



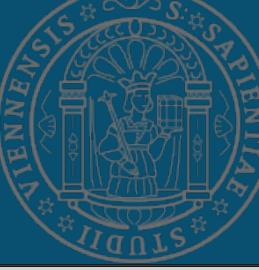
universität
wien

A genome wide RNAz screen for structured RNA elements in mouse

Bernhard Thiel, Andrea Tanzer,
Ivo Hofacker

tbi

A genome wide RNAz screen for structured RNA elements in mouse



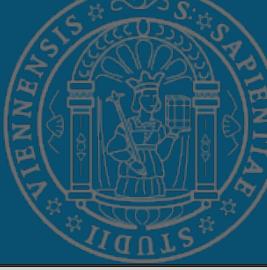
Overview

► Introduction

RNAz, LocARNA

- Our pipeline
- Preliminary results

A genome wide RNAz screen for structured RNA elements in mouse

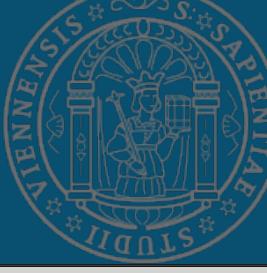


Introduction »

RNAz

- Uses alifold
- A machine learning approach
- Based on
 - Structure Conservation Index
 - Thermodynamic Stability

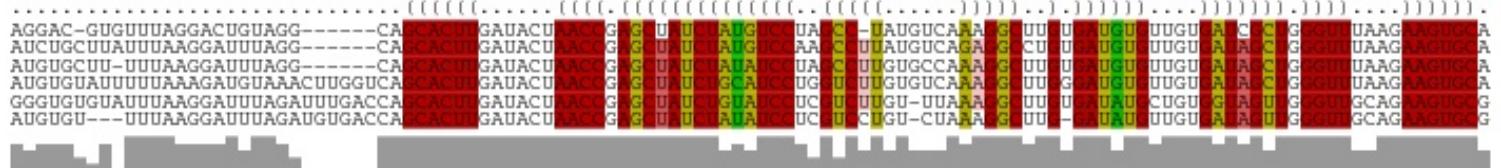
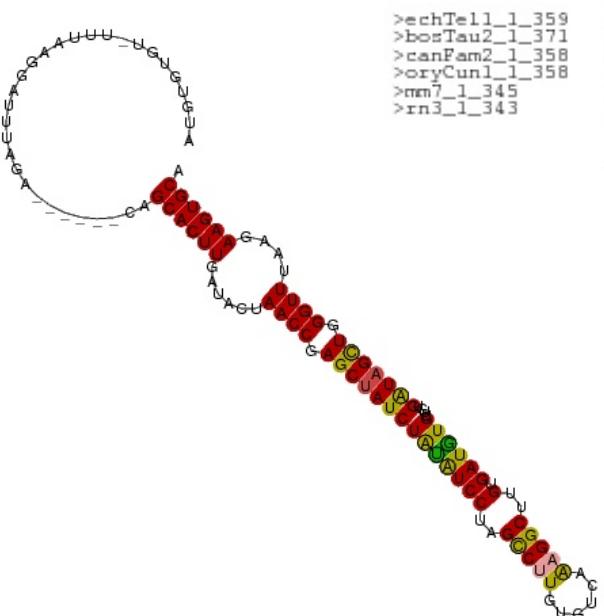
A genome wide RNAz screen for structured RNA elements in mouse



Introduction »

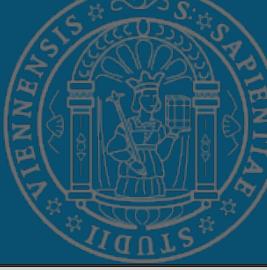
RNAz

```
>echTel1_1_359/80-200
AGGAC-GUGUUUAGGACUGUAGG-----CACGCACUUGAUACUAACCGAGCUAUCUAUGGUCCUAGCC-U AUGUCAAAGGCUUGUGAUGUGAUCGCUGGGUUUAAGAAGUGCA
.(-.-(((((.((..((.----)....)....)).)))).((((.((.(((.((((((.(((-.((....))).)).))).)).))....))).))).). (-32.60)
>bosTau2_1_371/80-200
AUCUGCUUAAAAAGGAUUUAGG-----CACGCACUUGAUACUAACCGAGCUAUCUAUGGUCCUAGGCCAAAGGCUUGUGAUGUGUUGUGAUAGCUGGGUUUAAGAAGUGCA
...(((((.....))).-----) (((((.....((((.((((((.....))).)).))).))).....))).))).). (-40.20)
>canFam2_1_358/80-200
AUGUGCUU-UUUAAGGAUUUAGG-----CACGCACUUGAUACUAACCGAGCUAUCUAUCCUAGGCCUUGUGCCAAAGGCUUGUGAUGUGUUGUGAUAGCUGGGUUUAAGAAGUGCA
.((.((.-(((((.((.----.))).))).))) (((.(((((.((.(((((.....))).)).))).))).....))).))).). (-40.30)
>oryCun1_1_358/80-200
AUGUGUAUUUUAAAAGAUGUAACUUGGUCAGCACUUGAUACUAACCGAGCUAUCUGCAUCCUGGUUCUUGUGUAAAGGCUUGGGAUUGUGAUAGCUGGGUUUAAGAAGUGCA
.....(((((.....((((.((.(((((.((.(((.((.....))).)).))).))).....))).))).))).). (-38.70)
>mm7_1_345/80-200
GGGUGUGUAUUUAAGGAUUUAGAUUUGACCAGCACUUGAUACUAACCGAGCUAUCUGUAUCCUCGUCCUUGU-UAAAAGGCUUGUGUAUGCUGUGGUAGUUGGGUUGCAGAACGUGCG
.((.((.((.((.((.((.....))).))).).((.((.((.((.((.((.((.....))).)).))).....))).))).))).). (-32.90)
>rn3_1_343/80-200
AUGUGU---UUUAAGGAUUUAGAUGUGACCAAGCACUUGAUACUAACCGAGCUAUCUAUCCUCGUCCUUGU-CUAAAGGCUUG-GAU AUGUUGUGAUAGUUGGGUUGCAGAACGUGCG
. ....(((((.....((((.((.(((((.((.((.....))).)).-))).))).....))).))).)).. (-32.90)
>consensus
AUGUGUG_UUUUAAGGAUUUAGA_____ CAGCACUUGAUACUAACCGAGCUAUCUAUCCUAGGCCUUGUGUAAAGGCUUGUGAUGUGUUGUGAUAGCUGGGUUUAAGAAGUGCA
.....(((((.....((((.((.(((((.((.((.....))).)).))).))).....))).))).))).. (-31.53 = -30.48 + -1.05)
```



Mean pairwise identity:	80.69
Mean single sequence MFE:	-36.27
Consensus MFE:	-31.53
Energy contribution:	-30.48
Covariance contribution:	-1.05
Combinations/Pair:	1.28
Mean z-score:	-2.64
Structure conservation index:	0.87
SVM decision value:	1.57
SVM RNA-class probability:	0.96
Prediction:	RNA

A genome wide RNAz screen for structured RNA elements in mouse

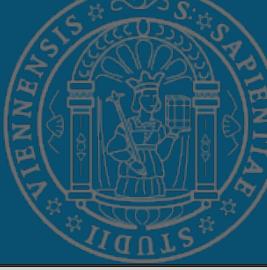


Introduction »

RNAz 2.0

- Dinucleotide shuffled z-scores
- Windows > 400nt
- >10 species

A genome wide RNAz screen for structured RNA elements in mouse

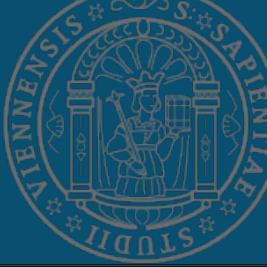


Introduction »

Challenges

- Highly dependent on alignment quality

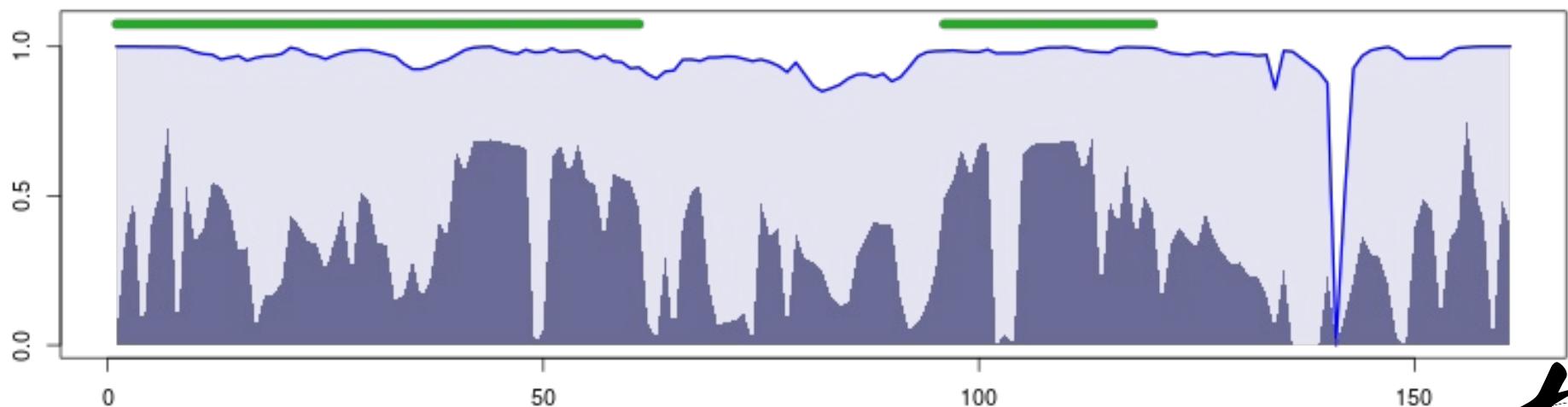
A genome wide RNAz screen for structured RNA elements in mouse



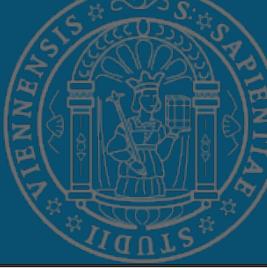
Introduction »

LocARNA

- Structure based multiple sequence alignment
- LocARNA-P
 - Column wise reliabilities for boundary prediction

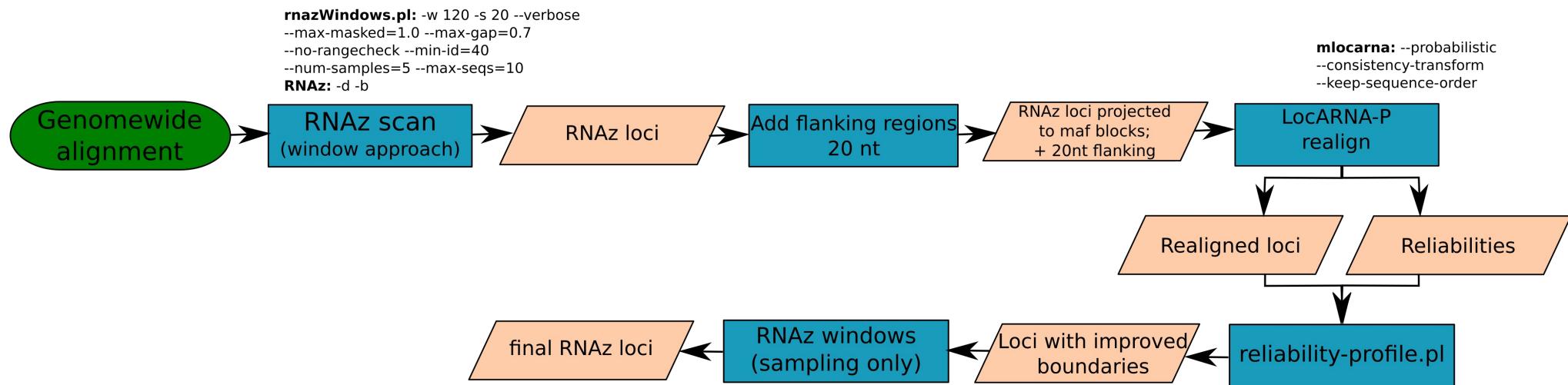


A genome wide RNAz screen for structured RNA elements in mouse

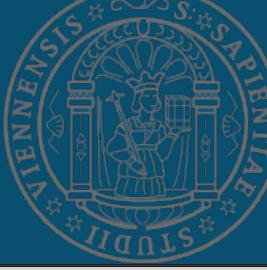


Our Pipeline »

Flowchart



tbi



A genome wide RNAz screen for structured RNA elements in mouse

Preliminary results » False Positive Rate

- Reduced False Positive Rate

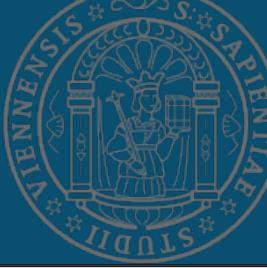
Approach	FDR	Method
Preliminar experiment 1	58%	MultiZ alignments; Multiperm; Low confidence results; Alignments filtered; Chr 11
Preliminar Experiment 2	72%	MultiZ alignments; Multiperm; Low confidence results; Alignments unfiltered; Chr 11
Experiment 2 Loci realigned	26%	LocARNA realigned; Boundaries predicted; Multiperm; Low confidence results;
Genome wide alignment	71%	MultiZ alignments; Multiperm; Low confidence results; Alignments unfiltered;
Genome wide Loci realigned	53%	LocARNA realigned; Boundaries predicted; Multiperm+flanking regions; Low c. results;
Washietl et al.	50%	ENCODE region (1% of human genome); High confidence Predictions; MultiZ alignment; filtered
Smith et al.	22%	High Confidence Results; Multiperm; EPO Alignments

Washietl et al. *Genome Res.* 2007. 17: 852-864

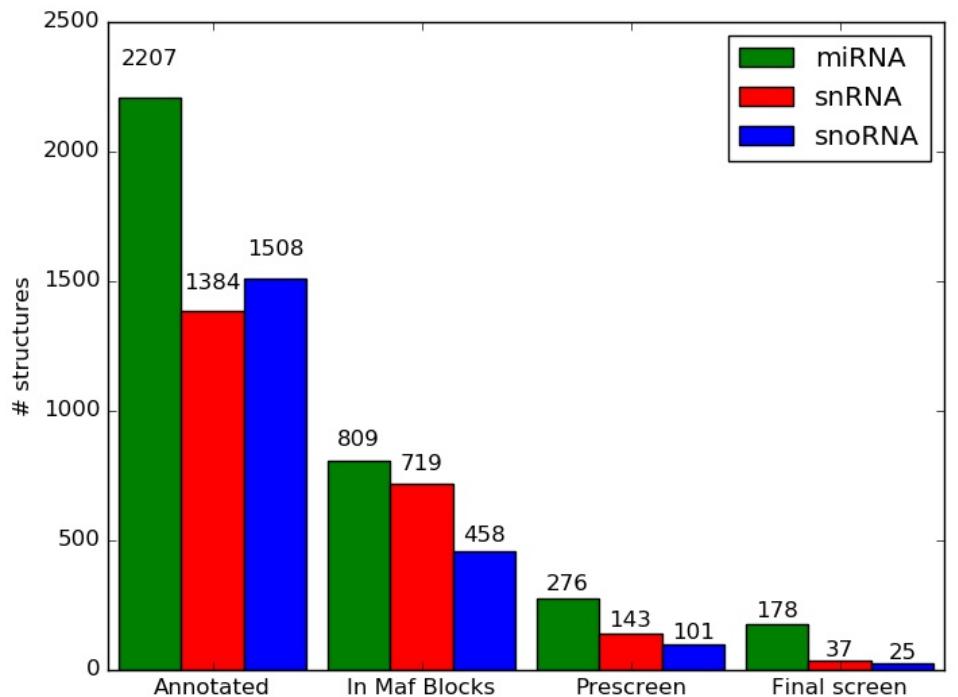
Smith et al. *Nucleic Acids Res.* 2013. 41(17): 8220-36.

The logo for the European Molecular Biology Laboratory (EMBL) Translational Bioinformatics (tbi) group, consisting of the lowercase letters "tbi" in a stylized, italicized font.

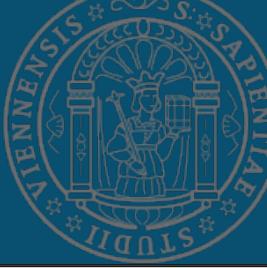
A genome wide RNAz screen for structured RNA elements in mouse



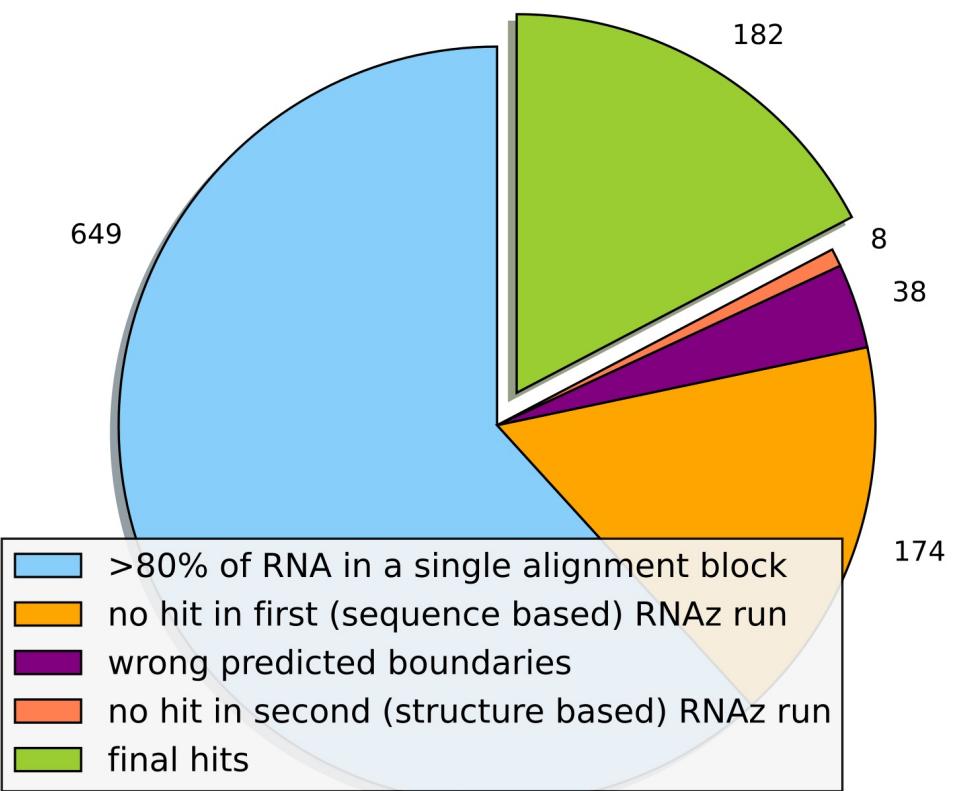
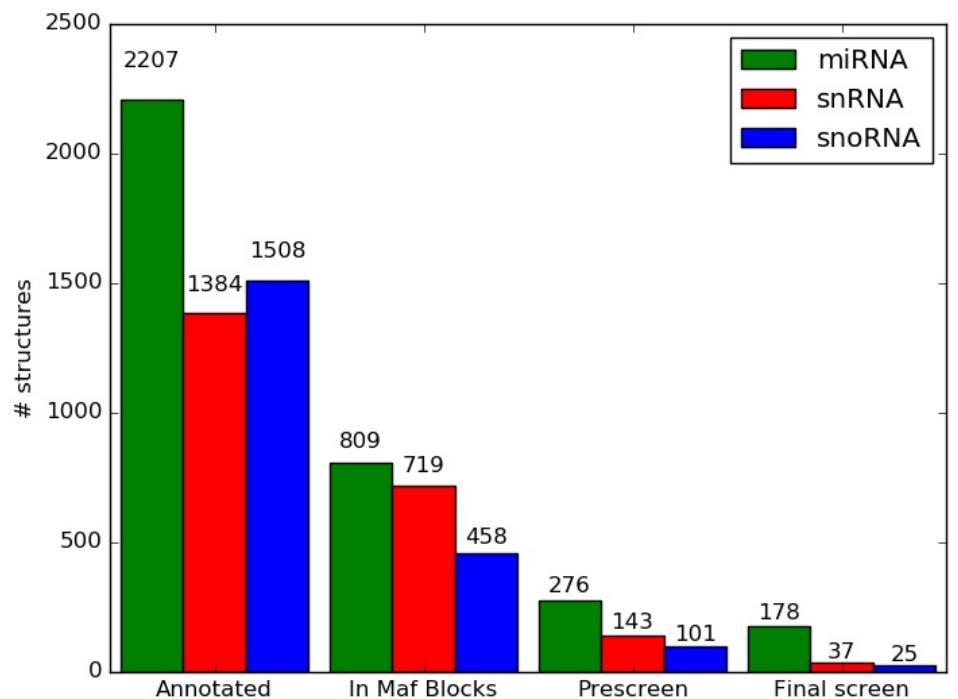
Preliminary results » Annotated RNAs



A genome wide RNAz screen for structured RNA elements in mouse



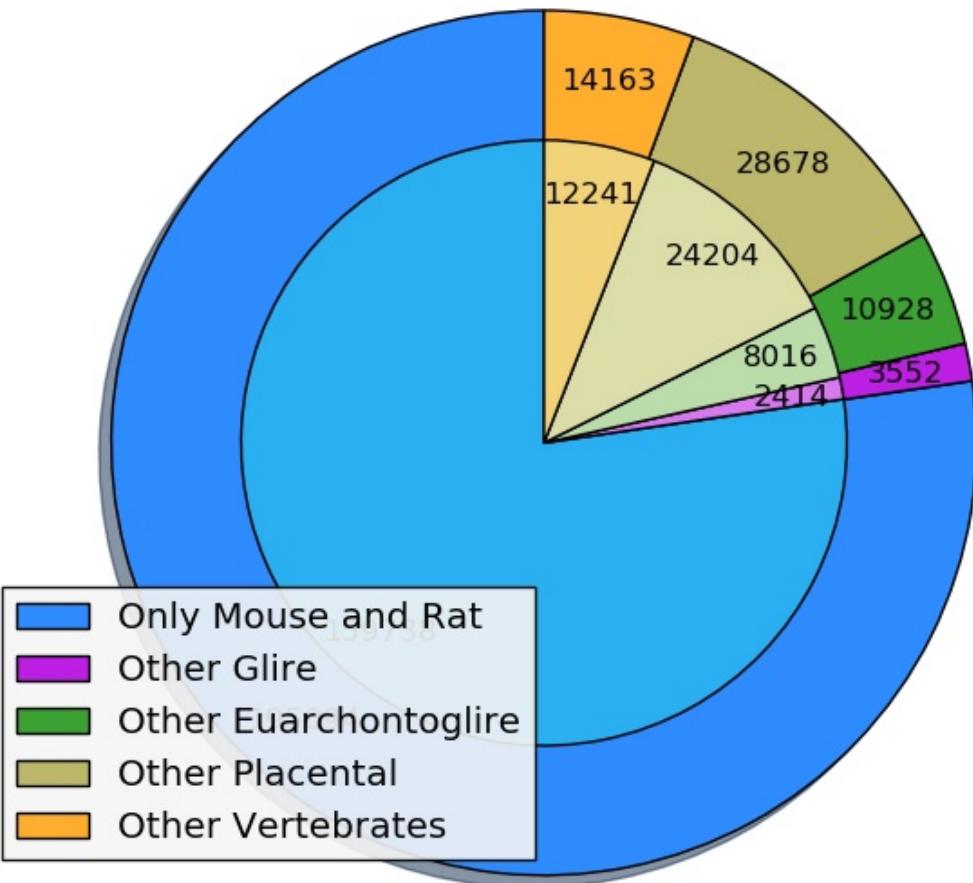
Preliminary results » Annotated RNAs

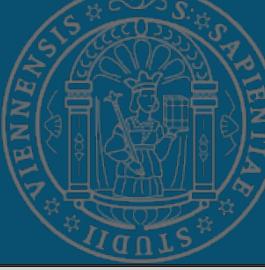


A genome wide RNAz screen for structured RNA elements in mouse



Preliminary results » Phylogenetic information

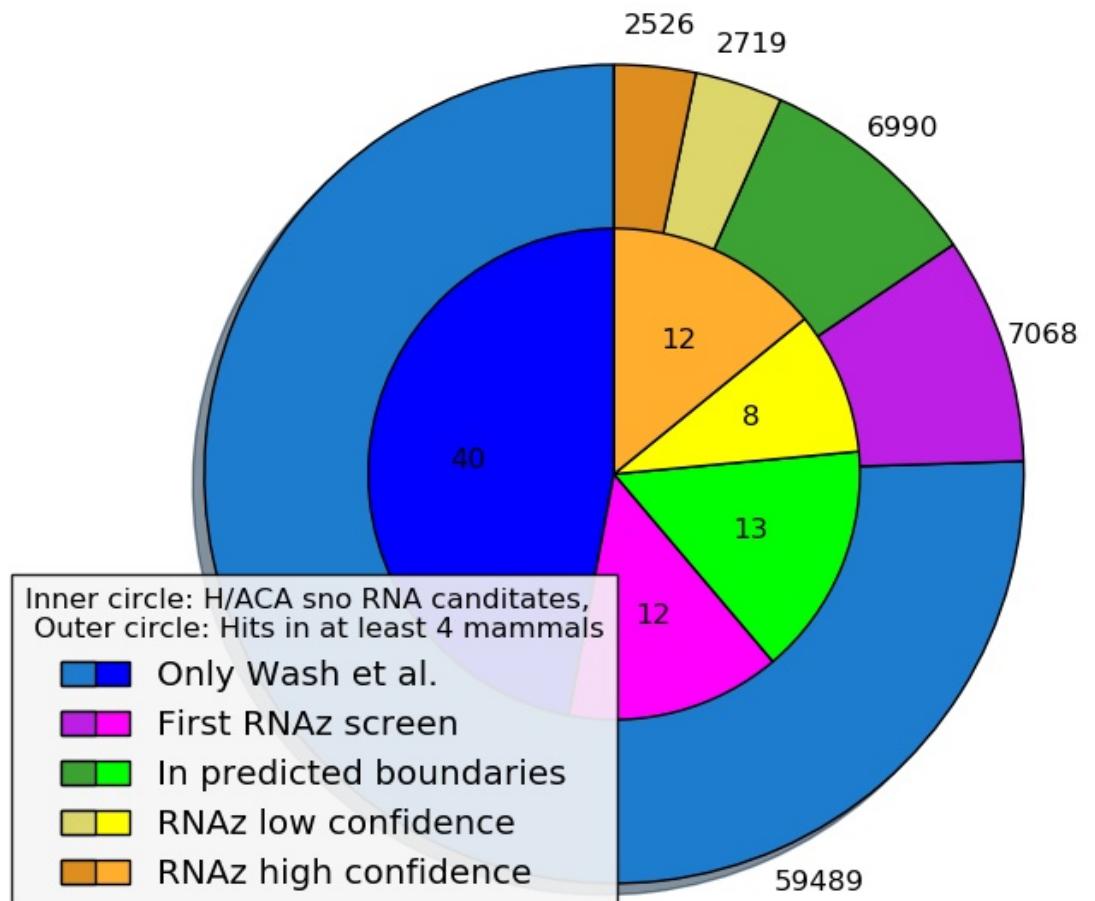




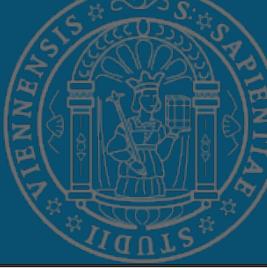
A genome wide RNAz screen for structured RNA elements in mouse

Preliminary results »

Comparing to Washietl et al.

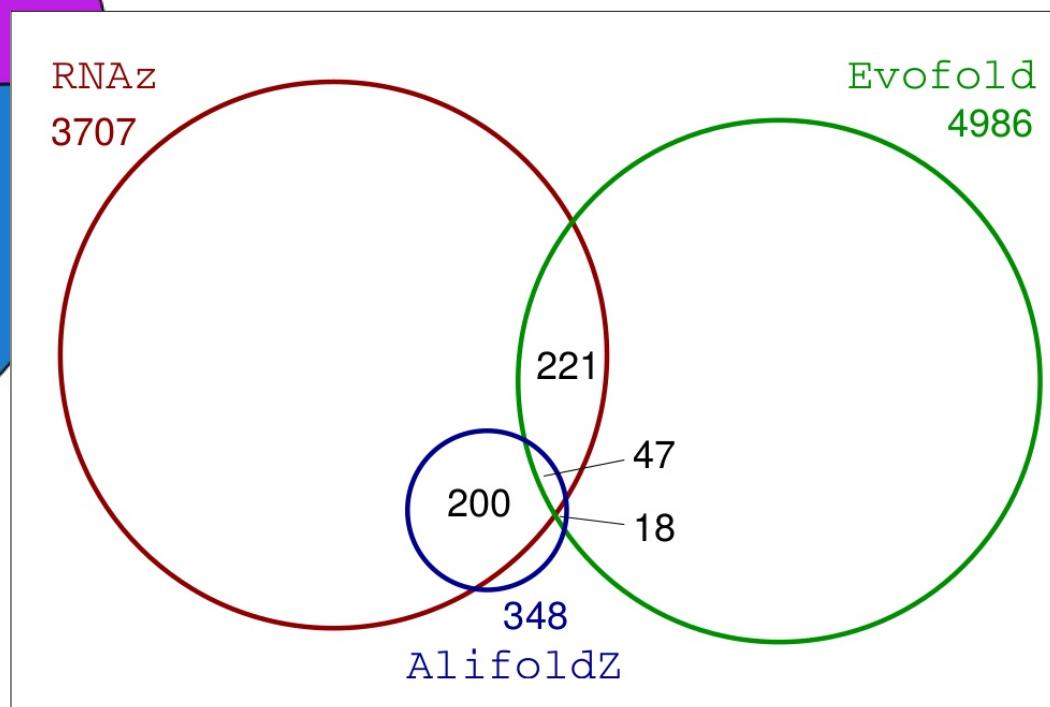
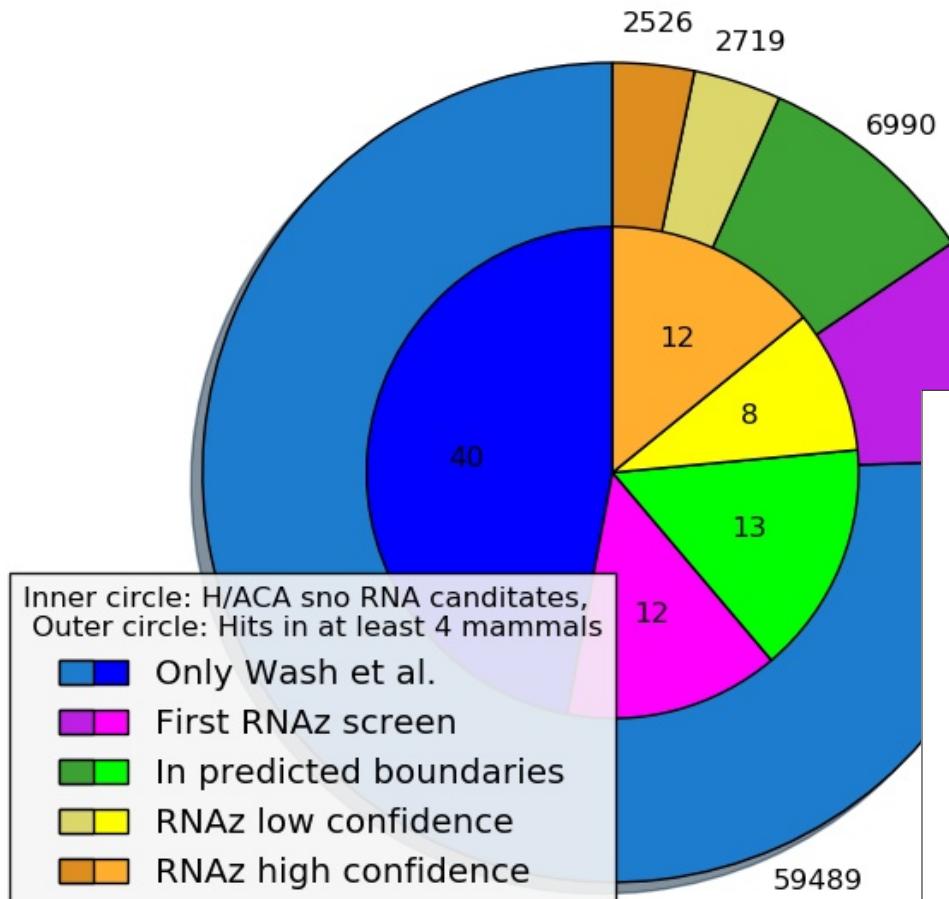


A genome wide RNAz screen for structured RNA elements in mouse

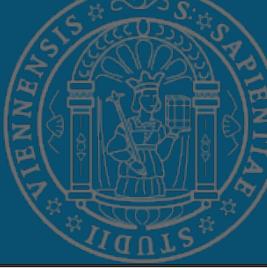


Preliminary results »

Comparing to Washietl et al.



A genome wide RNAz screen for structured RNA elements in mouse



Acknowledgements

Andrea Tanzer

Ivo Hofacker

Colleagues at the TBI

