bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

New observations on non-coding RNAs involved in the dual translation system in zebrafish development.

Timo M. Breit[1]*, Johanna F. B. Pagano[1], Pjotr L. van der Jagt[2], Ellis Mittring[2], Wim A. Ensink[1], Marina van Olst[1], Selina van Leeuwen[1], Wim de Leeuw[1], Ulrike Nehrdich[3], Herman P. Spaink[3], Han Rauwerda[1], and Rob J. Dekker[1]

[1] RNA Biology & Applied Bioinformatics research group, Swammerdam Institute for Life Sciences, Faculty of Science, University of Amsterdam, Amsterdam 1090 GE, the Netherlands

[2] BSc student internship, Biomedical Sciences, University of Amsterdam

[3] Department of Molecular Cell Biology, Institute of Biology, Leiden University, Gorlaeus Laboratories - Cell Observatory, Leiden 2333 CE, the Netherlands

* To whom correspondence should be addressed. Tel: +31 20 5257058; Fax: +31 20 5257762; Email: t.m.breit@uva.nl

Abstract

Cellular translation relies heavily on the involvements of several types of non-coding RNAs. In previous studies we have identified a dual translation system in zebrafish development, involving maternal-type and somatic-type rRNAs, snoRNAs, and snRNAs. In this study we focused on several remaining non-coding RNAs involved in the translation system; tRNAs, RNase P, and SRP RNA. Even though our studies have been limited in extent, for all three types of non-coding RNA we were able to identify a maternal-specific type, with substantial sequence differences as compared to the somatic-type variant. Hence, these RNA types complement the previously discovered RNA types in the unique dual translation system in zebrafish development.

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

The oocyte is an intriguing and unique cell type for which the translation system functions quite differently as compared to somatic cells. Whereas in somatic cells translation generally quickly follows transcription, in oocytes mRNA is accumulated for later use in (early) embryogenesis. During this waiting period, these maternal mRNAs need to be silenced in order to maintain a normal cellular metabolism in oocytes (Tadros and Lipshitz 2005). Besides the production of mRNA from a specific set of maternal genes, mRNAs in mature oocytes are also atypical as they have no, or very short poly(A) tails (Graindorge et al. 2006; Weill et al. 2012; Subtelny et al. 2014), which likely play a role in the mRNA storage process.

Another distinction for oocyte RNA was discovered decades ago in Xenopus, where the expressed variant of 5S rRNA in oocytes differs from that in somatic cells. This maternal-type 5S RNA variant is gradually replaced throughout embryonic development by a somatic-type variant (Wegnez et al. 1972; Brown et al. 1977; Wormington and Brown 1983; Guinta et al. 1986; Komiya et al. 1986). Starting from this long-standing observation, our continuous studies reported in several manuscripts (Locati et al. 2017a, 2017b, 2018; Pagano et al. 2019b, 2019a) show the existence of a specialized maternal-type translation system in zebrafish far beyond just 5S rRNA. One by one, we have revealed that also the translation-involved non-coding RNAs 5.8S, 18S, 28S, snoRNA, and snRNA are implicated in this maternal-type translation system (Figure 1 and Table 1). The translation system however, has several other non-coding RNA components (Figure 1), i.e. miRNAs, tRNAs, RNase P, and SRP RNA. Here we report on several additional observations with respect to the latter non-coding RNAs, with the exception of miRNA, and their possible involvement in the maternal-specific translation system in zebrafish.

The here presented observations originate from initial analyses and preliminary results of available data and/or new experiments. Research funding constraints only allowed for limited investigation of these translation-involved non-coding RNA types. However, since these observations nicely complete the discovery of a dual translation system in zebrafish, we felt that this warranted sharing our preliminary findings on the three “missing” non-coding RNA types in this manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

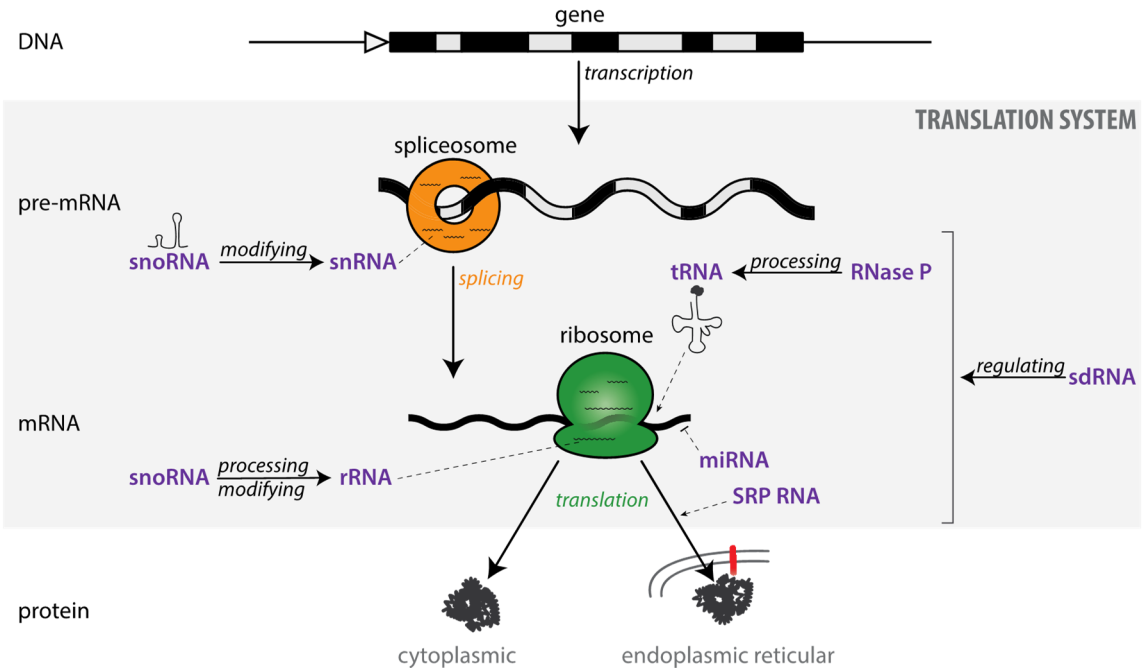

Figure 1. The translation system in the context of the central dogma.
Schematic representation of the translation system in the central dogma of molecular biology, with a focus on the main involved non coding RNA types (purple)

Table1. Main RNA types involved in the translation system

	Sub-member	Size (~)	RNA polymerase	Primary target	Main process
RNA processing & modifying					
snoRNA	H/ACA box	91 nt	Pol II	RNA	RNA pseudouridylation
	C/D box	141 nt	Pol II	RNA	RNA methylation
	U3	215 nt	Pol II or III	45S	cleaving
	U8	131 nt	Pol II or III	45S	cleaving
	snoZ30	104 nt	Pol II	U6	RNA methylation
snRNA	U1	164 nt	Pol II	pre-mRNA	splicing
	U2	191 nt	Pol II	pre-mRNA	splicing
	U4	141 nt	Pol II	pre-mRNA	splicing
	U5	118 nt	Pol II	pre-mRNA	splicing
	U6	107 nt	Pol III	pre-mRNA	splicing
RNase P		300 nt	Pol III	tRNA, lncRNA, rRNA, mRNA	cleaving
Primary translation process					
rRNA	5S	120 nt	Pol III	mRNA	translation
	5.8S	160 nt	Pol I	mRNA	translation
	18S	1.900 nt	Pol I	mRNA	translation
	28S	4.200 nt	Pol I	mRNA	translation
tRNA	61 anticodons	80 nt	Pol III	mRNA	translation
Coding					
mRNA	many	variable	Pol II		translation
Regulating RNA					
sdRNA	tRNA halves	40 nt		retrotransposons, DNA	DNA processing
miRNA	many	22 nt	Pol II	mRNA	degradation
Auxiliary					
SRP RNA		297 nt	Pol III	signal sequence of forming protein	translocating

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tRNAs

Background

Transfer RNAs (tRNAs) are central in the physical translation of the genetic code of mRNA into proteins in the context of ribosomes (Rich and RajBhandary 1976; Rodnina and Wintermeyer 2011). Yet, tRNAs are increasingly recognized as important players in a wide range of cellular processes (Schimmel 2017; Schorn and Martienssen 2018). tRNA biogenesis encompasses: translation of the precursor-tRNA, removal of the 5' leader sequence, trimming of the 3' trailer sequence, addition of CAA nucleotides, splicing-out of possible intronic sequences, and numerous nucleotide modifications (Phizicky and Hopper 2010; Kirchner and Ignatova 2015). Mature tRNAs are about 80 nucleotides long and have a characteristic cloverleaf secondary structure, as well as a L-shaped tertiary structure (Rich and RajBhandary 1976). In eukaryotes there are 61 possible anticodons and thus, in principle, 61 possible tRNAs (Alberts et al. 2015). However, due to the "Wobble principle", which allows an amino acid to be coded by a set of anticodons (n = 2, 4, or 6), named isoacceptors (Goodenbour and Pan 2006) via effectively 'ignoring' the third base of the codon, a cell does not need all 61 tRNAs for proper translation. The minimum number of (anti)codons is 31, and although the exact number is different for each species, in general about 44 different anticodon tRNA genes are present in the genome of a species (Marck and Grosjean 2002). At the same time, each anticodon tRNA is usually present by multiple copies in the genome and if these copies of one tRNA anticodon have differences in their body sequence, they are called isodecoders (Goodenbour and Pan 2006).

One of the reasons for many isodecoders is the possibility to use different sets of isodecoders for each tRNA anticodon for different cell types, for instance in the case of pathologies (Rogers et al. 2010; Gingold et al. 2014; Kirchner and Ignatova 2015). Differential tRNA expression in combination with codon bias (Quax et al. 2015) and in relation to protein folding (Marín et al. 2017) is currently an emerging field of research.

Zebrafish tRNA genomic organization

Given the existence of a maternal-specific ribosome in zebrafish development (Locati et al. 2017a, 2017b), the question was whether, similarly, a maternal-specific tRNA repertoire is used in early embryogenesis. To address this, an inventory was made (cf. M&M (Locati et al.

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2017a)) of the tRNA genes starting from the genomic tRNA database GtRNAdb (release 2010) (Chan and Lowe 2009), which predicts a staggering 9,904 tRNA genes in the zebrafish genome (Zv9).

The first notable observation was the fact that the zebrafish genome contains all 61 types of anticodon tRNA (Supplemental Table ST1). Thus far, we have been unable to find any other species that has all the anticodon tRNAs present in its genome. For each anticodon, several tRNA genes are present in the zebrafish genome, ranging from minimal one (Asp-ATC) gene to maximal 825 (Asn-GTT) identical copies, as well as isodecoders. Furthermore, anticodon tRNAs appear to be organized in large clusters, small clusters, and/or solitary genes (Figure 2A, Supplemental Table ST1). The largest clusters per anticodon tRNA are primarily (~56%) found on the right arm of chromosome 4, but are also present on several other chromosomes (Figure 2A). Conversely, for each group of isoacceptors, one anticodon tRNA does not appear in a big gene cluster (n>16) (Supplemental Table ST1). These are the anticodon tRNA genes that are also often absent in genomes of other species (Marck and Grosjean 2002). The anticodons of these low-presence tRNA genes distribute evenly over the genetic code (Figure 2B) and involve a C or U base at their 3rd position, which is in line with the possible presence of inosine in the tRNA anticodon (Agris et al. 2018).

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A

		tRNA gene distribution																								
		Chromosome																								
Amino acid	Anticodon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Glu	TTC	0	0	0	120	5	1	0	4	0	0	1	0	0	1	16	0	0	0	0	0	0	13	1	0	0
	CTC	3	0	0	104	6	0	3	2	0	0	0	0	0	0	30	2	0	0	0	0	1	11	1	0	0
Ser	GCT	0	0	4	233	26	0	1	5	0	1	5	0	0	0	0	0	24	0	0	1	2	24	0	0	0
	TGA	0	0	8	119	6	0	1	1	0	0	1	0	1	0	0	0	0	0	0	16	1	3	0	0	0
	CGA	0	0	7	47	1	1	0	2	1	0	0	0	3	1	0	0	0	0	0	2	0	1	0	0	0
	AGA	0	0	17	232	6	0	0	2	0	0	0	0	1	0	0	1	1	0	0	15	0	5	0	0	0
Ala	TGC	3	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	357	0	0	0	0
	GGC	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	CGC	0	0	0	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	106	0	0	0	0
	AGC	0	0	54	9	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	5	0	0	1	15
Gly	TCC	0	0	0	21	2	2	0	2	0	0	0	0	0	0	2	0	0	0	0	0	3	5	1	0	0
	GCC	0	0	14	459	39	0	0	76	0	8	0	0	4	1	0	0	0	0	0	0	0	36	0	0	2
	CCC	0	1	0	147	22	0	0	4	0	0	0	0	0	0	1	0	0	0	0	0	0	19	0	0	0
Met	CAT	1	2	3	227	13	0	0	3	54	0	0	2	0	0	0	0	3	0	2	2	3	205	0	0	0

B

First letter	Second letter: U		C		A		G		Third letter
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC		UCC		UAC		UGC		C
	UUA	Leu	UCA		UAA	STOP	UGA	STOP	A
	UUG		UCG		UAG	STOP	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lys	AGA	Arg	A
	AUG	Met	ACG		AAG		AGG		G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glu	GGA		A
	GUG		GCG		GAG		GGG		G

C

Isodecoder	Sequence	Read count Egg	Read count Adult Male
SerGCT-99	AAGCCGATGGACTGCTAATCCATTGTGCTCTGCACGCGTGGC	14147	1
SerGCT-351	AAGGCGATGGACTGCTAATCCATTGTGCTTTGCACGCATGCG	3305	0
SerGCT-121	AAGGCAATGGACTGCTAATCCATTGTGCTCTGCACGTGTGGG	9946	1
SerGCT-368	AAGGCGATGGACTGCTAATCTATTGTGATCTGCATGCGTGCG	8547	0
SerGCT-328	AAGGCGATGGACTGCTAATCCATTGTGCTCTGCACGCATGAG	1097	0
SerGCT-324	AAGGCGATGGACTGCTAATCCATTGTGCTCTGCACACGTGGG	2101	1
SerGCT-519	AAGGTGATGGACTGCTAATCCATTGTGCTCTGCACGCATGGG	3585	2
SerGCT-195	AAGGCGATAGACTGCTAATCCATTGTGCTCTGCACGCGTGGG	2	4114
SerGCT-369	AAGGCGATGGACTGCTAATCTATTGTGCTCTGCACGCGTGGG	1	2301
SerGCT-2264	GAGGCGATGGACTGCTAATCCATTGTGCTCTGCACGCGTGGG	0	1107
SerGCT-355	AAGGCGATGGACTGCTAATCCATTGTGCTTTGCATGCATGGG	0	1044
SerGCT-350	AAGGCGATGGACTGCTAATCCATTGTGCTTTGCACACGTGGG	0	725
SerGCT-202	AAGGCGATGAACTGCTAATCCATTGTGCTCTGCACGCGTGGG	0	700
SerGCT-193	AAGGCGATAGACTGCTAATCCATTGTGCTCTGCACGCGGGAG	0	691
SerGCT-344	AAGGCGATGGACTGCTAATCCATTGTGCTCTGCATGCGTGGG	2	1971
SerGCT-330	AAGGCGATGGACTGCTAATCCATTGTGCTCTGCACGCGGGGG	0	667
SerGCT-80	AACGTGATGGACTGCTAATCCATTGTGCTCTGCACACATGG	0	627
SerGCT-323	AAGGCGATGGACTGCTAATCCATTGTGCTCTGCAAGCGTGGG	0	572
SerGCT-198	AAGGCGATCGACTGCTAATCCATTGTGCTCTGCACGCGTGGG	0	297

D

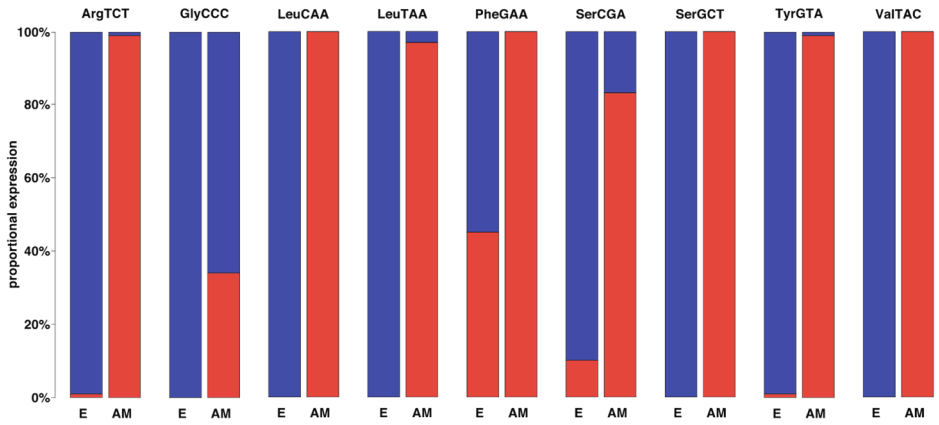

Figure 2: Maternal-type tRNA variant characteristics

A: The genomic distribution of some exemplary zebrafish tRNAs (complete table in Supplemental Table ST1).

B: The genetic code table. For the highlighted codons, the associated anticodons tRNA genes are present in the genome with just a few copies.

C: Alignment of exemplary tRNA SerGCT isodecoders and their differential expression in egg and adult male zebrafish.

D: A barplot showing the proportional, average expression of maternal-type (blue) and somatic-type (red) variants for some exemplary anticodon tRNAs in egg (E) and adult male (AM) zebrafish (Supplemental Table ST1).

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Zebrafish tRNA maternal expression

As with 5S rRNA and snRNA, tRNAs are quite difficult to sequence by standard next-generation sequencing approaches, likely due to their strong secondary structure, 3' amino acid, and many RNA modifications (Zheng et al. 2015). We therefore employed our RT-PCR-qSeq approach (cf. M&M (Pagano et al. 2019a)) with PCR primer pairs hybridizing to the 20 most 5' terminal and 3' terminal bases of nearly all anticodon tRNAs (Supplemental Table ST1). Regrettably, this approach leaves only about 40 nucleotides for isodecoder identification. For almost all investigated anticodon tRNAs, expression data was obtained by analyzing egg and adult male tail samples (cf. M&M (Pagano et al. 2019a)). An initial analysis clearly showed that there is differential expression between different isodecoders of each analyzed anticodon tRNA (Figure 2C and 2D). There are undoubtedly tRNA genes per anticodon which transcripts are either present in egg or adult tail. This demonstrates that also maternal-type tRNA genes exist. At the same time, we generally could not find obvious sequence signatures for maternal-type tRNAs. This however could be caused by the fact that we just examined the core of about 40 nucleotides for each tRNA gene, whereas there are often differences between isodecoders within the remaining parts of their tRNA body sequence, which are used in our method for PCR-primer binding, the 5' and 3' precursor tRNA parts, and/ or intronic sequences of precursor tRNA.

What did stand out was that it seems as if the maternal-type tRNA genes are organized in large clusters, mostly on chromosome 4, whereas somatic-type genes are mostly found as single gene or in a small cluster. Yet, although the low-presence tRNA isoacceptor genes (Figure 2B) are organized like somatic-type genes, their transcripts actually are primarily present in egg (Supplemental Table ST1).

Conclusion

Within the set of genes belonging to an anticodon tRNA, there are maternal-type and somatic-type isodecoders. The maternal-type isodecoders are mainly organized in large clusters on chromosome 4, which also harbor the maternal-type rRNA genes, yet without obvious sequence signatures. It would be interesting to determine whether the maternal-type and somatic-type tRNA genes line up with the reported "core" and "peripheral" tRNA genes proposed in *Drosophila* (Rogers et al. 2010). Also, maternal-type tRNAs might uniquely encompass all possible anticodons (n=61), which raises the question whether the

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

zebrafish maternal-type ribosome might not allow for Wobble translation. An intriguing link might also be found in the codon identity that is involved in mRNA stability regulation and translation efficiency during the maternal-to-zygotic transition (Bazzini et al. 2016). Our tRNA observations must of course be confirmed by comprehensive experiments and analyses, but given these compelling preliminary observations, we postulate that tRNAs are a contributing part of the maternal-specific translation system.

RNase P

Background

Another type of ribonucleoprotein (RNP) involved in the cellular translation system is the RNase P enzyme. In vertebrates the RNase P enzyme consists of a non-coding RNA subunit (rpph1) and several proteins (reviewed in Klemm et al. 2016). Initially discovered for its role as a ribozyme that removes the 5' leader of a precursor tRNA in the maturation process, it now is recognized to be also involved in rRNA and mRNA cleavage, as well as chromatin remodeling (reviewed in Jarrous 2017; Esakova and Krasilnikov 2010).

Zebrafish RNase P genomic organization

Given the role of the RNase P enzyme in tRNA maturation and rRNA processing, the question was whether a maternal-specific RNase P enzyme exists in zebrafish. For this we focused on the non-coding RNA subunit of the RNase P enzyme. A rpph1 gene (Ribonuclease P RNA component H1) of 307 nucleotides is annotated in the zebrafish genome on chromosome 2 (position 37,826,734). A standard BLAST analysis (Madden 2002) revealed a similar (96%) gene copy (rpph1-sim), as well as a dissimilar (80%) gene copy (rpph1-dis), ~10 kb and ~5 Mb downstream respectively (Figure 3 A, Supplemental Table ST2).

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

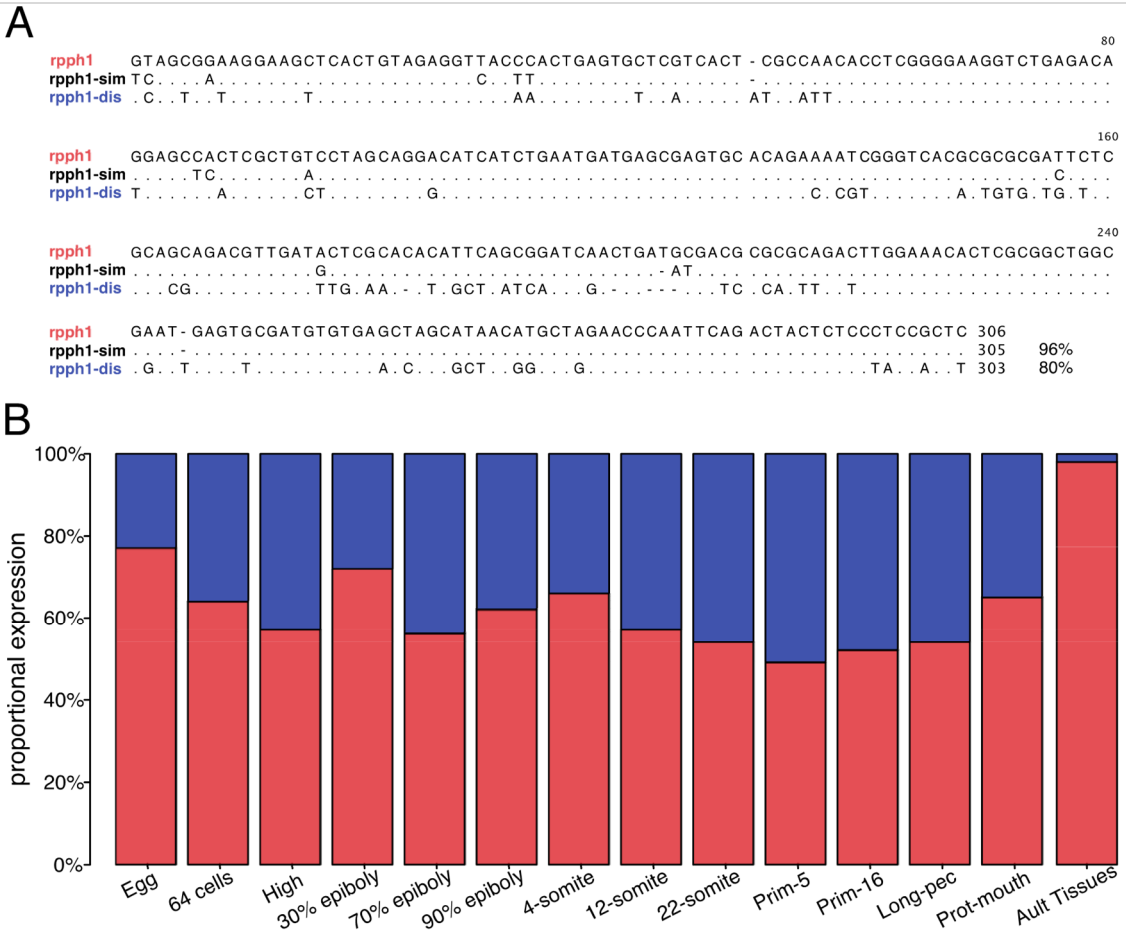

Figure 3: Maternal-type and somatic-type RNase P RNP variants

A: Sequence alignment of the three zebrafish RNase P RNP (rpph1) gene variants. Identical nucleotides or gaps are indicated as dots and dashes, respectively.

B: Bar plot showing the maternal-type (blue) and somatic-type (red) rpph1 proportional expression during zebrafish developmental stages (Supplemental Table ST2).

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Zebrafish RNA P maternal expression

To determine possible differential gene expression, we mapped the RNAseq reads from zebrafish eggs, a developmental embryonic series, and mature samples to the three rpph1 variant sequences (cf. M&M (Locati et al. 2017b), and Supplemental Table ST2).

The annotated rpph1 gene is the most prominently expressed variant in virtually all samples, in contrast to the similar copy, rpph1-sim, which is hardly ever expressed (Supplemental Table ST2). The dissimilar copy, rpph1-dis, is mainly expressed during oogenesis and embryogenesis, which is in line with it being the maternal-specific RNase P RNA gene. In contrast to most other somatic-specific non-coding RNAs, which are hardly present in egg, rpph1 transcripts are also quite prominently present in egg. Yet, it is important to emphasize that the analysis shows relative expression and the absolute expression of RNase P RNA is very low in egg and embryonic samples, as compared to low-proliferative tissues like brain and liver (Supplemental Table ST2). Additionally, an inspection of the mapped reads (data not shown) revealed other nucleotide differences as compared to the rpph1 reference sequences, suggesting the possible presence of one or more variants.

Conclusion

There are three rpph1 variants in close proximity located on chromosome 2; the original rphh1 is expressed in all samples tested, one (similar) variant is hardly present, and one (dissimilar) variant is present virtually just in egg and during embryogenesis, which makes this the maternal-specific variant. However, all rrph1 expression levels in egg and embryonic samples are rather low compared to that of the ubiquitous rphh1 variant in somatic tissues. Given the numerous sequence differences between the two rpph1 variants, it is likely that most, if not all, protein binding sites are affected, which may have a profound effect on the RNase P RNP as a whole. For example, a study involving mutations in a RNase P protein in *drosophila* showed that these alterations induced complete sterility by triggering several DNA damage checkpoints in the oogenesis (Molla-Herman et al. 2015). It is impossible to assert which task(s) the maternal-type rpph1 has in the oogenesis and embryogenesis, other than it is related to a part of the maternal-specific translation system.

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SRP-RNA

Background

The signal recognition particle (SRP) is a ribonucleoprotein responsible for the translocation of the mRNA-ribosome translation complex to the endoplasmic reticulum (ER). The SRP recognizes the ER signals sequence at the N-terminus of an integral membrane protein or secretory protein that is being synthesized. Direct binding of the SRP to the signal sequence peptide and indirect binding to ribosome stalls protein synthesis, after which the whole complex is guided to the ER. Binding to the SRP receptor in the ER membrane releases the SRP from the complex and translation resumes by synthesizing the protein through the ER membrane via a protein translocator (Walter and Blobel 1980; Lütcke 1995; Keenan et al. 2001; Akopian et al. 2013; Alberts et al. 2015).

The SRP consists of several proteins and a small non-coding RNA; SRP-RNA (also known as 7SL RNA), which is produced by RNA polymerase III (Walter and Blobel 1982; Leung and Brown 2010). This SRP-RNA forms the backbone of the SRP, is about 300 nucleotides long, and has a highly conserved secondary structure with a small (Alu) domain and a large (S) domain (Walter and Blobel 1983). The Alu domain is responsible for translation retardation and the S domain recognizes the signal sequence and also binds to the ER SRP receptor (Lütcke 1995; Wild et al. 2019).

Zebrafish SRP-RNA genomic organization

In the zebrafish genome, SRP-RNA genes can be found on three chromosomes; 3, 11, and 20 with a standard BLAST analysis (Madden 2002). On chromosome 11, 67 complete copies are organized mainly in three gene clusters with a mutual homology of 91% and up. Both chromosome 3 and chromosome 20 contain one copy of the SRP-RNA gene with 93% and 94% homology to the SRP-RNA consensus sequence of the genes on chromosome 11 (Figure 3A, Supplemental Table ST3). The position of the gene copy on chromosome 20 is conspicuous, as is it approximately 10 kb downstream of the maternally expressed tRNA-LeuCAA cluster. The copy on chromosome 3 is located just 448 bp from the start of the *ER membrane protein complex subunit 10* gene.

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

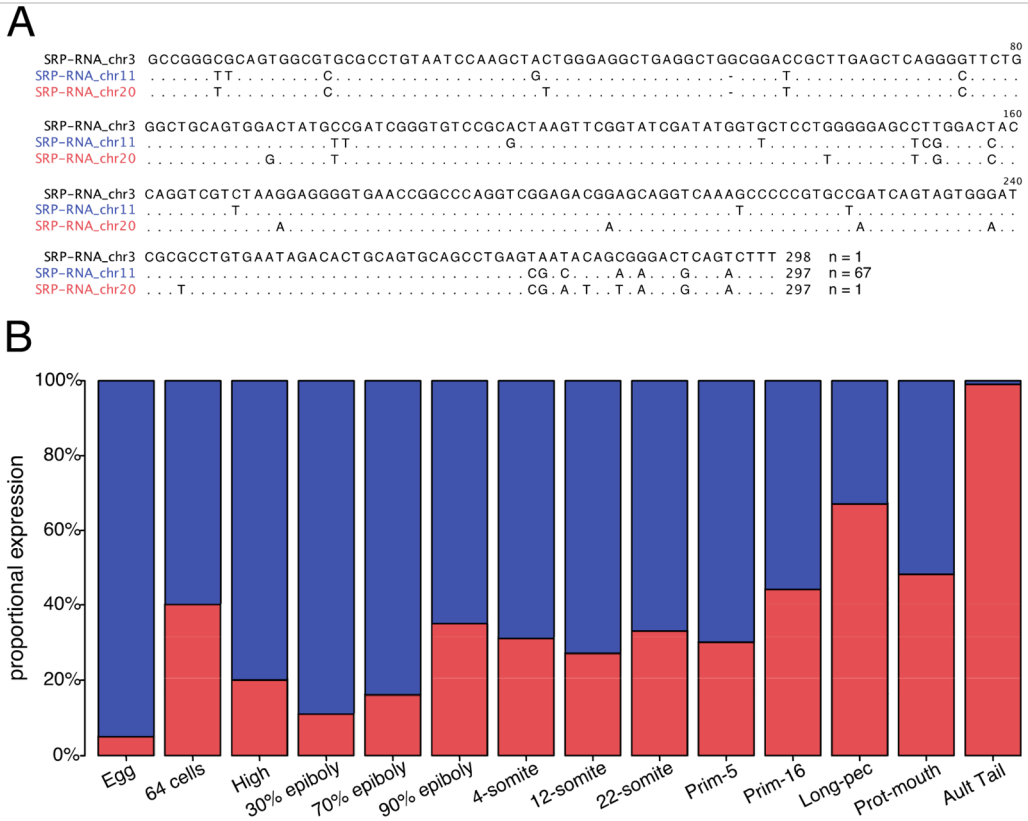

Figure 3: Maternal-type and somatic-type SRP-RNA variants

A: Sequence alignment of the zebrafish SRP-RNA gene variants. Identical nucleotides or gaps are indicated as dots and dashes, respectively.

B: Bar plot showing the maternal-type (blue) and somatic-type (red) SRP-RNA proportional expression during zebrafish developmental stages (Supplemental Table ST3).

Zebrafish SRP-RNA maternal expression

The homology between the different SRP-RNA gene copies is relatively high, but there are sufficient sequence differences between the three chromosomal versions to identify them (Figure 4A). Mapping the RNAseq reads from egg, the embryonic development series, and adult samples to the three SRP-RNA variant sequences (cf. M&M (Locati et al. 2017b)) showed a clear differential expression (Figure 4B, Supplemental Table ST3). Transcripts from the SRP-RNA gene clusters on chromosome 11 are predominantly present in egg plus early embryogenesis and are during embryogenesis replaced by those from the single SRP-RNA gene on chromosome 3, which then becomes the almost exclusive variant expressed in adult tissues. The single copy on chromosome 20 is hardly expressed in any of the samples. The resulting expression patterns are highly similar to maternal-type rRNA (similar to 67 copies on chromosome 11,), somatic-type rRNA (similar to single copy on chromosome 3),

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and unexpressed-type rRNA (similar to single copy on chromosome 20) (Locati et al. 2017a, 2017b).

Conclusion

The SRP-RNA seems to be a part of the maternal-specific translation system as there is clear differential expression between the gene copies. Intriguingly, the SRP-RNA genomic organization seems to follow the typical trend that maternal variants are present in large numbers and in clusters, whereas the somatic variants are often solitary gene copies.

The sequence differences between the maternal chromosome 11 SRP-RNA copies and the somatic (chromosome 3) variant (~7%) are both in the Alu domain, as well as in the S domain. As several SRP proteins bind directly to the SRP-RNA, the sequence differences might affect binding of them, so the maternal-type variant SRP might end up being significantly different from the somatic-type variant.

Altogether, zebrafish has a maternal-specific SRP to guide the mRNA-ribosome complex to the ER. As to why also this part of the translation system is present in a maternal-type variant remains elusive for now, although it seems possible that is a, yet unidentified element of the ribosome filter hypothesis (Mauro VP, 2002, 2007, and 2016), which seems related at least to the 45S rRNA maternal- and somatic types (Locati et al. 2017b). In any case, as virtually no research is done into the zebrafish SRP-RNA, the discovered dual system provided an excellent opportunity to study this.

Conclusion

In this manuscript, we presented the preliminary results of our initial gene-expression analyses of the non-coding RNA types involved in the zebrafish translation system, other than we previously reported on (Locati et al. 2017a, 2017b, 2018; Pagano et al. 2019b, 2019a). Alike rRNA (Locati et al. 2017a, 2017b), snoRNA (Pagano et al. 2019b), and snRNA (Pagano et al. 2019a), the investigated tRNA, RNase-P and SRP-RNA all seem to have a maternal-specific variant. Together, a complete maternal-specific translation system emerges, which is dedicated to oogenesis and (early) embryogenesis in zebrafish. Especially the existence of a maternal-specific SRP-RNA points to a complete system, as it targets the initial part of the produced protein. Besides the maternal-specific expression profiles of the RNA-processing snoRNA (Pagano et al. 2019b) and RNase P (both of which are indirect non-

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coding RNA actors in the translation system cf. Figure 1), all translation-involved non-coding RNA types show rather absolute expression in egg, virtual absent expression in adult tissue, and decreasing expression during embryogenesis.

Putting the findings of the here presented studies in context with our previous finding of non-coding RNAs that are involved in the cellular translation system, leads to a number of intriguing reflections.

Maternal-type proteins: All translation-involved non-coding RNAs function as ribonucleoprotein particles (RNPs), but often depend on RNA-RNA interactions. The nucleotide differences between maternal-type and somatic-type RNA variants are mostly located at the protein binding sites of the implicated non-coding RNAs. This raises the question whether the translation RNP-associated proteins, for instance between the maternal-type and somatic-type ribozymes, are also dissimilar. A quick search did not show maternal-specific expression of the protein-coding genes of these RNPs, however, as these genes contain introns, possible different proteins could be orchestrated via differential splicing.

An "unexpressed-type" translation system: One of the first steps in the performed analyses of each non-coding RNA type was to characterize their complete genomic presence. Once we identified all the complete gene copies of a RNA type, their expression patterns in zebrafish development were determined. Besides the expressed maternal-type and somatic type RNAs, we discovered several apparently functional genes that were not, or extremely low expressed. For this reason, we labeled these gene copies as "unexpressed". Unexpressed-type gene copies have been identified for: rRNAs, snoRNAs, snRNAs, RNase and SRP-RNA (for tRNAs we were unable to perform a complete genomic characterization).

Because we assume that these variants have to be expressed somewhere, several research attempts were made to find transcripts of these unexpressed-type non-coding RNAs. As we speculated that this might actually be a paternal-type variant, we analyzed zebrafish testis samples, as well as various other tissues (unpublished results). The unexpressed-type variant turned out to be not expressed in testis and the results of the other tissues were inconclusive given the overwhelming presence of maternal- and somatic-type transcripts in the RNAseq data. It may be that this particular variant is only expressed in certain low-frequency cell types, for instance stem cells, or alternatively it is only expressed in response to stimuli such as stress, hormones, or temperature. In any case, determining where this

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

unexpressed translation system variant is expressed will be extremely informative for unraveling the relevance of the involved (sequence) differences between all three translation systems.

Genomic organization: One of the most obvious distinctions between the maternal-type and somatic-type RNA variants is their particular genomic organization. There are almost always many maternal-specific gene copies, whereas of the somatic-type gene copies there are often just a few. The maternal loci have a hallmark organization of repeated and clustered genes with a preference for chromosome 4 (Table 1). Notably, this chromosome is identified as a gender chromosome in zebrafish (Anderson et al. 2012; Nagabhushana and Mishra 2016). The somatic-type gene copies are virtually always solitarily organized. For 5S rRNA we could show a connection between a specific retrotransposon family and the maternal-type tandem arrays (Locati et al. 2017a). Given the absence of complete retrotransposons in the somatic-type 5S rRNA locus, we suggested an involvement of this specific retrotransposon family in achieving and maintaining the differential copy number of the two types of 5S rDNA loci. Although we did not fully analyze this phenomenon in the other translation-involved non-coding RNA loci, it was clear that several of them showed similar retrotransposon hallmarks such as; partial gene sequences and genomic spaces in otherwise tightly-packed tandem arrays (unpublished results). It appears as if there is a specific retrotransposon family involved in almost each of the mentioned non-codon RNAs. This could hint to a universal mechanism to maintain a proper number of functional gene copies (Kojima 2003; Dupuis-Sandoval et al. 2015). In any case, identifying the exact retrotransposons associated with the relevant non-coding RNA loci might elucidate the genome organization of these RNAs, as well as their possible non-canonical functions (Buzdin et al. 2007).

Proliferation-specific translation system: One of the nagging issues for all of the non-coding RNA research is the lingering question as to why there is an apparent need in zebrafish for a dual (or even triple) different translation system. In an early stage of our research, we named the newly-discovered alternative system “maternal-type” due to its transcription origin in oocytes. As a logic consequence, the other translation system became “somatic-type”. However, this decision came back to haunt us, when we much later discovered that testis tissue, presumably (immature) sperm cells, contain a considerable amount of maternal-type translation-involved non-coding RNA. Another option could be that our

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

maternal-type translation system actually is related to stem cells in general, not just the totipotent fertilized egg cell. Recently there have been several reports on alternative snRNAs specifically expressed in stem cells (Vazquez-Arango et al. 2016; Kim et al. 2017; Schmidt et al. 2018). In line with these observations is the report on the existence of two distinct human translation programs that operate during either cell proliferation or differentiation (Gingold et al. 2014). This aligns with a large body of research that implicates (variant) nucleolus components in cancer cells (Derenzini et al. 1998; Lo et al. 2006; Jorjani et al. 2016). Given that the first phase in zebrafish embryogenesis of about eight hours is primarily dedicated to proliferation, after which is switch is made to mainly differentiation, it could well be that our maternal-type translation system is in reality a proliferation-specific translation system.

SUPPLEMENTAL TABELS

ST1.xlsx: tRNA

ST2.xlsx: RNase

ST3.xlsx: SRP

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Agris PF, Eruysal ER, Narendran A, Väre VYP, Vangaveti S, Ranganathan S V. Celebrating wobble decoding: Half a century and still much is new. *RNA Biol.* 2018;15(4-5):537–553.

Akopian D, Shen K, Zhang X, Shan S. Signal Recognition Particle: An Essential Protein-Targeting Machine. *Annu Rev Biochem*. 2013;82:693–721.

Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, Walter P. *Molecular Biology of THE CELL*. 6th ed. 2015.

Anderson JL, Marí A, Braasch I, Amores A, Hohenlohe P, Batzel P, Postlethwait JH. Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. *PLoS One.* 2012; 7(7):

Bazzini AA, del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, Yao J, Khokha MK, Giraldez AJ. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J.* 2016;35: 2087–2103.

Brown DD, Carroll D, Brown RD. The isolation and characterization of a second oocyte 5S DNA from xenopus laevis. *Cell*. 1977;12: 1045-56

Buzdin A, Gogvadze E, Lebrun MH. Chimeric retrogenes suggest a role for the nucleolus in LINE amplification. *FEBS Lett.* 2007;581: 2877–2882.

Chan PP, Lowe TM. GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res*. 2009;37.

Derenzini M, Trerè D, Pession A, Montanaro L, Sirri V, Ochs RL. Nucleolar function and size in cancer cells. *Am J Pathol*. 1998; 152: 1291–7.

Dupuis-Sandoval F, Poirier M, Scott MS. The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip Rev RNA.* 2015;6: 381–397.

Esakova O, Krasilnikov AS. Of proteins and RNA: The RNase P/MRP family. *RNA.* 2010;16: 1725–1747.

Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, Christophersen NS, Christensen LL, Borre M, Sørensen KD, et al. A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell.* 2014;158: 1281–1292.

Goodenbour JM, Pan T. Diversity of tRNA genes in eukaryotes. *Nucleic Acids Res.* 2006;34: 6137–6146.

Graindorge A, Thuret R, Pollet N, Osborne HB, Audic Y. Identification of post-transcriptionally regulated Xenopus tropicalis maternal mRNAs by microarray. *Nucleic Acids Res.* 2006;34: 986–995.

Guinta DR, Tso JY, Narayanswami S, Hamkalo BA, Korn LJ. Early replication and expression of oocyte-type 5S RNA genes in a Xenopus somatic cell line carrying a translocation. *Proc Natl Acad Sci*. 1986;83: 5150-5154

Jarrous N. Roles of RNase P and Its Subunits. *Trends Genet*. 2017;33: 594–603.

Jorjani H, Kehr S, Jedlinski DJ, Gumienny R, Hertel J, Stadler PF, Zavolan M, Gruber AR. An updated human snoRNAome. *Nucleic Acids Res.* 2016;44: 5068–5082.

Keenan RJ, Freymann DM, Stroud RM, Walter P. The Signal Recognition Particle. *Annu Rev Biochem*. 2001;70:755-75.

Kim YD, Lee J, Kim HS, Lee MO, Son MY, Yoo CH, Choi JK, Lee SC, Cho YS. The unique spliceosome signature of human pluripotent stem cells is mediated by SNRPA1, SNRPD1, and PNN. *Stem Cell Res.* 2017;22: 43–53.

Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and disease. *Nat Rev Genet*. 2015;16: 98–112

Klemm BP, Wu N, Chen Y, Liu X, Kaitany KJ, Howard MJ, Fierke CA. The diversity of ribonuclease P: Protein and RNA catalysts with analogous biological functions. *Biomolecules.* 2016;6(2): 27.

Kojima KK. Cross-Genome Screening of Novel Sequence-Specific Non-LTR Retrotransposons: Various Multicopy RNA Genes and Microsatellites Are Selected as Targets. *Mol Biol Evol.* 2003;21: 207–217.

Komiya H, Hasegawa M, Takemura S. Differentiation of oocyte- and somatic-type 5S rRNAs in animals. *J Biochem*. 1986;100: 369–74.

Leung E, Brown JD. Biogenesis of the signal recognition particle. *Biochem Soc Trans*. 2010; 38(4):1093-8.

Lo SJ, Lee CC, Lai HJ. The nucleolus: Reviewing oldies to have new understandings. *Cell Res*. 2006;16: 530–538.

Locati MD, Pagano JFB, Abdullah F, Ensink WA, van Olst M, van Leeuwen S, Nehrdich U, Spaink HP, Rauwerda H, Jonker MJ, et al. Identifying small RNAs derived from maternal- and somatic-type rRNAs in Zebrafish Development. *Genome.* 2018;61:371-378.

Locati MD, Pagano JFB, Ensink WA, van Olst M, van Leeuwen S, Nehrdich U, Zhu K, Spaink HP, Girard G, Rauwerda H, et al. Linking maternal and somatic 5S rRNA types with different sequence-specific non-LTR retrotransposons. *RNA.* 2017a;23: 446–456.

Locati MD, Pagano JFB, Girard G, Ensink WA, van Olst M, van Leeuwen S, Nehrdich U, Spaink HP, Rauwerda H, Jonker MJ, et al. Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA*. 2017b; 23: 1188–1199.

Lütcke H. Signal Recognition Particle (SRP), a Ubiquitous Initiator of Protein Translocation. *Eur J Biochem.* 1995;228: 531–550.

Madden T. Chapter 16: The BLAST Sequence Analysis Tool. *NCBI Handbook* 2002;1–15.

Marck C, Grosjean H. tRNomics: Analysis of tRNA genes from 50 genomes of eukarya, archaea, and bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA.* 2002;8: 1189–1232.

Marín M, Fernández-Calero T, Ehrlich R. Protein folding and tRNA biology. *Biophys Rev*. 2017; 9: 573–588.

Molla-Herman A, Valles AM, Ganem-Elbaz C, Antoniewski C, Huynh J-R. tRNA processing defects induce replication stress and Chk2-dependent disruption of piRNA transcription. *EMBO J.* 2015;34: 3009–3027.

Nagabhushana A, Mishra RK. Finding clues to the riddle of sex determination in zebrafish. *J Biosci*. 2016;41: 145–55.

Pagano JFB, Dekker RJ, Ensink WA, van Olst M, Bos A, van Leeuwen S, de Leeuw WC, Nehrdich U, Spaink HP, Rauwerda H, et al. 2019a. An alternative spliceosome defined by distinct snRNAs in early zebrafish embryogenesis. *Not Publ*.

Pagano JFB, Locati MD, Ensink WA, van Olst M, van Leeuwen S, de Leeuw WC, Nehrdich U, Spaink HP, Rauwerda H, Jonker MJ, et al. 2019b. Maternal- and Somatic-type snoRNA Expression and Processing in Zebrafish Development. *Not Publ*.

Phizicky EM, Hopper AK. tRNA biology charges to the front. *Genes Dev*. 2010; 24:1832-60.

Quax TEF, Claassens NJ, Söll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol Cell*. 2015;59: 149–161.

Rich A, RajBhandary UL. Transfer RNA: Molecular Structure, Sequence, and Properties. *Annu Rev Biochem*. 1976;45: 805–860.

Rodnina M V., Wintermeyer W. The ribosome as a molecular machine: the mechanism of

bioRxiv preprint doi: https://doi.org/10.1101/869651; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tRNA–mRNA movement in translocation. *Biochem Soc Trans.* 2011;59: 658–662.

Rogers HH, Bergman CM, Griffiths-Jones S. The evolution of tRNA genes in Drosophila. *Genome Biol Evol*. 2010;2: 467–477.

Schimmel P. The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat Rev Mol Cell Biol.* 2017;19: 45–58.

Schmidt D, Reuter H, Hüttner K, Ruhe L, Rabert F, Seebeck F, Irimia M, Solana J, Bartscherer K. The Integrator complex regulates differential snRNA processing and fate of adult stem cells in the highly regenerative planarian Schmidtea mediterranea. *PLoS Genet*. 2018;14: e1007828.

Schorn AJ, Martienssen R. Tie-Break: Host and Retrotransposons Play tRNA. *Trends Cell Biol*. 2018;793:806.

Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*. 2014;508, 66–71.

Tadros W, Lipshitz HD. Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in Drosophila. *Dev Dyn*. 2005; 232:593-608.

Tang DTP, Glazov EA, McWilliam SM, Barris WC, Dalrymple BP. Analysis of the complement and molecular evolution of tRNA genes in cow. *BMC Genomics*. 2009;10: 1–13.

Vazquez-Arango P, Vowles J, Browne C, Hartfield E, Fernandes HJR, Mandefro B, Sareen D, James W, Wade-Martins R, Cowley SA, et al. Variant U1 snRNAs are implicated in human pluripotent stem cell maintenance and neuromuscular disease. *Nucleic Acids Res.* 2016;44: 10960–10973.

Walter P, Blobel G. Disassembly and reconstitution of signal recognition particle. *Cell*. 1983; 34, 525-533.

Walter P, Blobel G. Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc Natl Acad Sci*. 1980;77: 7112–7116.

Walter P, Blobel G. Signal recognition particle contains a 7 S RNA essential for protein. *Nature*. 1982;299:691-8.

Wegnez M, Monier R, Denis H. Sequence heterogeneity of 5 S RNA in Xenopus laevis. *FEBS Lett*. 1972;25:13-20.

Weill L, Belloc E, Bava FA, Méndez R. Translational control by changes in poly(A) tail length: Recycling mRNAs. *Nat Struct Mol Biol*. 2012;19: 577–585.

Wild K, Juaire KD, Soni K, Shanmuganathan V, Hendricks A, Segnitz B, Beckmann R, Sinning I. Reconstitution of the human SRP system and quantitative and systematic analysis of its ribosome interactions. *Nucleic Acids Res.* 2019;47: 3184–3196.

Wormington WM, Brown DD. Onset of 5 S RNA gene regulation during Xenopus embryogenesis. *Dev Biol*. 1983;99:248-57.

Zheng G, Qin Y, Clark WC, Dai Q, Yi C, He C, Lambowitz AM, Pan T. Efficient and quantitative high-throughput tRNA sequencing. *Nat Methods*. 2015;12:835-837