**Additional File 2: Supplementary Methodology, Tables, and Figures – Mitochondrial Import of Malat1 Regulates Cardiac Mitochondrial Function in Type 2 Diabetes Mellitus through Interaction with MicroRNA-23b**

Quincy A. Hathaway, PhD1,2, Andrew D. Taylor, BS1,2, Amina Kunovac, BS1,2, Mark V. Pinti, BS2,3, Mackenzie S. Newman, PhD4, Chris C. Cook, MD5, Michael T. Winters, BS6, Emily Westemeier, BS6, Garrett K. Fink, BS1, Andrya J. Durr, BS1,2, Danielle L. Shepherd, PhD1,2, Aaron R. Robart, PhD7, Ivan Martinez, PhD6, and John M. Hollander, PhD1,2

**1**Division of Exercise Physiology, West Virginia University School of Medicine, Morgantown, WV, USA. **2**Mitochondria, Metabolism & Bioenergetics Working Group, West Virginia University School of Medicine, Morgantown, WV, USA. **3**West Virginia University School of Pharmacy, Morgantown, WV, USA. 4Department of Physiology and Pharmacology, West Virginia University School of Medicine, Morgantown, WV, USA. **5**Cardiovascular and Thoracic Surgery, West Virginia University School of Medicine, Morgantown, WV, USA, **6**Department of Microbiology, Immunology & Cell Biology, West Virginia University Cancer Institute, School of Medicine, Morgantown, WV, USA, **7**Department of Biochemistry, West Virginia University School of Medicine, Morgantown, WV, USA.

**Running Title:***Malat1 Regulates Mitochondrial Translation*

**Corresponding Author:**

John M. Hollander, Ph.D., F.A.H.A.

Division of Exercise Physiology

West Virginia University School of Medicine

PO Box 9227

1 Medical Center Drive

Morgantown, WV 26506

Tel: 1-(304) 293-3683

Fax: 1-(304) 293-7105

Email: jhollander@hsc.wvu.edu

**Supplemental – Materials and Methods:**

*Mitochondrial Isolations*

Mitochondrial subpopulations were isolated for analyses as previously described 1, with modifications by our laboratory 2-4.

*HL-1 Cell Culture*

Cells were supplemented with Claycomb media (Sigma Aldrich, St. Louis, MO) containing 10% FBS, glutamine (2 mM), penicillin/streptomycin, and norepinephrine (0.1 mM). Cells were seeded and transfection occurred at 70-80% confluence. Lipofectamine 3000 (Thermo Fisher, Waltham, MA) was used to transfect cells, per manufacturer’s instructions.

*Malat1 and mmu-MiRNA-23b-3p Knockdown*

Malat1 Silencer® Select (n253517, Thermo Fisher) and mmu-miRNA-23b-3p Anti-miR™ miRNA Inhibitor (AM10711, Thermo Fisher) were used to knockdown Malat1 and miRNA-23b expression, respectively. The Malat1 siRNA was designed to target the second exon of Malat1 transcript variant 4 (Malat1-204). A scramble control, Silencer™ Select Negative Control No. 1 siRNA (4390843, Thermo Fisher) was also employed. Immortalized HL-1 murine cardiomyocytes were cultured and preserved as described above. Lipofectamine 3000 (Thermo Fisher) was used to transfect cells, per manufacturer’s instructions.

*Malat1 Fluorescent In Situ Hybridization (FISH)*

A modified version of the StellarisⓇ RNA FISH Protocol for Adherent Cells was used. StellarisⓇ FISH Probes against Malat1 and Gapdh (Biosearch Technologies, Inc., Petaluma, CA) were utilized in the hybridization, and NucBlueTM Fixed Cell Stain ReadyProbes reagent (Thermo Fisher) was used to stain nuclei. Following MitoTracker™ Deep Red FM (Thermo Fisher) incubation, coverglass was washed with 1 mL 1X Phosphate Buffered Saline (PBS) and 1 mL StellarisⓇ RNA FISH Fixation Buffer (3.7% (vol./vol.) formaldehyde in 1X PBS) was then added to the 6-well dishes and incubated at room temperature for 10 minutes. The cells were washed twice with 1mL 1X PBS and permeabilized with 1mL 70% ethanol for 16 hours at 4oC. The ethanol was eluted and the cells were incubated with StellarisⓇ RNA FISH Wash Buffer A (Biosearch Technologies) (10% (vol./vol.) formamide in Wash Buffer A) for 5 minutes at room temperature.

The cells were transferred, on the coverglass, to a humidified chamber and incubated with 100µl of StellarisⓇ RNA FISH Hybridization Buffer (Biosearch Technologies) (10% (vol./vol.) formamide in Hybridization Buffer) containing either a Malat1 (#SMF-3008-1) or Gapdh (#SMF-3002-1) probe (StellarisⓇ FISH Probes with QuasarⓇ 570 Dye) for 16 hours at 37oC in darkness. Each StellarisⓇ FISH Probe has approximately 80 hybridization sites on the target RNA. The coverglass was then incubated with 1 mL StellarisⓇ RNA FISH Wash Buffer A for 30 minutes at 37oC in darkness, the buffered was aspirated, and one drop of NucBlueTM Fixed Cell Stain ReadyProbes reagent (Thermo Fisher) (DAPI stain) was added. The cells were then incubated for 30 minutes at 37oC in darkness. The NucBlueTM stain was removed, 1 mL StellarisⓇ RNA FISH Wash Buffer B (Biosearch Technologies) was added, and the cells were allowed to incubate at room temperature for 5 minutes. A drop of Vectashield Mounting Medium (Vector Laboratories, San Diego, CA) was added to a microscope slide, and the coverglass was mounted. The mounted cells were sealed with clear nail polish and stored at 4oC.

*Crosslinking Immunoprecipitation (CLIP)*

After crosslinking, cells were centrifuged (1,200 x g) for 7 minutes and subsequently frozen (-80oC) for further use. 75 µL per sample of Dynabeads™ Protein G (Thermo Fisher) were prepared through 3 washes with NP-40 buffer (20mM Tris (pH 8.0), 137mM NaCl, 10% Glycerol, 1% Triton X100, 2mM EDTA, 0.1mM PMSF). After the third wash, the supernatant on the beads was discarded and 5 µg of anti-GFP was added, along with NP-40 buffer, to reach a total volume of 100 µL. Beads were incubated overnight at 4oC. The next day, each sample was resuspended in NP-40 buffer to a total volume of 1 mL, and protein concentrations determined through the Bradford Method 5.

100 µg of protein from each sample was added to a new tube followed by RNAse I treatment (1:500 dilution, Ambion™ RNase I, Thermo Fisher) for 3 minutes at 37oC, which allowed for partial digestion of RNA bound to PNPase. The reaction was stopped by incubating the samples on ice. The supernatant on the beads solution was then removed and the protein mixture was added and incubated at 4oC for 4 hours. Following the incubation, beads were washed 3 times in NP-40 buffer, final supernatant removed, and a mixture of water/NuPAGE™ LDS Sample Buffer (4X) (Thermo Fisher) was added to the beads. The beads solution was heated at 70oC for 15 minutes and assessed through immunoblotting (see *Western Blotting* below) and sequenced (see *High-Throughput Sequencing Crosslinking Immunoprecipitation (HITS-CLIP)* below).

*Western Blotting*

Using the Bradford Method, protein concentrations were normalized. Primary antibodies (1:1000 dilution) implemented in the study included the following: anti-GFP ((B-2): sc-9996, SCBT, Dallas, TX), anti-PNPase ((C-5): sc-271973, SCBT), anti-PNPase ((D-1): sc-271479, SCBT), anti-ND2 (Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 2 (M-16): sc-20496, SCBT), anti-ND4 (Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 4 (A-16): sc-20499-R, SCBT), anti-GAPDH (ab8245, Abcam, Cambridge, MA), anti-VDAC1 (Voltage-Dependent Anion Channel, SAB5201374, Sigma), anti-COX IV (ab16056, Abcam), and anti-HDAC6, ((H-300): sc-11420, SCBT). Rabbit anti-mouse IgG (H&L) horseradish peroxidase (HRP) conjugate 1:10,000 (ab6728, Abcam) and goat anti-rabbit (H&L) HRP conjugate 1:10,000 (ab6721, Abcam) were used as secondary antibodies when appropriate. Normalization of protein content was through GAPDH (cytoplasmic) and VDAC (mitochondrial) expression. Chemiluminescence quantified with Radiance Chemiluminescent Substrate (Azure Biosystems, Dublin, CA), per manufacturer’s instructions and imaged using the G:Box Bioimaging system (Syngene, Frederick, MD). Images were taken with GeneSnap/GeneTools software (Syngene).

*RNA Isolation/Quantitative PCR*

Each sample was homogenized in QIAzol lysis reagent before proceeding. Total RNA was isolated for each sample and RNA to be analyzed through qPCR was converted to cDNA through the High-Capacity RNA-to-cDNA™ Kit (Thermo Fisher), per manufacturer’s instructions. Differential gene expression was assessed for mt-Nd3 and mt-Atp6 in mouse cytoplasmic and mitochondrial fractions, as well as Malat1-204 and Neat1-202 in mitochondrial fractions. Malat1 expression was also measured in HL-1 cells following PNPase CLIP (as described above). Experiments were performed on the Applied Biosystems 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA), using 2X SYBR Green Master Mix. Quantification was achieved using the 2-ΔΔCT method 6, standardized to U6 and Gapdh expression where appropriate.

*Long non-coding RNA (LncRNA) Sequencing*

Briefly, RNA was selected for lncRNA analysis using a Ribo-Zero rRNA Removal Kit (Illumina, San Diego, CA). Samples were run on the HiSeq 2500 (Illumina) in 51bp paired-end reads achieving ~30 million clustered reads per sample. Generated Fastq files were processed through HISAT2 7, 8 under standard parameters, with the exception of quality scoring control specified under “--phred33.” The reference genome used in generating BAM files included fasta formatted DNA from Ensemble release 95 for human (GRCh38) and mouse (GRCm38). Differential gene expression was performed in the R (v3.5.3) environment, through DESeq2 9. Visualization was performed through the packages ggplot2 10, limma 11, vidger 12, and EnhancedVolcano 13. Additional programs, such as Seqmonk (v1.45.4) and Cytoscape (v3.7.2) 14 were implemented to display data.

*Small ncRNA Sequencing*

Small RNA was selected using the NEXTFLEX® Small RNA-Seq Kit v3 (Bioo Scientific, Austin, TX). Samples were run on the HiSeq 2500 in Rapid Run (Illumina) in 64bp single-end reads on two lanes achieving ~12 million clustered reads per sample. Adapter trimming through cutadapt 15 (5- TGGAATTCTCGGGTGCCAAGG -3) allowed for isolation of genomic reads. Generated Fastq files were processed through Bowtie 16 using the parameters, “-n 0 -l 10 -m 100 --best --strata --chunkmbs 250.” The reference genome used in generating BAM files included fasta formatted DNA from Ensemble release 98 for human (GRCh38). Differential gene expression was assessed in the R (v3.5.3) environment, as indicated above.

*High-Throughput Sequencing Crosslinking Immunoprecipitation (HITS-CLIP)*

RNA was selected for using the NEXTFLEX® Small RNA-Seq Kit v3 (Bioo Scientific). Samples were run on the MiSeq in 36bp single-end reads achieving ~1 million clustered reads per sample. Adapter trimming through cutadapt and sequence alignment was performed as indicated above. Generated Fastq files were processed through Bowtie using the parameters, “-n 0 -l 10 -m 500 --best --strata --chunkmbs 250.” The reference genome used in generating BAM files included fasta formatted DNA from Ensemble release 98 for human (GRCh38) and mouse (GRCm38). Differential gene expression was assessed in the R (v3.5.3) environment, as indicated above.

*Mitochondrial Targeting Sequences*

To determine if lncRNA that were found in the mitochondrion and associated with PNPase had sequence or secondary structure homology, RNAfold through the ViennaRNA Package was employed 17. RNAfold was used in conjunction with LncFinder 18 and seqinR 19 to evaluate folding parameters of both the full length lncRNA, as well as isolated regions (12-20 nucleotides) shown to bind to PNPase in CLIP analyses. Isolated regions were identified for 50 human and 50 mouse lncRNA labeled “positive” (found in CLIP Seq) and 100 lncRNA labeled “negative” (randomly generated sequences 20). The isolated regions from genes identified within the PNPase CLIP were flanked by surrounding nucleotides within the genome to achieve 60 nucleotide sequences implemented in the subsequent RNAfold and machine learning applications. The length of 60 nucleotides for each target was selected to include enough information about the surrounding region to generate an appropriate secondary structure, but also succinct enough to be used in computational models for designing RNA secondary structures 21. The randomly generated sequences also consisted of 60 nt. The sequences directly identified in the CLIP Seq, as well as flanking regions, are provided for each positive control gene **(Additional File 1)**.

*Machine Learning for RNA Sequences*

Using 10-fold cross validation, classification of “positive” and “negative” 60 nt isolated regions was assessed through Classification and Regression Trees (CART) in randomForest 22 and rfUtilities 23. Machine learning algorithms were performed on 105 features, which included primary sequence and secondary structure information **(Additional File 1)**. Briefly, *rf.crossValidation* was used with a randomForest object to provide 10-fold cross validation. An unsupervised learning approach was applied to only the positive 60 nt regions through *rf.unsupervised*, with the identification of four specific clusters. As a confirmatory measure, Support Vector Machines (SVM), using the *svm\_cv* function of LncFinder, was implemented to determine 10-fold cross validation of full length and isolated regions.

*RNA-RNA and RNA-DNA Prediction Software*

NPInter (v4.0) 24 provided information for experimentally validated RNA-RNA interactions which were used to build models in Cytoscape (v3.7.2). IntaRNA 2.0 25, executed under standard parameters, was implemented to assess the binding affinities of mt-Rnr1-201, Malat1, and miRNA-23b-3p in mouse. The Regulatory Genomics Toolbox (RGT) function Triplex Domain Finder (TDF) was implemented to characterize triplex formation between MALAT1 and the mitochondrial genome in human and mouse 26.

**Supplemental Tables and Table Legends**

**Table S1: Primer Design**

|  |  |  |  |
| --- | --- | --- | --- |
| **Assay** | **Primer Name** | **Forward (5’ – 3’)** | **Reverse (5’ – 3’)** |
| Figure 1: Mitochondrial Purity | mt-Nd3 | AATGCGGATTCGACCCTAC | GAATTGCTCATGGTAGTGGAAG |
| mt-Atp6 | GATTCCCAATCGTTGTAGCC | GGAAAGAATGGAGACGGTTG |
| Figure S2: LncRNA mitochondrial validation | Malat1-204 | TGTGGCAAGAATCAAGCAAG | TTGAGGTTTGGGCTGGTAAC |
| Neat1-202 | ATTGGCCAGAAGACAACAGG | TCTCATAGCAAGCGCAAGAC |
| Figure 6: PNPase HL-1 CLIP | Malat1-204 | TGTGGCAAGAATCAAGCAAG | TTGAGGTTTGGGCTGGTAAC |

**Table S1: Primers designed for experiments.** Primers were designed through Primer3 and NCBI Primer-BLAST.

**Table S2: Machine Learning Performance**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Positives** | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** |
| **Classification and Regression Cross-Validation (CV)** |
| User’s Accuracy (Positives) | 70.00 | 100.00 | 0 | 0 | 0 |
| User’s Accuracy (Negatives) | 80.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| Producer’s Accuracy (Positives) | 70.85 | NA | NA | NA | NA |
| Producer’s Accuracy (Negatives) | 73.35 | 100.00 | 83.30 | 83.30 | 83.30 |
| CV Kappa | 0.44 | 0.67 | 0.30 | 0.25 | 0.32 |
| CV OoB Error | 0.28 | 0.09 | 0.15 | 0.18 | 0.18 |
| CV Error Variance | 7.88E-4 | 1.06E-3 | 2.07E-4 | 1.04E-4 | 3.59E-3 |
| **Classification and Regression Model** |
| User’s Accuracy (Positives) | 71.60 | 63.45 | 23.80 | 20.80 | 32.00 |
| User’s Accuracy (Negatives) | 72.60 | 97.9 | 98.90 | 97.90 | 95.25 |
| Producer’s Accuracy (Positives) | 72.75 | 88.90 | 80.00 | 71.40 | 62.55 |
| Producer’s Accuracy (Negatives) | 71.80 | 90.7 | 85.50 | 83.00 | 84.10 |
| Model Kappa | 0.45 | 0.67 | 0.32 | 0.25 | 0.32 |
| Model OoB Error | 0.28 | 0.09 | 0.15 | 0.18 | 0.18 |
| Model Error Variance | 2.39E-4 | 8.27E-5 | 1.85E-4 | 1.01E-4 | 9.57E-5 |
| **Support Vector Machines Cross-Validation (CV)** |
| Sensitivity | 0.76 | 0.67 | 0.23 | 0.38 | 0.38 |
| Specificity | 0.67 | 0.93 | 0.96 | 0.94 | 0.89 |
| Accuracy | 0.72 | 0.88 | 0.83 | 0.83 | 0.79 |
| F-Measure | 0.73 | NA | NA | NA | NA |
| Kappa | 0.43 | 0.60 | 0.23 | 0.37 | 0.27 |

**Table S2: Machine Learning performance for Classification and Regression Trees (CART) and Support Vector Machines (SVM).** Performance was determined for entire dataset (100 positives and 100 negatives) as well as for each cluster determined through unsupervised learning (25 positives and 100 negatives). Positives = 50 human and 50 mouse ncRNAs determined to be bound to PNPase, or 25 genes found in each of the four clusters, Negatives = 100 randomly generated sequences, CV = 10-fold cross validation, OoB = out of bag.

**Supplemental Figure Legends**

**Figure S1: Constructs designed for transfection into HL-1 cells.** Sequences were inserted into a pcDNA3.1+N-eGFP vector and transfected into cells. FL = full length PNPase open reading frame (ORF) in pcDNA3.1+N-eGFP, KH = exon 23 removed from full length PNPase in pcDNA3.1+N-eGFP, S1 = C-terminus removed from full length PNPase in in pcDNA3.1+N-eGFP.

**Figure S2: Sample Distances calculated by Poisson Distance.** (**A**) Sample distances of human atrial tissue (ND, n = 6: T2DM, n = 6) for cytoplasmic and mitochondrial samples. (**B**) Sample distances of mouse whole heart (WT, n = 6: DbDb, n =6) for cytoplasmic and mitochondrial samples. (**C**) Sample distances of crosslinking immunoprecipitation (CLIP) samples for Ago2 (n = 4), Fxr1 (n = 4), and PNPase (n = 4). T2DM = type 2 diabetes mellitus, ND = non-diabetic, DbDb = FVB/NJ *Leprdb\_+/+* mice, WT = FVB/NJ wild-type mice, Mito = mitochondrial, Cyto = cytoplasmic, Argonaute-2 (Ago2), Fragile X mental retardation syndrome-related protein 1 (Fxr1), and Polynucleotide Phosphorylase (PNPase).

**Figure S3: Distribution of reads for crosslinking immunoprecipitation (CLIP) samples.** The size of the reads within each sample (20 – 40 bps) and the stratification of small non-coding RNA (ncRNA) across groups in **(A)** mouse Ago2, Fxr1, and PNPase and **(B)** human PNPase CLIP. **(C)** Heatmap displaying the top 200 genes with the lowest Padj value in human mitochondria, with changes displayed as the fold change deviation from the mean of the data. Values derived from sequencing are considered statistically significant when *Padj* < 0.05 or –Log10Padj = 1.30. Significance was determined through the Wald test in the R environment for sequencing comparisons. DbDb = FVB/NJ *Leprdb\_+/+* mice, WT = FVB/NJ wild-type mice, Argonaute-2 (Ago2), Fragile X mental retardation syndrome-related protein 1 (Fxr1), and Polynucleotide Phosphorylase (PNPase).

**Figure S4. Malat1 fluorescent *in situ* hybridization in HL-1 cells.** **(A)** Confocal microscopy for multiple cells depicting the nucleus (DAPI, blue, emission 460 nm), probe (Malat1 or Gapdh, red, emission 570 nm), and merged image following fluorescent *in situ* hybridization. (B) Confocal microscopy for multiple cells depicting the nucleus (DAPI, blue, emission 460 nm), probe (Malat1 or Gapdh, red, emission 570 nm), mitochondria (MitoTracker®, green, emission 665 nm), and merged image following fluorescent *in situ* hybridization. Areas of overlay between the probe (red) and mitochondria (green) are depicted as yellow. Malat1 = metastasis associated lung adenocarcinoma transcript 1, Gapdh = glyceraldehyde 3-phosphate dehydrogenase.

**Figure S5: RNAfold of full length lncRNA.** RNAfold was implemented for human (MALAT1-201, NEAT1-201) and mouse (Malat1-204, Neat1-202 (partial)) lncRNA. Circled regions provide a zoomed in area where the reads from the PNPase CLIP (arrow indicating) are found. MALAT1-201 = metastasis associated lung adenocarcinoma transcript 1, transcript variant 1, NEAT1-201 = nuclear paraspeckle assembly transcript 1, transcript variant 1, Malat1-204 = metastasis associated lung adenocarcinoma transcript 1, transcript variant 4, Neat1-202 = nuclear paraspeckle assembly transcript 1, transcript variant 2.

**Figure S6: Mitochondrial RNA targeting sequence.** (**A**) RNAfold of a 60 base-pair region of human MALAT1, NEAT1, and KCNQT1OT, surrounding the reads (circled) found to bind to PNPase. (**B**) Supervised machine learning, through Classification and Regression Trees (CART), indicating the Producer’s accuracy and Out of Bag (OoB) error for both the model and 10-fold cross validation of positively identified ncRNAs and randomly generated sequences. (**C**) The average generated area under the curve (AUC) for prediction of ncRNA found in the PNPase CLIP. (**D**) Unsupervised learning to discover clustering (n = 4 clusters generated) patterns of ncRNAs bound to PNPase. (**E**) A representative ncRNA, sharing attributes of its respective cluster were selected and RNAfold was implemented to reveal structural and compositional differences. MALAT1 = metastasis associated lung adenocarcinoma transcript 1, NEAT1 = nuclear paraspeckle assembly transcript 1, KCNQ1OT1 = KCNQ1 Opposite Strand/Antisense Transcript 1, CV = 10-fold cross validation, Positive = 50 human and 50 mouse ncRNAs determined to be bound to PNPase, Negative = 100 randomly generated sequences, Zfp469 = zinc finger protein 469, HELLPAR = HELLP associated long non-coding RNA, MFE = minimum free energy, GC = GC content, Polynucleotide Phosphorylase (PNPase).

**Figure S7: LncRNA sequence and secondary structure parameters differentially expressed between positive and negative groups.** (**A**) The log distance of nucleotides found in unpaired configurations of dot-bracket secondary structure notation, or acguD, between groups. (**B**) The log distance of unpaired nucleotides (acgu) subtracted by paired nucleotides (ACGU) in the RNA secondary structure, or acgu-ACGU, between groups. (**C**) Minimum free energy of folding and (**D**) GC content between positive and negative groups. (**E**) Expression of mitochondrial mRNA transcripts in human type 2 diabetic (n = 6) compared to non-diabetic (n = 6). (**F**) Through IntaRNA 2.0, Malat-204 and miRNA-23b-3p were mapped to the mitochondrial mt-Nd4 region to determine regions of potential binding. The circled areas correlated to the predicated interactions and the CLIP data from mice showing the binding regions of Ago2 and Fxr1. Values derived from sequencing are considered statistically significant when *Padj* < 0.05 or –Log10Padj = 1.30. Differences between molecular and biochemical assays were considered statistically different if *P* ≤ 0.05, denoted by **\***. Significance was determined through the Wald test in the R environment for sequencing comparisons. A two-sided Student’s t-test was implemented for determining significance of qPCR. All data are presented as the mean ± standard error of the mean (SEM). Positive = 50 human and 50 mouse ncRNAs determined to be bound to PNPase, Negative = 100 randomly generated sequences. Malat1-204 = metastasis associated lung adenocarcinoma transcript 1, transcript variant 4, mt-Nd4 = mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 4, Argonaute-2 (Ago2), Fragile X mental retardation syndrome-related protein 1 (Fxr1), and Polynucleotide Phosphorylase (PNPase).

**References**

1. Palmer JW, Tandler B, Hoppel CL. Biochemical properties of subsarcolemmal and interfibrillar mitochondria isolated from rat cardiac muscle. *J Biol Chem* 1977;**252**:8731-8739.

2. Baseler WA, Dabkowski ER, Williamson CL, Croston TL, Thapa D, Powell MJ, Razunguzwa TT, Hollander JM. Proteomic alterations of distinct mitochondrial subpopulations in the type 1 diabetic heart: contribution of protein import dysfunction. *Am J Physiol Regul Integr Comp Physiol* 2011;**300**:R186-200.

3. Baseler WA, Dabkowski ER, Jagannathan R, Thapa D, Nichols CE, Shepherd DL, Croston TL, Powell M, Razunguzwa TT, Lewis SE, Schnell DM, Hollander JM. Reversal of mitochondrial proteomic loss in Type 1 diabetic heart with overexpression of phospholipid hydroperoxide glutathione peroxidase. *Am J Physiol Regul Integr Comp Physiol* 2013;**304**:R553-565.

4. Dabkowski ER, Baseler WA, Williamson CL, Powell M, Razunguzwa TT, Frisbee JC, Hollander JM. Mitochondrial dysfunction in the type 2 diabetic heart is associated with alterations in spatially distinct mitochondrial proteomes. *Am J Physiol Heart Circ Physiol* 2010;**299**:H529-540.

5. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;**72**:248-254.

6. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;**25**:402-408.

7. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015;**12**:357-360.

8. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* 2016;**11**:1650-1667.

9. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550.

10. Wickham H. ggplot2: Elegant Graphics for Data Analysis. *Use R* 2009:1-212.

11. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.

12. McDermaid A, Monier B, Zhao J, Liu B, Ma Q. Interpretation of differential gene expression results of RNA-seq data: review and integration. *Brief Bioinform* 2018.

13. Blighe K, Rana S, Lewis S. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling., 2018.

14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**:2498-2504.

15. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnetjournal* 2011;**17**.

16. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;**10**:R25.

17. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Algorithms Mol Biol* 2011;**6**:26.

18. Han S, Liang Y, Ma Q, Xu Y, Zhang Y, Du W, Wang C, Li Y. LncFinder: an integrated platform for long non-coding RNA identification utilizing sequence intrinsic composition, structural information and physicochemical property. *Brief Bioinform* 2018.

19. Charif D, Lobry JR. SeqinR 1.0-2: a contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman HE, Vendruscolo M, eds. Structural approaches to sequence evolution: Molecules, networks, populations. New York: Springer Verlag, 2007:207-232.

20. Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 2000;**28**:1102, 1104.

21. Levin A, Lis M, Ponty Y, O'Donnell CW, Devadas S, Berger B, Waldispuhl J. A global sampling approach to designing and reengineering RNA secondary structures. *Nucleic Acids Res* 2012;**40**:10041-10052.

22. Liaw A, Wiener M. Classification and Regression by randomForest. *R News* 2002;**2**:18-22.

23. Evans JS, Murphy MA, Holden ZA, Cushman SA. Modeling species distribution and change using Random Forests. In: Drew CA, Wiersma YF, Huettmann F, eds. Predictive species and habitat modeling in landscape ecology: concepts and applications. New York: Springer, 2011:139-159.

24. Teng X, Chen X, Xue H, Tang Y, Zhang P, Kang Q, Hao Y, Chen R, Zhao Y, He S. NPInter v4.0: an integrated database of ncRNA interactions. *Nucleic Acids Res* 2019.

25. Mann M, Wright PR, Backofen R. IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res* 2017;**45**:W435-W439.

26. Kuo CC, Hanzelmann S, Senturk Cetin N, Frank S, Zajzon B, Derks JP, Akhade VS, Ahuja G, Kanduri C, Grummt I, Kurian L, Costa IG. Detection of RNA-DNA binding sites in long noncoding RNAs. *Nucleic Acids Res* 2019;**47**:e32.