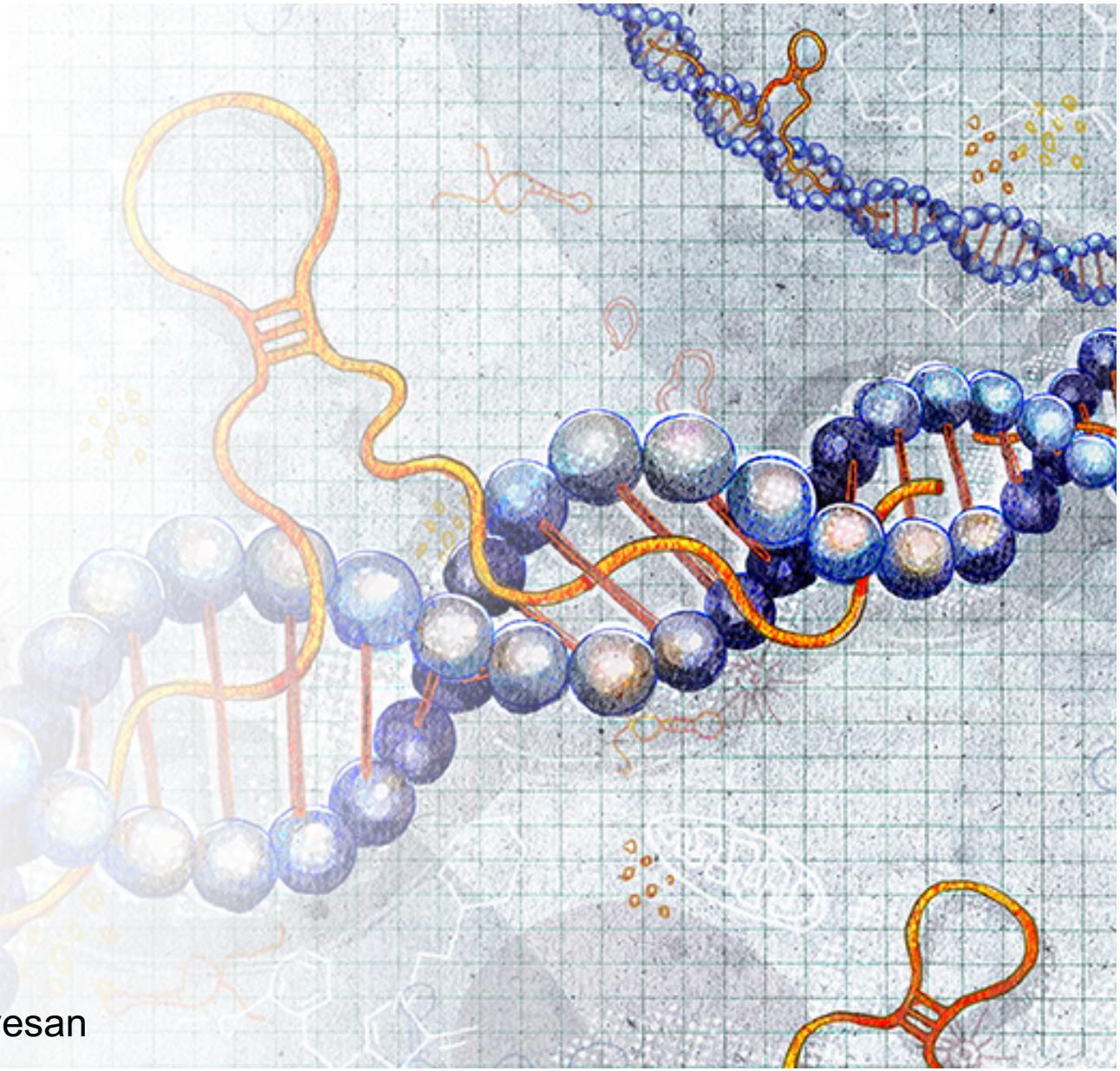
UNIVERSITÀ DEGLI STUDI DI TRIESTE
Dipartimento di Scienze della Vita

Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells



Gabriele Codotto, Lisa Migli and Erica Piovesan

ORGANIZATION

1. Introduction

2. Paper presentation:

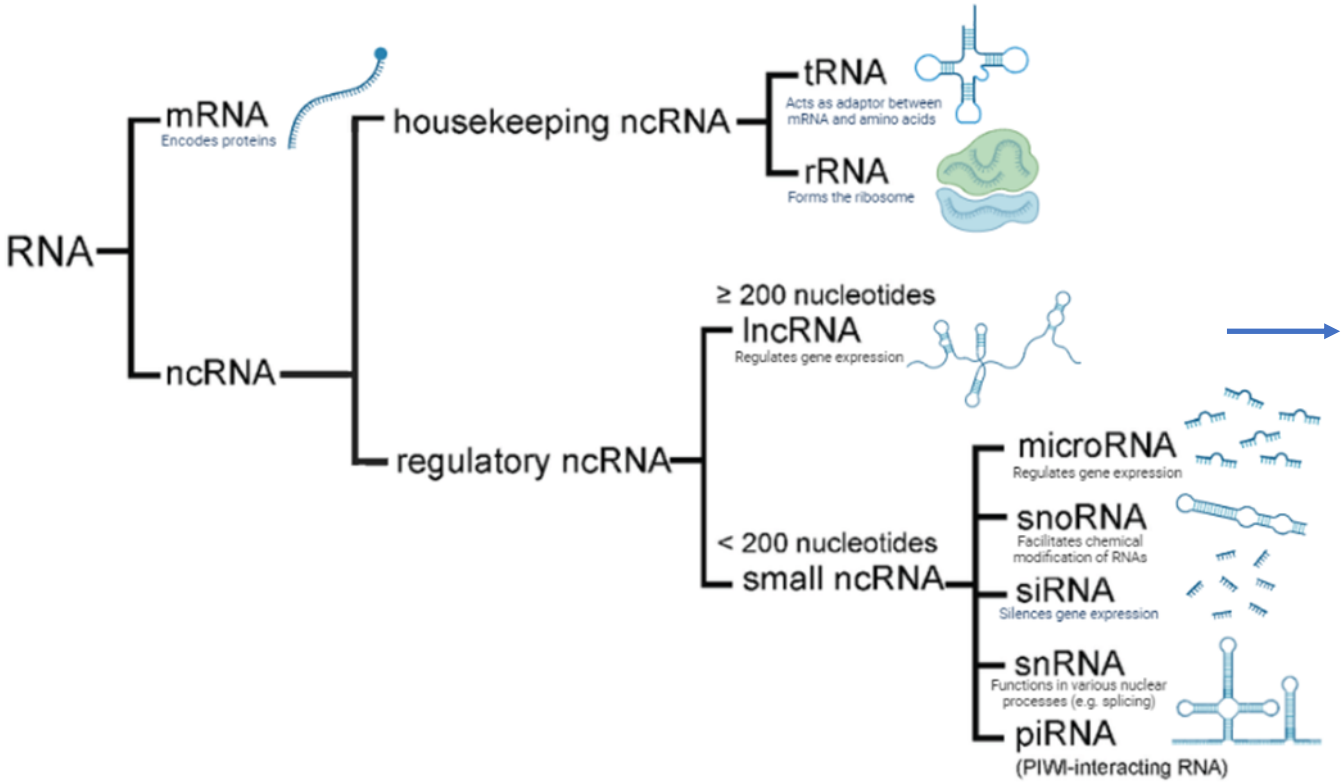
a. Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
b. Additional studies to address lncRNA ribosomal interaction, localization and function;
c. Studying the features of ribosomally-associated lncRNA;
d. Conclusions.

ROADMAP

- Introduction;
- Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
- Additional studies to address lncRNA ribosomal interaction, localization and function;
- Studying the features of ribosomally-associated lncRNA;
- Conclusions.

What are lncRNAs

- Less than 2% of mammalian genome is transcribed into mRNA
- A major portion is transcribed into ncRNA among which there are lncRNA



LncRNA features:

- Lack of protein-coding sequence;
- Transcribed by RNA polymerase II;
- Spliced into long transcripts (> 200 nt);
- They can be capped and polyadenylated;
- They contain abundant small ORF sequences.

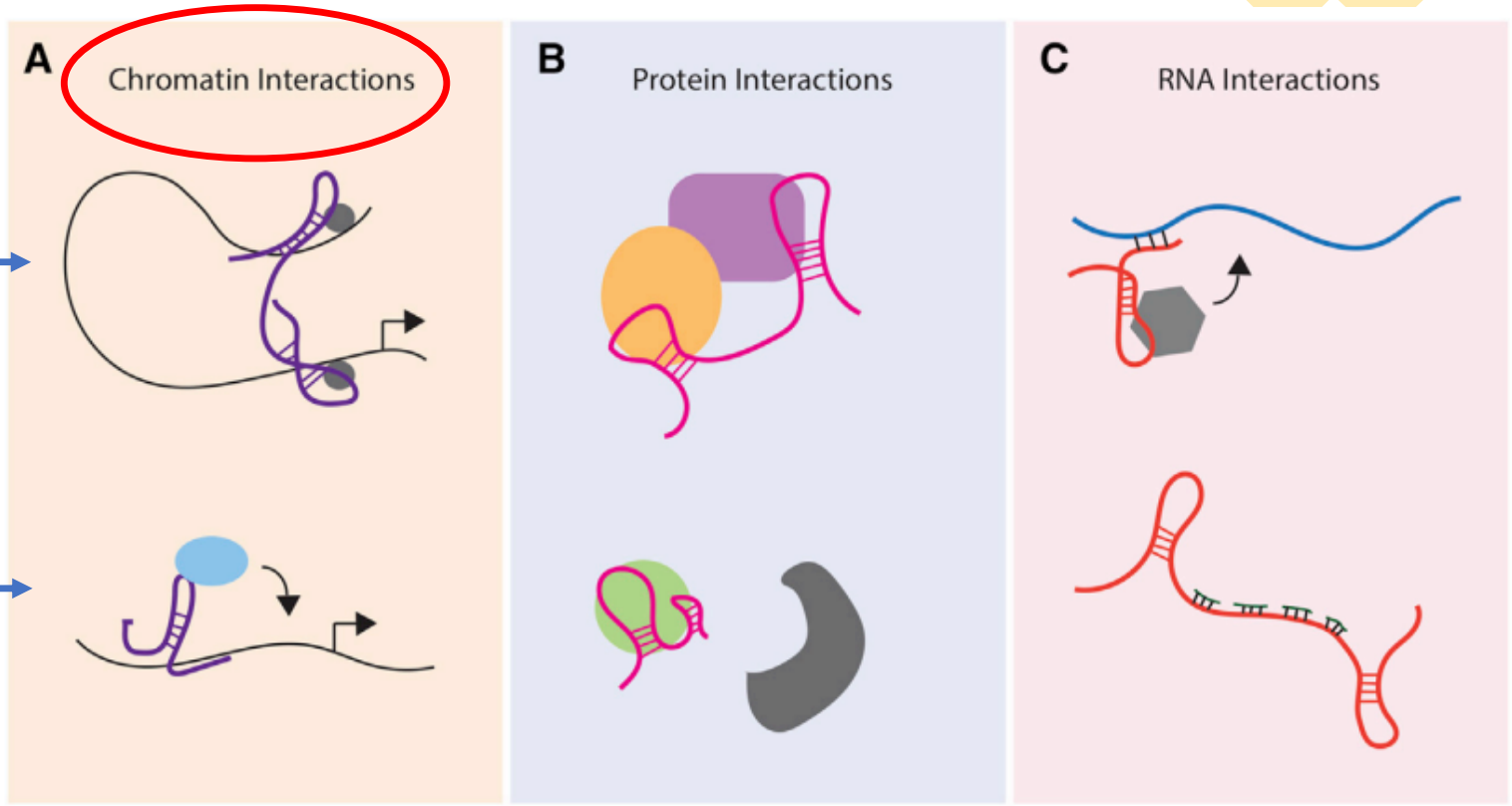
LncRNAs interact with chromatin

XIST

Chromatin-bound lncRNAs can regulate gene expression by controlling local chromatin architecture

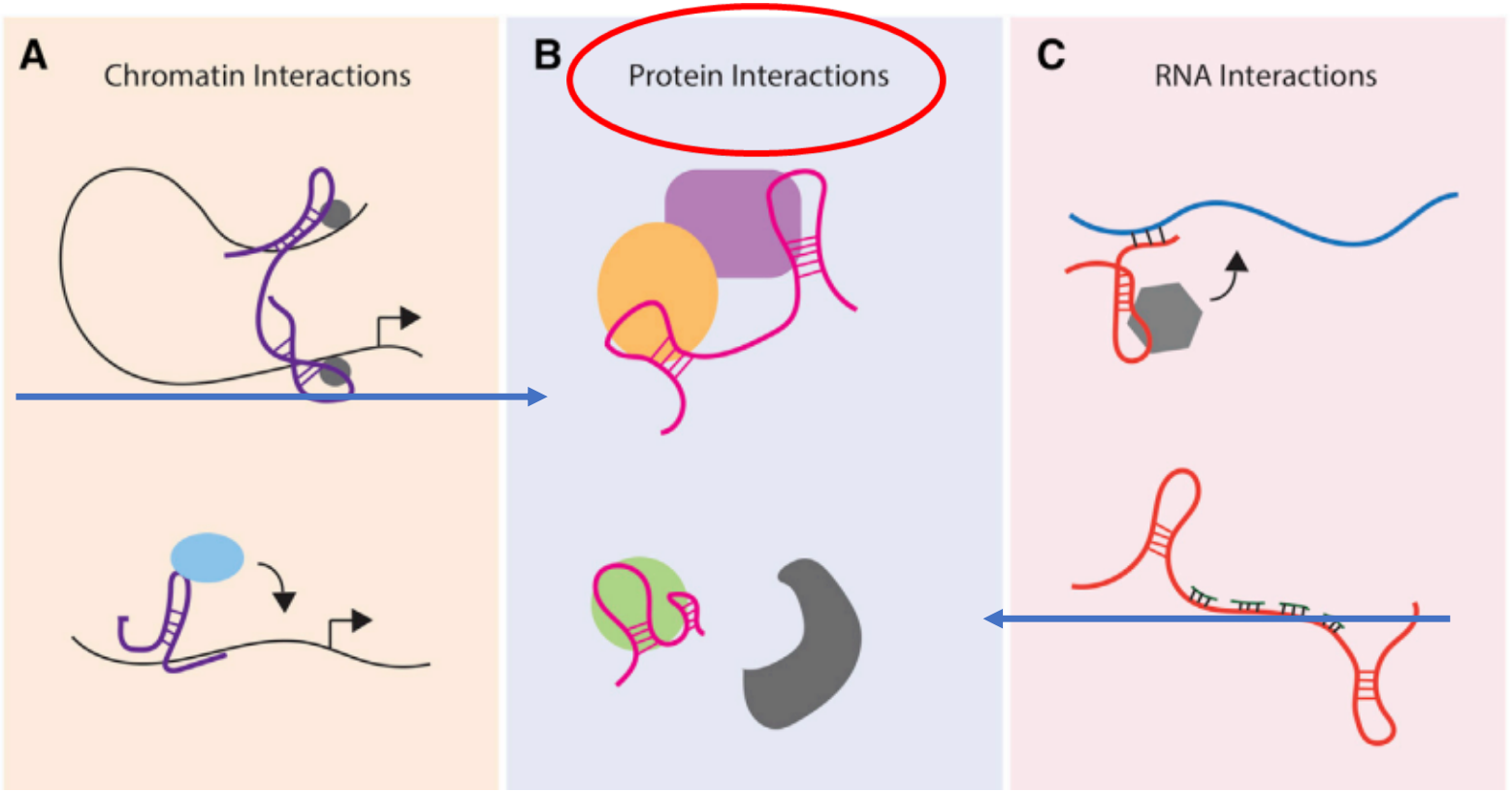
or by directing the recruitment of regulatory molecules to specific loci

HOTAIR-PRC2



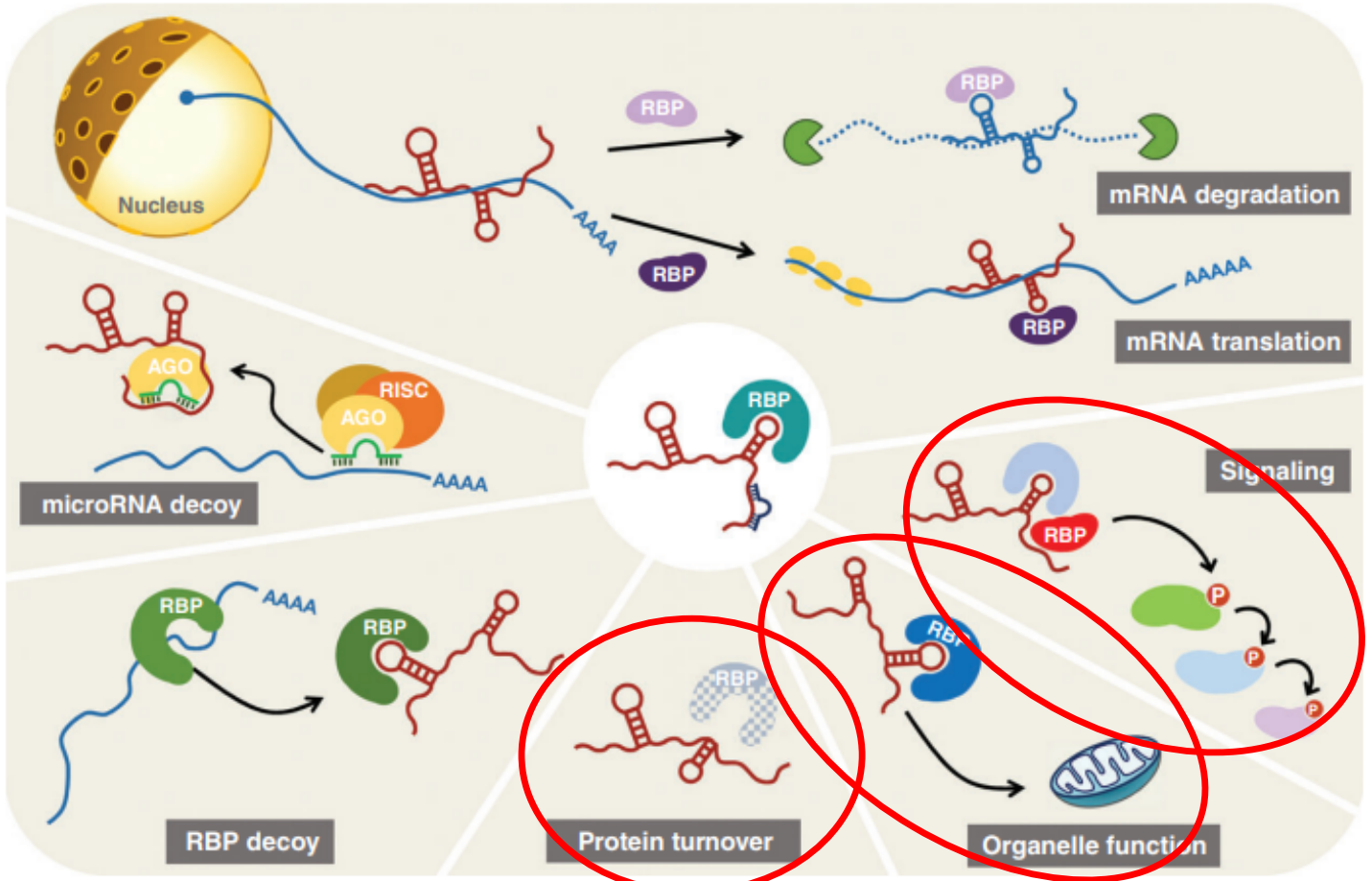
LncRNAs modulate protein function

lncRNA interactions with multiple proteins can promote the assembly of protein complexes



or can impair protein-protein interactions

LncRNAs modulate protein function: examples

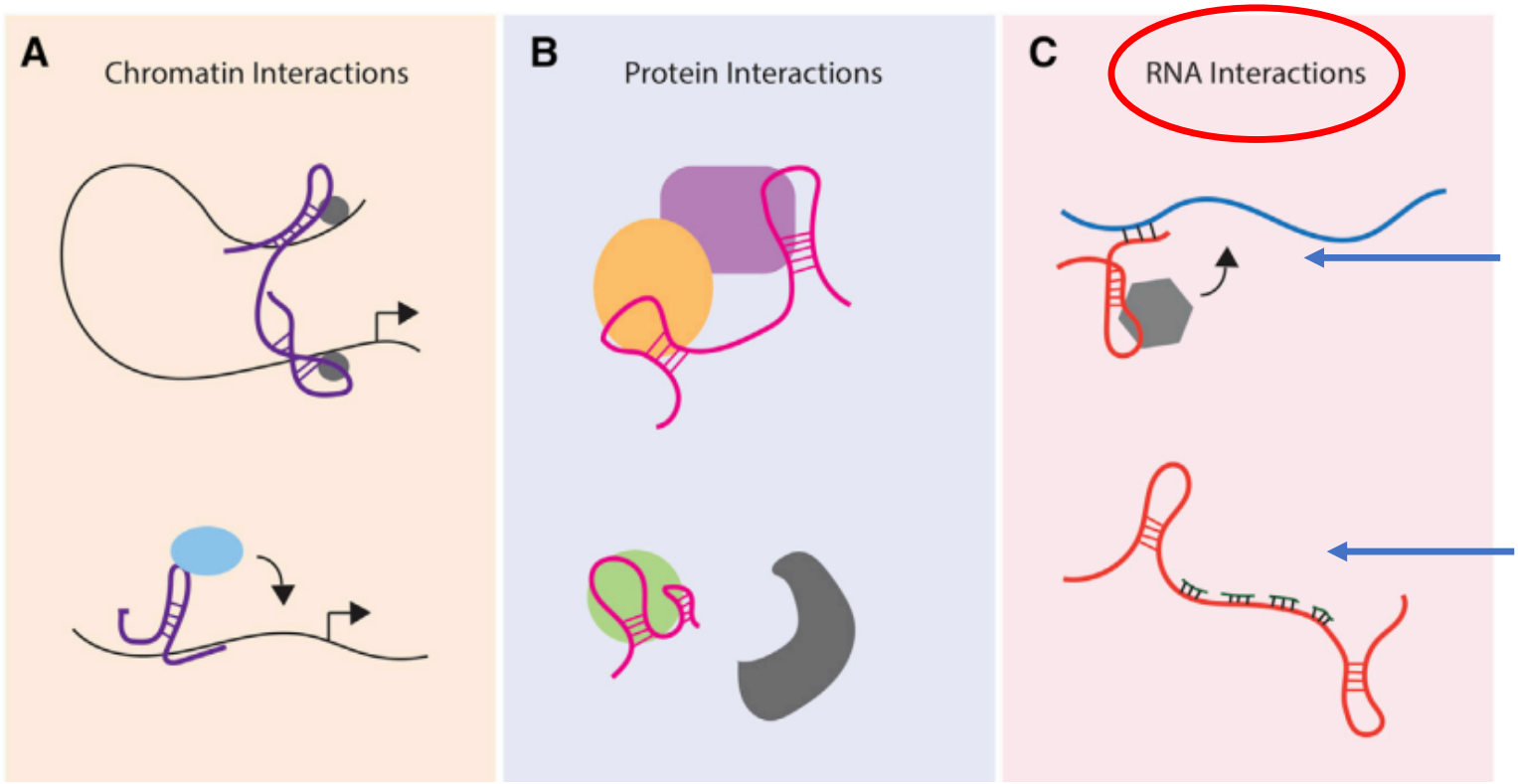


NKILA - NF-κβ/Iκβ complex

LincRNA-p21 – HIF1A

7SL - SRP

LncRNAs modulate RNA metabolism



mRNA interactions with lncRNA can recruit protein machinery involved in multiple aspects of mRNA metabolism to affect splicing, mRNA stability, or translation

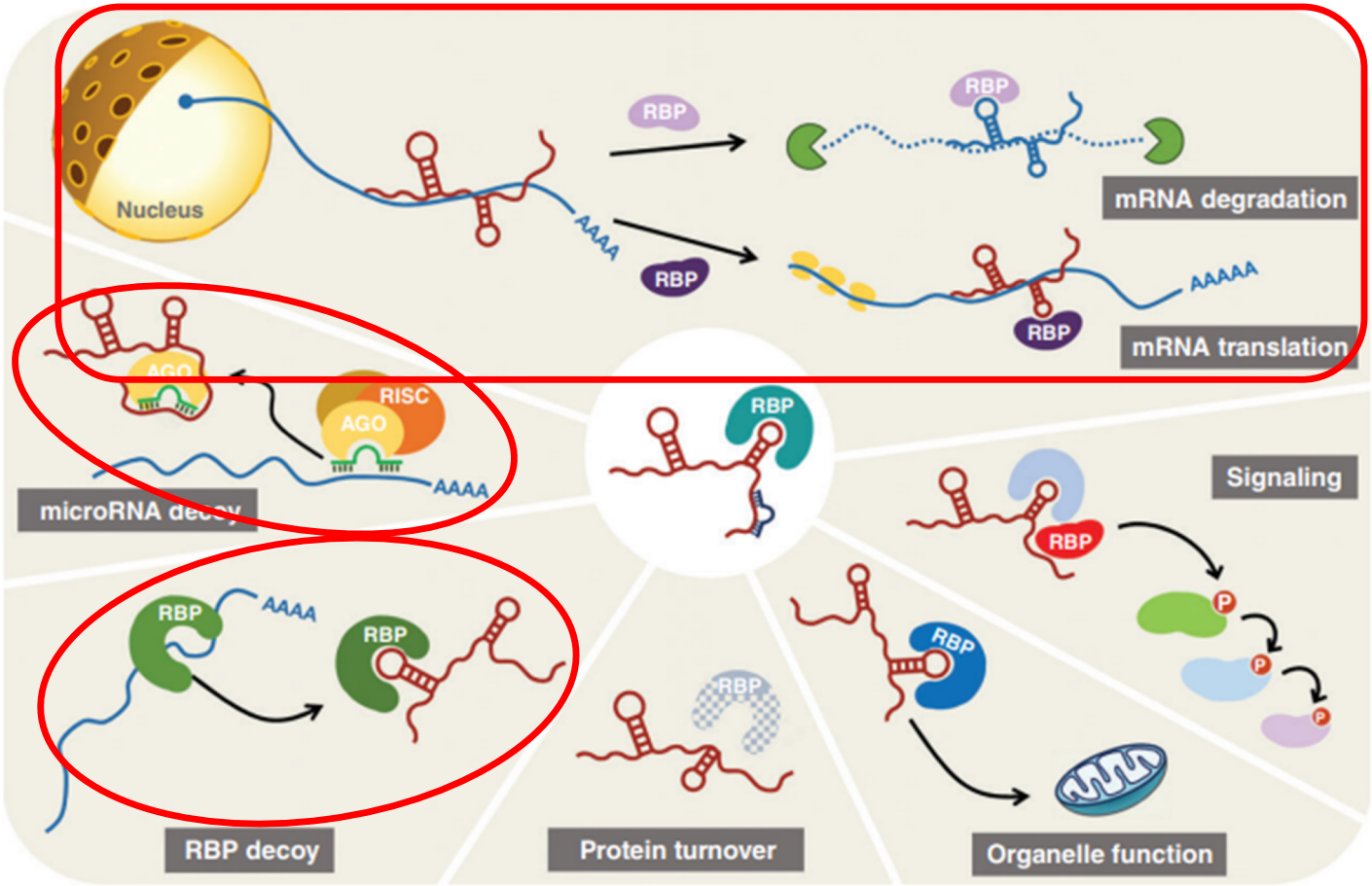
or can sequester miRNA away from target mRNA

LncRNAs modulate RNA metabolism: examples

TINCR – STAT1

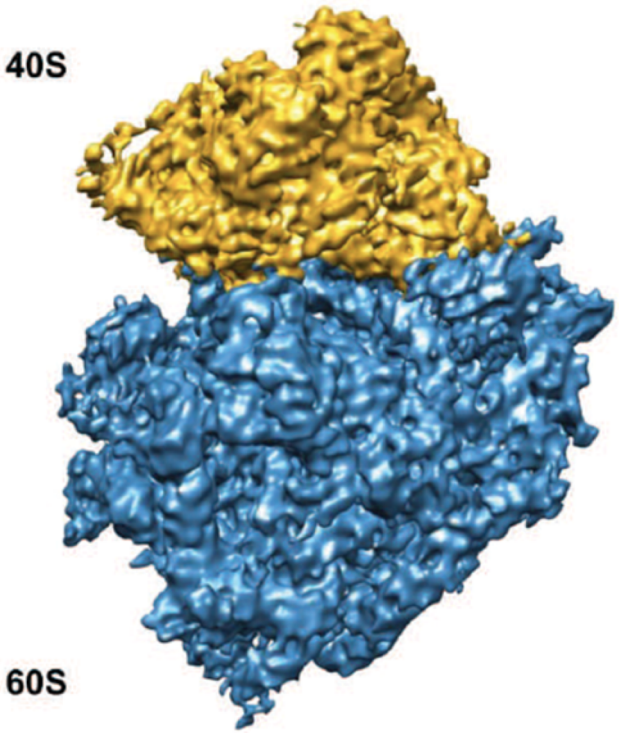
PTENP1

Gadd7/TDP-43



LncRNAs functions depend on RNA physical interactions → studying lncRNAs subcellular localization and its changes is a crucial step toward elucidating functions and mechanisms of newly discovered lncRNAs

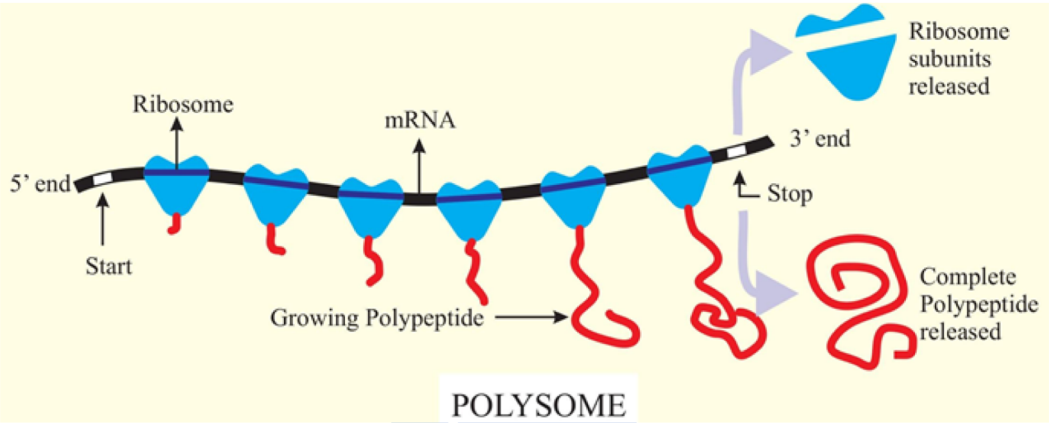
Main characteristics of ribosomes and polysomes



The ribosome is the translational machinery of cells
It is a large riboprotein complex which comprises four ribosomal RNAs and more than 80 proteins

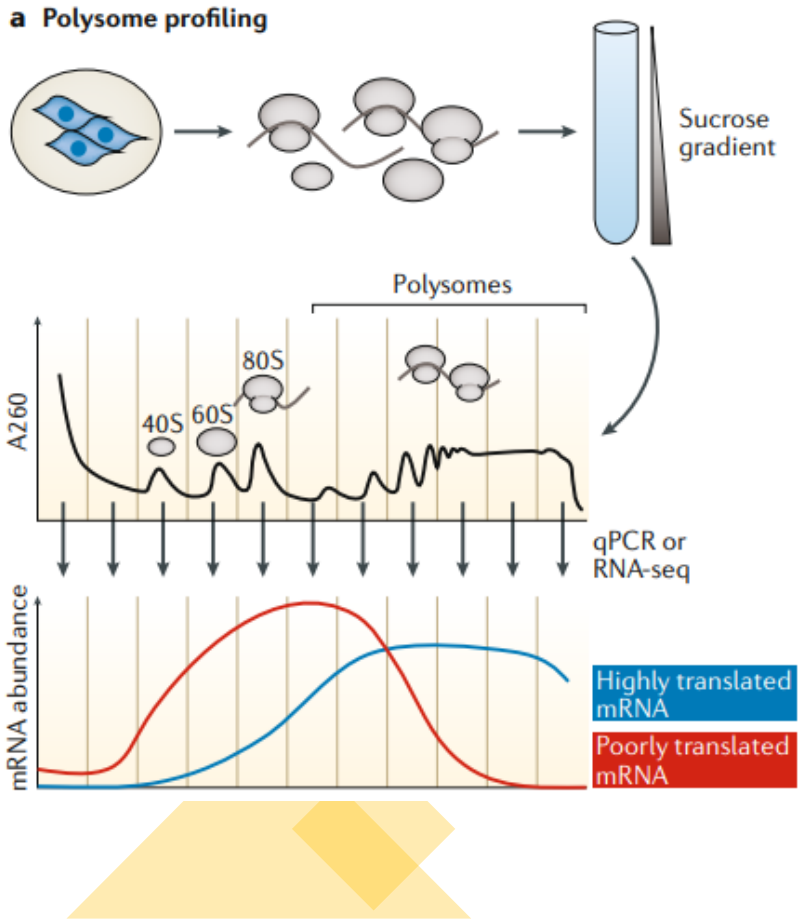
The 80S ribosome has a molecular weight of 4.3 MDa while in bacteria it has a weight of 2.3 MDa

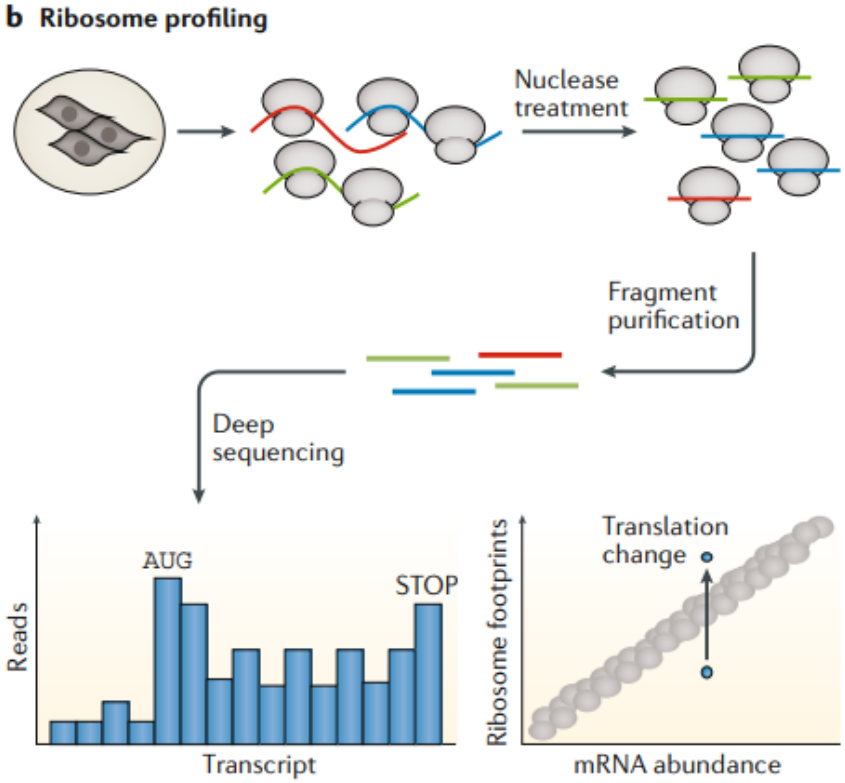
In translation phase multiple ribosome can bind to the same mRNA to form a polysome



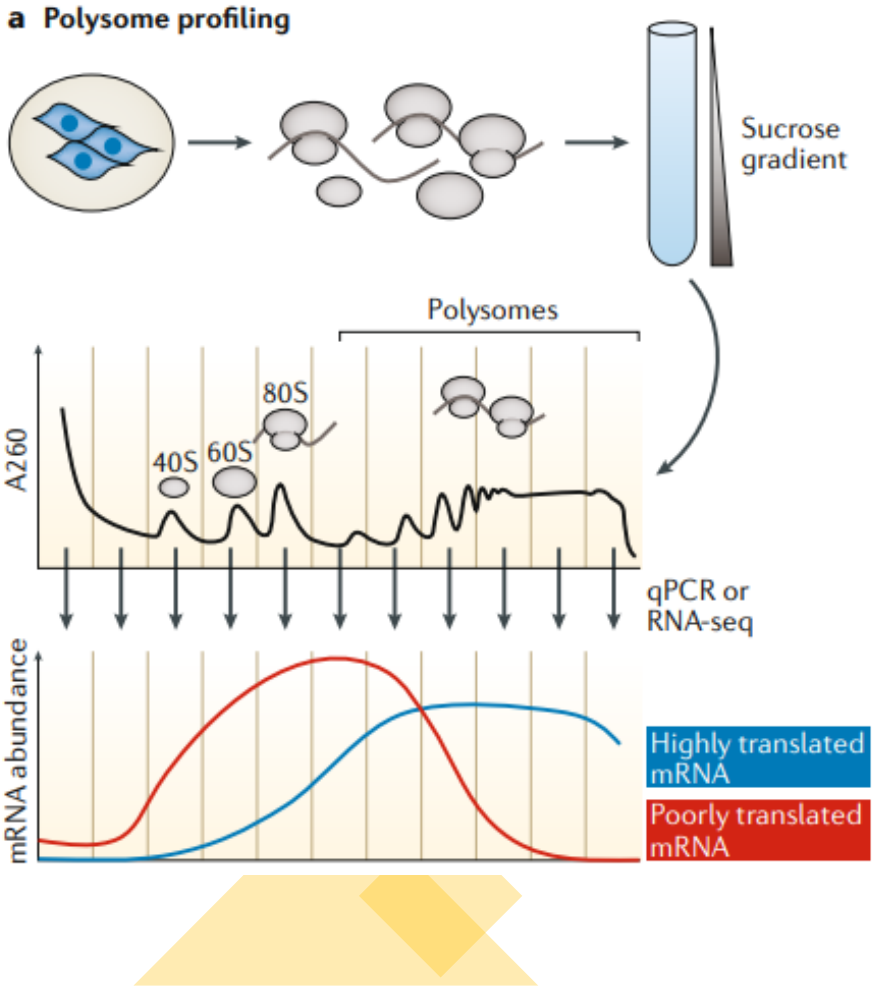
Cytoplasmic lncRNAs bind to ribosomes

Two techniques to analyze the translatome





Polysome profiling: an overview



Steps:

a) Treatment of lysates with the drugs that immobilize elongating ribosomes on mRNA

b) Ultracentrifugation in sucrose gradient

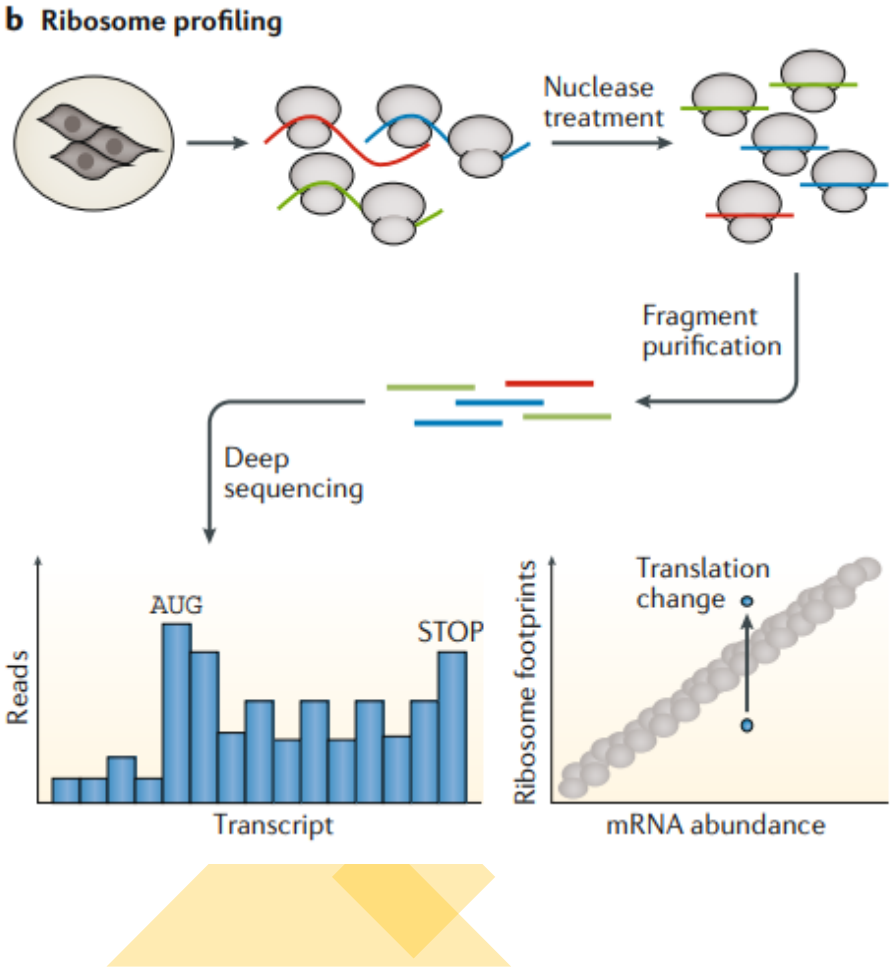
c) RNA is isolated from each fraction

d) qPCR / RNA-seq to determine transcripts abundance

! Limitations:

- Time-consuming (expecially the ultracentrifugation step)
- Multiple steps in the workflow could introduce errors

Ribosome profiling: an overview



Steps:

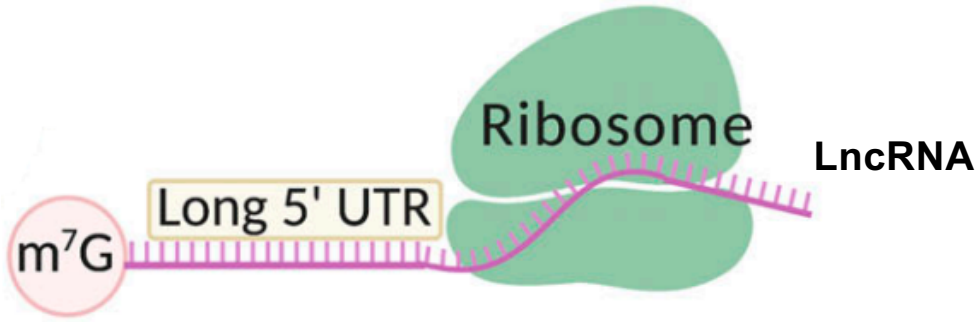
a) Ribosomes are immobilized on mRNA and isolated by density centrifugation
b) Nucleases to digest unprotected mRNA
c) The ~30-nucleotide ribosome-protected mRNA fragments are sequenced
d) Mapping of fragments to the transcriptome

! Limitation:
No information on the number of ribosomes that are present per single lncRNA transcript

First evidence of the association between lncRNA and ribosomes

Independent studies estimated ~20-40% of the cellular lncRNAs to interact with ribosomes using multiple approaches such as the aforementioned ribosome profiling and polysome profiling.

In particular, Van Heesch et al. performed subcellular RNA-seq on nuclei, cytosol and mono- and polyribosomes separated by polysome profiling and the resulting data confirmed that most lncRNAs are strongly enriched in the cytosol and in complexes that contain multiple ribosomes.



Why do lncRNAs associate with ribosomes?

1. These lncRNAs might encode for proteins; → This possibility was eliminated
2. It's possible that one of the processes that keep lncRNAs at ribosomes is non-sense mediated decay (NMD); → Unlikely
3. These lncRNAs might have functional roles in regulating translation:

Most attractive hypothesis →

(1) The lncRNAs occupy the ribosomes to keep them in a poised state until specific stimulatory cues are received

(2) LncRNAs could regulate translation of specific protein-coding transcripts by sequence specific pairing

↓

The antisense lncRNA of Uchl1 has been shown to regulate the association of sense Uchl1 with active polysomes in mice

↓

This regulatory mechanisms has been found also in bacteria

↓

Out of 25 antisense lncRNAs expressed in this data, only 3 had both partners expressed and showed subcellular co-localization

It's unlikely that a similar mechanism is abundant in human cells

PAPER PRESENTATION

Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells

JOANA CARLEVARO-FITA,[1,2,3] ANISA RAHIM,[4] RODERIC GUIGÓ,[1,2,3] LEAH A. VARDY,[4,5] and RORY JOHNSON[1,2,3]

[1]Centre for Genomic Regulation (CRG), 08003 Barcelona, Spain
[2]Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain
[3]Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), 08003 Barcelona, Spain
[4]A*STAR Institute of Medical Biology, Singapore 138648, Singapore
[5]School of Biological Sciences, Nanyang Technological University, 637551 Singapore

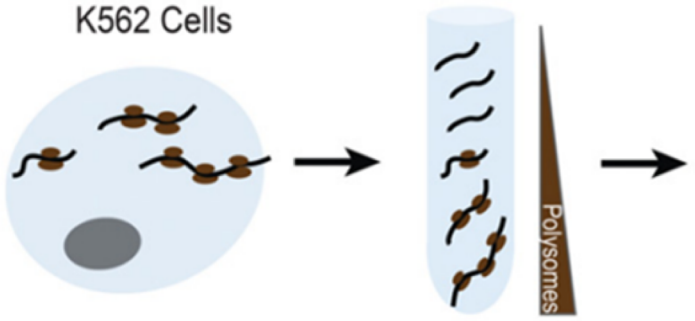
ROADMAP

- Introduction;
- Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
- Additional studies to address lncRNA ribosomal interaction, localization and function;
- Studying the features of ribosomally-associated lncRNA;
- Conclusions.

Creation of a comprehensive and quantitative map of cytoplasmic lncRNA

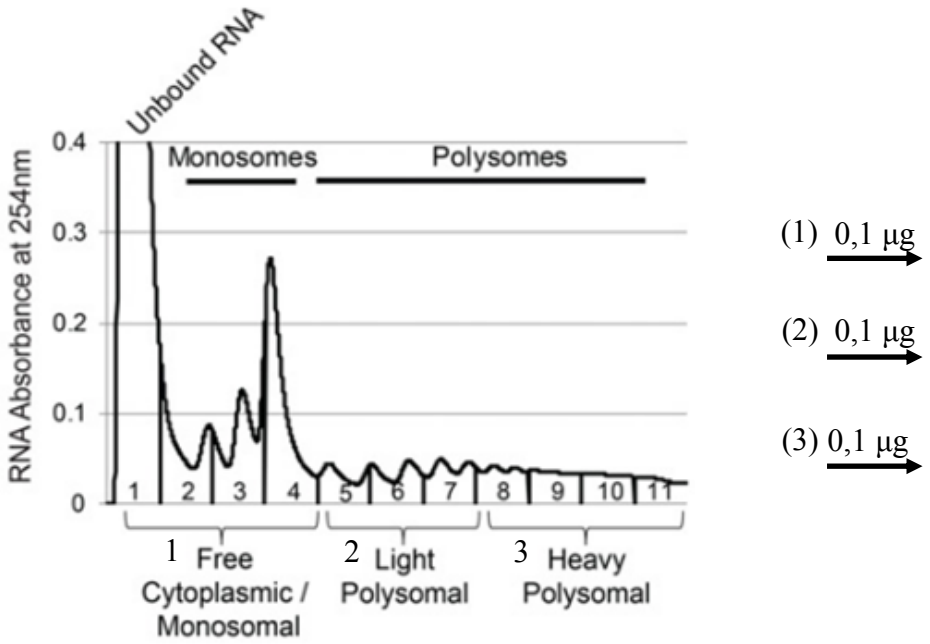
Experimental strategy: polysome profiling

Aim: identify ribosome bound RNAs and distinguish transcripts bound to single or multiple ribosomes



K562 cells were incubated with cycloheximide → cell pellets were lysed → extracts were centrifuged to remove the nuclei → the supernatants were further centrifuged and loaded onto linear sucrose gradient

Twelve fractions were collected from the top of the gradient



(1) Non-translated cytoplasmic RNAs (free mRNA, 40S and 60S ribosomal subunits and mRNA bound by a single ribosome)

(2) Complexes co-fractioning with two to six ribosomes

(3) High molecular weight complexes co-fractioning with more than six ribosomes

Creation of a comprehensive and quantitative map of cytoplasmic lncRNA

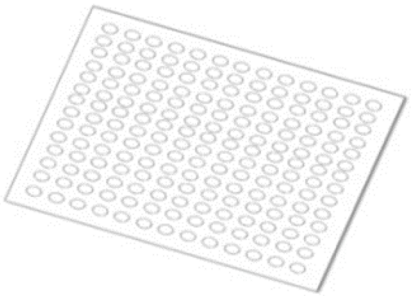
Experimental strategy: microarray hybridization

Aim: estimation of the relative amounts of cytoplasmic lncRNA in the three fractions

(1) 0,1 µg

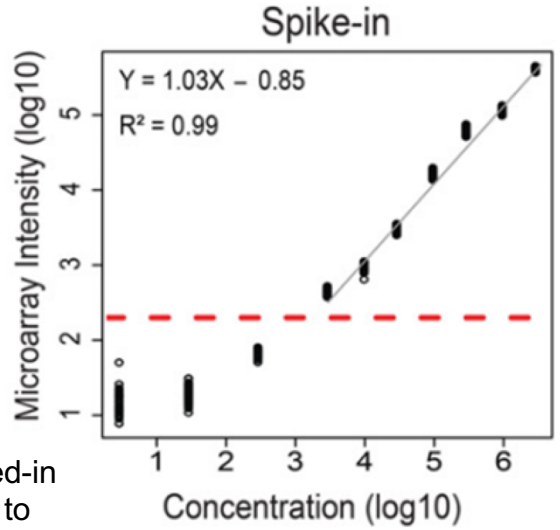
(2) 0,1 µg

(3) 0,1 µg



Custom microarrays probing the entire Gencode v7 lncRNA catalog

14.700 lncRNAs
2.796 mRNAs

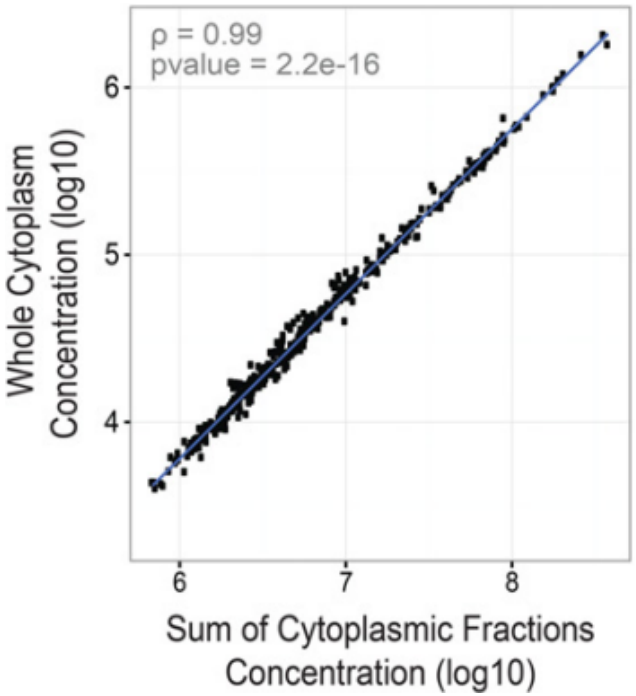


Data were normalized to spiked-in synthetic external RNA added to samples at known concentrations

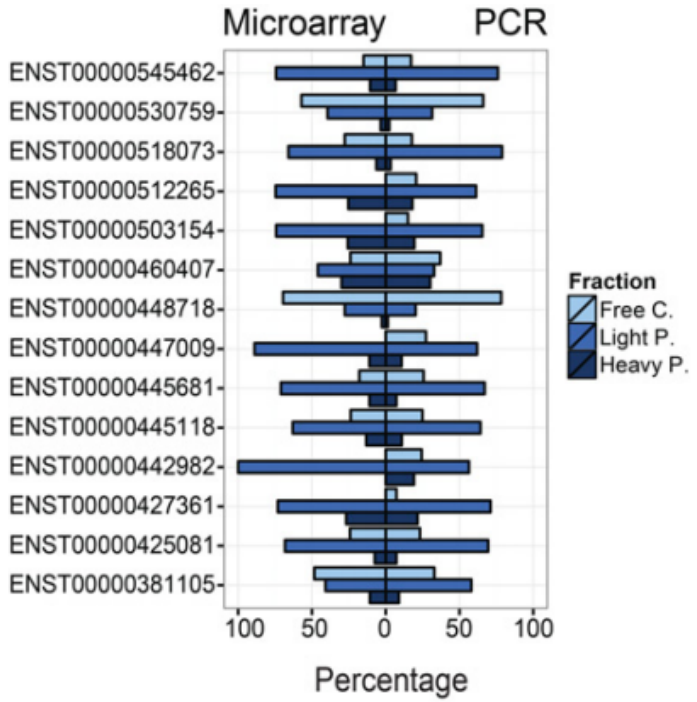
Dashed red line represents the defined detection threshold where regression ceases to be linear, only probes above this threshold were considered detected

LncRNA transcripts and protein-coding genes were considered to be present in a sample when more than half of their probes were detected above the cutoff

How to demonstrate the validity of this approach?



Close correlation between the sum of the estimated concentrations across the three cytoplasmic fractions and the concentration of a separate hybridization of total cytoplasmic RNA from the same cells



Quantitative PCR carried out on the same samples also supported the microarray estimation

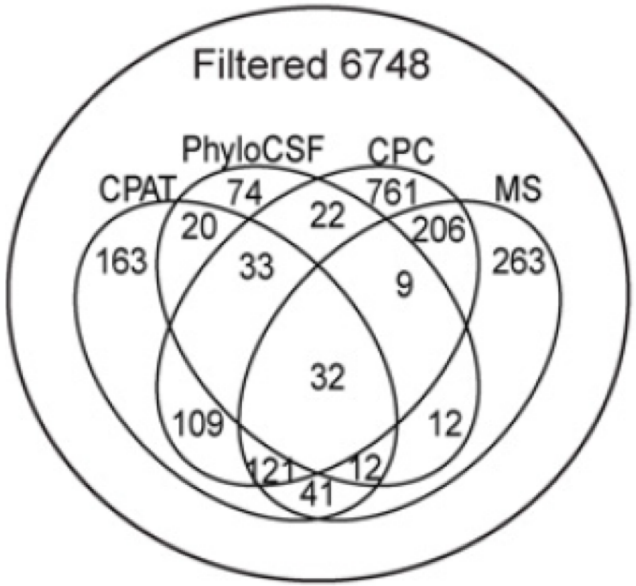
These data support the validity of the use of microarray to estimate the relative concentrations of lncRNA in the three cytoplasmic fractions

Creation of a high confidence lncRNA catalog

A stringent filtering step has been done in order to remove protein-coding transcripts in the Gencode v7 lncRNA catalog :

I. Removal of lncRNA that could be unannotated extensions of protein-coding genes or pseudogenes;

II. Fitration of the remaining genes with different computational methods for identifying protein-coding sequence.



This resulted in a dataset of 13.358 lncRNA transcripts, among which:

- 9.008 are classified as noncoding lncRNA (6.748 genes), also referred to as «filtered lncRNAs»;
- The 1.868 remaining genes (4.350 transcripts) are named «potential protein-coding RNAs»;

Then, considering the analysis of K562 cells extracts via microarray, they detected:

- 345 filtered lncRNAs in the cytoplasm + 292 in the nucleus, the latter based on ENCODE data;
- 755 mRNA in the cytoplasm.

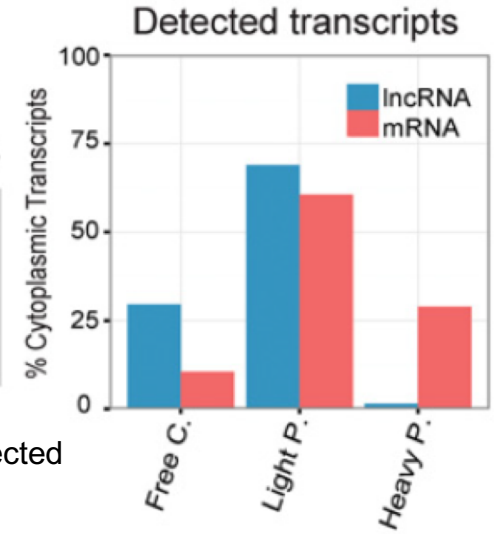
In the picture: CPAT = Coding Potential Assessment Tool; PhyloCSF = a comparative genomics method; CPC = Coding Potential Calculator; MS = Mass Spectrometry

Classification of the cytoplasmic lncRNA found in K562 cells

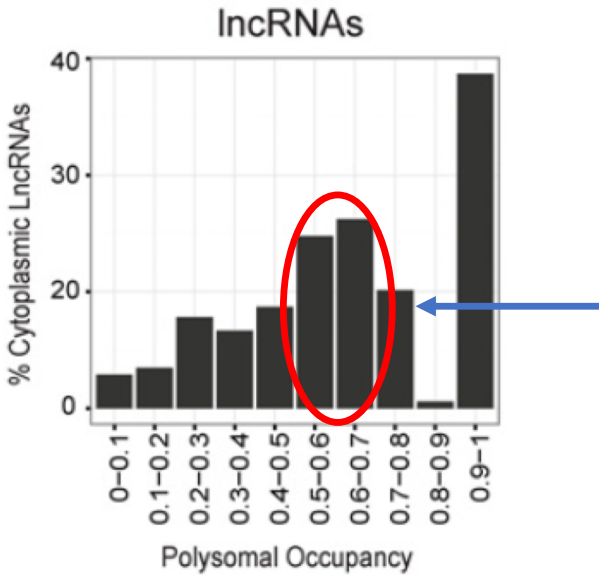
According to their maximal ribosomal association, the 345 cytoplasmic lncRNAs are classified into 3 groups:

- Free cytoplasmic (102, 29%)
- Light polysomal (238, 69%)
- Heavy polysomal (5, 1.4%)

	Transcripts	% Cytoplasmic Transcripts	Genes
Free Cytoplasmic	102	29.6	62
Light Polysome	238	69.0	143
Heavy Polysome	5	1.4	5
Nucleus	292		255
Not Present	8371		6420
Potential Coding	4350		1867

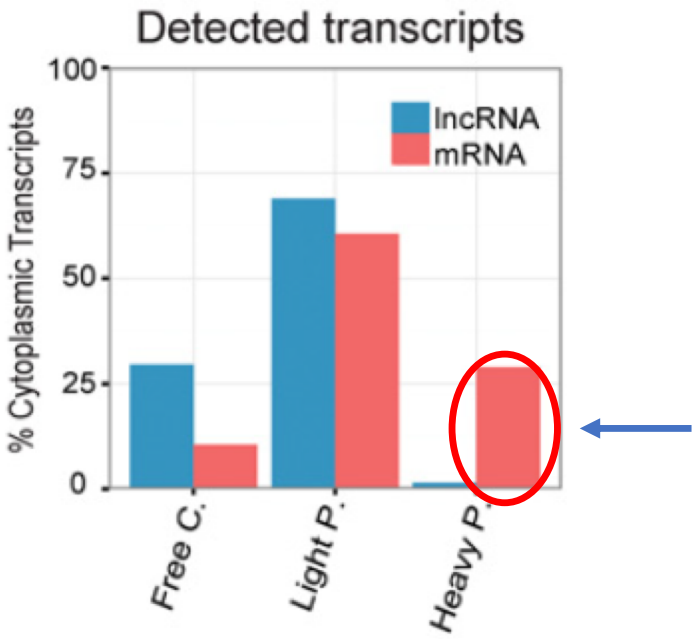


70.4% of lncRNA transcripts are detected in light or heavy polysomal fractions

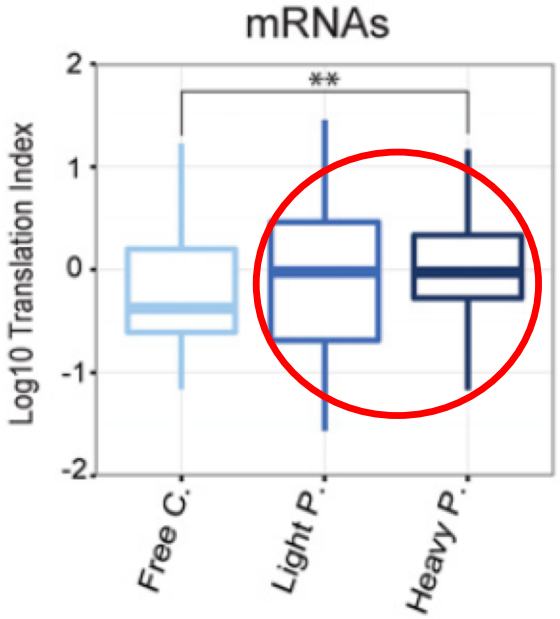


- Polysomal occupancy: the ratio of polysomal (light + heavy fractions) to total cytoplasmic RNA
- LncRNA span the entire range, with peaks between 50% and 60%
- Almost ¼ of lncRNAs examined had > 90% signal detected in polysomal fractions

What evidence support this classification approach?



I. 29% of protein coding mRNAs are classified as heavy polysomal → actively translated



Translation index: ratio of peptide expression to mRNA concentration

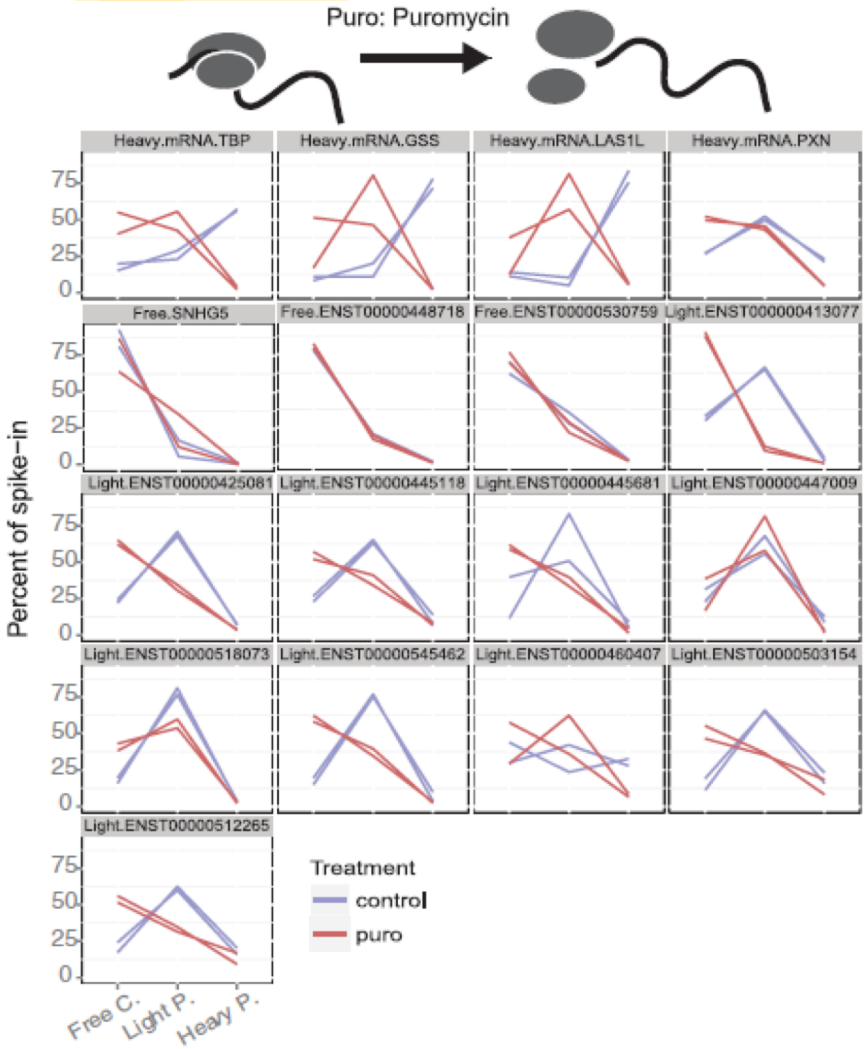
II. Both light + heavy polysomal mRNAs have a high translation index → increased polysomal occupancy of mRNAs

Potential protein-coding transcripts have a ribosome-association profile similar to filtered lncRNAs → they are not efficiently translated → the stringency of lncRNA filtering is confirmed

ROADMAP

- Introduction;
- Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
- Additional studies to address lncRNA ribosomal interaction, localization and function;
- Studying the features of ribosomally-associated lncRNA;
- Conclusions.

How did they exclude false positives?



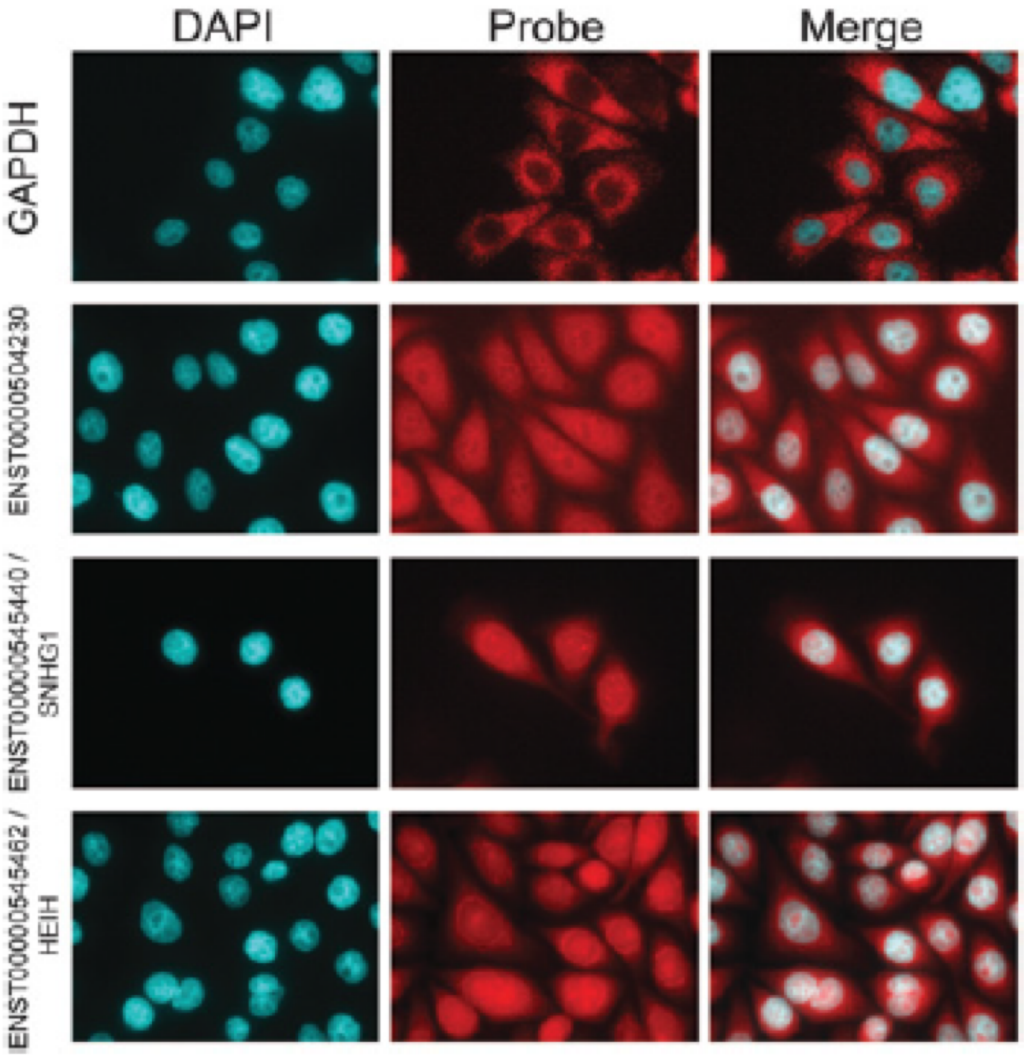
Experimental strategy:
a new polysome profiling experiment on cells treated with puromycin, followed by volume-normalized qRT-PCR.

Results:

- In response to puromycin, ribosome-bound transcripts relocalize to the lighter polysome + free cytoplasmic fractions;
- The free cytoplasmic lncRNAs are unaffected by puromycin treatment.

Conclusions → in the majority of cases, co-sedimentation in polysome profiling reflects a genuine physical interaction between lncRNA and ribosomes.

Additional validation using FISH



Experimental strategy:
Flourescence In Situ Hybridization (FISH) in adherent HeLa cells to visualize the localization of lncRNA at subcellular resolution.

Results:

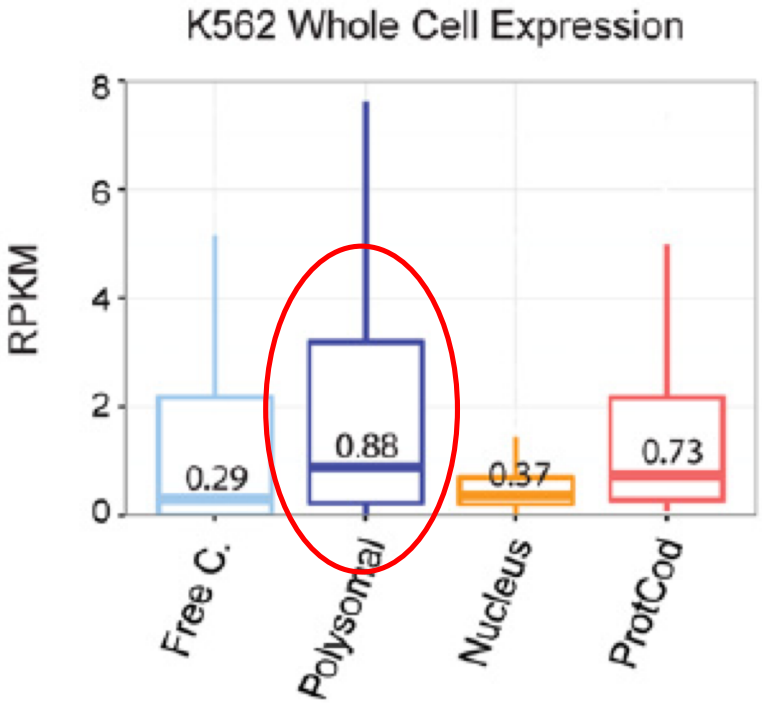
Three lncRNAs, which are expressed and cytoplasmically localized both in K562 and in HeLa, shows diffuse and pronounced cytoplasmic and perinuclear stainings

Conclusions → lncRNAs localization to the cytoplasm (and, possibly, the ER) supports their localization on translating polysomes → microarray data of ribosomal recruitment of lncRNAs are confirmed

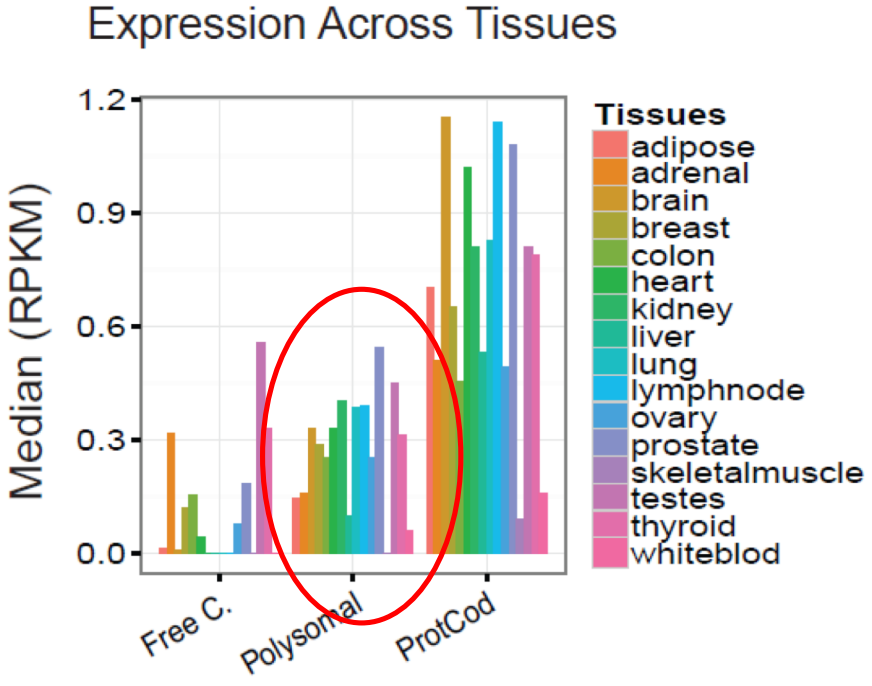
Comparing expression profiles of ribosome-associated and free cytoplasmic lncRNAs

Using K562 RNA-seq data from ENCODE, they observed that:

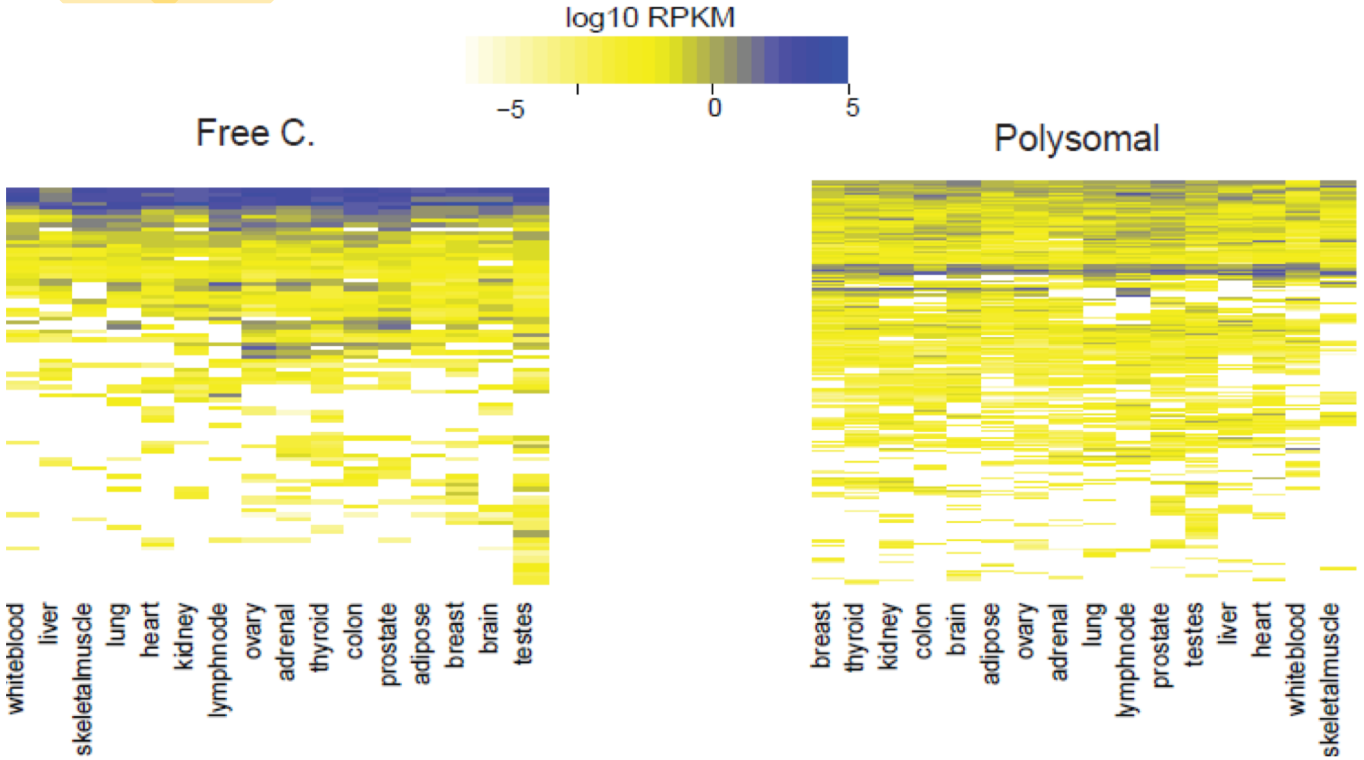
Polysomal lncRNAs have the highest median whole cell expression values, exceeding free cytoplasmic…

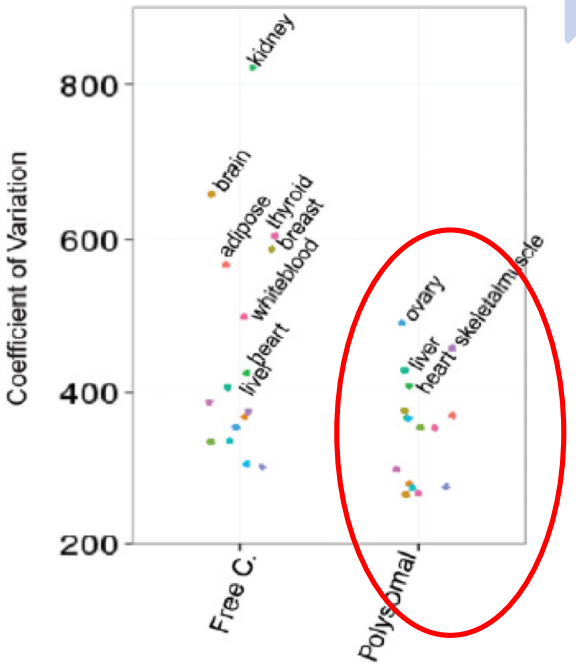


…with a similar trend across tissues



Differences in expression variability between free cytoplasmic and polysomal lncRNAs





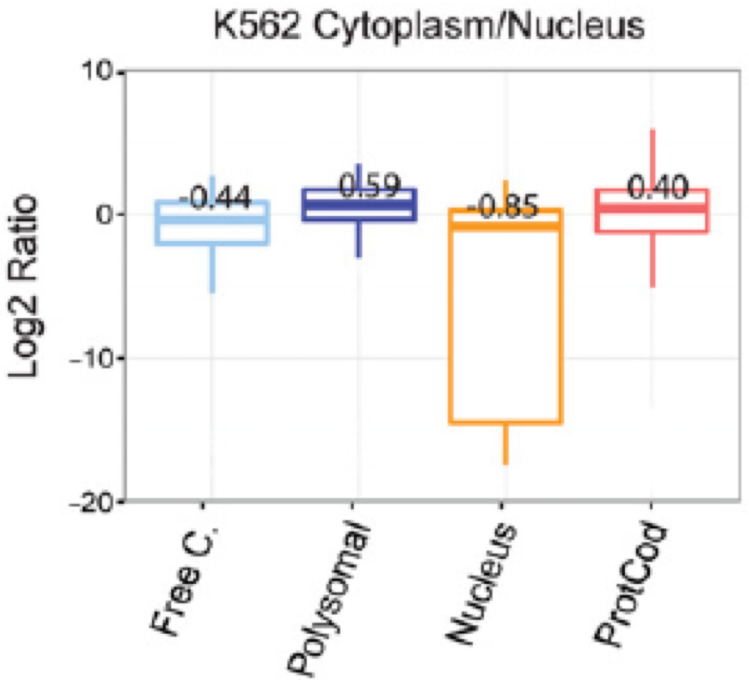
Observations:

- Some free cytoplasmic transcripts achieve higher abundance
- The % of transcripts expressed per tissue is lower than the one of polysomal

Conclusions → polysomal lncRNAs tend to have a lower but more homogeneous expression in human tissues

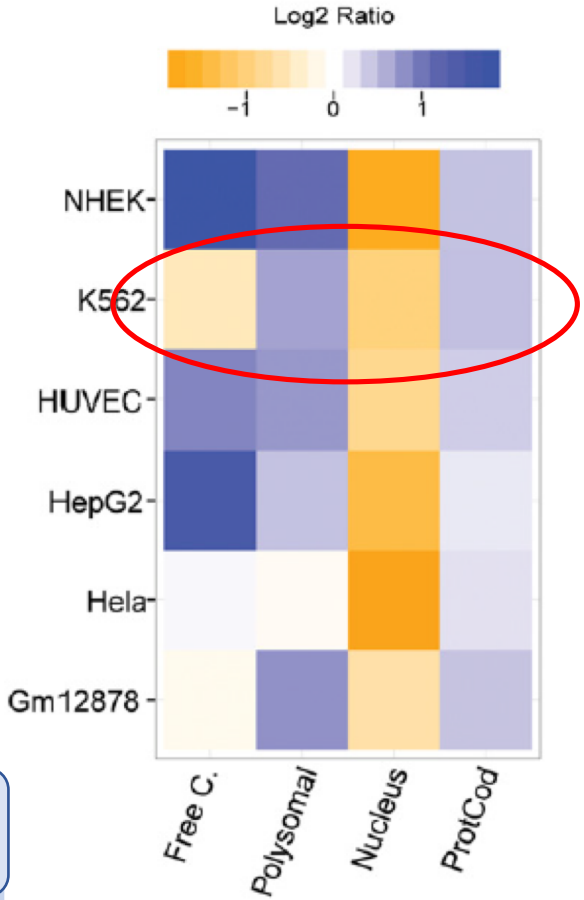
Additional evidences of subcellular localization and expression profiles of lncRNAs

Comparisons between polysome profiling and ENCODE RNA-seq data showed that:



#1 transcripts classified by ENCODE data as polysomal show elevated cytoplasmic-nuclear ratios, exceeding protein-coding mRNAs (ProtCod)

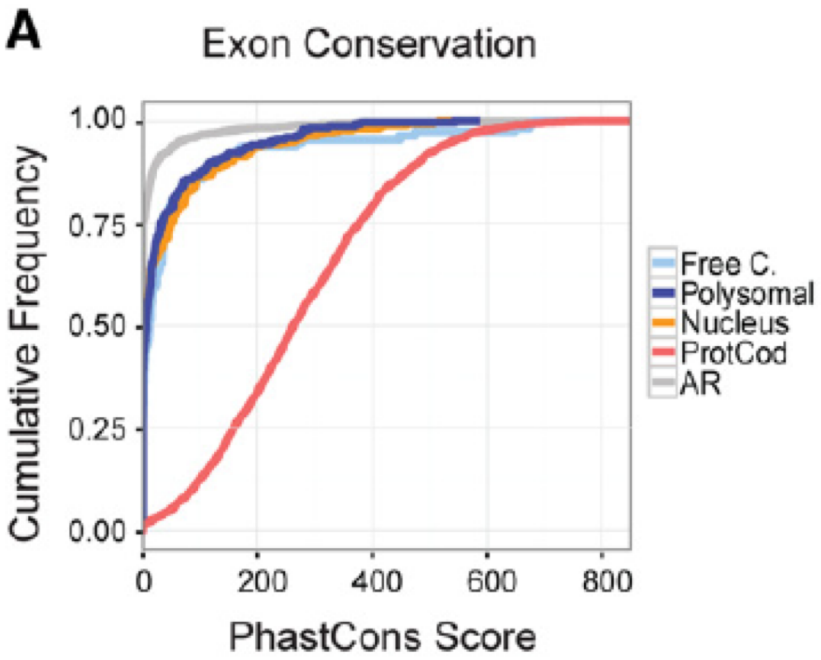
#2 The subcellular localization of lncRNAs observed in K562 cells is maintained across different cell types

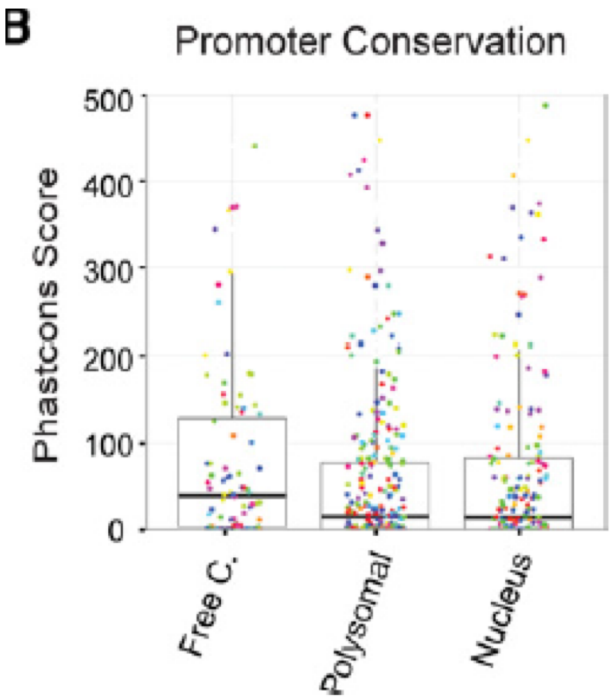


Conclusions → free + polysomal lncRNAs have median cytoplasmic specificity, exceeding that of protein-coding mRNAs

Evidence for conserved function of cytoplasmic lncRNAs

PhastCons measures of exonic and promoters conservation showed that:





The promoters of free cytoplasmic transcripts are more conserved than those of polysomal/nuclear ones

- Protein-coding exons have highly elevated conservation
- Free cytoplasmic + polysomal/nuclear lncRNAs exhibit similar rates of nonneutral evolution

Conclusions → cytoplasmic lncRNAs undergo a weak but nonneutral purifying evolutionary selection

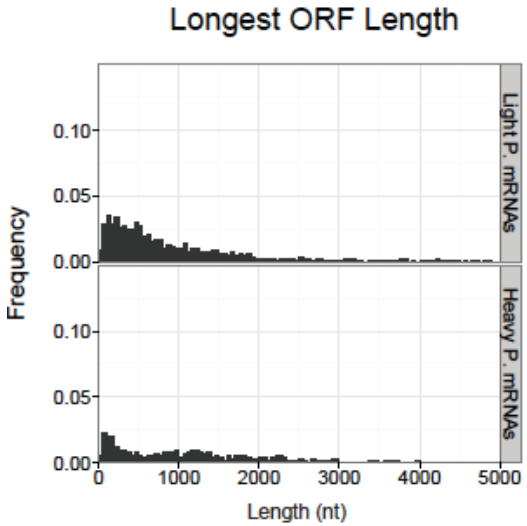
ROADMAP

- Introduction;
- Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
- Additional studies to address lncRNA ribosomal interaction, localization and function;
- Studying the features of ribosomally-associated lncRNA;
- Conclusions.

Ribosomally bound lncRNAs can be defined by mRNA-like 5' regions

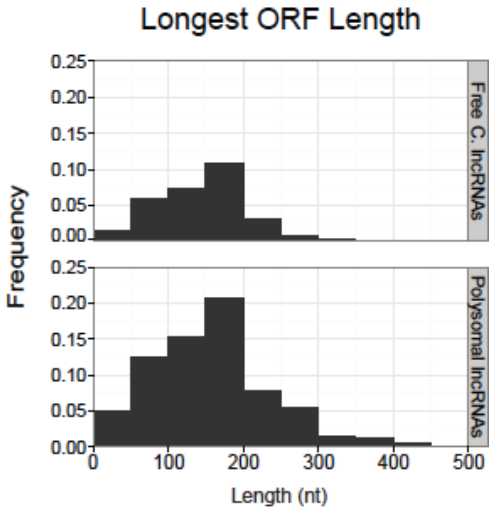
LncRNAs contain presence of small ORFs, that can serve as decoy for ribosomes

In mRNAs, ORF length influences the number of ribosomes that can bind and the ribosomal fraction in which they sediment



Association of ORF coverage with polysome density for mRNA transcripts

Association of ORF coverage with polysome density for lncRNA transcripts



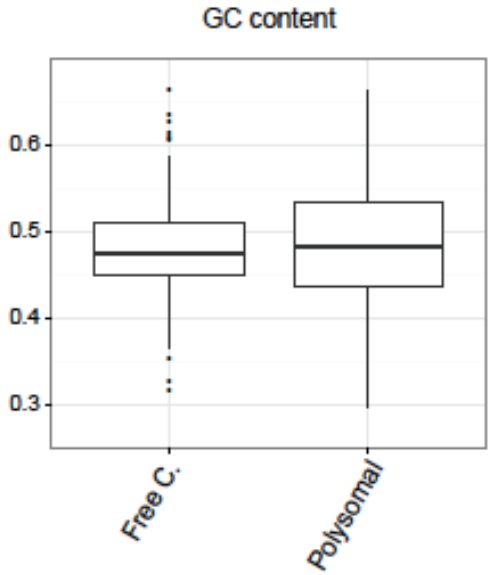
Conclusions → For lncRNAs we could not find evidence that ORF correlates with ribosomal recruitment

Can GC content distinguish free and ribosome-associated lncRNAs?

Experimental strategy:
analysing GC content by previously published data

Results:
usually GC content is a discriminator of coding and non coding RNAs

Hypothesis:
it has been assumed that features that influence mRNA recognition by ribosomes could also apply to lncRNAs



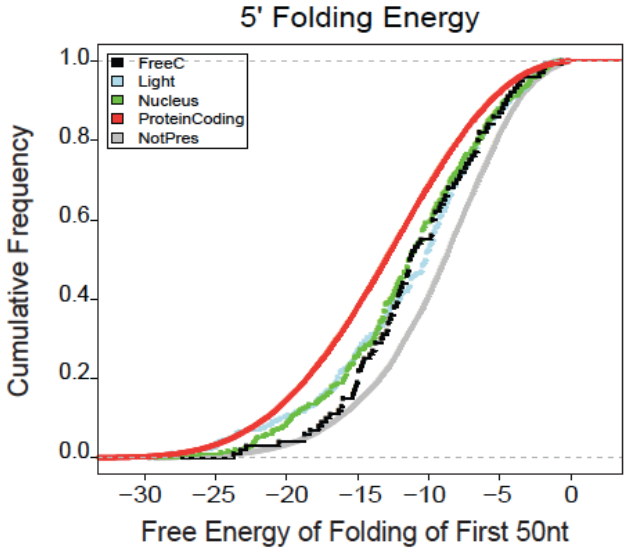
GC content of free cytoplasmic and polysomal lncRNA transcripts

Conclusions → it is not seen a clear disparity in structural propensity between ribosome associated and free cytoplasmic transcripts either

Differences in free energy folding in mRNAs and lncRNAs

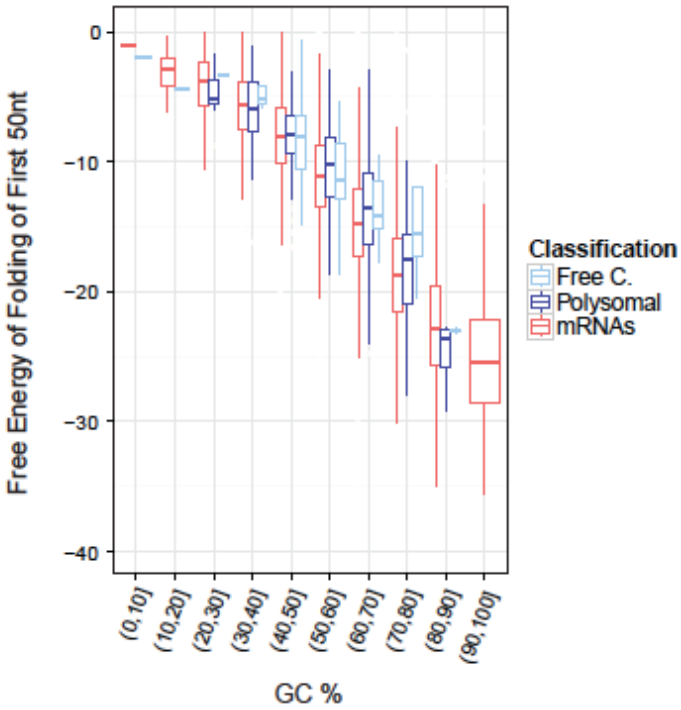
Experimental strategy:
to compare the 5' folding energy it has been used Vienna programme and estimated the free energy folding of 50 nt of lncRNA

Results:
there are differences in the free energy folding of the first 50 nucleotides comparing mRNAs and lncRNAs, these differences disappear when we take into account variation in GC content between mRNAs and lncRNAs



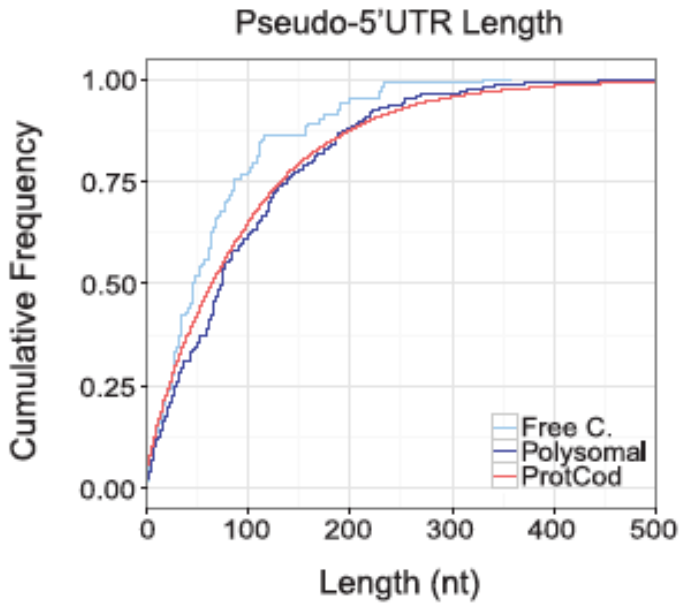
Cumulative frequency comparing free energy of folding of each group, while mRNA have more stable folding on average than expressed lncRNA

Conclusions → GC content may influence propensity of RNA folding



Boxplot showing free energy of folding of mRNAs, polysome-associated and free cytoplasmic lncRNAs divided into bins according to their GC content.

Length of pseudo-5’ UTR as a discriminator of RNAs



Cumulative distribution of pseudo-5’UTR lengths for each set of transcripts

lncRNAs don’t have identifiable ORFs and 5’ UTR, they contain pseudo ORFs

Pseudo-5’ UTR is the region between the transcriptional start site and the first AUG codon of the first ORF

The length of pseudo-5’ UTRs does distinguish ribosome from non-ribosome-associated lncRNAs

Similar to mRNAs polysomal lncRNAs have longer 5’UTR regions than free cytoplasmic and similar to the 5’ UTR for protein-coding genes, this feature may contribute positively to ribosomal recognition of lncRNA

Which is the role of capping in lncRNAs ?

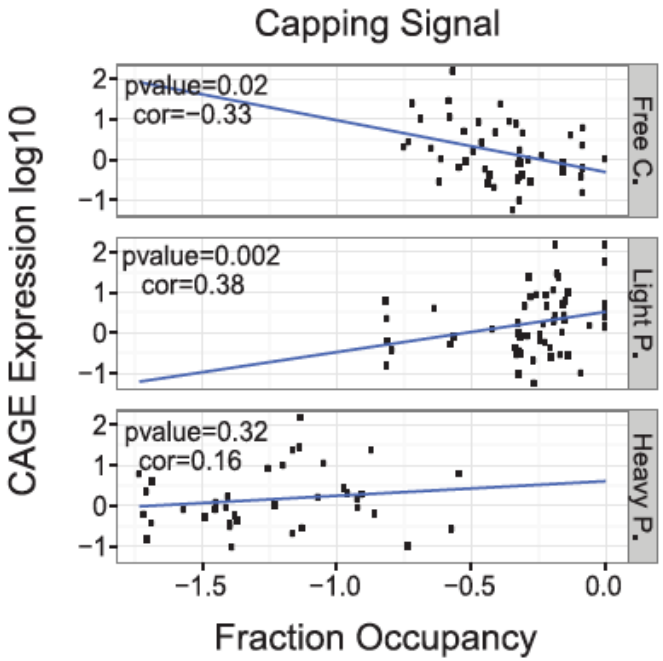
Experimental strategy:

studying the 5’ methyl-guanosine cap with CAGE data

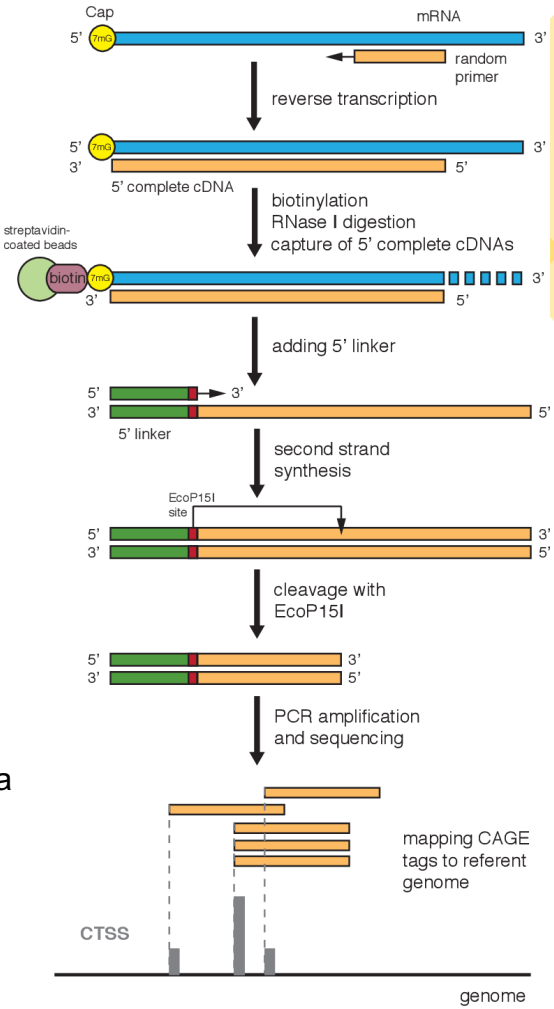
Hypothesis: mRNA require capping recognition for scanning, a similar function can be shared by lncRNAs

Results:

there is a positive relationship between capping and recruitment to light polysomes, on the other hand there is a negative correlation between capping and free cytoplasmic concentration



Capping efficiency calculated by normalizing K562 cells cytoplasmic CAGE tag expression to K562 cytoplasmic expression from RNA-seq data

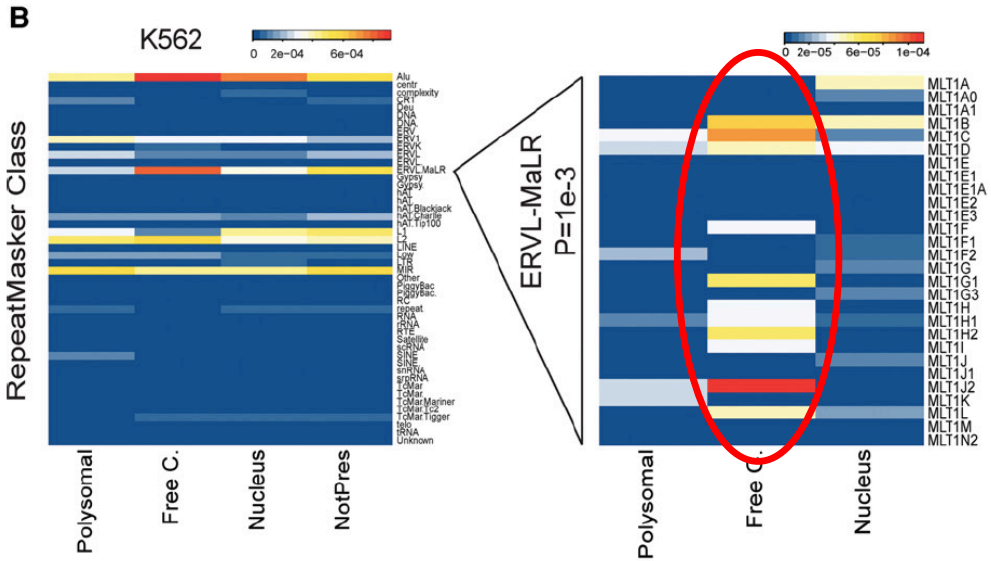


Conclusions → capping of lncRNA is a driver of ribosomal recruitment

Endogenous retroviral fragments are negatively correlated with ribosomal recruitment

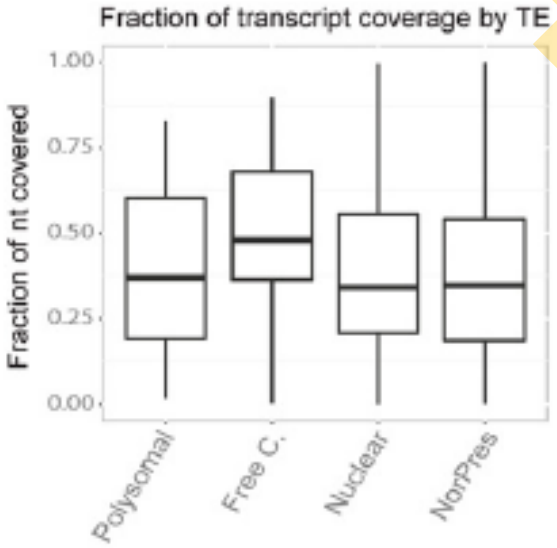
Transposable elements (TEs) contribute functional sequence to lncRNA, in particular there is an excess of TE-derived sequence within free cytoplasmic lncRNAs compared to polysomal

Investigation of TEs correlates with the subcellular localization of host transcript and calculates the insertion frequency of TE classes across lncRNAs



Heat maps showing the mean of the fractional overlap for RepeatMasker-defined classes

Fraction of each transcript covered by annotated transposable elements



Presence of Alu is correlated to elevated transcript expression

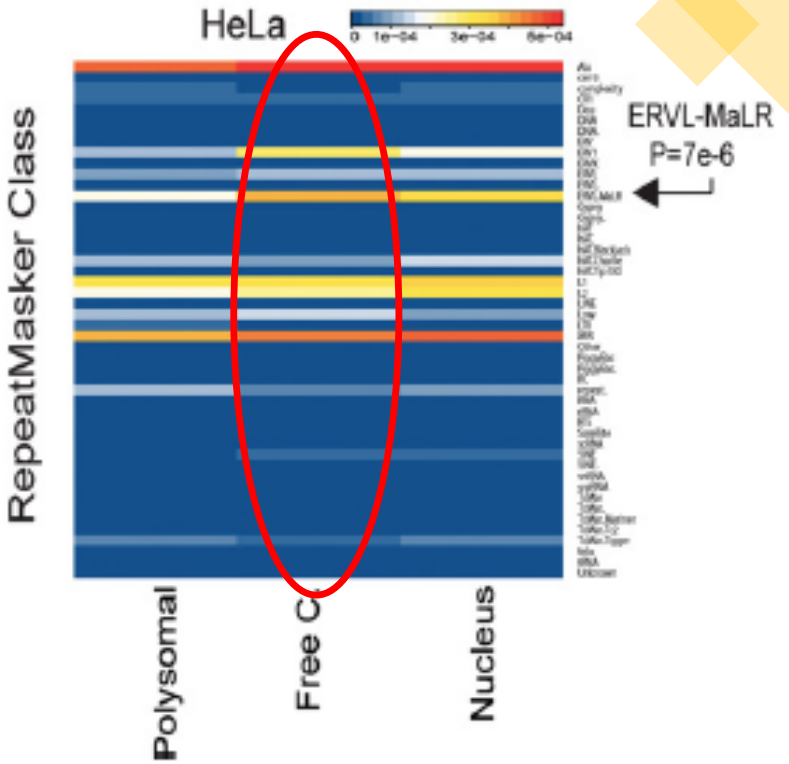
Identification of endogenous retrovirus class ERVL-MaLR, twofold enriched in free cytoplasmic lncRNAs compared to other expressed lncRNAs

Study of the role of ERVL-MaLR

There's no difference in the lenght of ERVL-MaLR insertions between lncRNA classes
Enrichment of ERVL-MaLR class elements in free cytoplasmic lncRNAs is independent of cell type

Experimental strategy:
Using studies of ribosome footprinting data from HeLa cells it has been observed that ERVL-MaLR class of TEs are specifically depleted from ribosome bound lncRNAs

Conclusions → together these data show that ERVL-MaLR fragments can influence lncRNA trafficking in the cell



Heat map showing fractional overlap for RepeatMasker classes in HeLa cells derived from ribosome footprinting experiment

How to study the stability of lncRNAs

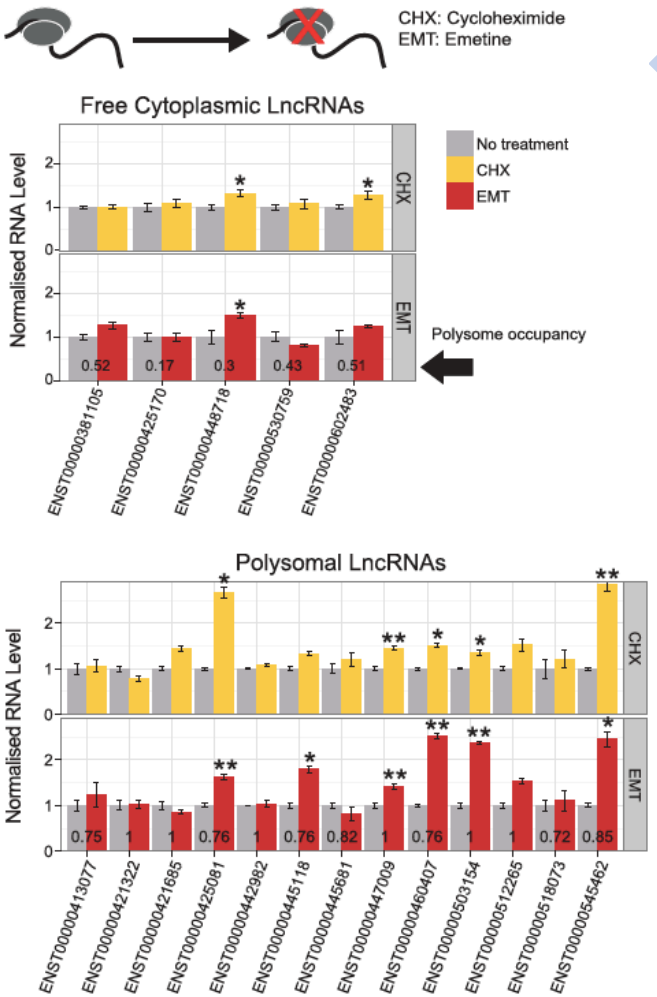
LncRNAs on the ribosomes are subject to degradation by the nonsense-mediated decay (NMD) pathway (data not shown)

Experimental strategy: It was tested if stalling of ribosomal elongation influenced lncRNA stability

Workflow: Two drugs (emetine and cycloheximide), known to be ribosome stalling-drugs, were used to asses stabilization of lncRNAs

Results: Out of six polysomal lncRNAs that respond to EMT, five respond also to CHX, other transcripts seem to be unaffected by the treatments

Conclusions →
Degradation of some cytoplasmic lncRNAs may be triggered by a translation-dependent mechanism



Changes in lncRNA stability in response to drug-induced ribosome stalling.
K262 cells were treated with emetine and cycloheximide to block translation, transcript levels were quantified to assess degradation rate of RNAs

ROADMAP

- Introduction;
- Assessment of the ribosome-associated lncRNA population of stringently filtered lncRNAs;
- Additional studies to address lncRNA ribosomal interaction, localization and function;
- Studying the features of ribosomally-associated lncRNA;
- Conclusions.

Conclusions

- An important population of lncRNAs is present in mammalian cytoplasm, supported by the association of lncRNAs and ribosomes;
- LncRNAs are classified according to their fraction of maximum detection:
- lncRNA of light polysomal fraction have mRNA 5' features: long pseudo-5' UTR and a cap structure;
- GC content or ORFs do not influence ribosomal interaction;
- repetitive sequence features negatively correlate with ribosomal recruitment.
- Stabilization of cytoplasmic lncRNAs in response to translation inhibitors cycloheximide and emetine as like as enrichment for mRNA-like 5' of light polysomal lncRNAs are important features for ribosomal engagement;
- Low association between lncRNA and heavy polysomes can be seen, further more there is no correlation between capping and heavy polysomal recruitment, while light polysomal lncRNAs show direct engaging by ribosomes;
- The RNA degradation-promoting activity of the ribosome can have a crucial role as the final destination of cytoplasmic lncRNAs involved in lncRNA lifecycle