Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN

Scienze della Terra, della Vita e dell'Ambiente

Ciclo XXXII

Settore Concorsuale: 05/B1 – ZOOLOGIA E ANTROPOLOGIA

Settore Scientifico Disciplinare: BIO/05 - ZOOLOGIA

Heredity and biogenesis of mitochondria: evolutionary and functional aspects.

Presentata da: **Dott. Elisabetta Punzi**

Coordinatore Dottorato
Prof. Giulio Viola

Supervisore
Prof. Marco Passamonti

Esame finale anno 2020

Abstract

Nearly all animal species inherit mitochondrial DNA in a matrilineal way: the mtDNA of the progeny originates entirely from the oocyte, whereas paternal mtDNA carried by spermatozoa is actively degraded pre- or post-fertilization by a variety of means. This well-known pattern goes by the name of “strictly maternal inheritance” of mitochondria or SMI. It is not currently known why SMI is so prevalent and what evolutionary advantage it confers exactly. Some bivalve species present a different mitochondrial inheritance pattern called doubly uniparental inheritance or DUI, the only known stable exception to SMI. Despite constituting an excellent system to study mitochondrial evolution and inheritance, its molecular background is poorly understood and the evolutionary equilibria it participates to even less so. The present thesis aims at uncover some of the molecular basis for DUI in *R. philippinarum* (Bivalvia, Veneridae) and to map some features of its basic biology.

Summary

Introduction

Mitochondria

Mitochondria are fundamental cellular organelles found in almost all eukaryotic organisms that originated through an endosymbiotic event (Margulis & Sagan, 1986). They are the target of almost a thousand proteins in yeast and 1100-1400 in human (Calvo & Mootha, 2010; Morgenstern et al., 2017) that are employed in the most diverse biological process, such as calcium homeostasis, heme synthesis, β-oxidation of free fatty acids and, most importantly, ATP production through oxidative phosphorylation or OXPHOS (Michel et al., 2012). The latter is performed thanks to the concerted action of almost a hundred proteins, some of which are encoded in the mitochondrial DNA (mtDNA). Generally, in animals, mtDNA is a small (16 kb) circular molecule that encodes for 13 protein-coding genes involved in OXPHOS and 24 structural RNAs, with little to no introns (Gissi, Iannelli, & Pesole, 2008), making it a very compact genome if compared to the nuclear one.

However, there is ample variation to this paradigm. Several animals have been reported to lack specific mitochondrial genes, such as *atp6* lacking in several ctenophores (Kohn et al., 2012) and chaetognaths (Miyamoto, Machida, & Nishida, 2010) and *atp8* being lost in several taxonomic lineages (Bernt, Braband, Schierwater, & Stadler, 2013), or to have duplicated genes, such as in some cephalopods (Kawashima et al., 2013). The most interesting deviation from the rule, however, is the presence of protein-coding genes with non-OXPHOS functions. For instance, in the octocoral *Sarcophyton glaucum* mtDNA, there is a homolog of *mutS,* a homolog of the bacterial component of the mismatch repair pathway, whose origin is thought to be due to horizontal gene transfer (Bilewitch & Degnan, 2011) and, notably, a gene called *humanin* has been detected in human mtDNA, coding for a 24 amino acid peptide which functions as neuro- and cytoprotector(Lee, Yen, & Cohen, 2013). Some of these newly-discovered genes have yet to be identified and are collectively known under the name of ORFans. The best characterized ORFans are *gau*, an ubiquitous open reading frame (ORF) encoded on the complementary strand of *cox1* of eukaryotic mitochondria, and a complex of ORFs found in male and female mitotypes in bivalves with DUI (see next paragraph). In particular, in the venerid

Ruditapes philippinarum (Milani, Ghiselli, Maurizii, Nuzhdin, & Passamonti, 2014), the mytilid *Mytilus edulis* (Ouimet et al., 2019) and the unionid *Venustaconcha ellipsiformis* (Breton et al., 2011), the presence of the protein product has been confirmed experimentally, leading to the hypothesis that these proteins, which display remarkable conservation levels (Milani, Ghiselli, Guerra, Breton, & Passamonti, 2013), are functional.

Mitochondria are in communication with other organelles through a variety of mechanisms that ensures that the interplay between them is finely tuned. Mitochondria are tightly linked to the endoplasmic reticulum by means of contact sites called mitochondria-associated ER membranes (MAMs), which coordinates several functions such as calcium uptake, apoptosis regulation and phospholipid synthesis (Xia et al., 2019). Vesicle trafficking is the main agent of the interaction between mitochondria and peroxisomes, fundamental for ROS balance maintenance and to perform immune responses (Schrader, Costello, Godinho, & Islinger, 2015). Finally, the nucleus maintains a complex dialogue with mitochondria through several means. Aside from the anterograde and retrograde cascade signaling, several noncoding RNAs comprehending long and short noncoding RNAs have been found to regulate nucleus-mitochondria interplay. Intriguingly, this involves not only RNA encoded in the nuclear genome (see for example mitoMiRs, Bandiera, Matégot, Girard, Demongeot, & Henrion-Caude, 2013), but also in the mitochondrial one. Recent findings evidenced new classes of small noncoding RNAs transcribed from the mtDNA in humans (Ro et al., 2013) and in the clam *Ruditapes philippinarum* (Pozzi, Plazzi, Milani, Ghiselli, & Passamonti, 2017). The latter is of particular significance, as the targets of the so called small mitochondrial RNAs or smithRNAs are nuclear genes and some of them appear to have been empirically validated (unpublished results).

One of the peculiarities of mitochondria is their mode of inheritance. In fact, in almost all the animal kingdom, mitochondria are inherited solely from the mother, an inheritance pattern known as strictly maternal inheritance or SMI. Despite the uniformity of the outcome across taxa, the resulting state of homoplasmy is achieved through different molecular processes (K. Sato & Sato, 2017) which involve degrading mtDNA during spermatogenesis as in *Drosophila* (DeLuca & O'Farrell, 2012) and *Oryzias latypes*

(Nishimura et al., 2006), degradation of paternal mitochondria after fecundation through autophagy as in *Caenorhabditis elegans* (M. Sato & Sato, 2011) and subsequent to ubiquitin marking as in mammals (Sutovsky et al., 1999), or even preventing paternal mitochondria entrance in the egg as in *Cricetulus griseus* (Pickworth & Change, 1969) (for a more thorough review, see chapter 1).

DUI

About a hundred bivalve species pertaining to the super-class Autolamellibranchia (Gusman, Lecomte, Stewart, Passamonti, & Breton, 2016) present a characteristic pattern of mitochondrial inheritance known as doubly uniparental inheritance or DUI. In DUI species, two different mtDNA lineages are present: the F-type and the M-type, inherited through the eggs and the sperm respectively. Upon entering the egg, male mitochondria face a different fate according to the zygote sex. In males, they keep their aggregate form and multiply, so that males are homoplasmic for the M-type in the germline and heteroplasmic in soma. In females, instead, paternal mitochondria disperse and their DNA becomes undetectable, making them mostly homoplasmic for F-mtDNA in both the germline and somatic tissues.

The origin of DUI is still shrouded in mystery. As Theologidis, Fodelianakis, Gaspar, & Zouros, 2008 detail, its presence in the bivalve phylogenetic tree is scattered, with entire families inheriting their mitochondria through SMI and families whose species present either SMI or DUI. At the time being, two opposing hypotheses have been formulated on DUI origin. One hypothesis presupposes a single origin at the base of Autolamellibranchia radiation, at the beginning of the Ordovician, with DUI being lost in several taxa and retained in others (Zouros, 2013). The other one postulates that DUI originated several times through viral infection (Milani et al., 2014), which would explain the scattered distribution. This second hypothesis is supported by another feature often found in DUI mitochondrial genomes, which are the already mentioned ORFans. The *in silico* analysis of DUI ORFans found no clear homology with proteins in databases, except for distant similarities with viral proteins (Milani et al., 2013). What is more, their sequences are not similar to each other, as one would expect with a single-origin DUI. The role of these ORFans is still source of debate, one possibility being that they are involved in the selfish behavior that their host mitochondria display, allowing

them to invade the female or male germline and acting as meiotic drivers (Milani, Ghiselli, & Passamonti, 2016; Milani, Ghiselli, Pecci, Maurizii, & Passamonti, 2015).

Aim of the thesis

With this thesis, I intended to shed light on the molecular processes behind DUI in the venerid species *Ruditapes philippinarum* with a twofold approach. The first step was to assess similarities and differences between the transcriptome of *R. philippinarum* and its congeneric SMI species *R. decussatus* with regards to the pathways known to be involved in SMI maintenance (Chapter 1). In order to gain insight about its structure and function, I moved forward trying to produce *R. philippinarum* M-ORFan RPHM21 through cell-free protein expression and in yeast (Chapter 2). Finally, given the discovery of smithRNAs in this species, I set to elucidate the miRNA environment and its links to mitochondria (Chapter 3).

Chapter one

Lose it or keep it: how bivalves can provide insights into mitochondrial inheritance mechanisms

Received: 17 July 2017 | Revised: 2 December 2017 | Accepted: 9 January 2018

DOI: 10.1002/jez.b.22788

RESEARCH ARTICLE

WILEY JEZ-B MOLECULAR AND DEVELOPMENTAL EVOLUTION

Lose it or keep it: (how bivalves can provide) insights into mitochondrial inheritance mechanisms

Elisabetta Punzi | **Liliana Milani** | **Fabrizio Ghiselli** | **Marco Passamonti**

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

Correspondence
Fabrizio Ghiselli, Department of Biological, Geological, and Environmental Sciences, University of Bologna, Via Selmi 3, 40126 Bologna, Italy.
Email: fabrizio.ghiselli@unibo.it

Elisabetta Punzi and Liliana Milani contributed equally to this work.

Funding information
Grant sponsors: Italian Ministry of Education, University and Research (MIUR) FIR Programme (RBFR13T97A); MIUR SIR Programme (RBSI14G0P5); Canziani bequest.

Abstract

The strictly maternal inheritance (SMI) is a pattern of mitochondrial inheritance observed across the whole animal kingdom. However, some interesting exceptions are known for the class Bivalvia, in which several species show an unusual pattern called doubly uniparental inheritance (DUI) whose outcome is a heteroplasmic pool of mtDNA in males. Even if DUI has been studied for long, its molecular basis has not been established yet. The aim of this work is to select classes of proteins known to be involved in the maintenance of SMI and to compare their features in two clam species differing for their mitochondrial inheritance mechanism, that is, the SMI species *Ruditapes decussatus* and the DUI species *Ruditapes philippinarum*. Data have been obtained from the transcriptomes of male and female ripe gonads of both species. Our analysis focused on nucleases and polymerases, ubiquitination and ubiquitin-like modifier pathways, and proteins involved in autophagy and mitophagy. For each protein group of interest, transcription bias (male or female), annotation, and mitochondrial targeting (when appropriate) were assessed. We did not find evidence supporting a role of nucleases/polymerases or autophagic machinery in the enforcement of SMI in *R. decussatus*. On the other hand, ubiquitinating enzymes with the expected features have been retrieved, providing us with two alternative testable models for mitochondrial inheritance mechanisms at the molecular level.

KEYWORDS
autophagy, nucleases, polymerases, transcriptomics, ubiquitin

1 | INTRODUCTION

In animals, the mitochondrial genome (mtDNA) is usually transmitted to the progeny exclusively by the female parent. Despite strictly maternal inheritance (SMI) being nearly ubiquitous across eukaryotes, its underlying molecular mechanism is widely variable, suggesting recurrent loss and restoration and/or several independent origins (Birky, 1995). Paternal inheritance can be prevented by mtDNA elimination by nucleases either during spermatogenesis or after fertilization; alternatively, paternal mitochondria can be selectively degraded after entering the oocyte through proteasomal action or mitophagy. In the fish *Oryzias latypes*, the copy number of nucleoids (i.e., mtDNA–protein complexes) decreases during spermatogenesis. Once the spermatozoon enters the oocyte, an unknown endonuclease degrades the remaining mtDNA molecules, leaving paternal mitochondria with no genomic content, yet morphologically intact (Nishimura et al., 2006). In spermatozoa of *Drosophila melanogaster*, the two mitochondria extend by the exceptionally long tail (1,800 μm); in this species, nucleoids are completely degraded during spermatogenesis in a proximal-distal way, from the neck to the end to the tail (DeLuca & O'Farrell, 2012). Endonuclease G was initially thought to be the main effector of this degradation; however, recent research revealed the essential role of the mitochondrial polymerase Tamas in nucleoid elimination (Yu, O'Farrell, Yakubovich, & DeLuca, 2017). A second mechanism ensures the complete clearance of paternal nucleoids: during *D. melanogaster* spermatid individualization, an actin structure called "investment cone" progresses along the sperm tail axoneme and collects trimmed nucleoids in a distal "waste bag." Subsequently, paternal mitochondria are degraded through autophagy soon after fertilization, between mitotic cycles 1 and 9 (Politi et al., 2014). The autophagic process involves the formation of a double-membrane vesicle that wraps the targeted structure and fuses with a lysosome, causing the degradation of the target. Autophagy has been extensively studied when occurring in response to starvation (Pfeifer and Scheller, 1975)—a process named also non-selective autophagy—but it performs a number of other selective tasks as well, such as pexophagy (i.e., selective degradation of peroxisomes via autophagy; Oku & Sakai, 2016), and mitophagy (i.e., mitochondrial autophagy; Lemasters, 2014).

The pioneering work of Sutovsky's research group highlighted the importance of the ubiquitination pathway in sperm mitochondria

elimination in cows and pigs. Ubiquitin (Ub) is a highly conserved peptide of 76-amino acids that is covalently bonded to lysine residues of proteins (Ciehanover, Hod, & Hershko, 1978), determining their sorting, degradation, or signal transduction, depending on the ubiquitination pattern (Swatek & Komander, 2016). Ubiquitination occurs as a three-step process involving Ub-activating (E1), Ub-conjugating (E2), and Ub-ligating (E3) enzymes. Tag specificity and selectivity are achieved by the high diversity of the E3 Ub-ligases (Hershko and Ciechanover, 1998). Ub moieties can be removed by a deubiquitinating enzyme (DUB), making ubiquitination a highly dynamic tagging system. During spermatogenesis in cows and pigs, the 30 kDa inner membrane protein prohibitin is diubiquitinated. After fertilization, mitochondrial membranes undergo a structural rearrangement that brings ubiquitinated prohibitins on the outer membrane, causing them to be exposed to recognition by zygotic/embryonic ubiquitination machinery. Such machinery, in turn, adds more Ub moieties to prohibitin and marks the switch from the di-Ub recognition signal to a poly-Ub degradation one (Sutovsky et al., 2000). Subsequently, paternal mitochondria are targeted to proteolytic destruction by the conjoint action of proteasome and autophagy/lysosome system (Sutovsky et al., 2000; Sutovsky, Mc Cauley, Sutovsky, & Day, 2003; Rojansky, Cha, & Chan, 2016). Further work by May-Panloup et al. (2003) and Luo et al. (2013), determined that vital sperm of mice and men has a very low nucleoid content, suggesting a process of mtDNA copy number reduction during spermatogenesis.

Lastly, autophagy and ubiquitination are the main processes responsible for the clearance of paternal mitochondria in *Caenorhabditis elegans* as well (Sato & Sato, 2011): upon entering the oocyte, sperm mitochondria and other structures of paternal origin called membranous organelles (MOs) are degraded through autophagy. MOs have been found to be ubiquitinated before and after fertilization, similarly to what happens in mammalian paternal mitochondria; however, no sign of ubiquitination has been detected on *C. elegans* paternal mitochondria.

1.1 | The exception to SMI

The only known evolutionarily stable exception to the common SMI is represented so far by the doubly uniparental inheritance of mitochondria or DUI (Skibinski, Gallagher, & Beynon, 1994a, b; Zouros, Ball, Saavedra, & Freeman, 1994a; Zouros, Oberhauser Ball, Saavedra, & Freeman 1994b). This mitochondrial inheritance mechanism has been found in ~100 species of bivalve molluscs (Gusman, Lecomte, Stewart, Passamonti, & Breton, 2016) and features two different mtDNAs, the F-type and the M-type, with high intraspecific divergence, and sex-specific inheritance. The distribution of the two mitochondrial genomes within an individual depends on its sex: females are homoplasmic for F-type mtDNA, whereas males carry the M-type mtDNA in the germline and both mitochondrial genomes in the soma, with varying proportions depending on species and tissue (Ghiselli, Milani, & Passamonti, 2011; Obata, Sano, & Komaru, 2011; Milani, Ghiselli, Iannello, & Passamonti, 2014a).

One of the most interesting peculiarities of DUI mtDNAs is that they contain a novel lineage-specific ORF (one in the F-type, one in the M-type) that, according to in silico prediction, might have had a viral origin (Milani, Ghiselli, Guerra, Breton, & Passamonti, 2013; Milani, Ghiselli, Maurizii, Nuzhdin, & Passamonti 2014b; Milani, Ghiselli, & Passamonti, 2016). Moreover, females of DUI species differ in offspring sex ratio, which can be either male-biased, female-biased, or balanced. This is a feature that appears to be mostly dependent on the maternal genotype, but not immune to paternal influence (Ghiselli et al., 2012; Kenchington, MacDonald, Cao, Tsagkarakis, & Zouros, 2002; Saavedra, Reyero, & Zouros, 1997; Yusa, Breton, & Hoeh, 2013). Observations in early embryos of Mytilus and the venerid *Ruditapes philippinarum* (both with DUI) revealed that sperm mitochondria show two different distribution patterns across blastomeres: aggregated or dispersed (Cao, Kenchington, & Zouros, 2004; Milani, Ghiselli, & Passamonti, 2012). In Mytilus, the two patterns have been associated with male and female embryos, respectively. However, differences in the aggregation pattern cannot account completely for the aforementioned distribution of mtDNA in tissues, and additional active mechanisms such as paternal mitochondria degradation in females and preferential replication in males (i.e., meiotic drive) have been proposed (Ghiselli et al., 2011, Milani, Ghiselli, Pecci, Maurizii, & Passamonti, 2015; Milani et al., 2016).

A further point of relevance concerns the evolutionary inception of DUI. It is not clear whether DUI had a single origin or arose several times throughout its evolutionary history. In the first case, DUI might be the result of a single event happened at the origin of the Autolamellibranchia superclass, more than 400 million years ago (Zouros, 2013). However, its distribution across the bivalve phylogenetic tree is not homogenous: for instance, within Pteriomorphia, mytilids have DUI, whereas ostreids and pectinids do not (Doucet-Beaupré et al., 2010), and among Veneridae, the two lineage-specific mtDNAs have been found in *R. philippinarum* (Passamonti & Scali, 2001) and *Meretrix lamarckii* (Bettinazzi, Plazzi, & Passamonti, 2016), whereas no evidence was found in *Ruditapes decussatus* (Ghiselli et al., 2017) and *Callista chione* (Plazzi, Cassano, & Passamonti, 2015). Besides being the result of incomplete sampling, this scattered distribution may also be imputed to false negatives due to the technical difficulties in the detection of the two different DUI mitochondrial genomes (see Theologidis, Fodelianakis, Gaspar, & Zouros, 2008 and Ghiselli et al., 2017 for a thorough discussion of this issue). In any case, if DUI had a single origin, several loss events have to be assumed to explain its scattered distribution across bivalves (Zouros, 2013).

That said, a multiple-origin hypothesis might be more parsimonious. Recent works proposed that the mitochondrial lineage-specific ORFs found in several bivalve species may play a role in DUI emergence and establishment (Breton et al., 2011b; Milani et al., 2013, 2014b, 2015, 2016). According to this hypothesis, the endogenization of viral sequences in mtDNA might be the trigger for DUI evolution; such viral sequences might have provided the recipient mtDNA with the ability to invade the germ line (e.g. through meiotic drive), thus producing a selfish element (Milani et al., 2015, 2016). Although such ORFs share some common features, their alignments were possible only among sequences of closely related species (Breton et al., 2011a; Milani et al., 2013): this may be due either to their fast evolution making their homology undetectable, or to several independent

endogenization events. As a matter of fact, a hypothesis featuring multiple viral origins of DUI may explain its scattered distribution across bivalves.

Being the only known stable exception to SMI, DUI provides a unique chance to study mitochondrial inheritance mechanisms by comparing two naturally occurring systems in two relatively close species. As mentioned before, it is well known that SMI maintenance, despite resulting in the same outcome, is achieved through the most diverse mechanisms (Birky, 1995, 2001; Sato & Sato, 2013). Similarly, it is conceivable that, at a molecular level, DUI relies on a machinery that differs from one taxon to another. So it seems legitimate to hypothesize that *R. philippinarum* may share a more similar machinery with a congeneric SMI species such as *R. decussatus*, rather than with other DUI species outside Veneroida. Of course, since the eventual mitochondrial distribution pattern between a SMI and a DUI species is completely distinct, there must be difference, but such difference can reside virtually in a single protein (Zouros, 2013).

Summarizing, the process of paternal mitochondria degradation in animals comprises two temporally distinct steps: degradation of sperm mtDNA and/or labeling of paternal mitochondria occurs during spermatogenesis, whereas degradation of nucleoids and/or recognition and degradation of paternal mitochondria happens after fertilization.

The sequences encoding the machinery for the first step have to be necessarily transcribed during spermatogenesis; the second step, instead, can comprehend sequences transcribed during oogenesis and accumulated into the oocyte, or by the zygote genome after maternal-zygotic transition, or both.

In order to uncover the molecular outline of mitochondrial inheritance, transcriptomic data from mature gonads of the SMI species *R. decussatus* and the DUI species *R. philippinarum* were analyzed, taking into account presence, transcription patterns, and mitochondrial targeting of all proteins belonging to pathways known to be involved in SMI achievement. Due to the nature of the available data, our research focused on the first step. Previous data (Ghiselli et al., 2012, Milani et al., 2013) show that, in *R. philippinarum* gonads, some sequences involved in the ubiquitination pathway are transcribed with a male bias, and in situ hybridization found some Ub-related transcripts localized in gametogenic cells, hinting at a possible implication of Ub system in DUI. A proteomic analysis on the DUI species *Mytilus edulis* (Diz et al., 2013) yielded similar results. Our analysis of transcripts belonging to nucleases/polymerases, autophagy and mitophagy, and ubiquitination pathway are consistent with pre-existing data, and allowed us to propose a model of SMI mechanism in *R. decussatus* and its modification in *R. philippinarum*.

2 | MATERIALS AND METHODS

2.1 | Dataset

RNA-Seq libraries were prepared from ripe gonads of twelve individuals (six females and six males) of *R. philippinarum* from the Pacific coast of USA (Puget Sound, WA), and twelve individuals (six females and six males) of *R. decussatus* from the Northern Adriatic Sea (Goro, Italy), following the protocols of Mortazavi, Williams, McCue, Schaeffer, and Wold (2008) with the modifications reported in Ghiselli et al. (2012). Raw reads and de novo assemblies of *R. philippinarum* and *R. decussatus* are available on NCBI (BioProjects PRJNA68513 and PRJNA170478, respectively). Details about sequencing, de novo assembly, and differential transcription analysis are described in Ghiselli et al. (2012), while statistics on the assemblies can be found in Supplementary data file S1. Differential transcription between males and females is expressed as the binary logarithm of the fold change of the transcription level [$\log_2$(FC)]; male-biased transcripts are defined as those for which $\log_2(FC) < -1$, whereas female-biased those for which $\log_2(FC) > 1$.

In order to perform a comparative analysis of the two transcriptomes, the de novo assemblies were annotated with a transcriptome annotation pipeline for non-model organisms (Ghiselli et al., in preparation; detailed information, data and scripts can be found at the following link: https://osf.io/2gdqe/?view_only=f0b2cde926db43719f3d705012c4eeaa).

Mitochondrial targeting of all the sequences belonging to both transcriptomes was assessed with TargetP (Emanuelsson, Brunak, von Heijne, & Nielsen, 2007).

2.2 | Data analysis

Following the literature on the subject, we narrowed our research to some "protein groups of interest" defined as follows: Ub-proteasome system (UPS) and Ub-like modifiers, mitophagy/autophagy, nucleases/DNA polymerases (Table 1). FPKM data of the all the retrieved sequences can be found in Supplementary data file S2 and S3.

Autophagy and mitophagy pathways rely on an evolutionarily conserved core machinery, and this has allowed us to compile lists of orthologs including all the proteins known to belong to these pathways. The sequences of the proteins included in such lists were used as queries in the searches against the transcriptomes of the two clam species. Conversely, proteins belonging to the groups of nucleases, DNA polymerases, and the UPS are part of multiple gene families varying in size and evolutionary history. As such, a gene-to-gene relationship with other species orthologs cannot be established. For this reason, we had to follow two different methods to retrieve loci of interest.

Orthologous sequences belonging to autophagy and mitophagy pathways in *Homo sapiens* and in the oyster *Crassostrea gigas* (the only bivalve species available) were downloaded from the KEGG database (Kanehisa and Goto, 2000). In order to present the most comprehensive results possible, proteins involved in both autophagy and mitophagy were retained in both datasets. These sequences were used as queries in a BLASTP (Camacho et al., 2009) search against databases built from *R. decussatus* and *R. philippinarum* transcriptomes. We filtered out the hits with an E-value above 1E-50, and we checked the remaining sequences. If a sequence showed similarity for orthologs in both *C. gigas* and *H. sapiens*, it was retained only if the similarity with the bivalve species had a stronger support (i.e., a lower *E*-value). If a sequence showed similarity with a *C. gigas* sequence, but did not have any hit against human orthologs, it was kept as well; the opposite cases—similarity with *H. sapiens* but not with *C. gigas*—were regarded as possible contaminants and discarded. The KO (KEGG Orthology)

TABLE 1 Overall sequences retrieved for each protein group of interest, comprehensive of transcription bias, orthology, and mitochondrial target

	Rde	Rph			Rde	Rph	
Ubiquitination				**Autophagy** (Total KO ids in KEGG: 100)			
Total sequences	778	728		KO ids	62	50	45 KO ids in common
Female biased	48	18		Total loci	124	92	
Male biased	28	20		Female biased	16	0	
mt target	38	42		Male biased	7	2	
Orthologs	471 (394)	450 (387)	381 Clusters in common	Orthologs	87 (59)	65 (59)	59 Clusters in common
Nucleases				**Mitophagy** (Total KO ids in KEGG: 57)			
Total sequences	277	230		KO ids	27	24	22 KO ids in common
Female biased	25	12		Total loci	46	35	
Male biased	16	4		Female biased	6	0	
mt target	24	13		Male biased	4	0	
Orthologs	154 (127)	131 (113)	109 Clusters in common	Orthologs	31 (29)	31 (29)	29 Clusters in common
Polymerases							
Total loci	284	266					
Female biased	19	7					
Male biased	14	6					
mt target	19	18					
Orthologs	173 (148)	147 (129)	128 Clusters in common				

Note: Orthologs, number of sequences that have one or more orthologs in the other species' transcriptome; in parentheses the number of sequences that have at least one ortholog with the same annotation and thus that belong to the clusters in common; clusters in common, ortholog clusters whose sequences have the same annotation in both species; KO ids, total KO identifiers with at least a corresponding sequence in the species – correspondence addressed in detail in Tables 3 and 4; Rde, *R. decussatus*; Rph, *R. philippinarum*.

identifier reported for the selected *C. gigas* and *H. sapiens* sequences was associated with each hit, so that exact correspondence with the KEGG reference pathways could be traced (Tables 2 and 3, and Supplementary data file S7–S10).

For UPS and nucleases/polymerases, instead, GO terms featuring the terms "ubiquitin," "proteasome," "nuclease," and "DNA polymerase" were selected from the GO database (Balakrishnan, Harris, Huntley, Van Auken, & Cherry, 2013; downloaded on 12 October 2016) and manually curated (Supplementary data file S4–S6). Sequences annotated with such GO terms were then extracted from the two transcriptomes (Supplementary data file S11–S16). Additionally, prohibitin sequences belonging to *C. elegans*, *Xenopus tropicalis*, *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus*, *Pongo abelli*, and *H. sapiens* were downloaded from UniProtKB (The UniProt Consortium, 2017) and were used to perform a local BLASTP search, which unanimously retrieved the two evolutionarily conserved subunits of prohibitin in both species.

3 | RESULTS AND DISCUSSION

3.1 | Nucleases and polymerases

We retrieved 277 sequences in *R. decussatus* and 230 sequences in *R. philippinarum* which were annotated with GO terms related to nuclease activity or polymerase activity (Table 1, Supplementary data file S4–S5, and S11–S14). These sequences were mostly involved in DNA repair (GO:0006281, "DNA repair", 56 occurrences in *R. decussatus* and 65 in *R. philippinarum*), but sequences involved in RNA retrotranscription were not uncommon (GO:0006278, "RNA-dependent DNA biosynthetic process," 37 and 24 occurrences, respectively), either annotated with transposon activity (according to BLASTP annotation, 23 and 14, respectively) or telomere maintenance (GO:0000723 "telomere maintenance" and child terms, 11 and 21 occurrences, respectively). The biological functions uncovered by the annotation are expected, given the high proliferation activity of cells in gametogenic gonads—obviously requiring both polymerases and nucleases—and the physiological quality-check role of telomere maintenance in mitosis and meiosis. If any endonuclease or polymerase were to enter male mitochondria in order to reduce mitochondrial nucleotide content during *R. decussatus* spermatogenesis as it happens in *O. latypes*, we expect that the candidate sequence would have both a male-biased transcription and a mitochondrial targeting presequence (Table 2). Regarding nucleases, several sequences possessing either one or the other feature have been retrieved, but none shows both (Figure 1A). As for polymerases, the great majority of sequences do not display a sex bias (Figure 1A), with only one female-biased contig per species and one strongly male-biased contig in *R. philippinarum* (−8.18397 $\log_2$(FC)), annotated as a "DNA polymerase nu-like," an error-prone polymerase involved in DNA damage repair.

TABLE 2 Summary of the assessed features of the proteins belonging to the pathways under study

	Endonucleases	Polymerases	Autophagy	Mitophagy	Ubiquitination
Proposed mode of action	Degrade mtDNA during spermatogenesis		Degrade mitochondria during spermatogenesis and/or after fertilization		Tags paternal mitochondria for degradation during spermatogenesis and/or after fertilization
Did we retrieve all the sequences necessary to enforce this pathway?	Yes	Yes	Almost all	Dubious	Yes
Is a transcriptional bias necessary? Male or female?	Yes – Male	Yes – Male	Yes – Could be both	Yes – Could be both	Yes – Male
Did we find sequences with such bias?	Yes, but lacking a mitochondrial presequence	Yes, but lacking a mitochondrial presequence	No	No	Yes
Does the resulting protein(s) have to enter the mitochondria (i.e., is a mitochondrial presequence necessary)?	Yes	Yes	No	No	No
Did we find sequences with the mitochondrial presequence?	Yes, but lacking a transcriptional bias	Yes, but lacking a transcriptional bias	N/A	N/A	N/A
Did we find sequences/groups of sequences with all the needed characteristics?	No	No	No	No	Yes

Details in the main text.

Our results are not consistent with a mechanism of nucleoid number reduction similar to that of *O. latypes* and some mammals; however, it has to be noted that mitochondrial targeting assessment is especially prone to false negatives due to the presence of import signals other than presequences, or to transcript length biases. More extensive research has to be performed to rule out the involvement of endonucleases in SMI enforcement in *R. decussatus*.

3.2 | Autophagy and mitophagy

Because of its high level of conservation across eukaryotes, autophagy is a particularly suitable pathway for transcriptomics studies in nonmodel species, so we were able to assess the completeness of autophagic supramolecular complexes by extracting autophagy-related orthologs from the two studied transcriptomes. The core components of autophagy are mostly present in both *R. decussatus* and *R. philippinarum*—for instance, GABARAP, an ortholog of yeast LC3, whose detection has been often used as a proxy for autophagy taking place (e.g. Kraft, Peter, & Hofmann, 2010; Jin and Klionsky, 2014). Moreover, most of the functional annotation of the sequences involved in both autophagy and mitophagy is in common between the two clam species (Table 1). Autophagy has been proved fundamental both for male and female gametogenesis, with roles ranging from regulation of signaling between follicle cells and oocytes in *Drosophila*, to correct acrosome formation in mouse spermatozoa (Barth, Hafen, & Köhler, 2012; Kanninen, de Andrade Ramos, & Witkin, 2013; Wang et al., 2014; Agnello, Chiarelli, Martino, Bosco, & Roccheri, 2016). In *R. philippinarum* transcriptome, only two sequences out of 92 display a sex-biased transcription; in *R. decussatus*, instead, there are 22 sex-biased sequences out of 124, representing almost one-fifth of the total number of sequences involved in autophagy in this species (Figure 1B). These sequences code mainly for regulatory enzymes and display predominantly a female bias (16 female-biased vs. 6 male-biased sequences; see Table 3 and Supplementary data file S7 and S8).

While these data suggest that autophagy-related genes are active at this stage in gonads, thus enabling the autophagy process, the same cannot be easily said for mitophagy: a core machinery for autophagy has been established with a wide consensus, whereas the molecular actors determining selective autophagy are more debated. A central mitophagic trigger mechanism revolves around the serine/threonine-protein kinase PINK1, which, upon attachment to the outer membrane of depolarized mitochondria, recruits the E3 Ub ligase Parkin for their degradation through mitophagy (Durcan & Fon 2015). Other Parkin-independent pathways have been defined as well; for instance, hypoxia triggers mitophagy through activation of Nix/Fundc1 pathway (Campello, Strappazzon, & Cecconi, 2014; Georgakopoulos, Wells, & Campanella, 2017). Even if roughly half of the sequences involved in the mitophagy pathway are present in both species, most of the fundamental ones are missing in both species (i.e., Parkin and the initiators of hypoxia-induced mitophagy FOXO3, Fundc1, Bnip3, and Bnip3L/Nix; see Table 4), while Ambra1, an effector of a hypothesized Parkin-independent mitophagy pathway, and PINK1 are present only in *R.*

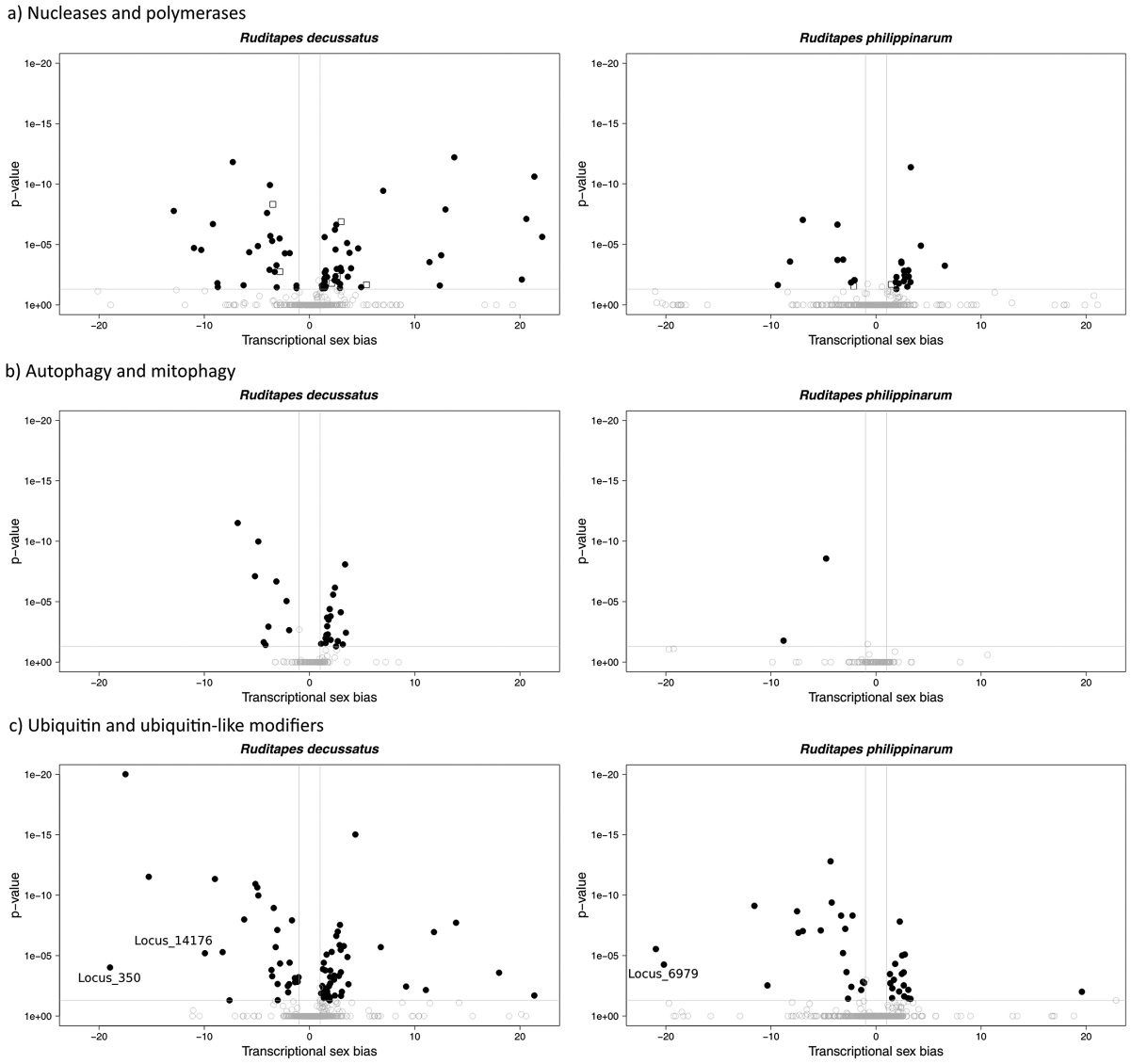


FIGURE 1 Distribution of *Ruditapes decussatus* and *Ruditapes philippinarum* loci according to the statistical significance and the transcriptional sex bias, expressed as the binary logarithm of the fold change of the transcription level. The horizontal gray line marks the significance threshold ($P = 0.05$), whereas the vertical gray lines mark the transcriptional sex bias threshold (see Materials and methods). (A) Loci annotated as nucleases and polymerases; the loci represented with an empty square possess a mitochondrial presequence; (B) loci annotated as belonging to autophagy and/or mitophagy pathway; (C) loci annotated as belonging to the ubiquitination or ubiquitination-like pathways

philippinarum. Moreover, a female-biased transcription of the retrieved mitophagy-associated genes in *R. decussatus*—even if weak (Figure 2)—point out to an inhibition of mitophagy rather than an activation (for a review on mitophagy regulation refer to Hamacher-Brady and Brady, 2016).

We can hypothesize at least two different mechanisms for SMI enforcement through mitophagy/autophagy (Table 2). On one hand, mitophagy could have a role in reducing nucleoid number during spermatogenesis. As data do not point out male-biased transcription of any of the sequences, it appears that this mechanism is not put in place in *R. decussatus*. On the other hand, male mitochondria could be digested after fertilization, as in studied mammals and *C. elegans*. If this is the case, we could reasonably expect an accumulation of autophagy- and mitophagy-related transcripts in oocytes, resulting in a female bias. However, with the exception of the already discussed bias regarding regulatory sequences, no other strong female bias has emerged. Still, this mechanism could take place after the maternal-zygotic transition and be due to zygotic transcripts (Schier, 2007), but in order to further elucidate this point, different developmental stages should be assessed for the presence of this pathway.

3.3 | Ubiquitination and Ub-like modifiers

We retrieved 778 and 728 ubiquitination-related sequences in *R. decussatus* and *R. philippinarum*, respectively (Table 1, Supplementary data file S6 and S15–S16, and Figure 1C). As the name of the

TABLE 3 Proteins Involved in Autophagy in *R. decussatus* (Rde) and *R. philippinarum* (Rph)

KOid	Name	Rde	Rph	KOid	Name	Rde	Rph
K00914	PIK3C3	o	o	K08270	DDIT4	x	x
K00922	PIK3CA_B_D	o	x	K08331	ATG13	o	o
K01110	PTEN	o	o	K08333	PIK3R4	o	o
K01363	CTSB	o	o	K08334	BECN	o	o
K01365	CTSL	o	o	K08336	ATG12	o	x
K01379	CTSD	o	o	K08337	ATG7	o	o
K02158	BAD	x	x	K08339	ATG5	o	o
K02161	BCL2	x	x	K08341	GABARAP	o	o
K02649	PIK3R1_2_3	o	x	K08342	ATG4	o	o
K02833	HRAS	x	x	K08343	ATG3	o	o
K03175	TRAF6	o	o	K08491	STX17	x	x
K03237	EIF2S1	o	o	K08509	SNAP29	x	o
K04345	PKA	o	o	K08512	VAMP8	x	x
K04366	RAF1	x	x	K08803	DAPK	x	x
K04368	MAP2K1	o	o	K08852	ERN1	o	x
K04369	MAP2K2	x	x	K08860	EIF2AK3	o	x
K04371	MAPK1_3	o	o	K10802	HMGB1	x	x
K04382	PPP2C	o	o	K11248	SH3GLB1	o	x
K04427	MAP3K7	o	o	K15464	BNIP3	x	x
K04440	JNK	o	o	K16172	IRS1	x	x
K04456	AKT	o	o	K16184	AKT1S1	x	x
K04526	INS	x	x	K16185	RRAGA_B	x	o
K04570	BCL2L1	x	o	K16186	RRAGC_D	o	x
K04688	RPS6KB	o	o	K16196	EIF2AK4	o	o
K04724	CFLAR	x	x	K17445	IRS3	x	x
K04958	ITPR1	o	o	K17446	IRS4	x	x
K05087	IGF1R	x	x	K17589	RB1CC1	o	x
K06068	PRKCD	o	o	K17603	ZFYVE1	x	x
K06276	PDPK1	o	o	K17606	IGBP1	o	o
K06528	LAMP1_2	x	x	K17888	ATG10L	x	x
K07187	IRS2	x	x	K17889	ATG14L	o	x
K07198	PRKAA	o	o	K17890	ATG16L1	o	o
K07203	MTOR	o	o	K17906	ATG2	o	o
K07204	RAPTOR	o	o	K17907	ATG9	o	x
K07206	TSC1	o	o	K17908	WIPI	o	o
K07207	TSC2	o	x	K17985	AMBRA1	o	x
K07208	RHEB	o	o	K18052	PRKCQ	x	x
K07298	STK11	o	o	K18082	MTMR3_4	o	o
K07359	CAMKK2	x	x	K18086	MTMR14	o	x
K07827	KRAS	o	o	K19330	RUBCN	o	x
K07828	NRAS	x	x	K19730	ATG101	o	o
K07829	RRAS	x	x	K20402	DEPTOR	x	x
K07830	RRAS2	o	x	K20868	ATG16L2	x	x
K07831	MRAS	x	o	K21245	SUPT20H	o	x
K07897	RAB7A	o	o	K21246	NRBF2	x	o
K07898	RAB7B	x	x	K21247	TP53INP2	x	x
K07920	RAB33B	o	o	K21248	VMP1	o	o
K08266	MLST8	o	o	K21249	UVRAG	o	x
K08268	HIF1A	x	x	K21250	PRAP1	x	x
K08269	ULK2	o	o	K21357	ULK1	x	x

KOids, KEGG Orthology entries; Name, common name of the ortholog group; o, presence; x, absence.

TABLE 4 Proteins Involved in Mitophagy in *R. decussatus* (Rde) and *R. philippinarum* (Rph)

KOid	Name	Rde	Rph	KOid	Name	Rde	Rph
K02833	HRAS	x	x	K08341	GABARAP	o	o
K03097	CSNK2A	x	x	K08860	EIF2AK3	o	x
K03115	CSNK2B	o	o	K09105	TFE3	x	x
K04374	ATF4	x	x	K09455	MITF	o	o
K04440	JNK	o	o	K11839	USP8	o	o
K04448	JUN	x	o	K11851	USP30	o	o
K04451	TP53	x	x	K14381	SQSTM1	o	o
K04551	UBB	x	x	K15485	BCL2L13	x	x
K04570	BCL2L1	x	o	K15590	TFEB	x	x
K04684	SP1	x	x	K15637	PGAM5	o	o
K04735	RELA	x	x	K17454	E2F1	x	x
K05410	TBK1	o	o	K17771	TOM7	x	x
K05704	SRC	o	o	K17907	ATG9	o	x
K06030	MFN2	o	o	K17969	FIS1	o	x
K07827	KRAS	o	o	K17985	AMBRA1	o	x
K07828	NRAS	x	x	K17987	NBR1	o	o
K07829	RRAS	x	x	K19945	TBC1D17	x	x
K07830	RRAS2	o	o	K19946	OPTN	o	o
K07831	MRAS	x	x	K20168	TBC1D15	o	o
K07870	RHOT1	o	o	K21343	USP15	o	o
K07871	RHOT2	x	x	K21347	TAX1BP1	o	o
K07897	RAB7A	o	o	K21348	CALCOCO2	x	x
K07898	RAB7B	x	x	K21356	MFN1	x	x
K08268	HIF1A	x	x	K21357	ULK1	x	x
K08334	BECN	o	o	K21361	CITED2	x	x
K08339	ATG5	o	o				

KOids, KEGG Orthology entries; Name, common name of the ortholog group; o, presence; x, absence.

pathway itself suggests, it is one of the most ubiquitous mechanism for routinely protein quality control within cells. As such, several E1, E2, E3, and DUBs were retrieved (Table 5). Moreover, ubiquitination covers specialized roles during gametogenesis, especially in males (for thorough reviews see: Richburg, Myers, & Bratton, 2014; Suresh, Lee, Kim, & Ramakrishna, 2016). In mammals, one of such roles is to provide sperm mitochondria with degradation signals by di-ubiquitinating the mitochondrial membrane protein prohibitin (Sutovsky et al., 2000). A similar pattern of prohibitin ubiquitination, even if with a slightly different timing, appears to extend to species outside the mammalian taxon: for instance, in the crayfish *Procambarus clarkii* prohibitin, Ub,

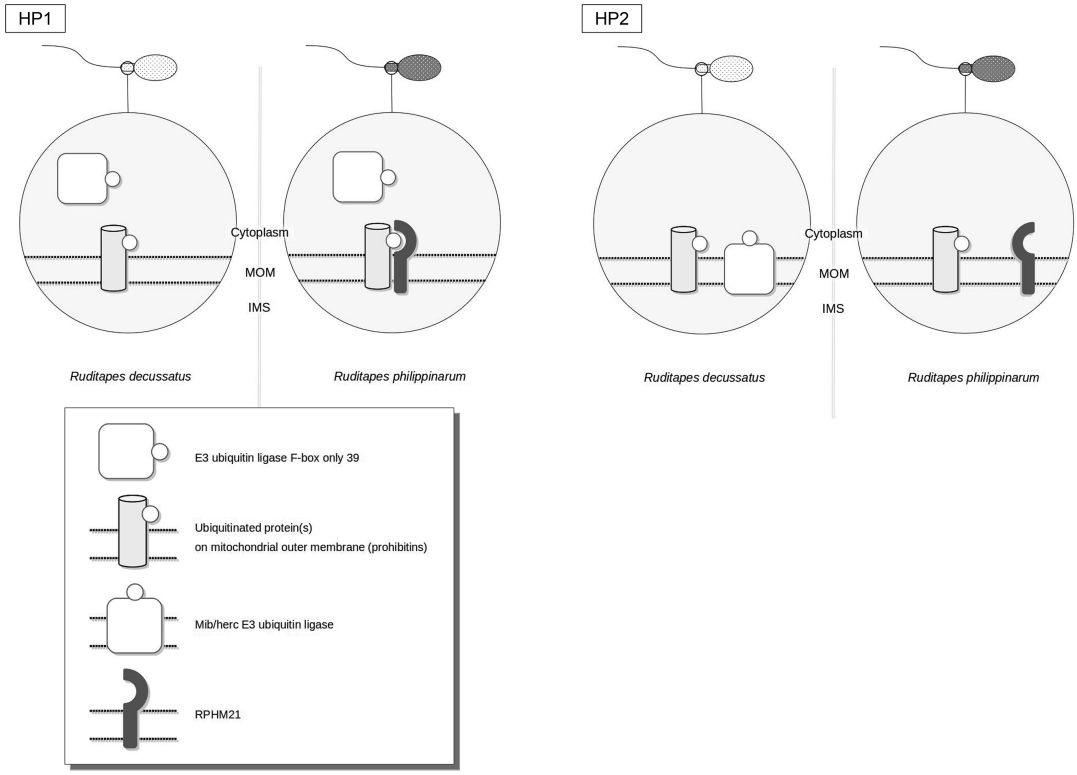


FIGURE 2 Representation of the ubiquitination state in mitochondria of spermatozoa in both clam species, according to the two hypotheses presented in this study. Hypothesis 1 (HP1): During spermatogenesis of both species, prohibitins or other proteins on the mitochondrial outer membrane are ubiquitinated by a "F-box only 39" E3. RPHM21 then masks the recognition/degradation signal in *R. philippinarum*, hindering mitochondria destruction after fertilization. Hypothesis 2 (HP2): RPHM21 is involved in processes other than masking the recognition signal, such as gonad differentiation or determination of the mitochondria aggregation pattern (Milani et al., 2014b). The ubiquitinating enzyme is a transmembrane mib/herc E3 in *R. decussatus* and is absent in *R. philippinarum*. MOM, mitochondrial outer membrane; IMS, intermembrane space

and mitochondria co-localize in late spermatogenesis (Dong, Hou, & Yang, 2015).

Prohibitins have been retrieved in the analyzed clam species as well. Given the evolutionary conservation of ubiquitination of prohibitin during spermatogenesis, it is conceivable that they might play a role in paternal mitochondria recognition as in mammals.

Given the high substrate specificity of E3 Ub ligases and their high recurrence in the two transcriptomes (258 in *R. decussatus* and 237 in *R. philippinarum* according to GO term annotation—see Table 5), we expect the candidate sequences to show a strongly male-biased transcription level, if not a male-specific transcription (Table 2).

TABLE 5 Estimate of the number of enzymes involved in the ubiquitination pathway according to the GO annotation

	Ruditapes decussatus	Ruditapes philippinarum
E1 – Ub-activating enzymes	5	5
E2 – Ub-conjugating enzymes	7	7
E3 – Ub-ligases	258	237
Deubiquitinating enzymes	57	61
Proteasome	153	144

In order to explain the different mitochondrial inheritance outcomes between the two species investigated here, we propose two hypotheses (Figure 2). There is some speculation in such hypotheses, but they are all consistent with the available data and can be useful to guide future experiments and research by providing candidate targets for further investigation.

3.3.1 | Hypothesis 1

Effectiveness of degradation through ubiquitination relies on the recognition of Ub moieties linked to the target. If the ubiquitination signal is persistent in both species, it has to be masked in *R. philippinarum* in order to achieve DUI. A candidate for this role is RPHM21, a protein encoded by a male-specific mitochondrial ORF transcribed and translated during spermatogenesis, localized in sperm mitochondria and nuclei, and in embryos (Milani et al., 2014b, 2015, 2016). Its main putative features are two transmembrane helices, a binding site for Ub, and domains involved in cytoskeleton interactions. As already hypothesized in Milani et al. (2014b), RPHM21 might prevent the recognition of the degradation signal on the male mitochondria by binding to ubiquitinated mitochondrial proteins (for instance, prohibitin dimers) through their Ub binding site. Indeed, male mitochondria are not degraded before the 32-blastomere stage in all

R. philippinarum embryos observed, irrespective of the aggregation pattern (Milani et al., 2014b), so RPHM21 protection mechanism could delay degradation of sperm mitochondria independently from the sex of the embryos. If this is the case, the E3 Ub ligase performing this task may be conserved in both species and show a male-biased transcription. Such features, indeed, apply to two sequences (identified as Locus_350 in *R. decussatus* and Locus_6979 in *R. philippinarum*, see Figure 1C and Figure 2-HP1) that belong to the same ortholog cluster, both undetectable in female gonads—designating them as male specific—and both annotated as "F-box only protein 39," a substrate recognition component of the SCF (Skp1/Cullin/F-box) complex, a family of modular E3 ligases.

3.3.2 | Hypothesis 2

On the other hand, if the membrane protein carrying the male recognition signal is unmasked also in *R. philippinarum* (i.e., no masking by RPHM21 or other factors), then the difference between the two species could lie instead in the ubiquitination pattern. Hence, ubiquitination in the SMI species could be performed by an E3 Ub ligase whose ortholog is either absent or transcriptionally downregulated/silenced in *R. philippinarum*, resulting in a male-biased sequence in *R. decussatus* lacking an ortholog in the other species. This description delineates the characteristics of several *R. decussatus* male-biased sequences. Although most of them are either involved in cell cycle maintenance or have a relatively weak male bias, the most transcriptionally biased one is a sequence (identified as Locus_14176, see Figure 1C) containing a mib/herc2 domain (a Ub ligase domain; PF06701) also annotated with GO:0016020 "membrane" and GO:0016021 "integral component of membrane". Studies suggest that transmembrane E3 substrates are preferentially transmembrane proteins themselves (Bauer, Bakke, & Morth, 2016). This E3 might ubiquitinate a male recognition protein on the mitochondrial outer membrane of the SMI species *R. decussatus* targeting sperm mitochondria for degradation.

4 | CONCLUSIONS

We can detail the process of paternal mitochondria degradation in animals as composed of two steps: (1) during spermatogenesis—degradation of nucleoids and/or marking of paternal mitochondria as means to distinguish them from maternal ones; and (2) after fertilization—degradation of nucleoids or paternal mitochondria.

The sequences encoding the machinery for the first step have to be necessarily transcribed during spermatogenesis; the second step, instead, can comprehend sequences transcribed during oogenesis, or after maternal-zygotic transition, or both. The transcriptomic data here analyzed, portraying late gametogenesis of the two bivalve species *R. decussatus* and *R. philippinarum*, allowed us to hypothesize which processes and genes might be involved in the first step, and which might be the molecular similarities and differences underlying the two different inheritance outcomes (Figure 2).

We propose two hypotheses (Figure 2): (1) the degradation signal present on the mitochondrial outer membrane (which could be represented by ubiquitinated prohibitins) is masked in the zygote (e.g., by RPHM21), so the enzyme responsible for such degradation labeling must be present in both *R. decussatus* and *R. philippinarum*. Two male-specific ortholog sequences annotated as "F-box only protein 39," an E3 Ub ligase, show characteristics which are compatible with this hypothesis; (2) the difference lies in the labeling pattern being absent or delayed in the DUI species. A transmembrane E3 Ub ligase with a strong male bias, retrieved in *R. decussatus* and with no apparent ortholog in *R. philippinarum*, is a good candidate to perform this task.

As for the second step, that is degradation of paternal mitochondria after fertilization, it may involve proteins transcribed after the maternal-zygotic transition, so further research involving developing embryos is needed to clarify this point.

Future perspectives include immunological analyses on sperm and zygotes of both species, and investigating localization and interaction among prohibitin/Ub and the other suggested candidate proteins will help defining the described mechanisms.

CONFLICT OF INTEREST

None.

ORCID

Liliana Milani http://orcid.org/0000-0001-5052-2075
Fabrizio Ghiselli http://orcid.org/0000-0002-1680-8616

REFERENCES

Agnello, M., Chiarelli, R., Martino, C., Bosco, L., & Roccheri, M. C. (2016). Autophagy is required for sea urchin oogenesis and early development. *Zygote*, *24*, 918–926.

Balakrishnan, R., Harris, M. A., Huntley, R., Van Auken, K., & Cherry, J. M. (2013). A guide to best practices for Gene Ontology (GO) manual annotation. *Database*, *2013*, bat054.

Barth, J. M. I., Hafen, E., & Köhler, K. (2012). The lack of autophagy triggers precocious activation of Notch signaling during *Drosophila* oogenesis. *BMC Developmental Biology*, *12*, 35.

Bauer, J., Bakke, O., & Morth, J. P. (2016). Overview of the membrane-associated RING-CH (MARCH) E3 ligase family. *New Biotechnology*, [Internet]. Retrieved from https://doi.org/10.1016/j.nbt.2016.12.002

Bettinazzi, S., Plazzi, F., & Passamonti, M. (2016). The complete female- and male-transmitted mitochondrial genome of *Meretrix lamarckii*. *PLoS One*, *11*, e0153631.

Birky, C. W. Jr. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 11331–11338.

Birky, C. W. Jr (2001). The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models. *Annual Review of Genetics*, *35*, 125–148.

Breton, S., Ghiselli, F., Passamonti, M., Milani, L., Stewart, D. T., & Hoeh, W. R. (2011a). Evidence for a fourteenth mtDNA-encoded protein in the female-transmitted mtDNA of marine Mussels (Bivalvia: Mytilidae). *PLoS One*, *6*, e19365.

Breton, S., Stewart, D. T., Shepardson, S., Trdan, R. J., Bogan, A. E., Chapman, E. G., … Hoeh, W. R. (2011b). Novel protein genes in animal mtDNA: A new sex determination system in freshwater mussels (Bivalvia: Unionoida)? *Molecular Biology and Evolution*, *28*, 1645–1659.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421.

Campello, S., Strappazzon, F., & Cecconi, F. (2014). Mitochondrial dismissal in mammals, from protein degradation to mitophagy. *Biochimica et Biophysica Acta*, *1837*, 451–460.

Cao, L., Kenchington, E., & Zouros, E. (2004). Differential segregation patterns of sperm mitochondria in embryos of the blue mussel (*Mytilus edulis*). *Genetics*, *166*, 883–894.

Ciehanover, A., Hod, Y., & Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communication*, *81*, 1100–1105.

Diz, A. P., Dudley, E., Cogswell, A., MacDonald, B. W., Kenchington, E. L. R., Zouros, E., & Skibinski, D. O. F. (2013). Proteomic analysis of eggs from *Mytilus edulis* females differing in mitochondrial DNA transmission mode. *Molecular and Cellular Proteomics*, *12*, 3068–3080.

DeLuca, S. Z., & O'Farrell, P. H. (2012). Barriers to male transmission of mitochondrial DNA in sperm development. *Developmental Cell*, *22*, 660–668.

Dong, W.-L., Hou, C.-C., & Yang, W.-X. (2015). Mitochondrial prohibitin and its ubiquitination during crayfish *Procambarus clarkii* spermiogenesis. *Cell Tissue Research*, *359*, 679–692.

Doucet-Beaupré, H., Breton, S., Chapman, E. G., Blier, P. U., Bogan, A. E., Stewart, D. T., & Hoeh, W. R. (2010). Mitochondrial phylogenomics of the Bivalvia (Mollusca): Searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. *BMC Evolutionary Biology*, *10*, 50.

Durcan, T. M., & Fon, E. A. (2015). The three 'P's of mitophagy: PARKIN, PINK1 and post-translational modifications. *Genes & Development*, *29*, 989–999.

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocol*, *2*, 953–971.

Georgakopoulos, N. D., Wells, G., & Campanella, M. (2017). The pharmacological regulation of cellular mitophagy. *Nature Chemical Biology*, *13*, 136–146.

Ghiselli, F., Milani, L., Chang, P. L., Hedgecock, D., Davis, J. P., Nuzhdin, S. V., & Passamonti, M. (2012). *De Novo* assembly of the Manila clam *Ruditapes philippinarum* transcriptome provides new insights into expression bias, mitochondrial doubly uniparental inheritance and sex determination. *Molecular Biology and Evolution*, *29*, 771–786.

Ghiselli, F., Milani, L., & Passamonti, M. (2011). Strict sex-specific mtDNA segregation in the germ line of the DUI species *Venerupis philippinarum* (Bivalvia: Veneridae). *Molecular Biology and Evolution*, *28*, 949–961.

Ghiselli, F., Milani, L., Iannello, M., Procopio, E., Chang, P. L., Nuzhdin, S. V., & Passamonti, M. (2017). The complete mitochondrial genome of the grooved carpet shell, *Ruditapes decussatus* (Bivalvia, Veneridae). *PeerJ*, *5*, e3692.

Gusman, A., Lecomte, S., Stewart, D. T., Passamonti, M., & Breton, S. (2016). Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ*, *4*, e2760.

Hamacher-Brady, A., & Brady, N. R. (2016). Mitophagy programs: Mechanisms and physiological implications of mitochondrial targeting by autophagy. *Cellular and Molecular Life Sciences*, *73*, 775–795.

Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annual Review of Biochemistry*, *67*, 425–479.

Jin, M., & Klionsky, D. J. (2014). Regulation of autophagy: Modulation of the size and number of autophagosomes. *FEBS Letters*, *588*, 2457–2463.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*, 27–30.

Kanninen, T. T., de Andrade Ramos, B. R., & Witkin, S. S. (2013). The role of autophagy in reproduction from gametogenesis to parturition. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *171*, 3–8.

Kenchington, E., MacDonald, B., Cao, L., Tsagkarakis, D., & Zouros, E. (2002). Genetics of mother-dependent sex ratio in blue mussels (Mytilus spp.) and implications for doubly uniparental inheritance of mitochondrial DNA. *Genetics*, *161*, 1579–1588.

Kraft, C., Peter, M., & Hofmann, K. (2010). Selective autophagy: Ubiquitin-mediated recognition and beyond. *Nature Cell Biology*, *12*, 836–841.

Lemasters, J. J. (2014). Variants of mitochondrial autophagy: Types 1 and 2 mitophagy and micromitophagy (Type 3). *Redox Biology*, *2*, 749–754.

Luo, S-M., Ge, Z-J., Wang, Z-W., Jiang, Z-Z., Wang, Z-B., Ouyang, Y-C., … Sun, Q.-Y. (2013). Unique insights into maternal mitochondrial inheritance in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 13038–13043.

May-Panloup, P., Chrétien, M-F., Savagner, F., Vasseur, C., Jean, M., Malthièry, Y., & Reynier, P. (2003). Increased sperm mitochondrial DNA content in male infertility. *Human Reproduction*, *18*, 550–556.

Milani, L., Ghiselli, F., Guerra, D., Breton, S., & Passamonti, M. (2013). A comparative analysis of mitochondrial ORFans: New clues on their origin and role in species with doubly uniparental inheritance of mitochondria. *Genome Biology and Evolution*, *5*, 1408–1434.

Milani, L., Ghiselli, F., Iannello, M., & Passamonti, M. (2014a). Evidence for somatic transcription of male-transmitted mitochondrial genome in the DUI species *Ruditapes philippinarum* (Bivalvia: Veneridae). *Current Genetics*, *60*, 163–173.

Milani, L., Ghiselli, F., Maurizii, M. G., Nuzhdin, S. V., & Passamonti, M. (2014b). Paternally transmitted mitochondria express a new gene of potential viral origin. *Genome Biology and Evolution*, *6*, 391–405.

Milani, L., Ghiselli, F., & Passamonti, M. (2012). Sex-linked mitochondrial behavior during early embryo development in *Ruditapes philippinarum* (Bivalvia Veneridae) a species with the Doubly Uniparental Inheritance (DUI) of mitochondria. *Journal of Experimental Zoology B: Molecular and Developmental Evolution*, *318*, 182–189.

Milani, L., Ghiselli, F., & Passamonti, M. (2016). Mitochondrial selfish elements and the evolution of biological novelties. *Current Zoology*, *62*, 687–697.

Milani, L., Ghiselli, F., Pecci, A., Maurizii, M. G., & Passamonti, M. (2015). The expression of a novel mitochondrially-encoded gene in gonadic precursors may drive paternal inheritance of mitochondria. *PLoS One*, *10*, e0137468.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*, 621–628.

Nishimura, Y., Yoshinari, T., Naruse, K., Yamada, T., Sumi, K., Mitani, H., … Kuroiwa, T. (2006). Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers. *Proceedings of the National Academy of Sciences United States of America*, *103*, 1382–1387.

Obata, M., Sano, N., & Komaru, A. (2011). Different transcriptional ratios of male and female transmitted mitochondrial DNA and tissue-specific expression patterns in the blue mussel, *Mytilus galloprovincialis*. *Development, Growth and Differentiation*, *53*, 878–886.

Oku, M., & Sakai, Y. (2016). Pexophagy in yeasts. *Biochimica et Biophysica Acta*, *1863*, 992–998.

Passamonti, M., & Scali, V. (2001). Gender-associated mitochondrial DNA heteroplasmy in the venerid clam *Tapes philippinarum* (Mollusca Bivalvia). *Current Genetics*, *39*, 117–124.

Pfeifer, U., & Scheller, H. (1975). A morphometric study of cellular autophagy including diurnal variations in kidney tubules of normal rats. *Journal of Cell Biology*, *64*, 608–621.

Plazzi, F., Cassano, A., & Passamonti, M. (2015). The quest for Doubly Uniparental Inheritance in heterodont bivalves and its detection in *Meretrix lamarckii* (Veneridae: Meretricinae). *Journal of Zoological Systematics and Evolutionary Research*, *53*, 87–94.

Politi, Y., Gal, L., Kalifa, Y., Ravid, L., Elazar, Z., & Arama, E. (2014). Paternal mitochondrial destruction after fertilization is mediated by a common endocytic and autophagic pathway in *Drosophila*. *Developmental Cell*, *29*, 305–320.

Richburg, J. H., Myers, J. L., & Bratton, S. B. (2014). The role of E3 ligases in the ubiquitin-dependent regulation of spermatogenesis. *Seminars in Cell Development Biology*, *30*, 27–35.

Rojansky, R., Cha, M.-Y., & Chan, D. C. (2016). Elimination of paternal mitochondria in mouse embryos occurs through autophagic degradation dependent on PARKIN and MUL1. Elife [Internet] 5. Retrieved from https://doi.org/10.7554/eLife.17896

Saavedra, C., Reyero, M. I., & Zouros, E. (1997). Male-dependent doubly uniparental inheritance of mitochondrial DNA and female-dependent sex-ratio in the mussel *Mytilus galloprovincialis*. *Genetics*, *145*, 1073–1082.

Sato, M., & Sato, K. (2011). Degradation of paternal mitochondria by fertilization-triggered autophagy in *C. elegans* embryos. *Science*, *334*, 1141–1144.

Sato, M., & Sato, K. (2013). Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochimica et Biophysica Acta*, *1833*, 1979–1984.

Schier, A. F. (2007). The maternal-zygotic transition: Death and birth of RNAs. *Science*, *316*, 406–407.

Skibinski, D. O., Gallagher, C., & Beynon, C. M. (1994a). Mitochondrial DNA inheritance. *Nature*, *368*, 817–818.

Skibinski, D. O., Gallagher, C., & Beynon, C. M. (1994b). Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. *Genetics*, *138*, 801–809.

Suresh, B., Lee, J., Kim, K-S., & Ramakrishna, S. (2016). The importance of ubiquitination and deubiquitination in cellular reprogramming. *Stem Cells International*, *2016*, 6705927.

Sutovsky, P., Mc Cauley, T. C., Sutovsky, M., & Day, B. N. (2003). Early degradation of paternal mitochondria in domestic pig (*Sus scrofa*) is prevented by selective proteasomal inhibitors lactacystin and MG(1321). *Biology of Reproduction*, *68*, 1793–1800.

Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., & Schatten, G. (2000). Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. *Biology of Reproduction*, *63*, 582–590.

Swatek, K. N., & Komander, D. (2016). Ubiquitin modifications. *Cell Research*, *26*, 399–422.

Theologidis, I., Fodelianakis, S., Gaspar, M. B., & Zouros, E. (2008). Doubly uniparental inheritance (DUI) of mitochondrial DNA in *Donax trunculus* (Bivalvia: Donacidae) and the problem of its sporadic detection in Bivalvia. *Evolution*, *62*, 959–970.

The UniProt Consortium. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, *45*, D158–D169.

Wang, H., Wan, H., Li, X., Liu, W., Chen, Q., Wang, Y., … Li, W. (2014). Atg7 is required for acrosome biogenesis during spermatogenesis in mice. *Cell Research*, *24*, 852–869.

Yusa, Y., Breton, S., & Hoeh, W. R. (2013). Population genetics of sex determination in *Mytilus* mussels: Reanalyses and a model. *Journal of Heredity*, *104*, 380–385.

Yu, Z., O'Farrell, P. H., Yakubovich, N., & DeLuca, S. Z. (2017). The mitochondrial DNA polymerase promotes elimination of paternal mitochondrial genomes. *Current Biology*, [Internet]. Retrieved from https://doi.org/10.1016/j.cub.2017.02.014

Zouros, E. (2013). Biparental inheritance through uniparental transmission: The doubly uniparental inheritance (DUI) of mitochondrial DNA. *Evolutionary Biology*, *40*, 1–31.

Zouros, E., Ball, A. O., Saavedra, C., & Freeman, K. R. (1994a). Mitochondrial DNA inheritance. *Nature*, *368*, 818.

Zouros, E., Oberhauser Ball, A., Saavedra, C., & Freeman, K. R. (1994b). An unusual type of mitochondrial DNA inheritance in the blue mussel Mytilus. *Proceedings of National Academy of Sciences of the United States of America*, *91*, 7463–7467.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Punzi E, Milani L, Ghiselli F, Passamonti M. Lose it or keep it: (how bivalves can provide) insights into mitochondrial inheritance mechanisms. *J Exp Zool (Mol Dev Evol)*. 2018;330:41–51. https://doi.org/10.1002/jez.b.22788

Chapter two

Cloning and *in vitro production* of RPHM21

In collaboration with the teams Miamor (PI: Stéphen Manon) and SysTEMM (PI: Marie-France Giraud), IBGC, Bordeaux

1. Introduction

In almost all animal species, mitochondria are inherited exclusively from the mother, a pattern that goes under the name of strictly maternal inheritance or SMI. However, a growing group of species more or less distantly related in the class Bivalvia constitutes an exception to this rule. In fact, in these species males transmit their mitochondrial DNA to their progeny, a peculiar inheritance pattern known as doubly uniparental inheritance or DUI. It is called "doubly uniparental" and not simply "biparental" because the two mitochondrial DNA constitute two separate lineages that have evolved independently for as much as 200 million years (Zouros, 2013). This evolutionary distance is well reflected in their diverging sequences, that can bear an amino acid p-distance as high as 51% (Zouros 2013). The sex-specific mtDNAs are inherited independently, each one through the gametes of the corresponding sex. After fertilization, a curious phenomenon is observed (Cao, Kenchington, & Zouros, 2004): the paternal (M) mitochondria can be tightly packed and remain in one cell during the first cellular divisions, entering blastomere 4d, from which germ cell originate, or can be dispersed, distributing randomly in the zygote. In *Mytilus* these two patterns were associated with the sex of the offspring: male in the former case and female in the latter (Cao et al., 2004).

Probably one of the most striking peculiarities of DUI is that each sex-specific mtDNA contains an ORFan (open reading frame having no detectable sequence similarity to other known proteins) of yet unknown function (Milani et al., 2013). There is some evidence concerning the transcription and as well the translation of the supernumerary ORF in M mtDNA of *Ruditapes philippinarum* (Ghiselli et al., 2013; Milani, Ghiselli, Maurizii, Nuzhdin, & Passamonti, 2014), in female (F) mtDNA of *Venustachonca ellipsiformis (Breton, Beaupré*, Stewart, Hoeh, & Blier, 2007; Breton et al., 2009), and in F mtDNA of *Mytilus edulis* (Ouimet et al., 2019). In *R. philippinarum*, the M-mtDNA-specific protein, known as RPHM21, is translated in the germline during gametogenesis, and its expression levels progress along with the spermatogenesis. In mature spermatozoa, it is found in mitochondria and in nuclei (Milani, Ghiselli, Maurizii, et al., 2014). Recently, the male mtDNA has been found in early female gametogenesis as well,

defying the idea that females are homoplasmic for F-mtDNA and hinting at the possibility that RPHM21 might be present during early stages of oogenesis (Ghiselli et al., 2019).

What is the function of the ORFans, ORF with unknown functions and no detectable homology, present in DUI mtDNA? The attempts to solve this puzzle have just added more mystery to it (Milani et al., 2013). The *in silico* predictions on their structure, domains and ultimately functions have revealed a complex situation. The ORFans do not bear much sequence similarity to each other, to the point that, made exception for a few very closely related species, they cannot be aligned to each other. However, they do not resemble anything else either, making a clear-cut prediction of their function impossible. These proteins do not seem to share any obvious evolutionary history, and Milani, Ghiselli, Guerra, et al. (2013), using multiple *in silico* approaches for a comparative analysis of DUI mitochondrial ORFans, proposed their origin through viral endogenization. This could also explain the scattered distribution of DUI in the bivalve phylogenetic tree.

The molecular mechanisms underlying DUI are still unknown; however, many structural and functional features of M and F mtDNAs were proposed as candidates for a role in mitochondrial inheritance and germ line establishment/differentiation (Ghiselli et al. 2013; Milani, Ghiselli, Guerra, et al. 2013; Zouros 2013). Among these candidates, the novel lineage-specific ORFs found so far in DUI species belonging to the families Unionidae (Breton et al. 2009, 2011a), Mytilidae (Breton et al. 2011b), and Veneridae (Ghiselli et al. 2013) were proposed. The existence of the translation product was verified in the unionid *Venustaconcha ellipsiformis* (Breton et al. 2009, 2011a) and in *R. philippinarum* (Milani L., Ghiselli F., Pecci A., Maurizii M.G., Passamonti M. 2015).

In this study, to gain some information on the function of the ORFans, I attempt to produce RPHM21, the protein encoded by the male-specific ORFan of *R. philippinarum* M-mtDNA, through an *in vitro* cell-free protein expression system and in yeast.

2. Materials and methods

2.1. Plasmid construction

A pEX-A128 plasmid containing the *rphm21* gene (flanked by ApaI and XhoI restriction sites) was purchased from Eurofins genomics. The sequence of *rphm21* was reencoded using http://genomes.urv.es/OPTIMIZER/ in order to respect the codon usage of both *E. coli* and *S. cerevisiae* (Annex 1). The gene was amplified by PCR using forward primer 5'-GGGGGGCATATGGTCTGGGTCGCCGTCGCC-3' and reverse primer 5'-GGGGGGCCCGGGTTAGTTGGAGTCGGGGTCGTCC-3' for subsequent insertion in pIVEX and forward primer 5'-GGGGGGGGCCCGTCTGGGTCGCCGTCGCC-3' and reverse primers 5'-GGGGGGCTCGAGTTAGTTGGAGTCGGGGTCGTCC-3' and 5'-GGGGGGCTCGAGTTAATGATGATGATGATGATGGTTGGAGTCGGGGTCGTCC-3' for insertion in pESC-His plasmids without and with a hexahistidine tag respectively, in frame with the mitochondrial tag of yeast COXIV. All PCRs were performed with Phusion® High-Fidelity DNA Polymerase (New England Biolabs Inc.). The fragments amplified were inserted in pIVEX2.3-MCS and pIVEX2.4d plasmids (Roche) using restriction sites *NdeI/SmaI* and in pESC-His (Agilent Technologies) plasmids using restriction sites *ApaI/XhoI* (Annex 2) and FastDigest restriction enzymes (ThermoScientific). Ligation was verified by PCR.

2.2. Cloning

The plasmids were cloned in *Escherichia coli* DH5α strains as in Pope & Kent, 1996. Briefly, 0.1-1 ng plasmid DNA was mixed with cells, left for approx. 40 minutes on ice, heat shocked for two minutes at 42°C and two minutes on ice. After the heat shock, 0.5 ml antibiotic-free LB growth medium (1% yeast extract, 1.6% BactoTryptone, 1% NaCl, pH 7.5) was added and the cells were left for up to an hour at 37°C to allow bacterial recovery. Cells were spread on LB-agar plates supplemented with ampicillin or carbenicillin 0.1% and grown at 37°C. Following the persistent difficulties in cloning,

LB medium was supplemented with 1% glucose to suppress gene transcription. Transformation was verified with PCR.

Positive colonies were transferred in 2-3 ml liquid LB growth medium supplemented with ampicillin or carbenicillin and 1% glucose and left to grow overnight in a rotating incubator at 28°C or 37°C and 180 rpm. Plasmids were purified from the liquid cultures using Monarch® Plasmid Miniprep Kit (New England Biolabs Inc.). Final DNA concentration was quantified on NanoDrop (ThermoScientific). The fragment sequence was verified by sequencing (Eurofins Technologies).

2.3. Cell-free protein expression system (CFPS)

In vitro protein synthesis was performed as in Larrieu et al., 2017 and Simonyan et al., 2017.

Briefly, the *in vitro* synthesis was performed in a 100 µL dialysis chamber, separated from a feeding reservoir (1700 µL) by a dialysis membrane (MW 10,000). The system was set up in an inverted microcentrifuge tube. Both chambers contain 0.1 M HEPES (pH 8.0), 1 mM EDTA, 0.05% NaN3, 2% PEG 8000, 151 mM potassium acetate, 7.1 mM magnesium acetate, 0.1 mg/ml folinic acid, 2 mM DTT, 1 mM NTP mix, 0.5 mM amino acid mix, 1 mM RDEWCM mix, 20 mM PEP, 20 mM acetylphosphate, protease inhibitors cocktail (Complete, Roche). The dialysis chamber was added with components for the synthesis: 35% (w/v) S30 *E. coli* BL21(DE3) lysate, 35% S30 buffer (10 mM TRIS-acetate, pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate, 0.5 mM DTT), 0.04 mg/ml pyruvate kinase (Sigma), 15 µg/mL pIVEX2.3MCS-RPHM21 or pIVEX2.4b-RPHM21 plasmid, 0.5 mg/ml tRNAs mix (Roche), 6 units/ml T7 RNA polymerase, 3 units/ml RNasin® Ribonuclease Inhibitors (Promega). The synthesis was made for at least 20 hours under agitation. A pIVEX2.3-MCS bearing another gene was used as positive control.

The reaction mix was centrifuged at 10,000g. The resulting pellet was washed twice in water, whereas the supernatant was precipitated in trichloroacetic acid (TCA) 0.3M and washed twice in ice-cold acetone. The fractions were resuspended in Laemmli loading buffer with 2% β-mercaptoethanol and examined through Western blot (see below).

2.4. mRNA synthesis control

In order to verify whether problems with *in vitro* protein synthesis were due to a transcriptional issue, an mRNA synthesis control was performed. Three 20 µl mix was prepared containing 200 ng of pIVEX2.3 MCS-RPHM21 digested with EcoRV, 0.5 mM NTP mix, 2mM spermidine, 0.8 µg T7 RNA polymerase, 40mM Tris-Cl pH 7.9, $MgCl_2$ 6 mM, 0.5 µl RNasin® Ribonuclease Inhibitors (Promega), 2 mM DTT. Each mix was left incubating at 28°C for 0, 30 or 60 minutes respectively. The synthesis was verified in agarose electrophoresis gel with RiboRuler High Range RNA Ladder (ThermoScientific) as a marker.

2.5. Western blotting

Extracted protein were precipitated in 0.3M TCA and washed twice in acetone, then solubilized in the Laemmli buffer containing 2% β-mercaptoethanol except where noted, and optionally heated at 70°C for 15 min or at 90°C for 5 min prior to gel loading on a 12.5% acrylamide SDS-PAGE. Western blotting was done according to Laemmli, 1970. Proteins were transferred to PVDF membrane (Amersham International, Buckinghamshire, UK) in a liquid system. Nonspecific protein-binding sites were blocked with 5% dried skimmed milk, 3% bovine serum albumin (BSA), and 0.1% Tween-20 (Sigma) in TBS (50 mM Tris-Cl, pH 7.6; 150 mM NaCl) or PBS (137 mM NaCL, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH_2PO_4) for 30 minutes at 4°C, and subsequently washed with 0.1% Tween TBS or PBS. To recognize RPHM21 protein, a specific antiserum produced in rabbit (anti-RPHM21; Davids Biotechnologie) was used, diluted 1:80,000 with 0.1% Tween TBS, overnight at 4 °C. The used antibodies are as follows: anti-RPHM21 as in Milani, Ghiselli, Maurizii, Nuzhdin, & Passamonti, 2014, peroxidase-coupled anti-hexahistidine diluted 1:10,000, anti-PGK diluted 1:5,000, anti-porin diluted 1:10,000, anti-ATP synthase subunit β diluted 1:10,000, anti-COX2 diluted 1:5,000, anti-carboxypeptidase diluted 1:2,000, anti-dolicholphosphate mannosyltransferase diluted 1:2,000. After abundant

rinsing, membranes were incubated with secondary antibodies conjugated with horseradish peroxidase at the dilution of 1:5,000 for 1 h 30 min at RT, except for the antibodies anti-hexahistidine which were already coupled to the enzyme. The washed membranes were detected with ECL Western Blotting Detection Reagents (Roche) and exposed to Hyperfilm ECL (GE Healthcare).

2.6. RPHM21 production in yeast

pESCHis-RPH21 was introduced in the yeast haploid strain W303-1B (*mat a, ade1, his3, leu2, trp1, ura3*) by means of "One step transformation" as in Chen, Yang, & Kuo, 1992.

Briefly, one solid colony or 1.5 ml liquid culture in stationary phase was mixed with 90 µl LiAc (LiAc 0.2N, PEG 3350 40%) solution, 10 µl DTT 1M, 50 µg salmon sperm carrier DNA, 1 mg plasmid, yeast (1 solid colony or 1,5 ml liquid culture in stationary phase). The mix was incubated for 30 minutes at 45°C and spread on synthetic medium supplemented with glucose (yeast nitrogen extract 1%, KH_2PO_4 0.1%, $(NH_4)_2SO_4$ 0.12%, 0.2% Drop mix, 0.01% auxotrophic markers, glucose 2%, pH 5.5).

Yeast positive to the plasmid was grown on synthetic medium whose carbon source was 2% glucose or 2% lactate. RPHM21 production was induced by addition of 0.2% galactose and was tested through sampling at 0, 2, 4 and 6 hours from induction time.

2.7. Yeast mitochondria isolation

Yeast mitochondria isolation was performed as in (Simonyan et al., 2017).

Mitochondria were isolated from yeast cells expressing RPHM21 or not. Washed and concentrated yeast cells were incubated in the presence of 0.5 M β-mercaptoethanol for 15 min and washed twice with 0.5 M KCl. Cells were then suspended in a 20 mM phosphate buffer (pH 6.8) containing 1.35 M sorbitol and 1 mg/mL zymolyase 20T in order to digest the cellular wall. Digestion was extended for 20–35 min and followed by yeast observation through a microscope. Upon digestion completion, spheroplasts (i.e. yeast cells without the cell wall) were washed twice in a 10 mM maleate-TRIS buffer (pH 6.8) containing 1.1 M sorbitol, 0.4 M mannitol, 1 mM EGTA, 0.1% BSA. Cells were

resuspended in a 10 mM maleate-TRIS buffer containing 0.6M mannitol, 1 mM EGTA, 0.2% BSA, and homogenized with three-second passes in a Waring Blender. Cell debris were removed by a 15 min, 900×g centrifugation and mitochondria-enriched fraction was recovered through a 15 min, 17,000×g centrifugation. Mitochondria were resuspended and gently homogenized in a glass-Teflon potter in the same buffer without BSA (10 mM maleate-tris buffer (pH 6.8) containing 0.6 M mannitol, 1 mM EGTA), and the same cycle of centrifugation was done to recover the mitochondrial pellet.

2.8. Yeast total protein extraction

Yeast total protein extraction was performed according to (Egner, Mahé, Pandjaitan, & Kuchler, 1995). Yeast was centrifuged and the washed pellet was resuspended in a NaOH 1.85M/ β-mercaptoethanol 7.5% solution and incubated on ice for 10 min. One volume 50% trichloroacetic acid was added and the mix was incubated for an additional 10 min. After centrifugation, 50 µl 5% SDS with the addition of 6 µl Tris Base 1M were used to resuspend the pellets. The samples were incubated at 42°C for 15 min and centrifuged for 15 min at 15,000 rpm, then the supernatants were used for subsequent analysis.

2.9. Sucrose gradients

A sucrose gradient was used to better separate subcellular fractions and localize RPHM21-6His as in (Meisinger, Sommer, & Pfanner, 2000). A continuous gradient of 15-65% OptiPrep™ and MES solution (MES 10 mM, sorbitol 0.2M, EDTA 2mM, PMSF 1 mM, Complete protease inhibitor cocktail tablet) was created using a density gradient fractionator. 1 mg proteins were deposited on the surface of the gradient and ultracentrifuged in a Beckman SW41 Ti swinging-bucket rotor at 28,000 rpm for 14 h. Fourteen fractions were recuperated from each gradient and the proteins were precipitated with 0.1 volumes TCA, centrifugated and the pellets were washed with ice-cold acetone and analyzed through Western blot.

2.10. Crosslinking

Purified mitochondria, obtained in “Yeast mitochondria isolation”, were washed twice in crosslinking solution (50 mM triethanolammine, 0.6M mannitol, 2 mM EDTA, Complete protease inhibitor cocktail tablet, pH 8) in a 1:1 volume ratio and resuspended in the same solution. DSP 100 or 200 µl diluted in dimethilsulphoxide was added and the mix was incubated 30 min at 30°C. The cross-linking reaction was stopped by addition of TRIS 10mM pH 7.5, which acted for 15 min at room temperature. Samples were mixed with Laemmli loading buffer without β-mercaptoethanol and examined through Western blot.

2.11. Oxygen consumption assays

To gain insights on the state and integrity of mitochondria, respiration measurements through oxygraphy were performed. Oxygen consumption rates were measured with a Clark electrode in a 1 ml thermostatically regulated chamber at 28°C in the respiration buffer (0.65 M mannitol/0.3 mM EGTA/3 mM Tris-phosphate/10 mM Tris-maleate, pH 6.75). 300 µg mitochondria were used for the assays. NADH 3 mM was used as electron donor, whereas the measurements of state 3 and 4 and decoupled oxygen consumption rates were performed by adding 100 µM ADP and 3 µM carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) respectively.

3. Results

3.1. Cloning in E. coli and cell-free protein synthesis

Cloning the *rphm21* gene in both pIVEX and pESC plasmid has proven a very difficult task. The number of transformed colonies was lower by two orders of magnitude than the norm and almost all of them were found to be false positives through PCR (figs.1 and 2). This was not dependent on the strain, as DH5α is reliably and routinely used for transformation (Kostylev, Otwell, Richardson, & Suzuki, 2015) and the positive control did not show any issue regarding the colony number.

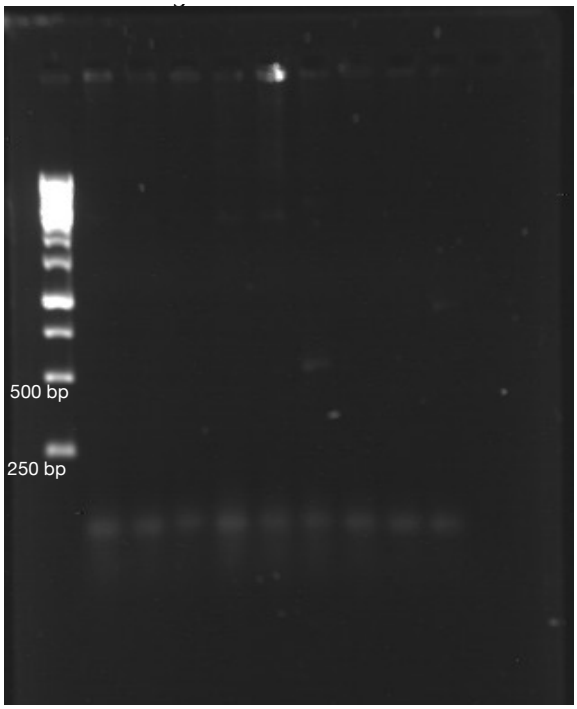


***Fig1**:*
PCR of E.coli colonies selected for the presence of pIVEX-RPHM21 plasmid. None of them bear the gene.

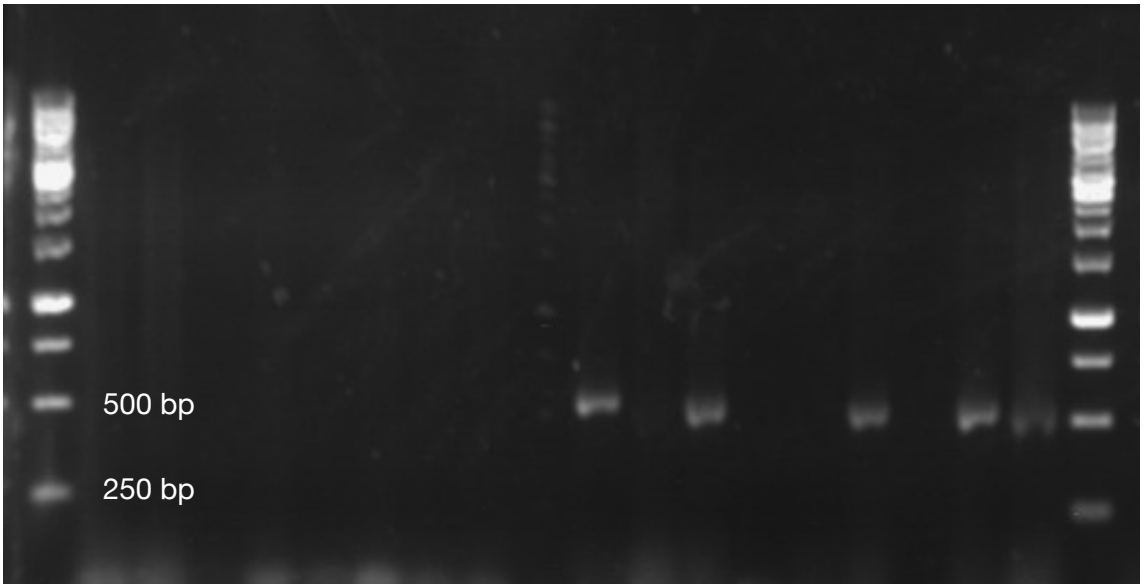


***Fig. 2:** PCR of E.coli colonies selected for the presence of pESC-RPHM21 with and without 6His. The number of true positives is much lower than expected.*

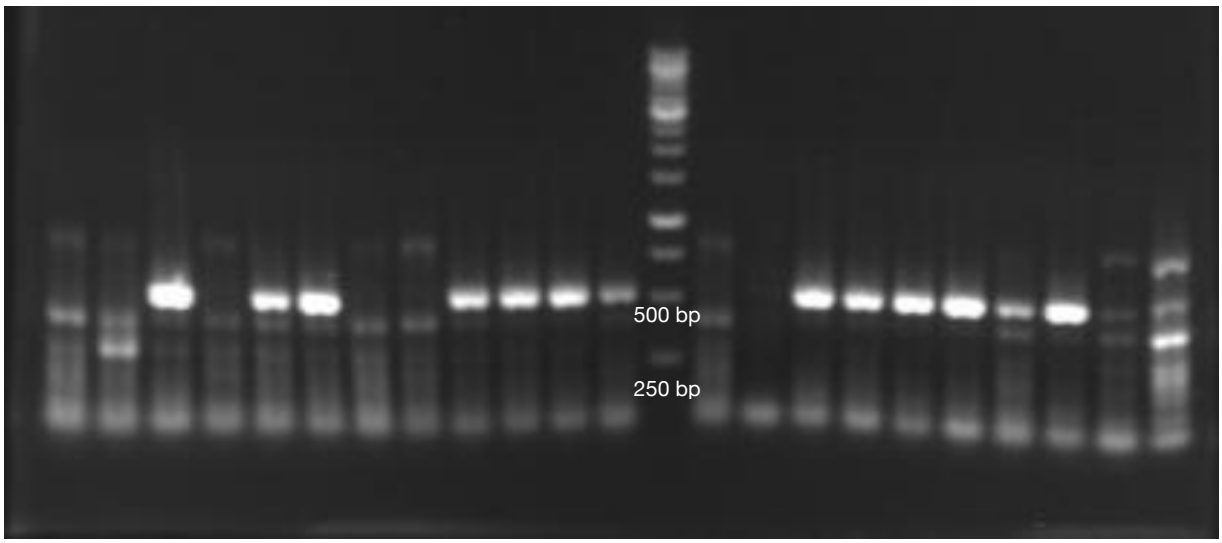


Fig. 3: *PCR of E.coli colonies transformed with pIVEX-RPHM21 and grown on 1% glucose. Each lane is the mix of ten colonies.*

Supplying LB with glucose increased the number of colonies growing on selective medium from less than ten to more than a hundred (fig. 3), making it possible to recover the plasmid with the correct gene sequence. Indeed, after numerous attempts, we were finally able to recover a pIVEX bearing the correct sequence of RPHM21 to use in *in vitro* cell-free protein synthesis (CFPS). Despite CFPS is a robust technique for protein synthesis (Gregorio, Levine, & Oza, 2019), no result was obtained.

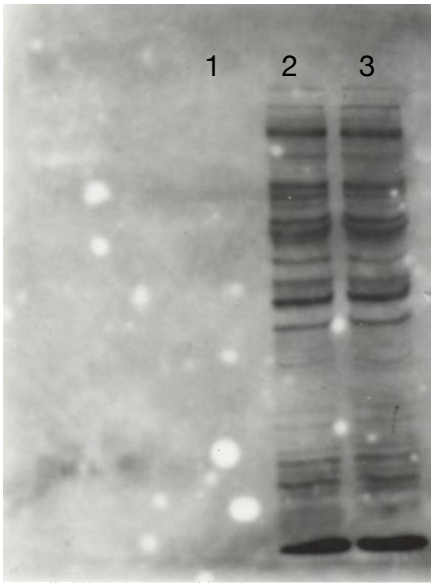


Fig. 4: *Western blot of RHM21 CFPS pellet and supernatant (lanes 1-2) and positive control (lane 3). Antibody anti-RPHM21.*

Fig. 4 shows a Western blot of RPHM21 CFPS pellet and supernatant compared to a control production. RPHM21 was detected using antibodies validated in (Milani, Ghiselli, Maurizii, et al., 2014). Being these polyclonal antibodies, they show a quite intense cross-reactivity against *E.coli* proteins, thus producing a large amount of bands. However, it is clear that no extra band ascribable to RPHM21 is detectable in RPHM21 supernatant with respect to control supernatant.

Re-sequencing of the plasmid used for CFPS and digestion control (fig. 5) demonstrated that the sequence was correct and harbored at the right position within the plasmid. Adding on this, mRNA production control was performed to test whether the issue lied in the transcription step, but the mRNA was produced at a steady rate and had a proper length (fig. 6).

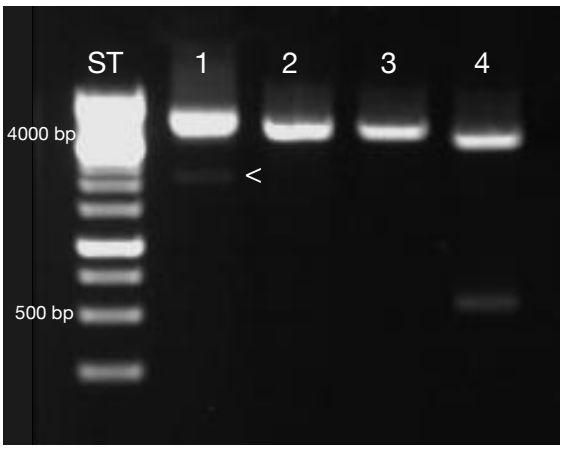


Fig. 5: *Digestion control of pIVEX-RPHM21 used for CFPS. From left to right: molecular standard, undigested, digested with Nde I, with SmaI, with both enzymes. In the first lane the supercoiled form is present as a faint band (arrow head), whereas in the fourth lane the lower band (516 bp) is the gene.*

Fig. 6: *mRNA production control of pIVEX-RPHM21 used for CFPS. From left to right: mRNA production at 0', 30' and 60' from reaction start.*

3.2. RPHM21 production in yeast

Due to cross-reactivity issues with *S. cerevisiae* proteins, it was not possible to use the antibody anti-RPHM21 transformed in yeast. In order to produce RPHM21 in yeast, target it to the mitochondrial matrix and identify it with certainty, a chimeric sequence comprising a mitochondrial presequence, RPHM21 gene and a hexahistidine tag was cloned into a pESC-His plasmid and the latter used to transform yeast.

In pESC, RPHM21 is under the transcriptional control of a galactose-inducible promoter. As clearly visible in fig. 7, pESC induction with galactose triggers RPHM21 production. No cross-reactivity nor leaky transcription was present, indicating a strong and reliable promoter. It is worth noting that the apparent weight of RPHM21 is higher than expected (25 kDa instead than19.5) because of the mitochondrial presequence and the histidine tag.

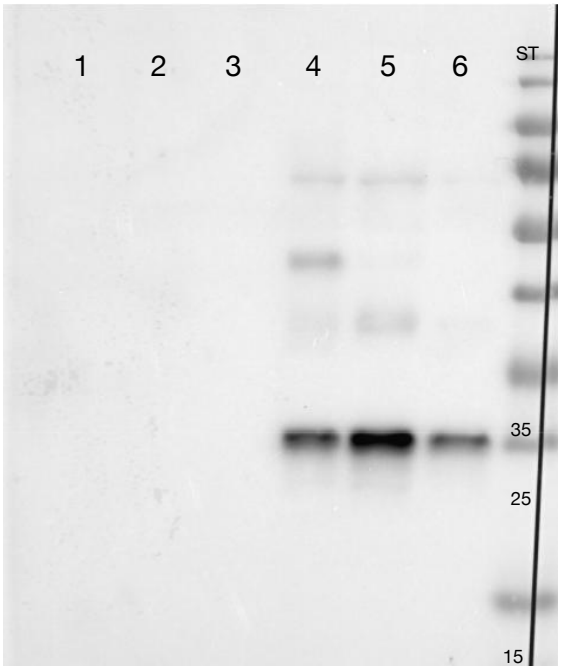


Fig. 7: *Western blot of pESC-RPHM21-His protein extracts grown on Lac (first three lanes) and Lac + Gal (lanes 4, 5 and 6) and revealed with an Ig@6His. The three lanes correspond to different treatments in Laemmli sample buffer before chargement. From left to right: no heating, heated at 70°C for 15 minutes, heated at 90°C for 5 minutes. ST:protein standard*

To verify that RPHM21 had been correctly targeted to mitochondria, a sucrose density gradient was performed and fractions were tested for presence of mitochondrial markers through Western blots. In fig. 8a, c and d, it is possible to see how most of the mitochondrial markers - porin, ATP synthase subunit beta, cytochrome oxydase 2, carboxypeptidase (CPY) and dolichol phosphate mannosyl transferase (DPM1) - localize in fractions 11-13, identifying these as the ones containing mitochondria. RPHM21 (fig 8b) is localized in fractions 11-12, visible as a faint band in lane 11 at about 25 kDa and in lane 12 as a light smear. Being the input quantity spread over several lanes, the signal is quite weak if compared to the previous blot, but is nonetheless present.

Yeast transformed with pESC-RPHM21-6His was grown on synthetic media with lactate as a sole carbon source. This allows to highlight possible mitochondrial problems caused by proteins targeted to this subcellular compartment, as lactate is a non fermentable carbon source for yeast (Turcotte, Liang, Robert, & Soontorngun, 2010). As shown in fig. 9, growth rates were not notably different between carbon sources. The presence of an additional and preferred carbon source boosted yeast growth in both cases, as expected. It is apparent that the massive production of a mitochondrially-targeted protein does not hinder yeast growth nor mitochondrial respiratory functions, conclusion further confirmed by oxygraphic analysis (Annex 3).

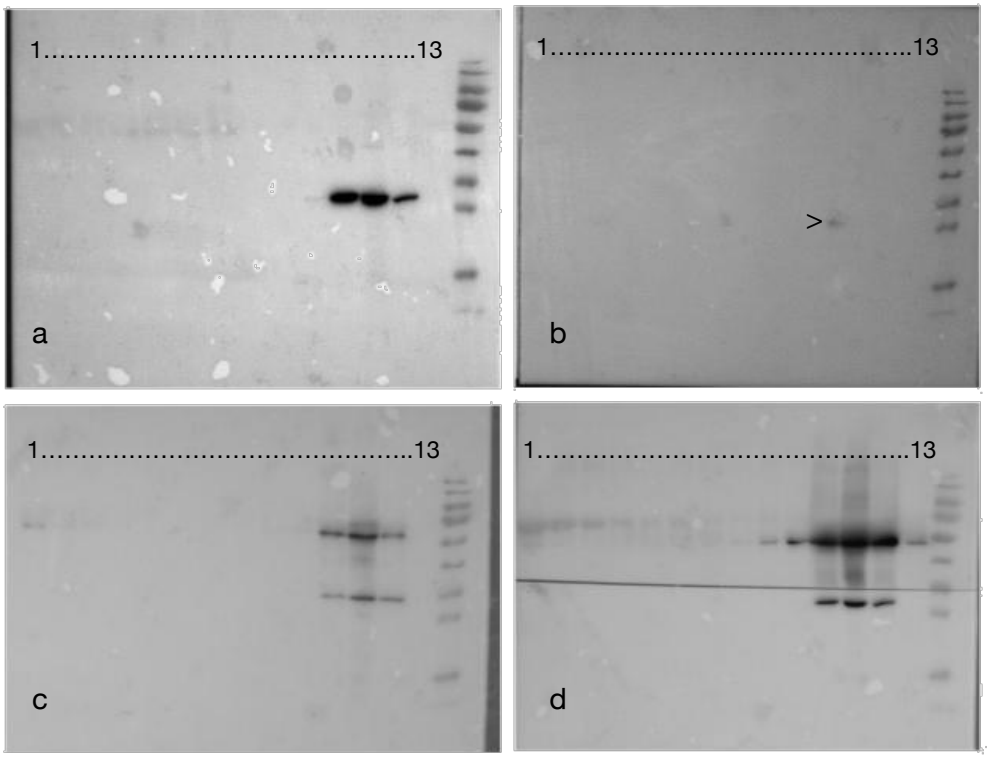


Fig. 8: *Western blot of yeast extract gradients. Each lane has been loaded with a fraction of yeast cellular extract, the topmost of the gradient being at the left and the bottom of the gradient at the right. Protein detected: a) porin and PGK, b) RPHM21 (arrow head), c) ATP synthase subunit beta and COX2, d) CPY and DPM1. Details in text.*

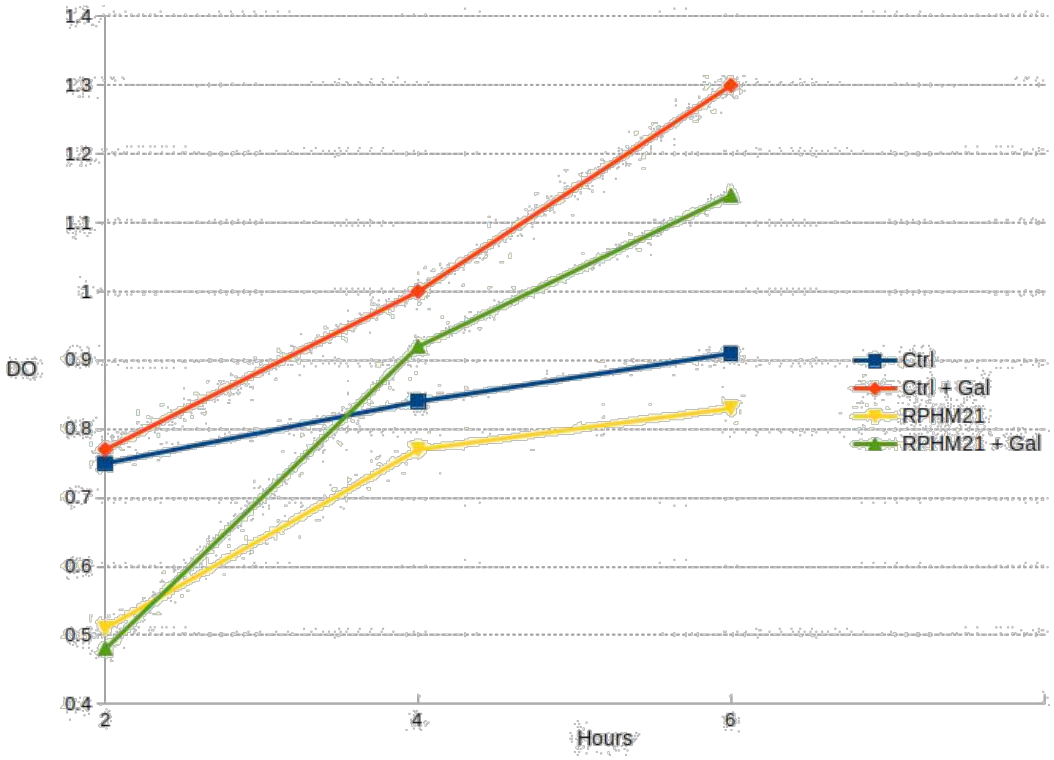


Fig. 9: *Growth monitoring of yeast on synthetic medium and lactose. Legend: ctrl: control (yeast transformed with empty plasmid); ctrl + Gal: control induced with galactose; RPHM21: yeast transformed with pESC-RPHM21-6His; RPHM21 + Gal: yeast transformed with pESC-RPHM21-6His induced with galactose.*

During the previous growth control, Gal-induced PHM21 population was sampled at 0, 2, 4 and 6 hours post-induction and the expression was verified through a Western blot. After 6 hours, protein production has reached a level suitable for the subsequent analysis, and the mitochondrial presequence starts to be cleaved, as shown by the faint band at about 20 kDa.

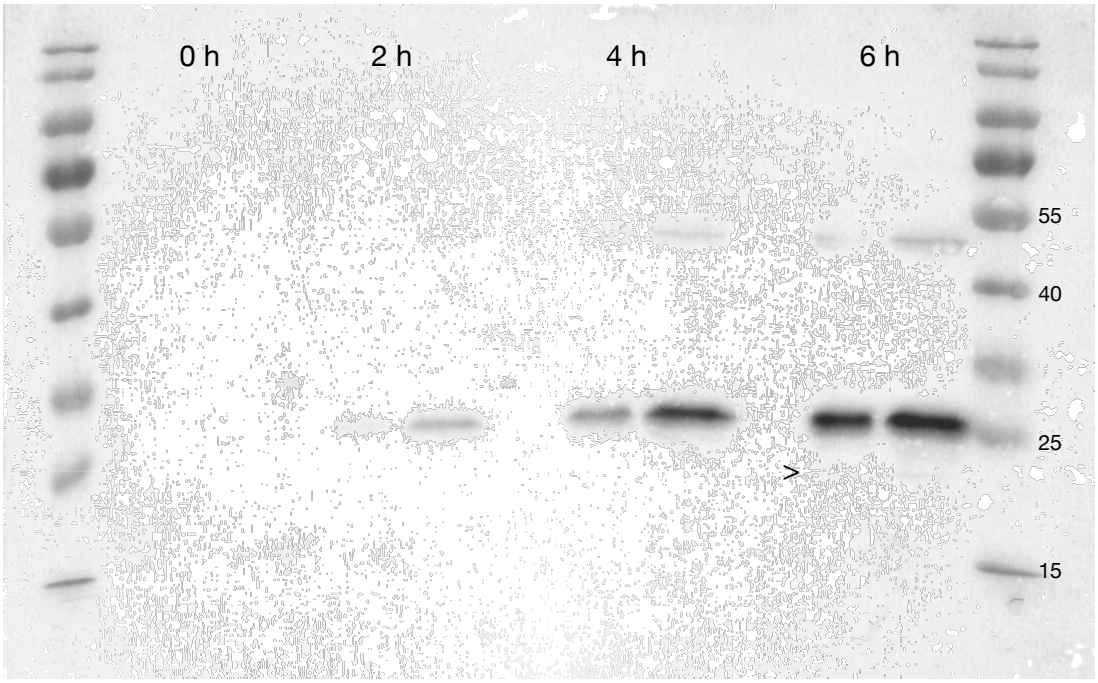


Fig. 10: *Western blot of pESC-RPHM21-6His induction at 0, 2, 4 and 6 hours post-induction. Every time has been loaded twice, the second lane has double the quantity of the first for the same time. At 6h post induction it is possible to see some cleavage of the protein (arrow head).*

Having verified that RPHM21 is indeed produced and properly targeted to the mitochondrion without impeding normal yeast growth, we proceeded to investigate its protein-protein interactions through cross-linking. Dithiobis(succinimidyl propionate) (DSP) is a homobifunctional cross-linking reagent containing a cleavable disulfide spacer. Since it bonds to amines, such as lysine functional groups and protein N termini, and is cell membrane permeable, it is widely used as a generic cross-linking reagent (Sinz, 2018). As it creates disulfide bonds, the Laemmli loading buffer for crosslinked samples was devoid of β-mercaptoethanol. The Western blot in fig. 11 shows the result of the cross-linking: upon use of DSP, the protein hardly enters the stacking gel, which suggest the presence of RPHM21 in the form of inclusion bodies (lanes 2-4). The

situation improves slightly after TCA precipitation, instead getting almost completely stuck at the interface with the stacking gel (lanes 5-7). Interestingly, the controls show the same issue, hinting at the formation of spurious disulfide bridges that were released by the use of β-mercaptoethanol in the previous Western blots.

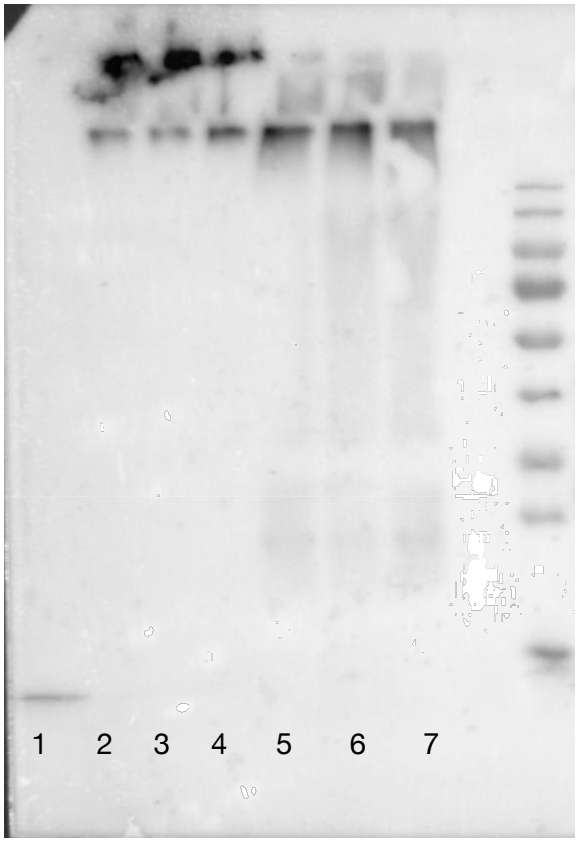


Fig. 11: *Western blot of crosslinking. From 1 to 7: control, RPHM21, RPHM21 + DSP 100 uM, RPHM21 + DSP 200 uM, RPHM21 + TCA, RPHM21 + DSP 100 uM + TCA, RPHM21 + DSP 200 uM + TCA.*

4. Discussion

Heterologous protein expression has become one of the cornerstones of biochemistry, both in fundamental and applied research. Here we report the first attempt to synthesize a mitochondrial ORFan in two distinct heterologous systems in order to uncover its function.

The first major hurdle encountered in this process was the cloning of RPHM21-bearing plasmids in *E. coli*. We were faced with an anomalously low colony number and, among these, a high rate of false positives, suggesting a very high toxicity of RPHM21 in a bacterial context. This was rather unexpected, as RPHM21 in pIVEX is under T7 promoter transcriptional control and the *E. coli* strain used for cloning, DH5α, was specifically chosen because it lacks the T7 RNA polymerase gene, thus theoretically being completely unable to show leaky transcription. The addiction of glucose to both liquid and solid LB medium, which grants a global transcriptional repression by inducer exclusion (Inada, Kimata, & Aiba, 1996) coupled with a decrease in temperature of *E. coli* growth from 37°C to 28°C helped solving this issue (Fig. 3) and retrieving the plasmid for the subsequent cell-free protein expression system (CFPS) step.

The most striking aspect, however, was the complete impossibility to produce RPHM21 in an *E. coli*-based CFPS. Cell-free protein expression has been specifically designed to address the shortcomings of protein production in a living organism, such as cytotoxicity (Rosenblum & Cooperman, 2014). The controls performed to check plasmid structure (fig. 5), mRNA production (fig.6) and CFPS itself (fig. 4) showed that the issue actually lied in the protein produced.

On the other hand, this extreme effect was not observed upon yeast transformation and induction. As a matter of fact, RPHM21 expression tightly followed galactose induction and did not visibly alter metabolic functions in yeast (fig. 9-10). The protein was correctly expressed and targeted in mitochondria (fig. 8). However, issues concerning the correct folding of RPHM21 surfaced when a cross-linking agent was added. Cross-linking helps uncover protein-protein noncovalent interactions by stabilizing them through the dual covalent bonding of the cross-linking agent (Tang & Bruce, 2009). In our case, however, yeast extracts supplemented with DSP showed a RPHM21 aggregate

too voluminous even to enter the stacking portion of the acrylamide gel. TCA precipitation reduces the problem, but the aggregate does not get past the stacking-running gel interface. It is apparent that RPHM21 does not fold properly in the mitochondrion, instead creating an amorphous aggregate, useless for the purposes of protein-protein interaction investigation. The formation of an aggregate might be due to an excessive translation caused by the strong, on-off Gal-inducible promoter, and in the future it could be circumvented using a more tunable promoter. Ironically, as a means to produce aggregate-forming proteins, it has been suggested to use CFSP (Tang & Bruce, 2009).

The high level of toxicity shown by RPHM21 in *E. coli* and in the CFSP is peculiar. If it impairs CFSP, its toxicity must be tied either to transcription or to translation. It could be a nuclease, or a DNA/RNA binding protein, or a protein that prevents translation by interacting with the ribosome. In those cases, the production of a small amount of RPHM21 would impair the proteic production in a negative feedback loop.

It has been suggested that RPHM21 could be a meiotic driver (Milani et al., 2015; Ghiselli et al. 2019). This would mean that, at least in the past, it had the ability to passively tweak their probability to be inherited at the expenses of a competitor (in our case, the female mitochondrion) that does not bear it. If brought to an extreme, the meiotic driver can be an "ultra-selfish" element that promotes its own transmission through the destruction of the competitor (Bravo Núñez, Nuckolls, & Zanders, 2018).

These "killer" drivers have two *modi operandi*, so to say. The first is the so-called "killer-target" drive system: the driver is a *trans* acting element that interacts with all the meiotic products and becomes destructive only when it comes in contact with a second meiotic product, its target. The target can be a protein or a locus, but, most importantly, its localization is restricted to the meiotic product that does not inherit the locus. If a meiotic product does not have a target, the "killer" driver has no effect.

The second type is the "poison-antidote". The killer produces both a poison, that kills indiscriminately, and an antidote, that protects the killer from self-destruction. In this case, the two loci must be very tightly linked in order for the killer not to be killed.

If RPHM21 is an ultra-selfish element, then it could be a "poison-antidote" type, that when placed on a plasmid it was decoupled from its antidote factor and was able to reproduce the effect that had on competitor mitochondria.

Of course, since DUI has been a stable inheritance pattern, this means that either the driver has become fixed in the population, that exclusively the antidote has become fixed instead or that its in an equilibrium with its competitor, possibly through an arms race. Since there are two mtDNA types, the driver has not become fixed. If, on the other hand, the antidote is fixed instead, the killer is perfectly neutralized. But, if this were the case, why would we see *R. philippinarum* M-mtDNA in female gonads, as (Ghiselli et al., 2019) have observed?

The answer could lie in the functioning of the poison-antidote system. If the system, or the competitor, becomes capable of producing the antidote, then the antidote production can work as a fitness signaling system. In other words, when the competitor is too damaged to produce the antidote, then the killer driver is able to destroy it. In this way, the killer's role is exapted by the system – in this case, the gonads – to selectively destroy the unfit competitors – in this case, the female mitochondria that are not fit enough to get into the eggs and be passed to the offspring.

The third option, the arms race, would be verified by populations where one of the two mitochondria has successfully invaded the other's inheritance route, making the population homoplasmic. Of course, this is an instance that could go unnoticed, and until now there has not been such an observation.

In conclusion, producing RPHM21 or one of the other ORFans in an extraneous system could reveal itself very demanding. However, an attempt to substitute the *E. coli* CFSP with a eukaryotic-based one could be the key to its production, and, ultimately, to its function.

Annexes

Annex 1 – orf21 sequence adapted to *E.coli* and *S.cerevisiae* codon usage

---

>RPHM21

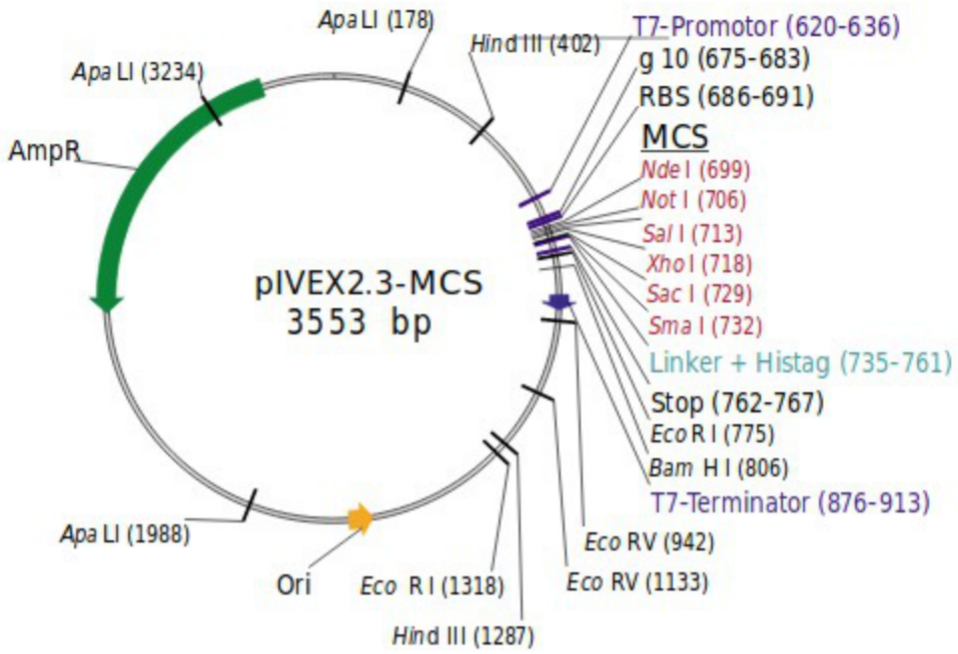
ATGTGGGTCGCCGTCGCCTTCATCCTGTCCTTCATCGCCTCCGACCTGTCCTGCCAGAT
CTCCATCTTCGACGTCTTCTTCTCCTGGGTCGAGTCCCTGGTCACCCTGTTCCTGAAG
GAGTTCTTCTCCGGCAACATCTATGTCGTCTCCTATACCTTCAAGGTCTTCTGGCTGAT
GATCTTCTTCTCCGTCAAGGGCAACCCCTGCGAGTTCACCGAGACCTCCTCCCCCCTG
CCCTCCTCCTCCTCCTCCTCCTCCGTCCCCTCCTCCTCCAAGCCCCCCAAGCAGGTCTA
TTCCGCCCCCATCATCATCTCCGGCTCCAAGGAGGACTTCGACTATCTGATGTCCCTGT
CCAAGGAGAACCTGCTGTTCAAGGTCATCGTCCTGGACTCCAAGGAGAACCAGAACT
TCAAGTCCTATATCGTCTATTTCTGGGAGAAGGTCGACCTGCCCTGCGAGAACCCCTC
CAAGGTCGTCACCGTCCTGATCATCGCC ATGGACGACCCCGACTCCAACTAA

>Mitochondrial presequence used in pESC-RPHM21

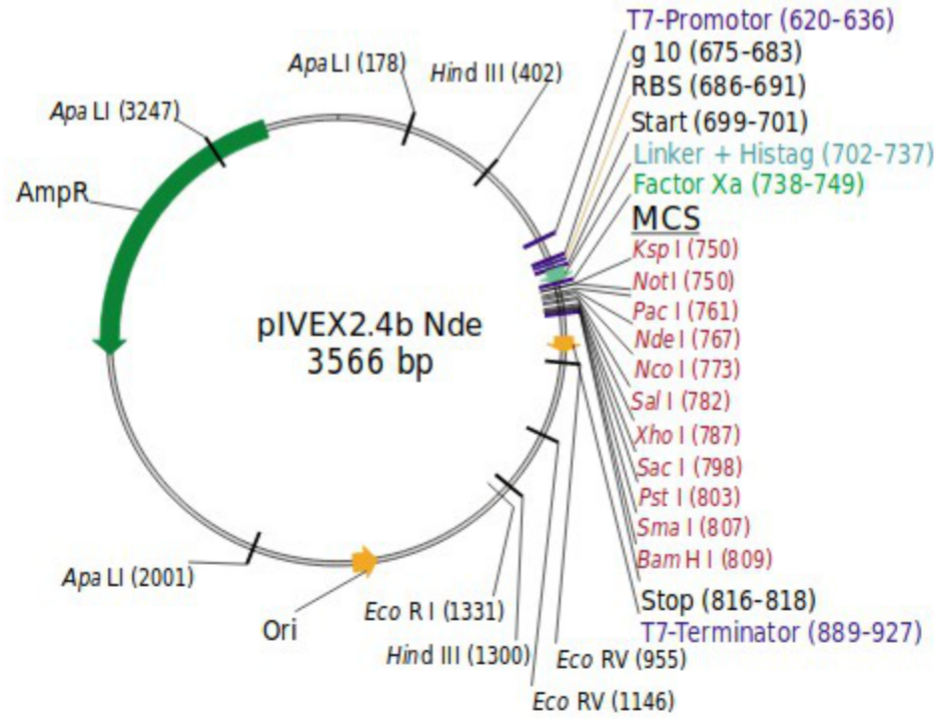
ATGCTTTCACTACGTCAATCTATAAGATTTTTCAAGCCAGCCACAAGAACTTTGTG
TAGCTCT AGATATCTGCTT

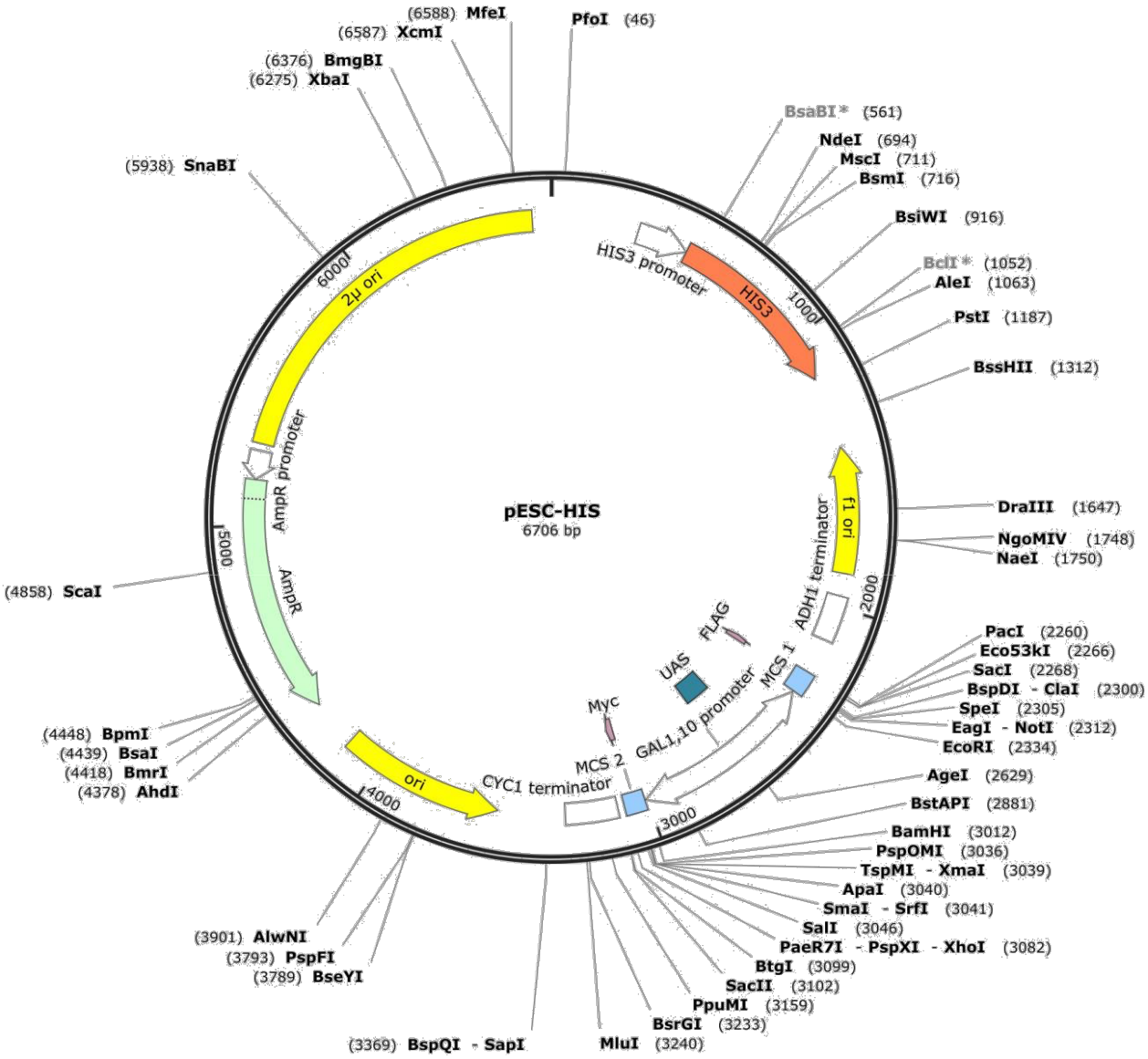
Annex 2 – Plasmids used for cloning and transformation

pIVEX2.3-MCS vector



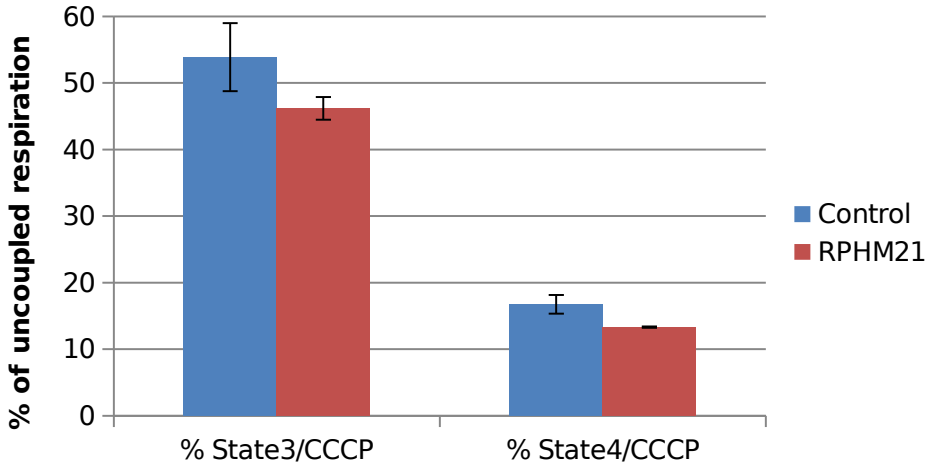
pIVEX2.4b vector




pESC-HIS
6706 bp
HIS3 promoter
HIS3
f1 ori
ADH1 terminator
FLAG
UAS
MCS 1
Myc
GAL1,10 promoter
MCS 2
CYC1 terminator
ori
AmpR
AmpR promoter
2u ori
1000
2000
3000
4000
5000
6000
(6588) MfeI
(6587) XcmI
(6376) BmgBI
(6275) XbaI
(5938) SnaBI
(4858) ScaI
(4448) BpmI
(4439) BsaI
(4418) BmrI
(4378) AhdI
(3901) AlwNI
(3793) PspFI
(3789) BseYI
(3369) BspQI - SapI
MluI (3240)
BsrGI (3233)
PpuMI (3159)
SacII (3102)
BtgI (3099)
PaeR7I - PspXI - XhoI (3082)
SalI (3046)
SmaI - SrfI (3041)
ApaI (3040)
TspMI - XmaI (3039)
PspOMI (3036)
BamHI (3012)
BstAPI (2881)
AgeI (2629)
EcoRI (2334)
EagI - NotI (2312)
SpeI (2305)
BspDI - ClaI (2300)
SacI (2268)
Eco53kI (2266)
PacI (2260)
NaeI (1750)
NgoMIV (1748)
DraIII (1647)
PstI (1187)
BssHII (1312)
AleI (1063)
BclI* (1052)
BsiWI (916)
BsmI (716)
MscI (711)
NdeI (694)
BsaBI* (561)
PfoI (46)

Annex 3

Measurements of uncoupled respiration of W303 yeast expressing or not RPHM21



Chapter three

Preliminary study on miRNA in *R.philippinarum* gonads

1. Introduction

MicroRNAs (miRNAs) are single-stranded, noncoding RNAs measuring about 22 nucleotides in length and regulate gene expression at the post-transcriptional level. They are widespread in animals and plants and, although the exact molecular processes in which they are involved are to be elucidated yet, they are implicated in a miriad of physiological and pathological processes, from cancer (Tutar, 2014) to development (Emde & Hornstein, 2014; Khuu, Nirvani, Utheim, & Sehic, 2016) to spermatogenesis (Chen, Li, Guo, Zhang, & Zeng, 2017). The basic mechanism behind miRNA functioning is the imperfect pairing to the UTR of a mRNA, either condemning the transcript to cleavage or making its translation impossible. In virtue of their imperfect pairing to the so-called seed sequence, their *in silico* prediction is plagued by a high rate of false positives (Riffo-Campos, Riquelme, & Brebi-Mieville, 2016), making it difficult to study in non-model organisms.

Despite a growing body of annotations and research, much is still needed to have even a slightly less nebulous picture of this fundamental biological process. For instance, in miRBase (http://www.mirbase.org/) *Homo sapiens* alone has more reported miRNA that the entire taxon of Lophotrochozoa, which is scantily represented. The varied family of noncoding RNAs still represents a source of scientific novelty whose involvement in biological processes seems to be ever growing (Cech & Steitz, 2014; Hsiao, Sun, & Tsai, 2017).

It is in the Lophotrochozoa clade that a new kind of small noncoding RNA has been recently uncovered. The smithRNA are a new class of noncoding RNAs which, akin to miRNAs, seem to possess the ability to inhibit translation or degrade its target, but, unlike miRNAs, they're transcribed into the mitochondrion and act on nuclear transcripts (Pozzi, Plazzi, Milani, Ghiselli, & Passamonti, 2017, unpublished data). This peculiar mechanism of action is put in place in a species, the venerid *Ruditapes philippinarum*, which has at least another surprising feature. Unlike the vast majority of animals, it doesn't inherit its mitochondria exclusively from the mother, but, in what it's called doubly uniparental inheritance or DUI, its zygotes receive mitochondrial DNA from both parents(Hoeh, Blakley, & Brown, 1991; Skibinski, Gallagher, & Beynon, 1994). In those species who follow DUI (around a hundred of bivalves) the two inherited mtDNAs form

two radically different lineages, that can differ as much as 50% in the two sexes (Breton, Beaupré, Stewart, Hoeh & Blier, 2007). Once entered the egg, the fate of paternal mtDNA, then, is to become undetectable in females and to multiply in males. It is unclear how paternal mitochondria could succeed in avoid degradation, but it is thought that an ORF located on the paternal mtDNA which codes for a protein of unknown function might play a role.

In order to simultaneously broaden the knowledge of miRNA and *R. philippinarum* biology, we conducted a preliminary study of gonads miRNA, extending the repertoire of lophotrochozoan miRNA.

2. Materials and Methods

2.1. Sampling and library preparation

miRNA libraries were prepared by (Pozzi et al., 2017) from six individuals collected during the reproductive season during summer 2016, stored in artificial seawater until sex assessment through microscopic examination and gonad homogenization. RNA was extracted using TRIzol (Thermo Fisher Scientific) and the libraries were prepared by Macrogen Inc using TruSeq Small RNA Library Preparation Kit (Illumina).

mRNA libraries were prepared during the reproductive season in summer 2015 from 15 individuals according to (Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008) and sequenced in 2-lanes, paired-ends 150bp in an Illumina HiSeq 2500.

The occurrence of two different libraries prepared with different animals means that it is not possible to correlate the transcription levels of miRNA and mRNA.

2.2. Assembly, annotation and differential expression analysis

The miRNA libraries were trimmed of adapter sequences with Trimmomatic (Bolger, Lohse, & Usadel, 2014), retaining those reads that had an average Phred score of at least 25 and were long at least 18 nucleotides. Prior to any analysis, two filtering steps were done. The first made use of Kraken (Wood & Salzberg, 2014) with a custom database with prokaryota and human genomes, to filter out contaminants. The second step was to align the reads to a yet unannotated *R. philippinarum* genome with Bowtie2 (Langmead & Salzberg, 2012). The extant information on the genome, unfortunately, didn't let perform a characterization of the genomic environment of miRNAs, but it allowed for a stringent filter devoid of the risks of false positives.

The mRNA libraries were trimmed with Trimmomatic as well. The assembly was performed with Trinity v2.6.6 (Haas et al., 2013) with default options and –min_khmer_cov 2. The ORF prediction and 3'UTR extraction were made with ExUTR

(Huang & Teeling, 2017), with the conditions of retrieving ORFs minimum 20 aminoacids long and 3'UTRs with a minimal length of 20 nucleotides.

2.3. miRNA identification

MiRNA identification was done with miRDeep2 (Friedländer, Mackowiak, Li, Chen, & Rajewsky, 2012; Mackowiak, 2011), an automated pipeline which manages a number of steps. The process starts with the identification and folding of potential pre-miRNA sequences by RNAfold (Lorenz et al., 2011), which determines the presence of a hairpin-loop both from a structural and an energetical point of view with the aid of randfold (Bonnet, Wuyts, Rouzé, & Van de Peer, 2004). Predicted pre-miRNA with a score above 10 and a significant randfold p-value were retained for further analysis.

Quantification was performed mapping the small RNA libraries on the aforementioned genome and counting the mapped reads with the quantifier.pl script from miRDeep2. Differentially expressed miRNAs were normalized with the trimmed mean of M-values method(Robinson, McCarthy, & Smyth, 2010) and established with edgeR (Robinson et al., 2010).

2.4. miRNA target identification and Gene Ontology enrichment

Targets of the miRNA were predicted using MiRanda (Betel, Koppal, Agius, Sander, & Leslie, 2010) and RNAhybrid (Krüger & Rehmsmeier, 2006). In order to identify potential targets, a conservative approach was adopted. Hits reported by miRanda had to have a strict seed binding of Δgduplex $\leq$ -10 kcal/mole (option "-en -10") and an exact seed match and an A in position 1 (option "-strict").

Gene Ontology (GO) annotation was retrieved with PANNZER2 (Törönen, Medlar, & Holm, 2018). Annotation that scored more than 0.4 PPV were retained for further analysis. GO enrichment was performed with topGO(Alexa & Rahnenfuhrer, 2019).

3. Results and discussion

3.1. Identification of miRNAs in *R. philippinarum*

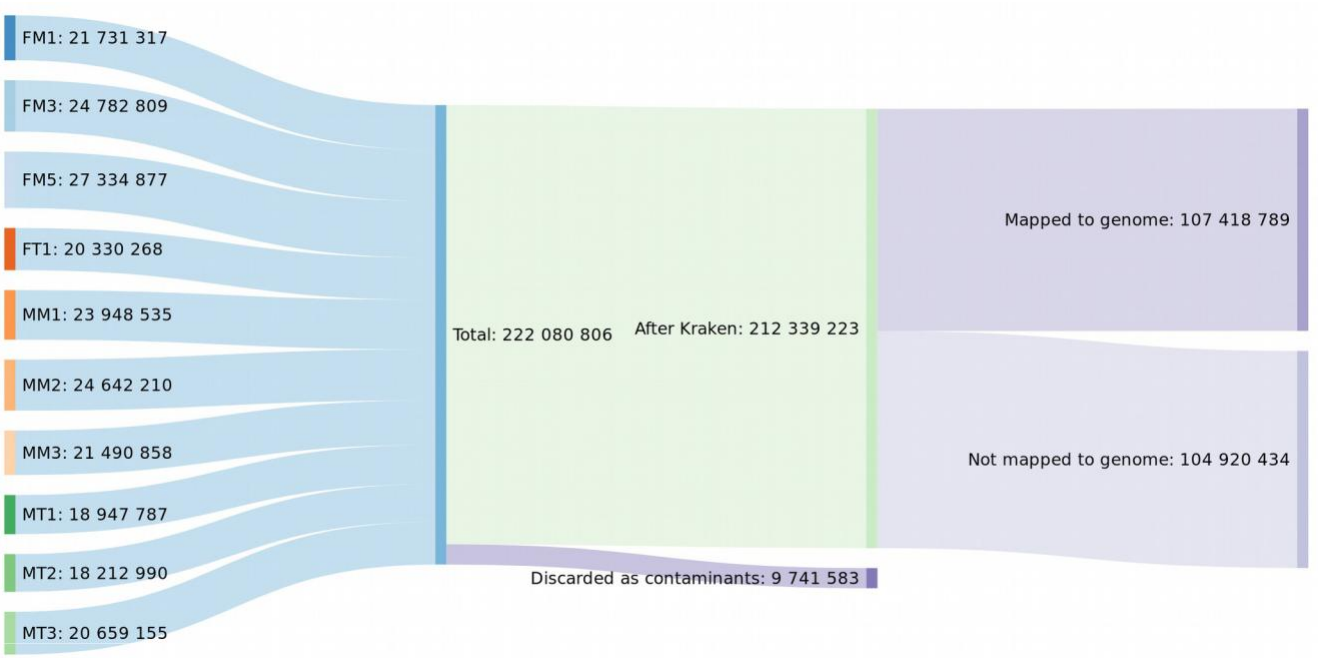


Fig. 1: *Schematics of reads distributions across samples, treatments and filtering steps.*

About 220M reads were gathered from six individuals, three males and three females (Fig. 1). In addition to contaminant filtering, a genome mapping was felt necessary, despite the genome not being completely assembled yet, as an added layer of caution; about 107M reads were deemed fit for the subsequent steps. The pipeline mirDeep2 (Friedländer et al., 2012) was used to identify 279 miRNA candidates, further narrowed down to 171 according to multiple stringency filters (Annex 1). These filters are a mirDeep2 score > 10, which reflects the probability that the sequence is a genuine miRNA, and a significant randfold p-value, which expresses the propensity of the sequence to assume the three-dimensional hairpin conformation typical of miRNA precursors. All of these showed a match to at least a known miRNA in miRBase, a number of which to miRNA of the bivalve *Crassostrea virginica* (Xu et al., 2014). The number of putative miRNAs identified through this study is well aligned to the miRNA range discovered in other invertebrates, i.e. 60-238 miRNA (Berezikov, 2011).

The length of the identified mature miRNA spans from 18 to 25 nucleotides, albeit the majority (154/171) is from 21 to 23 nucleotides.

3.2. Identification of sex-associated miRNA

Differential expression analysis revealed 22 miRNAs more expressed in females and just 9 more expressed in males ($p < 0.05$, Annexes 2 and 3) . What's more, female-biased miRNAs are both more transcribed relatively to their male counterpart (thus reaching 3.11 log_2 fold-change in the case of tig00018338_22836 and tig00011614_15613 versus a maximum of -1.3 for the male-biased tig00030723_29204 and tig00005293_7460) and in terms of counts per million reads.

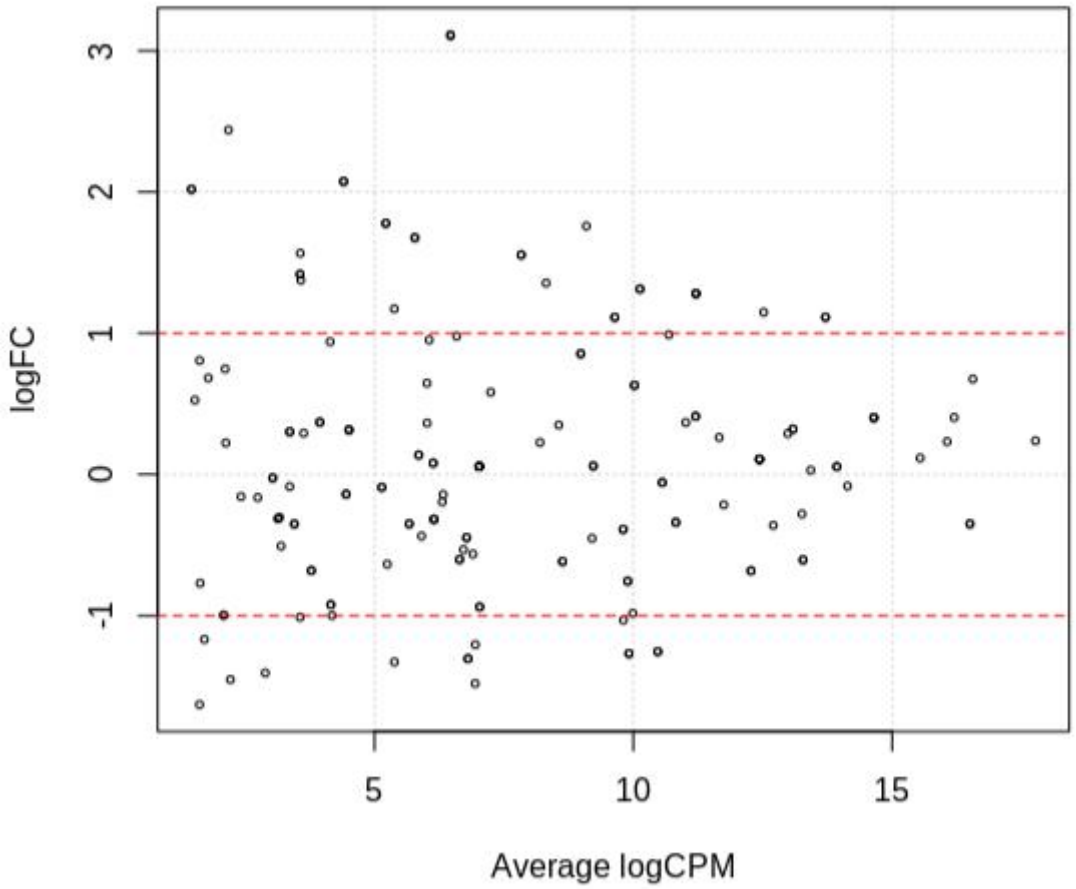


Fig. 2: *Transcription levels of microRNA in Ruditapes philippinarum. Fold change values greater than 1 correspond to female-biased transcription, whereas fold changes smaller than -1 correspond to male-biased transcription. Fold change is plotted against the expression level, estimated by the counts per million (CPM) reads mapped and normalized via the trimmed mean of M-values (TMM) method.*

3.3. Identification of miRNA targets

Since there is no reference genome available for *R. philippinarum*, in order to recover the 3'UTR from its transcriptome we made use of the exUTR pipeline (Huang & Teeling, 2017), whose first step is an annotation through BLAST rather than an alignment. In the *R. philippinarum* gonads transcriptome, 11946 3'UTR longer than 20 nucleotides were identified.

The potential miRNA targets were identified through a dual approach. miRanda retrieved 6128 targets, at least one for each miRNA. However, since *in silico* prediction of targets is known for retrieving a high number of false positives (Liu, Li, & Cairns, 2014; Riffo-Campos et al., 2016)despite setting miRanda as stringently as possible, RNAhybrid was choosen as an additional method to retrieve targets and alleviate this issue. RNAhybrid retrieved 9063 possible targets, and the targets actually identified by both methods were 1518. These were the targets that were kept.

After this process, four miRNA were found not to interact with any target and were excluded from further research. In total, 3411 miRNA-mRNA interactions were noted. On average, miRNA target 20.4 3'UTRs, and the mRNAs are targeted on average by 2.24 miRNA each.

3.4. Gene Ontology annotation and enrichment

ID–BP	**Annotation**	**p**
GO:0060174	limb bud formation	2.1e-06
GO:0034087	establishment of mitotic sister chromatin	7.4e-06
GO:0032332	positive regulation of chondrocyte differentiation	1.5e-05
GO:0061036	positive regulation of cartilage development	1.9e-05
GO:0034085	establishment of sister chromatid cohesion	4.1e-05
GO:0098856	intestinal lipid absorption	6.0e-05
GO:0070966	nuclear-transcribed mRNA catabolic process	8.3e-05
GO:1904480	positive regulation of intestinal absorption	8.3e-05
GO:0032330	regulation of chondrocyte differentiation	8.6e-05
ID–MF	**Annotation**	**p**
GO:0001227	transcriptional repressor activity, RNA polymerase II- specific	3.5e-05
GO:0004686	elongation factor-2 kinase activity	8.2e-05

Table 1 – Significant enriched GO terms among the targets of the female-biased miRNA. BP: Biological Process. MF: Molecular Function.

ID-BP	Annotation	p
GO:0006956	complement activation	3.9e-06
GO:0072376	protein activation cascade	2.9e-05
ID-CC	**Annotation**	**p**
GO:0005579	membrane attack complex	1.9e-06
GO:0046930	pore complex	7.0e-06
GO:0070822	Sin3-type complex	5.7e-05
GO:0034245	mitochondrial DNA-directed RNA polymerase complex	0.00021

Table 2 - Significant enriched GO terms among the targets of the male-biased miRNA. BP: Biological Process.CC: Cellular component

A first GO enrichment analysis, concerning miRNAs as a whole, didn't retrieve any significantly enriched term. When split according to sex bias, however, some terms appeared to be overrepresented.

Regarding the annotation of the female-biased miRNAs targets, there's a clear pattern of GO terms associated with growth and differentiation ("limb bud formation", "positive regulation of chondrocyte differentiation", "positive regulation of cartilage development") and activities linked to cell division ("establishment of mitotic sister chromatin", "establishment of sister chromatid cohesion", "elongation factor-2 kinase activity") which could be linked to gonad and zygote formation.

In males, there isn't an overall pattern that emerges at first, but other, less significant terms (not shown) share the theme of immune response represented here by "GO:0006956 -

complement activation“. Another GO term, “Sin3-type complex”, which denotes a histone deacetylase complex, evokes the epigenetic transcriptional regulation and has actually been linked with the modulation of sperm motility (Parab et al., 2015). It is of particular interest, given our model species’ peculiar mtDNA transmission route, the regulation of a mitochondrial RNA polymerase from a miRNA, as suggested by the occurrence of GO: 0034245. The sudden appearance of a selfish element imposes a narrow evolutionary window to find a countermeasure; miRNAs, being small molecules evolving in intergenic regions, could be plastic enough to block the spread of selfish elements.

3.5. MiRNAs with opposite bias share mRNA targets

Transcript	BLAST hit	e-value
TRINITY_DN56637_c3_g2_i1	Zinc finger and SCAN domain-containing protein 2	4e-37
TRINITY_DN56637_c3_g2_i2	Zinc finger and SCAN domain-containing protein 2	1e-36
TRINITY_DN59888_c2_g1_i6	Mitochondrial import inner membrane translocase subunit TIM14	8e-40
TRINITY_DN63196_c0_g1_i2	Lysine histidine transporter-like 3	4e-12
TRINITY_DN64283_c1_g2_i4	BTB/POZ domain-containing protein KCTD7	2e-26

Next, we assessed whether there were UTRs targeted by both male- and female-biased miRNA. We found five targets which interact with one male-biased miRNA, tig00009357_12789, and with eight female-biased miRNAs (Annex 4). Some of them have a clear significance in a developmental context.

TRINITY_DN56637_c3_g2_i1 and TRINITY_DN56637_c3_g2_i1 show similarity to zinc-finger and SCAN domain-containing proteins (zSCANs), DNA-binding proteins which are

expressed in human and murin embryonic stem cells (ESC), where they inhibit cellular differentiation and are essential to maintain ESC in a pluripotent state (Wang et al., 2007). zSCANs are part of a wider transcriptional network involved in ESC differentiation regulation (Yu, Kunarso, Hong, & Stanton, 2009). They are both targeted by 21 miRNAs each, well over the 2.24 miRNA average target, suggesting that their transcriptional activation plays a crucial role in gametogenesis and/or zygote development. The two transcripts differ by just two amino acids, so this could be a computational artifact and they could be the same sequence.

The other three targets are proteins important for fundamental processes such as mitochondrial import, amino acid transport and control of excitability in neurons. Further analyses are needed to understand if this regulatory pattern is typical of gametogenesis or it is a constitutive regulation.

Annexes

Annex 1 – sequences of all the retrieved miRNA

>tig00007430_10460
TACCCTGTAGATCCGAATTTGT

>tig00040972_42614
AACCCGTAGATCCGAACTTGT

>tig00036442_37460
AACCCGTAGATCCGAACTTGT

>tig00003725_5387
TGCCATTTTTATCAGTCACTGT

>tig00022992_26204
TCCCTGAGACCATAACTTGAGGAC

>tig00043088_45018
TCACAACCTGCTTGAATGAGGAC

>tig00032658_32203
TTACCCTGTTGAACCGAGCGTGT

>tig00033893_33950
AATTGCACTTGTCCCGGCCTGC

>tig00033893_33954
AATTGCACTTGTCCCGGCCTGC

>tig00033893_33946
AATTGCACTTGTCCCGGCCTGC

>tig00022279_25624
TGGACGGAGAACTGATAAGGGT
>tig00040972_42616
TGAGGTAGTAGGTTGTATAGT

>tig00036442_37462
TGAGGTAGTAGGTTGTATAGT

>tig00000144_96
TGAGATCATTTTGAAAACTGAT

>tig00004039_5761
TAATCTCAGCTGGTAATTCTGA

>tig00039553_41208
TAATCTCAGCTGGTAATTCTGA

>tig00000144_94
TGAGATCATTGCGAAAACTGATT

>tig00004934_7142
TACTGGCCTGTAAAATCCCAAA

>tig00039619_41246
TACTGGCCTGTAAAATCCCAAA

>tig00035924_36898
TATCACAGCTGGCTTGAGTGAGC

>tig00002865_4003
TATCACAGCTGGCTTGAGTGAGC

>tig00046835_48003
CGGGACTACGTCAACTACTTGC

>tig00009146_12568
TATCACAGCCAGCTTTGATGAGC

>tig00042746_44511
TATCACAGCCAGCTTTGATGAGC

>tig00002606_3764
TTTTGATTGTTGCTCAGAAAGCCG

>tig00015212_19450
CTTGGCACTGGCGGAATAATCAC

>tig00032525_32084
CTTGGCACTGGCGGAATAATCAC

>tig00042746_44510
TATCACAGCCAGCTTTGATGAGC

>tig00009146_12569
TATCACAGCCAGCTTTGATGAGC

>tig00009146_12565
TATCACAGCCAGCTTTGATGAGC

>tig00042746_44514
TATCACAGCCAGCTTTGATGAGC

>tig00039553_41210
TGAGTATTACATCAGGTACTGA

>tig00004039_5759
TGAGTATTACATCAGGTACTGA

>tig00039553_41206
TAATATCAGCTGGTAATCCTGAG

>tig00005276_7436
TTCGTTGTCGTCGAAACCTGCCT

>tig00035924_36900
TATCACAGCCTGCTTTGATGAGC

>tig00002865_4001
TATCACAGCCTGCTTTGATGAGC

>tig00001505_2321
TTGGTCCCCTTCAACCAGCTGT

>tig00038994_40525
TGAACACAGCTGGTGGTATCTTTT

>tig00007285_10168
TGAACACAGCTGGTGGTATCTTTT

>tig00006731_9639
TGAACACAGCTGGTGGTATCTTTT

>tig00002047_3150
TGACTAGATCCACACTCATCCA

>tig00009146_12572
TATCACAGCCTGCTTGGATCAGT

>tig00042746_44507
TATCACAGCCTGCTTGGATCAGT

>tig00044404_46109
ATTTGGCACTTGTGGAATAATC

>tig00044815_46468
ATTTGGCACTTGTGGAATAATC

>tig00002700_3830
CAGCTCATCATCGGGTAGCCT

>tig00009357_12787
GATCACAGCCTGCTTTGATGAGC

>tig00009146_12573
TGAAAGACATGGGTAGTGAGATG

>tig00042746_44506
TGAAAGACATGGGTAGTGAGATG

>tig00007285_10164
AACCACTTTCTGCACTCCAGAA

>tig00015187_19438
GAGCTGCCAAATGAAGGGCTGT

>tig00002700_3828
GAGCTGCCAAATGAAGGGCTGT

>tig00003027_4179
TGGAATGTAAAGAAGTATGAGT

>tig00001505_2319
TGGAATGTAAAGAAGTATGAGT

>tig00009357_12783
AATCACAGTCTGCTTTGGTGAGC

>tig00009357_12789
GATCACAGTCTGCTTTGGTGAGT

>tig00044427_46114
CTGCCGTAATTGGACTGGCATA

>tig00009146_12563
TATCACAGCCAGCTTTGATGACA

>tig00042746_44516
TATCACAGCCAGCTTTGATGACA

>tig00046325_47611
GTGAGCAAAGTTTCAGGTGTCA

>tig00009391_12838
GTGAGCAAAGTTTCAGGTGTCA

>tig00038996_40533
AACCACTTTCTGCACTCCAGAA

>tig00022816_26041
TGAATTATGGATTTGGAACATA

>tig00012899_16967
TGAATTATGGATTTGGAACATA

>tig00046413_47752
TGAATTATGGATTTGGAACATA

>tig00026075_28172
TGAGATTCAACTCCTCCAACTG

>tig00019400_23642
TGAGATTCAACTCCTCCAACTG

>tig00033893_33966
TATTGCACTTGCCCCGGCCTTT

>tig00024536_27281
TATTGCACTTGCCCCGGCCTTT

>tig00009391_12834
GTGAGCAAAGTTTCAAGTGTGT

>tig00046325_47615
GTGAGCAAAGTTTCAAGTGTGT

>tig00009357_12785
GATCACAGCCTGCTTTGGTGAGC

>tig00007389_10319
TAGGACATAACTCTGTCAAGGT

>tig00039617_41244
TAATGCCCCGTGAAATCCTAA

>tig00002933_4064
CGTGATAGGTCTTGCATTGCTG

>tig00044182_45882
CCAGATCTAACTCTTCCAGCTCA

>tig00009391_12836
GTGAGCAAAGTTTCAGGTGTCGG

>tig00046325_47613
GTGAGCAAAGTTTCAGGTGTCGG

>tig00037685_38963
TGAGTTATGGATTTGGAACATA

>tig00041823_43653
TGAAATATGGATTTGGAACATG

>tig00007538_10573
TGACTTATGTATTTGGAACATA

>tig00019259_23484
TGAATTATGGATTTGAAACATA

>tig00017930_22399
TGAATAATGGATTTGGAACATG

>tig00016616_21089
CGAATTATGGATTTGGAACATG

>tig00017051_21591
TGAATTATAGATTTGGAACATG

>tig00033375_33253
GTTGTGACCGTTGTAATGGGT

>tig00020741_24610
GTTGTGACCGTTGTAATGGGT

>tig00017408_21972
TCGCTGTTGACACAGGTAGAGT

>tig00012922_16977
TCGCTGTTGACACAGGTAGAGT

>tig00038222_39680
ACGTAGATTTAAGTTGATGTCC

>tig00042141_43937
ACGTAGATTTAAGTTGATGTCC

>tig00012741_16805
ACGTAGATTTAAGTTGATGTCC

>tig00037331_38467
ACGTAGATTTAAGTTGATGTCC

>tig00006607_9511
ACGTAGATTTAAGTTGATGTCC

>tig00040588_42225
ACGTAGATTTAAGTTGATGTCC

>tig00002031_3076
ACGTAGATTTAAGTTGATGTCC

>tig00032840_32545
TCGATAACGACGATCCGAGCAC

>tig00016904_21483
TCGATAACGACGATCCGAGCAC

>tig00045767_47178
TGGAAGACTTGTGATTTTGTTGT

>tig00009357_12781
GATCACAGCCTGCTTTGATATTC

>tig00046090_47523
TGTCATGGAGTTGCTCTCTTTA

>tig00005691_8087
TGAAATATGGATTTGGAACATA

>tig00005293_7460
TGAATTATGGATTTGGGACATG

>tig00030723_29204
TGAATTATGGATTTGGGACATG

>tig00010653_14386
TTGCATAGTCACAAAAGTGATC

>tig00036478_37525
TTGCATAGTCACAAAAGTGATC

>tig00004684_6715
AGAACTGTGTATGGACATCAGT

>tig00002933_4066
TGTGATAGGCATGCATTGTTG

>tig00007285_10166
TAAATGCATAATCTGGTATGTG

>tig00038994_40527
TAAATGCATAATCTGGTATGTG

>tig00038996_40531
TAAATGCATAATCTGGTATGTG

>tig00046835_48001
TCGGGACATTGTCAATTCCATG

>tig00000078_64
TGAATAATGGATTTGGAACATA

>tig00005719_8141
TGTTTCATTTACATATTTCATT

>tig00005185_7354
TCTTTGGTTATCTAGCTGTATGA

>tig00035538_36428
TAGCACCATTTGAAATCAGATT

>tig00002147_3262
TAGCACCATTTGAAATCAGATT

>tig00018436_22944
TGTTTCATTTACATATTTCATT

>tig00003786_5451
GGTCCTTCTCTGGGTCTGAGGACT

>tig00009940_13595
CAGAATGATGAAATAAGAGATC

>tig00036728_37699
TGTTTCATTTACATATTTCATT

>tig00002949_4087
TGTTTCATTTACATATTTCATT

>tig00022970_26156
TTTGGCACTGTTGTGGTGCTGT

>tig00004456_6163
TTCCCGGCCGATGCACCA

>tig00004238_6005
TGTTTCATTTACATATTTCATT

>tig00024386_27245
TGTTTCATTTACATATTTCATT

>tig00034791_35451
TGTTTCATTTACATATTTCATT

>tig00006732_9642
TGTTTCATTTACATATTTCATT

>tig00013067_17200
TATTTTACGAGTGATGGCTGTC

>tig00038904_40470
TATTTTACGAGTGATGGCTGTC

>tig00011219_14944
GTGCATTGTAGTTGCATTGCA

>tig00011219_14946
GTGCATTGTAGTTGCATTGCA

>tig00014638_18919
TCAGCTGTCATGATGCTTTCT

>tig00004570_6350
TGTTTCATTGGCATGGGGACTGC

>tig00041297_43066
TATTCTTCGACTAAAAGCTGCC

>tig00041297_43068
TATTCTTCGACTAAAAGCTGCC

>tig00018338_22836
TTCATACGTATGAACGAAAACCT

>tig00011614_15613
TTCATACGTATGAACGAAAACCT

>tig00008213_11520
TGAGACAGTGTGTCCTCCCTCA

>tig00033932_34112
TGAGACAGTGTGTCCTCCCTCA

>tig00003382_4592
TATTCTTCGACTAAAAGCTGCC

>tig00024577_27309
CTTACCCTGTAAATCGGAGAAGT

>tig00032658_32205
CTTACCCTGTAAATCGGAGAAGT

>tig00033103_32907
TAGCACTTGCCTGTGCTTGGGA

>tig00004867_7032
TAGCACTTGCCTGTGCTTGGGA

>tig00026997_28551
TAGCACTAGCTTGTGCTTGGGA

>tig00035235_36062
TAGCACTAGCTTGTGCTTGGGA

>tig00026807_28465
ATCGGACAGTAAAACCCTTGCT

>tig00034060_34406
GGATTTTCAACGACGCCGT

>tig00042149_43965
GGATTTTCAACGACGCCGT

>tig00042150_43978
GGATTTTCAACGACGCCGT

>tig00037544_38816
GGATTTTCAACGACGCCGT

>tig00035538_36430
TAGCACCATTTGAAATCAGTGC

>tig00018987_23304
TGGCGCCGTGGAAACATCCTTC

>tig00031437_30247
TGGCGCCGTGGAAACATCCTTC

>tig00006212_9025
TAAGGCACGCGGTGAATGCCA

>tig00006212_9023
TAAGGCACGCGGTGAATGCCA

>tig00006859_9764
CCTAGAACTTACTTGTGCTAAT

>tig00008213_11522
TGAGACAGTGTGTCCTCCCATG

>tig00033932_34114
TGAGACAGTGTGTCCTCCCATG

>tig00008771_12147
TCAGCAGTTGTACCACTGATTTG

>tig00033862_33922
GAACTTTGCCCGGGACATAACTCTA

>tig00038975_40494
ATTGAACGAACGTTTCCGAGAG

>tig00024163_27129
ATTGAACGAACGTTTCCGAGAG

>tig00038053_39505
TAGCACTTTTTGATGCTGGGAT

>tig00046835_48005
TGGATTTCCCAAGATCCGTGAT

>tig00003578_5059
CGTAGATTTAAGTTGATGTCCT

>tig00041876_43709
ACACCACATATGATTTACTGTCT

>tig00011321_15094
TGATTGTGGATTGTTTGGTGTGCCT

>tig00006524_9445
TGATTGTGGATTGTTTGGTGTGCCT

>tig00026082_28173
TGATTGTGGATTGTTTGGTGTGCCT

>tig00005158_7344
ATGCGTAGGCGTTGTGCACAGA

>tig00006132_8875
ATGCGTAGGCGTTGTGCACAGA

>tig00006214_9061
ATCCAAACGTTGAGTACGGCACT

>tig00003841_5539
ATCCAAACGTTGAGTACGGCACT

>tig00014190_18489
TAGCACATGTCTGTGCTGGGAA

>tig00008991_12444
ATTGTAGAAAACGACACAATAT

>tig00036498_37556
ACACCGCATATGATTTACTGTCT

>tig00001115_1616
ATTGTAGAAAACGACACAATAT

>tig00039419_41039
ATCCACTCCTTTTGGCTGTCTATG

>tig00013310_17595
ATCCACTCCTTTTGGCTGTCTATG

>tig00034356_34913
AGAAGGATGATTGCAGAACATTTGC

>tig00004144_5895
TTTTGTTGAAATAAGGACACCATAA

>tig00005579_7958
CGGAAGAAACCTTCGTCTAGAG

>tig00003750_5407
CGGAAGAAACCTTCGTCTAGAG

>tig00001171_1725
GAAACGCTTAGACTTTGATCTGGAG

>tig00002047_3148
TCACCGGGTAAACATTCATCCGC

>tig00002011_3051
TTGGTCACACCAGCTAGGTGAC

>tig00033674_33723
TGCCGTAGAACCGACTTCCCGG

>tig00038752_40284
GAGATCCGTTAAGTCTGTTGGATA

>tig00003541_5008
TGCCGTAGAACCGACTTCCCGG

>tig00038459_39939
TGCGCTGTCGTCCGGCGGTCAT

>tig00037516_38772
TATACATAAAGACTTGGGATTGG

>tig00020699_24581
TGCACCAAAATGATGCCAGGTGA

>tig00015836_20047
TGCGAGCGAAACTAGGGGCGAG

>tig00026112_28182
CATCACTATAGGCCTAGGACCCCT

>tig00034312_34780
ACGTGTTGTCGTTGTTTTT

>tig00010196_13882
ACGTGTTGTCGTTGTTTTT

>tig00033263_33094
GACAAGGGGCATCACTCTGAAAGCT

>tig00008612_11961
TCACATCATATACTTCAAGGATTGA

>tig00020643_24536
AAGTCCGCAGTAGTACAGTCGGTT

>tig00031748_30849
AATAATGGCATGTTTTTGTGATGTG

>tig00031745_30846
GAGGACTTTATCGGATGTTTGAAGA

>tig00021406_25064
TTCCAAAAGGTCTAGGACCCCT

>tig00033375_33251
CTAAGTAGTGATGCCGCGGGT

>tig00040026_41554
AAATAGTTGATTGTATACCATGCCT

>tig00037784_39183
GAGGACTTTATCGGATGTTTGAAGA

>tig00041725_43472
TTAATTGAGAGACTTTGTACT

>tig00033565_33485
AGGTTTAACATTGCTGTTCGTCGCA

>tig00035221_36016
TTTCTTCATATCTCTTCGCGATACT

>tig00035221_36014
TTTCTTCATATCTCTTCGCGATACT

>tig00043176_45054
TTTTGTAGAAACATCCCGAGAT

>tig00035150_35819
GAGGCAGTTACAGATCCTGAAG

>tig00047976_48615
CCCGAGAGCTAGAGATTCCTGT

>tig00012631_16672
TTCCAATAGGTCTAGGACCCCT

>tig00040065_41596
TTTTGTAGAAACATCCCGAGAT

>tig00008991_12441
TTCCAAAAGGTCTAGGACCCCT

>tig00020472_24399
ATTTCTGGGACAACAAAGAGC

>tig00002542_3686
TTTCGGACCGCTCTGGACCTTC

>tig00036860_37899
TGAATTAGCTACCTGAATGGGC

>tig00036482_37532
TTTCGGACCGCTCTGGACCTTC

>tig00038459_39941
TGCGCTGTCGTTCGGCAGGT

>tig00036583_37607
TAGCACTTTTTGATGCTGGGAT

>tig00018886_23231
TTCCAATAGGTCTAGGACCCCT

>tig00018141_22599
AGATATGTTTGATATATTTGGT

>tig00033893_33952
AATTGCACTTGTCCCGGCCT

>tig00027965_28742
AAGCACTAAATGATGCTGGTGT

>tig00043706_45546
AGGCAAGATGTTGGCATAGCTGA

>tig00017184_21689
TCAGCAGTTGTACCACTGATTTG

>tig00019702_23850
TAGTAGTGCACACTGGCCTGGCCC

>tig00038053_39501
TGCACCAAGTCTTTTCTAACCCAT

>tig00036519_37570
TATTATGCTGTTATTCACGAAA

>tig00020050_24074
ACAGTTCCACTACATAGGCTGGCCT

>tig00008210_11507
TTACCACTCCGAGCATTAGCTG

>tig00038924_40472
ACAGTTCCACTACATAGGCTGGCCT

>tig00010667_14393
GAATGCCCCAAGAAGTTTCT

>tig00012253_16268
GAATGCCCCAAGAAGTTTCT

>tig00015257_19534
CTTGAGAACACCTGTTGGACATACA

>tig00036583_37603
TGCACCAAGTCTTTTCTAACCCAT

>tig00038053_39503
TAGCACTTTCTGATGCTGGGTT

>tig00020699_24583
TAGCACTTAGAAATGCTAGGGGC

>tig00033103_32905
TAGCACTTGCCTGTGCTTGGGA

>tig00027965_28744
AAGCACTAGTGCATGCTGGGAA

>tig00022242_25615
TGAAGTAACAAAGAACTTGCAGCCT

>tig00020699_24589
TAGCACTTATGTATGCCGGGGG

>tig00001393_2211
GTCGGTGTAAATGAAGAGACAGTGG

>tig00014095_18399
TCGGTGGGACTTTCGTTCGC

AGAAAATCGTTGGCTGTCCTCCAA

>tig00020699_24585
AAGCACTTATGTATGCTTGGGGG

>tig00038053_39507
TAGCACTATTTTATGCTGGGGC

>tig00020699_24587
TAGCACTTATGTATGCTGGGGG

>tig00038843_40389
CCAGATTATACTTTGTGGTGACATG

>tig00009023_12461
CTTGAGAACACCTGTTGGACATACA

>tig00036583_37605
TAGCACTTTCTGATGCTGGGTT

>tig00004039_5763
TAATATCAGCTGGTAATCCTGAG

>tig00001393_2213
GTCGGTGTAAATGAAGAGACAGTGG

>tig00004981_7171
CCAGATTATACTTTGTGGTGACATG

>tig00010133_13755
CCAGATTATACTTTGTGGTGACATG

>tig00036583_37609
TAGCACTATTTTATGCTGGGGC

>tig00038459_39937
TGCGCTGTCGTCCGGCGGGTAT

>tig00020084_24097
ACTGGTCAGGACAGGTATCAACC

>tig00032621_32193

>tig00012074_16092
ACTGGTCAGGACAGGTATCAACC

>tig00039642_41269
AAAACGAGAGCCCTGAATAGAAC

>tig00002274_3378
TTCGTACGTATGAACAAAAACCT

>tig00030871_29571
TTCGTACGTATGAACAAAAACCT

>tig00030869_29569
TTCGTACGTATGAACAAAAACCT

>tig00018323_22833
GCTGCTTTTAAGTTACTGTGGGAGC

>tig00035462_36374
CTTGAGAACACCTGTTGGACATACA

>tig00014452_18772
ATCTCGGAACACCTCTGTTTTGGA

>tig00033103_32903
TAGCACTTGCCTGTGCTTGGGA

>tig00031481_30364
GCTGCTTTTAAGTTACTGTGGGAGC

>tig00021751_25294
AGACGGGTTATGATATTACTGGAAT

>tig00018678_23082
ATTTCTGTGCATTTGAACTAGAACT

>tig00011614_15614
TTCATACGTATGAACGAAAACCT

>tig00012993_17103
CAAAATGATGATACTTGGGAC

>tig00012874_16928
ACATGCCGGACAGGCCTGACCT

>tig00013220_17464
TATACATAAAGACTTGGGATTGG

>tig00046641_47802
ACGAAAACACTGGAAGAATGCCGT

Annex 2 – miRNA that show a female-biased expression

Name	Sex bias (as log_2FC)	P-value
tig00033932_34112	2.07428864680212	0.00277748461280275
tig00006731_9639	1.28035731052507	0.00377101189636397
tig00007285_10168	1.28035727664026	0.00380525483985817
tig00038994_40525	1.28035724197999	0.00384060551036767
tig00007389_10319	1.75830164391802	0.00391006987964652
tig00039553_41206	1.14831108618447	0.0049525672542177
tig00007285_10164	1.31322596185594	0.00588848490957295
tig00038996_40533	1.3132262571068	0.0060594329883858
tig00009391_12838	1.11226903165296	0.0100799670231518
tig00046325_47611	1.11225886571698	0.0103058359094671
tig00004039_5761	1.11272790488518	0.0114928064764792
tig00039553_41208	1.11272842220751	0.0118422174988511
tig00009391_12836	1.5545237159567	0.0146221151599943
tig00046325_47613	1.55471068946221	0.0152808143001927
tig00002949_4087	1.67569247382539	0.0241306140686574
tig00044182_45882	1.35499623784478	0.0243284688100102
tig00005719_8141	1.6761788198337	0.0243874788517947
tig00008213_11522	1.41642671515223	0.0422422269349546
tig00033932_34114	1.41629569177854	0.042857060882117

Annex 3 – miRNA that show a male-biased expression

Name	Sex bias (as log_2FC)	P-value
tig00003027_4179	-1.26855290453997	0.00291319763114421
tig00001505_2319	-1.26854119957097	0.0030207584222971
tig00045767_47178	-1.4806410254912	0.00719941927378854
tig00044404_46109	-1.25449828351639	0.00806840700497169
tig00044815_46468	-1.25448089456768	0.00854213922896576
tig00009357_12789	-1.03149958476519	0.0137060161823656
tig00030723_29204	-1.30454717985568	0.0389657258392422
tig00041823_43653	-1.20544786122644	0.0394819436929184

Annex 4 – Exact female-biased miRNA targeting with the mRNAs also targeted by tig00009357_12789

miRNA	Target
tig00006731_9639	TRINITY_DN56637_c3_g2_i1
	TRINITY_DN56637_c3_g2_i2
tig00007285_10168	TRINITY_DN56637_c3_g2_i1
	TRINITY_DN56637_c3_g2_i2
tig00009391_12836	TRINITY_DN59888_c2_g1_i6
tig00009391_12838	TRINITY_DN59888_c2_g1_i6
tig00038994_40525	TRINITY_DN56637_c3_g2_i1
	TRINITY_DN56637_c3_g2_i2
tig00039553_41206	TRINITY_DN63196_c0_g1_i2
	TRINITY_DN64283_c1_g2_i4
tig00046325_47611	TRINITY_DN59888_c2_g1_i6
tig00046325_47613	TRINITY_DN59888_c2_g1_i6

Conclusions

During the course of my PhD, I was able to catch a glimpse of the basic reproductive biology of *R. philippinarum*. First of all, drawing a parallelism between a DUI and a SMI species allowed me to discriminate the DUI-related signal from the transcriptional background noise, making me favor a mitochondrial inheritance process that sees ubiquitination or ubiquitin-like modifiers as a primary agent rather than nucleases or autophagy. Specifically, I identified three possible E3 ubiquitin ligases, two in *R. philippinarum* and one in *R. decussatus*, as candidates to perform ubiquitination on male mitochondria, targeting them to destruction upon entering the egg. This process might involve the M-ORFan-encoded protein RPHM21. To approach the conundrum of this protein from another angle, I tried to produce it *in vitro* and in yeast, to no avail. The anomalous difficulty to clone the *rphm21* gene and the insurmountable problems in producing it in a CFPS prompted me to envision that this protein might be toxic to bacteria and an hindrance to bacterial molecular processes even at very low concentrations, and that this toxicity might be the symptom of a meiotic drive restored through separation from its coevolved genomic environment. Finally, given the recent discovery of smithRNAs in our model species, I deemed it timely to delve into the basic features of small noncoding RNA analysis in *R. philippinarum* gonads, compiling a list of newly identified miRNAs and associating to them an overview of their targets' function.

Bibliography

Alexa, A., & Rahnenfuhrer, J. (2019). topGO: Enrichment Analysis for Gene Ontology (Version 2.37.0). https://doi.org/10.18129/B9.bioc.topGO

Bandiera, S., Matégot, R., Girard, M., Demongeot, J., & Henrion-Caude, A. (2013). MitomiRs delineating the intracellular localization of microRNAs at mitochondria. *Free Radical Biology & Medicine*, *64*, 12–19. https://doi.org/10.1016/j.freeradbiomed.2013.06.013

Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. *Nature Reviews Genetics*, *12*(12), 846–860. https://doi.org/10.1038/nrg3079

Bernt, M., Braband, A., Schierwater, B., & Stadler, P. F. (2013). Genetic aspects of mitochondrial genome evolution. *Molecular Phylogenetics and Evolution*, *69*(2), 328–338. https://doi.org/10.1016/j.ympev.2012.10.020

Betel, D., Koppal, A., Agius, P., Sander, C., & Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biology*, *11*(8), R90. https://doi.org/10.1186/gb-2010-11-8-r90

Bilewitch, J. P., & Degnan, S. M. (2011). A unique horizontal gene transfer event has provided the octocoral mitochondrial genome with an active mismatch repair gene that has potential for an unusual self-contained function. *BMC Evolutionary Biology*, *11*(1), 228. https://doi.org/10.1186/1471-2148-11-228

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bonnet, E., Wuyts, J., Rouzé, P., & Van de Peer, Y. (2004). Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics (Oxford, England)*, *20*(17), 2911–2917. https://doi.org/10.1093/bioinformatics/bth374

Bravo Núñez, M. A., Nuckolls, N. L., & Zanders, S. E. (2018). Genetic Villains: Killer Meiotic Drivers. *Trends in Genetics*, *34*(6), 424–433. https://doi.org/10.1016/j.tig.2018.02.003

Breton, S., Beaupré, H. D., Stewart, D. T., Hoeh, W. R., & Blier, P. U. (2007). The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough? *Trends in Genetics*, *23*(9), 465–474. https://doi.org/10.1016/j.tig.2007.05.011

Breton, S., Beaupré, H. D., Stewart, D. T., Piontkivska, H., Karmakar, M., Bogan, A. E., … Hoeh, W. R. (2009). Comparative mitochondrial genomics of freshwater mussels (Bivalvia: Unionoida) with doubly uniparental inheritance of mtDNA: gender-specific open reading frames and putative origins of replication. *Genetics*, *183*(4), 1575–1589. https://doi.org/10.1534/genetics.109.110700

Breton, S., Ghiselli, F., Passamonti, M., Milani, L., Stewart, D. T., & Hoeh, W. R. (2011). Evidence for a fourteenth mtDNA-encoded protein in the female-transmitted mtDNA of marine Mussels (Bivalvia: Mytilidae). *PloS One*, *6*(4), e19365. https://doi.org/10.1371/journal.pone.0019365

Calvo, S. E., & Mootha, V. K. (2010). The Mitochondrial Proteome and Human Disease. *Annual Review of Genomics and Human Genetics*, *11*, 25–44. https://doi.org/10.1146/annurev-genom-082509-141720

Cao, L., Kenchington, E., & Zouros, E. (2004). Differential segregation patterns of sperm mitochondria in embryos of the blue mussel (Mytilus edulis). *Genetics*, *166*(2), 883–894.

Cech, T. R., & Steitz, J. A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, *157*(1), 77–94. https://doi.org/10.1016/j.cell.2014.03.008

Chen, D.-C., Yang, B.-C., & Kuo, T.-T. (1992). One-step transformation of yeast in stationary phase. *Current Genetics*, *21*(1), 83–84. https://doi.org/10.1007/BF00318659

Chen, X., Li, X., Guo, J., Zhang, P., & Zeng, W. (2017). The roles of microRNAs in regulation of mammalian spermatogenesis. *Journal of Animal Science and Biotechnology*, *8*, 35. https://doi.org/10.1186/s40104-017-0166-4

DeLuca, S. Z., & O'Farrell, P. H. (2012). Barriers to male transmission of mitochondrial DNA in sperm development. *Developmental Cell*, *22*(3), 660–668. https://doi.org/10.1016/j.devcel.2011.12.021

Egner, R., Mahé, Y., Pandjaitan, R., & Kuchler, K. (1995). Endocytosis and vacuolar degradation of the plasma membrane-localized Pdr5 ATP-binding cassette multidrug transporter in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *15*(11), 5879–5887. https://doi.org/10.1128/mcb.15.11.5879

Emde, A., & Hornstein, E. (2014). miRNAs at the interface of cellular stress and disease. *The EMBO Journal*, *33*(13), 1428–1437. https://doi.org/10.15252/embj.201488142

Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research*, *40*(1), 37–52. https://doi.org/10.1093/nar/gkr688

Ghiselli, F., Maurizii, M. G., Reunov, A., Ariño-Bassols, H., Cifaldi, C., Pecci, A., … Milani, L. (2019). Natural Heteroplasmy and Mitochondrial Inheritance in Bivalve Molluscs. *Integrative and Comparative Biology*, *59*(4), 1016–1032. https://doi.org/10.1093/icb/icz061

Gissi, C., Iannelli, F., & Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity*, *101*(4), 301–320. https://doi.org/10.1038/hdy.2008.62

Gregorio, N. E., Levine, M. Z., & Oza, J. P. (2019). A User's Guide to Cell-Free Protein Synthesis. *Methods and Protocols*, *2*(1). https://doi.org/10.3390/mps2010024

Gusman, A., Lecomte, S., Stewart, D. T., Passamonti, M., & Breton, S. (2016). Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ*, *4*, e2760. https://doi.org/10.7717/peerj.2760

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hoeh, W. R., Blakley, K. H., & Brown, W. M. (1991). Heteroplasmy suggests limited biparental inheritance of Mytilus mitochondrial DNA. *Science*, *251*(5000), 1488–1490. https://doi.org/10.1126/science.1672472

Hsiao, K.-Y., Sun, H. S., & Tsai, S.-J. (2017). Circular RNA - New member of noncoding RNA with novel functions. *Experimental Biology and Medicine (Maywood, N.J.)*, *242*(11), 1136–1141. https://doi.org/10.1177/1535370217708978

Huang, Z., & Teeling, E. C. (2017). ExUTR: a novel pipeline for large-scale prediction of 3'-UTR sequences from NGS data. *BMC Genomics*, *18*(1), 847. https://doi.org/10.1186/s12864-017-4241-1

Inada, T., Kimata, K., & Aiba, H. (1996). Mechanism responsible for glucose-lactose diauxie in Escherichia coli: challenge to the cAMP model. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms*, *1*(3), 293–301.

Kawashima, Y., Nishihara, H., Akasaki, T., Nikaido, M., Tsuchiya, K., Segawa, S., & Okada, N. (2013). The complete mitochondrial genomes of deep-sea squid (Bathyteuthis abyssicola), bob-tail squid (Semirossia patagonica) and four giant cuttlefish (Sepia apama, S. latimanus, S. lycidas and S. pharaonis), and their application to the phylogenetic analysis of Decapodiformes. *Molecular Phylogenetics and Evolution*, *69*(3), 980–993. https://doi.org/10.1016/j.ympev.2013.06.007

Khuu, C., Nirvani, M., Utheim, T. P., & Sehic, A. (2016). MicroRNAs: Modulators of Tooth Development. *MicroRNA (Shariqah, United Arab Emirates)*, *5*(2), 132–139.

Kohn, A. B., Citarella, M. R., Kocot, K. M., Bobkova, Y. V., Halanych, K. M., & Moroz, L. L. (2012). Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, Pleurobrachia bachei. *Molecular Phylogenetics and Evolution*, *63*(1), 203–207. https://doi.org/10.1016/j.ympev.2011.12.009

Kostylev, M., Otwell, A. E., Richardson, R. E., & Suzuki, Y. (2015). Cloning Should Be Simple: Escherichia coli DH5α-Mediated Assembly of Multiple DNA Fragments with Short End Homologies. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0137466

Krüger, J., & Rehmsmeier, M. (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Research*, *34*(Web Server issue), W451–W454. https://doi.org/10.1093/nar/gkl243

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, *227*(5259), 680–685. https://doi.org/10.1038/227680a0

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Larrieu, I., Tolchard, J., Sanchez, C., Kone, E. Y., Barras, A., Stines-Chaumeil, C., … Giraud, M.-F. (2017). Cell-Free Expression for the Study of Hydrophobic Proteins: The Example of Yeast ATP-Synthase Subunits. In J.-J. Lacapere (Ed.), *Membrane Protein Structure and Function Characterization: Methods and Protocols* (pp. 57–90). https://doi.org/10.1007/978-1-4939-7151-0_4

Lee, C., Yen, K., & Cohen, P. (2013). Humanin: a harbinger of mitochondrial-derived peptides? *Trends in Endocrinology & Metabolism*, *24*(5), 222–228. https://doi.org/10.1016/j.tem.2013.01.005

Liu, B., Li, J., & Cairns, M. J. (2014). Identifying miRNAs, targets and functions. *Briefings in Bioinformatics*, *15*(1), 1–19. https://doi.org/10.1093/bib/bbs075

Lorenz, R., Bernhart, S. H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms for Molecular Biology: AMB*, *6*, 26. https://doi.org/10.1186/1748-7188-6-26

Mackowiak, S. D. (2011). Identification of novel and known miRNAs in deep-sequencing data with miRDeep2. *Current Protocols in Bioinformatics*, *Chapter 12*, Unit 12.10. https://doi.org/10.1002/0471250953.bi1210s36

Margulis, L., & Sagan, D. (1986). *Origins of sex: three billion years of genetic recombination*. Retrieved from http://archive.org/details/originsofsexthre00marg

Meisinger, C., Sommer, T., & Pfanner, N. (2000). Purification of Saccharomcyes cerevisiae Mitochondria Devoid of Microsomal and Cytosolic Contaminations. *Analytical Biochemistry*, *287*(2), 339–342. https://doi.org/10.1006/abio.2000.4868

Michel, S., Wanet, A., De Pauw, A., Rommelaere, G., Arnould, T., & Renard, P. (2012). Crosstalk between mitochondrial (dys)function and mitochondrial abundance. *Journal of Cellular Physiology*, *227*(6), 2297–2310. https://doi.org/10.1002/jcp.23021

Milani, L., Ghiselli, F., Guerra, D., Breton, S., & Passamonti, M. (2013). A Comparative Analysis of Mitochondrial ORFans: New Clues on Their Origin and Role in Species with Doubly Uniparental Inheritance of Mitochondria. *Genome Biology and Evolution*, *5*(7), 1408–1434. https://doi.org/10.1093/gbe/evt101

Milani, L., Ghiselli, F., Iannello, M., & Passamonti, M. (2014). Evidence for somatic transcription of male-transmitted mitochondrial genome in the DUI species Ruditapes philippinarum (Bivalvia: Veneridae). *Current Genetics*, *60*(3), 163–173. https://doi.org/10.1007/s00294-014-0420-7

Milani, L., Ghiselli, F., Maurizii, M. G., Nuzhdin, S. V., & Passamonti, M. (2014). Paternally transmitted mitochondria express a new gene of potential viral origin. *Genome Biology and Evolution*, *6*(2), 391–405. https://doi.org/10.1093/gbe/evu021

Milani, L., Ghiselli, F., & Passamonti, M. (2016). Mitochondrial selfish elements and the evolution of biological novelties. *Current Zoology*, *62*(6), 687–697. https://doi.org/10.1093/cz/zow044

Milani, L., Ghiselli, F., Pecci, A., Maurizii, M. G., & Passamonti, M. (2015). The Expression of a Novel Mitochondrially-Encoded Gene in Gonadic Precursors May Drive Paternal Inheritance of Mitochondria. *PloS One*, *10*(9), e0137468. https://doi.org/10.1371/journal.pone.0137468

Miyamoto, H., Machida, R. J., & Nishida, S. (2010). Complete mitochondrial genome sequences of the three pelagic chaetognaths Sagitta nagae, Sagitta decipiens and Sagitta enflata. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *5*(1), 65–72. https://doi.org/10.1016/j.cbd.2009.11.002

Morgenstern, M., Stiller, S. B., Lübbert, P., Peikert, C. D., Dannenmaier, S., Drepper, F., … Warscheid, B. (2017). Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell Reports*, *19*(13), 2836–2852. https://doi.org/10.1016/j.celrep.2017.06.014

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628. https://doi.org/10.1038/nmeth.1226

Nishimura, Y., Yoshinari, T., Naruse, K., Yamada, T., Sumi, K., Mitani, H., … Kuroiwa, T. (2006). Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1382–1387. https://doi.org/10.1073/pnas.0506911103

Ouimet, P., Kienzle, L., Lubosny, M., Burzyński, A., Angers, A., & Breton, S. (2019a). The ORF in the control region of the female-transmitted Mytilus mtDNA codes for a protein. *Gene*, 144161. https://doi.org/10.1016/j.gene.2019.144161

Ouimet, P., Kienzle, L., Lubosny, M., Burzyński, A., Angers, A., & Breton, S. (2019b). The ORF in the control region of the female-transmitted Mytilus mtDNA codes for a protein. *Gene*, 144161. https://doi.org/10.1016/j.gene.2019.144161

Parab, S., Shetty, O., Gaonkar, R., Balasinor, N., Khole, V., & Parte, P. (2015). HDAC6 deacetylates alpha tubulin in sperm and modulates sperm motility in Holtzman rat. *Cell and Tissue Research*, *359*(2), 665–678. https://doi.org/10.1007/s00441-014-2039-x

Pickworth, S., & Change, M. C. (1969). Fertilization of Chinese hamster eggs in vitro. *Journal of Reproduction and Fertility*, *19*(2), 371–374. https://doi.org/10.1530/jrf.0.0190371

Pope, B., & Kent, H. M. (1996). High Efficiency 5 Min Transformation of Escherichia Coli. *Nucleic Acids Research*, *24*(3), 536–537. https://doi.org/10.1093/nar/24.3.536

Pozzi, A., Plazzi, F., Milani, L., Ghiselli, F., & Passamonti, M. (2017a). SmithRNAs: Could Mitochondria "Bend" Nuclear Regulation? *Molecular Biology and Evolution*, *34*(8), 1960–1973. https://doi.org/10.1093/molbev/msx140

Pozzi, A., Plazzi, F., Milani, L., Ghiselli, F., & Passamonti, M. (2017b). SmithRNAs: Could Mitochondria "Bend" Nuclear Regulation? *Molecular Biology and Evolution*, *34*(8), 1960–1973. https://doi.org/10.1093/molbev/msx140

Riffo-Campos, Á. L., Riquelme, I., & Brebi-Mieville, P. (2016). Tools for Sequence-Based miRNA Target Prediction: What to Choose? *International Journal of Molecular Sciences*, *17*(12). https://doi.org/10.3390/ijms17121987

Ro, S., Ma, H.-Y., Park, C., Ortogero, N., Song, R., Hennig, G. W., … Yan, W. (2013). The mitochondrial genome encodes abundant small noncoding RNAs. *Cell Research*, *23*(6), 759–774. https://doi.org/10.1038/cr.2013.37

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rosenblum, G., & Cooperman, B. S. (2014). Engine out of the chassis: Cell-free protein synthesis and its uses. *FEBS Letters*, *588*(2), 261–268. https://doi.org/10.1016/j.febslet.2013.10.016

Sato, K., & Sato, M. (2017). Multiple ways to prevent transmission of paternal mitochondrial DNA for maternal inheritance in animals. *Journal of Biochemistry*, *162*(4), 247–253. https://doi.org/10.1093/jb/mvx052

Sato, M., & Sato, K. (2011). Degradation of paternal mitochondria by fertilization-triggered autophagy in C. elegans embryos. *Science (New York, N.Y.)*, *334*(6059), 1141–1144. https://doi.org/10.1126/science.1210333

Schrader, M., Costello, J., Godinho, L. F., & Islinger, M. (2015). Peroxisome-mitochondria interplay and disease. *Journal of Inherited Metabolic Disease*, *38*(4), 681–702. https://doi.org/10.1007/s10545-015-9819-7

Simonyan, L., Légiot, A., Lascu, I., Durand, G., Giraud, M.-F., Gonzalez, C., & Manon, S. (2017). The substitution of Proline 168 favors Bax oligomerization and stimulates its interaction with LUVs and mitochondria. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1859*(6), 1144–1155. https://doi.org/10.1016/j.bbamem.2017.03.010

Sinz, A. (2018). Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein–Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angewandte Chemie International Edition*, *57*(22), 6390–6396. https://doi.org/10.1002/anie.201709559

Skibinski, D. O., Gallagher, C., & Beynon, C. M. (1994). Sex-limited mitochondrial DNA transmission in the marine mussel Mytilus edulis. *Genetics*, *138*(3), 801–809.
Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., & Schatten, G. (1999). Ubiquitin tag for sperm mitochondria. *Nature*, *402*(6760), 371–372. https://doi.org/10.1038/46466

Tang, X., & Bruce, J. E. (2009). Chemical Cross-Linking for Protein–Protein Interaction Studies. In M. S. Lipton & L. Paša-Tolic (Eds.), *Mass Spectrometry of Proteins and Peptides: Methods and Protocols* (pp. 283–293). https://doi.org/10.1007/978-1-59745-493-3_17

Theologidis, I., Fodelianakis, S., Gaspar, M. B., & Zouros, E. (2008). Doubly uniparental inheritance (DUI) of mitochondrial DNA in Donax trunculus (Bivalvia: Donacidae) and the problem of its sporadic detection in Bivalvia. *Evolution; International Journal of Organic Evolution*, *62*(4), 959–970. https://doi.org/10.1111/j.1558-5646.2008.00329.x

Törönen, P., Medlar, A., & Holm, L. (2018). PANNZER2: a rapid functional annotation web server. *Nucleic Acids Research*, *46*(W1), W84–W88. https://doi.org/10.1093/nar/gky350

Turcotte, B., Liang, X. B., Robert, F., & Soontorngun, N. (2010). Transcriptional regulation of nonfermentable carbon utilization in budding yeast. *FEMS Yeast Research*, *10*(1), 2–13. https://doi.org/10.1111/j.1567-1364.2009.00555.x

Tutar, Y. (2014). miRNA and cancer; computational and experimental approaches. *Current Pharmaceutical Biotechnology*, *15*(5), 429. https://doi.org/10.2174/1389201015050140828161335

Wang, Z.-X., Kueh, J. L. L., Teh, C. H.-L., Rossbach, M., Lim, L., Li, P., … Stanton, L. W. (2007). Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *25*(9), 2173–2182. https://doi.org/10.1634/stemcells.2007-0085

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Xia, M., Zhang, Y., Jin, K., Lu, Z., Zeng, Z., & Xiong, W. (2019). Communication between mitochondria and other organelles: a brand-new perspective on mitochondria in cancer. *Cell & Bioscience*, *9*. https://doi.org/10.1186/s13578-019-0289-8

Xu, F., Wang, X., Feng, Y., Huang, W., Wang, W., Li, L., … Zhang, G. (2014). Identification of conserved and novel microRNAs in the Pacific oyster Crassostrea gigas by deep sequencing. *PloS One*, *9*(8), e104371. https://doi.org/10.1371/journal.pone.0104371

Yu, H., Kunarso, G., Hong, F. H., & Stanton, L. W. (2009). Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. *The Journal of Biological Chemistry*, *284*(45), 31327–31335. https://doi.org/10.1074/jbc.M109.016162

Zouros, E. (2013). Biparental Inheritance Through Uniparental Transmission: The Doubly Uniparental Inheritance (DUI) of Mitochondrial DNA. *Evolutionary Biology*, *40*(1), 1–31. https://doi.org/10.1007/s11692-012-9195-2