Analiza ekspresije exosc10/Exosc10 u tkivima zebrice (Danio rerio) i embrionalnim fazama razvoja

Oreški, Leticija

Sveučilište u Zagrebu
Prehrambeno-biotehnološki fakultet
Diplomski studij Molekularna biotehnologija

Leticija Oreški
1028/MB

Analiza ekspresije *exosc10*/Exosc10 u tkivima zebrice (*Danio rerio*) i embrionalnim fazama razvoja

DIPLOMSKI RAD

Predmet: Molekularna biologija
Mentor: Doc. dr. sc. Igor Stuparević

Zagreb, 2019.

TEMELJNA DOKUMENTACIJSKA KARTICA

Diplomski rad

Sveučilište u Zagrebu
Prehrambeno – biotehnološki fakultet
Diplomski sveučilišni studij Molekularna biotehnologija
Zavod za kemiju i biokemiju
Laboratorij za biokemiju
Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Biotehnologija

Analiza ekspresije *exosc10*/Exosc10 u tkivima zebrice (*Danio rerio*) i embrionalnim fazama razvoja

Leticija Oreški, 1028/MB

Sažetak:

Cilj ovog rada bio je provesti analizu ekspresije gena *exosc10* na razini mRNA na 8 različitih tkiva odraslih ženskih i muških jedinki zebrice (*Danio rerio*) i 9 faza embrionalnog razvoja u različitim vremenskim intervalima nakon oplodnje, upotrebom metode PCR-a u realnom vremenu. Exosc10 egzoribonukleaza je katalitička podjedinica RNA egzosoma koji se nalazi u svim eukariotskim stanicama i ima ulogu u procesiranju, kontroli kvalitete i degradaciji RNA u jezgri i citoplazmi. Kako bi se ti rezultati povezali s razinom proteina, provedena je Western blot metoda na lizatima muških i ženskih tkiva zebrice, korištenjem 2 različita primarna antitijela za koja je predviđeno da rade na zebrici. Provedeno je i detaljno pretraživanje genomskih databaza *GeneBank* i *Ensembl* kako bi se konstruiralo filogenetsko stablo i matrica homologije. To je, uz pomoć sintenijske analize, pokazalo da za Exosc10 kod zebrice postoji ortolog kod čovjeka, da su ti protein 60% identični među kralješnjacima te da je on evolucijski konzerviran među eukariotima. *exosc10* je konstitutivno i diferencijalno eksprimiran u svim testiranim tkivima, bez značajne razlike između dva spola. Gonade oba spola i ranije faze embrionalnog razvoja pokazuju najveću ekspresiju. Za gonade su rezultati ekspresije na razini proteina u korelaciji s rezultatima na razini transkripta, ali bi se analiza trebala daje optimizirati kako bi se dobila preciznija interpretacija. Oba testirana antitijela su prvi puta korištena na zebrici te su se pokazala funkcionalnim. Konačno, analiza ekspresije gena *c1d*, *mpp6*, *dis3*, *dis3l* i *dis3l2* na razini mRNA je također provedena, s obzirom na to da ti geni kodiraju za podjedinice RNA egzosoma te su tako povezani sa *exosc10*. Ti geni su također konstitutivno i diferencijalno eksprimirani u svim testiranim tkivima, ali je primijećena značajna razlika između spolova za *c1d.*

Ključne riječi: Exosc10 egzoribonukleaza; zebrica (*Danio rerio*); PCR u realnom vremenu; Western blot, bioinformatika
Rad sadrži: 32 stranice, 11 slika, 6 tablica, 31 literaturni navod, 0 priloga
Jezik izvornika: engleski
Rad je u tiskanom i elektroničkom obliku pohranjen u knjižnici Prehrambeno biotehnološkog fakulteta Sveučilišta u Zagrebu, Kačićeva 23, 10 000 Zagreb
Mentor: Doc. dr. sc. Igor Stuparević
Pomoć pri izradi: Jovica Lončar, dr. sc.
Datum obrane: srpanj, 2019.

BASIC DOCUMENTATION CARD

Master thesis

University of Zagreb
Faculty of Food Technology and Biotechnology
University graduate study Molecular Biotechnology
Department of Chemistry and Biochemistry
Laboratory for Biochemistry
Scientific area: Biotechnical Sciences
Scientific field: Biotechnology

Analysis of *exosc10*/Exosc10 expression in zebrafish (*Danio rerio*) tissues and embryonic developmental stages

Leticija Oreški, 1028/MB

Abstract:

The aim of this thesis was to perform *exosc10* expression analysis at the level of mRNA in zebrafish (*Danio rerio*) using quantitative real-time PCR (qRT-PCR) with eight different tissues of both female and male fish and with embryonic developmental stages in nine different time intervals after fertilization. Exosc10 exoribonuclease is a catalytic subunit of the RNA exosome complex in eukaryotic cells and is important in processing, quality control and degradation of RNA in both the nucleus and cytoplasm of eukaryotes. To correlate this to protein expression levels, Western blot (WB) analysis was performed using both male and female zebrafish tissue lysates, using two different primary antibodies predicted to work with zebrafish. A thorough bioinformatics research of genomic databases *GeneBank* and *Ensembl* was performed to construct a phylogenetic tree and a homology matrix. All of this, along with synteny analysis, indicated that the protein is orthologous to its human counterpart, that it shares about 60% identity among vertebrates and that it is evolutionarily conserved in eukaryotes. *exosc10* is constitutively and differentially expressed in all of the tissues tested, without notable difference between two genders. Gonads of both genders and earlier embryonic developmental stages show the highest expression. For gonads, results for protein expression are in correlation with qRT-PCR results, but the analysis should be optimized for further interpretation. Both antibodies used were, for the first time, proved to work on zebrafish. Lastly, gene expression analysis at the level of mRNA for *c1d*, *mpp6*, *dis3*, *dis3l* and *dis3l2* was performed, since they work closely with *exosc10* within the exosome. All genes are differentially and constitutively expressed in all of the tissues tested, with notable difference between two genders for *c1d*.

Keywords: Exosc10 exoribonuclease; zebrafish (*Danio rerio*); qRT-PCR; Western blot; bioinformatics
Thesis contains: 32 pages, 11 figures, 6 tables, 31 references, 0 supplements
Original in: English
Thesis is in printed and electronic form deposited in the library of the Faculty of Food Technology and Biotechnology, University of Zagreb, Kačićeva 23, 10 000 Zagreb
Mentor: Igor Stuparević, Assistant Professor, PhD
Technical support and assistance: Jovica Lončar, Research Associate
Defence date: July, 2019

UNIVERSITÉ D'ORLÉANS
UNIVERSITÉ DE ZAGREB

U. F. R. Faculté des Sciences-Faculté de Nutrition et Biotechnologie-Faculté des Sciences

MASTER SCIENCES et TECHNOLOGIE
Mention : BIOLOGIE-BIOCHIMIE

Spécialité : Techniques Bio-Industrielles

Stage report

Analysis of *exosc10*/Exosc10 expression in zebrafish (*Danio rerio*) tissues and embryonic developmental stages

by

LETICIJA OREŠKI

(December, 2018 - May, 2019)

Organism (company): Institut Ruđer Bošković, Division for Marine and Environmental Research, Laboratory for Molecular Ecotoxicology & University of Zagreb, Faculty of Food Technology and Biotechnology, Laboratory for Biochemistry

Place of the stage: Bijenička 54, 10000 Zagreb & Pierottijeva 6, 10000 Zagreb

Supervisor of the stage: Igor Stuparević, Assistant Professor; Jovica Lončar, Research Associate; Tvrtko Smital, Senior Scientist

Referees: Steeve H. Thany, Professor; Reno Hrašćan, Associate Professor

Date of defense: 1 July, 2019

PRESENTATION OF THE HOST ORGANIZATION

The Ruđer Bošković Institute (RBI) was founded by Ivan Supek in 1950 and is currently regarded as Croatia's leading scientific institute in natural and biomedical sciences as well as marine and environmental research. It is the largest scientific research center in Croatia which employs more than 5% of scientists and contains 50% of all research facilities in Croatia. Today it consists of as many as 80 laboratories which have a multidisciplinary character and are involved in many international research projects.

Division for Marine and Environmental Research (DMER) is the largest interdisciplinary division of RBI. One of its laboratories is Laboratory for Molecular Ecotoxicology (LME), which focuses on understanding cellular defense mechanisms in aquatic organisms and is headed by Tvrtko Smital, PhD. Two groups perform their research in these laboratories; Tvrtko Smital's group, which is focused on molecular ecotoxicology and basic ecotoxicological research. Their goal is to understand:

- Molecular base and role of critical cellular defense and/or detoxification mechanisms in aquatic organisms (one of which being zebrafish, *Danio rerio*)
- Interactions of cellular defense mechanisms with both classical and emerging environmental contaminants

and Marta Popović's group, which focuses on DNA damage.

ABSTRACT

The aim of this Master thesis was to perform tissue-specific *exosc10* expression analysis at the level of mRNA in zebrafish (*Danio rerio*) using quantitative real-time PCR (qRT-PCR) analysis. This analysis was done with eight different tissues; liver, intestine, gonads, eyes, gills, brain, kidney and muscle of both female and male fish and with embryonic developmental stages in nine different time intervals after fertilization, which are referred to as "hpf" (hours post fertilization); 1, 4, 6, 12, 24, 48, 72, 96 and 120 hpf. The importance of this lies in function of Exosc10 exoribonuclease, which is a catalytic subunit of the RNA exosome complex in eukaryotic cells and is important in processing, quality control and degradation of nearly all classes of RNA in both the nucleus and cytoplasm of eukaryotic organisms. No such research has been conducted on zebrafish so far. To correlate this to protein expression levels, Western blot (WB) analysis was performed using both male and female zebrafish lysates of different tissues mentioned above, using two different primary antibodies predicted to work with zebrafish. A thorough bioinformatics research of genomic databases *GeneBank* and *Ensembl* was performed to obtain data for phylogenic and synteny analysis of *exosc10*/Exosc10. We constructed a phylogenetic tree and a homology matrix and all of this with synteny analysis indicated that the protein is orthologous to its human counterpart, that it shares about 60% identity among vertebrates and that it is evolutionarily conserved in eukaryotes. Gene expression analysis shows that *exosc10* is constitutively and differentially expressed in all of the tissues tested, without notable difference between two genders. Gonads of both genders and earlier embryonic developmental stages show the highest expression, latter being due to maternal transfer of transcripts. For gonads, results for protein expression are in correlation with qRT-PCR results, but the analysis should be optimized for further interpretation. Both antibodies used were, for the first time, proved to work on zebrafish. Lastly, tissue-specific gene expression analysis at the level of mRNA for 5 other genes which work closely with *exosc10* within the RNA exosome complex was also performed; *c1d*, *mpp6*, *dis3*, *dis3l* and *dis3l2*. Dis3 proteins being ribonucleases and C1d and Mpp6 being cofactors of the exosome. All genes are differentially and constitutively expressed in all of the tissues tested, with notable difference between two genders for C1d. Future work will focus in determining Exosc10 protein expression levels of embryonic developmental stages as well as protein expression levels of ribonuclease Dis3 and exosome cofactors C1d and Mpp6.

Keywords: Exosc10 exoribonuclease, zebrafish (*Danio rerio*), qRT-PCR, Western blot, bioinformatics

ABSTRAIT

L'objectif de cette thèse était de réaliser une analyse d'expression *exosc10* spécifique au tissu au niveau de l'ARNm chez le poisson zèbre (*Danio rerio*) à l'aide d'une analyse quantitative par PCR en temps réel (qRT-PCR). Cette analyse a été réalisée avec huit tissus différents; foie, intestin, gonades, yeux, branchies, cerveau, reins et muscles de poissons mâles et femelles et aux stades de développement embryonnaire à neuf intervalles de temps différents après la fécondation, appelés «hpf» (heures après la fécondation); 1, 4, 6, 12, 24, 48, 72, 96 et 120 hpf. L'importance de cet aspect réside dans l'exoribonucléase Exosc10, qui est une sous-unité catalytique du complexe exosome d'ARN dans les cellules eucaryotes et joue un rôle important dans le traitement, le contrôle de la qualité et la dégradation de presque toutes les classes d'ARN dans le noyau et le cytoplasme des organismes eucaryotes. Aucune recherche de ce type n'a encore été menée sur le poisson zèbre. Pour corréler cela aux niveaux d'expression des protéines, une analyse Western blot (WB) a été réalisée en utilisant des lysats de poisson zèbre mâles et femelles de différents tissus, en utilisant deux anticorps primaires différents supposés fonctionner avec le poisson zèbre. Une recherche bioinformatique approfondie des bases de données génomiques *GeneBank* et *Ensembl* a été réalisée afin d'obtenir des données pour l'analyse phylogénique et synténique de *exosc10*/Exosc10. Nous avons construit un arbre phylogénétique et une matrice d'homologie et tout ceci avec une analyse synténique indique que la protéine est orthologue à son équivalent humain, qu'elle partage environ 60% de l'identité chez les vertébrés et qu'elle est conservée au cours de l'évolution chez les eucaryotes. L'analyse de l'expression génique montre que *exosc10* est exprimé de manière constitutive et différentielle dans tous les tissus testés, sans différence notable entre les deux sexes. Les gonades des deux sexes et des stades de développement embryonnaire antérieurs montrent la plus haute expression, ce dernier étant dû au transfert maternel des transcrits. Pour les gonades, les résultats pour l'expression des protéines sont en corrélation avec les résultats de la qRT-PCR, mais l'analyse doit être optimisée pour une interprétation plus poussée. Pour la première fois, les deux anticorps utilisés ont démontré leur efficacité sur le poisson zèbre. Enfin, une analyse de l'expression génique spécifique des tissus au niveau de l'ARNm pour 5 autres gènes qui travaillent étroitement avec *exosc10* dans le complexe ARN exosome a également été réalisée; *c1d, mpp6, dis3, dis3l* et *dis3l2*. Les protéines Dis3 sont des ribonucléases et C1d et Mpp6 sont des cofacteurs de l'exosome. Tous les gènes sont exprimés de manière différentielle et constitutive dans tous les tissus testés, avec une différence notable entre les deux sexes pour C1d. Les travaux futurs consisteront à déterminer les niveaux d'expression de la protéine Exosc10 aux stades de développement embryonnaire ainsi que du taux d'expression de la ribonucléase Dis3 et des cofacteurs de l'exosome C1d et Mpp6.

Mots clés: Exosc10 exoribonucléase, poisson zèbre (*Danio rerio*), qRT-PCR, Western blot, bioinformatique

ACKNOWLEDGEMENTS

I would like to thank Igor Stuparević, PhD, for reaching out and giving me a great opportunity to work at RBI by introducing me to great people. I would also like to thank the whole team at Laboratory for Molecular Ecotoxicology for kindly making me a part of their lab for a while and for a truly great working environment; especially Jovica Lončar, PhD, for being a great mentor with a systematic and critical approach, for his patience with teaching and his general friendliness.

Great thanks to my parents, who have given me the possibility of graduating by directing me and my sisters early on and supporting us financially, which I know was mission impossible. ☺

Lastly, I would like to thank Matej for being my best friend and partner since the beginning; for all the late-night conversations, midnight yoghurts and words of comfort in hard times. It was all so much easier thanks to you.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Aim of the thesis

In order to obtain a more detailed insight into *exosc10*/Exosc10 in zebrafish as a model organism, the main goals of this study were to:

- perform phylogenetic and synteny analysis of *exosc10*/Exosc10 to position zebrafish gene/protein against its ortholog in human
- establish mRNA and protein expression profile of *exosc10*/Exosc10 in zebrafish tissues and embryonic developmental stages (liver, intestine, gonads, eyes, gills, brain, kidney, muscle); both male and female
- establish mRNA expression profile of *c1d, mpp6, dis3, dis3l* and *dis3l2*, that, besides *exosc10*, code for different subunits of the RNA exosome complex in eukaryotic cells
- compare tissue versus developmental stages expression profiles of *exosc10*/Exosc10 exoribonuclease in order to determine the relation of developmental stages to respective tissues and determine gender differences in expression patterns

1.2. Zebrafish (*Danio rerio*) as a model organism

Fish are the most numerous and phylogenetically diverse group of vertebrates, and today, scientists use them for research of fundamental processes in vertebrates, such as evolution, development and disease[1]. Over the last 40 years, zebrafish has become a popular and powerful model organism for research of vertebrate genetics, development, regeneration, and toxicology[2]. Zebrafish are small[3], tropical, freshwater fish species that originate from South Asia[2] and belong to a class of teleosts[3]. There are many reasons for their use in research; besides their small size[1,4], fertilization and development are external, which makes observations and manipulation of developing embryos easier[1]. Embryos develop rapidly (in 2±4 days), with a beating heart by 24 hours[3]. All developmental stages are easily accessible and transparent in color so that real-time imaging of developing pathologies is possible[5]. Zebrafish breed year-round, have a short generation time (3–5 months)[1] and are easy to raise[2]. A female can lay 200–300 eggs in one morning, every 5–7 days, if well maintained[4]. Two different genders are easy to distinct, as seen on *Figure 1*; the males are slender and torpedo-shaped, with black longitudinal stripes and usually a gold coloration on their underside and fins. Females are fat when carrying eggs and they have little, if any, gold on their undersides.

In contrast to other fish, which can be triploid or tetraploid, what makes genetic analysis difficult, zebrafish maintains the diploid state (genome size: 1700 Mbp)[3]. Zebrafish genome is therefore more complex than that of a human because zebrafish have 25 pairs of chromosomes compared to 23 pairs in humans[1,4], which is due to a whole-genome duplication (WGD) that occurred in fish, but not in mammals[4]. This is significant because where a mutation in a mammalian ortholog may cause embryonic lethality, a mutation in one of the zebrafish paralogs may give rise to viable animals[1,4], which allows better definition of gene function in mutant zebrafish[4]. The obvious disadvantage of the gene duplication, however, is that certain information may be hard to extrapolate to other vertebrates such as mammals[4]. Nonetheless, zebrafish are highly used for comparative studies with other fish and other vertebrates, such as mouse and human[3].

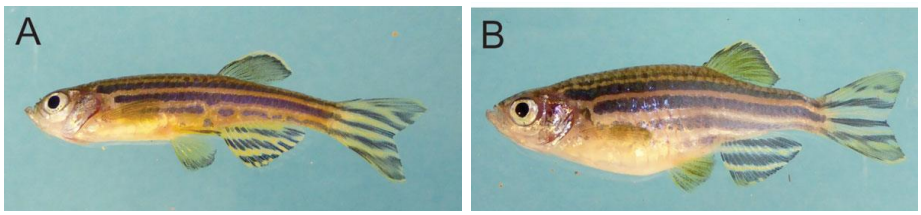


Fig. 1 Adult zebrafish with classical striping and short fins. (**A**) A male zebrafish; (**B**) a female zebrafish. The male has a thinner underside than the female, whose underside is round because she is carrying eggs[2].

1.3. Exosc10 exoribonuclease

Exosc10 (also called Rrp6 in *Saccharomyces cerevisiae* or PM/Scl-100 in *Homo sapiens*)[6,7] is a catalytic subunit of the multimeric exosome which can also function alone by binding to RNA via its C-terminal domain[7]. Exosc10 belongs to a large superfamily of nucleases, called DEDD[8,9]. All members of the DEDD superfamily share the catalytic DEDD domain, in which the active site consists of four conserved acidic amino acid residues (three aspartate and one glutamate, hence the name of the domain and the superfamily)[8]. Some DEDD family members include an additional conserved histidine or tyrosine residue and are designated DEDD-H or DEDD-Y, respectively. Human and yeast Exosc10 are members of the DEDD-Y family[9]. Exosc10 consists of the following domains: an N-terminal PMC2NT domain, a DEDD domain that catalyzes distributive 3'-5' exoribonuclease activity, an HRDC domain

(helicase and RNase D C-terminal domain)[10,6], a predicted HRDC2 domain, and a C-terminal domain (CTD)[10]. The PMC2NT domain mediates the interaction of Exosc10 with its partner protein, a nuclear cofactor C1d (in yeast called Rrp47)[10,8]. The conserved DEDD domain has an active site exposed to solvent which coordinates two divalent metal cations (preferably manganese) which are involved in activating a water molecule which attacks the last phosphodiester bond in the substrate[10,6]. The HDRC domain is crucial for full activity of the protein *in vitro*. Exosc10 protein is a distributive 3′–5′ exoribonuclease that cleave RNA by hydrolysis and binds near the 'cap' of the exosome[8]. Its C-terminal domain binds RNA and activates the nuclear RNA exosome[11]. Structured substrates are eligible for hydrolysis, provided there is a single stranded region upstream of the structure[8]. Human version of the protein is able to degrade more structured substrates than its yeast counterpart. Unlike yeast Rrp6, which is strictly nuclear, human EXOSC10 is also found in the nucleoplasm, nucleoli and cytoplasm[12].

1.4. The eukaryotic RNA exosome

Most RNAs in eukaryotic cells are produced as precursors which need to be processed at either 3' and/or 5' end to generate the mature transcript[13]. Additionally, many transcripts are degraded when recognized as faulty or as a part of normal recycling[13,14]. The eukaryotic RNA exosome is an essential, multi-subunit ribonuclease complex that participates in the processing, quality control and degradation of nearly all classes of RNA in both the nucleus and cytoplasm of eukaryotic organisms[10,15]. Its function relies on multiple cofactors, some of which will be discussed later[10,8]. Substrates of the exosome are versatile and include ribosomal RNA (rRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA)[14,16] and most of what we know today about the exosome has been studied in yeast *Saccharomyces cerevisiae*[14]. The eukaryotic exosome of yeast is composed of ten to eleven subunits, which can be divided into two major groups based on their structure and function. The first group consists of nine proteins with molecular masses of 20-50 kDa which form the exosome 'core' (Exo9)[8]. The 'core' is probably, evolutionarily speaking, the oldest and most conserved part of the exosome[10,8]. It has a central channel only wide enough for single-stranded RNA (ssRNA) to enter, but it has no ribonuclease activity[14]. Instead, it associates with Dis3 (also called Rrp44), an endoribonuclease and processive 3'-5' exoribonuclease and Exosc10, a distributive 3'-5' exoribonuclease. In this way, the 'core' coordinates Dis3 (Rrp44) and Exosc10 recruitment and modulates their activities[10,14]. A distributive enzyme

locates its target sites in a nucleic acid by a random, three-dimensional diffusion process; while a processive enzyme locates its recognition sites by one-dimensional scanning (sliding) process along the substrate molecule[17]. These catalytic subunits belong to the second group of exosome components and seem to influence each other's activities. Both Exosc10 and Dis3 (Rrp44) are large, multi-domain polypeptides with molecular masses around 100 kDa[8], but Exosc10 has an obligate partner, a cofactor named C1d (Rrp47)[18]. The 'core' is arranged in a two-layered ring in which the bottom layer is formed by six subunits and referred to as the 'hexamer', while the top layer consists of three RNA binding subunits and is called the 'cap'[14,8]. The size of this two-stacked ring is approximately 300 kDa[11]. Different isoforms of the exosome arise from different association combinations of the subunits with the 'core'[8]. This said, we can differentiate between cytoplasmic isoform of the exosome which consists of Exo9 complex and Dis3 (Rrp44) and a nuclear isoform, which consists of Exo9, Dis3 (Rrp44) and Exosc10[10,11]. An additional isoform has recently been hypothesized; a nucleolar isoform that contains Exo9 and Exosc10[10,19]. Subunit compositions of nuclear and cytoplasmic RNA exosomes from humans resemble that of yeast which are described above, but there are some notable differences[18]. In the yeast nucleus, the ubiquitous 11-subunit isoform of the exosome consisting of Exo9, Dis3 (Rrp44) and Exosc10 functions together with three conserved cofactors; C1d (Rrp47), Mtr4, and Mpp6. C1d (Rrp47) does not have enzymatic activity but together with Exosc10 forms a binding platform for recruiting Mtr4. Mtr4 is an essential nuclear helicase which assists the exosome by presenting it with suitably remodelled substrates that can be threaded with their unwound 3' end into the degradation core. Mpp6 is a small basic protein lacking recognizable domains, whose function is to promote the channelling of substrates from Mtr4 to the processive ribonuclease[20]. Graphical display of eukaryotic RNA exosome structure in yeast is shown on *Figure 2*. Recent studies have shown that proper termination and processing of ncRNAs by Exosc10 is especially important for genome-wide transcription regulation. In humans, loss of proper exosome activity may contribute to different pathologies such as autoimmune disease, cancer and neurological disorders. Additionally, inhibition of exosomal proteins can lead to a variety of diseases. Antibodies against many of the subunits of the human exosome, including Exosc10, have been characterized in autoimmune disorders affecting the connective tissue, specifically polymyositis (PM), scleroderma (Scl), and PM/Scl overlap syndrome. Other studies indicate that proper RNA processing is critical in neuromuscular development and maintenance. Furthermore, in a genome-wide drug-target screen, Exosc10 was found to be a potential target of the cell growth inhibitor 5-fluorouracil used in chemotherapy for the treatment of solid

tumors. These findings highlight the importance of understanding Exosc10 and other exosome components, to determine their role in human disease[6].

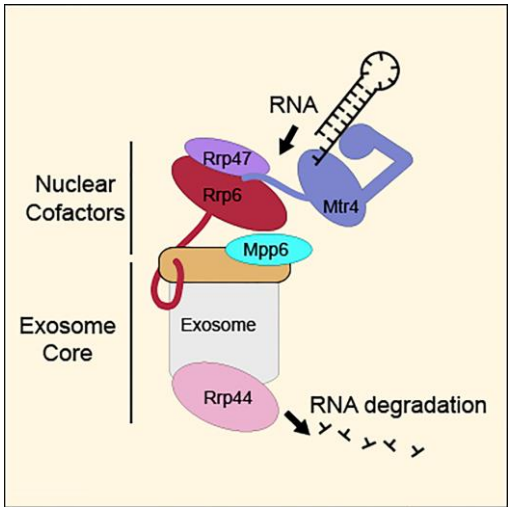


Fig. 2 Graphical display of eukaryotic RNA exosome structure in yeast. Exosome 'core' (Exo9) is shown in grey, Dis3 (Rrp44) ribonuclease in pink, ExoSC10 exoribonuclease in red, and nuclear cofactors C1d (Rrp47) in purple, Mpp6 in cyan and Mtr4 in blue[20].

1.5. Dis3 (Rrp44) exo- and endo- ribonuclease and its isoforms (Dis3l, Dis3l2)

Dis3 proteins are members of the RNR superfamily[8,21]. All Dis3 proteins have two cold-shock domains (CSDs), a central ribonuclease II domain (RNB) and a C-terminal S1 domain[10,8]. CSD and S1 non-specifically bind RNA, while RNB is a catalytic domain with processive exoribonuclease activity whose active site is built of four conserved aspartate residues[8,22]. Dis3 proteins attack 3′ ends of RNA and, once bound, they degrade it completely, releasing nucleoside 5′-monophosphates. The endonuclease activity of Dis3, however, resides in another domain, called the PIN domain (short for N-terminal Pilus-forming N-terminus), the active site of which is exposed to the solvent and consists of four acidic amino acid residues that coordinate two divalent metal cations (preferably manganese)[8]. Both structured and linear ssRNA are potential substrates of Dis3, and preference for substrates with phosphorylated 5′ termini has been reported. Although only one form of Dis3 is present in yeast, three different human homologues have been identified, Dis3, Dis3l and Dis3l2[14]. In all Dis3 homologues, two of the four amino acids of the PIN active site are mutated. As a result, Dis3 protein is an active endonuclease and Dis3l and Dis3l2 are not[8]. While Dis3 is largely limited to the

nucleus with only a small fraction in the cytoplasm, Dis3l and Dis3l2 are strictly cytoplasmic[14,8].

1.6. C1d (Rrp47), Mpp6 and other cofactors of the exosome

The RNA exosome associates with both nuclear and cytoplasmic cofactors to mediate quality control, processing and/or decay of a wide variety of transcripts[21]. The nuclear exosome interacts with the TRAMP complex to promote surveillance and degradation of aberrant RNA. It may also interact with the Nuclear Exosome Targeting (NEXT) complex to promote degradation of promoter upstream transcripts or with the NRD complex, which recruits the exosome to degrade or process certain snRNA and snoRNA polymerase II transcripts. Another complex it associates with is the SKI complex, which participates in 3'-5' mRNA degradation by associating with the translation apparatus. Many other cofactors have been identified in both human and yeast, but their function is less clear[10]. In the nucleus, the exosome associates with C1d and Mpp6, proteins which were studied in this work[21]. M-phase Phosphoprotein 6 (Mpp6) is important for the proper processing of the 3' ends of 5.8S rRNAs[12]. It binds to a 'cap' protein, Rrp40, and is required for the ability of Mtr4 (another cofactor) to extend the trajectory of an RNA entering the exosome core, which means that it promotes the channelling of substrates from the nuclear helicase to the processive ribonuclease, Exosc10[20]. Mpp6 was found to be an RNA-binding protein required for proper surveillance of mRNA and rRNA precursors, and the degradation of cryptic noncoding RNAs[12]. Both Mpp6 and C1d help to define targets which need to be processed and/or degraded by the exosome[10,14]. They are nuclear proteins enriched in nucleoli which have the ability of binding RNA. C1d binds structured substrates, while Mpp6 shows strong preference for pyrimidine-rich sequences. However, both proteins interact directly with Exosc10 and they are at least partly redundant[8].

2. MATERIALS & METHODS

2.1. Materials

2.1.1. Chemicals

Table 2.1.1. List of chemicals used in the study.

Compound	Catalog Number	Manufacturer
Nuclease free water	#R0582	Fermentas
Agarose	A9539	Sigma-Aldrich
Tris	1930809	Kemika
GeneRuler DNA Ladder Mix	#SM0333	Fermentas
Loading dye (6x)	#R0611	Fermentas
Absolute ethanol	505655	Kemika
TRIzol Reagent	15596-018	Thermo Fischer Scientific
Chloroform 99%	C24-32	Sigma-Aldrich

2.1.2. Commercial kits

Table 2.1.2. List of commercial kits used in the study.

Kit	Catalog number	Manufacturer
RNeasy Plus Mini Kit	74134	Qiagen
Monarch RNA Cleanup Kit (50 µg)	#T2040S	New England Biolabs
High Capacity cDNA Reverse Transcription Kit with Rnase Inhibitor	4374966	Applied Biosystems
Power SYBR Green PCR master mix	4367659	Applied Biosystems
DC™ Protein Assay Kit II	5000112	Bio-Rad

2.1.3. Enzymes

Table 2.1.3. List of enzymes used in the study.

Enzyme	Catalog number	Manufacturer
Hot Start *Taq* DNA Polymerase	MO495	New England Biolabs
Reverse transcriptase	4374966	Applied Biosystems
Dnase I, 2000 U/mL	M0303S	New England Biolabs

2.1.4. Instruments

Table 2.1.4. List of instruments used in the study.

Instrument	Manufacturer
T100 Thermal Cycler	Bio-Rad
qRT-PCR System AB 7300	Applied Biosystems
Eppendorf Thermomixer 5436	Eppendorf
Micro 120 (table centrifuge)	Hettich
Universal 32 R (centrifuge)	Hettich
Bio Vortex V1	Kisker
Tecan Infinite M200 microplate reader	Tecan
Electrophoresis power supply EPS 500/400	Pharmacia
BioSpec-nano	Shimadzu
ChemiDoc™ XRS+ System	Bio-Rad

2.1.5. qRT-PCR primers

Table 2.1.5. List of quantitative real time PCR primers used in the study.

Gene name	Forward 5'→3' (F) Reverse 5'→3' (R)	Tm [°C] (F/R)	GC % (F/R)	Amplicon length [bp]
exosc10	(F) GATGAGGCCTCCGGTGTAAG (R) TACCTCCACCCTTGCGATTC	60/59	60/55	103
c1d/rrp47	(F) ACGCCCTCAACTCAATGTTC (R) TATGTCCTGATTCTTTCCAGTTCT	58/57	50/37.5	97
mpp6	(F) CTCAAGGCCAAAGAAAACCACA (R) CTCCACGTCAGGATTGAAGC	60/59	45.45/55	105
dis3/rrp44	(F) AATGAGCACCACAGGGAGAC (R) GTACCATTTAGCAGCGACACG	60/60	55/52.38	96
dis3l	(F) TGGGGGAAAACGGTTAGAGC (R) TAAGTGGTTGCCCTGGACCT	60/61	55/55	105
dis3l2	(F) GTCTGCAGGGCTCAAGAGAG (R) TATCGGCTGACCCATCAGGA	60/60	60/55	104
ef1α	(F) CCTGGGAGTGAAACAGCTGATC (R) GCTGACTTCCTTGGTGATTTCC	60/60	55/50	100

- final concentration of all primers was 300 nM

2.1.6. Antibodies

Table 2.1.6. List of antibodies used in the study.

Antibody (ab)	Catalog number	Manufacturer	Epitope sequence	Epitope [aa]
EXOSC10 (B-8) mouse monoclonal primary ab	SC-374595	Santa Cruz	N.A.	1-300
Anti-EXOSC10 rabbit polyclonal primary ab	ab50558	Abcam	LDVPPALADFIHQQR-C	231-245
Goat anti-rabbit IgG (H+L), HRP conjugate secondary ab	SA00001-2	Proteintech Europe	N.A.	N.A.
m-IgGκ BP-HRP secondary ab	sc-516102	Santa Cruz	N.A.	N.A.

- N.A. = not available

2.2. Methods

2.2.1. Phylogenetic and synteny analysis

All protein sequences were retrieved from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) and the Ensembl (https://www.ensembl.org/index.html) databases. Zebrafish *exosc10* sequence was used to perform searches across the genomes of other vertebrate and invertebrate species using blastx algorithm. The alignment for the homology matrix was done using MegAlign 7.0 software from DNA Star Lasergene package (DNA Star, Madison, USA). Phylogenetic tree was built in PhyML 3.1 software using the maximum likelihood method (Guindon and Gascuel, 2003) and graphically edited in TreeGraph 2.0 software. BioEdit software version 7.0 was used for sequence editing, alignment display and calculation of sequence identities (Hall, 1999). Orthology predictions using synteny relationships between zebrafish and other genes were

made using a conserved synteny browser synchronized with genomes from the Ensembl database, Genomicus (http://www.genomicus.biologie.ens.fr/genomicus-92.01/cgi-bin/search.pl, Louis et al., 2013).

2.2.2. Total RNA isolation and cDNA synthesis for qRT-PCR

Total RNA was isolated from zebrafish tissues using either the Qiagen RNeasy Plus Mini Kit (Qiagen, Basel, Switzerland), according to manufacturer's protocol, or TRIZol reagent (Thermo Fischer Scientific). RNA quantity and purity ($OD_{260/280}$ and $OD_{260/230}$ - to check for proteins and alcohols, respectively) were determined using a BioSpec-nano spectrophotometer (Shimadzu), while integrity and possible genomic DNA (gDNA) contaminations were determined by agarose gel electrophoresis on a 1% agarose gel. gDNA contaminations were removed using DNase I (NEB, Massachusetts, USA). Cleanup after DNase I treatment was done using Monarch RNA Cleanup Kit (NEB, Massachusetts, USA). 1 µg of total RNA was transcribed to cDNA using High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems, Darmstadt, Germany) according to manufacturer's protocol.

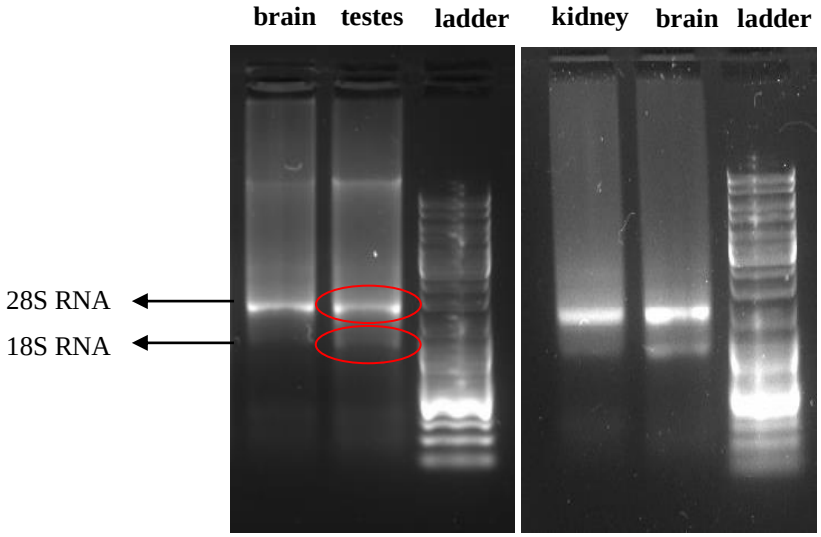


Fig. 3 RNA quality control for male zebrafish tissue samples after total RNA isolation (left) and total RNA isolation + DNase treatment (right) using gel electrophoresis (1% agarose gel). Far right – DNA ladder. 200 ng of RNA was loaded per lane.

The expected signal of two bright bands, upper being twice more intense than lower, corresponding to 28S and 18S RNA, respectively, can be seen for both samples. Samples which showed gDNA contamination (presence of another band on the top of the gel) were treated with DNase I. Another isolation was performed for degraded samples if the 2 bands mentioned were not visible and there was a smear on the bottom of the gel.

2.2.3. Quantitative real-time PCR analysis (qRT-PCR)

For the purpose of gene expression analysis, specific primers were designed using NCBI primer tool and the Lasergene software. Primers were purchased from Metabion (*exosc10*) and Macrogen (for all other genes). Real time PCR quantification (qRT-PCR) was performed on Applied Biosystems 7300 analyzer using Power SYBR Green PCR master mix (Applied Biosystems, Foster City, CA, USA) to analyze gene expression. Total reaction volume was 10 µL with: 5 µL of SYBR Green, 0.5 µL of each primer, 2.5 µL of nuclease free water and 1.5 µL of cDNA template. After initial heating to 95 °C for 10 minutes (polymerase activation), 40 cycles of amplification were carried out with denaturation at 95 °C for 15 seconds, annealing and elongation at 60 °C for 1 minute, all followed by the melting curve analysis: 95 °C for 15 seconds, 60 °C for 1 minute, 95 °C for 15 seconds, and finally, 60 °C for 15 seconds. Data were obtained and analyzed using Sequence Detector Software SDS 1.4 (Applied Biosystems). The obtained Ct values were exported to a Microsoft Excel spreadsheet and GraphPad Prism 5 software for further analysis. To generally compare gene expression between tissues and embryonic developmental stages, relative quantification was used as a method of choice. Target genes were therefore normalized to *ef1α* (housekeeping gene) to process qRT-PCR data (described in detail by Müller et al., 2002 and Simon, 2003, according to equation: MNE = $((E_{ref})^{Ct_{ref,\ mean}})/((E_{target})^{Ct_{target,\ mean}})$, where MNE stands for mean normalized expression, E_{ref} for housekeeping gene efficiency, E_{target} for target gene efficiency, $Ct_{ref,mean}$ for mean Ct value of housekeeping gene, and finally, $Ct_{target,mean}$ for mean Ct value of target gene. All data is presented as gene of interest expression relative to housekeeping gene expression multiplied by the factor of 10^6.

2.2.4. Western blotting

Adult zebrafish (male and female) were dissected to obtain different tissues (liver, intestine, gonads, eyes, gills, brain, kidney, muscle) for protein isolation for Western blot. Samples were obtained pooling tissues from two fish for each gender together to obtain a sufficient amount of material. Tissue lysis was performed in RIPA buffer containing 1% SDS (NaCl 150 mM, EDTA 1 mM, Tris 25 mM, NP-40 0.8%) with protease and phosphatase inhibitors. After centrifugation, supernatants were homogenized for 30 seconds, briefly sonicated and centrifuged at 1,000 *g* for 10 minutes at 4 °C. Total protein concentration in obtained lysate was quantified according to the manufacturer's instruction using the DC protein assay kit (Bio-Rad laboratories, CA, USA). 5 ug of each sample was separated by electrophoresis in homemade 5-18 % gradient gel or Mini-Protean TGX gel (BioRad, USA) and then transferred to the PVDF membrane (Millipore, MA, USA or BioRad, USA) by wet or semi-wet transfer. Blocking was performed for 2 hours in blocking solution containing low-fat milk (5%), 50 mM Tris, 150 mM NaCl and 0.05% Tween 20. Subsequently, membranes were washed and incubated for 1 hour with primary antibodies (1 : 3 000), followed by secondary antibodies (1 : 100 000). The proteins were visualized by chemiluminescence (Abcam, Cambridge, UK) and protein size was estimated by the use of protein marker (Thermo-Fischer Scientific, MA, USA). Histone H2B was used as a loading control (and to check for sample degradation).

3. RESULTS

3.1. Phylogenetic and syntenyc analysis of *exosc10*/Exosc10

Exosc10 is a catalytic subunit of the RNA exosome complex which can also function alone by binding to RNA via its C-terminal domain[7]. The eukaryotic RNA exosome complex is an essential, multi-subunit RNase complex which is involved in the processing, quality control and degradation of nearly all classes of RNA[10,15]. Substrates of the exosome are versatile and include many different types of RNA[14,16]. However, most of what we know today about the exosome and its function has been studied in yeast *Saccharomyces cerevisiae*[14]. Some research has been performed on human and mouse cell lines as well, but little is known about the exosome and its components in zebrafish (*Danio rerio*). Nonetheless, zebrafish is a popular model organism for research of vertebrate genetics, developmental biology, regeneration, and toxicology[2] and is greatly used for comparative studies with other vertebrates[3]. In order to gain a better understanding of zebrafish *exosc10*/Exosc10, phylogenetic and synteny analysis of *exosc10* was performed and a homology matrix was constructed. mRNA and protein expression profile of *exosc10* in both male and female zebrafish tissues and embryonic developmental stages were established as well. Additionally, mRNA expression profile of *c1d, mpp6, dis3, dis3l* and *dis3l2*, who, besides *exosc10*, code for different subunits of the RNA exosome, was also examined.

Our phylogenetic analysis of Exosc10 exoribonuclease included vertebrate and invertebrate phyla and prokaryotes. In total, 32 sequences were identified, and the resulting phylogenetic tree is presented in *Figure 4*. Exosc10 wasn't found in *Pinctada fucata* (a type of mollusk), so it is not depicted here. The phylogenetic analysis shows a distinguishable separation of the vertebrate and invertebrate Exosc10. Within vertebrates, another subdivision can be distinguished that includes Exosc10 of tetrapods and Exosc10 of teleosts. Invertebrate Exosc10 forms a distinct group which is separated from vertebrate and prokaryotic proteins. Also, bacterial proteins form a separate cluster. A cluster of Mate proteins was used as a control, which we refer to as the 'outgroup' and it is clearly separated from the rest of the tree. It is visible that the outgroup has clearly separated from the rest of the tree, which means that all other Exosc10 protein sequences share a significant similarity.

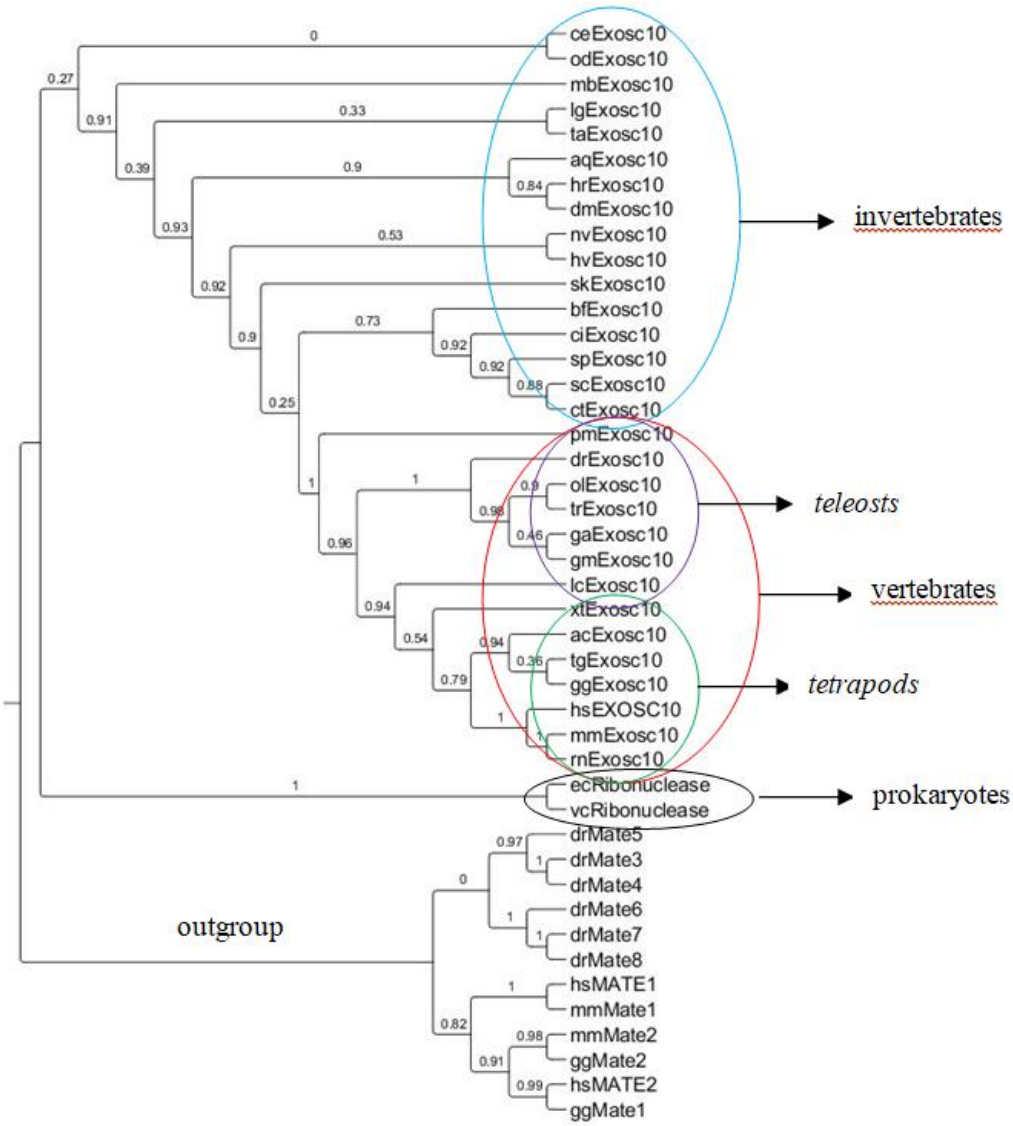


Fig. 4 Phylogenetic analysis of Exosc10 in vertebrates, invertebrates and prokaryotes.

The numbers indicate the possibility of branching, 0 being least possible and 1 being highly likely.

Species abbreviations: ce-*Caenorhabditis elegans*, od-*Oikopleura dioica*, mb-*Monosiga brevicollis*, lg-*Lottia gigantea*, ta-*Trichoplax adhaerens*, aq-*Amphimedon queenslandica*, hr-*Helobdella robusta*, dm-*Drosophila melanogaster*, nv-*Nematostella vectensis*, hv-*Hydra vulgaris*, sk-*Saccoglossus kowalevskii*, bf-*Branchiostoma floridae*, ci-*Ciona intestinalis*, sp-*Strongylocentrotus purpuratus*, sc-*Saccharomyces cerevisiae*, ct-*Capitella teleta*, pm-*Petromyzon marinus*, dr-*Danio rerio*, ol-*Oryzias latipes*, tr-*Takifugu rubripes*, ga-*Gasterosteus aculeatus*, gm-*Gadus morhua*, lc-*Latimeria chalumnae*, xt-*Xenopus tropicalis*, ac-*Anolis carolinensis*, tg-*Taeniopygia guttata*, gg-*Gallus gallus*, hs-*Homo sapiens*, mm-*Mus musculus*, rn-*Rattus norvegicus*, ec-*Escherichia coli*, vc-*Vibrio cholera*.

On homology matrix depicted on *Figure 5*, we can see the percentage of identity between proteins from fourteen different organisms. All of the organisms were chosen carefully, either being close relatives of zebrafish (ol and lc) or also being model organisms (xt, ce, dm, sc and spom) or taxonomically close groups of species (hs, mm, rn, gg, tg and ac – all vertebrates). It is clear that drExosc10 has the highest percentage of identity to olExosc10 (71.4% identical) and that it is least identical in comparison to ceExosc10 (only 25.3%). drExosc10 shares as much as 60.5% identity with hsExosc10.

Percent identity													
	mm	**rn**	**gg**	**tg**	**Ac**	**xt**	**dr**	**ol**	**lc**	**ce**	**Dm**	**sc**	**spom**
hs	87.1	87.2	72.5	81.3	61.8	65.4	60.5	61.8	68.7	26.1	33.9	28.6	26.7
mm		95.5	70.5	78.8	61.1	64.3	60.2	61.2	66.9	26.0	34.4	28.4	27.2
rn			70.6	78.8	61.7	64.6	60.8	62.0	67.8	26.1	34.3	28.5	27.4
Gg				90.1	64.0	63.2	60.0	60.9	68.3	27.8	34.1	27.7	28.1
Tg					83.5	71.4	69.1	69.6	76.6	34.9	42.4	36.3	33.7
Ac						55.3	53.0	52.6	57.0	29.7	31.7	27.5	28.2
Xt							58.6	59.6	63.9	26.1	31.7	29.0	26.9
Dr								71.0	61.2	25.3	32.4	26.2	27.7
Ol									63.2	26.9	34.4	27.6	28.2
Lc										26.7	33.2	28.7	27.9
Ce											25.6	24.7	22.4
Dm												24.5	24.7
Sc													33.6

Fig. 5 Protein homology matrix.

Species abbreviations: hs-*Homo sapiens*, mm-*Mus* musculus, rn-*Rattus* norvegicus, gg-*Gallus gallus*, tg-*Taeniopygia guttata*, ac-*Anolis carolinensis*, xt-*Xenopus tropicalis*, dr-*Danio rerio*, ol-*Oryzias latipes*, lc-*Latimeria chalumnae*, ce-*Caenorhabditis elegans*, dm-*Drosophila melanogaster*, sc-*Saccharomyces cerevisiae*, spom-*Schizosaccharomyces pombe*.

Zebrafish *exosc10* is located on chromosome 23 at 14.77 Mbp, while its human ortholog is located on chromosome 1 at 11.07 Mbp. Both zebrafish and human have only one copy of *exosc10*. It is also visible that there was an insertion in inverted orientation on zebrafish chromosome 23 downstream from *exosc10*. This insert is a duplicate of a part of zebrafish chromosome 11 (not shown here), but this part contains 'a' versions of *NDRG3*, *PHF20* and *SULF2*.

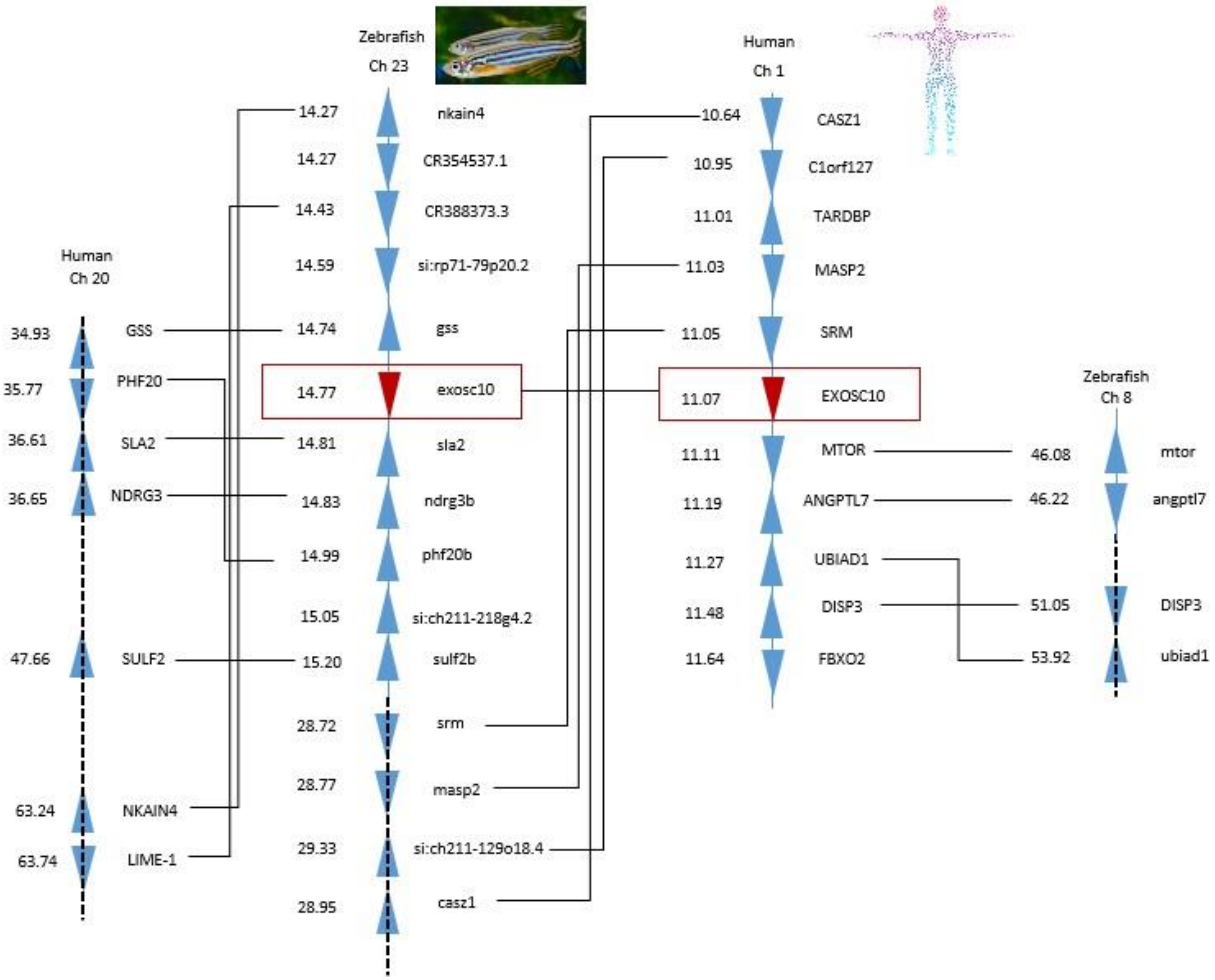


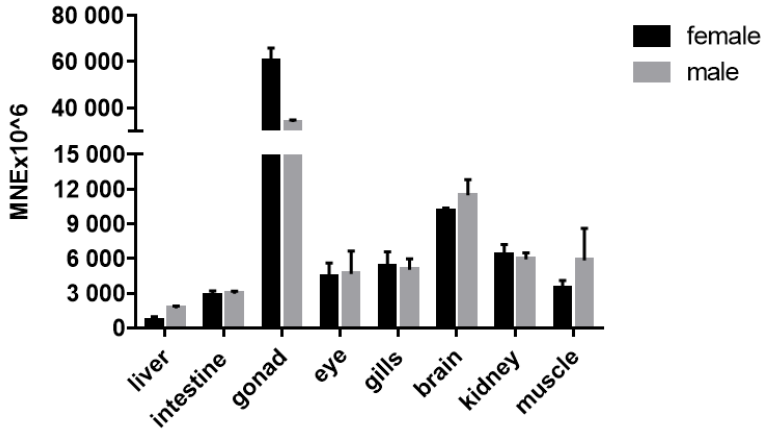
Fig. 6 Synteny analysis of *exosc10* in zebrafish (*Danio rerio*) and human (*Homo sapiens*). Blue line – continuous chromosome segment; black dashed line – discontinuous chromosome segment. Numbers next to the gene names represent megabase pair (Mbp) of gene location on the chromosome. Orthologs are connected with black lines.

Altogether, we constructed a phylogenetic tree and a homology matrix and all of this together with synteny analysis indicated that the protein is orthologous to its human counterpart, that it is approximately 60% identical among vertebrates and that it is evolutionarily conserved in eukaryotes.

3.2. *exosc10* expression profile

On *Figure 7 A*, a notable expression of *exosc10* is visible in all tissues tested and there are no differences between female and male tissues. However, the highest expression of *exosc10* is in both female and male gonads. On *Figure 7 B*, earlier stages of embryonic development, up to 6 hours post fertilization ('hpf'), show the highest expression and are followed by a gradual decrease of expression with time. This pattern was expected because this happens with many other transcripts in zebrafish. The expression of *exosc10* in ovaries is around 60 000 MNE x 10^6, while 1 hour post fertilization it is surprisingly similar; around 50 000 MNE x 10^6, which could mean that transcripts from ovaries are transferred to early embryos, and as the embryo develops and cells proliferate, the expression level seems to drop.

A)



B)

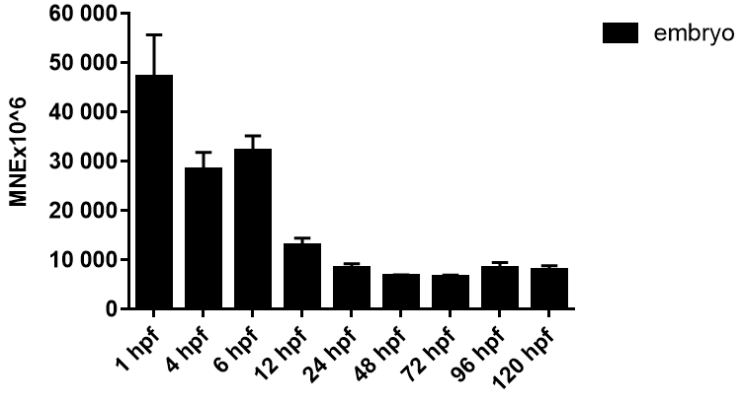


Fig. 7 (A) *exosc10* expression profile in female and male zebrafish. (B) *exosc10* expression profile in embryonic developmental stages of zebrafish. Data are depicted as MNE (mean normalized expression) +/- SE (standard error) from three independent experiments (independent biologocial samples).

3.3. Exosc10 protein expression profile

After determining the *exosc10* expression profile at the level of mRNA, we also wanted to determine Exosc10 protein level. Currently, no antibody for zebrafish Exosc10 is available on the market so two antibodies, based on their epitopes, were carefully chosen among commercially available antibodies and purchased for the purposes of this experiment. Abcam polyclonal and Santa Cruz monoclonal primary antibody were therefore used in further research. The specificity of these two antibodies was first tested by performing Western blot analysis on ovary tissue lysates in three different concentrations. This specific lysate was used because we expected the highest concentration of Exosc10 in gonads due to qRT-PCR results which have shown that gonads have the highest expression of *exosc10*.

On *Figure 8,* good separation of the protein ladder is visible. One expected band the size of Exosc10 (100 kDa) is visible on both blots in all lanes which contain the sample. Also, bands have different intensity for different tissue lysate concentrations, the most intense being the one with the greatest mass of sample (80 µg), and the least intense being the one with the smallest amount of sample (8 µg). On the right blot, there is an inversion and the signal is too strong, which is why the bands are white and the background is black, and not vice versa. H2B levels (15 kDa) are unfiormous and even dimers of the histone are visible (at 25 kDa).

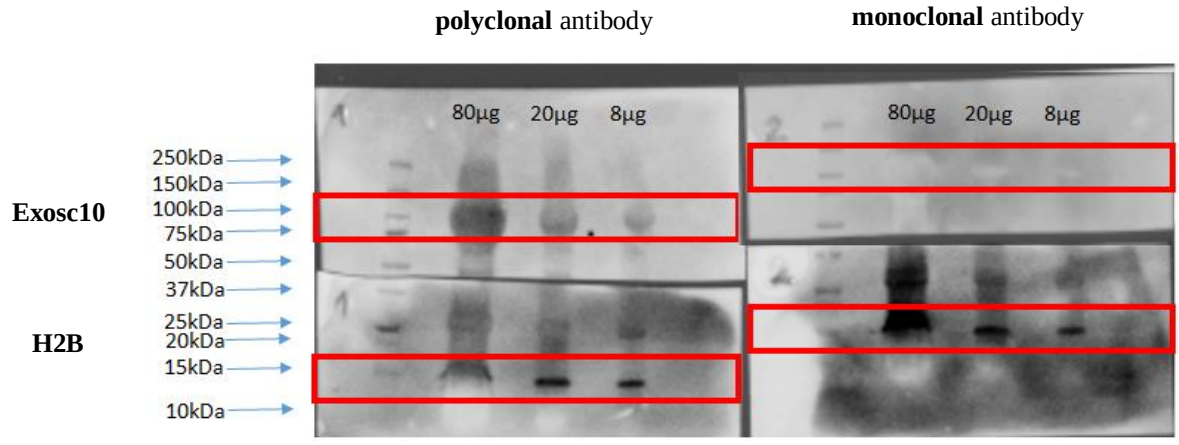


Fig. 8 Western blot analysis of Exosc10 in ovaries. Far left on the blot – protein ladder.

Polyclonal antibody manufactured by Abcam was chosen for further research. Western blot analysis was then performed on all tissue lysates to determine the Exosc10 protein expression in different zebrafish tissues. Male intestine and kidney samples weren't analysed with the rest due to size limitations of the gel (number of wells). Results obtained performing this analysis are depicted on *Figure 9*. Again, good separation of the protein ladder is visible.

However, many different bands are visible and some are even as intense as that of Exosc10, but the band size of that of Exosc10 is however very intense in some samples (top of the blot). Presence of many different bands is probably due to polyclonality of the antibody used, making it less specific for Exosc10. Nonetheless, this problem could be solved by additionally washing the membrane. The band size of that of Exosc10 shows high intensity in female brain, ovaries and both female and male liver, which is followed by eyes, gills and sceletal muscle. The band size of Exosc10 also differs in liver versus ovaries and brain and this could be due to modifications of the protein, making it more or less mobile during gel electrophoresis. H2B levels are uneven for female intestine and both female and male muscle (bottom of the blot). Situation for ovaries, testes and male brain is similar.

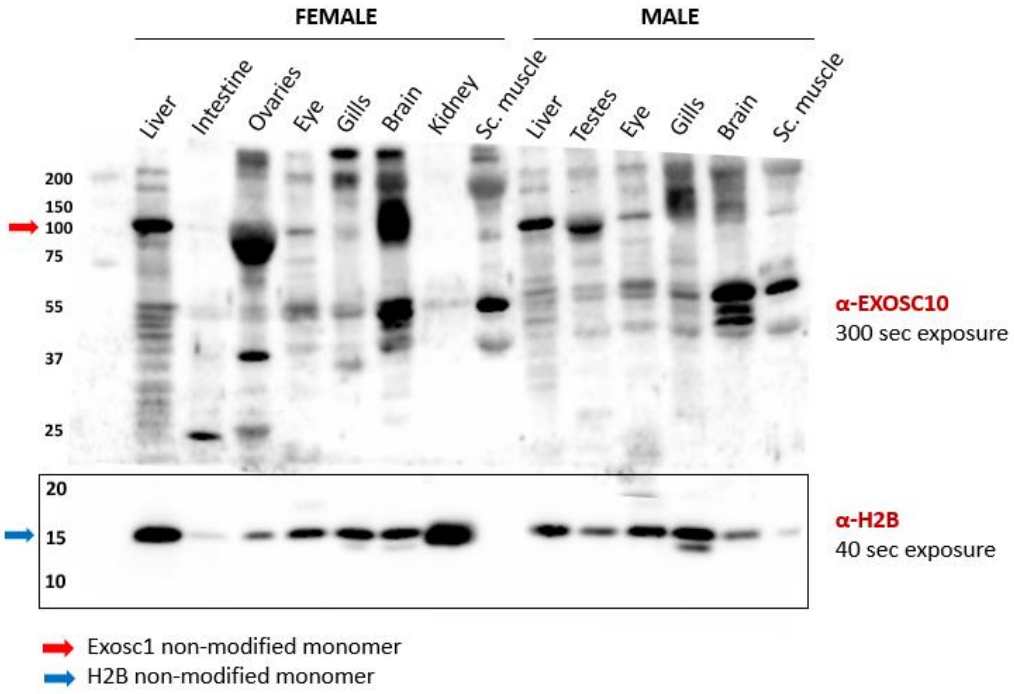


Fig. 9 Western blot analysis of Exosc10 in female and male zebrafish tissue lysates. 5 µg of sample were loaded per each lane, analysis on a 5-18% gradient gel. Left – protein ladder. Only Anti-EXOSC10 rabbit polyclonal primary antibody was used in this analysis (Abcam, Cambridge, UK).

Another Western blot analysis using polyclonal antibody manufactured by Abcam was performed. Only liver, gonads, brain and kidney tissue lysates from both female and male zebrafish were used. Results obtained performing this analysis are depicted on *Figure 10*. Again, good separation of the protein ladder is visible. Many different bands present on *Figure 9* are eliminated due to different blotting conditions (Abcam primary antibody was diluted 1:6,000 instead of 1:3,000 and more washing steps). The band size of Exosc10 is visible in both female and male liver and gonads. No such band is present in brain and kidney samples. H2B band is visible in all male tissues and liver, ovaries and brain of female tissues, while for female kidney sample degradation of the tissue is visible.

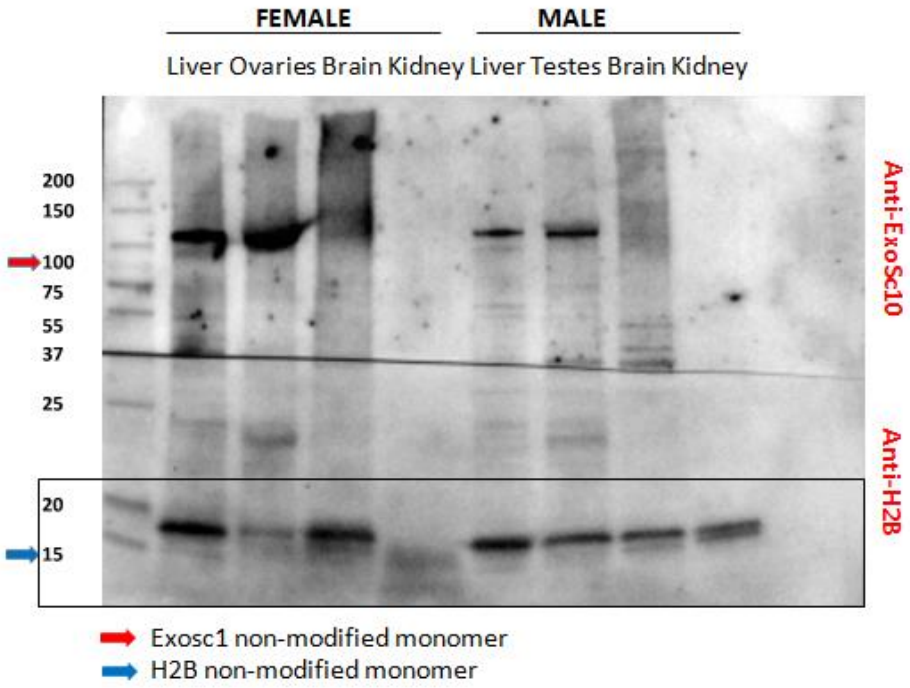
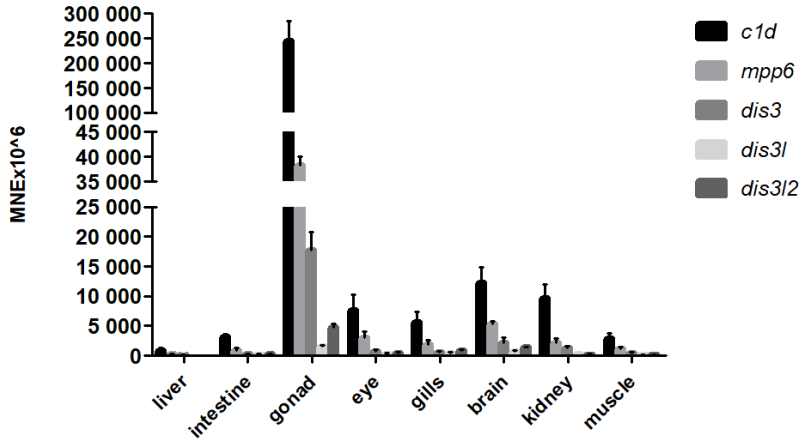


Fig. 10 Western blot analysis of Exosc10 in female and male zebrafish tissue lysates. 5 µg of sample were loaded per each lane, analysis on a 4-20% gradient gel. Left – protein ladder. Only Anti-EXOSC10 rabbit polyclonal primary antibody was used in this analysis (Abcam, Cambridge, UK).

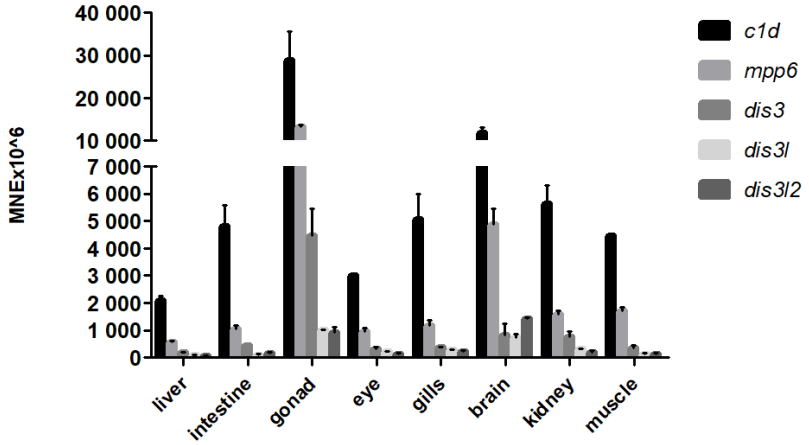
3.4. *c1d, mpp6, dis3, dis3l* and *dis3l2* expression profiles

Additionally, further analysis of expression profile for *c1d*, *mpp6, dis3, dis3l* and *dis3l2* was performed. This was done because according to current literature, some other exosome subunits could also be involved in meiosis and cellular differentiation[12]. Therefore, all genes researched code for either subunits or cofactors of the RNA exosome complex which then work closely with Exosc10 exoribonuclease. Our results show that *c1d* has the highest expression profile and that its expression surpasses that of *mpp6, dis3, dis3l* and *dis3l2* in all female and male tissues and all embroynic developmental stages and is the highest in both female (*Figure 11 A*) and male (*Figure 11 B*) gonads of zebrafish and earlier embryonic developmental stages up to 6 hpf (*Figure 11 C*). On *Figure 11 C* a gradual decrease of expression profile of all genes with time is visible. However, *c1d* is approximately 8 times upregulated in female versus male tissues (250 000 MNE x 10^6 versus 30 000 MNE x 10^6), and has a 4 times higher expression than *exosc10* in female tissues (250 000 MNE x 10^6 versus 60 000 MNE x 10^6). Even though all of the other exosome components tested have much lower expression level, their expression is nonetheless notable, and all of the components tested are differentially expressed in all tissues and embryonic developmental stages. The expression level of *c1d* in earlier embroynic developmental stages (1 hpf, 85 000 MNE x 10^6) is approximately 3 times higher than that of *c1d* in male gonads (30 000 MNE x 10^6). The second highest expression in all female and male tissues and embroynic developmental stages has *mpp6*, followed by *dis3* transcript. Lastly, in contrast to *exosc10* expression profile, expression profile of all of the genes discussed in this section shows a clear difference among the two genders.

A)



B)



C)

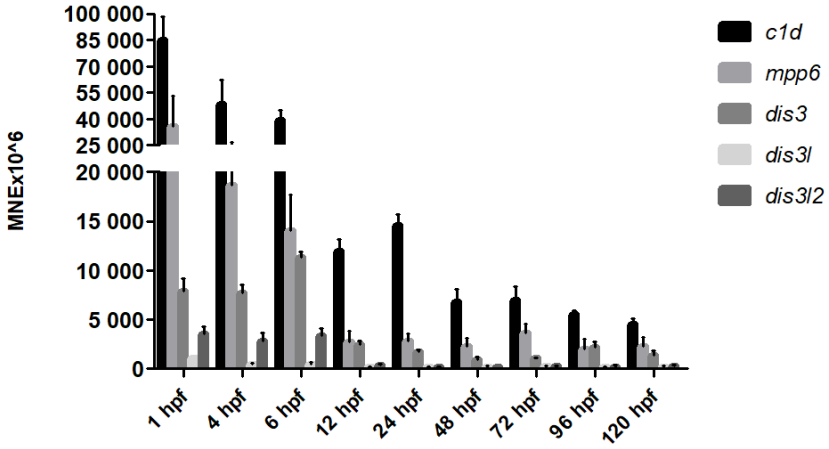


Fig. 11 (A) *c1d*, *mpp6*, *dis3*, *dis3l* and *dis3l2* expression profile in female (B) and male zebrafish (C) and embryonic developmental stages. Data are depicted as MNE (mean normalized expression) +/- SE (standard error) from three independent experiments (independent biologocial samples)

4. DISCUSSION

4.1. Phylogenetic and synteny analysis of *exosc10*/Exosc10

On *Figure 4*, we can see that the outgroup is clearly separated from the rest of the tree, meaning that all 32 Exosc10 protein sequences share a significant similarity. The phylogenetic analysis shows a distinguishable separation of the vertebrate and invertebrate Exosc10. Within vertebrates, another subdivision can be distinguished that includes Exosc10 of tetrapods and Exosc10 of teleosts. Invertebrate Exosc10 forms a distinct group which is separated from vertebrate and prokaryotic proteins. Also, bacterial proteins form a separate cluster, which was as expected because these are the only prokaryotic organisms used in tree construction and literature findings suggest that their ribonuclease complex is quite different than that of eukaryotes. They create a link between Exosc10, as a part of the eukaryotic RNA exosome complex, to a prokaryotic version of this complex, called RNase D, but several differences are known to exist between the structures of these proteins: RNase D includes a structured HRDC2 domain that limits access to the active site[10]. Thus, bacterial proteins (prokaryotes) diverged from eukaryotic ones, vertebrates separated from invertebrates and among vertebrates the mammals are grouped together. Among invertebrates, there is a little more disorder, but this can be due to quality of sequences currently available in the database.

On homology matrix depicted on *Figure 5*, it is clear that drExosc10 has the highest percentage of identity to olExosc10 (71.4% identical), which was expected because these are both fish species. However, drExosc10 is also highly similar to tgExosc10 (69.1% identity), which wasn't expected because that is a species which belongs to a class of birds. drExosc10 is least identical to ceExosc10 (only 25.3%), which is expected because it is a nematode, which belongs to invertebrates. Also, drExosc10 share only 26.2% identity with scExosc10 and 27.7% with spomExosc10. drExosc10 shares as much as 60.5% identity with hsEXOSC10. All of this, taken together, indicates that this protein is about 60% conserved among vertebrates and it confirms that this protein is evolutionarily highly conserved among eukaryotic organisms[23].

Phylogenetic analysis and genetic mapping suggest that a chromosome doubling event, probably by whole genome duplication (WGD), occurred in teleost fish species[24]. This probably occurred early in the teleost lineage after it split from the tetrapod lineage (in an ancestor of teleosts)[1,25], and only a subset of the duplicates have been retained in modern

teleost genomes today[25]. WGD therefore resulted in many duplicate genes (paralogs) to those found in mammals. However, only some of these gene duplications remain today, because some of the duplicated genes have a new function and others no longer express in the same tissues as the original genes (orthologs)[4]. On *Figure 6*, it can be seen that both zebrafish and human have only one copy of *exosc10*, and even though fish-specific WGD occurred in teleosts, only one copy of the gene remained today. It is also visible that there was an insertion in inverted orientation on zebrafish chromosome 23 downstream form *exosc10*. This insert is a duplicate of a part of zebrafish chromosome 11 (not shown here), but this part contains 'a' versions of *NDRG3, PHF20* and *SULF2*. Therefore, synteny analysis confirmed that zebrafish Exosc10 is orthologous to human Exosc10.

4.2. *exosc10* expression profile

Diploid yeast cells which lack Rrp6 fail to initiate premeiotic DNA replication normally and cannot undergo efficient meiosis[23]. That is why differential regulation of Rrp6 in various cellular states, such as meiosis, needs further investigation[12]. Current literature suggests that Rrp6 is not only essential for efficient meiosis and gametogenesis (sporulation) in various yeast strains, but also that EXOSC10 (its human ortholog) has essential function in germ cell proliferation, specially in male germline [7]. Our results suggest that Exosc10 has a vital role in zebrafish reproduction as well, because the highest expression level of *exosc10* was observed for both male and female gonads and in earlier phases of embryonic development (up to 6 hpf), as seen on *Figure 7*. No notable differences between the two genders have been observed. Recent research has shown that besides the genetic factors, which are inherited, some non-genetic factors, such as hormones, nutrients, different immune-relevant factors and proteins are transferred from mother to offspring as well [26,27]. We therefore propose that the gradual decrease of *exosc10* expression with time in embryonic developmental stages is a consequence of maternally transfered transcripts. However, all tissues have a significant expression level of *exosc10*, which was as expected because Exosc10 functions as an essential ribonuclease of eukaryotic cells where it is a part of the eukaryotic RNA exosome complex which is essential for the processing, quality control and degradation of nearly all classes of RNA in both the nucleus and cytoplasm of eukaryotic organisms[10,15]. Nonetheless, Exosc10 protein level varies depending on the tissue. Another potential reason for all tissues showing significant expression of *exosc10* was recently published in 2018 by Domingo-Prim et al. whose research suggests that not only is the exosome involved in RNA processing, quality

control and degradation, but that it is also necessary for the repair of DNA double-strand breaks (DSBs) in human cells. That is because RNA clearance is an essential step in homologous recombination (HR). When DSBs occur in a cell, lncRNAs are produced. Depletion of Exosc10 leads to increased lncRNAs, which means that the catalytic activity of Exosc10 contributes to the HR pathway by degrading lncRNAs and maintaining RNA homeostasis at DSBs[28].

4.3. Exosc10 protein expression profile

First, we checked the specificity of antibodies (Abcam polyclonal and SC monoclonal) by performing preliminary analysis on ovary tissue lysates in different concentrations. This was done on this specific tissue because we expected the highest concentration of Exosc10 there, which was based on qRT-PCR results. On *Figure 8*, one expected band corresponding to Exosc10 on both blots and a good dose-response ratio, is visible. This indicated that both the monoclonal and poylclonal antibody tested are specific for Exosc10. However, when we used Santa Cruz primary antibody, the concentration of the antibody was too high and that is why there is an inversion on the blot. However, H2B levels, which served as a control of sample loading and tissue degradation, are unfiormous and even dimers of the histone are visible. *Figure 9* represents a somewhat optimized anaylsis using different female and male tissues. Here we can see a non-uniformous loading control (H2B) at the bottom of the membrane, because H2B levels are uneven. However, Exosc10 is present in some samples on the top of the membrane. H2B levels are uneven for female intestine and both female and male muscle samples due to tissue degradation. Situation for ovaries, testes and male brain is similar but these tissues are less degraded. Exosc10 shows very high protein expression in female brain, ovaries and both female and male liver, followed by eyes, gills and skeletal muscle. *Figure 10* represents even more optimized analysis using only some female and male tissue lysates and diluting the primary antibody three times in comparison with the last analysis. Here, loading control (H2B) is more uniform and Exosc10 is present in both liver and gonads, while it seems not present in brain and kidney. For female kidney sample degradation is present, so it is questionable whether no expression of Exsoc10 is the case in that tissue. For male kidney and both female and male brain samples, no expression of Exosc10 seems to be the case, but to interpret intestine and kidney samples, new tissues are needed. This was expected because this antibody is polyclonal, and is therefore more likely to be less specific against Exosc10

than monoclonal antibody, which was also tested but not further optimized. To gain a better perspective, the whole analysis should be optimized and we suggest the following: tissue lysis should be performed faster by faster homogenization and reduced homogenization time. Everything should be done on ice and only 0.5% SDS should be used so that inhibitors work better. Also, maybe some of the steps could be reduced (i.e. sonication) and samples should not be used more than once after freezing the aliqouts to prevent vast degradation. Primary Abcam antibody can be even more diluted (1:6 000). Since concentration of proteins in lysate was determined previously to analysis and an equal amount of protein was loaded, problems we encountered are probably due to tissue degradation during either dissection, lysis, or re-freezing. SC primary antibody also needs optimization. Altogether, to gain a better understanding of Exosc10 protein expression in different zebrafish tissues, at least one, or even two independent protein isolations from all tissues should be performed. The most recent research published by Luz et al. earlier this year focuses on differential expression of RNA exosome subunits in an amphibian, *Lithobates catesbeianus*. Their research has shown that different subunits of the exosome can function independently, besides being a part of the exosome complex. Also, the exosome seems to be constitutively expressed in testes of this species and different subunits of the exosome, one of them being Rrp6, are not expressed at similar levels in different tissues. These findings suggest that the exosome composition and activity could be tissue specific and that its activity could be modulated in various ways[29], which is in line with the research we performed for Exosc10 on zebrafish tissues as well. In our work, only Exosc10 protein expression profile was examined, so another protein expression analyses for different subunits of the exosome could also be performed in zebrafish. In order to do that, different antibodies for other exosome components, besides Exosc10, should also be tested.

4.4. *c1d, mpp6, dis3, dis3l* and *dis3l2* expression profiles

Additionally to Exosc10, current findings suggest that some other exosome subunits are also essential for proper meiosis and cellular differentiation[12]. That is why we performed further analysis of expression profile for *c1d*, *mpp6*, *dis3*, *dis3l* and *dis3l2*, which, as mentioned earlier, are all either subunits or cofactors of the RNA exosome complex which work closely with Exosc10 exoribonuclease. As expected, C1d, as crucial cofactor of Exosc10 and its obligate partner, has the highest expression profile, along with Exosc10[18]. Its expression profile surpasses that of *mpp6*, *dis3*, *dis3l* and *dis3l2* in all female and male tissues and all

embroynic developmental stages and is the highest in both female (*Figure 11 A*) and male (*Figure 11 B*) gonads of zebrafish and earlier embryonic developmental stages as seen on *Figure 11 C* (up to 6 hpf). On *Figure 11 C* a gradual decrease of expression profile of all genes with time is present, as also seen on *Figure 7* for *exosc10*, which is a consequence of already mentioned maternal transfer. Even though all of the other exosome components tested have much lower expression level, their expression is nonetheless notable, and all of the components tested are differentially expressed in all tissues and embryonic developmental stages. Again, because of maternal transfer, the expression level of *c1d* in earlier embroynic developmental stages (1hpf) is approximately 3 times higher than that of *c1d* in male gonads. Some toxic substances, like brombisphenols, have a great imapact on early developmental stages of zebrafish, which is in correlation with *c1d* epxression. Exposure of zebrafish embroynic developmental stages to tetrabrombisphenol causes significant down-regulation of *c1d* expression[30]. The second highest expression in all female and male tissues and embroynic developmental stages has *mpp6*, followed by *dis3* ribonuclease gene. This could possiby be due to redundant function of Exosc10 and Dis3. Dis3 could out-compete other ribonucleases, such as Exosc10 to maintain proper RNA and nucleotide levels. For example, in the absence of Dis3, Exosc10 may become more active[31]. When Dis3 (Rrp44) is engaged with a substrate that occupies the channel, Exosc10 cannot bind its substrate. It is likely that exosome cofactors influence this by competing for interactions with the 'core' by modulating activities of Exosc10 and/or Dis3 (Rrp44). Therefore, although Dis3 (Rrp44) and Exosc10 are not in direct contact, their activities appear co-dependent when associated with Exo9[10]. Mpp6 is possibly this highly expressed because its function as a cofactor is to promote the channeling of substrates which are remodeled by Mtr4 helicase to Exosc10. Both Mpp6 and C1d (Rrp47) help to define targets which need to be processed and/or degraded by the exosome[10,14]. They are nuclear proteins enriched in nucleoli which have the ability of binding RNA. C1d (Rrp47) binds structured substrates, while Mpp6 shows strong preference for pyrimidine-rich sequences. However, both proteins interact directly with Exosc10 and they are at least partly redundant[8]. Proteins that co-purify with yeast nuclear exosome complexes, primarily through interactions with Exosc10, are C1d (Rrp47) and Mpp6. Both of these nuclear proteins are involved in RNA binding and are also conserved in higher eukaryotes[14]. Lastly, in contrast to *exosc10* expression profile, expression profile of all of the genes discussed in this section shows a clear difference among the two genders.

5. CONCLUSIONS

The most important conclusions derived from results obtained in this study can be summarized as follows:

1. Phylogenetic and synteny analysis of *exosc10*/Exosc10 confirmed that zebrafish *exosc10*/Exosc10 is orthologous to human *EXOSC10*/EXOSC10 and that it is evolutionarily highly conserved from chordates to mammals with about 60% identity among vertebrates.
2. *exosc10*, *c1d*, *mpp6*, *dis3*, *dis3l* and *dis3l2* are all differentially and constitutively expressed in all zebrafish tissues tested in both genders and embryonic developmental stages with significant differences in expression levels;
3. Exosc10 exoribonuclease has a potentially vital role in zebrafish reproduction system, it is maternally transferred from female to offspring;
4. Two primary antibodies, Santa Cruz monoclonal and Abcam polyclonal, are specific for Exosc10 and function in zebrafish tissue lysates;

6. REFERENCES

1. Gris, D. Public Access NIH Public Access. **185**, 974–981 (2013).
2. Meyers, J. R. Zebrafish: Development of a Vertebrate Model Organism. *Curr. Protoc. Essent. Lab. Tech.* **16**, 1–26 (2018).
3. Wixon, J. Featured organism: Danio rerio, the zebrafish. *Yeast* **17**, 225–231 (2000).
4. Hill, A. J., Teraoka, H., Heideman, W. & Peterson, R. E. Zebrafish as a model vertebrate for investigating chemical toxicity. *Toxicol. Sci.* **86**, 6–19 (2005).
5. Lieschke, G. J. & Currie, P. D. Animal models of human disease: Zebrafish swim into view. *Nat. Rev. Genet.* **8**, 353–367 (2007).
6. Midtgaard, S. F. *et al.* Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. *Proc. Natl. Acad. Sci.* **103**, 11898–11903 (2006).
7. Jamin, S. P. *et al.* EXOSC10/Rrp6 is post-translationally regulated in Male germ cells and controls the onset of spermatogenesis. *Sci. Rep.* **7**, 1–15 (2017).
8. Chlebowski, A., Lubas, M., Jensen, T. H. & Dziembowski, A. RNA decay machines: The exosome. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**, 552–560 (2013).
9. Januszyk, K., Liu, Q. & Lima, C. D. Activities of human RRP6 and structure of the human RRP6 catalytic domain. *Rna* **17**, 1566–1577 (2011).
10. Januszyk, K. & Lima, C. D. The eukaryotic RNA exosome. *Curr. Opin. Struct. Biol.* **24**, 132–140 (2014).
11. Wasmuth, E. V & Lima, C. D. The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome. *Nucleic Acids Res.* **45**, 846–860 (2017).
12. Fox, M. J. & Mosley, A. L. Rrp6: Integrated roles in nuclear RNA metabolism and transcription termination. *Wiley Interdiscip. Rev. RNA* **7**, 91–104 (2016).
13. Feigenbutz, M., Garland, W., Turner, M. & Mitchell, P. The exosome cofactor Rrp47 is critical for the stability and normal expression of its associated exoribonuclease Rrp6 in *Saccharomyces cerevisiae*. *PLoS One* **8**, 1–17 (2013).
14. Sloan, K. E., Schneider, C. & Watkins, N. J. Comparison of the yeast and human nuclear exosome complexes. *Biochem. Soc. Trans.* **40**, 850–855 (2012).
15. Stuparevic, I., Mosrin-Huaman, C., Hervouet-Coste, N., Remenaric, M. & Rahmouni, A. R. Cotranscriptional recruitment of RNA exosome cofactors Rrp47p and Mpp6p and two distinct Trf-Air-Mtr4 polyadenylation (TRAMP) complexes assists the exonuclease Rrp6p in the targeting and degradation of an aberrant messenger

ribonucleoprotein particle (mRN. *J. Biol. Chem.* **288**, 31816–31829 (2013).

16. Van Dijk, E. L., Schilders, G. & Pruijn, G. J. M. Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways. *Rna* **13**, 1027–1035 (2007).
17. Feng, S., Parrish, D. D. & Lambert, M. W. A processive versus a distributive mechanism of action correlates with differences in ability of normal and xeroderma pigmentosum group A endonucleases to incise damaged nucleosomal DNA. *Carcinogenesis* **18**, 279–286 (1997).
18. Zinder, J. C. & Lima, C. D. Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.* **31**, 88–100 (2017).
19. Wasmuth, E. V., Januszyk, K. & Lima, C. D. Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature* **511**, 435–439 (2014).
20. Falk, S., Bonneau, F., Ebert, J., Kögel, A. & Conti, E. Mpp6 Incorporation in the Nuclear Exosome Contributes to RNA Channeling through the Mtr4 Helicase. *Cell Rep.* **20**, 2279–2286 (2017).
21. Kowalinski, E. *et al.* Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. *Mol. Cell* **63**, 125–134 (2016).
22. Luan, S., Luo, J., Liu, H. & Li, Z. Regulation of RNA decay and cellular function by 3′-5′ exoribonuclease DIS3L2. *RNA Biol.* **16**, 160–165 (2019).
23. Lardenois, A. *et al.* Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. *Proc. Natl. Acad. Sci.* **108**, 1058–1063 (2010).
24. Amores, A. *et al.* Zebrafish hox clusters and vertebrate genome evolution. *Science (80-.).* **282**, 1711–1714 (1998).
25. Woods, I. G. *et al.* The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res.* **15**, 1307–1314 (2005).
26. Zhang, S., Wang, Z. & Wang, H. Maternal immunity in fish. *Dev. Comp. Immunol.* **39**, 72–78 (2012).
27. Swain, P. & Nayak, S. K. Role of maternally derived immunity in fish. *Fish Shellfish Immunol.* **27**, 89–99 (2009).
28. Domingo-Prim, J. *et al.* EXOSC10 is required for RPA assembly and controlled DNA end resection at DNA double-strand breaks. *Nat. Commun.* **10**, 2135 (2019).
29. Luz, J. S. *et al.* Differential expression of RNA exosome subunits in the amphibian *Lithobates catesbeianus* during reproductive and non-reproductive periods. *BMC Res.*

Notes **12**, 1–7 (2019).

30. Liu, H. *et al.* Pharmacokinetics and effects of tetrabromobisphenol a (TBBPA) to early life stages of zebrafish (*Danio rerio*). *Chemosphere* **190**, 243–252 (2018).
31. Hou, D., Ruiz, M. & Andrulis, E. D. The ribonuclease Dis3 is an essential regulator of the developmental transcriptome. *BMC Genomics* **13**, (2012).

Izjava o izvornosti

Izjavljujem da je ovaj diplomski rad izvorni rezultat mojeg rada te da se u njegovoj izradi nisam koristila drugim izvorima, osim onih koji su u njemu navedeni.

Leticija Oreški

ime i prezime studenta