PhD in Aquatic Biosciences // No. 46 - 2022

Ribosomal RNA, ribose methylation, and box C/D snoRNAs during embryonic development of teleosts zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*)

Sowmya Ramachandran

FACULTY OF BIOSCIENCES AND AQUACULTURE

NORD University

www.nord.no

Ribosomal RNA, ribose methylation, and box C/D snoRNAs during embryonic development of teleosts zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*)

Sowmya Ramachandran

A thesis for the degree of
Philosophiae Doctor (PhD)

PhD in Aquatic Biosciences no. 46 (2022)
Faculty of Biosciences and Aquaculture

PhD in Aquatic Biosciences no. 46 (2022)

Sowmya Ramachandran
Ribosomal RNA, ribose methylation, and box C/D snoRNAs during embryonic development of teleosts zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*)

ISBN: 978-82-93165-45-3

Print: Trykkeriet NORD

Nord University
N-8049 Bodø
Tel: +47 75 51 72 00
www.nord.no

Preface

This doctoral thesis is submitted in fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Faculty of Biosciences and Aquaculture (FBA), Nord University. The different studies compiled in this dissertation are original research studies conducted at Nord University, Bodø, Copenhagen University, Denmark and INRAe Rennes, France, over a period of three years. This PhD project is a part of the *InnControl* project (grant #275786) funded by the Research Council of Norway. The views in the main thesis text reflect the interpretation of the candidate alone and can not be considered as the views of any members of the supervisory committee.

The core project team consisted of the following members:

Sowmya Ramachandran, MSc, FBA, Nord University: PhD fellow

Igor Szczepan Babiak, Professor, FBA, Nord University: primary supervisor

Henrik Nielsen, Professor, Copenhagen University: co- supervisor

Steinar Daae Johansen, Professor, FBA, Nord University: co- supervisor

Bodø, October 2022

Sowmya Ramachandran

Dedication

I dedicate this thesis to Appa, my father, my favourite human, my superhero, my cheerleader and my devil's advocate. It's devastating to have lost you mere weeks before finishing my PhD. I feel for the three years that I could not come visit but I know I will forever have your voice be my conscience. Thank you for trusting my choices, always pushing me to do better and loving me so fiercely. I miss you every minute!

Acknowledgements

When I first applied for this PhD, I was quite possibly at rock bottom. I was fighting cancer, exhausted with life and hope was all I had left. I was offered two PhD positions on the same day and I can say with certainty that destiny brought me to Bodø. Truly and undoubtedly the best decision of my life! I am now in full-remission, got to work with my brilliant mentors, found my children, my family and a country to call home, after a decade of nomadic wandering. The support I have received has been incredible and I would like to take this opportunity to thank the many people who have made this monumental feat possible.

First and foremost to **Igor**, for believing in my abilities and choosing me for this project. Your guidance, calm supervision and genuine well-wishes have been vital to my progress. One cannot hope for a better supervisor, mentor and pillar of support than what I found in you. I am especially grateful to you for nurturing my teaching skills and providing me with every possible opportunity for self-development. Knowing that I can count on your advice at any point in time has been a real strength that I'm deeply grateful for. I will forever remember our long discussions and the scientific sensibilities imparted by you. Thank you for everything and I look forward to continuing our work together for a long time!

To **Henrik**, I could never have got here without the amount of wisdom and sheer inspiration that you have imparted. Working with you has been a joy, an honour and an experience I will cherish for a lifetime. Words pale in comparison to how grateful I am for your faith in my abilities. I will miss the sensory overload of excitement and nerves I experienced in the moments leading up to every single meeting with you. Thank you for being the pillar holding my PhD together!

To **Steinar**, I am constantly amazed at your enthusiasm for every aspect of academic research, your tireless efforts to elevate any scientific discussion, your sharp and quick humour with the perfect mix of sceptical cynicism. Every conversation and discussion

with you will be remembered fondly and every Watson-Crick base-paired structure figure will always remind me of our structural adventures. Thank you for the unforgettable journey into the RNA world!

To the **Research Council of Norway** for funding this study and providing me with this opportunity of a lifetime, together with **Nord University**. Thank you for making this PhD possible.

To **Nicolai**, my short time working with you at Panum serves as a reminder of all the reasons why I chose research. I would say you're the best research partner, peer, co-author and confidant in the world but I admit to being particularly biased. It has been a singular privilege having you show me the ropes and I look forward to renewing our research partnership in the near future. Thank you for your motivation and strength, it mattered more than you know.

To **Julien** and the team at **INRAe**, Rennes, for adding a whole new dimension to this scientific discourse and all your efforts in shaping the project into what it is today. Thank you for your continued collaboration.

To **Purush**, for all your support and guidance with the early stages of the project, especially in helping me learn about zebrafish embryos and staging as much as I did. To **Cesilie**, **Heidi**, **Joanna**, **Silje**, and the entire team at **Mørkvedbukta** research station for training and supporting me at the amazing zebrafish facility. To **Tor Erik**, for your unwavering technical assistance and incredible support. To **Terhi**, **Felix** and all my **groupmates**, I appreciate each and every one of you. Thank you for your support!

To the **faculty** and the **FBA administration**, particularly **Ann-Merete**, **Irene**, **Randi** and **Hilde**, I appreciate all your hard work. To **Kjetil**, **Jarle**, **Marit** and **Mette** for always taking the time to meet and discuss my research and my career, thank you all for your continued support and good wishes.

To all five cohorts of my **bachelor students**, I never knew I could be a teacher until I tried, and all your bright and eager groups have made teaching the one thing I aspire to for the future. Thank you for all your love, feedback and support.

To all my **friends** and **colleagues** at FBA, for giving me fond memories, interesting lunch hours and the quiet camaraderie of suffering the same fate together. I consider myself lucky to have been a part of such a supportive and uplifting PhD community. To **Kristine** and **Jeanett** for being splendid advisors to me and for all the guidance throughout my time here.

To **Kevin Ashton**, **Lotti Tajouri**, **Tony Badrick**, **Henning Boldt**, **Stine Petterson**, **Lars Martin Jakt**, **Monica Brinchmann**, **Joost Raeymaekers** and the many others who have inspired, shaped and propelled my education – special thanks to all of you.

To **Amma** for being an institution, a wall, my in-house therapist, my reality-check and entertainment. I could not have survived this without you. After everything we have been through this year, your quiet strength has kept me going in so many ways. I love you so so much!

To my children **Edward** and **George**, for being the best little boys that a PhD parent could dream of! It has been a rough few years but you two have kept me anchored to life. You are my joy, my sunshine, my gravity and my whole world. Amma loves you the most!

If you are lucky in life, you'd have a supportive work partner, if you are even luckier, you'd have a supportive life partner – I am lucky enough to have found both in the same person. To **William**, we have never needed words to communicate so I'll just say this – this is as much your milestone as it is mine. Look! we made it this far 😊 Thank you for keeping me going and here's to new adventures together!

Contents

Abbreviations

2'-*O*-Me	ribose methylation (RNA modification)
2D TLC	two-dimensional thin layer chromatoraphy
Box C	'CUGA' sequence, conserved element of SNORDs
Box D	'RUGAUGA' sequence, conserved element of SNORDs
bp	base pairs
ch	chromosome
cryo-EM	cryogenic electron microscopy
DNA	deoxy-ribonucleic acid
DNAzyme	DNA based ribozyme
gas5	growth arrest-specific 5
GREM	glucocorticoid receptor element mimic
Hpf	hours post fertilization
KD	gene knockdown (manipulation of gene/protein expression)
KO	gene editing by knockout of specific genomic regions
lncRNA	long non-coding RNA
LC-MS	liquid chromatography mass spectrometry
L(E)CA	last (eukaryotic) common ancestor
LSU	ribosomal large subunit
mRNA	messenger RNA
MS	mass spectrometry
MTases	methyltransferases
MYA	million years ago
MZT	maternal-zygotic transition
ncRNA	non-coding RNA
nt	nucleotide
piRNA	PIWI protein-interacting RNA

PTC	peptidyl transferase center
PTM	post-translational modifications
RNA	ribonucleic acid
RP	ribosomal proteins
RT-PE	reverse transcriptase based primer extension
rDNA	ribosomal DNA
rRNA	ribosomal RNA
'S' in rRNA subunit	Svedberg constant/Sedimentation rate
SAM	S-adenosyl-L-methionine
SELEX	Systematic Evolution of Ligands by EXponential enrichment
smORFs	small open reading frames
SNORA	box H/ACA snoRNA
SNORD	box C/D snoRNA
snoRNA	small nucleolar RNA
snoRNP	small nucleolar ribonucleoprotein
Sp.	species
SSU	ribosomal small subunit
tRNA	transfer RNA
TS-WGD	Teleost-specific whole genome duplication
UTR	untranslated region
WGD	whole genome duplication
Ψ	Pseudouridine (nucleotide type). Also represents pseudouridylation (RNA modification)

List of figures and tables

Figures

Tables

Abstract

Zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*) are prominent model organisms used to study human diseases, development and molecular mechanisms, owing to both organisms' high degree of genomic similarity with the human genome. Both organisms are particularly suited to laboratory inhabitation, with transparent embryos, frequent spawning, short generation turnaround and availability of established mutant lines being the most attractive reasons for their popularity as models in the fields of developmental and molecular genetics. Genomic and transcriptomic studies on these organisms have resulted in abundant data and insights. However, rRNA which exists at the core of the molecular machinery, has not been well-studied in these organisms, with zebrafish lacking defined rRNA secondary structures and medaka lacking even a complete rRNA reference sequence. Recent research on zebrafish revealed the presence of an early embryonic development ribosomal subtype, in addition to the ribosomal subtype expressed from late embryonic development onwards, and more conserved with other species. This duality has not previously been investigated at an epitranscriptomic level. Ribose methylation or 2'-*O*-Me rRNA modification profiling would aid in further comparison of the two rRNA subtypes at an epitranscriptomic level. 2'-*O*-Me modifications are guided by box C/D snoRNA (SNORDs) which are encoded in the introns of protein-coding or non-coding genes called host genes. The general objective of this study was to investigate the prevalence of 2'-O-Me in zebrafish and medaka rRNA during embryonic development with respect to the developmentally specialised ribosomal type. The specific objectives included (i) mapping 2'-*O*-Me sites in both the early and late subtypes of rRNA during zebrafish embryonic development; (ii) investigating the presence or absence of the specialised early-embryonic rRNA type in medaka; (iii) comparing the methylation profiles and guiding SNORDs of both organisms; and (iv) studying the phylogenetic variations of the prominent SNORD host gene *gas5* (which hosts up to 10 known SNORDs), across chosen teleost species. In this study, RiboMeth-seq was used to study zebrafish and medaka 2'-*O*-Me profiles and a SNORD-based non-experimental phylogenetic approach was used to study the *gas5*

host gene in multiple teleost species. This study resulted in the first ever rRNA 2'-*O*-Me maps of zebrafish and medaka rRNA, resolved rRNA secondary structures and the comparison of rRNA between both organisms.

The first part of the study describing the 2'-*O*-Me profile of zebrafish rRNA confirmed that the two rRNA types, in addition to different sequence lengths and composition, were also structurally different. The 2'-*O*-Me profile also differed between the two types of rRNA, supported by the identification of plausible SNORDs for most of the modified sites. The late-type rRNA had 98 high-scoring 2'-*O*-Me modified sites and was structurally more similar to mouse and human rRNA. The early-type rRNA with 97 modified sites was therefore proposed to be the specialised additional rRNA type. All teleost fish evolved from a common ancestor which underwent an additional round of whole genome duplication which could have contributed to the dual ribosomal types in zebrafish which belongs to the Teleostei infraclass. The second part of this study describing the 2'-*O*-Me profile of medaka rRNA served to further explore this hypothesis. Medaka being a teleost fish that diverged from zebrafish ~115 mya, was a distant enough relative to test whether the presence of an additional ribosomal type could be present in all teleost fish. Our results revealed that medaka embryos expressed only one type of rRNA throughout development, which was more similar to the late type rRNA of zebrafish. Our findings indicate that the specialised early type rRNA is not a common feature shared by all teleost fish but whether it is a feature shared by closer relatives of zebrafish remains to be seen. While compiling the list of plausible SNORDs for medaka 2'-*O*-Me sites, the much studied SNORD78 hosted by *gas5* in human, mouse and zebrafish, was observed to be absent in medaka along with the modification associated with it. To investigate this further, we employed a phylogenetic approach to compare the *gas5* host gene and particularly the SNORDs located in the introns of this gene across 19 teleost species. While SNORD78 was only missing in 2 of the chosen teleost species including medaka, multiple copies of SNORD77/80 were found in all species along with 'orphan' SNORDs in some species. This study highlights the unexplored area of SNORD-host gene relationship studies on

teleost species and the potential impact on evolutionary connections between teleosts that can be gleaned from many such studies. Functional SNORD studies on multiple teleost species are needed to fully understand the functions of specific SNORDs and the modifications guided by them.

Abstrakt – Sammendrag på norsk

Sebrafisk (*Danio rerio*) og medaka (*Oryzias latipes*) er framragende modellorganismer i studer av blant annet human sykdom, utviklingsbiologi, og for molekylære biologiske mekanismer generelt, dette på grunn av klare biologiske likhetstrekk til mennesket. Både sebrafisk og medaka er spesielt godt tilpasset studier i laboratorier på grunn av egenskaper som transperrante embryoer, hyppig egglegging, kort generasjonstid, og tilgjengelighet for etablering av mutante linjer. Slike egenskapene er blant de viktigste årsakene til den store populariteten som sebrafisk og medaka har fått som modellsystem innen utviklingsbiologi og molekylærgenetikk. Genom og transkriptom studier har resultert i store data-samlinger som også har bidratt til en betydelig innsikt. Men ribosomalt RNA (rRNA), som transkriberes og modnes i kjernen, har ikke blitt særlig godt studert i disse organismene. Sebrafisk mangler rapporterte definerte sekundære rRNA-strukturer, og medaka mangler til og med en fullstendig rRNA-referansesekvens. Nyere forskning på sebrafisk har vist tilstedeværelsen av en spesialisert ribosomal type i tidlig embryonal utvikling, dette i tillegg til den ribosomale typen uttrykt i sen embryonal utvikling og videre utover. Den sistnevnte synes også i større grad å være konservert med ribosomer i andre arter. To typer ribosom med ulike roller i utvikling har aldri tidligere vært observert hos fisk. Denne dualiteten i ribosom-organisering ble derfor studert på et epitranskriptomisk nivå. De generelle målene for dette PhD prosjektet omfattet derfor å forstå omfanget av variasjon mellom de to typene og for å stille hypoteser om den biologiske rollen i sebrafisk så vel som andre relaterte teleost-arter. De spesifikke målene inkluderte (i) kartlegging av 2'-*O*-Me-steder i både tidlig- og sen-embryonale rRNA-typer i løpet av embryonal utvikling av sebrafisk; (ii) undersøke tilstedeværelsen eller fraværet av den spesialiserte tidlig-embryonale rRNA-typen i medaka; (iii) å sammenligne metyleringsprofilene og de korresponderende SNORD RNA for sebrafisk og medaka; og (iv) studere de fylogenetiske variasjonene til SNORD koded av vertsgenet *gas5* (som kan ha opp til 10 SNORD RNA), i et sett utvalgte teleost-arter. I denne PhD studien ble RiboMeth-seq brukt til å studere 2'-*O*-Me profiler i sebrafisk og medaka, og en ikke-eksperimentell

fylogenetisk tilnærming ble brukt for å studere *gas5*-codet SNORD i flere teleost-arter. Dette resulterte i den første kartleggingen av 2'-*O*-Me seter i rRNA hos sebrafisk og medaka, generering av rRNA sekundære strukturer, og en sammenligning av rRNA mellom disse organismene.

Den første delen av studien, som beskriver 2'-*O*-Me-profilen til sebrafisk-rRNA, bekreftet at de to rRNA-typene var forskjellige både i primær og sekundær struktur. 2'-*O*-Me-profilen var også ulik mellom de to rRNA typene, en observasjon støttet av identifiseringen av plausible SNORD-er for de fleste av de modifiserte stedene. Den sene typen rRNA hadde 98 2'-*O*-Me modifiserte seter og var strukturelt mer lik den vi finner i mus og humant rRNA. Den tidlige typen rRNA med 97 modifiserte seter ble derfor foreslått å være den spesialiserte ekstra rRNA varianten. Alle teleoster har utviklet seg fra en felles stamfar som har gjennomgått en ekstra runde med hel-genom duplisering, og som dermed kunne ha bidratt til de to ribosomale typene i sebrafisk. Denne hypotesen ble videre utfordret i den andre delen av studien, som beskriver 2'-*O*-Me-profilen til medaka rRNA. Medaka, som er en teleost som divergerte fra sebrafisk ~115 mya, er en fjernt beslektet art som egner seg godt til å teste hypotesen av en ekstra ribosomal type. Resultatene våre viste at medaka-embryoer bare uttrykte en type rRNA gjennom utviklingen, og som var mer lik sen-type rRNA av sebrafisk. Funnene våre indikerer derfor at den spesialiserte tidlige typen rRNA ikke er en felles for alle teleoster, men om dette deles av arter som står nærmere sebrafisk gjenstår å se. Mens vi sammenfattet listen over plausible SNORD-er for medaka 2'-*O*-Me-seter, ble den mye studerte SNORD78 fra *gas5* genet i mennesker, mus og sebrafisk observert å være fraværende i medaka, og dette sammen med modifikasjonen forbundet med SNORD78. For å undersøke dette videre valgte vi en fylogenetisk tilnærming for å sammenligne *gas5* vertsgenet og SNORD-ene lokalisert i intronene til dette genet i 19 teleost-arter. Mens SNORD78 bare manglet i 2 av de valgte teleost-artene, inkludert medaka, ble det funnet flere kopier av SNORD77/80 i alle arter. Hos noen arter var slike SNORD77/80 assosiert med "foreldreløse" (orphant) SNORD-er. Denne studien understreker at SNORD-verts-gen forholdsstet ennå er lite utforsket i teleostarter,

inkludert ulike evolusjonære aspekter. Funksjonelle studier av SNORD som involverer flere teleost-arter er derfor nødvendige for å bedre kunne forstå funksjonene til spesifikke SNORD-er og de modifikasjonene som er assosiert med disse.

1 Introduction

During the quest to fully sequence the human genome, we were finally able to grasp the sheer depth of the genetic code and subsequently the multitude of complex interactions that make us the unique individuals that we are. In the nearly twenty years since the first comprehensive version of the human genome was made available from the human genome project, this quest has resulted in unprecedented technological advancements which were previously deemed improbable. We have also been left with many unanswered questions and more data than we know how to interpret. We can now sequence genomes in hours, fragment, amplify, modify or remove parts of a genome, and quantitatively study the resulting effects. Beyond studying merely genomic interactions, decades of research efforts to study DNA and RNA biology and the various regulatory mechanisms involved in biological processes at a molecular level, have resulted in advanced investigative methods and insights into the complex relationship between DNA, RNA and protein, in great detail.

Table 1: Prominent 'Omics' of biological research

Omics	Focus area	Broad categories
Genomics	DNA and genes	structural, functional and comparative
Epigenomics	Transcriptional regulation	chemical modifications of DNA and histones
Transcriptomics	mRNA, rRNA and tRNA	gene expression and risk-prediction
Epitranscriptomics	Post-transcriptional regulation	chemical modifications of RNA
RNomics	Non-coding RNA	regulatory mechanisms and structure
Proteomics	Proteins	structure, binding and interactions

Omics research has been heavily invested upon, both in terms of time and resources **(Table 1**). This research project is mainly concerned with epitranscriptomics and RNomics. A good deal of our knowledge and advancement has come through research initiatives into many RNA types (**Table 2**), which are instrumental in facilitating the journey from genes to proteins and beyond.

Table 2: [1]Types of RNA

RNA type	Full name	Size	Location	Function	Illustration
[2]**mRNA**	messenger RNA	[3]>400 nt	nucleus/cytoplasm	DNA to protein intermediary	
[5]**rRNA**	ribosomal RNA	[4]100 to >4000 nt	nucleolus/cytoplasm/ mitochondria	protein synthesis	
[5]**tRNA**	transfer RNA	~80 nt	cytoplasm/mitochondria	protein synthesis	
[6]**snRNA**	small nuclear RNA	150 nt	nucleus/cajal body	splicing	
[6]**snoRNA**	small nucleolar RNA	60 to 300 nt	nucleolus	rRNA modification	
[7]**miRNA**	micro RNA	~22 nt	nucleus/cytoplasm	gene regulation	
[7]**piRNA**	PIWI-interacting RNA	~27 nt	nucleus/cytoplasm	RNA silencing	
[7]**lncRNA**	long non-coding RNA	>200 nt	nucleus/chromatin	epigenetic regulation / scaffolding	various

Illustrations created using bioRender.com

[1] Non-exhaustive overview of the most prominently studied RNA types
[2] involved in protein coding information
[3] varies depending on the gene
[4] varies between rRNA subunits of different species
[5] housekeeping - translation
[6] housekeeping/ regulatory – modification and splicing
[7] regulatory – small/long

1.1 Ribosomes, rRNA and translational control

The value of almost every gene is inherently linked to the proteins it codes for. This gene-protein relationship is the basis of the biological concept known as the central dogma (**Figure 1**), which is divided into transcription and translation. Translation is the process by which the information encoded in the messenger RNA (mRNA) is converted into a meaningful output for the organism in the form of proteins (Schuller and Green 2018). Translation has been studied in great detail, including the various regulatory mechanisms affecting the translation machinery (Ramakrishnan 2002; Simonović and Steitz 2009). Ribosomes form the core of the entire translation machinery and are composed of ribosomal proteins (RPs) along with ribosomal RNA (rRNA), which are critical to protein synthesis (Decatur and Fournier 2003; Zhou et al. 2015). Ribosomes are present in all living organisms and all ribosomal genes must therefore have evolved from a common ancestor. The intrinsic value of ribosomes in facilitating protein synthesis was discovered in the late 20th century when the large subunit (LSU) rRNA was confirmed to be the catalyst for peptide bond formation (a peptidyl transferase), thus earning ribosomes the name 'ribozymes'. This critical functional role explained why the region of LSU rRNA, now known as the peptidyl transferase centre (PTC), is one of the most conserved rRNA sections in the last common ancestors of multiple organisms. Ribosome biogenesis has been extensively studied and, while a lot of the studies have contributed to a general understanding of the process, many mechanisms are still unclear (Baßler and Hurt 2019; Kumar 2021). Eukaryotic ribosome biogenesis is mediated by hundreds of assembly factors and over 75 snoRNAs, and occurs mostly in the nucleolus (Henras et al. 2008; Aubert et al. 2018).

rRNA has long been considered the unit of evolutionary distance between any two species (Mindell and Honeycutt 1990; Hillis and Dixon 1991; Olsen and Woese 1993). Majority of the ribosomal mass is made up of rRNA but the proportion can vary between genetic systems. rRNA subunits are named using Svedberg Units (S), which is the measure of rate of sedimentation. The genes encoding all the rRNA subunits are

located within the nucleolus (Bachellerie et al. 2002; Elliot and Ladomery 2016). These rDNA sequences are located as tandem repeats in the nucleolus organiser regions (NORs) within the nucleolus (McClintock 1934), and the size and number of nucleoli present in the cells of a particular species is dependent on ribosome biogenesis activity (Pederson 2011). rDNA has also been shown to shape nucleolus-associated chromatin domains or NADs (Picart-Picolo et al. 2019). During ribosome biogenesis in eukaryotes, rRNA synthesis can be described in five stages: pre-RNA and RP synthesis, pre-rRNA base modifications, pre-rRNA folding, pre-rRNA assembly with RPs, and pre-rRNA processing (Baßler and Hurt 2019).

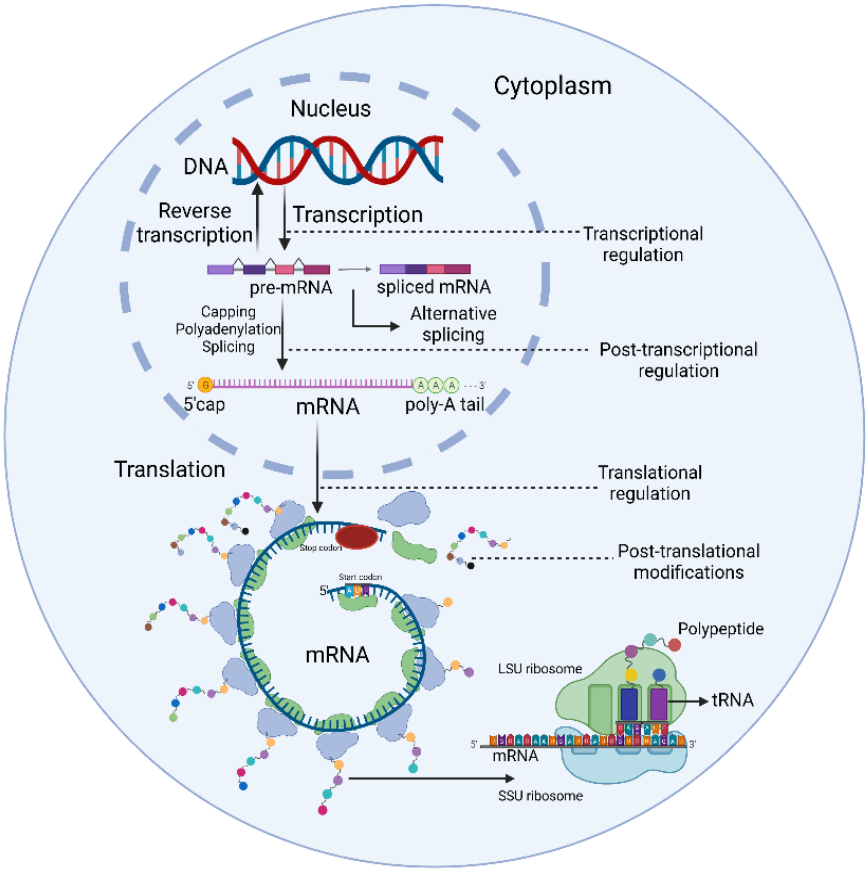


Figure 1: Central dogma of biology. Overview of the steps involved in the central dogma, including transcription, translation and intermediate regulatory steps. Illustration created using bioRender.com

rRNA begins as a pre-rRNA, which is transcribed by polymerase I into 28S, 18S, and 5.8S rRNA (**Figure 2**), which later undergo extensive post-transcriptional processing (Tschochner and Hurt 2003; Henras et al. 2008). The 5S rRNA is transcribed from a separate gene by a different enzyme, RNA polymerase III (Thompson et al. 2003). Before the different subunits are separated, the external and internal spacer sequences split the pre-rRNA molecules (Engel et al. 2013; Fernández-Tornero et al. 2013). At this stage, the rRNA base modifications occur (Monaco et al. 2018), without which the rRNA assembly and folding would not be feasible (Zemora and Waldsich 2010). rRNA folding (Rodgers and Woodson 2021) is supervised by assembly proteins and, while a lot remains unknown about the early stages of this process, a canonical SNORD (box C/D small nucleolar RNA) has been shown to function as a chaperone in the assembly process (Aquino et al. 2021). rRNA processing in the nucleolus is aided by small nucleolar RNA (snoRNA) complexes to stabilise the interaction between rRNA and ribosomal proteins and to protect the binding sites (Woodson 2011).

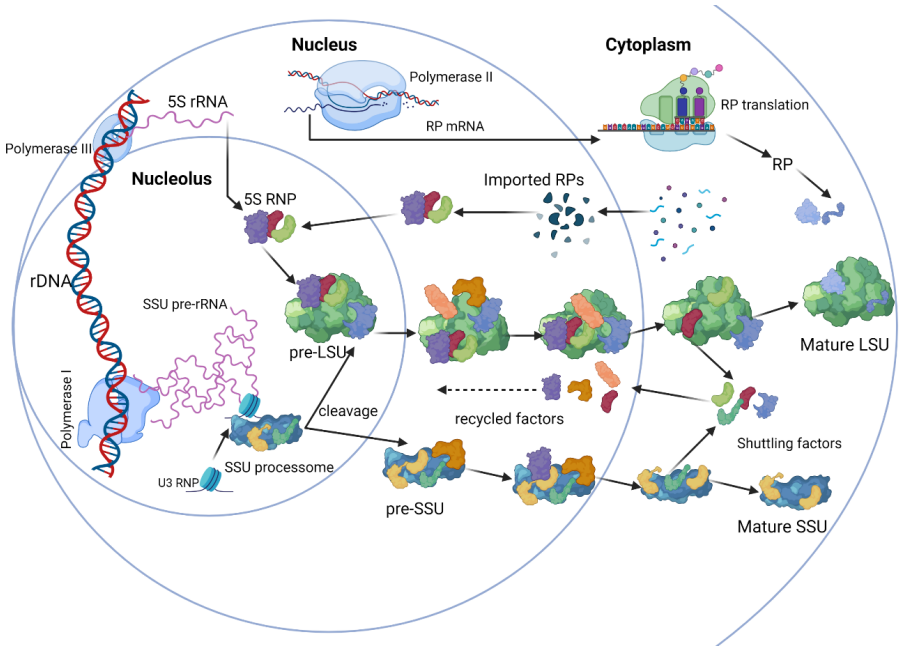


Figure 2: Illustrated representation of ribosome biogenesis – representing the ribosomal SSU and LSU from rDNA to a mature subunit. Created using bioRender.com

Most of the events associated with translational control occur during translation initiation and proceeds in several steps (Sonenberg and Hinnebusch 2009; Hinnebusch 2014, 2017). The mature rRNA contains specialised regions and the conformational changes in these rRNA regions are critical to protein synthesis (Ning et al. 2014; Sloan et al. 2017), making the rRNA structure itself a factor of translational control. A part of the small subunit is termed the decoding centre, which is necessary to read the genetic information encoded in the mRNA, while the large ribosomal subunit (LSU) contains the peptidyl transferase centre (PTC), where the enzyme peptidyl transferase is the LSU rRNA itself (Stahl et al. 2002). The canonical translation factors such as eukaryotic initiation, elongation and termination factors (eIFs, eEFs and eTFs), are involved in differential control of translational activity for a subset of mRNAs in addition to globally regulating translation (Dalla Venezia et al. 2019). Further modifications to the rRNA such as pseudouridylation (Ψ), which occur during the maturation phase and these modifications, capable of regulating tRNA and mRNA translational access, contribute to a physical layer of gene expression control (Sharma and Lafontaine 2015; Penzo and Montanaro 2018). Such modifications are mostly guided by the direct binding of a snoRNA antisense element to the region of interest to be modified.

1.2 Functionally specialised ribosomes

In eukaryotes, genes coding for rRNA are among the most conserved. Scientists have attempted to explain the functions of ribosomes in different ways. The 'abundance model' proposes that the initiation rates of different transcripts are dependent on the degree of availability of free ribosomes. This theory is limited, as it would mean that all proteins are unidirectionally regulated and that the total ribosomal content is prioritised over structure and composition. The 'specialised ribosome' model proposes that the structure and composition of ribosomes exhibit stoichiometry and these variations have functionally specific roles in translation. This would mean that ribosome heterogeneity is possible, along with specialised ribosomes. Studying ribosome specialisation could hold the key to explaining mRNA specificity and the

regulatory benefits of said ribosomes. It is not sufficient to just demonstrate altered RP stoichiometry but to identify the changes that can exclusively be attributed to ribosomes, in order to prove functionally specialised ribosomes (Emmott et al. 2019).

Evolution of ribosomes has contributed to changes in the ribosomal composition and structure, which align with the varying complexity of proteins and preferred translation mechanism as a result (Gay et al. 2022). For many years, ribosomes were solely regarded as highly structurally conserved housekeepers, but recent research suggests that the slightly modified ribosomes (presence or absence of specific rRNA or RP sections), could selectively translate specific sets of mRNAs. That the ribosome is not just a framework machinery but an intricate and dynamic complex which can influence translational regulation, has been inferred from studies demonstrating that post-translational modifications (PTMs) can alter the structural stability and interaction of the modified substrate, and in turn the translational output (Simsek and Barna 2017; Jansson et al. 2021). However, a complete functional understanding of this process is yet to be achieved (Gay et al. 2022).

When a ribosome with altered composition can influence translational output, either by effecting changes in translational fidelity or favouring translation initiation of specific mRNAs, in addition to its core duties in translation, it is said to be functionally specialised (Fujii et al. 2018; McMahon et al. 2019). Unlike transcriptional control which involves a large transcription machinery, PTM of RPs is a single step process. This additional quick layer of regulation could link metabolic conditions to translation states in the cells (Lee et al. 2015). Additionally, every organism contains multiple rDNA repeats with probable mutations. Given that the non-functional mutations will be selected against, if enough functional ribosomal changes were evolutionarily selected for over time, it could potentially give rise to the ribosome code. This is mainly due to the fact that regulatory proteins are encoded by poorly-translated and scarce mRNA, making them susceptible to regulation by specialised ribosomes (Ferretti and Karbstein 2019). Due to the potentially stable intermediates of ribosomal degradation, it is also

possible that what is believed to be specialised ribosomes could just be partially degraded, non-functional ribosomes (Ferretti and Karbstein 2019).

1.3 Ribose 2'-*O*-methylation

Ribose 2'-*O*-methylation or 2'-*O*-Me, is one of the most prevalent RNA modifications and has been documented in tRNA, rRNA, snRNA and mRNA. Unlike other known methylation modifications which occur on specific bases (Kumar and Mohapatra 2021), 2'-*O*-Me occurs at the second carbon of ribose and is independent of the base (**Figure 3**). 2'-*O*-Me in bacterial (Persson 1997) and eukaryotic (Towns and Begley 2012) tRNA, is mostly catalysed by stand-alone enzymes known as methyltransferases (MTases), while modifications in some archaeal species are guided by snoRNA (Grosjean et al. 2008). MTases act by transferring a methyl group from a donor molecule to the intended target, including ribose residues (Schubert et al. 2003; Boschi-Muller and Motorin 2013). The most common methyl donor is the S-adenosyl-L-methionine or SAM (**Figure 3**) and MTases are generally classified into six types based on the methylation target, five of which use SAM as the donor (Schubert et al. 2003).

A methylation profile/map is a list of all identifiable modified residues in a specific RNA type for a given species. 2'-*O*-Me profiles can include information on the methylation stoichiometry. The tRNA methylation profiles of some bacteria (Boccaletto et al. 2018), archaea (Takeda et al. 1994; Renalier et al. 2005; Kuratani et al. 2008), yeast (Droogmans and Grosjean 1991; Wilkinson et al. 2007; Kotelawala et al. 2008), and vertebrates (Glasser et al. 1992; Païs de Barros et al. 1996; Sprinzl and Vassilenko 2005), have been studied. The presence of such modifications in specific tRNA sites in archaea contribute to tRNA stability that is critical to tRNA function (Kawai et al. 1991). There are limited data available for 2'-*O*-Me in snRNAs, of which spliceosomal snRNAs are the most studied (Krogh et al. 2017b). The presence of 2'-*O*-Me in the cap-proximal nucleotides of mRNA has been uncovered (Smietanski et al. 2014) while emerging evidence indicates the possible presence of ribose 2'-*O*-Me in the internal nucleotides of mRNA (Dai et al. 2017). Further studies on snRNA and mRNA modifications are

necessary to fully understand the structural and functional implications of 2'-*O*-Me on these RNA types. Contrastingly in rRNA, most eukaryotic and archaeal 2'-*O*-Me modifications depend on snoRNA guides (Lane and Tamaoki 1967; Lestrade 2006; Henras et al. 2008; Graziadei et al. 2016; Yang et al. 2016), with very few exceptions guided by stand-alone enzymes (Bonnerot et al. 2003; Lapeyre and Purushothaman 2004; Tomkuvienė et al. 2017; Chou et al. 2017). Although many functional roles of this chemical modification are being uncovered, rRNA 2'-*O*-Me modifications have been shown to influence RNA stability, translational regulation, development, and cancer (Monaco et al. 2018).

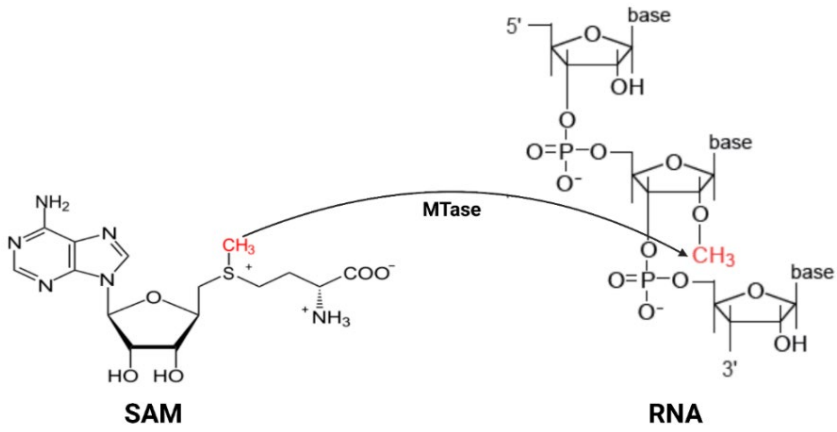


Figure 3: Representation of the ribose 2'-O-Me modification

While 2'-*O*-Me was originally classified as a nuclear event, it has now been demonstrated in yeast as a co-transcriptional event (Koš and Tollervey 2010). This chemical modification affects the rRNA by rendering each site where the 2'-OH is methylated, resistant to alkaline hydrolysis, and at low dNTP concentrations by inhibiting the retro-transcription of RNA to cDNA (Monaco et al. 2018). This resistance to alkaline hydrolysis facilitated the development of high-throughput methods to study

these modifications in more detail across various species (Krogh and Nielsen 2019). The implications of 2'-*O*-Me modifications on RNA structure are not yet resolved. While it is known to contribute to the stabilisation of nucleotide conformation and restrict strand conformation and flexibility by restricting the rotational freedom of 3'-phosphates, further investigation of this field has remained limited (Monaco et al. 2018).

rRNA 2'-*O*-Me profiles have been studied in many species. It has been shown that bacterial rRNA has the least number of modifications with only four highly conserved sites established (Sergeeva et al. 2015). Predictive mapping of 2'-*O*-Me in the rRNA of archaeal species has shown the presence of 2'-*O*-Me occurring in clusters in functionally critical regions, although these remain to be experimentally validated (Dennis et al. 2015; Seistrup et al. 2016). The first complete rRNA 2'-*O*-Me map was produced in yeast, *Saccharomyces cerevisiae*, and 55 2'-*O*-Me sites have been described (Klootwijk and Planta 1973; Birkedal et al. 2015). Following this, the 2'-*O*-Me profiles of protist (*Euglena gracilis*), plant (*Arabidopsis thaliana*), worm (*Caenorhabditis elegans*) and frog (*Xenopus laevis*) rRNA have been partially mapped (Maden 1986, 1988; Barneche et al. 2001; Higa et al. 2002; Schnare and Gray 2011) while in human (Krogh et al. 2016) and mouse (Hebras et al. 2020), the rRNA 2'-*O*-Me maps are complete. Based on these 2'-*O*-Me profiles, some sites have been shown to be largely conserved across species. More importantly, the variation in methylation levels at some sites under specific conditions as evidenced in mouse (Hebras et al. 2020) and human (Krogh et al. 2016), support ribosome heterogeneity and ribosome specialisation.

1.4 snoRNA and rRNA modifications

rRNA biogenesis is aided by snoRNAs, which also guide site-specific modifications of rRNA such as 2'-*O*-Me and pseudouridylation (Ψ) (Yip et al. 2013). To achieve these modifications, snoRNA act as the antisense guide RNA and hybridise directly to the corresponding rRNA sites. Categorised as non-coding RNA, snoRNA are abundantly

present and up to 200 known snoRNAs have been identified depending on the species (Reichow et al. 2007). snoRNA-like molecules have also been observed in archaea suggesting the ancient nature of snoRNA evolution (Omer et al. 2003). Two unique types of snoRNA exist, to serve as guides to each type of site-specific modification. The Box H/ACA snoRNA or SNORA (**Figure 4**) guides the pseudouridylation modification and the Box C/D snoRNA or SNORD guides the 2'-*O*-Me modification (**Figure 4**). Both types are easily recognisable due to their distinct structural composition (Elliot and Ladomery 2016). SNORAs are made up of 2 hairpin loop structures connected by a hinge made up of the H box which is characterised by the sequence profile 5'-ANANNA-3' where N can be any nucleotide and 3nt from the 5' end of the SNORA exists the conserved 5'-ACA-3' sequence. The guide sequences for the pseudouridylated sites exist in the inner internal loops of the hairpins (Omer et al. 2003).

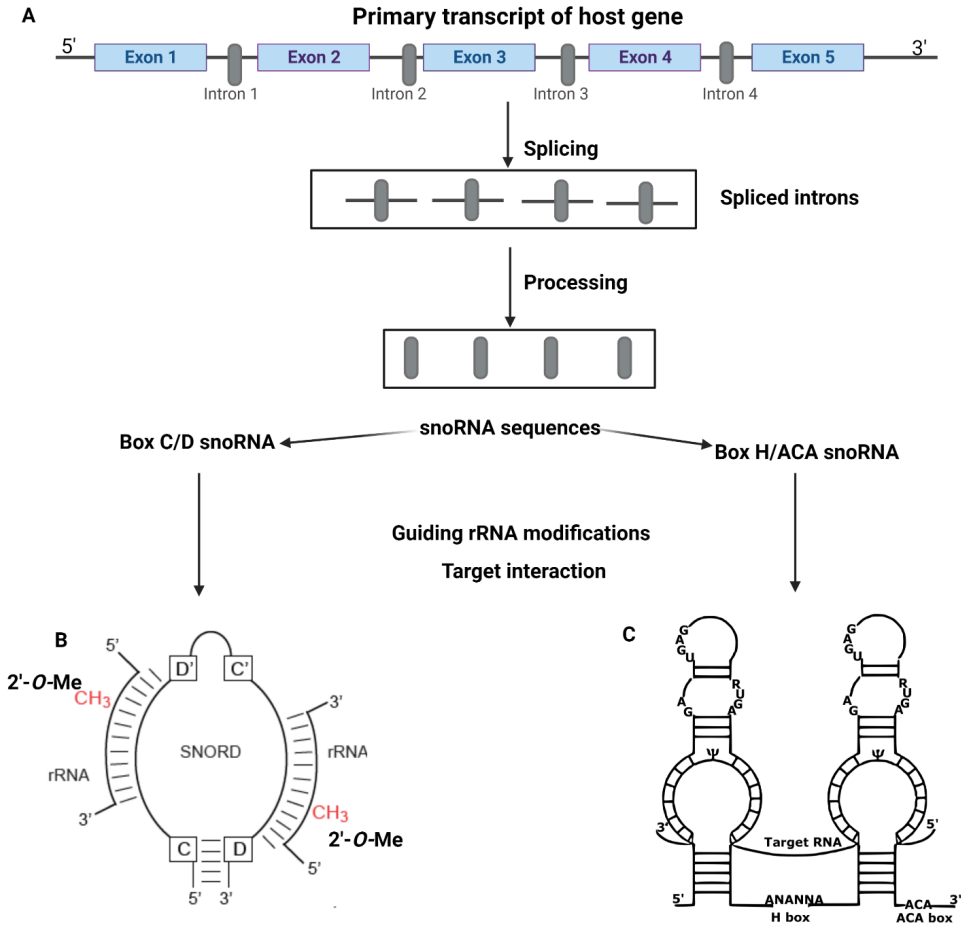


Figure 4: snoRNA-rRNA interaction – Illustration describing the relationship between a host gene and its snoRNA which guides a rRNA modification. A. snoRNA encoded in the introns of a host gene, schematic representation not to scale. B. Box C/D snoRNA (SNORD) interaction with target rRNA resulting in a 2'-OH methylation modification. C. Box H/ACA snoRNA (SNORA) interaction with target rRNA resulting in pseudouridylation. Created using bioRender.com.

SNORDs are characterised by a set of short sequences termed the 'D box' and the 'C box'. D boxes are short, only 4nt in length, and are often 5'-CUGA-3'. C boxes are longer, 7nt in length, and their sequence is mostly either 5'-AUGAUGA-3' or 5'-GUGAUGA-3'.

In addition to the well conserved C and D boxes, SNORDs also have less conserved C' and D' boxes in between the C box and D box. The C boxes and their respective D boxes are structurally separated by a stem loop. SNORDs guide 2'-*O*-Me rRNA modifications as the antisense element to the rRNA methylation site is always found upstream of the D and D' boxes (**Figure 4**), and these boxes can each guide a methylation at different sites in the rRNA (Bachellerie et al. 2002; Dieci et al. 2009; Jorjani et al. 2016).

snoRNAs have varied and unique functions. In addition to guiding ribose methylations, a few SNORDs have been observed to guide rRNA acetylation (Sharma et al. 2015, 2017). Many canonical SNORDs do not appear to guide any known rRNA modifications and are called 'orphan SNORDs' (Bratkovič et al. 2020). Increasing number of research studies report other functional roles for some orphan SNORDs as well as guiding SNORDs such as chromatin remodelling, polyadenylation and regulation of gene expression (Falaleeva et al. 2017; Bratkovič et al. 2020; Bergeron et al. 2020). It has been observed that, contrary to their name, some functionally specialised SNORDs function outside of the nucleolus (Holley et al. 2015; Rimer et al. 2018). In early ribosomal processing, the U3 snoRNA helps to cleave and remove the spacers in the transcript. Erratic RNA cleavage has been observed as a result of depleting U3 (Dragon et al. 2002). snoRNAs often exist in snoRNPs (small nucleolar ribonucleic protein) complexes and can be quite large. For example, the U3 snoRNP consists of up to 28 proteins in addition to the U3 snoRNA and has been shown to sediment at the same rate as a fully formed ribosome (Dragon et al. 2002).

Most of these non-canonical functions have not been described for SNORAs which could be due to the lack of methods that can accurately detect and quantify SNORAs (Bergeron et al. 2020). The links between SNORDs and human diseases are being uncovered more and more indicating the importance of SNORDs in regulatory mechanisms (Cavaillé 2017; Liang et al. 2019; Schaffer 2020; Aquino et al. 2021). Ribosome activity is directly affected by the presence and number of chemical modifications. The increased hydrophobicity from 2'-*O*-Me and additional H bonding

from pseudouridylation affects ribosome structure and stability. Links of snoRNA loss to diseases such as cancer, Prader-Will syndrome and X-linked dyskeratosis congenita, have been established and some of these appear unconnected to ribosome function (Mannoor et al. 2012).

1.5 SNORDs and host genes

The SNORD-host gene relationship is an important aspect in studying the conservation of SNORDs and the resulting ribose methylations across species. In eukaryotic genomes, SNORDs are found to be encoded in the intronic regions (Dieci et al. 2009) of specific genes (**Figure 4**). These genes that encode SNORDs are therefore termed 'host genes'. Over a hundred known host genes have been identified, of which some genes are non-protein coding genes/genes of unknown function. Some host genes appear to host only a single known SNORD while some others may encode over 10 SNORDs (Bachellerie et al. 2002). **Table 3** highlights the most notable host genes and the maximum number of SNORDs hosted by each in zebrafish.

Although the SNORD-host gene relationship is well conserved, the non-protein coding host genes tend to be poorly conserved overall with a well-conserved underlying SNORD architecture (Deogharia and Majumder 2018). It was initially proposed that the expression of SNORDs was directly proportional to the host gene transcription (Dieci et al. 2009; Yang 2015); however, emerging evidence suggests that SNORD expression could be preserved even upon degradation of host gene transcripts brought on by nonsense mediated decay (Lykke-Andersen et al. 2016; Kufel and Grzechnik 2019). This supports the further observations in human and mouse cancer cells, where a limited correlation is seen between SNORD expression and host gene transcripts (Boivin et al. 2018; Warner et al. 2018; McCann et al. 2020).

Table 3: Top five genes that host the most SNORDs in zebrafish

Host gene	Gene name	No. of SNORDs
gas5	Growth arrest specific-5	10
arl2	ADP-ribosylation factor-like 2	8
rpl13a	Ribosomal protein L13a	5
Nop56	Nucleolar protein 56	4
rpl17a	Ribosomal protein L17a	4

There is also evidence of tissue-specific snoRNA expression (Cavaillé et al. 2000; Jorjani et al. 2016), which has been an area of limited research activity in the past. In an attempt to resolve this limitation, new snoRNA-sensitive techniques of RNA-seq have been described (Boivin et al. 2018), which could aid in SNORD expression studies in correlation to host gene expression and tissue specificity. Such a study of the human snoRNome was recently published (Fafard-Couture et al. 2021), which shows that many snoRNAs are not uniformly expressed across all tissue types and that a subgroup of snoRNAs possess distinct gene organisation which could functionally contribute to both tissue-specific and housekeeping roles. Most of the known information about snoRNA has been gleaned from studies on yeast but this is still quite an incomplete understanding. Therefore, host genes play a key role in understanding the evolution of snoRNA, its conservation and in speculating the potential function of said snoRNA based on the host-gene function (Richard and Kiss 2006).

1.6 Notable snoRNA studies

Advances in investigative strategies and increased research interest have led to the identification of hundreds of novel snoRNAs (Jorjani et al. 2016) and some of these have been studied to affect not just ribosome biogenesis but also to possess regulatory functions over cellular processes. The base-pairing dependent guiding has made it easier to create databases and search strategies using the consensus motifs together with the guide regions, thereby leading to the identification of other snoRNA that not only relate to rRNA but also to snRNA. Many such strategies have been employed on

the transcriptomes of various organisms creating a putative map of snoRNA evolution and host-gene conservation, as well as identifying the inactive or orphan snoRNAs (Danny Bergeron et al. 2020). Several tools and databases have become available in the recent years which have facilitated mapping and comparative genomic studies of snoRNA across multiple species (Lestrade 2006; Makarova and Kramerov 2011; Zhang et al. 2012; Yoshihama et al. 2013; Bartschat et al. 2014). This area of research is actively being updated with new findings and tools.

In vertebrates, most SNORDs are encoded in the introns, not exons, and are therefore processed out of intron lariats. Some SNORDs encoded in unusually stable lariats have been shown to affect snoRNA maturation instead of guiding a modification (Talross et al. 2021). Contrarily, most yeast snoRNAs are independently transcribed and processed by exo- or endonucleases with only seven snoRNAs encoded in introns (Vincenti et al. 2007).

In plants, snoRNA distribution in several species have been mapped where 20 monocot specific SNORD families were identified while no SNORAs were similarly identified. No such specificity was observed in SNORDs and SNORAs of dicots. This could either be a result of extreme divergence, true gene loss or incomplete genome assemblies. The snoRNA target prediction reaffirms the conserved nature of the targets of conserved snoRNAs (Bhattacharya et al. 2016). From an evolutionary perspective, snoRNAs pre-date eukaryotes due to the presence of similar molecules in archaea (Omer et al. 2003). Another study aimed to trace back snoRNA families to the last eukaryotic common ancestor (LECA) and also investigated whether snoRNAs retained their intronic location across the eukaryote tree. The results confirmed that snoRNA families could be traced back to the LECA and that some LECA snoRNAs guided positionally conserved modifications in archaea. The intronic position of snoRNAs, though conserved, was not stable over large timescales of evolution (Hoeppner and Poole 2012).

Investigative studies on the co-evolution of snoRNA and their targets in vertebrates showed that a large number of species did not have a complete 28S rRNA reference

sequence available and found it impossible to therefore complete LSU target guiding snoRNA studies (Kehr et al. 2014). This study managed to assign putative functions for 9 orphan SNORDs. Their framework (Interactive conservation index or ICI) makes it possible to investigate if potential snoRNA mobility and conservation of snoRNA function are indeed unrelated.

A nearly complete set of snoRNAs has been annotated for the giant oocyte nucleus of *Xenopus tropicalis* (Deryusheva et al. 2020). A comparison of these data with nine other vertebrate species including zebrafish, mouse and human, revealed that not all rRNA modifications were evolutionarily conserved and variations in patterns could be seen even between closely related species. It was also shown that the exact modified nucleotide position within functionally important regions could vary in different species. This led to the hypothesis of cross-talking modifications as a driving factor in the evolutionary conservation of snoRNAs (Deryusheva et al. 2020).

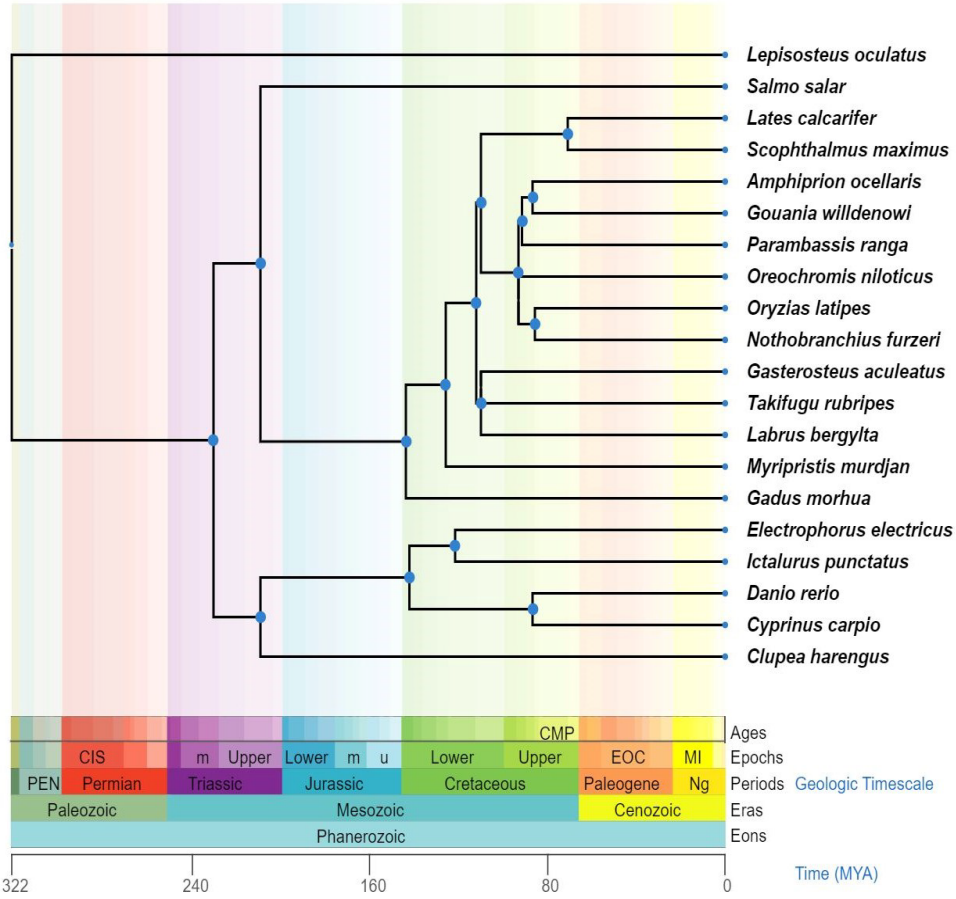
Offering an alternative dimension to snoRNA studies, a new class of dual-initiation promoters in snoRNA host genes has been uncovered which selectively generate snoRNA and host genes with varying spatio-temporal expression (Nepal et al. 2020). snoRNA genes in zebrafish are transcribed from non-canonical YC-initiation in addition to the canonical YR-initiation. These findings suggest that there exists an ancient evolutionary shared promoter architecture based on the dual-initiation promoter genes (Nepal et al. 2020). From the current knowledge, fish species appear to be a particularly attractive group of organisms to study snoRNA in evolutionary contexts which also happens to be a largely unexplored area of research in prominent fish species.

1.7 Teleostei infraclass and its evolutionary significance

In the animal kingdom, under the sub-phylum vertebrata, the Actinopterygii class covers all the ray-finned fish species which makes up over 50% of all known vertebrate species. Actinopterygii is further classified into two sub-classes: Chondrostei and

Neopterygii. Neopterygii further diverged into Teleostei infraclass 310 mya and the Holostei infraclass 275 mya (**Figure 5**). While holosteans broadly consist of bowfins and gars, teleosts make up over 96% of all known fish species (Helfman et al. 2010). The diversity of teleost species across habitat, size, behaviour, reproductive systems, nutrition and value to civilization, renders it very pertinent to scientific inquiry. Teleosts are entwined in human civilization far more intrinsically than we may realise. Some teleost species such as Atlantic salmon (*Salmo salar*), Eurasian carp (*Cyprinus carpio*), Nile tilapia (*Oreochromis niloticus*), European bass (*Dicentrarchus labrax*), brown trout (*Salmo trutta*) and turbot (*Scophthalmus maximus*) are among the most prominently farmed in aquaculture industry (Naylor et al. 2021). A large proportion of the global human population relies heavily on cultured fish as a food source.

Some other teleost species are used as accessory fish in the aquaculture industry but are not directly consumed by humans, e.g. Ballan wrasse (*Labrus bergylta*) is used to remove lice from salmon in farms. Apart from farmed fish, thousands of teleost species such as Atlantic herring (*Clupea harengus*), Atlantic cod (*Gadus morhua*) are caught wild and consumed in different regions of the world. In contrast, some teleost species have a high economic value due to their purely ornamental nature. Clownfish, goldfish, guppies, clingfish, glassfish, rice fish and many others are highly sought after for their small size, colourful appearance and aquarium friendly life cycle. In addition to their economic and commercial value to humans, over 26,000 known teleost fish species are key players in marine and freshwater ecosystems, both as predators and as prey (Helfman et al. 2010).



on timetree.org.

Figure 5: Evolutionary divergence and timeline of the fish species described in this study beginning with the common node of *Actinopterygii*. This figure has been generated using the TimeTree algorithm

Table 4: Summary of the teleost species included in the present study.

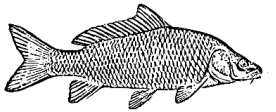
Scientific name	Common name	Superorder	Location	Avg adult size	Illustration
Clupea harengus	Atlantic herring	Clupeomorpha	Atlantic ocean	45 cm	
Danio rerio	zebrafish	Ostariophysi	South Asia	2-3 cm	
Cyprinus carpio	Eurasian carp	Ostariophysi	Southern Europe West Asia	40-80 cm	
Electrophorus electricus	electric eel	Ostariophysi	South America	200 cm	
Ictalurus punctatus	channel catfish	Ostariophysi	North America	30-60 cm	

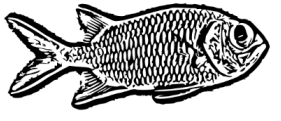
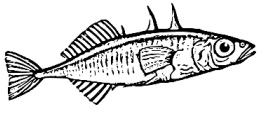
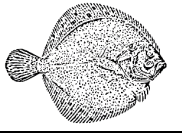
Table 4, contd.					
Salmo salar	Atlantic salmon	Protacanthopterygii	North Atlantic	100-120 cm	
Gadus morhua	Atlantic cod	Paracanthopterygii	Arctic North Atlantic	100-140 cm	
Amphiprion ocellaris	clown anemonefish	Percomorpha	West Pacific East Indian Ocean	9-11 cm	
Myripristis murdjan	pinecone soldierfish	Acanthopterygii	Indo-Pacific	18-25 cm	
Gasterosteus aculeatus	stickleback	Acanthopterygii	North Atlantic NW Pacific	5-10 cm	
Scophthalmus maximus	turbot	Acanthopterygii	Baltic Mediterranean NE Atlantic	50-80 cm	

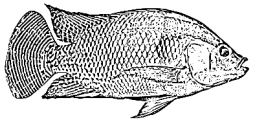
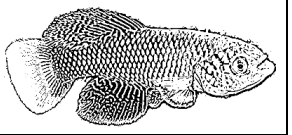
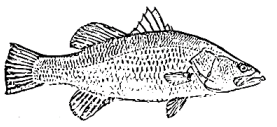
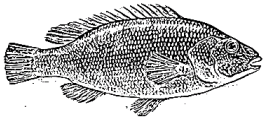
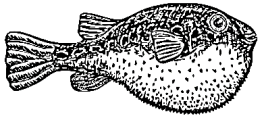
Table 4, contd.					
Parambassis ranga	Indian glassyfish	Acanthopterygii	South Asia	6-8 cm	
Gouania willdenowi	blunt-snouted clingfish	Acanthopterygii	Mediterranean	5 cm	
Oreochromis niloticus	Nile tilapia	Acanthopterygii	North Africa	35 cm	
Oryzias latipes	Japanese medaka	Acanthopterygii	East Asia	5-6 cm	
Nothobranchius furzeri	turquoise killifish	Acanthopterygii	SE Africa	5-6 cm	
Lates calcarifer	Barramundi perch	Acanthopterygii	Indo-West Pacific	60-120 cm	

Table 4, contd.					
Labrus bergylta	Ballan wrasse	Acanthopterygii	East Atlantic	50 cm	
Takifugu rubripes	fugu	Acanthopterygii	Sea of Japan East China Yellow sea	<80 cm	

All teleost species share a common ancestor that underwent an additional (third) round of whole genome duplication (WGD) known as teleost-specific whole genome duplication event or TS-WGD (Postlethwait et al. 1998; Hoegg et al. 2004). A sub-set of teleost species have also undergone further fourth- and fifth- rounds of whole genome duplication ((Xu et al. 2014; Lien et al. 2016). Naturally, these evolutionary events and their resulting traits shared by such a large group of organisms, is a profound field of research. The teleost species studied as part of this project are described in **Table 4**.

1.8 Animal model organisms

Several animal model organisms have been critical to understanding the genetic and evolutionary aspects of cellular functions. Model organisms can be loosely defined as those species whose genomes or cellular pathways exhibit resemblances to the human genome or functional pathways. Different organisms have been used as valuable models to study several aspects of human diseases and development, irrespective of their evolutionary distance from human (**Figure 6**). Although many teleost species could be feasible models for the study of human diseases and development, only a few are predominantly being used for specific areas of research. For example, turquoise killifish (*Nothobranchius furzeri*) is mainly used to study aging cells and the pathways associated with it (Hu and Brunet 2018), while the electric eel has been used in neuroscience research, particularly in connection with the study of neuronal impulses associated with paralysis and other neuromuscular disorders (Catania 2019; de Santana et al. 2019). However, the most prominent model organisms are medaka and zebrafish.

The Japanese rice fish or medaka (*Oryzias latipes*), has been a model research organism for over a century (Temminck 1846). This freshwater fish is also capable of surviving in brackish waters. Medaka are small fish, typically 5-6 cm long and belong to the Beloniformes order. Medaka are integumentary, pelvic fin brooders (Hilgers and Schwarzer 2019) and have a short generation turnaround, drastically reducing the time needed to perform multi-generational studies. Its size and aquarium-friendly

disposition makes medaka a favourable laboratory organism. Housing and maintaining medaka is cost and space efficient and they can survive on dry feed as well as tolerate a wide range of temperatures and pH.

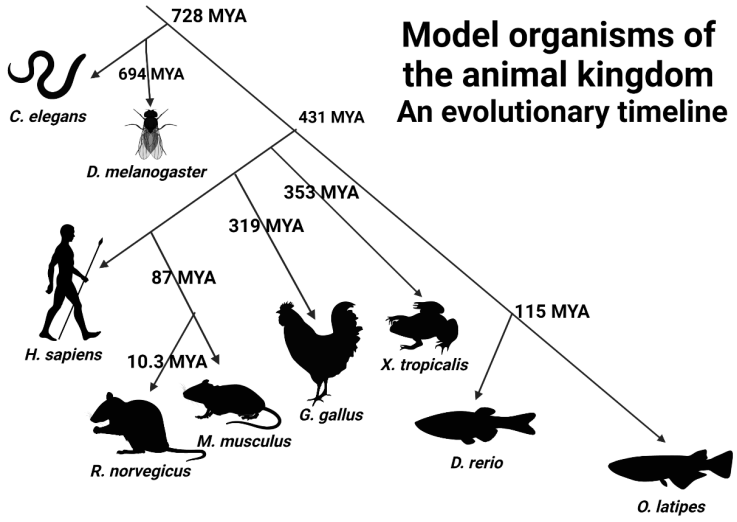


Figure 6: Timeline of divergence of species prominently used as model organisms for human biology research. The times specified are in accordance with the evolutionary distance to medaka. The timeline was generated using the data from timetree species divergence and the illustration was made using bioRender.com.

Most importantly, medaka embryos are optically transparent making it an invaluable tool for studying embryonic development in greater detail. With recent technological advancements in genome sequencing, genome editing and epitranscriptomics, medaka is being increasingly used in research particularly as a model in toxicological studies and also as a successful model for various human diseases (Wittbrodt et al. 2002; Shima and Mitani 2004; Hilgers and Schwarzer 2019). Medaka have a genetic XY sex determination system (Matsuda et al. 1998) and studies on the mating behaviour and mate selection patterns of medaka suggest behavioural correlation with humans

(Okuyama et al. 2017). A draft genome of medaka is available with a more comprehensive full-genome sequence expected in the future (Kasahara et al. 2007).

Zebrafish (*Danio rerio*) have been referred to as 'aquatic mice' due to the amount of research conducted on the species and is the most recognised and favoured teleost model organism. Zebrafish also has a short generation turnaround, optically transparent embryos, frequent spawning, cost and space-efficient husbandry and maintenance (Dooley and Zon 2000). Over 500 mutant phenotypes have been described in zebrafish and many of these are reminiscent of various human disease phenotypes (Rafferty and Quinn 2018; Kim and Zhang 2020; Kroll et al. 2021). Unsurprisingly, zebrafish has been used extensively as a model for understanding the molecular pathology of many human diseases including hematopoietic, renal, cardiovascular and developmental disorders (Rahman Khan and Sulaiman Alhewairini 2019; Patton et al. 2021; Zang et al. 2022). The dawn of the 21st century saw the release of the completed zebrafish whole genome sequence and with it came an explosive interest in genomic and transcriptomic studies in zebrafish. The zebrafish genome has been fully-sequenced and is made up of 1.37 billion base containing over 25,000 coding genes, 6,000 non-coding genes, 59,000 transcripts and 25 chromosomes (Howe et al. 2013).

1.9 Embryonic development of zebrafish and medaka

The profiling, classification and imaging of each embryonic developmental stage has been painstakingly documented for both zebrafish and medaka (Kimmel et al. 1995; Iwamatsu 2004). While the adults of both species are quite comparable in many aspects, one of the key differences is the method of reproduction. Zebrafish breed in groups and scatter their eggs (Breder and Rosen 1966). In both zebrafish and medaka, external fertilization is observed where the entire embryonic development occurs independently of the parents. Medaka has a slightly different mode where the fertilised eggs are attached *via* filaments to the *cloaca* of the female. These embryos can remain attached for several hours and even days until a suitable deposition surface

is encountered by the swimming female, usually aquatic vegetation (Yamamoto 1975). Another major difference is the timeline of embryonic development. The zebrafish embryos hatch approximately 3 days post fertilisation at 28.5°C while medaka embryos hatch approximately 10 days post fertilisation at 26°C.

Despite the varying developmental timeline, the two species do follow a similar succession of developmental events and categorised developmental stages. This makes comparative studies on embryonic development between the two species, highly feasible. The embryonic development in zebrafish is documented in more detail (Kimmel et al. 1995) than that of medaka (Iwamatsu 2004); therefore, the named zebrafish stages and images are used as the principal template and the corresponding time taken to achieve the comparable developmental stage in medaka is accordingly highlighted. Every described zebrafish developmental stage does not have a directly corresponding stage in medaka development. The most comparable stages, both in terms of structure and major developmental events have therefore been chosen for more detailed comparison of embryonic development between zebrafish and medaka (**Figure 7**).

1.9.1 Unfertilised egg

Oocyte maturation in zebrafish consists of six stages, culminating in the mature, translucent egg (Selman et al. 1993). The mature unfertilised egg is the maternal control used throughout this study. In zebrafish, the unfertilised eggs are obtained by gently stripping a fertile female to release eggs. These eggs are spherical with a smooth surface and are released by the females prior to fertilization. In contrast, the mature unfertilised eggs in medaka contain filaments or villi distributed over the entire egg surface. Tens of eggs are clumped together by means of longer filaments. This poses challenges to the sampling and extraction of unfertilised eggs from medaka. Although the yolk is transparent and spherical, medaka eggs are characterised by the presence of oil droplets in the cortical alveoli (Iwamatsu 2004).

1.9.2 1-cell stage

The first categorised stage of the zygote period that follows fertilisation is the 1-cell stage which is otherwise referred to as 'zygote'. As the name suggests, during this stage, the embryo is a distinct single cell with a 'hump' of yolk-free cytoplasm on one end of the initially smooth zygote (Kimmel et al. 1995). This cytoplasmic 'hump' marks the animal pole of the developing embryo and is typically observed at 0.2 and 0.3 hours post-fertilization (hpf). The corresponding 1 cell stage in medaka is observed between 0.6 and 0.8 hpf and is further described in two stages. The activated egg stage is marked by cortical alveolar exocytosis, thinning and hardening of the chorion while the blastodisc stage or the more prominent 1 cell stage is characterised by the displacement of the oil droplets to the vegetal pole and distinct grooves on the blastodisc to indicate the subsequent blastomeres (Iwamatsu 2004).

1.9.3 32-cell stage

The 32-cell stage is equally well described in both species. The embryo undergoes a series of active cell divisions during early embryonic development and each early division can be tracked without difficulty. The cells are arranged in a 4*8 orientation at the 32-cell stage in zebrafish which is completed at 1.75 hpf (Kimmel et al. 1995). In medaka, the central cells are divided horizontally while the marginal ones are divided meridionally giving rise to an inner layer of 8 clustered cells surrounded by an outer ring-like layer of 24 cells. This stage is reached 2.7 hpf (Iwamatsu 2004).

1.9.4 1000-cell stage

The 1k-cell stage in zebrafish is visually complicated to distinguish from the immediate next stage or high stage (Kimmel et al. 1995). This development stage is typically observed at 3 hpf and is characterised by the presence of 11 distinct tiers of cells and the presence of a yolk syncytial layer around the blastodisc margin. In medaka, the 1k-cell stage or the early blastula stage is observed around 6.5 hpf and is characterised by the presence of around 1000 cells, a cortical syncytial layer and a high blastodisc (Iwamatsu 2004).

1.9.5 50% epiboly

Early gastrulation in zebrafish starts with the 50% epiboly stage where the margin of the blastoderm is mostly equidistant from both the animal and the vegetal poles (Kimmel et al. 1995). The blastoderm, though not completely thickened, is observed to be of a uniform thickness. This stage is reached 5.1 hpf. Towards this timepoint, comparative staging of zebrafish and medaka embryos becomes complicated. When considering the 50% epiboly stage as the first stage of early gastrulation, the corresponding stage in medaka happens to be the pre-early gastrula stage which occurs at 8.5 hpf. However, at this stage, the medaka embryo is only at 20% epiboly (Iwamatsu 2004; Furutani-Seiki and Wittbrodt 2004). When considering the actual epiboly itself, the comparative 50% epiboly is achieved at mid-gastrulation, almost 13 hpf. Mid-gastrulation in zebrafish is described as the 75% epiboly stage (Kimmel et al. 1995; Furutani-Seiki and Wittbrodt 2004). To simplify the comparison for this overview, the respective 50% epiboly stages of both species have been illustrated.

1.9.6 Bud stage

The bud stage affords an easier comparison between zebrafish and medaka development due to the characteristic formation of a bud. 100% epiboly is synonymous with the bud stage in zebrafish (Kimmel et al. 1995). The ‘bud’ in the zebrafish embryo corresponds to the tail bud and this stage also exhibits the presence of a thickened neural plate. The bud stage develops at 10 hpf. The late neurula stage is the comparative bud stage in medaka, although the ‘bud’ in this case refers to the formation of the optic bud. This stage is also referred to as the 0-somite stage and is characterised by the presence of enlarged Kupffer’s vesicles. In medaka, the bud stage is reached around 25 hpf (Iwamatsu 2004).

1.9.7 12-somite stage

Although the 12-somite stage is not singularly described for zebrafish, the 14-somite stage is well described. The somite stages are identified by simply counting the number of somites or segments observed along the developing spinal column. The 12-somite

stage in zebrafish is marked by tail elongation, division of the brain into the 4 subdivisions, neural crest migration and a prominent otic placode. This occurs approximately 15 hpf (Kimmel et al. 1995). In medaka, the 12-somite stage is when the tubular heart forms, and similarly to the zebrafish stage, the brain subdivisions are observed. This stage is observed at 33 hpf (Iwamatsu 2004).

1.9.8 Prim-5 stage

Prim-5 marks the beginning of the pharyngula period during the embryonic development of zebrafish and is observed at 24 hpf. Around 30 somites, heartbeat and pigmentation can be observed during this stage (Kimmel et al. 1995; Furutani-Seiki and Wittbrodt 2004). This stage is once again difficult to perform a direct comparison on medaka as no Prim stages have been described in medaka embryogenesis. However, the 30-somite stage in medaka is also characterised by pigmentation of the retina and has been chosen as the comparative stage for Prim-5. In medaka, this stage is observed at 64 hpf and exhibits the liver bulge along with pancreatic eminence ((Iwamatsu 2004).

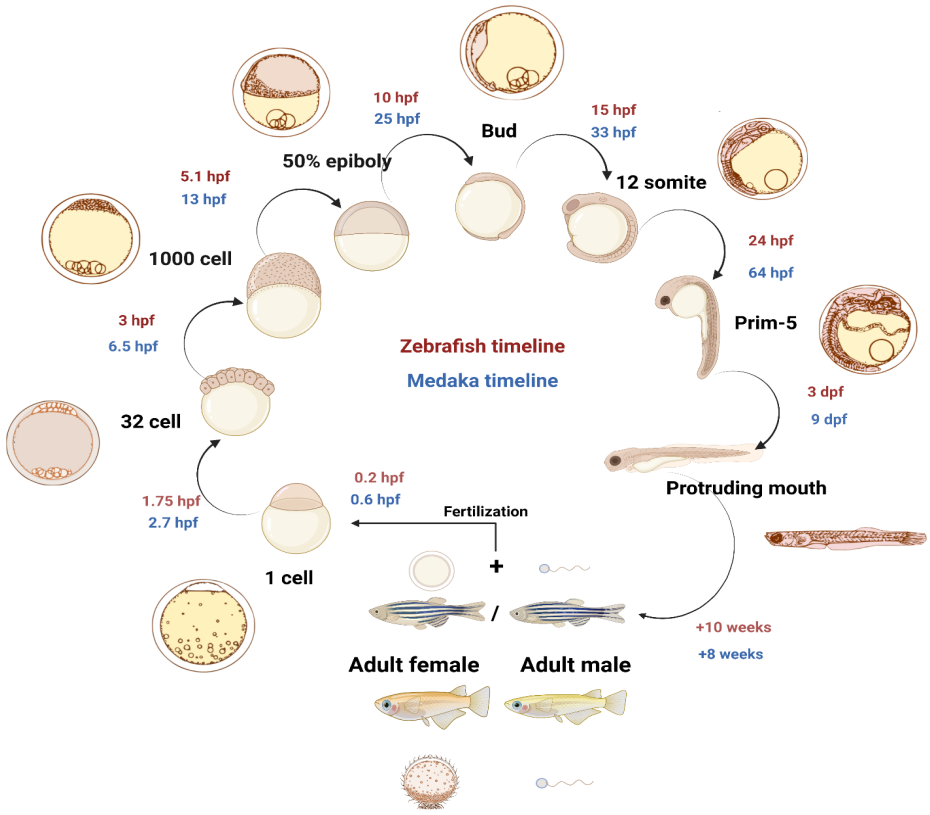


Figure 7: Embryonic development of zebrafish and medaka – A comparative timeline of zebrafish and medaka embryonic development. The inner ring and outer ring represent zebrafish and medaka embryo illustrations respectively. Created with biorender.com

1.9.9 Protruding-mouth stage

The protruding-mouth stage is the final stage of embryonic development in zebrafish during which yolk resorption is completed. This stage is characterised, as the name suggests, by the protruding and wide open mouth as well as the development of gill slits and gill filaments. The swim bladder appears to be darkened due to the increased melanin production and the organism is preparing to eat and breathe on its own. This stage is usually seen at 72 hpf (Kimmel et al. 1995). The protruding-mouth stage is not officially described in medaka and owing to the high timeline variations from zebrafish,

the correspond medaka stage is quite challenging to pin-point. The most probable ones are between day 8 and day 9 post fertilisation, also known as the hatching stages. In this time period, the swim bladder and gall bladder develop visibly, the gill development is observed, the mouth moves along with active eye movement and the organism is similarly being prepared to hatch and survive (Iwamatsu 2004).

Following hatching, the larvae of both organisms continue to grow and develop to attain sexual maturity. Zebrafish reach sexual maturity between 10 and 12 weeks post hatching, regardless of sex (Westerfield 2000) while medaka, under laboratory conditions can achieve sexual maturity around 8 weeks after hatching (Shima and Mitani 2004).

1.10 Ribosomal modifications in a developmental context

Availability of advanced quantitative investigative approaches such as RiboMeth-seq (Birkedal et al. 2015) has facilitated better insight into sub-stoichiometric ribose methylation which indicates that smaller ribosomal populations can have heterogenous modification profiles (Hebras et al. 2020). Mouse rRNA 2'-*O*-Me has been studied in a developmental context and the developing tissues showed differential methylation, as opposed to adult tissues, were most sites appeared to be fully methylated (Hebras et al. 2020). SNORD78 expression during mouse development could be regulated by alternative splicing of the *gas5* host gene (Hebras et al. 2020). Besides this, no other study has been described prior to this study, that investigates ribose methylation in a developmental context.

In zebrafish, it has been shown that two specialised types of rRNA are expressed at varying levels during early embryonic development and these are structurally divergent from one another (Locati et al. 2017a). Studies of ribosomal modifications during development have led us to believe that differentiating cells provide the best opportunity for ribosomes to exhibit specialisation (Locati et al. 2017b). The presence of a development-related ribosomal variant in zebrafish warrants intentional research

of ribosome heterogeneity, ribosomal modification and snoRNA expression in a developmental context. A survey of available literature revealed that ribose methylation had never been profiled in any fish species before and we intended to fill this very gap in our understanding of embryonic development beginning with the model organisms, zebrafish and medaka.

2 Study Objectives

The overall objective of this PhD project was to **investigate the prevalence of 2'-*O*-Me in zebrafish and medaka rRNA during embryonic development with respect to the developmentally specialised ribosomal type.**

The specific study objectives were:

1) To map the 2'-*O*-Me sites in both the early- and late-embryonic rRNA types during zebrafish embryonic development and explore the methylation stoichiometry across developmental stages – **Paper I**

2) To investigate the presence or absence of the additional ribosomal type in another teleost model organism, medaka – **Paper II**

3) To compare the ribose methylation profile of rRNA during the embryonic development of the two zebrafish types and medaka – **Paper II**

4) To compile and assign plausible SNORDs to most of the modified sites in zebrafish and medaka rRNA – **Paper I and Paper II**

5) To study the phylogenetic variations of *gas5*, a prominent SNORD host gene, between various teleost species – **Paper III**

3 List of papers

Paper I

The shift from early to late types of ribosomes in zebrafish development involves changes at a subset of rRNA 2'-O-Me sites. RNA

Sowmya Ramachandran*, Nicolai Krogh*, Tor Erik Jørgensen, Steinar Daae Johansen, Henrik Nielsen, Igor Babiak

**Contributed equally*

RNA. 2020 Dec;26(12):1919-1934. doi: 10.1261/rna.076760.120

Paper II

The study of 2'-O-Me rRNA modifications in medaka using RiboMeth-seq reveals a high degree of conservation with human rRNA and the absence of the specialized early-rRNA found in zebrafish

Sowmya Ramachandran, Nicolai Krogh, Thuy-Thao-Vi Nguyen, Steinar Daae Johansen, Julien Bobe, Tor Erik Jørgensen, Henrik Nielsen and Igor Babiak

Manuscript

Paper III

Preliminary identification and comparison of Box C/D snoRNA sequences hosted by the teleost *gas5* gene

Sowmya Ramachandran, Henrik Nielsen, Steinar Daae Johansen and Igor Babiak

Manuscript

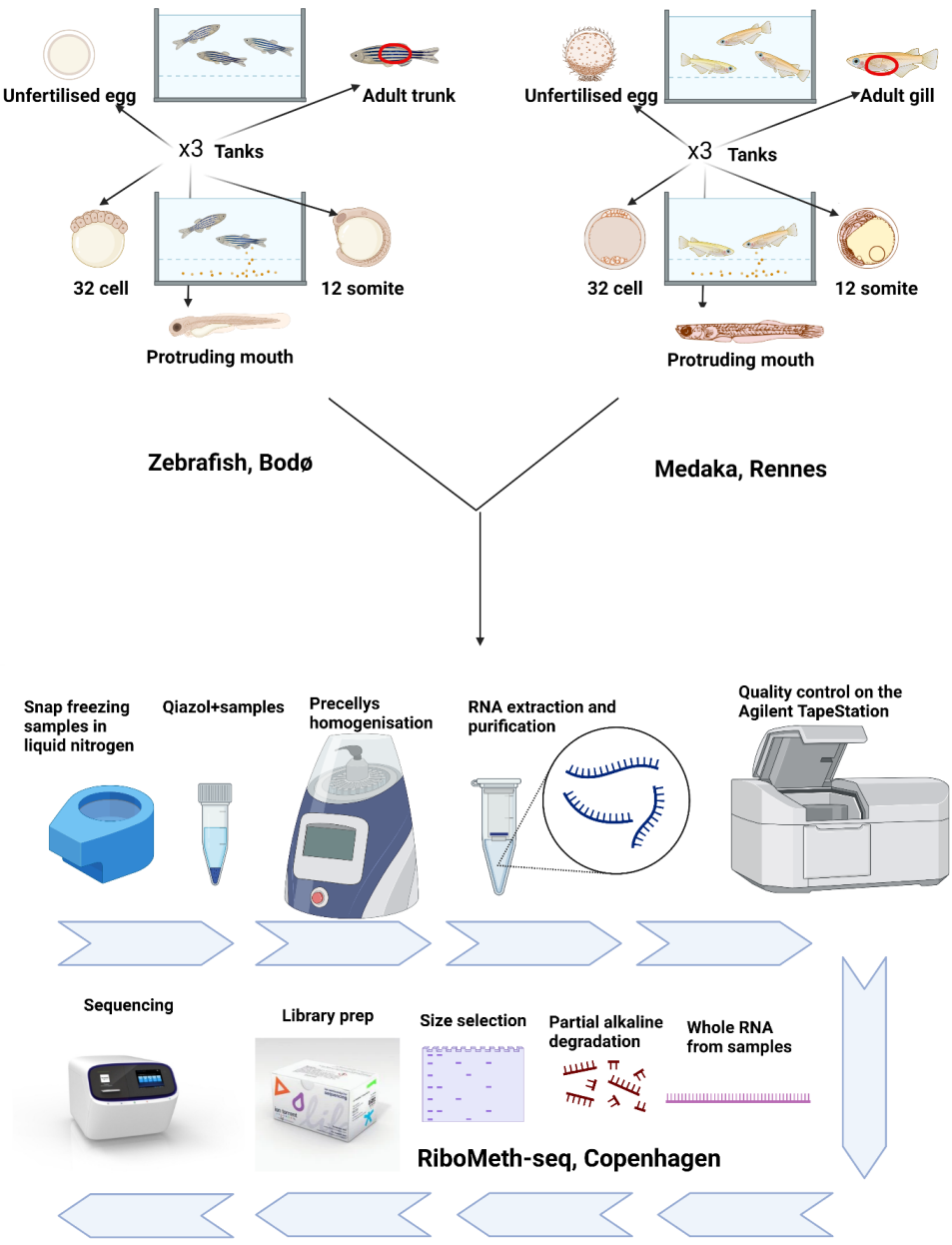


Figure 8: Experimental workflow – sampling of zebrafish and medaka in Norway and France respectively, RNA extraction and RiboMeth-seq workflow in Denmark. Created with bioRender.com

4 General Discussion

4.1 Dual ribosomes – not a teleost factor?

The presence of additional ribosomal types has been studied for decades in a variety of species. The 5S rDNA alone has been extensively studied since the first report on the 5S rRNA sequence heterogeneity in *Xenopus laevis* (Wegnez and Monier 1972). The rDNA sequence repeats are either absent or poorly represented in the currently available genomes due to the technological challenges of analysing such repetitive sequences (Shaw and Brown 2012). The presence of unique oocyte-type and somatic-type 5S rRNA in zebrafish was only described recently (Locati et al. 2017a), followed by the study which uncovered the presence of distinct early and late embryonic type rRNA of every rRNA subunit in zebrafish (Locati et al. 2017b), making it the first fish species in which the presence of two types of rRNA was documented. In **Paper I**, we were able to establish that the two rRNA types in zebrafish were different in both rRNA structure and ribose methylation profile in addition to varying methylation stoichiometry across developmental stages.

The zebrafish results led us to consider whether the two distinct rRNA types could be a direct result of TS-WGD and therefore shared by all teleost species. The results from **Paper II,** indicated otherwise as medaka, another teleost model organism, only possessed one type of rRNA across development. Although this alone is not sufficient to disregard the hypothesis of TS-WGD contributing to ribosomal duality in zebrafish, it certainly confirms that ribosomal duality is not an overarching feature of every teleost species (**Paper II**). Upon closer evaluation of the taxonomic relatives of zebrafish, *Cyprinus carpio* (Eurasian carp) has been shown to have undergone the most recent, additional, fourth round of WGD following TS-WGD (Xu et al. 2014) and belongs to the family Cyprinidae, which also contains zebrafish. Furthermore, the phylogenetic analysis of *gas5* in **Paper III** shows all the species of the Ostariophysi super-order grouped closer together than the rest of the selected teleosts. Whether the ribosomal duality was specially evolved in zebrafish or is a common feature of other species in

the *Danio* genus, or the Cyprinidae family, or the whole of the Ostariophysi super-order, remains to be seen. The duality could also be potentially explained by other factors such as the loci of rDNA clusters in zebrafish as two distinct somatic and maternal specific rDNA clusters have been previously described (Locati et al. 2017b; Ortega-Recalde et al. 2019).

WGD events have been studied to better understand the evolution of vertebrates, and the study of 'Ohnologues' (the retained paralogs from WGD events) is of significance. All vertebrates evolved from a common ancestor that underwent two rounds of WGD and the ohnologues from this event have been shown to affect cell signalling, development and gene regulation. Since the genomes of most organisms exhibit lineage-specific rearrangement, a newer method to study ohnologues can help overcome this limitation (Singh et al. 2015), using which it has been shown that teleost fish have retained more ohnologues from the second round of WGD than mammals (Singh and Isambert 2019). The common ohnologues retained from the second WGD and TS-WGD events in teleosts could be of great value if similar analyses are performed on carps which underwent an additional WGD event. The resulting list of ohnologues would be intriguing targets for studies on vertebrate evolution and on the functional significance of these ohnologues. Of the teleost lineages considered in **Paper III**, in addition to the common carp, the genomes of salmonids including Atlantic salmon also underwent an additional fourth round of WGD termed the Ss4R (Lien et al. 2016; Varadharajan et al. 2018). The number of rDNA and SNORD host gene ohnologues that have been retained from each one of these WGD events could point to a deeper evolutionary reason behind the two types of rRNA seen in zebrafish and insights into the possibility of such duality in other vertebrates.

4.2 Is medaka a better model for human embryogenesis than zebrafish?

Medaka was the first vertebrate in which sex-linked inheritance was described (Aida 1921), following which the sex determination in medaka has been extensively studied

(Kondo et al. 2009). Early research on sex determination in medaka effectively established that medaka follows the XX/XY sex determining system, similar to human and many other vertebrates (Yamamoto 1953; Tuzuki et al. 1966; Satoh and Egami 1972; Hamaguchi 1987; Matsuda et al. 1998). Most of the studies on sex determination in medaka, having originally been performed on organisms which had not been subjected to selective breeding, depict a fairly accurate comparison to its wild counterparts.

While it has been shown that all zebrafish developed female-like gonads which produced early-stage oocytes (Takahashi 1977), subsequent experiments on the sex determination mechanism in laboratory AB line zebrafish have established the importance of genetics in sex determination but were unable to conclusively determine the mechanism behind zebrafish sex determination (Streisinger et al. 1981; Liew et al. 2012). Although it was initially proposed that wild zebrafish followed a ZZ/ZW chromosomal sex determination system with heterogametic females and homogametic males (Sharma et al. 1998), which could explain the female-bias seen in the domesticated lines by factoring in the potential loss of sex determining loci during the selective breeding process, a thorough evaluation of the ZZ/ZW system in zebrafish later concluded that the molecular variation between the Z and W chromosomes were minimal (Wilson et al. 2014) and that the original findings could have been a result of region-specific variation in the specific wild zebrafish population used in the original study (Sharma et al. 1998). Despite this unresolved ambiguity in a fundamental and critical developmental mechanism (Kossack and Draper 2019), zebrafish continues to be favoured as a model for developmental studies over medaka.

The original study on dual zebrafish ribosomes (Locati et al. 2017) and our subsequent findings of divergent ribose methylation profiles between the two rRNA types in **Paper I**, confirm yet another critical difference between human and zebrafish embryonic development. Our findings form **Paper II** further confirm the similarity between human and medaka developmental machinery at a ribosomal level. This would suggest that

medaka, being the more similar model organism to human in terms of both sex determination as well as ribosomal system, has obvious advantages to be preferred as a developmental model for human over zebrafish. Moreover, the presence of a late-embryonic rRNA type specific modification which is conserved in medaka (**Paper I and Paper II**), could be suggestive of an adaptation to selective breeding or a compensatory mechanism to make up for the loss of a sex-determination system in domesticated zebrafish. Whether these considerations could be clarified to some extent by repeating the **Paper I** experiment on wild zebrafish embryonic samples remains to be seen. It could also be argued from the **Paper III** findings that medaka is however missing a critical SNORD from a prominent host gene, which is otherwise present in both zebrafish and human. The loss of a SNORD in medaka may not be enough of a disadvantage compared to the presence of an extra ribosomal type in zebrafish to be considered a better model for human embryogenesis, however this cannot be confirmed until the sex determination system of zebrafish and the function of SNORD78 (which is missing in medaka) is fully known.

Whether zebrafish or medaka is a better model organism has been of debate, particularly in the recent years with the emerging twin studies on both organisms. Both organisms possess certain advantages over the other as a more suitable model depending on the nature of the research question (Teame et al. 2019; Hilgers and Schwarzer 2019). Zebrafish and medaka are increasingly being used simultaneously in twin studies to observe and analyse the similarities and variations between experimental outcomes of both organisms (Furutani-Seiki and Wittbrodt 2004). The results from **Paper III** further highlight the diversity and specialisations between several teleost species and proposes that no single teleost organism could be the 'perfect' model organism but by combining two prominent model organisms such as zebrafish and medaka, the research perspectives would broaden and facilitate a deeper understanding of several aspects of human disease and development (Signore et al. 2009; Chowdhury et al. 2022).

4.3 Secondary structure of rRNA

While the primary objectives of this study did not emphasize on structural resolution of rRNA, in **Paper I**, the first complete rRNA secondary structures for zebrafish were resolved based on the available human rRNA structures (Bernier et al. 2014). The purpose of resolving the rRNA secondary structure was mainly to compare the extent of structural variation between the early and late rRNA types, despite the apparent differences in sequence length and composition between the two types. The structural resolution provided for an added advantage to the ribose methylation profile analysis in verifying the sites of novel modifications with respect to their potential location in known structurally conserved regions. Subsequently, the rRNA secondary structures for medaka were similarly resolved in **Paper II**, facilitating a direct structural comparison of medaka rRNA with the two zebrafish rRNA types. The structural comparison further supported our hypothesis that the early-type rRNA in zebrafish was indeed the additional specialised type, owing to the structural similarity of medaka rRNA with human, mouse and zebrafish late-type rRNA. Owing to the lack of rRNA references sequences for the teleost species studied in **Paper III**, no such structural resolution and comparison could be carried out.

Studies have shown that the structure of rRNA is critical to its function and is therefore more conserved than the sequence of rRNA (Gutell et al. 1994). This structural conservation has been validated on many levels with the most recent being the inference that the conserved rRNA structures are not restricted to specific lineages and that although the structure is conserved over sequence, rapidly evolving sequences are rarely connected to such conserved structures (Seemann et al. 2022). The structure of the PTC is extremely conserved (Welch et al. 1997) and no structural variations have been observed in this region in both the zebrafish rRNA types as well as the medaka rRNA. The so-called expansion segments are often the regions of highest variation in rRNA subunits between species (Ramesh and Woolford 2016). The 9es3a segment on the medaka SSU rRNA appeared to be shorter than that of both the zebrafish rRNA

types and presented as the only apparent structural variation among the SSU rRNA of both species. In **Paper I** we proposed the potential for the GC-rich elongation of this expansion segment, coupled with the modifications at sites U286 and U287 first described in both the SSU rRNA types of zebrafish, to be involved in the folding of the 9es3 expansion segment. However, U287 in medaka is unmodified suggesting a link between the length of the expansion segment and the number of modifications observed in the site (**Paper II**). In human SSU rRNA, while the length of the 9es3a expansion segment falls between that of medaka and both the zebrafish types, both the U286 and U287 sites in the 9es3 segment are unmodified (Bernier et al. 2014). The most striking aspect of the zebrafish LSU rRNA in **Paper I** was the modification at C3916 which was only observed in the late rRNA type. This site is located close to the 79es31 expansion segment and is also seen to be modified in medaka LSU rRNA in **Paper II**. This modification could also serve a similar purpose as the one proposed for SSU U286 since both the zebrafish late-type LSU rRNA and medaka LSU rRNA possess shorter 79es31 segments than that of the zebrafish early-type, mouse and human LSU rRNA, all of which do not contain a modification at this site. This modification could just as well be inversely related to the length of the neighbouring expansion segment and is another small example on the long list of structural differences in rRNA between species which requires a larger repository of resolved structures in various species for further comparative studies.

4.4 SNORD and host gene observations

Being a non-coding gene, *gas5* has been proposed to have newly evolved a functional exon . Additionally, the 13 *gas5* exons, through alternative splicing, give rise to multiple mature RNA isoforms. In comparing the *gas5* exons of human and mouse, despite the poor sequence conservation (Smith and Steitz 1998), the presence and architecture of the gene itself with respect to the encoded SNORDs, is functionally valuable for this ncRNA (Carninci and Hayashizaki 2007). In human *gas5*, it has been shown that some *gas5* isoforms are capable of folding into secondary structures to compete with the

DNA binding sites of glucocorticoid receptors (Kino et al. 2010). Classified as a tumour suppressor gene (Smith and Steitz 1998), *gas5* transcripts have been studied with respect to various types of cancers (Yang et al. 2020). It has been proposed as a diagnostic and prognostic marker for several cancer types (Li et al. 2017; Gao et al. 2017).

The exon closest to the 3'- end in human *gas5* contains the element that serves as a mimic of the glucocorticoid receptor element (Chandler et al. 1983), and has been labelled the glucocorticoid receptor element mimic (GREM). GREM has been studied to be shared only among the newest and closest primate relatives of human (Hudson et al. 2014), which explains the absence of the GREM element in mice or other vertebrates. This exon with acquired functionality, is seen to flank the intron containing SNORD81 in both human and mouse *gas5* (Goustin et al. 2019). Expressing the GREM sequence alone has been shown to induce apoptosis in breast cancer cells (Pickard and Williams 2016), although further evidence is required to validate these findings.

In **Paper I**, we first observed the absence of SNORD81 and the modification guided by it in zebrafish. This was divergent from the mouse *gas5* where the GREM alone is absent while SNORD81 and its modification are both present (Goustin et al. 2019). Medaka *gas5* was also missing SNORD81 and the modification it guides (**Paper II**), which led us to question whether the absence of SNORD81 and its flanking 3'-exon was evolutionarily conserved in teleosts. The *gas5* annotations in **Paper III** revealed that SNORD81 is unilaterally absent from every teleost species selected for phylogeny, supporting the evidence of the acquired functionality being evolutionarily very young. Furthermore, many canonical 'orphan' SNORDs were present in the *gas5* gene of some species in **Paper III**. While this could be perceived as an important specialisation, the orphan SNORDs as well as the additional copies of SNORD77/80 have been treated with a degree of caution since **Paper I** initially described the host gene of SNORD119 in zebrafish which also hosted an orphan SNORD that we named SNORDX. Upon closer examination of the host gene, we found ten copies of the same SNORD within the same

host genes. This is not an isolated finding as additional SNORD copies are often found in the same host gene as evidenced by SNORD77/80 copies in teleost *gas5* (**Paper III**). The function of the duplicate SNORDs, if any, is currently unknown. However, the presence of duplicates further complicates potential functional studies based on SNORD KO as the guide will have to successfully KO every copy of a particular SNORD.

The medaka results from **Paper II** also revealed the surprising absence of SNORD78. SNORD78 is a highly-studied SNORD due to its documented involvement in oncogenesis (Martens-Uzunova et al. 2015) as this SNORD is observed to be up-regulated in cancer cells. SNORD78 KD in lung cancer cells resulted growth arrest, reduced proliferation and increased apoptosis (Zheng et al. 2015). Disrupting SNORD78 expression in zebrafish through KD has previously been shown to be lethal in zebrafish embryos (Higa-Nakamine et al. 2012). However, KO of SNORD78 in human embryonic kidney cells resulted in no comparable phenotypes to that of previous SNORD78 KD studies (Hebras et al. 2020). Consideration must be given to the off-target effects of KD methods and cell-type specific functional implications, before evaluating the merits of SNORD78 KD phenotypes. The absence of this SNORD in medaka is intriguing. Of the teleost species studied in **Paper III**, SNORD78 was only absent in fugu (*Takifugu rubripes*) in addition to medaka, indicating that the SNORD78 could be evolutionarily lost in a subset of teleost species which are closely related to both medaka and fugu under the superorder Acanthopterygii.

4.5 Methodological considerations

4.5.1 Detection of 2'-*O*-Me

Over the last 50 years, 2'-*O*-Me modifications in RNA have been detected using many different methods (Krogh and Nielsen 2019). The 'classical' methods include gas chromatography (Abbate and Rottman 1972), homochromatography (Sardana and Fuke 1980), mass spectrometry or MS (Qiu 1999), Liquid chromatography-MS or LC-MS (Takeda et al. 1994), two-dimensional thin layer chromatography or 2D-TLC (Cavaillé et al. 1999), RNA fingerprinting (Maden 1986, 1988), RNase H (Yu et al. 1997),

DNA based ribozyme or DNAzyme (Buchhaupt et al. 2007), and reverse transcriptase dependent primer extension or RT-PE (Maden 2001). While each of these methods possessed unique merits, the common underlying issues were that all of these methods required high input volumes, were not feasible for non-abundant RNA types, false-positive rates, laborious quantification processes (if at all possible), and the lack of high-throughput feasibility (Motorin and Marchand 2018).

Sequencing-based high-throughput methods for detection of 2'-*O*-Me have been proposed and published recently. These include RiboMeth-seq, CLIP-seq, 2'-OMe-Seq, RibOxi-Seq and Nm-Seq. A comprehensive comparison of the advantages and disadvantages of each of these methods has been reviewed by Krogh and Nielsen 2019. In addition to solving the high-throughput issue of the classical methods, these sequencing-based methods only require a fraction of the RNA input compared to the classical methods. 2'-*O*-Me is the only rRNA modification currently capable of being detected by such methods. All of the sequencing-based methods are suited for rRNA; however, the most suited for our study was the RiboMeth-seq (Birkedal et al. 2015; Krogh et al. 2017a). This was due to the potential for quantification of methylation along with the lowest input volume, making it an ideal candidate for our small embryo samples and in studying developmental variations.

While RiboMeth-seq is an excellent method to study ribose methylation, low-scoring sites are often unable to be picked up, owing to some degree of background. In the experimental setup for **Paper I**, this was resolved in a few different ways: firstly, by making a list of known sub-stoichiometric sites in human and mouse rRNA and comparing the zebrafish methylation scores in those sites; secondly, by the use of RT-PE studies to validate these sites; and finally, by assigning plausible SNORDs to these sites from a dedicated RNA-seq. For **Paper II**, owing to the comparative nature of the study, only the first method of identification was employed. Assigning plausible SNORDs to most sites in medaka rRNA remains incomplete due to the lack of dedicated transcriptomic data and no RT-PE studies were carried out. Although the results served

sufficient for a simple comparison between medaka and zebrafish ribose methylome, the medaka results from **Paper II** do not have the same level of completion as the zebrafish results in **Paper I**. It must also be noted that the complete list of plausible SNORDs is not available for every modified site in human and mice despite the more extensive research on these species (Krogh et al. 2017a; Hebras et al. 2020). More recently, rRNA modifications have been successfully profiled at a single-molecule level using a novel method involving nanopore sequencing technology which showed that the modification profiles seldom varied with changes in the translation conditions (Bailey et al. 2022). This method could prove quite valuable in studying the relationship between RNA modifications and RNA function.

4.5.2 Theoretical phylogeny vs targeted sequencing-based phylogeny

Teleosts in particular have been subject to great interest in phylogenetic analysis of various genes or genomic regions owing to their evolutionary history. Of the thousands of species in the Teleostei infraclass, no phylogenetic study of tens or even hundreds of species is ever going to be a true representation of the evolutionary trends in teleosts. Furthermore, taxonomic classifications of many teleost species have been subject to changes and genomic and transcriptomic studies are contributing to re-classification of some species into new or different orders and families. This makes phylogenetic studies of prominent and known teleost species evermore important a tool to elucidate evolutionary relationships between them.

Some studies have successfully performed large-scale phylogenetic analysis of select teleost species. The approaches adopted thus far have varied greatly from a dedicated transcriptomic sequencing of over a 100 species (Hughes et al. 2018), DNA from tissue biopsies of over 200 fish species (Near et al. 2012), comparing known coding-sequences of select species with the updated versions of their genome (Mikalsen et al. 2020), extracting DNA from museum specimens of species significantly described in previous phylogenies (Eytan et al. 2015) or by using a combination of amino acid sequences and coding-nucleotide sequences for previously phylogenetically described

genes (Takezaki 2021). Standard multiple alignment methods have largely suited these studies since the focus was on the conservation of the coding regions.

As with any research involving multiple species, the experimental setup to study each of the species considered in **Paper III** would have been tedious and enormous. Between our colleagues and partner labs, we did possess the ability to procure and sequence RNA from most of the selected species, which would have validated most of our observations. However, triplicates from each species for 20 species, factoring for the effects of various housing conditions and laboratory biases, would have still been too high a number to warrant such a monumental effort. Added to it the cost of prepping and sequencing 60 libraries and the downstream bioinformatic analysis, propelled us to seek out the alternative non-experimental approach that we have used. The downside to this approach is unfortunately the lack of uniformity in the quality of the genomic data between species, inability to analyse replicates and evaluate gene expression data in the sampled species. Naturally, we also did not have the ability to study **Paper III** in the context of embryonic development.

For the teleost phylogeny of the *gas5* gene in **Paper III**, most alignment tools could not align the SNORDs with each other despite the presence of distinctive C and D boxes. This was in part due to the poor quality of the genome sequences from some species and was also due to the extreme sequence length variations of the gene between the different species. The manual alignment thus provided the most accurate input for the phylogenetic tree. Many bioinformatic resources are available to study snoRNA (Lestrade 2006; Zhang et al. 2012; Yoshihama et al. 2013) and, in our attempt to employ an established snoRNA annotation pipeline (Bartschat et al. 2014), we found the scripts and instructions too complicated, often with no technical support and inadequate site maintenance. Being primarily molecular biologists with limited expertise in advanced bioinformatic scripting and debugging, we chose the most manual, simple and straight-forward approaches wherever possible.

In **Paper III**, we introduced a non-experimental phylogenetic approach targeting conserved elements in a non-coding SNORD host gene, *gas5*, across 19 teleost and 1 holostean species. Prior studies which relied solely on computational methods and automatic ncRNA annotation (Zhang et al. 2009, 2010) have been critiqued harshly (Makarova and Kramerov 2011) owing to stark discrepancies in the results compared to established methods with validated results (Makarova and Kramerov 2009). Therefore, we ventured into **Paper III** with caution and strived for consistency of nomenclature across all the chosen species. The *gas5* gene was uniquely indisposed to the phylogenetic methods employed by other such studies on teleosts due to the fact that *gas5,* being a host gene has a highly conserved genetic architecture of C and D boxes in spite of being poorly conserved overall. Studies on mouse and human *gas5* genes have confirmed the presence of the conserved architecture while our own results from **Paper I** and **Paper II** served to re-affirm this conservational pattern. Although the presence of a plausible SNORD for a known site of modification is not a true indicator of the modification itself, the absence of a SNORD from a known and conserved SNORD-host gene pair, could be indicative of the loss of the subsequent modification. For this reason, it was neither necessary nor relevant to fully annotate the *gas5* gene of all species. The list of plausible SNORDs and the genes that host them in zebrafish (**Paper I**) and medaka (**Paper II**), though incomplete, offers many interesting candidate SNORD-host gene pairs, including some protein-coding genes, for further phylogenetic analysis. For those genes which are better-conserved and annotated in the genomes of most of our species set, a full gene annotation and a more comprehensive phylogenetic analysis could be relevant and offer a better comparison to the results observed from the *gas5* phylogeny in **Paper III.**

4.6 Limitations

Despite providing novel inputs to ribosomal research and teleost evolutionary research, this study has had its pitfalls and limitations – some very obvious and others less so. In any project involving the life cycle of a live organism, the timeline and feasibility of the

work relies entirely on the predictability of the organism and therefore the primary limitations had to do with the zebrafish and medaka life-cycles themselves. During the course of an experiment, if the need arises to sample further stages or perform functional analysis on the same group/generation of fish as the preliminary samples, it is highly unlikely that the same fish would still be reproductively active and in most cases would have already been euthanised to make space for the next generation. The laboratory lines are also not entirely representative of the wild zebrafish and medaka, arousing uncertainty about the data generated from lab-based organisms being compared with a plethora of wild fish species in a phylogenetic study.

Additionally, the capacity of the laboratory in terms of housing space and technician resources played a big hand in the design of the functional studies. With zebrafish, even though housing and maintenance are not as space and resource consuming as mice or other mammals, the strain does significantly increase with the increased complexity of the experiment. For a knockout experiment, to test multiple guides, one tank of fish has to be set aside for each guide, monitored and genotyped before selecting the most successful and proceeding with further generations. Typically the F2 generation takes close to one year to achieve. The more knockouts that are planned, the space requirement increases exponentially and the genotyping workload even more so. Therefore we were unable to perform knock outs of multiple SNORDs simultaneously.

The COVID19 pandemic which was tumultuous to most research work, also affected this project in terms of time and planning. Being a collaborative effort of research groups in three different countries, the changing travel and work regulations of each of them provided new and unpredictable challenges in planning and executing the second half of the project. While we had to abandon our original visions and plans for the duration of a 3-year PhD period, we managed to adapt to the challenges and made the necessary changes to keep the study moving forward. Much of the functional work

on SNORDs in currently ongoing and this PhD project will be succeeded by more studies on the dual zebrafish ribosomes.

The next big limitation to the study is the usage of visualisation and bioinformatics software. Most of the structural visualisations and analysis have been conducted manually instead of utilising various available tools and software. This was a conscious choice for the most part. The most important reason being the study design's need for depth and accuracy which most tools are presently unable to achieve. Most of the tools and software available to predict the secondary structure of RNA can only accurately predict short RNA motifs and small RNA. When applied to large and complex RNA structures such as rRNA subunits, these tools struggle to output a complete secondary structure.

Finally, potential for improvement in the experimental set-up itself has to be noted. While not a major factor, more groups of zebrafish and medaka control lineages could have been used to eliminate any residual bias arising from sampling individuals that potentially share common parents. Other closely related *Danio* and *Oryzias* species could have been used to study and eliminate any potential species-specific factors and to investigate whether the presence of dual ribosomes is a common occurrence across all species in the genus. Moreover, newly described medaka and zebrafish hybrids, *Latio* and *Reripes* (Gert et al. 2021), could have been similarly studied to offer better insights into the origin and function of the dual ribosomal type. Unfortunately, these hybrids were not available to us at the time of our experimental setup.

5 Conclusions

The main conclusion from this PhD study is that the **dual ribosomal types seen in zebrafish during embryonic development, is not a teleost-specific feature**. Based on the main study objectives, the following conclusions are made:

i. The early- and late- type rRNA in zebrafish differ in methylation stoichiometry as well as rRNA structure. The early-type rRNA is the most divergent from other known species and is proposed as the additional specialised type.

ii. The additional ribosomal type is absent in medaka where only one type of rRNA is present throughout embryonic development, thereby confirming that the specialised ribosomal type is not a direct result of TS-WGD.

iii. Most of the novel modified sites found in zebrafish appear to be conserved in medaka including the late-type specific modification in zebrafish LSU. The zebrafish late-type rRNA was more similar to medaka rRNA than the early-type rRNA.

iv. Plausible SNORDs were assigned to most of the zebrafish sites while the medaka list remains fairly incomplete due to the poor annotation of the medaka genome. *gas5* was the gene that hosted the highest number of SNORDs in both species although medaka *gas5* was missing the prominent SNORD78.

v. A non-experimental phylogenetic approach was successfully applied to compare the *gas5* SNORDs of 19 teleost and 1 holostean fish species. A number of orphan SNORDs and duplicate SNORD copies were observed in the *gas5* gene of some species.

6 Future Perspectives

Validating the presence of two individual types of rRNA in zebrafish during embryonic development, has opened the floodgate to new questions and hypotheses while also adding to the current knowledgebase. The most important questions arising from this project naturally pertain to the additional specialised ribosomal type itself. What could be the purpose of having a specialised ribosomal type particularly expressed during early embryonic development in zebrafish? Are there more species that possess similar dual ribosomes, and if so, are they closely related to zebrafish? Was this duality common in an early group of species that was later evolutionarily selected against? Or is this part of specialised evolution in a specific group of species? Why is this ribosomal type not expressed during early embryonic development in mice or medaka?

Do many host genes serve a greater purpose beyond hosting SNORDs? Is any phenotypic effect seen from KD or KO of specific SNORDs, the effect of the absence of the subsequent methylation itself or is it an inadvertent result of disrupting the host gene and its unknown functions? Can a vertebrate organism survive the absence of developmentally dynamic methylations and the SNORDs that guide it? Irrespective of the evolutionary significance, why would these two ribosomal types express different ribose methylation profiles if the same SNORDs are available to both of them? Is the specialised ribosome methylated by an independent set or a subset of SNORDs? Does this speak to the nature of the SNORD host genes itself or does the SNORD-host gene relationship precede the modifications on the ribosome?

Answering each one of these questions is a monumental endeavour in itself and the sum of these would result in a near-complete comprehension of the 2'-*O*-Me in rRNA including the specific function of each modification and the SNORDs that guide these modifications. A good place to begin answering these questions would be to investigate the presence of the specialised early developmental-type rRNA in other *Danio* species as well as Cyprinids and even distantly-related organisms of the Ostariophysi superorder. This would also facilitate a better phylogenetic comparison

of the species which potentially possess the additional ribosomal type. Since the majority of 2'-OH methylations in rRNA are guided by snoRNA, altering the levels of snoRNA or targeted snoRNA specific KO experiments seem to be the preferred approach to understanding the functional significance of these modifications, although the opinion still persists that gain-of-function studies are more valuable. In our ongoing experiments, we are presently working to study the effects of a few interesting SNORDs by KO of said SNORDs in zebrafish embryos. A more detailed phylogenetic comparison of several SNORDs is also being studied which can hopefully help elucidate the evolutionary conservation patterns of more SNORDs.

7 References

Abbate J, Rottman F (1972) Gas chromatographic method for determination of 2'-O-methylation in RNA. Anal Biochem 47:378–388. https://doi.org/10.1016/0003-2697(72)90131-5

Aida T (1921) On the inheritance of color in a fresh-water fish, *aplocheilus latipes* temmick and schlegel, with special reference to sex-linked inheritance. Genetics 6:554–573. https://doi.org/10.1093/genetics/6.6.554

Aquino GRR, Hackert P, Krogh N, et al (2021) The RNA helicase Dbp7 promotes domain V/VI compaction and stabilization of inter-domain interactions during early 60S assembly. Nat Commun 12:6152. https://doi.org/10.1038/s41467-021-26208-9

Aubert M, O'Donohue M-F, Lebaron S, Gleizes P-E (2018) Pre-Ribosomal RNA Processing in Human Cells: From Mechanisms to Congenital Diseases. Biomolecules 8:123. https://doi.org/10.3390/biom8040123

Bachellerie JP, Cavaillé J, Hüttenhofer A (2002) The expanding snoRNA world. Biochimie 84:775–90. https://doi.org/10.1016/s0300-9084(02)01402-5

Bailey AD, Talkish J, Ding H, et al (2022) Concerted modification of nucleotides at functional centers of the ribosome revealed by single-molecule RNA modification profiling. Elife 11:. https://doi.org/10.7554/eLife.76562

Barneche F, Gaspin C, Guyot R, Echeverría M (2001) Identification of 66 box C/D snoRNAs in Arabidopsis thaliana: extensive gene duplications generated multiple isoforms predicting new ribosomal RNA 2'-O-methylation sites. J Mol Biol 311:57–73. https://doi.org/10.1006/jmbi.2001.4851

Bartschat S, Kehr S, Tafer H, et al (2014) snoStrip: a snoRNA annotation pipeline. Bioinformatics 30:115–116. https://doi.org/10.1093/bioinformatics/btt604

Baßler J, Hurt E (2019) Eukaryotic Ribosome Assembly. Annu Rev Biochem 88:281–306. https://doi.org/10.1146/annurev-biochem-013118-110817

Bernier CR, Petrov AS, Waterbury CC, et al (2014) RiboVision suite for visualization and analysis of ribosomes. Faraday Discuss 169:195–207. https://doi.org/10.1039/C3FD00126A

Birkedal U, Christensen-Dalsgaard M, Krogh N, et al (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. Angewandte Chemie - International Edition 54:451–455. https://doi.org/10.1002/anie.201408362

Boccaletto P, Machnicka MA, Purta E, et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46:D303–D307. https://doi.org/10.1093/nar/gkx1030

Boivin V, Deschamps-Francoeur G, Couture S, et al (2018) Simultaneous sequencing of coding and noncoding RNA reveals a human transcriptome dominated by a small number of highly expressed noncoding genes. RNA 24:950–965. https://doi.org/10.1261/rna.064493.117

Bonnerot C, Pintard L, Lutfalla G (2003) Functional Redundancy of Spb1p and a snR52-Dependent Mechanism for the 2′-O-Ribose Methylation of a Conserved rRNA Position in Yeast. Mol Cell 12:1309–1315. https://doi.org/10.1016/S1097-2765(03)00435-0

Boschi-Muller S, Motorin Y (2013) Chemistry enters nucleic acids biology: Enzymatic mechanisms of RNA modification. Biochemistry (Moscow) 78:1392–1404. https://doi.org/10.1134/S0006297913130026

Bratkovič T, Božič J, Rogelj B (2020) Functional diversity of small nucleolar RNAs. Nucleic Acids Res 48:1627–1651. https://doi.org/10.1093/nar/gkz1140

Breder CM, Rosen DE (1966) Modes of reproduction in fishes. T.F.H. Publications., Neptune City, N.J

Buchhaupt M, Peifer C, Entian K-D (2007) Analysis of 2′-O-methylated nucleosides and pseudouridines in ribosomal RNAs using DNAzymes. Anal Biochem 361:102–108. https://doi.org/10.1016/j.ab.2006.11.001

Carninci P, Hayashizaki Y (2007) Noncoding RNA transcription beyond annotated genes. Curr Opin Genet Dev 17:139–144. https://doi.org/10.1016/j.gde.2007.02.008

Catania KC (2019) The Astonishing Behavior of Electric Eels. Front Integr Neurosci 13:. https://doi.org/10.3389/fnint.2019.00023

Cavaillé J (2017) Box C/D small nucleolar RNA genes and the Prader-Willi syndrome: a complex interplay. Wiley Interdiscip Rev RNA 8:e1417. https://doi.org/10.1002/wrna.1417

Cavaillé J, Buiting K, Kiefmann M, et al (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc Natl Acad Sci U S A 97:14311–6. https://doi.org/10.1073/pnas.250426397

Cavaillé J, Chetouani F, Bachellerie J-P (1999) The yeast Saccharomyces cerevisiae YDL112w ORF encodes the putative 2′-O-ribose methyltransferase catalyzing the formation of Gm18 in tRNAs. RNA 5:S1355838299981475. https://doi.org/10.1017/S1355838299981475

Chandler VL, Maler BA, Yamamoto KR (1983) DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. Cell 33:489–499. https://doi.org/10.1016/0092-8674(83)90430-0

Chou H-J, Donnard E, Gustafsson HT, et al (2017) Transcriptome-wide Analysis of Roles for tRNA Modifications in Translational Regulation. Mol Cell 68:978-992.e4. https://doi.org/10.1016/j.molcel.2017.11.002

Chowdhury K, Lin S, Lai S-L (2022) Comparative Study in Zebrafish and Medaka Unravels the Mechanisms of Tissue Regeneration. Front Ecol Evol 10:. https://doi.org/10.3389/fevo.2022.783818

Dai Q, Moshitch-Moshkovitz S, Han D, et al (2017) Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 14:695–698. https://doi.org/10.1038/nmeth.4294

Dalla Venezia N, Vincent A, Marcel V, et al (2019) Emerging Role of Eukaryote Ribosomes in Translational Control. Int J Mol Sci 20:1226. https://doi.org/10.3390/ijms20051226

Danny Bergeron, Fafard-Couture É, Scott MS (2020) Small nucleolar RNAs: Continuing identification of novel members and increasing diversity of their molecular mechanisms of action. Biochem Soc Trans 48:645–656. https://doi.org/10.1042/BST20191046

de Santana CD, Crampton WGR, Dillman CB, et al (2019) Unexpected species diversity in electric eels with a description of the strongest living bioelectricity generator. Nat Commun 10:4000. https://doi.org/10.1038/s41467-019-11690-z

Decatur WA, Fournier MJ (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. J Biol Chem 278:695–8. https://doi.org/10.1074/jbc.R200023200

Dennis PP, Tripp V, Lui L, et al (2015) C/D box sRNA-guided 2′-O-methylation patterns of archaeal rRNA molecules. BMC Genomics 16:632. https://doi.org/10.1186/s12864-015-1839-z

Deogharia M, Majumder M (2018) Guide snoRNAs: Drivers or Passengers in Human Disease? Biology (Basel) 8:1. https://doi.org/10.3390/biology8010001

Deryusheva S, Talhouarne GJS, Gall JG (2020) “Lost and Found”: SnoRNA Annotation in the Xenopus Genome and Implications for Evolutionary Studies. Mol Biol Evol 37:149–166. https://doi.org/10.1093/molbev/msz209

Dieci G, Preti M, Montanini B (2009) Eukaryotic snoRNAs: A paradigm for gene expression flexibility. Genomics 94:83–88. https://doi.org/10.1016/j.ygeno.2009.05.002

Dooley K, Zon LI (2000) Zebrafish the perfect model. Genetics of disease 10:252–256

Dragon F, Gallagher JEG, Compagnone-Post PA, et al (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature 417:967–970. https://doi.org/10.1038/nature00769

Droogmans L, Grosjean H (1991) 2′-O-methylation and inosine formation in the wobble position of anticodon-substituted tRNA-Phe in a homologous yeast in vitro system. Biochimie 73:1021–1025. https://doi.org/10.1016/0300-9084(91)90143-O

Elliot D, Ladomery M (2016) Molecular Biology of RNA, 2nd edn. Oxford University Press, Oxford

Emmott E, Jovanovic M, Slavov N (2019) Ribosome Stoichiometry: From Form to Function. Trends Biochem Sci 44:95–109. https://doi.org/10.1016/j.tibs.2018.10.009

Engel C, Sainsbury S, Cheung AC, et al (2013) RNA polymerase I structure and transcription regulation. Nature 502:650–655. https://doi.org/10.1038/nature12712

Eytan RI, Evans BR, Dornburg A, et al (2015) Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. BMC Evol Biol 15:113. https://doi.org/10.1186/s12862-015-0415-0

Fafard-Couture É, Bergeron D, Couture S, et al (2021) Annotation of snoRNA abundance across human tissues reveals complex snoRNA-host gene relationships. Genome Biol 22:172. https://doi.org/10.1186/s13059-021-02391-2

Falaleeva M, Welden JR, Duncan MJ, Stamm S (2017) C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: Old dogs show new tricks. Bioessays 39:. https://doi.org/10.1002/bies.201600264

Fernández-Tornero C, Moreno-Morcillo M, Rashid UJ, et al (2013) Crystal structure of the 14-subunit RNA polymerase I. Nature 502:644–649. https://doi.org/10.1038/nature12636

Ferretti MB, Karbstein K (2019) Does functional specialization of ribosomes really exist? https://doi.org/10.1261/rna

Fujii K, Susanto TT, Saurabh S, Barna M (2018) Decoding the Function of Expansion Segments in Ribosomes. Mol Cell 72:1013-1020.e6. https://doi.org/10.1016/j.molcel.2018.11.023

Furutani-Seiki M, Wittbrodt J (2004) Medaka and zebrafish, an evolutionary twin study. Mech Dev 121:629–637. https://doi.org/10.1016/j.mod.2004.05.010

Gao Q, Xie H, Zhan H, et al (2017) Prognostic Values of Long Noncoding RNA GAS5 in Various Carcinomas: An Updated Systematic Review and Meta-Analysis. Front Physiol 8:. https://doi.org/10.3389/fphys.2017.00814

Gay DM, Lund AH, Jansson MD (2022) Translational control through ribosome heterogeneity and functional specialization. Trends Biochem Sci 47:66–81. https://doi.org/10.1016/j.tibs.2021.07.001

Gert KR, Quio LEC, Novatchkova M, et al (2021) Reciprocal zebrafish-medaka hybrids reveal maternal control of zygotic genome activation timing. bioRxiv 2021.11.03.467109. https://doi.org/10.1101/2021.11.03.467109

Glasser A-L, el Adlouni C, Keith G, et al (1992) Presence and coding properties of 2'- *O* -methyl-5-carbamoylmethyluridine (ncm ^{5} Um) in the wobble position of the anticodon of tRNA ^{Leu} (U*AA) from brewer's yeast. FEBS Lett 314:381–385. https://doi.org/10.1016/0014-5793(92)81510-S

Goustin A, Thepsuwan P, Kosir M, Lipovich L (2019) The Growth-Arrest-Specific (GAS)-5 Long Non-Coding RNA: A Fascinating lncRNA Widely Expressed in Cancers. Noncoding RNA 5:46. https://doi.org/10.3390/ncrna5030046

Graziadei A, Masiewicz P, Lapinaite A, Carlomagno T (2016) Archaea box C/D enzymes methylate two distinct substrate rRNA sequences with different efficiency. RNA 22:764–772. https://doi.org/10.1261/rna.054320.115

Grosjean H, Gaspin C, Marck C, et al (2008) RNomics and Modomics in the halophilic archaea Haloferax volcanii: identification of RNA modification genes. BMC Genomics 9:470. https://doi.org/10.1186/1471-2164-9-470

Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol Rev 58:10–26. https://doi.org/10.1128/mr.58.1.10-26.1994

Hamaguchi S (1987) The structure of the germinal dense bodies (nuages) during differentiation of the male germ line of the teleost, Oryzias latipes. Cell Tissue Res 248:. https://doi.org/10.1007/BF00218205

Hebras J, Krogh N, Marty V, et al (2020) Developmental changes of rRNA ribose methylations in the mouse. RNA Biol 17:150–164. https://doi.org/10.1080/15476286.2019.1670598

Helfman G, Collette BB, Facey DE, Bowen BW (2010) The diversity of fishes: biology, evolution and ecology, 2nd edn. Aquaculture International 18:707–708. https://doi.org/10.1007/s10499-009-9264-4

Henras AK, Soudet J, Gérus M, et al (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. Cellular and Molecular Life Sciences 65:2334–2359. https://doi.org/10.1007/s00018-008-8027-0

Higa S, Maeda N, Kenmochi N, Tanaka T (2002) Location of 2'-O-methyl nucleotides in 26S rRNA and methylation guide snoRNAs in Caenorhabditis elegans. Biochem Biophys Res Commun 297:1344–1349. https://doi.org/10.1016/S0006-291X(02)02377-X

Higa-Nakamine S, Suzuki T, Uechi T, et al (2012) Loss of ribosomal RNA modification causes developmental defects in zebrafish. Nucleic Acids Res 40:391–398. https://doi.org/10.1093/nar/gkr700

Hilgers L, Schwarzer J (2019) The untapped potential of medaka and its wild relatives. Elife 8:. https://doi.org/10.7554/eLife.46994

Hillis DM, Dixon MT (1991) Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. Q Rev Biol 66:411–453

Hinnebusch AG (2017) Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. Trends Biochem Sci 42:589–611. https://doi.org/10.1016/j.tibs.2017.03.004

Hinnebusch AG (2014) The Scanning Mechanism of Eukaryotic Translation Initiation. Annu Rev Biochem 83:779–812. https://doi.org/10.1146/annurev-biochem-060713-035802

Hoegg S, Brinkmann H, Taylor JS, Meyer A (2004) Phylogenetic Timing of the Fish-Specific Genome Duplication Correlates with the Diversification of Teleost Fish. J Mol Evol 59:190–203. https://doi.org/10.1007/s00239-004-2613-z

Hoeppner MP, Poole AM (2012) Comparative genomics of eukaryotic small nucleolar RNAs reveals deep evolutionary ancestry amidst ongoing intragenomic mobility. BMC Evol Biol 12:. https://doi.org/10.1186/1471-2148-12-183

Holley CL, Li MW, Scruggs BS, et al (2015) Cytosolic accumulation of small nucleolar RNAs (snoRNAs) is dynamically regulated by NADPH oxidase. J Biol Chem 290:11741–8. https://doi.org/10.1074/jbc.M115.637413

Howe K, Clark MD, Torroja CF, et al (2013) The zebrafish reference genome sequence and its relationship to the human genome. Nature 496:498–503. https://doi.org/10.1038/nature12111

Hu C-K, Brunet A (2018) The African turquoise killifish: A research organism to study vertebrate aging and diapause. Aging Cell 17:e12757. https://doi.org/10.1111/acel.12757

Hudson WH, Pickard MR, de Vera IMS, et al (2014) Conserved sequence-specific lincRNA–steroid receptor interactions drive transcriptional repression and direct cell fate. Nat Commun 5:5395. https://doi.org/10.1038/ncomms6395

Hughes LC, Ortí G, Huang Y, et al (2018) Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proceedings of the National Academy of Sciences 115:6249–6254. https://doi.org/10.1073/PNAS.1719358115

Iwamatsu T (2004) Stages of normal development in the medaka Oryzias latipes. Mech Dev 121:605–618

Jansson MD, Häfner SJ, Altinel K, et al (2021) Regulation of translation by site-specific ribosomal RNA methylation. Nat Struct Mol Biol 28:889–899. https://doi.org/10.1038/s41594-021-00669-4

Jorjani H, Kehr S, Jedlinski DJ, et al (2016) An updated human snoRNAome. Nucleic Acids Res 44:5068–82. https://doi.org/10.1093/nar/gkw386

Kasahara M, Naruse K, Sasaki S, et al (2007) The medaka draft genome and insights into vertebrate genome evolution. Nature 447:714–719. https://doi.org/10.1038/nature05846

Kawai G, Ue H, Yasuda M, et al (1991) Relation between functions and conformational characteristics of modified nucleosides found in tRNAs. Nucleic Acids Symp Ser 49–50

Kehr S, Bartschat S, Tafer H, et al (2014) Matching of soulmates: Coevolution of snoRNAs and their targets. Mol Biol Evol 31:455–467. https://doi.org/10.1093/molbev/mst209

Kim BH, Zhang G (2020) Generating Stable Knockout Zebrafish Lines by Deleting Large Chromosomal Fragments Using Multiple gRNAs. G3 Genes|Genomes|Genetics 10:1029–1037. https://doi.org/10.1534/g3.119.401035

Kimmel CB, Ballard WW, Kimmel SR, et al (1995) Stages of Embryonic Development of the Zebrafish. Developmental Dynamics 203:253–310. https://doi.org/10.1002/aja.1002030302

Kino T, Hurt DE, Ichijo T, et al (2010) Noncoding RNA Gas5 Is a Growth Arrest– and Starvation-Associated Repressor of the Glucocorticoid Receptor. Sci Signal 3:. https://doi.org/10.1126/scisignal.2000568

Klootwijk J, Planta RJ (1973) Analysis of the Methylation Sites in Yeast Ribosomal RNA. Eur J Biochem 39:325–333. https://doi.org/10.1111/j.1432-1033.1973.tb03130.x

Kondo M, Nanda I, Schmid M, Schartl M (2009) Sex Determination and Sex Chromosome Evolution: Insights from Medaka. Sexual Development 3:88–98. https://doi.org/10.1159/000223074

Koš M, Tollervey D (2010) Yeast Pre-rRNA Processing and Modification Occur Cotranscriptionally. Mol Cell 37:809–820. https://doi.org/10.1016/j.molcel.2010.02.024

Kossack ME, Draper BW (2019) Genetic regulation of sex determination and maintenance in zebrafish (Danio rerio). pp 119–149

Kotelawala L, Grayhack EJ, Phizicky EM (2008) Identification of yeast tRNA Um $_{44}$ 2′-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA ^{Ser} species. RNA 14:158–169. https://doi.org/10.1261/rna.811008

Krogh N, Birkedal U, Nielsen H (2017a) RiboMeth-seq: Profiling of 2′-O-Me in RNA. pp 189–209

Krogh N, Jansson MD, Häfner SJ, et al (2016) Profiling of 2′-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. Nucleic Acids Res 44:7884–7895. https://doi.org/10.1093/nar/gkw482

Krogh N, Kongsbak-Wismann M, Geisler C, Nielsen H (2017b) Substoichiometric ribose methylations in spliceosomal snRNAs. Org Biomol Chem 15:8872–8876. https://doi.org/10.1039/C7OB02317K

Krogh N, Nielsen H (2019) Sequencing-based methods for detection and quantitation of ribose methylations in RNA. Methods 156:5–15. https://doi.org/10.1016/j.ymeth.2018.11.017

Kroll F, Powell GT, Ghosh M, et al (2021) A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. Elife 10:. https://doi.org/10.7554/eLife.59683

Kufel J, Grzechnik P (2019) Small Nucleolar RNAs Tell a Different Tale. Trends in Genetics 35:104–117. https://doi.org/10.1016/j.tig.2018.11.005

Kumar S, Mohapatra T (2021) Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression. Front Cell Dev Biol 9:. https://doi.org/10.3389/fcell.2021.628415

Kumar V (2021) Ribosomal biogenesis in eukaryotes. In: Emerging Concepts in Ribosome Structure, Biogenesis, and Function. Elsevier, pp 129–150

Kuratani M, Bessho Y, Nishimoto M, et al (2008) Crystal Structure and Mutational Study of a Unique SpoU Family Archaeal Methylase that Forms 2'-O-Methylcytidine at Position 56 of tRNA. J Mol Biol 375:1064–1075. https://doi.org/10.1016/j.jmb.2007.11.023

Lane BG, Tamaoki T (1967) Studies of the chain termini and alkali-stable dinucleotide sequences in 16 s and 28 s ribosomal RNA from L Cells. J Mol Biol 27:335–348. https://doi.org/10.1016/0022-2836(67)90024-1

Lapeyre B, Purushothaman SK (2004) Spb1p-Directed Formation of Gm2922 in the Ribosome Catalytic Center Occurs at a Late Processing Stage. Mol Cell 16:663–669. https://doi.org/10.1016/j.molcel.2004.10.022

Lee SJ, Swanson MJ, Sattlegger E (2015) Gcn1 contacts the small ribosomal protein Rps10, which is required for full activation of the protein kinase Gcn2. Biochemical Journal 466:547–559. https://doi.org/10.1042/BJ20140782

Lestrade L (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res 34:D158–D162. https://doi.org/10.1093/nar/gkj002

Li W, Li N, Shi K, Chen Q (2017) Systematic review and meta-analysis of the utility of long non-coding RNA GAS5 as a diagnostic and prognostic cancer biomarker. Oncotarget 8:66414–66425. https://doi.org/10.18632/oncotarget.19040

Liang J, Wen J, Huang Z, et al (2019) Small Nucleolar RNAs: Insight Into Their Function in Cancer. Front Oncol 9:. https://doi.org/10.3389/fonc.2019.00587

Lien S, Koop BF, Sandve SR, et al (2016) The Atlantic salmon genome provides insights into rediploidization. Nature 533:200–205. https://doi.org/10.1038/nature17164

Liew WC, Bartfai R, Lim Z, et al (2012) Polygenic Sex Determination System in Zebrafish. PLoS One 7:e34397. https://doi.org/10.1371/journal.pone.0034397

Locati MD, Pagano JFB, Ensink WA, et al (2017a) Linking maternal and somatic 5S rRNA types with different sequence-specific non-LTR retrotransposons. RNA 23:446–456. https://doi.org/10.1261/rna.059642.116

Locati MD, Pagano JFB, Girard G, et al (2017b) Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. https://doi.org/10.1261/rna

Lykke-Andersen S, Chen Y, Ardal BR, et al (2016) Corrigendum: Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes Dev 30:1128–34. https://doi.org/10.1101/gad.281881.116

Maden BEH (2001) Mapping 2'-O-Methyl Groups in Ribosomal RNA. Methods 25:374–382. https://doi.org/10.1006/meth.2001.1250

Maden BEH (1988) Locations of methyl groups in 28 S rRNA of Xenopus laevis and man. J Mol Biol 201:289–314. https://doi.org/10.1016/0022-2836(88)90139-8

Maden BEH (1986) Identification of the locations of the methyl groups in 18 S ribosomal RNA from Xenopus laevis and man. J Mol Biol 189:681–699. https://doi.org/10.1016/0022-2836(86)90498-5

Makarova JA, Kramerov DA (2011) SNOntology: Myriads of novel snornas or just a mirage? BMC Genomics 12

Makarova JA, Kramerov DA (2009) Analysis of C/D box snoRNA genes in vertebrates: The number of copies decreases in placental mammals. Genomics 94:11–19. https://doi.org/10.1016/j.ygeno.2009.02.003

Mannoor K, Liao J, Jiang F (2012) Small nucleolar RNAs in cancer. Biochim Biophys Acta 1826:121–8. https://doi.org/10.1016/j.bbcan.2012.03.005

Martens-Uzunova ES, Hoogstrate Y, Kalsbeek A, et al (2015) C/D-box snoRNA-derived RNA production is associated with malignant transformation and metastatic progression in prostate cancer. Oncotarget 6:17430–17444. https://doi.org/10.18632/oncotarget.4172

Matsuda M, Matsuda C, Hamaguchi S, Sakaizumi M (1998) Identification of the sex chromosomes of the medaka, <i>Oryzias latipes</i>, by fluorescence in situ hybridization. Cytogenet Genome Res 82:257–262. https://doi.org/10.1159/000015113

McCann KL, Kavari SL, Burkholder AB, et al (2020) H/ACA snoRNA levels are regulated during stem cell differentiation. Nucleic Acids Res 48:8686–8703. https://doi.org/10.1093/nar/gkaa612

McClintock B (1934) The relation of a particular chromosomal element to the development of the nucleoli in Zea mays. Zeitschrift für Zellforschung und mikroskopische Anatomie 21:294–326

McMahon M, Contreras A, Holm M, et al (2019) A single H/ACA small nucleolar RNA mediates tumor suppression downstream of oncogenic RAS. Elife 8:. https://doi.org/10.7554/eLife.48847

Mikalsen S-O, Tausen M, í Kongsstovu S (2020) Phylogeny of teleost connexins reveals highly inconsistent intra- and interspecies use of nomenclature and misassemblies in recent teleost chromosome assemblies. BMC Genomics 21:223. https://doi.org/10.1186/s12864-020-6620-2

Mindell DP, Honeycutt RL (1990) Ribosomal RNA in Vertebrates: Evolution and Phylogenetic Applications. Annu Rev Ecol Syst 21:541–566

Monaco P, Marcel V, Diaz J-J, Catez F (2018) 2'-O-Methylation of Ribosomal RNA: Towards an Epitranscriptomic Control of Translation? Biomolecules 8:106. https://doi.org/10.3390/biom8040106

Motorin Y, Marchand V (2018) Detection and Analysis of RNA Ribose 2'-O-Methylations: Challenges and Solutions. Genes (Basel) 9:642. https://doi.org/10.3390/genes9120642

Naylor RL, Hardy RW, Buschmann AH, et al (2021) A 20-year retrospective review of global aquaculture. Nature 591:551–563. https://doi.org/10.1038/s41586-021-03308-6

Near TJ, Eytan RI, Dornburg A, et al (2012) Resolution of ray-finned fish phylogeny and timing of diversification. Proceedings of the National Academy of Sciences 109:13698–13703. https://doi.org/10.1073/pnas.1206625109

Nepal C, Hadzhiev Y, Balwierz P, et al (2020) Dual-initiation promoters with intertwined canonical and TCT/TOP transcription start sites diversify transcript processing. Nat Commun 11:. https://doi.org/10.1038/s41467-019-13687-0

Ning W, Fei J, Gonzalez RL (2014) The ribosome uses cooperative conformational changes to maximize and regulate the efficiency of translation. Proceedings of the National Academy of Sciences 111:12073–12078. https://doi.org/10.1073/pnas.1401864111

Okuyama T, Yokoi S, Takeuchi H (2017) Molecular basis of social competence in medaka fish. Dev Growth Differ 59:211–218

Olsen GJ, Woese CR (1993) Ribosomal RNA: a key to phylogeny. The FASEB Journal 7:113–123

Omer AD, Ziesche S, Decatur WA, et al (2003) RNA-modifying machines in archaea. Mol Microbiol 48:617–29. https://doi.org/10.1046/j.1365-2958.2003.03483.x

Ortega-Recalde O, Day RC, Gemmell NJ, Hore TA (2019) Zebrafish preserve global germline DNA methylation while sex-linked rDNA is amplified and demethylated during feminisation. Nat Commun 10:3053. https://doi.org/10.1038/s41467-019-10894-7

Païs de Barros JP, Keith G, el Adlouni C, et al (1996) 2'-O-methyl-5-formylcytidine (f5Cm), a new modified nucleotide at the "wobble" of two cytoplasmic tRNAs Leu (NAA) from bovine liver. Nucleic Acids Res 24:1489–96. https://doi.org/10.1093/nar/24.8.1489

Patra Bhattacharya D, Canzler S, Kehr S, et al (2016) Phylogenetic distribution of plant snoRNA families. BMC Genomics 17:. https://doi.org/10.1186/s12864-016-3301-2

Patton EE, Zon LI, Langenau DM (2021) Zebrafish disease models in drug discovery: from preclinical modelling to clinical trials. Nat Rev Drug Discov 20:611–628. https://doi.org/10.1038/s41573-021-00210-8

Pederson T (2011) The Nucleolus. Cold Spring Harb Perspect Biol 3:a000638–a000638. https://doi.org/10.1101/cshperspect.a000638

Penzo M, Montanaro L (2018) Turning Uridines around: Role of rRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function. Biomolecules 8:38. https://doi.org/10.3390/biom8020038

Persson B (1997) The spoU gene of Escherichia coli, the fourth gene of the spoT operon, is essential for tRNA (Gm18) 2'-O-methyltransferase activity. Nucleic Acids Res 25:4093–4097. https://doi.org/10.1093/nar/25.20.4093

Picart-Picolo A, Picault N, Pontvianne F (2019) Ribosomal RNA genes shape chromatin domains associating with the nucleolus. Nucleus 10:67–72. https://doi.org/10.1080/19491034.2019.1591106

Pickard MR, Williams GT (2016) The hormone response element mimic sequence of GAS5 lncRNA is sufficient to induce apoptosis in breast cancer cells. Oncotarget 7:10104–10116. https://doi.org/10.18632/oncotarget.7173

Postlethwait JH, Yan YL, Gates MA, et al (1998) Vertebrate genome evolution and the zebrafish gene map. Nat Genet 18:345–9. https://doi.org/10.1038/ng0498-345

Qiu F (1999) Selective detection of ribose-methylated nucleotides in RNA by a mass spectrometry-based method. Nucleic Acids Res 27:20e–220. https://doi.org/10.1093/nar/27.18.e20

Rafferty SA, Quinn TA (2018) A beginner's guide to understanding and implementing the genetic modification of zebrafish. Prog Biophys Mol Biol 138:3–19. https://doi.org/10.1016/j.pbiomolbio.2018.07.005

Rahman Khan F, Sulaiman Alhewairini S (2019) Zebrafish (*Danio rerio*) as a Model Organism. In: Current Trends in Cancer Management. IntechOpen

Ramachandran S, Krogh N, Jørgensen TE, et al (2020) The shift from early to late types of ribosomes in zebrafish development involves changes at a subset of rRNA 2′-O-Me sites. RNA 26:1919–1934. https://doi.org/10.1261/RNA.076760.120

Ramakrishnan V (2002) Ribosome Structure and the Mechanism of Translation. Cell 108:557–572. https://doi.org/10.1016/S0092-8674(02)00619-0

Ramesh M, Woolford JL (2016) Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. RNA 22:1153–1162. https://doi.org/10.1261/rna.056705.116

Reichow SL, Hamma T, Ferré-D'Amaré AR, Varani G (2007) The structure and function of small nucleolar ribonucleoproteins. Nucleic Acids Res 35:1452–64. https://doi.org/10.1093/nar/gkl1172

Renalier M-H, Joseph N, Gaspin C, et al (2005) The Cm56 tRNA modification in archaea is catalyzed either by a specific 2′-O-methylase, or a C/D sRNP. RNA 11:1051–63. https://doi.org/10.1261/rna.2110805

Richard P, Kiss T (2006) Integrating snoRNP assembly with mRNA biogenesis. EMBO Rep 7:590–2. https://doi.org/10.1038/sj.embor.7400715

Rimer JM, Lee J, Holley CL, et al (2018) Long-range function of secreted small nucleolar RNAs that direct 2′-O-methylation. J Biol Chem 293:13284–13296. https://doi.org/10.1074/jbc.RA118.003410

Rodgers ML, Woodson SA (2021) A roadmap for rRNA folding and assembly during transcription. Trends Biochem Sci 46:889–901. https://doi.org/10.1016/j.tibs.2021.05.009

Sardana MK, Fuke M (1980) A rapid procedure to determine the content of 2′-O-methylation in RNA by homochromatography. Anal Biochem 103:285–288. https://doi.org/10.1016/0003-2697(80)90611-9

Satoh N, Egami N (1972) Sex differentiation of germ cells in the teleost, *Oryzias latipes* , during normal embryonic development. Development 28:385–395. https://doi.org/10.1242/dev.28.2.385

Schaffer JE (2020) Death by lipids: The role of small nucleolar RNAs in metabolic stress. J Biol Chem 295:8628–8635. https://doi.org/10.1074/jbc.AW120.011105

Schnare MN, Gray MW (2011) Complete Modification Maps for the Cytosolic Small and Large Subunit rRNAs of Euglena gracilis: Functional and Evolutionary Implications of Contrasting Patterns between the Two rRNA Components. J Mol Biol 413:66–83. https://doi.org/10.1016/j.jmb.2011.08.037

Schubert HL, Blumenthal RM, Cheng X (2003) Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci 28:329–335. https://doi.org/10.1016/S0968-0004(03)00090-2

Schuller AP, Green R (2018) Roadblocks and resolutions in eukaryotic translation. Nat Rev Mol Cell Biol 19:526–541. https://doi.org/10.1038/s41580-018-0011-4

Seemann SE, Mirza AH, Bang-Berthelsen CH, et al (2022) Does rapid sequence divergence preclude RNA structure conservation in vertebrates? Nucleic Acids Res 50:2452–2463. https://doi.org/10.1093/nar/gkac067

Seistrup KH, Rose S, Birkedal U, et al (2016) Bypassing rRNA methylation by RsmA/Dim1during ribosome maturation in the hyperthermophilic archaeon *Nanoarchaeum equitans*. Nucleic Acids Res gkw839. https://doi.org/10.1093/nar/gkw839

Selman K, Wallace RA, Sarka A, Qi X (1993) Stages of oocyte development in the zebrafish,Brachydanio rerio. J Morphol 218:203–224. https://doi.org/10.1002/jmor.1052180209

Sergeeva OV, Bogdanov AA, Sergiev PV (2015) What do we know about ribosomal RNA methylation in Escherichia coli? Biochimie 117:110–118. https://doi.org/10.1016/j.biochi.2014.11.019

Sharma KK, Sharrna OP, Tripathi NK (1998) Female heterogamety in Danio rerio (Cypriniformes: Cyprinidae). Proceedings-National Academy of Sciences India Section B 68:123–126

Sharma S, Lafontaine DLJ (2015) 'View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification. Trends Biochem Sci 40:560–575. https://doi.org/10.1016/j.tibs.2015.07.008

Sharma S, Langhendries J-L, Watzinger P, et al (2015) Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. Nucleic Acids Res 43:2242–58. https://doi.org/10.1093/nar/gkv075

Sharma S, Yang J, van Nues R, et al (2017) Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. PLoS Genet 13:e1006804. https://doi.org/10.1371/journal.pgen.1006804

Shaw P, Brown J (2012) Nucleoli: Composition, Function, and Dynamics. Plant Physiol 158:44–51. https://doi.org/10.1104/pp.111.188052

Shima A, Mitani H (2004) Medaka as a research organism: past, present and future. Mech Dev 121:599–604. https://doi.org/10.1016/j.mod.2004.03.011

Signore IA, Guerrero N, Loosli F, et al (2009) Zebrafish and medaka: model organisms for a comparative developmental approach of brain asymmetry. Philosophical Transactions of the Royal Society B: Biological Sciences 364:991–1003. https://doi.org/10.1098/rstb.2008.0260

Simonović M, Steitz TA (2009) A structural view on the mechanism of the ribosome-catalyzed peptide bond formation. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1789:612–623. https://doi.org/10.1016/j.bbagrm.2009.06.006

Simsek D, Barna M (2017) An emerging role for the ribosome as a nexus for post-translational modifications. Curr Opin Cell Biol 45:92–101

Singh PP, Arora J, Isambert H (2015) Identification of Ohnolog Genes Originating from Whole Genome Duplication in Early Vertebrates, Based on Synteny Comparison across Multiple Genomes. PLoS Comput Biol 11:e1004394. https://doi.org/10.1371/journal.pcbi.1004394

Singh PP, Isambert H (2019) OHNOLOGS v2: a comprehensive resource for the genes retained from whole genome duplication in vertebrates. Nucleic Acids Res. https://doi.org/10.1093/nar/gkz909

Sloan KE, Warda AS, Sharma S, et al (2017) Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol 14:1138–1152. https://doi.org/10.1080/15476286.2016.1259781

Smietanski M, Werner M, Purta E, et al (2014) Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. Nat Commun 5:3004. https://doi.org/10.1038/ncomms4004

Smith CM, Steitz JA (1998) Classification of *gas5* as a Multi-Small-Nucleolar-RNA (snoRNA) Host Gene and a Member of the 5'-Terminal Oligopyrimidine Gene Family Reveals Common Features of snoRNA Host Genes. Mol Cell Biol 18:6897–6909. https://doi.org/10.1128/MCB.18.12.6897

Sonenberg N, Hinnebusch AG (2009) Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell 136:731–745. https://doi.org/10.1016/j.cell.2009.01.042

Sprinzl M, Vassilenko KS (2005) Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res 33:D139-40. https://doi.org/10.1093/nar/gki012

Stahl G, McCarty GP, Farabaugh PJ (2002) Ribosome structure: revisiting the connection between translational accuracy and unconventional decoding. Trends Biochem Sci 27:178–183. https://doi.org/10.1016/S0968-0004(02)02064-9

Streisinger G, Walker C, Dower N, et al (1981) Production of clones of homozygous diploid zebra fish (Brachydanio rerio). Nature 291:293–296. https://doi.org/10.1038/291293a0

Takahashi H (1977) Juvenile hermaphroditism in the zebrafish, Brachydanio rerio. Hokkaido University Fisheries Research Bulletin. 28:57–65

Takeda N, Nakamura M, Yoshizumi H, Tatematsu A (1994) Detection of ribose-methylated nucleotides in Pyrodictium occultum tRNA by liquid chromatography—frit-fast atom bombardment mass spectrometry. J Chromatogr B Biomed Sci Appl 660:223–233. https://doi.org/10.1016/0378-4347(94)00299-1

Takezaki N (2021) Resolving the Early Divergence Pattern of Teleost Fish Using Genome-Scale Data. Genome Biol Evol 13:. https://doi.org/10.1093/gbe/evab052

Talross GJS, Deryusheva S, Gall JG (2021) Stable lariats bearing a snoRNA (slb-snoRNA) in eukaryotic cells: A level of regulation for guide RNAs. Proceedings of the National Academy of Sciences 118:. https://doi.org/10.1073/pnas.2114156118

Teame T, Zhang Z, Ran C, et al (2019) The use of zebrafish (Danio rerio) as biomedical models. Animal Frontiers 9:68–77. https://doi.org/10.1093/af/vfz020

Temminck CJ (1846) Fauna Japonica. Poisson 10:173–269

Thompson M, Haeusler RA, Good PD, Engelke DR (2003) Nucleolar Clustering of Dispersed tRNA Genes. Science (1979) 302:1399–1401. https://doi.org/10.1126/science.1089814

Tomkuvienė M, Ličytė J, Olendraitė I, et al (2017) Archaeal fibrillarin–Nop5 heterodimer 2′-*O* -methylates RNA independently of the C/D guide RNP particle. RNA 23:1329–1337. https://doi.org/10.1261/rna.059832.116

Towns WL, Begley TJ (2012) Transfer RNA Methytransferases and Their Corresponding Modifications in Budding Yeast and Humans: Activities, Predications, and Potential Roles in Human Health. DNA Cell Biol 31:434–454. https://doi.org/10.1089/dna.2011.1437

Tschochner H, Hurt E (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13:255–263. https://doi.org/10.1016/S0962-8924(03)00054-0

Tuzuki E, Egami N, Hyodo Y (1966) Multiplication and sex-differentiation of germ cells during development in the medaka, <I>Oryzias latipes</I>. Japanese Journal of Ichthyology 13:176–182. https://doi.org/10.11369/jji1950.13.176

Varadharajan S, Sandve SR, Gillard GB, et al (2018) The Grayling Genome Reveals Selection on Gene Expression Regulation after Whole-Genome Duplication. Genome Biol Evol 10:2785–2800. https://doi.org/10.1093/gbe/evy201

Vincenti S, de Chiara V, Bozzoni I, Presutti C (2007) The position of yeast snoRNA-coding regions within host introns is essential for their biosynthesis and for efficient splicing of the host pre-mRNA. RNA 13:138–150. https://doi.org/10.1261/rna.251907

Warner WA, Spencer DH, Trissal M, et al (2018) Expression profiling of snoRNAs in normal hematopoiesis and AML. Blood Adv 2:151–163. https://doi.org/10.1182/bloodadvances.2017006668

Wegnez M, Monier R (1972) Sequence heterogeneity of 5 S RNA in *Xenopus laevis*. FEBS Lett 25:13–20. https://doi.org/10.1016/0014-5793(72)80443-5

Welch M, Majerfeld I, Yarus M (1997) 23S rRNA Similarity from Selection for Peptidyl Transferase Mimicry. Biochemistry 36:6614–6623. https://doi.org/10.1021/bi963135j

Westerfield M (2000) The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio), 4th edn. University of Oregon Press, Eugene

Wilkinson ML, Crary SM, Jackman JE, et al (2007) The 2'-O-methyltransferase responsible for modification of yeast tRNA at position 4. RNA 13:404–413. https://doi.org/10.1261/rna.399607

Wilson CA, High SK, McCluskey BM, et al (2014) Wild Sex in Zebrafish: Loss of the Natural Sex Determinant in Domesticated Strains. Genetics 198:1291–1308. https://doi.org/10.1534/genetics.114.169284

Wittbrodt J, Shima A, Schartl M (2002) Medaka — a model organism from the far east. Nat Rev Genet 3:53–64. https://doi.org/10.1038/nrg704

Woodson SA (2011) RNA Folding Pathways and the Self-Assembly of Ribosomes. Acc Chem Res 44:1312–1319. https://doi.org/10.1021/ar2000474

Xu P, Zhang X, Wang X, et al (2014) Genome sequence and genetic diversity of the common carp, Cyprinus carpio. Nat Genet 46:1212–1219. https://doi.org/10.1038/ng.3098

Yamamoto T (1975) The Medaka, Oryzias latipes, and the Guppy, Lebistes reticularis. In: Handbook of Genetics. Springer US, Boston, MA, pp 133–149

Yamamoto T-O (1953) Artificially induced sex-reversal in genotypic males of the medaka (Oryzias latipes). Journal of Experimental Zoology 123:571–594. https://doi.org/10.1002/jez.1401230309

Yang J, Sharma S, Watzinger P, et al (2016) Mapping of Complete Set of Ribose and Base Modifications of Yeast rRNA by RP-HPLC and Mung Bean Nuclease Assay. PLoS One 11:e0168873. https://doi.org/10.1371/journal.pone.0168873

Yang L (2015) Splicing noncoding RNAs from the inside out. Wiley Interdiscip Rev RNA 6:651–60. https://doi.org/10.1002/wrna.1307

Yang X, Xie Z, Lei X, Gan R (2020) Long non-coding RNA GAS5 in human cancer (Review). Oncol Lett 20:2587–2594. https://doi.org/10.3892/ol.2020.11809

Yip WSV, Vincent NG, Baserga SJ (2013) Ribonucleoproteins in archaeal pre-rRNA processing and modification. Archaea 2013:614735. https://doi.org/10.1155/2013/614735

Yoshihama M, Nakao A, Kenmochi N (2013) snOPY: a small nucleolar RNA orthological gene database. BMC Res Notes 6:426. https://doi.org/10.1186/1756-0500-6-426

Yu YT, Shu MD, Steitz JA (1997) A new method for detecting sites of 2'-O-methylation in RNA molecules. RNA 3:324–31

Zang L, Torraca V, Shimada Y, Nishimura N (2022) Editorial: Zebrafish Models for Human Disease Studies. Front Cell Dev Biol 10:. https://doi.org/10.3389/fcell.2022.861941

Zemora G, Waldsich C (2010) RNA folding in living cells. RNA Biol 7:634–641. https://doi.org/10.4161/rna.7.6.13554

Zhang X, Huang B, Zhang L, et al (2012) SNObase, a database for S-nitrosation modification. Protein Cell 3:929–33. https://doi.org/10.1007/s13238-012-2094-6

Zhang Y, Liu J, Jia C, et al (2010) Systematic identification and evolutionary features of rhesus monkey small nucleolar RNAs. BMC Genomics 11:61. https://doi.org/10.1186/1471-2164-11-61

Zhang Y, Wang J, Huang S, et al (2009) Systematic identification and characterization of chicken (Gallus gallus) ncRNAs. Nucleic Acids Res 37:6562–6574. https://doi.org/10.1093/nar/gkp704

Zheng D, Zhang J, Ni J, et al (2015) Small nucleolar RNA 78 promotes the tumorigenesis in non-small cell lung cancer. Journal of Experimental & Clinical Cancer Research 34:49. https://doi.org/10.1186/s13046-015-0170-5

Zhou X, Liao W-J, Liao J-M, et al (2015) Ribosomal proteins: functions beyond the ribosome. J Mol Cell Biol 7:92–104. https://doi.org/10.1093/jmcb/mjv014

Paper I

The shift from early to late types of ribosomes in zebrafish development involves changes at a subset of rRNA 2′-*O*-Me sites

SOWMYA RAMACHANDRAN,[1,3] NICOLAI KROGH,[2,3] TOR ERIK JØRGENSEN,[1] STEINAR DAAE JOHANSEN,[1] HENRIK NIELSEN,[1,2] and IGOR BABIAK[1]

[1]Genomics Group, Faculty of Biosciences and Aquaculture, Nord University, 8049 Bodø, Norway
[2]Department of Cellular and Molecular Medicine, University of Copenhagen, 2400 Copenhagen, Denmark

ABSTRACT

During zebrafish development, an early type of rRNA is gradually replaced by a late type that is substantially different in sequence. We applied RiboMeth-seq to rRNA from developmental stages for profiling of 2′-*O*-Me, to learn if changes in methylation pattern were a component of the shift. We compiled a catalog of 2′-*O*-Me sites and cognate box C/D guide RNAs comprising 98 high-confidence sites, including 10 sites that were not known from other vertebrates, one of which was specific to late-type rRNA. We identified a subset of sites that changed in methylation status during development and found that some of these could be explained by availability of their cognate SNORDs. Sites that changed during development were enriched in the novel sites revealed in zebrafish. We propose that the early type of rRNA is a specialized form and that its structure and ribose methylation pattern may be an adaptation to features of development, including translation of specific maternal mRNAs.

Keywords: box C/D snoRNA; zebrafish development; RiboMeth-seq; ribose methylation; ribosomal RNA

INTRODUCTION

Zebrafish (*Danio rerio)* is a well-established animal model for studying aspects of biology, in particular developmental biology. This is due to optical transparency of zebrafish embryos that facilitate easy and reliable identification of developmental stages, short generation time (~3–4 mo), and year-round spawning (Dooley and Zon 2000; Segner 2009). Many transcriptomics and epitranscriptomics studies favor zebrafish as a model since genomic tools and whole-genome data are readily accessible. Zebrafish belongs to Teleostei infraclass, which evolved ~340 million years ago from a common ancestor (Amores et al. 2011). In Teleostei, the genome has undergone two rounds of whole-genome duplication (Meyer and Schartl 1999; Wolfe 2000). Many human and zebrafish genes are orthologous to each other with ~65% of zebrafish genes containing at least one human orthologue and ~70% of human genes containing at least one zebrafish orthologue. Therefore, studies using zebrafish genetics are widely used to understand development, human diseases, and metabolism (Vilella et al. 2009; Howe et al. 2013).

Synthesis of cellular proteins is conducted by ribosomes that are ribonucleoprotein particles (RNPs). In higher eukaryotes, they are composed of ~80 ribosomal proteins and four species of ribosomal RNA (rRNA). rRNA is the most abundant RNA in terms of mass and has a high nucleotide modification rate (~2%) compared to mRNA. rRNA in eukaryotes is arranged into two subunits: a small subunit (SSU), which functions as the decoding center and facilitates the translocation of the tRNA/mRNA pair through the ribosome and a large subunit (LSU), which contains the peptidyltransferase center responsible for catalyzing peptide bond formation. During ribosome biogenesis, two pre-rRNAs are transcribed—one common to 18S (SSU), 5.8S, and 28S (LSU) rRNA, and one exclusively for 5S rRNA. Ribosome assembly is facilitated by ~200 assembly factors. In addition, it involves a large number of small nucleolar RNPs in which small nucleolar RNAs (snoRNAs) are responsible for guiding enzymes to introduce

[3]These authors contributed equally to this work.
Corresponding authors: hamra@sund.ku.dk, igor.s.babiak@nord.no
Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.076760.120.

nucleotide modifications, with pseudouridine (Ψ) and 2′-*O*-methyl (2′-*O*-Me) as the most abundant (Watkins and Bohnsack 2012). Ribose methylation is introduced by the generic methyltransferase, Fibrillarin, guided to the target RNA by box C/D snoRNAs (SNORDs) through base-pairing with the target. The methyl is introduced at the nucleotide base paired to the fifth nucleotide upstream of box D ("+5") (Cavaille et al. 1996; Kiss-Laszlo et al. 1996). 2′-*O*-Me is believed to be important for ribosome biogenesis and translational fidelity. It provides a potential layer of regulation and thus supports the notion of specialized ribosomes (Gilbert 2011; Xue and Barna 2012; Shi and Barna 2015; Guo 2018; Ferretti and Karbstein 2019). The study of ribosome heterogeneity due to ribose methylation has become feasible by advancements in sequencing-based profiling methods (Krogh and Nielsen 2019) and was recently demonstrated in cancer cell lines (Krogh et al. 2016) and during mouse development (Hebras et al. 2020).

A recent study by Locati et al. (2017) revealed that zebrafish express two different sets of rRNA from clusters at two distinct genomic loci (Fig. 1A, left). The two types of rRNA were referred to as maternal- and somatic-type rRNAs, respectively, and they were transcribed, processed and modified (Fig. 1A, middle) at different levels throughout development (Fig. 1A, right). In the absence of direct evidence that the somatic type arises exclusively from somatic cell lineages (germ cell lineage has not been investigated) and the lack of confirmation that the maternal type is maternal-specific (the absence of de novo zygotic transcription has not been demonstrated), we have chosen to name these subtypes of rRNA in the context of developmental advancement: early- and late-rRNA, respectively. The primary sequences of the two subtypes are considerably different. Early-SSU rRNA is 1939 nt whereas the late form is 1889 nt, and the calculated similarity using BioEdit based on ClustalW alignment is 91.3%. Early-LSU rRNA is 4270 nt and late-LSU rRNA is 4106 nt, and they are 87.0% similar. These differences imply that the variable segments in rRNA display structural differences between the two subtypes. The separate chromosomal locations of the gene clusters, the structural differences, and their distinct expression patterns during development suggest that the two subtypes of rRNA have unique properties.

Here, we address the extent to which the differences between the two types of rRNA extend to their modification patterns, specifically with respect to 2′-*O*-Me sites. We used RiboMeth-seq (Birkedal et al. 2015), a sequencing-based method for mapping and estimation of modification stoichiometry, applied to zebrafish whole-cell RNA purified across selected developmental stages from unfertilized eggs to an adult sample. The method also provided low-coverage information on the expression of SNORDs that we supplemented by dedicated small RNA-seq of selected developmental stages. We identified 98 high-confidence 2′-*O*-Me sites, one of which was specific to late-rRNA. Several sites showed hypomethylation in early developmental stages in contrast to the adult stage, where most sites appeared close to fully methylated. Strikingly, sites that were specific to zebrafish compared to human were overrepresented among these hypomethylated sites. Structural and phylogenetic considerations suggest that early-rRNA is more divergent than late-rRNA, and it thus appears that rRNA sequence and structure, as well as ribose methylations are components that contribute to ribosome heterogeneity that may underlie specialization of ribosomes in eggs and early stages compared to adult zebrafish.

RESULTS

Zebrafish rRNA has 98 high-confidence 2′-*O*-Me sites, 10 of which are novel compared to human, including a late-rRNA specific site

To map the 2′-*O*-Me landscape in zebrafish early- and late-rRNA, RiboMeth-seq was applied to four developmental stages: unfertilized eggs, the 32-cell, 12-somite, and protruding-mouth, as well as tail from 1-yr old (adult) fish (Fig. 1B). Initially, the reads obtained from RiboMeth-seq were used for scoring the percentage of early- and late-rRNA in individual samples by mapping reads to the sequences of both types of rRNA in parallel and using a set of validated single-nucleotide differences to calculate the relative expression of the two. To minimize bias in the estimation, we chose four isolated (>30 nt distance to other variants) positions (T24, A55, C1633, and C1728) in SSU and seven positions (C1218, G1351, T1523, C2096, A2201, C3196, and C3780) in LSU (early-rRNA numbering). The single-nucleotide polymorphism (SNP) analyses of the two parallel mappings yielded complementary results at each selected position, that is, the sum of the matched and the sum of the expected mismatched reads were essentially similar. The analysis revealed similar results for SSU and LSU, with almost exclusive expression of early-rRNA in unfertilized egg and 32-cell stage, ∼20% of late-rRNA in the 12-somite stage, ∼75% in the protruding-mouth stage, and almost exclusive expression of late-rRNA in adult tail (Fig. 1C). These results are consistent with previous observations (Locati et al. 2017).

Next, we selected the reference sequences of either early- or late-rRNA based on the main subtype in each sample (Fig. 1C). RiboMeth-seq scores (RMS-scores) were calculated based on early-rRNA in unfertilized egg, 32-cell, and 12-somite stages, and on late-rRNA in the protruding-mouth stage and adult tail. Importantly, the majority (87/98) of 2′-*O*-Me sites was in regions of the rRNA conserved between early- and late-rRNA in the scoring interval of six nucleotides on either side of the queried position and mapping the reads to either of the sequences turned out not to change the calculated RMS-scores significantly.

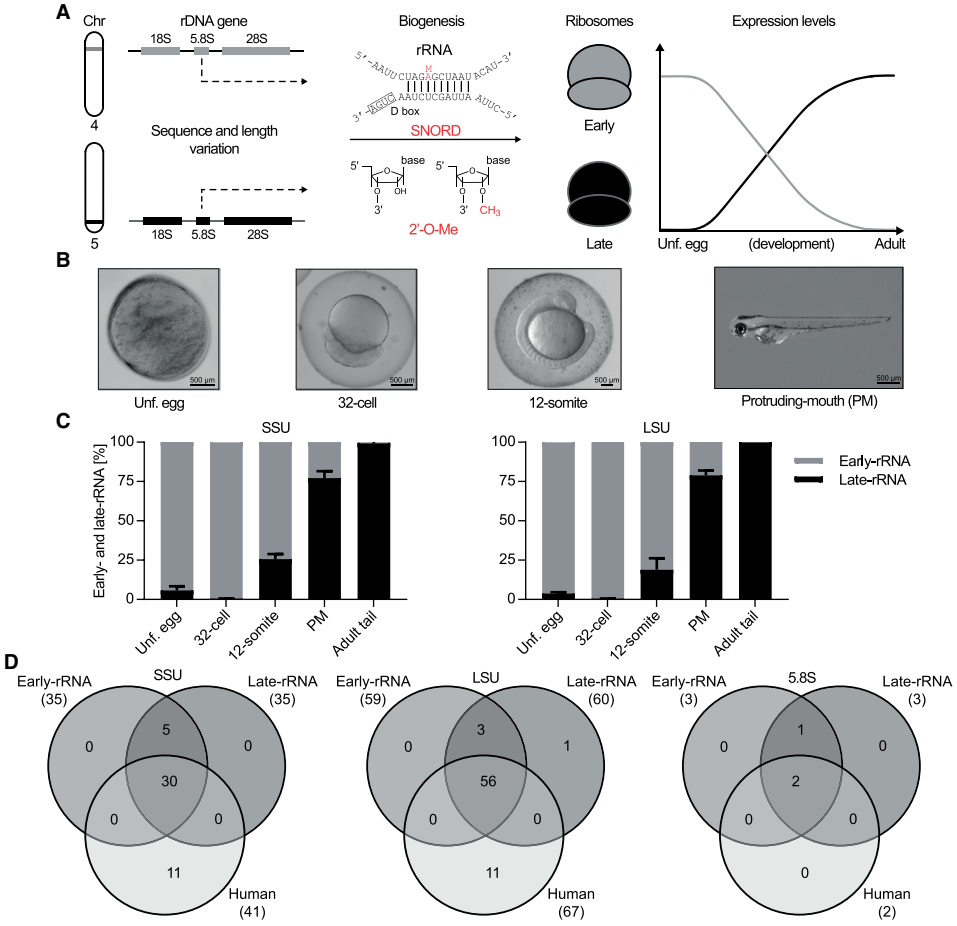


FIGURE 1. Expression and modification by ribose methylation of early- and late-rRNA types of rRNA in zebrafish. (*A*) Schematic illustration of genomic localization, 2′-*O*-Me modification, and the prevalence of expressed early- and late-rRNA. The rRNA biogenesis symbolized by the arrow involves multiple steps and several types of modifications, but only ribose methylation, relevant for the present study, is illustrated. Here, an example of a canonical interaction between a box C/D guide RNA (SNORD) and its rRNA target is shown with indication of the methyl group introduced in red. (*B*) Bright field microscope images of the developmental stages analyzed in the present study: (Unf. Egg) unfertilized egg, the 32-cell stage, the 12-somite stage, and (PM) protruding-mouth stage. (*C*) Estimated relative expression (±SEM) of early- and late-SSU rRNA (*left*) and LSU rRNA (*right*) based on SNP analysis of reads from RiboMeth-seq. The result is based on the analysis of four diagnostic nucleotides in SSU and seven diagnostic nucleotides in LSU, respectively. $n = 3$. (*D*) Venn-diagrams depicting the conservation of ribose-methylated sites in SSU, LSU, and 5.8S between human rRNA and zebrafish early- and late-rRNA subtypes, respectively.

We defined high-confidence 2′-*O*-Me sites as sites with RMS-score >0.75, additionally confirmed with presence of primer extension stop signal at limiting dNTP concentration and/or presence of a plausibly assigned SNORD. In total, 104 2′-*O*-Me sites had RMS-score >0.75. A subset of 24 sites including those with relatively low or variable scores (within or between stages) was then subjected to the primer extension method (Supplemental Fig. S1). 20 sites were confirmed (Table 1) and four sites were excluded due to lack of primer extension signal (Supplemental Table S1). Another two of these sites were excluded because we were unable to assign a plausible SNORD (Supplemental Table S1). The resulting number of high-confidence sites total 98 with 35 located in SSU, 60 in LSU, and three in 5.8S (Fig. 1D; Supplemental Table S2). Strikingly, all sites were found in both early- and late-rRNA, except LSU-

TABLE 1. Summary of methylated sites supported by primer extension analysis

rRNA	Human	Early	Late	Cons.	Px #[a]
18S	U286	U333	U296	Z	Px 1.1/1.2
18S	U287	U334	U297	Z	Px 1.1/1.2
18S	C346	C392	C356	Z	Px 1.2
18S	U354	U400	U364	X	Px 1.2
18S	A1031	A1095	A1052	X	Px 4.1
18S	C1272	C1339	C1294	X	Px 19.1
18S	U1288	U1355	U1310	X	Px 19.1
18S	A1637	A1702	A1656	Z	Px 6.1
28S	C2791	C2239	C2223	X	Px 9.1
28S	A2802	A2250	A2234	X	Px 9.1
28S	G2817	G2265	G2249	Z	Px 9.1
28S	C3866	C3147	C3035	X	Px 10.1
28S	G3878	G3159	G3047	X	Px 10.1
28S	U3904	U3185	U3073	X	Px 10.1
28S	C3916	C3197	C3085	Z	Px 10.1
28S	G3923	G3204	G3092	X	Px 10.1
28S	C4032	C3273	C3167	X	Px 11.1
28S	G4198	G3496	G3346	X	Px 12.1
28S	U4272	U3570	U3420	Z	Px 12.1
28S	U4276	U3574	U3424	X	Px 12.1

(Cons.) Conservation. (X) Methylated sites conserved between zebrafish and human rRNA. (Z) Novel methylation identified in zebrafish.
[a]Primer extension oligos are listed in Supplemental Table S5.

C3916, which was late-rRNA specific. This implies that the mechanism for instalment of ribose methylation functions equally well when rRNA is expressed in the context of amplified, extrachromosomal rDNA (early-rRNA) and from chromosomally integrated rDNA copies (Locati et al. 2017). Our analysis does not rule out the existence of additional sites in specialized tissues or sites with low methylation stoichiometry.

To allow easy comparison with human rRNA, we annotated zebrafish 2′-*O*-Me sites according to their corresponding human nucleotide counterparts using the reference sequence from the snoRNA-LBME-db (Lestrade and Weber 2006; for zebrafish nucleotide numbering, see Supplemental Tables S1–S3). In comparison with human, 10/98 sites were only found in zebrafish (five in SSU, one in 5.8S, and four in LSU [Fig. 1D; Supplemental Table S3). Conversely, 21 sites in human rRNA were not detected in zebrafish (Fig. 1D). The absence of methylation at these sites in zebrafish is consistent with sequence divergence at target sites (12/21) and absence of cognate SNORDs (17/21) as evidenced by our small RNA-seq and consistent with the small nucleolar RNA orthogonal database (snoopy; Yoshihama et al. 2013). For the remaining four sites (SSU-A159, SSU-C174, LSU-A2388, and LSU-C2811), the corresponding SNORDs were likely conserved due to a functional second antisense element.

The methylation stoichiometry at a subset of sites change during development

The RiboMeth-seq method has proven to yield estimates of methylation stoichiometry that are consistent with mass spectrometry and RP-HPLC (Krogh and Nielsen 2019). Applications of RiboMeth-seq previously revealed that cultured cells have a high proportion of sites (~1/3) that are fractionally methylated (Krogh et al. 2016; Erales et al. 2017) in contrast to cells in differentiated tissues that are fully or close to fully methylated at almost all rRNA sites. Interestingly, a recent analysis in mice showed that several sites (~25/108) were increasingly methylated during development (Hebras et al. 2020). Figure 2A shows the methylation profile across zebrafish developmental stages. In adult tail samples, the vast majority (76/98) of methylated sites were fully or close to fully methylated as defined by an RMS-score >0.90 with 16 sites scoring in the 0.80–0.90 range and only six sites with a score <0.80 (Fig. 2A; Supplemental Data 1). Samples from embryonic developmental stages showed significantly lower degrees of methylation at 12 sites (Fig. 2A,B). By comparing these to RiboMeth-seq of mouse tissues (Hebras et al. 2020), 4/12 sites (SSU-Um354, SSU-Gm436, 5.8S-Um14, and LSU-Gm3923) behaved similarly between the two species by showing an increasing level of methylation during development. Interestingly, zebrafish novel sites were overrepresented among the remainder of the varying sites (5/8 sites). Hence, half of the zebrafish novel sites displayed varying methylation levels during development. It should be noted that RiboMeth-seq profiling has high background levels and thus may underestimate effects in the low-scoring range.

Unsupervised hierarchal clustering analysis of the RMS-scores across all samples showed the 12-somite and the 32-cell stages to group closest to unfertilized egg, whereas the protruding-mouth stage grouped together with the adult tail sample (Supplemental Fig. S2). The clustering was mainly driven by two sets of sites changing the most: SSU-U354, -C1272, LSU-C3916, and -A4560 for the first group, and SSU-C346, -G436, LSU-G2817, -G3923, -U4272, 5.8S-U8, and -U14, for the second group (Supplemental Fig. S2). Thus, it appeared that the methylation pattern at a subset of sites followed the developmental timeline and the transition from high levels of early-rRNA during the initial onset of development to high levels of late-rRNA later in development and in adult fish.

Structural comparison suggests that rRNA predominantly expressed early in development is derived

The core structure of rRNA is highly conserved. However, rRNA from different species shows considerable differences in more peripheral structural elements, in particular the

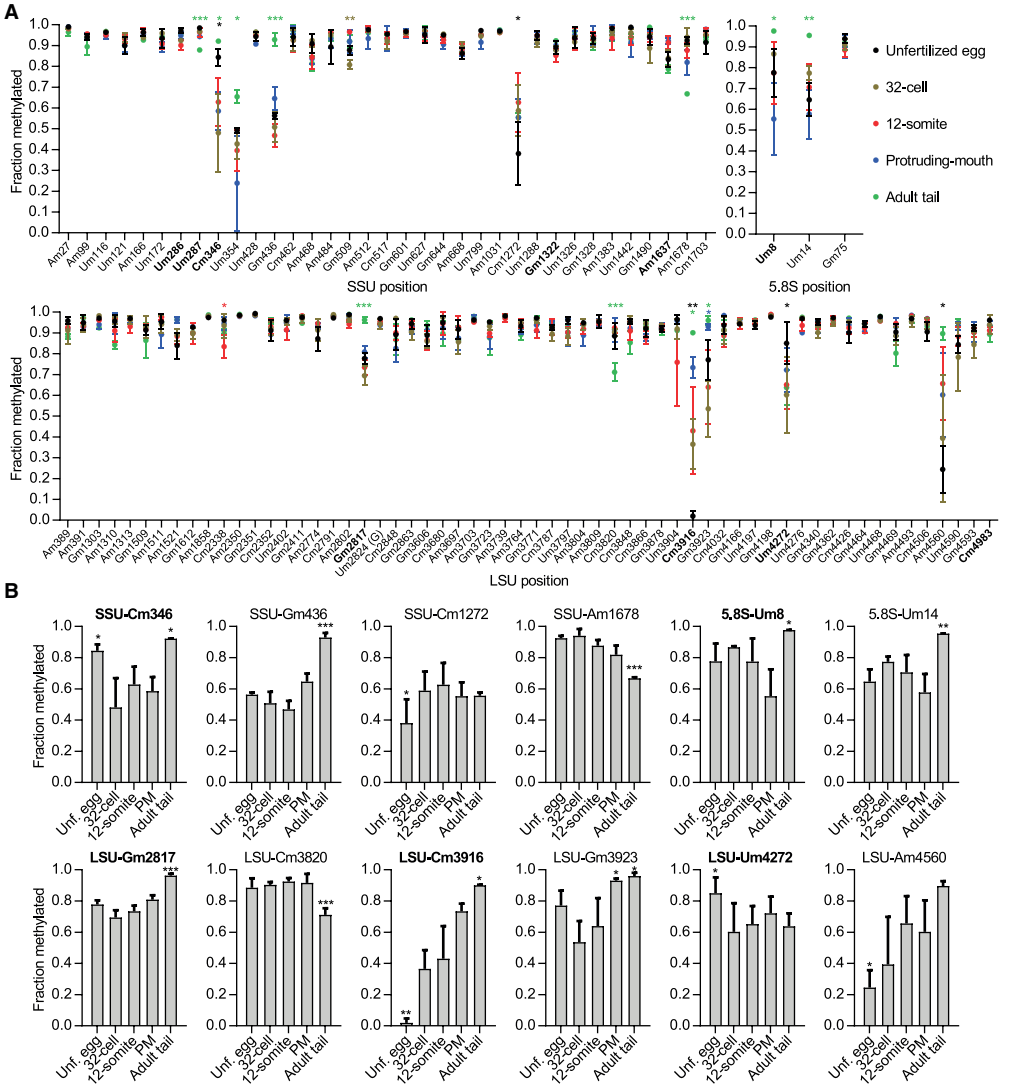


FIGURE 2. Ribosome-wide profiling of 2′-*O*-Me in zebrafish developmental stages and adult fish. (*A*) Graph depicting fraction methylated at all methylated sites in five developmental stages. Sites (*x*-axis) are numbered according to human rRNA to allow comparisons and novel sites found in this study are highlighted in bold. The corresponding nucleotide positions for zebrafish early- and late-rRNA can be found in Supplemental Table S2. Data from different stages are indicated by color and asterisks indicate the levels of statistical significance when comparing the sample in question to all other samples. Error bars indicate the standard deviation, $n = 3$. (*B*) Examples of differentially methylated sites between the five stages analyzed. Zebrafish novel methylations are highlighted in bold. Asterisks and error bars were used in the same way as in *A*, although coloring was omitted.

expansion segments (Gerbi 1996; Ramesh and Woolford 2016). At the sequence level, early-rRNA was found to be 86.5% (SSU) and 69.9% (LSU) similar to human, whereas late-rRNA was slightly more similar, with 90.0% (SSU) and 70.6% (LSU). For a detailed comparison of human and zebrafish rRNA, see (Locati et al. 2017). Because one of

the functions that have been ascribed to SNORDs is to assist in rRNA folding, we set out to specifically compare the secondary structures of zebrafish and human rRNA and relate this to 2′-*O*-Me sites. Structures were drawn based on human rRNA from the RiboVision database (Bernier et al. 2014) aided by sequence alignments (Fig. 3; Supplemental Data 2, 3). Compared to human rRNA, the zebrafish base-pairing scheme differed the most in expansion segments, typically with human rRNA having the longest extensions, as expected (Fig. 3A; Supplemental Figs. S3–S6). As an example, zebrafish LSU expansion segment 7 in H25 lacked ~400 nt, primarily with shorter helixes 25ES7a, "b", "d", "e" and entirely lacking "f", "g", and "h" (Fig. 3A). In contrast to regions that lacked 2′-*O*-Me sites, other structural

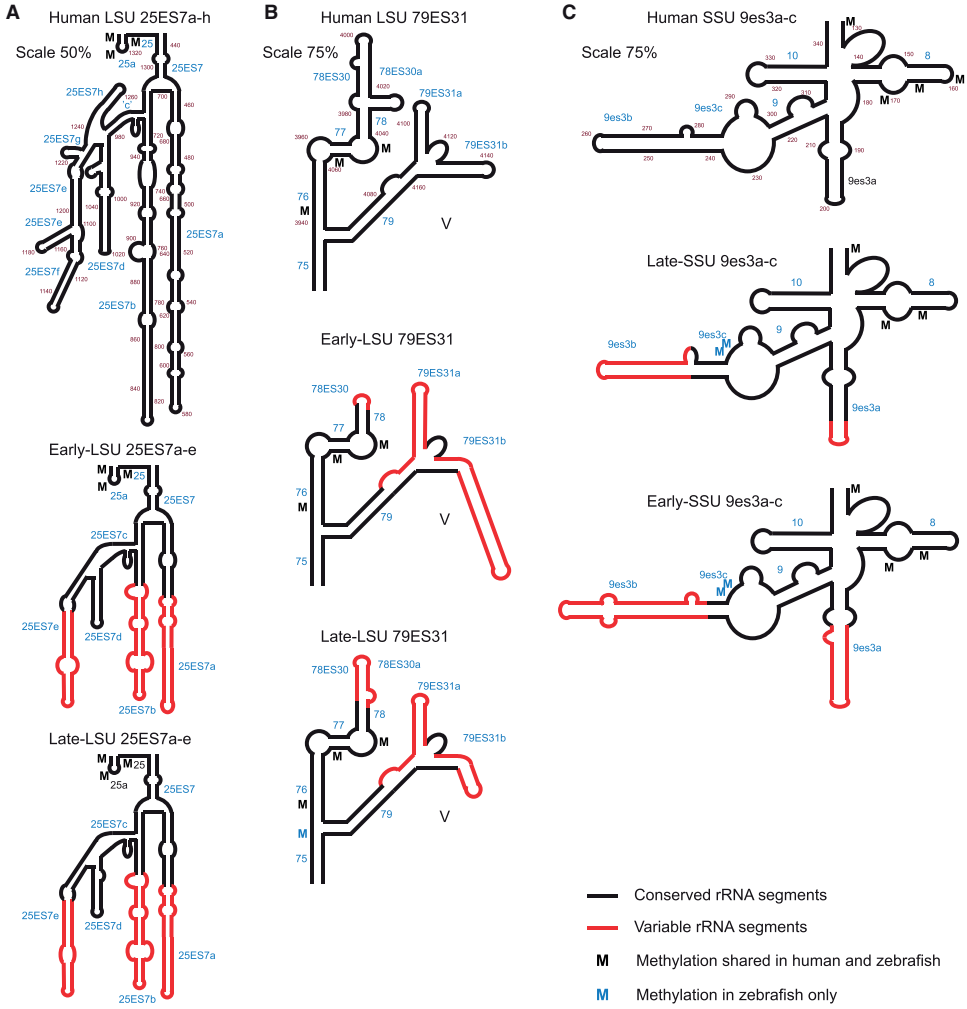


FIGURE 3. Structural comparison of zebrafish early- and late-rRNA with human rRNA. The scaling compared to Supplemental Figures S3–S5 is indicated. (*A*) Helix 25 in LSU and expansion segment 25ES7a–h showing the deletion of 25ESf, "g," and "h" and shortening of 25ESa, "b," "d," and "e." (*B*) Helixes 75-79 and the two expansion segments 78ES30 and 79ES31 in Domain V in LSU with deletion and reduction of 78ES30 in early- and late-LSU, respectively, and elongation of 79ES31b in early LSU. (*C*) Part of the 5′ domain in SSU focusing on helix 9es3-ac. Segments of high (black) and low (red) degree of conservation, helix numbering (blue), conserved (M; black), and novel zebrafish 2′-*O*-Me sites (M; blue) are indicated.

elements where early- and late-rRNA structures differed had strikingly placed zebrafish-specific 2′-*O*-Me sites. In the structure comprising H75-79, H78ES30 was shorter and H79ES31b was considerably longer in early-rRNA than the human counterpart, whereas these structural elements were much more similar to human in late-rRNA (Fig. 3B). The late-rRNA specific methylation LSU-C3916 was found at the three-way junction that organizes these helices (Supplemental Fig. S6) suggesting a role in the folding or conformational flexibility of late-rRNA in zebrafish. In early-SSU rRNA, the "a" part of the expansion segment in h9 (9es3a) was extended by a 28 nt long GC-rich sequence with ~14 nt on either side of the loop compared to the human counterpart (Fig. 3C; Supplemental Figs. S3, S4). In contrast, this helix was only extended by three G–C pairs in late-SSU rRNA immediately adjacent to the loop and thus almost identical to the human structure. Furthermore, the "b" and "c" parts of 9es3b-c in early-SSU rRNA had an insertion of a 14 nt G-rich long sequence compared to the human counterpart, and was rearranged so the C-rich 5′ part of the helix was swapped with the G-rich 3′ part. In late-SSU rRNA, the insertion of the G-rich sequence was only 6 nt-long and the rearrangement seemed less dramatic (Fig. 3C; Supplemental Figs. S3, S4). Zebrafish methylation sites SSU-U286 and -U287 were found in an internal loop at the base of 9es3b-c suggesting their involvement in folding of this expansion segment, to avoid interference with the folding of the conserved core. Altogether, the sequence and structure comparison revealed that late-rRNA resembled human rRNA the most.

Assigned SNORDs conform to the SNORD-target RNA interaction rules

Two sets of sequencing data were used to study the SNORD guides, a low-coverage data set that came along with the RiboMeth-seq analysis, and a dedicated small RNA-seq data set for which we used RNA isolated from early developmental stages. SNORDs were identified by running the FASTA file generated from the sequencing through snoScan (Schattner et al. 2005) or by manual searching for antisense elements matching a known methylation site. A plausible SNORD conforming to the SNORD-target RNA interaction rules could be assigned to 93/96 sites (Supplemental Table S4; Supplemental Data 4), not considering the two methylations in the A-loop of LSU rRNA known to be installed by a stand-alone methyltransferase (Lapeyre and Purushothaman 2004) and allowing +5 and +6 double modifications at three sites (SSU-U287, LSU-G2351, and LSU-G4198). Thus, there are currently three novel methylation sites (SSU-A1322, SSU-A1637, and LSU-G2817) lacking an assigned SNORD (Supplemental Table S4). This set of SNORDs constitutes a minimal set of expressed SNORDs that could account for the observed methylations.

The ribose methylation system in zebrafish appeared to follow strict rules sufficient to explain the experimental observations. First, the majority of SNORDs guided only a single methylation with only three SNORDs using antisense elements associated with both box D and box D′ (Fig. 4A) and one SNORD (SNORD30) guiding two sites (SSU-A1383 and LSU-A3804). Second, the rules for base-pairing interaction between the antisense element in the SNORD and the target sequence in rRNA appeared to be very strict with a minimum of nine base-pairings from position 2–10 upstream of box D (or D′) (Fig. 4B). The base-pairings were almost exclusively Watson–Crick, with few G–U pairs and rarely tolerated mismatches at the second position upstream of box D (Supplemental Fig. S7). Third, the box C and box D of almost all SNORDs complied with the consensus sequences 5′-RUGAUGA and 5′-CUGA, respectively (Fig. 4C, upper panels). Box D′ was less conserved whereas Box C′ was difficult to identify in most SNORDs and appeared almost completely degenerated (Fig. 4C, lower panels). Thus, zebrafish SNORDs and SNORD-target interactions follow the rules of other vertebrate species, such as human and mouse.

Based on the established rules for methylation guiding, we propose that the methylation of LSU-2824, which in human is Um and in zebrafish Gm (Fig. 4D), is a remarkable example of coevolution of SNORD34 and the rRNA target involving three compensatory base pair changes in order to preserve the methylation. LSU-2824 is located in an otherwise highly conserved region of the ribosome with three additional conserved methylations, one in each of H47 and H60, and two in H61 (Fig. 4D, right). In human, there is an additional methylation 13 nucleotides upstream of LSU-2824 (LSU-C2811). In zebrafish, a G–U pair close to the base of H61 differs from a G–C pair in human (Fig. 4D, right) and the G is methylated. It would be of interest to see if these two species-specific modifications play similar roles in organizing the neighboring four-way junction. Only two other instances of compensatory base pair changes in the SNORD:target interaction were found (Supplemental Fig. S8). Another observation that is plausibly explained by applying the consensus rules is the methylation at LSU-C3916 that is exclusively found in late-rRNA (Fig. 2B). Here, we hypothesize that a C rather than an A immediately 5′ of the methylated C at position +5 prevents the formation of consensus base-pairing with SNORD202 in early-rRNA and thus the introduction of the 2′-*O*-Me (Fig. 4E).

SNORD expression of a subset of fractionally methylated sites is delayed and correlates with increased methylation in the course of development

We used the low-coverage RNA-seq from RiboMeth-seq to estimate SNORD levels and correlate with 2′-*O*-Me stoichiometry. In our hands, these estimates are consistent with RT-qPCR and Northern blot analysis, probably

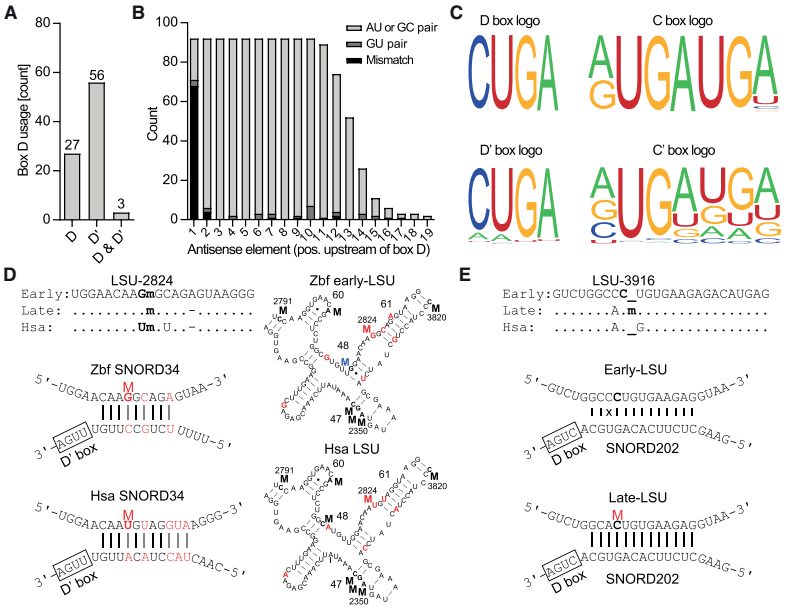


FIGURE 4. Analysis of SNORDs identified in this study. (*A*) Distribution of SNORDs guiding methylations in rRNA using box D, box D′, or both. (*B*) Base pair interaction between the antisense element in the SNORD and the rRNA target in relation to the distance from box D (or D′). (*C*) Sequence logo of the box D, D′, C, and C′ of SNORDs guiding methylations in rRNA generated by WebLogo (Crooks et al. 2004). (*D*, *upper left*) sequence comparison of human and zebrafish early- and late-rRNA containing a conserved methylation at LSU-2824 with indications of identical (.), deleted (-), and methylated (m) nucleotides. (*Lower left*) Human and zebrafish SNORD34 and the interaction with the target sequence surrounding LSU-2824. Nucleotides that differ between the two are labeled in red. (*Right*) 2D structures of a conserved LSU rRNA region between human and zebrafish highlighting methylations in H47, H48, H60, and H61. Methylation at LSU-2824 is in red and the zebrafish-specific methylation in blue. (*E*, *upper*) sequence comparison of human and zebrafish early- and late-rRNA highlighting LSU-Cm3916 only found in late-rRNA. (*Lower*) Drawings of the SNORD base-pairings with early- and late-rRNA, respectively. The presence of a mismatch (x) with early-rRNA may prevent pairing and methylation. (Hsa) *Homo sapiens*, (Zbf) zebrafish.

because they are based on sequencing of alkaline fragments that reduce the bias due to adapter ligation to the mature ends of the SNORDs. The overall SNORD levels in unfertilized eggs and the 32-cell stage were very low compared to the 12-somite and protruding-mouth stages, and the adult tail (Fig. 5A). A likely explanation is that rDNA transcription and thus rRNA biogenesis is shut-down in unfertilized eggs at the time of harvesting and that transcription of late-rRNA similarly is low in the 32-cell stage (Fig. 1C). This provides an interesting opportunity for correlating SNORD production and methylation stoichiometry during development. To this end, we identified seven 2′-*O*-Me sites with increasing methylation during development (methylation pattern from being fractional in the 12-somite and protruding-mouth stage to fully methylated in the adult tail sample) and their cognate SNORDs based on the results in Figure 2A,B; Supplemental Data 1. When the expression levels of these SNORDs were compared to those of the SNORDs guiding the 83 2′-*O*-Me sites that showed no change in methylation, we observed a significantly delayed expression of the seven SNORDs (Fig. 5B). Notably, the expression of these SNORDs was not a simple reflection of host gene expression. Three of the host genes encode additional SNORDs that followed the expression of the main group of SNORDs and guide modifications of sites that are fully methylated throughout development. The uncoupling of SNORD expression from host gene expression can, for example, occur by alternative splicing of the primary transcript and non-sense mediated decay and has previously been documented (Lykke-Andersen et al. 2014; Hebras et al. 2020).

Focused analysis of a host gene reveals an unanticipated multitude of SNORDs

Our SNORD search was focused on discovery of a minimal set that could explain the observed methylation patterns. However, it is known from SNORD inventories in many organisms that several SNORDs are represented by more than one family member and that SNORDs without a

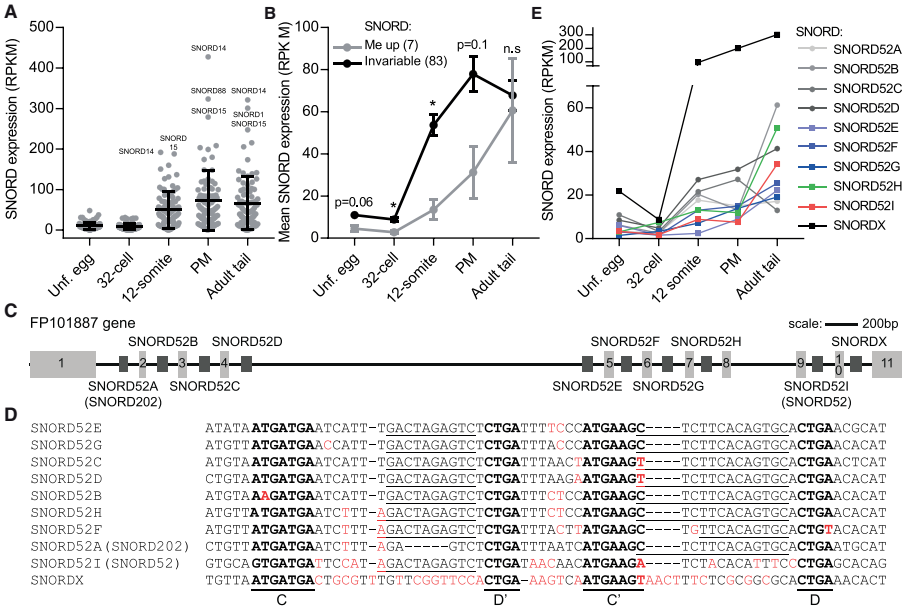


FIGURE 5. SNORD expression levels during zebrafish development. (*A*) Relative expression levels of all SNORDs identified in this study that guides methylation in rRNA. (*B*) Cumulated expression of SNORDs guiding invariable 2′-*O*-Me sites (Invariable) and sites that exhibit increased methylation during development (Me up). This group comprised SSU-C346, -U354, -G436, LSU-C3916, -G3923, -A4560, and 5.8S-U14 and the cognate SNORDs, SNORD-200, -90, -100, -202, -111, -119, and -71. Asterisks indicate significant differences between the two groups of methylation sites within a given developmental stage. (*C*) Schematic of the poorly annotated *FP101887* host gene with indication of 11 exons based on comparison to the longest annotated transcript that appear to encode a lncRNA. The introns host 10 SNORDs of which nine are related. Two SNORDs are given double names (see text for explanation). (*D*) Alignment of the SNORDs encoded by the *FP101887* gene. Conserved SNORD elements are in bold and the antisense elements are underlined. Nucleotides deviating from the most abundant nucleotide at any given position in the alignment are labeled in red to highlight the relatedness of the SNORDs. (*E*) Expression data extracted from RiboMeth-seq analysis (low coverage RNA-seq).

known target ("orphans") as well as SNORD pseudogenes exist. Supplemental Data 4 lists additional members of the SNORD families in our minimal data set as revealed by a naïve BLAST analysis. Because methylation of LSU-C3916 was specific to late-rRNA and guided by an apparently novel SNORD (SNORD202), we conducted a thorough analysis of the host gene of this SNORD, *FP101887*. The annotated transcripts from the gene do not appear to encode proteins, but on the other hand, the transcripts were not detected in a survey of zebrafish lncRNA (Ulitsky et al. 2011). In addition to the SNORD202 sequence used in our minimal set, the gene has two annotated SNORD202 family members that diverge slightly in sequence, but all appear to be bona fide guide RNAs. By aligning the gene sequence and the longest of seven annotated transcripts, we derived the exon–intron structure (Fig. 5C; Supplemental Fig. S9) and searched for SNORD motifs in the intronic regions. Surprisingly, this revealed ten SNORDs, of which nine are related in sequence (Fig. 5D). Seven of these have antisense elements consistent with guiding methylations at LSU-U3904 (upstream of the D′ box) and LSU-C3916 (upstream of the D box). One has the antisense element for LSU-C3916 only, and, conversely, one has the antisense element for LSU-U3904 only. These are the two family members that we detected for our minimal set with the latter corresponding to SNORD52. All of the family members appear to be expressed at low to medium levels and all increased their expression during development as evidenced by our low-coverage RNA-seq analysis (Fig. 5E). The expression levels of individual members appeared related to the stability of the short terminal stem (Supplemental Fig. S10) formed by sequences flanking boxes C and D that pair to form a kink-turn stabilized by binding of the zebrafish homolog of the Snu13/15.5k_NHPX/L7Ae protein (Watkins and Bohnsack 2012). The antisense element associated with box D and required for guiding methylation of LSU-C3916 was not found in any of the SNORD52 homologs

aligned in an evolutionary analysis of 29 vertebrate SNORDs (Kehr et al. 2014). Thus, the 3′ half of the seven double-guiding SNORDs in *FP101887* and SNORD202 are formally a novel SNORD family. However, to maintain transparency in the literature, we label all of these SNORDs as SNORD52 family members. The SNORD52-unrelated SNORD encoded by the last annotated intron in *FP101887* (labeled SNORDX in Fig. 5C) is among the most highly expressed SNORDs in zebrafish. It has at least three plausible targets in rRNA, but none of these were found to be methylated. Thus, it was excluded from our minimal set and awaits further analysis.

DISCUSSION

2′-*O*-Me sites and assigned SNORDs in the zebrafish

We applied RiboMeth-seq to whole-cell RNA from zebrafish and uncovered 98 high-confidence sites defined as methylated sites that were further supported by primer extension analysis (20 sites) and/or assignment of a SNORD guide (or an enzyme in the case of the two A-loop sites) conforming to consensus (95 sites). Apart from one site that was specific to late-rRNA, all sites were methylated in both early- and late-rRNA, albeit to varying degrees for a subset of sites. Ten of the sites were not found in human and mouse (Krogh et al. 2016; Hebras et al. 2020).

We identified 86 SNORDs responsible for 93 of the 96 RNA-guided ribose methylations. Together, these SNORDs constitute a minimal set that was sufficient to explain the observed methylation patterns in RiboMeth-seq experiments. 41 of the SNORDs were also listed in the small nucleolar RNA Orthogonal database (Yoshihama et al. 2013), and only 12 were annotated in the ZFIN database (www.zfin.org). Two of the SNORDs in our set were described separately (Makarova and Kramerov 2009). Altogether, we have updated the information on zebrafish 2′-*O*-Me sites in rRNA and cognate SNORDs to the level of other major model organisms. All of the SNORDs in the minimal set were encoded in intronic regions within host genes (Supplemental Data 4). The base-pairing interactions between the SNORDs and their targets almost all (89/98) conformed to strict consensus rules. The main exception was three cases of double modifications at positions +5 and +6, a phenomenon that has also been observed in several other systems. As in human (Krogh et al. 2016) and mouse (Hebras et al. 2020), but different from yeast (Birkedal et al. 2015) and many species of archaea (Dennis et al. 2015), 83/86 SNORDs appeared to guide a single methylation, predominantly using the box D′-associated antisense element (two-thirds of the cases). Five of the SNORDs appeared novel and were numbered SNORD200-204 (high numbers used in order not to interfere with human numbering). These SNORDs returned no hits when queried in the Rfam database (Kalvari et al. 2018) and we were unable to detect related SNORDs by BLAST searches with one exception. SNORD201 is related to human orphan SNORD101 in the human snoRNA atlas (Jorjani et al. 2016). The sequence similarity does not include the antisense element and thus, we maintain the SNORD201 numbering emphasizing the functional SNORD of the two until a more comprehensive phylogenetic analysis can form the basis for defining the family to which these SNORDs belong.

SNORD inventories from several organisms have revealed that many SNORDs are represented by several genomic copies that may differ in sequence and that many SNORDs without known targets ("orphans") exist. In Supplemental data 4 we have listed 33 additional SNORDs belonging to 16 families (one to five additional copies) revealed by BLASTN analysis of the zebrafish genome. The additional copies were typically found in the same host gene as the SNORD from the minimal set. However, this is clearly not the full story as revealed by our in-depth analysis of the *FP101887* host gene. Here we found a total of 10 SNORDs, seven more than annotated for the gene. Nine of these belong to the same family, with seven double guides and two in which the antisense element associated with either the D box (SNORD52) or D′ box (SNORD202) was degenerate. One possibility is that the double-guiding SNORDs gave rise to two single-guiding SNORDs following an evolutionary trend toward single-guiding SNORDs. For transparency in the literature, we assigned all of these nine SNORDs to the SNORD52 family, using double names SNORD52A/SNORD202 and SNORD52I/SNORD52 in this paper to highlight the novelty of the box D-associated antisense element guiding methylation at LSU-C3916. Clearly, a full inventory of zebrafish must await a dedicated analysis, as is the case for most other model organisms.

Early-rRNA is structurally different from late-rRNA and may have adopted functions specific to early development

Genome duplication is an important source of paralogues in evolution. However, there are only a few examples that comprise rRNA genes, which is surprising considering that sequence variants of rRNA genes have been suggested to be of functional importance (Kurylo et al. 2018; Parks et al. 2018). The few well-described examples comprise rRNA genes in the malaria parasite, *Plasmodium falciparum* (Gunderson et al. 1987; Mercereau-Puijalon et al. 2002), and the parasitic trypanosomatids (Liu et al. 2016; Shalev-Benami et al. 2017; Rajan et al. 2020). In the latter organisms, several studies have addressed variable ribose methylation in rRNA of the blood stream form of the parasite propagating in humans and the procyclic form propagating in the insect host (Liu et al. 2016; Shalev-Benami et al. 2017; Rajan et al. 2020). In the case of

Trypanosoma brucei, the 2′-*O*-Me sites were recently mapped by several sequencing-based approaches, including RiboMeth-seq, and it was argued that differences in methylation could be part of adaptation to differences in nutrients and temperature in the two hosts (Rajan et al. 2020). The teleost-specific whole-genome duplications may have given rise to the two genomic rDNA clusters described in zebrafish (Locati et al. 2017). Sequence and secondary structure comparisons revealed that late-rRNA was more similar to rRNA from other vertebrates than the early-rRNA subtype (Fig. 3), suggesting that the cluster encoding early-rRNA was more diverged, and likely acquired new properties. It was recently reported that snoRNAs U3, U8, and snoZ30 also exist in early- and late-expressed versions (Pagano et al. 2020). Finally, it is possible that the genome duplication also gave rise to new SNORDs, but due to the low sequence constraints of SNORDs it was not possible to establish if the zebrafish-specific SNORDs are derived from more ancestral vertebrate SNORDs by duplication and sequence drift.

Temporal concurrence of the early-rRNA subtype and maternal mRNA, as well as their similar rate of decline in the course of zebrafish development, suggests a functional link. Maternal mRNA, stockpiled in an oocyte and utilized for translation well beyond the time point of zygotic genome activation, has unique features ensuring its stability and mechanisms of activation and degradation. These include short or no poly(A) tail, and the subsequent cytoplasmic polyadenylation regulates translational maturation and efficiency, transcript stability, and controls maternal-to-zygotic transition (Vasudevan et al. 2006; Subtelny et al. 2014; Winata and Korzh 2018; Winata et al. 2018; Vastenhouw et al. 2019). Also, maternal mRNA is enriched in epitranscriptomics modifications, such as N^6-methyladenosine (m^6A), which functions in destabilization (clearance), but also affect the translational efficiency (Huang et al. 2017; Aanes et al. 2020), and 5-methylcytosine (m^5C), which enhances the stability of maternal mRNA (Yang et al. 2019). Finally, maternal mRNA contains translation-dependent information, in particular codon usage that influences translation efficiency and transcript decay (Bazzini et al. 2016; Mishima and Tomari 2016). Together, the features of maternal mRNA support the possibility of specialized translational machinery evolved in parallel. Ribosomal heterogeneity is well-established and has been associated with differential transcript preferences and translation efficiencies, although hard evidence for functionally specialized ribosomes is scarce (Genuth and Barna 2018; Guo 2018; Ferretti and Karbstein 2019).

Fractional modification at certain 2′-*O*-Me sites is a feature of development

The RiboMeth-seq analysis revealed that most sites were fully or close to fully methylated in adult tissues, leaving only six sites as fractionally modified (RMS-score <0.80). This is in line with our observations from human and mouse (Hebras et al. 2020) and in contrast to the early analyses of cell cultures in which one-third of sites were fractionally modified (Krogh et al. 2016). Fractional, or substoichiometric methylation, is an important phenomenon because it implies that the ribosome population is heterogeneous. However, fractional modification could be a passive phenomenon without functional consequences. One possibility is that fractional modification is caused by recently acquired SNORDs that have not yet become optimized for targeting. By comparison of fractional sites to a careful phylogenetic analysis of the origin of their cognate SNORDs there is no support in humans (Kehr et al. 2014; Krogh et al. 2016) or yeast (Birkedal et al. 2015; Canzler et al. 2018) for such an explanation. Another trivial explanation is that the methylation stoichiometry reflects SNORD availability. Although this appears to be the case for some SNORDs, there is no global correlation between SNORD levels and methylation stoichiometry in humans (Krogh et al. 2016), mouse (Hebras et al. 2020), or in the present study. Finally, methylation could reflect SNORD access and thus reflect partitioning between different rRNA folding pathways. If this was the case, the variation in methylation in studies like the present, would imply considerable changes to ribosome biogenesis during development. Based on these arguments, we propose that fractional methylation in at least some cases is a nontrivial phenomenon that should be subjected to functional studies.

An important find in the present study was that 8/98 2′-*O*-Me sites were fractionally modified during zebrafish development, but fully modified in the adult fish. Conversely, 2/5 sites that were fractional in adult were close to fully methylated in all investigated developmental stages. Thus, changes to the level of methylation appear to be a feature of zebrafish development similar to what was previously observed in the mouse (Hebras et al. 2020). Strikingly, species-specific sites appear particularly enriched among sites that display developmental variation. A significant fraction (6/10) of the zebrafish-specific sites found in the present study conformed to this notion. Similarly, among the 21 sites in mouse that were not found in zebrafish, 12 displayed methylation differences between developmental stages and adult mouse. These constituted a large fraction of the sites (12/20) that showed variation. Fractional methylation implies heterogeneity at the level of the ribosome and may be related to a regulatory transition, a hallmark of early development. Thus, we suggest that a subset of ribose methylations in rRNA constitute an adaptation to specific features of organismal development in vertebrates. Other studies have noted changes in SNORD levels during development, for example, in the classical developmental model *Dictyostelium discoideum* (Aspegren et al. 2004) and in *Drosophila*

melanogaster (Angrisani et al. 2015). It will be of interest to study if these changes in SNORDs are paralleled by changes in ribosomal RNA methylation patterns and affect development.

The components of box C/D snoRNPs have multiple functions. Fibrillarin is the methyltransferase responsible for RNA-guided ribose methylation (Cavaille et al. 1996; Kiss-Laszlo et al. 1996), but it has additional catalytic functions, including histone methylation (Tessarz et al. 2014; Iyer-Bierhoff et al. 2018), and a key role as a structural component in nucleolar phase transitions (Feric et al. 2016). Proteins that combine with RNAs for targeting, such as Fibrillarin and box C/D RNAs, constitute powerful and highly adaptable systems. Several of the box C/D snoRNPs are involved in ribosome biogenesis without introduction of methylations, for example, U3 snoRNP that is essential for early cleavages of pre-rRNA (Beltrame and Tollervey 1995; Marmier-Gourrier et al. 2011). Those that conform to the methylation paradigm may in fact also be diverse in function. In some cases, the SNORD may carry the main function in chaperoning the folding of rRNA. In other cases, the methylation may carry the main function in stabilizing the ribosome or mediating translation. For this reason, it is important to experimentally establish the methylation pattern in key model organisms and to relate this to the specific biology of the organism. In this study, we have provided the foundation for such work in zebrafish. We suggest that some zebrafish modifications are adaptations to structural alterations in rRNA (Fig. 3) and that others may impact mRNA recruitment and translation during development. These observations open the possibility for functional studies involving manipulations of selected SNORDs using antisense oligonucleotides that interfere with host gene splicing or processing of the SNORD (e.g., using morpholinos; Higa-Nakamine et al. 2012) or by CRISPR–Cas9 KO of SNORD sequences. Both strategies are technically challenging because the SNORDs are located close to splice sites and because the effects of manipulating the SNORD sequences on the often complex splicing of the host gene transcripts are unknown.

MATERIALS AND METHODS

Zebrafish maintenance and sampling

The AB zebrafish line embryos and adults were sourced from the zebrafish facility belonging to Nord University, Norway. The experimental procedures and husbandry were performed in agreement with the Norwegian Regulation on Animal Experimentation (The Norwegian Animal Protection Act, No. 73 of 20 December 1974). This was certified by the National Animal Research Authority, Norway, General License for Fish Maintenance and Breeding no. 17. The maintenance of zebrafish was done using an Aquatic Habitats recirculating system (Pentair) and followed the standard protocol (Westerfield 2000). The fish were fed SDS zebrafish-specific diet (Special Diet Services) according to the manufacturers' instruction.

Sampling was performed in triplicates, and the source fish originated from three different tanks to eliminate tank-specific variations. To collect unfertilized eggs, two sexually mature females were randomly chosen from each experimental tank, sedated using 50 mg/L MS-222 (Tricaine; Sigma Aldrich) buffered with equal parts of sodium bicarbonate ($NaHCO_3$), and the unfertilized eggs were obtained by manual stripping. The eggs were then washed in ice-cold PBS, drained and immediately snap-frozen for later RNA purification (Presslauer et al. 2017). To collect 32-cell, 12-somite and protruding-mouth stages, parental fish were free-spawning in their respective tanks without prior isolating males and females. Embryos were collected within 10 min after the fertilization and further incubated in Petri dish placed in a cell culture incubator at 28.5°C. Their development was visually tracked under light microscopy. The required embryonic stages were identified according to Kimmel et al. (1995), snap-frozen in liquid nitrogen and stored at −80°C for later RNA purification. To collect samples of trunk from adult individuals (here referred to as "tail" sample), single random male and female were chosen from each experimental tank. These individuals were euthanatized with 200 mg/L MS-222 and the sample filets were dissected and snap-frozen for later RNA purification. Images of zebrafish developmental stages were taken with a Zeiss Axio Zoom V16 microscope using Zeiss Zen image analysis software.

Purification of whole-cell RNA

RNA samples (<50 embryos per replicate) were treated with 1 mL Qiazol (Qiagen) and an adequate amount of lysis beads, and homogenized in the Precellys 24 system (Bertin instrumentation) at 5000 RPM for 2 × 20 sec with a 20 sec gap. Unclear lysates were centrifuged at 13,500 RPM for 10 min at 4°C to remove debris. Otherwise, RNA was purified according to the manufacturer. Subsequently, RNA pellets were dried, resuspended in RNase free water, and 1 µL was used to measure quantity (average of >400 ng/µL) and quality (RIN > 9.6) using an Agilent TapeStation 2200 (Agilent). For long-term storage, RNA was repelleted and kept in 70% ethanol at −80°C. Prior to use, the integrity of the RNA was assessed on a denaturing 1% agarose gel.

RiboMeth-seq and SNP analysis

The initial mapping and quantitation of 2′-*O*-Me in rRNA was achieved using RiboMeth-seq on biological triplicates except for the adult tail sample which was conducted in duplicate. RiboMeth-seq was essentially performed as previously described (Birkedal et al. 2015; Krogh et al. 2016). Whole-cell RNA was subjected to partial alkaline degradation, purified on a 10% UPAG. Prior to cDNA synthesis using SuperScript IV RT (ThermoFisher Scientific), adaptor ligation was done using a tRNA ligase. The resulting libraries were sequenced on Ion PI Chips (v3) using an Ion Proton semiconductor sequencer with default analysis parameters except that –trim-adapter-cutoff=0 was added to the "PreBaseCaller for calibration" and "BaseCaller" to include adapters in the FASTQ file generated using the FastqCreator on the Torrent server. Reads were subsequently sorted based on barcodes in the 5′adapter (cDNA) and this adapter was removed

using a Python script. Here, it is important to keep in mind that the analysis is critically dependent on recording of read ends and that the sequence primarily serves to map the read. The generic Ion P1B (3′adapter) was subsequently removed and untrimmed reads or reads shorter than 15 were discarded using cutadapt v2.0. The trimmed reads (median length of ∼34 nt with >95% of reads being between 25 and 45 nt long) were mapped separately to the zebrafish early- or late-rRNA (Locati et al. 2017) and SNORDs (Supplemental Data 5) using Bowtie2 v2.3.4.1 with -k 10. Prior to read-end counting multiple mapped reads were removed allowing only the best mapped reads using a Python script. The "fraction methylated" (RMS-score) was calculated as previously ("score C" in (Birkedal et al. 2015). Subsequently, at a few sites, the RMS-score was manually corrected based on an inherent problem that arises when the commercial RNA oligo used in the first adapter ligation is less than full-length. If the adapter is missing a nucleotide at its 5′ end and ligated to a library RNA fragment with an identical nucleotide at its 3′ end, the barcode is thus intact and the library nucleotide will be removed during data processing. Such errors are easily detected and dealt with by excluding the 3′ end data set from the calculation in the affected replicate and at the problematic site only (see Krogh et al. 2017). Corrected sites are indicated in Supplemental Data 1.

SNP analysis was performed using Samtools v1.3.1. To calculate relative expression of early- and late-rRNAs, isolated SNPs with >30 nt distance to other SNPs were picked based on differences between early- and late-rRNA described in Locati et al. (2017). In total, four positions (T24, A55, C1633, and C1728) in SSU and seven positions (C1218, G1351, T1523, C2096, A2201, C3196, and C3780) in LSU (early-rRNA numbering) were used in the analysis.

Detection of 2′-*O*-Me by primer extension

A subset of 2′-*O*-Me sites with low (RMS < 0.75) or inconsistent scores and sites without a plausible SNORD were further assessed by the high/low dNTP-concentration primer extension method (Maden 2001). Primers were designed based on the predicted 2′-*O*-Me sites and a list of all the primers can be found in Supplemental Table S5. Reverse transcription of 1 µg of whole-cell RNA from appropriate developmental stages were performed in 20 µL 1× RT buffer at 42°C for 60 min supplemented with 1 µL AMV RT (Promega, 20 U) at low and high dNTP concentrations (0.01 mM and 1 mM, respectively). The cDNA generated were separated on an 8% UPAG together with a sequencing ladder. Dried gels were exposed to Phosphor Imager Screens and scanned using a Typhoon Biomolecular Imager (Amersham) to visualize the radioactive signals from the probes. Images were analyzed using Fiji software.

Identification of SNORDs by small RNA-seq

Embryos from different zebrafish developmental stages (32-cell, oblong, and 15% epiboly) were collected as described above, snap-frozen in liquid nitrogen and stored at −80°C. Whole-cell RNA from each of the stages were extracted from ∼30 zebrafish embryos using QIAzol lysis reagent (QIAGEN) as described above. The quality of the RNA was assessed by an Agilent TapeStation 2200 and all samples used for library construction had a RIN value >9. One µg of whole-cell RNA from each stage was subjected to rRNA depletion using the RiboMinus Eukaryote System v2 (Invitrogen), and small RNA libraries were constructed using Ion Total RNA-seq Kit v2 (ThermoFisher Scientific) with minor modifications. In short, the rRNA-depleted RNA samples were enriched for small RNA (<200 nt) using the magnetic bead clean-up module supplied with the kit or the Monarch RNA Cleanup Kit (New England Biolabs), adapters diluted 1:2 were ligated to the RNA for 2 h and the RNA subsequently reverse transcribed using Superscript IV (ThermoFisher Scientific). The cDNA was purified using the magnetic bead clean-up module, without size-selection. Amplification of cDNA and purification were performed according to the manufacturer's instructions. Manual template preparation of libraries was carried out on the Ion OneTouch 2 System (ThermoFisher Scientific) and subsequently sequenced on Ion 540 chips, the Ion GeneStudio S5 System. Reads were automatically trimmed, low quality reads discarded using default settings on the Torrent server, and a FASTQ file generated using the FastqCreator plugin.

SNORD search and rRNA interaction prediction

The FASTQ files from the small RNA-seq of the three stages were converted to FASTA files using a Perl script, merged and subsequently used as the basis of the SNORD search and rRNA interaction prediction. Initially, SNORDs were identified by running the merged FASTA file through snoScan (Schattner et al. 2005) against zebrafish early- and late-rRNA reference sequences (Locati et al. 2017). Only the top-ranked SNORDs conforming to the consensus rules for SNORD-target rRNA interaction were picked for further analyses (Krogh et al. 2016). A few SNORDs were manually identified by searching for the predicted sequence of an antisense element based on the neighboring sequence of the 2′-*O*-Me site in question. Subsequently, SNORDs were assigned to the methylated sites identified by RiboMeth-seq and primer extension analysis and named in accordance with the human SNORDs (HGCN). The generated list of SNORDs with predicted rRNA interactions was aligned against the zebrafish genome (ENSEMBL, GRCz11) using BLASTN, with search sensitivity set to normal (http://www.ensembl.org/Danio_rerio/Tools/Blast) to extract genomic location, host gene information (Supplemental Table S4; Supplemental Data 4, 5), to validate the SNORD sequences, and extract information on potential homologs. By mapping reads from the RiboMeth-seq analyses against the SNORD sequences, the list of SNORDs was subsequently filtered based on expression and thus only expressed SNORDs were considered in this study (Supplemental Data 5).

Statistical analyses

The RMS-score from RiboMeth-seq results are expressed as mean ± SD. SNORD expression data as mean ± SEM. Correlations were analyzed using Spearman's Rank correlation. Statistical analyses were performed using Microsoft Excel software and GraphPad Prism 7. Comparison of two groups was analyzed by Student's unpaired *t*-test (two-tailed) and statistically significant differences between groups are indicated as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. The unsupervised hierarchical cluster analysis and associated heatmap was generated in R using the pheatmap package with

the complete linkage method and otherwise default settings (https://CRAN.R-project.org/package=pheatmap).

DATA DEPOSITION

Sequencing data from RiboMeth-seq and small RNA-seq are deposited at the NCBI Gene Expression Omnibus database and accessible through GSE151797.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We wish to thank Cesilie Røtnes Amundsen and Purushothaman Kathiresan, Nord University, for their assistance in quality assessments and sampling of the embryos. The study was funded by the Research Council of Norway, InnControl project (grant #275786) and The Danish Cancer Society (grant R167-A10943-17-S2; to N.K.).

Received June 12, 2020; accepted September 6, 2020.

REFERENCES

Aanes H, Engelsen D, Manaf A, Alemu EA, Vågbø CB, Martín L, Lerdrup M, Hansen K, Mathavan S, Winata C, et al. 2020. N^6-methyladenosine dynamics during early vertebrate embryogenesis. bioRxiv doi:101101/528422

Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. 2011. Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics* **188:** 799–808. doi:10.1534/genetics.111.127324

Angrisani A, Tafer H, Stadler PF, Furia M. 2015. Developmentally regulated expression and expression strategies of *Drosophila* snoRNAs. *Insect Biochem Mol Biol* **61:** 69–78. doi:10.1016/j.ibmb.2015.01.013

Aspegren A, Hinas A, Larsson P, Larsson A, Soderbom F. 2004. Novel non-coding RNAs in Dictyostelium discoideum and their expression during development. *Nucleic Acids Res* **32:** 4646–4656. doi:10.1093/nar/gkh804

Bazzini AA, Del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, Yao J, Khokha MK, Giraldez AJ. 2016. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J* **35:** 2087–2103. doi:10.15252/embj.201694699

Beltrame M, Tollervey D. 1995. Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J* **14:** 4350–4356. doi:10.1002/j.1460-2075.1995.tb00109.x

Bernier CR, Petrov AS, Waterbury CC, Jett J, Li F, Freil LE, Xiong X, Wang L, Migliozzi BL, Hershkovits E, et al. 2014. RiboVision suite for visualization and analysis of ribosomes. *Faraday Discuss* **169:** 195–207. doi:10.1039/C3FD00126A

Birkedal U, Christensen-Dalsgaard M, Krogh N, Sabarinathan R, Gorodkin J, Nielsen H. 2015. Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew Chem Int Ed Engl* **54:** 451–455. doi:10.1002/anie.201408362

Canzler S, Stadler PF, Schor J. 2018. The fungal snoRNAome. *RNA* **24:** 342–360. doi:10.1261/rna.062778.117

Cavaille J, Nicoloso M, Bachellerie JP. 1996. Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* **383:** 732–735. doi:10.1038/383732a0

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14:** 1188–1190. doi:10.1101/gr.849004

Dennis PP, Tripp V, Lui L, Lowe T, Randau L. 2015. C/D box sRNA-guided 2′-*O*-methylation patterns of archaeal rRNA molecules. *BMC Genomics* **16:** 632. doi:10.1186/s12864-015-1839-z

Dooley K, Zon LI. 2000. Zebrafish: a model system for the study of human disease. *Curr Opin Genet Dev* **10:** 252–256. doi:10.1016/S0959-437X(00)00074-5

Erales J, Marchand V, Panthu B, Gillot S, Belin S, Ghayad SE, Garcia M, Laforets F, Marcel V, Baudin-Baillieu A, et al. 2017. Evidence for rRNA 2′-*O*-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. *Proc Natl Acad Sci* **114:** 12934–12939. doi:10.1073/pnas.1707674114

Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP. 2016. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165:** 1686–1697. doi:10.1016/j.cell.2016.04.047

Ferretti MB, Karbstein K. 2019. Does functional specialization of ribosomes really exist? *RNA* **25:** 521–538. doi:10.1261/rna.069823.118

Genuth NR, Barna M. 2018. The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life. *Mol Cell* **71:** 364–374. doi:10.1016/j.molcel.2018.07.018

Gerbi SA. 1996. Expansion segments: regions of variable size that interrupt the universal secondary structure of ribosomal RNA. In *Ribosomal RNA, structure, evolution, processing, and function in protein biosynthesis* (ed. Zimmermann RA, Dahlberg AE), pp. 71–87. CRC Press, Boca Raton, FL.

Gilbert WV. 2011. Functional specialization of ribosomes? *Trends Biochem Sci* **36:** 127–132. doi:10.1016/j.tibs.2010.12.002

Gunderson JH, Sogin ML, Wollett G, Hollingdale M, de la Cruz VF, Waters AP, McCutchan TF. 1987. Structurally distinct, stage-specific ribosomes occur in *Plasmodium*. *Science* **238:** 933–937. doi:10.1126/science.3672135

Guo H. 2018. Specialized ribosomes and the control of translation. *Biochem Soc Trans* **46:** 855–869. doi:10.1042/BST20160426

Hebras J, Krogh N, Marty V, Nielsen H, Cavaille J. 2020. Developmental changes of rRNA ribose methylations in the mouse. *RNA Biol* **17:** 150–164. doi:10.1080/15476286.2019.1670598

Higa-Nakamine S, Suzuki T, Uechi T, Chakraborty A, Nakajima Y, Nakamura M, Hirano N, Suzuki T, Kenmochi N. 2012. Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res* **40:** 391–398. doi:10.1093/nar/gkr700

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496:** 498–503. doi:10.1038/nature12111

Huang C, Shi J, Guo Y, Huang W, Huang S, Ming S, Wu X, Zhang R, Ding J, Zhao W, et al. 2017. A snoRNA modulates mRNA 3′ end processing and regulates the expression of a subset of mRNAs. *Nucleic Acids Res* **45:** 8647–8660. doi:10.1093/nar/gkx651

Iyer-Bierhoff A, Krogh N, Tessarz P, Ruppert T, Nielsen H, Grummt I. 2018. SIRT7-dependent deacetylation of fibrillarin controls histone H2A methylation and rRNA synthesis during the cell cycle. *Cell Rep* **25:** 2946–2954. doi:10.1016/j.celrep.2018.11.051

Jorjani H, Kehr S, Jedlinski DJ, Gumienny R, Hertel J, Stadler PF, Zavolan M, Gruber AR. 2016. An updated human snoRNAome. *Nucleic Acids Res* **44:** 5068–5082. doi:10.1093/nar/gkw386

Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI. 2018. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res* **46:** D335–D342. doi:10.1093/nar/gkx1038

Kehr S, Bartschat S, Tafer H, Stadler PF, Hertel J. 2014. Matching of Soulmates: coevolution of snoRNAs and their targets. *Mol Biol Evol* **31:** 455–467. doi:10.1093/molbev/mst209

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn* **203:** 253–310. doi:10.1002/aja.1002030302

Kiss-Laszlo Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T. 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85:** 1077–1088. doi:10.1016/S0092-8674(00)81308-2

Krogh N, Nielsen H. 2019. Sequencing-based methods for detection and quantitation of ribose methylations in RNA. *Methods* **156:** 5–15. doi:10.1016/j.ymeth.2018.11.017

Krogh N, Jansson MD, Hafner SJ, Tehler D, Birkedal U, Christensen-Dalsgaard M, Lund AH, Nielsen H. 2016. Profiling of 2′-*O*-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res* **44:** 7884–7895. doi:10.1093/nar/gkw482

Krogh N, Kongsbak-Wismann M, Geisler C, Nielsen H. 2017. Substoichiometric ribose methylations in spliceosomal snRNAs. *Org Biomol Chem* **15:** 8872–8876. doi:10.1039/C7OB02317K

Kurylo CM, Parks MM, Juette MF, Zinshteyn B, Altman RB, Thibado JK, Vincent CT, Blanchard SC. 2018. Endogenous rRNA sequence variation can regulate stress response gene expression and phenotype. *Cell Rep* **25:** 236–248. doi:10.1016/j.celrep.2018.08.093

Lapeyre B, Purushothaman SK. 2004. Spb1p-directed formation of Gm2922 in the ribosome catalytic center occurs at a late processing stage. *Mol Cell* **16:** 663–669. doi:10.1016/j.molcel.2004.10.022

Lestrade L, Weber MJ. 2006. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res* **34:** D158–D162. doi:10.1093/nar/gkj002

Liu Z, Gutierrez-Vargas C, Wei J, Grassucci RA, Ramesh M, Espina N, Sun M, Tutuncuoglu B, Madison-Antenucci S, Woolford JL, et al. 2016. Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit. *Proc Natl Acad Sci* **113:** 12174–12179. doi:10.1073/pnas.1614594113

Locati MD, Pagano JFB, Girard G, Ensink WA, van Olst M, van Leeuwen S, Nehrdich U, Spaink HP, Rauwerda H, Jonker MJ, et al. 2017. Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA* **23:** 1188–1199. doi:10.1261/rna.061515.117

Lykke-Andersen S, Chen Y, Ardal BR, Lilje B, Waage J, Sandelin A, Jensen TH. 2014. Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes Dev* **28:** 2498–2517. doi:10.1101/gad.246538.114

Maden BE. 2001. Mapping 2′-*O*-methyl groups in ribosomal RNA. *Methods* **25:** 374–382.

Makarova JA, Kramerov DA. 2009. Analysis of C/D box snoRNA genes in vertebrates: the number of copies decreases in placental mammals. *Genomics* **94:** 11–19. doi:10.1016/j.ygeno.2009.02.003

Marmier-Gourrier N, Clery A, Schlotter F, Senty-Segault V, Branlant C. 2011. A second base pair interaction between U3 small nucleolar RNA and the 5′-ETS region is required for early cleavage of the yeast pre-ribosomal RNA. *Nucleic Acids Res* **39:** 9731–9745. doi:10.1093/nar/gkr675

Mercereau-Puijalon O, Barale JC, Bischoff E. 2002. Three multigene families in *Plasmodium* parasites: facts and questions. *Int J Parasitol* **32:** 1323–1344. doi:10.1016/S0020-7519(02)00111-X

Meyer A, Schartl M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11:** 699–704. doi:10.1016/S0955-0674(99)00039-3

Mishima Y, Tomari Y. 2016. Codon usage and 3′ UTR length determine maternal mRNA stability in zebrafish. *Mol Cell* **61:** 874–885. doi:10.1016/j.molcel.2016.02.027

Pagano JFB, Locati MD, Ensink WA, van Olst M, van Leeuwen S, De Leeuw WS, Nehrdich U, Spaink HP, Rauwerda H, Jonker MJ, et al. 2020. Maternal- and somatic-type snoRNA expression and processing in zebrafish development. bioRxiv doi:101101/858936

Parks MM, Kurylo CM, Dass RA, Bojmar L, Lyden D, Vincent CT, Blanchard SC. 2018. Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci Adv* **4:** eaao0665. doi:10.1126/sciadv.aao0665

Presslauer C, Tilahun Bizuayehu T, Kopp M, Fernandes JM, Babiak I. 2017. Dynamics of miRNA transcriptome during gonadal development of zebrafish. *Sci Rep* **7:** 43850. doi:10.1038/srep43850

Rajan KS, Zhu Y, Adler K, Doniger T, Cohen-Chalamish S, Srivastava A, Shalev-Benami M, Matzov D, Unger R, Tschudi C, et al. 2020. The large repertoire of 2′-*O*-methylation guided by C/D snoRNAs on *Trypanosoma brucei* rRNA. *RNA Biol* **17:** 1018–1039. doi:10.1080/15476286.2020.1750842

Ramesh M, Woolford JL Jr. 2016. Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. *RNA* **22:** 1153–1162. doi:10.1261/rna.056705.116

Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33:** W686–W689. doi:10.1093/nar/gki366

Segner H. 2009. Zebrafish (Danio rerio) as a model organism for investigating endocrine disruption. *Comp Biochem Physiol C Toxicol Pharmacol* **149:** 187–195. doi:10.1016/j.cbpc.2008.10.099

Shalev-Benami M, Zhang Y, Rozenberg H, Nobe Y, Taoka M, Matzov D, Zimmerman E, Bashan A, Isobe T, Jaffe CL, et al. 2017. Atomic resolution snapshot of Leishmania ribosome inhibition by the aminoglycoside paromomycin. *Nat Commun* **8:** 1589. doi:10.1038/s41467-017-01664-4

Shi Z, Barna M. 2015. Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. *Annu Rev Cell Dev Biol* **31:** 31–54. doi:10.1146/annurev-cellbio-100814-125346

Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. 2014. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508:** 66–71. doi:10.1038/nature13007

Tessarz P, Santos-Rosa H, Robson SC, Sylvestersen KB, Nelson CJ, Nielsen ML, Kouzarides T. 2014. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505:** 564–568. doi:10.1038/nature12819

Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147:** 1537–1550. doi:10.1016/j.cell.2011.11.055

Vastenhouw NL, Cao WX, Lipshitz HD. 2019. The maternal-to-zygotic transition revisited. *Development* **146:** dev161471. doi:10.1242/dev.161471

Vasudevan S, Seli E, Steitz JA. 2006. Metazoan oocyte and early embryo development program: a progression through translation regulatory cascades. *Genes Dev* **20:** 138–146. doi:10.1101/gad.1398906

Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E. 2009. EnsemblCompara GeneTrees: complete, duplication-aware

phylogenetic trees in vertebrates. *Genome Res* **19:** 327–335. doi:10.1101/gr.073585.107

Watkins NJ, Bohnsack MT. 2012. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA* **3:** 397–414. doi:10.1002/wrna.117

Westerfield M. 2000. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* University of Oregon Press, Eugene, OR.

Winata CL, Korzh V. 2018. The translational regulation of maternal mRNAs in time and space. *FEBS Lett* **592:** 3007–3023. doi:10.1002/1873-3468.13183

Winata CL, Lapinski M, Pryszcz L, Vaz C, Bin Ismail MH, Nama S, Hajan HS, Lee SGP, Korzh V, Sampath P, et al. 2018. Cytoplasmic polyadenylation-mediated translational control of maternal mRNAs directs maternal-to-zygotic transition. *Development* **145:** dev159566. doi:10.1242/dev.159566

Wolfe K. 2000. Robustness–it's not where you think it is. *Nat Genet* **25:** 3–4. doi:10.1038/75560

Xue S, Barna M. 2012. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* **13:** 355–369. doi:10.1038/nrm3359

Yang Y, Wang L, Han X, Yang WL, Zhang M, Ma HL, Sun BF, Li A, Xia J, Chen J, et al. 2019. RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. *Mol Cell* **75:** 1188–1202 e1111. doi:10.1016/j.molcel.2019.06.033

Yoshihama M, Nakao A, Kenmochi N. 2013. snOPY: a small nucleolar RNA orthological gene database. *BMC Res Notes* **6:** 426. doi:10.1186/1756-0500-6-426

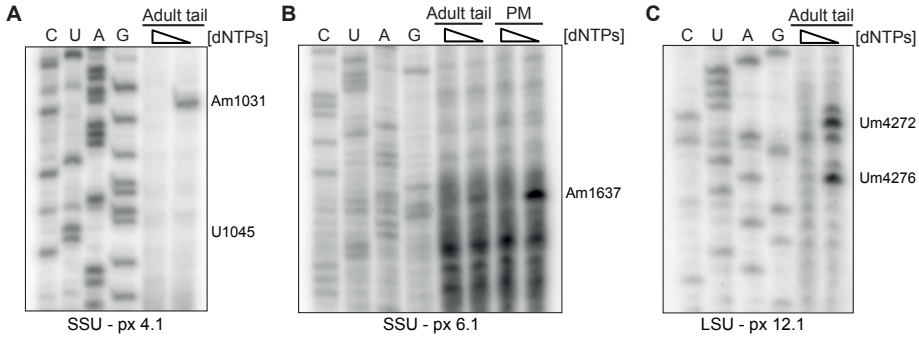

A
Adult tail
C U A G
[dNTPs]
Am1031
U1045
SSU - px 4.1
B
Adult tail PM
C U A G
[dNTPs]
Am1637
SSU - px 6.1
C
Adult tail
C U A G
[dNTPs]
Um4272
Um4276
LSU - px 12.1

Figure S1

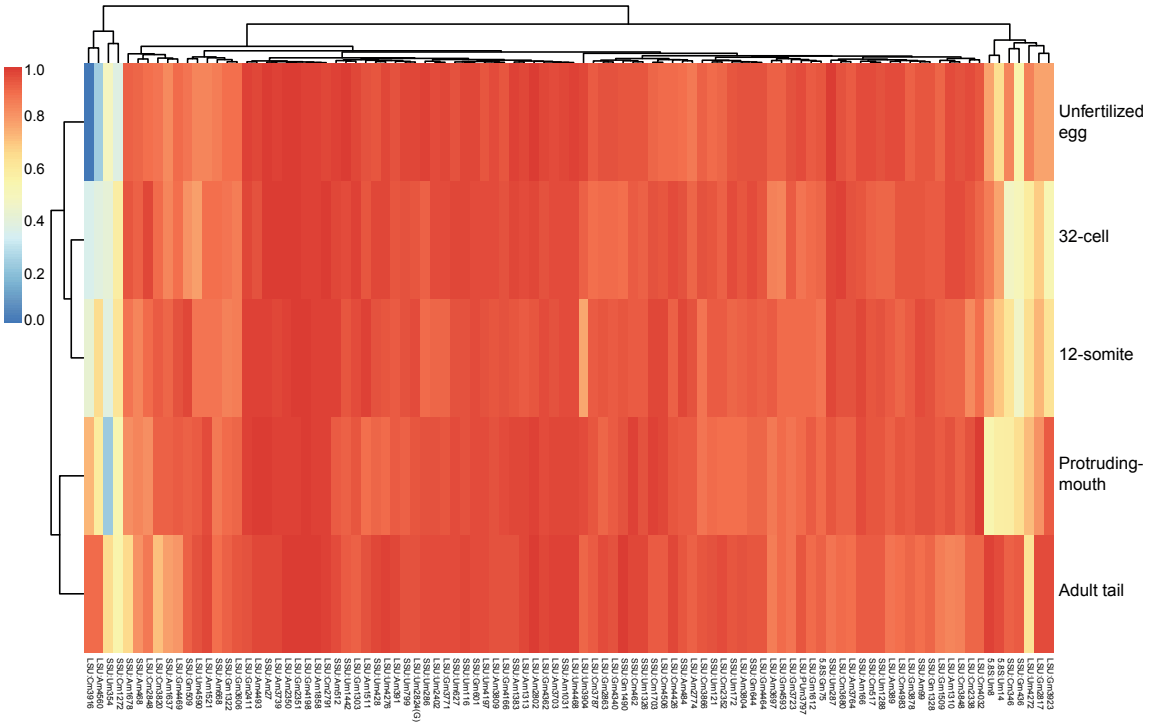
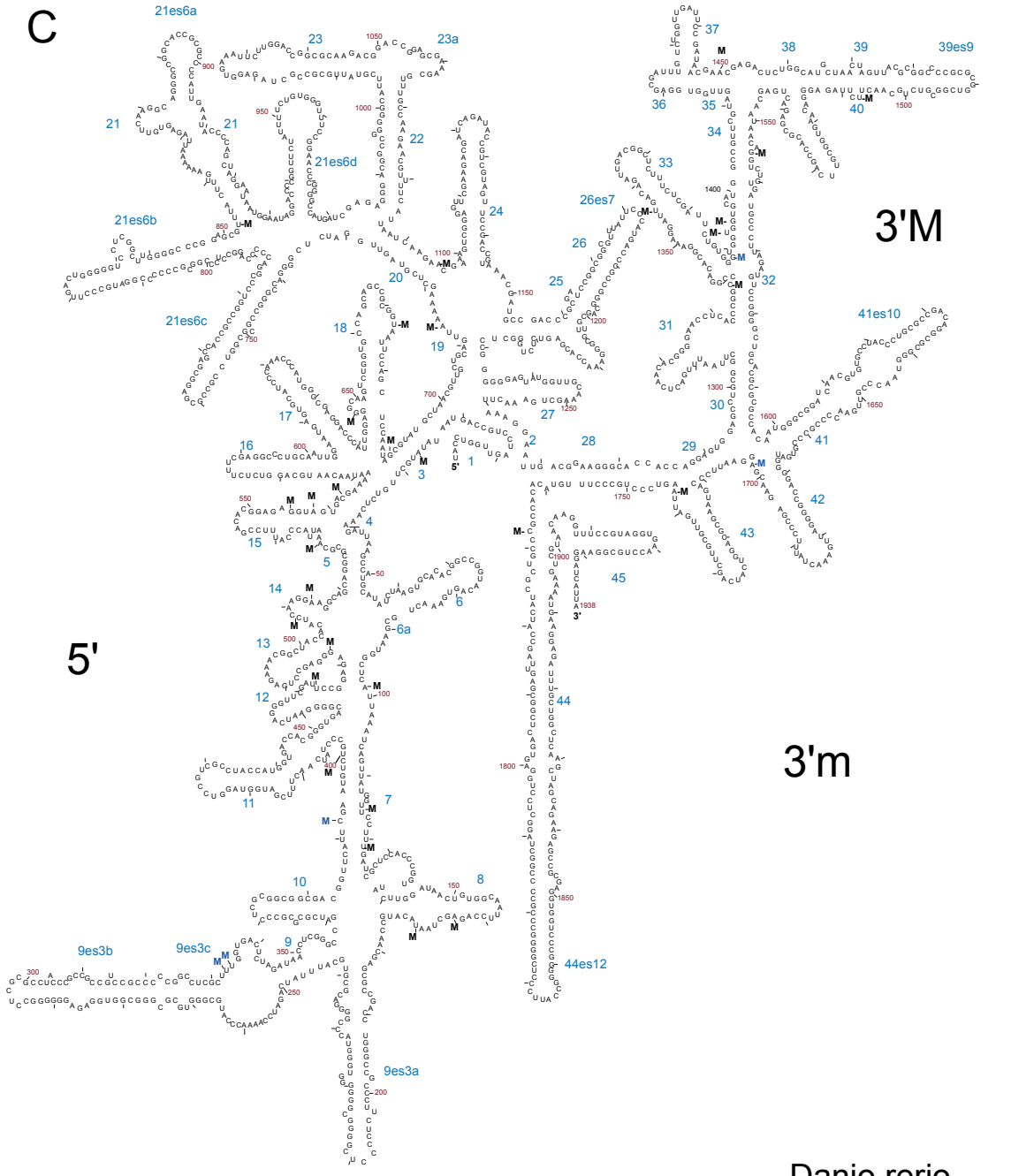

1.0
0.8
0.6
0.4
0.2
0.0
Unfertilized egg
32-cell
12-somite
Protruding-mouth
Adult tail

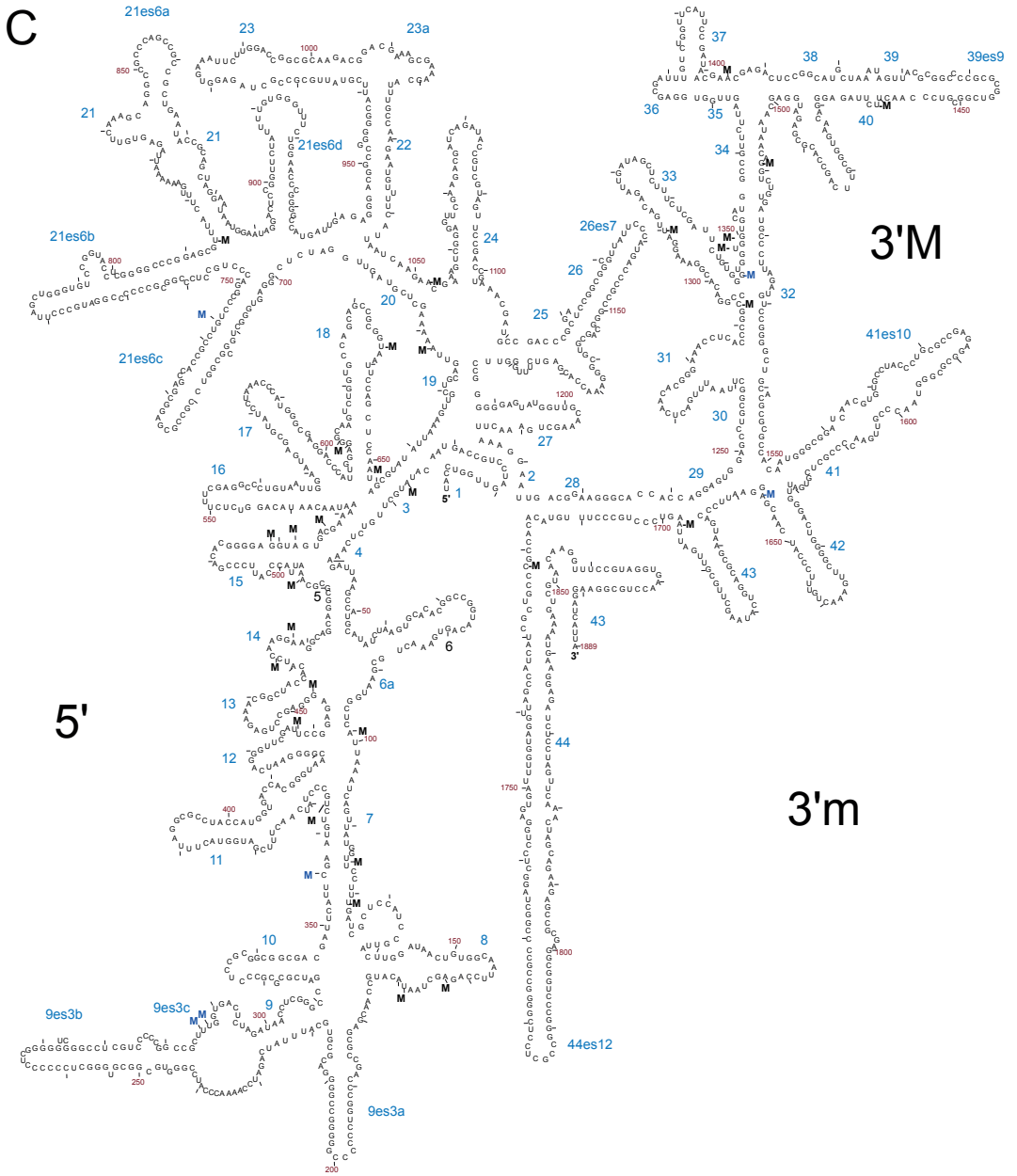
Figure S2



Danio rerio

early embryonic small subunit ribosomal RNA

Figure S3



Danio rerio

late embryonic small subunit ribosomal RNA

Figure S4

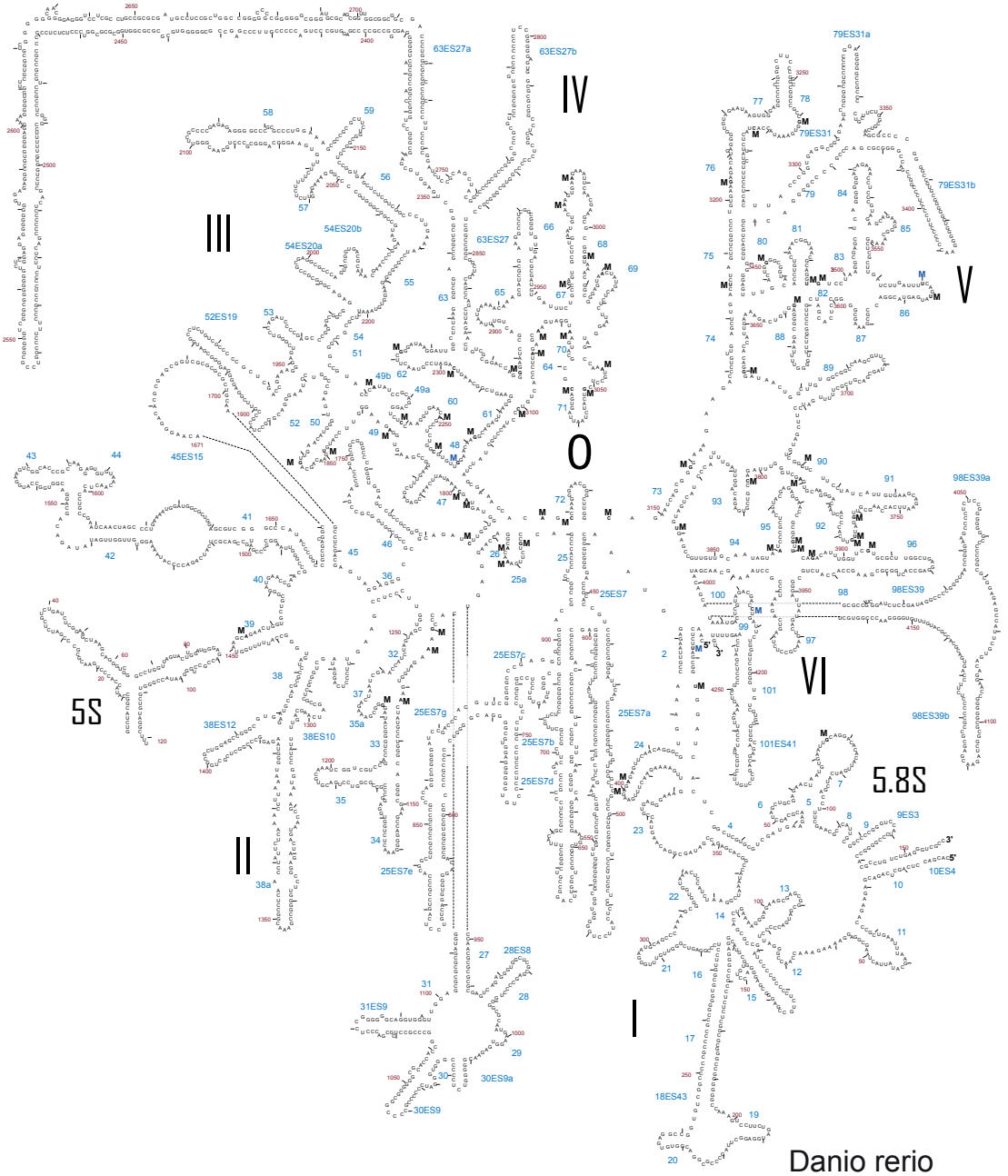
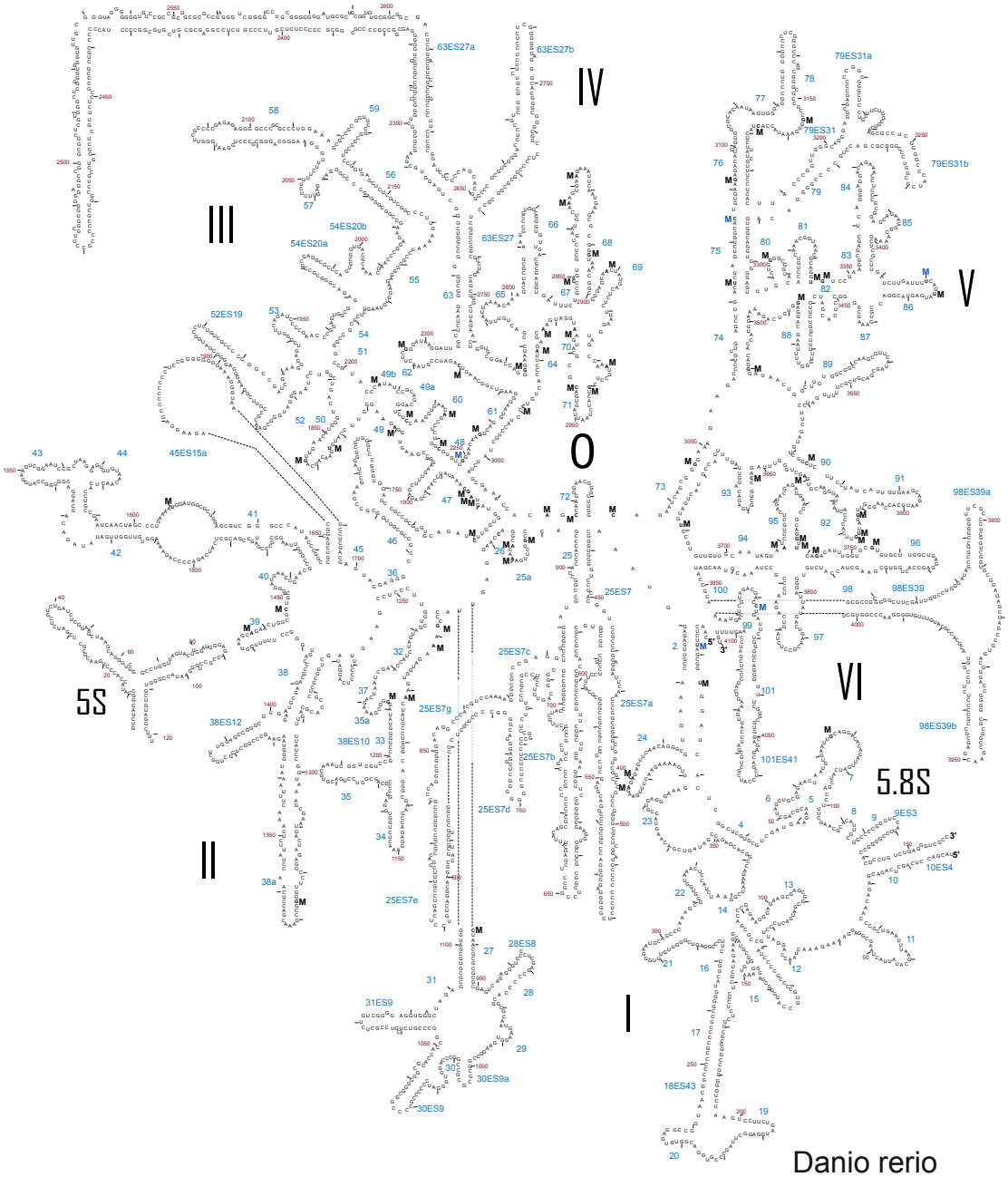

I
II
III
IV
V
VI
5S
5.8S
Danio rerio
early embryonic large subunit ribosomal RNA

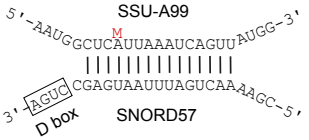
Figure S5



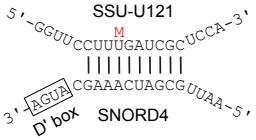
late embryonic large subunit ribosomal RNA

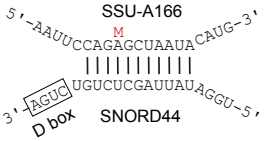
Figure S6

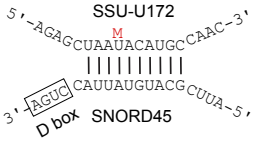

SSU-A27
5'-AGUAAUAUAUGCUUGUCUCAA-3'
M
3'-AGUCUAUAUACGAACAGUACC-5'
D' box
SNORD27

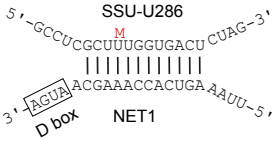

SSU-A99
5'-AAUGGCUCAUUAAAUCAGUUAUGG-3'
M
3'-AGUCCGAGUAAUUUAGUCAAAAGC-5'
D box
SNORD57


SSU-U116
5'-GUUAUGGUUCCUUUGAUCGC-3'
M
3'-AGUCACCAAGGAAACUCUUG-5'
D box
SNORD42

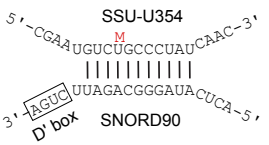

SSU-U121
5'-GGUUCCUUUGAUCGCUCCA-3'
M
3'-AGUACGAAACUAGCGUUAA-5'
D' box
SNORD4


SSU-A166
5'-AAUUCCAGAGCUAAUACAUG-3'
M
3'-AGUCUGUCUCGAUUAUAGGU-5'
D box
SNORD44


SSU-U172
5'-AGAGCUAAUACAUGCCAAC-3'
M
3'-AGUCCAUUAUGUACGCUUA-5'
D box
SNORD45


SSU-U286
5'-GCCUCGCUUUGGUGACUCUAG-3'
M
3'-AGUAACGAAACCACUGAAAUU-5'
D box
NET1


SSU-C346
5'-GGUUCAUUCGAAUGUCUGCCCU-3'
M
3'-AGUCCUAAGCUUACAGACACAA-5'
D' box
SNORD200

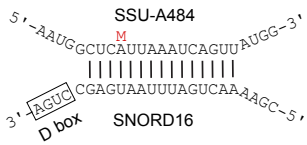

SSU-U354
5'-CGAAUGUCUGCCCUAUCAAC-3'
M
3'-AGUCUUAGACGGGAUACUCA-5'
D' box
SNORD90


SSU-U428
5'-GGUUCGAUUCCGGAGAGGGA-3'
M
3'-CGUCUCUAAGGCCUCUUUCG-5'
D' box
SNORD68


SSU-G436
5'-UCCGGAGAGGGAGCCUGAG-3'
M
3'-AGUCAUCUCCCUCGGUUUA-5'
D' box
SNORD100

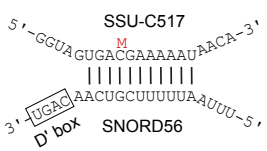

SSU-C462
5'-UACCACAUCCAAGGAAGGCAGC-3'
M
3'-AGACUGUAGGUUCCUUCCACGA-5'
D box
SNORD14


SSU-A468
5'-AUCCAAGGAAGGCAGCAGGCG-3'
M
3'-AGUAUUCCUUCCGUCGUACCU-5'
D box
SNORD83


SSU-A484
5'-AAUGGCUCAUUAAAUCAGUUAUGG-3'
M
3'-AGUCCGAGUAAUUUAGUCAAAAGC-5'
D box
SNORD16

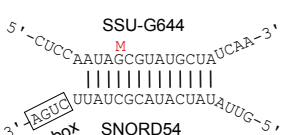

SSU-G509
5'-CACGGAGAGGUAGUGACGAAAAAUAAC-3'
M
3'-AGUCUUCUCCAUCACUGUUUUUUGGGU-5'
D box
SNORD11

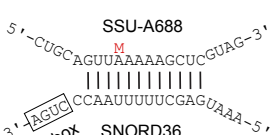

SSU-A512
5'-GAGAAGGUAGUGACGAAAAAUAAC-3'
M
3'-AGUCACCAUCACUGCUUAUUGACA-5'
D box
SNORD70


SSU-C517
5'-GGUAGUGACGAAAAAUAACA-3'
M
3'-UGACAACUGCUUUUUAAUUU-5'
D' box
SNORD56

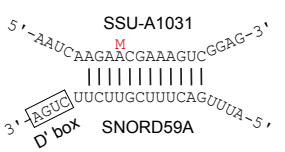

SSU-G601
5'-CAUUGGAGGGCAAGUCUGGUGC-3'
M
3'-AGUCUCUCCCGUUCAAACAGUA-5'
D' box
SNORD103

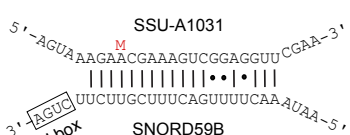

SSU-U627
5'-AGCCGCGGUAAUUCCAGCUCCAA-3'
M
3'-AGACUGCCAUUAAGGUCGAUAAU-5'
D' box
SNORD65


SSU-G644
5'-CUCCAAUAGCGUAUGCUAUCAA-3'
M
3'-AGUCUUAUCGCAUACUAUAUUG-5'
D box
SNORD54


SSU-A688
5'-CUGCAGUUAAAAAGCUCGUAG-3'
M
3'-AGUCCCAAUUUUUCGAGUAAA-5'
D box
SNORD36


SSU-U799
5'-AGUAAGCGUUUACUUUGAAAAAAUU-3'
M
3'-AGUCACGCAAAUGAAACUUUUCUUA-5'
D box
SNORD105


SSU-A1031
5'-AAUCAAGAACGAAAGUCGGAG-3'
M
3'-AGUCUUCUUUGCUUUCAGUUUA-5'
D' box
SNORD59A

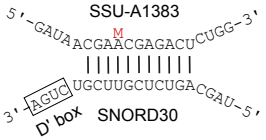

SSU-A1031
5'-AGUAAAGAACGAAAGUCGGAGGUUCGAA-3'
M
3'-AGUCUUCUUUCAGUUUUCAAAUAA-5'
D' box
SNORD59B

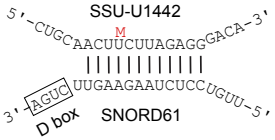

SSU-C1272
5'-CACCCGGCCCGGACACGGAAA-3'
M
3'-AGUCACCGGGUCUGUGCACCG-5'
D' box
SNORD66


SSU-U1288
5'-GGAAAGGAUUGACAGAUUGACG-3'
M
3'-AGUCACCUAACUGUCUAAUACA-5'
D box
SNORD110

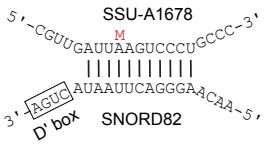

SSU-U1326
5'-GUGGGUGGUGGUGCAUGGCCG-3'
M
3'-AGUCGACCAUCACGUACUUCA-5'
D box
SNORD33


SSU-G1328
5'-GGGUGGUGGUGCAUGGCCG-3'
M
3'-AGUC UCACCACGUACUAAA-5'
D' box
SNORD32


SSU-A1383
5'-GAUAACGAACGAGACUCUGG-3'
M
3'-AGUC UGCUUGCUCUGACGAU-5'
D' box
SNORD30


SSU-U1442
5'-CUGCAACUUCUUAGAGGGACA-3'
M
3'-AGUC UUGAAGAAUCUCCUGUU-5'
D box
SNORD61

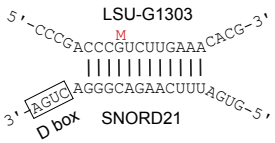

SSU-G1490
5'-CAAUAACAGGUCUGUGAUGCC-3'
M
3'-AGUC AUGUCCAGACACUUUCA-5'
D' box
SNORD25

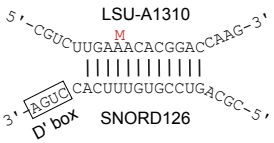

SSU-A1678
5'-CGUUGAUUAAGUCCCUGCCC-3'
M
3'-AGUC AUAAUUCAGGGAACAA-5'
D' box
SNORD82

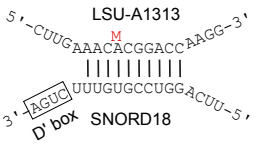

SSU-C1703
5'-AGAGACCGCCCGUCGUCAA-3'
M
3'-AUAC AGGCGGGCAGCUCUU-5'
D' box
SNORD43

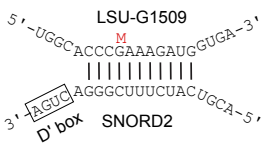

LSU-A389
5'-CUUUGAAGAGAGAGUUCAACA-3'
M
3'-AGUC CUUCUCUCUUAAGACUU-5'
D' box
SNORD26


LSU-A391
5'-UUGAAGAGAGAGUUCAACAG-3'
M
3'-AGUC UCUCUCUCAAGUACGC-5'
D' box
SNORD81

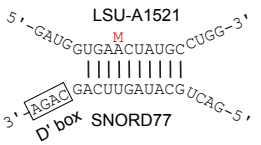

LSU-G1303
5'-CCCGACCCGUCUUGAAACACG-3'
M
3'-AGUC AGGGCAGAACUUUAGUG-5'
D box
SNORD21


LSU-A1310
5'-CGUCUUGAAAACACGGACCAAG-3'
M
3'-AGUC CACUUUGUGCCUGACGC-5'
D' box
SNORD126


LSU-A1313
5'-CUUGAAACACGGACCAAGG-3'
M
3'-AGUC UUUGUGCCUGGACUU-5'
D' box
SNORD18


LSU-G1509
5'-UGGCACCCGAAAGAUGGUGA-3'
M
3'-AGUC AGGGCUUUCUACUGCA-5'
D' box
SNORD2

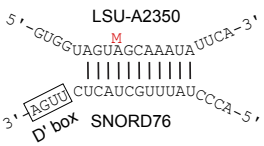

LSU-A1511
5'-AUCCCCGAAAGAUGGUGAACUA-3'
M
3'-AGUC AGCUUUCUACCACUCUAU-5'
D' box
SNORD51


LSU-A1521
5'-GAUGGUGAACUAUGCCUGG-3'
M
3'-AGAC GACUUGAUACGUCAG-5'
D' box
SNORD77

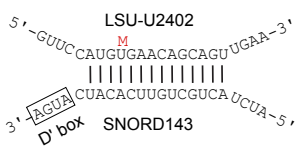

LSU-G1612
5'-GGGCAUAGGGGCGAAAGACUAAU-3'
M
3'-AGUC GAUCCUCGCUUUCUGUAAC-5'
D box
SNORD80

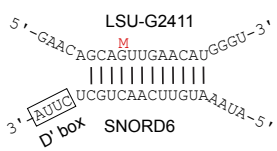

LSU-A1858
5'-UGGUAAGCAGAACUGGUGCUG-3'
M
3'-AGUC AUCGUCUUGACCAAUUU-5'
D' box
SNORD38


LSU-C2338
5'-GCGCAGAUCUUGGUGGUAGU-3'
M
3'-UGAC GCUAGAACCACCUUUA-5'
D box
SNORD24

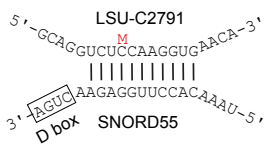

LSU-A2350
5'-GUGGUAGUAGCAAAUAUUCA-3'
M
3'-AGUU CUCAUCGUUUAUCCCA-5'
D' box
SNORD76

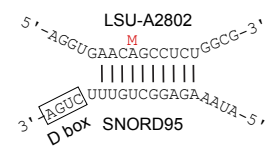

LSU-C2352
5'-GGUAGUAGCAAAUAUUCAAAC-3'
M
3'-AGUC UAUCGUUUAUAAGAAGU-5'
D' box
SNORD24


LSU-U2402
5'-GUUCCAUGUGAACAGCAGUUGAA-3'
M
3'-AGUA CUACACUUGUCGUCAUCUA-5'
D' box
SNORD143

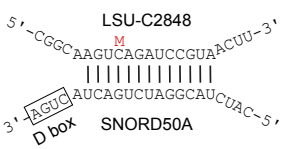

LSU-G2411
5'-GAACAGCAGUUGAACAUGGGU-3'
M
3'-AUUC UCGUCAACUUGUAAAUA-5'
D' box
SNORD6


LSU-A2774
5'-CCGUACCCAUAUCCGCAGCAG-3'
M
3'-AGUC AGGGUAUAGGCGUUCAA-5'
D box
SNORD99

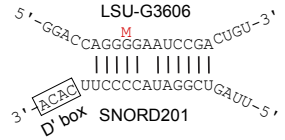

LSU-C2791
5'-GCAGGUCUCCAAGGUGAACA-3'
M
3'-AGUC AAGAGGUUCCACAAAU-5'
D box
SNORD55


LSU-A2802
5'-AGGUGAACAGCCUCUGGCG-3'
M
3'-AGUC UUUGUCGGAGAAAUA-5'
D box
SNORD95


LSU-G2824
5'-UGGAACAAGGCAGAGUAA-3'
M
3'-AGUU UGUUCCGUCUUUUU-5'
D' box
SNORD34

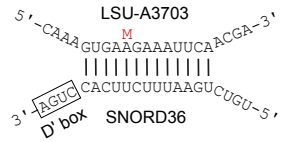

LSU-C2848
5'-CGGCAAGUCAGAUCCGUAACUU-3'
M
3'-AGUC AUCAGUCUAGGCAUCUAC-5'
D box
SNORD50A


LSU-G2863
5'-GUAACUUCGGGAUAAGGAUUG-3'
M
3'-AGUC UAAGCCCUAUUUCCUCA-5'
D' box
SNORD50B


LSU-G3606
5'-GGACCAGGGGAAUCCGACUGU-3'
M
3'-ACACUUCCCCAUAGGCUGAUU-5'
D' box
SNORD201

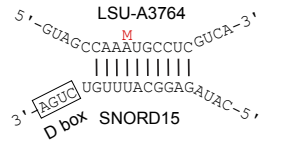

LSU-C3680
5'-UUUCUGCCCAGUGCUCUGA-3'
M
3'-AGUCUCGGGUCACGAUUUG-5'
D' box
SNORD88

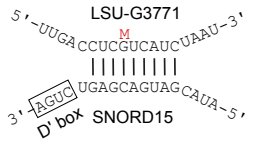

LSU-A3697
5'-GAAUGUCAAAGUGAAGAAAUUC-3'
M
3'-AGUCCAGUUUCACUUCUUACAA-5'
D' box
SNORD37


LSU-A3703
5'-CAAAGUGAAGAAAUUCAACGA-3'
M
3'-AGUCCACUUCUUUAAGUCUGU-5'
D' box
SNORD36

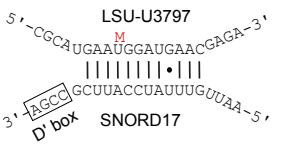

LSU-G3723
5'-CGUUGCGGGUAAACGGCGGGAGUAACU-3'
M
3'-AGACUACCCAUUUGCCGCUUUCACACU-5'
D' box
SNORD87


LSU-A3739
5'-CGGGAGUAACUAUGACCUCU-3'
M
3'-AUACACAUUGAUACUGCUUA-5'
D' box
SNORD46


LSU-A3764
5'-GUAGCCAAAUGCCUCGUCA-3'
M
3'-AGUCUGUUUACGGAGAUAC-5'
D box
SNORD15

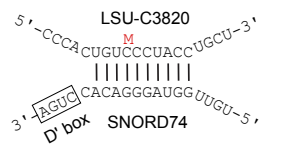

LSU-G3771
5'-UUGACCUCGUCAUCUAAU-3'
M
3'-AGUCUGAGCAGUAGCAUA-5'
D' box
SNORD15


LSU-C3787
5'-AUUAGUGACGCGCAUGAAUG-3'
M
3'-AGUCAACUGUGCGUACCCCU-5'
D' box
SNORD10


LSU-U3797
5'-CGCAUGAAUGGAUGAACGAGA-3'
M
3'-AGCCGCUUACCUAUUUGUUAA-5'
D' box
SNORD17

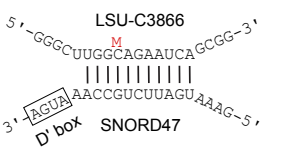

LSU-A3804
5'-AUGGAUGAACGAGAUUCCCACU-3'
M
3'-AGUCUGCUUGCUCUGACGAUUU-5'
D' box
SNORD30


LSU-A3809
5'-UGGCCGAGAUUCCCACUGUCCCU-3'
M
3'-AGUCUCUACUAGGGUAACAAAGU-5'
D' box
SNORD79

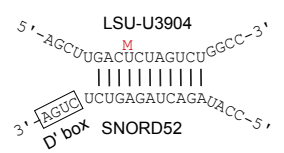

LSU-C3820
5'-CCCACUGUCCCUACCUGCU-3'
M
3'-AGUCCACAGGGAUGGUUGU-5'
D' box
SNORD74

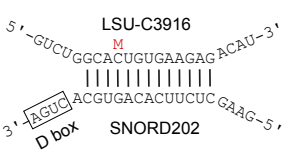

LSU-C3848
5'-AACCACAGCCAAGGGAACGGG-3'
M
3'-AGUCUGUCGGUUCUCUUCGCA-5'
D box
SNORD53A

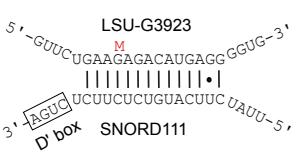

LSU-C3848
5'-AACCACAGCCAAGGGAACGGG-3'
M
3'-AGUCUGUCGGUUCUCUUCUCA-5'
D box
SNORD53B

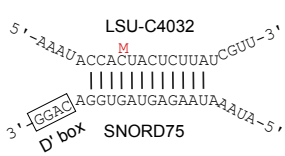

LSU-C3866
5'-GGGCUUGGCAGAAUCAGCGG-3'
M
3'-AGUAAACCGUCUUAGUAAAG-5'
D' box
SNORD47


LSU-G3878
5'-AUCAGCGGGGAAAGAAGACCC-3'
M
3'-AGACGACCCCUUUUUUCCAGU-5'
D' box
SNORD12

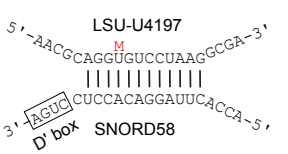

LSU-U3904
5'-AGCUUGACUCUAGUCUGGCC-3'
M
3'-AGUCUCUGAGAUCAGAUACC-5'
D' box
SNORD52

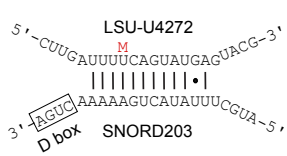

LSU-C3916
5'-GUCUGGCACUGUGAAGAGACAU-3'
M
3'-AGUCACGUGACACUUCUCGAAG-5'
D box
SNORD202


LSU-G3923
5'-GUUCUGAAGAGACAUGAGGGGUG-3'
M
3'-AGUCUCUUCUCUGUACUUCUAUU-5'
D' box
SNORD111


LSU-C4032
5'-AAAUACCACUACUCUUAUCGUU-3'
M
3'-GGACAGGUGAUGAGAAUAAAUA-5'
D' box
SNORD75

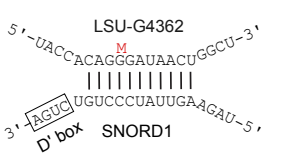

LSU-G4166
5'-UUUGACUGGGGCGGUACACC-3'
M
3'-AGUCGGACCCCGCCAUUAUC-5'
D' box
SNORD31


LSU-U4197
5'-AACGCAGGUGUCCUAAGGCGA-3'
M
3'-AGUCCUCCACAGGAUUCACCA-5'
D' box
SNORD58


LSU-U4272
5'-CUUGAUUUUCAGUAUGAUGUACG-3'
M
3'-AGUCAAAAAGUCAUAUUUCGUA-5'
D box
SNORD203


LSU-U4276
5'-AUUUUCAGUAUGAGUACGGAC-3'
M
3'-AGUUUGUCAUACUUAUGAACU-5'
D box
SNORD41


LSU-G4340
5'-AAGCGGGAGGUGUCAGAAAAGUUA-3'
M
3'-AGUAUGCUCCACAGUCUUUUUCAU-5'
D' box
SNORD60


LSU-G4362
5'-UACCACAGGGAUAACUGGCU-3'
M
3'-AGUCUGUCCCUAUUGAAGAU-5'
D' box
SNORD1

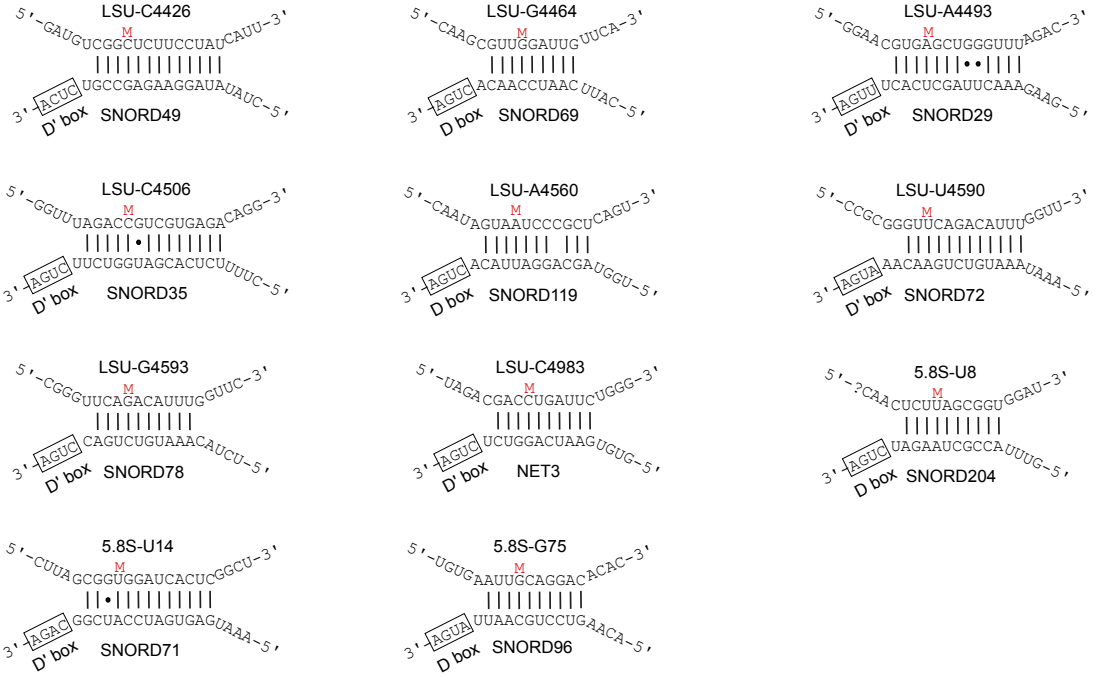

LSU-C4426
5'-GAUGUCGGCUCUUCCUAUCAUU-3'
M
3'-ACUCUGCCGAGAAGGAUAUAUC-5'
D' box
SNORD49
LSU-G4464
5'-CAAGCGUUGGAUUGUUCA-3'
M
3'-AGUCACAACCUAACUUAC-5'
D box
SNORD69
LSU-A4493
5'-GGAACGUGAGCUGGGUUUAGAC-3'
M
3'-AGUUUCACUCGAUUCAAAGAAG-5'
D' box
SNORD29
LSU-C4506
5'-GGUUUAGACCGUCGUGAGACAGG-3'
M
3'-AGUCUUCUGGUAGCACUCUUUUC-5'
D' box
SNORD35
LSU-A4560
5'-CAAUAGUAAUCCCGCUCAGU-3'
M
3'-AGUCACAUUAGGACGAUGGU-5'
D box
SNORD119
LSU-U4590
5'-CCGCGGGUUCAGACAUUUGGUU-3'
M
3'-AGUAAACAAGUCUGUAAAUAAA-5'
D' box
SNORD72
LSU-G4593
5'-CGGGUUCAGACAUUUGGUUC-3'
M
3'-AGUCCAGUCUGUAAACAUCU-5'
D' box
SNORD78
LSU-C4983
5'-UAGACGACCUGAUUCUGGG-3'
M
3'-AGUCUCUGGACUAAGUGUG-5'
D' box
NET3
5.8S-U8
5'-?CAACUCUUAGCGGUGGAU-3'
M
3'-AGUCUAGAAUCGCCAUUUG-5'
D box
SNORD204
5.8S-U14
5'-CUUAGCGGUGGAUCACUCGGCU-3'
M
3'-AGACGGCUACCUAGUGAGUAAA-5'
D' box
SNORD71
5.8S-G75
5'-UGUGAAUUGCAGGACACAC-3'
M
3'-AGUAUUAACGUCCUGAACA-5'
D box
SNORD96

Figure S7

SNORD11 - SSU509

```
Early: 5'CACGGAGAGGUAGUGACGAAAAAUAAC
              ||||||||||||•|||||
       3'AGUCUUCUCCAUCACUGUUUUUU

Late:  5'CACGGGGAGGUAGUGACGAAAAAUAAC
              •||||||||||||•|||||
       3'AGUCUUCUCCAUCACUGUUUUUU

Hsa:   5'CCCGGGGAGGUAGUGACGAAAAAUAAC
              •||||||||||||
       3'AGUCUUCUCCAUCACUGAUUUCU
```

SNORD34 - LSU2824

```
Early: 5'UGGAACAAGGCAGAGUAAGGG
              |||||||||•
       3'AGUUUGUUCCGUCUUUUU

Late:  5'UGGAACAAGGCAG-GUAAGGG
              ||||||||-•
       3'AGUUUGUUCCGUC-UUUUU

Hsa:   5'UGGAACAAUGUAG-GUAAGGG
              |||||||| |||
       3'AGUUUGUUACAUC-CAUCAAC
```

SNORD50A - LSU2848

```
Early: 5'CGGCAAGUCAGAUCCGUAACUUC
              ||||||||||||
       3'AGUCAUCAGUCUAGGCAUCUACC

Late:  5'CGGCAAGUCAGAUCCGUAACUUC
              ||||||||||||
       3'AGUCAUCAGUCUAGGCAUCUACC

Hsa:   5'CGGCAAGCCGGAUCCGUAACUUC
              ||||•||||||||
       3'AGUCUUCGGUCUAGGCAUUUUCA
```

SNORD41 - LSU4276

```
Early: 5'AUUUUCAGUAUGAGUACGGAC
              ||||||||•|||
       3'AGUCUGUCAUACUUAUGAACU

Late:  5'AUUUUCAGUAUGAGUACGGAC
              ||||||||•|||
       3'AGUCUGUCAUACUUAUGAACU

Hsa:   5'AUUUUCAGUACGAAUACAGAC
              ||||||||||||
       3'AGUCGGUCAUGCUUAUGCGCU
```

Figure S8

TCATAAAAATGACTTACATTCATGTAATCTTGTATTTACATTCAGGCAGA
TATTCCGCTACCCTGCCCTTCTGCAGCATTAGGACCCTTTTAAATGTTCC
TGCGTGCACGCTGCCATGATCCTTCATTATGCGGATGTTTGTGGGCTAGG
AGACAGTGACAAATTTGCTTCTTTCGTATTAGTTAAAATATCTTTTTGGT
TTTATATTATATATACACGTTTACTTTGAAAATGTCGATAGTCTGTGCGT
TTACCTGAAAGCAGTATCTGTAAAATAACTGACGTTGGCTGCTCACGGCC
TCTCCCTCGTGGTCGTAATTTTATTTAGTTAAGCATTTTTATCGAATATG
ACTTGATAATTTTTTAGTAAGCTTTTGAAAATGCTTGCTGCTTATAAATA
TTGTTGTGTAGAGAACCAAAGTCCAAGAAACGGTGCGAAAGGGCAAACGC
AGATTTCTTGTAAGTACAAGAACACTTTTTTTTTTTTTTTTTAAAGAG
GATTTTAGATTGATGTTAAAATATATATAAACATGCTAAATGACTGGAGA
TTTTTTAATAAAAATATATGATTTAAATGTTGTTTTACTGCCTGTTATGA SNORD52A (SNORD202)
TGAATCTTTAGAGTCTCTGATTTAATCATGAAGCTCTTCACAGTGCACTG
AATGCATGCAGTTCTAGTCATTTTTGATTTTTGTTGTGGCTTGAATGTAC
CTAAATGTTCAAATAACCTTTACAGGTTGGATGGCAGGCCGTTGACGCCA
ACTGTTGACAACCATTATGTAAGTATTACTTTTCCAATGAAAACCTTTTT
TCGAAGACAGCATTTTGCCCTTATTGCATGTAAAGATGAATCATTTGACT SNORD52B
AGAGTCTCTGATTTCTCCATGAAGCTCTTCACAGTGCACTGAACACATGC
AGTACTGCTTATGATGGATTTGTTTTGGCTCGAATGTTCCTAATTGTAAT
TTTTCACAGGTTGGATGAGAGAAATTGAGGCTGGTAGACGATGCCAACAG
TTGACGACCAATTGGTAAGAATTAATTTTTCAACCAAAAGTTTCATTCCA
AGACGGCAAATTATCTTTTAGCCAATTATTGCATGTAATGATGAATCATT SNORD52C
TGACTAGAGTCTCTGATTTAACTATGAAGTTCTTCACAGTGCACTGAACT
CATGCTGCATTGCTTACAATGGATTCAGTTTTTGGCTTGAATATTCCTAA
ATAAGTTACCTTTACAGGTTGGATGGAAGAAATTGAGGCCAGTAGATGAT
GCCAACAGTTGACGATCAGTTGGTAAGAATTACCTTTCCAATAGAAAAAA
AAAAATTCAAAGAAAGCACATTGTCCTTTAGCCAATTACTGCCTGTAATG SNORD52D
ATGAATCATTTGACTAGAGTCTCTGATTTAAGAATGAAGTTCTTCACAGT
GCACTGAACACATGCAGTACTGTTTATGATGGATCTTGTTCTTGGCTTGA
ATATTCCTAAATGTTAAAATACCTTCTCAGGGGCCTGTTTCAGAAAGGAG
GTTAACTGAAAACTCAGAGTATTTTAACCCTGAAATGAGAGAAACTCTGG
GTTTTCCGTTTCAAAATGGCAGGTTTGTTAAACTCGAGAAAGCAGGGTAA
GTCAAGCCTGTTTCTGAAAGAAAGGTAACTTTAACTCAGAGTCAGTTACT
GTGGTAACTTACTCTGTGAATCTAACCTGGTCAAAAGCAGGTTTTATTCT
CTAAACTCAGAGTTTCTGTCTGTCTCCTCCCCTTTTTTAAAGATGAAGCG
GTATTTCTCGTCTTAGCCTTACGTTTCTACCCACCTATTTTAATGCTCAT
TTTGGATATGTGCATAAAAATGATTGATAGAAACGTCAAGATGCACATAA
CTTTTTAAAATATGCATAAAAAACGTGCATAACTGAGTAGGATAAACTTT
TATTCGATAGAAAATGTGCGCATAAACTACGATGGAAACACTTTTACTGA
ACAAATTACAGTATGTACATTAAAAAAAAGATCATGATCATATGATAATG
AAAATGTGTGTAAATGAACAAACCAGCAGGCTGAGCACACTGTAACATGT
CTTATTGTGTTAGTCATTCTAAAATGCCTTAACTGTTTCAGTATTTGTGT
ATATTATTAATTACCTCCGAACGTCTGGTGCGTGTCAAGAGTCTCCGCGT
CTCATGGCTTCAGACGCCCCCACGTGTTCATTGTGTTTCAGCATTGTCTT
CTGACATTGAGGCGCAAGTACATTAATTAGGCTAAATAAAGAATTAATTG
ACGCAGCTTCTTCTACCACAGTAAATTCTGTTTGATATTTGGTGCCAGTT
TAATCAGGAAGTGACGATTTTTTTCTCTTTGACTCGTTGGATGGAATTTT
ATTCGCATCTTTTATGCAATATTCCAGTTTTGCACATACATTTATGTAAT
ATATTTGGATGGAAACATAGCTATTGCCTACATTTTAGTCACACACAGAA
GAACTGTACGAGAATGGCTTGTCCTTTTTTAATAAATAATTAAATTAAAT
TCACCTTCGACATGCACTGTAACTAGATTAATGGGATTATTATTCTTTCT
AAAAATTATTTTTAGTACTGCTTACCTTCTAAATGTTATATTAACCTCTA
TCACTTGATCAATAAAAAAGGCAGACACACAAGTGCACTGTTAAATCATT
TAAAACTAATAAATGTGTAAAAAAGTTTTGAAAAAAAAATAAACGTCACA
AATTACATTACTTATTTTATTATACTTAAATATTTAAATACTTAAAAAGT
GAAATTAGGCATTAACTTGAGCAGCTGTTTTCCCACGCTAACTGTCTTTT
CTTCACTGATGGTTGCTTCTTTTCAAATATATACGCTCATATTTGCTATA
ACTTTGCATGAGCAAATCAAGTTCAGCTGGAGAAAGAAATGGTGATCTCT
TTTATAAGATATTGCCATAGTCACTCGTAATGTCTGCGCTCCATTGATAA

TGCCTTTTTAAAGTCATGGTGTATGTGCTTAACTCTGAGTTGGTCTACTC
GAAGTTGATTGACCCAACTCAGATCAGCTATTCTGAAACCGAAAACTCTG
AGTTTTTAATCTCTCGGTTAATCAACTCAGAGTTCAAGTTTAAACTCTGA
GTTGTTTGAACCTCCTTACTGAAACAGGCCCCAGGTTGGATCAAAGAAAC
TGAGGCTGGTAGTCGATGTCAACAGTTGACAATAAAATGGTAAGAATTGC
TTTTCCAATGGAAAAAAAAAATCAATTCAAAGAAAGCACGTTGTCCCAGA
GTTTCTCATCTATGTTCCTGGAGGACCACTAACACTGCATTGTTTGAATG
ACCTCTTTGTCTGTCACATCCATCACCGGTTTTTCAGTCACTGCTAATAA
GCTGATGATCTGAATACGTTGTGTTTGGTTAAGGAAACATGTAAAATGTG
CAGGGCTGGTGGTCCTCCAGGAAGGTGGTTGAGAAACCCTGCTTTAGCCA
ATTATTGCATATAATGATGAATCATTTGACTAGAGTCTCTGATTTTCCCA SNORD52E
TGAAGCTCTTCACAGTGCACTGAACGCATGCGGTACTGCTTATGGTGGAT
TTATTTTTGGCTTAAATATTCCTAAATGTTAAAATACCTTCACAGGTTAG
ATGAGTGAAATCAGTTGCCTTGTAGCAGATGCCATCTTTCAGCAAAAAAA
AGAGGTAATTACATAAAATGAAAGCCTTTTGTACATTATTGCATGTCATG SNORD52F
ATGAACCATTTGACTAGAGTCTCTGATTTACCCATGAAGCTCTTCACAGT
GCACTGAACACATGCTCTACAAAAAAAATGAGTGGTTTGCTGTGGCTTAA
ATGACCCTGAATGCTTAACTTTCACAGTTTGGATGAGAAAAACAGAGGCT
GCCAGGTAGCTGATGCCAACAATAGGCGACCAATGGGTAAGAATGGGTTT
TCGATAAATCTTTCAAAGATAGCATTTTTTCATTTATGCCATGTTATGAT SNORD52G
GAATCTTTAGACTAGAGTCTCTGATTTACTTATGAAGCTGTTCACAGTGC
ACTGTACACATGCAATACAAAATAATGAGTGGTTTGTTGTTACTTGAATG
ATCCTGAATGCTTAAATAACTTTCACAGTTTGGATCAAACTGAAGCTGCT
AGATAGCTGATGCCAACAATACACAGCCAATAGGTGAGAATGAGTTTAAG
CAAAATCTTTTAAAGAAATGCCATGTTATGATGAATCTTTAGACTAGAGT SNORD52H
CTCTGATTTCTCCATGAAGCTCTTCACAGTGCACTGAACACATGCAGTAC
TGCTTAAGACTAATTTGAAGTGGCTTGAATGTTTCTAAAATGTAATAACT
TTTATAGGTTGAATGGAACAAAAGCCTGCTGGCTGATTCTAAGTTGACAT
CCAATGGGTAAGAAGGATCGTCAATGAAAGCTTCACAAATTAGGATAAAT
TTACACAGCAATTTTTTTTATATATTTTTTGTAGTTTGCGCTTCTGAAAA
GTTTCACTTGTAAAAGTCAACACTGTCAAAAAGATCCTCATTCACAATTC
AGTGAAAAAAGCTGAAATTGAGCATTTCAGACTGGTAGTTGGTTTAAATG
TTGCACGTGAATTTTTTTTTTCAAGCTTACACTCATCGTAACCATTATTT
CCAATAAATTGATCTCTTGAATAGAAACGTATGTGAATTTTCAGGCCCCA
AGTCACATGTAAATGAACAAAAATATTTTTATTCATTTTCAAAAGCATTT
GTCCATTAATGGTCAGTTGCAGGTAATGAGTATTCATAAAACTAGAGTTT
CTTTCTTTTCAAGGTTTTCAACTACAGGGTTCGAAGCCAACATGGATGTC
TTGGATTCATGAATCTGCTTAAAGGTTGGCGTTTTGACCTAACTGAGCTG
ATTCTCTATTGTGCAGTGATGATTCCATAGACTAGAGTCTCTGATAACAA SNORD52I (SNORD52)
CATGAAGATCTACACATTTCTCTGAGCACAGCTTTTTTATTATGAGCATG
TATTTGTTACTAAGAACTTTTGATTATTGGATTTTATTTTGTTTTTAGGT
TGACGAGATGCCAGTGAAGGGGTGCTGCATACACATGAATGCGGAGCCGT
GTAAATACTGGTTCAGTAAACCTAGCTCTTTTTCTTGTGTTAATGATGAC SNORDX
TGCGTTTGTTCGGTTCCACTGAAAGTCAATGAAGTAACTTTCTCGCGGCG
CACTGAAACACTGACAGACAATGCATGAACTTTTATTTCATGTCTTGTTG
TCTAATACAATTTTGTTTTCTTTCAAGGATGTGACCGAAGAGGAGCTGTG
ACTGGTTTGCTGCACAATGGTGAATTAAAATGATATACATCAATTTAAAT
GCTTGTTTAGTGGCATTCAATAGACTGTTTGAGTTAATTTTCTCATTTTC
AGTTTTATAAACTGCTTTTTCAACAGAGTATTGCTTTTAATAAACATCAG
AACTAAAGTGAAAAGTTGTCCATGTTTGTTTTTTGCATGAAATAAAAATA
TTGCGTATGGTTTTAACTTTTAGTCAATATAACATTCCTAGGCAGAAATT

Figure S9. Danio rerio FP101887. Exons (in blue) annotated by alignment with transcript FP101887.1-204. The sequences of the SNORDs are underlined from box C to box D (both included).

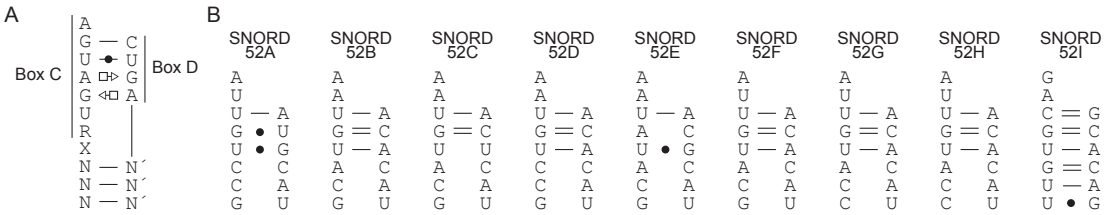

A
Box C
Box D
B
SNORD 52A
SNORD 52B
SNORD 52C
SNORD 52D
SNORD 52E
SNORD 52F
SNORD 52G
SNORD 52H
SNORD 52I

Figure S10

List of previously published theses for PhD in Aquaculture / PhD in Aquatic Biosciences, Nord University

No. 1 (2011)
PhD in Aquaculture
Chris André Johnsen
Flesh quality and growth of farmed Atlantic salmon (*Salmo salar* L.) in relation to feed, feeding, smolt type and season
ISBN: 978-82-93165-00-2

No. 2 (2012)
PhD in Aquaculture
Jareeporn Ruangsri
Characterization of antimicrobial peptides in Atlantic cod
ISBN: 978-82-93165-01-9

No. 3 (2012)
PhD in Aquaculture
Muhammad Naveed Yousaf
Characterization of the cardiac pacemaker and pathological responses to cardiac diseases in Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-02-6

No. 4 (2012)
PhD in Aquaculture
Carlos Frederico Ceccon Lanes
Comparative Studies on the quality of eggs and larvae from broodstocks of farmed and wild Atlantic cod
ISBN: 978-82-93165-03-3

No. 5 (2012)
PhD in Aquaculture
Arvind Sundaram
Understanding the specificity of the innate immune response in teleosts: Characterisation and differential expression of teleost-specific Toll-like receptors and microRNAs
ISBN: 978-82-93165-04-0

No. 6 (2012)
PhD in Aquaculture
Teshome Tilahun Bizuayehu
Characterization of microRNA during early ontogeny and sexual development of Atlantic halibut (*Hippoglossus hippoglossus* L.)
ISBN: 978-82-93165-05-7

No. 7 (2013)
PhD in Aquaculture
Binoy Rajan
Proteomic characterization of Atlantic cod skin mucosa – Emphasis on innate immunity and lectins
ISBN: 978-82-93165-06-04

No. 8 (2013)
PhD in Aquaculture
Anusha Krishanthi Shyamali Dhanasiri
Transport related stress in zebrafish: physiological responses and bioremediation
ISBN: 978-82-93165-07-1

No. 9 (2013)
PhD in Aquaculture
Martin Haugmo Iversen
Stress and its impact on animal welfare during commercial production of Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-08-8

No. 10 (2013)
PhD in Aquatic Biosciences
Alexander Jüterbock
Climate change impact on the seaweed *Fucus serratus*, a key foundational species on North Atlantic rocky shores
ISBN: 978-82-93165-09-5

No. 11 (2014)
PhD in Aquatic Biosciences
Amod Kulkarni
Responses in the gut of black tiger shrimp *Penaeus monodon* to oral vaccine candidates against white spot disease
ISBN: 978-82-93165-10-1

No. 12 (2014)
PhD in Aquatic Biosciences
Carlo C. Lazado
Molecular basis of daily rhythmicity in fast skeletal muscle of Atlantic cod *(Gadus morhua)*
ISBN: 978-82-93165-11-8

No. 13 (2014)
PhD in Aquaculture
Joanna Babiak
Induced masculinization of Atlantic halibut (*Hippoglossus hippoglossus* L.): towards the goal of all-female production
ISBN: 978-82-93165-12-5

No. 14 (2015)
PhD in Aquaculture
Cecilia Campos Vargas
Production of triploid Atlantic cod: A comparative study of muscle growth dynamics and gut morphology
ISBN: 978-82-93165-13-2

No. 15 (2015)
PhD in Aquatic Biosciences
Irina Smolina
Calanus in the North Atlantic: species identification, stress response, and population genetic structure
ISBN: 978-82-93165-14-9

No. 16 (2016)
PhD in Aquatic Biosciences
Lokesh Jeppinamogeru
Microbiota of Atlantic salmon (*Salmo salar L.*), during their early and adult life
ISBN: 978-82-93165-15-6

No. 17 (2017)
PhD in Aquatic Biosciences
Christopher Edward Presslauer
Comparative and functional analysis of microRNAs during zebrafish gonadal development
ISBN: 978-82-93165-16-3

No. 18 (2017)
PhD in Aquatic Biosciences
Marc Jürgen Silberberger
Spatial scales of benthic ecosystems in the sub-Arctic Lofoten-Vesterålen region
ISBN: 978-82-93165-17-0

No. 19 (2017)
PhD in Aquatic Biosciences
Marvin Choquet
Combining ecological and molecular approaches to redefine the baseline knowledge of the genus Calanus in the North Atlantic and the Arctic Oceans
ISBN: 978-82-93165-18-7

No. 20 (2017)
PhD in Aquatic Biosciences
Torvald B. Egeland
Reproduction in Arctic charr – timing and the need for speed
ISBN: 978-82-93165-19-4

No. 21 (2017)
PhD in Aquatic Biosciences
Marina Espinasse
Interannual variability in key zooplankton species in the North-East Atlantic: an analysis based on abundance and phenology
ISBN: 978-82-93165-20-0

No. 22 (2018)
PhD in Aquatic Biosciences
Kanchana Bandara
Diel and seasonal vertical migrations of high-latitude zooplankton: knowledge gaps and a high-resolution bridge
ISBN: 978-82-93165-21-7

No. 23 (2018)
PhD in Aquatic Biosciences
Deepti Manjari Patel
Characterization of skin immune and stress factors of lumpfish, *Cyclopterus lumpus*
ISBN: 978-82-93165-21-7

No. 24 (2018)
PhD in Aquatic Biosciences
Prabhugouda Siriyappagouder
The intestinal mycobiota of zebrafish – community profiling and exploration of the impact of yeast exposure early in life
ISBN: 978-82-93165-23-1

No. 25 (2018)
PhD in Aquatic Biosciences
Tor Erik Jørgensen
Molecular and evolutionary characterization of the Atlantic cod mitochondrial genome
ISBN: 978-82-93165-24-8

No. 26 (2018)
PhD in Aquatic Biosciences
Yangyang Gong
Microalgae as feed ingredients for Atlantic salmon
ISBN: 978-82-93165-25-5

No. 27 (2018)
PhD in Aquatic Biosciences
Ove Nicolaisen
Approaches to optimize marine larvae production
ISBN: 978-82-93165-26-2

No. 28 (2019)
PhD in Aquatic Biosciences
Qirui Zhang
The effect of embryonic incubation temperature on the immune response of larval and adult zebrafish (*Danio rerio*)
ISBN: 978-82-93165-27-9

No. 29 (2019)
PhD in Aquatic Biosciences
Andrea Bozman
The structuring effects of light on the deep-water scyphozoan *Periphylla periphylla*
ISBN: 978-82-93165-28-6

No. 30 (2019)
PhD in Aquatic Biosciences
Helene Rønquist Knutsen
Growth and development of juvenile spotted wolffish (*Anarhichas minor*) fed microalgae incorporated diets
ISBN: 978-82-93165-29-3

No. 31 (2019)
PhD in Aquatic Biosciences
Shruti Gupta
Feed additives elicit changes in the structure of the intestinal bacterial community of Atlantic salmon
ISBN: 978-82-93165-30-9

No. 32 (2019)
PhD in Aquatic Biosciences
Peter Simon Claus Schulze
Phototrophic microalgal cultivation in cold and light-limited environments
ISBN: 978-82-93165-31-6

No. 33 (2019)
PhD in Aquatic Biosciences
Maja Karoline Viddal Hatlebakk
New insights into Calanus glacialis and C. finmarchicus distribution, life histories and physiology in high-latitude seas
ISBN: 978-82-93165-32-3

No. 34 (2019)
PhD in Aquatic Biosciences
Arseny Dubin
Exploration of an anglerfish genome
ISBN: 978-82-93165-33-0

No. 35 (2020)
PhD in Aquatic Biosciences
Florence Chandima Perera Willora Arachchilage
The potential of plant ingredients in diets of juvenile lumpfish (*Cyclopterus lumpus*)
ISBN: 978-82-93165-35-4

No. 36 (2020)
PhD in Aquatic Biosciences
Ioannis Konstantinidis
DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication
ISBN: 978-82-93165-36-1

No. 37 (2021)
PhD in Aquatic Biosciences
Youngjin Park
Transcriptomic and cellular studies on the intestine of Atlantic salmon
Discovering intestinal macrophages using omic tools
ISBN: 978-82-93165-34-7

No. 38 (2021)
PhD in Aquatic Biosciences
Purushothaman Kathiresan
Proteomics of early embryonic development of zebrafish (Danio rerio)
ISBN: 978-82-93165-37-8

No. 39 (2021)
PhD in Aquatic Biosciences
Valentin Kokarev
Macrobenthic communities of sub-Arctic deep fjords: composition, spatial patterns and community assembly
ISBN: 978-82-93165-38-5

No. 40 (2021)
PhD in Aquatic Biosciences
Aurélien Delaval
Population genomics of a critically endangered data-deficient elasmobranch, the blue skate *Dipturus batis*
ISBN: 978-82-93165-39-2

No. 41 (2021)
PhD in Aquatic Biosciences
Isabel Sofía Abihssira García
Environmental impact of microplastics in relation to Atlantic salmon farming
ISBN: 978-82-93165-40-8

No. 42 (2022)
PhD in Aquatic Biosciences
Yousri Abdelmutalab Ahmed Abdelhafiz
Insights into the bacterial communities of Nile tilapia – core members and intergenerational transfer
ISBN: 978-82-93165-41-5

No. 43 (2022)
PhD in Aquatic Biosciences
Fredrik Ribsskog Staven
Interaction studies on lumpfish exposed to Atlantic salmon: behavioural observations and the underlying physiological and neurobiological mechanisms
ISBN: 978-82-93165-42-2

No. 44 (2022)
PhD in Aquatic Biosciences
Solveig Lysfjord Sørensen
Influence of feed ingredients and additives on mucosal health with focus on the intestine of Atlantic salmon (*Salmo salar*)
ISBN: 978-82-93165-43-9

No. 45 (2022)
PhD in Aquatic Biosciences
Apollo Marco Dalonos Lizano
Examining challenges in species-level taxonomy among Calanus copepods in the Northern seas using genome and transcriptome data
ISBN: 978-82-93165-44-6

Zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*), are model organisms for human biology and diseases. Even though an abundance of information about the DNA and RNA of zebrafish and medaka is readily available, a functional understanding of many biological mechanisms is lacking, especially with respect to the early stages of embryonic development. Ribosomal RNA (rRNA) structure and composition are highly conserved across most organisms with variable modifications such as ribose methylations. Zebrafish has been shown to possesses two types of rRNA, one of which is only observed during early embryonic development. The purpose of this study was to investigate the structure and modification profile of the two rRNA types and to compare the results with that of medaka, a distant evolutionary relative of zebrafish. We used the sequencing-based RiboMeth-seq method to study the ribose methylation (*2'-O*-Me) profile in zebrafish and medaka during select stages of embryonic development. The results showed that the two rRNA types in zebrafish differed in structure and modification profile while medaka only expressed one rRNA type that was most similar to the rRNA of other known vertebrates. We also propose a new investigative strategy to study the RNA that guide *2'-O*-Me modifications. Overall, this study leads to question whether the two rRNA types seen in zebrafish is restricted to a specific sub-group of fish and if medaka could indeed be the better model to study early embryonic development.

NORD University

ISBN: 978-82-93165-45-3
Trykk: Trykkeriet, Nord universitet
www.nord.no