Allotopic Expression of mRNAs as a Novel Gene Therapy for Encephalomyopathies

by

Nicole Marie Kotchey

Bachelors of Science, University of Pittsburgh, 2002

Submitted to the Graduate Faculty of

Molecular Pharmacology in partial fulfillment

of the requirements for the degree of

Master in Science

University of Pittsburgh School of Medicine

2007

UNIVERSITY OF PITTSBURGH SCHOOL OF MEDICINE

Department of Pharmacology

This thesis was presented

by

Nicole Marie Kotchey

It was defended on

August 20, 2007

and approved by

Donald B. DeFranco Ph.D., Professor, Pharmacology

Gregg Homanics Ph.D., Associate Professor, Anesthesiology and Pharmacology

Ferruccio Galbiati Ph.D., Assistant Professor, Pharmacology

Thesis Director: Michael Palladino Ph.D., Assistant Professor, Pharmacology

Allotopic Expression of mRNAs as a Novel Gene Therapy for Encephalomyopathies

Nicole Marie Kotchey,

MS

University of Pittsburgh, 2007

Mutation of the *mtATP6* gene, which encodes an essential subunit of the F_0F_1-ATP synthase (Complex V) in mitochondria, is known to cause a group of related encephalomyopathies. The ATP synthase acts as a hydrogen ion transporter that couples ion dissipation with ATP production. Diseases including NARP (neuropathy ataxia and retinitis pigmentosa) and MILS (maternally inherited Leigh's syndrome) are caused by missense mutations in the *ATP6* gene. *Drosophila melanogaster*, the common fruit fly, has a mitochondrial *ATP6* missense mutation that models NARP/MILS diseases. Our aim is to develop a transgenic strategy where allotopic expression of a mitochondrial-targeted *ATP6* mRNA may serve as a potential gene therapy for these devastating mitochondrial diseases. Mitochondria in metazoans are known to import nuclear encoded *5SrRNAs,* which are thought to be essential for mitochondrial protein synthesis. We utilized a cluster of 100 individual *5S rRNA* genes found at 56F region of the right arm of chromosome 2 in *Drosophila melanogaster*. Sequence comparisons revealed 17 groups of genomic variants and 14 processed *rRNA* counterparts. Identifying which, if any, of the known *5S rRNAs* are competent for mitochondrial import was integral to our proposed gene therapy approach. A protocol was developed that utilizes gradient and percoll centrifugation steps to isolate highly purified mitochondria that lack detectable cytosolic contamination. RT-PCR and cloning were used to determine which *5S rRNAs* were expressed and localized to the mitochondria. The cytoplasmic and mitochondrial derived clones and gDNA control clones

support the assertion that, at least under normal *in vivo* conditions, ~ 60 % of the identified *5S rRNA* genes are not expressed and are likely pseudogenes. One variant, *5S rRNA* III, is predominantly expressed and localized to the mitochondria. Also, 8 novel and 3 possible *5S rRNA* gene isoforms not currently categorized in sequence databases have been discovered. Clones capable of expressing chimeric *rRNA::mRNAs* in cells and *in vivo* were generated. These constructs could later be used to assess the ability of *5S rRNA* to direct mitochondrial import of “passenger” mRNAs.

TABLE OF CONTENTS

PREFACE viii

1.0 INTRODUCTION 1

2.0 IDENTIFICATION OF *5S RNAS* EXPRESSED AND LOCALIZED TO THE MITOCHONDRIA 6

2.1 MATERIALS AND METHODS 7

2.2 RESULTS 12

3.0 VALIDATION OF METHODIOLOGIES USED IN 2.0 17

3.1 MATERIALS AND METHODS 17

3.2 RESULTS 19

4.0 DISCUSSION 22

APPENDIX-A FIGURES 26

APPENDIX-B TABLES 36

BIBLIOGRAPHY 39

LIST OF TABLES

Table 1. List of Primers and their Sequences 36

Table 2. *Drosophila Melanogaster 5S rRNA* Sequences 37

Table 3. Results of *5S rRNA* Localization with Novel Variant Determination 38

LIST OF FIGURES

Figure 1. Proposed *5S rRNA* Import Pathway ... 26

Figure 2. *5S rRNA* Homology Between Species ... 27

Figure 3. 17 Genomic Variants of 5S rDNA ... 28

Figure 4. Secondary Structure of *5S rRNA* III ... 29

Figure 5. Purification of Mitochondria and Cytoplasm ... 30

Figure 6. Scheme for RNA Purification and *5S rRNA* Determination ... 31

Figure 7. PCR to Check Primers and System Conditions ... 32

Figure 8. PCR of cDNA ... 33

Figure 9. 5S rDNA Sequence Templates (-) Minus Primers ... 34

Figure 10. Constructs using 5S rDNA III ... 35

PREFACE

The author wants to acknowledge the love and support of her mother, father, and brother, without which, even this little achievement would not be possible. To my angel who is always there to put a smile on my face and make a bad or painful day better.

I would like to thank Dr. Michael Palladino for all his help and support.

I would like to thank Dr. Donald DeFranco for his kindness and understanding.

I would like to thank my thesis committee members for giving up their busy time to ask poignant questions making my thesis more exact.

Nomenclature used: *mtATP6* gene, ATP Synthase, $sesB^1$/ANT1 (stress-sensitive B) / (Adenine Nucleotide Transporter 1), encephalomyopathies, NARP (neuropathy ataxia and retinitis pigmentosa), MILS (maternally inherited Leigh's syndrome), missense mutations, progressive neuromuscular impairment, myodegeneration, allotopic expression, transgene, gene therapy

1.0 INTRODUCTION

Mitochondrial diseases are a collection of relatively common genetic disorders affecting 10-15 people per 100,000 individuals (DiMauro and Schon, 2003). This impairment of respiration or the electron transport chain has been linked to multiple neurodegenerative diseases including: Huntington's, Alzheimer's, and Parkinson's diseases, ALS, and mitochondrial encephalomyopathies (Orth and Schapira, 2001). Mutations, directly affecting mitochondrial genes or altering nuclear genes that encode proteins imported into the mitochondria, have been found to cause a similar group of encephalomyopathies. Along with prominent neurological and muscular dysfunction, patients experience marked complications in other organs or organ systems including hepatic, digestive, cardiac, and endocrine (Schapira and Cock, 1999; DiMauro and Schon, 2003).

ATP6 is one of 13 protein coding genes and 24 RNA coding genes (2 *rRNAs* and 22 *tRNAs*) in the mitochondrial genome of *Drosophila* (Lewis et al., 1995) and humans (Entelis et al, 2002). The *ATP6* gene encodes an essential subunit of the F_0F_1-ATP synthase (Complex V). Under normal conditions Complex V acts as hydrogen ion transporter that couples hydrogen ion dissipation down its electrochemical gradient to rotary ATP production from ADP and inorganic phosphate (Mitchell and Moyle, 1967; Boyer 1997; Stock et al., 2000). Mutations affecting *ATP6* cause a group of related encephalomyopathies. Diseases such as NARP (neuropathy ataxia and retinitis pigmentosa) and MILS (maternally inherited Leigh's syndrome) are caused

by missense mutations in the *ATP6* gene with hypotheses that oxidative damage and/or bioenergetic impairment due to respiratory chain dysfunction lead to the pathogenesis found in these mitochondrial encephalomyopathies (Orth and Schapira, 2001).

In 2006, Celotto et al, discovered the *ATP6*^{1} mitochondrial mutation as a maternally inherited enhancer in their *sesB*^{1}/ (ANT1) stock. *sesB*^{1} animals are stress-sensitive and are paralyzed by mechanical stress. The *sesB*^{1} strain is adult viable and previous studies in the laboratory have shown that conditional paralysis mutants are enriched for those with neurodegeneration (Palladino et al, 2002). A maternally inherited enhancer of *sesB*^{1} was discovered, genetically isolated from the *sesB*^{1} mutation and named MAT^{BS} (maternal isolate, bang sensitive; Celotto et al, 2006). In studying the mitochondria genome of MAT^{BS} one missense mutation was revealed: G-A transition leading to a glycine to glutamate codon change at 116 of the *ATP6* gene. Further research showed that the glycine residue at 116 was invariant among many species. This was a key finding, the first pathogenic mitochondrial DNA mutation to be isolated in a genetic model system (Celotto et al, 2006). *ATP6*^{1} results in neural dysfunction, progressive myodegeneration, locomotor impairment, and decreased lifespan. Furthermore, *ATP6*^{1} enhances many phenotypes of *sesB*^{1} including locomotor impairment and longevity defects. For example, reduced median lifespan compared to wild type flies in *sesB*^{1}, *ATP6*^{1}, and double mutant were 31%, 67%, and 76% respectively (Celotto et al, 2006).

Encephalomyopathies are a devastating group of disorders that can affect humans at any stage and age of life. According to DiMauro and Schon (2003), they are caused by over 150 mutations from point or missense mutations to deletions of portions of the mitochondrial genome. The main factor determining when and to what degree a person will get an encephalomyopathy is the percent heteroplasmy of the mutation. In humans, for example, the

ATP6 missense mutation L156R is aphenotypic under 70% mutant heteroplasmy (30% normal). However, mutant heteroplasmy of 70-90% typically results in the adult onset NARP, whereas, greater than 90% mutant heteroplasmy results in the more severe childhood onset MILS. The $ATP6^{1}$ mutant in *Drosophila* was characterized by clonal analysis and found to have a high mutant heteroplasmy (~ 98 ± 2%) suggesting this mutant may model the more severe MILS encephalomyopathy (Celotto et al, 2006). Like many mitochondrial diseases efficacious treatments are not currently available.

For my thesis research, we aimed to develop a transgenic strategy that utilizes allotopic expression of RNA as a potential gene therapy for the *ATP6* loss-of-function mutation. Several lines of evidence suggest mitochondria have an RNA import pathway. Since *ATP6* is a mitochondrial-encoded gene, we propose a carrier RNA molecule will be necessary to transport the wild-type *ATP6* mRNA from a nuclear source. We hypothesized that the *5S rRNA* might serve such a role since the mitochondria contain their own protein synthesis machinery and previous data suggested that *5S rRNAs* are imported into mitochondria from the nucleus (Magalhaes et al, 1998; Entelis et al, 2001).

Mitochondria, remnant of their symbiotic origin, have their own protein synthesis machinery and the mtDNA encodes the *12S rRNA* and *16S rRNA* but the *5S rRNA* can be encoded in either the nucleus or mitochondria and there is controversy as to whether the *5S rRNA* is a functional part of a mitochondrial ribosome (Entelis et al, 2001). For example, the mtDNA of species such as plants, algae and some protozoans encode for the *5S rRNA* gene and *5S rRNAs* have been found in mitochondrial ribosomes of these species (Entelis et al, 2001). In eukaryotes, including *Drosophila*, fungi, and mammals, the *5S* genes are nuclear and some believe that *5S rRNAs* are not incorporated into mitochondrial ribosomes of these species (Entelis et al, 2001).

Magalhaes et al, 1998 and Yoshionari et al, 1994 used synthetic and natural *5S rRNAs* to show that 5S can enter into the mitochondria, but this does not address whether it is part of the mitochondrial ribosome. These are two separate issues whether the *5S rRNA* is a functional part of the mitochondrial ribosome and if the *5S rRNA* can be imported from the nucleus into the mitochondria. *5S rRNA* import into the mitochondria is of principle interest as a vehicle to target passenger *mRNAs* for a viable gene therapy approach.

Previous studies in *Drosophila* estimated there to be between 100-165 *5SrRNA* genes found in a single tandem array at region 56F on the right arm of chromosome two (Spradling and Rubin, 1981). Early *5S rRNA* transcripts are 135 nucleosides (NS) with 3' post-transcriptional cleavage of the last 15 nucleosides to produce a mature 120 NS rRNA. Each gene in the array encodes a 135 bp pre-rRNA, a poly-T region thought to direct RNAPol III where to terminate transcription, and a "spacer" region of approximately 238bp (Spradling and Rubin, 1981).

Many researchers disagree whether a negatively charged RNA molecule can cross through hydrophobic membranes with an electrical potential to enter the mitochondria (Entelis et al, 2001). However, *5S rRNA* has been shown to enter mammalian mitochondria even into the internal matrix using an assay with highly purified mitoplast preparations (Magalhaes et al, 1998). Natural and synthetic *5S rRNAs* and other RNAs are known to be imported, including *tRNAs* (natural and synthetic) in eukaryotes and synthetic in mammalian cells (Magalhaes et al, 1998; Entelis et al, 2001). Nuclear encoded ribozymes, including *RNase H*, *RNase P*, and *RNase MRP* have also been found in the mitochondria of Humans (Entelis et al, 2001). The combined ability of mitochondria from yeast to mammalian culture cells to import RNA provides evidence for a mitochondrial RNA import pathway.

The presence of an RNA import pathway suggests an avenue that could be exploited for gene therapy of mitochondrial-based diseases. Mitochondria are polyploidy with an estimated 3-30 genomes per mitochondria. However, these organelles are the proverbial "power house" of the cell providing most of the energy the cell needs to carry on daily functions via the citric acid cycle and the electron transport chain, processes known to produce unstable free radicals and superoxides. As a result mutation rates in the mitochondrial genome are much higher than in the nuclear genome. When mutations compromise mitochondrial function one of several terrible diseases collectively referred to as mitochondrial encephalomyopathies results. Unfortunately, these diseases lack a useful treatment and only symptomatic care is provided. Gene therapy is effectively used for other diseases where no successful drug therapy is available, such as cystic fibrosis. However, with current technology we are not able to manipulate the mitochondrial genome. Allotopic RNA gene therapy would utilize existing gene therapy mechanisms of administration that utilize viruses to integrate a gene into the nuclear DNA. This gene could encode a fusion *rRNA::mRNA* where the rRNA is a carrier that would be recognized by the RNA import pathway and the mRNA is a WT gene referred to a passenger (Figure 1). Should the structural requirements of RNA import be identified this would open up the possibility for development of treatments for any mitochondrial encephalomyopathy.

2.0 IDENTIFICATION OF *5S RNAS* EXPRESSED AND LOCALIZED TO THE MITOCHONDRIA

To reiterate the most important aspect of my thesis research, we aimed to develop a transgenic strategy that utilized allotopic expression of RNA as a potential gene therapy for the *ATP6* loss-of-function mutation. Since *ATP6* is a mitochondrial-encoded gene, we hypothesized that a small RNA molecule would be a sufficient carrier to transport the wild-type *ATP6 mRNA* from a nuclear source. We propose that *5S rRNA* might serve this role since the mitochondria contain their own protein synthesis machinery and previous data suggested that *5S rRNAs* are imported from the nucleus (Yoshionari et al, 1994; Magalhaes et al, 1998; Entelis et al, 2002).

However, we first had to characterize all of the known and annotated *5S rRNAs* because there was very little evidence for a specific sequence or sequences found imported into mitochondria of flies or humans. Benhamou and Jordan 1975, proposed a sequence for the most transcribed cytosolic *5S rRNA* that we latter named variant III. They further suggested that many of the other 5S genes were likely pseudogenes, DNA sequences that resemble functional genes but do not express any RNA or protein. This information was a helpful start but for our project we needed specific *5S rRNAs* that had entered the mitochondria, which required us to develop mitochondrial purification protocols and an assay for *5S rRNA* expression. Only these individual *5S rRNAs* that have entered the mitochondria would be the best vehicle to target passenger *mRNAs* for a feasible gene therapy approach.

2.1 MATERIALS AND METHODS

5S rRNA Configurations. After searching for individual *Drosophila melanogaster 5S rRNA* sequences within NCBI Genome, D.m.5SJ01122, D.m.5SX87880, and D.m.5SM25016 were found (Figure 2). BLAST searches with any one of these sequences sent us to the 56F region of *Drosophila melanogaster* chromosome 2 and to the 100 individual annotated genomic 5S rDNA sequences in Flybase (http://flybase.bio.indiana.edu). DNASTAR software (Lasergene) was used to analyze the 100 sequences by homology into 17 genomic variants (Figure 3). The predicted secondary structures of the 14 (120 NS) *5S rRNAs* were made utilizing mfold (http://www.bioinfo.rpi.edu/applications/mfold). *5S rRNA* III with 48 isoforms is shown as a representative (Figure 4).

Drosophila stocks and culture. Standard cornmeal molasses fly media was used. Canton S (CS) wild type stocks were mated to expand colonies in glass jars at room temperature.

Isolation of Pure Mitochondria and Cytoplasm. All steps were performed on ice or at 4°C; centrifugation was with a Sorvall SS-34 rotor in a Sorvall RC6 Plus Centrifuge (Thermoelectron corp.). All instruments and centrifuge tubes were pre-cleaned with RNase Displace a DNA-RNase remover (Fisher, #04-355-138), solutions were filter sterilized, and filter tips were used. Approximately 2.0 g adult *Drosophila* CS flies were placed in 10 ml cold Buffer A [250 mM sucrose (Fisher, 55-3), 1 mM EDTA (Fisher, BP120-500), 50 mM Tris-HCl pH 7.4 (Fisher, BP152-5)] with 1 mM DTT and 1:100 Protease Inhibitor cocktail (Roche) added before use. The flies were homogenized 30 times with each pestle of Kontes Glass Dounce Homogenizer (OD = 27.22 mm) pestle A and B 21.38 mm and 21.44 mm, respectively. The mixture was centrifuged at 1000xg for 10 min at 4°C. The supernatant was carefully poured into a new 50 ml tube and centrifuged at 10,000xg for 15 min. The supernatant was carefully poured

into a new 50 ml tube and is now called the crude-cytoplasm. The second pellet (crude mitochondria) was resuspended, washed and centrifuged in 30 ml Buffer A at 10,000xg 3 times 10 min each. The crude-cytoplasm was spun 15 min at 60,000xg to pellet any broken mitochondria. The supernatant (~7 ml) was now mitochondria-free-cytoplasm and was carefully poured into a new 50 ml tube, labeled, and kept on ice. The mitochondrial pellet was resuspended in 1 ml Buffer B [250 mM sucrose, 1 mM EGTA (Calbiochem, #324626), 10 mM Tris-HCl pH 7.4] and carefully laid on top of a continuous percoll gradient of [2.2 ml (2.5 M sucrose), 6.65 ml (100% percoll (MP Biomedical, #19536)), 12.25 ml (10 mM Tris-HCl pH 7.4), 42 μl (0.5 M EDTA)]. The gradient was centrifuged at 60,000xg for 45 min and the mitochondrial layer was the thin red band 2/3 of the way down the tube right above the dense clear percoll layer. The mitochondria were red in this procedure because the CS flies have red eyes and the red eye pigments contaminate all aqueous layers. To remove any contaminating percoll, the mitochondria were resuspended, washed and centrifuged in 30 ml Buffer B (10,000xg, 10 min) 3 times. After the last wash the purified mitochondria were kept on ice until RNA extraction (Mitochondrial Protocol modified from Gasnier et al, 1993) (Figure 5).

LiCl/Urea RNA extraction from Mitochondria and Cytoplasm. Ten ml of 3 M LiCl (Fisher, L120-500)/6M Urea (Fisher, U15-500) stored at -30°C was added to each sample and homogenized (30X pestle A and B). The samples were placed at -30°C overnight to precipitate RNA. The samples were centrifuged at 10,000xg for 10 min, a pellet appeared and the supernatant was carefully decanted. The pellet was resuspended in 1/10 the starting volume 1X TE/0.01%SDS (1ml if started with 10 ml LiCl/Urea). Two phenol (pH 4.2 ± .2) (Fisher, BP1751-400) /chloroform (EMD, CX105514) extractions were performed: equal amounts were added, mixed well, and spun 5 min at 10,000xg. After the second extraction, the aqueous (top)

layer was removed to a new 1.5 ml eppendorf tube (Figure 6).

Ethanol precipitation of RNA. The RNA was precipitated from a 1 ml extraction with 1/40th the volume (25 µl) 4 M NaCl (Fisher, 5671-3) and 2.5 times the volume (2.5 ml) ethanol (Pharmco, #11ACS200). The precipitated RNA was stored in this form at -80°C. The RNA was centrifuged at 10,000xg for 10 min and the pellet resuspend in 50 µl DNA/RNase free water (Acros, 327390010). One µl of RNase inhibitor (Invitrogen, #155-18-012) was added. The RNA was DNase treated (2 µl of DNase at 4.5 µg/µl for 1 hr at 37°C) to remove any DNA contamination. As before, the RNA was Phenol/Chloroform extracted, precipitated and a spectrophotometer was used to quantify the amount of RNA. One µl of RNase inhibitor was added and the amount of RNA was adjusted with DNA/RNase free water such that the concentration was 5 µg/µl (Figure 6).

Reverse-Transcriptase-PCR. A 20 µl final volume reaction consisted of: 2 µl (10 mM dNTP), 2 µl (2 µM RT Primer), 2 µl (100 mM DTT), 4 µl (5X Buffer), 1 µl (RNase inhibitor), 1µl (5 µg RNA), 6 µl (DNA/RNase free water), and 2 µl RT ssIII (Invitrogen MMLVRT). Minus RT reactions were performed with 2 µl of H_2O instead of the ssIII. The reaction tubes, with the proper controls, were incubated at 45 °C for 1.5 hr then the temperature was increased to 55 °C for 1.5 hr in a Thermal Cycler (MJ Research PTC-200 Peltier Thermal Cycler) (Figure 6).

Hot Start-Touchdown PCR. This following protocol was used for all rRNAs (*5S, 28S*, *16S*) and gDNA control reactions. Starter 50 µl PCR reactions consisted of 10 µl (5X Buffer {at 1X contains 1.5 mM $MgCl_2$}), 4 µl (10 mM dNTPs (Fisher, BP, 2564-4)), 4 µl (2 µM Forward Primer), 4 µl (2 µM Reverse Primer), 4.5 µl (50 mM $MgCl_2$), 3 µl (cDNA), 20 µl (DNA/RNase free water), and 0.5 µl (Go-Taq Polymerase, Promega). PCR conditions were: 1) 94^0C → 2 min, 2) 94^0C → 30 sec, 3) 56^0C → 30 sec, 4) 72^0C → 30 sec, 5) Repeat steps 2-4 (10 times), 6) 94^0C

→ 30 sec, 7) 53^{0}C → 30 sec, 8) 72^{0}C → 30 sec, 9) Repeat steps 6-8 (10 times), 10) 94^{0}C → 30 sec, 11) 50^{0}C → 30 sec, 12) 72^{0}C → 30 sec, 13) Repeat steps 10-12 (20 times), 14) 72^{0}C → 6 min, 15) 4^{0}C → hold. Approximately 1.5 min into step (1) the program was paused and 0.5 µl of Taq was added to each tube. A 2.5% agarose gel [6 g Quick Dissolve Agarose (Gene Pure LE, #E-3199-500), 1.5 g Agarose Low Melt (Fisher, BP-1360-100), 300 ml 1X TBE, 15 µl 1% ethidium bromide] was electrophoresed to pre-stain PCR products and the results were observed on a Bio Doc-It™ Imaging System. PCR conditions were confirmed using gDNA and *5S, 28S*, and *16S* primers to insure one condition per set of reactions (Figure 7) (PCR conditions modified from Nishimura et al, 1982)

PCR-Primers. Three primers were synthesized to PCR amplify the 14 different ~120 bp *5S rRNA* products known to exist in the wild type fly: **5S F**, **5S R Deg**, and **5S R 3371** (IDT, Inc). Controls to amplify the *28SrRNA* (**28S F** and **28S R**) and *16SrRNA* (**16S F** and **16S R**) were also synthesized (IDT, Inc.) (Table 1).

Reverse Transcriptase-PCR Primers. The RT primer for the cytoplasmic control was **28S RT**; likewise the mitochondrial RT control was **16S RT**. The same 5S primers (**5S R Deg**) and (**5S R 3371**) used for regular PCR were used as RT primers (Table 1).

Cloning. The TOPO TA Cloning® Kit (Invitrogen, 45-0641) was used and contained the PCR 2.1 TOPO® Vector. Vendor's instructions were followed. Briefly, in a 6 µl reaction [1 µl (PCR Product), 1 µl (1.2 M NaCl 0.6 M $MgCl_2$), 1 µl (TOPO® vector), and 3 µl (H_2O)] were mixed and incubated 5 min at room temperature. The 1-2 µl of the TOPO® Cloning Reaction was transformed into One Shot® Competent Cells. *E.coli* were plated on LB-ampicillin plates [50 µg/ml AMP (VWR, #7177-48-2)] with 160 µl of 10 µg/ml X-gal (Fisher, AB-0479) added and placed at 37°C overnight. Individual colonies were selected that were white or light blue and

analyzed for insert by amplifying colonies in 3ml LB (50 μg/ml AMP) media overnight in a 37°C incubator with shaking. Plasmid DNA was purified using Wizzard® Plus SV Miniprep Kit (Promega) and eluted in 100 μl nuclease free H_2O. The presence of a *5S rRNA* insert was verified by an *Eco* R1 digest, (the restriction site flanked both sides of the cloning site by ~10 bp). Electrophoresis with a 2.5% agarose gel separated the restriction fragment from the vector and ~ 140 bp fragment was the ideal product.

Big Dye Sequencing Reaction. 20 μL Sequence Reactions consisted of: 4 μL (5x Buffer), 3 μL (Template), 2.5 μL (2μM Primer), 1.2 μL (Big Dye V3.1 (Applied Biosystems, #4337455)), and 9.3 μl (H_2O). Using the following conditions: 1) 96^{0}C → 30sec, 2) 55^{0}C → 30sec, 3) 60^{0}C → 30sec, 4) Repeat steps 1-3 (35 times), 5) 72^{0}C → 7 min final extension, 6) 4^{0}C → hold. The DNA was cleaned with magnetic beads, which bind to DNA, so excess primers and nucleotides can be removed as to not interfere with the sequencing signal. Briefly, 10 μL of Clean Sequence (Agencourt Bioscience Corp., #009145) and 80 μL 80% ethanol were added to each reaction and mixed. The mixture from each reaction was added to its own well of a skirted 96 well plate (PGC Scientific, #62-6042-53) and placed on top of a 96 well magnet for 5 min to allow the beads to sufficiently bind DNA. The beads in each well were washed twice with 200 μl 80% ethanol then all the liquid was removed with a vacuum aspirator. The plate was removed from the magnet, and 50 μl of H_2O was added and mixed. The plate was sent to a University of Pittsburgh sequencing facility for analysis.

NCBI BLAST. In order to determine exactly what gene was amplified in the mitochondria 28S +RT lane of figure 8, 2 μl of the PCR of the RT-reaction was used along with the TOPO TA Cloning® Kit to create clones with 100 bp fragment inserts (See Cloning above for more information). Two clones were chosen for sequencing and upon analysis were not 100%

identical. One fragment of 107 nt 28S-E named for its (position in the skirted plate) and the other was 28S-F contained 119 nt. A NCBI BLASTN with parameters nr/nt was individually performed on each sequence.

2.2 RESULTS

An extensive literature search was performed to determine which *5S rRNA* genes were expressed and if there was any information about localization of these RNAs. The result of the query suggests that little information is known. In addition, the search through Flybase revealed 100 annotated genes of the estimated 165 were found in the 56 F region on the right arm of D.m. chromosome 2. Performing a BLAST on the entire *Drosophila* Genome for 5S rDNAs revealed the 56F region plus many other less homologous single regions on other chromosomes that were smaller than 135bp. Papers published more than 3 decades ago estimated the total number of genes to be ~200 from techniques such as 5S rRNA-DNA hybridization (Procunier and Tartof, 1975; Procunier and Dunn, 1978).

Comparisons of known 120 NT 5S rDNA sequences from yeast, *Drosophila*, and humans suggested that the *Drosophila* sequences identified were of sufficient homology to described *5S rRNAs* (Figure 2). Red shaded letters represent identical sequence while white shaded lettering reveals variations among the species. Next, we examined predicted secondary structures of these 120 NS rRNAs utilizing mfold (www.bioinfo.rpi.edu/applications/mfold). This helped us to determine the predicted stability of these *5S rRNA* molecules and examine structural homology between flies and mammals. Figure 4 depicts *5S rRNA* variant III with a highly stable secondary structure with a ΔG of -49.9 kcal/mol. This variant is the most prevalent isoform of the 5S genes

containing 48/100 of the known isoforms found in Flybase (Table 2). We also wanted to determine from the literature if any other laboratory had worked on *5S rRNA* transport into the mitochondria. With potentially 100 copies of the gene we hoped to narrow down exactly what *5S rRNAs* were already used in experiments. Previous data from at least three published sources (Yoshionari et al, 1994; Magalhaes et al, 1998; and Entelis et al, 2001) suggested that *5S rRNAs* are imported into the mitochondria from the nucleus.

We performed an alignment using MegAlign software from Lasergene to categorize the 100 genomic 5S rDNAs found in Flybase. Based on sequence homology, 17 groups of genomic variants (A- Q), and 14 processed RNA counterparts (Roman Numerals: I-XIV) were established (Figure 3 & Table 2). In figure 3, the red shading represent identical sequence while yellow lettering reveals variations among the 17 genomic groups. Additionally, the black dotted line is the predicted 3' post-transcriptional cleavage site of the *5S rRNAs* fifteen NT to the left of the 3'end (Spradling and Rubin, 1981). Table 2, reveals the numerical breakdown of the 100 genomic 5S rDNAs into the 14 modified post-transcriptional 120 NS RNA variants.

With, in theory, 100 gene copies resulting in 17 genomic variants and 14 processed *5S rRNA* counterparts and three published sources showing *5S rRNAs* can be imported into the mitochondria from the nucleus; we needed to be able to efficiently purify the mitochondria on a large scale from the rest of the cytoplasm in the wild type fly. The nucleus needed to be avoided because 5S rDNAs do not contain any introns and any nuclear contamination would confound later PCR steps. A protocol modified from Gasnier et al, (1993), which utilized both gradient and percoll centrifugation steps to isolate highly purified mitochondria, was used (Figure 5). The goal was to obtain mitochondria and cytosol fractions lacking any detectable contamination from the other fraction. Figure 6 pictorially shows the stepwise events the purified mitochondrial and

cytosolic fractions underwent. The purified fractions were re-homogenized with 6 M Urea/ 3 M LiCl and the RNA was extracted with Phenol (pH 4.0) /Chloroform and precipitated in NaCl and ethanol. The RNA was DNase treated and then extracted again (See M&M for more details).

Before performing the RT-PCR reaction, optimal temperature and salt conditions were determined using WT-gDNA template. From preliminary PCR reactions, we learned that the RT-reaction would have to be placed at higher temperatures, as was the Touchdown PCR's protocol optimized in order to get the appropriate products from the highly structured RNA. Amplifying all of the variants without introducing any bias was crucial in determining which variants, if any, entered the mitochondria. A degenerate primer (**5S R Deg**) was made to determine 13 of the 14 *5S rRNAs*. One *5S rRNA*, variant IX, had its own reverse primer (**5S R 3371**) because the last ~20 NS sequence did not correspond well with the other 13 (Figure 3). As noted from Figure 7, the *5S* PCR product was ~120 bp. The *16S* and *28S* control products were 170 and just over 200 bp, respectively. These are consistent with the predicted lengths of 122, 175, and 208 bp, respectively. We were unable to produce an RT-PCR product using the reverse primer (**5S R 3371**) for *5S rRNA* XI.

In the PCR of the RT-reaction (±*5S*, ±*16S*, ±*28S*) the 208 bp *28S rRNA* was the cytoplasmic control, the 175 bp *16S rRNA* was the mitochondrial control and the *5SrRNAs* are around 120 bp (Figure 6). The nuclear control were the – RT lanes, which did not amplify during the RT reaction but could amplify a product if there was any nuclear/mitochondrial DNA present during PCR. Figure 8 revealed *16S rRNA* (mitochondrial) contamination from the cytoplasmic prep lanes ± RT. Importantly, there is no *28S rRNA* (cytoplasmic) contamination from the mitochondrial PCR lanes ± RT. More significant is that the –RT lanes for *5S rRNA* for both cytoplasm and mitochondria were clean from contamination. The Mito +*28S* lane

contained a 100 bp product that was the incorrect size for the *28S*. This band was cloned, sequenced and further analysis revealed PCR recombination from two non-contiguous pieces of the *28S* gene.

After we were satisfied that the mitochondrial and cytoplasmic purification protocol were optimized as seen from the RT and cDNA PCR, the next step was cloning and sequencing both mitochondrial and cytoplasmic *5S rRNAs*. For each fraction (mitochondria or cytoplasm) three independent TOPO® Cloning Reactions (A, B, and C) were performed using a different cDNA PCR sample. Cyto (A, B, and C) and Mito (A, B, and C) were individually transformed into their own vial of One Shot® Competent Cells and plated individually onto LB-Ampicillin (50 μg/ml AMP, 10 mg/ml X-GAL) plates. Individual colonies were selected from all three plates that were white or light blue and analyzed for insert by amplifying colonies in 3ml LB (50 μg/ml AMP) media overnight in a 37°C incubator with shaking. Plasmid DNA was purified and the insert was confirmed by use of the PCR 2.1 TOPO® Vector which had flanking *Eco* RI restriction sites. When 10 μl of plasmid was cut in an *Eco* RI digestion ~140 bp fragment was created (120 bp-*5S* +20 bp *Eco* RI). The digest was electrophoresed on a 2.5% agarose gel with a 100 bp marker. White colony selection and the *Eco* RI digest of the insert insured there was a 5S rDNA in the vector to sequence.

We performed the sequencing reaction and the cleanup steps of greater than 100 positive clones each from the cytoplasm and mitochondria. A sequencing primer **pCRII-TOPOR** was used to obtain insert sequence and SeqMan software program from Lasergene was used to place the *5S* sequences into contigs based on homology. Figure 9, the template for determining which variant the sequence corresponded started at base 17 and ended at base 97. These were the first base after the **5S F** primer and first base before the **5S R Deg** respectively. From analysis of

figure 9, the sequences of variants I & IV are identical between bases 17-97, so any positive clones from these two variants were indistinguishable. Interestingly, the cytoplasm contained 105 identical *5S rRNA* III clones while the mitochondria consisted of 119 *5S rRNA* III clones. The mitochondria also contained 2 *5S rRNA* VI clones. Our sequence analysis revealed over 100 (+) positive *5S rRNA* III clones from both the mitochondria and the cytoplasm. Variant III was the predominant *5S rRNA* expressed and the best vehicle to target passenger *mRNAs* to use in gene therapy.

In conclusion, RT-PCR, PCR of the cDNA, then TOPO cloning and sequencing was used to determine which of the 12 *5S rRNA* groups are localized to the mitochondria (Note: Only 13 of the 14 processed rRNAs could be individually determined with the primers synthesized: variants I & IV have NS differences only in the areas that annealed to the primer sequence so expression of variants I and IV are combined). Our *in vivo* data are consistent with earlier *in vitro* work suggesting expressed *5S RNAs* may be largely homogenous (Benhamou and Jordan, 1975).

3.0 VALIDATION OF METHODIOLOGIES USED IN 2.0

In order to validate our results that *5S rRNA* III was not only the most prevalent *5S rRNA* in the cytosol but also the most prevalent imported into the mitochondria, we had to verify that our assay was capable of detecting numerous variants and were unbiased. This was our goal in using the degenerate *5S* reverse primer, however, we needed confirmation. We performed the same amplification, cloning, and sequencing procedures with the same primers and conditions only we used purified genomic DNA as the template. We utilized a clonal analysis, as before, and hypothesized that we would now recover a mixture of 5S rDNAs clones.

3.1 MATERIALS AND METHODS

Preparation of CS gDNA. A QIAamp DNA Mini Kit (QIAGEN, 51306) was used to purifying Genomic DNA. Vendor's instructions were followed but modified by the Palladino Laboratory specifically for *Drosophila*. Briefly, 10 flies were homogenized in 80 µl cold 1X PBS in a 1.5 ml eppendorf tube. One hundred eighty µl buffer ATL was added and mixed then 20 µl Proteinase K was added the solution was vortexed and placed at 56°C for 3 hrs with occasional vortexing. The tube was centrifuged briefly then 4 µl of 100 mg/ml RNase A was added, vortexed and placed at room temperature for 2 min. Two hundred µl buffer AL was added to the

tube vortexed to mix and incubated at 70 °C for 10 min. Two hundred μl of 100% ethanol was added, vortexed to mix and briefly centrifuged. The mixture was added to a QIAamp Spin Column and centrifuged at 8000RPM for 1 min. Five hundred μl of AW1 wash buffer was added and the column was centrifuged at 8000RPM for 1 min. Five-hundred μl of AW2 wash buffer was added and the column was centrifuged at 14000RPM for 3 min. The spin column was placed in a clean 1.5 ml eppendorf tube 200 μl buffer AE was added and placed at RT for 2 min. Then the column was centrifuged at 8000RPM for 1 min.

PCR of gDNA. The same exact PCR Reactions and Conditions used for the *5S* cDNAs were used for the gDNA. Three μl of purified gDNA were used in the PCR reaction. The forward primer was **5S F** and the reverse primer was **5S R Deg** in a 50 μl reaction and the same Hot Start-Touchdown program as before.

Cloning and Sequencing. The same cloning method using the TOPO TA Cloning® Kit and then transforming into One Shot® Competent Cells was used on the PCR products from gDNA. As before, clones with inserts were verified by loss of β-galactosidase activity and restriction enzyme digest and gel electrophoresis of purified plasmid DNA. The inserts were sequenced and analyzed as before in 2.1 Material and Methods.

5S rDNA-ATP6 (20 bp) Construct. A 20 bp 5' sequence tag of *Drosophila melanogaster* WT ATP6 mtDNA was appended by PCR onto the 120 bp 5S rDNA III (see Table I and Figure 10). Briefly, PCR 2.1 TOPO® Vector plasmid DNA containing the 5S rDNA III insert was linearized with *Age* I restriction enzyme. Two PCR reactions were run with the linearized vector template and primers **Eco5S F** (1a) & **XbaAT5S R** (1b) or **EcoAT5S F** (2a) & **Xba5S R** (2b) to create a 160 bp or larger PCR Fragment (119 bp-*5S* rDNA III + 20 bp ATP6 + 10 bp *Xba* I + 10 bp *Eco* RI). Five μl of the PCR reactions from 1a, 1aCT, 1b, and 1bCT were

each digested with *Eco* RI and *Xba* I at 37°C for 20 min. The digest was electrophoresed on a 2.5% agarose gel with a 100 bp marker and 140 bp fragment from an *Eco* RI digestion of the PCR 2.1 TOPO® Vector (120 bp-*5S r*DNA III +20 bp *Eco* RI). One hundred fifty one bp is the proposed size of the digested PCR fragment (119 bp-*5S* rDNA III + 20 bp ATP6 + 6 bp *Xba* I + 6 bp *Eco* RI).

Cloning and Sequencing. The two inserts, 1a and 1b, were cloned using the same method as 2.1 into the DES TOPO® Vector (pMT/V5-His TOPOV) a *Drosophila* Expression System and then transforming into One Shot® Competent Cells. As before, clones with inserts were verified by loss of β-galactosidase activity and restriction enzyme digest and gel electrophoresis of purified plasmid DNA. The inserts were sequenced and analyzed as before in 2.1 Material and Methods.

3.2 RESULTS

The categories we used to characterize the sequences after the gDNA control was completed were (+) "positive sequence" and "variants not found in Flybase." The (+) positive sequence exactly matched one of the 12 *5S rRNA* variants between the two primers. Variants not found in Flybase, were approximately 10-20% of each group's total sequenced 5S clones.

By reviewing the gDNA control data from Table 3, the (+) sequences were: 37 *5S rRNA* I or VI clones, 29 *5S rRNA* III clones, 1 *5S rRNA* VII clone, and 1 *5S rRNA* XIV clone. This genomic data is starting to mirror the expected amounts of *5S rRNAs* in the 100 characterized from Flybase, i.e. the variants with the larger amounts of copies are starting to be represented. The observed lane for gDNA was calculated by [(observed = gDNA/Total#)*100]. The observed

was primary data, and the total # of (+) clones is 96. This proposes if we sequenced 200-500 more colonies the observed values would become closer to the expected gDNA values given in Table 3, because some of the variants with only 1 copy could become represented. The gDNA clones also helped confirming that our technique was unbiased since the genomic control PCR was able to pick up 5 of the 12 *5S rRNA* variants 68/96 (+) clones or ~71%.

An interesting occurrence was that 10-20% of the total *5S* clones sequenced in each category were variants not currently established in Flybase. While searching through the mitochondrial and cytoplasmic sequences, notes were taken on where and the frequency of the variation of the individual sequences from *5S rRNA* III. After searching through the gDNA controls a number of the same variations were observed. Guidelines were set up in order to classify the variations based if they were found in 2 or more independent clones (i.e. the clones were found in 2 or more different PCR reactions) or found in 2 or more clones. From Table 3, the light blue area represented the original I to XIV variants from Flybase except XI (**&**) which used a different reverse primer in the analysis. The tan area represents variants XV to XXII found in 2 or more independent clones (#). These variants were found in gDNA and cytoplasmic or mitochondria sequencing, a few were independent clones only from gDNA. The last group of variants consisted of 2 or more clones (*****) and was found in mitochondrial variants XXIII to XXV.

After determining that *5S rRNA* III was imported into the mitochondria, two constructs were made using 5S rDNA III. Only mitochondrial clonal inserts were chosen that contained sequence that were identical in all 135 bp to 5S rDNA III. By PCR, a 20 bp 5' sequence tag of *Drosophila melanogaster* WT ATP6 mtDNA was appended onto the 120 bp 5S rDNA III (see Table I and Figure 10). In figure 10, the restriction digest was electrophoresed on a 2.5%

agarose gel with a 100 bp marker and 140 bp fragment from an *Eco* RI digestion of the PCR 2.1 TOPO® Vector (120 bp-*5S r*DNA III +20 bp *Eco* RI). The digested PCR product was a ~ 151 bp Fragment (119 bp-*5S* rDNA III + 20 bp ATP6 + 6 bp *Xba* I + 6 bp *Eco* RI), CT1a and CT 1b did not contain any DNA in the PCR reaction. In the two reverse primers (**XbaAT5S R** and **Xba5S R**), the last base of the 5S rDNA (A) was deleted. This was to insure the 3' post-transcriptional processing of the 135 bp genomic to 120 NS RNA would not occur and prevent deletion of the 20 NS ATP6. Also, since we are unsure how *5S rRNA* enters into the mitochondria, we created the construct with the 20 bp ATP6 on either side of the *5*S rDNA (Figure 10). Whatever the mechanism of RNA import into the mitochondria we did not want it blocked by *ATP6*-mRNA secondary structure tagged onto the *5S rRNA*.

We characterized the 100 *5S rRNAs* into 14 groups or variants by sequence homology, and determined that *5S rRNA* III was the most prevalent species in both the cytoplasm and mitochondria. We will use this variant for the rest of the experiments involved in development of constructs for mRNA target gene therapy. Our PCR protocol had low bias i.e., a high copy number in the genome lead to high frequency in clones. Also our method amplified based on the amount of a certain 5S variant in the genome. We used a genomic control to insure the number of 5S rDNAs were observed in the same relative pattern as in Flybase. Our analysis provided good evidence that many variants are pseudogenes, as they were not found in the control or experimental groups. Eight novel and 3 possible 5S genes were identified that are not currently categorized in Flybase. Finally, two 5S-ATP6 constructs were cloned into DES vectors, sequenced and are ready for mRNA expression assays both *in vivo* and *in vitro*.

4.0 DISCUSSION

Our goal was to develop a transgenic strategy where allotopic expression of a mitochondrial-targeted *ATP6* mRNA may serve as a potential gene therapy for devastating mitochondrial diseases. We utilized a cluster of 100 individual *5S rRNA* genes found at 56F region of the right arm of chromosome 2 in *Drosophila melanogaster*. Sequence comparisons revealed 17 groups of genomic variants and 14 processed *rRNA*s. A multi-centrifugation and percoll gradient protocol was developed to isolate highly purified mitochondria, which lacked cytosolic contamination. RT-PCR, cloning, and sequencing were used to determine which *5S rRNAs* were expressed and localized to the mitochondria. The numerical results of the cytoplasmic, mitochondrial, and gDNA control clones support the assertion that, at least under normal *in vivo* conditions, ~ 60 % of the identified *5S rRNA* genes are not expressed and are likely pseudogenes. *5S rRNA* III, the primary expressed variant, was localized to the mitochondria. In addition to this new find, 8 novel and 3 possible *5S rRNA* isoforms not presently categorized in sequence databases have been discovered. Clones were generated capable of expressing chimeric *rRNA::mRNAs* in cells and *in vivo* (Figure 10). These constructs could later be used to assess the ability of *5S rRNA* to direct mitochondrial import of "passenger" mRNAs (Figure 2).

This research required developing an assay to both purify mitochondrial RNA and quantify isoform expression. Importantly, we validated our assay by performing the degenerate amplification and cloning protocol with purified genomic DNA, which demonstrated recovery of

a mixture of 5S rDNAs (Chapter 3.0). These control data confirmed our earlier results by demonstrating our assay, PCR and cloning methods, were capable of detecting numerous variants and did not reflect a simple bias in the amplification protocol.

Our data demonstrate that of the ~ 100 genes the cytosol is largely homogeneous with variant III being the major *5S rRNA* present (Chapter 2.0). These data are consistent with previous reports (Benhamou and Jordan, 1975). Our data also revealed the surprising finding that mitochondrial *5S rRNAs* are also relatively homogeneous and that the major isoform present is variant III (Chapter 2.0). Additionally, we identified 8 novel and 3 possible genes not currently categorized in Flybase and found data in support of the assertion that, at least under normal conditions, ~ 60 % of the identified *5S rRNA* genes are not expressed and are likely pseudogenes (Chapter 3.0).

The idea that over half of the identified *5S rRNA* genes are likely pseudogenes becomes plausible when the results of the genomic control are examined. If 100 (+) gDNA clones were sequenced and it was expected that all of the 5S variants were transcribed, then the observed total would be very close to the expected total for each variant in table 3. However, by looking at the gDNA data, only 5 different variants (I, III, IV, VI, XIV) were cloned. Variants I & IV had the most copies at 37, although from our data, they never appeared in the mitochondria or cytoplasm. From the 12 distinguishable *5S rRNA* genes, the mitochondria contained two individual variants (III, VI) the most remarkable being type III while the cytoplasm only contained variant III. This suggests that even if the other genes are capable of being transcribed the 48 copies of variant III are fully capable of the cell's translation needs under normal everyday conditions. Likewise, from our data, *5S rRNA* III may be the major variant imported into mitochondria. The other variants could have developed mutations over the years from DNA

replication or transposition in which they now differ from isoform III explaining inactivity. Another explanation is variant III is transcribed during stable 22°C adult conditions. The other variants could be actively transcribed during embryogenesis /development, stress, or disease.

There are two controversies concerning *5S rRNAs* presented briefly in the introduction of this paper. First, if *5S rRNAs* can be imported from the nucleus into the mitochondria. Previous data from at least three published sources (Yoshionari et al, 1994; Magalhaes et al, 1998; and Entelis et al, 2001) suggested that this is in fact what occurs. Our data supports the *5S rRNA* import hypothesis (Table 3). In 2001, Entalis et al, were able to import 5S into "energized" (ATP was present) mitochondria with IDPs (import directing proteins) through the GIP (general import pore). Using a human *in vitro* based assay, they found that an intact pre-protein channel, IDPs, and ATP were required for *5S rRNA* import. This seems like an evolutionary logical solution, to use a mitochondria import pathway that is already intact and functioning and adapt it for RNA use. The second controversial topic concerning *5S rRNAs* was whether it is a functional component of mitochondrial ribosomes. The consensus from articles is that mtDNA of species such as plants, algae and some protozoans encode for the *5S rRNA* gene and *5S rRNAs* have been found in mitochondrial ribosomes of these species (Entelis et al, 2001). In eukaryotes, including *Drosophila*, fungi, and mammals, the *5S* genes are nuclear and some believe that *5S rRNAs* are not incorporated into mitochondrial ribosomes of these species (Entelis et al, 2001). One hypothesis is that *Leishmania tarentolae*, *Chlorokybus atmophyticus* and *Coffea arabica* have larger genomes in general (20,992 nt, 201,763 nt, and 155,189 nt), while most species including: *Drosophila melanogaster*, *Neurospora crasse*, and *Homo sapiens* have smaller and more compact genomes (19,517 nt, 7,050 nt, and 16,571 nt).

Future directions for this project include determining the utility of *5S rRNA* III as a vehicle to target passenger RNAs suggesting a viable gene therapy approach. By PCR, we appended a 20 bp 5' sequence tag of ATP6, onto the 120 bp 5S rDNA III (see Table I and Figure 10). *Eco* RI and *Xba* I were constructed in the primers to create restriction sites in order to place the construct into expression vectors, namely pUAST and "DES TOPO" (pMT/V5-His TOPOV) a *Drosophila* Expression System. With more evidence pointing toward a RNA import pathway, our vector constructs could be transfected into *Drosophila* S2 cells and the localization of the *ATP6::5SRNA* chimera into the mitochondria determined by RT-PCR.

In conclusion, we characterized ~100 *5S rRNAs* into 14 groups or variants by sequence homology and determined that variant III was the most prevalent species in *Drosophila* mitochondria. This suggests that variant III may serve as a viable carrier of "passenger" mRNA that may be of therapeutic value. We aim to use the *5S rRNAs* to target 5S::*ATP6* mRNA into mitochondria and assess the ability of allotopic mRNA expression to rescue the *mtATP6*[1] mitochondrial mutation *in vivo*. We will generate transgenic flies expressing the chimeric 5S::*ATP6* mRNA. The long-term goal is to elucidate details of mitochondrial RNA import and determine the feasibility of exploiting this function as an avenue of gene therapy for mitochondrial disease.

APPENDIX-A FIGURES

Figure 1. Proposed *5S rRNA* Import Pathway

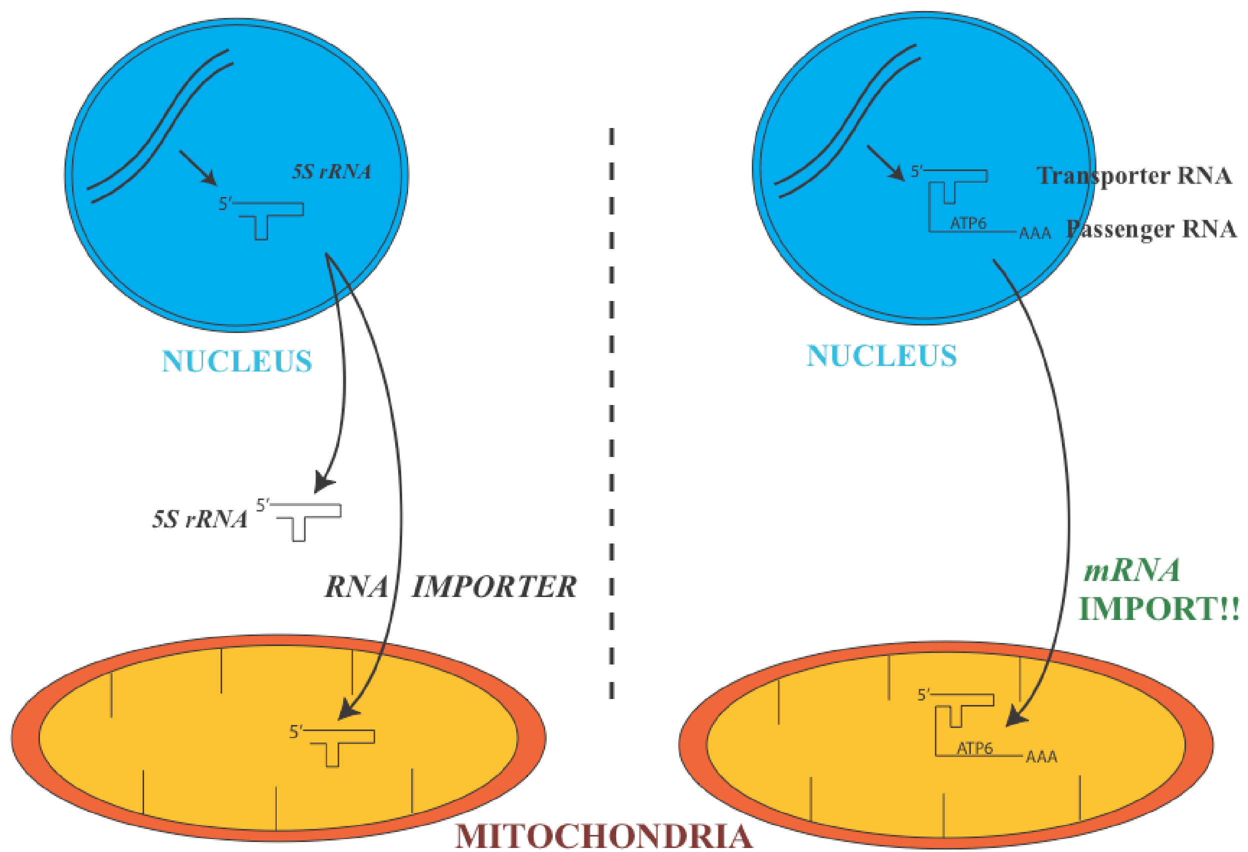


Figure 1 depicts a RNA Import Pathway where *5S rRNA* transcribed in the nucleus can be imported into the mitochondria. For mRNA allotopic expression gene therapy, the gene would encode a fusion *5SrRNA::mRNA*. The *5S rRNA* is a transporter that would be recognized by the RNA import pathway and the mRNA is a WT gene (D.m. *ATP6*) referred to as a passenger RNA.

Figure 2. ***5S rRNA* Homology Between Species**

```
D.m. 5S X87880  GCCAACGACCATACCACGCTGAATACATCGGTTCTCGTCCGATCACCGAAATTAAGCAGCGT 62
D.m. 5S J01122  GCCAACGACCATACCACGCTGAATACATCGGTTCTCGTCCGATCACCGAAATTAAGCAGCGT 62
D.m. 5S M25016  GCCAACGACCATACCACGCTGAATACATCGGTTCTCGTCCGATCACCGAAATTAAGCAGCGT 62
H.s. 5S X71800  GTCTACGGCCATACCACCCTGAACGCGCCCGATCTCGTCTGATCTCGGAAGCTAAGCAGGGT 62
S.c. 5S X67579  GGTTGCGGCCATATCTACCAGAAAGCACCGTTTCCCGTCCGATCAACTGTGTTAAGCTG-GT 61
S.p. 5S K00768  GTCTACGGCCATACCTAGGCGAAAACACCAGTTCCCGTCCGATCACTGCAGTTAAGCGTCTG 62

D.m. 5S X87880  CGGGCGCGGTTAGTACTTAGATGAGGGACCGCTTGGGAACACCGCGTGTTGTTGGCC      119
D.m. 5S J01122  CGGGCGCGGTTAGTACTTAGATGGGGGACCGCTTGGGAACACCGCGTGTTGTTGGCCTC    121
D.m. 5S M25016  CGCGGGCGGTTAGTACTTAGATGGGGGACCGCTTGGGAACACCGCGTGTTGTTGGCCT     120
H.s. 5S X71800  CGGGCCTGGTTAGTACTTGGATGGGAGACCGCCTGGGAATACCGGGTGCTGTAGGCTTT    121
S.c. 5S X67579  AGAGCCTGACCGAGTAGTGTATGGGTGACCATACGCGAAACTCAGGTGCTGCAATCT      118
S.p. 5S K00768  AGGGCCTCGTTAGTACTATGCTTGGAGACAACATGGGAATCCGGGGTGCTGTAGGCT      119
```

Figure 2 represents a comparison of known 120 NT 5S rDNA sequences from yeast, *Drosophila* and humans, ensuring that the known *Drosophila* sequences (D.m.5SJ01122, D.m.5SX87880, and D.m.5SM25016) showed sufficient homology to described *5S rRNAs*. Red shaded letters represent identical sequence while white shaded lettering reveals variations among the species. S.c. is the yeast *Saccharomyces cerevisiae* and S.p. is yeast *Sichizosaccharomyces pombe*, D.m. is *Drosophila melanogaster* and H.s. is *Homo sapiens*.

Figure 3. 17 Genomic Variants of 5S rDNA

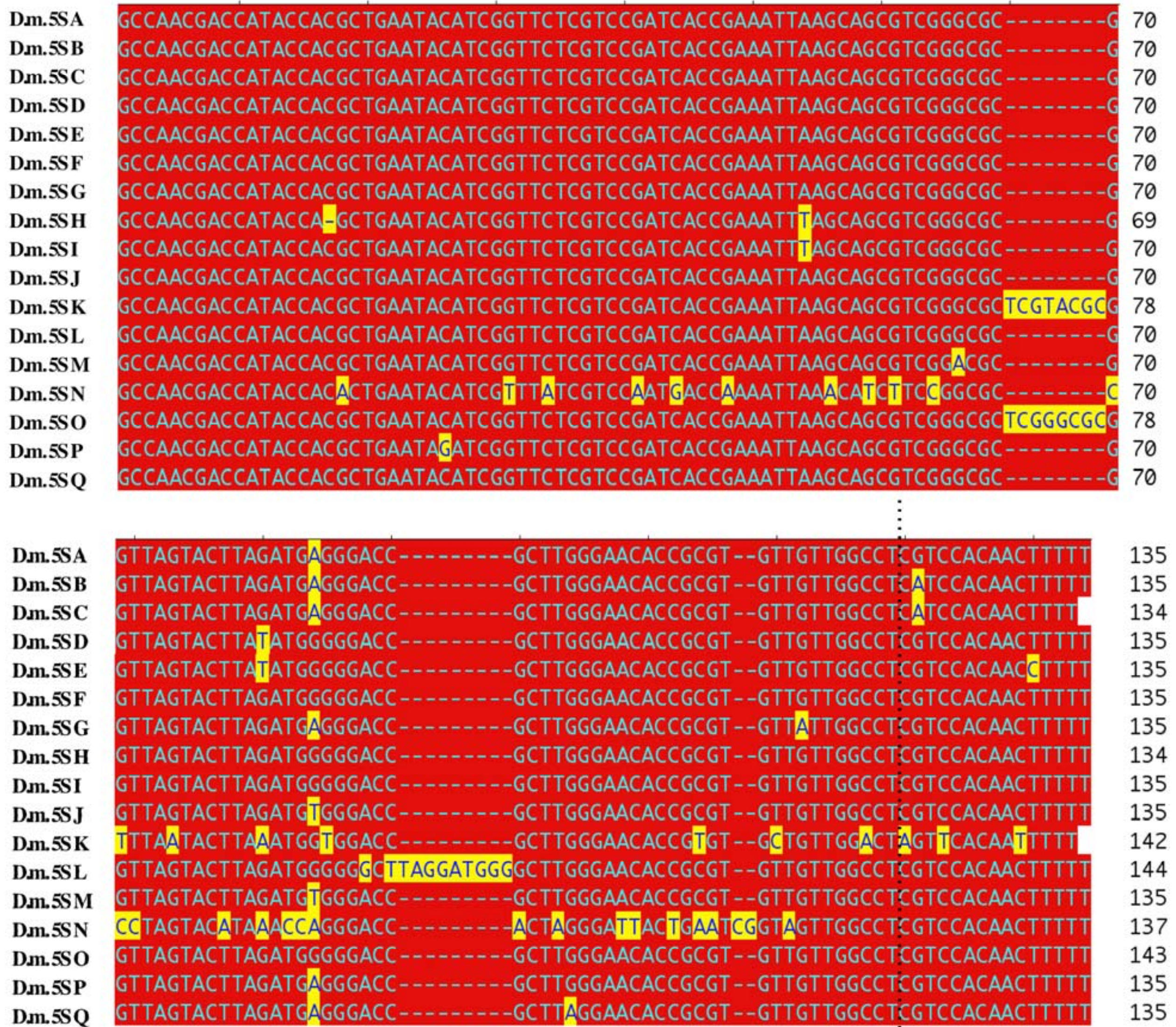


Figure 3 represents 17 groups of 5S rDNA genomic variants (A- Q). The red shading represent identical sequence while yellow lettering reveals variations among the 17 genomic groups. The black dotted line is the predicted 3' post-transcriptional cleavage site of the *5S rRNAs* (the 14 processed RNA counterparts are Roman Numerals: I-XIV found in Table 2).

Figure 4. Secondary Structure of *5S rRNA* III

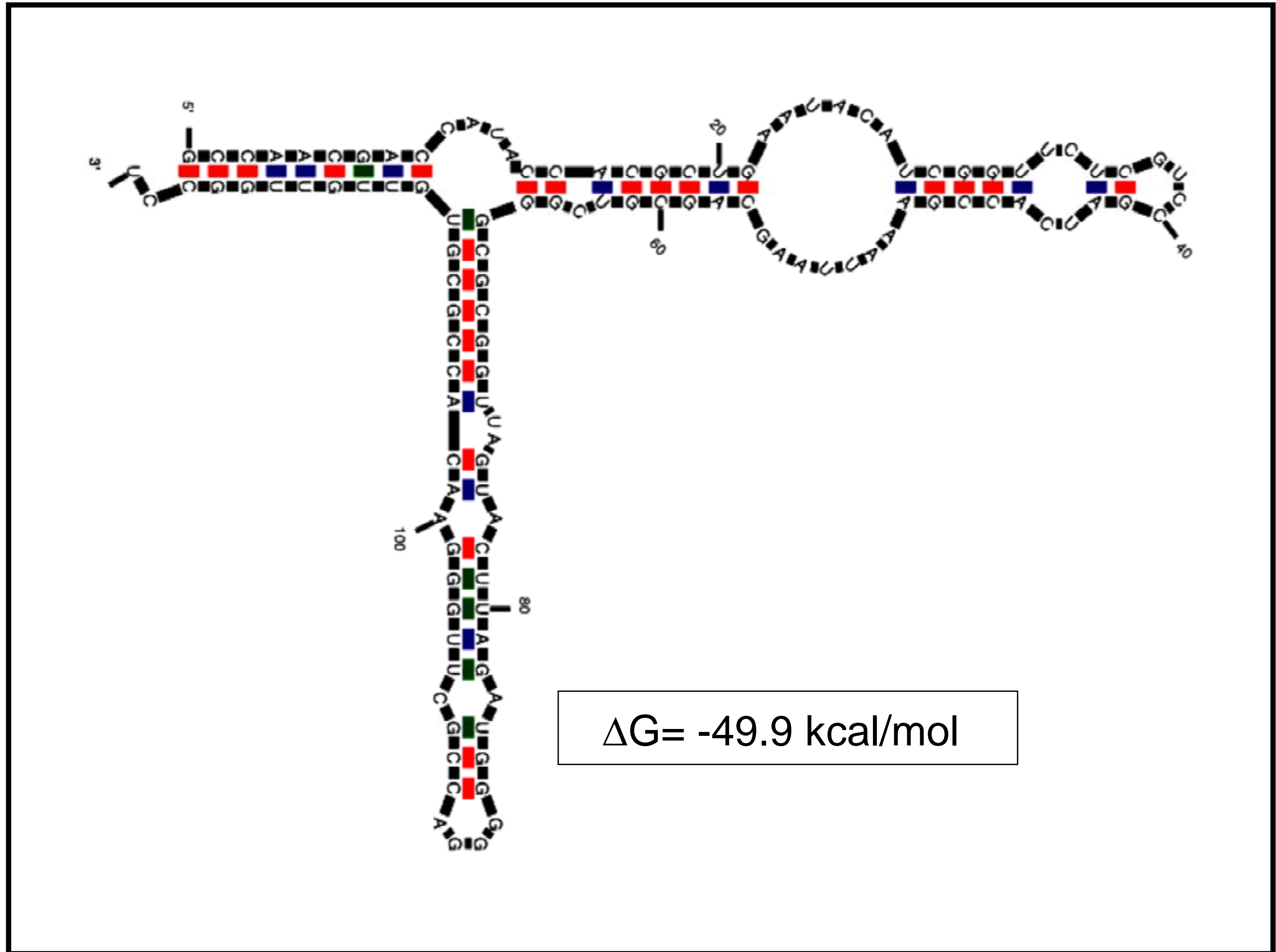
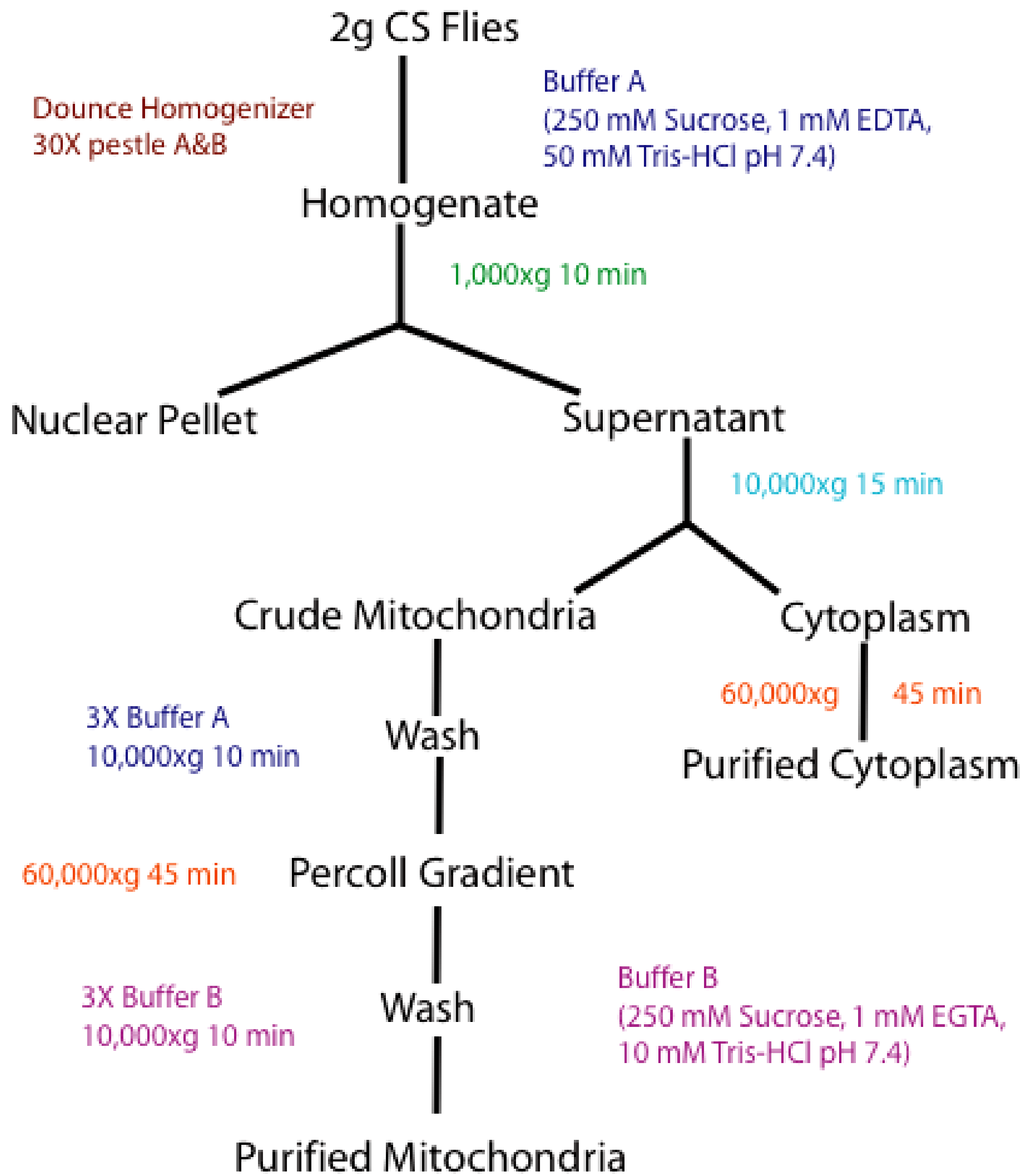


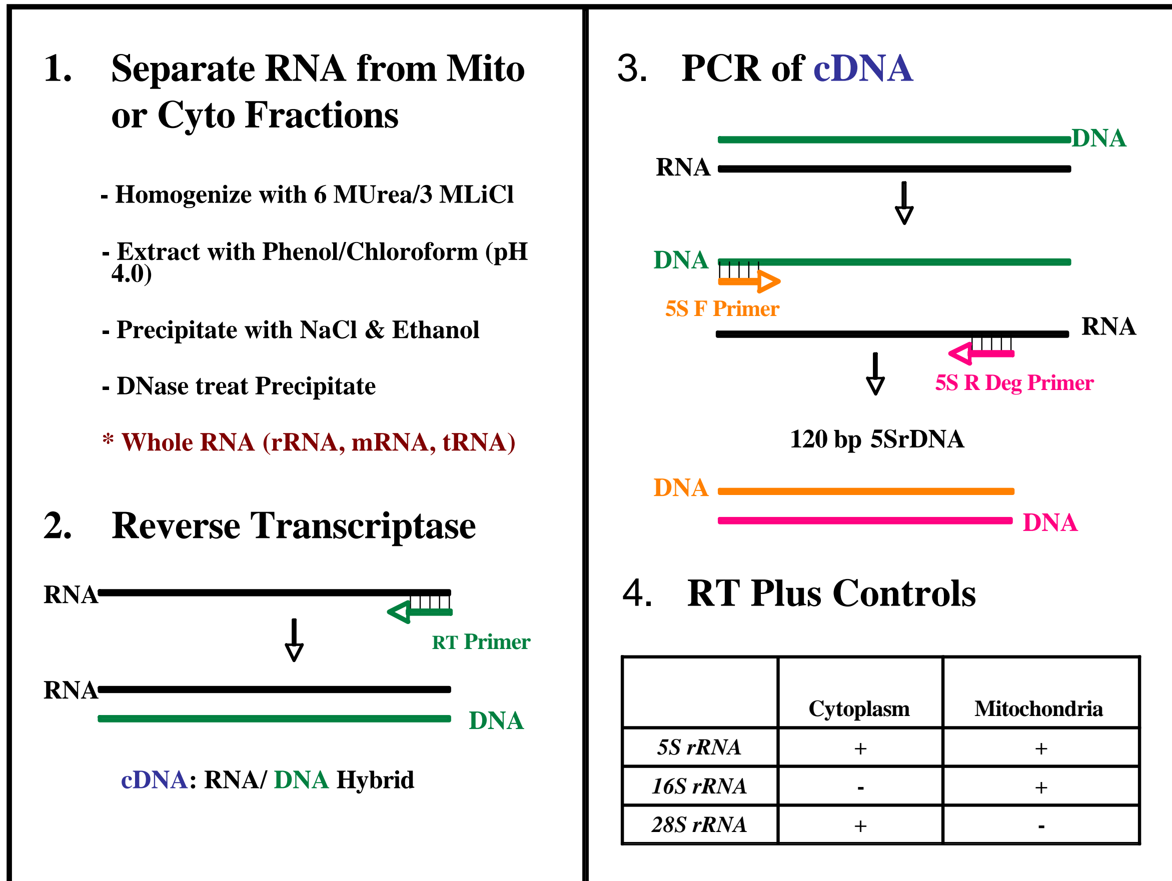
Figure 4 depicts *5S rRNA* variant III with a highly stable secondary structure with a ΔG of -49.9 kcal/mol. This variant is the most prevalent isoform of the 5S genes containing 48/100 of the known copies found in Flybase (Table 2). Our findings also indicate it was the most common variant found in the D.m. cytoplasm and mitochondria (see Table 3).

Figure 5. Purification of Mitochondria and Cytoplasm



In **figure 5**, a protocol modified from Gasnier et al, (1993), which utilized both gradient and percoll centrifugation steps to isolate highly purified mitochondria, was used. The goal was to obtain mitochondria and cytosol fractions lacking any detectable contamination from the other fraction.

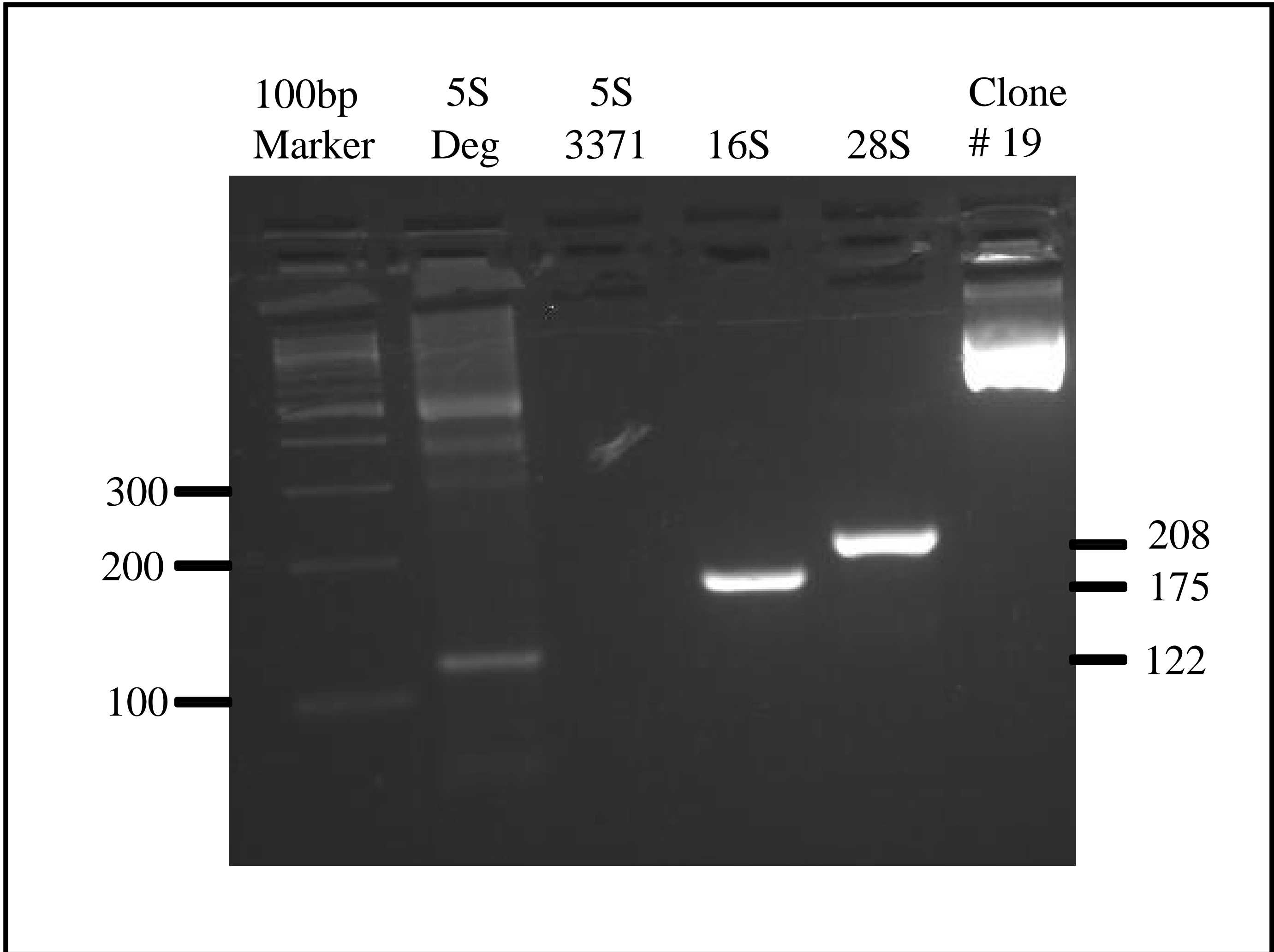
Figure 6. Scheme for RNA Purification and *5S rRNA* Determination



	Cytoplasm	Mitochondria
5S rRNA	+	+
16S rRNA	-	+
28S rRNA	+	-

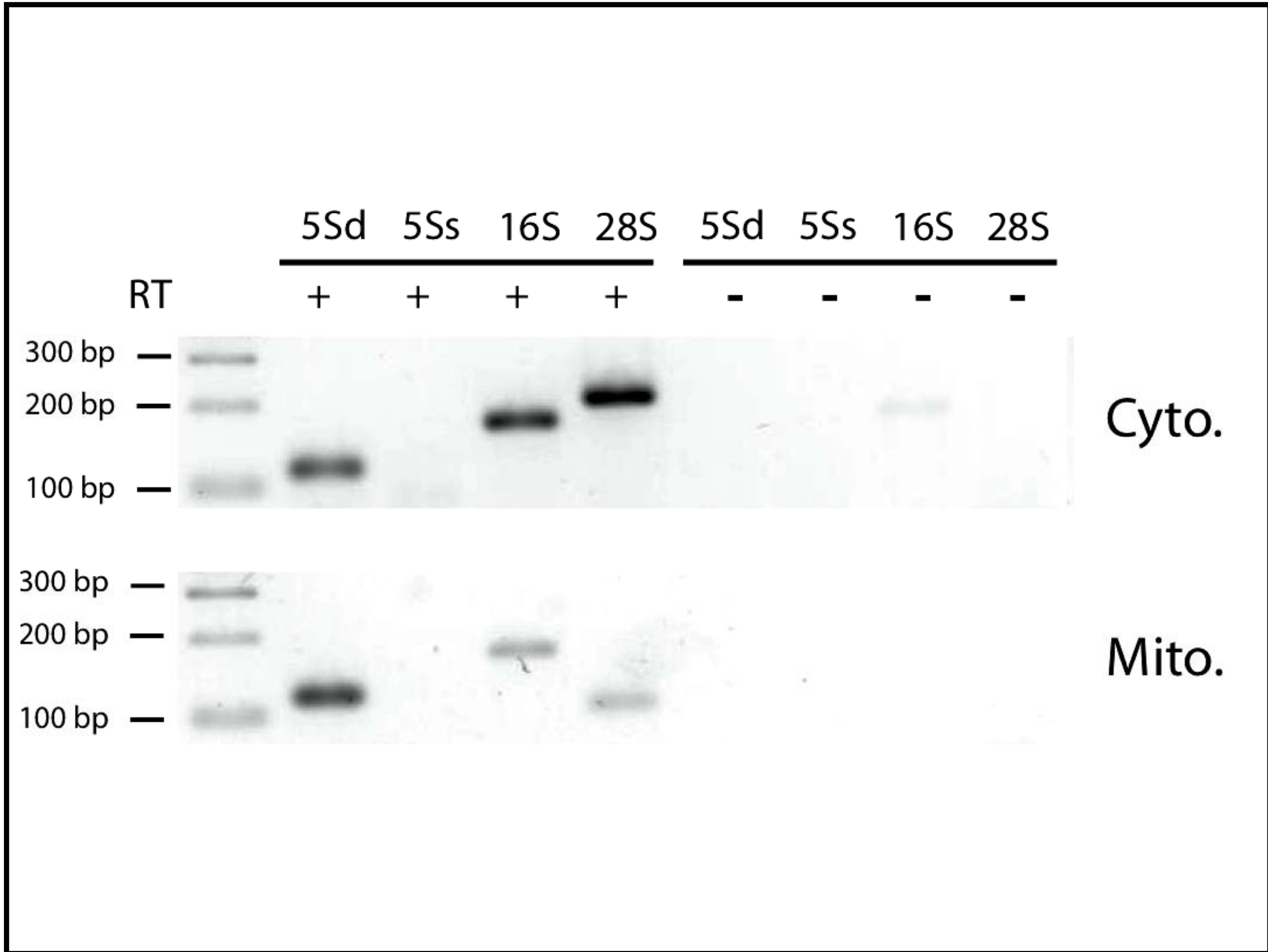
Figure 6 describes the stepwise events the purified mitochondrial and cytosolic fractions underwent. 1.) The purified fractions were re-homogenized with 6 M Urea/ 3 M LiCl and the RNA was extracted with Phenol (pH 4.0) /Chloroform and precipitated in NaCl and ethanol. The RNA was DNase treated and then extracted again (See M&M for more details). 2.) Pictorially depicts the Reverse Transcriptase reaction using RNA and a RT primer resulting in cDNA. 3.) Pictorially describes the PCR reaction using cDNA as a template and resulting in double stranded 5S rDNA. 4.) Describes the controls of the RT /PCR reactions.

Figure 7. PCR to Check Primers and System Conditions



In **figure 7**, 20 µl was electrophoresed on a 2.5% agarose gel with a 100 bp marker. The *5S* PCR product was around 120 bp, the *16S* 170 bp and the *28S* just over 200 bp, which corresponds to the predicted lengths of 122, 175, and 208 bp, respectively. We were unable to produce a product from the reverse primer (**5S R 3371**) for *5S rRNA* XI.

Figure 8. PCR of cDNA

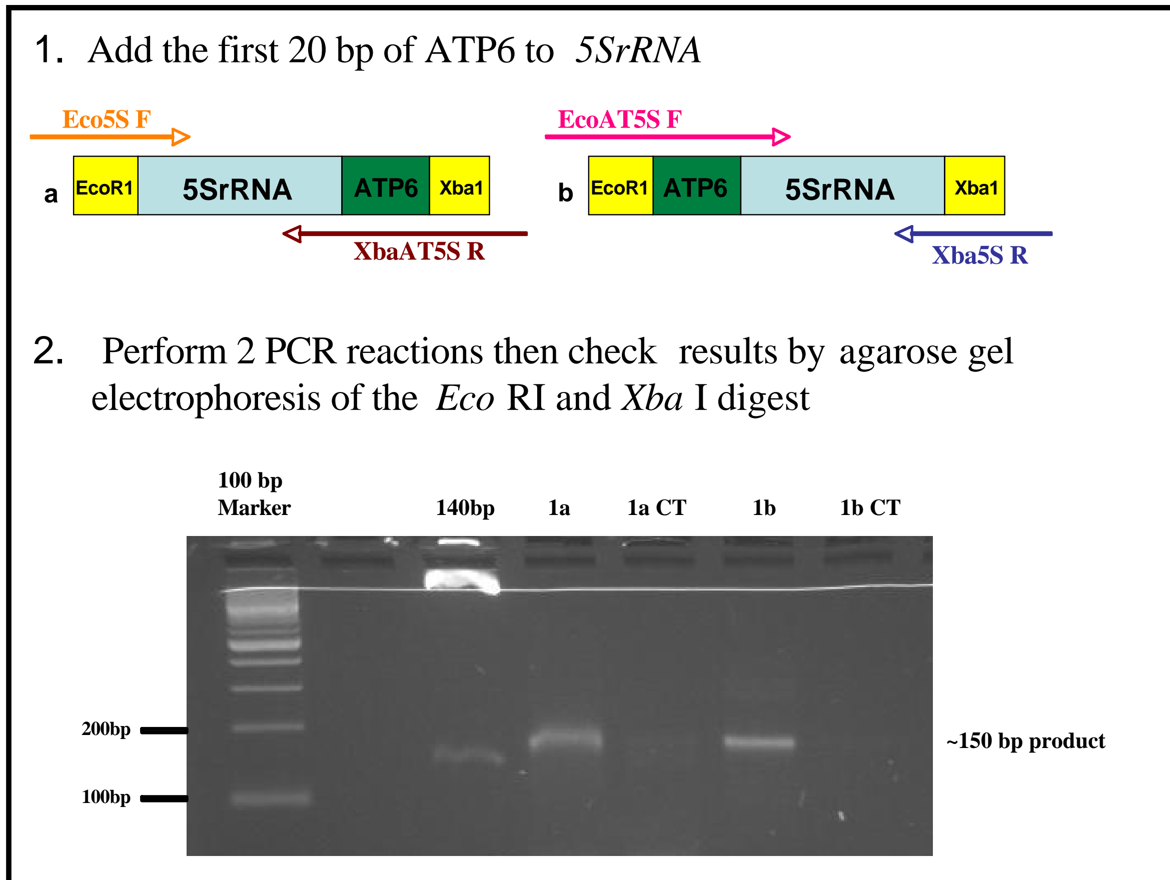


In **figure 8**, 20 µl was electrophoresed on a 2.5% agarose gel with a 100 bp marker. The 208 bp *28S rRNA* was the cytoplasmic control, the 175 bp *16 SrRNA* was the mitochondrial control and *5S rRNAs* are around 120 bp. There is *16 SrRNA* (mitochondrial) contamination from the cytoplasmic prep lanes $\pm$ RT. Importantly, there is no *28S rRNA* (cytoplasmic) contamination from the mitochondrial PCR lanes $\pm$ RT. More significant is that the –RT lanes for *5S rRNA* for both cytoplasm and mitochondria were clean from contamination. The Mito +*28S* lane contained a 100 bp product that was the incorrect size for the *28S*. This band was cloned, and sequenced and further analysis revealed PCR recombination from two non-contiguous pieces of the *28S* gene.

Figure 9.   5S rDNA Sequence Templates (-) Minus Primers

Figure 9 is the template for determining which variant the sequenced clone corresponded. The template started at base 17 and ended at base 97, these were the first base after the **5S F** primer and first base before the **5S R Deg** respectively. In this template *5S rRNA* I & IV were identical when the primer sequences were removed.

Figure 10. Constructs using 5S rDNA III



In **figure 10**, 5 μl digest was electrophoresed on a 2.5% agarose gel with a 100 bp marker and 140 bp fragment from an *Eco* RI digestion of the PCR 2.1 TOPO® Vector (120 bp-5S rDNA III +20 bp *Eco* RI). The PCR product was ~ 151 bp fragment (119 bp-5S rDNA III + 20 bp ATP6 + 6 bp *Xba* I + 6 bp *Eco* RI), and the PCR reactions for 1a CT and 1b CT did not contain any DNA.

APPENDIX-B TABLES

Table 1. List of Primers and their Sequences

Primer	Sequence
5S F	5'- GCC AAC GAC CAT ACC A -3'
5S R Deg	5'- AGK CCA AYA RCA CRC GGT GTT CC -3'
5S R 3371	5'- AGG CCA ACT ACC GAT TCA GTA ATC -3'
16S F	5'- GTT ACT TTA GGG ATA ACA GCG -3'
16S R	5'- AAC CAA CCT GGC TTA CAC CGG -3'
16S RT	5'- TTT GAT ATT TGG TCC TTT CGT AC -3'
28S F	5'- CAG GTT GAA GTC AGG CGA AAC C -3'
28S R	5'- GAC CCT AAG GCC TCT AAT CAT TCG C -3'
28S RT	5'- CCA TTT AAA GTT TGA GAA TAG GTT AAG -3'
PCRII-TOPO R	5'- GCC CAA TAC GCA AAC CG -3'
Eco5S F (1a)	5'- ACT GCG AAT TCG CCA ACG ACC ATA CCA CGC T -3'
EcoAT5S F (2a)	5'- ACT GCG AAT TC**A TG**A TAA CAA ATT TAT TTT CGC CAA CGA CCA TAC CAC GCT -3'
Xba5S R (2b)	5'- ACA AGT CTA GAG GCC AAC AAC ACG CGG TGT T -3'
XbaAT5S R (1b)	5'- ACA AGT CTA GAG AAA ATA AAT TTG TTA T**CA T**GG CCA ACA ACA CGC GGT GTT -3'

In **table 1**, are the names of the primers and their sequences used in this Thesis. The bold letters in the sequences stand for the ATG =AUG start codons for the *Drosophila* WT *ATP6* mRNA constructs.

Table 2. ***Drosophila Melanogaster 5S rRNA*** **Sequences**

Genomic Variants		rRNA Variants	
Group	#	Group	#
A	22		
B	2		
C	1	I	25
D	2		
E	1	II	3
F	48	III	48
G	5	IV	5
H	3	V	3
I	7	VI	7
J	2	VII	2
K	1	VIII	1
L	1	IX	1
M	1	X	1
N	1	XI	1
O	1	XII	1
P	1	XIII	1
Q	1	XIV	1
Total	100	Total	100

In **table 2,** based on sequence homology, 17 groups of genomic variants (A- Q), and 14 processed RNAs (Roman Numerals: I-XIV) were established. The numerical breakdown of the 100 genomic 5S rDNAs into the 14 modified post-transcriptional 120 NS RNA variants is also shown.

Table 3. Results of *5S rRNA* Localization with Novel Variant Determination

5S rRNA	Cytoplasmic	Mitochondrial	gDNA	Expected gDNA	Observed gDNA
#I or IV	0	0	37	25+5=30	38.5
II	0	0	0	3	0.0
#III	105	119	29	48	30.2
V	0	0	0	3	0.0
#VI	0	2	1	7	1.0
VII	0	0	1	2	1.0
VIII	0	0	0	1	0.0
IX	0	0	0	1	0.0
X	0	0	0	1	0.0
&XI	0	0	0	1	0.0
XII	0	0	0	1	0.0
XIII	0	0	0	1	0.0
XIV	0	0	1	1	1.0
#XV	0	4	2	0	2.1
#XVI	0	0	4	0	4.2
#XVII	0	2	2	0	2.1
#XVIII	0	0	5	0	5.2
#IXX	0	0	3	0	3.1
#XX	1	1	1	0	1.0
#XXI	0	0	4	0	4.2
#XXII	0	0	2	0	2.1
*XXIII	0	3	0	0	0.0
*XXIV	0	2	0	0	0.0
*XXV	0	2	0	0	0.0

In **table 3**, the light blue represented the original I to XIV variants from Flybase except XI (**&**) which used a different reverse primer in the analysis. The tan area represents variants are XV to XXII found in 2 or more independent clones (i.e. independent PCRs) (#). Normally the variants were found in both gDNA and the cytoplasmic or mitochondria sequencing, a few were independent clones only from the gDNA. The last group of variants consisted of 2 or more clones (*) and was found in the mitochondrial variants XXIII to XXV.

BIBLIOGRAPHY

Benhamou J, Jordan BR (1975) Nucleotide Sequence of *Drosophila melanogaster* 5S RNA: Evidence for a General 5S Model. FEBS LETTERS **62(2)**: 146-149.

Boyer PD (1997) The ATP synthase-a splendid molecular machine. Annu Rev Biochem **66**: 717-749.

Celotto AM, Frank AC, McGrath, SW, Fergestad T, Van Voorhies WA, Buttle KF, Mannella CA, Palladino MJ (2006) Mitochondrial Encephalomyopathy in *Drosophila*. The Journal of Neuroscience **26(3)**: 810-820.

DiMauro S, Schon EA (2003) Mitochondrial respiratory-chain diseases. N Engl J Med **348**: 2656-2668.

Entelis NS, Kolesnikova OA, Dogan S, Martin RP, Tarassov IA (2001) 5S rRNA and tRNA Imports into Human Mitochondria. The Journal of Biological Chemistry **276**: 45642-45653.

Entelis NS, Kolesnikova OA, Kazakova H, Brandina I, Kamenski P, Martin RP, Tarassov IA (2002) Import of Nuclear Encoded RNAs into yeast and Mitochondria: Experimental Approaches and Possible Biomedical Applications. The Journal of Genetic Engineering **24**: 191-213.

Gasnier F, Rousson R, Lerme F, Vaganay E, Louisot P. Gateau-Roesch O (1993) Use of Percoll Gradients for Isolation of Human Placenta: Mitochondria Suitable for Investigating Outer Membrane Proteins. Analytical Biochemistry **212**: 173-178.

Lewis DL, Farr CL, Kaguni LS (1995) *Drosophila melanogaster* Mitochondrial DNA: Completion of the Nucleotide Sequence and Evolutionary Comparisons. Insect Mol. Biol. **4(4)**: 263-278.

Magalhaes PJ, Andreu AL, Schon EA (1998) Evidence for the Presence of 5S rRNA in Mammalian Mitochondria. Molecular Biology of the Cell **9**: 2375-23832.

Mitchell P, Moyle J (1967) Chemiosmotic hypothesis of oxidative phosphorylation. Nature **213**: 137-139.

Nishimura M, Douce R, Akazawa T (1982) Isolation and Characterization of Metabiolically Competent Mitochondria from Spinach Leaf Protoplasts. Plant Physiol **69**: 916-290.

Orth M, Schapira HA (2001) Mitochondria and degenerative disorders. Am J Med Genet **106**: 27-36.

Palladino MJ, Hadley TJ, Ganetzky B (2002) Temperature-sensitive paralytic mutants are enriched for those causing neurodegeneration in Drosophila. Genetics **161**: 1197-1208.

Procunier JD, Dunn JR. (1978) Genetic and Molecular Organization of the 5S locus and Mutants In *D. melanogaster*. Cell **15**: 1087-1093.

Procunier JD, Tartof KD (1975) Genetic analysis of the 5S RNA in *Drosophila melanogaster*. Genetics **81**: 515-523.

Schapira HA, and Cock HR (1999) Mitochondrial myopathies and encephalomyophathies. Eur J Clin Invest **29**: 886-898.

Spradling AC, and Rubin GM (1981) Drosophila Genome Organization: Conserved and Dynamic Aspects. Annual Reviews Genetics **15**: 219-264.

Stock D, Gibbons C, Archaga I, Leslie AG, Walker JE (2000) The rotary mechanism of ATP synthase. Curr Opin Struc Biol **10**: 672-679.

Yoshionari S, Koike T, Yokogawa T, Nishikawa K, Ueda T, Miura K, Watanabe K. (1994) Existence of nuclear encoded 5S-rRNA in bovine mitochondria. FEBS Letters **338**: 137-142.