



Metodi funzionali per valutare la conservazione evolutiva delle sequenze regolatrici

Relatore: Prof. Paolo Provero

Candidato: Stefano Gilotto

Universita' degli studi di Torino

Dipartimento di Biotecnologie Molecolari e Scienze per la Salute

Il miglioramento delle tecniche di analisi ha portato alla luce una grande mole di dati. Questo ha posto la necessita' di prevedere quali sequenze siano attive, e quale sia la loro funzione. La conservazione evolutiva delle sequenze e' uno degli elementi su cui ci si basa maggiormente per applicare queste previsioni.

Nonostante esistano numerose evidenze a favore dell'uso della conservazione, alcuni studi hanno individuato sequenze la cui funzionalità è conservata, ma non la loro sequenza.

Sono presentate delle tecniche per poter prevedere queste sequenze attraverso analisi di carattere funzionale.

Fisher et al.

Ipotesi: la funzione regolatrice di regioni non-coding puo' essere conservata, in assenza di conservazione di sequenza.

Oggetto di studio e' il *locus* del gene *Ret*, i cui esoni sono ben conservati, al contrario delle regioni limitrofe. Si confronta l'espressione delle regioni regolatrici tra i teleosti (Zebrafish) e i mammiferi (uomo).

Le regione regolatrici sono definite attraverso il confronto di sequenze tra specie evolutivamente vicine.

- ZCS-Zebrafish Conserved Sequences:
10 regioni conservate tra zebrafish e pesce palla
- HCS-Human Conserved Sequences:
13 regioni conservate tra uomo e alcuni mammiferi

Le regioni regolatrici sono espresse in embrioni di zebrafish, attraverso costrutti transgenici con geni reporter.

- 9 su 10 ZCS hanno un'espressione simile al gene *Ret*
- 11 su 13 HCS hanno un'espressione simile al gene *Ret*, anche in tessuti non presenti nei mammiferi o anatomicamente diversi.

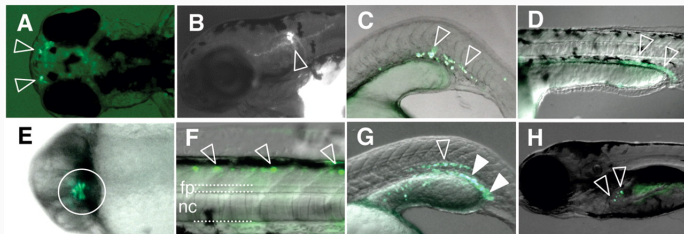


Figure 1: Espressione delle ZCS e HCS

Le regioni regolatrici umane sono quindi funzionalmente analoghe a quelle di zebrafish.

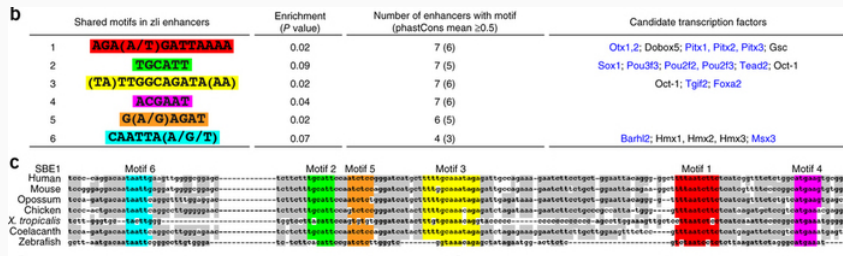
Yao et al.

Si possono individuare nuovi enhancers, basandosi su diverse caratteristiche funzionali legate alla loro espressione.

In particolare si ricercano regioni che regolino lo sviluppo della nella zli-*zona limitans intrathalamica*, prendendo come punto di partenza il gene *Ssh* e il relativo enhancer SBE1.

Si cercano enhancers espressi in tessuti e momenti simili a SBE1.
Vengono individuati 52 possibili enhancers e ne sono selezionati 7 di particolare rilevanza.
Per comprendere quali siano le caratteristiche che li accomunino si cercano motifs e TFBS su questi.

Vengono individuati 6 motivi, che presentano sequenze capaci di potersi legare a TF noti.



I motifs sono soggetti ad analisi per confermare la loro funzione di TFBS, tutti con esito positivo

- saggio *in vitro* con reporter luciferasi per enhancers con motifs deleti
- ChIPSeq sui motifs e i TF
- saggio *in vivo* con reporter LacZ per motifs mutati

La delezione di SBE1 non elimina l'espressione di *Ssh* nella zli, suggerendo la presenza di un altro enhancer.

Basandosi sulle modificazioni istoniche arricchite in SBE1, viene individuato **SBE5**.

Tutte le verifiche relative ai TFBS applicate a SBE1 danno esito positivo per SBE5. La delezione di SBE1 e SBE5 **silenzia** *Ssh* nella zli.

Vengono usati i motifs individuati come base per una ricerca in *S.kowalevskii*.

skSBE1: omologia di sequenza presente solo a livello dei 6 motifs
Costrutti transgenici di skSBE1 in embrioni di topo hanno un pattern di espressione molto simile a mmSBE1. Lo stesso vale per mmSBE1 e mmSBE5 in *S.kowalevskii*.

Blow et al. (2010)

Il numero di enhancers attivi nello sviluppo del embrionale del cuore e' molto basso rispetto agli altri tessuti.

ChipSeq con p300, coattivatore trascrizionale, espresso quasi ovunque nell'embrione di topo.

- 3597 regioni nel cuore
- 2759, 2786 e 3839 regioni nel prosencefalo, mesencefalo e negli arti

Conservazione delle sequenze:

- 84% delle regione predette nel tessuto cardiaco non sovrappongono quelle degli altri tessuti
- 6% delle regioni predette nel tessuto cardiaco si sovrappongono a regioni 'ultra-conserved' in PhastCons. In prosencefalo, mesencefalo e negli arti si ha il 44%, 39% e 30% rispettivamente.
- le regioni del prosencefalo sono 7 volte piu' presenti tra quelle conservate tra mammiferi e pesci

Test in vivo di 130 possibili enhancers predetti in cuore, usando un saggio transgenico in embrioni di topo:

- 81 su 130 sono enhancers attivi solo nel cuore durante lo sviluppo
- arricchiti 13 volte nelle regioni di 10kb a monte dei geni definiti 'heart-development' in Gene Ontology.
- arricchiti 14 volte delle regioni di 10kb a monte di 1000 geni espressi durante lo sviluppo embrionale del cuore.

Conclusione

Conclusione

La conservazione evolutiva delle sequenze rimane una grande risorsa per la predizione della funzionalità delle sequenze. Ad oggi sono state osservate poche eccezioni come quelle presentate, ma questo può essere dovuto al fatto che non le stiamo ricercando.

Ringraziamenti

- Prof. Paolo Provero
- Elena Grassi e il team dell'Unità Computazionale di Bioinformatica
- Gli altri