Title

Investigating the Regulation of Mitochondrial RNA and a foray into Mitochondrial Viruses

Author

Begeman, Adam

Publication Date

2024

Peer reviewed|Thesis/dissertation

Investigating the Regulation of Mitochondrial RNA and a foray into Mitochondrial Viruses

By

Adam Begeman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Samantha C Lewis, Chair

Professor Kathleen Collins

Professor Nicholas Ingolia

Professor Britt A Glaunsinger

Fall 2024

Abstract

Investigating the Regulation of Mitochondrial RNA and a foray into Mitochondrial Viruses

by

Adam Begeman

Doctor of Philosophy in Molecular & Cellular Biology

University of California, Berkeley

Assistant Professor Samantha C Lewis, Chair

Mitochondria are essential organelles responsible for a wide array of cellular functions including fatty acid oxidation, immune signaling, and most importantly cellular respiration. They maintain many copies of their own unique genomes which encode for essential electron transport chain proteins, as well as ribosomal and transfer RNAs required for gene expression. Transcription of the mitochondrial genome initiates bidirectionally and produces two polycistronic transcripts that then must undergo processing and maturation in discrete structures adjacent to the mitochondrial DNA called RNA granules. Mitochondrial transcription is initiated heterogeneously and asynchronously, meaning mitochondrial gene expression is regulated by fine-tuned post transcriptional processing and degradation machinery encoded in the nuclear genome. Unlike nuclear gene expression, all steps of the mitochondrial central dogma happen within the confines of the mitochondrial matrix, making it unclear how mature mitochondrial RNAs and their translation machinery are spatially organized across dynamic mitochondrial networks that are constantly undergoing fission and fusion. In the first part of this work, we report that processed mitochondrial RNAs are consolidated into translational hubs distal to either mitochondrial DNA or RNA processing granules in human cells. We found that the highly conserved helicase SUV3 contributes to the distribution of processed RNA within mitochondrial networks, and that perturbations in this pathway lead to an accumulation of dsRNA and a reorganization of mature transcripts into translationally-repressed mesoscale bodies. This reorganization was found to be downstream of dsRNA accumulation and in part dependent on mitochondrial translation. This work reveals that, just as mitochondrial transcription is regulated at nucleoids and RNA processing is regulated at RNA granules, translation by mitoribosomes occurs within defined RNA domains that are dynamically remodeled for quality control.

Similarly to mitochondrial gene expression, the organization of nuclear gene expression machinery into cytoplasmic RNA granules serves to regulate both the where and when of cytoplasmic translation. These RNA and protein rich bodies contain all the necessary components needed for protein production and are remodeled during cellular stress into

translationally suppressed stress granules. While cytoplasmic RNA granules serve as important regulatory hubs, allowing for rapid and local responses to changing cellular needs, they also serve as an attractive site for viral replication due to the concentration of gene expression machinery. While work has been studying how viruses may co-opt the cytoplasmic granules for viral replication, relatively little is known about how viruses may adapt replication or infection strategies to exploit the gene expression systems within the mitochondria. Given our new understanding that mitochondria also organize their gene expression systems into translation hubs that are remodeled during stress, we wondered if there was any evidence of viral species exploiting them for their replication. In the second part of this work, we searched thousands of publicly available RNA sequencing runs for novel viral species and report the discovery 763 new viral sequences belonging to the family *Mitoviridae*, a family of (+)ssRNA viruses that have previously been suggested to interact with mitochondrial gene expression machinery. The identified sequences fill in existing gaps in known mitovirus diversity, and allowed us to further expand the virus family, including previously uncharacterized clades and classes. Using this expanded diversity, we were able to annotate new mitovirus specific protein motifs, and identify hallmarks of mitochondrial translation such as mitochondrial specific codons and codon usage. This work expands the known diversity of mitochondrial viruses, and provides strong evidence for their infection, and use of mitochondria for viral replication.

Together, this work provides novel insights into the regulation of mitochondrial gene expression and how it is reorganized in response to transcriptional stress. It then investigates a family of (+)ssRNA viruses termed *Mitoviridae,* and provides evidence of their direct co-opting of mitochondrial gene expression machinery for their replication and life cycle. This work provides the foundation to not only explore the functional consequences of the spatial regulation of mitochondrial gene expression, but also to provide evidence of a viral species that may co-opt these translation hubs for their own replication

Table of Contents

To my partner Annie, the best part of my PhD

Acknowledgments

As I complete my graduate studies and thesis I would like to express my deepest appreciation to those who have helped me achieve this milestone. The collective support and guidance I received along the way has been instrumental to my success.

First, I would like to thank my family. To my parents, from a young age you have always supported my intellectual pursuits; always believing I could do anything I put my mind to. Your steadfast encouragement and belief laid the foundation for my success. Thank you for being accommodating and understanding of my not so conventional work hours at times. You have remained an integral part of my life despite the distance. I will forever cherish the memories made on your visits to Berkeley. To Grace, thank you for always being a phone call away. Your continuous love and support has meant the world to me. Thank you for always believing in me, especially at times when I was not sure. I continue to aspire to be the person you see me as. To Mimi and Papa, thank you for always supporting my ambitions and loving me unconditionally.

To my amazing partner and fiancée Annie, none of this would have been possible without you. I have struggled writing this section in particular because nothing I put into words seems to capture just how unbelievably grateful I am for your unending love and support. Thank you for being the person I wanted to call with exciting experimental results, and thank you even more for being the person I knew I could lean on when everything seemed to fail. Thank you for your undying patience and endless understanding. But most of all, thank you for taking this journey with me. You were the best part of these four and a half years and I cannot wait for the lifetime ahead of us. I know it has not always been the easiest, but I am confident in the future we are building for ourselves. To my future mother-in-law, Kerri Maloney, thank you for always being willing to host our cats at your "pet hotel" and for the countless loads of laundry you helped with throughout the years. The support you have provided for both me and our family during this time is invaluable.

To my first friend at UC Berkeley, Rachel Jansen, thank you for being there from the very beginning and every day since. Our immediate connection during interviews became the foundation of one of my most cherished friendships. Thank you for the sage advice you've shared over lunches and coffees, and for your continued belief in me. You will always be my first call when I cannot see a path forward, or when I need a good dose of encouragement. Our lunch dates, pickleball matches, wine nights, and cocktail excursions have been the highlight of my time here. If that wasn't enough, you also introduced me to your amazing fiancé, Dillon McCutcheon. Dillon, thank you for accepting me with open arms and for being a much needed reprieve during times of stress. I am forever grateful to have found a friend with whom hanging out is always a good time regardless of what we are doing. Thank you for being the reason I went to the gym on the days when I wouldn't have otherwise and for your dedication to being an

exceptional dungeon master. Finding someone who I can joke with about anything while also being able to talk through anything with is incredibly special and something I will never take for granted.

To my best friend in the world, Colin Phipps, thank you for always being there. I know that no matter what is happening, no matter where we are, I can always count on you; this has meant the world to me and my success. You continue to be my best friend regardless of where we are in the world and I am forever grateful for your continued friendship. I am unbelievably lucky to have you in my life and am endlessly appreciative of our friendship. Whether it is over the phone, on Discord, or in person, we always seem to lift one another up. At this point, you are more family than friend; a sentiment shared by the rest of the Begeman clan.

To one of my oldest friends Will Holtzmann, thank you for your continued friendship. From yearbook in middle school, to high school AP classes, and then to undergrad at DU, you have been a constant and cherished companion. We have been on parallel paths most of our lives, and although you took the easy way out by pursuing a PhD in Physics, I continue to value your support and understanding nonetheless. You have been an amazing outlet for me throughout my PhD. Thank you for being someone who is both always willing to talk through an experimental frustration and someone who will explore dungeons online with me when I have needed a mental break from work.

To my very first scientific mentor, Scott Horowitz, thank you for taking the chance on me. You gave me the structure and support that has made me the scientist I am today. I will always be in your debt for providing me the opportunity to succeed, and for trusting me, a young undergraduate student who knew nothing, with your research. My admissions to and success at UC Berkeley would not have been possible without your mentorship, belief, and support. You gave me the foundation and structure that has made me the scientist I am today. You fostered my drive, creativity, and passion which has served me so well as I have progressed in my science career. Thank you for everything.

To John Smolka, you have been an invaluable mentor and friend. Thank you from the bottom of my heart for the countless hours and vast amount of energy you dedicated to my training. You took me under your wing when I joined the Lewis Lab and, through what I can only imagine was a painstaking effort, helped turn me into the scientist I am today. Thank you for always making time to talk through odd experimental results, offer guidance, and your willingness to explore and discuss quirky experimental designs. What started as mid-day coffee runs to keep us sharp in the lab soon turned into late night cocktail escapades as our friendship grew. You've become a true friend in every sense of the word and I look forward to all the years to come.

To my mentor, Sam Lewis, your guidance and mentorship through my graduate studies has been pivotal to my success. Thank you for always pushing me to be a better version of myself both scientifically and professionally. From the very beginning, you supported my ambition and provided me with the environment, framework, and guidance to thrive. I am incredibly proud of what we have accomplished together and honored to be the first student graduating from the lab; I cannot wait to see how the Lewis Lab evolves over the years to come. To my fellow Lewis Lab members, thank you for making the lab a place I always enjoyed coming to; the comradery and collaboration has been a gift that I will always cherish.

Thank you to my committee members, Kathy Collins, Nick Ingolia, and Britt Glaunsinger, for your invaluable feedback and scientific expertise. Kathy, you have been an amazing mentor to me since I first arrived on campus four and a half years ago. You continue to be someone I know I can turn to for sound and sage advice. Thank you for shepherding me through my rotations, qualifying exam, and thesis research alike. Nick, thank you for always finding the time to consider the nuances of my work and help me interrogate my experimental ideas and direction. Your deep scientific curiosity and passion are inspiring. Britt, you may not remember but you were the first person who interviewed me for this program. While many easily wrote that casual hour off, you sat with me and took a genuine interest in what I had done in undergrad and in what I wanted to accomplish during graduate school. Our conversation demonstrated that Berkeley was more than a prestigious name. Thank you for the kindness you displayed all those years ago, and for bringing that same kindness and attention to every thesis meeting. Thank you all for your contributions to my academic and scientific growth.

Chapter 1: Introduction

Mitochondria are endosymbiotic organelles best known for their role in ATP production for the cell. Beyond oxidative phosphorylation, mitochondria also participate in a variety of cellular functions including fatty acid oxidation, iron-sulfur cluster biogenesis, calcium buffering, amino acid and nucleic acid metabolism, immune signaling, and apoptotic signaling (*1*, *2*). These double membrane bound organelles maintain hundreds to thousands of copies of their own genome that encodes essential electron transport chain proteins which are translated using dedicated mitochondrial specific ribosomes (*3*). This means that mitochondria maintain a completely separate gene expression system from the nuclear genome that coexists within the same cell. the coordination and regulation of mitochondrial genome expression is in part maintained by dynamic fission and fusion events between mitochondrial segments that distribute gene products (*4*).

Due to the role of mitochondria in cellular homeostasis and metabolism, mitochondrial dysfunction is the etiology of many human diseases. Misregulation of mitochondrial metabolism has been linked to a wide range of cancers and metabolic shifts linked to tumor progression and metastasis (*5*, *6*). Damage and decreased function of mitochondria over time is also linked to degenerative disorders including Parkinson's disease, Charcot-Marie Tooth syndrome, and optic neuropathy (*7*). Mutations or perturbations in mitochondrial genome maintenance are also linked to rare genetic disorders (*8*). Central to its role as the energy producing organelle and signaling hub is the mitochondria's ability to maintain its own gene expression system, the regulation of which remains an active field of study.

The Mitochondrial Central Dogma

Human mitochondrial DNA (mtDNA) is a circular double stranded molecule of 16.6 kb that encodes 13 electron transport chain (ETC) proteins, 22 mitochondrial specific transfer RNAs (tRNAs), and 2 ribosomal RNAs (rRNAs) (*9*). The vast majority of the genes are located on one strand of the mitochondrial genome termed the "heavy" strand, with only one gene, ND6, and 6 of the tRNAs located on the other termed the "light" strand (*9*). Mitochondrial DNA is packaged into nucleoprotein complexes called nucleoids that usually contain one to two copies of the mitochondrial genome (*9–11*). A major structural protein of the nucleoid is TFAM, a high mobility group (HMG) box domain protein that binds to and slides along mtDNA nonspecifically and is thought to play a role in regulating the accessibility of mtDNA for the purposes of replication and transcription (*9*, *12–14*).

Mitochondrial DNA is replicated asynchronously among the nucleoid population by a suite of dedicated proteins that are encoded in the nuclear genome and imported to the matrix. While identifying all the components involved in mtDNA replication is ongoing, there is a set of

well characterized essential factors needed for mtDNA synthesis termed the "minimal" replisome. This includes the DNA polymerase gamma holoenzyme (POLG1/2), the helicase TWINKLE, and the mitochondrial single stranded DNA binding protein SSBP1 (*9*). The spatial and temporal regulation of mtDNA replication is still being elucidated. While it has been shown that mtDNA replication is linked with mitochondrial division sites marked by ER contacts, and there is some work suggesting that nucleoid availability mediated by TFAM occupancy may play a role, we still lack a complete understanding of how individual mtDNA molecules are selected for replication or what regulates the number of mtDNA copies (*9*, *13*, *15*).

Mitochondrial DNA is transcribed bidirectionally to generate two polycistronic transcripts originating from the light and heavy strand promoters located in the noncoding region of the genome called the D-loop (*3*). Transcription is initiated by mitochondrial transcription factor A (TFAM) and mitochondrial transcription factor B2 (TFB2M) recruiting POLRMT to the promoter regions (*3*, *16*). Once initiated, POLRMT requires transcription elongation factor (TEFM) for proper processivity (*3*, *16*). The polycistronic nature of the mtDNA transcripts necessitates processing and maturation of the RNA into its coding and functional components (*3*, *17*). This is mainly done by endonucleolytic excision of the tRNAs by the RNase P complex on the 5' end and RNaseZ at the 3' end (*3*, *17*). All the coding transcripts, except for ND6, are then polyadenylated by mtPAP (*3*). This processing, as well as tRNA modifications, amino acetylation, and mito ribosome assembly, are thought to occur within a ribonuclear protein compartment localized adjacent to the mtDNA nucleoid termed the RNA granule (*18*). Marked by the G-rich RNA sequence binding factor 1 (GRSF1) and four Fas-activated serine/threonine kinase (FASTK) proteins, RNA granules are macromolecular processing centers required for proper mitochondrial gene expression (*18*, *19*). The molecular events that precipitate the initiation of transcription have been reconstituted *in vitro* but much less is known about spatiotemporal regulation in living cells. However, the lifecycle of the processed RNAs and how the genic RNAs are shuttled from RNA granules to mitoribosomes remains obscure.

Mitochondrial translation occurs on dedicated mitochondrial ribosomes using tRNAs entirely contained within the mitochondrial matrix. Mitoribosomes are proteinaceous compared to cytosolic ribosomes in metazoans, with the two RNA components encoded by mtDNA (*20*, *21*). These mitoribosomes and the 22 mtDNA encoded tRNAs constitute a complete gene expression system with distinct frequencies of codon usage from that of cytosolic gene expression (*22*). Mammalian mitochondrial coding sequences lack 5' and 3' untranslated regions, thus the loading of mRNAs onto the mitoribosomes requires alternative mechanisms and protein factors (*20*, *21*). Recently it was shown that the leucine-rich penticopeptide rich domain containing protein (LRPPRC) along with its binding partner stem-loop interacting RNA-binding protein (SLIRP) are involved in loading coding RNAs onto mitoribosomes through binding interactions with the poly(A) tails and the ribosome protein subunit mS39 (*20*). However, despite

understanding this laundry list of proteins involved in translation initiation and elongation, we still lack clarity on where or when mitochondrial translation happens in cells.

Finally, the turnover of mitochondrial RNAs is mediated by the nuclease complex PNPase and the RNA helicase SUV3 (SUPV3L1) (*23*). These two enzymes are thought to work in conjunction with each other, oftentimes localized to RNA granules but also suggested to form "D-foci" (*18*, *23*). This complex is particularly important for suppressing the accumulation of double stranded RNA (dsRNA) which forms readily in the mitochondrial matrix due to the bidirectionality of mtDNA transcription creating complementary RNA sequences (*18*, *23*). Loss or mutation of SUV3 or PNPase results in decreased transcription, RNA processing defects, and an accumulation of dsRNA (*24–26*). In the case of PNPase knock down, this dsRNA has been shown to escape the mitochondrial matrix where it can then activate the innate immune response through the MDA5/MAVS axis (*25*). Inborn errors in SUV3 via truncation or loss of function mutation cause neurodegenerative disease in humans despite not resulting in dsRNA release or IFN activation directly, suggesting a potential mitochondrial gene expression specific axis to the disease (*25*, *27*). Because mitochondria maintain hundreds to thousands of copies of their genomes that transcribe asynchronously across the entire network, it is thought that mtRNA turnover is the driving force of gene regulation (*23*, *28*).

The Importance of Mitochondrial Network Dynamics

While canonically we may think of mitochondria as individual discrete organelles, in reality, they exist as a dynamic and interconnected network in which fission and fusion of the mitochondrial outer and inner membranes allow for efficient exchange of material (*4*). This constant mixing of the matrix components creates an even distribution of mtDNA and gene products across the network and allows for healthy mitochondrial areas to complement deficient or damaged ones (*4*, *29*). Fission of the mitochondrial membranes is mediated by dynamin-related protein 1 (DRP1) which oligomerizes around the mitochondrial tubule at sites of ER contact (*4*, *30*). This constriction event then allows for a GTP dependent "power stroke" resulting in complete fission of both the outer and inner mitochondrial membranes (*4*, *31*). The recruitment of DRP1 to sites of fission is mediated by outer mitochondrial membrane proteins mitochondrial fission factor (Mff), mitochondrial fission 1 (Fis1), and mitochondrial dynamics protein (MiD)-49 and 51 (*4*). Mitochondrial fusion of membranes is mediated by two separate protein complexes dedicated for both outer and inner membrane fusion. The outer membrane fusion involves the dynein-associated GTPases mitofusion 1/2 (MFN1/2), whereas the inner mitochondrial membrane fusion event is mainly mediated by optic atrophy 1 (OPA1) (*4*).

Because the mitochondrial network is an interconnected network, turnover of damaged mitochondrial segments can be challenging. This removal of segments of the mitochondrial network, or mitophagy, is classified as either PINK1-Parkin dependent or independent depending

on how the phagophore is recruited to the mitochondrial membrane (*32*). The PINK1 dependent pathway is the most well characterized pathway of mitochondrial turnover and involves the stabilization of PINK1 on the outer membrane in a damage dependent manner, which in turn recruits the E3 ubiquitin (Ub) ligase parkin whose Ub chains mark the mitochondrial object for degradation (*32*). Smaller mitochondrial derived vesicles (MDVs) can also play a role in mitochondrial maintenance but these are able to be directly degraded by the lysosome after budding off of the mitochondria (*32–34*). All forms of mitochondrial turnover generally involve fission events that remove damaged or depolarized mitochondria segments from the greater network (*32*, *35*). Decreased mitochondrial dynamics or mitophagy is linked to a plethora of neurodegenerative and aging related disorders, including Alzheimer's, and Parkinson's disease (*4*, *7*, *32*). While a lot of work has been done attempting to characterize the process of mitophagy in this disease context, it remains unclear how the maintenance or degradation of mitochondrial segments is targeted in such a dynamic and constantly mixing environment.

Mitochondrial networks asynchronously replicate, transcribe, and translate their protein products and thus exhibit some level of heterogeneity in gene expression across segments of the mitochondrial network (*36*, *37*). Membrane dynamics and content mixing play a homogenizing role in the mitochondrial network to mitigate the differences in gene products across regions of the network. However in cell types with protracted network structures, such as neurons, or in disease contexts with more diffuse mitochondrial genetic material, such as mtDNA depletion syndromes, this maintenance of homogeneity becomes difficult (*4*, *7*, *29*, *32*, *37*). Heterogeneity in mitochondrial ETC function is an attractive avenue for explaining selective turnover of damaged mitochondrial segments, but outside of some correlative work linking membrane potential to mitophagy factor recruitment (*35*), we lack a molecular mechanism for how this heterogeneity might arise.

The Problem of Mitochondrial Genetic Engineering

Mitochondria maintaining their own central dogma means that they are also susceptible to genetic disorders that target their DNA. Mutations or deletions in mitochondrial DNA can lead to impaired ETC function and are responsible for a range of diseases including MELAS syndrome, Leber hereditary optic neuropathy, and Kearns-Sayre syndrome (*38*). Because mitochondria contain many copies of their genomes, the proportion of the mutated mtDNA can be quite variable, and often mitochondrial networks will contain both wild type and damaged mtDNA, a situation called heteroplasmy (*38*).

Developing strategies to treat mitochondrial DNA mutation or delegation disorders has been limited by the challenges editing the mitochondrial genome. For the nuclear genome, this is easily achieved using powerful CRISPR based genetic engineering strategies to efficiently edit or insert DNA fragments at will (*39*). These CRISPR systems utilize small guide RNAs to direct the

nucleases to the desired loci (*39*). Delivery of the Cas proteins and gRNA components to the nucleus generally relies on either lipid based or viral infection systems (*39*). Unfortunately despite years of research, CRISPR based strategies remain ineffective for mammalian mitochondrial genome editing (*40*). Successful genetic manipulation of the mitochondrial genome has relied entirely on protein based nucleases and base editors which dramatically limit the breadth of what is possible (*40*). The first major barrier to bona fide mitochondrial genome editing is that there is currently no way to deliver exogenous nucleic acids into the mitochondrial matrix (*40*). Overcoming this nucleic acid delivery problem would not only be a breakthrough in studying mitochondrial biology through gene knock-out/knock-in studies but also provide an entirely new therapeutic modality for mitochondrial related diseases.

The (+)ssRNA Virus Family: *Mitoviridae*

Of the known eukaryotic RNA viruses, positive strand (+) ssRNA viruses are distinct in their ability to remodel host endomembranes into invaginated viral replication organelles (ROs) derived from endoplasmic reticulum (ER), Golgi, plasma membrane, double membrane bound vesicles, or the outer mitochondrial membrane (*41*). These viral ROs serve to concentrate viral RNA, proteins, and host factors to facilitate the formation of viral replication complexes (VRCs), as well as shield the viral genome from host antiviral sensing mechanisms (*41*). Beyond RO, some viruses have the ability to enter membrane-bound host organelles to directly access their biosynthetic potential, promote the formation of organelle-derived replication vesicles, or sequester organelle-localized proteins in the cytoplasm for their own benefit (*41–43*). This ability to infect and co-op host gene expression and replication machinery has made viruses attractive vehicles for genetic engineering.

Previous studies identified a Family of (+)ssRNA viruses that may co-opt the mitochondrial gene expression system of fungal, plant, or invertebrate hosts for their own replication, named *Mitoviridae* (*44–46*). While a handful of mitovirus species have been identified as fungal pathogens, these viruses remain severely understudied and the extent to which they interact with host mitochondria in a cellular context remains unclear (*44*). Their genomes are composed of a single RNA with a single open reading frame (ORF) encoding for an RNA dependent RNA polymerase (RdRp) (*44*). Like all (+)ssRNA viruses, they are thought to self-replicate via a dsRNA intermediate generated by their own RdRp (*44*). Mitoviruses were originally discovered and isolated from the pathogenic fungi Dutch elm disease fungus (Ophiostoma ulmi) and chestnut blight fungus (Cryphonectria parasitica), where mitovirus infection imposed hypovirulence (*45*, *46*). These original studies identified small dsRNA elements present within the mitochondrial fraction of cellular fractionation experiments that encoded an RdRp only when the ORF was translated using the fungal mitochondrial codon table (*45*, *46*). Subsequent follow up experiments using the type species Cryphonectria mitovirus 1 (CpMV-1) demonstrated that the mitovirus infection was maternally inherited in the fungal

progeny, consistent with mitochondrial biology, and capable of horizontal transfer within a fungal population (*47*). These studies also found that the mitovirus infection could be transferred to even vegetative incompatible fungi species via protoplast fusions, in which the mitovirus infection was seemingly transferred via mitochondrial fusion events (*47*, *48*). Some studies have also reported altered mitochondrial morphology in mitovirus infected fungal samples as visualized by electron microscopy (EM), but the consistency and phenotype of this observation is fairly heterogeneous (*49*).

Mitoviruses are reported to interact with the mitochondrial ribosomes within the mitochondrial matrix, making them attractive targets for a potential naturally occurring platform for mitochondrial RNA transfection. Understanding mitoviruses are able to infect the mitochondrial matrix and co-opt its gene expression systems would be a step forward in potentially developing bona fida mitochondrial gene editing at will.

Focus of this Work

This work aims to fill the gaps of knowledge in two distinct arenas of mitochondrial nucleic acid regulation. The first part investigates the spatial regulation of mature mitochondrial RNAs and how they are remodeled during stress. We demonstrate the functional significance of the distribution of processed mitochondrial RNAs and identify a new regulatory axis for mitochondrial gene expression during transcript processing stress. The second part of this work investigates the breath and host interactions of mitoviruses using high throughput metagenomic approaches. These findings greatly expand the known diversity of the family *Mitoviridae* and provide evidence of mitoviruses directly interacting with mitochondrial ribosomes for their life cycles.

References

1. N. Pfanner, B. Warscheid, N. Wiedemann, Mitochondrial proteins: from biogenesis to functional networks. *Nat. Rev. Mol. Cell Biol.* **20**, 267–284 (2019).
2. L. Koklesova, M. Samec, A. Liskova, K. Zhai, D. Büsselberg, F. A. Giordano, P. Kubatka, O. Golunitschaja, Mitochondrial impairments in aetiopathology of multifactorial diseases: common origin but individual outcomes in context of 3P medicine. *EPMA J.* **12**, 27–40 (2021).
3. O. Rackham, A. Filipovska, Organization and expression of the mammalian mitochondrial genome. *Nat. Rev. Genet.* **23**, 606–623 (2022).
4. W. Chen, H. Zhao, Y. Li, Mitochondrial dynamics in health and disease: mechanisms and potential targets. *Signal Transduct. Target. Ther.* **8**, 1–25 (2023).
5. S. Vyas, E. Zaganjor, M. C. Haigis, Mitochondria and Cancer. *Cell* **166**, 555–566 (2016).
6. D. C. Wallace, Mitochondria and cancer. *Nat. Rev. Cancer* **12**, 685–698 (2012).
7. S. Xu, X. Zhang, C. Liu, Q. Liu, H. Chai, Y. Luo, S. Li, Role of Mitochondria in Neurodegenerative Diseases: From an Epigenetic Perspective. *Front. Cell Dev. Biol.* **9** (2021).
8. P. F. Chinnery, "Primary Mitochondrial Disorders Overview" in *GeneReviews®*, M. P. Adam, J. Feldman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, A. Amemiya, Eds. (University of Washington, Seattle, Seattle (WA), 1993; http://www.ncbi.nlm.nih.gov/books/NBK1224/).
9. M. Falkenberg, Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays Biochem.* **62**, 287 (2018).
10. C. Kukat, C. A. Wurm, H. Spåhr, M. Falkenberg, N.-G. Larsson, S. Jakobs, Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc. Natl. Acad. Sci.* **108**, 13534–13539 (2011).
11. E. V. Kakudji, S. C. Lewis, Mitochondrial nucleoids. *Curr. Biol.* **34**, R1067–R1068 (2024).
12. N. A. Bonekamp, M. Jiang, E. Motori, R. G. Villegas, C. Koolmeister, I. Atanassov, A. Mesaros, C. B. Park, N.-G. Larsson, High levels of TFAM repress mammalian mitochondrial DNA transcription in vivo. *Life Sci. Alliance* **4** (2021).
13. C. Brüser, J. Keller-Findeisen, S. Jakobs, The TFAM-to-mtDNA ratio defines inner-cellular nucleoid populations with distinct activity levels. *Cell Rep.* **37**, 110000 (2021).
14. H. B. Ngo, G. A. Lovely, R. Phillips, D. C. Chan, Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.* **5**, 3077 (2014).
15. S. C. Lewis, L. F. Uchiyama, J. Nunnari, ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. *Science* **353** (2016).
16. H. S. Hillen, Y. I. Morozov, A. Sarfallah, D. Temiakov, P. Cramer, Structural Basis of Mitochondrial Transcription Initiation. *Cell* **171**, 1072-1081.e10 (2017).
17. M. Jedynak-Slyvka, A. Jabczynska, R. J. Szczesny, Human Mitochondrial RNA Processing and Modifications: Overview. *Int. J. Mol. Sci.* **22**, 7999 (2021).
18. V. J. Xavier, J.-C. Martinou, RNA Granules in the Mitochondria and Their Organization under Mitochondrial Stresses. *Int. J. Mol. Sci.* **22**, 9502 (2021).
19. T. Rey, S. Zaganelli, E. Cuillery, E. Vartholomaiou, M. Croisier, J.-C. Martinou, S. Manley, Mitochondrial RNA granules are fluid condensates positioned by membrane dynamics. *Nat. Cell Biol.* **22**, 1180–1186 (2020).
20. S. Aibara, V. Singh, A. Modelska, A. Amunts, Structural basis of mitochondrial translation. *eLife* **9**, e58362 (2020).
21. F. Wang, D. Zhang, D. Zhang, P. Li, Y. Gao, Mitochondrial Protein Translation: Emerging

Roles and Clinical Significance in Disease. *Front. Cell Dev. Biol.* **9** (2021).
22. W. Jia, P. G. Higgs, Codon Usage in Mitochondrial Genomes: Distinguishing Context-Dependent Mutation from Translational Selection. *Mol. Biol. Evol.* **25**, 339–351 (2008).
23. A. Barchiesi, C. Vascotto, Transcription, Processing, and Decay of Mitochondrial RNA in Health and Disease. *Int. J. Mol. Sci.* **20**, 2221 (2019).
24. X. Zhu, X. Xie, H. Das, B. G. Tan, Y. Shi, A. Al-Behadili, B. Peter, E. Motori, S. Valenzuela, V. Posse, C. M. Gustafsson, B. M. Hällberg, M. Falkenberg, Non-coding 7S RNA inhibits transcription via mitochondrial RNA polymerase dimerization. *Cell* **185**, 2309-2323.e24 (2022).
25. A. Dhir, S. Dhir, L. S. Borowski, L. Jimenez, M. Teitell, A. Rötig, Y. J. Crow, G. I. Rice, D. Duffy, C. Tamby, T. Nojima, A. Munnich, M. Schiff, C. R. de Almeida, J. Rehwinkel, A. Dziembowski, R. J. Szczesny, N. J. Proudfoot, Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* **560**, 238–242 (2018).
26. P. Clemente, A. Pajak, I. Laine, R. Wibom, A. Wedell, C. Freyer, A. Wredenberg, SUV3 helicase is required for correct processing of mitochondrial transcripts. *Nucleic Acids Res.* **43**, 7398–7413 (2015).
27. S. L. van Esveld, R. J. Rodenburg, F. Al-Murshedi, E. Al-Ajmi, S. Al-Zuhaibi, M. A. Huynen, J. N. Spelbrink, Mitochondrial RNA processing defect caused by a mutation in two siblings with a novel neurodegenerative syndrome. *J. Inherit. Metab. Dis.* **45**, 292–307 (2022).
28. E. McShane, M. Couvillion, R. Ietswaart, G. Prakash, B. M. Smalec, I. Soto, A. R. Baxter-Koenigs, K. Choquet, L. S. Churchman, A kinetic dichotomy between mitochondrial and nuclear gene expression processes. *Mol. Cell* **84**, 1541-1555.e11 (2024).
29. K. Nakada, K. Inoue, T. Ono, K. Isobe, A. Ogura, Y.-I. Goto, I. Nonaka, J.-I. Hayashi, Inter-mitochondrial complementation: Mitochondria-specific system preventing mice from expression of disease phenotypes by mutant mtDNA. *Nat. Med.* **7**, 934–940 (2001).
30. J. R. Friedman, L. L. Lackner, M. West, J. R. DiBenedetto, J. Nunnari, G. K. Voeltz, ER Tubules Mark Sites of Mitochondrial Division. *Science* **334**, 358–362 (2011).
31. J. S. Chappie, J. A. Mears, S. Fang, M. Leonard, S. L. Schmid, R. A. Milligan, J. E. Hinshaw, F. Dyda, A pseudo-atomic model of the dynamin polymer identifies a hydrolysis-dependent powerstroke. *Cell* **147**, 209 (2011).
32. A. Picca, J. Faitg, J. Auwerx, L. Ferrucci, D. D'Amico, Mitophagy in human health, ageing and disease. *Nat. Metab.* **5**, 2047–2061 (2023).
33. V. Soubannier, P. Rippstein, B. A. Kaufman, E. A. Shoubridge, H. M. McBride, Reconstitution of Mitochondria Derived Vesicle Formation Demonstrates Selective Enrichment of Oxidized Cargo. *PLOS ONE* **7**, e52830 (2012).
34. R. F. Roberts, M. Y. Tang, E. A. Fon, T. M. Durcan, Defending the mitochondria: The pathways of mitophagy and mitochondrial-derived vesicles. *Int. J. Biochem. Cell Biol.* **79**, 427–436 (2016).
35. T. Kleele, T. Rey, J. Winter, S. Zaganelli, D. Mahecic, H. Perreten Lambert, F. P. Ruberto, M. Nemir, T. Wai, T. Pedrazzini, S. Manley, Distinct fission signatures predict mitochondrial degradation or biogenesis. *Nature* **593**, 435–439 (2021).
36. L. Chatre, M. Ricchetti, Large heterogeneity of mitochondrial DNA transcription and initiation of replication exposed by single-cell imaging. *J. Cell Sci.* **126**, 914–926 (2013).
37. L. Chen, M. Zhou, H. Li, D. Liu, P. Liao, Y. Zong, C. Zhang, W. Zou, J. Gao, Mitochondrial heterogeneity in diseases. *Signal Transduct. Target. Ther.* **8**, 1–27 (2023).
38. R. W. Taylor, D. M. Turnbull, Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.* **6**, 389–402 (2005).
39. J. Y. Wang, J. A. Doudna, CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
40. K. Lim, Mitochondrial genome editing: strategies, challenges, and applications. *BMB Rep.* **57**, 19 (2024).

41. G. Wolff, C. E. Melia, E. J. Snijder, M. Bárcena, Double-Membrane Vesicles as Platforms for Viral Replication. *Trends Microbiol.* **28**, 1022–1033 (2020).
42. D. J. Miller, M. D. Schwartz, P. Ahlquist, Flock House Virus RNA Replicates on Outer Mitochondrial Membranes in Drosophila Cells. *J. Virol.* **75**, 11664–11676 (2001).
43. I. H. Chen, Y. W. Huang, C. H. Tsai, The functional roles of the cis-acting elements in Bamboo mosaic virus RNA genome. *Front. Microbiol.* **8**, 645 (2017).
44. B. I. Hillman, G. Cai, "The Family Narnaviridae. Simplest of RNA Viruses." in *Advances in Virus Research* (Academic Press Inc., 2013; https://pubmed.ncbi.nlm.nih.gov/23498906/)vol. 86, pp. 149–176.
45. J. J. Polashock, B. I. Hillman, "A small mitochondrial double-stranded (ds) RNA element associated with a hypovirulent strain of the chestnut blight fungus and ancestrally related to yeast cytoplasmic T and W dsRNAs" (1994).
46. H. J. Rogers, K. W. Buck, C. M. Brasier, A mitochondrial target for double-stranded RNA in diseased isolates of the fungus that causes Dutch elm disease. *Nature* **329**, 558–560 (1987).
47. S. Shahi, A. Eusebio-Cope, H. Kondo, B. I. Hillman, N. Suzuki, Investigation of Host Range of and Host Defense against a Mitochondrially Replicating Mitovirus. *J. Virol.* **93** (2019).
48. J. J. Polashock, P. J. Bedker, B. I. Hillman, Movement of a small mitochondrial double-stranded RNA element of Cryphonectria parasitica: Ascospore inheritance and implications for mitochondrial recombination. *Mol. Gen. Genet.* **256**, 566–571 (1997).
49. K. Shafik, M. Umer, H. You, H. Aboushedida, Z. Wang, D. Ni, W. Xu, Characterization of a Novel Mitovirus Infecting Melanconiella theae Isolated From Tea Plants. *Front. Microbiol.* **12**, 3358 (2021).

Chapter 2: Investigating the Spatial Regulation of Mitochondrial Gene Expression

Portions of this chapter were preprinted as:

A spatial atlas of mitochondrial gene expression reveals dynamic translation hubs and remodeling in stress

Adam Begeman[1], John A. Smolka[1], Ahmad Shami[1], Tejashree Pradip Waingankar[1], Samantha C Lewis[1,2,3,4*]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA USA
[2]Innovative Genomics Institute, Berkeley, CA, USA
[3]Helen Wills Neuroscience Institute, Berkeley, CA USA
[4]Department of Nutritional Sciences and Toxicology, University of California, Berkeley, CA USA

*Corresponding author: samlewis@berkeley.edu

Citation:

Begeman A, Smolka JA, Shami A, Pradip Waingankar T, Lewis SC. (2024). A spatial atlas of mitochondrial gene expression reveals dynamic translation hubs and remodeling in stress. bioRxiv. https://doi.org/10.1101/2024.08.05.604215

Abstract

Mitochondrial genome expression is important for cellular bioenergetics. How mitochondrial RNA processing and translation are spatially organized across dynamic mitochondrial networks is not well understood. Here, we report that processed mitochondrial RNAs are consolidated with mitoribosome components into translation hubs distal to either nucleoids or processing granules in human cells. During stress, these hubs are remodeled into translationally repressed mesoscale bodies containing messenger, ribosomal, and double-stranded RNA. We show that the highly conserved helicase SUV3 contributes to the distribution of processed RNA within mitochondrial networks, and that stress bodies form downstream of proteostatic stress in cells lacking SUV3 unwinding activity. We propose that the spatial organization of nascent chain synthesis into discrete domains serves to throttle the flow of genetic information in stress to ensure mitochondrial quality control.

Introduction

Mitochondria are endosymbiotic organelles with essential roles in energy production, metabolic regulation, and the innate immune response (*1*). Mitochondria have their own genome encoding 13 respiratory chain complex proteins but rely on nuclear genes to control mitochondrial DNA (mtDNA) replication, transcription, and transcript processing (*2*). Each mammalian cell contains hundreds to thousands of mitochondrial genomes that are individually packaged into complexes termed mitochondrial nucleoids, the units of mtDNA inheritance and sites of transcription (*2*). Constitutive processing granules associate with nucleoids to render the polycistronic mitochondrial RNA into mature messenger, transfer, ribosomal, and non-coding RNAs (*3*). Mitochondrial DNA replication and transcription are thought to be coupled. While most mtDNAs are tightly packaged and likely inaccessible, the sole mitochondrial RNA polymerase, POLRMT, both primes replication and executes processive transcription on a permissible subset of the nucleoid population (*2*). Because each nucleoid contains only 1-2 copies of mtDNA, many nucleoid complexes are distributed throughout the mitochondrial syncytium, where they are asynchronously replicated and transcribed (*4*). In contrast to the nuclear genome, for which DNA replication, transcription, processing and translation occur in distinct compartments, all steps of mitochondrial gene expression co-occur within the innermost compartment of each mitochondrion, the matrix.

The spatial organization of mitochondrial gene expression across dynamic mitochondria that fuse, divide, and are motile in the cytoplasm is not well understood. The core component of the nucleoid complex is mtDNA binding protein TFAM, which has regulatory roles in replication initiation and in transcription via its control of mtDNA compaction, and thus accessibility (*5*, *6*). The RNA processing granules comprise several Fas-activated serine/threonine kinases (FASTK family), as well as G-rich sequence factor 1 protein (GRSF1) which interacts with mitochondrial RNaseP to stimulate primary transcript processing (*3*, *7*, *8*). The leucine-rich pentatricopeptide repeat protein LRPPRC forms a complex with RNA-binding protein SLIRP to bind and stabilize mitochondrial RNAs and is required for their loading to the mitoribosome (*9*). How the processed and polyadenylated transcripts navigate between the processing granules and mitoribosome loading has been obscure, though previous work established that some mitochondria translate more than others within the same cell, suggesting

that RNA localization within mitochondrial networks may be regulated to tune electron transport chain (ETC) biogenesis (*10–13*).

Content mixing, facilitated by cycles of membrane fusion and fission, permits the distribution of nascent ETC components, as well as nucleic acids throughout mitochondrial networks. This is particularly important in cells harboring deleterious mtDNA mutations (*14*). Due to its multicopy nature, mutant and wildtype mtDNA often co-exists within cells; complementation of individual organelles via distribution of wildtype gene products from sites of translation confers resilience to ETC dysfunction (*15*). Similarly, distribution of mtDNA gene products can buffer the effects of mtDNA depletion (*16*).

Defects in mtDNA expression cause mitochondrial dysfunction and are linked to cancer, aging and neurodegeneration (*1*). Mitochondrial RNA degradation specifically has emerged as a point of focus (*17*). In homeostasis, the ATP-dependent helicase SUV3 and the mitochondrial PNPase complex function in a linear pathway to unwind and degrade nucleic acids in the matrix, suppressing the persistence of double-stranded RNA that can form by complementarity of H- and L-strand transcripts (*17*, *18*). During stress, accumulation and egress of dsRNA to the cytosol triggers antiviral IFN-1 and pro-inflammatory pathways (*17*, *19*). While both SUV3 and the PNPase complex are required to suppress accumulation of dsRNA in human mitochondria, only PNPase defects lead to mitochondrial leakage and the activation of an inflammatory cascade (*19*, *20*). These findings indicated that while PNPase down-regulation precipitates pro-inflammatory type I IFN responses, it is SUV3 that acts as the upstream gatekeeper of dsRNA accumulation in mitochondria. Consistently, while the PNPase complex is not well conserved between model organisms, SUV3 is highly evolutionarily conserved in sequence and function amongst all eukaryotes, highlighting its essentiality (*21*).

Inborn errors in SUV3 cause neurodegenerative disease in humans despite no evidence of dsRNA release, suggesting an intrinsic mitochondrial stress response upstream of PNPase-dependent organelle permeabilization (*20*). Consistently, a defect in any of multiple steps of mitochondrial gene expression, from mtDNA synthesis, to transcription, to RNA processing, triggers a complex integrated stress response and suppression of mitoribosome translation by unclear mechanism(s) (*22*, *23*). These findings suggest foundational and functional links between the regulation of the mitochondrial central dogma within dynamic networks and quality control of the mitochondrial proteome. Here, we defined how mitochondrial gene expression is spatially organized at sub-organellar scales into regulatory hubs that are amenable to stress-induced remodeling to protect mitochondria during elevated proteostasis burden.

Results

Mitochondrial mRNA is excluded from nucleoids and processing granules

Mitochondrial DNA and polycistronic RNA are packaged into distinct nucleoprotein complexes, dedicated to mtDNA synthesis and transcript processing, respectively. However, it is unclear whether the translation of messenger RNAs occurs at defined sites in mitochondrial networks or is coordinated with nucleoids and/or the mitochondrial RNA processing granules (MRGs). Thus, we first sought to directly visualize the spatial distribution of mitochondrial messenger RNAs relative to mtDNA nucleoid complexes or MRGs (Figure 2.1A). We refined a

method of fluorescence in situ hybridization (mtRNA-FISH) using fluorophore-conjugated probe sets complementary to processed mRNA, ribosomal RNA, or tRNAs (Figure 2.2A). We validated mitochondrial RNA-FISH signals by confirming their RNaseA-sensitivity and dependence on active transcription by POLRMT (Figure 2.2B-C). We then simultaneously imaged mitochondria, mtDNA, mtRNA and RNA-binding protein GRSF1, a well characterized marker of the total MRG population in mammalian cells, at high spatial resolution, using Airyscan confocal microscopy in IMR90 non-immortalized human fibroblasts. We found that mitochondrial messenger RNAs encoding subunits of Complex I, Complex IV and the mitochondrial ATP Synthase were focally distributed, while in contrast diffuse, ubiquitous mt-tRNA and ribosomal RNA signals marked all mitochondria (Figure 2.1B-D**;** Figure 2.3). None of the RNA species we examined colocalized with dsDNA puncta, and linescan analysis indicated that RNA foci were independent of nucleoid positioning along mitochondrial tubules (Figure 2.1B-D, far right). As expected, GRSF1-positive MRGs were evenly distributed among mitochondria and intersected with nucleoids significantly more often than expected by random, as previously reported for other cell types (Figure 2.1E-G). Surprisingly, messenger RNA puncta were nearly twice as abundant as GRSF1 puncta along mitochondria (Figure 2.1F), and the majority did not colocalize with GRSF1 immunofluorescence signals (Figure 2.1G). These observations suggested an order to the distribution of processed mRNAs in mitochondrial networks, beyond their relationship with MRGs.

To distinguish mRNA foci from the polycistron we compared RNA-FISH labeling to pulse-labeling of nascent RNA using the click chemistry-compatible nucleoside analog 5-Ethynyluridine (EU) (Figure 2.4A-B) (*24*). GRSF1-positive MRGs colocalized with EU-labeled RNA puncta (Figure 2.4C-E), but not with processed RNA signals reported by FISH.

To rigorously quantify the spatial distributions of nascent and processed RNA across entire mitochondrial networks relative to nucleoids and MRGs, we developed and implemented an image analysis pipeline to systematically map and compare punctate fluorescent signals along filamentous mitochondria at the cellular scale (Figure 2.4F-G). We segmented mitochondria using a machine learning approach, skeletonized them, and extracted linescans to computationally identify peaks of fluorescence intensity in each channel along every mitochondrion. We then used this information from thousands of mitochondria in dozens of cells to generate average fluorescence intensities along a typical linescan segment, in essence, a virtual representation of nucleic acid organization in a typical mitochondrion. We validated our approach by demonstrating that dsDNA and the mtDNA nucleoid marker protein TFAM were highly correlated in these data, as were the EU and GRSF1 intensities (Figure 2.4H, left, right). We found that dsDNA and GRSF1 intensities were spatially linked to a lesser although still significant degree, consistent with our previous observation (Figure 2.4H, middle). With this tool in hand, we developed a spatial atlas of ribosomal RNA, tRNAs, and mRNA relative to mtDNA and MRGs. Strikingly, RNA-FISH signals not only failed to correlate with nucleoid or MRG markers, but were in fact significantly anti-correlated (Figure 2.1H). These findings indicate that not only are mRNA puncta distinct from nucleoids and MRGs, but they are also surprisingly excluded from those complexes.

We then used CRISPR Cas9 technology to ask whether RNA distribution was dependent on mtDNA copy number by knocking out POLG, the catalytic subunit of the sole mitochondrial

DNA polymerase, using three guide RNAs targeting its second exon (Figure 2.5A-B). While the abundance of total mitochondrial RNA scaled to mtDNA nucleoid content Figure 2.5C-D), ND4-FISH remained punctate, even within the subset of mitochondria that lacked nucleoids altogether (Figure 2.5D). Taken together, we conclude that mitochondria contain focal assemblies of processed RNA that are independent of either the mtDNA nucleoids where mtDNA replication and transcription occur, or MRGs, the nexus of nascent polycistron processing.

Mitochondrial mRNA marks punctate translation hubs

Mitochondrial translation rates are known to vary among individual mitochondria, though whether there may be a spatial relationship between mitochondrial mRNA distribution and newly synthesized translation products has been unclear (*11*, *13*). Based on our observations of RNA-FISH signals, we hypothesized that mitochondrial mRNA puncta define microscopically visible domains where translation occurs. Thus, we sought to examine the relative localization of mitoribosomes, mRNA, and nascent protein synthesis, taking ND4 as a representative mRNA (Figure 2.6A). We visualized mitoribosome assemblies via indirect immunofluorescence with an antibody against MRPL23, a component of the hydrophobic peptide exit tunnel of the large subunit (mt-LSU) (*25*). Strikingly, MRPL23 signals were punctate, and the majority colocalized with ND4-FISH (Figure 2.6B-C; Figure 2.7A-E).

We performed translation imaging by pulse-labeling cells with L-homopropargylglycine (HPG), a methionine analog that is readily recognized by the mitochondrial $tRNA^{Met}$ and incorporated into growing peptide chains and detected using copper click chemistry (*26*). After a 15-minute pulse, we observed that focal HPG signals colocalized with MRPL23, which was further supported by iterative linescan analysis (Figure 2.6D). HPG labeling was sensitive to the mitoribosome-specific peptidyl-transferase inhibitor chloramphenicol (CAP), as well as the selective POLRMT inhibitor IMT1B (Figure 2.6E; Figure 2.7F), validating that these observations reflect the output of steady-state gene expression (*27*). Consistent with our earlier observations, HPG-labeled translation domains were not spatially linked with mtDNA nucleoids marked by TFAM, or mtRNA processing granules marked by GRSF1 (Figure 2.7G-H). HPG signals were spatially linked to ND4 puncta, as well as local peaks of RNR2-FISH and tRNA-FISH intensity (Figure 2.6F). In contrast to the relationship between the processed RNAs and nucleoids, ND4-, RNR2- and tRNA-FISH signals correlated well with HPG via iterative linescan (Figure 2.6G). Moreover, the majority of ND4 puncta overlapped with HPG, to an extent significantly above random chance (Figure 2.6H). Taken together, we conclude that mitochondrial protein synthesis occurs in domains or hubs that are marked by mRNA puncta.

Translation hubs are dynamic and remodeled when mitochondrial fission is defective

We next used a series of pulse-chase experiments to ask whether the translation hubs reported by HPG were dynamic and responsive to network remodeling. We first labeled cells with HPG for 15 minutes; chased in HPG-free medium for 5, 15, 30, or 60 minutes; and then subsequently analyzed the density, size, and average fluorescence intensity of translation hubs after fixation (Figure 2.8A). Consistent with our previous observations, after 5 min of chase, HPG fluorescence intensity was initially restricted to punctate domains (Figure 2.8A, top row). Quantification of HPG signals during the series of extended chase experiments revealed that HPG signals progressively and dramatically decreased with lengthening chase time (Figure 2.8B-D), consistent with nascent chain degradation and/or the distribution away from those sites

throughout mitochondria. This observation was important, because it indicated that in fibroblasts, translation hubs serve as point sources for ETC component proteins.

We then sought to assess whether all mitochondria are capable of translation under our experimental conditions. We pulsed cells with HPG for 15, 30, or 60 minutes while holding the chase constant at 5 minutes and again measured features of HPG labeling among mitochondria (Figure 2.8E). We found that the number of HPG-labeled domains per mitochondrion increased in a manner directly proportional to pulse length (Figure 2.8F), though heterogeneity across the entire mitochondrial network persisted (Figure 2.7I). The size and fluorescence intensity of individual HPG hotspots increased as well (Figure 2.8G-H). These observations indicate that local translation shapes protein distribution within mitochondria, as peptide synthesis occurs at discrete sites and protein products then diffuse along the inner membrane, are distributed to other mitochondrial fragments via membrane dynamics, or are degraded.

Mitochondrial content mixing is in large part dependent on fission, fusion, and motility dynamics governed by dynamin family guanosine triphosphatases (GTPases) (*28*). DRP1, a cytosolic dynamin-related protein, forms helical scission assemblies around mitochondria mediated by its interactions with receptors in the outer membrane and by the close apposition of membrane contact sites with other organelles (*29*). To test the idea that compartmentalized gene expression in translation hubs shapes the flow of genetic information in mitochondrial networks at the cellular level, we next examined the role of mitochondrial fusion-fission cycles. We used transient overexpression of mCherry-tagged DRP1^{K38A}, a dominant mutation in the GTP-binding pocket that disrupts GTP hydrolysis, to decrease mitochondrial fission rate and assess the impact on the distribution of nascent translation products (Figure 2.8I) (*29*, *30*). Transient overexpression of mCherry-DRP1^{K38A} caused significant mitochondrial elongation as compared with control cells as predicted (Figure 2.7J), above and beyond baseline mitochondrial elongation caused by cycloheximide pre-treatment. We found that suppressing mitochondrial fission led to increased size and fluorescence intensity of both RNA- and HPG-enriched translation hubs in a subset of mitochondria, consistent with defective content mixing Figure 2.8J-M). Thus, when mitochondrial fission is perturbed, newly synthesized ETC proteins labeled by HPG fail to be distributed, reducing network homogenization. These data confirm that dynamic translation hubs and fusion/fission cycles shape the distribution of nascent peptides across mitochondrial networks.

Mitochondrial RNA is remodeled into translationally repressed liquid-like mesoscale bodies during stress

We hypothesized that the organization of the matrix into dynamic translation hubs could serve to facilitate a rapid and local translational response to perturbation. Previous work has shown that pathogenic variants that disrupt mitochondrial protein synthesis converge on activation of a complex integrated stress response (ISR), characterized in part by shutoff of mitoribosome translation and accumulation of aberrant double-stranded mitochondrial RNA (*22*, *23*, *31*, *32*). During ISR activation, release of mitochondrial dsRNA into the cytosol triggers a pro-inflammatory transcriptional response, because the nucleic acid is recognized as foreign (*17*). Similarly, outer membrane permeabilization elicited by mitochondrial poisons also permits nucleic acid egress, which contributes to an mtDNA-triggered innate immune response (*33–35*). Given that context, we sought a means by which we could test whether translation hubs may be

remodeled to facilitate translational inhibition during dsRNA accumulation - without triggering inflammatory cascades that might confound our imaging-based approach.

SUV3 is an essential mitochondrial ATP-dependent RNA helicase that unwinds both double-stranded RNA (dsRNA) and RNA:DNA hybrids; it is also important for the nucleolytic degradation of aberrant, proteotoxic dsRNAs by the mitochondrial PNPase complex (Figure 2.9A) (*17*, *36*). In human cells, SUV3 silencing causes dsRNA accumulation (*36*). Importantly, multiple studies in cells and in patients with inborn errors in SUV3 have shown that it functions upstream of the PnPase complex in a linear pathway, such that loss of SUV3 function alone is insufficient for dsRNA release to the cytosol or detectable activation of innate immunity (*19*, *20*). Thus, we used depletion and transient overexpression of wildtype and mutant SUV3 isoforms to examine a potential role for translation hubs in mediating mitochondrial stress responses to dsRNA accumulation.

Transient overexpression of $SUV3^{WT}$-HA in control IMR90 cells revealed its localization to a multitude of discrete puncta coincident with a subset of endogenous GRSF1 foci at MRGs (Figure 2.10A), consistent with a role in suppressing hybridization between complementary endogenous RNAs. We next depleted SUV3 from cells by CRISPR Cas9 technology using three guide RNAs targeting exon 1 and examined mitochondrial dsRNA levels over a 7-day time course using indirect immunofluorescence with the anti-dsRNA antibody J2. While dsRNA was undetectable in control cells, we found that cells depleted of SUV3 continuously accumulated dsRNA, which coalesced into distinct foci within a subset of malformed, swollen mitochondria (Figure 2.9B). We then examined ssRNA distribution, finding that RNR2-FISH signals were no longer distributed throughout mitochondrial networks, but collapsed into distended boli while rendering large areas of the mitochondrial network devoid of ribosomal RNA (Figure 2.10B, top). To verify that this phenotype was specifically due to the lack of SUV3 helicase activity, we reintroduced either $SUV3^{WT}$-HA or $SUV3^{G207V}$-HA, a dominant, catalytically dead mutant allele, into the cells (Figure 2.10B-C). Transient overexpression of $SUV3^{WT}$-HA, but not $SUV3^{G207V}$-HA, rescued RNR2 localization, indicating that loss of SUV3 function not only leads to dsRNA accumulation, but remodeling of ssRNA distribution as well.

We next asked whether the accumulated dsRNA may seed the formation of the larger ssRNA-enriched boli. We examined fixed cells depleted of SUV3 by indirect immunofluorescence to simultaneously visualize TOM20, dsRNA via J2, and RNR2-FISH (Figure 2.9C). Indeed, we observed that RNR2-FISH signals had coalesced into boli surrounding the dsRNA explaining the distended appearance of the mitochondrial membranes. Co-labeling of cells with ND4-FISH and immunodetection of dsDNA revealed the re-organization of mRNA into the enlarged boli as well (Figure 2.11). Despite this, nucleoids remained distributed suggesting that mitochondrial DNA and RNA may be positioned in mitochondria by distinct mechanisms.

Previous studies have posited that mitochondrial ribonucleo-protein complexes may exhibit properties of phase-separated condensates (*37–39*). Given our observations that processed RNAs are excluded from nucleoids, and the striking remodeling of RNA in cells depleted of SUV3, we next sought to examine the dynamics of RNA in live cells. We labeled RNA in live IMR90 cells with the vital dye SYTO RNASelect and co-stained with the vital dye

Mitotracker. Consistent with our observations in fixed cells, RNASelect labeled many discrete puncta in control cells, while it labeled prevalent mesoscale structures when SUV3 was limiting (Figure 2.9D**;** Figure 2.12A), that maintained their membrane potential as reported by both Mitotracker and tetramethylrhodamine ethyl ester (TMRE) staining (Figure 2.12B).

We next asked whether these mesoscale bodies exhibited properties of liquid-like membraneless RNA bodies, similar to Balbiani bodies or RNP granules induced by viral infection, by assessing RNA dynamics and propensity for content exchange by time-lapse microscopy and fluorescence recovery after photobleaching (FRAP) (*37*, *40*). RNA boli labeled by RNASelect-labeled exhibited dynamic, fluctuating morphologies over time, though they persisted as discrete domains within mitochondria, with infrequent fusion or fission (Figure 2.9D; Figure 2.12A, bottom). Indeed, we also observed that multiple RNASelect-labeled domains would often co-persist within the same mitochondrion. We employed FRAP to determine whether these structures exchanged contents, finding that RNASelect intensity recovered on average to 40% of the pre-bleach intensity after background correction over a time period of 5 minutes (Figure 2.9E-F). This was less and slower recovery than previously reported for MRGs, but much more recovery than seen for solid mitochondrial aggregates composed exclusively of protein previously described in yeasts (*37*, *41*). Additional biochemical studies are needed to determine whether these structures could be bona fide biocondensates. Given their dynamic nature, stress-specific context, and evidence of content exchange, we will refer to them here as "mitochondrial RNA stress bodies" (MSB).

To determine the relationship between MSB formation and mitochondrial translation, we next used RNA-FISH and HPG pulse-labeling to assess protein synthesis in cells with the MSB phenotype. Relative to control cells, SUV3 depletion led to a near total loss of detectable HPG incorporation, concurrent with MSB formation (Figure 2.9G). Indeed, the intensity of HPG labeling in SUV3-depleted cells was comparable to control cells incubated with chloramphenicol (CAP) (Figure 2.9H). These findings demonstrate critical functional dependencies between dsRNA accumulation, the spatial organization of RNA in the matrix, and steady state mitoribosome translation, by which MSB formation is linked to translational inhibition (Figure 2.9I).

Mesoscale body formation is linked to proteostasis

The reorganization of RNA into MSBs prompted us to investigate the fates of the pre-existing MRGs during that process. We visualized the localization of endogenous GRSF1 in IMR90 cells labeled with RNR2-FISH by indirect immunofluorescence during SUV3 depletion (Figure 2.13). Unlike control cells, we observed a significant increase in GRSF1 colocalization with processed RNA in MSBs (Figure 2.13). In addition, we noted that while MSBs were marked by GRSF1, separate small GRSF1 puncta remained distributed throughout mitochondrial networks that did not colocalize with RNR2-FISH signals, consistent with a continued role in binding to and processing polycistronic RNA at MRGs and similar to the distribution we noted for nucleoids. Thus, while MSB's contain dsRNA, ssRNA, and a typical MRG protein, these structures are spatially distinct from MRGs and remain so over time.

Finally, we considered whether MSB formation is a response to proteotoxicity downstream of defective mtRNA processing or e.g. dsRNA accumulation. We reasoned that, if

MSBs form in response to an RNA processing defect alone, suppression of mitoribosome translation would have little effect on their formation or dynamics, as the RNA would still be produced. In contrast, if quality control were regulated at the protein level, only cells that actually translate aberrant messages would trigger MSB formation. Thus, we asked whether the preemptive arrest of mitoribosomes by CAP could suppress MSB formation. In distinct experiments, we performed a CAP pulse-chase analysis in cells either before (Figure 2.14A) or after (Figure 2.14B) transfection with sgSUV3-Cas9 RNP complexes, and subsequently analyzed the size and fluorescence intensity of MSBs. Preemptive mitoribosome arrest via CAP dramatically reduced the size of FISH-labeled MSBs relative to sgSUV3 cells treated with DMSO only (Figure 2.14B). In contrast, turning off the mitoribosome after SUV3 depletion failed to suppress MSB size or fluorescence intensity. To further define the contribution of proteotoxicity to MSB formation, we asked whether dsRNA accumulation and MSB formation are separable. We incubated cells with CAP, induced SUV3 depletion, then fixed them and detected dsRNA and RNR2-FISH. Consistent with our previous observation, CAP treatment suppressed the MSB phenotype; moreover, the cells still accumulated dsRNA (Figure 2.14C). This experiment demonstrated that MSB formation is a response to a proteostatic stress, as it relies on active translation by mitoribosomes. Consistently, we found that, in the absence of CAP, in early stages of MSB formation the dsRNA is apparent before RNR2 remodeling into boli (Figure 2.14D). Taken together, these data suggest a model in which homeostatic mitochondrial translation hubs enriched in processed RNA are remodeled into MSB structures in coordination with translational inhibition to protect proteostasis (Figure 2.14E).

Discussion

Our data indicate that within human mitochondrial networks, processed RNAs are translated in ribonucleo-protein hubs distinct from mtDNA nucleoids or MRGs, and these hubs are remodeled in response to RNA pre-processing stress, concurrent with suppression of mitoribosome translation. We propose that the spatial organization of mtDNA nucleoids, MRGs, and processed messages at these sites provides a means of throttling the flow of genetic information to ensure quality control of the mitochondrial proteome. Such a process of remodeling and translational suppression may be particularly important for mitochondria, as once expressed, core components of the electron transport chain complexes are long-lived (*42*).

It will be important to determine the fundamental molecular mechanisms by which mitoribosome shutdown occurs. Our analyses highlight the importance of the sub-mitochondrial organization of gene expression in homeostasis, and how that organization is remodeled in stress via sequestration of mtRNA and proteins within context-specific mitochondrial RNA stress bodies. Liquid-like condensates have important functions in regulating gene expression in the cell nucleus and cytoplasm, where they provide a means to locally concentrate sets of proteins and RNAs for regulation in 4-dimensions (*40*, *43–45*). Both mitochondrial nucleoids and MRGs have been suggested to have biophysical properties of condensates; how those properties may contribute to mtDNA or RNA integrity remains to be discerned (*37*, *38*).

Compartmentalization of biochemical processes is a unifying principle in biology. In eukaryotes, the nuclear envelope provides a means to control the entry and egress of transcription machinery and products, which may be processed and are ultimately translated in

distinct and spatially segregated compartments. In the cytosol, ribonucleoprotein granules, such as P-bodies that form during animal development, and yeast stress granules, serve as hubs to organize post-transcriptional regulation of gene expression. These well-characterized bodies consolidate translationally repressed mRNAs to regulate where and when gene expression occurs, which is particularly important in stress. Within mitochondria, the mitochondrial genome, immature polycistronic transcripts, and processed RNAs in various stages share the milieu of the matrix compartment. In this capacity, remodeling of the matrix into the structures described here could function to modulate the accessibility of mitoribosome loading or accessory factors to mRNAs that would normally shuttle between these compartments, tuning production of ETC proteins to sub-cellular cues. This could be particularly important in highly polarized cells such as neurons, in which mitochondrial functions may need to be specialized for the cell soma, dendrite, and axonal compartments (*46*). Given that SUV3 loss-of-function causes neurodegenerative disease in humans, it will be important to investigate whether MSBs exhibit further features of biocondensates and to refine our understanding of the conditions under which they form, as constitutive MSBs may constitute a form of pathological inclusion (*20*, *47*, *48*). Beyond SUV3, LRPPRC and/or SLIRP are likely to mediate both translation hub activity and MSB formation, given their key roles in protecting mitochondrial RNAs from degradation as well as in mitoribosome loading.

Our findings connect the spatial organization of the steps of mitochondrial genome expression in the matrix to the kinetics of mitoribosome translation and overall mitochondrial network morphometrics. This connection has implications for understanding the cellular pathology and complex stress responses underlying metabolic dysregulation. Developing approaches that can suppress or modulate the flow of genetic information in mitochondria may hold promise in the treatment of rare human diseases caused by defects in mitochondrial gene expression, as well as in the context of ISR activation during cancer and neurodegeneration.

Materials and Methods

Plasmids

The mCherry-DRP1^{K38A} plasmid was generated via QuickChange mutagenesis from mCherry-Drp1, a gift from Gia Voeltz (Addgene #49152) (*30*). To generate SUV3-HA and SUV3G207V-HA, the human SUV3 cDNA was synthesized (Twist Bioscience) with or without the G207V amino acid substitution, including an HA tag appended to the C-terminus of the sequence, and cloned into the pTwist CMV-driven mammalian expression vector.

Mammalian cell growth, transfection, and vital dyes

Human IMR90 cells (ATCC #CCL-186) were grown in high-glucose Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% heat inactivated fetal bovine serum (FBS) and 1% penicillin/streptomycin at 37 degrees Celsius in a humidified 5% CO^2 chamber. Prior to imaging, cells were seeded onto glass-bottom 35 mm dishes (Mattek) and cultured for 24-48 hours. Transient plasmid transfections were performed using Lipofectamine 2000 according to the manufacturer protocol (Thermo Fisher), and imaged 24 hours later unless otherwise noted. For live cell imaging, 1 mL of conditioned cell media was removed and saved and then, 50 nM Mitotracker Deep Red (Thermo Fisher), 500 nM Tetramethylrhodamine, Ethyl Ester (TMRE)

(Thermo Fisher), or 5 uM SYTO RNAselect (Thermo Fisher) was added directly to the dish for 30 minutes and then replaced with the reserved conditioned media just prior to imaging.

Cell fixation, antibodies, and immunofluorescence

Cells were seeded as described above. Cells were then fixed in pre-warmed (37 degrees Celsius) 4% paraformaldehyde solution (PFA) in DPBS for 30 minutes at room temperature while protected from light. Glass-bottom dishes were then gently washed with room temperature DPBS and cells were permeabilized in 0.1% TritionX-100 diluted in DPBS for 20 minutes. Dishes were washed with TBST blocking buffer (TBS pH 7.5, 0.1% Tween-20) containing 1% bovine serum albumin (BSA), and primary antibodies were added at 1:1000 dilution in the same buffer and incubated for 1 hour at room temperature, or overnight at 4 degrees Celsius. Dishes were then rinsed with blocking buffer and a solution containing secondary antibodies at 1:2000 dilution was added for 1 hour at room temperature. Dishes were rinsed with the blocking buffer and imaged in DPBS at room temperature. We employed the following antibodies to detect endogenous proteins and nucleic acids: rabbit anti-TOM20 (Proteintech, 11802-1-AP), mouse anti-TOM20 (SantaCruz Biotechnology, SC17764), mouse anti-dsDNA (Abcam, ab27156), rabbit anti-GRSF1 (Sigma, HPA036984), rabbit anti-TFAM (Abcam, ab176558), rabbit anti-MRPL23 (Sigma, HPA050406), mouse anti-DRP1 (Abnova, H00010059), rabbit anti-HA (Invitrogen, 71-5500), mouse anti-J2 (Sigma, MABE1134), donkey anti-rabbit AlexaFluor Plus 488 highly cross-adsorbed conjugate (ThermoFisher, A32790), goat anti-mouse AlexaFluor 405 conjugate (ThermoFisher, A31553), donkey anti-mouse AlexaFluor Plus 405 highly cross-adsorbed conjugate (ThermoFisher, A48257), goat anti-rabbit AlexaFluor 405 conjugate (ThermoFisher, A31556), goat anti-mouse AlexaFluor Plus 647 highly cross-adsorbed conjugate (ThermoFisher, A32728).

Mitochondrial RNA fluorescence *in situ* hybridization (FISH)

Labeling of mitochondrial RNA via fluorescence *in situ* hybridization was performed using a modified version of the Stellaris RNA-FISH protocol (Biosearch Technologies). All probe sequences are available in Supplementary Table 1 at https://doi.org/10.1101/2024.08.05.604215. Cells were seeded onto plates as described above. When FISH was combined with copper click chemistry to detect nascent proteins, then HPG labeling and the click reaction were performed prior to the FISH protocol and the cells were fixed and immunolabeled as described above. Otherwise, the cells were fixed in prewarmed (37 degrees Celsius) 3% PFA, 1.5% glutaraldehyde in DPBS for 10 minutes at room temperature. The fixation reaction was then quenched with a solution containing a 1:10 dilution of 1 M glycine in DPBS and washed with DPBS. To quench auto-fluorescence associated with glutaraldehyde, dishes were then incubated with 10 mg/mL sodium borohydride diluted in DPBS for 5 minutes at room temperature. Cells were then washed with DPBS and permeabilized in 0.1% TritionX-100, 1 uM dithiothreitol (DTT), 0.1% SDS in DPBS for 20 minutes at room temperature. Following permeabilization, dishes were washed with DPBS and incubated in Stellaris wash buffer A, prepared according to manufacturer's protocol (Biosearch Technologies), for 5 minutes. Wash buffer A was then aspirated and 200 uL of RNA-FISH probe set containing Stellaris hybridization buffer, prepared according to the manufacturer's protocol, was then added directly to the center of the glass bottom dish. Dishes were then incubated in a hybridization chamber at 37 degrees Celsius for 4 hours. Following incubation, the hybridization buffer was removed by aspiration and dishes were incubated in Stellaris wash buffer A at 37 degrees Celsius for 30 minutes. Dishes were then

washed with Stellaris wash buffer B, immunolabeled, and imaged as described above. For IMT1B control 10 uM final concentration was used for 96 hours prior to labeling and for RNaseA 100 ug/mL was used at room temperature for 1 hour after fixation and permeabilization but prior to FISH labeling.

EU Click reaction for labeling of nascent RNA

Labeling of nascent mitochondrial RNA was performed using ethynyl-uridine (EU) incorporation and resulting click reaction. Cells were first seeded and cultured as described above. On the day of the experiment, 1 mL of conditioned media was removed and set aside, then 1 uL of 1 mM triptolide was then added directly to cells in the imaging dish at a final concentration of 1 uM and samples were incubated for 30 minutes at 37 degrees Celsius. Next 1 uL of 500 mM EU was added to the sample media to a final concentration of 500 uM, and incubated for 1 hour. EU- and triptolide-containing media was then replaced with the reserved1 mL of conditioned media and incubated for a final 5 minutes. Following fixation and permeabilization as described above, EU incorporation was detected via Click-iT Plus Edu Alexa Fluor 647 imaging kit according to manufacturer's instructions, and then immunolabeled and imaged as described above.

HPG Click reaction for labeling of nascent peptide synthesis

Metabolic labeling of proteinactive mitochondrial translation was performed using a modified version of the manufacturer's Click-IT Homopropargylglycine (HPG) protocol (Thermo Fisher). Cells were first seeded and cultured as described above. Cells were then washed with warm (37 C) PBS and then incubated in methionine-free DMEM for 10 minutes before the addition of drugs to suppress either cytosolic protein translation (50 ug/mL Cycloheximide) or mitochondrial protein translation (50 ug/mL Chloramphenicol). Cells were then pulsed with HPG (50 uM final concentration) added directly to the drug containing methionine-free DMEM and incubated for 5-60 minutes as indicated in the text. Cells were then chased with methionine free DMEM lacking HPG as described. Following fixation and permeabilization as described above, samples then underwent a click chemistry reaction to render HPG fluorescent via the covalent addition of AlexaFluor probes following the manufacturer's protocol (Thermo Fisher).

Live Cell Imaging and FRAP

For FRAP assays, cells were seeded on poly-d-lysine-coated glass bottom 35 mm imaging dishes and cluttered for 1 to 2 days in normal growth conditions and then labeled with vital dyes as described above. The 148 (0.268um^2) pixel bleaching ROIs were placed on representative RNAselect bolus in sgSUV3 condition and 5 time points were taken prior to bleaching to establish baseline. A different cell and frame of view was chosen for each FRAP experiment and samples exchanged after 2 hours of imaging. Recovery was monitored for 500 frames at 0.59 seconds per frame for a total of ~300 seconds.

Microscopy and image acquisition

All images were acquired using a Zeiss LSM 980 with Airyscan 2 laser scanning confocal microscope, equipped with 405, 588, 561, and 639 nanometer laser lines and Fast Airyscan detector array. Images were acquired using an inverted 63x/1.4 NA oil objective. All live imaging was done in a humidified chamber at 37 C and in the presence of 5% CO^2. Airyscan processing and maximum intensity image projection was performed using Zeiss ZEN Blue

software version 3.7 (Carl Zeiss). Image brightness and/or contrast were linearly adjusted in ZEN Blue or FIJI (*49*).

Image Analysis

FIJI, Arivis, and Python were used. All Z stack images were 3D Airyscan processed and maximum intensity projections were generated using Zeiss ZEN Blue software version 3.7 (Carl Zeiss) and saved as czi files. These projections were then converted to Arivis sis file format using the Arivis SIS batch converter (ver 4.1.0) for subsequent analysis in Arivis Vision4D (ver 4.1). All graphs and other visualization of quantification was created using Graphpad Prism 10 (ver 10.1.0) All Arivis segmentation pipelines, machine learning trainings and models, as well as all custom analysis code for FIJI, and Python available at https://github.com/TheLewisLab.

Segmentation

Segmentation of the mitochondrial network was achieved using the Arivis machine learning image trainer to design a machine learning model using the Fluorescence and EM Robust training dataset for all channels in which mitochondria were labeled. The trainer was trained using, on average, 3 representative images from the image set, classifying mitochondria and background signal until sufficient segmentation of the mitochondrial network was achieved and artifacts minimized. For Mitotracker deep red, RNR2 and tRNA signals, an intensity threshold segmenter was used to define mitochondrial objects. The resulting segments were then filtered for size, using a 0.100 - 0.200 um^2 cutoff.

For punctate signals (ND4, GRSF1, dsDNA, TFAM, MRPL23), the Arivis blob finder method was used. For more continuous signal, domains of enrichment were defined using the intensity-based threshold segmenter. All objects were then filtered by their proximity to the mitochondrial network segment using the Arivis compartmentalization function with a 60% object intersection cutoff. Object intersections within the mitochondria were also determined using the Arivis compartmentalization function with a 20% cutoff. Relevant object information, such as area, intensity measurements, or intersection data, was then exported as a csv for downstream analysis. The resulting mitochondrial network segmentation was used as an image mask and exported from Arivis as OME TIFFs for downstream analysis.

For HPG variable pulse and variable chase intensity thresholding, different intensity values were chosen for the variable pulse (2500 gray value) and variable chase (1000 gray value) due to the differing dynamic range in captures between datasets.

Object intersection analysis

To determine the mitochondrial object intersection percentages above expected by random chance, object intersection percentages of segmented objects were compared to simulated data using the average sizes, number, and mitochondrial area data extracted from Arivis segmentation above. In short, a Monte Carlo simulation was performed 100 times using the extracted object data to determine how often the two objects would overlap at least 20% by area given their respective sizes, abundances, and the compartment area of the mitochondrial network. This simulation was run for each frame of view and compared to the extracted object intersection values generated from the Arivis segmentation pipeline.

Automated linescan extraction

To quantify fluorescence intensity along linescans from skeletonized mitochondrial networks, we generated a custom Fiji macro, as follows. In short, mitochondria in micrographs were masked as described above and the mask was skeletonized using a binary image threshold via the Skeletonize 3D Fiji Package (*50*). The 'Analyze Skeleton' feature was then applied to systematically map object branch points, endpoints, and junctions within the mitochondrial network. This function was performed iteratively 5 times, each time pruning branches of less than 8 pixels in length to remove artificial branches generated during the skeletonization. After the skeleton was pruned of artificial branches the junction pixels were removed as well and the Analyze skeleton getShortestPathPoints() function used to generate polylines across the mitochondrial network. These lines were then added to the ROI manager, and using the Fiji ProfilePlot function, fluorescence intensity linescans in all channels were generated for each line. The resulting csv files were then used for downstream analysis.

Linescan analysis and peak calling

Peak calling and analysis of the fluorescence intensity linescans was performed using custom Python code available at https://github.com/TheLewisLab. Peaks of fluorescence intensity were called using the scipy find_peaks function for each extracted linescan (*51*). For each peak called in a given fluorescence channel, the 0.5 um on either side was extracted and averaged to form the average signal over a given fluorescence peak. Random average line scans were generated by randomly selecting a peak position across the same line scan dataset and averaging the 1 um surrounding that point. Averages, confidence intervals, and relevant statistics were performed using the native python packages.

Generation and validation of CRISPR KOs

Guide RNA sequences used to target POLG and SUV3 are available in Supplementary Table 2 at https://doi.org/10.1101/2024.08.05.604215. To generate knockout cells via CRISPR Cas9 gene editing, we implemented a modified version of the Lipofectamine CRISPRMax Cas9 Reagent lipofectamine protocol (ThermoFisher). Cells were transfected with Cas9 ribonucleoparticles pre-complexed with synthetic guide RNAs (Synthego) generated bySynthego Gene Knockout Kit made up of 2NLS-Cas9 and validated synthetic multi guide RNAs. Briefly, cells were seeded to 15-20% density the day prior to transfection in 6 well plates. To prepare the transfection mixture, 2 uL of 20 uM NLS-Cas9 were complexed with 4 uL of 10 uM sgRNAs in 14 uL Optimem (Invitrogen) for a final volume of 20 uL, then mixed. In parallel, 10 uL of Cas9 Plus reagent was diluted into 70 uL of Optimem, mixed, and subsequently added to the complexed Cas9/sgRNA complexes. For transfection, 6 uL of the Lipofectamine CRISPRMax Cas9 Reagent was diluted into 94 uL of Optimem, mixed, and added to the Cas9/sgRNA/Cas9 Plus solution. The mixture was allowed to incubate at room temperature for 10-minute incubation at room temperature before added dropwise to the plated cells. The media was replaced 24 hours later.

Acknowledgements

The authors thank the members of the Lewis laboratory, as well as Drs. James Hurley, Gary Karpen, and Michael Rape for helpful comments.

This work was supported by a NSF Graduate Research Fellowship to AB, and N.I.H. grants R00GM129456 and R35GM147218 to SL. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. National Institutes of Health or the U.S. National Science Foundation.

Figures

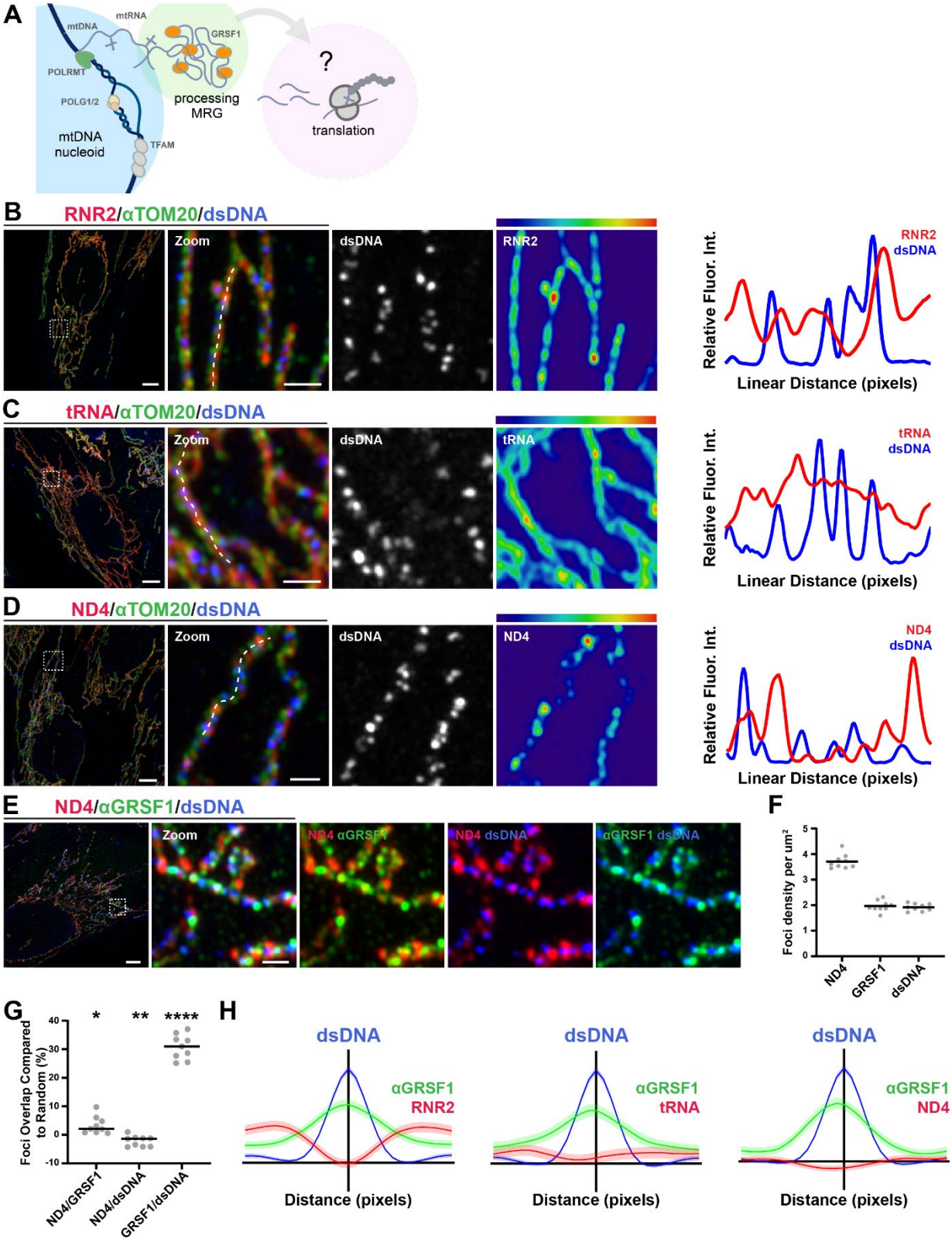

A
mtDNA
mtRNA
GRSF1
POLRMT
POLG1/2
mtDNA nucleoid
TFAM
processing MRG
?
translation
B
RNR2/αTOM20/dsDNA
Zoom
dsDNA
RNR2
Relative Fluor. Int.
RNR2
dsDNA
Linear Distance (pixels)
C
tRNA/αTOM20/dsDNA
Zoom
dsDNA
tRNA
Relative Fluor. Int.
tRNA
dsDNA
Linear Distance (pixels)
D
ND4/αTOM20/dsDNA
Zoom
dsDNA
ND4
Relative Fluor. Int.
ND4
dsDNA
Linear Distance (pixels)
E
ND4/αGRSF1/dsDNA
Zoom
ND4 αGRSF1
ND4 dsDNA
αGRSF1 dsDNA
F
Foci density per um²
ND4
GRSF1
dsDNA
G
Foci Overlap Compared to Random (%)
*
**
****
ND4/GRSF1
ND4/dsDNA
GRSF1/dsDNA
H
dsDNA
αGRSF1
RNR2
Distance (pixels)
dsDNA
αGRSF1
tRNA
Distance (pixels)
dsDNA
αGRSF1
ND4
Distance (pixels)

Figure 2.1 Processed mitochondrial RNA is excluded from nucleoids and MRGs. (A) Overview of mammalian mitochondrial gene expression pathway. (B) Representative images and linescans of fixed IMR90 cells immunolabeled with antibodies against TOM20 (green), dsDNA (blue), and RNA-FISH targeting mitoribosomal component RNR2 (red), (C) mt-transfer RNAs (red), or (D) ND4 messenger RNA (red). Scale bars 5 μm; 1 μm in zoom. (E) IMR90 cell immunolabeled with antibodies against GRSF1 (green), dsDNA (blue), and RNA-FISH targeting ND4 mRNA (red). Scale bars 5 μm; 1 μm in zoom. (F) Comparison of the density of ND4-FISH, anti-GRSF1, and anti-dsDNA foci normalized to mitochondrial area. (G) The frequency of total ND4-FISH, anti-GRSF1, and anti-dsDNA foci that overlap by 20% or more in maximum intensity projections in pairwise comparison. Dotted line represents the frequency of overlap expected by random chance, given the foci density along mitochondrial tubules. (H) Iterative linescan analyses of GRSF1 and RNA localization relative to mtDNA nucleoids marked by anti-dsDNA immunofluorescence and RNA-FISH labeling: RNR2 (n=1392 nucleoids from 21 cells); tRNA (n= 1020 nucleoids from 13 cells); ND4 (n=1211 nucleoids from 26 cells). (****P<0.0001, **P<0.01, *P<0.05, one-sided t-test).

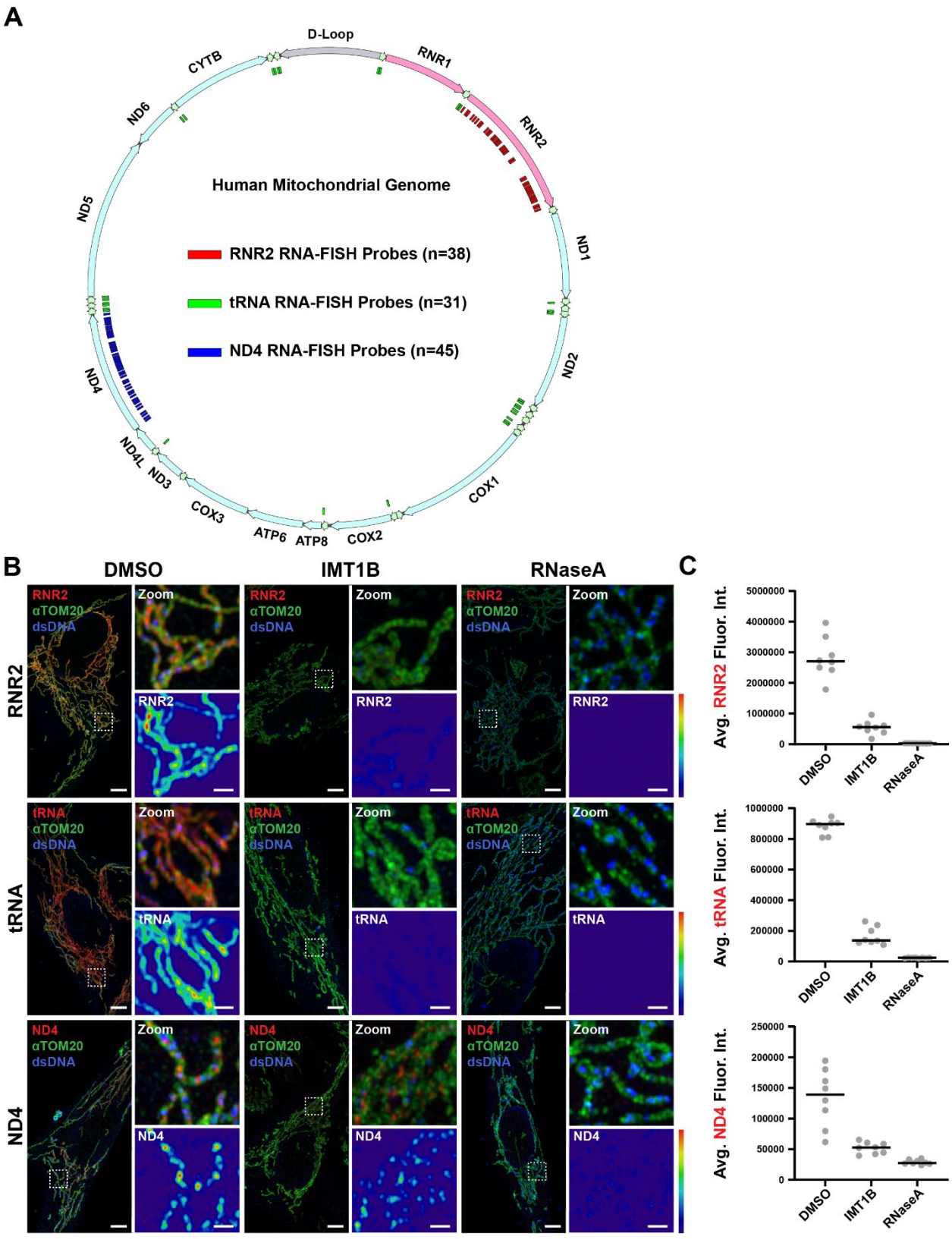

A
D-Loop
RNR1
RNR2
ND1
ND2
COX1
COX2
ATP8
ATP6
COX3
ND3
ND4L
ND4
ND5
ND6
CYTB
Human Mitochondrial Genome
RNR2 RNA-FISH Probes (n=38)
tRNA RNA-FISH Probes (n=31)
ND4 RNA-FISH Probes (n=45)
B
DMSO
IMT1B
RNaseA
RNR2
tRNA
ND4
αTOM20
dsDNA
Zoom
C
Avg. RNR2 Fluor. Int.
Avg. tRNA Fluor. Int.
Avg. ND4 Fluor. Int.
DMSO
IMT1B
RNaseA

Figure 2.2 Specifically visualizing mitochondrial DNA encoded RNAs via RNA-FISH. (A) Schematic of where the RNA-FISH probes are located on each mitochondrial RNA. (B) Representative images of fixed IMR90 cells treated with DMSO (left), 10 uM IMT1B for 96 hours (middle), or 100 ug/mL RNaseA for 1 hour prior to FISH labeling(right) immunolabeled for TOM20 (green) and dsDNA (blue), and mtRNA-FISH against either RNR2 (top), tRNAs (middle), or ND4 (bottom) (red). Scale bars 5 μm; 1 μm in zoom. (C) Quantification of average RNA intensity per mitochondrial area in each condition for each RNA in (B).

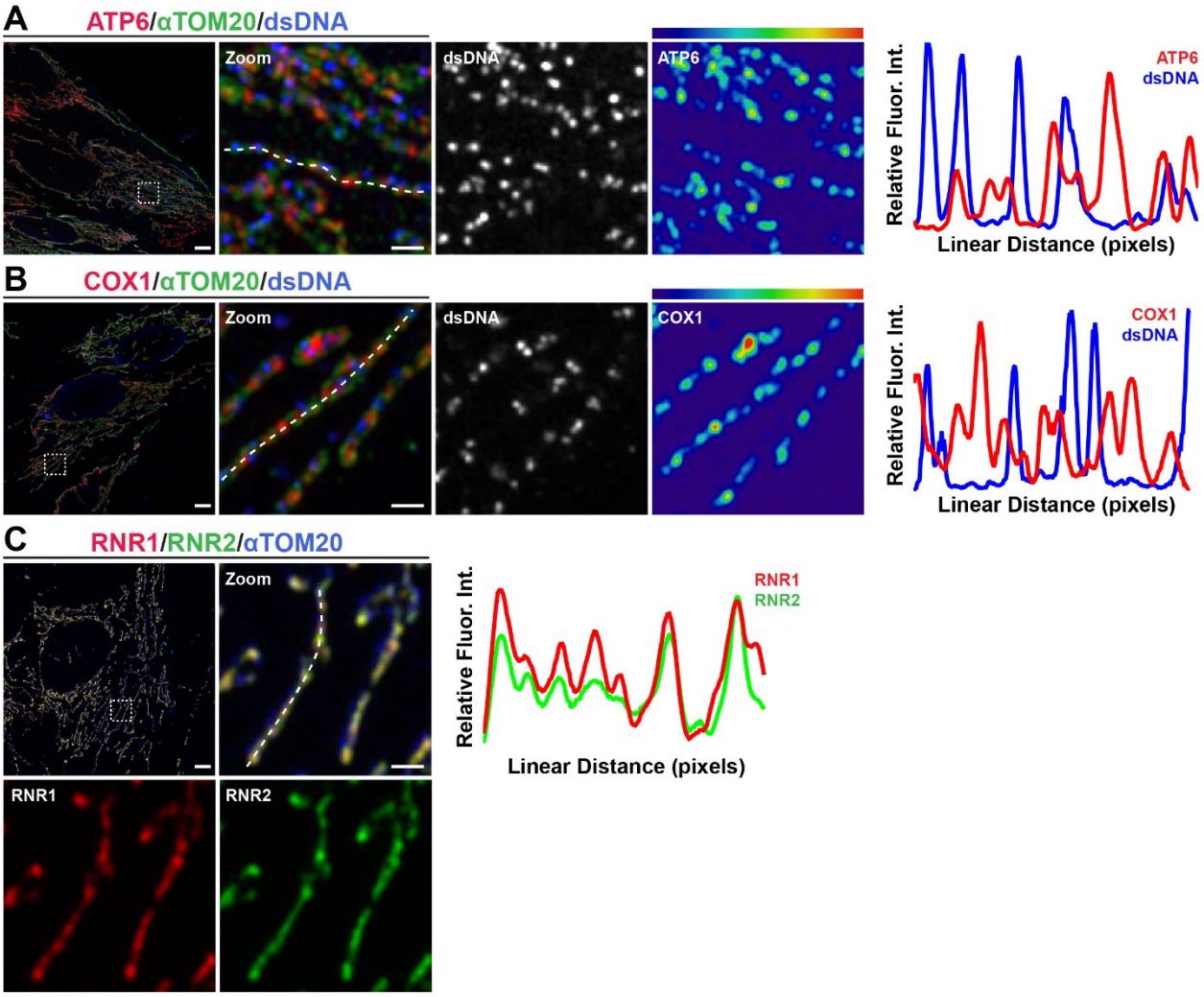


Figure 2.3 RNA-FISH of other mitochondrial DNA encoded RNAs. (A) Representative images and linescans of fixed IMR90 cells immunolabeled with antibodies against TOM20 (green), dsDNA (blue), and RNA-FISH targeting AT6 messenger RNA (red), (B) COX1 messenger RNA (red), or (C) mitoribosomal components RNR1 (red) and RNR2 (green). Scale bars 5 μm; 1 μm in zoom.

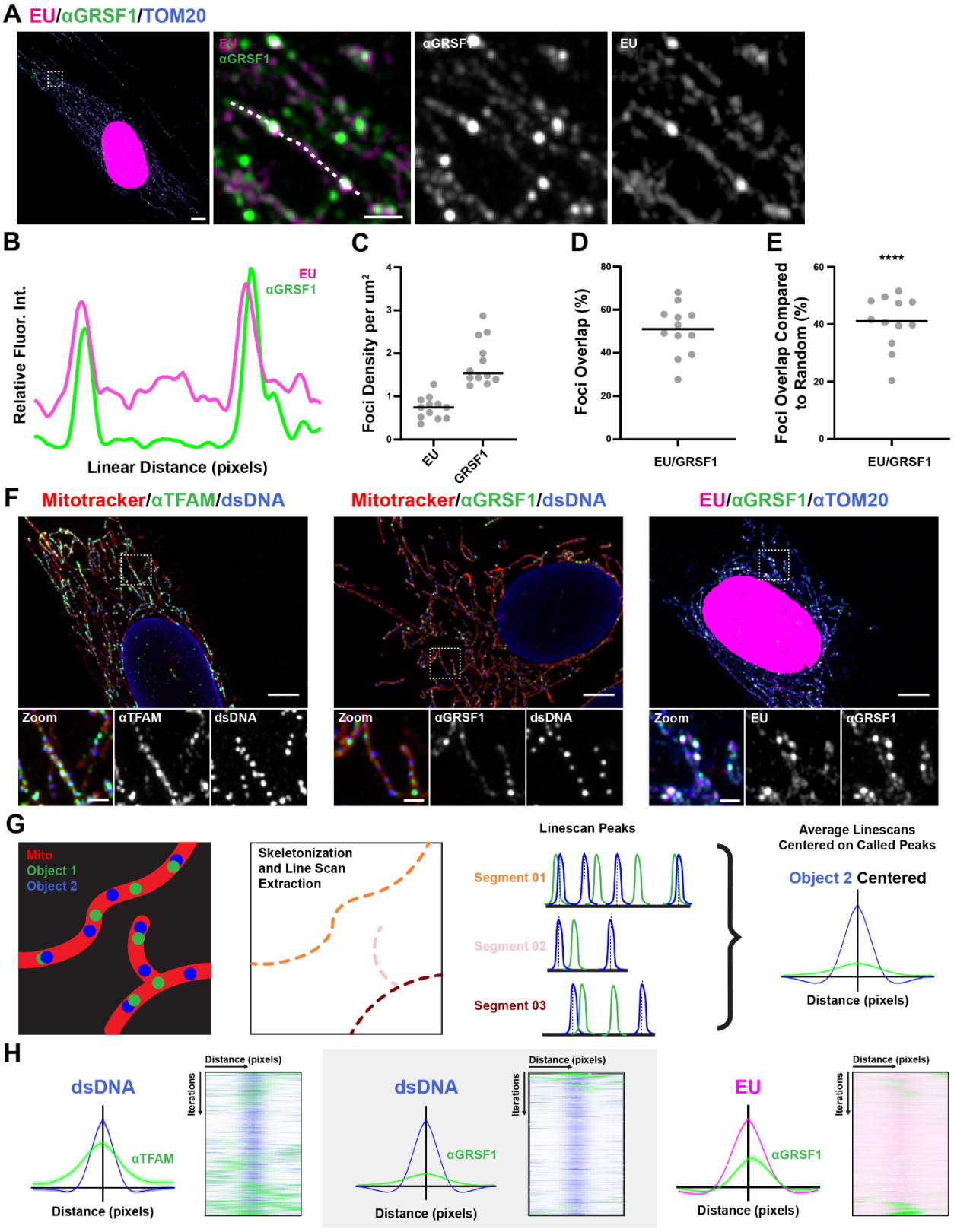

A EU/αGRSF1/TOM20
EU
αGRSF1
αGRSF1
EU
B
Relative Fluor. Int.
EU
αGRSF1
Linear Distance (pixels)
C
Foci Density per um²
EU
GRSF1
D
Foci Overlap (%)
EU/GRSF1
E
****
Foci Overlap Compared to Random (%)
EU/GRSF1
F
Mitotracker/αTFAM/dsDNA
Mitotracker/αGRSF1/dsDNA
EU/αGRSF1/αTOM20
Zoom
αTFAM
dsDNA
Zoom
αGRSF1
dsDNA
Zoom
EU
αGRSF1
G
Mito
Object 1
Object 2
Skeletonization and Line Scan Extraction
Linescan Peaks
Segment 01
Segment 02
Segment 03
Average Linescans Centered on Called Peaks
Object 2 Centered
Distance (pixels)
H
dsDNA
αTFAM
Distance (pixels)
Distance (pixels)
Iterations
dsDNA
αGRSF1
Distance (pixels)
Distance (pixels)
Iterations
EU
αGRSF1
Distance (pixels)
Distance (pixels)
Iterations

Figure 2.4 Labeling nascent mitochondrial transcription with the metabolic label EU and Validating Iterative linescan analysis. (A) Representative image of IMR90 cells in which nascent mtRNA was pulse-labeled with EU (magenta) then fixed and immunolabeled with antibodies against TOM20 (blue) and GRSF1 (green). Scale bars 5 μm; 1 μm in zoom. (B) Linescan of (A). (C) Comparison of the density of EU and anti-GRSF1 foci normalized to mitochondrial area. (D) The percent of EU and GRSF1 foci overlapping by at least 20%. (E) The frequency of EU and anti-GRSF1 foci overlap greater than expected by random chance. (F) (Left, Middle) Representative images of IMR90 cells labeled with Mitotracker Deep Red, fixed and immunolabeled with antibodies against dsDNA (blue), (Left) TFAM (green), or (Middle) GRSF1 (green). (Right) Representative image of IMR90 cells in which nascent mtRNA was pulse-labeled with EU (magenta) then fixed and immunolabeled with antibodies against TOM20 (blue) and GRSF1 (green). Scale bars 5 μm; 1 μm in zoom. (G) Schematic of how the average line scan centered around an object of interest is generated from segmented mitochondrial images. (H) Iterative linescan analysis and heatmap profile of (Left) dsDNA and TFAM (n=884 dsDNA foci from 18 cells), (Middle) dsDNA and GRSF1 (n=3799 dsDNA foci from 23 cells), and (Right) EU and GRSF1 (n=1647 EU foci from 26 cells).

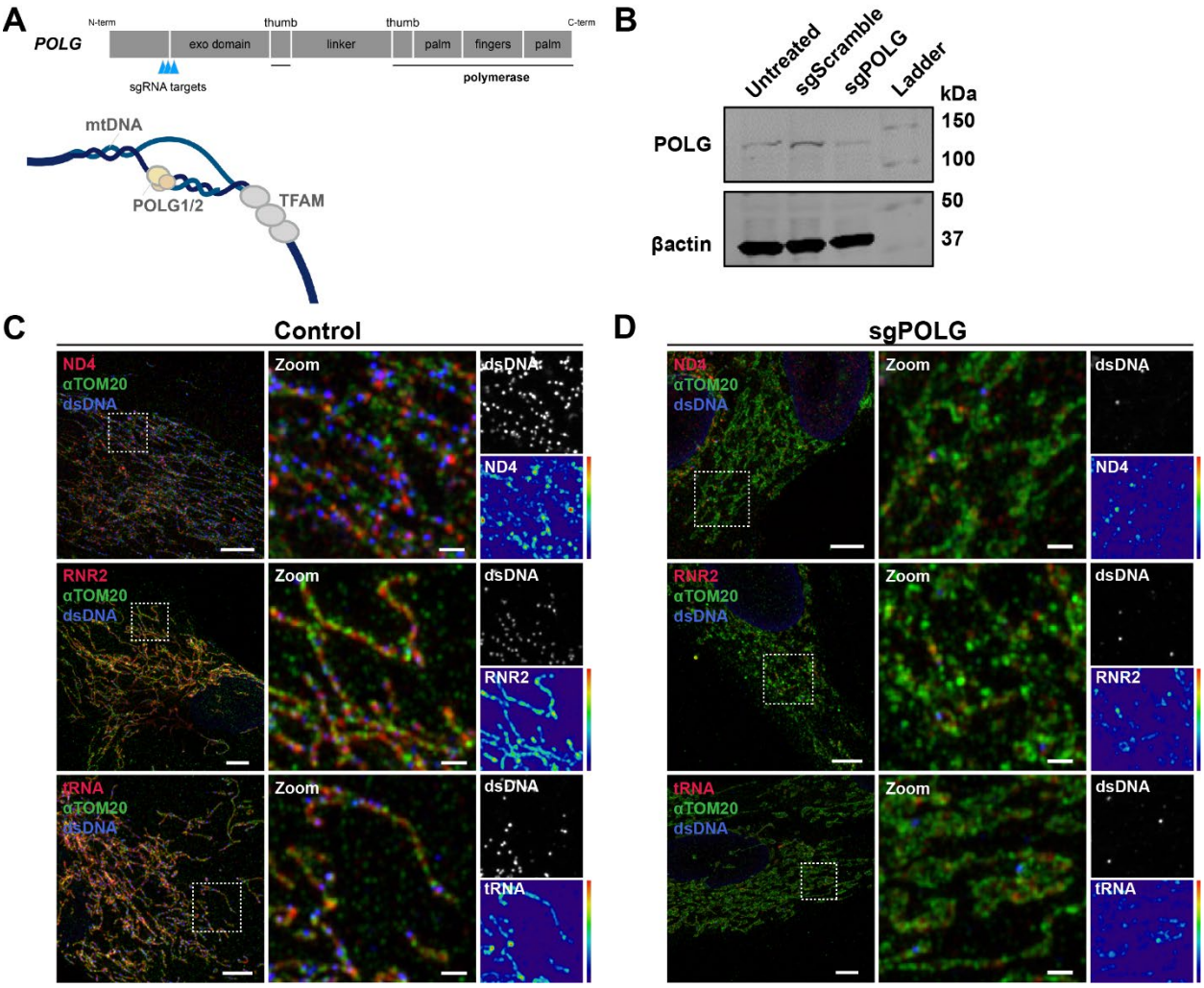


Figure 2.5 Mitochondrial RNA in POLG depleted cells. (A) Schematic of POLG and the sites of sgRNA editing. (B). Western blot to detect POLG in untreated, Scramble sgRNA cells, and POLG sgRNA cells. (C) Representative fixed control IMR90 cells immunolabeled with antibodies against TOM20 (green) and dsDNA (Blue), and mtRNA-FISH against ND4 (top), RNR2 (middle), tRNA (bottom) (red). Scale bars 5 μm; 1 μm in zoom. (D) Representative fixed 1 week sgPOLG IMR90 cells immunolabeled with antibodies against TOM20 (green) and dsDNA (Blue), and mtRNA-FISH against ND4 (top), RNR2 (middle), tRNAs (bottom) (red). Scale bars 5 μm; 1 μm in zoom.

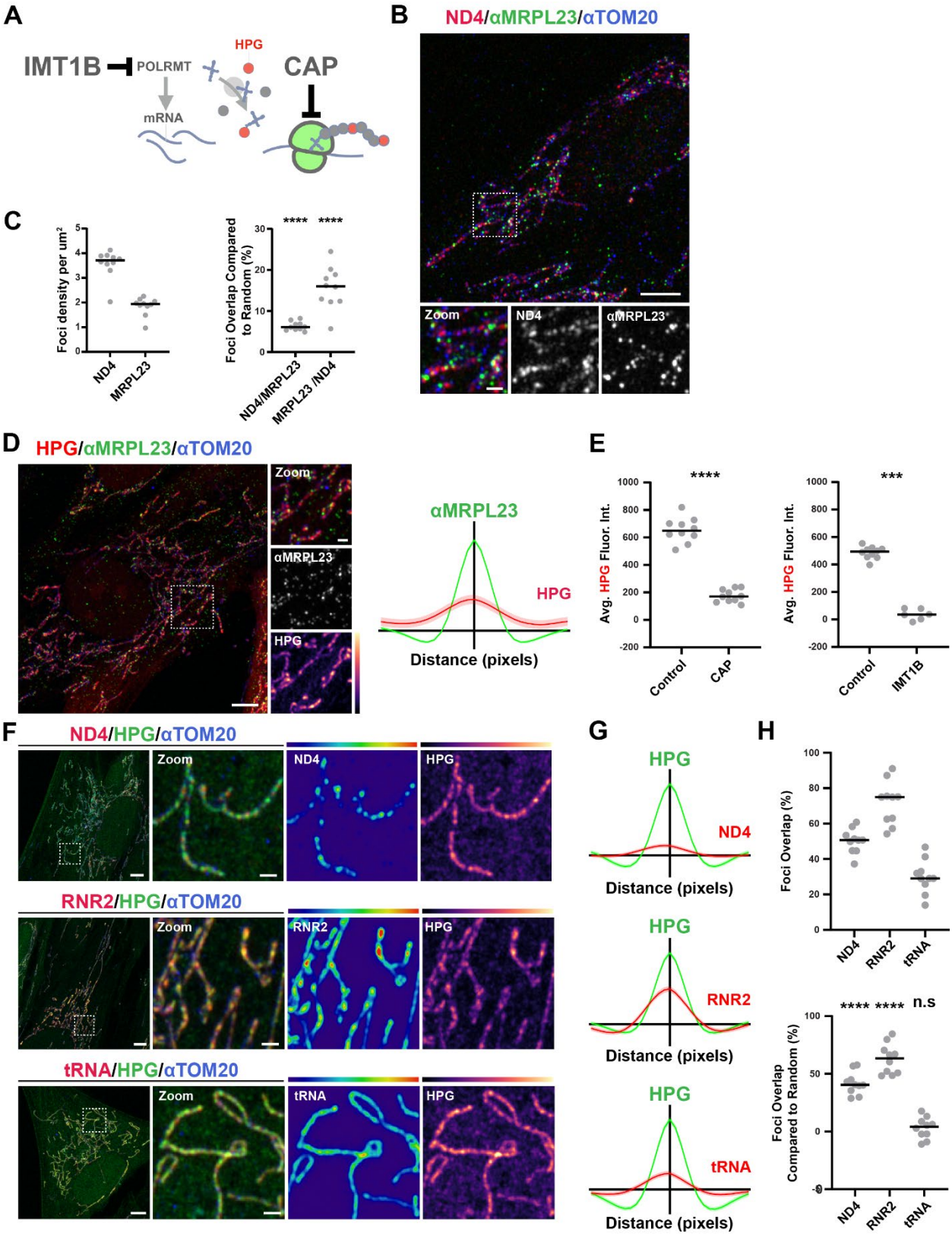

A
IMT1B
POLRMT
HPG
CAP
mRNA
B
ND4/αMRPL23/αTOM20
Zoom
ND4
αMRPL23
C
Foci density per um²
ND4
MRPL23
Foci Overlap Compared to Random (%)
****
****
ND4/MRPL23
MRPL23 /ND4
D
HPG/αMRPL23/αTOM20
Zoom
αMRPL23
HPG
Distance (pixels)
E
Avg. HPG Fluor. Int.
****
Control
CAP
***
IMT1B
F
ND4/HPG/αTOM20
RNR2/HPG/αTOM20
tRNA/HPG/αTOM20
G
HPG
ND4
RNR2
tRNA
Distance (pixels)
H
Foci Overlap (%)
Foci Overlap Compared to Random (%)
n.s

Figure 2.6 Mitochondrial RNA enrichment marks local translation hubs. (A) Metabolic labeling approach for imaging and manipulating mitochondrial translation. (B) Representative image of fixed IMR90 cells immunolabeled with antibodies against TOM20 (blue), MRPL23 (green), and RNA-FISH targeting ND4 mRNA (red). Scale bars 5 μm; 1 μm in zoom. (C) (Left) Comparison of the density of ND4-FISH and anti-MRPL23 foci normalized to mitochondrial area. (Right) The frequency of ND4-FISH and anti-MRPL23 foci overlap greater than expected by random chance. ****P<0.0001, one-sided t-test. (D) (Left) Representative image of IMR90 cell fixed and immunolabeled to detect MRPL23 (green) and TOM20 (blue) after a 15 minute pulse of 50 μM HPG (red). Scale bars 5 μm; 1 μm in zoom. (Right) Iterative linescan analysis of MRPL23 (n= 4048 from 40 cells) and ND4-FISH fluorescence intensity. (E) (Left) Average HPG fluorescence intensity in segmented mitochondria in control cells versus following 20 min pulse of 50 ug/mL Chloramphenicol (CAP). (Right) Average HPG fluorescence intensity in segmented mitochondria in control cells versus after a 48 hour pulse with 10 uM IMT1B, a mitochondrial RNA polymerase inhibitor. ****P<0.0001, ***P<0.001, Mann-Whitney test. (F) Visualization of HPG fluorescence intensity (green) within mitochondria relative to RNA-FISH (red) targeting ND4 mRNA (top), RNR2 (middle), or mt-tRNAs (bottom). At right, RNA-FISH and HPG are shown color-coded for intensity. Scale bars 5 μm; 1 μm in zoom. (G) Iterative linescans analyses of RNA fluorescence intensities (red) relative to HPG foci (green): ND4 (n=3870 HPG foci from 44 cells); RNR2 (n= 4052 nucleoids from 36 cells); tRNA (n=4522 nucleoids from 41 cells). (H) (top) The proportion of segmented RNA foci that overlap with HPG foci by 20% or more in maximum intensity projections. (Bottom) As above, adjusted relative to the frequency of overlap expected by random chance. ****P<0.0001, one-sided t-test.

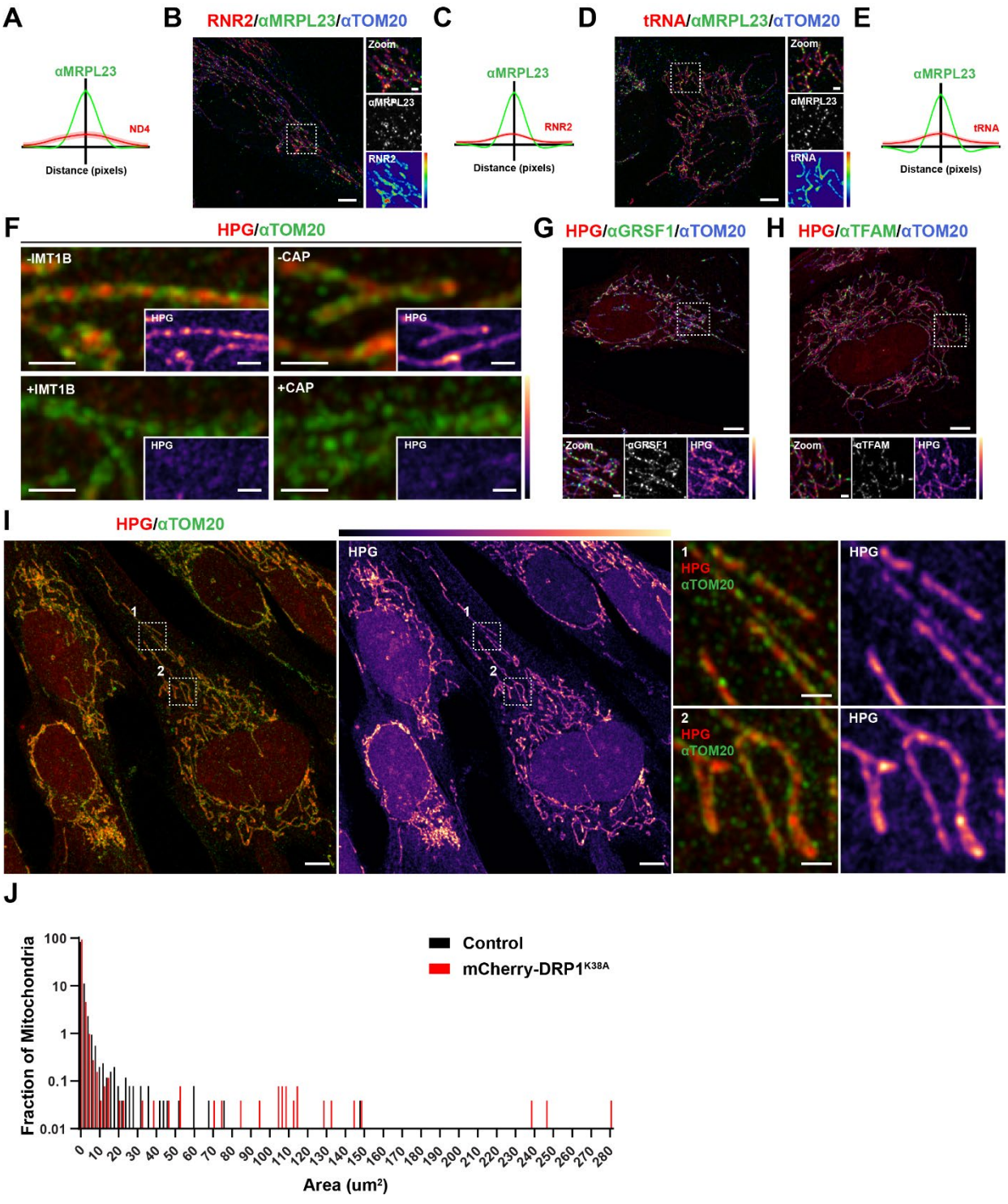

A
αMRPL23
ND4
Distance (pixels)
B
RNR2/αMRPL23/αTOM20
Zoom
αMRPL23
RNR2
C
αMRPL23
RNR2
Distance (pixels)
D
tRNA/αMRPL23/αTOM20
Zoom
αMRPL23
tRNA
E
αMRPL23
tRNA
Distance (pixels)
F
HPG/αTOM20
-IMT1B
HPG
-CAP
HPG
+IMT1B
HPG
+CAP
HPG
G
HPG/αGRSF1/αTOM20
Zoom
αGRSF1
HPG
H
HPG/αTFAM/αTOM20
Zoom
αTFAM
HPG
I
HPG/αTOM20
1
2
HPG
1
HPG
αTOM20
HPG
2
HPG
αTOM20
HPG
J
Control
mCherry-DRP1K38A
Fraction of Mitochondria
100
10
1
0.1
0.01
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280
Area (um²)

Figure 2.7 Extended RNA-FISH and HPG data. (A) Iterative linescan analysis of ND4 (red) relative to MRPL23 (green) (n=944 MRPL23 foci from 29 cells). (B) Representative image of fixed IMR90 cell immunolabeled with antibodies against TOM20 (blue) and MRPL23 (green), and mtRNA-FISH targeting RNR2 (red). Scale bars 5 μm; 1 μm in zoom. (C) Iterative linescan analysis of RNR2 (red) relative to MRPL23 (green) (n=2265 MRPL23 foci from 30 cells). (D) Representative image of fixed IMR90 cell immunolabeled with antibodies against TOM20 (blue) and MRPL23 (green), and mtRNA-FISH targeting tRNA (red). Scale bars 5 μm; 1 μm in zoom. (E) Iterative linescan analysis of tRNA (red) relative to MRPL23 (green) (n=2201 MRPL23 foci from 35 cells). (F) Representative images of control IMR90 cells as well as cells treated with IMT1B (left) or chloramphenicol (right) during HPG pulse-labeling then subsequently fixed, immunolabeled for TOM20 (green), and subjected to Copper-click cycloaddition of AlexaFluor647 to HPG-alkyne (red). Scale bars 1 μm. (G) IMR90 cell pulse-labeled with HPG (red) then fixed and subjected to immunolabeling against GRSF1 (green) and TOM20 (blue). Scale bars 5 μm; 1 μm in zoom. (H) IMR90 cell pulse-labeled with HPG (red) then fixed and subjected to immunolabeling against TFAM (green) and TOM20 (blue). Scale bars 5 μm; 1 μm in zoom. (I) IMR90 cell pulse-labeled for 1 hour with HPG (red) then fixed and subjected to immunolabeling against TOM20 (green). Scale bars 5 μm; 1 μm in zoom. (J) Histogram of mitochondrial areas in control cells (black) and mCherry-DRP1K38A expressing cells (red).

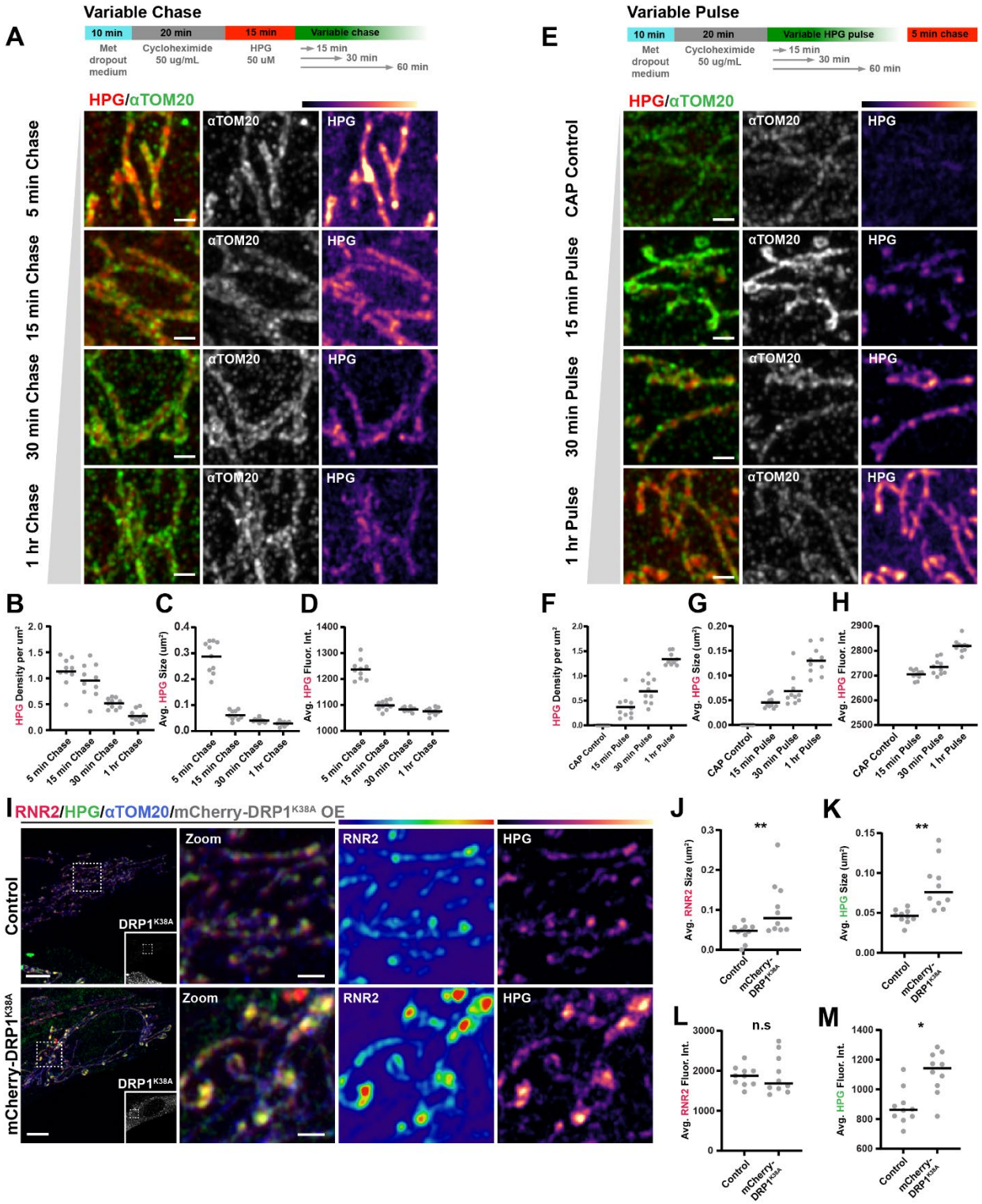

A
Variable Chase
10 min
20 min
15 min
Variable chase
Met dropout medium
Cycloheximide 50 ug/mL
HPG 50 uM
15 min
30 min
60 min
HPG/αTOM20
5 min Chase
15 min Chase
30 min Chase
1 hr Chase
αTOM20
HPG
E
Variable Pulse
10 min
20 min
Variable HPG pulse
5 min chase
Met dropout medium
Cycloheximide 50 ug/mL
15 min
30 min
60 min
HPG/αTOM20
CAP Control
15 min Pulse
30 min Pulse
1 hr Pulse
αTOM20
HPG
B
HPG Density per um²
5 min Chase
15 min Chase
30 min Chase
1 hr Chase
C
Avg. HPG Size (um²)
D
Avg. HPG Fluor. Int.
F
HPG Density per um²
CAP Control
15 min Pulse
30 min Pulse
1 hr Pulse
G
Avg. HPG Size (um²)
H
Avg. HPG Fluor. Int.
I
RNR2/HPG/αTOM20/mCherry-DRP1K38A OE
Control
mCherry-DRP1K38A
DRP1K38A
Zoom
RNR2
HPG
J
Avg. RNR2 Size (um²)
**
Control
mCherry-DRP1K38A
K
Avg. HPG Size (um²)
**
L
Avg. RNR2 Fluor. Int.
n.s
M
Avg. HPG Fluor. Int.
*

Figure 2.8 Mitochondrial translation hubs are dynamic. (A) (Top) Overview of HPG labeling time course with variable chase times. (Bottom) Representative images of IMR90 cells pulse labeled for 15 minutes with HPG and chased with unlabeled methionine for 5, 15, 30, or 60 minutes. Scale bar 1 μm. (B) Density of thresholded HPG domains normalized to mitochondrial area in each condition. (C) Average size of HPG-labeled domains in each condition. (D) Average above-threshold fluorescence intensity of HPG-labeled domains in each condition. (E) (Top) Overview of HPG labeling time course with variable pulse times. (Bottom) Representative images of IMR90 cells pulse labeled for 15, 30, or 60 minutes with HPG followed by a constant chase in unlabeled methionine for 5 minutes, relative to control cells incubated in HPG for 15 minutes concurrent with 50 uM CAP. Scale bar 1 μm. (F-H) Quantification of the number of thresholded HPG objects per segmented mitochondrial area in each condition. (I) Representative images of IMR90 cells that were transiently transfected with mCherry-DRP1K38A (greyscale), pulse labeled for 15 minutes with 50 uM HPG (green), immunolabeled with an antibody against TOM20 (blue), and RNR2-FISH (red). Scale bar 5 μm; 1μm in zoom. (J) Average size of thresholded RNR2 signal intensity. (K) Average size of thresholded HPG signal intensity. (L) Average fluorescence intensity of RNR2 signals per thresholded mitochondrion. (M) Average fluorescence intensity of HPG per thresholded mitochondrion.(**P<0.01,*P<0.05, Mann-Whitney Test).

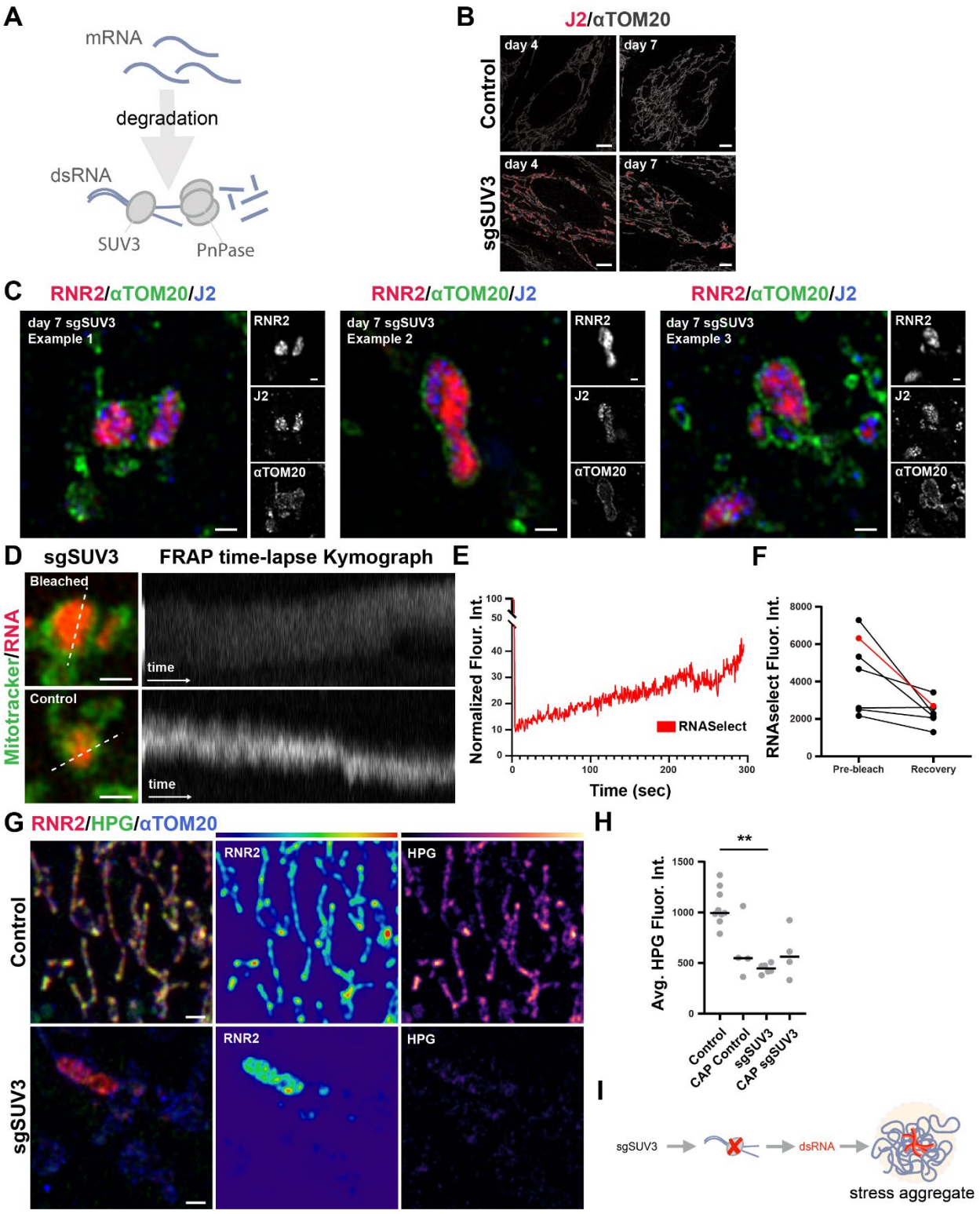

A
mRNA
degradation
dsRNA
SUV3
PnPase
B
J2/αTOM20
day 4
day 7
Control
sgSUV3
C
RNR2/αTOM20/J2
day 7 sgSUV3
Example 1
Example 2
Example 3
RNR2
J2
αTOM20
D
sgSUV3
FRAP time-lapse Kymograph
Mitotracker/RNA
Bleached
Control
time
E
Normalized Flour. Int.
RNASelect
Time (sec)
F
RNAselect Fluor. Int.
Pre-bleach
Recovery
G
RNR2/HPG/αTOM20
HPG
H
Avg. HPG Fluor. Int.
**
Control
CAP Control
sgSUV3
CAP sgSUV3
I
sgSUV3
dsRNA
stress aggregate

Figure 2.9 Mitochondrial RNAs remodel into translationally-repressed mesoscale bodies during stress. (A) Schematic of mitochondrial transcript unwinding and degradation by SUV3 and PnPase. (B) Representative images of cells immunolabeled for TOM20 (greyscale) and dsRNA (J2; red) at 4 days versus 7 days of SUV3 depletion. Scale bar 5 μm. (C) Representative images and kymographs from live cells labeled with Mitotracker Deep Red (green) and SYTO RNASelect (red), during photobleach and recovery (Top) or unbleached control (Bottom). Scale bar 1 μm. (D) FRAP intensity for representative bleach and recovery of SYTO RNASelect over 300 seconds. (E) Comparison of pre-bleach and recovery fluorescence intensities for X mesoscale bodies from Y cells over a 300 second interval. (F) Three representative images of cells fixed and immunolabeled with antibodies against dsRNA (J2; blue), TOM20 (green), and RNA-FISH targeting RNR2 (red). Scale bar 1 μm; 1 μm in zoom. (G) Representative image of control (Top) and SUV3-depleted (Bottom) cells pulse-labeled with 50 uM HPG for 15 min (green), fixed and immunolabeled to detect TOM20 (blue), and RNA-FISH against RNR2 (red). Scale bar 1 μm. (H) Average HPG fluorescence intensity in segmented mitochondria in the presence or absence of 50 ug/mL CAP. (**P<0.01, Kruskal-Wallis test (**P<0.01), followed by Dunn's multiple comparisons. (I) Diagram of proposed impact of SUV3 depletion on dsRNA accumulation and single-stranded RNA reorganization.

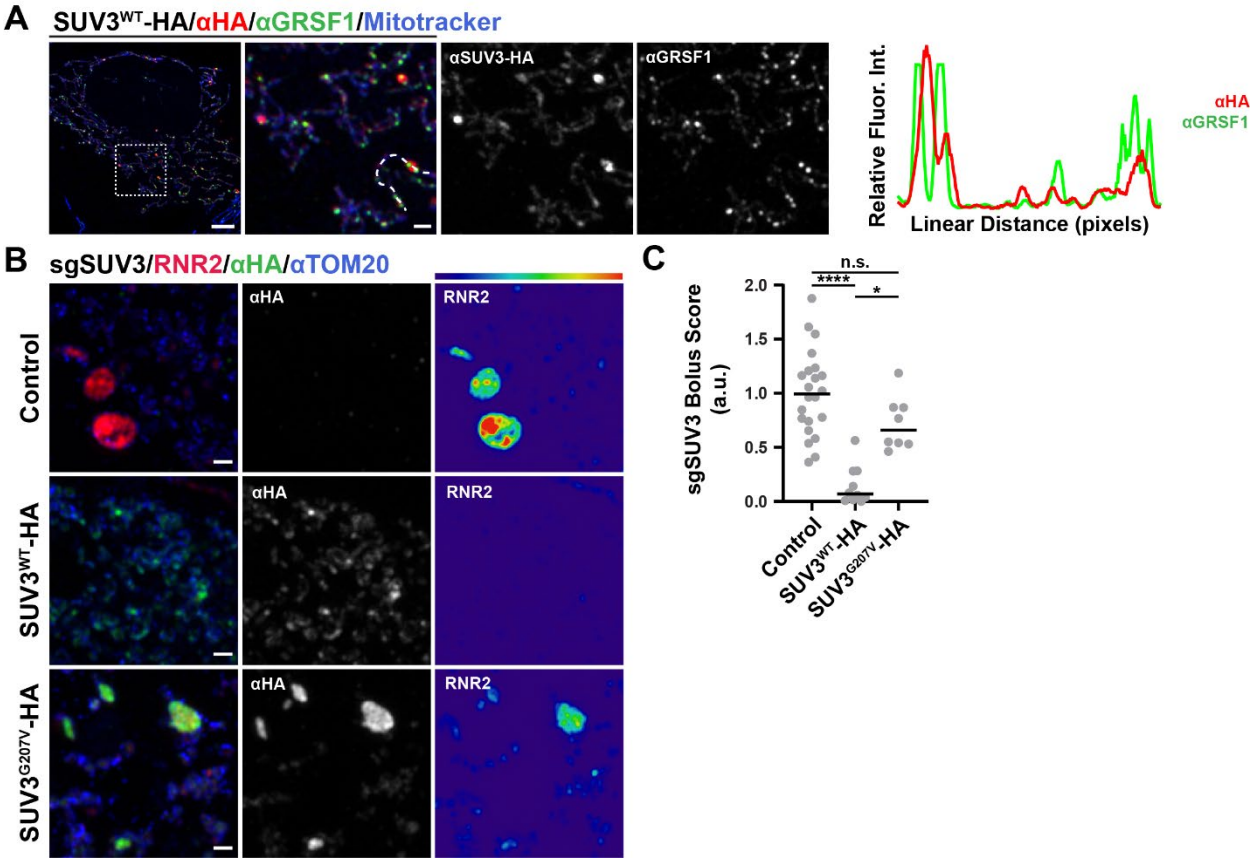


Figure 2.10 SUV3-HA expression and extended sgSUV3 data. (A) Representative image and linescan of fixed IMR90 cell transiently transfected with SUV3-HA and immunolabeled with antibodies against the HA tag (red), GRSF1 (green), and TOM20 (blue). Scale bars 5 μm; 1 μm in zoom. (B) Representative images of fixed 1 week sgSUV3 IMR90 cells either not expressing (Top), expressing a wild type HA tagged SUV3 (Middle), or expressing a G207V point mutant HA tagged SUV3 (Bottom), immunolabeled with antibodies against the HA tag (green) and TOM20 (blue), and mtRNA-FISH against RNR2 (red). Scale bars 1 μm. (C) Quantification of the strength of the SUV3 depletion phenotype (as measured by the ratio of the total RNR2 boluses area over the total mitochondrial area with sgSUV3 cells not expressing an HA tagged SUV3 protein set to 1). (****P<0.0001,*P<0.05 Kruskal-Wallis test (****), followed by Dunn's multiple comparison).

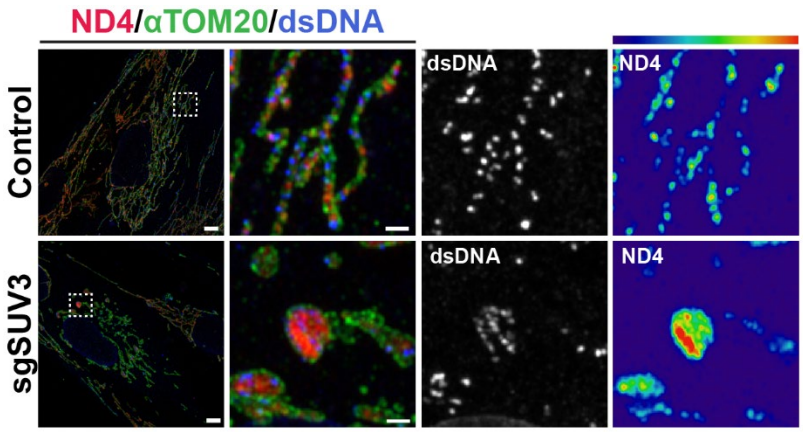


Figure 2.11 Messenger RNA localization in cells depleted of SUV3. Representative image of fixed IMR90 cells 1 week after transient transfection with CRISPR RNP complexes made up of Cas9 sgRNAs against SUV3, immunolabeled with antibodies against TOM20 (green) and dsDNA (blue), and mtRNA-FISH against ND4 (red). (Top) cell not exhibiting phenotype of SUV3 depletion, (Bottom) cell exhibiting phenotype of SUV3 depletion. Scale bars 5 μm; 1 μm in zoom.

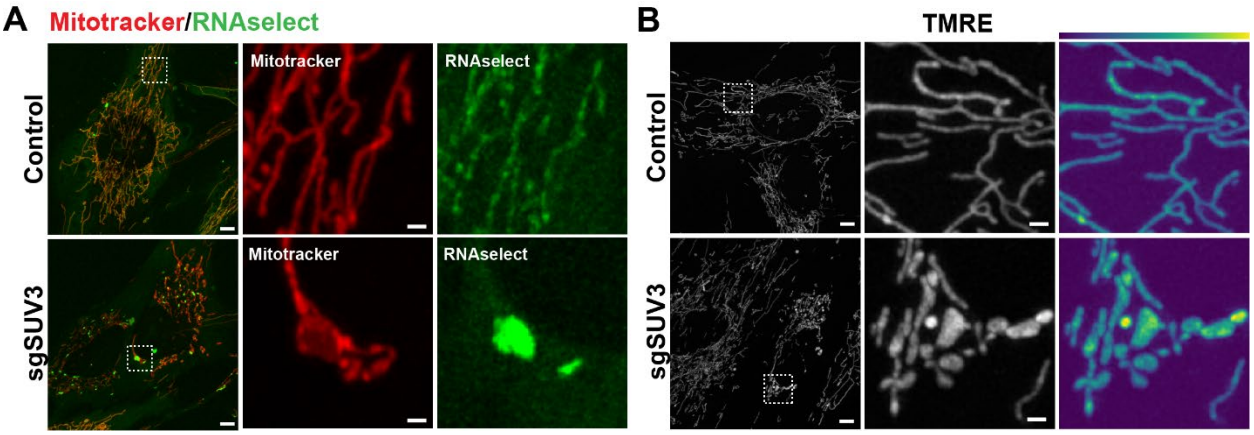


Figure 2.12 Membrane potential is preserved in cells exhibiting MSB structures. (A) Representative live image of control cell (top) and 1 week sgSUV3 cell (bottom) labeled with mitotracker deep red (red) and RNAselect (green). Scale bars 5 μm; 1 μm in zoom. (B) Representative live image of control cell (top) and 1 week sgSUV3 cell (bottom) labeled with TMRE. Scale bars 5 μm; 1 μm in zoom.

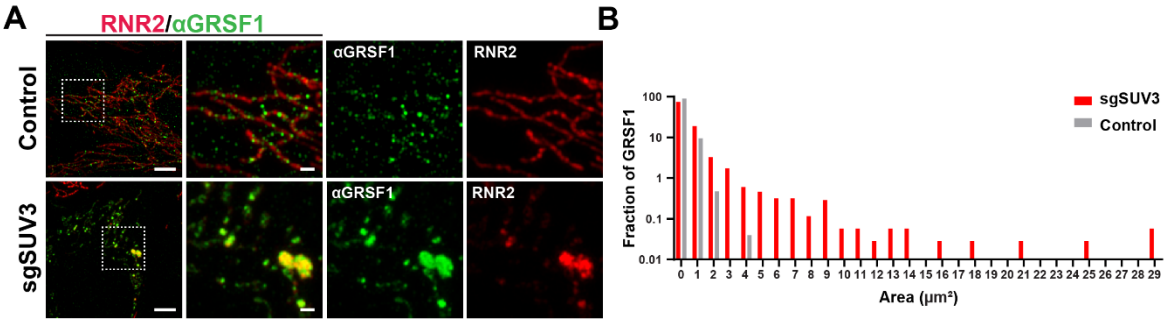


Figure 2.13 GRSF1 is enriched with MSBs. (A) Representative images and linescans of control (Top) and 1 week sgSUV3 (Bottom) IMR90 cells, fixed and immunolabeled with an antibody against GRSF1 (green), and mtRNA-FISH against RNR2 (red). Scale bars 5 µm; 1 µm in zoom. (B) Histogram of the GRSF1 object surface areas in control (gray) and sgSUV3 cells (red) (bins of 1 um2).

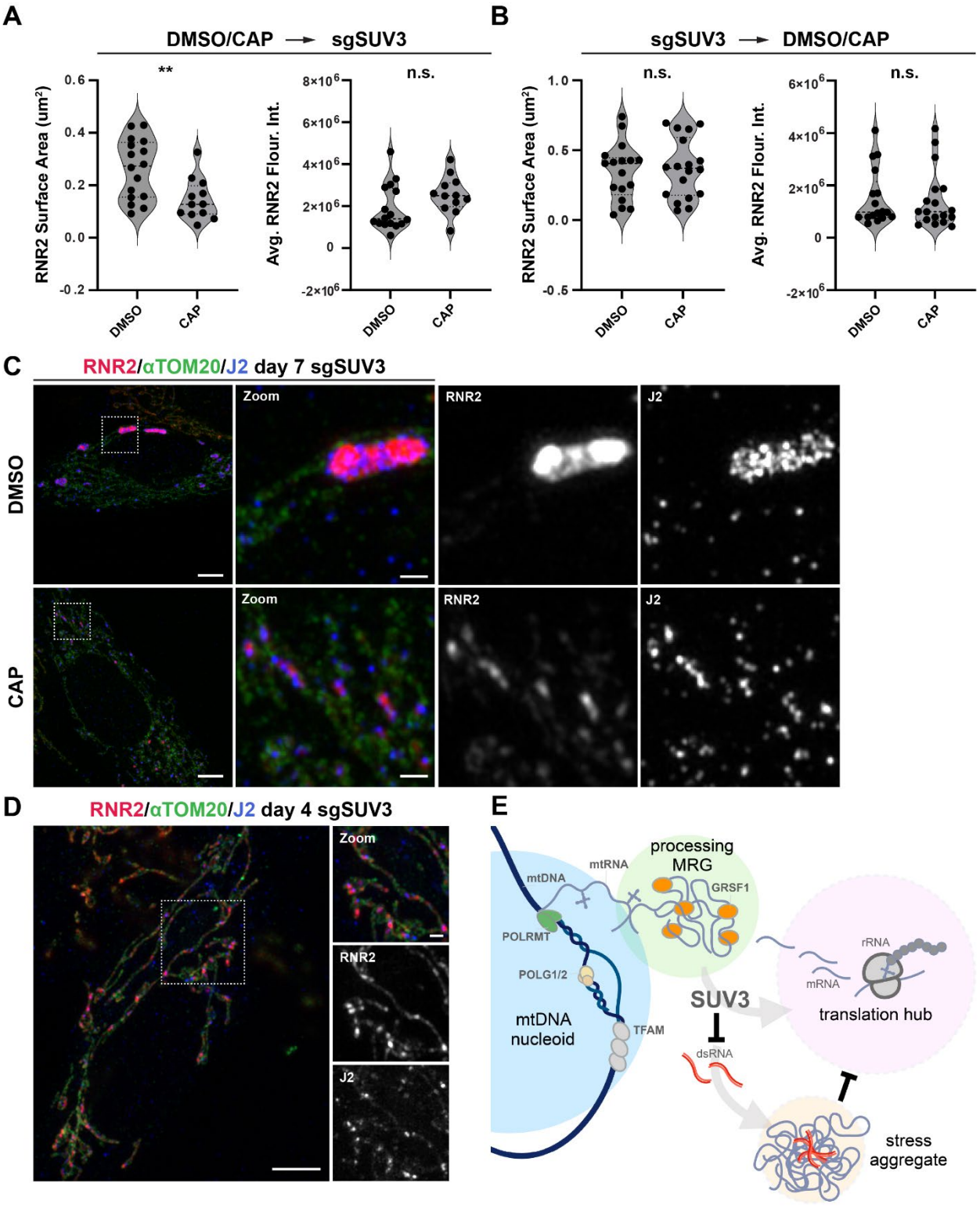

A
DMSO/CAP → sgSUV3
RNR2 Surface Area (um²)
**
Avg. RNR2 Flour. Int.
n.s.
DMSO
CAP
B
sgSUV3 → DMSO/CAP
n.s.
n.s.
C
RNR2/αTOM20/J2 day 7 sgSUV3
DMSO
CAP
Zoom
RNR2
J2
D
RNR2/αTOM20/J2 day 4 sgSUV3
E
mtDNA
mtRNA
processing
MRG
GRSF1
POLRMT
POLG1/2
TFAM
mtDNA
nucleoid
SUV3
dsRNA
rRNA
mRNA
translation hub
stress
aggregate

Figure 2.14 Mitochondrial RNA remodeling in stress is proteoprotective. (A) Average size and fluorescence intensity of segmented RNR2-FISH signals in IMR90 cells when SUV3 depletion is induced in cells already translationally repressed by incubation in 50 ug/mL CAP. (B) Average size and fluorescence intensity of segmented RNR2-FISH signals in IMR90 cells incubated in 50 ug/mL CAP 7 days after transfection with sgSUV3 RNP complexes. (**P<0.01, Mann-Whitney Test). (C) Representative images of control (Top) and 50 uM CAP-treated (Bottom) cells 7 days after transfection with sgSUV3 RNP complexes,, fixed and immunolabeled to detect dsRNA (J2; blue), TOM20 (green), and RNA-FISH against RNR2 (red). Scale bar 5 μm; 1μm in zoom. (D) Representative image of cells 4 days after transfection with sgSUV3 RNP complexes, fixed and immunolabeled to detect dsRNA (J2; blue), TOM20 (green), and RNR2-FISH against RNR2 (red). Scale bar 5 μm; 1μm in zoom. (E) Proposed model for organization of mitochondrial gene expression into translational hubs. SUV3-dependent dsRNA accumulation, translational suppression, and mesoscale body formation protect the mitochondrial proteome during stress.

References

1. A. Suomalainen, J. Nunnari, Mitochondria at the crossroads of health and disease. *Cell* **187**, 2601–2627 (2024).
2. C. M. Gustafsson, M. Falkenberg, N.-G. Larsson, Maintenance and Expression of Mammalian Mitochondrial DNA. *Annu. Rev. Biochem.* **85**, 133–160 (2016).
3. H. Antonicka, F. Sasarman, T. Nishimura, V. Paupe, E. A. Shoubridge, The Mitochondrial RNA-Binding Protein GRSF1 Localizes to RNA Granules and Is Required for Posttranscriptional Mitochondrial Gene Expression. *Cell Metab.* **17**, 386–398 (2013).
4. S. C. Lewis, L. F. Uchiyama, J. Nunnari, ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. *Science* **353**, aaf5549 (2016).
5. N.-G. Larsson, J. Wang, H. Wilhelmsson, A. Oldfors, P. Rustin, M. Lewandoski, G. S. Barsh, D. A. Clayton, Mitochondrial transcription factor A is necessary for mtDNA maintance and embryogenesis in mice. *Nat. Genet.* **18**, 231–236 (1998).
6. N. A. Bonekamp, M. Jiang, E. Motori, R. G. Villegas, C. Koolmeister, I. Atanassov, A. Mesaros, C. B. Park, N.-G. Larsson, High levels of TFAM repress mammalian mitochondrial DNA transcription in vivo. *Life Sci. Alliance* **4** (2021).
7. A. A. Jourdain, M. Koppen, M. Wydro, C. D. Rodley, R. N. Lightowlers, Z. M. Chrzanowska-Lightowlers, J.-C. Martinou, GRSF1 Regulates RNA Processing in Mitochondrial RNA Granules. *Cell Metab.* **17**, 399–410 (2013).
8. A. Ohkubo, L. V. Haute, D. L. Rudler, M. Stentenbach, F. A. Steiner, O. Rackham, M. Minczuk, A. Filipovska, J.-C. Martinou, The FASTK family proteins fine-tune mitochondrial RNA processing. *PLOS Genet.* **17**, e1009873 (2021).
9. B. Ruzzenente, M. D. Metodiev, A. Wredenberg, A. Bratic, C. B. Park, Y. Cámara, D. Milenkovic, V. Zickermann, R. Wibom, K. Hultenby, H. Erdjument-Bromage, P. Tempst, U. Brandt, J. B. Stewart, C. M. Gustafsson, N.-G. Larsson, LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. *EMBO J.* **31**, 443–456 (2012).
10. L. Chatre, M. Ricchetti, Large heterogeneity of mitochondrial DNA transcription and initiation of replication exposed by single-cell imaging. *J. Cell Sci.* **126**, 914–926 (2013).
11. M. Zorkau, C. A. Albus, R. Berlinguer-Palmini, Z. M. A. Chrzanowska-Lightowlers, R. N. Lightowlers, High-resolution imaging reveals compartmentalization of mitochondrial protein synthesis in cultured human cells. *Proc. Natl. Acad. Sci.* **118**, e2008778118 (2021).
12. C. Estell, E. Stamatidou, S. El-Messeiry, A. Hamilton, In situ imaging of mitochondrial translation shows weak correlation with nucleoid DNA intensity and no suppression during mitosis. *J. Cell Sci.* **130**, 4193–4199 (2017).
13. R. Yousefi, E. F. Fornasiero, L. Cyganek, J. Montoya, S. Jakobs, S. O. Rizzoli, P. Rehling, D. Pacheu-Grau, Monitoring mitochondrial translation in living cells. *EMBO Rep.* **22**, e51635 (2021).
14. S. Vidoni, C. Zanna, M. Rugolo, E. Sarzi, G. Lenaers, Why Mitochondria Must Fuse to Maintain Their Genome Integrity. *Antioxid. Redox Signal.* **19**, 379–388 (2013).
15. K. Nakada, K. Inoue, T. Ono, K. Isobe, A. Ogura, Y.-I. Goto, I. Nonaka, J.-I. Hayashi, Inter-mitochondrial complementation: Mitochondria-specific system preventing mice from expression of disease phenotypes by mutant mtDNA. *Nat. Med.* **7**, 934–940 (2001).
16. G. Elachouri, S. Vidoni, C. Zanna, A. Pattyn, H. Boukhaddaoui, K. Gaget, P. Yu-Wai-Man, G. Gasparre, E. Sarzi, C. Delettre, A. Olichon, D. Loiseau, P. Reynier, P. F. Chinnery, A.

Rotig, V. Carelli, C. P. Hamel, M. Rugolo, G. Lenaers, OPA1 links human mitochondrial genome maintenance to mtDNA replication and distribution. *Genome Res.* **21**, 12–20 (2011).
17. A. Dhir, S. Dhir, L. S. Borowski, L. Jimenez, M. Teitell, A. Rötig, Y. J. Crow, G. I. Rice, D. Duffy, C. Tamby, T. Nojima, A. Munnich, M. Schiff, C. R. de Almeida, J. Rehwinkel, A. Dziembowski, R. J. Szczesny, N. J. Proudfoot, Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* **560**, 238–242 (2018).
18. R. J. Szczesny, L. S. Borowski, L. K. Brzezniak, A. Dmochowska, K. Gewartowski, E. Bartnik, P. P. Stepien, Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance. *Nucleic Acids Res.* **38**, 279–298 (2010).
19. M. R. Krieger, M. Abrahamian, K. L. He, S. Atamdede, H. Hakimjavadi, M. Momcilovic, D. Ostrow, S. D. Maggo, Y. P. Tsang, X. Gai, G. F. Chanfreau, D. B. Shackelford, M. A. Teitell, C. M. Koehler, Trafficking of mitochondrial double-stranded RNA from mitochondria to the cytosol. *Life Sci. Alliance* **7** (2024).
20. S. L. van Esveld, R. J. Rodenburg, F. Al-Murshedi, E. Al-Ajmi, S. Al-Zuhaibi, M. A. Huynen, J. N. Spelbrink, Mitochondrial RNA processing defect caused by a mutation in two siblings with a novel neurodegenerative syndrome. *J. Inherit. Metab. Dis.* **45**, 292–307 (2022).
21. R. J. Szczesny, L. S. Borowski, M. Malecki, M. A. Wojcik, P. P. Stepien, P. Golik, RNA Degradation in Yeast and Human Mitochondria. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1819**, 1027–1034 (2012).
22. M. M. Foged, E. Recazens, S. Chollet, M. Lisci, G. E. Allen, B. Zinshteyn, D. Boutguetait, C. Münch, V. K. Mootha, A. A. Jourdain, Cytosolic N6AMT1-dependent translation supports mitochondrial RNA processing. bioRxiv [Preprint] (2024). https://doi.org/10.1101/2024.07.02.601698.
23. C. B. Jackson, A. Marmyleva, R. Awadhpersad, G. Monteuuis, T. Mito, N. Zamboni, T. Tatsuta, A. E. Vincent, L. Wang, T. Langer, C. J. Carroll, A. Suomalainen, De novo serine biosynthesis is protective in mitochondrial disease. bioRxiv [Preprint] (2023). https://doi.org/10.1101/2023.03.23.533952.
24. C. Y. Jao, A. Salic, Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc. Natl. Acad. Sci.* **105**, 15779–15784 (2008).
25. E. Lavdovskaia, E. Hanitsch, A. Linden, M. Pašen, V. Challa, Y. Horokhovskyi, H. P. Roetschke, F. Nadler, L. Welp, E. Steube, M. Heinrichs, M. M.-Q. Mai, H. Urlaub, J. Liepe, R. Richter-Dennerlein, A roadmap for ribosome assembly in human mitochondria. *Nat. Struct. Mol. Biol.*, 1–11 (2024).
26. Y. Kimura, H. Saito, T. Osaki, Y. Ikegami, T. Wakigawa, Y. Ikeuchi, S. Iwasaki, Mito-FUNCAT-FACS reveals cellular heterogeneity in mitochondrial translation. *RNA* **28**, 895–904 (2022).
27. N. A. Bonekamp, B. Peter, H. S. Hillen, A. Felser, T. Bergbrede, A. Choidas, M. Horn, A. Unger, R. Di Lucrezia, I. Atanassov, X. Li, U. Koch, S. Menninger, J. Boros, P. Habenberger, P. Giavalisco, P. Cramer, M. S. Denzel, P. Nussbaumer, B. Klebl, M. Falkenberg, C. M. Gustafsson, N.-G. Larsson, Small-molecule inhibitors of human mitochondrial DNA transcription. *Nature* **588**, 712–716 (2020).
28. B. Westermann, Mitochondrial fusion and fission in cell life and death. *Nat. Rev. Mol. Cell Biol.* **11**, 872–884 (2010).

29. E. Smirnova, L. Griparic, D.-L. Shurland, A. M. van der Bliek, Dynamin-related Protein Drp1 Is Required for Mitochondrial Division in Mammalian Cells. *Mol. Biol. Cell* **12**, 2245–2256 (2001).
30. J. R. Friedman, L. L. Lackner, M. West, J. R. DiBenedetto, J. Nunnari, G. K. Voeltz, ER Tubules Mark Sites of Mitochondrial Division. *Science* **334**, 358–362 (2011).
31. I. Hochberg, L. A. M. Demain, J. Richer, K. Thompson, J. E. Urquhart, A. Rea, W. Pagarkar, A. Rodríguez-Palmero, A. Schlüter, E. Verdura, A. Pujol, P. Quijada-Fraile, A. Amberger, A. J. Deutschmann, S. Demetz, M. Gillespie, I. A. Belyantseva, H. J. McMillan, M. Barzik, G. M. Beaman, R. Motha, K. Y. Ng, J. O'Sullivan, S. G. Williams, S. S. Bhaskar, I. R. Lawrence, E. M. Jenkinson, J. L. Zambonin, Z. Blumenfeld, S. Yalonetsky, S. Oerum, W. Rossmanith, W. W. Yue, J. Zschocke, K. J. Munro, B. J. Battersby, T. B. Friedman, R. W. Taylor, R. T. O'Keefe, W. G. Newman, Bi-allelic variants in the mitochondrial RNase P subunit PRORP cause mitochondrial tRNA processing defects and pleiotropic multisystem presentations. *Am. J. Hum. Genet.* **108**, 2195–2204 (2021).
32. G. Wang, E. Shimada, C. M. Koehler, M. A. Teitell, PNPASE and RNA trafficking into mitochondria. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1819**, 998–1007 (2012).
33. K. McArthur, L. W. Whitehead, J. M. Heddleston, L. Li, B. S. Padman, V. Oorschot, N. D. Geoghegan, S. Chappaz, S. Davidson, H. San Chin, R. M. Lane, M. Dramicanin, T. L. Saunders, C. Sugiana, R. Lessene, L. D. Osellame, T.-L. Chew, G. Dewson, M. Lazarou, G. Ramm, G. Lessene, M. T. Ryan, K. L. Rogers, M. F. van Delft, B. T. Kile, BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science* **359**, eaao6047 (2018).
34. M. Tigano, D. C. Vargas, S. Tremblay-Belzile, Y. Fu, A. Sfeir, Nuclear sensing of breaks in mitochondrial DNA enhances immune surveillance. *Nature* **591**, 477–481 (2021).
35. A. P. West, G. S. Shadel, Mitochondrial DNA in innate immune responses and inflammatory pathology. *Nat. Rev. Immunol.* **17**, 363–375 (2017).
36. L. S. Borowski, A. Dziembowski, M. S. Hejnowicz, P. P. Stepien, R. J. Szczesny, Human mitochondrial RNA decay mediated by PNPase–hSuv3 complex takes place in distinct foci. *Nucleic Acids Res.* **41**, 1223–1240 (2013).
37. T. Rey, S. Zaganelli, E. Cuillery, E. Vartholomaiou, M. Croisier, J.-C. Martinou, S. Manley, Mitochondrial RNA granules are fluid condensates positioned by membrane dynamics. *Nat. Cell Biol.* **22**, 1180–1186 (2020).
38. M. Feric, A. Sarfallah, F. Dar, D. Temiakov, R. V. Pappu, T. Misteli, Mesoscale structure–function relationships in mitochondrial transcriptional condensates. *Proc. Natl. Acad. Sci.* **119**, e2207303119 (2022).
39. Q. Long, Y. Zhou, H. Wu, S. Du, M. Hu, J. Qi, W. Li, J. Guo, Y. Wu, L. Yang, G. Xiang, L. Wang, S. Ye, J. Wen, H. Mao, J. Wang, H. Zhao, W.-Y. Chan, J. Liu, Y. Chen, P. Li, X. Liu, Phase separation drives the self-assembly of mitochondrial nucleoids for transcriptional modulation. *Nat. Struct. Mol. Biol.* **28**, 900–908 (2021).
40. C. Roden, A. S. Gladfelter, RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* **22**, 183–195 (2021).
41. L. Ruan, J. T. McNamara, X. Zhang, A. C.-C. Chang, J. Zhu, Y. Dong, G. Sun, A. Peterson, C. H. Na, R. Li, Solid-phase inclusion as a mechanism for regulating unfolded proteins in the mitochondrial matrix. *Sci. Adv.* **6**, eabc7288 (2020).
42. S. Krishna, R. Arrojo e Drigo, J. S. Capitanio, R. Ramachandra, M. Ellisman, M. W. Hetzer, Identification of long-lived proteins in the mitochondria reveals increased stability

of the electron transport chain. *Dev. Cell* **56**, 2952-2965.e9 (2021).
43. P. A. Sharp, A. K. Chakraborty, J. E. Henninger, R. A. Young, RNA in formation and regulation of transcriptional condensates. *RNA* **28**, 52–57 (2022).
44. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
45. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
46. G. Pekkurnaz, X. Wang, Mitochondrial heterogeneity and homeostasis through the lens of a neuron. *Nat. Metab.* **4**, 802–812 (2022).
47. A. Zbinden, M. Pérez-Berlanga, P. De Rossi, M. Polymenidou, Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. *Dev. Cell* **55**, 45–68 (2020).
48. S. Spannl, M. Tereshchenko, G. J. Mastromarco, S. J. Ihn, H. O. Lee, Biomolecular condensates in neurodegeneration and cancer. *Traffic* **20**, 890–911 (2019).
49. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
50. I. Arganda-Carreras, R. Fernández-González, A. Muñoz-Barrutia, C. Ortiz-De-Solorzano, 3D reconstruction of histological sections: Application to mammary gland tissue. *Microsc. Res. Tech.* **73**, 1019–1029 (2010).
51. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Chapter 3: Uncovering Novel Mitovirus Sequences in Public Sequencing Databases

Portions of this chapter were published as:

Metatranscriptomic analysis uncovers prevalent viral ORFs compatible with mitochondrial translation

Adam Begeman[1], Artem Babaian[2,3], Samantha C Lewis[1,4,5,6*]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA USA
[2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
[3]Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada
[4]Innovative Genomics Institute, Berkeley, CA, USA
[5]Helen Wills Neuroscience Institute, Berkeley, CA USA
[6]Department of Nutritional Sciences and Toxicology, University of California, Berkeley, CA USA

*Corresponding author: samlewis@berkeley.edu

Citation:

Begeman A, Babaian A, Lewis SC. (2023). Metatranscriptomic analysis uncovers prevalent viral ORFs compatible with mitochondrial translation. mSystems8:e01002-22.https://doi.org/10.1128/msystems.01002-22

Abstract

RNA viruses are ubiquitous components of the global virosphere, yet relatively little is known about their genetic diversity or the cellular mechanisms by which they exploit the biology of their diverse eukaryotic hosts. A hallmark of (+)ssRNA (positive single-stranded RNA) viruses is the ability to remodel host endomembranes for their own replication. However, the subcellular interplay between RNA viruses and host organelles that harbor gene expression systems, such as mitochondria, is complex and poorly understood. Here we report the discovery of 763 new virus sequences belonging to the family *Mitoviridae* by metatranscriptomic analysis, the identification of previously uncharacterized mitovirus clades, and a putative new viral class. With this expanded understanding of the diversity of mitovirus and encoded RNA-dependent RNA polymerases (RdRps), we annotate mitovirus-specific protein motifs and identify hallmarks of mitochondrial translation, including mitochondrion-specific codons. This study expands the known diversity of mitochondrial viruses and provides additional evidence that they co-opt mitochondrial biology for their survival.

Importance

Metatranscriptomic studies have rapidly expanded the cadre of known RNA viruses, yet our understanding of how these viruses navigate the cytoplasmic milieu of their hosts to survive remains poorly characterized. In this study, we identify and assemble 763 new viral sequences belonging to the *Mitoviridae*, a family of (+)ssRNA viruses thought to interact with and remodel host mitochondria. We exploit this genetic diversity to identify new clades of *Mitoviridae*, annotate clade-specific sequence motifs that distinguish the mitoviral RdRp, and reveal patterns of RdRp codon usage consistent with translation on host cell mitoribosomes. These results serve as a foundation for understanding how mitoviruses co-opt mitochondrial biology for their proliferation.

Introduction

RNA viruses are ubiquitous and prevalent components of the eukaryotic virosphere. However, the genetic diversity of eukaryotic RNA viruses is poorly described, due in part to sparse sampling and biases towards pathogens that impact human health or commercial agriculture. Metagenomics is a powerful approach to characterize viral biodiversity and has been used to substantially expand the known set of RNA viruses, particularly unculturable viruses from polar, marine, or microbiome contexts (*1*–*5*). However, there are still substantial gaps in our knowledge of viral biodiversity, particularly the holobionts of fungi and metazoans. Positive single-stranded RNA viruses, or (+)ssRNA, in particular, exhibit a distinct ability to restructure the cytoplasm of host cells to facilitate their propagation. They remodel host endomembranes into invaginated viral replication organelles (ROs) derived from the

endoplasmic reticulum, Golgi, or the outer mitochondrial or plastid membranes (*6–10*). Viral ROs concentrate viral RNA, proteins, and host factors to facilitate the formation of viral replication complexes (VRCs), as well as shield viral genome replication from host antiviral-sensing mechanisms (*11*, *12*).

Beyond RO, some viruses have the ability to enter membrane-bound host organelles to directly access their biosynthetic potential, promote the formation of organelle-derived replication vesicles, or sequester organelle-localized proteins in the cytoplasm for their own benefit (*9*, *13*, *14*). Of the organelles linked to (+)ssRNA viral replication, mitochondria and chloroplasts maintain distinct chromosomes and gene expression machinery for the transcription, processing, and translation of organellar genes, including dedicated organellar ribosomes (*15*, *16*). These features endow them with a unique capacity for nucleic acid metabolism and protein production (*15–18*). Mitochondria in particular are attractive targets for viral replication given their central roles in host defense mechanisms; the cloister of viral nucleic acid inside the mitochondrion itself may provide an opportunity to evade immune activation pathways at the outer mitochondrial membrane surface (*19–21*).

Metatranscriptomic sequencing studies are rapidly expanding the catalog of Earth's RNA virome, yet our grasp of the strategies employed by viral proteins and RNA to hijack host organelles lags. Of particular interest are the viral families that have reported association with the mitochondria. Mitochondria and plastids are interesting targets for viral proliferation due to their dedicated gene expression pathway and incredibly dynamic membrane restructuring machinery. There have been recent advances giving insight into organelle replication and inheritance of nucleic acids, but few have investigated how these systems may be co-opted or perturbed by viral pathogens. We sought to gain a better understanding of the viral species that may interact with mitochondrial biology. As a first pass, we sought to exploit the unique endosymbiosis of mitochondria and their unique gene translation machinery.

We focused our study on the *Mitoviridae*, a family of (+)ssRNA viruses identified in association with fungal, plant, or invertebrate host mitochondria (*17–21*). Mitoviral genomes are composed of a capsidless single RNA with one open reading frame (ORF) encoding an RNA-dependent RNA polymerase (RdRp) (see Figure 3.3A) (*22*). The presence of Mitovirus is linked to mitochondrial proteome remodeling in fungi and plants (*23–25*). While a handful of species (and many genomic fragments) have been identified in environmental samples, the mitoviral replicative cycle and the extent to which these viruses interact with host mitochondria in vivo remain poorly understood (*23*, *26*, *27*). The distribution and prevalence of *Mitoviridae* among Earth ecosystems are also unclear.

To determine the extent of mitoviral diversity and their evolutionary history, we searched metatranscriptomic data sets for evidence of mitochondria-associated (+)ssRNA viruses. In

contrast to previous studies which focused on outer mitochondrial membrane remodeling by RNA viruses, we sought to identify viral genomes and/or genomic fragments that may co-opt the biosynthetic potential of mitochondrial matrix contents, such as the mitochondrial translational machinery, as a means of propagation.

Materials and Methods

Searching SRA with Serratus

Lenarviricota protein query sequences

Mitovirus and other *Lenarviricota* nucleotide sequences were downloaded from GenBank with queries "txid186768[Organism:exp]" (N = 2,364, date: 12 October 2020) and ": txid2732407[Organism:exp] NOT txid186768[Organism:exp]" (N = 4,878, date: 12 October 2020), respectively, and hypothetical coding sequences were removed. To search short-nucleotide reads for mitoviral RdRP with a translated-nucleotide search using a standard genetic code, query CDS sequences were translated into amino acids using the standard genetic code to enable stop codon read-through (*transeq –Table 0*, EMBOSS 6.6.0).

Sequence Read Archive search space

Sequence Read Archive (SRA) sequencing runs were accessed from the SRA website using the search term: *'"VIRAL METAGENOME" OR "VIROME" OR "VIROMIC" OR "VIRAL RNA" OR "METATRANSCRIPTOMIC" NOT "METAGENOMIC" NOT amplicon[All Fields] AND "platform illumina"[Properties] AND cluster_public[prop]'* on 25 October 2020, returning 60,327 runs, which were randomly sampled to 1,000 runs (See Table S1 at https://doi.org/10.1128/msystems.01002-22).

Serratus search

Short-read sequencing runs were aligned against the above *Lenarviricota* protein query using the Serratus cloud-computing architecture (v0.2.0) (*5*) in protein mode. Reference architecture was 300 downloads (r5.xlarge) instances, 500 align (c5n.xlarge) instances, and 20 merges (c5.xlarge) instances. Translated-nucleotide search mode was run with DIAMOND (version 2.0.1) and parameters "*—unal 0 k 1 –b 0.2*". Processing of all sequencing runs was attempted at least twice, and 981/1,000 (98.1%) were completed.

Mitovirus discovery pipeline

Contig assembly

Selected SRA data sets from Serratus search were downloaded, unpacked, and all paired reads split using the SRA toolkit program fasterq-dump (See Table S1 at https://doi.org/10.1128/msystems.01002-22). Sequencing reads were checked for sequencing barcodes using Trimmomatic sequencing adapter library (*28*) and trimmed accordingly. Contigs for each SRA experiment were then assembled using SPADES v.3.14.1 in RNA mode with

default options (*29*, *30*). For SRA experiment, ERR2195693 and ERR2809108 normal SPADES were used with default options due to both experiments containing unpaired reads.

Contig identification

The assembled contigs for each SRA experiment were then independently aligned with a reference set of viral RdRps derived from the NCBI protein database (see above) using the BLASTX functionality in Diamond v.2.0.6 (*31*), using the "--sensitive" tag, searching in all six frames, and requiring a minimum ORF of 300 amino acids. The contigs derived from each SRA project were searched using the vertebrate mitochondrial codon table (NCBI code 2), the fungal mitochondrial codon table (NCBI code 4), and the invertebrate mitochondrial codon table (NCBI code 5). Contigs that aligned against the mitovirus reference RdRps were pulled out using the AlignmentBreakup.py python file (https://github.com/TheLewisLab/Mitovirus-Code).

Mitovirus RdRp identification and confirmation

ORFs were identified using NCBI OrfFinder v.0.4.3 allowing for alternative start codons for codon tables 2 and 5due to the wide species diversity of mitochondrial start codons for those tables. Sequences with more than one ORF or ORFs less than 300 amino acids long were discarded. The resulting putative RdRps were then identified using BLASTP against the entire NCBI non-redundant protein database (accessed 23 February 2021). For sequences that aligned in multiple codon tables, fungal mitochondrial codon table 4 was used as mitoviruses are thought to mainly infect fungal hosts. The top BLASTP hit by E-value was then used to make the final taxonomy assignment for each putative viral sequence. While assembling the list of reference mitovirus sequences, we noticed a number of reference mitovirus sequences that had the highest percent identity to viruses other than mitoviruses and corrected them for the purpose of this study (See Table S2 at https://doi.org/10.1128/msystems.01002-22).

Read mapping and viral abundance

To estimate the abundance of each mitovirus in the metagenomic sample, sequencing reads were mapped back onto the assembled mitovirus contigs using Bowtie2 (*32*). A bowtie2 index was generated for all mitovirus genome segments found in each SRA sample, and bowtie2 v.2.4.5 was then used to map the original reads back onto the assembled genomes using default settings. The abundance of each mitovirus genome was calculated as mapped reads per kilobase per million total reads in the SRA experiment independent of any quality metrics or filtering. Estimates of viral abundance are reported in Table S2 at https://doi.org/10.1128/msystems.01002-22.

RdRp clustering and phylogenetic analysis

RdRp amino acid sequences from this study and the RdRp amino acid sequences from NCBI-reported mitoviruses and their closest evolutionary neighbors, narnavirus, levivirus, and ourmiavirus, were compared by the Enzyme Function Initiative-Enzyme Similarity Tool (EFI-

EST) (*33*, *34*) to generate protein similarity networks with an E-value cutoff of 1×10^{-5} for class-level classification and of 1×10^{-60} for family-level classification consistent with previous studies (*4*, *35*, *36*). Reference sequences were downloaded from the NCBI protein database in May 2021 excluding partial protein sequences of fewer than 300 amino acids. RdRp clustering was represented using the Cytoscape organic layout (*37*). Sequences from the mitovirus cluster as well as from 10 randomly selected representative sequences from each clade noted above were then extracted and aligned using Clustal Omega multiple sequence alignment (*38*). The resulting alignment was then used to build a phylogenetic tree using FastTree v.2.1.11 with automatic determination of the substitution model and 1,000 ultrafast bootstraps and JTT+CAT model (*39*). The resulting tree was visualized using iTOL with clade branch lengths less than 0.9 collapsed, rooted at the most recent evolutionary neighbor between mitovirus sequences and the outgroups mentioned above (*40*).

Codon usage analysis

Codon usage was calculated using the Python file codonfrequencyanalysis.py (GITHUB) using mitochondrial coding sequences in the NCBI reference sequence database using a mitochondrion and each organism filter tag. The prokaryotic viruses and their viral host coding sequences were obtained from the NCBI nucleotide database using corresponding organism tag and filtering for complete coding sequences. Nuclear codon usage was obtained using the HIVE Codon usage table search engine (*41*). Putative mitovirus codon usage correlation against the reference codon usage was calculated using Pearson's R2 linear correlation formula as provided by the scipy python package. Heatmaps were created using the native R v.4.1.2 heatmap function with row hierarchical clustering based on the codon usage correlation values.

Protein motif discovery

Protein motifs were identified using the de novo online motif discovery platform MEME with default parameters (*42*). The mitovirus-specific protein motifs were defined using the discriminative discovery mode in which the closest evolutionary neighbor viral RdRps from narnaviruses, ourmiaviruses, and leviviruses were designated as the outgroup (*42*). The percent occupancy at each motif location and the representative sequence alignment with evolutionary neighbors were visualized using clustal omega multiple protein alignment and Jalview (*38*, *43*).

Alphafold structural prediction

Alphafold mitovirus RdRp structural prediction of representative putative mitovirus RdRp sequence ERR3412979_288_4 was generated using the Google Collaboratory notebook distribution of AlphaFold and ColabFold with default parameters (*44*, *45*). The mitovirus structural prediction and the location of protein motifs were visualized using PYMOL (*46*).

RNA structural prediction

Putative mitoviral RNA structures were predicted using the RNAfold stand-alone binaries with minimum free-energy calculations on the last 100 nucleotides of the 3′ end of each mitovirus whose ORF contained a "stop" codon (*47*). Predicted RNA structures were then analyzed using MEME motif enrichment with a non-standard alphabet corresponding to the dot-bracket notation of the predicted RNA structures (*42*, *47*).

Results

Discovery of novel Mitovirus spp. and genomic fragments

To gain a comprehensive understanding of the origins and evolution of mitoviruses, we used the Serratus viral discovery platform (*5*) to search sequencing reads from public metagenomic sequencing data sets hosted on the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) for mitoviral genomic sequences (see SRA Runs Searched; Table S1 at https://doi.org/10.1128/msystems.01002-22; Figure 2.1) and identified 763 putative mitoviral genomes or genomic fragments of 900 base pairs or longer, sufficient to encode the mitoviral RdRp (Figure 3.2). We used FastTree to perform phylogenetic analyses of these candidate mitoviruses, comparing them to previously reported mitoviral genomic fragments, using representative ourmiavirus, narnavirus, and levivirus sequences as outgroups (Figure 3.3B). These evolutionary neighbors of *Mitoviridae* infect plants (ourmiaviruses), fungi (naraviruses), or bacteria (leviviruses) (*22*).

We scanned each assembled sequence contig for ORFs in all six translational contexts using the fungal, invertebrate, or vertebrate mitochondrial codon tables for ORFs that aligned with the amino acid sequence of a reference mitoviral RdRp using BLASTP. We assigned positive hits to taxonomy on the basis of BLASTP hits against the NCBI non-redundant protein database. Due to the non-traditional codon tables, existing viral discovery tools were not applicable.

While all candidates aligned with a previously reported mitoviral RdRp, the majority shared less than 45% sequence identity to a reference mitovirus (See Table S2 at https://doi.org/10.1128/msystems.01002-22). A subset of divergent sequences was phylogenetically clustered; indeed, we uncovered two previously undescribed and phylogenetically distinct clades of sufficient divergence to constitute new viral families. Despite this genetic diversity, AT content among the mitoviral candidates was significantly higher than in ourmiaviruses, narnaviruses, or leviviruses, consistent with the reported reference mitoviral genomes and the mitochondrial genome content of potential eukaryotic hosts from these SRA projects (Figure 3.3C). Overall this search expanded known mitovirus diversity by approximately 50% (Figure 3.3D) (*48*, *49*). Interestingly, mitoviral sequences associated with metazoan hosts cells tended to group in the phylogeny, potentially consistent with expansion into naive host niches. These findings imply that *Mitoviridae* are both prevalent and abundant within

the global RNA virome and that substantial uncharacterized genetic diversity exists within this group.

Evolutionary relationships among mitoviral clades

We next examined the potential sequence and function space of the mitoviral RdRp sequences we identified using sequence similarity networks (SSNs). SSNs are useful tools to study the relationships between large sets of protein sequences that may be hard to root in traditional phylogenetic trees when there exist large gaps in characterized genetic diversity among sampled sequences. For SSN analysis, we extracted all RdRps from the phylum *Lenarviricota* which is made up of the family *Mitoviridae* and their closest evolutionary neighbors, *Narnaviridae*, *Leviviridae*, and *Ourmiaviridae* from the NCBI protein database, and aligned them pairwise using EFI-EST (*33*). The resulting alignments were then used to build SSNs using EFI-EST, implemented, and visualized in Cytoscape (*33*, *37*). Furthermore, we noted a distinct cluster formed exclusively by sequences without clear homologs in the NCBI-nr database, suggesting a novel phylogenetic group (Figure 3.4A).

We next compared these results to an SSN constructed using an E-value cutoff of 1×10^{-60}, which has been used to split viruses into family-level assignments (*4*, *35*, *36*) (Figure 3.4B). We identified 17 as yet uncharacterized family-level clusters. This analysis also suggested that the family *Mitoviridae* may consist of two major clades that are well represented in known mitovirus diversity (Figure 3.3B and 3.4B). The results from the phylogenetic classification and SSN analysis suggest there exist a large number of previously unidentified family-level clusters that make up the family *Mitoviridae*, many of which seem to be associated with metazoan hosts.

Discovery of conserved structural motifs in the mitoviral RdRp

Viral RdRps can generally be identified by five evolutionarily conserved structural motifs located in the core of the RdRp (*34*, *38*). These protein motifs play a central role in catalysis and therefore retain high structural similarity across all five Baltimore groups of viral RdRps (*34*, *38*). To characterize the conserved catalytic domains within mitoviral RdRps, as well as interrogate unique structural motifs, we used the sequence motif discovery platform MEME (*50*). We searched all 763 newly identified mitoviral peptide sequences for conserved domains and successfully identified the five conserved general RdRp catalytic motifs (Figure 3.5C). These motifs retained all catalytically required amino acids such as the characteristic DX2-5D and GDD motifs and were spatially organized consistent with previously identified viral RdRp domains (Figure 3.5A and 3.5C).

Using the Google Research Colaboratory distribution of AlphaFold and ColabFold (*51*, *52*), we were then able to get a structural prediction of a representative mitovirus RdRp (Figure 3.5B and 3.6). By mapping the conserved general RdRp catalytic motifs onto the structural prediction, we observed that all motifs fall within the catalytic pocket of the RdRp, with the

highly conserved acids pointing inward (Figure 3.6B), suggesting that mitoviral RdRps are indeed catalytically active.

Next, we sought to identify structural motifs unique to the *Mitoviridae* family. To do so, we used the discriminatory mode of MEME to search for conserved protein motifs enriched in the mitoviral RdRp sequence set relative to their closest evolutionary neighbor, the *Narnaviridae* (*19*). We uncovered five unique highly conserved structural motifs near the core of the RdRp (Figure 3.5A and 3.5D). To confirm that these sequence motifs are unique to the *Mitoviridae* family, we performed simple enrichment analysis in the MEME suite (*50*), in which the mitovirus evolutionary neighbors, narnaviruses, leviviruses, and ourmaiviruses, are searched for occurrences of the mitovirus-specific motifs. This analysis revealed no significant matches, indicating these protein motifs are truly unique to mitovirus RdRps. A representative multiple sequence alignment between these four groups is shown in Figure 3.7 . Mapping these motifs onto the ColabFold structural prediction revealed they are located in the core of the RdRp (Figure 3.5B and 3.6). Mitovirus-specific motifs 1, 2, and 3 are all facing the interior of the RdRp core with highly conserved amino acids indicating a possible role in catalysis (Figure 3.6C). In contrast, motifs 4 and 5 are solvent exposed on the exterior of the RdRp, consistent with a possible role in cofactor recruitment (Figure 3.6D and 3.6E).

Evidence for mitoribosomal translation of the mitoviral RNA-dependent RNA polymerase

Mitoribosomes function in the mitochondrial matrix and employ a codon table distinct from their nuclear counterpart (*15*). These non-canonical codons are decoded using tRNAs encoded in the mitochondrial genome (*15*, *53*). While the exact encoding of amino acids varies from species to species, one hallmark of mitochondrial translation is the decoding of the UGA trinucleotide as tryptophan instead of a “stop” codon (*15*, *53*). This distinction can be leveraged to identify proteins that are truncated when translated on cytoplasmic ribosomes at full length when translated in the mitochondrial compartment (Figure 3.8A). We wondered to what extent mitoviruses use the mitochondrial-specific codon and its effect on the mitovirus RdRp. We found that roughly 55% of previously reported mitovirus and 48% of our new putative mitovirus sequences decoded UGA as tryptophan instead of “stop” with the majority encoding this tRNA multiple times (See Table S3 at https://doi.org/10.1128/msystems.01002-22; Figure 3.8B). We estimate that this decoding is required for the expression of the full-length RdRp (Figure 3.8C and 3.8D).

The encoding of amino acids between all synonymous triplicate nucleotide codons is not equal, and this codon usage bias (CUB) can often be useful in tracking evolutionary history, differential gene expression, and even virus–host interactions (*50*, *54–58*), making CUB potentially useful to restrict the breadth of potential hosts (*51*, *54*, *57*) (Figure 3.9). We probed whether codon usage among mitoviral sequences was more similar to that of mitochondrially encoded or nuclear-encoded gene products. To test this, we calculated the Pearson’s correlation

coefficients between the codon usage frequency of each putative mitoviral RdRp or mitoviral RdRp fragment and a range of different host mitochondrial and nuclear codon usage frequencies as defined in Table S4 at https://doi.org/10.1128/msystems.01002-22.

We found that the codon usage frequencies of putative novel mitoviral RdRps and RdRp fragments correlate significantly more ($P < 0.0001$) with that of the fungal mitochondria than the nuclear transcripts of its hosts (Figure 3.10A and 3.10C; Figure 3.11A). The majority of identified mitoviruses have been identified in association with fungal hosts. Of the previously identified 46 reference mitovirus sequences, 45.6% were discovered in the fungal Ascomycota phylum, 19.6% in the Basidiomycota phylum, 17.4% in Mucoromycota phylum, 4.3% in plants, and 13.0% assembled from non-specific metagenomic data. Consistently, the majority of the mitovirus sequences that we report here display a pattern of codon usage most similar to fungi and specifically with the fungal mitochondrial codon usage table. We also here report a set of mitovirus sequences that when translated display a high codon usage correlation with the invertebrate mitochondrial table.

In contrast, codon usage among narnaviral RdRps, which exclusively replicate within the cytoplasm of the host cell, was significantly more correlated ($P < 0.0001$) with nuclear codon usage than it was with mitochondrial codon usage (Figure 3.9A or Figure 3.11A). Interestingly, codon usage in the mitoviral RdRps also showed a strong correlation with both the plant mitochondrial and nuclear codon frequencies (Figure 3.10C; Figure 3.11A). This could suggest that mitovirus RdRps are poised to replicate both in the mitochondria and cytoplasm of the plant host. However, plant mitochondrial codon usage is extremely similar to both fungal mitochondrial codon usage and plant nuclear codon usage (Figure 3.10B; Figure 3.11C), suggesting that this high codon usage correlation could also be indicative of this close association. We also identified a subset of RdRps that showed a significant correlation with the metazoan mitochondrial codon usage table ($P < 0.0001$) (Figure 3.10C; Figure 3.11B). Interestingly, this subset is also enriched for mitovirus sequences assembled from animal sequencing projects (Figure 3.10C in blue). This supports recent reports identifying mitoviruses in invertebrate metagenomic samples (*52*). Taken together, these findings strongly suggest that mitovirus RdRps utilize the mitochondrial ribosomes and their unique subset of tRNAs for mitoviral translation.

Discussion

Inspired by the recent success in searching metagenomic sequencing data sets for novel virus species (*1–3*, *59*, *60*), here we identified 763 new putative mitovirus sequences and sequence fragments from publicly available metagenomic profiles of samples isolated from a wide array of geographic locales and ecological environments (Figure 3.2). This study increases the number of known mitovirus sequences, with an approximate 50% increase in diversity. Our

findings underscore the scarcity of knowledge about eukaryotic ssRNA viruses and the understudied *Mitoviridae* family in particular. Furthermore, this study serves as an initial foray into exploring how mitoviruses may exploit the unique organelle biology of host cells for their propagation.

Here, we expand the understanding of the evolutionary relationships among ssRNA viruses and the relationships between new and previously identified mitoviruses. A number of our newly identified mitoviruses cluster into distinct family-level organizations, suggesting a broader genetic diversity amongst the clade than had been appreciated in previous studies (*22*). Indeed, SSN and phylogenetic analyses indicated that the family *Mitoviridae* actually consists of two distinct major clades, with evidence suggesting many other underrepresented family-level clusters (Figure 3.3B and Figure 3.4B).

Previous reports on mitoviruses tend to rely primarily on crude mitochondrial fractionations as evidence for their mitochondrial localization and do not show a functional relationship between the mitovirus and the mitochondrial gene expression systems (*26*, *27*). It has also been shown mitoviruses are able to horizontally transfer between fungal species by presumable mitochondrial fusion during protoplast fusion events (*61*, *62*). However, there still lacks a direct link between the mitovirus life cycle and mitochondrial biology. Here, through codon usage correlations and mitochondrial codon analysis, we provide evidence linking mitoviruses to the mitochondrial gene expression systems. Our data suggest that not only do the majority of described mitoviruses rely on mitochondrial ribosomes for RdRp translation but mitoviral codon usage parallels that of the host cell, suggesting an evolutionary adaptation to hijack the mitochondrial gene expression system (Figure 3.8 and 3.10).

It is well documented that (+)ssRNA virus remodel endogenous host membranes to form ROs and recruit necessary host factors that make up their VRC (*6*, *7*, *9–11*). Interfacing with host factors requires functional structural motifs within the viral RdRp to facilitate membrane remodeling and protein recruitment (*63*). The field currently lacks robust cytological data as to the subcellular localization of mitoviruses within intact host cells, and therefore knowledge about how they accomplish this process is limited. Given their unique association with mitochondria, we wondered if *Mitoviridae* may utilize specific protein motifs or structures that make them distinct from all other previously characterized RNA viruses that replicate in the cytoplasm. Here, we identify five previously undescribed evolutionary conserved protein motifs unique to the polymerases of family *Mitoviridae*, likely implicated in that process (Figure 3.5). Structural predictions of a typical mitovirus RdRp suggest that two of these identified protein motifs map to the surface of the folded protein, rendering them accessible to host cofactors or protein recognition domains (Figure 3.6D and 3.6E). Beyond their potential role in catalyst or host factor recruitment, these novel mitovirus-specific motifs will also serve as useful for future studies

looking to identify new mitoviruses. While outside the scope of this study, future experiments addressing the function of these conserved mitovirus-specific domains would be of great interest.

Beyond their interest as a virus that may exploit mitochondrial-specific gene expression equipment, mitoviruses also represent a potentially exciting development for efforts towards mitochondrial transgenesis. A notable barrier to extant methods for genome engineering in the mitochondria has been the lack of tools to introduce endogenous nucleic acid into the mitochondrial matrix (*64*, *65*). Just as the discovery of plasmids allowed for exogenous gene expression, and the characterization of the yeast 2-μm plasmid supported the expansion of yeast genetic editing, a better understanding of mitovirus biology may yield fruitful insights for manipulating organellar nucleic acids.

Through searching publicly available sequencing data, we have expanded the known mitovirus diversity and shed some light on their interactions with mitochondria. While there is much to still uncover surrounding the molecular and cell biology of mitoviruses, this study serves as the first foray into an understudied world.

Acknowledgements

The authors thank Dr. John Smolka, Dr. Iswar Hariharan, Dr. Samuel Diaz-Munoz, and Dr. Jason Stajich for helpful discussions. We are grateful to the biology community for data sharing.

This work was supported by the Shurl and Kay Curci Foundation, National Institutes of Health grants 5T32GM007232-38 and R00GM129456 to Samantha C. Lewis, and a National Science Foundation Graduate Research Fellowship to Adam Begeman. Computing resources were provided by the University of British Columbia Community Health and Wellbeing Cloud Innovation Centre, powered by AWS. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Figures

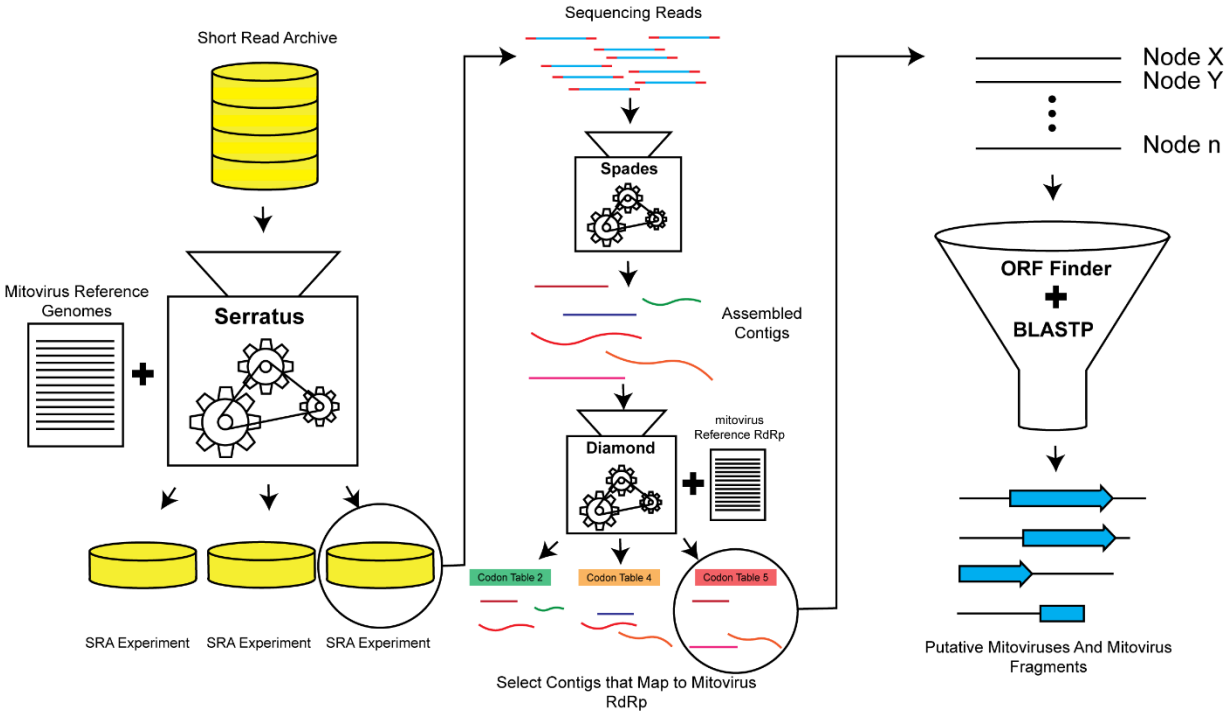


Figure 3.1 Schematic of computational pipeline used to discover putative new mitovirus sequences and sequence fragments.

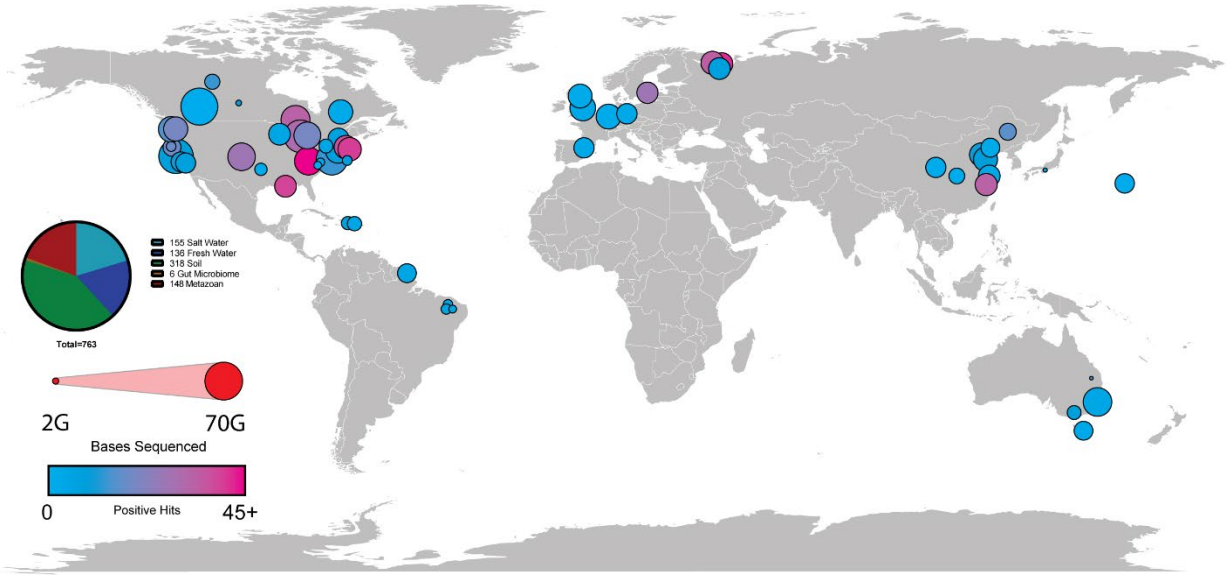


Figure 3.2 Diversity of sequencing sample collection sites. Map of all sample collection sites that resulted in the identification of new putative mitovirus. The size of each circle corresponds to the total number of bases sequenced at each site while the color reflects the number of new putative mitoviruses to come from that sample. A large amount of North American and Western European samples is a product of acquisition bias for samples in these regions.

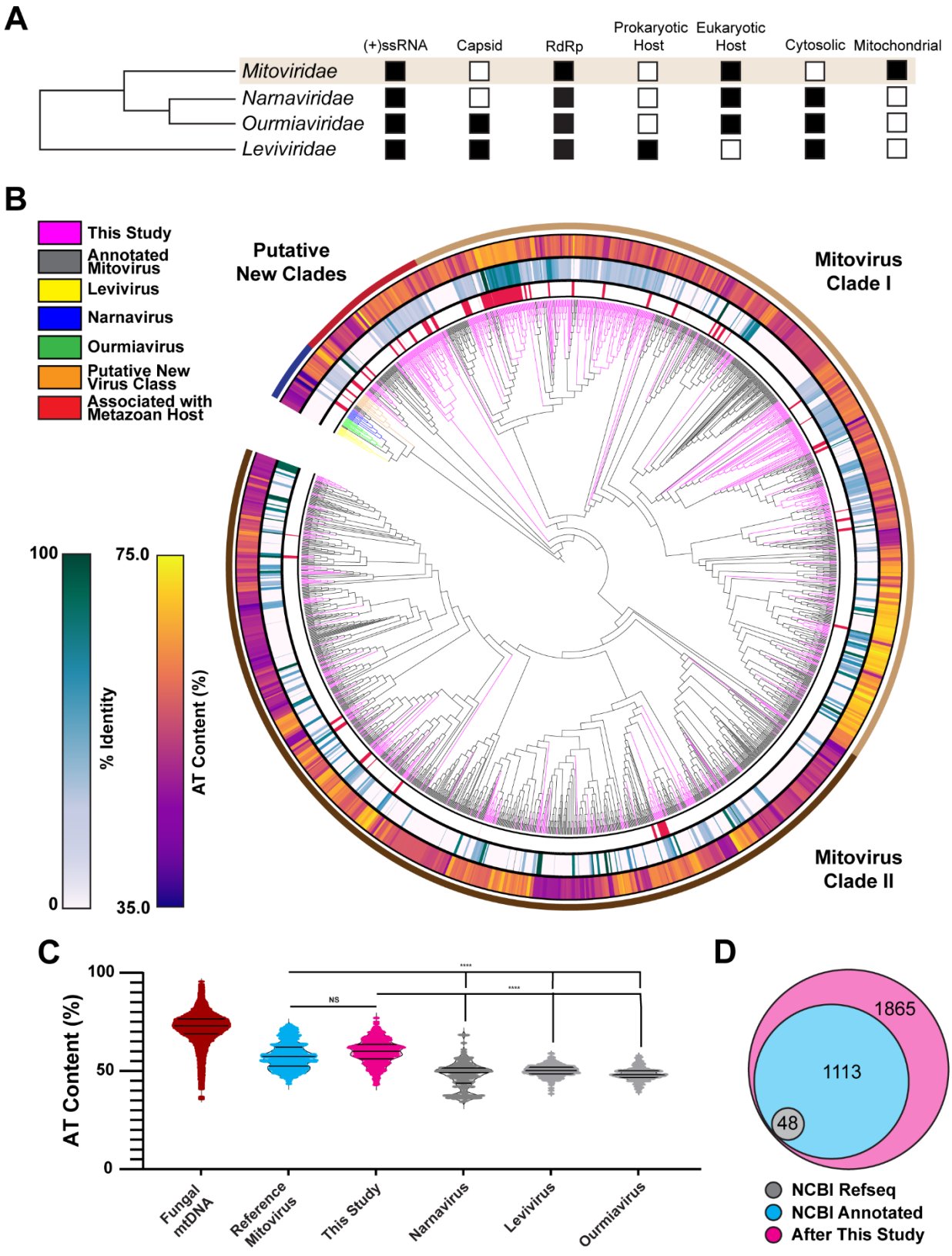

A
(+)ssRNA
Capsid
RdRp
Prokaryotic Host
Eukaryotic Host
Cytosolic
Mitochondrial
Mitoviridae
Narnaviridae
Ourmiaviridae
Leviviridae
B
This Study
Annotated Mitovirus
Levivirus
Narnavirus
Ourmiavirus
Putative New Virus Class
Associated with Metazoan Host
Putative New Clades
Mitovirus Clade I
Mitovirus Clade II
100
0
% Identity
75.0
35.0
AT Content (%)
C
AT Content (%)
100
50
0
NS
****
****
Fungal mtDNA
Reference Mitovirus
This Study
Narnavirus
Levivirus
Ourmiavirus
D
1865
1113
48
NCBI Refseq
NCBI Annotated
After This Study

Figure 3.3 Discovery and characterization of novel putative mitovirus sequences. (A) Characteristics of *Mitoviridae* and its closest evolutionary neighbors: *Narnaviridae*, *Ourmiaviridae*, and *Leviviridae*. (B) Phylogenetic tree of new putative mitovirus sequences (magenta), existing mitovirus sequences as annotated by NCBI (black), a putative new viral clade (orange), and 10 representative sequences from each of the mitovirus's closest evolutionary neighbors, the narnaviruses (blue), ourmiaviruses (green), and leviviruses (yellow). The AT content of each sequence, the percent identity to known mitovirus sequences, and whether the sequences were assembled from an animal sequencing project are represented by the concentric rings. (C) AT content of the new putative mitovirus sequences, reference mitoviruses, the mitovirus's closest evolutionary neighbors, and fungal mitochondrial DNA. Statistical tests run: Mann–Whitney test, **** corresponds to P-value < 0.0001. (D) Number of mitovirus sequences before and after this study.

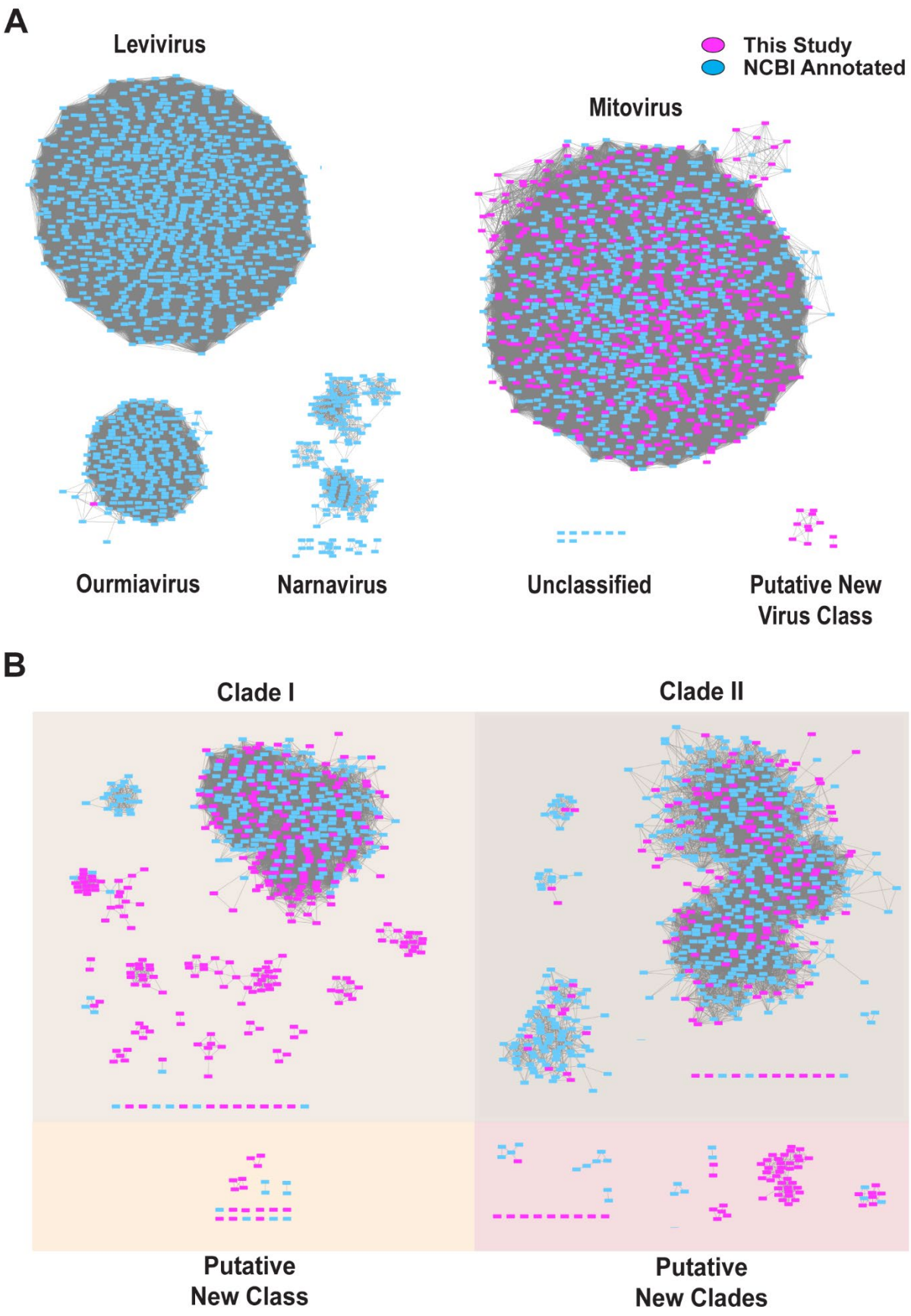

A
Levivirus
This Study
NCBI Annotated
Mitovirus
Ourmiavirus
Narnavirus
Unclassified
Putative New
Virus Class
B
Clade I
Clade II
Putative
New Class
Putative
New Clades

Figure 3.4 Sequence similarity networks of reference and putative new mitovirus RdRps. (A) Clade-level sequence similarity network of reference mitovirus and closest evolutionary neighbor RdRps (cyan), and putative new mitovirus RdRps (magenta) generated using EFI-EST with an E-value cutoff $1 \times 10{-}5$. (B) Family-level sequence similarity network of just mitovirus sequences generated using EFI-EST with an E-value cutoff of $1 \times 10{-}60$.

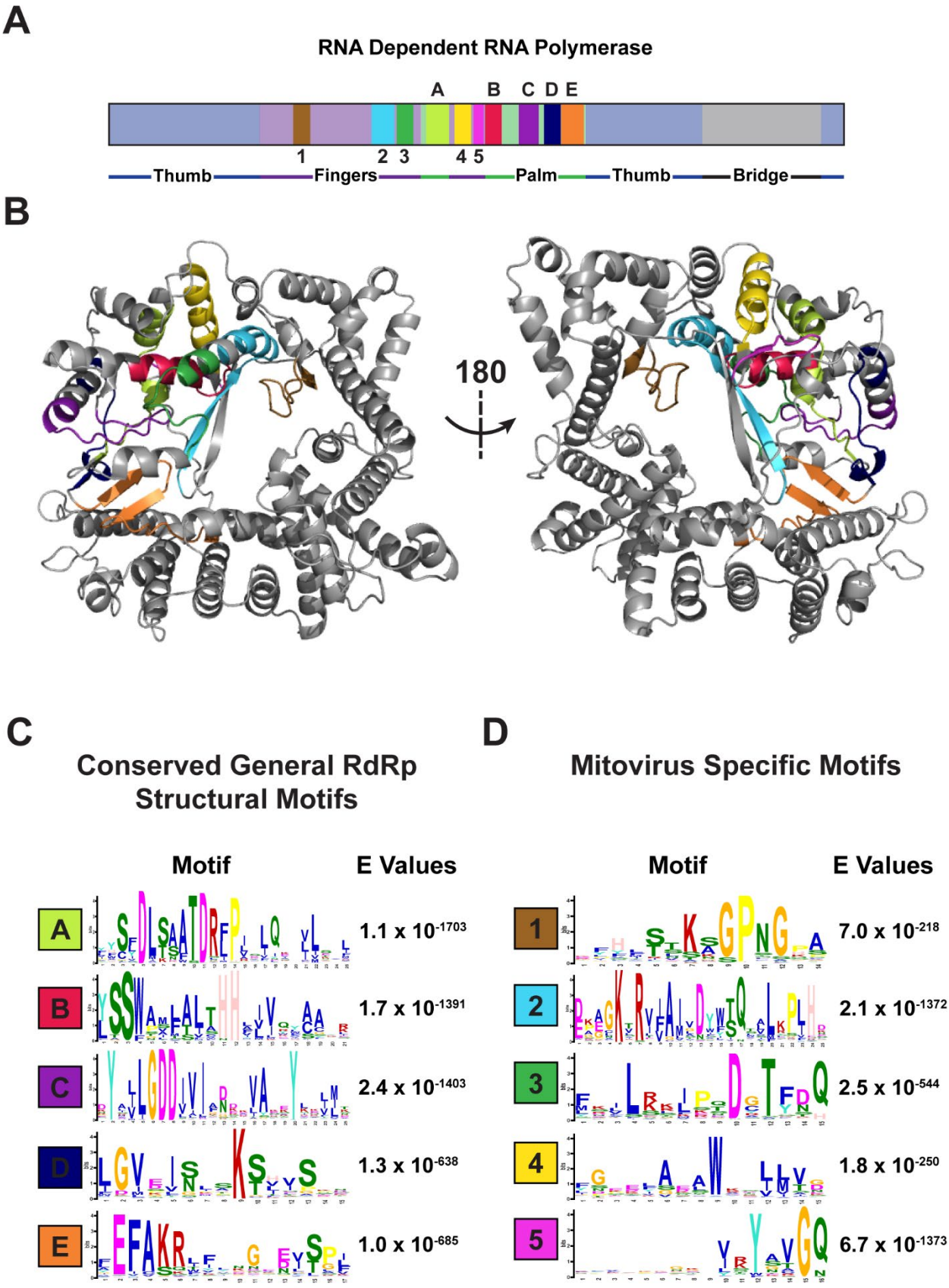

A
RNA Dependent RNA Polymerase
A B C D E
1 2 3 4 5
Thumb
Fingers
Palm
Thumb
Bridge
B
180
C
Conserved General RdRp Structural Motifs
Motif
E Values
A
1.1 x 10-1703
B
1.7 x 10-1391
C
2.4 x 10-1403
D
1.3 x 10-638
E
1.0 x 10-685
D
Mitovirus Specific Motifs
Motif
E Values
1
7.0 x 10-218
2
2.1 x 10-1372
3
2.5 x 10-544
4
1.8 x 10-250
5
6.7 x 10-1373

Figure 3.5 Mitovirus-conserved protein motifs. (A) Location of each identified protein motif relative to known RNA-dependent RNA polymerase domains. (B) Alphafold structural prediction and corresponding motif location on representative mitovirus RdRp. (C) Conserved general RdRp catalytic motifs discovered using standard MEME motif discovery platform *(42)*. (D) Conserved mitovirus unique protein motifs as reported by MEME discriminatory mode using narnavirus RdRps as outgroup.

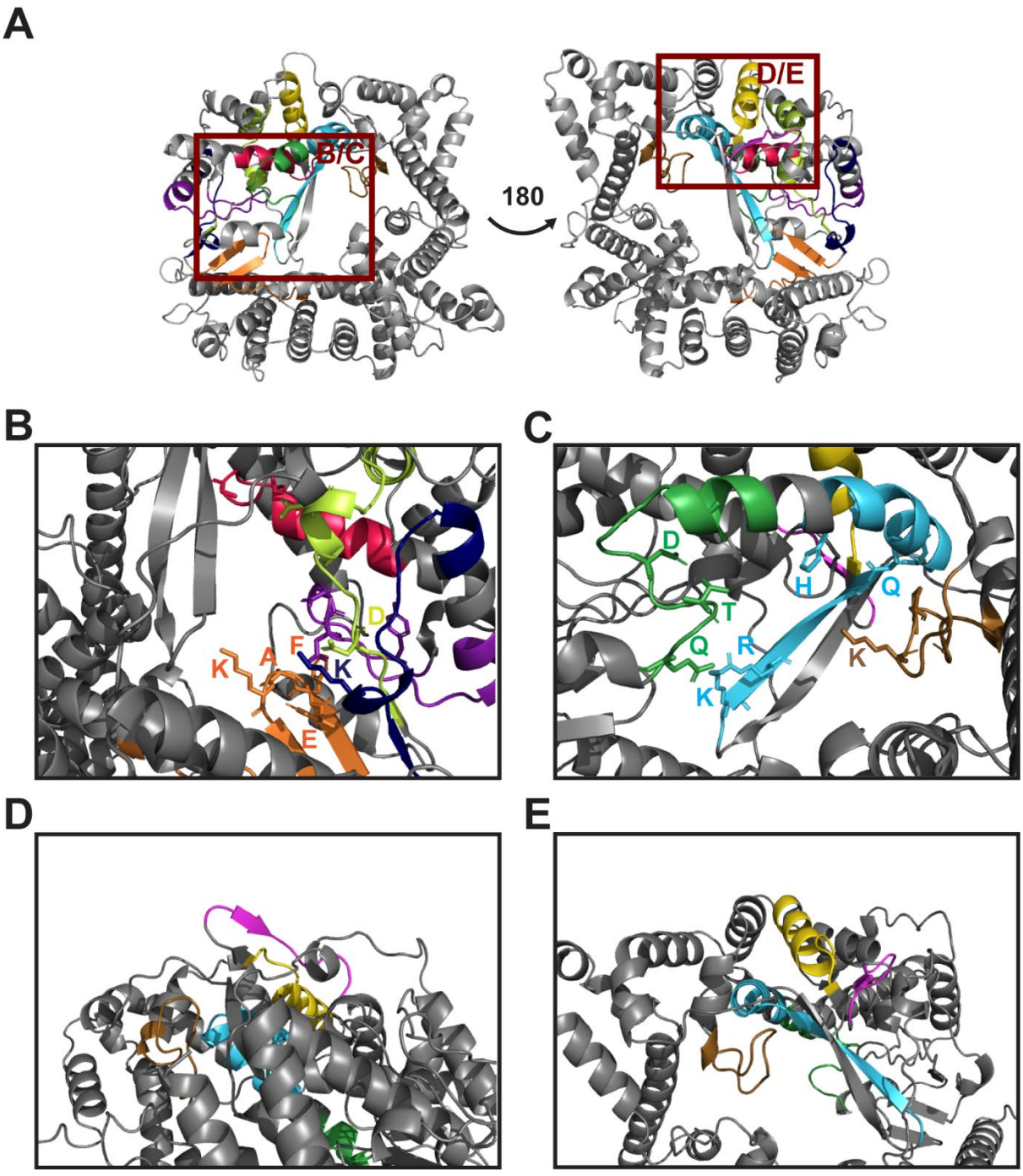


Figure 3.6 Alphafold Prediction of representative mitovirus RdRp. (A) Location of each subpanel and conserved protein motif (Figure 3.2) on representative mitovirus RdRp structural prediction. (B) Closer look at predicted catalytic pocket of mitoviral RdRp with highly conserved/catalytically essential amino acids shown. (C) Closer look at mitovirus specific protein motifs located within predicted catalytic pocket. (D/E) Closer look at mitovirus specific protein motifs predicted to be located on the surface of the RdRp structure. Structural predictions done on putative mitovirus ERR3412979_288_4 using Colabfold and visualized in PyMOL.

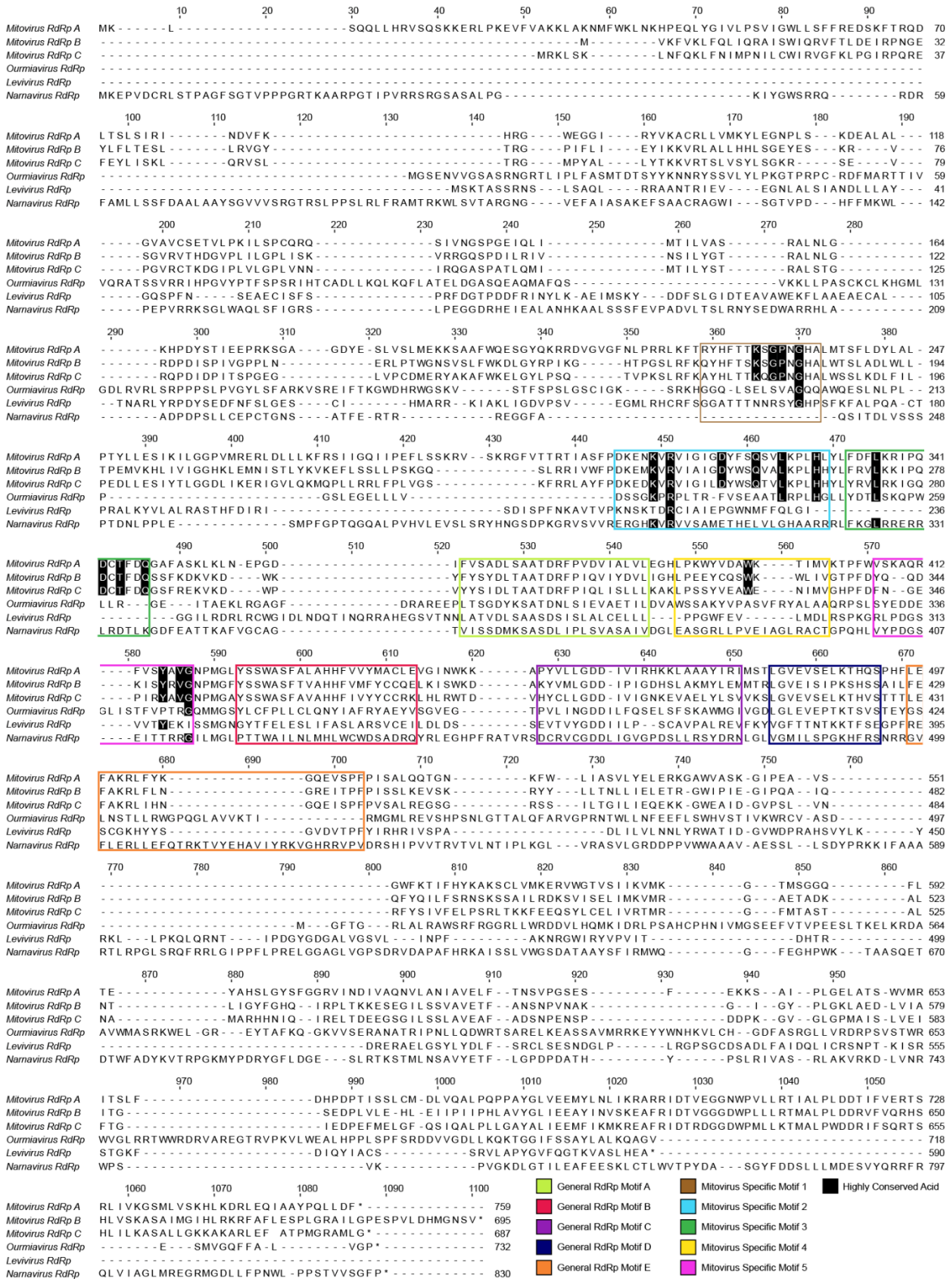

Mitovirus RdRp A
Mitovirus RdRp B
Mitovirus RdRp C
Ourmiavirus RdRp
Levivirus RdRp
Narnavirus RdRp
General RdRp Motif A
General RdRp Motif B
General RdRp Motif C
General RdRp Motif D
General RdRp Motif E
Mitovirus Specific Motif 1
Mitovirus Specific Motif 2
Mitovirus Specific Motif 3
Mitovirus Specific Motif 4
Mitovirus Specific Motif 5
Highly Conserved Acid

Figure 3.7 Representative multiple sequence alignment of mitoviral RdRp and closest evolutionary neighbors. Multiple sequence alignment using Clustal Omega was performed on three mitovirus RdRps and representative RdRps from mitoviral closest evolutionary neighbors the ourmiaviruses, leviviruses, and narnaviruses. Motifs correspond to motifs in Figure 3.4, Black amino acids are highly conserved amino acids from motif analysis.

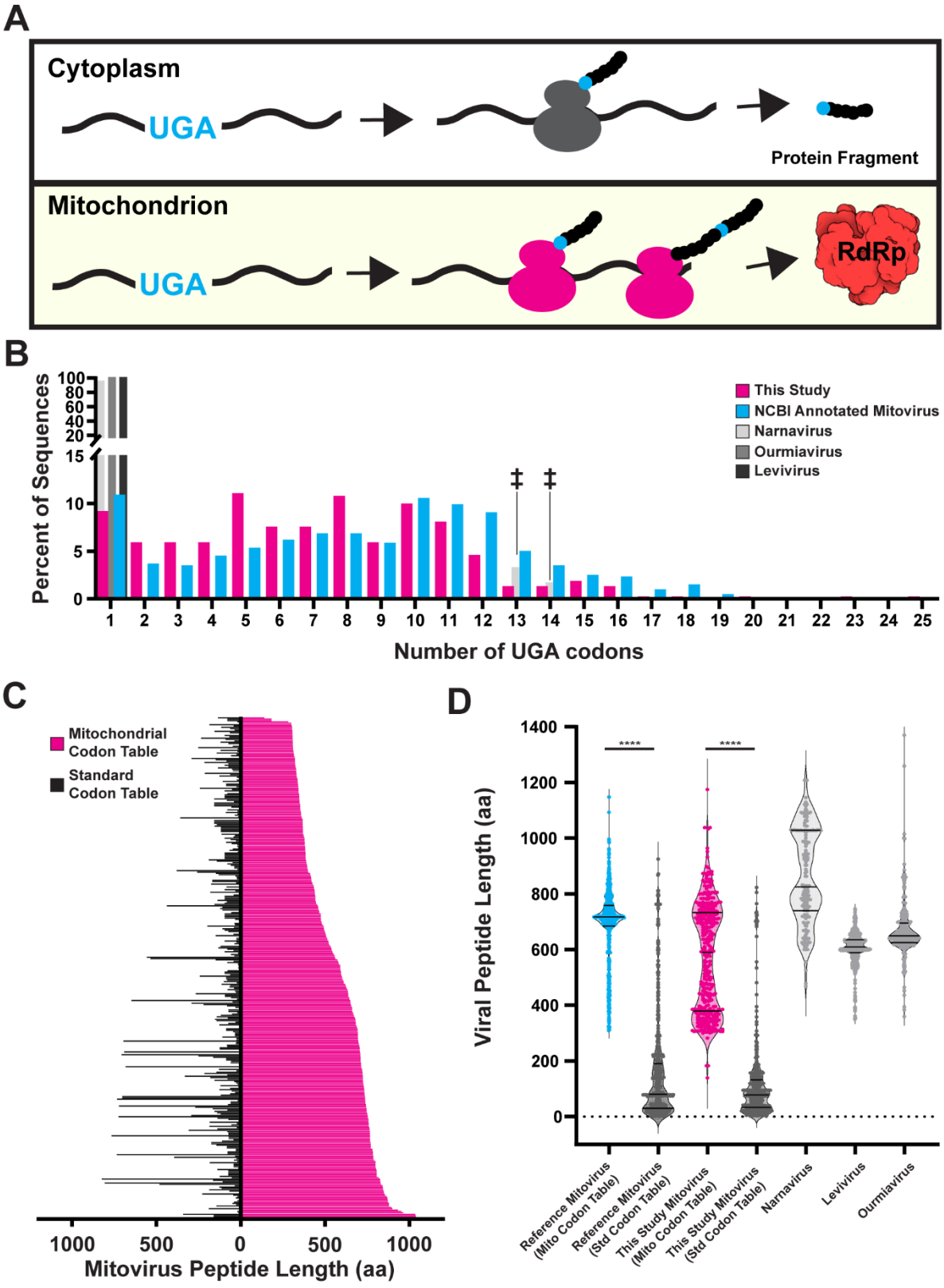

A
Cytoplasm
UGA
Protein Fragment
Mitochondrion
UGA
RdRp
B
Percent of Sequences
100
80
60
40
20
15
10
5
0
This Study
NCBI Annotated Mitovirus
Narnavirus
Ourmiavirus
Levivirus
‡
‡
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Number of UGA codons
C
Mitochondrial Codon Table
Standard Codon Table
1000 500 0 500 1000
Mitovirus Peptide Length (aa)
D
Viral Peptide Length (aa)
1400
1200
1000
800
600
400
200
0
****
****
Reference Mitovirus (Mito Codon Table)
Reference Mitovirus (Std Codon Table)
This Study Mitovirus (Mito Codon Table)
This Study Mitovirus (Std Codon Table)
Narnavirus
Levivirus
Ourmiavirus

Figure 3.8 Analysis of mitovirus non-canonical codon usage. (A) Transcripts using mitochondrial-specific UGA codon will only produce full-length products if translated on mitochondrial ribosomes. (B) Number of UGA codons in putative new mitovirus sequences (magenta), reference mitovirus sequences (cyan), and closest evolutionary neighbors (grays). ‡ represent misannotated narnavirus sequences (see Methods and Table S2 at https://doi.org/10.1128/msystems.01002-22). (C) Length of putative mitovirus peptides if translated using either the mitochondrial codon table (magenta) or standard cytosolic codon table (black). (D) Violin plot of data in B, including reference mitoviruses (cyan), and closest evolutionary neighbors (grays). Statistical tests run: Mann–Whitney test, **** corresponds to P-value < 0.0001.

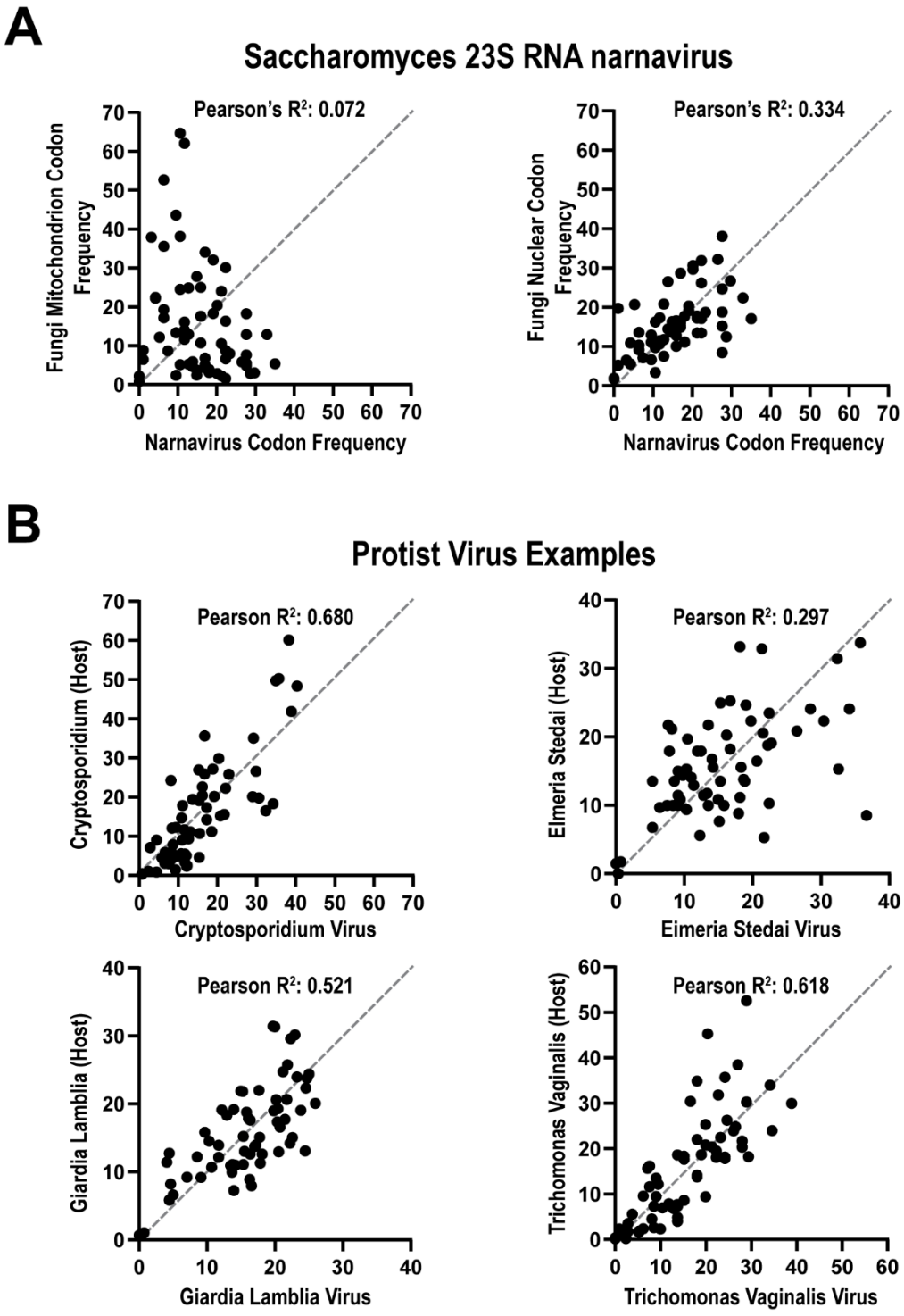


Figure 3.9 Narnavirus and protist virus codon usage correlations. (A) Codon usage correlation between representative narnavirus and fungal mitochondrial codon usage (left) and fungal nuclear codon usage (right). (B) Example codon usage correlation between four protist viruses and their respective host codon usage. All correlation values are Pearson's linear R2.

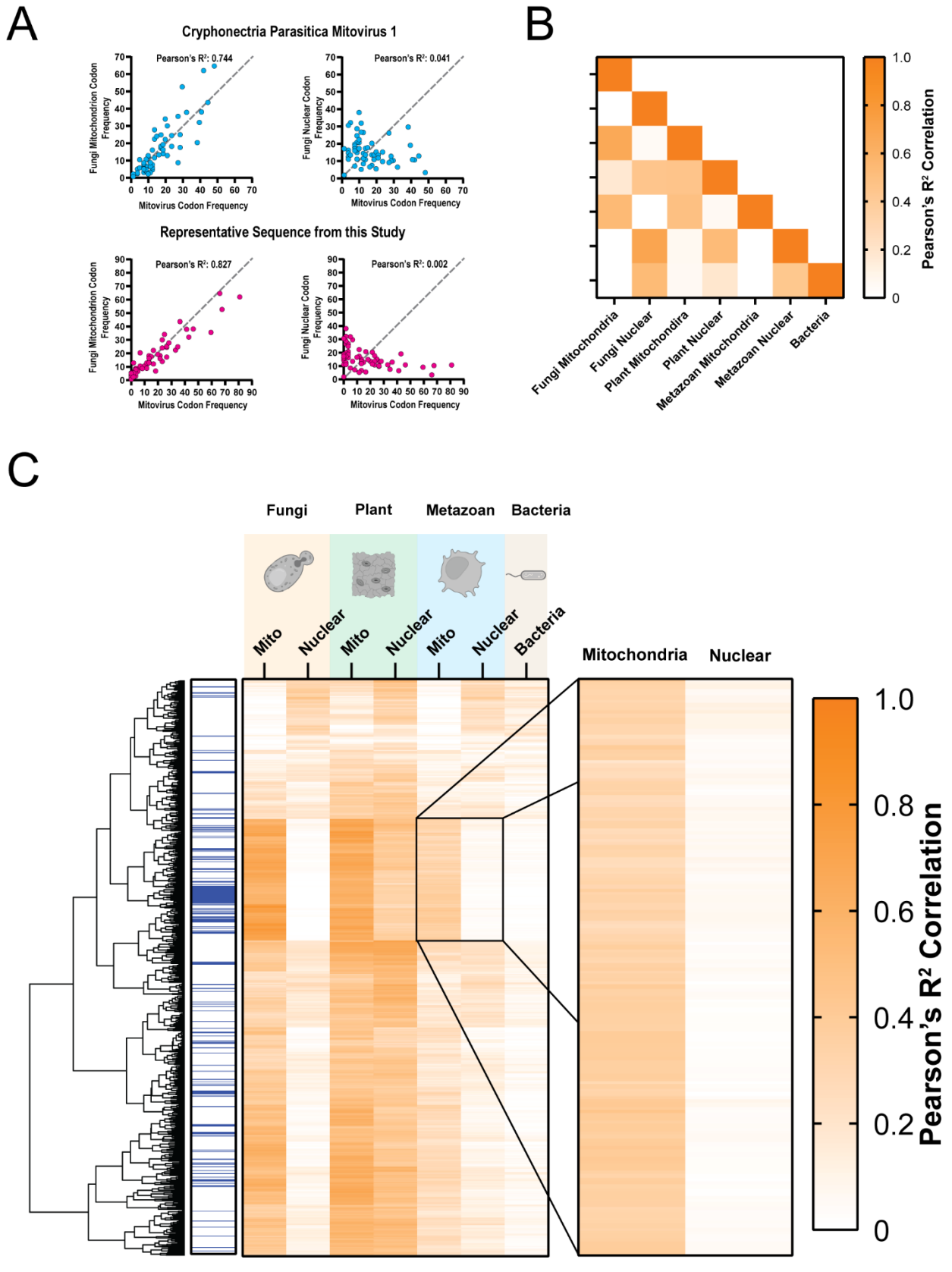
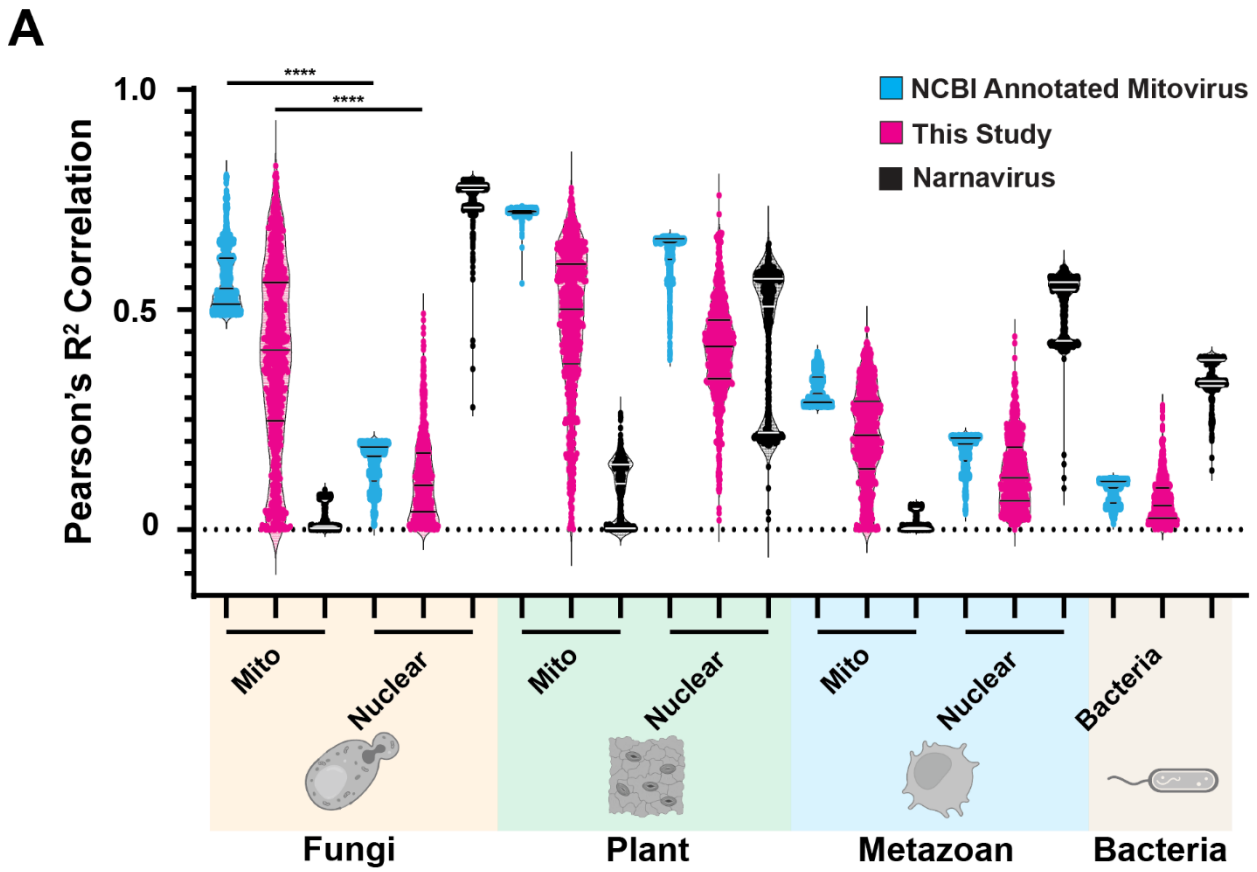
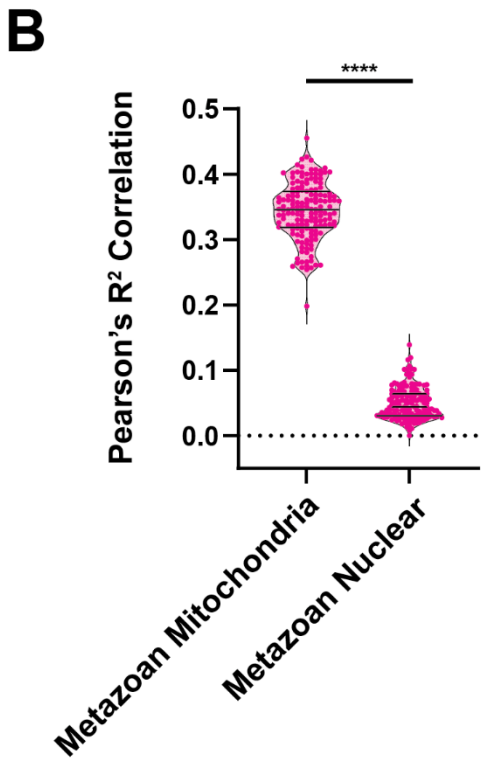

A
Cryphonectria Parasitica Mitovirus 1
Pearson's R²: 0.744
Pearson's R²: 0.041
Fungi Mitochondrion Codon Frequency
Fungi Nuclear Codon Frequency
Mitovirus Codon Frequency
Representative Sequence from this Study
Pearson's R²: 0.827
Pearson's R²: 0.002
B
Pearson's R² Correlation
Fungi Mitochondria
Fungi Nuclear
Plant Mitochondira
Plant Nuclear
Metazoan Mitochondria
Metazoan Nuclear
Bacteria
C
Fungi
Plant
Metazoan
Bacteria
Mito
Nuclear
Mitochondria
Nuclear
Pearson's R² Correlation

Figure 3.10 Codon usage bias of mitovirus sequences. (A) Codon usage correlation of representative putative new mitovirus (magenta) and reference mitovirus (cyan) between fungal nuclear codon usage and fungal mitochondrial codon usage. (B) Heatmap of codon usage correlation values between every reference data set used. (C) Heatmap of codon usage correlation values for each putative new mitovirus open reading frame (row) and both the mitochondrial and nuclear codon usage for fungal, plant, metazoan, and bacteria (columns). Sequences associated with animal sequencing studies are highlighted in blue. A subset of sequences with high metazoan codon usage correlation was called out. All correlation values are Pearson's linear R2.


A
Pearson's R² Correlation
1.0
0.5
0
****
****
NCBI Annotated Mitovirus
This Study
Narnavirus
Mito
Nuclear
Mito
Nuclear
Mito
Nuclear
Bacteria
Fungi
Plant
Metazoan
Bacteria


B
Pearson's R² Correlation
****
0.5
0.4
0.3
0.2
0.1
0.0
Metazoan Mitochondria
Metazoan Nuclear

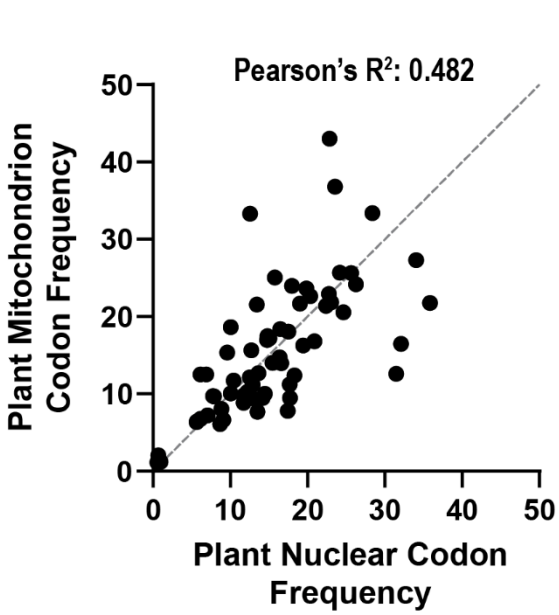

C
Pearson's R²: 0.482
Plant Mitochondrion Codon Frequency
50
40
30
20
10
0
0
10
20
30
40
50
Plant Nuclear Codon Frequency

Figure 3.11 Extended codon usage analysis. (A) Violin plot of codon usage correlation values from Figure 3.4C for reference mitoviruses (cyan), new putative mitoviruses (magenta), and narnaviruses (black). (B) Violin plot of codon usage correlation values of called out mitoviruses in Figure 3.4C. (C) Codon usage correlation between plant mitochondrial codon usage and plant nuclear codon usage. Statistical tests run: (A) Two-Way ANOVA with Tukey's multiple comparison post hoc, (B) Mann-Whitney, **** corresponds to P value < 0.0001.

References

1. Y. Z. Zhang, Y. M. Chen, W. Wang, X. C. Qin, E. C. Holmes, Expanding the RNA Virosphere by Unbiased Metagenomics. *Annu. Rev. Virol.* **6**, 119–139 (2019).
2. K. B. Gilbert, E. E. Holcomb, R. L. Allscheid, J. C. Carrington, Hiding in plain sight: New virus genomes discovered via a systematic analysis of fungal public transcriptomes. *PLoS ONE* **14**, e0219207 (2019).
3. Y. I. Wolf, S. Silas, Y. Wang, S. Wu, M. Bocek, D. Kazlauskas, M. Krupovic, A. Fire, V. V. Dolja, E. V. Koonin, Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nat. Microbiol.* **5**, 1262–1270 (2020).
4. M. J. Tisza, D. V. Pastrana, N. L. Welch, B. Stewart, A. Peretti, G. J. Starrett, Y. Y. S. Pang, S. R. Krishnamurthy, P. A. Pesavento, D. H. McDermott, P. M. Murphy, J. L. Whited, B. Miller, J. Brenchley, S. P. Rosshart, B. Rehermann, J. Doorbar, B. A. Ta'ala, O. Pletnikova, J. C. Troncoso, S. M. Resnick, B. Bolduc, M. B. Sullivan, A. Varsani, A. M. Segall, C. B. Buck, B. A. Ta'ala, O. Pletnikova, J. C. Troncoso, S. M. Resnick, B. Bolduc, M. B. Sullivan, A. Varsani, A. M. Segall, C. B. Buck, Discovery of several thousand highly diverse circular DNA viruses. *eLife* **9** (2020).
5. R. C. Edgar, J. Taylor, V. Lin, T. Altman, P. Barbera, D. Meleshko, D. Lohr, G. Novakovsky, B. Buchfink, B. Al-Shayeb, J. F. Banfield, M. de la Peña, A. Korobeynikov, R. Chikhi, A. Babaian, Petabase-scale sequence alignment catalyses viral discovery. *Nat. 2022 6027895* **602**, 142–147 (2022).
6. S. Miller, J. Krijnse-Locker, Modification of intracellular membrane structures for virus replication. *Nat. Rev. Microbiol. 2008 65* **6**, 363–374 (2008).
7. I. F. De Castro, L. Volonté, C. Risco, Virus factories: biogenesis and structural design. *Cell. Microbiol.* **15**, 24–34 (2013).
8. I. F. de Castro, J. J. Fernández, D. Barajas, P. D. Nagy, C. Risco, Three-dimensional imaging of the intracellular assembly of a functional viral RNA replicase complex. *J. Cell Sci.* **130**, 260–268 (2017).
9. G. Wolff, C. E. Melia, E. J. Snijder, M. Bárcena, Double-Membrane Vesicles as Platforms for Viral Replication. *Trends Microbiol.* **28**, 1022–1033 (2020).
10. C. Harak, V. Lohmann, Ultrastructure of the replication sites of positive-strand RNA viruses. *Virology* **479–480**, 418–433 (2015).
11. A. Shulla, G. Randall, (+) RNA virus replication compartments: a safe home for (most) viral replication. *Curr. Opin. Microbiol.* **32**, 82–88 (2016).
12. E. M. Scutigliani, M. Kikkert, Interaction of the innate immune system with positive-strand RNA virus replication organelles. *Cytokine Growth Factor Rev.* **37**, 17–27 (2017).
13. D. J. Miller, P. Ahlquist, Flock House Virus RNA Polymerase Is a Transmembrane Protein with Amino-Terminal Sequences Sufficient for Mitochondrial Localization and Membrane Insertion. *J. Virol.* **76**, 9856–9867 (2002).
14. I. H. Chen, Y. W. Huang, C. H. Tsai, The functional roles of the cis-acting elements in Bamboo mosaic virus RNA genome. *Front. Microbiol.* **8**, 645 (2017).
15. M. Falkenberg, Mitochondrial DNA replication in mammalian cells: Overview of the pathway, *Essays in Biochemistry*. **62** (2018)pp. 287–296.
16. Y. Yagi, T. Shiina, Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* **5**, 61 (2014).
17. R. Zoschke, R. Bock, Chloroplast Translation: Structural and Functional Organization,

Operational Control, and Regulation. *Plant Cell* **30**, 745 (2018).
18. E. C. Koc, M. E. Haque, L. L. Spremulli, Current Views of the Structure of the Mammalian Mitochondrial Ribosome. doi: 10.1002/ijch.201000002.
19. R. B. Seth, L. Sun, C.-K. Ea, Z. J. Chen, Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-κB and IRF3. *Cell* **122**, 669–682 (2005).
20. S. E. Weinberg, L. A. Sena, N. S. Chandel, Mitochondria in the Regulation of Innate and Adaptive Immunity. *Immunity* **42**, 406–417 (2015).
21. A. P. West, G. S. Shadel, Mitochondrial DNA in innate immune responses and inflammatory pathology. *Nat. Rev. Immunol.* **17**, 363–375 (2017).
22. B. I. Hillman, G. Cai, "The Family Narnaviridae. Simplest of RNA Viruses." in *Advances in Virus Research* (Academic Press Inc., 2013; https://pubmed.ncbi.nlm.nih.gov/23498906/)vol. 86, pp. 149–176.
23. Y. Park, X. Chen, Z. K. Punja, Molecular and Biological Characterization of a Mitovirus in Chalara elegans (Thielaviopsis basicola). doi: 10.1094/PHYTO-96-0468 (2006).
24. M. Wu, L. Zhang, G. Li, D. Jiang, S. A. Ghabrial, Genome characterization of a debilitation-associated mitovirus infecting the phytopathogenic fungus Botrytis cinerea. *Virology* **406**, 117–126 (2010).
25. D. Di Silvestre, L. Tadini, A. Trotta, L. Valledor, G. H. Salekdeh, J. V. Jorrin Novo, Editorial: A Systems View of Plant Cellular Communication. *Front. Plant Sci.* **13**, 674 (2022).
26. H. J. Rogers, K. W. Buck, C. M. Brasier, A mitochondrial target for double-stranded RNA in diseased isolates of the fungus that causes Dutch elm disease. *Nature* **329**, 558–560 (1987).
27. J. J. Polashock, B. I. Hillman, "A small mitochondrial double-stranded (ds) RNA element associated with a hypovirulent strain of the chestnut blight fungus and ancestrally related to yeast cytoplasmic T and W dsRNAs" (1994).
28. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114 (2014).
29. A. Prjibelski, D. Antipov, D. Meleshko, A. Lapidus, A. Korobeynikov, Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinforma.* **70**, e102 (2020).
30. E. Bushmanova, D. Antipov, A. Lapidus, A. D. Prjibelski, rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* **8**, 1–13 (2019).
31. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND, *Nature Methods*. **12** (2014)pp. 59–60.
32. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods 2012 94* **9**, 357–359 (2012).
33. J. A. Gerlt, J. T. Bouvier, D. B. Davidson, H. J. Imker, B. Sadkhin, D. R. Slater, K. L. Whalen, Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks, *Biochimica et Biophysica Acta - Proteins and Proteomics*. **1854** (2015)pp. 1019–1037.
34. J. A. Gerlt, Genomic Enzymology: Web Tools for Leveraging Protein Family Sequence-Function Space and Genome Context to Discover Novel Functions. *Biochemistry* **56**, 4293–4308 (2017).
35. S. Kraberger, K. Schmidlin, R. S. Fontenele, M. Walters, A. Varsani, Unravelling the Single-Stranded DNA Virome of the New Zealand Blackfly. *Viruses 2019 Vol 11 Page 532*

11, 532 (2019).
36. R. S. Fontenele, C. Lacorte, N. S. Lamas, K. Schmidlin, A. Varsani, S. G. Ribeiro, Single Stranded DNA Viruses Associated with Capybara Faeces Sampled in Brazil. *Viruses 2019 Vol 11 Page 710* **11**, 710 (2019).
37. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
38. F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, M. F, P. YM, L. J, B. N, G. T, M. N, B. P, T. ARN, P. SC, F. RD, L. R, F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47** (2019).
39. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* **5**, e9490 (2010).
40. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49** (2021).
41. A. Alexaki, J. Kames, D. D. Holcomb, J. Athey, L. V. Santana-Quintero, P. V. N. Lam, N. Hamasaki-Katagiri, E. Osipova, V. Simonyan, H. Bar, A. A. Komar, C. Kimchi-Sarfaty, Codon and Codon-Pair Usage Tables (CoCoPUTs): Facilitating Genetic Variation Analyses and Recombinant Gene Design. *J. Mol. Biol.* **431**, 2434–2441 (2019).
42. T. L. Bailey, M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li, W. S. Noble, MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* **37** (2009).
43. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
44. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nat. 2021 5967873* **596**, 583–589 (2021).
45. M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold - Making protein folding accessible to all. *bioRxiv*, 2021.08.15.456425 (2021).
46. Schrödinger LLC, The PyMOL Molecular Graphics System, Version\sim1.8 (2015).
47. R. Lorenz, S. H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. Hofacker, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 1–14 (2011).
48. A. T. Peterson, Defining viral species: making taxonomy useful. *Virol. J.* **11**, 131 (2014).
49. P. Simmonds, M. J. Adams, M. Benk, M. Breitbart, J. R. Brister, E. B. Carstens, A. J. Davison, E. Delwart, A. E. Gorbalenya, B. Harrach, R. Hull, A. M. Q. Q. King, E. V. Koonin, M. Krupovic, J. H. Kuhn, E. J. Lefkowitz, M. L. Nibert, R. Orton, M. J. Roossinck, S. Sabanadzovic, M. B. Sullivan, C. A. Suttle, R. B. Tesh, R. A. Van Der Vlugt, A. Varsani, F. M. Zerbini, Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol. 2017 153* **15**, 161–168 (2017).
50. G. M. Jenkins, E. C. Holmes, The extent of codon usage bias in human RNA viruses and its

evolutionary origin. *Virus Res.* **92**, 1–7 (2003).
51. E. H. M. M. Wong, D. K. Smith, R. Rabadan, M. Peiris, L. L. M. M. Poon, Codon usage bias and the evolution of influenza A viruses. Codon Usage Biases of Influenza Virus. *BMC Evol. Biol.* **10**, 1–14 (2010).
52. M. Shi, X. D. Lin, J. H. Tian, L. J. Chen, X. Chen, C. X. Li, X. C. Qin, J. Li, J. P. Cao, J. S. Eden, J. Buchmann, W. Wang, J. Xu, E. C. Holmes, Y. Z. Zhang, Redefining the invertebrate RNA virosphere. *Nature* **540**, 539–543 (2016).
53. S. Anderson, A. T. Bankier, B. G. Barrell, M. H. L. De Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden, I. G. Young, "Sequence and organization of the human mitochondrial genome" (6, 1981).
54. F. Chen, P. Wu, S. Deng, H. Zhang, Y. Hou, Z. Hu, J. Zhang, X. Chen, J. R. Yang, Dissimilation of synonymous codon usage bias in virus–host coevolution due to translational selection. *Nat. Ecol. Evol. 2020 44* **4**, 589–600 (2020).
55. P. M. Sharp, G. Matassi, Codon usage and genome evolution. *Curr. Opin. Genet. Dev.* **4**, 851–860 (1994).
56. J. R. Powell, E. N. Moriyama, Evolution of codon usage bias in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7784–7790 (1997).
57. M. Castells, M. Victoria, R. Colina, H. Musto, J. Cristina, Genome-wide analysis of codon usage bias in Bovine Coronavirus. *Virol. J.* **14**, 1–7 (2017).
58. W. H. Campbell, G. Gowri, Codon usage in higher plants, green algae, and cyanobacteria, *Plant Physiology*. **92** (1990)pp. 1–11.
59. J. R. Brum, M. B. Sullivan, Rising to the challenge: Accelerated pace of discovery transforms marine virology, *Nature Reviews Microbiology*. **13** (2015)pp. 147–159.
60. M. Sangiovanni, I. Granata, A. S. Thind, M. R. Guarracino, From trash to treasure: Detecting unexpected contamination in unmapped NGS data. *BMC Bioinformatics* **20**, 1–12 (2019).
61. N. Suzuki, C. Cornejo, A. Aulia, S. Shahi, B. I. Hillman, D. Rigling, In-Tree Behavior of Diverse Viruses Harbored in the Chestnut Blight Fungus, Cryphonectria parasitica. *J. Virol.* **95** (2021).
62. S. Shahi, A. Eusebio-Cope, H. Kondo, B. I. Hillman, N. Suzuki, Investigation of Host Range of and Host Defense against a Mitochondrially Replicating Mitovirus. *J. Virol.* **93** (2019).
63. Q. Wang, F. Mu, J. Xie, J. Cheng, Y. Fu, D. Jiang, A Single ssRNA Segment Encoding RdRp Is Sufficient for Replication, Infection, and Transmission of Ourmia-Like Virus in Fungi. *Front. Microbiol.* **0**, 379 (2020).
64. P. A. Gammage, C. T. Moraes, M. Minczuk, Mitochondrial Genome Engineering: The Revolution May Not Be CRISPR-Ized, *Trends in Genetics*. **34** (2018)pp. 101–110.
65. P. Silva-Pinheiro, M. Minczuk, The potential of mitochondrial genome engineering. *Nat. Rev. Genet. 2021*, 1–16 (2021).