile in una cellula vivente, sia considerata un modello rappresentativo per uno stato di malattia, e che quindi si ricerchi il gene responsabile per questa alterazione. Se nulla si conosce a riguardo di questo gene, né del suo prodotto genico, un approccio genetico classico (Figura 2, sinistra) prevede:

- l'introduzione di mutazioni casuali (random) nella linea cellulare prescelta (step 1);
- la selezione di cellule mutate che evidenziano l'alterazione fenotipica desiderata (step 2);
- l'identificazione del gene mutato nel corredo genico delle cellule alterate, e l'identificazione del prodotto genico (enzima, recettore ecc.) che sarà il bersaglio molecolare per un progetto di ricerca terapeutica.

Al termine del lavoro il prodotto genico identificato viene sovraespresso in maniera tale da poter mettere a punto un saggio biologico HTS per lo screening di librerie o collezioni di composti allo scopo di identificarne uno o più che interagiscano col bersaglio nella maniera desiderata (inibizione, attivazione ecc.).
Un approccio chimico-genetico (Figura 2, destra) prevede invece:

- lo screening parallelo di una libreria di composti (uno per pozzetto) sulla linea cellulare prescelta (step 1);
- la selezione di cellule mutate che evidenziano l'alterazione fenotipica desiderata (step 2) e l'identificazione del composto chimico responsabile per l'alterazione;
- l'identificazione del gene mutato nel corredo genico delle cellule alterate, e l'identificazione del prodotto genico (enzima, recettore ecc.) che sarà il bersaglio molecolare per un progetto di ricerca terapeutica.

Oltre agli stessi risultati ottenibili attraverso l'approccio classico, si identifica anche la struttura di un composto provocante l'alterazione fenotipica in questione e capace di penetrare all'interno della cellula; questa struttura può essere molto utile per determinare in maniera razionale la strategia chimica del progetto di ricerca terapeutica indirizzato verso l'interazione con il prodotto genico identificato.
L'approccio chimico-genetico rappresenta un'enorme opportunità per i chimici. Innanzitutto, la selezione della diversità chimica da testare (step 1) richiede capacità di scelta fra ciò che è commercialmente disponibile, richiede il disegno razionale di prodotti da sintetizzare e l'assemblaggio finale di una libreria "diversa" (per massimizzare le possibilità di trovare molecole attive su ogni possible target) ma anche dotata di caratteristiche chimico-fisiche adeguate (permeabilità attraverso membrane cellulari, stabilità ecc.). A seguito di tutto ciò sta prendendo piede la cosiddetta "diversity-oriented synthesis" (DOS) [21] che per l'appunto ha come obiettivo quello di identificare piattaforme chimico/tecnologiche in grado di preparare librerie di composti adatte allo screening chimico-genetico in formati e quantità tali da permetterne il saggio in centinaia, se non migliaia di test e la conservazione (senza degradazione) per mesi, se non per anni. Esempi recenti dell'uso di questo approccio [22-27] sono da suggerire al lettore interessato: la maggior parte di essi utilizza un'affidabile piattaforma tecnologica messa a punto dal gruppo di Schreiber [28, 29]. Una variante di questo approccio, chiamata "reverse chemical genetics" [20, 30], permette di identificare con facilità e rapidità il meccanismo di azione molecolare di un composto (spesso di origine naturale) ad azione terapeutica efficace.

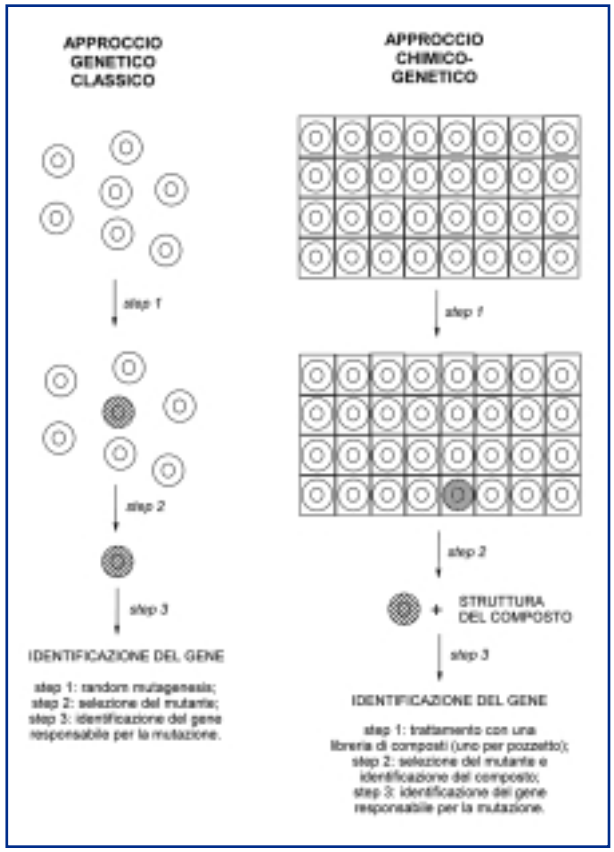

Figura 2 - Genetica e chimica genetica

Proteomica

Insieme all'ormai conosciuto termine genoma, che definisce il corredo genico di un organismo, si è recentemente sviluppato il termine proteoma, che indica in maniera analoga l'insieme delle proteine codificate ed espresse dal genoma di un organismo. La proteomica [31, 32], cioè l'insieme delle tecnologie e delle discipline che studiano il proteoma, mira sia a catalogare il proteoma umano come si è già fatto per il genoma, sia a valutare (ad esempio attraverso lo studio differenziale di tessuti sani e malati) la rilevanza di proteine poco o punto caratterizzate come nuovi bersagli molecolari in diagnostica ed in terapia. Tecniche analitiche quali la spettrometria di massa [33, 34], spesso accoppiata alla cromatografia [35], o l'elettroforesi capillare [36] e perfino di recente la spettroscopia Nmr [37], sono essenziali per la caratterizzazione di campioni proteomici, e la loro applicazione nella proteomica diverrà sempre più comune e rilevante; inutile sottolineare l'impatto che un buon

ricercatore analitico, o un gruppo di lavoro può avere sulla riuscita di progetti di questo tipo. Una discussione a parte merita l'approccio informatico e strutturistico alla caratterizzazione del proteoma umano: qui tecniche di chimica strutturistica e di chemioinformatica, insieme alla bioinformatica, sicuramente giocano un ruolo di primo piano nel generare, ad esempio, nuovi programmi per la ricerca di database proteici allo scopo di identificare il numero e la natura delle famiglie di proteine all'interno del proteoma umano, nel valutare la loro propensione al folding in determinati intorni (e quindi nel prevedere la loro struttura tridimensionale), nell'identificare partner per le cosiddette protein-protein interactions, e così via; come nel caso della "functional genomics", studi analoghi utilizzano sequenze proteiche derivanti da geni mutati ad interesse terapeutico e le confrontano con le proteine non mutate corrispondenti permettendo di valutare l'impatto di mutazioni genotipiche sulla struttura della proteina in questione e sulla sua capacità di espletare le proprie normali funzioni. Il lettore interessato è rinviato ad un ridotto ma significativo numero di lavori molto recenti [38-42] che potranno essergli di aiuto nella comprensione di molti aspetti legati alla cosiddetta "structural proteomic".

Un'analisi proteomica classica permette di determinare la presenza e l'abbondanza di una proteina specifica in un campione; ciò però a volte non è direttamente collegabile all'attività della proteina stessa, se ad esempio essa è un enzima che deve essere attivato (fosforilazione da parte di kinasi ecc.). La cosiddetta "functional proteomics" [43] permette invece di misurare l'attività di una proteina, o di una classe di proteine in tessuti sani o malati attraverso la sua interazione con molecole traccianti (probe). Più in dettaglio, questi probe:

- modificano irreversibilmente il bersaglio attraverso la formazione di legami covalenti;
- agiscono specificamente sul bersaglio, o sulla classe di proteina selezionata;
- sono dotati di un frammento strutturale che facilita e rende automatica la determinazione quantitativa del bersaglio nel suo stato attivo attraverso metodi di detezione comuni (fluorescenza, radioattività ecc.).

Il disegno razionale di questi probe, a partire da strutture note ad attività conosciuta, richiede notevole esperienza chimica per poter creare un ibrido strutturale mantenente l'attività biologica e la specificità necessaria, ma anche contenente un frammento adatto per la quantificazione dell'attività del bersaglio. Nel caso l'informazione strutturale necessaria al disegno razionale dei probe sia assente (o se si vuole ottimizzare una struttura perfettibile), il disegno di librerie combinatoriali improntate all'ottenimento di probe aventi tutte le caratteristiche sopra citate è un obiettivo ancor più ambizioso che richiede competenze chimiche e tecnologiche di grande rilievo. Il lettore è rinviato ad una serie di lavori di alto profilo pubblicati recentemente [44-46] per "toccare con mano" la veridicità di quanto testé affermato.

Conclusioni

Recentemente un noto commentatore scientifico ha riportato il seguente giudizio su una rivista internazionale: "The most likely hurdle on the road to an effective use of chemical genetics is the cultural difficulty of getting biologists and chemists to collaborate" [47]. Come negare che almeno un fondo di verità (e magari un poco più di un fondo…) si nasconda in questa frase? Come, però, negare che è proprio combattendo questa tendenza, più o meno naturale, che i chimici troveranno la possibilità di partecipare come primattori (e non come supporto più o meno qualificato) al moderno processo di scoperta e di sviluppo di nuovi farmaci per il benessere dell'umanità? Noi pensiamo di aver espresso il nostro parere sull'argomento, e lasciamo al lettore il giudizio su quanto da noi articolato in questo breve contributo.

Bibliografia

[1] International Human Genome Sequencing Consortium, *Nature,* 2001, **409,** 860.
[2] J.C. Venter *et al., Science,* 2001, **291,** 1304.
[3] T.R. Hughes *et al., Cell,* 2000, **102,** 109.
[4] S.M. Albelda, D. Sheppard, *Am. J. Respir. Cell Mol. Biol.,* 2000, **23,** 265.
[5] N. Kaminski, *Am. J. Respir. Cell Mol. Biol.,* 2000, **23,** 705.
[6] N.W. Paton *et al., Bioinformatics,* 2000, **16,** 548.
[7] J. Wang, *Nucl. Acids Res.,* 2000, **28,** 3011.
[8] T.A. Weaver, *Drug Disc. Today,* 2000, **5**(suppl.), 36.
[9] S.L. Madden *et al., Drug Disc. Today,* 2000, **5,** 415.
[10] G.G. Lennon, *Drug Disc. Today,* 2000, **5,** 59.
[11] A.C. Levken *et al., Physiol. Genomics,* 2000, **2,** 37.
[12] K. Durick *et al., Genome Res.,* 1999, **9,** 1019.
[13] J.C. Corton, A.J. Stauber, *Toxicol. Sci.,* 2000, **58,** 217.
[14] S.W. Burchiel *et al., Toxicol. Sci.,* 2001, **59,** 193.
[15] D.G. Robertson *et al., Toxicol. Sci.,* 2000, **57,** 326.
[16] A.D. Roses, *Human Mol. Gen.,* 2001, **10,** 2261.
[17] H.L. McLeod, W.E. Evans, *Annu. Rev. Pharmacol. Toxicol.,* 2001, **41,** 101.
[18] S.L. Schreiber, *Bioorg. Med. Chem.,* 1998, **6,** 1127.
[19] C.M. Crews, U. Splittgerber, *Trends Biochem. Sci.,* 1999, **24,** 317.
[20] B.R. Stockwell, *Nat. Rev. Genet.,* 2000, **1,** 116.
[21] S.L. Schreiber, *Science,* 2000, **287,** 1964.
[22] T.U. Mayer *et al., Science,* 1999, **289,** 971.
[23] R.T. Peterson *et al., Proc. Natl. Acad. Sci. USA,* 2000, **97,** 12965.
[24] C.M. Grozinger *et al., J. Biol. Chem.,* 2001, **276,** 38837.
[25] D.R. Spring *et al., J. Am. Chem. Soc.,* 2002, **124,** 1354.
[26] S. Ding *et al., J. Am. Chem. Soc.,* 2002, **124,** 1594.
[27] H. Kubota *et al., Chem. Biol.,* 2002, **9,** 265.
[28] H.E. Blackwell *et al., Chem. Biol.,* 2001, **8,** 1167.
[29] P.A. Clemons *et al., Chem. Biol.,* 2001, **8,** 1183.
[30] B.R. Stockwell *et al., J. Am. Chem. Soc.,* 1999, **121,** 10662.
[31] V. Hatzimanikatis *et al., Biotech. Prog.,* 1999, **15,** 312.
[32] D.F. Hunt, *J. Proteome Res.,* 2002, **1,** 15.
[33] A. Schevchenko *et al., Analyt. Chem.,* 2000, **72,** 2132.
[34] R. Aebersold, D.R. Goodlett, *Chem. Rev.,* 2001, **101,** 269.
[35] Y. Shen *et al., Analyt. Chem.,* 2001, **73,** 3011.
[36] S. Locke, D. Figeys, *Analyt. Chem.,* 2000, **72,** 2684.
[37] A. Yee *et al., Proc. Natl. Acad. Sci. USA,* 2002, **99,** 1825.
[38] S. Sunyaev *et al., Curr. Opin. Struct. Biol.,* 2001, **11,** 125.
[39] Y.L Wolf *et al., Genome Res.,* 1999, **9,** 17.
[40] L. Lo Conte *et al., Nucleic Acids Res.,* 2000, **28,** 257.
[41] S. Govindarajan *et al., Proteins,* 1999, **35,** 408.
[42] Z. Wang, J. Moult, *Hum. Mutat.,* 2001, **17,** 263.
[43] Y. Liu *et al., Proc. Natl. Acad. Sci. USA,* 1999, **96,** 14694.
[44] D. Greenbaum *et al., Chem. Biol.,* 2000, **7,** 569.
[45] V.M. Hoang *et al., Biochemistry,* 2001, **40,** 4077.
[46] N. Winssinger *et al., Angew. Chem. Int. Ed.,* 2001, **40,** 3152.
[47] T. Gura, *Nature,* 2000, **407,** 282.