Outlining the maturational pathway of mitochondrial short non-coding RNAs using *in silico* and *in vitro* approaches

Alessandro Formaggioni*, Federico Plazzi, Marco Passamonti

University of Bologna. Department of Biological, Geological and Environmental Sciences. Bologna.

alessand.formaggion2@unibo.it ale97_28

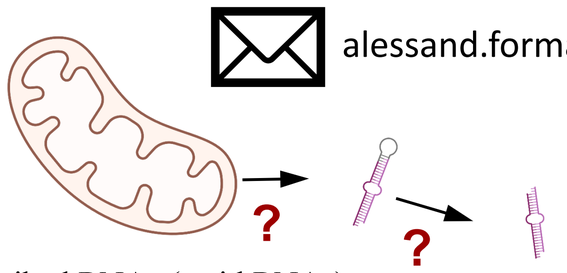
PRIN 2020

Donazione Canziani

More at:

References

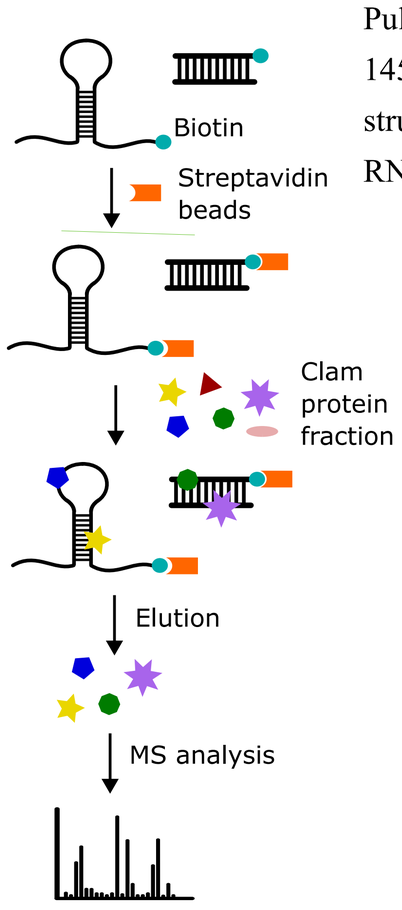
1. Passamonti et al.; 2020; *Sci. Rep.*;10(1):8219
2. Farberov et al.; 2023; *iScience,* 26(10):107723
3. Pozzi et al.; 2020; *GBE*; 14(2):evac023



Intro

The Small Mitochondrial Highly Transcribed RNAs (smithRNAs) are a new class of small interfering RNAs transcribed in the mt-genome regulating nuclear targets. Their effects were *in vivo* validated in the Manila clam, *Ruditapes philippinarum*. Moreover, *in silico* analysis predicted that smitRNAs are widespread across metazoans[1]. The biogenesis of smithRNAs is not well defined yet. We analyzed publicly available CLIP-seq data to detect which proteins may interact with smithRNAs. Moreover, synthetized smithRNA were pulled down with *R. philippinarum* proteins.

In vitro analysis using RNA-protein pull-down



Pull-down analysis with biotinylated smithRNAs (122nca and 145t), one nuclear miRNA (let7), and their preliminary structures (pre_), to identify which proteins interact with these RNAs in *R. philippinarum*.

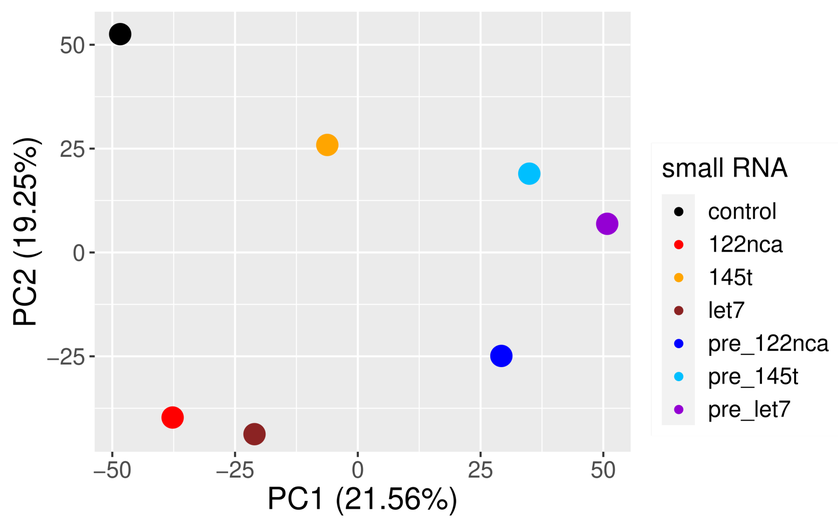


Figure 1: The PCA is calculated based on the intensity of the proteins detected by the mass spectrometer in each sample. Preliminary RNAs (pre_122nca, pre_145t, pre_let7) cluster together, as well as two mature small RNAs (122nca, let7).

We selected proteins that were detected at two-fold higher levels in a pulled-down sample compared to the control sample. These proteins are significantly associated with specific GO terms: let-7 and 145t interact with isomerases (cycliophilin and rotamase).

Tab. 1A: : GO terms enriched in each small RNA with the respective p-value. 'X' when the GO term is not significant.

Description	let7	145t	122nca
Cycliophilin type peptidyl-prolyl cis-trans isomerase	5.03E-07	2.00E-06	X
RNA recognition motif	X	8.46E-08	X
mRNA processing	X	X	1.52E-02
Rotamase	4.51E-06	9.55E-06	X

Preliminary structures were more associated with spliceosome activity, in particular in pre_145t and pre_122nca:

Tab. 1B: : GO terms enriched in each small RNA with the respective p-value. 'X' when the GO term is not significant.

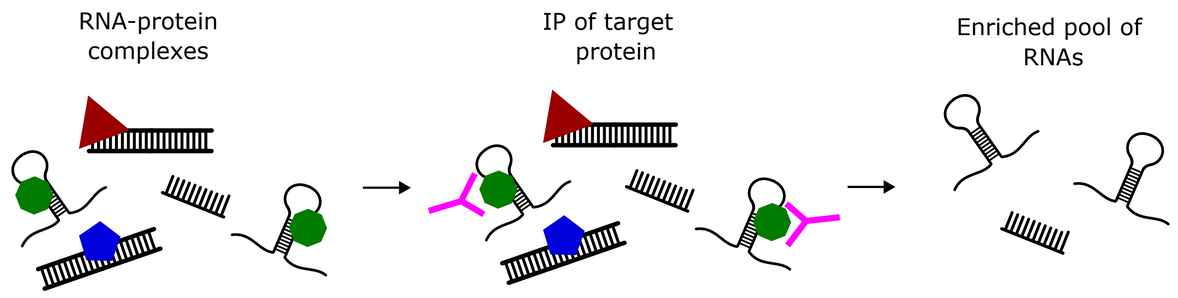
Description	pre_let7	pre_145t	pre_122nca
RNA recognition motif	3.82E-07	6.60E-07	4.34E+07
mRNA splicing	X	7.05-E07	5.58E-07
Nucleotide-binding alpha-beta plait domain	2.39E-06	3.89E-07	2.52E-07
Spliceosome	X	1.04-E07	4.61E-07

Many GO terms were associated with splicesome activity (cycliophylin, mRNA processing, mRNA splicing, splicing). miRNAs were reported to have a role in the modulation of splicing[2]

Conclusions

In silico and *in vitro* data suggest that smithRNAs could interact with proteins belonging to the microRNA pathway. In particular, we showed that AGO2 preferentially interacts with tRNAs (where most smithRNAs are located[1,3]) in the mt-genome of three different species. Mt-tRNAs were also predicted to interact with the Microprocessor complex (DROSHA+DGCR8). *In vitro* analysis showed some similarities between the nuclear miRNA let7 and the two smithRNAs, with the splicesome involved in their maturation.

In silico analysis using CLIP-seq libraries



In vertebrates: DROSHA+DGCR8 and AGO2 participate respectively to the first and last step of the miRNA pathway. Analyzing CLIP-seq libraries of these proteins, all three proteins were predicted to interact with mt-RNAs (Fig. 2). The most enriched mt-RNAs are all located in tRNA regions (tRNA-Val, tRNA-Phe, tRNA-Met), this signal is conserved in homologous regions in *M. musculus*.

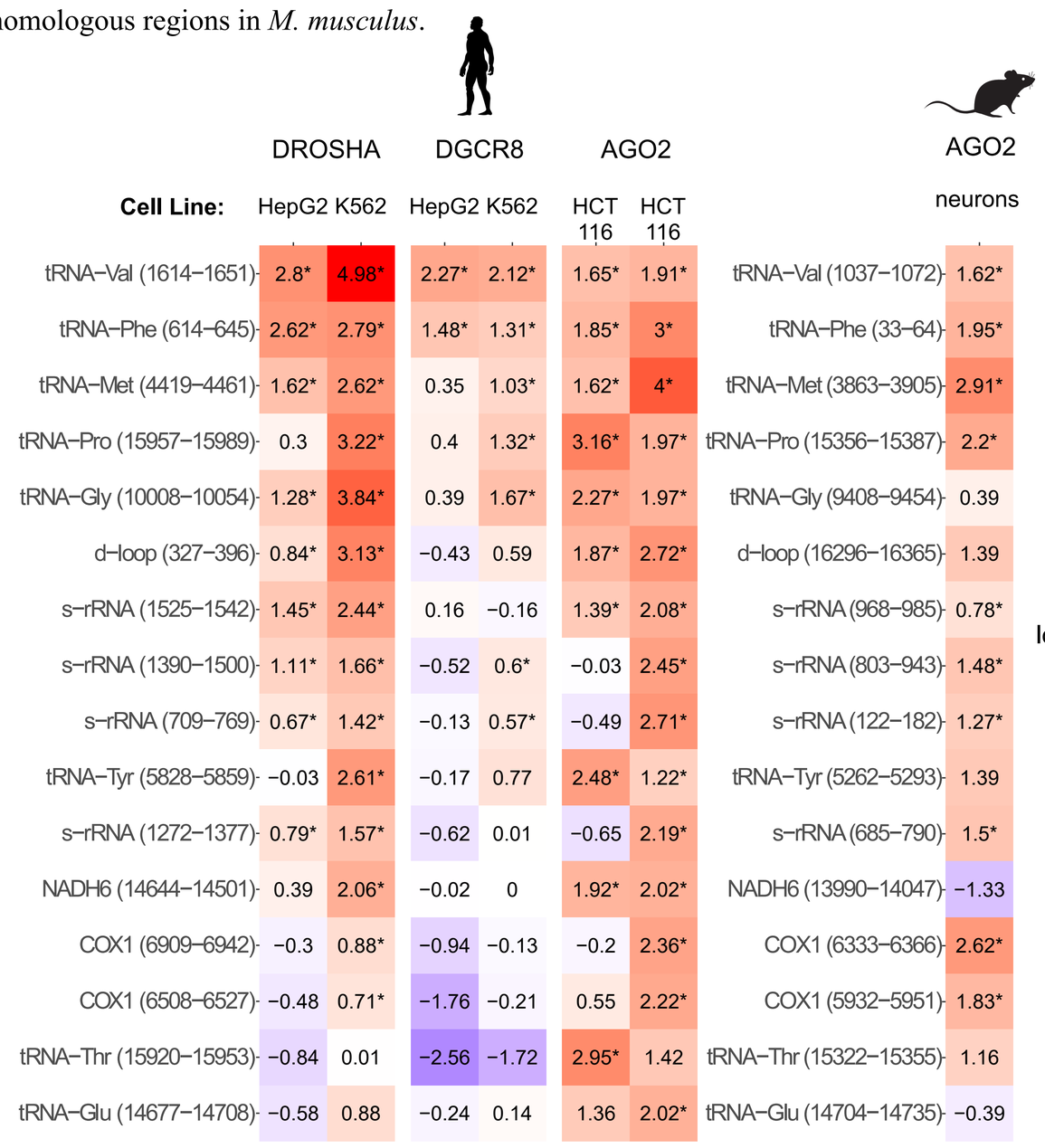


Figure 2: all the mt-regions that showed a logFC (log_2(n° mapping reads IP library / n° mapping reads control library)) > 2 in at least one of the samples were reported in the figure. Boxes are colored in shades of reds when logFC > 0 and in shades of blue when logFC < 0. Significant logFC values are marked with *.

In *C. elegans*: we analyzed the RIP-seq libraries of 4 argonaute proteins (CSR1, WAGO1, HRDE1, ALG1). Only CSR1 was predicted to ineract with mt-RNAs (Tab. 2A), which are located in coding or unassigned regions (UR).

Tab. 2A: significantly DE mt-small RNAs in CSR-1 IP libraries vs control libraries.

Region	logFC	logCPM	P-value
UR:13315-13358	3.87	11.90	1.94E-04
CYTB:4882-4926	3.79	8.28	2.67E-05
CYTB:5250-5291	2.61	6.26	6.00E-04
CYTB:5293-5340	3.57	6.24	5.13E-05
NADH5:12350-12390	1.94	4.56	1.38E-02
NADH1:1794-1843	3.39	4.59	8.84E-05

In *D. melanogaster:* we analyzed RIP-seq libraries of 5 argonaute proteins (AGO1-3, PIWI and Aubergine). Singificanlty enriched mt-regions were found only in the AGO2-IP libraries. Each region is located in a different tRNA. tRNA-Met is by far the most enriched (Tab. 2B).

Tab. 2B: significantly DE mt-small RNAs in AGO2 IP libraries vs control libraries.

Region	logFC	logCPM	P-value
tRNA-Met:167-187	9.95	9.21	2.20E-09
tRNA-Leu1:3009-3032	1.22	5.30	4.30E-03
tRNA-Ala:5982-6001	2.00	5.05	1.59E-04
tRNA-Asn:6124-6151	1.03	6.02	3.52E-03
tRNA-Ser:6190-6218	2.38	7.03	4.29E-05
tRNA-Thr:9936-9968	1.66	4.32	2.40E-03
tRNA-Leu2:12712-12734	0.80	5.39	2.60E-03