# STRING 'n' RNAs

#### Alexander Junge

Center for non-coding RNA in Technology and Health, University of Copenhagen

> 30<sup>th</sup> TBI Winterseminar Bled, Slovenia February 17, 2015



Conclusion

# non-coding RNAs (ncRNAs)

- RNAs not coding for proteins exist in all kingdoms of life
- Fulfill diverse set of biological functions
- Functions rely on **interactions** with proteins, coding RNAs and other non-coding RNAs



### ncRNA-Protein Interactions

# Example: ribosomal RNAs (rRNAs)

- Ribosome apparatus consists of two subunits
- Each subunit composed of rRNA and different proteins
- Catalytic core made of rRNAs
  ⇒ Ribosome is ribozyme
- Proteins form scaffold



#### ≈80 proteins Ribosomal RNAs

PDB Molecule of the Month series by David

S. Goodsell, October 2000

## ncRNA-mRNA Interactions

# Example: microRNAs (miRNAs)

- Short, single-stranded RNA molecules of  $\approx$  22 nucleotides
- Basepairing with complementary target mRNAs results in
- 1. Translation inhibition or
- 2. mRNA degradation



Ameres & Zamore, Nature Reviews Molecular Cell Biology, 2013

#### Current Understanding of ncRNAs and Proteins

#### Interactions of ncRNAs are often not well characterized.

#### Current Understanding of ncRNAs and Proteins

#### Interactions of ncRNAs are often not well characterized.

What about proteins?

#### Introduction

## The STRING Database (http://string-db.org/)

- Protein-protein interaction networks for 1133 organisms
- Integrates known and predicted protein-protein interactions
  - Experiments
  - Text mining
  - Genomic context
  - Databases
- Interactions are scored according to their reliability



Franceschini et al., Nucleic Acids Res., 2013

#### Introduction

## The STRING Database (http://string-db.org/)

- Protein-protein interaction networks for 1133 organisms
- Integrates known and predicted protein-protein interactions
  - Experiments
  - Text mining
  - Genomic context
  - Databases
- Interactions are scored according to their reliability
- No such resource for ncRNA-target interactions



Franceschini et al., Nucleic Acids Res., 2013

#### Main Project Goals

Aim: Integrate ncRNAs into STRING by

- 1. **Collecting** ncRNA-RNA and ncRNA-protein interactions from various sources
- 2. **Unifying** ncRNA-target interactions and protein-protein interactions

 $\rightarrow$  Facilitate understanding of the cell's complex interaction network

#### The Project Started with a Couple of Hackathons

#### **Code Written At A Hackathon**

http://www.willa.me/2013/11/the-six-most-common-species-of-code.html

#### Let's see how far we have come...

#### Sources for RNA-target Interactions



Experimentally validated interactions from

- StarBase [Li et al., Nucleic Acids Res., 2014]
- miRTarBase [Hsu et al., Nucleic Acids Res., 2014]
- NPInter [Yuan el al., Nucleic Acids Res, 2014]

Predicted miRNA-target interactions from

- miRanda [John et al., PLoS Biol., 2005]
- PicTar [Krek et al., Nature Genetics, 2005]
- TargetScan [Garcia et al., Nat Struct Mol Biol., 2011]

miRNA	Target mRNA	Raw Score
hsa-miR-4685-5p	ENSP00000407818	16.8
hsa-miR-548at-3p	ENSP00000406043	13.2
hsa-miR-205-3p	ENSP00000407858	2.0
:	•	•

 $\rightarrow$  translate raw scores into reliability score  $\in [0, 1]$ 

miRNA	Target mRNA	Raw Score
hsa-miR-4685-5p	ENSP00000407818	16.8
hsa-miR-548at-3p	ENSP00000406043	13.2
hsa-miR-205-3p	ENSP00000407858	2.0
		•
•	•	:

 $\rightarrow$  translate raw scores into reliability score  $\in [0, 1]$ 







#### STRING RNA - a Sneak Peek

HOME RESEARCH RESOURCES PUBLICATIONS ABOUT PEOPLE EVENTS JOBS NEWS CONTACT			
String RNA      Searching for one RNA or protein synonyms in local ncRNA database and in STRING      ncRNA or protein name(*):    (examples: ±1 ±2)      miR-155    organism(*):      Homo sapiens    ;      Search    Reset	Search - single name - Multiple names Help Download		
The following ncRNAs and proteins appear to match your input. Please review the list and check exposed one, then click "Continue" to proceed. You also can directly click the expected one. ncRNAs:			
ncRNA *mR-155*: > hse-mR-155-3p basemR-155-5p <- Back Continue ->	_		
© Center for non-coding RNA in Technology and Health			

#### STRING RNA - a Sneak Peek



## Summary

- Integrate different sources of ncRNA-target interactions
- Convert raw interaction scores into confidences
- Provide research community with more complete picture
  - Protein-protein interactions
  - ncRNA-protein interactions
  - ncRNA-RNA interactions

Next:

• Text mining to find RNA-RNA- and RNA-protein-interactions

#### Acknowledgements



CENTERFO RBIOLOGI CALSEQU ENCEANA LYSIS CBS



Novo Nordisk Foundation Center for Protein Research, University of Copenhagen

- Jan Christian Refsgaard
- Alberto Santos Delgado
- Lars Juhl Jensen

Center for Biological Sequence Analysis, Technical University of Denmark

- Christian Garde
- Chris Workman

Center for non-coding RNA in Technology and Health, University of Copenhagen

- Xiaoyong Pan
- Ferhat Alkan
- Christian Anthon
- Jan Gorodkin

My funding is provided by





#### Acknowledgements



CENTERFO RBIOLOGI CALSEQU ENCEANA LYSIS CBS

Novo Nordisk Foundation Center for Protein Research, University of Copenhagen

- Jan Christian Refsgaard
- Alberto Santos Delgado
- Lars Juhl Jensen

Center for Biological Sequence Analysis, Technical University of Denmark

- Christian Garde
- Chris Workman



Center for non-coding RNA in Technology and Health, University of Copenhagen

- Xiaoyong Pan
- Ferhat Alkan
- Christian Anthon
- Jan Gorodkin

Thank you for your attention!