CHARACTERIZATION OF OPEN READING FRAMES IN THE TOBACCO MITOCHONDRIAL TRANSCRIPTOME

by

Benjamin Thomas Grimes

A Thesis Submitted to the
Faculty of the Graduate School at
Middle Tennessee State University
in Partial Fulfillment
of the Requirements for the Degree of
Master of Science
in Biology

Murfreesboro, TN

August 2013

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and support of many people. I would like to thank Dr. Bruce Cahoon for mentoring me through my years in graduate school. His patience and understanding, even in the face of adversity, helped me immeasurably. I am grateful to my committee, Dr. Matthew Elrod-Erickson and Dr. Rebecca Seipelt-Thiemann for their invaluable support and advice. I am also grateful to the MTSU Department of Biology for granting me scholarship funding and support over multiple projects. I would also like to thank my parents, Michael and Dixie, and my two elder sisters, Hope and Christa, for their incredible support and encouragement throughout my entire time here at MTSU. Lastly, I would like to thank my friend Alicia for her support of me through these years.

ABSTRACT

Initial annotation of the tobacco mitochondrial genome in 2005 identified numerous identifiable protein coding, tRNA, and rRNA genes plus 119 open reading frames (ORFs) with no clearly stated function. The purpose of this study was to explore expression of putative open reading frames in the *Nicotiana tabacum* mitochondrial genome. I surmised that if these gene products are transcribed and translated they may have a functional role in the tobacco mitochondria. In this study, these ORFs were screened for expression using RNA deep sequencing, qRT-PCR, and polysome analysis. Twenty-five of the 119 were found to have steady-state RNA amounts above background and 15 of those were found to be polysome associated. These ORFs were also bioinformatically compared to other genes. Some of the transcribed ORFs were found to be fragments of mitochondrial proteins, some were not identifiable while some were homologous to plastid or nuclear coding genes. Five of the expressed ORFs have been linked to cytoplasmic male sterility in other plants. The 10 remaining ORFs found to be polysome-associated have the potential to be novel mitochondrial proteins.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Page

Chapter I

Introduction

Gene expression is an incredibly complex and highly regulated process in eukaryotic cells. Understanding how and why cells and their associated organelles express certain genes has been one of the primary foci of modern molecular genetics. There are many techniques that scientists use to uncover the mysteries of gene expression such as quantitative real-time reverse transcriptase PCR, microarrays, or northern blotting; however, recent advances in sequencing technology has allowed us to visualize the entire RNA profile or transcriptome of a cell or organelle's genome. Thus, RNA-sequencing (RNA-seq) has become an exciting and invaluable tool in uncovering previously unknown expression data (reviewed in Costa et al. 2010). In this study, the tobacco mitochondrial transcriptome is analyzed in order to characterize open reading frames that were shown by RNA-seq to be expressed. The following introduction contains information to aid in the understanding of plant mitochondrial evolution, genomics, transcription, and translation biology – all of which are central to this study.

Mitochondrial Biology and Evolution

Mitochondria are semiautonomous organelles that reside in the cytoplasm of eukaryotic cells performing oxidative phosphorylation within their inner membranes; this pathway fuels aerobic respiration and provides the necessary amounts of adenosine triphosphate to maintain the energy demands of the cell. Mitochondria are also responsible for synthesizing important vitamins and cofactors for the cell as well as the metabolism of fats and amino acids (Huang et al. 2012). At one time, these organelles were once free living alpha-proteobacteria that integrated themselves into a eukaryotic ancestor through a process called endosymbiosis (reviewed in Gray 1999). Over the

course of evolution, this endosymbiont became progressively integrated with the host cell by contributing its genetic information to the nucleus through a process called endosymbiotic gene transfer (EGT); thus, a mutual relationship was formed -- the mitochondria became dependent on nuclear genes for function while the host cell became dependent on the exceptional and efficient energy production of its new symbiont (Adams et al. 2000). In fact, this dependence is evident when seen through the scope of eukaryoric evolution. Recent studies suggest that the rise of complex eukaryotic genomes could not have been achieved without the energy supply mitochondria provided to early evolving eukaryotic cells; a larger and more complex genome would have naturally produced a larger proteome, creating an enormous energy demand to maintain an increasingly elaborate cellular infrastructure (Lane and Martin 2010). Continuing along the vein of mitochondrial evolution, these organelles streamlined their double stranded, circular genomes yet retained the ability to create their own proteins, rRNAs, and tRNAs (Buchanan et al. 2000). Interestingly, all known mitochondrial genomes encode four common genes: apocytochrome b (cob), cytochrome oxidase subunit I (cox1), 26S rRNA, and 18S rRNA (Feagin, 2000). However, mitochondrial DNA (mtDNA) does not code for all the functional RNA and proteins needed for the organelle to function properly. This downsizing of mitochondrial genomes has been attributed to a biological phenomenon called “reductive evolution” (Khachane et al. 2007). Organellar genes that perform tasks redundant to that of the host experience a lack of selection and are lost throughout the course of evolution (Khachane et al. 2007; Gray 2001) For example, the original multi-subunit prokaryotic-like RNA polymerase has been replaced by a single-subunit bacteriophage-like RNA polymerase (Gray 1998). Genes for this

phage-like mitochondrial RNA polymerase (mtRNAP) are encoded in the nucleus, translated on cytoplasmic ribosomes, and transported into the mitochondria through membrane translocons in both plant and animal cells (Tracy and Stern 1995; Kuhn et al. 2004). Conversely, RNA transcript products made from mtDNA do not leave the organelle and are synthesized on their prokaryotic-like ribosomes (Buchanan et al. 2000). Another important characteristic of mitochondria is that they cannot be created de novo -- all mitochondria arise from the division or fission of pre-existing organelles (Scott and Logan 2011). This unique characteristic makes mitochondria extremely dynamic, allowing them to move, grow, and respond to cellular energy demands. Division and fission is also attributed to the chimeric nature of the mitogenome as will be discussed later in this chapter.

As we turn our lens toward the subject of this study, plant mitochondria, it is worth noting that cytoplasmic inheritance of mtDNA occurs almost exclusively through the maternal germ line in most multicellular organisms as the sperm contributes little of its cytoplasm to the zygote (Lodish et al. 2008).

Plant Mitochondria: The Genome

Plant mitochondrial genomes vary greatly from their animal counterparts -- the most noticeable differences are in their genome size and conformations. Animal mitochondrial genomes are circular and average between 15,000 to 18,000 nucleotides while plant mitochondrial genomes can range from 200,000 to over 2.6 million nucleotides and may exist as a multipartite structure (Lodish et al. 2008). Plant mitogenomes are present in heterogeneous populations of both linear and circular subgenomic dsDNA molecules that, when combined, form a complete genome in the

form of a theoretical "master circle" (Backert et al. 1997; Kosa et al. 2006; Arrieta-Montiel and Mackenzie 2011); furthermore, genetic variance between mitochondria of different plant species, especially flowering plants like tobacco, is incredibly large. To understand the reasons behind the dynamic characteristics of plant mitochondria, our focus must shift to the genetic programming of this organelle: the genome.

Full sequencing of plant mitochondrial genomes have elucidated the reasons for their large size. Marienfield et al. (1999) examined the entire genome of Arabidopsis and discovered that plant mitochondria underwent frequent recombination events in angiosperms as they evolved from bryophytes; furthermore, plant mitochondria are also exposed to and recombine with both chloroplast and nuclear genomes, which contribute to frequent capture/recapture of sequences. Although, even with this data, the size of the intergenic regions could not be fully explained until Lilly and Harvey in 2001 proposed that expanding, short degenerate repeats filled the "gaps" between genes in cucurbit mtDNA. This study was prompted after many scientists were mystified by their inability to find angiosperm mtDNA homology within existing bioinformatic databases (Clifton et al. 2004). Nevertheless, even with these large intergenic sections in a consistent state of flux, the genes that code for important mitochondrial proteins are conserved (reviewed in Palmer 1990).

The genomes themselves are quite remarkable in the fact that they may never exist in vivo as a full master molecule but rather exist as a multipartite structure of circular and linear subgenomic DNA molecules (Fauron et al. 1995; Backert et al. 1997). Frequent recombination of these subgenomic circles happens as a consequence of their recurring repeated sequences, possibly mediated by a DNA strand transfer enzyme seen

in *Arabidopsis* called AtRecA with the aid of double strand breaks in the genome (Fauron et al. 1995; Khazi et al. 2003;). RecA is an enzyme found in Escherichia coli that performs a myriad of tasks on DNA which have been found by Khazi et al. 2003 in Arabadopsis to target mitochondrial genomic sequences. This is one of many mechanisms controlling recombination in plant mitochondria. Genetic variation in plant mitogenomes is one of the most important and unique characteristics that sets these organelles apart.

Perhaps the most important factor influencing plant mitogenomic recombination is it's propensity to undergo double strand breaks (DSB), as mentioned earlier. Double strand breaks have been suggested to contribute to plant mitochondria's large size, conformations, foreign gene content, and general variability (Manchekar et al. 2006). DSBs trigger DNA repair mechanisms that can cause mutations and genome instability. The repair mechanisms use mainly two paths when dealing with DSBs: non-homologous end joining and homologous recombination (reviewed by Huertas 2010; Arrieta-Montiel and Mackenzie 2012). Non-homologous end joining or NHEJ seems to be one of the main mechanisms behind random, foreign pieces of DNA found in plant mitogenomes as in horizontal gene transfer. DNA repair mechanisms simply "stick" non-homologous DNA segments to the ends of the DSBs; moreover, there is no mechanism to make sure the foreign piece of DNA is checked for contiguity (McVey and Lee 2008). Homologous recombination is also used which simply use homologous stretches of DNA to repair the broken section (McVey and Lee 2008). Zaegel et al. 2006 suggested that the large repeated intergenic regions that contribute to the large size of plant mitochondria play a role as "hot spots" for double strand break events and aid in mitochondrial genomic

evolution. Recombination mechanisms found in the mitochondria of flowering plants seemed to be evolutionarily conserved as they favored the accumulation of genetic variation, likely increasing fitness (Bullerwell and Gray 2004). It is important to note that these recombination mechanisms have been shown to be under nuclear control (Hartmann et al. 2000). Several nuclear genes participate in controlling mitogenomic rearrangement and mutation rates; for example, the nuclear encoded MSH1, which has been shown to monitor and suppress mitogenomic recombination events, can also be suppressed which effectively increases mutation rates (Arrieta-Montiel et al. 2009). Studies using RNAi against MSH1 has caused extensive mitogenome rearrangement in *A. thaliana* (Sandhu et al. 2007). The MSH1 gene is also a homolog to the MutS gene found in *Escherichia coli*, giving further evidence to the evolution of mitochondria from bacterial ancestors and their subsequent transfer of genetic information to the nucleus of the host cell (Schofield and Hsieh 2003).

Another evolutionary driving force that contributes to the chimeric nature of plant mitochondrial genomes is their ability to readily uptake DNA through horizontal gene transfer (HGT). Richardson and Palmer 2006 showed that the mitochondria of the dicot Amborella trichopodia contained many gene sequences belonging to different species' mitogenomes. Mitochondria seem to be more prone to HGT than plastids and nuclei because they are considered to be a "discontinuous whole" as well stated by Scott and Logan 2010. With their highly recombinant DNA, propensity for genomic double strand breakage, and perpetual ability to undergo fusion and fission, these organelles set themselves apart from the rest of the cell regarding potential for genomic diversity

(Knoop et al. 2012). This diversity has large effect on the mitochondrial proteome as discussed later in this chapter.

Plant mitochondria express their own genes and in the case of tobacco, there are 36 genes total that have been annotated (Sugiyama et al. 2005). A closer look at the conserved genes in the plant mitochondrial genome shows, expectedly, that they code for essential proteins involved in oxidative phosphorylation, which include nine subunits of the NADH dehydrogenase complex, one subunit of apocytochrome b (cob), three subunits of cytochrome oxidase (cox1, cox2, and cox 3), and two subunits of the massive ATP synthase complex (atp6, atp8). As mentioned in the p

revious section, all known mitochondria do indeed code for four common genes; however, plant mitochondria exclusively contain genes for cytochrome c biogenesis (cytochrome c maturation or "ccm"), extra ribosomal proteins, and the mtt-b-like transporter protein (Fauron et al. 2004). In light of this, it is worth noting that cytochrome C has been shown to play a large role in signaling cell apoptosis in plant cells, highlighting this organelle's diverse involvement in cell maintenance (Balk and Leaver 2001).

Mitochondrial Transcription and mRNA Processing

As described above, and in the consistent nature of mitochondrial gene transfer with the nucleus, a phage-type (T7 and T3-like) nuclear encoded RNA polymerase (RNAP) is used by the mitochondria to transcribe its genes (reviewed in Hess and Borner 1999). The bacterial, multi-subunit RNAP, which was suspected to be associated with mitochondria early in its evolution within the eukaryotic cell, has been lost in all plant mitochondrial genomes and has since been replaced with one phage-like RNAP in

monocotyletons and two phage-like RNAPs in dicotyletons (Lang et al. 1997; Notsu et al. 2002; Hedtke et al. 1997). This RNAP nuclear gene family, RpoT, has been discovered in various plant species and its three genes, RpoTm, RpoTp, and RpoTmp have been traced and localized to mitochondria and plastids: RpoTm is found to exclusively localize to mitochondria, RpoTp is found to localize to chloroplasts, and RpoTmp is found to localize to both organelles (Weihe 2004). RpoTm is the standard polymerase for expressing genes in the mitochondria but it was found that RpoTmp was important in transcribing mitochondrial genes early in the development of the plant (Baba et al. 2004).

The molecular biology of true T7 bacteriophage RNAP is interesting in the fact that, as it operates within the context of viral biology, it does not need transcription factors to initiate transcription; however, the phage-like RNAP in mitochondria needs DNA binding transcription factors – such as mitochondrial transcription factor A (mtTFA) and mitochondrial transcription factor B (mtTFB) to help mediate transcription initiation (reviewed in Liere 2011). Although recently, Nayak et al. 2009 found that yeast phage-like RNAP did not need promoter binding elements to initiate transcription, but instead used its C-terminal loop to recognize promoters in vivo like its true T7 viral counterpart. Finally, mitochondrial RNA polymerases are not always encoded in the nucleus. Linear pieces of extrachromosomal DNA have been found in the mitochondria of higher plants and encode RNAPs related to viral polymerases (Liere and Borner 2012).

As mentioned above, the multi-subunit prokaryotic-like RNAP no longer exists in mitochondria; yet interestingly, sigma factors or proteins involved in directing prokaryotic RNAP to promoter sites are found to localize in mitochondria (Beardslee et al. 2002). Additionally, there have been nuclear gene products directed to the

mitochondria that act as transcription factors for expression of specific genes in Zea mays mitochondria (Newton et al. 1995). The merging of eubacterial-like RNAP initiation systems and simple phage-like RNAPs in mitochondrial transcription is interesting, as they seemingly do not possess compatible transcription initiation machinery.

Plant mitochondrial promoter biology has also proven to be interesting in comparison to animal mtDNA as plant mitochondrial genes possess multiple promoters. Since the genomes of plant mitochondria are so large and prone to repeated recombination events, multiple promoters protect transcription initiation sites for conserved genes (Kuhn et al. 2005). Transcription assays and sequence analyses have uncovered promoter motifs in both monocots and dicots (reviewed in Liere 2011). Monocots have a more straightforward promoter motif, CRTA, which is similar to the YRTA promoter for plastid nuclear encoded polymerase (NEP) with a hexanucleotide AT rich region 10 nucleotides further upstream – although the CRTA tetranucleotide promoter has been found to deviate slightly from this conserved sequence (Fey and Marechal-Drouard 1999). Dicots have a CRTA promoter that is part of a nine-nucleotide conserved sequence, CRTAaGaGA (transcription initiation site is underlined) and also possess an AT rich region upstream from the start site (Dobrowski et al. 1998b). It may be worth noting however, that recent studies have confirmed that the dicot-associated tetranucleotide motif deviates from the CRTA consensus to utilize both ATTA and RGTA sequences (Kuhn et al. 2005). These monocot and dicot promoter types are found at transcriptional start sites of mRNAs, tRNAs, and rRNAs (Binder and Brennicke 2002). As more mitogenomic data has become available from multiple species, it has become very common to see genes without obvious promoter motifs; furthermore, it has been

shown that the developmental mitochondrial RNA polymerase encoded by RpoTmp is gene specific rather than promoter specific (Kuhn et al. 2009). This opens the possibility of *cis*-acting elements directing transcription in this circumstance, highlighting the diverse nature of transcription initiation in plant mitochondria.

The tobacco transciptome data used in this study showed that the entire genome is actually expressed at some level (Figure 3.1). This is consistent with findings in *Arabidopsis* that showed transcription in the intergenic regions throughout the mitogenome resulted in large quantities of "junk" transcripts (Perrin et al. 2004; Holec et al. 2008). Even though there are unique promoter sequences commonly found just upstream of genes in the mitogenome, they can also be found scattered throughout large intergenic sequences and, as stated earlier, promoter motifs can be cryptic or not even used at all by RNA polymerase which results in massive indiscriminate low-scale expression (Dombrowski et al. 1998a; Hoffman and Binder 2002). This finding is coupled with the fact that transcription termination is not controlled well which produces long run-on transcripts; for example, Perrin et al (2004) discovered nascent *atp9* transcripts were extending far beyond their mature 3' ends. With this information at hand, scientists have arrived at the conclusion that transcription initiation (and overall expression control) is quite relaxed; thus, posttranscriptional mRNA maturation of genuine genes is pivotal in their preparation for translation (Binder et al. 2012).

Even though plant mitochondrial transcription is chaotic, the process itself is relatively well understood. After the mitochondrial RNAP binds to the promoter with the aid of transcription factors, it will begin transcribing either an oligocistronic or monocistronic nascent mRNA strand depending on the distance between the genes – if

the genes are between 3,000 and 8,000 nucleotides apart, they will most likely be transcribed as a cluster (Dombrowski et al. 1998a). Afterwards, the large run-through transcript is processed and edited to form translatable mRNAs: Pre-mRNA maturation takes place in the matrix with 5' end modification occurring through endonuclease activity or, at times, the pre-mRNA strand does not need 5' processing – the 5' end is simply the beginning of the reading frame (Binder et al. 2011). The 3' processing is much more nebulous in nature, whereas some of the transcripts have been proposed to be terminated with the help of mitochondrial transcription termination factors (mTERFs), other strands simply run on well past the end of the reading frame – the latter situation suggests that transcription termination in plant mitochondria is not critically important (Linder et al. 2005; Gagliardi and Binder 2007). Nevertheless, the 3' end is still important in protecting the RNA from exonuclease activity, as local cis acting elements aid in step loop formation (Kuhn et al. 2001). These cis elements, no doubt, are important for maintaining RNA steady state levels.

tRNA processing is undoubtedly the best understood in that RNaseP (5' endonuclease) and RNAseZ (3' endonuclease) have been characterized (Binder et al 2011; Canino et al. 2009; Kuntzmann et al. 1998). Moreover, it has been shown that mRNAs can be processed by these enzymes when they are co-transcribed with tRNAs, helping to piece together part of the puzzle on how mature mRNA transcripts are made in plant mitochondria. In addition, it is worth noting that Hinrichsen et al. (2009) showed that RNA processing was actually found to be independent of transcription -- the implications of this finding is still under investigation.

Plant Mitochondrial mRNA Stability

Understanding the importance of RNA synthesis versus degradation within the matrix is important when attempting to interpret transcriptome data. Since system-wide transcription is relatively relaxed, mitochondria must process a considerable amount of mRNA post-transcriptionally to maintain appropriate steady state RNA levels (Kuhn et al. 2009); moreover, it is important to maintain RNA levels post-transcriptionally to match appropriate stoichiometric levels of proteins needed to assemble mitochondrial machinery (Geige et al. 2000). Taken altogether, it would seem that controlled and specific degradation is crucial to maintaining steady state RNA levels since the organelle is usually overwhelmed with extraneous transcripts (Binder et al. 2012). Nevertheless, genuine transcripts needed by the organelle need to be protected from exonuclease activity. Degradation of transcripts is inhibited by 3' stem loop structures and contrary to cytosolic mRNA protection, polyadenylation seems to aid in destabilizing and degradation of the molecule (Gagliardi and Leaver 1999). Interestingly, these 3' structures are not found in many genuine gene transcripts giving evidence of other protective mechanisms such as RNA binding proteins (Binder et al. 2012). It has been found that a type of pentatricopeptide repeat (PPR) protein acts to protect the plant mitochondrial *atp6-orf79* transcript from exonucleolytic degradation by binding to the 3' and 5' ends of the message (Wang et al. 2006). Briefly, PPR proteins are part of a large protein family that associates with specific RNA sequences and have been described in great detail in association with chloroplast RNA maintenance; however, their association with mitochondria is currently under investigation. Scientists have mapped the 3' and 5' ends of all mitochondrial gene transcripts in *Arabidopsis* and found that conserved

sequences are rare (Forner et al 2007). Since PPR proteins require conserved RNA sequences for binding, it seems that PPR proteins are seldom required for transcript protection. Ultimately, it seems that much is still unknown when it comes to RNA stabilization and protection in the mitochondria.

Degradation, on the other hand, is better understood. As stated earlier, polyadenylation in the context of the mitochondria promotes RNA degradation (Gagliardi and Leaver 1999; Kuhn et al. 2001). In other cellular compartments, polyadenylation is achieved by a poly A polymerase (PAP) enzyme, but in plant mitochondria it is most likely added and maintained by polynucleotide phosphorylase (PNPase) (Kuhn et al. 2001; Martin and Keller 2004). This enzyme leads the way in 3' exonucleolytic processing along with other RNases within the matrix. This system of degradation is yet another example of how mitochondria show their similarity to prokaryotic cells, as this mechanism is seen in *E. coli* (Holec et al. 2008). *Endo*nucleolytic activity was thought to not play a role in mitochondrial RNA processing until it was shown that 5' ends of 18s and 5s rRNA extremities are formed in this way. As stated earlier, the tRNA processing enzymes are suspected to play a role in the maturation of 3' and 5' ends of all RNA species within the mitochondria (Canino et al. 2009)

Another important factor determining steady state levels is the influence of cytoplasmic transcript background. Leino et al. 2005 showed that genes in the mitochondria, especially random open reading frames, of *Brassica napus* (inherited from *Arabidopsis*) were freely transcribed and not degraded when in the presence of a male-sterile cytoplasm (CMS mutants). The same genes were rapidly degraded when analyzed

in fertile cell lines showing the control that cytoplasmic transcript background has over mitochondrial RNA degradation systems.

When analyzing a sequenced transcriptome, as in the case of tobacco in this study, it must be understood that a snapshot of expressed genes has been obtained. Therefore, depth of coverage maps are showing a brief representation of steady state RNA levels in the mitochondria. As seen in Figure 3.1, an overwhelming majority of expressed genes are represented by rRNAs. This is because ribosomal RNAs have low rates degradation rates due to their complex conformations (Kuhn et al. 2009). This causes transcriptome analyses to show very high amounts of 5s, 18s, and 26s rRNA components – even after ribosome digestion. Run-on transcript analyses have been performed on many of the genes in the plant mitogenome giving a true measurement of expression rates straight from the genome; Geige et al 2000 studied *Arabidopsis,* and discovered that rRNA run-on expression rates were not as high as previously thought and matched expression of many other important genes. Current data shows that unique mitochondrial transcription rates vary considerably between different genes – most likely due to different promoter strengths and unique stoichiometric needs of each gene (Leino et al. 2005).

Plant Mitochondrial mRNA Splicing

Mitochondrial transcripts contain introns that must be spliced out similar to eukaryotic nuclear pre-mRNAs; however, unlike the intricate spliceosome complexes used in nuclear intron processing, the mitochondria contain group I and II self-splicing introns (reviewed in Lang et al. 2007; Lodish et al. 2012). With these introns, the transcript itself performs the transesterification reaction to join exons together as well as

splicing the intronic RNA out of the message with the help of maturase proteins *in vivo* (Bonen 2008). These introns can also act as ribozymes *in vitro* without the aid of proteins as long as they are given appropriate magnesium concentrations (Mohr et al. 2006). Since group I and II introns are prone to migrating and reinserting themselves into different parts of the mitogenome, their mobility comes as no consequence to genuine genes because they can still execute splicing reactions regardless of the exons between; in addition, open reading frames that code for some maturase proteins are actually encoded within the intron (Bonen 2011). Group II introns, which comprise of most of the intron types in plant mitochondria, are susceptible to the usual mitogenomic rearrangements and must therefore act as *trans* intronic units at times, piecing together exons from different parts of the genome that have been shuffled (Bonen 2008).

It has been shown that the mechanism of these self-splicing introns can be inhibited if stabilizing proteins are unable to aid with the process (de Longevialle et al. 2007); therefore, the importance of protein structures to stabilize the three-dimensional splicing reaction is evident. One of these proteins, an RNA maturase encoded by the matR gene, is found in seed plant mitochondria. MatR has multiple promoters within that initiate transcription for both a maturase protein and a reverse-transcriptase-like protein that both aid in splicing mechanisms (Farre and Araya 1999). Discovery of this particular gene was exciting as it signaled the possibility that more maturase proteins would be found within the genome; however, this particular maturase is one of the unexpected few that has been found to reside in the plant mitogenome (Bonen 2011). This sparked a search for splicing machinery encoded elsewhere in the plant cell. Not surprisingly, it has been found that many nuclear encoded maturase and PPR proteins are

targeted to the mitochondria and are used in these unusual spliceosomes (Keren et al. 2009). Moreover, mitochondrial aminoacyl tRNA synthetases and DEAD box RNA helicases can aid in mRNA splicing in the matrix (Barkan 2004; Matthes et al. 2007) – these factors are also imported into the mitochondria, reinforcing the dependence of various mitochondrial processes on nuclear encoded genes.

It is important to understand group I and II intron systems within the scope of mRNA splicing evolution. These intron types -- found in mitochondria and chloroplasts -- are extremely important when piecing together the genesis of modern eukaryotic splicing systems. Group I/II introns contain complex stem loops that act with stunning similarity to small nuclear RNAs (Lambowitz and Zimmerly 2004; Bonen et al. 2011) used in spliceosome complexes within the nucleus (Bonen et al. 2011). A widely accepted theory proposes that these unique stem loop structures migrated into the eukaryotic cell during mitochondrial inclusion to aid in expressing its genetic system (Bonen et al. 2011). Interestingly, it has been proposed that the formation of the nuclear membrane was needed to maximize the benefit of these new mobile splicing elements in order to separate RNA maturation from premature translation; furthermore, the dynamic nature of these eventual snRNAs allowed the eukaryotic genome to increase in complexity due to exon-shuffling (Kolkman and Stemmer 2001).

Plant Mitochondrial mRNA Editing

Transcript editing is an important process leading to correctly translated proteins in plant mitochondria. Unlike mRNA splicing, which simply removes introns from immature mRNAs, editing actually changes the nucleotide sequence. Almost all mitochondrial editing is performed through the process of changing a cytosine base to

uracil (C→U); this is performed by removing the amine group, possibly by a deaminase enzyme, from the cytosine base to form a carbonyl group which creates uracil (Takenaka et al. 2008). Editing performs critical tasks such as: generating start and stop codons, enabling protein function by altering amino acid content, and restoring fertility in cases of cytoplasmic male sterility (CMS) (Hiesel et al. 1994; Gallagher et al. 2002; Kugita et al. 2003). It is important to note that not all editing results in a functional change – some are silent. Examples include editing sites that belong to the third codon position. Editing these sites may not change the original codon match and translation will not be affected (Kempken et al. 1995).

Thus far, four nuclear encoded RNA binding proteins used in plant mitochondrial mRNA editing have been found: OGR1, MEF1, MEF9, and MEF11 (Takenaka 2009). These genes encode PPR proteins that possess a wide array of abilities in RNA maintenance from splicing to editing. These proteins contain two important regions: a C-terminal glutamic acid (E) and an aspartic acid-tyrosine-tryptophan (DYW) region that associates with RNA editing and splice sites (Kotera et al. 2005). Little is known about the mechanisms PPR proteins perform the editing reaction, but Bruhs and Kempken 2012 proposed that they are involved in a much larger "editosome." Thus, the function of protein machinery in plant mitochondrial mRNA editing is still under investigation.

As mentioned earlier, it has been proposed that the formation of the nuclear membrane was needed to separate mRNA processing and editing to ensure premature translation did not take place; however, in the mitochondrial matrix there is no mechanism to prevent this from happening (Holec et al. 2008). Translation of unedited or unfinished edits can lead to malfunctioning or unusable proteins. Lu and Hanson 1994

studied the *atp6* gene in *Petunia* and discovered that protein products from completely edited transcripts were accumulating within the mitochondria. There have been studies that show polypeptides from unedited or partially edited transcripts accumulate in *Zea mays* (Phraener et al. 1996). The consequences of this are still under investigation, but from a gene regulation perspective, partially edited transcripts that have been translated can provide variation of gene products from only one gene – possibly to regulate RNA editing and provide an additional level of genetic control (Bruhs and Kempken 2012).

Plant Mitochondria: Translation and Protein Import

Mitochondria are mostly known for being ATP producing factories for the eukaryotic cell; however, they are also involved in a myriad of biochemical pathways from the metabolism of fats and amino acids to synthesizing vitamins and cofactors (Kruft er al. 2001). They contain approximately 40 genes that mainly code for proteins involved in the formation of translation machinery and the electron transport chain, but to maintain its complex biochemical role, it must import up to 3,000 nuclear gene products to complete its proteome (Knoop et al. 2011; Millar et al. 2001; Huang et al. 2011). Local translation in the mitochondria is vital for creating proteins involved in cytochrome c biogenesis and oxidative phosphorylation which are involved in apoptosis and ATP production; thus, maintenance of translation in the mitochondria is essential not only for the organelle but for the entire cell.

Mitochondria possess prokaryotic-like 70s ribosomes to translate locally transcribed mRNAs and their genome codes for three rRNAs used in the production of these ribosomes: 5s rRNA, 18s rRNA, and 26s rRNA (Lodish et al 2008). But before protein synthesis can begin, the organelle must import a complete set of tRNAs from the

cytosol (the mitogenome codes for some tRNAs but not a complete set). Import of tRNAs from the cytosol is a complex process that involves import through membrane translocons and specific target sequences on the tRNA needed by suspected targeting proteins (Duchene et al 2012). Mutations in these sequences that are required for mitochondrial targeting as well as interfacing with the aminoacyl tRNA synthetase enzyme (aaRS), can cause import failure (Salinas et al. 2005). Evolutionarily, the mitochondria streamlined protein synthesis by importing tRNAs. Depending on the needs of the organelle, it can populate the matrix with the optimal amount and type of tRNAs to achieve efficient translation (Geige and Brennicke 1999).

The journey of mitogenomic genes to mature proteins is well understood, but translation initiation in plant mitochondria has been a difficult mechanism to understand and may shed light into how translation is prioritized. Mitochondrial mRNAs seem to be indiscriminately translated regardless of whether they are mature transcripts; thus, mitochondrial proteins are post-translationally selected by the organelle for use -- as seen in *Zea mays* (Phreaner et al. 1996). There could be many reasons for this uninhibited way of translating mRNAs, but the best explanation may have to do with the absence of Shine-Dalgarno sequences that are customarily embedded slightly upstream of the start codon AUG in plant mitochondria which aid in binding the transcript to the ribosome (Geige and Brennicke 1999). It has also been shown that translation occurs from mRNAs that are lacking stop codons (Raczynska et al. 2006). In cytosolic mRNA messages, polyadenylated ends usually aid in translation termination, but in mitochondria, polyadenylation destabilizes the molecule – most likely causing polyadenylated mRNAs to never reach the ribosome because of degradation (Gagliardi and Leaver 1999). It is

interesting to note that nuclear transcripts depend on the poly-A tail for numerous reasons such as translation initiation and termination, but mitochondrial transcripts do not have this level of control; thus, this supports the overarching conclusion that translation is not controlled well here. Even though local translation is not tightly regulated, mitochondria do possess a measure of control on the proteome as seen when reacting to periods of stress (Chen et al. 2009). Several core proteins involved in programmed cell death were observed to up-regulate while proteins involved in ATP synthesis and the electron transport chain are down-regulated.

The majority of proteins that comprise the plant mitochondrial proteome must be imported as evident by the mere 40 proteins contributed by the mitogenome out of thousands needed (Millar et al. 2001). Therefore the targeting of proteins to the mitochondria is extremely important, as seen through the cell's complex and highly regulated import system. Plant mitochondria import proteins using two pathways: the general secretory pathway (Sec) and the twin-arginine translocation pathway (Tat) (Yen et al. 2002). In the Sec pathway, nuclear encoded proteins that are bound for the mitochondrial matrix contain a signal peptide at its N-terminus that is recognized by cytosolic chaperones and then transported to the translocon on the outer membrane (TOM) of the organelle. The protein is then threaded through the bilayer and the translocon on the inner membrane (TIM) where it reaches the matrix and the remaining signal peptide is cleaved by various peptidases (Lodish et al. 2008; Dalbey and Kuhn 2000). Heat shock and refolding proteins are then recruited bring the protein to its final active state. Conversely, the prokaryotic-like Tat pathway transports folded proteins across the lipid bilayer into the matrix. Proteins that cannot lose their conformation due

to complex associations with cofactors must be imported through Tat complexes – these multi-subunit complexes recognize target sequences on the mature protein and allow entry through the membrane (reviewed in Wickner and Schekman 2005; Yen et al. 2002).

Open Reading Frames in the Tobacco Chondriome

As discussed earlier, plant mitochondrial genomes undergo frequent recombination and mutation events. These events most likely produce novel and unique open reading frames scattered throughout the mitogenome. It is also common to find segments of nuclear and chloroplast homology within the plant mitogenome (Scott and Logan 2010). Briefly, open reading frames are regions of DNA that potentially code for a protein; they begin with a start codon (ATG) and the frame continues on until it reaches a termination codon. Start codons in mitochondria show similarities to their prokaryotic ancestors as they are initiated with a formylated methionine (f-met) (Reviewed in Smits et al. 2010). Stop codons in plant mitochondria are more nebulous in nature as they are terminated with a wider array of degenerate codons; furthermore, many mitochondrial reading frames are translated without stop codons as shown by Raczynska et al. 2006. One of the main questions that this experiment seeks to answer is whether transcribed open reading frames are always translated as implied by Chen et al. 2009.

The biological importance of mitochondrial open reading frames is evident when pertaining to cytomplasmic male sterility (CMS). CMS is a condition where the plant ceases to produce anthers or pollen (Leino et al. 2005). Thus far, it has been shown that CMS is controlled by unique open reading frames expressed from the mitochondrial genome (Chang et al. 2011). Control over cytoplasmic fertility in *Sorghum* has been shown to be controlled by chimeric open reading frames, orfs 25 and 265, expressed in

the mitochondria (Tang et al. 1996). The product of these ORFs produce a protein containing conserved sequences similar to other CMS-controlling open reading frames found in other species such as *Oryza* and *Triticum* (Ikagi et al. 1994; Rathburn and Hedgcoth 1991). These mitogenomic ORFs not only control CMS, but they also indirectly modulate RNA steady state levels since cytoplasmic message background effects RNA degradation (Leino et al. 2005). Maier et al. 2008 proposed that the simple expression of these open reading frames have a burdensome effect on RNA processing which can be detrimental to organellar processes.

Thesis Objectives and Hypothesis

The purpose of this study was to characterize open reading frames in the tobacco mitochondrial genome identified *in silico*. Next generation RNA-sequencing had uncovered information in tobacco's mitochondrial transcriptome that was intriguing. Preliminary RNA-sequence data showed that 25 of 119 ORFs throughout the transcriptome are potentially being transcribed. These results spark the main question of this study*:* **Are these RNA transcripts being translated into a functioning protein?** If so, this opens the door to characterizing previously unknown mitochondrial proteins. If they are not being translated, are these transcripts regulatory non-coding RNAs such as ribozymes or lncRNAs? Recent studies have only scratched the surface as to whether organellar genomes express RNA regulatory molecules as Hotto et al. 2011 recently found chloroplast-encoded antisense regulatory molecules.

The experimental design included employing two main techniques: qRT-PCR and polysomal analysis. qRT-PCR was used to confirm the RNA-seq data and required

designing 25 primer sets to specifically amplify the 25 ORFs in question, while polysome analysis was used to determine if ribosomes were attached to any of the ORF transcripts.

Hypothesis: Some open reading frames in the *Nicotiana tabacum* mitochondrial genome are transcribed and translated suggesting they may produce functional proteins.

Chapter II

Materials and Methods

Plant Material

For leaf tissue, Nicotiana tabacum, var. petit Havana seeds were grown in potting soil for ~42 days in a Percival PGC-10 incubator set for a 16 hour day/8 hour night cycle at 28°C. Seeds were first spread out in several small pots and allowed to grow for 14 days. Individual plants were then transplanted to larger pots to avoid overcrowding. After transplantation, 1/3 concentration Miracle Gro was fed to the plants once a week until harvest.

Root tissue was obtained by growing tobacco seeds in sterile Magenta boxes in a Percival PGC-10 incubator set for a 16 hour day/8 hour night cycle at 28 degrees. To prepare tobacco for root tissue growth, 2'x 2' paper towel squares were cut out and placed inside the incubation boxes. 1/3 concentration Miracle Gro was then added dropper-wise onto the towels until they were sufficiently damp. The Magenta boxes were then autoclaved on slow exhaust for 25 minutes and allowed to cool to room temperature. Seed placement was carried out with sterile forceps in a laminar hood. Tobacco seeds (var. petit Havana) were sterilized by placing them into a 2mL Eppendorf tube, adding 70% ethanol and inverting for 1 minute, washing with sterile water and vortexing for 1 minute, adding 50% bleach solution to the seeds and soaking for 15 minutes, and washing again with sterile water 3 times to ensure there was no bleach carryover. Seeds were then transferred into a sterile petri dish with a thin layer of water where they awaited placement in the Magenta boxes.

Deep Sequencing of the Tobacco Mitochondrial Transcriptome

Total RNA was extracted from 1 month old tobacco leaves using Qiagen's RNeasy kit. A sample of the total RNA was sent to the University of Illinois sequencing center. The center removed ribosomal RNAs using Ribo-Zero (Epicentre) and prepared total RNA library using Illumina's TruSeq RNAseq Sample Prep kit. Libraries were sequenced on one lane for 100 cycles from each end on a HiSeq2000 platform using a TruSeq SBS sequencing kit v.3 and analyzed with Casava 1.8. 163,836,382 sequence reads were reported with an average length of 260nt. Sequences were aligned to the tobacco mitochondrial genome (Genbank NC_006581) using DNAstar's Seqman NGen program to produce a depth of coverage map.

RNA Extraction and Purification for Transcription Analysis

100-200mg of leaf tissue was cut from healthy plants and frozen using liquid nitrogen and ground into a powder with a mortar and pestle. The powder was placed in 2ml centrifuge tubes and quickly resuspended in lysis buffer (Qiagen). RNA was then extracted with the Qiagen RNeasy Plant Extraction Kit using the instructions provided by the manufacturer except that liquid nitrogen was used during the ethanol precipitation step to maximize nucleic acid recapture. Precipitated RNA was solubilized in RNase-free water and stored at -80°C. RNA concentration was measured using a NanoDrop Lite (Thermo Scientific).

Root tissue was excised from healthy tobacco plants using a surgical blade and dissection microscope. Dissected root tissue was immediately placed in 2mL centrifuge

tubes where they were dipped in liquid nitrogen and placed on ice. The frozen root tissue was then ground into frozen powder using a mortar and pestle and solubilized in lysis buffer (Qiagen). RNA was extracted and analyzed as described above for the leaf tissue.

Polysomal RNA Isolation and Purification for Translation Analysis

Polysomal RNA was extracted and isolated using a protocol modified from Mayfield et al (1995). 200-400mg of leaf tissue was harvested from and ground into a powder with liquid nitrogen. The frozen powder was then quickly resuspended in extraction buffer (200mM Tris HCL, pH 9.0, 200mM KCL, 35mM MgCl2, 25mM EGTA, 200mM sucrose, 1% Triton-X 100, 2% BRIJ) with 0.5mg/ml heparin, 0.7% 2-mercaptoethanol, and 100mg/ml chloramphenicol added just prior to the extraction. The extracts were transferred to 2mL tubes and centrifuged at 13,000 RPM for 10min at 4°C to pellet cell debris. Supernatant was then transferred to new tubes where 7.5ul of 10% deoxycholate was added to solubilize cell membranes. Samples were centrifuged at 13,000 RPM for 10min at 4°C. The remaining supernatant was then layered on top of a two-step sucrose gradient (1.75M and 0.5M sucrose in 1x cushion buffer: 40mM Tris HCL, pH 9.0, 200mM KCL, 30mM MgCL2, 5mM EGTA) with 0.5mg/ml heparin, 0.7% 2-mercaptoethanol, and 100mg/ml chloramphenicol added just prior to creation of the layers. Polysomal RNA was pelleted by ultracentrifugation at 180,000xg for 120min at 4°C. After ultracentrifugation, the supernatant was removed by carefully pipetting and transferring the top layers to new centrifuge tubes on ice. The remaining polysomal RNA was extracted from pellets with Qiagen's Plant RNeasy Kit following the manufacturer's protocol. The nucleic acid remaining in the dense sucrose supernatant was precipitated

with 4 volumes of 95% ethanol (to dilute the sucrose) and 1/10 volume of sodium acetate. Tubes were vortexed, dipped in liquid nitrogen for 30 seconds, thawed on ice, and RNA pelleted by centrifugation at 30,000xg for 30min at 4°C. The nucleic acid from the supernatant was processed using Qiagen's Plant RNeasy Kit. All RNA captured by the RNeasy kit was re-precipitated with 3 volumes of 8M LiCl to remove residual heparin that has been shown to carry over through the extraction process and inhibit reverse transcriptase as shown by Del Prete et al. 2007. RNA concentrations were measured with a NanoDrop Lite (Thermo Scientific).

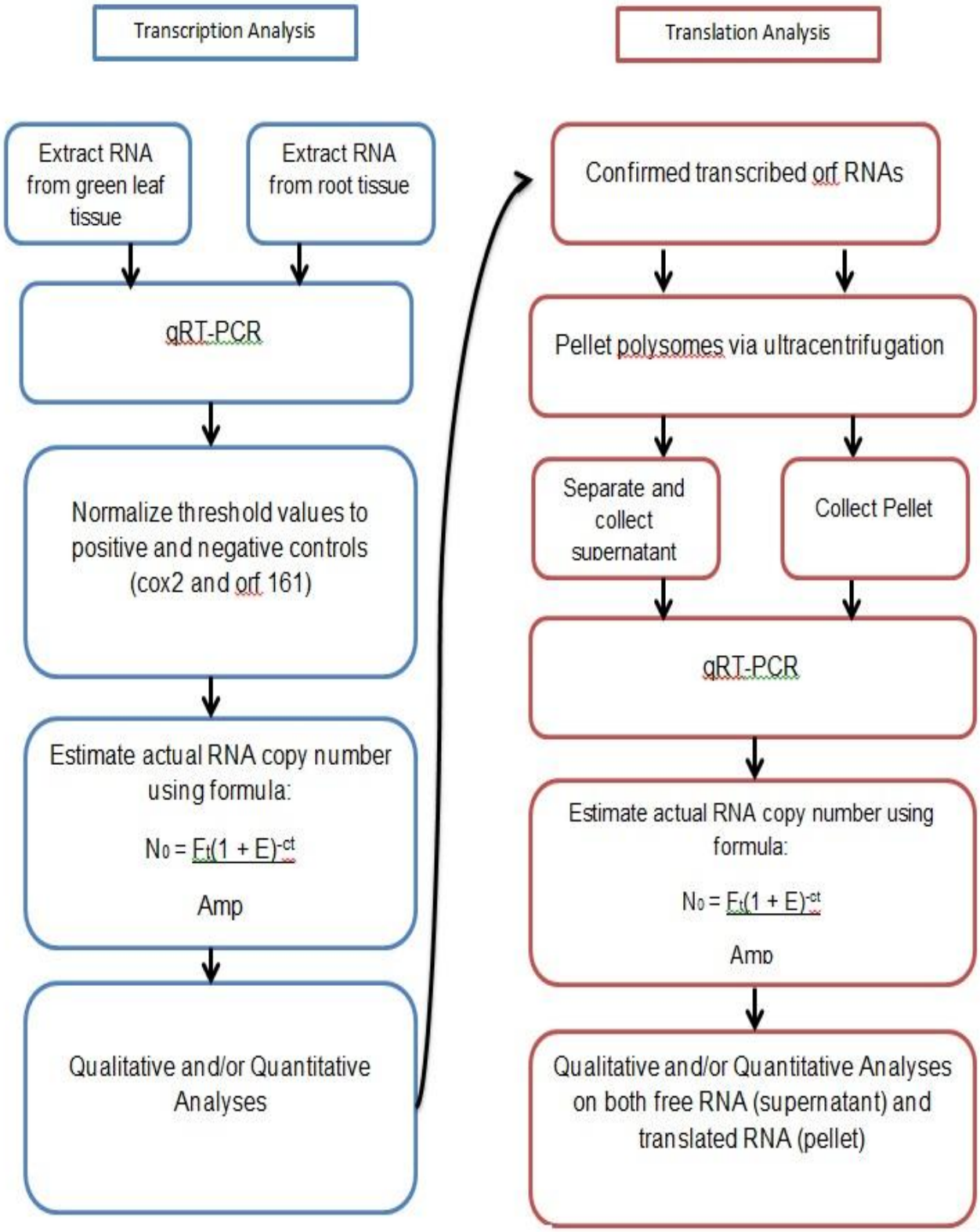

Figure 2.1: Experiment flow chart

EDTA Treatment of Polysomal Pellets

As a negative control, RNA extracts to be used for polysomal RNA isolation were incubated with 0.5M ethylenediaminetetraacetic acid, vortexed, and placed on ice for 10min to remove ribosomes. After incubation on ice, the EDTA/extract mixture was then layered on top of sucrose gradients to be ultracentrifuged at 180,000xg for 2h at 4°C. RNA was then extracted from the supernatant and pellet with the Qiagen RNeasy Plant Kit using the manufacturer's protocol.

Primer Design

All 25 primer pairs were prepared as described in Sharpe et al. (2008). Briefly, optimal annealing temperatures were tested on a CFX96 C1000 thermal cycler (Hercules, CA) programmed with eight incremental temperature points to form a gradient spanning from 50° to 65°C. To ensure the amplification of a single product, melt curves were inspected for each reaction and products were visualized on a 3% agarose gel stained with ethidium bromide. Single, dark bands located at the correct ladder position on the gel indicated good primer pair annealing temperature for all. Experimental primers were made for 17 open reading frames with amplicon sizes designed between 75-150 b.p. long for optimal qRT-PCR read runs. Elongation times for all reactions were set for 15 seconds to ensure full target amplification. The study's positive control (Cox2) and negative control (Orf 161) were designed under the same methods outlined above. Primers were purchased from Eurofins MWG Operon, Inc. (Huntsville, AL). Optimal annealing temperatures, primer pair efficiencies, and amplicon lengths can be found in Table 2.1.

Table 2.1: Optimal temperature ranges, sequences, amplicon sizes, and efficiencies for all primer sets used in this study.

Gene	Left Primer	Right Primer	**Anneal**	**Efficienc**	Amplico
Orf 133	tcctctggcattacgacct	caaaacttctaccccgagca	**53.0-62.5**	**99.5**	118
Orf 25	acaaattcgcaagctgatcc	cagcaattccccaatccta	**53.0-62.5**	**100**	83
Orf 222	tcaatcaggggctaatctg	cgaaagagccttggatcaac	**53.0-62.5**	**98.7**	84
Orf 216	gttcttggtggatgcaggac	agatccaggaatcccacctt	**53.0-62.5**	**100**	114
Orf 265b	agcgaaattggaaatgcaac	ggcttccctgtcgagaa	**53.0-62.5**	**100**	95
Orf 159b	caacctgtccgatgcttctt	aaccctatttcgcccaagtt	**53.0-62.5**	**100**	121
Orf 197	gaattccccaatcccagagt	tcctactcgcggtatgc	**53.0-62.5**	**103.6**	141
Orf 265a	ctcacccagggagtacgaaa	aaggtgcaccctcagtatgg	**53.0-62.5**	**100**	169
Orf 239	ggtcgcactttatggcattt	tgcgcagtctaccgtttc	**53.0-62.5**	**98.5**	144
Orf 147	cgcattctaggcacagatc	gggatctcttttctgcaacg	**53.0-62.5**	**106**	96
Orf177	gaggccacaagtcaacaac	ggttcccctgtgtcatcaa	**53.0-62.5**	**99**	99
Orf129b	gccttggtacggcctataca	gtactgcaacggttgggtc	**51.0-65.0**	**97**	88
Orf151	ccccctgaaaagtatcacga	ccaaagcatctatgggttgaa	**51.0-62.5**	**99**	124
Orf175	cgctgatcgtggataagac	cttcatcccggattcttcat	**55.9-62.5**	**105**	129
Orf134	gtgcggtttctgggaattta	cccctccaacaagaaaagg	**50.0-59.5**	**100**	98
Orf306	gcaagcacggttaagggata	tttgccgtccacaaaagaat	**53.0-62.5**	**100**	140
Orf138c	ttctccccttaggaccgact	tacagaagccttcgccaact	**53.0-62.5**	**98**	116
Orf144	tcgaatcggaacctttatgc	agctatcaatccccgctt	**53.0-62.5**	**100**	143
Orf118	gaagcggggattgatagctt	cctatgccagcccaaactaa	**51.0-64.1**	**104**	102
Orf160	gggttcatctctctcgacc	ccgccatagagaagagatc	**51.0-64.1**	**100**	80
Orf101d	gtttatccgggggagagatt	gaacccaattcctacggtga	**53.0-62.5**	**100**	119
Orf111c	ggcaggcaggcctatatttc	cacgtgaggggttattgctt	**53.0-62.5**	**104**	127
Orf125d	aggctgttatgggagacg	tcgaacacccccttaaaaga	**55.9-62.5**	**102**	84
Orfb	cgtcgacttcttgggaaaaa	tctttccattcctcgtgagc	**53.0-62.5**	**98**	82
Orf115	gaagtgcagcttgattgtcg	cccttgaacaaaaagccaat	**50.0-59.5**	**96**	100
Orf166b	tctggggtggtgcttctatc	agtgtttttgccccttatcg	**53.0-62.5**	**104**	81
BKGRND	gttgcaagtcttccgacgat	gaaagagttaagcgcctcca	**53.0-62.5**	**102**	146
POSCTR	gatctcaagacgcagcaaca	gcgacccaagatccatga	**51.0-65.0**	**100**	108

qRTPCR

qRT-PCR analysis was performed using a Bio-Rad CFX96 C1000 Thermal Cycler (Hercules, CA). PCR reactions were prepared with 5µl of Quanta PerfeCTa One-Step SYBR Green Mix (Gaithersburg, MD), 2ul 5pmol primers, 50ng RNA template, 0.1µl M-MLV reverse transcriptase (Promega, Madison, WI), and nuclease free water for a 10µl total reaction volume. Cycle programming started with a 30 minute reverse transcriptase step at 45° and then a 3 minute 95° Taq polymerase activation step. The normal PCR stage used a 95°C step for 30 seconds, a primer annealing temperature of 59° for 15 seconds, and an elongation step at 72°C for 15 seconds, for 39 cycles. A melt curve was created by the CFX software by programming the thermal cycler to heat from 65°C to 95°C, increasing 0.5°C every five seconds. Melt curve data was used in conjunction with DNA gel analysis to ensure single target amplification. Primer efficiency was also calculated by creating a ten-fold dilution set. Melt curves and amplification data was then used to determine percent efficiency -- primers were only used if they amplified at >95%.

DNase Cleanup of all RNA

All RNA samples were treated to remove DNA contamination. The DNase reaction was set up as follows: 12.5ul of contaminated RNA was placed in a tube with 12.5ul of DNase reaction mix (2.5ul DNase buffer, 1.5ul DNase, 0.5ul CaCl2, 8ul RNase-free water) and incubated for 30m at 37°C. After incubation, 1.25ul of EDTA was added and tubes were incubated for 10m at 75°C to deactivate the enzyme. The cleaned RNA was then precipitated with 2vol of ethanol and 1/10vol of sodium acetate. Precipitation

was aided by flash freezing tubes in liquid nitrogen for 20s and allowing to thaw on ice before centrifugation.

Calculation of Estimated Copy Numbers

The initial qRT-PCR runs were performed to confirm transcription shown by mitochondrial transcriptome deep sequencing. Three biological replicates were used for each tissue sample (roots and leaves) with two technical replicates for each biological replicate. Crossover threshold values (Ct) were used to determine initial RNA amounts by using the formula published by Alvarez et al. 2007. In this formula, the inverse crossover threshold is used as the exponent over the product of the fluorescence and efficiency of the reaction. This number is divided by the amplicon (Amp) or the size of the product in question. During the qPCR reaction, the instrument's sensitive camera detects fluorescence intensity; therefore, larger amplicon sizes will give off more light. The equation's denominator takes this variable into account and divides the initial RNA amount by the amplicon size. Since the number representing fluorescence (F_t) is a constant, an appropriate number was used to give meaningful values after calculation. In this circumstance 7×10^{10} was used for the fluorescence constant. Primer pair efficiencies were empirically determined by pooling RNAs from all conditions to be tested and a ten-fold dilution series prepared and run under optimal qPCR conditions. Software included with the CFX 1000 Real-Time PCR System was used to calculate the efficiency (1+E).

$$N_0 = \frac{F_t(1 + E)^{-ct}}{Amp}$$

Figure 2.2. Estimated copy number formula. N_0 = initial amount of mRNA, F_t = fluorescence, E = reaction efficiency, C_t = crossover threshold, Amp = amplicon size.

Statistical Analysis

All copy numbers estimates were normalized against the lowest threshold value of the six Cox2 positive control replicates. The normalization factor was calculated by dividing the Cox2 copy number estimate from each sample by the smallest Cox2 value. This normalization factor was then used to adjust the matching experimental biological and technical replicates across the entire data set. The normalized values were then used to apply standard descriptive statistics such as mean, median, mode, standard deviation, and standard error. To determine p-values from the normalized copy number data, two statistical hypothesis tests were used: t-test and Mann-Whitney rank sum. When both normality and equal variance passed, a t-test could be used; the Mann-Whitney rank sum test was used regardless of normality and equal variance to maintain statistical consistency.

Chapter III

Results

Open Reading Frame Selection and Homology Information

Deep sequencing of total RNA from tobacco leaves revealed discreetly transcribed regions in the tobacco mitochondrial genome (Figure 3.1) Most of the highly transcribed areas were associated with common mitochondrial protein coding genes. In addition, a number of transcribed regions spanned ORFs that were predicted *in silico* when the mitogenome was first sequenced (Sugiyama et al. 2005). In total, 25 open reading frames throughout the mitogenome were potentially transcribed (Figure 3.1 and Table 3.1). Candidate ORFs selected for this study did not overlap existing genes and showed expression levels with a depth of coverage ≥ 200. Predicted protein sequences for open reading frames of interest were searched against existing databases using the standard protein Basic Local Alignment Search Tool (BLASTP) provided by the website for the National Center for Biotechnology Information. Ten of the 25 open reading frames showed strong homology to known proteins in the electron transport chain and ribosome: ORFs 101d, 111c, 133, 25, 216, 265b, 159b, 134, 197, and B. Two others, ORFs 129b and 151, were related to mitoviral RNA polymerase and integrase genes (Figure 3.5). Three ORFs showing homology to conserved functional motifs originating in the chloroplast: ORFs 133, 101d, and 111c. Seven predicted ORF products showed weak to no homology against any existing proteins in the BLASTP database: ORFs 177, 175, 118, 125d, 144, 306, 222, and 239. The remaining 6 ORFs showed homology to other species' mitochondrial genomes but no conserved functional domains in their protein products: ORFs 138c, 160, 115, 166b, and 147.

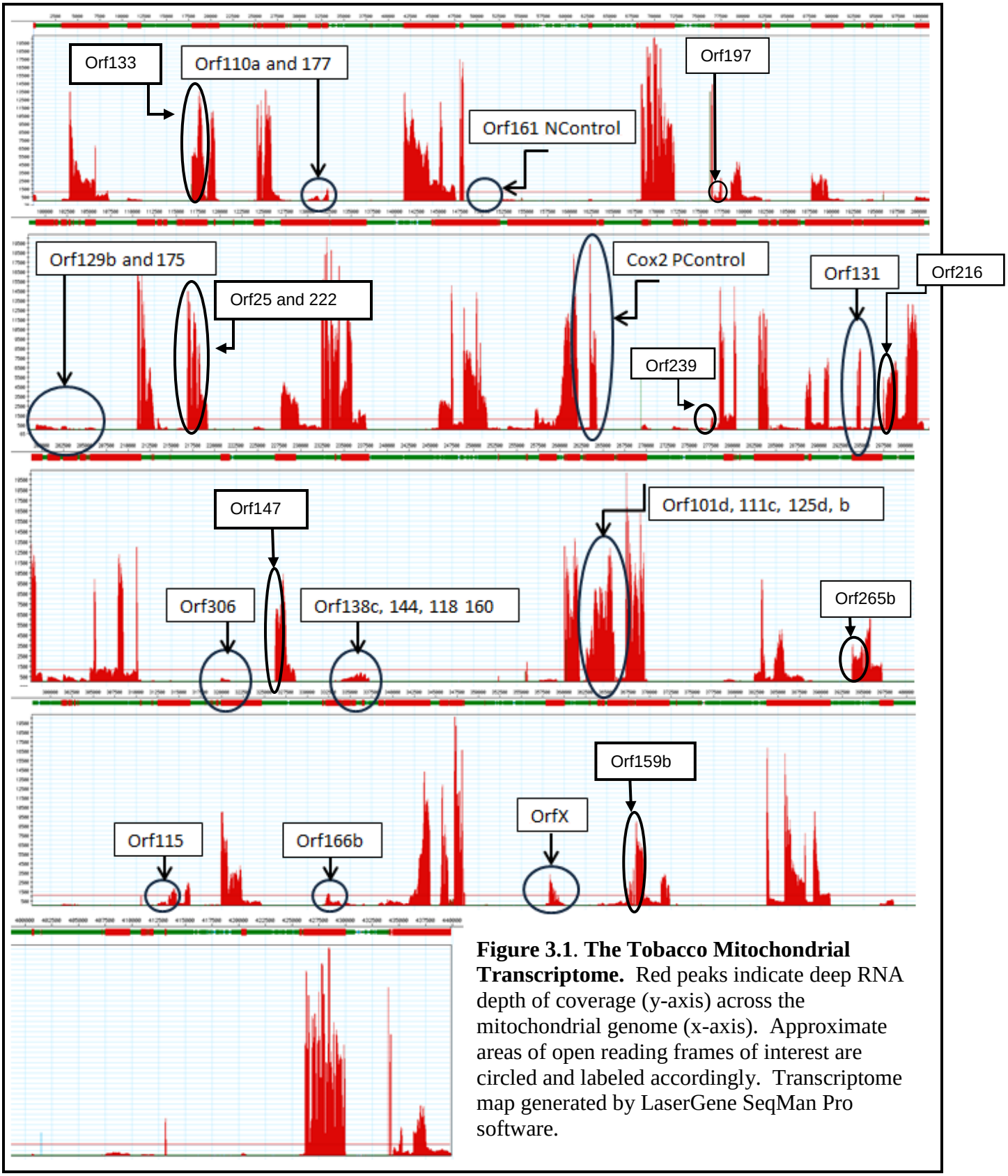

Figure 3.1. **The Tobacco Mitochondrial Transcriptome.** Red peaks indicate deep RNA depth of coverage (y-axis) across the mitochondrial genome (x-axis). Approximate areas of open reading frames of interest are circled and labeled accordingly. Transcriptome map generated by LaserGene SeqMan Pro software.

Table 3.1: List of open reading frames chosen for this study. Location of start and stop codons refer to Genbank accession NC_006581.1. Sequence length refers only to the putative coding region. ORF nomenclature is based on number of amino acids in final hypothetical protein product.

List of Open Reading Frames				
Orf Name	**Start**	**Stop**	**Sequence Length (bp)**	**Protein Length (aa)**
133	18091	18492	401	133
177	31683	32216	534	177
197	75297	74704	593	197
129b	100135	100524	390	129
151	102201	101746	456	151
175	103008	102481	528	175
25	113853	114449	596	198
222	114996	115664	668	222
239	171890	172609	719	239
134	187937	188341	405	134
216	191161	191811	650	216
306	215367	216287	921	306
147	221540	221983	443	147
138c	228518	228934	417	138
144	229441	229875	435	144
118	230082	229726	357	118
160	231140	231622	483	160
101d	254902	255207	306	101
111c	256039	256374	336	111
125d	257199	257576	377	125
orfb	257737	258207	473	156
265b	286975	287772	797	265
115	306597	306250	348	115
166b	325305	324805	515	166
159b	360762	360283	479	159

Table 3.2. Protein homology report for open reading frames that were found to be in a polysomal pellet. All ORFs were searched against the entire BLASTp database. Sixteen out of the 25 transcribed ORFs had polysome levels significantly higher than background. Reading frames with homology to identifiable proteins are shown in bold vertical print. ORFs not included in this table are: 134, 129b, 151, 175, 118, 125d, 144, 306, 222, and 239

BLASTp Sequence Homology Report (Greater than or equal to 90% ID)

Mitochondrial Protein Homology																
Species	**177**	**138c**	**160**	**101d**	**111c**	**115**	**166b**	**133**	**25**	**222**	**216**	**265b**	**159b**	**197**	**147**	**B**
				ndhB (chloroplast)	**ndhB (chloroplast)**			**rpl2 (chloroplast)**	**atp4 (Mitochondria)**		**rps1 (Mitochondria)**	**atp8 (Mitochondria)**	**rpl10 (Mitochondria)**	**nad2 (Mitochondria)**		**atp8 (Mitochondria)**
N. tabacum	x	x	x			x	x			x					x	
A. thaliana		x														
C. papaya		x	x			x										
M. guttatis		x	x													
V. vinifera		x	x				x								x	
F. vesca		x	x													
B. carinata		x														
B. juncea		x														
B. oleracea		x														
B. rapa		x														
B. napus		x														
C. lanatus		x				x										
R. communis		x	x			x										
M. pinnata		x														
V. radiata		x														
C. pepo		x				x										
D. carota		x	x												x	
B. macrocarpa						x										
B. vulgaris		x				x									x	
S. lycopersicum															x	
Z. Mays															x	
Nuclear Protein Homology																
N. benthamiana	x	x	x			x	x			x						
V. vinifera		x	x													
M. domestica		x														
S. lycopersicum		x				x										
A. thaliana		x														
M. truncatula		x														
B. hygrometrica			x													
R. communus															x	
L. usitatissimum			x													
C. melo															x	

Transcription Confirmation of Experimental ORFs

The initial identification of transcribed ORFs came from a single pass deep sequencing experiment so all results required confirmation from multiple replicates. The 25 putatively transcribed open reading frames were confirmed using qRT-PCR analyses from three independent replicates. Specific primers were designed *in silico* for each experimental ORF. Each primer set was thermally optimized as performed in Sharpe et al. 2008 and tested for reaction efficiency (Table 3.4). All ORF qRT-PCR results were compared to the Cox_2 mitochondrial protein coding gene as a positive control and a non-expressed ORF as a negative background control (Figure 3.2) compared to an area of low background expression that did not appear to be expressed; in this case, the RNA-seq information was accurate, as "background" expression was minimal (Figure 3.1, Table 3.3). Relaxed transcription of plant mitochondria create a noise or background present in the transcriptome that is always expressed (Kuhn et al. 2005, 2009); therefore, it was important to identify a baseline of gene expression to understand what meaningful transcription looks like in my study. The background control for this experiment not only provided an area of expression in a noise region, but also contained an open reading frame (Figure 3.2). Expression was observed for all ORFs examined.

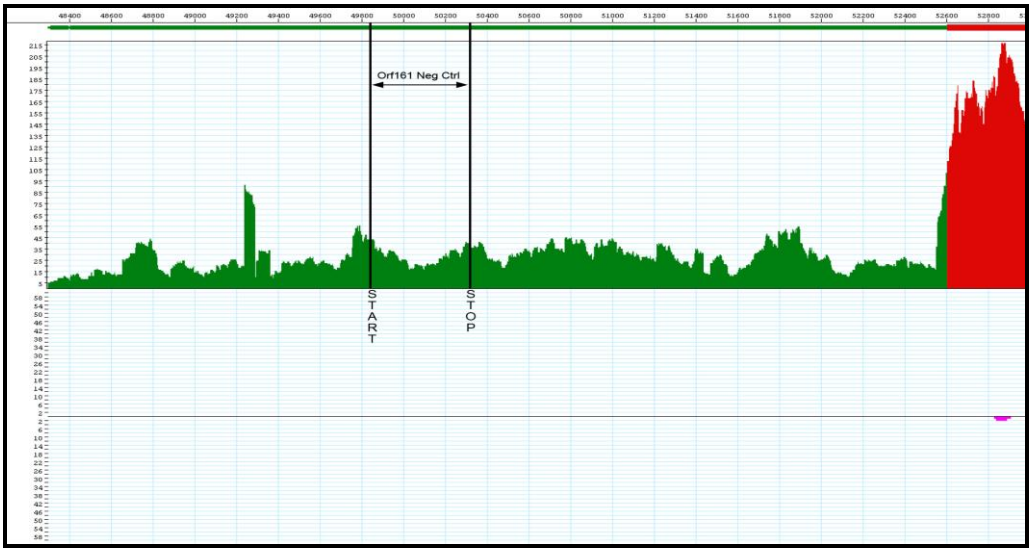

Figure 3.2: Depth of coverage of the negative control "background" reading frame. All peaks showing **green** are below 100 reads (y-axis) for that specific mitogenomic loci (x-axis).

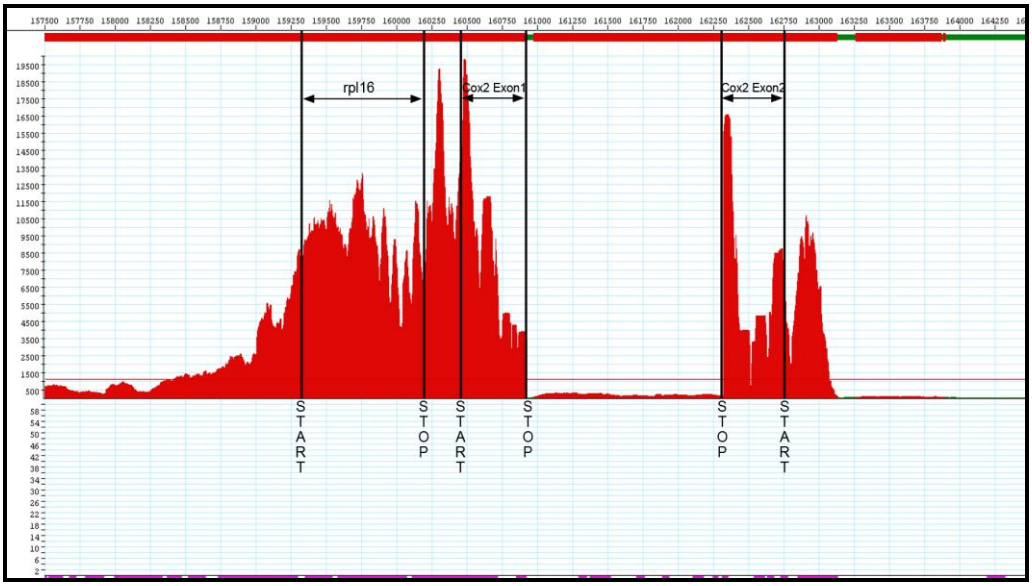

Figure 3.3: Depth of coverage of the positive control "Cox_2" and its two exons. All peaks showing **red** are above 100 reads (y-axis) for that specific mitogenomic loci (x-axis). Cox_2 exon 2 was used as the positive control for the entire experiment.

Leaf and root RNA samples were used to compare and contrast possible expression differences in each tissue. qRT-PCR results for all 25 open reading frames suggested they were all transcribed and confirmed the RNA-seq preliminary data. Since qRT-PCR data does not follow a normal distribution, the non-parametric Mann-Whitney rank sum statistical test was used to derive p-values by comparing all ORFs to the background transcript abundance readings (Tables 3.5 and 3.6). An overall look at transcript abundance of all experimental genes shows root expression levels are many times greater in 17 of the 25 ORFs (Figure 3.4); however, three experimental ORFs show leaf expression to be much higher: ORF 101d, 111c, and 133. Orf 101d and 111c both show strong homology to chloroplast nad2 genes while ORF 133 shows homology to the chloroplast ribosomal protein gene rpl2.

Table 3.3: Statistical hypothesis testing for transcription confirmation of all experimental genes taken from root and leaf tissue. The non-parametric Mann-Whitney rank sum test was used to derive p-value for each ORF in comparison to background expression. All genes are being transcribed.

Leaves					
Background Vs.	Mann-Whitney Rank Sum (p =)	Significantly Higher than Background?	Background Vs.	Mann-Whitney Rank Sum (p =)	Significantly Higher than Background?
Cox_2	0.002	Yes	OrfB	0.002	Yes
Orf177	0.002	Yes	Orf115	0.002	Yes
Orf129b	0.002	Yes	Orf166	0.002	Yes
Orf151	0.002	Yes	Orf133	0.1	Yes
Orf175	0.002	Yes	Orf25	0.1	Yes
Orf134	0.002	Yes	Orf222	0.1	Yes
Orf306	0.002	Yes	Orf216	0.1	Yes
Orf138c	0.002	Yes	Orf265b	0.1	Yes
Orf144	0.002	Yes	Orf159b	0.1	Yes
Orf118	0.002	Yes	Orf197	0.1	Yes
Orf160	0.002	Yes	Orf239	0.1	Yes
Orf101d	0.002	Yes	Orf147	0.1	Yes
Orf111c	0.002	Yes			
Orf125d	0.002	Yes			
Roots					
Cox_2	0.002	Yes	OrfB	0.002	Yes
Orf177	0.002	Yes	Orf115	0.002	Yes
Orf129b	0.002	Yes	Orf166	0.002	Yes
Orf151	0.002	Yes	Orf133	0.1	Yes
Orf175	0.002	Yes	Orf25	0.1	Yes
Orf134	0.015	Yes	Orf222	0.1	Yes
Orf306	0.002	Yes	Orf216	0.1	Yes
Orf138c	0.002	Yes	Orf265b	0.1	Yes
Orf144	0.002	Yes	Orf159b	0.1	Yes
Orf118	0.002	Yes	Orf197	0.1	Yes
Orf160	0.002	Yes	Orf239	0.1	Yes
Orf101d	0.002	Yes	Orf147	0.1	Yes
Orf111c	0.002	Yes			
Orf125d	0.002	Yes			

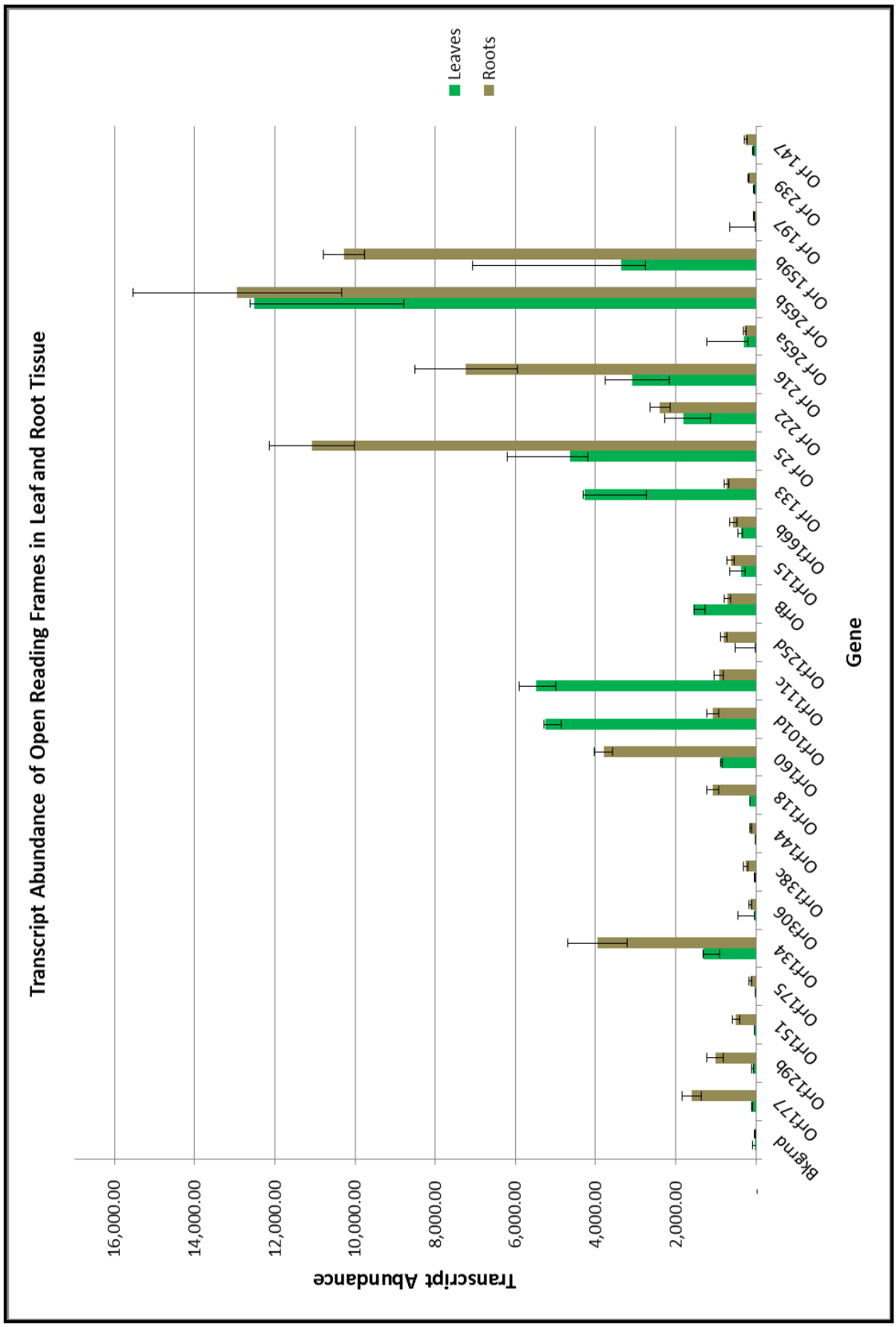

Figure 3.4: Overall qRTPCR analysis of 25 open reading frames analyzed in this study. Expression levels were obtained in both leaf and root tissue.

These three reading frames showed sequence homology to viral proteins. ORF 129b showed relation to mitoviral RNA polymerase. ORF 151 produces a protein related to viral integrase genes or enzymes that allow viruses to integrate genetic information into host DNA (Lodish et al. 2008). ORF 175 showed no homology to known genes (Figure 3.5).

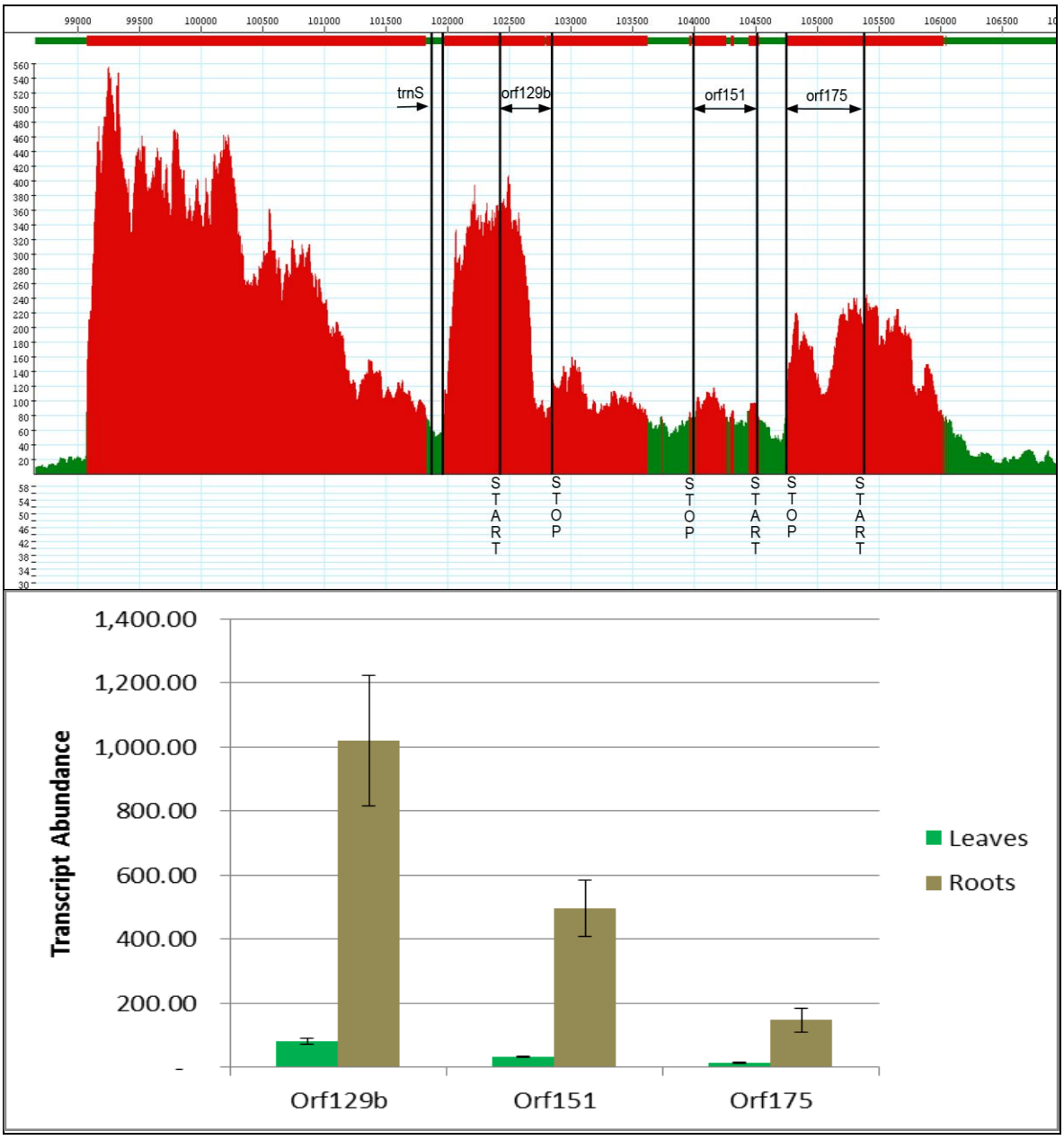

Figure 3.5: Deep Sequencing and qRTPCR Analysis of ORFs 129b, 151, and 175

ORF 110a was the only open reading frame tested that was not transcribed high enough above background levels to be considered for the study; however, the nearby overlapping ORF 177 showed considerable expression levels. ORF 177 only showed homology to nuclear DNA in *Nicotiana benthamiana*.

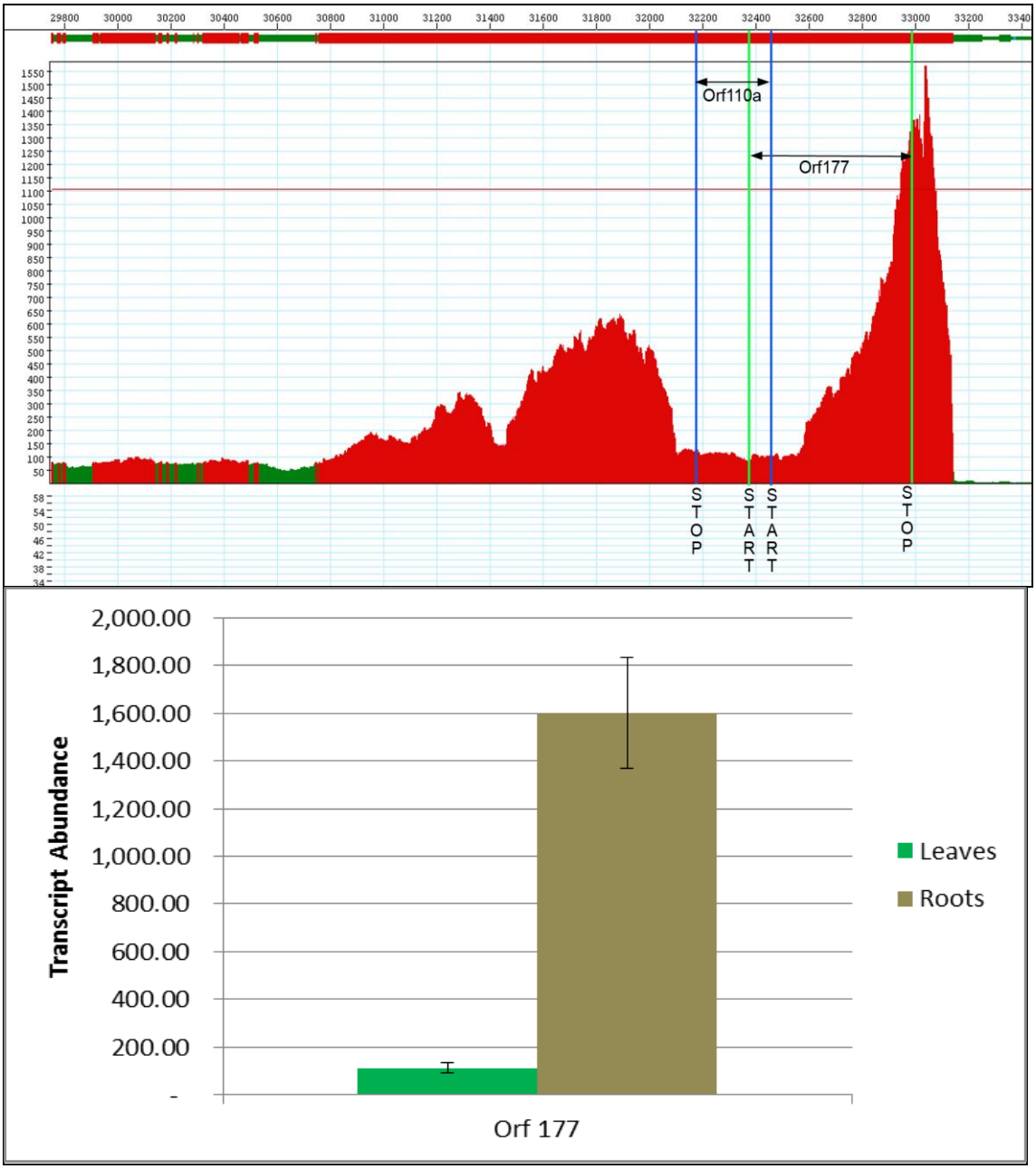

Figure 3.6: Deep Sequencing and qRTPCR Analysis of ORF 177

Preliminary RNA-seq data shows ORF 306 expressing low (<320 scan depth). This ORF did not show any homology to nuclear, chloroplast, and mitochondrial genes when searched against bioinformatics databases.

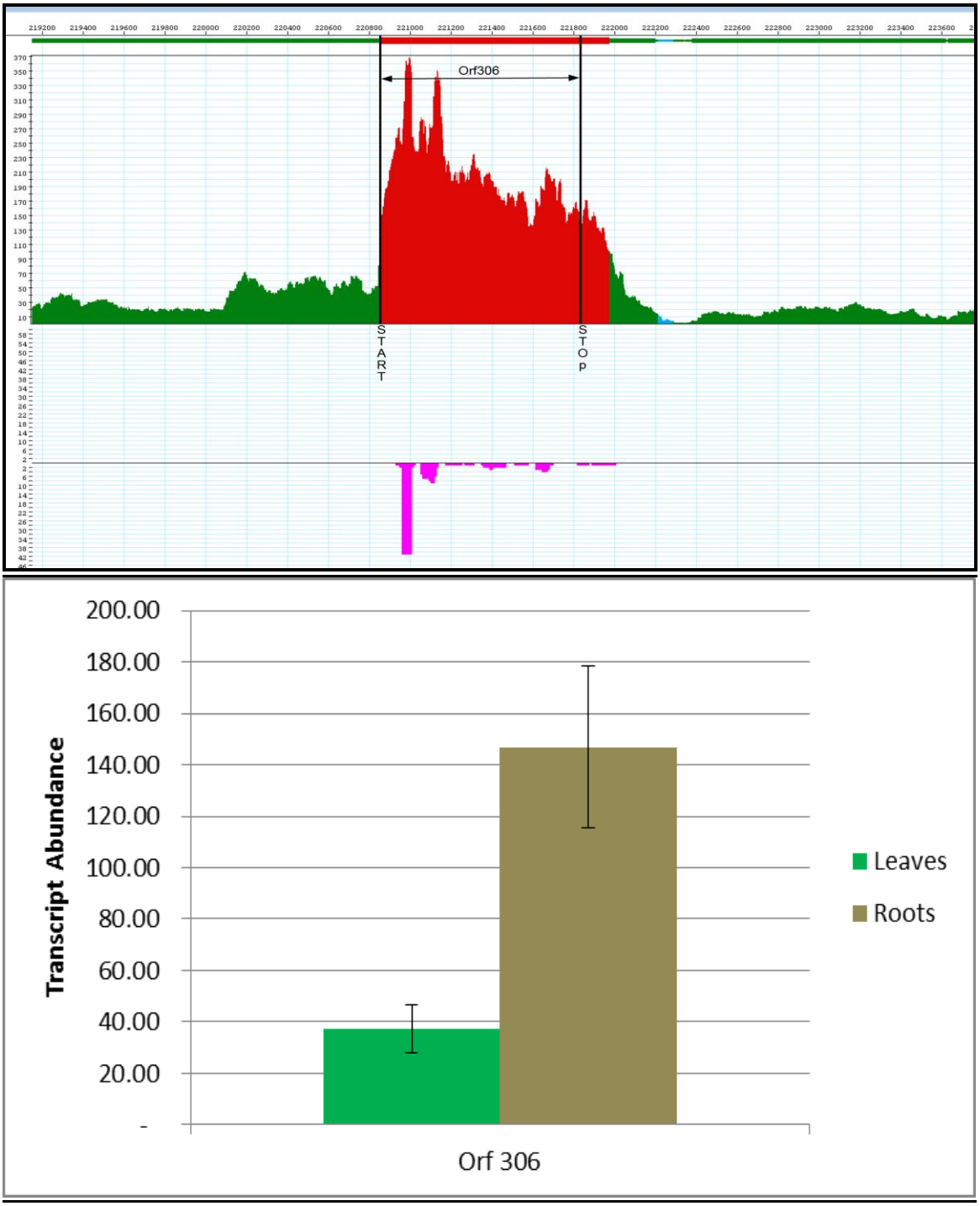

Figure 3.7: Deep Sequencing and qRTPCR Analysis of ORF 306

ORF 138c showed homology to mitochondrial and nuclear genes from many species while ORFs 144 and 118 did not show any clear nuclear, plastid, or mitochondrial homology; however, ORF 160 did show homology to several species' nuclear and mitochondrial genes. Overall, preliminary RNA-seq scan coverage for these 4 ORFs showed noticeable expression, but generally low compared to the other ORF candidates in this study (< 960 scan depth).

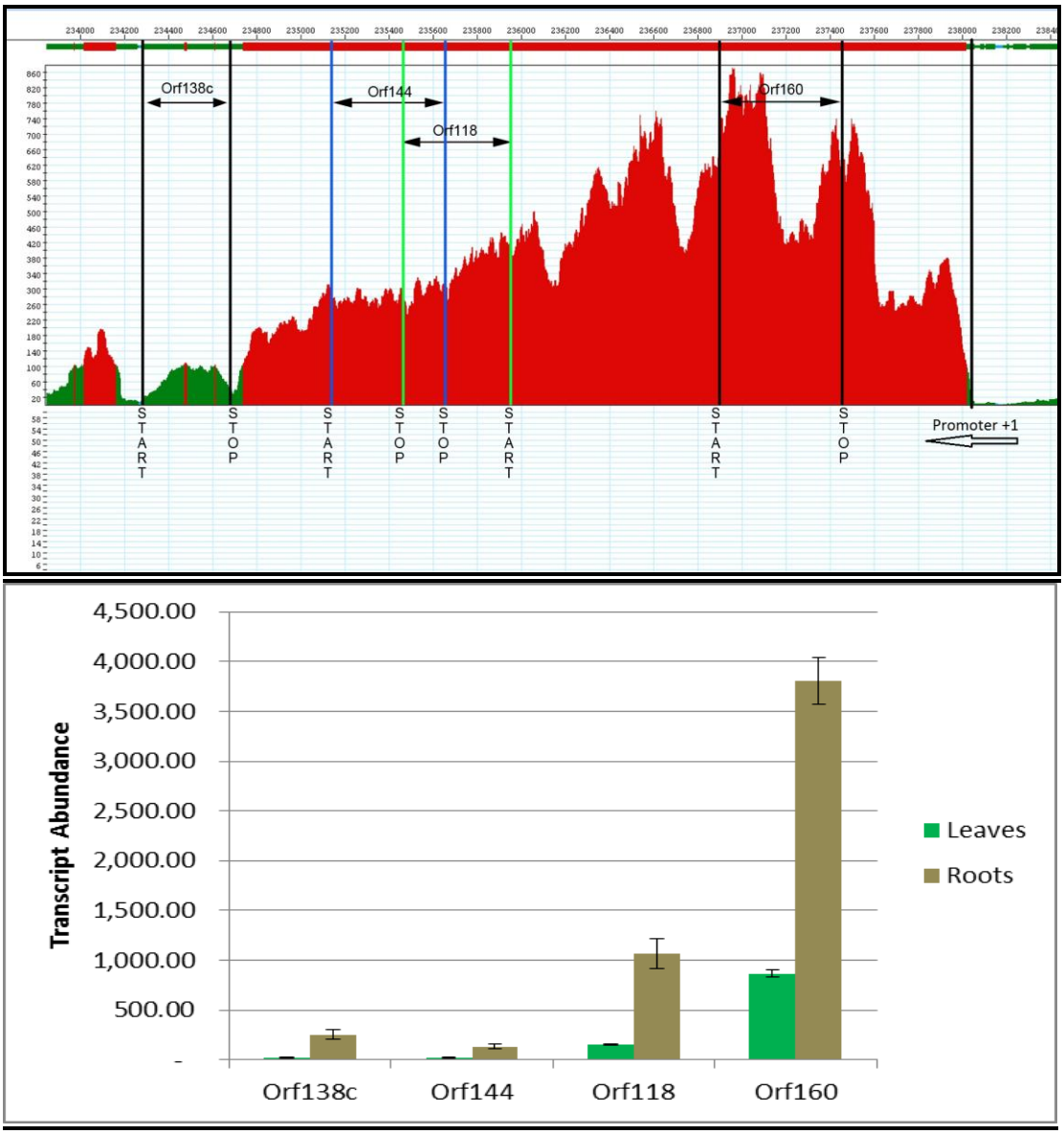

Figure 3.8: Deep Sequencing and qRTPCR Analysis of ORFs 138c, 144, 118, and 160

ORFs 101d and 111c were found to be chloroplast NADH complex genes. Their expression was considerably higher than most of the mitochondrial ORFs that were examined. Expectedly, expression of these genes was 5 fold higher in green leaf tissue where active chloroplasts reside. ORF 125d showed no conserved domains while ORFB, related to ATP synthase subunit 8, showed considerable expression levels.

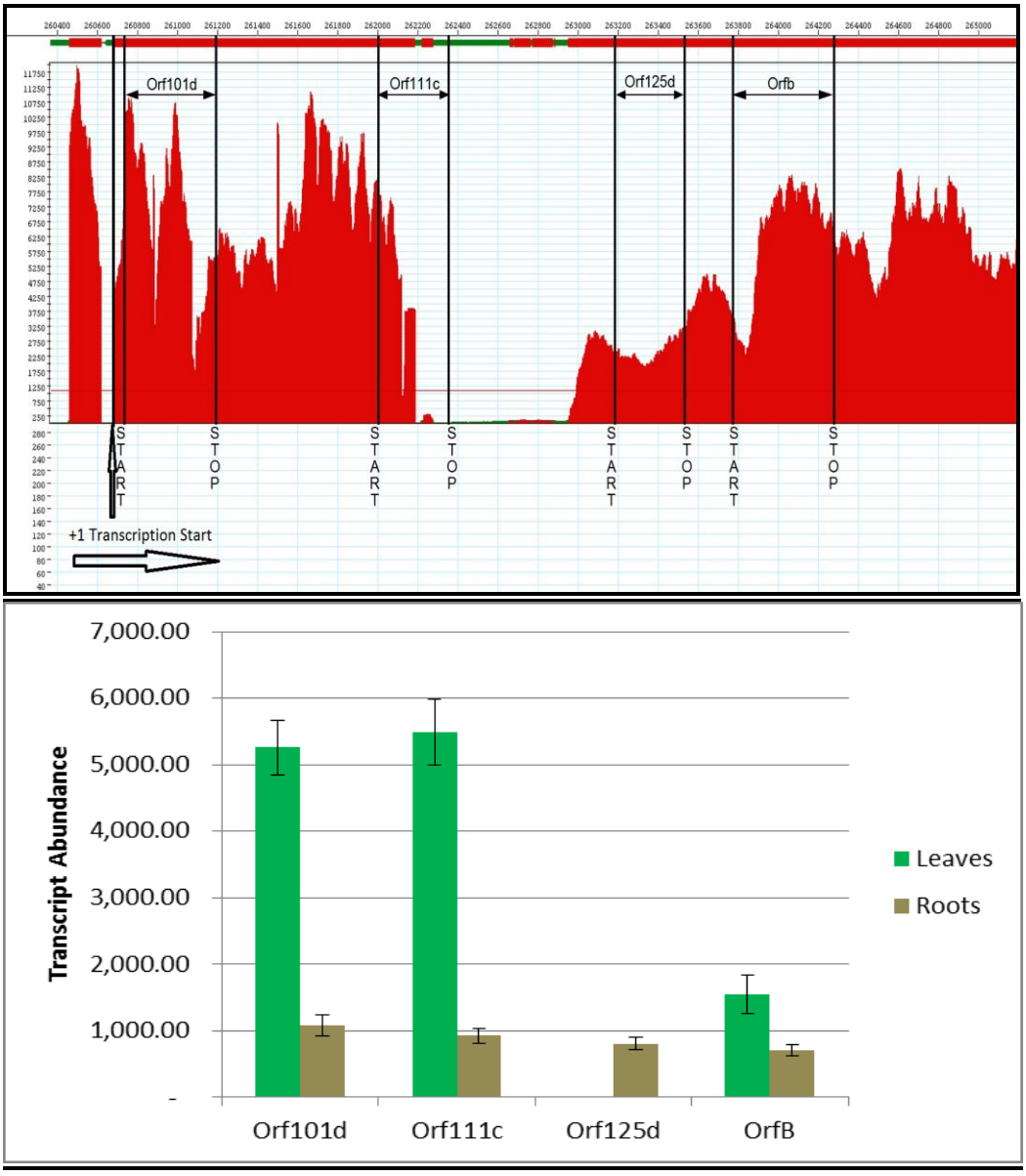

Figure 3.9: Deep Sequencing and qRTPCR Analysis of ORFs 101d, 111c, 125d, and ORFB

ORF 115 shows expression levels consistent with preliminary RNA-seq data. This open reading frame showed considerable protein homology to different species' nuclear and mitochondrial genomes.

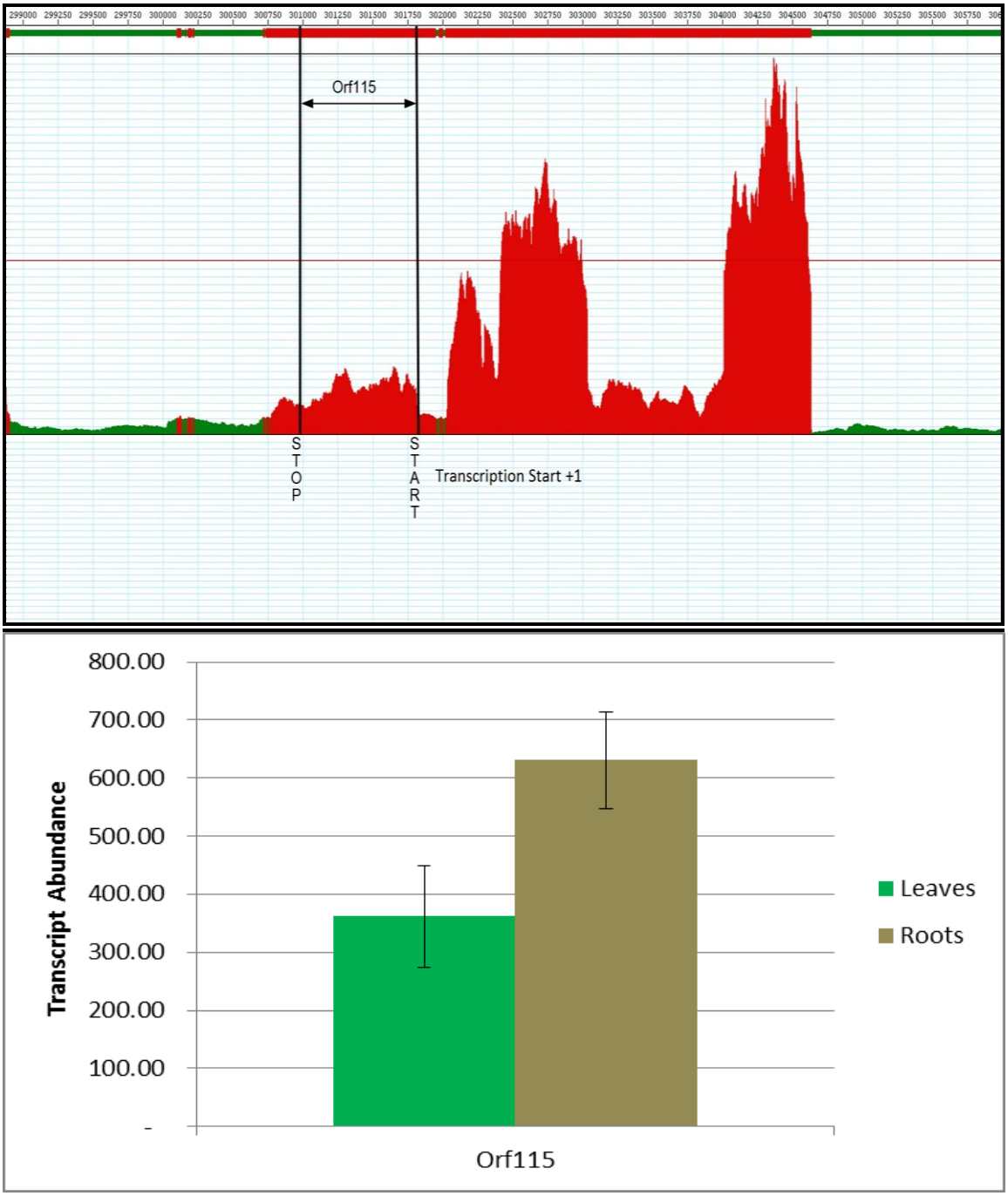

Figure 3.10: Deep Sequencing and qRTPCR Analysis of ORF 115

ORF 166b showed moderate expression levels and considerable protein homology to nuclear and mitochondrial genomes.

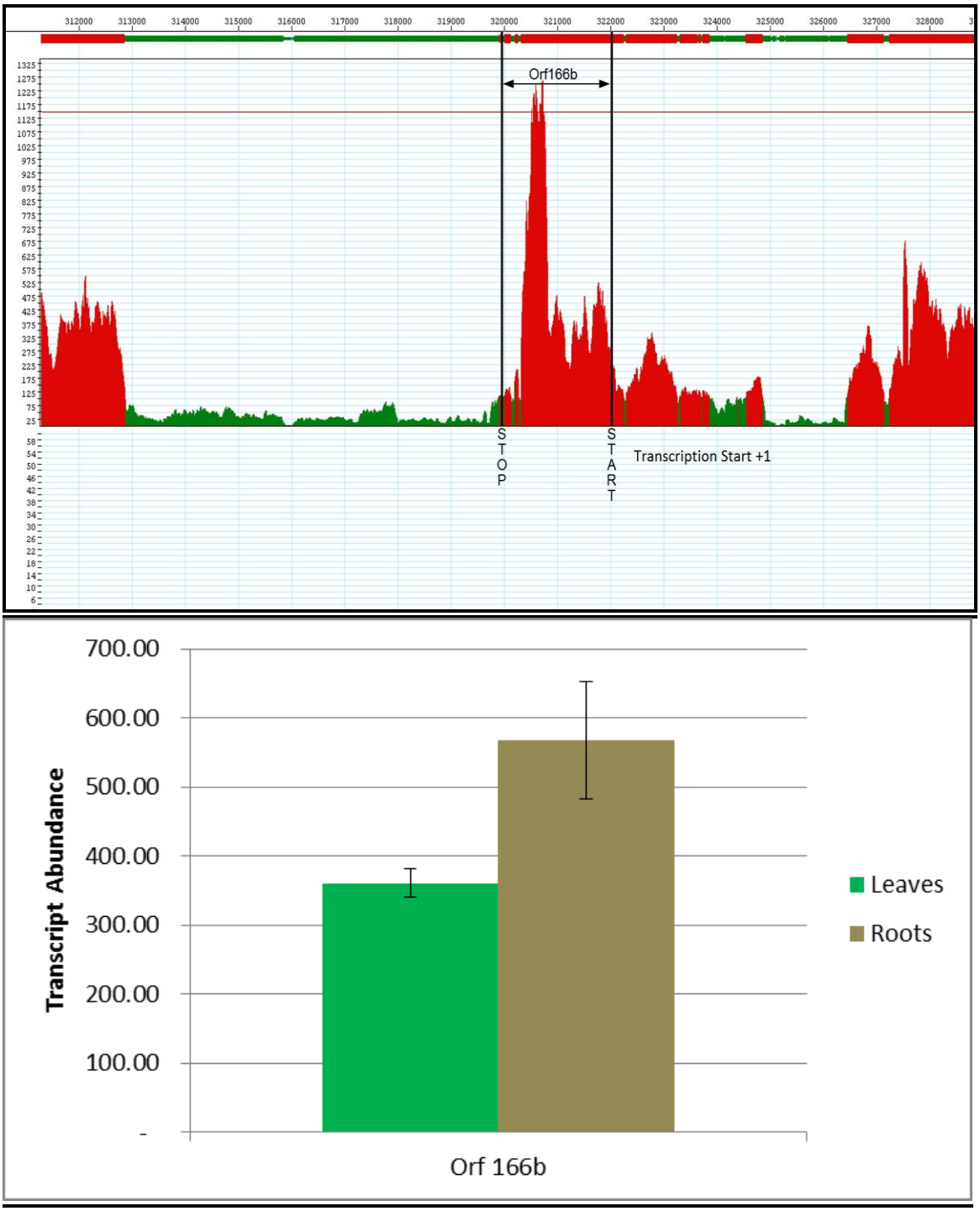

Figure 3.11: Deep Sequencing and qRTPCR Analysis of ORF 166b

ORF 133 showed high identity to the rpl2 chloroplast gene that codes for a ribosomal protein. Expression in the leaves was over 5 fold higher than in the roots. RNA-seq showed very deep scans, peaking at 13,500 scans within the reading frame.

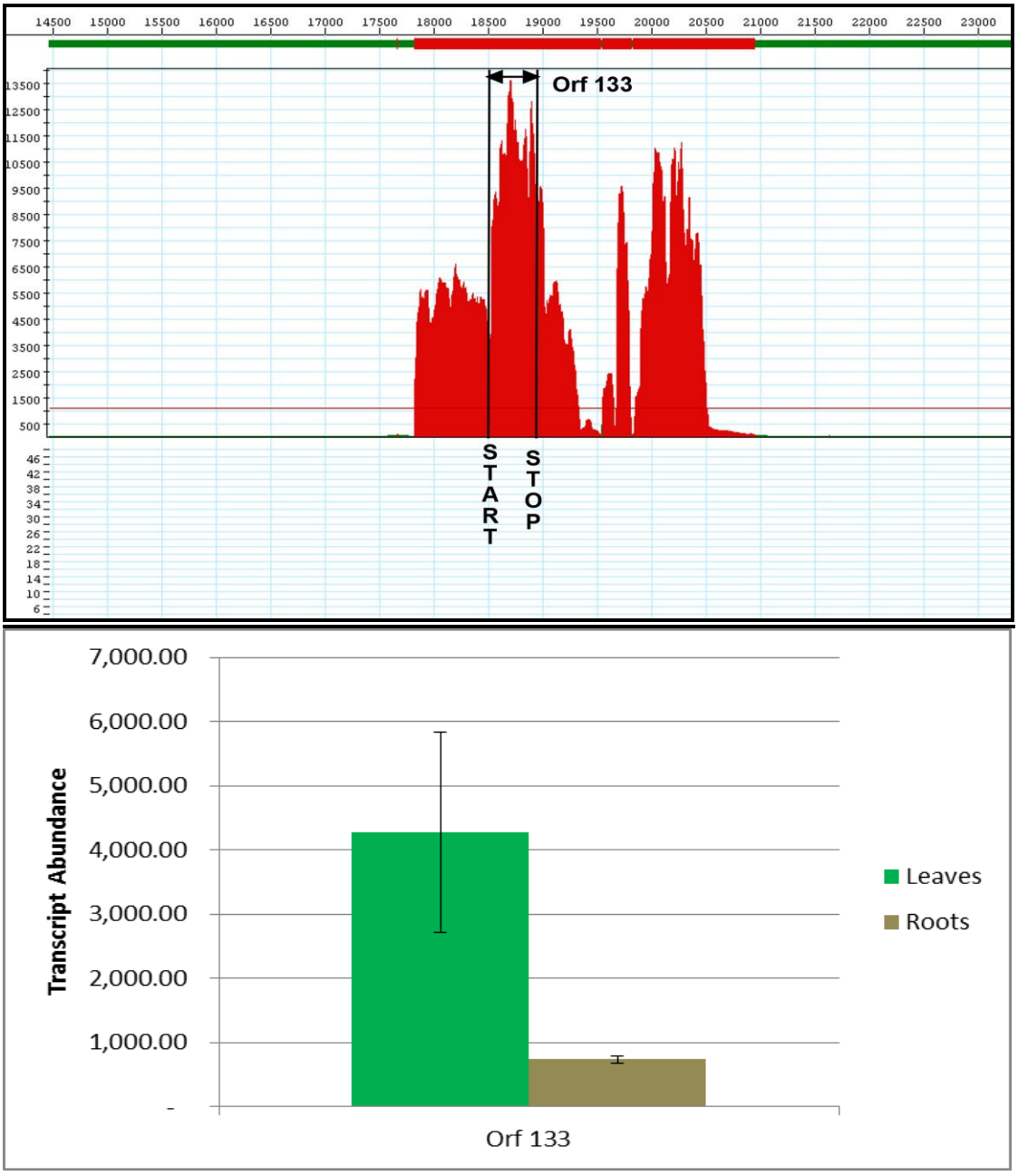

Figure 3.12: Deep Sequencing and qRTPCR Analysis of ORF 133

ORF 25 showed high identitiy to the ATP4 mitochondrial gene in multiple species with moderate expression in the leaves and exceptionally high expression in the roots. ORF 222, though close by, only showed homology to nuclear DNA in *Nicotiana benthamiana*. Both ORFs showed considerable depth of coverage on preliminary RNA-seq graphs

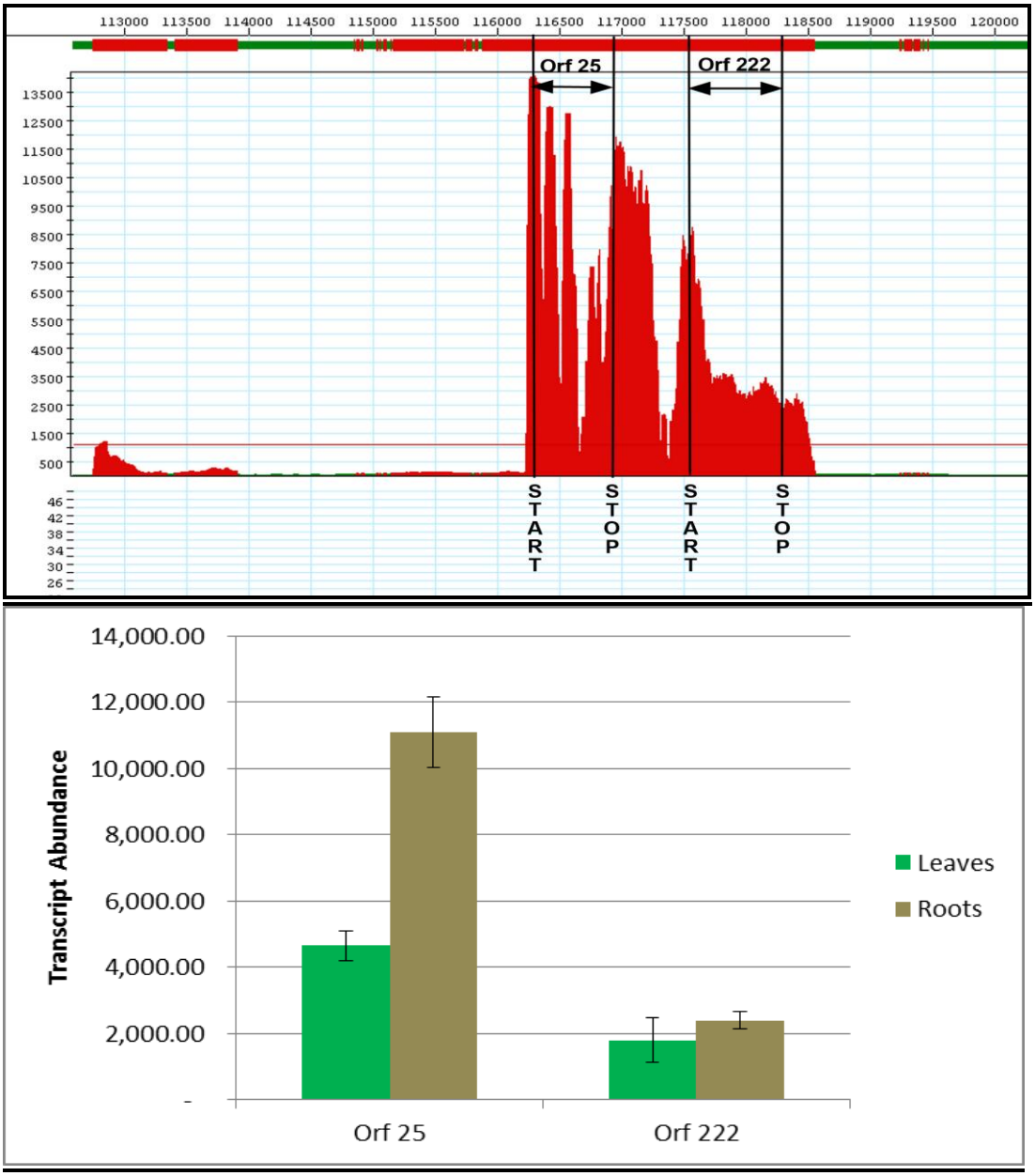

Figure 3.13: Deep Sequencing and qRTPCR Analysis of ORFs 25 and 222

ORF 216 showed high identity to the mitochondrial rps1 gene which codes for structural small subunit ribosomal proteins. As expected, genes related to construction and maintenance of organellar ribosomes are highly expressed.

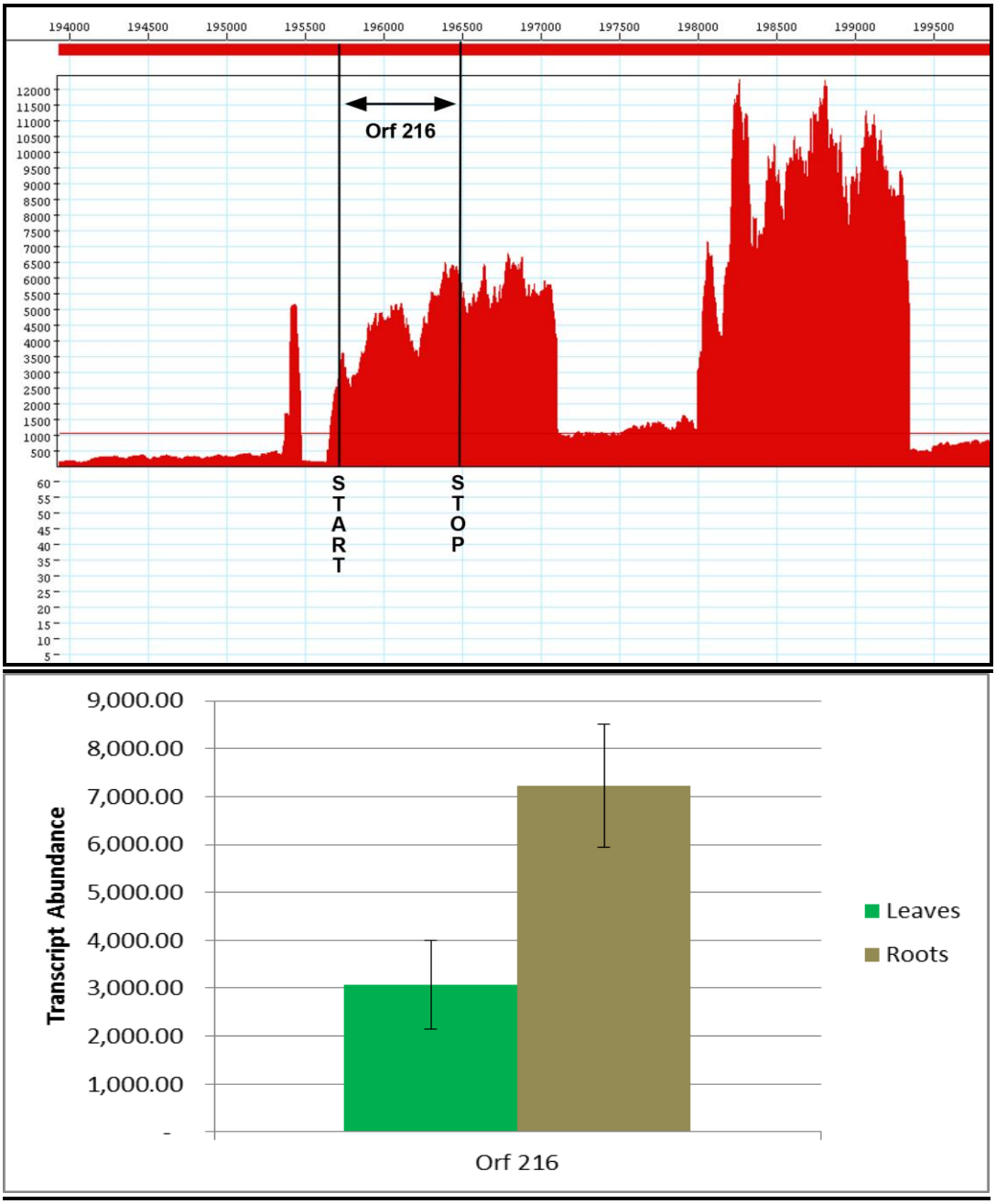

Figure 3.14: Deep Sequencing and qRTPCR Analysis of ORF 216

ORF 265b shows high identity to the mitochondrial ATP8 gene which codes for a protein subunit used in the ATP synthase supercomplex. Expression was moderate and nearly identical in both leaf and root tissue. This ORF also resides in a repeated region of the tobacco mitogenome.

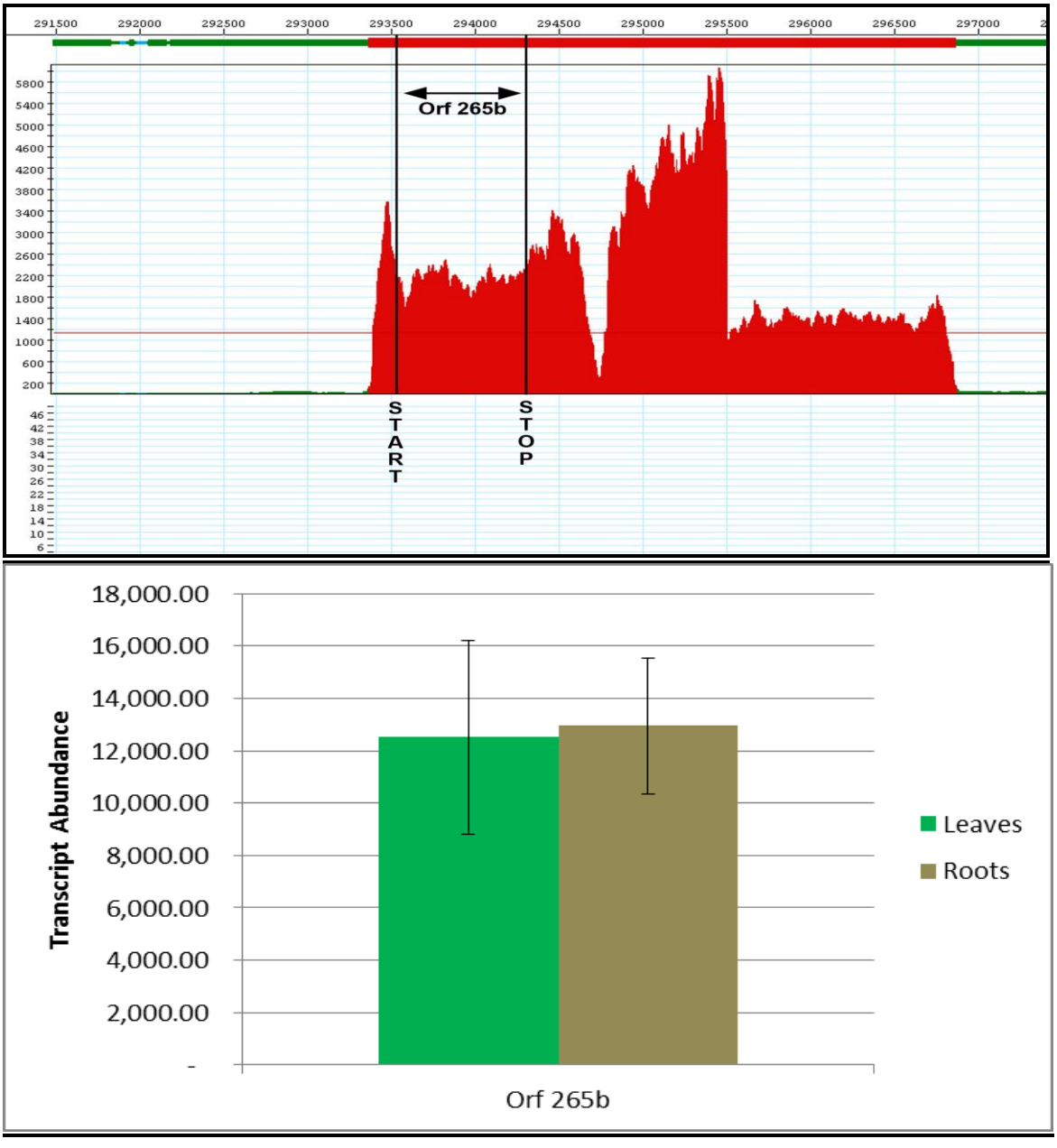

Figure 3.15: Deep Sequencing and qRTPCR Analysis of ORF 265b

ORF 159b is the only reading frame in this study that shows high identitiy to a *mitochondrial* ribosomal structure gene – in this case rpl10. Expression level was high in both tissues, but the root sample showed exceptionally high average copy numbers (>10,000 transcripts).

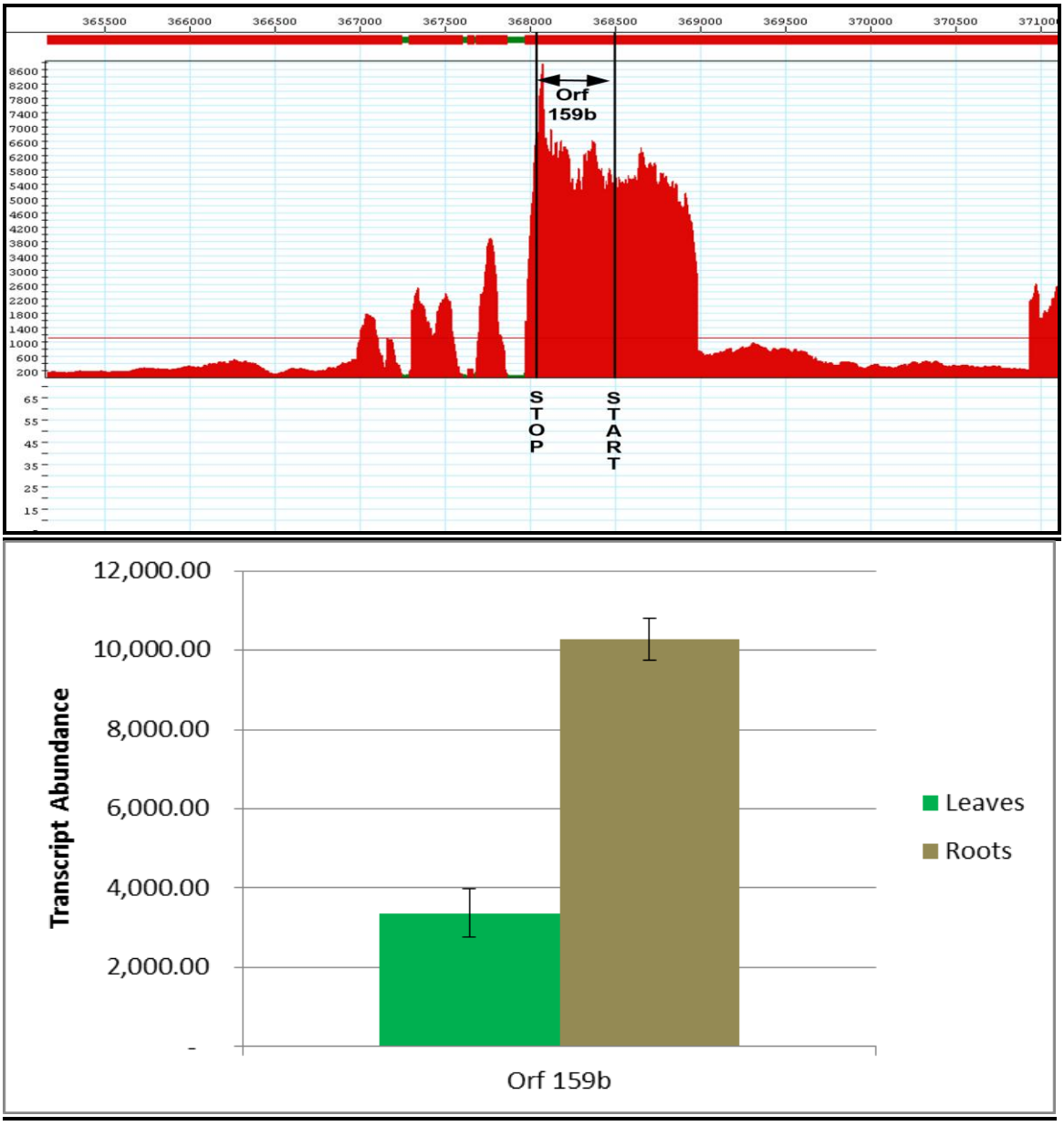

Figure 3.16: Deep Sequencing and qRTPCR Analysis of ORF 159b

ORF 197, in relation to deep-seq RNA scans, showed to be expressing minimally. It does, however, bear homology to conserved protein regions on the mitochondrial nad2 gene.

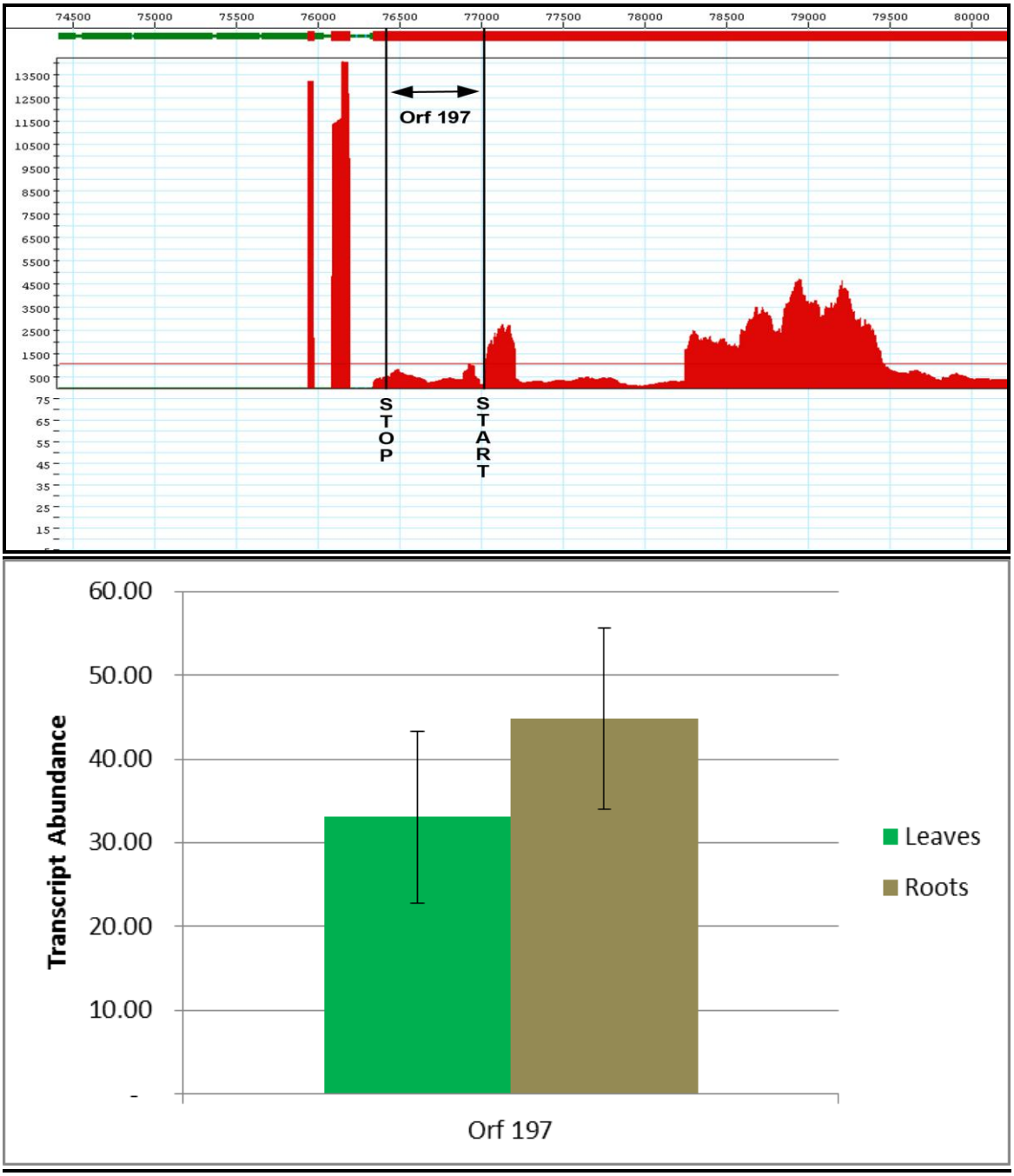

Figure 3.17: Deep Sequencing and qRTPCR Analysis of ORF 197

ORF 239 showed minimal expression; protein homology search did not yield sequence similarities.

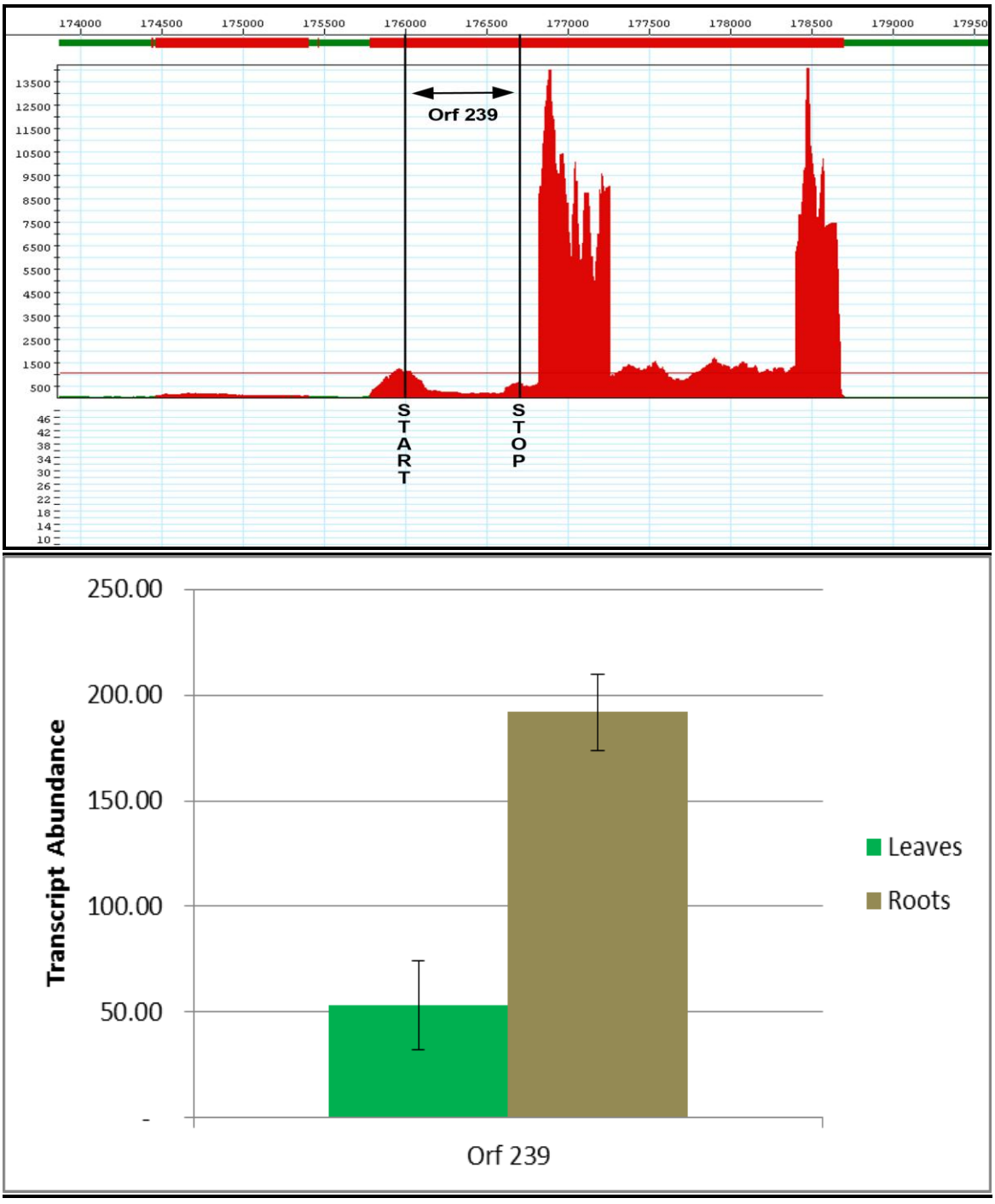

Figure 3.18: Deep Sequencing and qRTPCR Analysis of ORF 239

ORF 147 showed moderate total expression across the two tissues. Homology searches did not yield functional conserved regions in its protein product, but did show high similarity with other species' nuclear and mitochondrial genomes suggesting this ORF has been subject to numerous horizontal gene transfers.

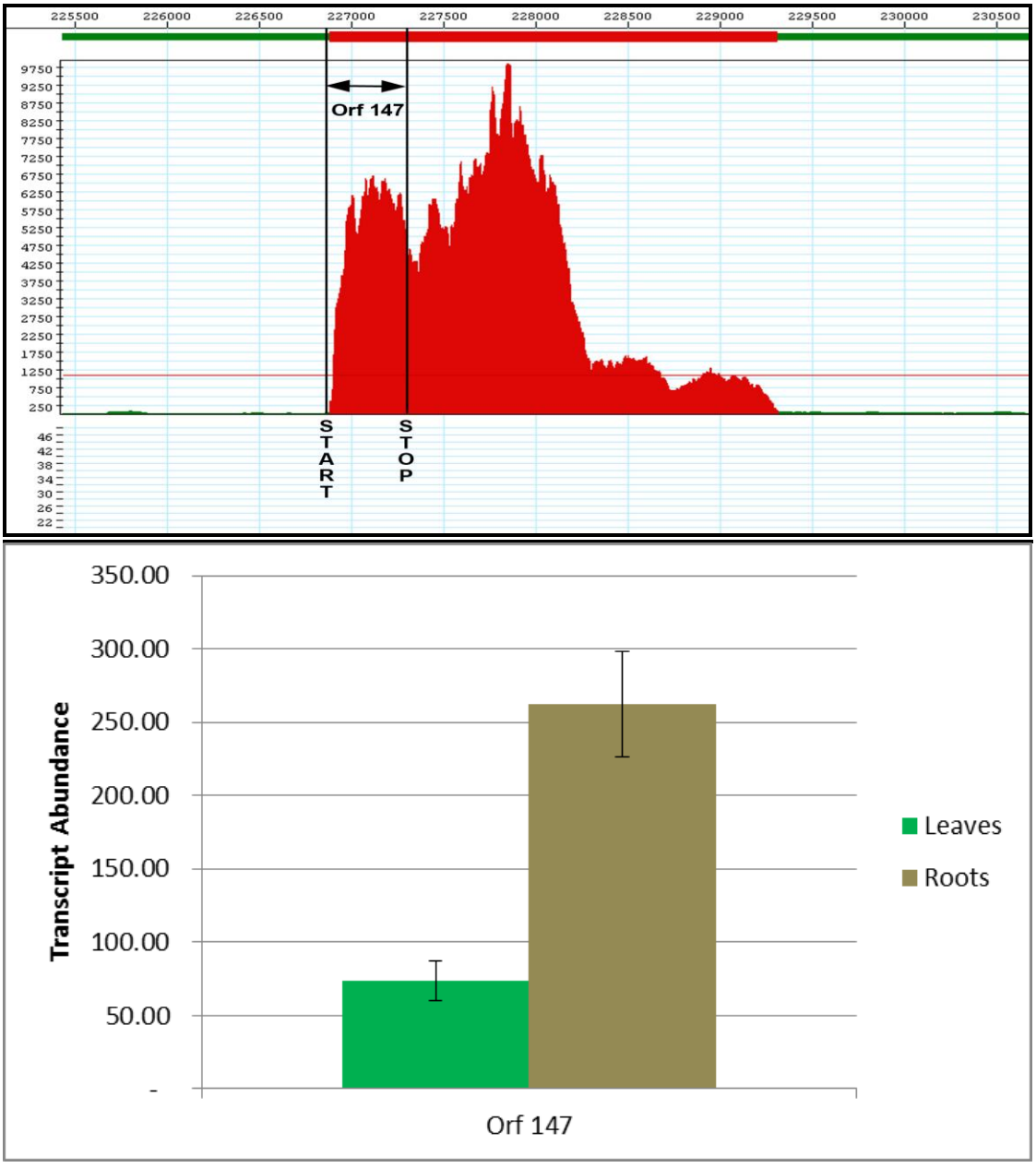

Figure 3.19: Deep Sequencing and qRTPCR Analysis of ORF 147

Polysome Analysis of Transcription-Confirmed Experimental ORFs

All the confirmed transcribed ORFs were subjected to polysomal analysis to test for evidence of translation. Cox2 was used as a control. Figure 3.21 demonstrates that Cox2 transcripts were clearly detected in polysome pellets and that transcripts treated with ethylenediaminetetraacetic acid (EDTA) released the Cox2 RNA messages from ribosomes to produce a negative control.

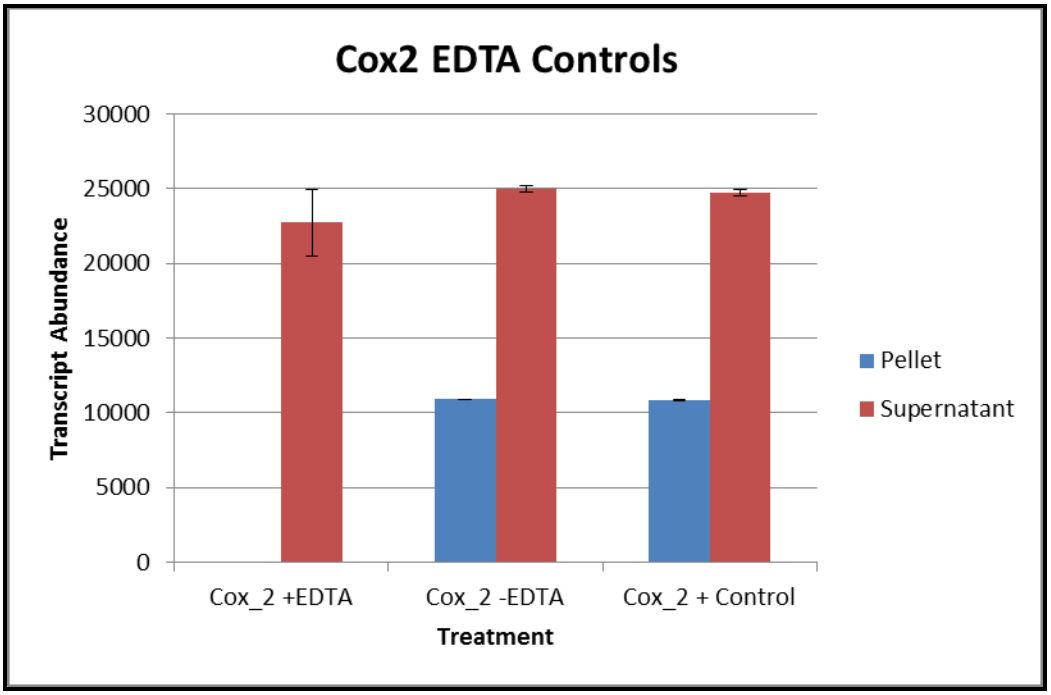

Figure 3.20: qRT-PCR of EDTA treated polysomal pellet and fraction. Assay uses +control Cox2 as gene target. Addition of EDTA successfully unbound RNA messages from ribosomes.

Polysomal analysis provides two crucial pieces of information when qRT-PCR is employed: transcript abundance as free RNA in the supernatant and ribosome-bound transcript abundance. Before interpreting the results shown in tables 3.6 and 3.7, it should be understood that comparisons between the supernatant and pellet are unreliable.

This is due to ribosomal RNA dominating the overall steady state pool of transcripts in the mitochondria; their complex confirmations and associations with proteins block degradation enzymes from acting on the 5' and 3' ends of the molecule (Giege et al. 2000; Kuhn et al. 2009). When collecting polysomes from tissue lysate via centrifugation, a tremendous amount of rRNA is collected in the pellet. This greatly increases the ratio of rRNA to mRNA when collecting total RNA for qRT-PCR. The supernatant, however, contains much less rRNA -- RNA taken from the supernatant contains much more mRNA mass per volume giving an unreliable representation of targeted transcript abundance compared to the pellet. The main goal of the polysome analysis step was to place emphasis on whether the transcript was present on ribosomes – not necessarily transcript abundance. Simply the presence of transcript bound to ribosomes presents the possibility of novel translated proteins from these open reading frames. Supernatant results reaffirmed that transcription of all experimental open reading frames was taking place with the exception of ORF 175 -- a chronically low expressing ORF seen throughout the study (Table 3.8). Pellet data showed that 15 experimental ORFs were indeed ribosome associated in some capacity (Table 3.9). Aside from the ORFs that possessed chloroplast homology (101c, 111c, and 133) native mitochondrial ORFs that possessed conserved domains which provide functionality to respiration and ribosome development were polysome associated; furthermore, ORFs *without* clear conserved domains were still polysome associated. These ORFs were conserved across many other species (Table 3.2, Table 3.8). Overall polysome analysis shows that it is clear that not all transcribed ORFs whose genesis arose in the mitochondria are potentially translated.

Table 3.4. Transcript abundance found in the supernatant (free mRNA). P-values were calculated comparing each experimental ORF to background expression using the Mann-Whitney Rank Sum statistical hypothesis test.

0%	2%	4%	6%	8%	10%	12%	14%	16%	18%	20%

Percent of total expression across all orfs

Supernatant			
Gene/Abundance	**Estimated Copy Number**	**Mann-Whitney Rank Sum**	**Transcribed?**
Cox2 (+Control)	24757	0.002	Y
Cox2 +EDTA	22734	n/a	Y
Background	20	n/a	Y
Orf177	212	0.002	Y
Orf129b	147	0.004	Y
Orf151	166	0.002	Y
Orf175	52	0.132	N
Orf134	2283	0.002	Y
Orf306	56	0.026	Y
Orf138c	117	0.004	Y
Orf144	64	0.026	Y
Orf118	340	0.002	Y
Orf160	736	0.002	Y
Orf101d	14456	0.002	Y
Orf111c	11137	0.002	Y
Orf125d	115	0.002	Y
Orfb	4140	0.002	Y
Orf115	507	0.002	Y
Orf166b	257	0.002	Y
Orf133	33295	0.002	Y
Orf25	11486	0.002	Y
Orf222	6809	0.002	Y
Orf216	10482	0.002	Y
Orf265b	41028	0.002	Y
Orf159b	19655	0.002	Y
Orf197	1818	0.002	Y
Orf239	2108	0.002	Y
Orf147	2606	0.002	Y

Table 3.5. Transcript abundance found in the polysomal pellet (ribosome-bound RNA). P-values were calculated comparing each experimental ORF to background expression using the Mann-Whitney Rank Sum statistical hypothesis test.

0%	3%	6%	9%	12%	15%	18%	21%	24%	27%	30%

Percent of total expression across all orfs

Pellet			
Gene/Abundance	**Estimated Copy Number**	**Mann-Whitney Rank Sum**	**Polysome?**
Cox2 (+Control)	10856	0.002	Y
Cox2 +EDTA	Undetectable	n/a	n/a
Background	134	n/a	n/a
Orf177	648	0.002	Y
Orf129b	289	0.002	N
Orf151	163	0.132	N
Orf175	0	0.31	N
Orf134	27	0.002	N
Orf306	1	0.002	N
Orf138c	400	0.002	Y
Orf144	4	0.009	N
Orf118	0	0.002	N
Orf160	2117	0.002	Y
Orf101d	45472	0.002	Y
Orf111c	5847	0.002	Y
Orf125d	60	0.065	N
Orfb	9679	0.002	Y
Orf115	338	0.009	Y
Orf166b	1003	0.002	Y
Orf133	17382	0.002	Y
Orf25	3803	0.002	Y
Orf222	2132	0.002	Y
Orf216	14845	0.002	Y
Orf265b	19692	0.002	Y
Orf159b	5407	0.002	Y
Orf197	231	0.589	N
Orf239	206	0.937	N
Orf147	346	0.004	Y

Chapter IV

Discussion

\

Mitochondria play a vital role in overall cell metabolism: from providing the cell with energy to synthesizing organic macromolecules, this organelle has evolved to power the many needs of the cell; thus, understanding expression of its genome is an important part of why and how mitochondria can achieve such diverse cellular tasks. Many mysteries about the mitochondria, especially plant mitochondria, have yet to be elucidated so this study attempted to shed light on one of those mysteries: *to what level do plant mitochondria express uncharacterized open reading frames (ORFs) in its genome?* Ultimately, this study seeks to uncover if these ORFs are possibly loaded onto ribosomes in preparation for protein assembly. Sugiyama et al. 2005 annotated the entire tobacco mitogenome, taking into account 119 unique ORFs. RNA-sequencing was an invaluable part of this project as it provided a snapshot of the entire tobacco mitogenome's expressed gene content, subsequently revealing 25 ORFs found to be transcribed considerably. Fifteen of the 25 ORFs were found to be associated with ribosomes revealing the possibility of novel functional mitochondrial proteins. In summary, these 25 reading frames were subjected to a two part study, falling in line with the central dogma of molecular biology: transcription and translation.

The Tobacco Mitochondrial Transcriptome

In this study, deep sequencing of the tobacco mitochondrial transcriptome detected long mono- and poly-cistronic transcripts that were mostly concentrated in parts of the genome where functionally significant genes occur. I also found that some regions without detectable genes are transcribed and that some regions with predicted open

reading frames are transcribed well above background. In addition, long stretches of the mitogenome appear to be non/low-transcribed intergenic regions.

Plant mitochondrial gene expression has been shown to be a relaxed and inefficient coordination of phage-type RNA polymerase (RpoTm) with trans-acting nuclear encoded factors (mtTFA, mtTFB) to promoter locations (Binder et al. 2012). Some YRTA promoter sequence motifs have been found upstream of important genes needed in the construction of respiratory proteins and rRNAs (Liere and Borner 2012). In order to retain expression of crucial mitogenomic genes, the organelle has ensured successful transcription initiation by retaining multiple promoters upstream of the gene (Dombrowski et al. 1998a). Other promoters lacking the canonical YRTA motif have also been found scattered throughout the mitogenome upstream of identifiable mitochondrial genes and in intergenic regions (Kuhn et al. 2009). The presence of multiple promoters causes the production of many mono and polycistronic transcripts (Dombrowski et al. 1998b). The overall result is long run-on transcripts found throughout plant mitochondrial transcriptomes and large quantities of cryptic transcripts from intergenic regions (Linder et al. 2005; Gagliardi and Binder 2007). Furthermore, plant mitochondria do not appear to possess post-transcriptional quality control mechanisms, so massive message accumulation can occur placing a remarkable amount of pressure on mtRNA degradation systems (Perrin et al. 2004; Holec et al. 2008; Binder et al. 2012).

My results suggest that multiple promoters exist and are grouped to ensure high transcription of specific regions of the mitogenome, as suggested by previous studies.

However, my data also suggests that transcription may not be as relaxed as is often insinuated and that large swaths of the mitogenome are free of promoters and may not be transcribed, or transcribed at very low rates.

Expression of Mitogenomic Open Reading Frames

All 25 open reading frames tested were transcribed in some capacity in both root and leaf tissue. Root tissue transcription was almost unanimously higher across all ORFs, except for ORFs possessing homology to chloroplast genes (Figure 3.4). Biological reasoning behind this root-specific increase in expression is likely because of the high energy demand in non-photosynthetic tissue as well as the nutrient dense environment that roots are present. Polysome analysis of 25 open reading frames showed 15 were found attached to ribosomes. These 15 ORFs could possibly be novel proteins; furthermore, some possessed homology characteristic to chimeric genes suggesting possible association with cytoplasmic male sterility gene systems (CMS).

There are conflicting hypotheses regarding the possible benefit expression of mitochondrial open reading frames provide. First, is that their origin through recombination events causes a burdensome effect on the organelle's RNA processing systems as proposed by Maier et al. 2008 -- after all, the organelle only codes for 36 proteins over a spread of 400,000 base pairs. It would seem that creating such high cryptic message amounts from the genome would prove difficult to properly process needed conserved genes. On the other hand, the expression of these genes may tend to favor the molecular evolution of the organelle. Beaudet et al. 2013 found that mitogenomic ORFs were in fact sophisticated mobile elements specializing in horizontal

gene transfer; these ORFs were actually responsible for making chimeric versions of existing mitochondrial genes – perhaps revealing a microcosmic glimpse of genomic evolution. In my study, many of the highly expressed ORFs had similarities to this observation showing retained conserved sequences that code for important functional areas on the final protein product. For example, my results show *ORF 25* and *265b* maintained homology to the ATP synthase apparatus – both were found to be highly expressed (Figure 3.14). It is very possible that some ORFs reviewed in my study are hybridized forms of the original genes which are adjusted by these mobile reading frames inserting and reinserting themselves all over the mitogenome (Sellem et al 1996). This type of evolutionary "tinkering" is in accordance with other viewpoints on mitochondrial evolution in plants and animals; DNA sequence repair mechanisms from massive intergenic double strand breakage give these ORFs numerous opportunities to engage in horizontal gene transfer (Maier et al. 2008; Christensen 2013). In the end, however, both arguments are likely valid as the complexity of recombination and post-transcriptional RNA processing in plant mitochondria support both hypotheses. Another intriguing possibility is that these expressed chimeric ORFs are remnants of ancient coding genes that became obsolete through reductive evolution and gene transfer to the nucleus (Khachane et al. 2007; Gray 2001). These genes perhaps slowly eroded over time as their use became usurped by imported nuclear encoded proteins – a process still actively operating.

Ultimately, these observations of chimeric open reading frames lead to one of the most well understood mechanisms that plant mitochondrial ORFs participate in:

cytoplasmic male sterility (CMS). ORFs can become CMS-associated genes that trigger expression events that lead to a sterile cytoplasm (Leino et al. 2005). These genes are commonly novel recombinant ORFs that also include parts of mitochondrial genes (Chang et al. 2011; Carlsson and Glimelius 2012); moreover, these mitochondrial ORF hybrids were found by Hanson and Bentolia 2004 to be frequently associated with genes for *atp* subunits *4, 6, 8,* and *9*. It is apparent that these chimeric ORFs can produce malfunctioning proteins, and when indiscriminately included into the construction of the ATP synthase apparatus, these delinquent "look-alike" hybrid proteins cause ATP production to become greatly diminished causing the cell to suffer from energy starvation leading to abandonment of pollen production (CMS) (Teixeira et al. 2005; Chase 2007). Three ORF chimeric genes, *ORF 25, b*, and *265b,* examined in my study were in fact homologous to *atp 4* and *8*. Possible involvements of these ORFs with two others (ORFs *216* and *159b*) as CMS-associated genes present an intriguing possibility. Yui et al. 2003 showed in tobacco that CMS can be induced by alterations in mitochondrial enzymes linked to the citric acid cycle – in this study, ORF 197 presented as a chimeric gene, homologous to *nad4*; however, it was not found to be ribosome associated in leaf tissue. It is also interesting to note that CMS-associated ORFs have been found to express in tissues other than flowers, such as leaves and roots (Yang et al. 2005). CMS can also have important implications in RNA steady state systems. Leino et al. 2005 showed that cytoplasmic message background alters degradation rates of mtRNA; thus, expressed chimeric ORFs could be a deliberate mode of RNA steady state modulation.

Open Reading Frame Transcription Organization and Processing

Many of the open reading frames in this study were either in close proximity to annotated genes or with other ORFs. It is likely that these ORFs are expressed in heterogenous populations of polycistronic and monocistronic transcripts as proposed by Binder et al. 2012. RNA-seq depth of coverage shows varying expression rates across genes that are in close proximity to one another reaffirming that polycistronic gene organization is not as common in mitochondria as compared to chloroplasts (Figure 3.1) (Barkan and Goldschmidt-Clermont 2000). Taking my qRT-PCR experimental data from the ORFs into account, it is difficult to determine the cistronic ratio the ORFs are transcribed in due to limitations in primer design and local fluctuation of expression; furthermore, the ubiquitous presence of promoters scattered throughout the mitogenome points to many monocistronic units.

When looking at plant mitochondrial transcription as a whole, it is tempting to think that the system is ruled by chaos when looking closely at transcriptome data. This is not the case, as mitochondria (and plastids) have recruited the tools of eukaryotic RNA maintenance and processing. Ever since early eukaryotic cells acquired the alpha-proteobacterial ancestor of the mitochondria, organellar gene exchange with the nucleus of the host cell has been taking place (Gray 2001). As the mitochondria evolved with the eukaryotic cell, the complexity of nuclear RNA maintenance influenced mitochondrial RNA maintenance; hence, communication with the nucleus is crucial to mitochondrial function in order to manage the persistent expression of the mitogenome (Hammani et al. 2011; Binder et al. 2012). Mitochondria are aided by specialized proteins encoded in the

nucleus called pentatricopetide repeat proteins (PPR). They are a diverse family of proteins that have been found to stabilize and process RNA in the chloroplast, nucleus, and mitochondria (Andres et al. 2006). PPR proteins are targeted to the mitochondria, binding to specific sites on the RNA molecule with the aid of their 35 amino acid sequence motif, aiding in transcription, translation, processing, and editing (Zoschke et al. 2012). These proteins also stabilize and protect important gene transcripts destined for protein expression. Transcription and processing of conserved open reading frames like the ones in my study likely recruit these proteins, but data on this mechanism has yet to be attained.

Non-Coding RNAs in Plant Mitochondria?

One of the original ideas behind what these ORFs could be was that they might act as non-coding regulatory RNAs (ncRNAs). Fujii et al. 2011 came to a similar conclusion in their analysis of the *Oryza sativa* mitochondrial transciptome; they found that many intergenic ORFs (iORFs) were found to be transcribed leading their thoughts to possible regulatory RNAs. Moreover, the possibility of mitochondrial ncRNAs in plants has been bolstered by cutting edge research. Ro et al. 2013 positively identified thousands of ncRNAs (mitosRNAs) in human mitochondria opening the door for the possibility of a new shared mitochondrial characteristic among eukaryotes. Chloroplasts also contain and express ncRNAs as described by Hotto et al. 2011. In my study, most of the ORFs chosen were considerably longer which can only qualify these as possible non-coding RNAs if they are to be part of the "long-coding" or lcRNA species. The approach to this experiment did not seek to directly answer whether or not these ORFs

could be classified as regulatory RNAs; however, data suggests that meaningful expression of ORFs in this study (ORFs with conserved sequence motifs) were found to be targeted to ribosomes – a trait not present in ncRNAs.

Expressed ORFs Not Associated with CMS

Some transcribed and/or translated ORFs cannot be linked to a specific process. None of the following ORFs were identified as having conserved gene motifs; however, most possessed homology with many other species' mitogenomes while some did not show any homology to any known genes. The 7 non-chimeric ORFs that were associated with polysomes are good candidates for protein expression analysis (ORFs 138c, 160, 115, 166b, 147, 177, and 222).

ORFs 101c, 111c, and 133 are Chloroplast Gene Fragments

Since organelle separation was outside the scope of this study, one technical difficulty was the inadvertent acquisition of chloroplast transcripts. These three ORFs are examples where chloroplast transcripts contaminated tissue extracts for both RNA-sequencing and qRT-PCR analysis. Sequence homology of these open reading frames point directly to functional genes found in the chloroplast genome. ORFs 101c and 111c are chloroplast *nad* genes; ORF 133 is homologous to *rpl2,* a rRNA gene. Their high expression in leaf tissue is simply because active chloroplasts reside there.

ORFs 138c, 160, 115, 166b, and 147

Each of these ORFs were found to have homology across many species. Although they did not specifically possess conserved motifs, their transcripts were found to be attached to ribosomes. This group of ORFs would be the best place to find novel polypeptides; further analysis can be performed on these ORFs to determine if post-translational mechanisms have degraded or preserved them. In the early 1990's, confirmations of novel plant mitochondrial proteins were already being made. Prioli et al. 1993 discovered *orf221* in *Zea mays* encoded a membrane bound protein through antibody detection and protein topology analysis. Similar analysis can be employed to these ORFs as they present an opportunity to discover new mitochondrial proteins conserved across several species.

ORFs 177 and 222

These two ORFs were unique as they were transcribed *and* attached to ribosomes, but they did not show homology to any known genes. In short, these two ribosome-bound ORFs are tobacco-specific. Orf 222, especially, showed high transcript abundance in the pellet. This could simply be due to long run on transcripts spanning from the nearby, but not overlapping, *nad4* gene. Post-translational modification could be taking place here, processing the run-on *nad4* protein by clipping off the sequence that included *orf222*.

ORFs 129b and 151

These two ORFs were being transcribed and showed homology to viral RNA polymerase and integrase genes; however, they were not loaded onto ribosomes

suggesting that they could be used as antiviral defense RNAs. Mitochondria do indeed use a viral RNA polymerase to transcribe its genes, but it is not encoded within its own genome; thus, the possibility of this RNA polymerase gene being used to transcribe mitochondrial genes is not likely

ORFs 175, 118, 125d, 306, 222, and 239

This group of ORFs confirmed that a measure of post-transcriptional control is exerted on indiscriminately transcribed non-conserved reading frames. There was no homology to known genes or conserved motifs in this entire group making it unlikely that they produce any type of useful protein. One reason for their absence from ribosomes is likely due to these RNA sequences lacking conserved *cis* elements to recruit stem-loop aiding proteins for endonuclease protection (Kuhn et al. 2001).

Works Cited

Adams, K. L., Daley, D. O., Qiu, Y. L., Whelan, J., & Palmer, J. D. (2000). Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature, 408*(6810), 354-357. doi: 10.1038/35042567

Arrieta-Montiel, M. P., & Mackenzie, S. A. (2011). Plant Mitochondrial Genomes and Recombination. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 65-82). New York, NY: Springer.

Arrieta-Montiel, M. P., Shedge, V., Davila, J., Christensen, A. C., & Mackenzie, S. A. (2009). Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics, 183*(4), 1261-1268. doi: 10.1534/genetics.109.108514

Baba, K., Schmidt, J., Espinosa-Ruiz, A., Villarejo, A., Shiina, T., Gardestrom, P., . . . Bhalerao, R. P. (2004). Organellar gene transcription and early seedling development are affected in the rpoT;2 mutant of Arabidopsis. *Plant J, 38*(1), 38-48. doi: 10.1111/j.1365-313X.2004.02022.x

Backert, S., Lurz, R., Oyarzabal, O. A., & Borner, T. (1997). High content, size and distribution of single-stranded DNA in the mitochondria of Chenopodium album (L.). *Plant Mol Biol, 33*(6), 1037-1050.

Balk, J., & Leaver, C. J. (2001). The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome c release. *Plant Cell, 13*(8), 1803-1818.

Barkan, A., & Goldschmidt-Clermont, M. (2000). Participation of nuclear genes in chloroplast gene expression. *Biochimie, 82*(6-7), 559-572.

Beardslee, T. A., Roy-Chowdhury, S., Jaiswal, P., Buhot, L., Lerbs-Mache, S., Stern, D. B., & Allison, L. A. (2002). A nucleus-encoded maize protein with sigma factor activity accumulates in mitochondria and chloroplasts. *Plant J, 31*(2), 199-209.

Beaudet, D., Nadimi, M., Iffis, B., & Hijri, M. (2013). Rapid Mitochondrial Genome Evolution through Invasion of Mobile Elements in Two Closely Related Species of Arbuscular Mycorrhizal Fungi. *PLoS One, 8*(4), e60768. doi: 10.1371/journal.pone.0060768

Binder, S., & Brennicke, A. (2003). Gene expression in plant mitochondria: transcriptional and post-transcriptional control. *Philos Trans R Soc Lond B Biol Sci, 358*(1429), 181-188; discussion 188-189. doi: 10.1098/rstb.2002.1179

Binder, S., Holzle, A., & Jonietz, C. (2011). RNA Processing and RNA Stability in Plant Mitochondria. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 107-130). New York, NY: Springer.

Bonen, L. (2008). Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion, 8*(1), 26-34. doi: 10.1016/j.mito.2007.09.005

Bonen, L. (2011). RNA Splicing in Plant Mitochondria. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 131-155). New York, NY: Springer.

Bruhs, A., & Kempken, F. (2011). RNA Editing in Higher Plant Mitochondria. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 157-175). New York, NY: Springer.

Buchanan, B. B., Gruissem, W., & Jones, R. L. (Eds.). (2001). *Biochemistry and Molecular Biology of Plants*. Rockville, MD: American Society of Plant Physiologists.

Bullerwell, C. E., & Gray, M. W. (2004). Evolution of the mitochondrial genome: protist connections to animals, fungi and plants. *Curr Opin Microbiol, 7*(5), 528-534. doi: 10.1016/j.mib.2004.08.008

Canino, G., Bocian, E., Barbezier, N., Echeverria, M., Forner, J., Binder, S., & Marchfelder, A. (2009). Arabidopsis encodes four tRNase Z enzymes. *Plant Physiol, 150*(3), 1494-1502. doi: 10.1104/pp.109.137950

Carlsson, J., & Glimelius, K. (2011). Cytoplasmic Male Sterility and Nuclear Encoded Fertility Restoration. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 469-492). New York, NY: Springer.

Chang, S., Yang, T., Du, T., Huang, Y., Chen, J., Yan, J., . . . Guan, R. (2011). Mitochondrial genome sequencing helps show the evolutionary mechanism of mitochondrial genome formation in Brassica. *BMC Genomics, 12*, 497. doi: 10.1186/1471-2164-12-497

Chen, X., Wang, Y., Li, J., Jiang, A., Cheng, Y., & Zhang, W. (2009). Mitochondrial proteome during salt stress-induced programmed cell death in rice. *Plant Physiol Biochem, 47*(5), 407-415. doi: 10.1016/j.plaphy.2008.12.021

Christensen, A. C. (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol*. doi: 10.1093/gbe/evt069

Clifton, S. W., Minx, P., Fauron, C. M., Gibson, M., Allen, J. O., Sun, H., . . . Newton, K. J. (2004). Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol, 136*(3), 3486-3503. doi: 10.1104/pp.104.044602

Costa, V., Angelini, C., De Feis, I., & Ciccodicola, A. (2010). Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol, 2010*, 853916. doi: 10.1155/2010/853916

Dalbey, R. E., & Kuhn, A. (2000). Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu Rev Cell Dev Biol, 16*, 51-87. doi: 10.1146/annurev.cellbio.16.1.51

de Longevialle, A. F., Meyer, E. H., Andres, C., Taylor, N. L., Lurin, C., Millar, A. H., & Small, I. D. (2007). The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 Intron 1 in Arabidopsis thaliana. *Plant Cell, 19*(10), 3256-3265. doi: 10.1105/tpc.107.054841

Dombrowski, S., & Binder, S. (1998). 3' inverted repeats in plant mitochondrial mRNAs act as processing and stablizing elements but do not terminate transcription In I. M. Moller, P. Gardestrom, K. Glimelius & E. Glaser (Eds.), *Plant Mitochondria: from Gene to Function* (pp. 9-12). Leiden, Netherlands: Backhuys Publications.

Dombrowski, S., Hoffman, M., Kuhn, J., Brennicke, A., & Binder, S. (1998). On mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In I. M. Moller, P. Gardestrom, K. Glimelius & E. Glaser (Eds.), *Plant Mitochondria: From Gene to Function* (pp. 165-170). Leiden, Netherlands: Backhuys, Publications.

Duchene, A. M., & Giege, P. (2012). Dual localized mitochondrial and nuclear proteins as gene expression regulators in plants? *Front Plant Sci, 3*, 221. doi: 10.3389/fpls.2012.00221

Farre, J. C., & Araya, A. (1999). The mat-r open reading frame is transcribed from a non-canonical promoter and contains an internal promoter to co-transcribe exons nad1e and nad5III in wheat mitochondria. *Plant Mol Biol*, *40*(6), 959-967.

Fauron, C., Allen, J. O., Clifton, S. W., & Newton, K. J. (2004). Plant Mitochondrial Genomes. In H. Daniell & C. Chase (Eds.), *Molecular Biology and Biotechnology of Plant Organelles* (pp. 151-177). Dordrecht, Netherlands: Kluwer Academic Publishers.

Fauron, C., Casper, M., Gao, Y., & Moore, B. (1995). The maize mitochondrial genome: dynamic, yet functional. *Trends Genet*, *11*(6), 228-235.

Feagin, J. E. (2000). Mitochondrial genome diversity in parasites. *Int J Parasitol*, *30*(4), 371-390.

Fey, J., & Marechal-Drouard, L. (1999). Compilation and analysis of plant mitochondrial promoter sequences: An illustration of a divergent evolution between monocot and dicot mitochondria. *Biochem Biophys Res Commun*, *256*(2), 409-414. doi: 10.1006/bbrc.1999.0349

Forner, J., Weber, B., Thuss, S., Wildum, S., & Binder, S. (2007). Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Res*, *35*(11), 3676-3692. doi: 10.1093/nar/gkm270

Fujii, S., Toda, T., Kikuchi, S., Suzuki, R., Yokoyama, K., Tsuchida, H., . . . Toriyama, K. (2011). Transcriptome map of plant mitochondria reveals islands of unexpected transcribed regions. *BMC Genomics, 12*, 279. doi: 10.1186/1471-2164-12-279

Gagliardi, D., & Binder, S. (2007). Expression of the Plant Mitochondrial Genome. In D. Logan (Ed.), *Plant Mitochondria* (pp. 50-95). Ames, IA: Blackwell Publishing.

Gagliardi, D., & Leaver, C. J. (1999). Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J, 18*(13), 3757-3766. doi: 10.1093/emboj/18.13.3757

Gallagher, L. J., Betz, S. K., & Chase, C. D. (2002). Mitochondrial RNA editing truncates a chimeric open reading frame associated with S male-sterility in maize. *Curr Genet, 42*(3), 179-184. doi: 10.1007/s00294-002-0344-5

Giege, P., & Brennicke, A. (1999). RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs. *Proc Natl Acad Sci U S A, 96*(26), 15324-15329.

Giege, P., Hoffmann, M., Binder, S., & Brennicke, A. (2000). RNA degradation buffers asymmetries of transcription in Arabidopsis mitochondria. *EMBO Rep, 1*(2), 164-170. doi: 10.1038/sj.embor.embor613

Gray, M. W. (1998). Rickettsia, typhus and the mitochondrial connection. *Nature, 396*(6707), 109-110. doi: 10.1038/24030

Gray, M. W., Burger, G., & Lang, B. F. (1999). Mitochondrial evolution. *Science, 283*(5407), 1476-1481.

Gray, M. W., Burger, G., & Lang, B. F. (2001). The origin and early evolution of mitochondria. *Genome Biol, 2*(6), REVIEWS1018.

Hanson, M. R., & Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell, 16 Suppl*, S154-169. doi: 10.1105/tpc.015966

Hartmann, C., Henry, Y., Tregear, J., & Rode, A. (2000). Nuclear control of mitochondrial genome reorganization characterized using cultured cells of ditelosomic and nullisomic-tetrasomic wheat lines. *Curr Genet, 38*(3), 156-162.

Hedtke, B., Borner, T., & Weihe, A. (1997). Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. *Science, 277*(5327), 809-811.

Henrichsen, C. N., Vinckenbosch, N., Zollner, S., Chaignat, E., Pradervand, S., Schutz, F., . . . Reymond, A. (2009). Segmental copy number variation shapes tissue transcriptomes. *Nat Genet, 41*(4), 424-429. doi: 10.1038/ng.345

Hess, W. R., & Borner, T. (1999). Organellar RNA polymerases of higher plants. *Int Rev Cytol, 190*, 1-59.

Hiesel, R., Combettes, B., & Brennicke, A. (1994). Evidence for RNA editing in mitochondria of all major groups of land plants except the Bryophyta. *Proc Natl Acad Sci U S A, 91*(2), 629-633.

Hoffmann, M., & Binder, S. (2002). Functional importance of nucleotide identities within the pea atp9 mitochondrial promoter sequence. *J Mol Biol, 320*(5), 943-950.

Holec, S., Lange, H., Canaday, J., & Gagliardi, D. (2008). Coping with cryptic and defective transcripts in plant mitochondria. *Biochim Biophys Acta, 1779*(9), 566-573. doi: 10.1016/j.bbagrm.2008.02.004

Hotto, A. M., Schmitz, R. J., Fei, Z., Ecker, J. R., & Stern, D. B. (2011). Unexpected Diversity of Chloroplast Noncoding RNAs as Revealed by Deep Sequencing of the Arabidopsis Transcriptome. *G3 (Bethesda), 1*(7), 559-570. doi: 10.1534/g3.111.000752

Huang, G. S., Millar, A. H., & Taylor, L. T. (2011). Plant Mitochondria. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 207-239). New York, NY: Springer.

Huertas, P. (2010). DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol, 17*(1), 11-16. doi: 10.1038/nsmb.1710

Kempken, F., Hofken, G., & Pring, D. R. (1995). Analysis of silent RNA editing sites in atp6 transcripts of Sorghum bicolor. *Curr Genet, 27*(6), 555-558.

Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., . . . Ostersetzer-Biran, O. (2009). AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. *RNA, 15*(12), 2299-2311. doi: 10.1261/rna.1776409

Khachane, A. N., Timmis, K. N., & Martins dos Santos, V. A. (2007). Dynamics of reductive genome evolution in mitochondria and obligate intracellular microbes. *Mol Biol Evol, 24*(2), 449-456. doi: 10.1093/molbev/msl174

Khazi, F. R., Edmondson, A. C., & Nielsen, B. L. (2003). An Arabidopsis homologue of bacterial RecA that complements an E. coli recA deletion is targeted to plant mitochondria. *Mol Genet Genomics, 269*(4), 454-463. doi: 10.1007/s00438-003-0859-6

Knoop, V., Volkmar, U., Hecht, J., & Grewe, F. (2011). Mitochondrial Genome Evolution in the Plant Lineage. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 3-29). New York, NY: Springer.

Kosa, P., Valach, M., Tomaska, L., Wolfe, K. H., & Nosek, J. (2006). Complete DNA sequences of the mitochondrial genomes of the pathogenic yeasts Candida orthopsilosis and Candida metapsilosis: insight into the evolution of linear DNA

genomes from mitochondrial telomere mutants. *Nucleic Acids Res, 34*(8), 2472-2481. doi: 10.1093/nar/gkl327

Kotera, E., Tasaka, M., & Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature, 433*(7023), 326-330. doi: 10.1038/nature03229

Kruft, V., Eubel, H., Jansch, L., Werhahn, W., & Braun, H. P. (2001). Proteomic approach to identify novel mitochondrial proteins in Arabidopsis. *Plant Physiol, 127*(4), 1694-1710.

Kugita, M., Yamamoto, Y., Fujikawa, T., Matsumoto, T., & Yoshinaga, K. (2003). RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res, 31*(9), 2417-2423.

Kuhn, K., Bohne, A. V., Liere, K., Weihe, A., & Borner, T. (2007). Arabidopsis phage-type RNA polymerases: accurate in vitro transcription of organellar genes. *Plant Cell, 19*(3), 959-971. doi: 10.1105/tpc.106.046839

Kuhn, K., Richter, U., Meyer, E. H., Delannoy, E., de Longevialle, A. F., O'Toole, N., . . . Whelan, J. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. *Plant Cell, 21*(9), 2762-2779. doi: 10.1105/tpc.109.068536

Kuhn, K., Weihe, A., & Borner, T. (2005). Multiple promoters are a common feature of mitochondrial genes in Arabidopsis. *Nucleic Acids Res, 33*(1), 337-346. doi: 10.1093/nar/gki179

Kunzmann, A., Brennicke, A., & Marchfelder, A. (1998). 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. *Proc Natl Acad Sci U S A, 95*(1), 108-113.

Lambowitz, A. M., & Zimmerly, S. (2004). Mobile group II introns. *Annu Rev Genet, 38*, 1-35. doi: 10.1146/annurev.genet.38.072902.091600

Lane, N., & Martin, W. (2010). The energetics of genome complexity. *Nature, 467*(7318), 929-934. doi: 10.1038/nature09486

Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., . . . Gray, M. W. (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature, 387*(6632), 493-497. doi: 10.1038/387493a0

Lang, B. F., Laforest, M. J., & Burger, G. (2007). Mitochondrial introns: a critical view. *Trends Genet, 23*(3), 119-125. doi: 10.1016/j.tig.2007.01.006

Leino, M., Landgren, M., & Glimelius, K. (2005). Alloplasmic effects on mitochondrial transcriptional activity and RNA turnover result in accumulated transcripts of Arabidopsis orfs in cytoplasmic male-sterile Brassica napus. *Plant J, 42*(4), 469-480. doi: 10.1111/j.1365-313X.2005.02389.x

Liere, K., & Borner, T. (2011). Transcription in Plant Mitochondria. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 85-105). New York, NY: Springer.

Liere, K., Weihe, A., & Borner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. *J Plant Physiol, 168*(12), 1345-1360. doi: 10.1016/j.jplph.2011.01.005

Lilly, J. W., & Havey, M. J. (2001). Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. *Genetics, 159*(1), 317-328.

Linder, T., Park, C. B., Asin-Cayuela, J., Pellegrini, M., Larsson, N. G., Falkenberg, M., . . . Gustafsson, C. M. (2005). A family of putative transcription termination factors shared amongst metazoans and plants. *Curr Genet, 48*(4), 265-269. doi: 10.1007/s00294-005-0022-5

Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Scott, M. P., Bretscher, A., . . . Matsudaira, P. (2008). *Molecular Cell Biology* (Vol. 6). New York, NY: W.H. Freeman and Company.

Lu, B., & Hanson, M. R. (1994). A single homogeneous form of ATP6 protein accumulates in petunia mitochondria despite the presence of differentially edited atp6 transcripts. *Plant Cell, 6*(12), 1955-1968. doi: 10.1105/tpc.6.12.1955

Maier, U. G., Bozarth, A., Funk, H. T., Zauner, S., Rensing, S. A., Schmitz-Linneweber, C., . . . Tillich, M. (2008). Complex chloroplast RNA metabolism: just debugging the genetic programme? *BMC Biol, 6*, 36. doi: 10.1186/1741-7007-6-36

Manchekar, M., Scissum-Gunn, K., Song, D., Khazi, F., McLean, S. L., & Nielsen, B. L. (2006). DNA recombination activity in soybean mitochondria. *J Mol Biol, 356*(2), 288-299. doi: 10.1016/j.jmb.2005.11.070

Marienfeld, J., Unseld, M., & Brennicke, A. (1999). The mitochondrial genome of Arabidopsis is composed of both native and immigrant information. *Trends Plant Sci, 4*(12), 495-502.

Martin, G., & Keller, W. (2004). Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. *RNA, 10*(6), 899-906.

Matthes, A., Schmidt-Gattung, S., Kohler, D., Forner, J., Wildum, S., Raabe, M., . . . Binder, S. (2007). Two DEAD-box proteins may be part of RNA-dependent high-molecular-mass protein complexes in Arabidopsis mitochondria. *Plant Physiol, 145*(4), 1637-1646. doi: 10.1104/pp.107.108076

McVey, M., & Lee, S. E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet, 24*(11), 529-538. doi: 10.1016/j.tig.2008.08.007

Millar, A. H., Sweetlove, L. J., Giege, P., & Leaver, C. J. (2001). Analysis of the Arabidopsis mitochondrial proteome. *Plant Physiol, 127*(4), 1711-1727.

Mohr, S., Matsuura, M., Perlman, P. S., & Lambowitz, A. M. (2006). A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc Natl Acad Sci U S A, 103*(10), 3569-3574. doi: 10.1073/pnas.0600332103

Nayak, D., Guo, Q., & Sousa, R. (2009). A promoter recognition mechanism common to yeast mitochondrial and phage t7 RNA polymerases. *J Biol Chem, 284*(20), 13641-13647. doi: 10.1074/jbc.M900718200

Newton, K. J., Winberg, B., Yamato, K., Lupold, S., & Stern, D. B. (1995). Evidence for a novel mitochondrial promoter preceding the cox2 gene of perennial teosintes. *EMBO J, 14*(3), 585-593.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., . . . Kadowaki, K. (2002). The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the

evolution of flowering plants. *Mol Genet Genomics, 268*(4), 434-445. doi: 10.1007/s00438-002-0767-1

Palmer, J. D. (1990). Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet, 6*(4), 115-120.

Perrin, R., Lange, H., Grienenberger, J. M., & Gagliardi, D. (2004). AtmtPNPase is required for multiple aspects of the 18S rRNA metabolism in Arabidopsis thaliana mitochondria. *Nucleic Acids Res, 32*(17), 5174-5182. doi: 10.1093/nar/gkh852

Raczynska, K. D., Le Ret, M., Rurek, M., Bonnard, G., Augustyniak, H., & Gualberto, J. M. (2006). Plant mitochondrial genes can be expressed from mRNAs lacking stop codons. *FEBS Lett, 580*(24), 5641-5646. doi: 10.1016/j.febslet.2006.09.010

Rathburn, H. B., & Hedgcoth, C. (1991). A chimeric open reading frame in the 5' flanking region of coxI mitochondrial DNA from cytoplasmic male-sterile wheat. *Plant Mol Biol, 16*(5), 909-912.

Richardson, A. O., & Palmer, J. D. (2007). Horizontal gene transfer in plants. *J Exp Bot, 58*(1), 1-9. doi: 10.1093/jxb/erl148

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

Salinas, T., Schaeffer, C., Marechal-Drouard, L., & Duchene, A. M. (2005). Sequence dependence of tRNA(Gly) import into tobacco mitochondria. *Biochimie, 87*(9-10), 863-872. doi: 10.1016/j.biochi.2005.04.004

Sandhu, A. P., Abdelnoor, R. V., & Mackenzie, S. A. (2007). Transgenic induction of mitochondrial rearrangements for cytoplasmic male sterility in crop plants. *Proc Natl Acad Sci U S A, 104*(6), 1766-1770. doi: 10.1073/pnas.0609344104

Schofield, M. J., & Hsieh, P. (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol, 57*, 579-608. doi: 10.1146/annurev.micro.57.030502.090847

Scott, I., & Logan, D. C. (2011). Mitochondrial Dynamics. In F. Kempken (Ed.), *Plant Mitochondria* (pp. 31-63). New York, NY: Springer.

Sellem, C. H., d'Aubenton-Carafa, Y., Rossignol, M., & Belcour, L. (1996). Mitochondrial intronic open reading frames in Podospora: mobility and consecutive exonic sequence variations. *Genetics, 143*(2), 777-788.

Sharpe, R. M., Dunn, S. N., & Cahoon, A. B. (2008). A plastome primer set for comprehensive quantitative real time RT-PCR analysis of Zea mays: a starter primer set for other Poaceae species. *Plant Methods, 4*, 14. doi: 10.1186/1746-4811-4-14

Smits, P., Smeitink, J., & van den Heuvel, L. (2010). Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies. *J Biomed Biotechnol, 2010*, 737385. doi: 10.1155/2010/737385

Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A., & Sugiura, M. (2005). The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol Genet Genomics, 272*(6), 603-615. doi: 10.1007/s00438-004-1075-8

Takenaka, M. (2010). MEF9, an E-subclass pentatricopeptide repeat protein, is required for an RNA editing event in the nad7 transcript in mitochondria of Arabidopsis. *Plant Physiol, 152*(2), 939-947. doi: 10.1104/pp.109.151175

Takenaka, M., & Brennicke, A. (2009). Multiplex single-base extension typing to identify nuclear genes required for RNA editing in plant organelles. *Nucleic Acids Res, 37*(2), e13. doi: 10.1093/nar/gkn975

Teixeira, R. T., Farbos, I., & Glimelius, K. (2005). Expression levels of meristem identity and homeotic genes are modified by nuclear-mitochondrial interactions in alloplasmic male-sterile lines of Brassica napus. *Plant J, 42*(5), 731-742. doi: 10.1111/j.1365-313X.2005.02407.x

Tracy, R. L., & Stern, D. B. (1995). Mitochondrial transcription initiation: promoter structures and RNA polymerases. *Curr Genet, 28*(3), 205-216.

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., . . . Liu, Y. G. (2006). Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell, 18*(3), 676-687. doi: 10.1105/tpc.105.038240

Weihe, A. (2004). The transcription of plant organelle genomes. In H. Daniell & C. D. Chase (Eds.), *Molecular Biology and Biotechnology of Plant Organellles* (pp. 213-237). Dordrecht, Netherlands: Kluwer Academic Publishers.

Wickner, W., & Schekman, R. (2005). Protein translocation across biological membranes. *Science, 310*(5753), 1452-1456. doi: 10.1126/science.1113752

Yang, J. H., Zhang, F., Yu, Q., Wang, T., Chen Z.J. (2005). Molecular identification of cytoplasmic male sterility associated gene orf 220 in leaf mustard (Brassica juncea var. multiceps Tsen et Lee). *Yi Chuan Xue Bao*, *32*(6): 594-599.

Yen, M. R., Tseng, Y. H., Nguyen, E. H., Wu, L. F., & Saier, M. H., Jr. (2002). Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol, 177*(6), 441-450. doi: 10.1007/s00203-002-0408-4

Appendices

Appendix A: Ct Values for Tobacco Leaves in Transcription Analysis

Gene/Replicate	Amplicon Size	Efficiency	Crossover Threhold	Copy #
Cox2_A	108	2	15.43	15,730.58
Cox2_A	108	2	15.45	15,514.01
Cox2_B	108	2	16.28	8,727.08
Cox2_B	108	2	16.21	9,160.97
Cox2_C	108	2	15.45	15,514.01
Cox2_C	108	2	15.43	15,730.58
Bkgrd_A	146	2.02	25.08	11.28
Bkgrd_A	146	2.02	26.16	5.28
Bkgrd_B	146	2.02	27.57	1.96
Bkgrd_B	146	2.02	27.21	2.52
Bkgrd_C	146	2.02	25.1	11.13
Bkgrd_C	146	2.02	25.35	9.33
Orf177_A	99	1.99	21.26	335.59
Orf177_A	99	1.99	21.3	326.48
Orf177_B	99	1.99	23.23	86.51
Orf177_B	99	1.99	23.5	71.84
Orf177_C	99	1.99	22.37	156.34
Orf177_C	99	1.99	22.73	122.04
Orf129b_A	88	1.97	23.3	117.36
Orf129b_A	88	1.97	23.74	87.09
Orf129b_B	88	1.97	23.75	86.50
Orf129b_B	88	1.97	23.7	89.48
Orf129b_C	88	1.97	23.05	139.04
Orf129b_C	88	1.97	22.46	207.43
Orf151_A	124	1.99	23.12	74.50
Orf151_A	124	1.99	23.4	61.44
Orf151_B	124	1.99	24.18	35.92
Orf151_B	124	1.99	24.28	33.53
Orf151_C	124	1.99	24	40.66
Orf151_C	124	1.99	23.36	63.16
Orf175_A	129	2.05	23.88	20.88
Orf175_A	129	2.05	23.95	19.86
Orf175_B	129	2.05	24.56	12.82
Orf175_B	129	2.05	24.2	16.60

Orf175_C	129	2.05	23.26	32.59
Orf175_C	129	2.05	23.35	30.55
Orf134_A	98	2	17.27	4,842.25
Orf134_A	98	2	17.69	3,619.22
Orf134_B	98	2	18.66	1,847.63
Orf134_B	98	2	21.68	227.77
Orf134_C	98	2	19.32	1,169.33
Orf134_C	98	2	19.96	750.37
Orf306_A	140	2	23.62	41.55
Orf306_A	140	2	23.39	48.74
Orf306_B	140	2	25.06	15.32
Orf306_B	140	2	24.33	25.40
Orf306_C	140	2	22.1	119.17
Orf306_C	140	2	22.08	120.83
Orf138c_A	116	2	23.47	55.65
Orf138c_A	116	2	23.61	50.50
Orf138c_B	116	2	24.82	21.83
Orf138c_B	116	2	25	19.27
Orf138c_C	116	2	24.11	35.71
Orf138c_C	116	2	24.19	33.78
Orf144_A	143	2	23.14	56.74
Orf144_A	143	2	23.64	40.12
Orf144_B	143	2	25.15	14.09
Orf144_B	143	2	25.37	12.09
Orf144_C	143	2	22.92	66.09
Orf144_C	143	2	23.22	53.68
Orf118_A	102	2	21.05	338.67
Orf118_A	102	2	21.11	324.88
Orf118_B	102	2	22.21	151.56
Orf118_B	102	2	22.25	147.42
Orf118_C	102	2	21.58	234.55
Orf118_C	102	2	21.69	217.33
Orf160_A	80	2	19.39	1,364.58
Orf160_A	80	2	19.23	1,524.63
Orf160_B	80	2	20.09	840.00
Orf160_B	80	2	19.79	1,034.16
Orf160_C	80	2	19.18	1,578.40

Orf160_C	80	2	19.26	1,493.25
Orf101c_A	119	2	16.06	9,225.13
Orf101c_A	119	2	16.18	8,488.85
Orf101c_B	119	2	16.87	5,261.85
Orf101c_B	119	2	17.13	4,394.11
Orf101c_C	119	2	15.57	12,956.19
Orf101c_C	119	2	16.14	8,727.51
Orf111c_A	127	2.04	15.77	7,733.74
Orf111c_A	127	2.04	15.94	6,850.98
Orf111c_B	127	2.04	15.95	6,802.31
Orf111c_B	127	2.04	16.02	6,471.16
Orf111c_C	127	2.04	15.26	11,125.03
Orf111c_C	127	2.04	15.46	9,646.62
Orf125d_A	84	2.02	25.13	18.94
Orf125d_A	84	2.02	25.07	19.75
Orf125d_B	84	2.02	25.1	19.34
Orf125d_B	84	2.02	26	10.27
Orf125d_C	84	2.02	25.37	16.00
Orf125d_C	84	2.02	25.39	15.77
OrfB_A	82	1.98	18.7	2,591.84
OrfB_A	82	1.98	17.85	4,632.06
OrfB_B	82	1.98	20.15	962.60
OrfB_B	82	1.98	20.61	703.04
OrfB_C	82	1.98	18.67	2,645.50
OrfB_C	82	1.98	18.1	3,904.88
Orf115_A	100	1.96	22.01	277.02
Orf115_A	100	1.96	22.31	226.37
Orf115_B	100	1.96	20.71	664.41
Orf115_B	100	1.96	21	546.62
Orf115_C	100	1.96	20.72	659.96
Orf115_C	100	1.96	20.96	561.53
Orf166_A	81	2.04	19.94	620.23
Orf166_A	81	2.04	19.91	633.64
Orf166_B	81	2.04	21.01	289.23
Orf166_B	81	2.04	20.71	358.21
Orf166_C	81	2.04	19.59	796.02
Orf166_C	81	2.04	19.81	680.46

Orf133_A	118	2	15.58	12,975.73
Orf133_A	118	2		
Orf133_B	118	2	17.39	3,700.56
Orf133_B	118	2		
Orf133_C	118	2	17.49	3,452.75
Orf133_C	118	2		
Orf25_A	83	2	16.46	10,023.73
Orf25_A	83	2		
Orf25_B	83	2	17.69	4,273.29
Orf25_B	83	2		
Orf25_C	83	2	16.90	7,388.84
Orf25_C	83	2		
Orf222_A	84	2	17.27	5,649.29
Orf222_A	84	2		
Orf222_B	84	2	19.61	1,115.79
Orf222_B	84	2		
Orf222_C	84	2	18.71	2,082.14
Orf222_C	84	2		
Orf216_A	114	2	16.73	6,052.36
Orf216_A	114	2		
Orf216_B	114	2	17.15	4,523.68
Orf216_B	114	2		
Orf216_C	114	2	18.04	2,441.04
Orf216_C	114	2		
Orf265a_A	169	2	19.06	811.98
Orf265a_A	169	2		
Orf265a_B	169	2	20.18	373.58
Orf265a_B	169	2		
Orf265a_C	169	2	21.23	180.43
Orf265a_C	169	2		
Orf265b_A	95	2	14.89	26,001.66
Orf265b_A	95	2		
Orf265b_B	95	2	15.44	17,759.66
Orf265b_B	95	2		
Orf265b_C	95	2	16.33	9,583.36
Orf265b_C	95	2		
Orf159b_A	121	2	16.33	7,524.13

Orf159b_A	121	2		
Orf159b_B	121	2	17.34	3,736.08
Orf159b_B	121	2		
Orf159b_C	121	2	17.28	3,894.73
Orf159b_C	121	2		
Orf197_A	141	2	22.42	94.79
Orf197_A	141	2		
Orf197_B	141	2	24.15	28.57
Orf197_B	141	2		
Orf197_C	141	2	23.97	32.37
Orf197_C	141	2		
Orf239_A	144	2	21.58	166.14
Orf239_A	144	2		
Orf239_B	144	2	23.41	46.73
Orf239_B	144	2		
Orf239_C	144	2	23.71	37.96
Orf239_C	144	2		
Orf147_A	96	2	22.42	139.22
Orf147_A	96	2		
Orf147_B	96	2	22.97	95.09
Orf147_B	96	2		
Orf147_C	96	2	23.08	88.11
Orf147_C	96	2		

Appendix B: Ct Values for Tobacco Roots in Transcription Analysis

Gene/Replicate	Ampicon Size	Efficiency	Crossover Threshold	Copy #
Cox2_A	108	2	13.93	44,492.79
Cox2_A	108	2	14.07	40,378.06
Cox2_B	108	2	13.64	54,398.68
Cox2_B	108	2	13.92	44,802.26
Cox2_C	108	2	14.17	37,674.06
Cox2_C	108	2	14.18	37,413.83
Bkgrd_A	146	2.02	24.29	19.67
Bkgrd_A	146	2.02	25.79	6.85
Bkgrd_B	146	2.02	24.21	20.80
Bkgrd_B	146	2.02	22.09	92.36
Bkgrd_C	146	2.02	23.96	24.80
Bkgrd_C	146	2.02	26.83	3.30
Orf177_A	99	1.99	18.59	2,107.40
Orf177_A	99	1.99	18.65	2,022.16
Orf177_B	99	1.99	18.48	2,273.11
Orf177_B	99	1.99	18.13	2,892.14
Orf177_C	99	1.99	19.89	861.47
Orf177_C	99	1.99	19.54	1,096.07
Orf129b_A	88	1.97	20.09	1,034.55
Orf129b_A	88	1.97	20.35	867.34
Orf129b_B	88	1.97	19.65	1,394.17
Orf129b_B	88	1.97	18.85	2,398.23
Orf129b_C	88	1.97	20.6	732.11
Orf129b_C	88	1.97	20.6	732.11
Orf151_A	124	1.99	20.67	402.11
Orf151_A	124	1.99	20.21	551.85
Orf151_B	124	1.99	19.31	1,025.15
Orf151_B	124	1.99	19.42	950.41
Orf151_C	124	1.99	20.93	336.24
Orf151_C	124	1.99	21.19	281.15
Orf175_A	129	2.05	21.64	104.26
Orf175_A	129	2.05	21.71	99.15
Orf175_B	129	2.05	20.14	306.02
Orf175_B	129	2.05	19.9	363.55

Orf175_C	129	2.05	21.65	103.51
Orf175_C	129	2.05	21.85	89.67
Orf134_A	98	2	16.88	6,345.25
Orf134_A	98	2	17.51	4,100.15
Orf134_B	98	2	16.41	8,788.87
Orf134_B	98	2	17.08	5,523.86
Orf134_C	98	2	18.07	2,781.14
Orf134_C	98	2	19.41	1,098.61
Orf306_A	140	2	21.22	219.32
Orf306_A	140	2	24.94	16.64
Orf306_B	140	2	21.8	146.72
Orf306_B	140	2	20.87	279.54
Orf306_C	140	2	21.56	173.27
Orf306_C	140	2	21.57	172.07
Orf138c_A	116	2	20.87	337.37
Orf138c_A	116	2	20.87	337.37
Orf138c_B	116	2	20.43	457.68
Orf138c_B	116	2	20.39	470.55
Orf138c_C	116	2	22.52	107.50
Orf138c_C	116	2	22.41	116.02
Orf144_A	143	2	22.38	96.09
Orf144_A	143	2	21.78	145.64
Orf144_B	143	2	20.96	257.12
Orf144_B	143	2	20.94	260.71
Orf144_C	143	2	22.58	83.65
Orf144_C	143	2	22.38	96.09
Orf118_A	102	2	19.07	1,336.04
Orf118_A	102	2	18.97	1,431.93
Orf118_B	102	2	18.66	1,775.18
Orf118_B	102	2	18.67	1,762.91
Orf118_C	102	2	20.12	645.26
Orf118_C	102	2	20.21	606.24
Orf160_A	80	2	17.38	5,496.29
Orf160_A	80	2	17.66	4,526.70
Orf160_B	80	2	17.56	4,851.60
Orf160_B	80	2	17.58	4,784.80
Orf160_C	80	2	18.01	3,551.58

Orf160_C	80	2	18.21	3,091.83
Orf101c_A	119	2	18.48	1,723.77
Orf101c_A	119	2	18.87	1,315.46
Orf101c_B	119	2	18.38	1,847.49
Orf101c_B	119	2	18.62	1,564.36
Orf101c_C	119	2	20.09	564.71
Orf101c_C	119	2	19.88	653.19
Orf111c_A	127	2.04	18.52	1,088.70
Orf111c_A	127	2.04	18.66	985.28
Orf111c_B	127	2.04	18.06	1,511.25
Orf111c_B	127	2.04	17.94	1,646.24
Orf111c_C	127	2.04	18.96	795.55
Orf111c_C	127	2.04	19.14	699.74
Orf125d_A	84	2.02	22.74	101.65
Orf125d_A	84	2.02	23.47	60.84
Orf125d_B	84	2.02	21.59	228.16
Orf125d_B	84	2.02	21.36	268.21
Orf125d_C	84	2.02	24.14	37.98
Orf125d_C	84	2.02	24.42	31.20
OrfB_A	82	1.98	17.2	7,221.15
OrfB_A	82	1.98	17.63	5,383.19
OrfB_B	82	1.98	17.11	7,679.02
OrfB_B	82	1.98	16.72	10,023.18
OrfB_C	82	1.98	18	4,180.94
OrfB_C	82	1.98	19.76	1,256.44
Orf115_A	100	1.96	19.14	1,911.09
Orf115_A	100	1.96	20.18	949.15
Orf115_B	100	1.96	18.6	2,748.52
Orf115_B	100	1.96	18.77	2,451.40
Orf115_C	100	1.96	20.19	942.78
Orf115_C	100	1.96	20.14	975.04
Orf166_A	81	2.04	17.3	4,073.56
Orf166_A	81	2.04	17.35	3,930.91
Orf166_B	81	2.04	17.43	3,712.98
Orf166_B	81	2.04	18.53	1,694.84
Orf166_C	81	2.04	17.85	2,752.19
Orf166_C	81	2.04	18.36	1,913.23

Orf133_A	118	2	19.34	957.77
Orf133_A	118	2		
Orf133_B	118	2	19.42	906.10
Orf133_B	118	2		
Orf133_C	118	2	19.63	783.36
Orf133_C	118	2		
Orf25_A	83	2	16.03	13,504.31
Orf25_A	83	2		
Orf25_B	83	2	16.06	13,226.39
Orf25_B	83	2		
Orf25_C	83	2	16.11	12,775.85
Orf25_C	83	2		
Orf222_A	84	2	17.98	3,453.53
Orf222_A	84	2		
Orf222_B	84	2	18.12	3,134.14
Orf222_B	84	2		
Orf222_C	84	2	18.69	2,111.21
Orf222_C	84	2		
Orf216_A	114	2	15.89	10,834.02
Orf216_A	114	2		
Orf216_B	114	2	16.54	6,904.31
Orf216_B	114	2		
Orf216_C	114	2	16.36	7,821.78
Orf216_C	114	2		
Orf265a_A	169	2	20.06	405.99
Orf265a_A	169	2		
Orf265a_B	169	2	20.01	420.30
Orf265a_B	169	2		
Orf265a_C	169	2	20.88	229.97
Orf265a_C	169	2		
Orf265b_A	95	2	15.17	21,414.73
Orf265b_A	95	2		
Orf265b_B	95	2	15.83	13,552.92
Orf265b_B	95	2		
Orf265b_C	95	2	16.07	11,475.87
Orf265b_C	95	2		
Orf159b_A	121	2	15.71	11,563.64

Orf159b_A	121	2		
Orf159b_B	121	2	15.49	13,468.56
Orf159b_B	121	2		
Orf159b_C	121	2	15.95	9,791.46
Orf159b_C	121	2		
Orf197_A	141	2	22.77	74.37
Orf197_A	141	2		
Orf197_B	141	2	22.90	67.96
Orf197_B	141	2		
Orf197_C	141	2	24.33	25.22
Orf197_C	141	2		
Orf239_A	144	2	21.14	225.38
Orf239_A	144	2		
Orf239_B	144	2	21.08	234.96
Orf239_B	144	2		
Orf239_C	144	2	21.15	223.83
Orf239_C	144	2		
Orf147_A	96	2	21.10	347.58
Orf147_A	96	2		
Orf147_B	96	2	20.76	439.95
Orf147_B	96	2		
Orf147_C	96	2	21.96	191.50
Orf147_C	96	2		

Appendix C: Ct Values for Polysomal Analysis: Supernatant

Gene/Replicate	Amplicon	Efficiency	Crossover Threshold	Copy#
Cox2_A	108	2	13.19	74,310.79
Cox2_A	108	2	13.54	58,303.06
Cox2_B	108	2	13.13	77,466.45
Cox2_B	108	2	13.43	62,922.31
Cox2_C	108	2	14.77	24,855.61
Cox2_C	108	2	14.75	25,202.58
Bkgrd_A	146	2.02	24.68	14.95
Bkgrd_A	146	2.02	25.19	10.44
Bkgrd_B	146	2.02	22.71	59.73
Bkgrd_B	146	2.02	22.89	52.63
Bkgrd_C	146	2.02	23.56	32.86
Bkgrd_C	146	2.02	23.38	37.29
Orf177_A	99	1.99	21.19	352.15
Orf177_A	99	1.99	21.22	344.96
Orf177_B	99	1.99	20.59	532.16
Orf177_B	99	1.99	20.93	421.14
Orf177_C	99	1.99	21.16	359.50
Orf177_C	99	1.99	21.35	315.44
Orf129b_A	88	1.97	23.26	120.59
Orf129b_A	88	1.97	23.27	119.77
Orf129b_B	88	1.97	21.7	347.27
Orf129b_B	88	1.97	22	283.35
Orf129b_C	88	1.97	22.05	273.90
Orf129b_C	88	1.97	21.94	295.11
Orf151_A	124	1.99	22.23	137.45
Orf151_A	124	1.99	22.18	142.26
Orf151_B	124	1.99	21.09	301.18
Orf151_B	124	1.99	20.83	360.19
Orf151_C	124	1.99	20.92	338.56
Orf151_C	124	1.99	21.04	311.72
Orf175_A	129	2.05	22.59	52.72
Orf175_A	129	2.05	23.07	37.35
Orf175_B	129	2.05	21.74	97.04
Orf175_B	129	2.05	21.91	85.89

Orf175_C	129	2.05	21.56	110.42
Orf175_C	129	2.05	21.62	105.77
Orf134_A	98	2	16.62	7,598.30
Orf134_A	98	2	16.84	6,523.64
Orf134_B	98	2	16.32	9,354.61
Orf134_B	98	2	17.07	5,562.28
Orf134_C	98	2	18.84	1,630.91
Orf134_C	98	2	18.83	1,642.25
Orf306_A	140	2	22.25	107.40
Orf306_A	140	2	23.83	35.92
Orf306_B	140	2	21.56	173.27
Orf306_B	140	2	21.7	157.25
Orf306_C	140	2	22.46	92.85
Orf306_C	140	2	22.72	77.54
Orf138c_A	116	2	22.46	112.06
Orf138c_A	116	2	22.68	96.22
Orf138c_B	116	2	21.33	245.26
Orf138c_B	116	2	21.08	291.67
Orf138c_C	116	2	21.88	167.52
Orf138c_C	116	2	21.22	264.70
Orf144_A	143	2	22.8	71.82
Orf144_A	143	2	22.83	70.34
Orf144_B	143	2	21.09	234.96
Orf144_B	143	2	21.25	210.30
Orf144_C	143	2	22.51	87.81
Orf144_C	143	2	22.52	87.20
Orf118_A	102	2	20.89	378.39
Orf118_A	102	2	20.6	462.64
Orf118_B	102	2	19.41	1,055.53
Orf118_B	102	2	19.69	869.32
Orf118_C	102	2	20.27	581.54
Orf118_C	102	2	20.62	456.27
Orf160_A	80	2	19.75	1,063.23
Orf160_A	80	2	19.81	1,019.92
Orf160_B	80	2	18.7	2,201.46
Orf160_B	80	2	18.58	2,392.40
Orf160_C	80	2	19.88	971.62

Orf160_C	80	2	19.82	1,012.88
Orf101c_A	119	2	14.11	35,643.56
Orf101c_A	119	2	14.08	36,392.51
Orf101c_B	119	2	13.78	44,804.43
Orf101c_B	119	2	13.72	46,707.08
Orf101c_C	119	2	15.4	14,576.47
Orf101c_C	119	2	15.64	12,342.56
Orf111c_A	127	2.04	14.18	24,027.12
Orf111c_A	127	2.04	14.03	26,739.07
Orf111c_B	127	2.04	13.83	30,837.01
Orf111c_B	127	2.04	13.59	36,591.78
Orf111c_C	127	2.04	15.34	10,508.26
Orf111c_C	127	2.04	15.06	12,830.02
Orf125d_A	84	2.02	22.06	163.96
Orf125d_A	84	2.02	22.12	157.18
Orf125d_B	84	2.02	22.08	161.67
Orf125d_B	84	2.02	22.06	163.96
Orf125d_C	84	2.02	21.43	255.33
Orf125d_C	84	2.02	21.81	195.46
OrfB_A	82	1.98	16.42	12,302.83
OrfB_A	82	1.98	16.17	14,593.90
OrfB_B	82	1.98	16.27	13,630.28
OrfB_B	82	1.98	16.18	14,494.55
OrfB_C	82	1.98	19.17	1,880.08
OrfB_C	82	1.98	18.65	2,681.89
Orf115_A	100	1.96	20.21	930.18
Orf115_A	100	1.96	19.82	1,209.33
Orf115_B	100	1.96	20.69	673.42
Orf115_B	100	1.96	20.46	786.14
Orf115_C	100	1.96	20.02	1,057.05
Orf115_C	100	1.96	20.76	642.43
Orf166_A	81	2.04	20.58	393.00
Orf166_A	81	2.04	20.58	393.00
Orf166_B	81	2.04	19.36	937.86
Orf166_B	81	2.04	19.58	801.71
Orf166_C	81	2.04	20.97	297.60
Orf166_C	81	2.04	20.81	333.56

Orf133_A	118	2	15.17	17,240.67
Orf133_A	118	2	15.13	17,725.37
Orf133_B	118	2	14.06	37,213.24
Orf133_B	118	2	12.99	78,126.72
Orf133_C	118	2	13.15	69,925.38
Orf133_C	118	2	13.07	73,912.38
Orf25_A	83	2	16.47	9,954.49
Orf25_A	83	2	16.43	10,234.35
Orf25_B	83	2	15.55	18,835.05
Orf25_B	83	2	16.01	13,692.82
Orf25_C	83	2	15.11	25,551.71
Orf25_C	83	2	15.18	24,341.53
Orf222_A	84	2	18.29	2,785.76
Orf222_A	84	2	18.07	3,244.67
Orf222_B	84	2	16.41	10,253.68
Orf222_B	84	2	16.75	8,100.83
Orf222_C	84	2	15.8	15,649.77
Orf222_C	84	2	15.73	16,427.83
Orf216_A	114	2	15.8	11,531.41
Orf216_A	114	2	16.33	7,986.13
Orf216_B	114	2	14.79	23,223.24
Orf216_B	114	2	14.78	23,384.77
Orf216_C	114	2	N/A	
Orf216_C	114	2	14.49	28,591.16
Orf265a_A	169	2	20.18	373.58
Orf265a_A	169	2	20.34	334.37
Orf265a_B	169	2	17.09	3,181.06
Orf265a_B	169	2	16.63	4,375.68
Orf265a_C	169	2	16.18	5,977.36
Orf265a_C	169	2	16.25	5,694.26
Orf265b_A	95	2	15.15	21,713.67
Orf265b_A	95	2	15.41	18,132.82
Orf265b_B	95	2	13.97	49,198.13
Orf265b_B	95	2	13.88	52,365.03
Orf265b_C	95	2	13.05	93,088.54
Orf265b_C	95	2	12.92	101,866.20
Orf159b_A	121	2	15.9	10,136.76

Orf159b_A	121	2	15.95	9,791.46
Orf159b_B	121	2	14.81	21,578.52
Orf159b_B	121	2	14.42	28,276.35
Orf159b_C	121	2	13.53	52,401.06
Orf159b_C	121	2	13.92	39,988.80
Orf197_A	141	2	20.52	353.76
Orf197_A	141	2	20.82	287.34
Orf197_B	141	2	17.2	3,532.86
Orf197_B	141	2	16.92	4,289.58
Orf197_C	141	2	17.02	4,002.32
Orf197_C	141	2	17.08	3,839.28
Orf239_A	144	2	20.61	325.44
Orf239_A	144	2	20.82	281.35
Orf239_B	144	2	16.81	4,532.99
Orf239_B	144	2	17.04	3,864.98
Orf239_C	144	2	16.91	4,229.43
Orf239_C	144	2	16.61	5,207.04
Orf147_A	96	2	21.03	364.86
Orf147_A	96	2	20.82	422.03
Orf147_B	96	2	17.09	5,599.99
Orf147_B	96	2	17.14	5,409.24
Orf147_C	96	2	17.2	5,188.89
Orf147_C	96	2	16.94	6,213.58
Cox_2 +EDTA	108	2	13.53	58,708.59
Cox_2 +EDTA	108	2	13.44	62,487.68

Appendix D: Ct Values for Polysomal Analysis: Pellet

Gene/Replicate	Amplicon	Efficiency	Crossover Threshold	Copy#
Cox2_A	108	2	15.97	10,819.03
Cox2_A	108	2	15.91	11,278.47
Cox2_B	108	2	14.58	28,354.38
Cox2_B	108	2	14.39	32,345.65
Cox2_C	108	2	15.24	17,944.87
Cox2_C	108	2	15.33	16,859.62
Bkgrd_A	146	2.02	23.28	40.01
Bkgrd_A	146	2.02	23.76	28.55
Bkgrd_B	146	2.02	19.49	574.66
Bkgrd_B	146	2.02	19.62	524.46
Bkgrd_C	146	2.02	20.59	265.17
Bkgrd_C	146	2.02	20.5	282.49
Orf177_A	99	1.99	20.46	581.96
Orf177_A	99	1.99	20.74	479.97
Orf177_B	99	1.99	18.97	1,622.49
Orf177_B	99	1.99	18.64	2,036.13
Orf177_C	99	1.99	19.38	1,223.64
Orf177_C	99	1.99	19.36	1,240.60
Orf129b_A	88	1.97	23.28	118.96
Orf129b_A	88	1.97	23.04	139.98
Orf129b_B	88	1.97	20.29	903.36
Orf129b_B	88	1.97	19.99	1,107.13
Orf129b_C	88	1.97	20.96	573.54
Orf129b_C	88	1.97	20.8	639.27
Orf151_A	124	1.99	23.25	68.12
Orf151_A	124	1.99	23.14	73.48
Orf151_B	124	1.99	20.02	628.93
Orf151_B	124	1.99	20	637.64
Orf151_C	124	1.99	21.01	318.23
Orf151_C	124	1.99	21.11	297.06
Orf175_A	129	2.05	33.52	0.02
Orf175_A	129	2.05	33.4	0.02
Orf175_B	129	2.05	26.9	2.39
Orf175_B	129	2.05	27.24	1.87

Orf175_C	129	2.05	28.63	0.69
Orf175_C	129	2.05	28.27	0.89
Orf134_A	98	2	25.5	16.13
Orf134_A	98	2	25.74	13.66
Orf134_B	98	2	22.76	107.74
Orf134_B	98	2	22.95	94.45
Orf134_C	98	2	23.51	64.06
Orf134_C	98	2	24.38	35.05
Orf306_A	140	2	31.16	0.22
Orf306_A	140	2	31.7	0.15
Orf306_B	140	2	25.79	9.23
Orf306_B	140	2	27.11	3.70
Orf306_C	140	2	27.41	3.00
Orf306_C	140	2	27.49	2.84
Orf138c_A	116	2	21.73	185.88
Orf138c_A	116	2	21.87	168.69
Orf138c_B	116	2	18.48	1,768.35
Orf138c_B	116	2	18.67	1,550.15
Orf138c_C	116	2	19.78	718.17
Orf138c_C	116	2	19.93	647.26
Orf144_A	143	2	30.25	0.41
Orf144_A	143	2	29.72	0.59
Orf144_B	143	2	24.43	23.20
Orf144_B	143	2	24.18	27.59
Orf144_C	143	2	25.99	7.87
Orf144_C	143	2	27.31	3.15
Orf118_A	102	2	33.71	0.05
Orf118_A	102	2	33.77	0.05
Orf118_B	102	2	28.16	2.45
Orf118_B	102	2	28.76	1.62
Orf118_C	102	2	30.51	0.48
Orf118_C	102	2	30.92	0.36
Orf160_A	80	2	18.92	1,890.10
Orf160_A	80	2	19.11	1,656.87
Orf160_B	80	2	17.17	6,357.50
Orf160_B	80	2	17.01	7,103.15
Orf160_C	80	2	18.04	3,478.49

Orf160_C	80	2	18.01	3,551.58
Orf101c_A	119	2	13.79	44,494.94
Orf101c_A	119	2	13.96	39,549.01
Orf101c_B	119	2	11.89	166,061.00
Orf101c_B	119	2	12.04	149,662.55
Orf101c_C	119	2	13.39	58,711.43
Orf101c_C	119	2	13.28	63,363.03
Orf111c_A	127	2.04	17.18	2,830.16
Orf111c_A	127	2.04	16.93	3,382.35
Orf111c_B	127	2.04	14.15	24,546.56
Orf111c_B	127	2.04	14.16	24,372.18
Orf111c_C	127	2.04	16	6,564.10
Orf111c_C	127	2.04	15.21	11,528.76
Orf125d_A	84	2.02	25.07	19.75
Orf125d_A	84	2.02	25.22	17.78
Orf125d_B	84	2.02	20.99	347.90
Orf125d_B	84	2.02	21.4	260.77
Orf125d_C	84	2.02	23.46	61.27
Orf125d_C	84	2.02	22.71	103.81
OrfB_A	82	1.98	16.8	9,490.13
OrfB_A	82	1.98	17.14	7,523.26
OrfB_B	82	1.98	15.2	28,309.78
OrfB_B	82	1.98	15.16	29,093.98
OrfB_C	82	1.98	16.09	15,413.61
OrfB_C	82	1.98	15.89	17,670.04
Orf115_A	100	1.96	22.55	192.61
Orf115_A	100	1.96	22.5	199.20
Orf115_B	100	1.96	19.85	1,185.16
Orf115_B	100	1.96	19.64	1,365.06
Orf115_C	100	1.96	21.01	542.95
Orf115_C	100	1.96	20.8	625.37
Orf166_A	81	2.04	19.28	992.90
Orf166_A	81	2.04	19.5	848.77
Orf166_B	81	2.04	17.55	3,408.53
Orf166_B	81	2.04	17.54	3,432.92
Orf166_C	81	2.04	19.21	1,043.71
Orf166_C	81	2.04	18.5	1,731.48

Orf133_A	118	2	14.83	21,822.50
Orf133_A	118	2	14.81	22,127.13
Orf133_B	118	2	13.69	48,092.64
Orf133_B	118	2	14.13	35,450.75
Orf133_C	118	2	14.7	23,880.22
Orf133_C	118	2	14.61	25,417.39
Orf25_A	83	2	16.85	7,649.41
Orf25_A	83	2	17.37	5,334.48
Orf25_B	83	2	16.86	7,596.57
Orf25_B	83	2	16.52	9,615.41
Orf25_C	83	2	17.97	3,519.44
Orf25_C	83	2	18.33	2,742.23
Orf222_A	84	2	17.98	3,453.53
Orf222_A	84	2	17.98	3,453.53
Orf222_B	84	2	17.55	4,652.71
Orf222_B	84	2	17.63	4,401.73
Orf222_C	84	2	18.51	2,391.76
Orf222_C	84	2	18.75	2,025.21
Orf216_A	114	2	14.48	28,790.03
Orf216_A	114	2	14.55	27,426.47
Orf216_B	114	2	14.34	31,723.88
Orf216_B	114	2	14.37	31,071.01
Orf216_C	114	2	16.18	8,861.17
Orf216_C	114	2	16.12	9,237.47
Orf265a_A	169	2	19.28	697.13
Orf265a_A	169	2	19.25	711.78
Orf265a_B	169	2	18.99	852.34
Orf265a_B	169	2	18.82	958.94
Orf265a_C	169	2	18.97	864.24
Orf265a_C	169	2	18.94	882.40
Orf265b_A	95	2	14.47	34,788.33
Orf265b_A	95	2	14.31	38,868.56
Orf265b_B	95	2	14.35	37,805.69
Orf265b_B	95	2	14.14	43,729.40
Orf265b_C	95	2	15.81	13,742.11
Orf265b_C	95	2	15.87	13,182.31
Orf159b_A	121	2	16.33	7,524.13

Orf159b_A	121	2	16.11	8,763.60
Orf159b_B	121	2	15.68	11,806.62
Orf159b_B	121	2	15.24	16,016.91
Orf159b_C	121	2	16.88	5,139.13
Orf159b_C	121	2	16.81	5,394.63
Orf197_A	141	2	22.12	116.70
Orf197_A	141	2	22.39	96.78
Orf197_B	141	2	20.14	460.36
Orf197_B	141	2	20.42	379.15
Orf197_C	141	2	19.57	683.42
Orf197_C	141	2	19.51	712.44
Orf239_A	144	2	24.39	23.69
Orf239_A	144	2	23.64	39.84
Orf239_B	144	2	20.26	414.79
Orf239_B	144	2	20.36	387.02
Orf239_C	144	2	19.64	637.48
Orf239_C	144	2	19.37	768.68
Orf147_A	96	2	21.85	206.67
Orf147_A	96	2	21.59	247.49
Orf147_B	96	2	19.66	943.06
Orf147_B	96	2	19.89	804.09
Orf147_C	96	2	20.08	704.87
Orf147_C	96	2	19.74	892.19
Cox_2 +EDTA	108	2	Undetectable	Undetectable
Cox_2 +EDTA	108	2	Undetectable	Undetectable

Appendix E: Statistical Hypothesis Testing of Transcription in Leaves

Background Vs.	Normality	Equal Variance	Pair-Wise T-Test	Mann-Whitney rank sum
Cox_2	Fail	Fail	n/a	0.002
Orf177	Fail	Fail	n/a	0.002
Orf129b	Pass	Fail	n/a	0.002
Orf151	Fail	Fail	n/a	0.002
Orf175	Pass	Fail	n/a	0.002
Orf134	Pass	Fail	n/a	0.002
Orf306	Pass	Pass	0.006	0.002
Orf138c	Pass	Pass	0.001	0.002
Orf144	Fail	Fail	n/a	0.002
Orf118	Fail	Fail	n/a	0.002
Orf160	Fail	Fail	n/a	0.002
Orf101c	Fail	Fail	n/a	0.002
Orf111c	Fail	Fail	n/a	0.002
Orf125d	Fail	Fail	n/a	0.002
OrfB	Fail	Fail	n/a	0.002
Orf115	Fail	Fail	n/a	0.002
Orf166	Fail	Fail	n/a	0.002
Orf133	Pass	Pass	0.051	0.1
Orf25	Pass	Pass	0.001	0.1
Orf222	Pass	Pass	0.054	0.1
Orf216	Pass	Pass	0.029	0.1
Orf265a	Pass	Pass	0.046	0.1
Orf265b	Pass	Pass	0.028	0.1
Orf159b	Pass	Pass	0.005	0.1
Orf197	Pass	Pass	0.052	0.1
Orf239	Pass	Pass	0.081	0.1
Orf147	Pass	Pass	0.007	0.1

Appendix F: Statistical Hypothesis Testing of Transcription in Roots

BKGRND VS:	Normality	Equal Variance	Pair-Wise t-test	Mann-Whitney rank sum
Cox_2	Pass	Fail	n/a	0.002
Orf177	Pass	Fail	n/a	0.002
Orf129b	Fail	Fail	n/a	0.002
Orf151	Pass	Fail	n/a	0.002
Orf175	Pass	Fail	n/a	0.002
Orf134	Pass	Pass	0.004	0.015
Orf306	Pass	Pass	0.001	0.002
Orf138c	Pass	Pass	0.002	0.002
Orf144	Pass	Fail	n/a	0.002
Orf118	Pass	Fail	n/a	0.002
Orf160	Pass	Pass	0.001	0.002
Orf101c	Fail	Fail	n/a	0.002
Orf111c	Fail	Fail	n/a	0.002
Orf125d	Fail	Fail	n/a	0.002
OrfB	Fail	Fail	n/a	0.002
Orf115	Fail	Fail	n/a	0.002
Orf166	Fail	Fail	n/a	0.002
Orf133	Pass	Pass	0.001	0.1
Orf25	Pass	Fail	n/a	0.1
Orf222	Pass	Fail	n/a	0.1
Orf216	Pass	Pass	0.005	0.1
Orf265a	Pass	Fail	n/a	0.1
Orf265b	Pass	Pass	0.008	0.1
Orf159b	Pass	Pass	0.001	0.1
Orf197	Pass	Pass	0.077	0.1
Orf239	Pass	Fail	0.001	0.1
Orf147	Pass	Pass	0.002	0.1

Appendix G: Statistical Hypothesis Testing of Transcript Abundance in the Polysomal Fraction

BKGRND VS:	Normality	Equal variance	Pair Wise t-test	Mann-Whitney rank sum
Cox_2	Fail	Fail	n/a	0.002
Orf177	Pass	Pass	0.001	0.002
Orf129b	Pass	Pass	0.022	0.004
Orf151	Pass	Fail	n/a	0.002
Orf175	Fail	Fail	n/a	0.132
Orf134	Pass	Fail	n/a	0.002
Orf306	Pass	Pass	0.023	0.026
Orf138c	Pass	Pass	0.025	0.004
Orf144	Pass	Pass	0.008	0.026
Orf118	Pass	Fail	n/a	0.002
Orf160	Pass	Fail	n/a	0.002
Orf101c	Fail	Fail	n/a	0.002
Orf111c	Pass	Fail	n/a	0.002
Orf125d	Pass	Pass	0.032	0.002
OrfB	Pass	Fail	n/a	0.002
Orf115	Fail	Fail	n/a	0.002
Orf166	Pass	Pass	0.001	0.002
Orf133	Fail	Fail	n/a	0.002
Orf25	Fail	Fail	n/a	0.002
Orf222	Fail	Fail	n/a	0.002
Orf216	Fail	Fail	n/a	0.002
Orf265b	Fail	Fail	n/a	0.002
Orf159b	Fail	Fail	n/a	0.002
Orf197	Fail	Fail	n/a	0.002
Orf239	Fail	Fail	n/a	0.002
Orf147	Fail	Fail	n/a	0.002

Appendix H: Statistical Hypothesis Testing of Transcript Abundance in the Pellet

BKGRND VS:	Normality	Equal variance	Pair-Wise t-test	Mann-Whitney rank sum
Cox_2	Pass	Pass	0.001	0.002
Orf177	Pass	Pass	0.001	0.002
Orf129b	Pass	Pass	0.019	0.132
Orf151	Fail	Fail	n/a	0.310
Orf175	Pass	Pass	0.001	0.002
Orf134	Pass	Pass	0.003	0.002
Orf306	Pass	Pass	0.001	0.002
Orf138c	Pass	Pass	0.008	0.009
Orf144	Pass	Pass	0.001	0.002
Orf118	Pass	Pass	0.001	0.002
Orf160	Fail	Fail	n/a	0.002
Orf101c	Fail	Fail	n/a	0.002
Orf111c	Pass	Fail	n/a	0.002
Orf125d	Pass	Pass	0.041	0.065
Orf115	Pass	Pass	0.003	0.009
Orf166	Pass	Pass	0.001	0.002
Orf133	Pass	Fail	n/a	0.002
Orf25	Fail	Fail	n/a	0.002
Orf222	Pass	Pass	0.001	0.002
Orf216	Pass	Pass	0.001	0.002
Orf265b	Pass	Pass	0.001	0.002
Orf159b	Fail	Fail	n/a	0.002
Orf197	Pass	Pass	0.22	0.589
Orf239	Pass	Pass	0.424	0.937
Orf147	Pass	Pass	0.009	0.004