Structured RNAs are everywhere in the mouse brain

Stefan E Seemann

Center for non-coding RNA in Technology and Health (RTH),
IBHV, University of Copenhagen

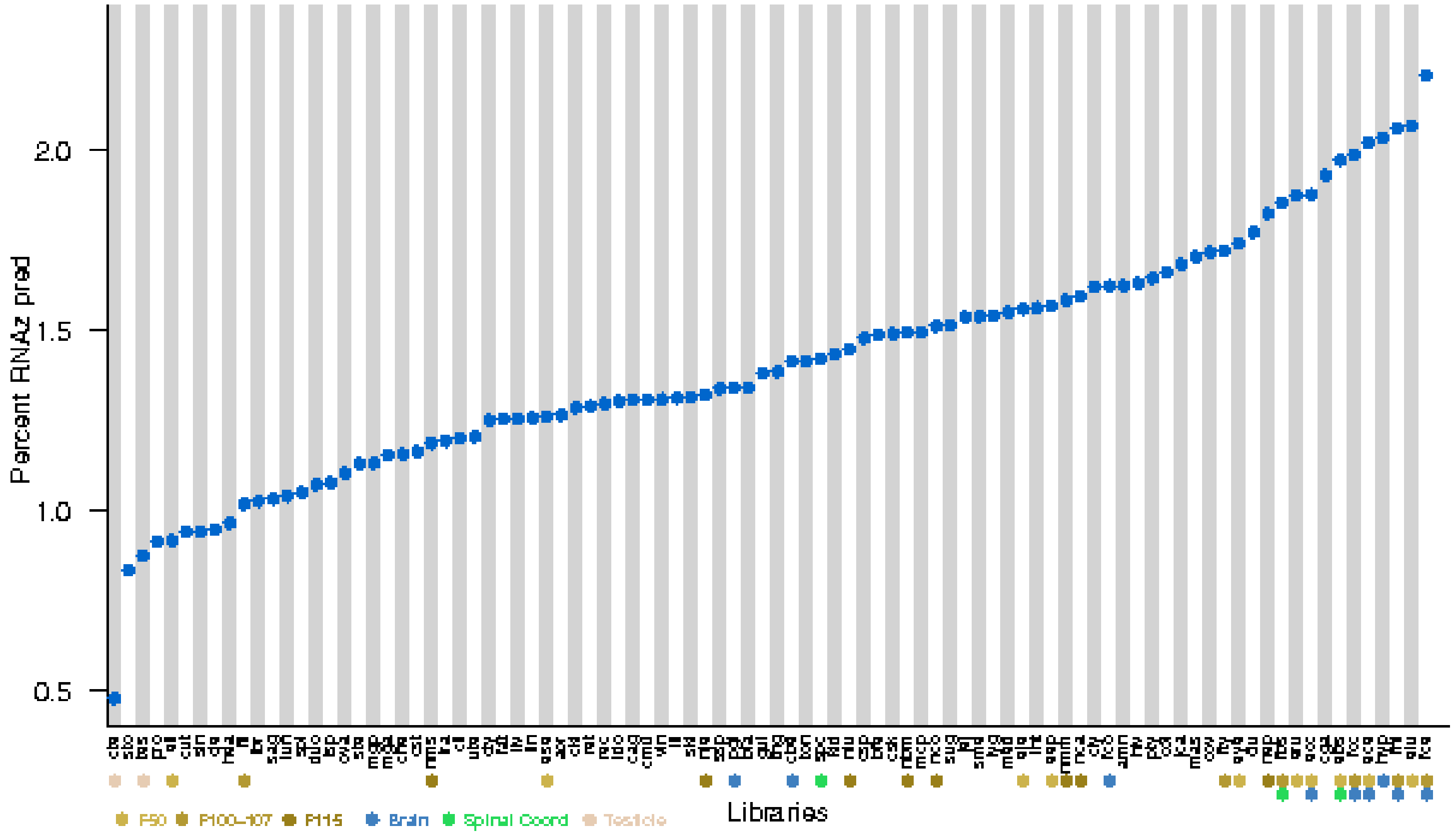
February 18, 2011

Structured RNAs enriched in pig brain

Expression of *in silico* predicted ncRNAs in Sino-Danish pig EST data

☐ 12 libraries comprising neuronal tissues (brain and spinal cord)

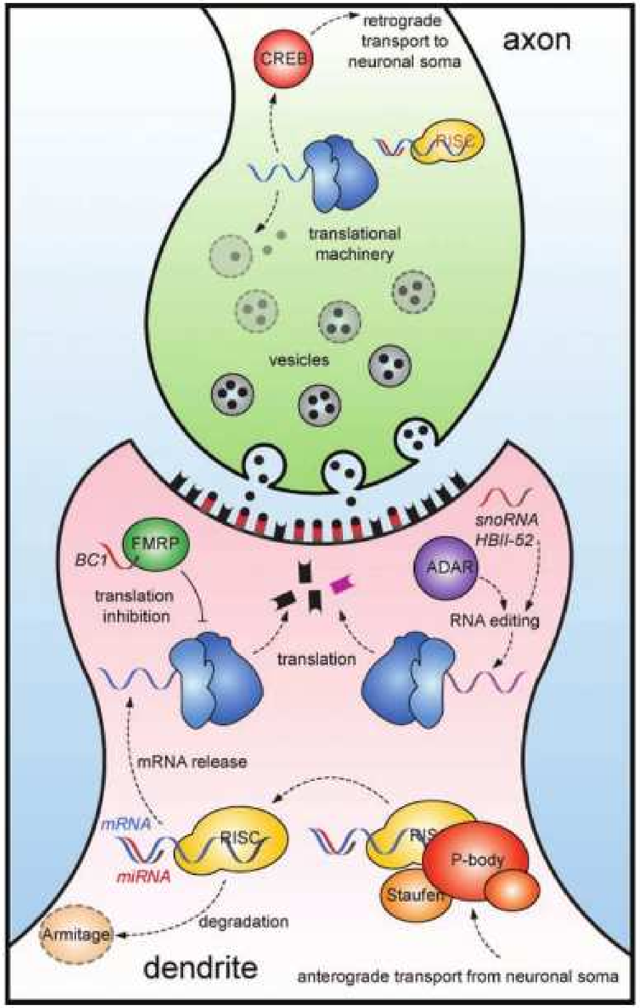
☐ most RNA motifs expressed in developmental tissues, esp. brain



Neuronal expression of structural ncRNAs

▷ Neuronal expression
RNA structure detection
Allen Brain Atlas
Spatial expression
RNA binding proteins
Neuroblast RNA-RNA interaction
Summary

Regulatory ncRNAs functioning in brain:



[Mercer 2008]

lincRNAs (large intervening ncRNAs)
HOTAIR – trans-acting gene regulation
Air and *H19* – direct the imprinting of neighboring genes

microRNAs
miR-124 – guides neuronal differentiation by targeting large number of mRNAs
miR-134 – regulates size of dendritic spines by inhibiting Limk1

snoRNAs (small nucleolar RNAs)
HBII-52 – modifies A-I RNA editing + alternative splicing of serotonin receptor

UTR structures

UTR structures function in *cis* or *trans*:

☐ targets of RNA binding proteins: selenocysteine insertion sequence (SECIS) element and nanos 3 UTR TCE

☐ targets of sRNAs: riboswitches

☐ localization of the mRNA: *Drosophila bicoid* 3'UTR regulatory element

☐ many 3' UTRs in developmental mouse are seperately expressed from mRNA

Goal of the study

Integration of

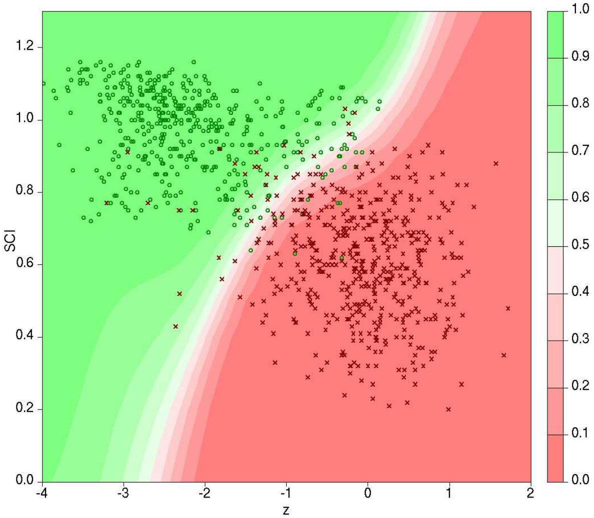
Allen Brain Atlas (ABA) *in situ* hybridization expression data

and

`CMfinder` and `RNAz` *in silico* predicted RNA structures

Genome-wide RNA structure detection

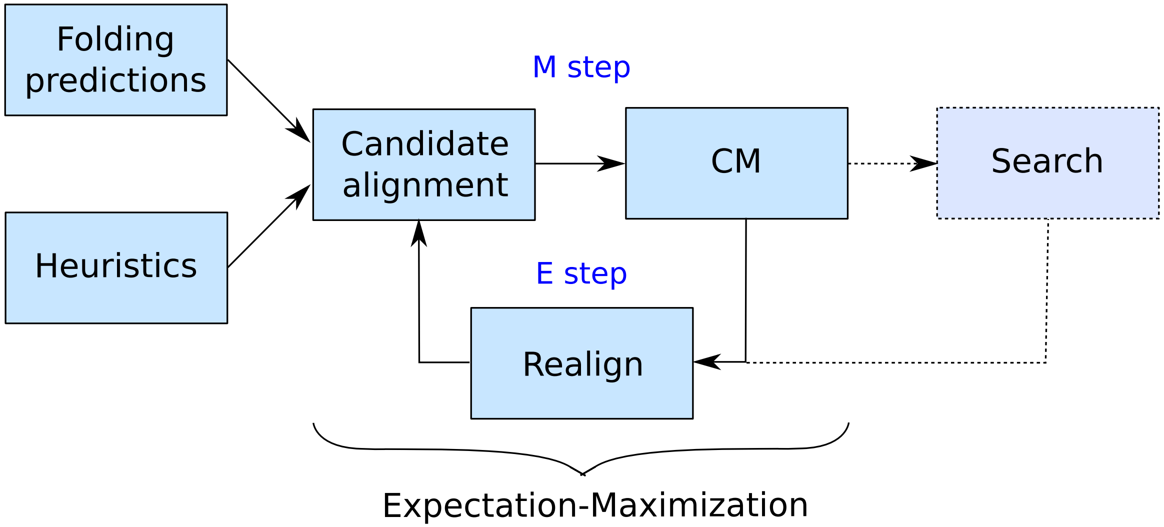
`RNAz` : RNA gene finding from alignments



A SVM is trained to classify based on SCI and z-score (plus number of sequences, mean pairwise identity, alignment length)

⟶ *RNA-class probability* as overall score

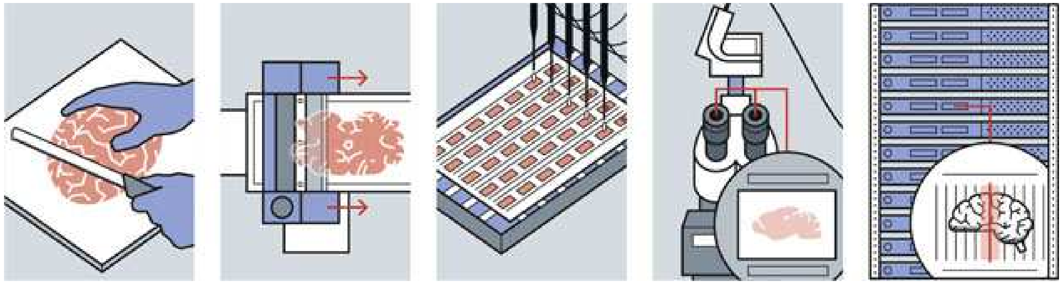
`CMfinder` : RNA motif finder from unaligned sequences



Allen Brain Atlas (ABA)

In situ hybridization (ISH) data of adult mouse brain ⇒ ABA *[Lein 2007]*

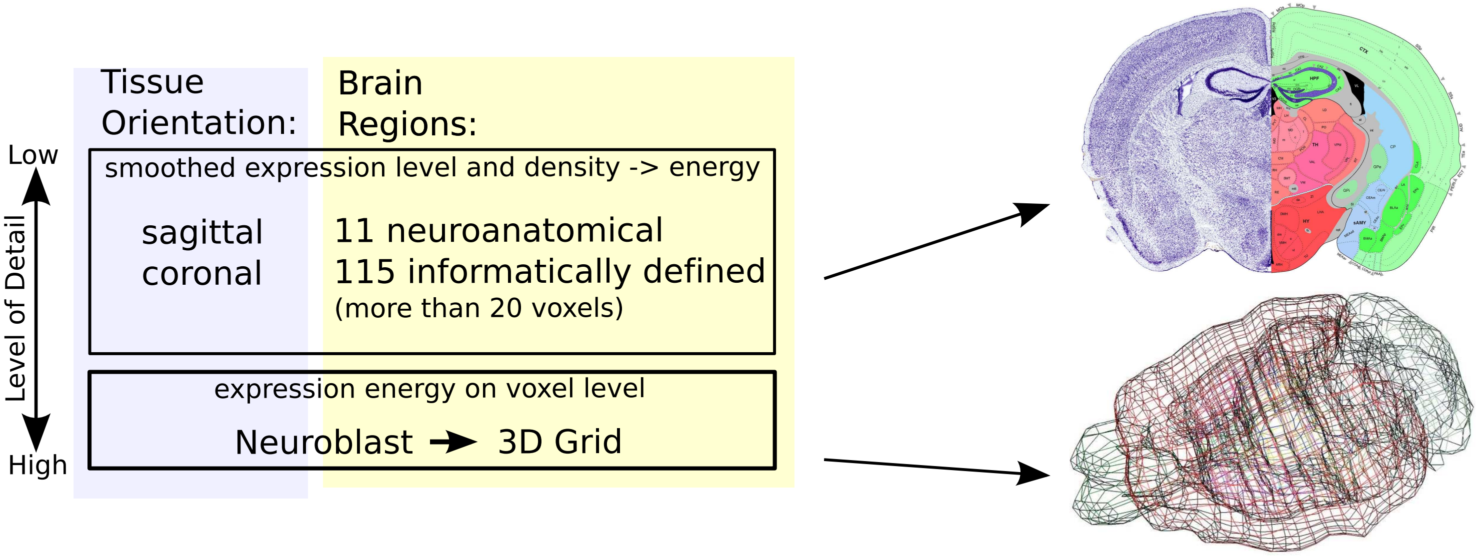
Photo: David Clugston



Displays expression of 21,000 RNA transcripts at cellular resolution

Allen Brain Atlas (ABA)

ISH images are mapped to neuroanatomical regions:



Tissue Orientation:

Brain Regions:

Low

Level of Detail

High

smoothed expression level and density -> energy

sagittal
coronal

11 neuroanatomical
115 informatically defined
(more than 20 voxels)

expression energy on voxel level

Neuroblast ➔ 3D Grid

16,900 riboprobes with cellular expression above background

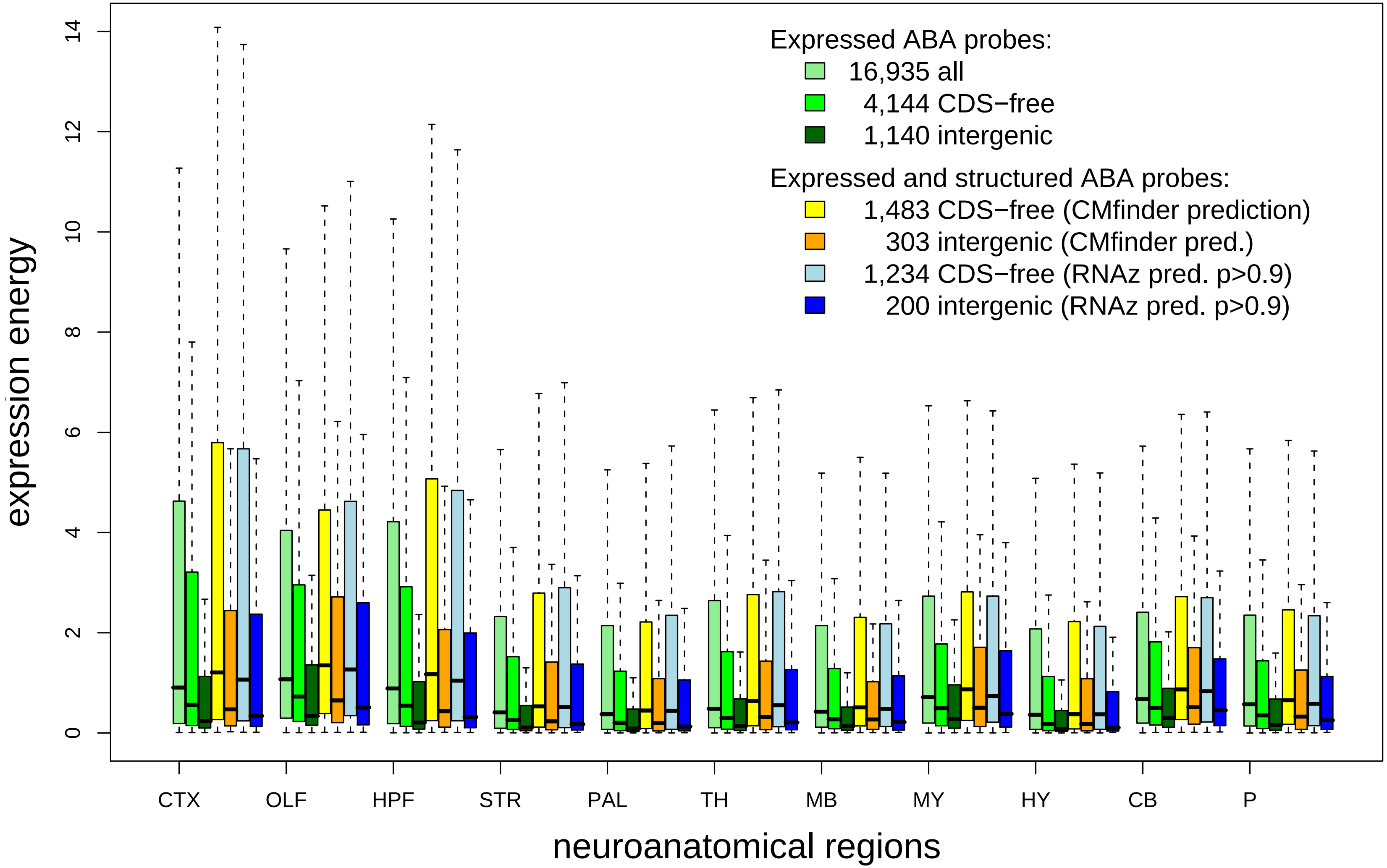
$\sim$ 1,000 probes lack significant protein-coding potential *[Mercer 2008]*

Structured ABA riboprobes

Neuronal expression
RNA structure detection
▷ Allen Brain Atlas
Spatial expression
RNA binding proteins
Neuroblast
RNA-RNA interaction
Summary

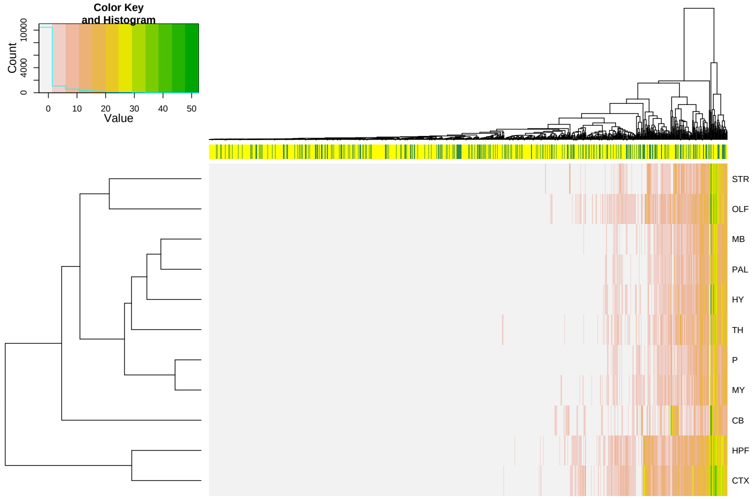
	Intergenic	Intron	5' UTR	3' UTR	Total
(1) Structured CDS-free RNAs					
CMfinder	303	139	17	1,024	1,483
RNAz $p > 0.9$	200	124	21	889	1,234
Total	425	213	31	1,442	2,111
2,025 unstructured CDS-free RNAs					
(2) Long non-coding RNAs (lncRNAs)					
CMfinder OR RNAz AND CRITICA	139	40			179
524 unstructured putative lncRNAs					
(3) Structured UTRs					
CMfinder			632	3,009	3,528
RNAz $p > 0.9$			132	844	960
Total			717	3,368	3,951
8,083 probes with unstructured UTRs					

Spatial expression analyses

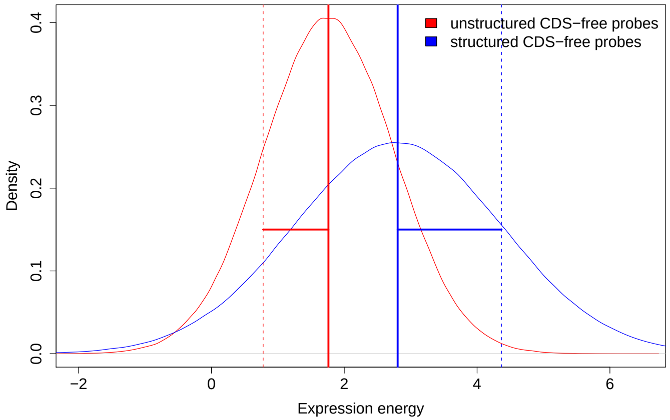


Spatial expression analyses

Statistical significant over-expression of transcripts with structured UTRs (green) in all 11 neuroanatomical regions.



Larger standard deviation of structured probes (blue) describes their higher spatial expression divergence in the brain.



Gene ontology analysis shows an significant overrepresentation of binding, transport and localization for probes with UTR structures.

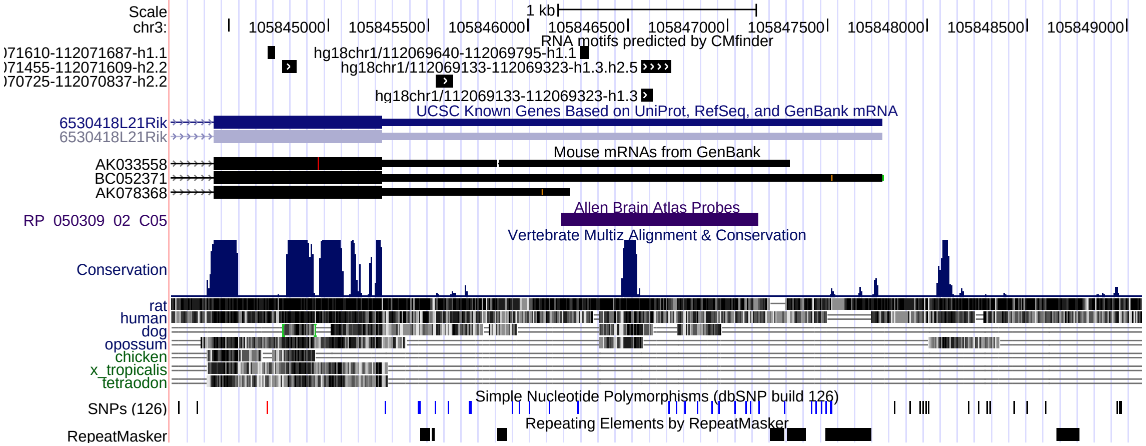
⇒ Hypothesis: 3' UTR structures are *cis*-functional, *e.g.*, support the binding of transport proteins for distributing the host RNA over long distances between nucleus and synapses

Structured UTRs and RNA binding proteins

Neuronal expression
RNA structure detection
Allen Brain Atlas
Spatial expression
▷ RNA binding proteins
Neuroblast
RNA-RNA interaction
Summary

- ☐ 3,951 probes host 8,421 `CMfinder` pred. structures (avg 69nts)
- ☐ predicted RNA structures are enriched in 3’ UTR adjacent to ORF
- ☐ 461 ABA transcripts code for RNA binding proteins in RBPDB
- ☐ similar expr. patterns of structured UTRs and binding proteins
- ☐ no enrichment of correlated expression pairs
- ☐ 90% UTR str./neighbourhood with binding motifs to 19 proteins

3’ UTR structure of *6530418L21Rik* hosts a binding motif to the zinc finger protein *Zfp365* ⇒ brain-wide correlated expression



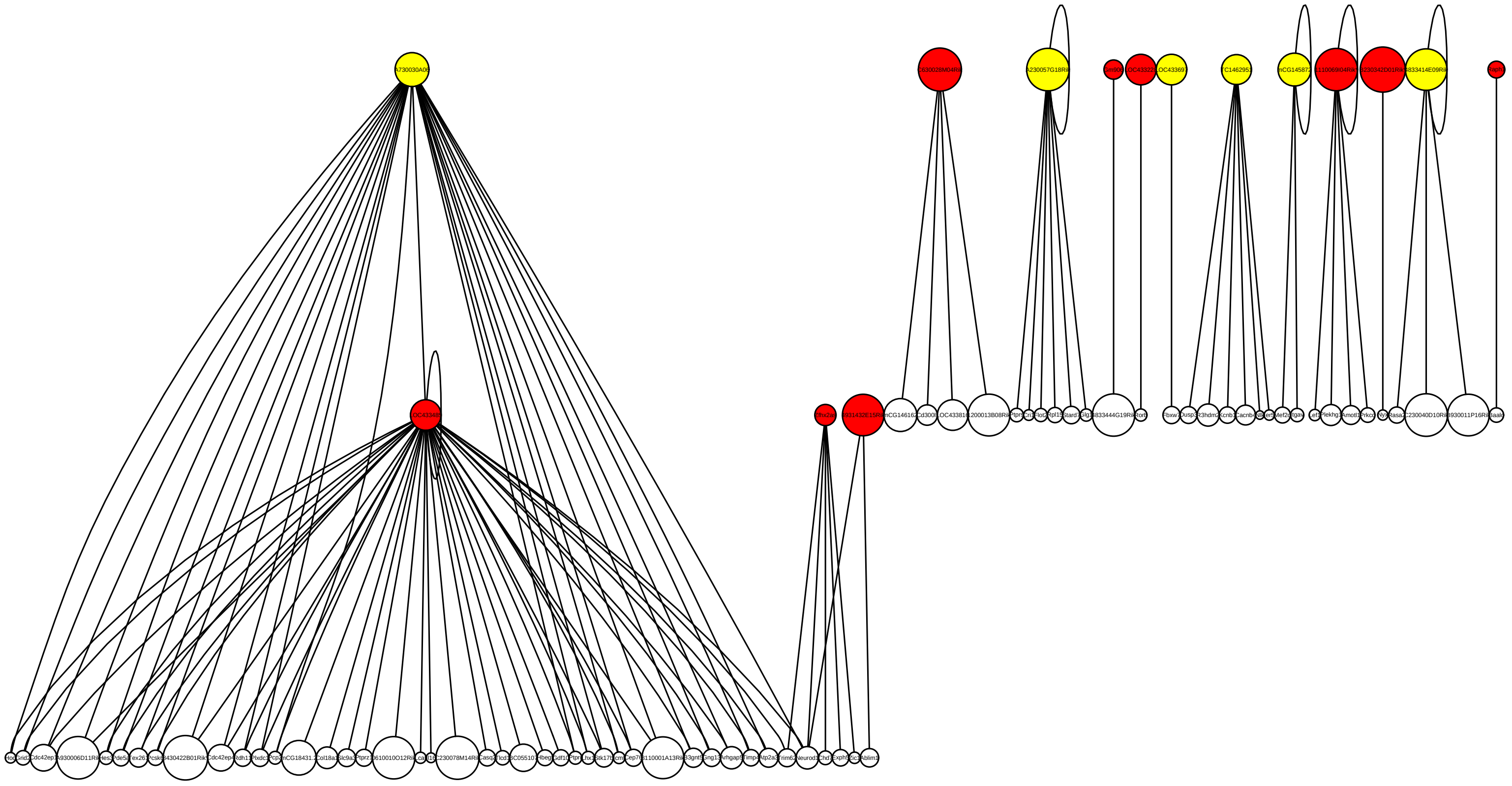
Does *6530418L21Rik* work as signalling molecule and its 3’ UTR structure as binding signal?

Correlated expression of structured RNAs

Neuronal expression
RNA structure detection
Allen Brain Atlas
Spatial expression
RNA binding proteins
▷ Neuroblast
RNA-RNA interaction
Summary

Correlation network of 15 putative lncRNAs with brain-wide expression patterns ($\rho_T = 0.8$).

(yellow - structured lncRNAs, red - unstructured lncRNAs, white - protein-coding RNAs)

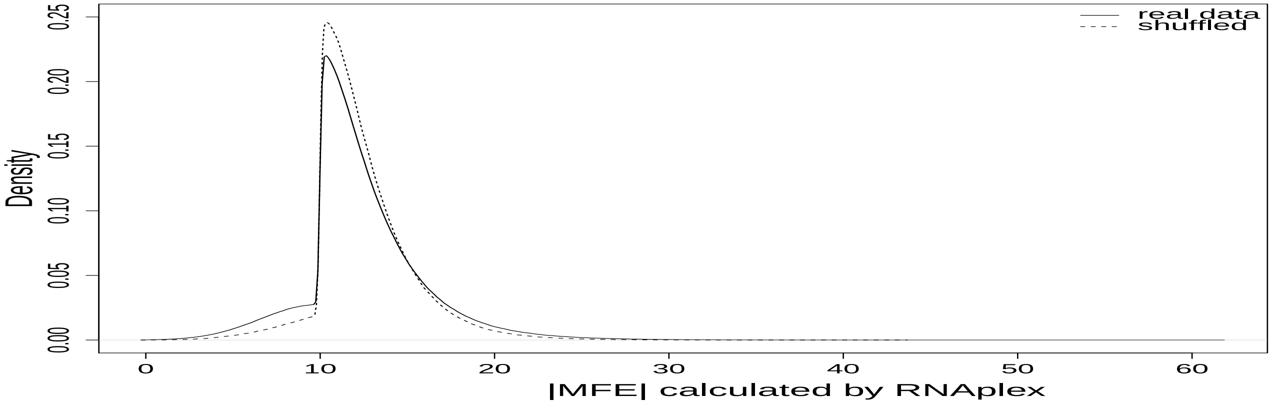


19 structured lncRNA candidates have spatially correlated expression to one transcript ⇒ more spatial patterns as CDS-free probes

Thermodynamic stable RNA-RNA interactions

How well is correlated expression described by RNA-RNA interactions?
→ Search for interaction sites between structured lncRNAs and UTRs of ABA transcripts:

- ☐ accessibility of transcripts by `RNAplfold`
- ☐ thermodynamic stable RNA-RNA interactions by `RNAplex`
- ☐ p-value estimates statistical significance of RNA duplex



- ☐ 13 structured lncRNA candidates have predicted interaction sites to 3' or 5' UTR of the correlated mRNAs
- ☐ predictions are not significant
- ☐ many significant interaction sites with MFE<-40 kcal/mol between lncRNAs and non-correlated expressed mRNAs

Summary

- ☐ *in silico* predicted RNA structures are enriched both in UTR regions and non-coding transcripts in all regions of mouse brain
- ☐ UTR structures may serve as binding motifs for RNA binding and transporting proteins
- ☐ transport of information stored in RNA from nucleus through long axons to synapses ⇒ component of neuronal functionality
- ☐ BUT, exonucleases with problems to initiate degradation close to stable stem structures ⇒ slower degradation of structured transcripts (?)
- ☐ low expression of lncRNA candidates compared to protein coding transcripts
- ☐ several structured lncRNA candidates for correlated expression and RNA-RNA interaction in different brain regions

Acknowledgement

Neuronal expression
RNA structure detection
Allen Brain Atlas
Spatial expression
RNA binding proteins
Neuroblast
RNA-RNA interaction
Summary

Thank you

- Allen Institute for Brain Science, Seattle
  Susan Suskin, Michael Hawrylycz
- University of Washington, Seattle
  Larry W Ruzzo, Zizhen Yao
- University of Copenhagen
  Jan Gorodkin and my colleagues