

Ricerca in banca dati di potenziali geni bersaglio di fattori trascrizionali

G.Lavorgna e E.Boncinelli

DIBIT-HSR - Via Olgettina, 58 - Milano

Uno dei problemi fondamentali della ricerca biologica e' la comprensione dei meccanismi fondamentali che permettono alle cellule differenziate, sebbene totipotenti, di poter selettivamente utilizzare solo una piccola e peculiare porzione del loro enorme potenziale genetico. Infatti, ognuna di esse e' caratterizzata da una particolare configurazione di geni attivi ed inattivi, che subiscono una stretta regolazione della loro espressione sia spaziale che temporale. Per es., le cellule della mano sono tali in quanto esprimono un 'set' di geni in larga misura differente da quello espresso da altri tipi cellulari, come, ad es., le cellule del cervello. Analogamente, le cellule neoplastiche presentano un caratteristico pannello d'espressione genica, che e' alterato rispetto alle cellule sane. Questo controllo dell'attivita' trascrizionale e' in larga misura affidato a delle proteine, i fattori trascrizionali (FT), che controllano, spesso in combinazione con altri FT, l'attivita' di svariate centinaia di geni. I FT esercitano la loro azione legandosi specificamente a piccole sequenze di DNA (i cosiddetti siti di legame), localizzate nelle sequenze regolatrici dei loro geni bersaglio, la cui attivita' viene, di conseguenza, diminuita o intensificata. Un metodo per poter analizzare il meccanismo di azione di un FT consiste quindi nella identificazione dei suoi geni-bersaglio. Purtroppo, questa procedura si rileva spesso complessa e costosa da un punto di vista sperimentale e non e' sempre coronata da successo. Un attraente approccio alternativo al problema consiste nel cercare i potenziali siti di legame dei FT direttamente negli enormi databases di sequenze di DNA accumulatisi nel corso degli anni (1). La natura elusiva e degenerata dei siti di legame rende comunque quest'approccio poco attuabile. Una ricerca nei databases per queste piccole sequenze porta, infatti, a dei risultati in cui la predominanza del 'rumore di fondo' (cioe' l'elevato numero di siti non biologicamente rilevanti) renderebbe virtualmente impossibile la formulazione di ipotesi testabili poi sperimentalmente. Abbiamo quindi sviluppato un software, [TargetFinder](#), che fosse in grado di diminuire notevolmente questo rumore di fondo, preservando allo stesso tempo i reali geni-bersaglio individuati nella ricerca in banca dati (2). L'euristica utilizzata dal programma si basa sull'incorporazione nella ricerca di semplici considerazioni che tengono conto del contesto biologico ove la ricerca stessa viene effettuata. Sara' illustrato come TargetFinder e' stato utilizzato dal nostro laboratorio per la ricerca di potenziali geni bersaglio del fattore trascrizionale Otx2 murino (3).

Referenze:

- 1) Lavorgna G, Boncinelli E, Wagner A, Werner T Detection of potential target genes in silico? Trends Genet 1998 Sep;14(9):375-6.
- 2) Lavorgna G, Guffanti A, Borsani G, Ballabio A, Boncinelli E TargetFinder: searching annotated sequence databases for target genes of transcription factors. Bioinformatics 1999 Feb;15(2):172-3.
- 3) Kelley C, Lavorgna G, Boncinelli, E, Mellon, P, Submitted.