IDENTIFICATION OF RNA TARGETS FOR THE
REGULATORY RNA-BINDING PROTEIN CELF1
by
CLIP-SoliD and RNAseq-SoliD
GENE

Transcription

Gene

pre-mRNA

Gene

mRNA_1

Gene

mRNA_2

Gene

mRNA_3

Gene

Protein_1

Gene

Protein_2

Gene

Protein_3

Gene

DEGRADATION
REGULATORY FACTORS:
- RNABPs (~600 annotated)
- small regulatory RNAs
CELF PROTEINS

CELF1
- CELF1_5
- CELF1_4
- CELF1_3
- CELF1_2
- CELF1_1
  - xCELF1-1A
  - xCELF1-1B
  - xCELF-1

CELF2
- CELF2_3
  - xCELF-2
  - CELF2_4
  - CELF2_2
  - CELF2_1
  - xCELF-2
CELF1
MOLECULAR FUNCTIONS :

- alternative splicing regulation

- mRNA deadenylation targeting factor
  
  translational silencing  mRNA degradation
CELF1 FUNCTIONS IN Xenopus

(Gautier-Courteille et al., Dev. 2004)
CELF1 FUNCTIONS IN MICE

growth retardation phenotype

spermatogenesis defect

(Kress et al., MCB. 2007)
Functional Inactivation

- Primary Molecular defects
  - Secondary (N) Molecular defects
    - Observable Phenotypes

  = DIRECT TARGETS

  = INDIRECT TARGETS

  Observable Phenotypes
Primary Molecular defects = \textbf{DIRECT TARGETS}

Identify RNA associated with CELF1

Secondary \textit{(N)} Molecular defects = \textbf{DIRECT + INDIRECT TARGETS}

Search for GENES with altered expression
- microarray analysis on CELF1 deprived cells
Primary Molecular defects = **DIRECT TARGETS**

- Identify RNA associated with CELF1

- define molecular landscape of CELF 1 target *IN VIVO*

- determine sequences constraint defining the specificity of CELF 1 binding
Primary Molecular defects = DIRECT TARGETS

Identify RNA associated with CELF1

- 3 hybrid screening
- SELEX
- GST pulldown
- Immunoprecipitation
- CLIP

<table>
<thead>
<tr>
<th>GENETIC</th>
<th>IN VITRO</th>
<th>IN VIVO</th>
</tr>
</thead>
</table>
Primary Molecular defects = **DIRECT TARGETS**

Identify RNA associated with CELF1 -CLIP-SolID

2500 nt
Primary Molecular defects = **DIRECT TARGETS**

Identify RNA associated with CELF1 -CLIP-SolID

![Diagram](image)
CLIP procedure

Proteins

RNA/proteins

RNA

DNA

A

B

C

D

Proteins

RNA/proteins

RNA

DNA
CLIP procedure
CLIP procedure

UV(254nm) → CLIP procedure → SOLiD sequencing → In house clonal sequencing

- 62% bacterial rRNA
- 26% Human sequences
- 15% no hit
**RNASEQ procedure**

1. **Hela Cells** → **TOTAL RNA** → **poly(A)+ RNA**
2. Partial digest **RNAs T1** → **T4 PNK**
3. **Adapters + RNA ligase**
4. **Reverse transcriptase, Rnase H** → size selection of cDNA (100-L<250) → **PCR**
5. **In house clonal sequencing**

- **MITOCHONDRIAL GENOME**
- **HUMAN GENOMIC SEQUENCES**
- **GENOMIC HUMAN**
- **HUMAN miRNAs**
- **SMALL RNAs**
- **RIBOSOMAL RNAs**
MAPPING (corona lite)

MOTIF SCORING

CLUSTER DEFINITION

DIFFERENTIAL ANALYSIS

BIOLOGICAL ANALYSIS
MAPPING
(corona lite)
RNAs associated with CELF1
-CLIP-SoliD

6 X 1/4 runs realised → 321 968 741 sequences

Phred quality analysis
Experimental quality analysis
RNAs associated with CELF1
-CLIP-SoliD

6 X 1/4 runs realised → 321 968 741 sequences
83 000 000 expected human sequences

mapping hg19 / corona lite
(from l=30, 3mm to l=24, 2 mm)

20 589 217 uniquely mapped (repeat masked)
4 X 1/4 runs realised  →  347 347 696 sequences

mapping hg19 / corona lite
(l=30, 3mm)

79 869 723 uniquely mapped (repeat masked)
The RNASEQ:
- is strand specific
- cover the exons
- does not map on intronic sequences

94% of the mapped tag are preceded by a G
The CLIPSEQ:
- is strand specific
- cover both intronic and exonic sequences
- cover specific regions of the genome

No enrichment for G residue upstream of sequence tag.

WHAT ARE THE ENRICHED REGIONS IN CLIP SEQ vs RNASEQ?
FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology

Anthony P. Fejes¹,*, Gordon Robertson¹, Mikhail Bilenky¹, Richard Varhol¹, Matthew Bainbridge² and Steven J. M. Jones¹,*
In mRNAs 3'UTR are the major binding sites for CELF1

DO PREVIOUSLY IDENTIFIED TARGET CONFIRM THE VALIDITY OF THE CELF1 CLIP?
Identification of CUG-BP1/EDEN-BP target mRNAs in Xenopus tropicalis

Antoine Graindorge¹, Olivier Le Tonquèze¹, Raphaël Thuret², Nicolas Pollet², H. Beverley Osborne¹ and Yann Audic¹,*

137 mRNAs (/ 3000 tested) with human orthologs are specifically enriched in CELF1 IP in XENOPUS

43 are identified as CELF1 targets in Hela cells
TRANSCRIPT-WIDE DISTRIBUTION OF CELF1 CLUSTERS

2788 enriched clusters

CELF1 CLIP vs RNA-SEQ

FINDPEAKS on exonic reads

5'UTR 2%
3'UTR 80%
CDS 18%

In mRNAs 3'UTR are the major binding sites for CELF1

ARE THE CELF 1 CLUSTERS ENRICHED IN POTENTIAL BINDING SITES FOR CELF1?
CUG-BP1/CELF1 requires UGU-rich sequences for high-affinity binding

Julien MARQUIS*1, Luc PAILLARD†, Yann AUDIC†, Bertrand COSSON†2, Olivier DANOS*, Christine LE BEC* and
H. Beverley OSBORNE†3

*Génethon, CNRS UMR 8115, 1 bis rue de l'Internationale 91002 Evry cedex 2, France, and †CNRS UMR 6061, Génétique et Développement, IFR 140 GFAS, Université de Rennes 1,
Faculté de Médecine, 2 Avenue Léon Bernard, CS 34317, 35043 Rennes Cedex, France

---

A

EDEN15 motif

---

CUGBP1 RRM2

Structural Insights into RNA Recognition by the Alternate-Splicing Regulator CUG-Binding Protein 1

Marianna Teplova,† Jikui Song,† Hai Yan Gao,† Alexei Teplov,† and Dinshaw J. Patel*

<table>
<thead>
<tr>
<th>word</th>
<th>count</th>
<th>expect</th>
<th>sigma2</th>
<th>score</th>
<th>rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>gtgtgt</td>
<td>1472</td>
<td>633.09</td>
<td>635.12</td>
<td>33.29</td>
<td>4096</td>
</tr>
<tr>
<td>tgtgtg</td>
<td>1674</td>
<td>850.57</td>
<td>798.05</td>
<td>29.15</td>
<td>4095</td>
</tr>
<tr>
<td>ttggtt</td>
<td>1429</td>
<td>878.34</td>
<td>765.65</td>
<td>19.9</td>
<td>4094</td>
</tr>
<tr>
<td>ttgltt</td>
<td>1402</td>
<td>878.34</td>
<td>765.65</td>
<td>18.92</td>
<td>4093</td>
</tr>
<tr>
<td>tatata</td>
<td>536</td>
<td>245.36</td>
<td>250.25</td>
<td>18.37</td>
<td>4092</td>
</tr>
<tr>
<td>cctccc</td>
<td>633</td>
<td>323.98</td>
<td>300.65</td>
<td>17.82</td>
<td>4091</td>
</tr>
</tbody>
</table>
Gapped Local Alignment of Motifs on TOP 100 CLUSTERS

5'UTR: 2%
CDS: 18%
3'UTR: 80%
TRANSCRIPT-WIDE DISTRIBUTION OF CELF1 CLUSTERS

CELF1 CLIP vs RNA-SEQ

FINDPEAKS on exonic reads

IS THERE ANY ENRICHMENT IN SOME BIOLOGICAL FUNCTIONS ASSOCIATED WITH CELF1 TARGETS?
mRNA stability and translation

- EIF4A2
- CNOT6
- DICER1
- XRN1
- HuR
- CPEB3
- CPEB2
- PTBP1
- CELF1

Splicing regulation

- SFRS1
- SF1
- TARDBP
- CNOT6 mRNA
- EIF4A2 mRNA
- CNOT6 mRNA
- DICER1 mRNA
- XRN1 mRNA
- HuR mRNA
- CPEB3 mRNA
- CPEB2 mRNA
- PTBP1 mRNA
- CELF1 mRNA
- TARDBP mRNA
- SFRS1 mRNA
CONCLUSIONS

- Genome wide landscape for CELF1 mRNAs targets
- Widespread association of CELF1 with mRNAs
- Enrichement for potential CELF1 binding motif in the target sequences
- CELF 1 is probably self regulatory
- CELF 1 binds (controls ?) a large number of RNA BP involved in cytoplasmic and/or nuclear regulation of gene expressions

PERSPECTIVES

WHAT ARE THE FUNCTIONAL CONSEQUENCES OF CELF 1 BINDING ?

- MOLECULAR CONSEQUENCES : - TRANSCRIPTOME
  - POLYSOMAL ANALYSIS
  - SPLICING

-
CNRS-UMR6061  Genetic and Development

Gene Expression and Development

Carole Gautier-Courteille
Serge Hardy
Vincent Legagneux
Olivier Le Tonquèze
Hubert Lerivray
Stephanie Mottier
Agnès Méreau
Marie Cibois
Luc Paillard
Maud Noiret
Gaëlla Boulanger
Yann Audic
Bernhard Gschloessl

former lab member

Antoine Graindorge
ET MAINTENANT ?
MAPPING -> GALAXY FRAMEWORK
QUANTIFICATION RNA SEQ -> GALAXY FRAMEWORK (RQUANT / RDIFF)
Secondary (N) Molecular defects = INDIRECT TARGETS

Search for GENES with altered expression - microarray analysis on CELF1 deprived cells

Hela Cells → SiRNA treatment → TOTAL RNA → ILLUMINA HT12V4 (DNAVISION)

PROTEINS

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTF</td>
<td>Si_CUGBP1</td>
<td>Si_Rd</td>
<td>NTF</td>
</tr>
</tbody>
</table>

CUGBP1

PCNA

Significance Analysis of Microarrays
CLIP procedure

UV(254nm)

[Diagram showing the CLIP procedure with steps involving RNA ligase, reverse transcriptase, RNase H, proteinase K, size selection of cDNA, PCR, and SOLiD sequencing.]
PC

GENOCLUSTER2

web server

GALAXY (PSU)

GALAXY (tubingen)