The Pet127 protein is a mitochondrial 5'-to-3' exoribonuclease from the PD-(D/E)XK superfamily involved in RNA maturation and intron degradation in yeasts

Karolina Łabędzka-Dmoch^{1,7}, Michał Rażew^{2,5,7}, Marta Gapińska², Jakub Piątkowski¹, Adam Kolondra¹, Hanna Salmonowicz^{1,3}, Joanna M. Wenda^{1,6}, Marcin Nowotny², and Paweł Golik^{1,4*}

Author affiliations:

¹ Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Warsaw, 02-106, Poland

² Laboratory of Protein Structure, International Institute of Molecular and Cell Biology, Warsaw, 02-109, Poland

³ Laboratory of Metabolic Quality Control, IMOL, Polish Academy of Sciences, Warsaw, 00-783, Poland

⁴ Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, 02-106, Poland

Author notes:

⁵ Present Address: Present Address: European Molecular Biology Laboratory, 71 Avenue des Martyrs, 38042 Grenoble, France

⁶ Present Address: Department of Molecular Biology and Institute for Genetics and Genomics in Geneva, Section of Biology, Faculty of Sciences, University of Geneva, 1211 Geneva, Switzerland

⁷ These authors contributed equally to this work

* To whom correspondence should be addressed. Tel: +48 (22) 592 3234; Fax: +48 (22) 658 4176;

Email: p.golik@uw.edu.pl

Running title: Pet127: a mitochondrial 5'-to-3' exoribonuclease

Keywords: Pet127, exoribonuclease, mitochondria, *Candida albicans*, introns, RNA degradation

ABSTRACT

Pet127 is a mitochondrial protein found in multiple eukaryotic lineages, but absent from several taxa, including plants and animals. Distant homology suggests that it belongs to the divergent PD-(D/E)XK superfamily which includes various nucleases and related proteins. Earlier yeast genetics experiments suggest that it plays a nonessential role in RNA degradation and 5' end processing. Our phylogenetic analysis suggests that it is a primordial eukaryotic invention that was retained in diverse groups, and independently lost several times in the evolution of other organisms. We demonstrate for the first time that the fungal Pet127 protein in vitro is a processive 5'-to-3' exoribonuclease capable of digesting various substrates in a sequence nonspecific manner. Mutations in conserved residues essential in the PD-(D/E)XK superfamily active site abolish the activity of Pet127. Deletion of the *PET127* gene in the pathogenic yeast *Candida albicans* results in a moderate increase in the steady state levels of several transcripts and in accumulation of unspliced precursors and intronic sequences of three introns. Mutations in the active site residues result in a phenotype identical to that of the deletant, confirming that the exoribonuclease activity is related to the physiological role of the Pet127 protein. Pet127 activity is, however, not essential for maintaining the mitochondrial respiratory activity in *C. albicans*.

INTRODUCTION

Mechanisms that assure the presence of correctly processed RNA molecules in tightly regulated quantities are a central feature of gene expression. Pathways that shape the transcriptome show great evolutionary diversity, yet they share the common theme of interplay between transcription and RNA degradation. In general, transcriptional activity in the known genomes extends beyond the annotated gene sequences, often covering the majority of the genome (Kapranov et al., 2007a; Kapranov et al., 2007b; Wade and Grainger, 2014; Tudek et al., 2015; Łabędzka-Dmoch et al., 2021). Intronic RNAs, as well as transcripts originating from noncoding intergenic sequences are quickly degraded by diverse ribonucleases.

Mitochondrial genes are transcribed as polycistronic primary transcripts (from two to dozens, depending on the species) that undergo extensive processing (Ojala et al., 1981; Lipinski et al., 2010; Turk et al., 2013; Kolondra et al., 2015; Łabędzka-Dmoch et al., 2021) by simple single-subunit polymerases of bacteriophage origin (Masters et al., 1987). In wild-type *Candida albicans* there are

eight primary transcripts transcribed from separate (but similar) promoters (Kolondra et al., 2015). Transcripts corresponding to the entire mtDNA sequence, including intergenic regions, can however be observed in RNA degradation impaired strains (Łabędzka-Dmoch et al., 2021), suggesting pervasive transcription encompassing intergenic regions. RNA degradation is also responsible for the varying final abundance of mitochondrial RNAs originating from a single primary transcript (Mercer et al., 2011; Turk et al., 2013; Kolondra et al., 2015; Shang et al., 2018; Łabędzka-Dmoch et al., 2021).

The enzymatic activities involved in RNA degradation include diverse ribonucleases (Houseley and Tollervey, 2009), with exoribonucleases playing the major role in mitochondria (Szczesny et al., 2012). The main ribonuclease of mitochondria is the mtEXO (mitochondrial degradosome) complex, exhibiting a 3'-to-5' exoribonucleolytic activity that is composed of the conserved Suv3 helicase, and either the hydrolytic Dss1 RNase in fungi (Dziembowski et al., 2003; Hoffmann et al., 2008) and in trypanosomes (Mattiaccio and Read, 2008), or the phosphorolytic polynucleotide phosphorylase (PNPase) in animals (Borowski et al., 2013) and plants (Holec et al., 2006). This activity is essential for the functioning of the organellar gene expression system, and its dysfunction results in the loss of respiratory capacity in *S. cerevisiae* (Dmochowska et al., 1995; Golik et al., 1995; Rogowska et al., 2006), *S. pombe* (Hoffmann et al., 2008), and *C. albicans* (Łabędzka-Dmoch et al., 2021). The yeast mtEXO complex shapes the 3' ends of protein coding transcripts that are protected from further degradation by the Rmd9 protein (Hillen et al., 2021), and degrades a variety of junk RNAs (Dziembowski et al., 2003; Rogowska et al., 2006; Łabędzka-Dmoch et al., 2021). In mammalian mitochondria it is also responsible for RNA degradation and surveillance (Szczesny et al., 2010; Dhir et al., 2018; Pietras et al., 2018a; Pietras et al., 2018b). The structural and enzymatic aspects of mtEXO activity have been extensively studied in yeast (Malecki et al., 2007; Razew et al., 2018).

In contrast, the nature and function of the mitochondrial 5'-to-3' exoribonuclease activity, related to the Pet127 protein is not as well understood. A loss-of-function mutation of the *S. cerevisiae* *PET127* gene was found in a screen for suppressors of impaired mitochondrial mRNA-specific translation (Haffter and Fox, 1992). Subsequent genetic and molecular studies indicated that Pet127p is involved in the processing of the 5' termini of mitochondrial transcripts (Wiesenberger and Fox, 1997; Chen et al., 1999; Islas-Osuna et al., 2002; Fekete et al., 2008). Overexpression of the Pet127p can partially suppress the phenotype related to the dysfunction of the mtEXO exoribonuclease complex (Wegierski et al., 1998). Involvement of the Pet127 protein in RNA degradation was also demonstrated in *S.*

pombe (Wiesenberger et al., 2007). Recently, *S. cerevisiae* Pet127p was shown to be a negative regulator of the mitochondrial RNA polymerase Rpo41p through protein-protein interactions that do not depend on Pet127p ribonuclease activity (Corbi and Amon, 2021).

In spite of these diverse functions, deletion of *PET127* in laboratory strains of *S. cerevisiae* does not lead to a loss of respiratory competence (Haffter and Fox, 1992). Likewise, in *S. pombe* Pet127p is not essential for respiratory competence, and its dysfunction has only a moderate phenotype at the elevated temperature (Wiesenberger et al., 2007). In contrast, overexpression of the *S. cerevisiae* *PET127* gene from a strong *ADC1* promoter leads to respiratory failure and loss of functional (*rho*⁺) mtDNA (Wiesenberger and Fox, 1997).

The 5'-to-3' exoribonuclease activity was ascribed to the Pet127 protein on the basis of genetic studies, mentioned above, investigating the changes in mitochondrial RNAs in the mutants. The first indications suggesting that the Pet127 protein may possess intrinsic ribonuclease activity came from *in silico* studies. Even though sequence homology is significant only for true orthologs of Pet127, application of remote homology detection algorithms revealed its similarity to the PD-(D/E)XK superfamily that includes nucleases such as Dxo1 or Rat1 (Steczkiewicz et al., 2012). Significantly, this similarity includes the presence of key catalytic residues in structurally conserved domains.

In order to study the exoribonuclease activity of Pet127p, we purified the protein and tested its activity towards different RNA substrates. As purification of the Pet127 protein of sufficient purity and solubility for enzymatic assays from more conventional model species proved to be challenging, we used the ortholog (TmPet127) from *Talaromyces marneffe*, a thermotolerant dimorphic ascomycetous fungus (Gauthier, 2015).

In order to provide a comprehensive overview of the phenotypic effect of *pet127* null mutation, we also constructed deletant and point catalytic mutant strains in the *Candida albicans* *PET127* gene ortholog (*CaPET127*), and analyzed their phenotypes on physiological and molecular level. In particular, we performed an RNA-seq analysis of the mitochondrial transcriptome in the mutants. Our previous studies (Kolondra et al., 2015; Łabędzka-Dmoch et al., 2021) confirmed the utility of *C. albicans* for mitochondrial transcriptomics, as unlike *S. cerevisiae*, this petite-negative yeast species has a stable mitochondrial genome that does not undergo frequent deletions or rearrangements, obfuscating the results in *S. cerevisiae*.

Our results indicate that even though the Pet127 protein is not essential for the functioning of *C. albicans* respiratory system, its 5'-to-3' exoribonuclease activity, confirmed *in vitro* using the *T. marneffei* ortholog, is involved in mitochondrial RNA metabolism, including degradation of intronic sequences.

RESULTS

Pet127 is an ancient eukaryotic protein lost in multiple independent lineages

Studies of the Pet127 protein concentrated on model yeasts belonging to the fungal phylum of Ascomycota. The absence of Pet127 orthologs in the genomes of multicellular animals and plants is apparent, and homologs from some unicellular Eukaryotes were identified previously (Wiesenberger et al., 2007). Using the wealth of recently obtained genomic sequences that provide new insights into the eukaryotic diversity and phylogeny (Katz and Grant, 2015; Lax et al., 2018; Burki et al., 2020), we looked for Pet127 orthologs among the currently recognized major taxa ('supergroups') (Burki et al., 2020). Orthology was determined using the reciprocal best hits criterion, with the *S. cerevisiae* Pet127 protein sequence (accession number NP_014660.1) as query. In particular, taxa used in recent phylogenomic studies of Eukaryotes (Katz and Grant, 2015; Lax et al., 2018) were checked for the presence of Pet127 orthologs. The results, shown in Figure 1A, reveal that Pet127 is present in several divergent lineages, and absent in others, without a clear phylogenetic pattern. In the TSAR supergroup it can be found in Alveolata (in ciliates and dinoflagellates) and in Rhizaria, whereas in Archaeplastida only in the Rhodophytes (a list of identified taxa possessing a Pet127 ortholog is provided in Supplemental Table S1). Remarkably, it is also present in some members of the basal clade Discoba, as well as in *Guillardia theta*, belonging to Cryptophyta in the supergroup Cryptista. The number of complete sequenced genomes in most Protist lineages is not sufficient to exclude the possibility that some other clades may also harbor the Pet127 protein, however, given the large number of known plant genomes, the lack of the *PET127* gene in Chloroplastida appears certain.

The situation in Amorphea, the supergroup that contains Opisthokonta, Amoebozoa, and related clades, is of particular interest (Figure 1B). Pet127 is entirely absent from Holozoa, the clade containing animals and related protist taxa. It is, however, present in the majority of Fungi, including Dikarya, the superphylum containing, among others, the extensively studied Ascomycota and Basidiomycota. It can also be found in *Fonticula*, a cellular slime mold related to Fungi, and in

Thecamonas, a zooflagellate classified in Apusomonada (Paps et al., 2013). Among Amoebozoa, orthologs of Pet127 are encoded in the genomes of several dictyostelid slime molds (and one protostelid species), but neither in physarids, nor in unicellular amoebae with known genomic sequences.

The maximum likelihood phylogenetic tree of Pet127 amino acid sequences (Supplemental Figure S1) is generally consistent with the proposed eukaryotic phylogeny, ruling out recent horizontal gene transfer as the explanation of the observed pattern of conservation. This suggests that Pet127 is an ancient eukaryotic protein that was probably present in the last common eukaryotic ancestor, and its absence in multiple lineages is most likely a result of several independent gene loss events.

Mitochondrial localization of the Pet127 protein in Fungi was demonstrated experimentally in yeast *S. cerevisiae* (Wiesenberger and Fox, 1997), and can be predicted for other fungal orthologs. *In silico* prediction of protein localization in nonmodel organisms, particularly in taxa that were not extensively studied experimentally, remains a challenge. However, the application of an ML algorithm designed to predict the localization of proteins in mitochondria and mitochondrion-related organelles in nonmodel organisms (Kume et al., 2018) indicates that Pet127 orthologs from dinoflagellates, Rhodophyta, and Discoba should also localize to mitochondria, suggesting that subcellular localization is also a conserved feature of this protein.

Alignment of Pet127 amino acid sequences (Figure 1C) shows that the core PD-(D/E)XK nuclease domain, including the key catalytic residues (Kosinski et al., 2005; Steczkiewicz et al., 2012) is conserved in diverse eukaryotic lineages, whereas the other regions show significant divergence, suggesting that the exoribonuclease activity of the protein is maintained by selective pressure. In the region responsible for the regulatory interaction with the Rpo41 mitochondrial RNA polymerase, corresponding to residues 48-215 in *S. cerevisiae* (Corbi and Amon, 2021), only a short stretch (residues 178-208) shows conservation (Supplemental Figure S2). This suggests that this interaction is either not universal, or it depends on those segments of the Rpo41 protein, located in the N-terminal domain, that are also highly divergent (Ringel et al., 2011; Yang et al., 2015).

Pet127 is not essential for the functioning of *C. albicans* mitochondrial respiratory system

In *C. albicans* the Pet127 ortholog (CaPet127) is encoded by the C1_11070W gene (orf19.2309), and is a protein of 705 amino acids sharing 24% sequence identity and 42% sequence similarity with the *S. cerevisiae* protein (accession number NP_014660.1). The homozygous deletion mutant

$\Delta Capet127/\Delta Capet127$ was constructed in the background of the BWP17 wild-type laboratory strain (Wilson et al., 1999) using two-step PCR based targeting (Walther and Wendland, 2008). The heterozygous $\Delta Capet127/CaPET127$ strain obtained in the first step was also used in respiratory growth tests.

Additionally, the $Capet127_{D375A}/Capet127_{D375A}$ point mutant strain was also constructed, encoding a form of the protein with the conserved D375 residue in the PD-(D/E)XK core nuclease motif (Figure 1C) changed to alanine. This residue corresponds to D378 in the *S. cerevisiae* protein (ScPet127), and to D388 in *Talaromyces marneffei* (TmPet127).

Respiratory competence of the deletant and point mutant strains was compared to that of the wild-type parental strain by assessing growth on solid media with non-fermentable (glycerol) and fermentable (glucose) carbon source (Figure 2). The test was performed at 30°C and 37°C. In all the tested conditions the growth of homozygous and heterozygous deletants, as well as the point mutants, was indistinguishable from that of the wild-type strain. We can therefore conclude that the loss of Pet127 activity does not affect the respiratory growth capacity of *C. albicans*.

Changes in steady state mRNA levels and increased intron accumulation in Pet127 deficient mutants

We performed a series of Northern blot experiments using RNA isolated from mitochondria of the homozygous $\Delta Capet127/\Delta Capet127$ deletant, as well as the $Capet127_{D375A}/Capet127_{D375A}$ point mutant strain in order to investigate the effect of Pet127 dysfunction on the major mitochondrial transcripts (Figure 3). While the changes in steady state levels and processing patterns of multiple transcripts are detectable, they are minor. None of the mRNAs and rRNAs show significantly decreased or largely aberrant expression, consistently with the observed lack of respiratory growth impairment. The size of the mature mRNAs and rRNAs in the mutant strains appears to be unchanged compared to the wild-type control. This contrasts with the observations made in *S. cerevisiae*, where several transcripts show decreased steady-state levels and increased length in *pet127 Δ* strains (Wiesenberger and Fox, 1997).

The three bicistronic transcripts encoding subunits of Complex I: NAD6-NAD1, NAD2-NAD3, and NAD4L-NAD5, as well as the monocistronic NAD4 mRNA all show increased steady state levels (two- to five-fold according to blot quantification) in the mutant strains (Figure 3A). In the case of NAD2, NAD3, and NAD4L probes, hybridization reveals that in mutant strains both bicistronic and

monocistronic forms are present, even though the downstream ORFs overlap with the upstream ones, and are presumably translated only as a part of the bicistron (Kolondra et al., 2015). Similarly, the steady state level of the mature mRNA encoding COX1 and, to a lesser extent, COX3 are slightly increased, whereas the mature COX2 mRNA level appears to be unchanged (Figure 3B). The bicistronic ATP6-ATP8 transcript shows slightly decreased steady-state level of the mature mRNA (Figure 3C), and both rRNAs are unaffected (Figure 3D). In the case of the COB transcript, the level of mature mRNA is unaffected, but there is a visible increase in the signal from the processing intermediates (indicated by arrows on Figure 3B).

In order to investigate the effect of Pet127 dysfunction on the mitochondrial transcriptome, we performed RNA-seq analysis on mitochondrial RNA from the homozygous $\Delta Capet127/\Delta Capet127$ deletant, as well as the *Capet127*_{D375A}/*Capet127*_{D375A} point mutant (two independent isolates that gave identical results), comparing them with the wild-type BWP17 strain. RNA sequencing libraries were prepared from isolated mitochondria and sequenced using the Ion Torrent Proton™ NGS System. We used the same RNA-seq workflow to describe the mitochondrial transcriptome of wild-type *C. albicans* (Kolondra et al., 2015), and to analyze the phenotype of mutants deficient in the mtEXO function (Łabędzka-Dmoch et al., 2021). The mtDNA sequence of *C. albicans* strain SC5314 (GenBank:AF285261.1), which is known to be identical to the mtDNA sequence of BWP17 (Kolondra et al., 2015), with one of the two identical 6.8 kb repeat regions removed, was used as the template for read mapping. 32 – 36 million reads were obtained from each sample, of which 92% – 93% mapped to the reference sequence.

Visualizing the distribution of coverage depth (Figure 4A) shows that the mutant strains do not show apparent differences from the wild-type, unlike the previously described mtEXO deficient mutants (Łabędzka-Dmoch et al., 2021), where the density distribution was significantly shifted towards higher coverage values. Counting reads mapping to the annotated primary transcripts (Kolondra et al., 2015) in the sense and antisense orientation shows that in all the analyzed strains the majority of reads (99.8%) map to the transcription units, with only 0.2% of reads corresponding to intergenic regions, regardless of the presence of the functional Pet127 protein. Of the reads mapping to the primary transcripts, only 0.01% are in the antisense (“mirror”) orientation, again with no differences between the wild-type and the mutants. This corresponds to the results obtained in wild-type *C. albicans* mitochondria in the previous studies, and is in stark contrast to the phenotype of

mtEXO mutants, where both intergenic region and antisense RNAs show significant increase (Kolondra et al., 2015; Łabędzka-Dmoch et al., 2021). These results suggest that general mitochondrial RNA degradation and surveillance is not impaired in strains lacking the Pet127 activity.

Coverage graphs of regions corresponding to different mitochondrial genes (Supplemental Figure S3) shows that apart from a moderate increase in coverage, consistent with the results of Northern hybridizations, there are no significant qualitative differences between the mutants and the wild-type. Notably, there is no evidence of any unprocessed 5' extensions that were observed in *S. cerevisiae* *pet127Δ* strains (Wiesenberger and Fox, 1997). The most apparent phenotype is related to an increase in reads mapping to introns, evident in the case of the second intron (i2) of the RNL gene (Figure 4B). Counting reads mapping to introns and exons of RNL, COB and COX1 genes (Figure 4C) shows that overall the proportion of intronic reads is increased about 1.5-fold in the deletant and mutant strains compared to wild-type. This is mostly due to the accumulation of the second intron of RNL (7-fold increase), the first intron of COB (1.7-fold increase), and the third intron of COX1 (1.8-fold increase), the remaining introns are mostly unaffected.

In order to verify these observations we performed semiquantitative RT-PCR of amplicons internal to the first intron of COB and the second intron of RNL, encompassing relevant exon-intron junctions, and across the introns from exonic primers (Figure 4D). In the case of the first intron of COB, the increase is apparent for the internal intronic amplicon and both exon-intron junctions, indicating that the unspliced precursor and, potentially also the excised intron are accumulated in the Pet127 deletants. Similar results were observed for the second intron of RNL, with additional evidence from the amplification across the intron from primers located in the second and third exon, showing that the amount of spliced product is unchanged, while the unspliced precursor accumulates (in the case of COB amplification across the intron was not successful).

In order to estimate the quantitative changes in transcript levels in Pet127 deficient strains, we calculated fold change of normalized read counts mapping to each gene compared to wild-type. The results (Figure 5) are consistent with the observations made by quantitatively analyzing the Northern blots, with a two- to four-fold increase in the expression of transcripts encoding the Complex I subunits (NAD genes) as the most prominent change.

Overall, the changes in the Northern hybridization and RNA-seq patterns are moderate, quantitative rather than qualitative, and consistent with the lack of observable impairment of

respiratory function in the deletant. In all the experiments, the D375A substitution resulted in a phenotype that was indistinguishable from that of the complete deletion, indicating that this mutation is sufficient to produce a nullomorph allele with regard to the postulated exoribonuclease activity.

Purified *T. marneffei* Pet127 protein is a progressive 5'-to-3' exoribonuclease.

Genetic and molecular studies described above in *C. albicans*, as well as those published earlier in *S. cerevisiae* (Wiesenberger et al., 2007; Fekete et al., 2008) and *S. pombe* (Wiesenberger et al., 2007) suggest that the Pet127 protein has a 5'-to-3' exoribonuclease activity. The presence of conserved sequence and structure motifs indicates that it belongs to the PD-(D/E)XK superfamily (Steczkiewicz et al., 2012). In order to investigate the enzymatic activity of Pet127 we attempted to purify the catalytically active protein from different fungal species. The best results were obtained using the Pet127 ortholog from the filamentous fungus *Talaromyces marneffei*. It is a protein of 849 amino acids (accession number XP_002153343.1) sharing 24% sequence identity and 40,5% sequence similarity with the *C. albicans* protein, and 26% identity and 44% similarity with the *S. cerevisiae* ortholog. The conserved region encompasses the putative nuclease domain (Figure 1C).

To determine Pet127 ribonucleolytic activity, we established the expression and purification protocol, then purified the full length wild type protein as well as its two point mutants D388A and Q386A in the predicted motif II of PD-(D/E)XK nuclease active site. These two residues correspond to D375 and Q373 in *C. albicans*, as well as to D236 and E234 of the extensively studied mammalian Dxo1 nuclease (Chang et al., 2012; Jiao et al., 2013), respectively. Its crystal structure (DOI: 10.2210/pdb4J7L/pdb) shows that both residues are responsible for coordination of the two magnesium ions in the active site (Jiao et al., 2013). Based on the sequence similarity of the catalytic domain within members of the PD-(D/E)XK nuclease superfamily, mutations of these residues in TmPet127 should also lead to inhibition of its ribonuclease activity. Northern blot and RNA-seq experiments, described above, demonstrated that the D375A mutation in *C. albicans* (corresponding to D388A in *T. marneffei*) has a phenotype indistinguishable from that of the complete deletion.

To investigate the exoribonuclease activity of Pet127, we conducted activity assays against substrates (sequences of the substrates are listed in Supplemental Table S2 and in the Figure 6) labelled with fluorescein either on 5' end (Figure 6A) or 3' end (Figure 6B and 6C). While 3'-labelled substrates were efficiently degraded by Pet127, the 5' – labelled substrate remained undigested. This indicates that the Pet127 protein has an intrinsic 5'-to-3' exoribonuclease activity, consistent with the

hypothesis formulated based on genetic studies and sequence analysis. In this experiment, Pet127 is active towards different substrates. The activity towards a substrate which comprised a 5'-terminal tract of Gs is slightly diminished in the comparison to an RNA of the same length without a 5'-terminal G (Figure 6B). This feature was also found within Dxo1/Rai1 hydrolases, members of the same PD-(D/E)XK superfamily (Doamekpor et al., 2020). Finally, we tested Pet127 activity towards a hairpin-structure-forming RNA, derived from the *C. albicans* ATP6/8 mRNA transcript (Figure 6C) (Kolondra et al., 2015). Pet127 is able to degrade this type of structured RNA, although with lower efficiency than for unstructured RNA. The single stranded end of this substrate is degraded quickly, and the remaining ds-RNA hairpin appears to be partially resistant to degradation. This suggests that the Pet127 5'-to-3' exoribonuclease activity is not sequence or structure-dependent, although it is partially inhibited by 5'-terminal G tracts and strong secondary structure regions. Degradation of all the substrates is complete, leaving only a 3'-FAM labelled nucleotide as the final product (Supplemental Figure S4), even at the earliest time points, suggesting that the exoribonuclease activity of Pet127 is processive.

In all the assays, the activity of D388A and Q386A TmPet127 variants was severely reduced (Figure 6). The D388A substitution removed any detectable activity of the protein, consistent with the experiments performed in the D375A *C. albicans* mutant, which was indistinguishable from the deletant in RNA experiments. The substitution at Q386 also severely diminished the exoribonuclease activity, although traces of the degradation product, corresponding to 6-7% of the wild-type activity, can be detected at later time points.

DISCUSSION

The PD-(D/E)XK protein superfamily is very large and extremely diverse, containing members with a number of different roles and activities, including both active nucleases and their inactive homologs (Steczkiewicz et al., 2012). It includes prokaryotic restriction endonucleases, proteins involved in DNA repair and recombination, tRNA intron splicing factors, and exoribonucleases involved in mRNA decapping and degradation in Eukaryotes (Poole and Stevens, 1995; Stevens and Poole, 1995; Belfort and Weiner, 1997; Ban and Yang, 1998; Aravind et al., 2000; Hickman et al., 2000; Kosinski et al., 2005; Orłowski and Bujnicki, 2008; Xiang et al., 2009; Steczkiewicz et al., 2012; Jiao et al., 2013; Doamekpor et al., 2020). Exoribonucleases containing the PD-(D/E)XK fold include both processive

(Xiang et al., 2009; Miki and Großhans, 2013) and distributive (Jiao et al., 2013; Doamekpor et al., 2020) enzymes. Identification of the PD-(D/E)XK motif and the associated conserved fold in the Pet127 protein (Steczkiewicz et al., 2012) was therefore not sufficient to infer its enzymatic activity.

A 5'-to-3' exoribonuclease activity was suggested for Pet127 based on a series of genetic studies conducted in model yeasts *S. cerevisiae*, where the *pet127Δ* strains accumulate several transcripts with longer, unprocessed 5' termini (Wiesenberger and Fox, 1997; Fekete et al., 2008). In the best studied case of the COB transcript, the 5' end of the mature mRNA is determined by the site of Cbp1 protein binding, which protects it from the 5'-to-3' exoribonuclease activity that is dependent on the product of the *PET127* gene (Fekete et al., 2008). Degradation dependent on the Pet127 protein activity was also shown to play a role in producing different steady-state levels of mature RNAs derived from a single primary transcript encompassing COB and tRNA-E (Krause, 2004). None of these studies, however, provided decisive proof of intrinsic ribonuclease activity in the Pet127 protein.

In this work we also purified the *T. marneffei* Pet127 protein and conducted ribonuclease activity assays. In order to verify whether the PD-(D/E)XK domain is responsible for Pet127 activity, we also purified mutein forms with substitutions in two amino acid residues that are crucial for the catalytic activity by coordinating magnesium cations in the active site of the enzyme, as evidenced from the study of mammalian Dxo1 exoribonuclease (Jiao et al., 2013). These two residues are strictly conserved in all the Pet127 orthologs we identified in the available sequence databases.

Our results (Figure 6) indicate that the purified TmPet127 protein does possess a readily detectable exoribonuclease activity *in vitro*, and that mutations in the residues predicted to be crucial for the PD-(D/E)XK domain active site completely, or nearly completely abolish this activity. We can thus conclude that Pet127 is a *bona fide* ribonuclease of the PD-(D/E)XK superfamily. Assays using 5' or 3' labelled substrates indicate that RNA degradation by the TmPet127 protein occurs from the 5' terminus, and the lack of any detectable intermediates even at the earliest time points suggests a processive exoribonucleolytic activity. Pet127 *in vitro* is capable of complete degradation of the substrate, leaving only the single labelled nucleotide as product, similar to the Dxo1 exoribonuclease (Chang et al., 2012; Jiao et al., 2013). Results of assays with different substrates indicate that the TmPet127 activity is not sequence-specific, although its activity towards a substrate containing multiple Gs at the 5' terminus is visibly reduced. It is even capable of degrading, albeit with a reduced efficiency, a substrate containing a region of strong hairpin secondary structure. Unlike another well

studied mitochondrial exoribonuclease, the Dss1 protein, where its intrinsic 3'-to-5' exoribonuclease activity is low, and it requires the action of the Suv3 RNA helicase to efficiently degrade even short unstructured substrates (Malecki et al., 2007; Razew et al., 2018), the Pet127 5'-to-3' exoribonuclease is highly active *in vitro* without any accessory proteins.

Most data on the physiological role of the Pet127 exoribonuclease come from the studies in model yeast *S. cerevisiae*. The most apparent phenotype of mutants lacking Pet127p in this species is associated with abnormal processing of 5' termini of multiple mRNA and rRNA transcripts, associated with visible decrease in the mRNA levels of COX2, ATP6, and ATP9 (Wiesenberger and Fox, 1997). In contrast, the loss of the Pet127 protein in *Schizosaccharomyces pombe* does not cause apparent RNA maturation defects, but results in a 1.5-4 fold increase in the steady state level of mature RNAs, most apparent for COX2 and COX3, suggesting that in this species the Pet127 exoribonuclease is involved in general RNA degradation rather than processing (Wiesenberger et al., 2007). In order to further investigate the biological function of Pet127, we performed genetic analysis in *Candida albicans*, which is an attractive mitochondrial model due to the stability of its mtDNA (no formation of cytoplasmic petites), and well-studied organellar transcriptome (Kolondra et al., 2015; Łabędzka-Dmoch et al., 2021).

The results of molecular phenotype investigations using Northern blot hybridizations (Figure 3) and RNA-seq transcriptome analysis (Figures 4 and 5) indicate that in *C. albicans*, unlike in *S. cerevisiae*, there are no apparent defects in RNA maturation in deletants and mutants lacking Pet127. The number of RNA-seq reads covering 5' regions of transcripts does increase, but there is no detectable extension of the mature RNAs upstream beyond the 5' terminus of transcripts observed in the wild-type strain. As in *S. pombe*, the steady-state levels of several mRNAs increase in the mutants by a similar amount (2—4 fold), but this increase does not involve all the transcripts (Figure 5). The effect is most apparent for the mRNAs of genes encoding the subunits of Complex I – the bicistronic NAD6-NAD1, NAD2-NAD3, and NAD4L-NAD5 mRNAs, as well as the monocistronic NAD4. The expression of tRNA-Asn, located in the first transcription unit just upstream of NAD6 is also increased. Several transcripts, mostly tRNAs and ATP6-ATP8 bicistronic mRNA show a slight (by about 25%) decrease in expression. Curiously, the transcripts that show visibly increased levels in the mutant strains are located close to the 3' end of their respective polycistronic primary transcription units (see Figure 3E). tRNA-Asn, NAD6, and NAD1 form the 3' terminal half of the first transcription unit (TU1) that starts

with RNL, NAD2 and NAD3 are downstream of COX1 in TU5, and NAD4 is transcribed as a part of TU8 downstream of tRNA-Leu and tRNA-Met-2. This suggests that degradation by the Pet127 5'-to-3' exoribonuclease probably occurs after the primary transcripts are processed into separate RNAs by the tRNA punctuation mechanism (Kolondra et al., 2015).

Another phenotype of Pet127 deficient mutants in *C. albicans* is related to the accumulation of intronic sequences, evidenced by RNA-seq and confirmed by RT-PCR (Figure 4). The effect is most pronounced for the second intron in RNL, with about 7-fold increase in the fraction of reads mapping to the intron among all reads mapping to the gene. A more moderate increase is also observed for the first intron of COB (1.7-fold), and the third intron of COX1 (1.8-fold). Semiquantitative RT-PCR analysis performed for the second intron of RNL and the first intron of COB confirms the RNA-seq results and additionally shows that the increase involves not only the excised intronic RNA, but also the unspliced precursor. These changes do not, however, significantly affect the levels of respective mature spliced RNAs. This indicates that in *C. albicans* the Pet127 exoribonuclease is involved in the degradation of introns and splicing intermediates, a role that was previously ascribed mostly to the main mitochondrial mtEXO ribonuclease (Golik et al., 1995; Dziembowski et al., 2003; Łabędzka-Dmoch et al., 2021). Based on the available data we cannot, however, explain why certain introns accumulate in mutants lacking Pet127, while others in the same gene do not.

In principle, observed changes in RNA steady-state levels could be attributed to changes in transcription, as well as to impaired degradation. Mitochondrial transcripts are, however, polycistronic, and changes in transcription would affect multiple RNAs transcribed from the same promoter, as well as multiple introns of the same gene. This is clearly not the case here, as, for example, the level of the COX2 transcript, transcribed from the same promoter as the visibly increased NAD1 and NAD6 RNAs, remains unchanged (see Figure 3E). Similarly, only the first of the two COB introns shows accumulation in the mutant strains. It is thus reasonable to ascribe the observed phenotype to changes in RNA degradation rather than transcription, using steady-state transcript levels as proxy.

These results naturally raise questions about the functional interplay between the known mitochondrial exoribonucleases. In the nucleus and cytoplasm, both the 5' and 3' exoribonucleolytic pathways participate in RNA turnover and surveillance (Houseley and Tollervey, 2009). In mitochondria, the 3'-to-5' exoribonucleolytic degradation pathway provided by the mtEXO complex is universally conserved, although the exoribonuclease subunit itself can be one of two different types: a

phosphorolytic PNPase in plants and animals, and a hydrolytic Dss1 ortholog in other lineages (Dziembowski et al., 2003; Holec et al., 2006; Hoffmann et al., 2008; Mattiaccio and Read, 2008; Borowski et al., 2013). The mtEXO activity is essential for the functioning of the mitochondrial genetic system (Dziembowski et al., 2003; Rogowska et al., 2006; Szczesny et al., 2010; Dhir et al., 2018; Pietras et al., 2018a; Pietras et al., 2018b; Łabędzka-Dmoch et al., 2021). In contrast, the 5'-to-3' activity of Pet127 is neither essential, nor universal. Nullomorphic *pet127* mutants of *S. cerevisiae* and *S. pombe* show no impairment of respiratory function at normal growth conditions, and only a moderate phenotype at elevated temperatures (Wiesenberger and Fox, 1997; Wiesenberger et al., 2007). Similarly, in *C. albicans* the loss of Pet127 does not produce any measurable growth impairment on respiratory media. In *S. cerevisiae*, overexpression of Pet127 partially rescues the phenotype of reduced mtEXO activity caused by the deletion of the gene encoding the Suv3 helicase component (Wegierski et al., 1998), suggesting that the two RNA degradation pathways are to a certain extent redundant.

A recent study (Corbi and Amon, 2021) demonstrated that in *S. cerevisiae* Pet127 acts as a negative regulator of the mitochondrial RNA polymerase (Rpo41) in a manner that is independent of its RNase domain, raising a possibility that the principal biological role of this protein could be regulatory rather than enzymatic. In our study, however, a single amino acid substitution (D375A in CaPet127, equivalent to D388A in TmPet127) that abolished the ribonucleolytic activity produced a phenotype indistinguishable from that of a complete deletion in RNA based assays, suggesting that it is the exoribonuclease activity which responsible for the phenotype in *C. albicans*. With respect to changes in the steady-state levels of mitochondrial transcripts this phenotype also resembles that of the *pet127Δ* deletant in *S. pombe*, suggesting that the involvement of Pet127 in general RNA degradation is the conserved function of this protein, at least in diverse Ascomycete lineages. The PD-(D/E)XK domain with its critical residues is also the part of the Pet127 protein sequence which is conserved among orthologs from diverse eukaryotic lineages. The primary conserved role of Pet127 thus appears to be related to its exoribonuclease activity rather than the non-enzymatic regulatory function.

The presence of Pet127 orthologs in diverse lineages, including early branching groups like Discoba (Figure 1), suggests that it is an ancient eukaryotic protein. Pet127 forms a distinct subfamily among PD-(D/E)XK nucleases, with significant similarity on the sequence level limited to true

orthologs. Its relationship to other members of the superfamily, like Dxo1 (Chang et al., 2012; Jiao et al., 2013; Doamekpor et al., 2020) or Rat1 (Xiang et al., 2009) becomes apparent only upon application of remote homology detection methods (Steczkiewicz et al., 2012). Mitochondrial localization of Pet127 is experimentally supported in Fungi, and predicted in organisms as diverse as dinoflagellates, Rhodophyta, and Discoba. There is no convincing evidence for the presence of any Pet127 homologs in Bacteria or Archaea (a few positive hits in Bacteria show very high similarity to fungal sequences and thus probably are a contamination artifact or a result of recent horizontal gene transfer). In Bacteria, the only common 5'-to-3' exoribonuclease is RNase J, also found in chloroplasts of plants (Halpert et al., 2019), which is unrelated to Pet127 and does not belong to the PD-(D/E)XK superfamily. Pet127 appears thus to be an early eukaryotic evolutionary invention, and its absence of Pet127 in multiple lineages, most apparent in Holozoa (including animals) and Chloroplastida (including higher plants) is most parsimoniously explained by independent gene loss events. Genetic studies in three different yeast species show that Pet127 is not essential for the functioning of mitochondria, and its loss does not produce an obvious phenotype under laboratory conditions. It is thus not surprising that it was easily lost in the evolution of several distinct evolutionary lineages. What is more perplexing is the actual biological function of Pet127 underlying its conservation in all the diverse eukaryotic groups that retained its orthologs. The organisms that either lost or retained Pet127 represent a multitude of diverse mitochondrial genome organizations (Gray et al., 2004; Wideman et al., 2020; Zardoya, 2020) and include both unicellular and multicellular organisms.

The results demonstrating the role of Pet127 in the degradation of intron-containing transcripts prompted us to investigate whether mitochondrial introns, particularly those belonging to Group I, could be linked to its biological function. Animal mitochondria are mostly intronless, with the exceptions mostly attributed to horizontal gene transfer events, and plant mitochondria are rich in Group II introns, but not in Group I (Haugen et al., 2005; Nielsen and Johansen, 2009; Huchon et al., 2015; Mukhopadhyay and Hausner, 2021). There are, however, no Group I introns in the mitochondrial genomes of Rhodophyta, and ciliate mtDNAs generally lack introns, yet both groups contain Pet127 orthologs. On the other hand, multiple Group I introns are present in the mitochondrial genomes of Choanoflagellata (Burger et al., 2003), which lack Pet127 like all other known Holozoa. The key to the primary biological function of Pet127 and its conservation in some, but not all, eukaryotic lineages must therefore lie elsewhere, and remains to be elucidated. Investigating the

primary function of this mysterious exoribonuclease will entail moving beyond experiments in model organisms under laboratory conditions, and require combining functional and evolutionary approaches. We believe that it shall reveal interesting insights into the origin and functioning of the eukaryotic cell architecture.

MATERIAL AND METHODS

Strains and media

Candida albicans strain BWP17(*arg4::hisG/arg4::hisG, his1::hisG/his1::hisG, ura3::imm434/ura3::imm434, iro1::imm434 /iro1::imm434*) (Wilson et al., 1999) was used as the wild-type control and the starting point for construction of mutants. The Δ *Capet127/* Δ *Capet127* strain (*arg4/arg4, his1::hisG/his1::hisG, ura3::imm434 /ura3::imm434, iro1::imm434 /iro1::imm434, pet127::HIS1/pet127::SAT*) was constructed using PCR based targeting (Walther and Wendland, 2008). The first allele was disrupted by integration of the *CaSAT1* gene with ~40 nt flanks homologous to the upstream and downstream sequence of *CaPET127*, amplified by PCR on the template of pFASAT1 (Walther and Wendland, 2008). In the second round, a deletion cassette was constructed by yeast recombinational cloning (Colot et al., 2006). The *CaHIS1* gene from the plasmid pFAHIS1 (Walther and Wendland, 2008), together with ~1 kb flanks upstream and downstream of *CaPET127*, all obtained by PCR, were used for *in vivo* recombination into the pRS426 vector in the *S. cerevisiae* (*MAT α , ade2-1, leu2-3, 112 ura3-,1 trp1-1, his3-11, 15 can1-100 [rho⁺ intronless]*) strain (Saint-Georges et al., 2002). Deletion cassettes were amplified on the template of recombined plasmids by PCR and introduced into *C. albicans* heterozygous strains by electroporation (Reuss et al., 2004). Genotypes of obtained knockouts were confirmed by PCR.

The *Capet127_{D375A}/* *Capet127_{D375A}* strain (*arg4/arg4, his1::hisG/his1::hisG, ura3::imm434 /ura3::imm434, iro1::imm434 /iro1::imm434, PET127::pet127_{D375A}/PET127:: pet127D375A, eno1::CaCAS9-SAT/ENO1*) was constructed using the *C. albicans* CRISPR/Cas9 Solo system as described previously (Vyas et al., 2015). The guide sequence 5'- CTTTTTATTGAGATCACAGT -3' started 23 nt upstream of the GAT encoding Asp (D). Aspartic acid to alanine substitution was performed by introduction of repair template bearing GAT (Asp)-GCA (Ala) mutation introducing SphI restriction site used for diagnostic RLFP and silent PAM mutation (TTG to TTA encoding leucine). To obtain repair template first the pET28aSumo vector with cloned CaPET127 was subjected to site

directed mutagenesis with the primers KL182 and KL183. Later the vector pET28aPET127_{D375A}Sumo was used as a PCR template with the primers KL184 and KL185 to amplify the donor DNA for homologous recombination.

Sequences of all primers used in the construction and verification of strains are listed in the Supplemental Table S2.

For isolation of mitochondria strains were grown in liquid YPGal medium (1% yeast extract, 2% peptone and 2% galactose) containing 80 g/l uridine at 37°C until logarithmic growth phase. Respiratory growth was tested on agar plates with either YPD (1% yeast extract, 2% peptone and 2% glucose) or YPG (1% yeast extract, 2% peptone and 2% glycerol).

Isolation of mitochondria and RNA extraction

RNA was obtained from purified mitochondria isolated from log-phase liquid cultures of *C. albicans* grown in YPGal as described previously (Kolondra et al., 2015). Isolation of mitochondrial RNA was performed by hot phenol procedure (Schmitt et al., 1990) or by FenoZol (A&A Biotechnology) extraction. RNA samples were treated with 0.6U of DNase I (Roche) per µg of RNA according to the manufacturer's protocol. DNase-treated RNA was phenol-extracted, precipitated, and resuspended in water as described previously (Kolondra et al., 2015).

Northern blot analysis [KLD, PG]

Northern hybridization was performed essentially as described previously (Malecki et al., 2008; Kolondra et al., 2015). Briefly, 2.5 µg of mitochondrial RNA was separated on a 1 % denaturing formaldehyde gel, transferred onto Nytran N nylon membrane (GE Healthcare) and hybridized with the appropriate probe. The probes and ³²P labeling protocols used to detect each transcript were described previously (Kolondra et al., 2015). The blots were scanned on the Typhoon FLA 9000 (GE) biomolecular imager, and quantitative analysis was performed in Fiji (Schindelin et al., 2012).

Transcriptome sequencing, mapping and analysis

The RNA-seq experiments were performed essentially as described previously (Łabędzka-Dmoch et al., 2021). Briefly, mitochondrial RNA-seq libraries were obtained using 400-500 ng RNA prepared from purified mitochondria, using the Ion Total RNA-Seq Kit v2 (ThermoFisher Scientific) according to the manufacturer's protocol. The libraries were sequenced on the Ion Torrent Proton™ NGS System according to the manufacturer's instructions. Raw sequencing data were processed using the Torrent Suite™ Software (Life Technologies). Barcode removal and quality trimming were performed in

Torrent Suite™ using default parameters (30% QC threshold, reads <25 nt rejected). The processed reads were exported as FASTQ files.

The complete mtDNA sequence of *C. albicans* strain SC5314 (GenBank:AF285261.1), with one of two identical copies of the inverted repeat region removed and additional annotations (Kolondra et al., 2015) was used as a reference. Reads were mapped to the reference sequence using BWA-mem (Li and Durbin, 2009), and SAMtools (Li et al., 2009) was used to manipulate the resulting alignments and to calculate coverage depth for each position in the reference sequence. Coverage graphs were obtained by visualizing BWA files obtained using bamCompare from the deepTools2 package (Ramírez et al., 2016) in pyGenomeTracks (Ramírez et al., 2018). The vioplot package in R was used to visualise coverage depth distributions. Reads mapping to chosen features were quantified using featureCounts (Liao et al., 2014) from the Rsubread package (Liao et al., 2019). Counts were normalized using the TMM method in the edgeR package (Robinson et al., 2010) for expression comparison between wild-type and mutants.

RT-PCR analysis

5µg of DNase-treated RNA was reverse transcribed using Maxima™ Reverse Transcriptase (Thermo Scientific) with random hexamers. 2 µl of tenfold diluted RT product was amplified in 20 µl with 0.5 U of Phusion polymerase (Thermo Scientific) and 10 pmol specific primers and analyzed by agarose gel electrophoresis. The number of amplification cycles was experimentally adjusted to stop the reaction before reaching the plateau phase. Sequences of all primers used in RT-PCR are listed in the Supplemental Table S2.

In silico analysis

Searching for Pet127 orthologs was performed using BLAST algorithm (Altschul et al., 1990) against amino acid (blastp) and translated nucleotide (tblastn) sequence databases at NCBI (NCBI, 2018), with the *S. cerevisiae* Pet127 protein sequence (Acc# NP_014660.1) as query. Identified sequences were then used to search (blastp) the *S. cerevisiae* protein sequences, and only those that gave ScPet127 as the top hit were retained (reciprocal best hits criterion for orthology). Additionally, sequences that had long stretches of 100% sequence identity with Pet127 proteins from taxa classified in different kingdoms were ruled out as contamination – this eliminated two hits in plant (tree) sequences with stretches of 100% identity with Pet127 proteins from xylophagous fungi, and one avian sequence with complete 100% identity with *S. cerevisiae*. Global alignment amino acid

sequence identity and similarity was calculated using the Needleman-Wunsch algorithm in STRETCHER from the EMBOSS suite (Rice et al., 2000). NommPred (Kume et al., 2018) was used to predict mitochondrial localization.

Amino acid sequences were aligned using muscle (Edgar, 2004) and analyzed in SeaView v. 5 (Gouy et al., 2021). For phylogenetic analysis the alignments were trimmed to remove noninformative regions in BMGE (Criscuolo and Gribaldo, 2010) using the BLOSUM30 similarity matrix. Phylogenetic tree was inferred using IQ-TREE (Nguyen et al., 2015) under the LG+F+R3 model. Branch support was assessed using the ultrafast bootstrap approximation (Hoang et al., 2018) with 1000 repeats and the nearest neighbor interchange (NNI) search.

Protein expression and purification

Codon optimized gene for *Talaromyces marneffe* Pet127 (TmPet127) was synthesized (BioBasic Inc.) and cloned into the expression vector pET28a-SUMO containing a His₆-SUMO (Small ubiquitin-related modifier) tag before the gene sequence. Mutations in the plasmid were introduced using QuikChange Site-Directed Mutagenesis kit (Agilent). Full length wild-type and point mutants were expressed in *E. coli* BL21 (DE3) Star strain. The bacteria cultures were grown in Luria broth (LB) medium at 37°C, induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside at OD₆₀₀ = 0.6 and grown overnight at 18°C. The cells were harvested by centrifugation and pellets were washed with phosphate-buffered saline (PBS) prior to purification. All proteins (wild type and mutants) were purified using the same protocol. The pellet from bacteria that expressed the desired protein, was lysed by sonication in buffer containing 50 mM Tris (pH 7.5), 500 mM NaCl, 10 mM imidazole, 5% (v/v) glycerol, and 5 mM 2-mercaptoethanol (buffer A).

The lysate was clarified by centrifugation at 186 000 × g at 4°C for 40 min and the supernatant was loaded onto a His-Trap HP column (GE Healthcare) equilibrated buffer A. Protein was eluted using a step gradient of imidazole (10 – 300 mM in buffer A). The fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions that contained Tm-Pet127 (buffer A containing 300 mM imidazole) were dialyzed in buffer B containing 20 mM Tris (pH 7.0), 100 mM NaCl, 5% (v/v) glycerol and 1 mM dithiothreitol (DTT). SUMO protease has been added to the protein solution in order to remove the His₆-SUMO tag. To remove nucleic acid contamination the protein was then loaded onto a HiTrap Heparin HP column (GE Healthcare) and eluted using a step gradient of NaCl (100mM – 1 M in buffer B). After confirming by SDS-PAGE the presence of Tm-Pet127 in

fractions eluted with 1 M NaCl in buffer B, the fractions were pooled together and concentrated to 0.5 ml. Concentrated fractions were loaded onto Superose 6 Increase size exclusion column, equilibrated previously with a buffer C that contained 20 mM Tris (pH 7.0), 1 M NaCl, 5% (v/v) glycerol, 1 mM MgCl₂ and 1 mM DTT. Fractions that contained Tm-Pet127 protein were concentrated using a 50k MWCO Amicon Ultra Centrifugal Filter Device (Millipore).

Exoribonuclease activity assays

6-carboxyfluorescein (FAM) labelled RNA substrates were synthesized by Future Synthesis. The sequences of the oligonucleotides are presented in the Supplemental Table S2. Exoribonuclease assays were performed with 90 nM protein concentration and 30 nM RNA in a reaction buffer that contained 20 mM Tris (pH 8.5), 50 mM KCl, 2 mM MgCl₂ and 0.5 mM DTT. The reaction was stopped at selected time-points by the addition of an equal volume of 95% formamide with Orange G dye and boiling the sample for 5 min at 95°C. Reaction products were resolved by 20% TBE-Urea PAGE and visualized with Typhoon Trio Imager (GE Healthcare).

ACCESSION NUMBERS

RNA-seq reads were deposited at NCBI Sequence Read Archive (SRA) repository under Accession number SRP346086 (<https://www.ncbi.nlm.nih.gov/sra/?term=SRP346086>).

SUPPLEMENTARY DATA

Supplementary Data are available.

ACKNOWLEDGEMENT

This work was supported by the National Science Centre of Poland [SYMFOPIA 00463, UMO-2015/19/B/NZ2/00201]. We thank Dr Anna Karnkowska for her useful comments on Protist taxonomy. We would also like to thank Dr Helena Kossowska for her advice and support in setting up and performing the RNA-seq experiments.

AUTHOR CONTRIBUTIONS

KŁD designed and supervised the genetic analysis, Northern blots, and RT-PCR in *C. albicans* performed by KŁD, HS, and JMW. MR, MG, and MN designed and performed the experiments

involving purification of the *T. marneffei* protein and enzymatic assays. AK performed the preliminary RNA-seq experiments. JP, and PG designed, performed, and analyzed the RNA-seq experiments presented in this paper (with KLD contributing to the experimental stage of the RNA-seq analysis). PG supervised the project and the preparation of the manuscript, to which all the authors contributed.

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Aravind L, Makarova KS, Koonin EV. 2000. SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res* **28**: 3417–3432.
- Ban C, Yang W. 1998. Structural basis for MutH activation in *E.coli* mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J* **17**: 1526–1534.
- Belfort M, Weiner A. 1997. Another bridge between kingdoms: tRNA splicing in archaea and eukaryotes. *Cell* **89**: 1003–1006.
- Borowski LS, Dziembowski A, Hejnowicz MS, Stepień PP, Szczesny RJ. 2013. Human mitochondrial RNA decay mediated by PNPase-hSuv3 complex takes place in distinct foci. *Nucleic Acids Res* **41**: 1223–1240.
- Burger G, Forget L, Zhu Y, Gray MW, Lang BF. 2003. Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc Natl Acad Sci U S A* **100**: 892–897.
- Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The New Tree of Eukaryotes. *Trends Ecol Evol* **35**: 43–55.
- Chang JH, Jiao X, Chiba K, Oh C, Martin CE, Kiledjian M, Tong L. 2012. Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat Struct Mol Biol* **19**: 1011–1017.
- Chen W, Islas-Osuna MA, Dieckmann CL. 1999. Suppressor analysis of mutations in the 5'-untranslated region of COB mRNA identifies components of general pathways for mitochondrial mRNA processing and decay in *Saccharomyces cerevisiae*. *Genetics* **151**: 1315–1325.

- Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci U S A* **103**: 10352–10357.
- Corbi D, Amon A. 2021. Decreasing mitochondrial RNA polymerase activity reverses biased inheritance of hypersuppressive mtDNA. *PLoS Genet* **17**: e1009808.
- Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology* **10**: 210.
- Dhir A, Dhir S, Borowski LS, Jimenez L, Teitell M, Rötig A, Crow YJ, Rice GI, Duffy D, Tamby C et al. 2018. Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* **560**: 238–242.
- Dmochowska A, Golik P, Stepień PP. 1995. The novel nuclear gene *DSS-1* of *Saccharomyces cerevisiae* is necessary for mitochondrial biogenesis. *Curr Genet* **28**: 108–112.
- Doamekpor SK, Gozdek A, Kwasnik A, Kufel J, Tong L. 2020. A novel 5'-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes. *Nucleic Acids Res* **48**: 349–358.
- Dziembowski A, Piwowarski J, Hoser R, Minczuk M, Dmochowska A, Siep M, van der Spek H, Grivell L, Stepień PP. 2003. The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. *J Biol Chem* **278**: 1603–1611.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
- Fekete Z, Ellis TP, Schonauer MS, Dieckmann CL. 2008. Pet127 Governs a 5' → 3'-Exonuclease Important in Maturation of Apocytochrome b mRNA in *Saccharomyces cerevisiae*. *J Biol Chem* **283**: 3767–3772.
- Gauthier GM. 2015. Dimorphism in fungal pathogens of mammals, plants, and insects. *PLoS Pathog* **11**: e1004608.

- Golik P, Szczepanek T, Bartnik E, Stepień PP, Lazowska J. 1995. The *S. cerevisiae* nuclear gene *SUV3* encoding a putative RNA helicase is necessary for the stability of mitochondrial transcripts containing multiple introns. *Curr Genet* **28**: 217–224.
- Gouy M, Tannier E, Comte N, Parsons DP. 2021. Seaview Version 5: A Multiplatform Software for Multiple Sequence Alignment, Molecular Phylogenetic Analyses, and Tree Reconciliation. *Methods Mol Biol* **2231**: 241–260.
- Gray MW, Lang BF, Burger G. 2004. Mitochondria of protists. *Annu Rev Genet* **38**: 477–524.
- Haffter P, Fox TD. 1992. Suppression of carboxy-terminal truncations of the yeast mitochondrial mRNA-specific translational activator *PET122* by mutations in two new genes, *MRP17* and *PET127*. *Mol Gen Genet* **235**: 64–73.
- Halpert M, Liveanu V, Glaser F, Schuster G. 2019. The Arabidopsis chloroplast RNase J displays both exo- and robust endonucleolytic activities. *Plant Mol Biol* **99**: 17–29.
- Haugen P, Simon DM, Bhattacharya D. 2005. The natural history of group I introns. *Trends Genet* **21**: 111–119.
- Hickman AB, Li Y, Mathew SV, May EW, Craig NL, Dyda F. 2000. Unexpected structural diversity in DNA recombination: the restriction endonuclease connection. *Mol Cell* **5**: 1025–1034.
- Hillen HS, Markov DA, Wojtas ID, Hofmann KB, Lidschreiber M, Cowan AT, Jones JL, Temiakov D, Cramer P, Anikin M. 2021. The pentatricopeptide repeat protein Rmd9 recognizes the dodecameric element in the 3'-UTRs of yeast mitochondrial mRNAs. *Proc Natl Acad Sci U S A* **118**: e2009329118.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol* **35**: 518–522.
- Hoffmann B, Nickel J, Speer F, Schafer B. 2008. The 3' ends of mature transcripts are generated by a processosome complex in fission yeast mitochondria. *J Mol Biol* **377**: 1024–1037.
- Holec S, Lange H, Kühn K, Alioua M, Börner T, Gagliardi D. 2006. Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Mol Cell Biol* **26**: 2869–2876.

- Houseley J, Tollervey D. 2009. The Many Pathways of RNA Degradation. *Cell* **136**: 763–776.
- Huchon D, Szitenberg A, Shefer S, Ilan M, Feldstein T. 2015. Mitochondrial group I and group II introns in the sponge orders Agelasida and Axinellida. *BMC Evolutionary Biology* **15**:
- Islas-Osuna MA, Ellis TP, Marnell LL, Mittelmeier TM, Dieckmann CL. 2002. Cbp1 Is Required for Translation of the Mitochondrial Cytochrome b mRNA of *Saccharomyces cerevisiae*. *J Biol Chem* **277**: 37987–37990.
- Jiao X, Chang JH, Kilic T, Tong L, Kiledjian M. 2013. A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol Cell* **50**: 104–115.
- Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL et al. 2007a. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**: 1484–1488.
- Kapranov P, Willingham AT, Gingeras TR. 2007b. Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* **8**: 413–423.
- Katz LA, Grant JR. 2015. Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. *Syst Biol* **64**: 406–415.
- Kolondra A, Labedzka-Dmoch K, Wenda JM, Drzewicka K, Golik P. 2015. The transcriptome of *Candida albicans* mitochondria and the evolution of organellar transcription units in yeasts. *BMC Genomics* **16**: 827.
- Kosinski J, Feder M, Bujnicki JM. 2005. The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. *BMC Bioinformatics* **6**: 172.
- Krause K. 2004. Analysis of transcription asymmetries along the tRNAE-COB operon: evidence for transcription attenuation and rapid RNA degradation between coding sequences. *Nucleic Acids Res* **32**: 6276–6283.
- Kume K, Amagasa T, Hashimoto T, Kitagawa H. 2018. NommPred: Prediction of Mitochondrial and Mitochondrion-Related Organelle Proteins of Nonmodel Organisms. *Evol Bioinform Online* **14**: 1176934318819835.

- Łabędzka-Dmoch K, Kolondra A, Karpińska MA, Dębek S, Grochowska J, Grochowski M, Piątkowski J, Hoang Diu Bui T, Golik P. 2021. Pervasive transcription of the mitochondrial genome in *Candida albicans* is revealed in mutants lacking the mtEXO RNase complex. *RNA Biol* 1–15.
- Lax G, Eglit Y, Eme L, Bertrand EM, Roger AJ, Simpson AGB. 2018. Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. *Nature* **564**: 410–414.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 GPDP. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
- Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**: e47.
- Lipinski KA, Kaniak-Golik A, Golik P. 2010. Maintenance and expression of the *S. cerevisiae* mitochondrial genome -- from genetics to evolution and systems biology. *Biochim Biophys Acta* **1797**: 1086–1098.
- Malecki M, Jedrzejczak R, Stepień PP, Golik P. 2007. *In vitro* reconstitution and characterization of the yeast mitochondrial degradosome complex unravels tight functional interdependence. *J Mol Biol* **372**: 23–36.
- Malecki M, Jedrzejczak R, Puchta O, Stepień PP, Golik P. 2008. *In vivo* and *in vitro* approaches for studying the yeast mitochondrial RNA degradosome complex. *Methods Enzymol* **447**: 463–488.
- Masters BS, Stohl LL, Clayton DA. 1987. Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell* **51**: 89–99.
- Mattiaccio JL, Read LK. 2008. Roles for TbDSS-1 in RNA surveillance and decay of maturation by-products from the 12S rRNA locus. *Nucleic Acids Res* **36**: 319–329.

- Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood AM, Haugen E, Bracken CP, Rackham O, Stamatoyannopoulos JA et al. 2011. The human mitochondrial transcriptome. *Cell* **146**: 645–658.
- Miki TS, Großhans H. 2013. The multifunctional RNase XRN2. *Biochem Soc Trans* **41**: 825–830.
- Mukhopadhyay J, Hausner G. 2021. Organellar Introns in Fungi, Algae, and Plants. *Cells* **10**: 2001.
- NCBI RC. 2018. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **46**: D8–D13.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.
- Nielsen H, Johansen SD. 2009. Group I introns: Moving in new directions. *RNA Biol* **6**: 375–383.
- Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**: 470–474.
- Orlowski J, Bujnicki JM. 2008. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. *Nucleic Acids Res* **36**: 3552–3569.
- Paps J, Medina-Chacón LA, Marshall W, Suga H, Ruiz-Trillo I. 2013. Molecular phylogeny of unikonts: new insights into the position of apusomonads and ancyromonads and the internal relationships of opisthokonts. *Protist* **164**: 2–12.
- Pietras Z, Wojcik MA, Borowski LS, Szewczyk M, Kulinski TM, Cysewski D, Stepień PP, Dziembowski A, Szczesny RJ. 2018a. Controlling the mitochondrial antisense - role of the SUV3-PNPase complex and its co-factor GRSF1 in mitochondrial RNA surveillance. *Mol Cell Oncol* **5**: e1516452.
- Pietras Z, Wojcik MA, Borowski LS, Szewczyk M, Kulinski TM, Cysewski D, Stepień PP, Dziembowski A, Szczesny RJ. 2018b. Dedicated surveillance mechanism controls G-quadruplex forming non-coding RNAs in human mitochondria. *Nat Commun* **9**: 2558.
- Poole TL, Stevens A. 1995. Comparison of features of the RNase activity of 5'-exonuclease-1 and 5'-exonuclease-2 of *Saccharomyces cerevisiae*. *Nucleic Acids Symp Ser* 79–81.

- Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. 2018. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun* **9**: 189.
- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160–5.
- Razew M, Warkocki Z, Taube M, Kolondra A, Czarnocki-Cieciura M, Nowak E, Labedzka-Dmoch K, Kawinska A, Piatkowski J, Golik P et al. 2018. Structural analysis of mtEXO mitochondrial RNA degradosome reveals tight coupling of nuclease and helicase components. *Nat Commun* **9**: 97.
- Reuss O, Vik A, Kolter R, Morschhäuser J. 2004. The SAT1 flipper, an optimized tool for gene disruption in *Candida albicans*. *Gene* **341**: 119–127.
- Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276–277.
- Ringel R, Sologub M, Morozov YI, Litonin D, Cramer P, Temiakov D. 2011. Structure of human mitochondrial RNA polymerase. *Nature* **478**: 269–273.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Rogowska AT, Puchta O, Czarnecka AM, Kaniak A, Stepień PP, Golik P. 2006. Balance between transcription and RNA degradation is vital for *Saccharomyces cerevisiae* mitochondria: reduced transcription rescues the phenotype of deficient RNA degradation. *Mol Biol Cell* **17**: 1184–1193.
- Saint-Georges Y, Bonnefoy N, di Rago JP, Chiron S, Dujardin G. 2002. A pathogenic cytochrome b mutation reveals new interactions between subunits of the mitochondrial bc1 complex. *J Biol Chem* **277**: 49397–49402.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.
- Schmitt ME, Brown TA, Trumpower BL. 1990. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **18**: 3091–3092.

- Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**: 6097–6100.
- Shang J, Yang Y, Wu L, Zou M, Huang Y. 2018. The *S. pombe* mitochondrial transcriptome. *RNA* **24**: 1241–1254.
- Steczkiwicz K, Muszewska A, Knizewski L, Rychlewski L, Ginalski K. 2012. Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res* **40**: 7016–7045.
- Stevens A, Poole TL. 1995. 5'-exonuclease-2 of *Saccharomyces cerevisiae*. Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1. *J Biol Chem* **270**: 16063–16069.
- Szczesny RJ, Borowski LS, Brzezniak LK, Dmochowska A, Gewartowski K, Bartnik E, Stepień PP. 2010. Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance. *Nucleic Acids Res* **38**: 279–298.
- Szczesny RJ, Borowski LS, Malecki M, Wojcik MA, Stepień PP, Golik P. 2012. RNA degradation in yeast and human mitochondria. *Biochim Biophys Acta* **1819**: 1027–1034.
- Tedersoo L, Sánchez-Ramírez S, Kõljalg U, Bahram M, Döring M, Schigel D, May T, Ryberg M, Abarenkov K. 2018. High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal diversity* **90**: 135–159.
- Tudek A, Candelli T, Libri D. 2015. Non-coding transcription by RNA polymerase II in yeast: Hasard or nécessité. *Biochimie* **117**: 28–36.
- Turk EM, Das V, Seibert RD, Andrulis ED. 2013. The mitochondrial RNA landscape of *Saccharomyces cerevisiae*. *PLoS One* **8**: e78105.
- Vyas VK, Barrasa MI, Fink GR. 2015. A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families. *Sci Adv* **1**: e1500248.
- Wade JT, Grainger DC. 2014. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat Rev Microbiol* **12**: 647–653.
- Walther A, Wendland J. 2008. PCR-based gene targeting in *Candida albicans*. *Nat Protoc* **3**: 1414–1421.

- Wegierski T, Dmochowska A, Jabłonowska A, Dziembowski A, Bartnik E, Stepień PP. 1998. Yeast nuclear *PET127* gene can suppress deletions of the *SUV3* or *DSS1* genes: an indication of a functional interaction between 3' and 5' ends of mitochondrial mRNAs. *Acta Biochim Pol* **45**: 935–940.
- Wideman JG, Monier A, Rodríguez-Martínez R, Leonard G, Cook E, Poirier C, Maguire F, Milner DS, Irwin NAT, Moore K et al. 2020. Unexpected mitochondrial genome diversity revealed by targeted single-cell genomics of heterotrophic flagellated protists. *Nature Microbiology* **5**: 154–165.
- Wiesenberger G, Fox TD. 1997. Pet127p, a membrane-associated protein involved in stability and processing of *Saccharomyces cerevisiae* mitochondrial RNAs. *Mol Cell Biol* **17**: 2816–2824.
- Wiesenberger G, Speer F, Haller G, Bonnefoy N, Schleiffer A, Schafer B. 2007. RNA Degradation in Fission Yeast Mitochondria Is Stimulated by a Member of a New Family of Proteins that Are Conserved in Lower Eukaryotes. *J Mol Biol* **367**: 681–691.
- Wilson RB, Davis D, Mitchell AP. 1999. Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J Bacteriol* **181**: 1868–1874.
- Xiang S, Cooper-Morgan A, Jiao X, Kiledjian M, Manley JL, Tong L. 2009. Structure and function of the 5'→3' exoribonuclease Rat1 and its activating partner Rai1. *Nature* **458**: 784–788.
- Yang X, Chang HR, Yin YW. 2015. Yeast Mitochondrial Transcription Factor Mtf1 Determines the Precision of Promoter-Directed Initiation of RNA Polymerase Rpo41. *PLoS One* **10**: e0136879.
- Zardoya R. 2020. Recent advances in understanding mitochondrial genome diversity. *F1000Res* **9**: F1000 Faculty Rev–270.

FIGURES

Figure 1. Pet127 is a conserved eukaryotic protein, lost in multiple independent lineages.

(A) Presence of Pet127 in different eukaryotic supergroups. Names in bold with an asterisk indicate that at least one member of the group carries a Pet127 ortholog. Supergroups and putative tree topology are drawn following recent phylogenomic studies (Lax et al., 2018; Burki et al., 2020), dashed lines indicate uncertain monophyly of a group. Excavata are no longer considered a monophyletic supergroup. (B) Presence of Pet127 orthologs in Opisthokonta and related taxa. Tree topology follows the conclusions of phylogenomic analyses of (Paps et al., 2013), fungal subkingdoms

are according to (Tedersoo et al., 2018). Ascomycete species, like *S. cerevisiae*, *S. pombe*, *C. albicans*, and *T. marneffeii*, belong to Dikarya. (C). Alignment of Pet127p amino acid sequences from selected eukaryotic lineages (the full list of taxa and accession numbers can be found in Supplemental Table S1). Detailed alignment is shown for the most conserved core nuclease domain, its location in the sequence of *Talaromyces marneffeii* Pet127 protein is marked on the schematic (grey box). Conservation is visualized as the sequence logo (Schneider and Stephens, 1990) below the alignment, with the total height of the position indicating the degree of conservation measured as the information content on a scale of 0 to 4.3 bits. Arrows indicate the key amino acids in the conserved nuclease core (Kosinski et al., 2005; Steczkiewicz et al., 2012): E340, and the residues forming the variant of the PD-(D/E)XK motif (here QD-DLK): Q386, D388, D401, L402, K403. Residues mutated in the described experiments are marked with asterisks. Numbering follows that of the *T. marneffeii* sequence.

Figure 2. The homozygous $\Delta Capet127$ strain and functional mutants are viable and do not show respiratory-deficiency. Growth on solid agar media containing glucose (YPD) or glycerol (YPG) as the carbon source. A series of 10x dilutions from an overnight YPD starter cultures of wild type (BWP17), heterozygous *CaPET127*/ $\Delta Capet127$ ($\Delta pet127HZ$), homozygous $\Delta Capet127$ ($\Delta Capet127$) and two independent point mutant *Capet127*_{D375A}/*Capet127*_{D375A} (D375A) strains were spotted on plates and incubated for 48h at 30°C or 37°C.

Figure 3. Changes in steady-state RNA levels in $\Delta Capet127$ and *Capet127*_{D375A} mutant strains. Northern blot analysis of mitochondrial mRNA and rRNA transcripts from wild-type (BWP17), homozygous $\Delta Capet127$ ($\Delta Capet127$) mutant and *Capet127*_{D375A} (D375A) catalytic mutants (two independent strains) **A.** mRNAs encoding subunits of Complex I. **B.** mRNAs encoding subunits of Complex III (COB) and Complex IV (COX). Arrows indicate the splicing intermediates of COB. **C.** mRNAs encoding subunits of the ATP synthase (Complex V). **D.** rRNAs of the small (RNS) and large (RNL) subunits of the mitoribosome. RNAs were prepared from purified mitochondria and separated by agarose/formaldehyde gel electrophoresis in denaturing conditions. Methylene blue staining of the small subunit mitochondrial rRNA in the blot is shown below each autoradiogram as a loading control. In panels A-D blot series [NAD3, RNS]; [NAD4, COB]; [RNL, COX1]; [COX2, NAD5] were prepared by stripping and re-hybridizing the same membrane, hence the same loading controls. **E.** Schematic map

of *C. albicans* mtDNA (without the second identical repeat region) showing the location of genes in primary transcription units (TU), according to (Kolondra et al., 2015).

Figure 4. Increased intron accumulation in $\Delta Capet127$ and $Capet127_{D375A}$ strains. **A.** The distribution of sites in the mtDNA reference sequence with varying coverage in the wild-type (BWP17), homozygous $\Delta Capet127$ mutant and $Capet127_{D375A}$ catalytic mutant. Width of the plot corresponds to the frequency of sites covered by the number of RNA-seq reads shown on the X axis. White circles and dark bars correspond to the median and the interquartile range (IQR), respectively. Coverage depth was calculated using the -depth option of SAMtools (Li et al., 2009) and visualised in R using the vioplot package. **B.** RNA-seq reads mapping to the intronic sequences in the RNL, COB, and COX1 genes, expressed as % of all reads mapping to the gene sequence (exons + introns). **C.** Coverage by RNA-seq reads of the fragment of the *C. albicans* mtDNA reference sequence encompassing the RNL gene and the downstream tRNA-Ala (tA(UGC)mt) in the wild-type (BWP17), homozygous $\Delta Capet127$ mutant and $Capet127_{D375A}$ catalytic mutant. BWA files obtained using bamCompare (Ramírez et al., 2016) were visualized in pyGenomeTracks (Ramírez et al., 2018). The depth coverage axis was set at the maximum value of 5000 reads to better visualize low-coverage regions, truncating the highest values. **D.** Semiquantitative RT-PCR analysis of the introns and intron-exon junctions in the homozygous $\Delta Capet127$ mutant (two independent repeats) compared to the wild-type (BWP17) strain. Fragments internal to the intron, encompassing the intron-exon junctions, and across intron were amplified following reverse transcription (+RT) of DNase-treated mitochondrial RNA. Amplification of reactions with the reverse transcriptase omitted (-RT) were used to control for DNA contamination of the RNA samples, and genomic DNA (gDNA) was used as a positive control.

Figure 5. Changes in steady-state levels of mitochondrial transcripts in $\Delta Capet127$ deficient strains estimated by RNA-seq. RNA-seq reads mapping to each gene were counted using featureCounts (Liao et al., 2014) and normalized using the TMM method in the edgeR package (Robinson et al., 2010). The heatmap presents log2 fold change of expression in the $\Delta Capet127$ mutant ($\Delta Capet127$) and $Capet127_{D375A}$ (D375A) mutant compared to the wild-type strain.

Figure 6 Purified *T. marneffei* Pet127 is a processive 5'-to-3' exoribonuclease. RNA degradation activity of *Tm*-Pet127 wt and point mutants against: **A.** 5' fluorescein (FAM) labeled single stranded RNA substrate (ssRNA), **B.** 5' G-rich ssRNA compared to a random sequence, **C.**

hairpin RNA structure within the mRNA ATP6/8 transcript from *C. albicans*. Degradation products were analyzed by 20% denaturing TBE-urea PAGE.

A

B

C

D

E

A **B** **C** **D**

Downloaded from [journal.cshlp.org](https://www.journal.cshlp.org) on April 26, 2023 - Published by Cold Spring Harbor Laboratory Press

RNA

A PUBLICATION OF THE RNA SOCIETY

The Pet127 protein is a mitochondrial 5'-to-3' exoribonuclease from the PD-(D/E)XK superfamily involved in RNA maturation and intron degradation in yeasts

Karolina Labedzka-Dmoch, Michal Razew, Marta Gapinska, et al.

RNA published online February 23, 2022

Supplemental Material <http://rnajournal.cshlp.org/content/suppl/2022/02/23/rna.079083.121.DC1>

P<P Published online February 23, 2022 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Open Access Freely available online through the *RNA* Open Access option.

Creative Commons License This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>