



# Metodi funzionali per valutare la conservazione evolutiva delle sequenze regolatrici

---

Relatore: Prof. Paolo Provero

Candidato: Stefano Gilotto

Università degli Studi di Torino

Dipartimento di Biotecnologie Molecolari e Scienze per la Salute

Miglioramento delle tecniche di analisi:

- grande mole di dati genomici

Miglioramento delle tecniche di analisi:

- grande mole di dati genomici
- necessità di poter prevedere la funzione delle sequenze

Miglioramento delle tecniche di analisi:

- grande mole di dati genomici
- necessità di poter prevedere la funzione delle sequenze
- conservazione evolutiva delle sequenze

- numerose evidenze a favore dell'utilizzo di modelli basati sulla conservazione

- numerose evidenze a favore dell'utilizzo di modelli basati sulla conservazione
- esistono sequenze la cui funzione è conservata, ma non la loro sequenza

- numerose evidenze a favore dell'utilizzo di modelli basati sulla conservazione
- esistono sequenze la cui funzione è conservata, ma non la loro sequenza

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

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

**Oggetto di studio:** il *locus* del gene *Ret*, i cui esoni sono ben conservati, al contrario delle regioni limitrofe.

**Metodo:** Si confronta l'espressione delle regioni regolatrici tra i teleosti (Zebrafish) e i mammiferi (Uomo).

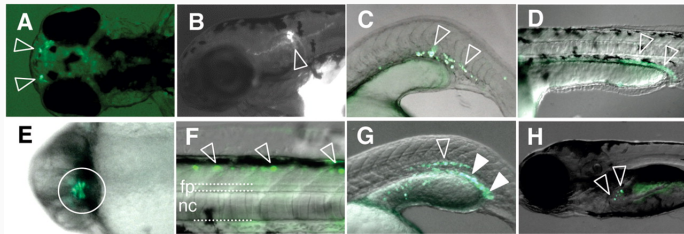


Le regione regolatrici sono definite attraverso il confronto di sequenze tra specie evolutivamente vicine.

- 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.

**Ipotesi:** il controllo dello sviluppo della *zli-zona limitans intrathalamica* precede la comparsa dei vertebrati.

**Oggetto di studio:** il sistema di controllo della *zli* e in particolare il gene *Ssh* e il relativo enhancer [SBE1](#).

**Metodo:** ricerca di nuovi enhancers attraverso caratteristiche funzionali

Si cercano enhancers espressi in tessuti e fasi dello sviluppo embrionale simili a SBE1.

- vengono individuati 52 enhancers

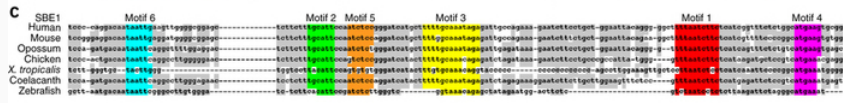
Si cercano enhancers espressi in tessuti e fasi dello sviluppo embrionale simili a SBE1.

- vengono individuati 52 enhancers
- si cercano motifs e TFBS su questi

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

**b**

	Shared motifs in zli enhancers	Enrichment (P value)	Number of enhancers with motif (phastCons mean $\geq 0.5$ )	Candidate transcription factors
1	AGA(A/T)GATAAAA	0.02	7 (6)	Otx1.2; Dobox5; Pitx1, Pitx2, Pitx3; Gsc
2	TGCATT	0.09	7 (5)	Sox1; Pou3f3; Pou2f2, Pou2f3; Tead2; Oct-1
3	(TA)TTGGCAGATA(AA)	0.02	7 (6)	Oct-1; Tgfr2; Foxa2
4	ACGAAT	0.04	7 (6)	
5	G(A/G)AGAT	0.02	6 (5)	
6	CAATTA(A/G/T)	0.07	4 (3)	Barhl2; Hmx1, Hmx2, Hmx3; Msx3



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.

- utilizzando le modificazioni istoniche arricchite in SBE1 è individuato [SBE5](#)

La delezione di SBE1 non elimina l'espressione di *Ssh* nella zli.

- utilizzando le modificazioni istoniche arricchite in SBE1 è individuato SBE5
- presenta gli stessi motivi di SBE1, ma non ci sono altre sequenze in comune

La delezione di SBE1 non elimina l'espressione di *Ssh* nella zli.

- utilizzando le modificazioni istoniche arricchite in SBE1 è individuato [SBE5](#)
- presenta gli stessi motivi di SBE1, ma non ci sono altre sequenze in comune
- le analisi dei TFBS operate su SBE5 danno esito positivo

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

- **skSBE1**: omologia di sequenza presente solo a livello dei 6 motifs

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

- **skSBE1**: omologia di sequenza presente solo a livello dei 6 motifs
- **skSBE1** in embrioni di topo presenta il pattern di espressione di **mmSBE1**

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

- **skSBE1**: omologia di sequenza presente solo a livello dei 6 motifs
- **skSBE1** in embrioni di topo presenta il pattern di espressione di **mmSBE1**
- **mmSBE1/5** in **embrioni** di *S.kowalevskii* presentano il pattern di espressione di **skSBE1**

**Ipotesi:** l'utilizzo della conservazione ha individuato pochi enhancers legati allo sviluppo cardiaco. Quindi questi potrebbero non essere conservati.

**Oggetto di studio:** sviluppo del cuore nell'embrione di topo.

**Metodo:** ChipSeq con p300, coattivatore trascrizionale.

- 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



### 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.

### 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 più 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

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.

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.

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 una conseguenza del grande utilizzo di modelli basati sulla conservazione.

# Ringraziamenti

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