Theophyllin aptamer switch example

Introduction

Artificial 5'UTR regions were designed which change their structure upon theophyllin binding to the integrated aptamer structure. This triggers the opening of a terminator structure allowing to fully transcribe the mRNA.

Sequenzes

The original sequence can be found in `RS3mut.fa`.

TASKS 1: Fold

1. Run RNAfold on the fasta input file RS3mut.fa
2. Include the soft constraint that models theophyllin binding with –motif="GAUACCAG&CCCUUGGCAGC,(...((((&)...)))...),-9.22"
3. Investigate the changes: ..*How did the MFE change?

The changes to the MFE will be very weak, what does that mean, were there changes at all and how can we find out?

TASKS 2: Predict partition function

1. Run RNAfold on the fasta input file RS3mut.fa
2. Include the soft constraint that models theophyllin binding with –motif="GAUACCAG&CCCUUGGCAGC,(...((((&)...)))...),-9.22"
3. Predict the partition function
4. Investigate the changes: *..How did the MFE change?* ..How does the structure ensemble look like? ..*Is the MFE strongly represented in the ensemble?

Aha, high percentage of structures exhibit the aptamer structure bound to the ligand. But whats the percentage?

TASKS 3: Calculate probability of aptamer structure

1. Use constraints to calculate the accessibility of RBS with and without sRNA
2. File `RS3_constraint.fa` contains the structure of the theophyllin aptamer, which can be used as hard constraint with `-C`

REMINDER:

prob(aptamer) = exp((pf - constraint_energy) / KT) KT = ((temperature + 273.15)*1.98717)/1000.0

References

`Wachsmuth, Manja, Sven Findeiß, Nadine Weissheimer, Peter F. Stadler, and Mario Mörl. 2013. "De Novo Design of a Synthetic Riboswitch That Regulates Transcription Termination." Nucleic Acids Research 41 (4): 2541-51. doi:10.1093/nar/gks1330.`

RNA toehold switch example

Introduction

The toehold switches are artificial designs which were built to regulate translation of a GFP target gene. A small RNA serves as trigger molecule to activate GPF expression by binding to the 5'UTR region. During the design approach, both sequenzes were optimized to adopt only one stable secundary structure when on their own. As soon as the sRNA binds the 5'UTR, a stable complex forms and triggers a strucutal change in the mRNA region. This opens up the hairpin sequestering the RBS and the start codon, thus allowing for ribosomes to bind and initiate translation.

Sequenzes

The 5'UTR can be found in `pAG_TS1_KS001.fa`, the sRNA in `pAG_TS1_AT001` and both together form `pAG_TS1_001`.

```
GGGUGAAUGAAUUGUAGGCUUGUUAUAGUUAUG AACAGAGGAG ACAUAAC AUG AACAAG
CCUAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
................................ (   RBS   ) ....... xxx <- start codon
...... coding region ...
```

TASKS

Constraint folding

Structure must change if 5' end is unpaired. Can be checked with hard constraints

```
cat pAG_TS1_KS001.fa pAG_TS1_KS001_bindingsite.fa | RNAfold -C
```

if `-C` option is removed, then switch will fold into sequestering stem again!

Predict cofolding

cofold with sRNA

```
RNAcofold < pAG_TS1_001.fa
```

To see dot-plot use `-p` option!

Predict binding site

RNAup to predict binding site with opening and binding energies

```
RNAup -b < pAG_TS1_001.fa
```

Calculate RBS accessibility

Use hard constraints to calculate the accessibility of RBS with and without sRNA

Without sRNA binding

```
cat pAG_TS1_KS001.fa | RNAfold -p
cat pAG_TS1_KS001.fa pAG_TS1_KS001_constraint.fa | RNAfold -p -C
```

prob(aptamer) = exp((pf - constraint_energy) / KT) KT = ((temperature + 273.15)*1.98717)/1000.0

e^((-27,13+19,33)÷(((37+273,15)×1,98717)÷1000)) = 0,000003189 = 00,00% of states are completely accessible in the full region in and between RBS and AUG.

With sRNA binding

```
cat pAG_TS1_001.fa | RNAcofold -p
cat pAG_TS1_001.fa pAG_TS1_001_constraint.fa | RNAcofold -p -C
```

2. Predict cofolding with sRNA using RNAcofold
3. Predict binding site using RNAup
4. Calculate RBS accessibility, use hard constraints to calculate the accessibility of RBS with and without sRNA

prob(aptamer) = exp((pf - constraint_energy) / KT) KT = ((temperature + 273.15)*1.98717)/1000.0

References

Green, Alexander A., Pamela A. Silver, James J. Collins, and Peng Yin. 2014. "Toehold Switches: De-Novo-Designed Regulators of Gene Expression." Cell 159 (4): 925-39. doi:10.1016/j.cell.2014.10.002.

Sequence conservation

Introduction

Non-coding RNAs are often better conserved in their secondary structure, than in their sequence. Because of this families of non-coding RNAs are often defined due to structural similarity. RNAfold predicts the structure of a single sequence. RNAalifold can be used to predict the minimum energy structure that is simultaneously formed by a set of aligned sequences. On the one hand this allows to determine if a set of sequences shares a common structure. On the other hand it can show if the structure of additional sequences fit to an already existing set of RNAs

The examples show some RNA families where the common secondary structure is better conserved than the sequence.

Examples

HACA_2 and HACA_18 are fungal H/ACA snoRNAs. They tend to have very extended loop regions

Tasks

Command line commands: The standard call is `RNAalifold < alignment.aln`

`RNAalifold --color --aln < alignment.aln` results in additional coloring of the structure and also outputs the sequence alignment with additional information. Many other options are available, check –help.

Look for the newly created .ps-files: alirna.ps (and aln.ps).

References

Canzler, Sebastian, Peter F. Stadler, and Jana Hertel. "Evolution of fungal U3 snoRNAs: Structural variation and introns." Non-Coding RNA 3.1 (2017): 3.)

Source

RF00546 is a Drosophilid H/ACA snoRNA family.
RF00006 are vault RNAs.
(Source: www.rfam.org)