uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

u Ottawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Jennifer Li-Pook-Than
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biology)
GRADE / DEGREE

Department of Biology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Group II Intron Splicing in Wheat Mitochondria during Embryo-to-Seedling Development

TITRE DE LA THÈSE / TITLE OF THESIS

Linda Bonen
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

J. Arnason

M. Paulin-Levasseur

D. Johnson

S. Zimmerly

B. Miki

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

GROUP II INTRON SPLICING IN WHEAT MITOCHONDRIA DURING EMBRYO-TO-SEEDLING DEVELOPMENT

Jennifer Li-Pook-Than

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
in partial fulfillment of the requirements for the
Ph.D. degree in the

Ottawa-Carleton Institute of Biology

Thèse soumise à la
Faculté des études supérieures et postdoctorales
Université d'Ottawa
en vue de l'obtention du doctorat

L'institut de biologie d'Ottawa-Carleton

Your file Votre référence
ISBN: 978-0-494-34135-3
Our file Notre référence
ISBN: 978-0-494-34135-3

GROUP II INTRON SPLICING IN WHEAT MITOCHONDRIA DURING EMBRYO-TO-SEEDLING DEVELOPMENT

ABSTRACT:

Plant mitochondrial transcripts undergo complex RNA processing, including intron removal and C-to-U editing. My focus has been on the excision of the highly-structured group II introns, mainly located in NADH dehydrogenase (*nad*) genes encoding for complex I of the respiratory chain. Their splicing mechanism is expected to be biochemically similar to that of nuclear spliceosomal introns whereby lariats are generated via a two-step transesterification pathway utilizing an unpaired adenosine (near the 3' end of the intron) as the first attacking nucleophile. However, certain plant mitochondrial introns deviate from classical group II features, one example being *nad1* intron 2.

To better understand the splicing mechanisms of aberrant plant mitochondrial introns, I examined their excised physical forms in germinating wheat embryos by employing strategies such as RNase H analysis, RT-PCR and cRT-PCR (RNA circularized with RNA ligase) to distinguish between lariat, linear and fully-circularized RNA species. For *nad1* intron 2, which lacks the bulged adenosine, I found a mixed population of precisely full-length linear molecules, as well as circular ones that included blocks of non-encoded nucleotides. Interestingly, in the case of the *cox2* intron, which contains a bulged adenosine, lariats were not predominant. Instead, I observed full-length linear introns that possessed non-encoded 3' terminal adenosines, as well as heterogeneous circular introns that lacked 3' nucleotide stretches. These mixed populations of excised introns suggest the use of multiple splicing mechanisms *in vivo* including hydrolytic splicing and potentially the use of alternative nucleophiles.

Based on northern blot analysis, my examination of high molecular weight precursors which contained one or multiple group II introns (*nad7, nad1, cox2, rps3*) suggests that transcription is outpacing the RNA processing events during wheat germination (6 hr to 2 day stages of development). This contrasted with abundant mRNAs, for both the respiratory chain genes *(nad7, cox1, cox2, atp6)* and ribosomal protein genes (*rps2, rps3, rps7),* in respective dormant embryos and 6 day seedlings. When *cox2* precursors were examined, they were found to be incompletely edited in germinating embryos when compared to seedlings at both splice

junctions and at internal exon sites. This supports the model that editing and splicing machinery are rate-limiting during the early stages of seedling development.

Although fully-spliced mRNAs showed complete editing at expected sites at all developmental stages, including in dormant seeds, exon editing positions up to 6 nt away from unspliced introns were unedited in *nad4* and *nad7,* while *nad2 and nad5* counterparts were fully edited. This suggests an inter-relationship between splicing and editing at junction sites that would include 1) an unspliced intron sterically hindering editing machinery, or 2) an editing recognition motif generated only when the intron is excised and consecutive exons are joined together. It will be of future interest to elucidate the plant mitochondrial nuclear-encoded proteins and/or small RNAs involved in group II intron splicing and editing.

ÉPISSAGE DES INTRONS NOMMÉS GROUPE II DANS LES MITOCHONDRIES DE BLÉ DURANT LA GERMINATION

RÉSUMÉ:

Avant la génération d'ARN messagers (ARNm), les transcrits de mitochondries de plantes subissent des changements d'ARN incluant l'épissage des séquences interposées (introns) et la conversion de cytosine en uridine selon l'édition (C-à-U). Ce travail est concentré sur l'épissage d'une catégorie d'introns dont la structure secondaire est distincte et complexe, appelés les introns de groupe II. La majorité de ces introns se retrouve dans les gènes qui codent pour la sous-unité I du complexe de la NADH déshydrogénase (*nad*) des plantes. Ce mécanisme de retrait a des similarités avec le système 'splicéosomal' utilisé par le noyau ou les 'lassos' sont crés avec une adénosine bombée nucléophilique (située dans la région 3' des introns), le premier de deux étapes de transestérification. Cependant, dans le cas des introns des mitochondries des plantes, certaines structures secondaires peuvent être anormales comparativement à la structure typique d'intron de groupe II: tel que l'intron 2 de *nad1*.

Le but de ce travail est de mieux comprendre le mécanisme de retrait des introns aberrants dans les mitochondries des plantes pendant la germination des graines. J'ai examiné les formes physiques des introns, après l'épissage, en utilisant des méthodes comme l'analyse RNase H, TR-RCP et cTR-RCP (circularisation avec la ligase d'ARN) pour distinguer parmi les espèces d'ARN 'lassos', circulaires et linéarisées dans les mitochondries du blé. Pour l'intron 2 de *nad1,* dont l'adénosine bombée est absente, j'ai trouvé un ensemble mixte de molécules linéaires (de longueur totale) et aussi d' introns circulaires contenant des insertions qui ne correspondaient pas à la sequence codante de l'intron. Dans le cas de l'intron de *cox2* qui contient l'adénosine bombée, au lieu d'une majorité de 'lassos' tel qu'attendu, j'ai observé des introns ayant des adénosines supplémentaires à l'extrémité 3' de l'intron, et aussi des molécules hétérogènes de cercles dépourvus des régions 3' de l'intron de *cox2*. Cet assortiment d'espèces d'introns épissés nous démontre l'utilisation des multiples mécanismes de retrait des introns, incluant l'épissage hydrolytique et la possibilité d'utilisation de nucléophiles alternatifs.

Mes analyses de northerns sur des précurseurs de grand poids moléculaire de gènes contenant un ou plusieurs introns de group II (*nad7, nad1, cox2, rps3*) nous suggèrent que la transcription est plus efficace que le mécanisme de modification d'ARN durant la germination

du blé. Ceci était contraire aux résultats d'ARNm pour les gènes de chaînes respiratoires *(nad7, cox1, cox2, atp6)* et les protèines ribosomales (*rps2, rps3, rps7*) qui étaient abondants dans les différentes étapes de développement du blé incluant l'embryon et la jeune pousse. L'étude des précurseurs de *cox2* a révélé qu'ils étaient moins édités dans les étapes préliminaires de développement mais plus complètement édités dans les étapes de pousse. Ceci nous indique que les mécaniques utilisées pour l'édition, comme l'épissage, sont moins disponibles dans les mitochondries avant le développement préliminaire de la pousse.

Les ARNm étaient complètement édités dans toutes les étapes de développement du blé étudiées, incluant dans les embryons. Cependant, les transcrits ayant un site d'édition moins de 6 nucléotides avant ou après un intron, n'étaient pas édités pour *nad4* et *nad7,* tandis qu'ils étaient complètement édités pour *nad2* et *nad5*. Quelques modèles pour expliquer la relation entre l'édition et l'épissage sont 1) la présence de la structure de l'intron à côté du site d'édition empêche l'accès à la mécanique d'édition, ou 2) le motif de séquence de reconnaissance par la mécanique d'édition apparait seulement après l'épissage quand les exons sont réunis. Des projets ultérieurs nous permettraient de mieux connaître les protéines (codées dans le noyau) ou les 'petits' ARNs situés dans les mitochondries de plantes qui participent aux mécanismes d'épissage des introns de groupe II et d'édition C-à-U.

Acknowledgments

First and foremost I would like to thank my supervisor Dr. Linda Bonen for her commitment to research and all her helpful advice. She is a devoted teacher and researcher and has been a source of inspiration to me. I'd also like to acknowledge my committee members, Drs. Thor Arnason, Douglas Johnson and Brian Miki, who have provided excellent feedback throughout the years.

I am grateful to my current labmates Sophie Calixte and Thomas Hazle whose assistance and friendship have been central during my studies. I also greatly appreciate the contribution of my past labmates N. Niknejad and J. Crosthwait, 4th year students (O. Antoniuk, M. Roy), former graduate students C. Carrillo, Y. Chapdelaine, D. Zhuo and S. Subramanian, as well as my University of Ottawa colleagues (CW, NM, JB, AT, SW, CA, RL). I would also like to acknowledge the University of Ottawa Department of Biology support staff for providing an enriching environment for research.

I am thankful to all my friends and extended family in Montreal. A special thanks to P. Sawbridge for all his support. This work is dedicated to my parents James and Irene, and my brother, Andrew for their love and constant encouragement and without whom this work could not have been accomplished.

Table of Contents

List of Appendices

List of Figures

List of Tables

List of Abbreviations

ATP	adenosine triphosphate
atp	adenosine triphosphate synthase subunit genes
bp	base pairs
BSA	bovine serum albumin
ccm	cytochrome c biogenesis subunit genes
cDNA	complementary DNA
CDS	coding sequence
cob	cytochrome bc1 oxidoreductase subunits genes
cox	cytochrome oxidase subunits genes
d	day(s)
dI- dVI	domains 1-6
ddH_2O	double distilled water
dH_2O	single distilled water
ddNTP	dideoxyribonucleotide triphosphate
dNTPs	deoxyribonucleotide triphosphates (G, A, T, C)
DTT	dithiolthreitol
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
hr	hour(s)
IEP	intron encoded protein
kb	kilobases
mat-R	plant mitochondrial intron-encoded maturase (reverse transcriptase)
mRNA	messenger RNA
mt	mitochondrial
mtt	membrane targeting and translocating genes
µL	microlitre
µg	microgram
mL	milliliter
mg	milligram
$MgCl_2$	magnesium chloride
MMLV	Moloney murine leukemia virus
NAD	nicotinamide adenine dinucleotide
nad	NADH dehydrogenase subunits genes
nt(s)	nucleotide(s) R = Purine A = adenosine G = guanosine Y = Pyrimidine T = thymidine U = uridine C = cytidine
-OH	hydroxyl group
ORF	open reading frame

PCR	polymerase chain reaction
RNase H	ribonuclease
rpl	ribosomal protein large subunit
rps	ribosomal protein small subunit
rRNA	ribosomal RNA
RT, cRT	reverse transcriptase, circularized reverse transcriptase (with RNA ligase)
sdh	succinate dehydrogenase subunit genes
SDS	sodium dodecyl sulphate
snRNA	small nuclear RNA
SSII	Superscript II
TBE	Tris-borate-EDTA buffer
Tris	tris(hydroxymethyl) aminomethane
tRNA	transfer RNA
U	units
V	volts

INTRODUCTION

CHAPTER 1: LITERATURE REVIEW

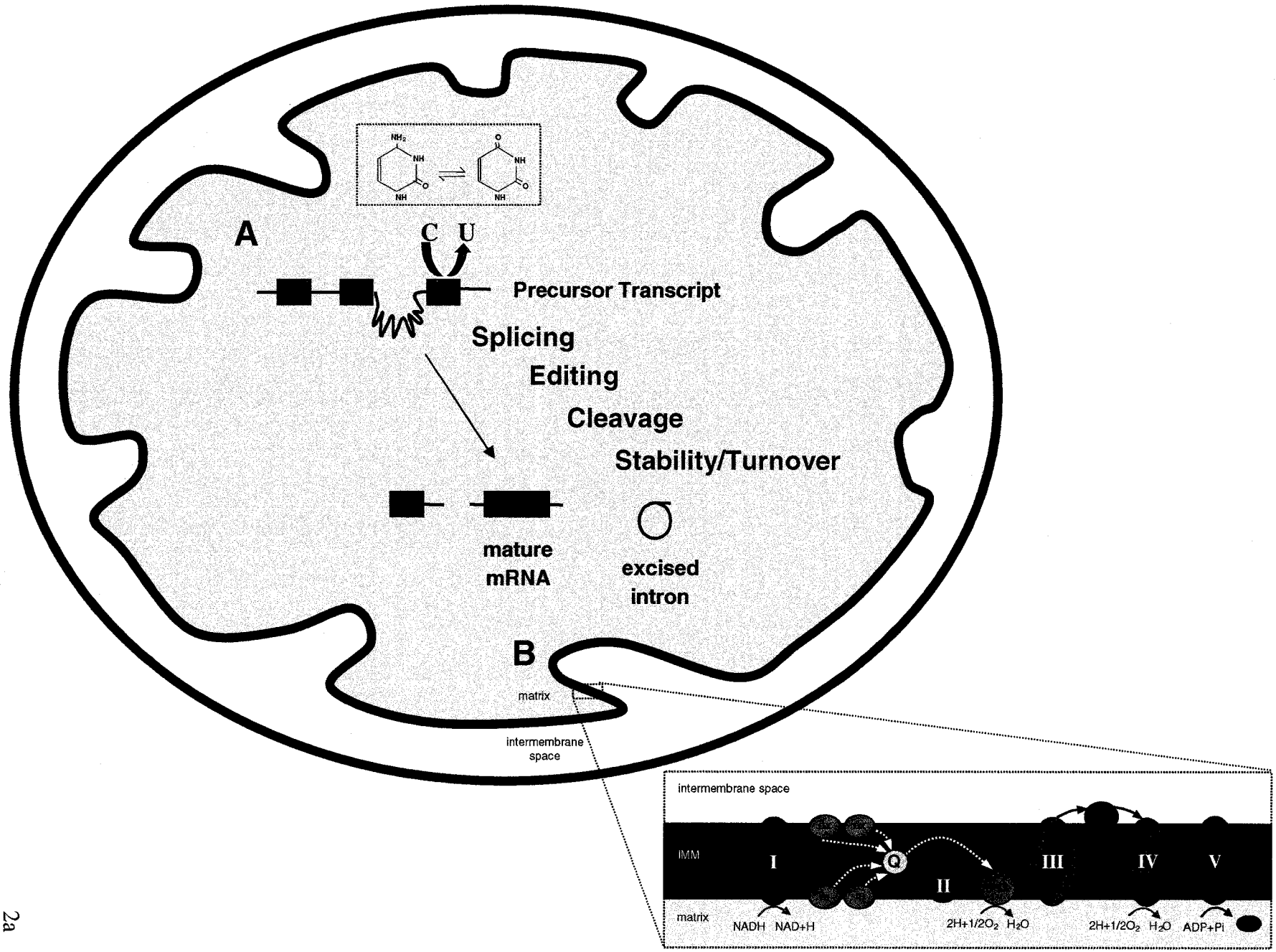
In our present era where complete genomes are being sequenced in groundbreaking numbers, the secrets to the complexity of life are not just found within coding sequences, but are being increasingly associated with non-coding RNA, one example being intron splicing. The discovery of introns by Sharp and Roberts in 1977 was pivotal in changing fundamental views about gene organization and expression in eukaryotes. The precise excision of these intragenic sequences is essential for the production of mature messenger RNA and a functional protein. Introns are ubiquitous in nature, present in the nuclear genes of almost all eukaryotes, organellar genes of plants, fungi and protists, as well as a few bacteria and archaebacteria genes. Of particular interest are the group II type introns which have ribozymic activity *in vitro*, are a member of the mobile genetic intron family and are thought to have evolutionary links to nuclear spliceosomal introns (Cech, 1993; Michel and Ferat, 1995). My research work is focused on the RNA processing events in plant mitochondria, such as group II intron removal and cytidine-to-uridine (C-to-U) editing (Figure 1.1), with particular attention on group II intron splicing during early wheat germination, a period of rapid mitochondrial biogenesis.

1.1 Group II introns

Group II introns are typically found in the mitochondria (as well as chloroplasts) of plants, in the mitochondria of fungi and protists, as well as in a few bacterial and archaebacterial genes (Michel and Ferat, 1995; Zimmerly et al., 2001; Bonen and Vogel, 2001). The distinct secondary structure of group II introns comprises a complex network of stem-loops arranged in six domains (dI-dVI). A key feature involved in conventional splicing is an unpaired adenosine (A) located 6-7 nt upstream of the 3' intron terminus in dVI, essential for lariat formation (Figure 1.2A). Splicing occurs via two transesterification steps, where in the first step the 2'-OH of this branchpoint adenosine attacks the 5' end of the intron resulting in a lariat intermediate (Figure 1.3). In the second transesterification step, the free 3'OH of the upstream exon acts as the second nucleophile and attacks the downstream exon, joining the exon pieces and releasing the lariat intron (Figure 1.3A). This autocatalytic reaction is dependent on high salt (Mg^{2+}) conditions for proper tertiary structure folding (reviewed in Lehmann and Schmidt, 2003).

Figure 1.1 Schematic of intracellular events within plant mitochondria.

[A] Post-transcriptional RNA processing, namely splicing, editing, cleavage and stability and turnover events occur in plant mitochondria to generate a mature mRNAs and excised introns. Of the mitochondrially-encoded genes undergoing these processes, the majority encode for proteins involved in **[B]** the electron transport chain and ATP production (red arrows). The oxidative phosphorylative pathway of respiration contains five complexes (I-V); complex I - NADH dehydrogenase (*nad*), complex II - succinate:CoQ oxidoreductase (*sdh*), complex III - Cytochrome bc1 oxidoreductase (*cob*), complex IV - cytochrome oxidase (*cox*), and complex V - ATP synthase (*atp*). Plants can also respire via alternative respiratory pathways (dotted arrows); the rotenone-insensitive pathway which uses external and internal NADH dehydrogenase proteins ($NADH_{ex/in}$) found on the outer and inner surface of the inner mitochondrial membrane (IMM), respectively, and can oxidize cytosolic NAD(P)H (Elhafez et al., 2006), and the cyanide-insensitive pathway, which bypasses complex IV (*cox*) via an alternate oxidase (*aox*) complex (Moller, 2001). Figure 1.1B was adapted from Moller, 2001.


A
C U
Precursor Transcript
Splicing
Editing
Cleavage
Stability/Turnover
mature
mRNA
excised
intron
B
matrix
intermembrane
space
intermembrane space
IMM
I
Q
II
III
IV
V
matrix
NADH NAD+H
2H+1/2O2 H2O
2H+1/2O2 H2O ADP+Pi

Bacterial group II introns are usually contained within transposons or plasmids, and based on the endosymbiont theory, their presence in α-proteobacteria are a speculated source for their proliferation in present-day organelles (Martinez-Abarca and Toro, 2000). There are also examples of group II introns in bacterial genes, such as those found in a bacterial ribonucleotide reductase gene (Liu et al., 2003). Plant mitochondrial group II introns in flowering plants are located mostly in protein coding genes, in contrast to chloroplast counterparts which are commonly found in tRNA genes. Based on the complete mitochondrial genomes of the flowering plants, *Beta vulgaris* (Kubo et al., 2000), *Arabidopsis thaliana* (Unseld et al., 1997), *Brassica napus* (Handa, 2003), *Nicotiana tabacum* (Sugiyama et al., 2005), *Oryza sativa* (Notsu et al., 2002), *Zea mays* (Clifton et al., 2004) and *Triticum aestivum* (Ogihara et al., 2005), a compiled list of cis- and trans-splicing introns is shown in Table 1.1. Of the 26 known group II introns characterized to date in the mitochondria of seed plants, 19 are positioned in the *nad* genes encoding proteins of complex I of the respiratory chain pathway. Of the remaining, one is located in the respiratory chain component *cox2* gene of complex IV, one in the cytochrome c biogenesis gene (*ccmFc*), and three are located in the ribosomal protein small and large subunit genes (*rps3, rps10 and rpl2*) in various plant species. A second group II intron was found in *cox2* of diverse seed plants including such genera as *Peperomia, Pinus* and *Gingko* (Qiu et al., 1998), while a novel group II intron was also found in *rps3* in several gymnosperms (Regina et al., 2005).

The majority of plant mitochondrial introns are cis-splicing, however a subset of group II introns are disrupted and require trans-splicing (Table 1.1, green blocks vs. yellow blocks). The position of the group II intron fracture occurs typically within domain IV (Qiu and Palmer, 2004). It is through the highly recombinogenic nature of plant mitochondrial genomes (Bonen and Brown, 1993) that these 'genes-in-pieces' are located at different positions throughout the genome and require coordinated transcription for the assembly of the distal exons and subsequent mRNA generation. Typically the configuration is bipartite (2 intron halves), but an interesting example observed in the mitochondria of *Oenothera* is of a tripartite *nad5* intron 3 with the breakage in dI (Knoop et al., 1997). A group II intron found in the *psaA* gene in chloroplast *Chlamydomonas reinhardtii* is also split into three pieces, with its respective precursors from distantly located exon 1 adjoined to the 5' intron, exon 2 with 3' intron and a separate non-coding intron RNA section *tscA* (Merendino et al., 2006). Trans-introns are found

Table 1.1 Summary of group II introns found in fully sequenced flowering plant mitochondria.

To date, seven flowering plant mitochondrial genomes have been completed. These include the monocots *Triticum aestivum (Ta,* Ogihara et al., 2005*)*, *Zea mays (Zm,* Clifton et al., 2004*)* and *Oryza sativa (Os,* Notsu et al., 2002*),* as well as eudicots *Beta vulgaris (Bv,* Kubo et al., 2000*),* *Nicotiana tabacum (Nt,* Sugiyama et al., 2005*)*, *Arabidopsis thaliana (At,* Unseld et al., 1997) and *Brassica napus (Bn,* Handa, 2003*)*. + and $+^{T}$ represent the presence of the group II intron in cis-form (green boxes) and trans-form (yellow boxes), respectively. The symbols ^{O}, ^{G} and ^{P} denote introns found in other genera (though not found in the present set of completed genomes) such as a *cox2* intron in *Peperomia* and *Pinus* (Qiu et al., 1998), a second *rps3* intron found in Gymnosperms (Regina et al., 2005) and a potato-specific *rps10* intron (Choury et al., 2005), respectively. Based on a phylogenetic perspective by Knoop (2004), an alternate nomenclature including the nucleotide position of the group II intron has also been provided in the last column.

Gene	Intron	*Ta*	*Zm*	*Os*	*Bv*	*Nt*	*At*	*Bn*	Alternate nomenclature* (Knoop, 2004)
nad1	1	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	*nad1*i394
	2	+	+	+	+	+	+	+	*nad1*i477
	3	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	*nad1*i669
	4	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+	+	*nad1*i728
nad2	1	+	+	+	+	+	+	+	*nad2*i156
	2	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	*nad2*i542
	3	+	+*	+	+	+	+	+	*nad2*i709
	4	+	+	+	+	+	+	+	*nad2*i1282
nad4	1	+	+	+	+	+	+	+	*nad4*i461
	2	+	+	+	+	+	+	+	*nad4*i976
	3	+	+	+	-	+	+	+	*nad4*i1399
nad5	1	+	+	+	+	+	+	+	*nad5*i239
	2	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	*nad5*i1455
	3	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	+^{T}	*nad5*i1477
	4	+	+	+	+	+	+	+	*nad5*i1872
nad7	1	+	+	+	+	+	+	+	*nad7*i140
	2	+	+	+	+	+	+	+	*nad7*i209
	3	+	+	+	+	+	+	+	*nad7*i676
	4	+	+	+	+	-	+	+	*nad7*i917
cox2	1	+	+	+	+	+	+	+	*cox2*i373
	2^{O}	-	-	-	-	-	-	-	*cox2*i691
rps3	1	+	+	+	-	+	+	+	N/A
	2^{G}	-	-	-	-	-	-	-	N/A
rps10	1^{P}	-	-	-	-	+	+	+	N/A
ccmFc	1	+	+	+	+	+	+	+	N/A
rpl2	1	-	-	+	-	+	+	+	N/A
total	26	22	22	23	20	23	24	24	

in the *nad1*, *nad2* and *nad5* genes (6 of the 26) of flowering plant mitochondria (Bonen, 1993), with the exception of *nad1* intron 4 which is cis-splicing in *Brassica* and Arabidopsis, but trans-splicing in the other sequenced angiosperms (Handa, 2003). Though *nad1* and *nad2* group II trans introns are mostly found in flowering plants, *nad5* trans introns emerge as early as in mosses (though none in the byrophyte *Marchantia*) (Malek and Knoop, 1998). In the case of the green alga *Mesostigma veride*, a common ancestor to all green plants, the two mitochondrial trans-splicing group II introns in *nad3* are suggested to have arisen independently (Turmel et al., 2002). This has interesting implications on the evolution of these introns-in-pieces, and the mechanisms affecting their splicing.

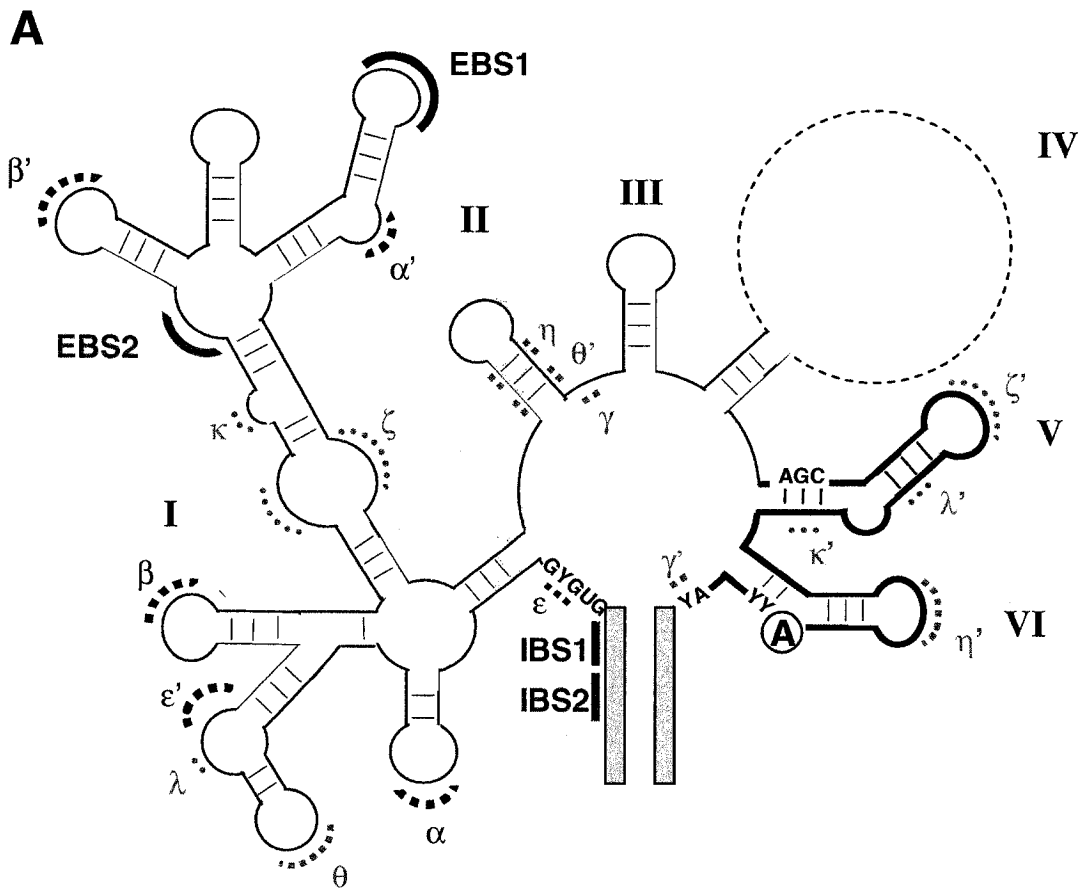
Notably, only one mitochondrial intron is shared at the same position in an early diverging plant, the bryophyte *Marchantia polymorpha*, and the flowering plants, namely *nad2* intron 3 (*nad2*i709; see Table 1.1 for alternate intron nomenclature as described by Knoop, 2004 and Dombrovska and Qiu, 2004). This denotes the highly dynamic nature of these retro-elements. Interestingly, a group I intron, *nad5*i573, is conserved in most mosses and liverworts, but in angiosperms it is absent, whereby *nad5* gained four group II introns, two of which are trans-splicing (Groth-Malonek et al., 2005). An example of a group I intron in a seed plant is *cox2* of *Peperomia* which is thought to be acquired via lateral gene transfer from a fungal donor (reviewed in Cho et al., 1998).

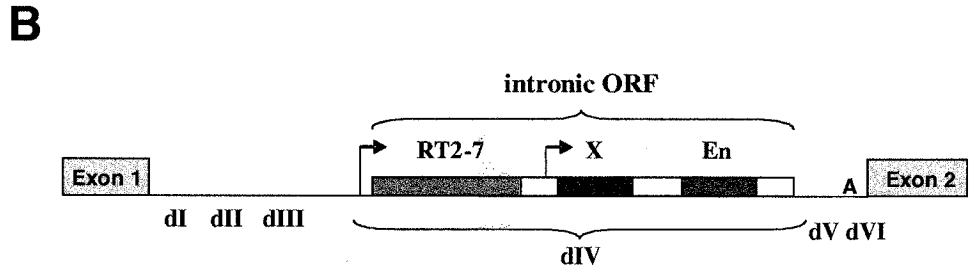
1.1.1 Splicing of group II introns

The crux of the auto-catalytic activity of group II introns lies within their distinct secondary and tertiary structural mechanics (Figure 1.2). Despite their overall large size and low sequence conservation, group II introns form a distinctive six domain (dI-dVI) secondary arrangement that undergoes inter- and intra-domain long range interactions generating a three dimensional configuration that is essential for ribozymic activity. An overall consensus 3D structural model has not yet been fully elucidated, however key components have been found that are crucial for self-splicing. One important feature is the base-pairing of exon binding sites (EBS) within dI, one of the largest of the six domains, with intron binding sites (IBS) found at the 3' end of the upstream exon (Figure 1.2). This base pairing, as well as interactions within positions in dI itself, such as α–α', β–β' and ε–ε', create the primary scaffold for other long range interactions with regions within dV (ζ–ζ', λ–λ' and κ–κ') and is altogether essential for

Figure 1.2 Schematic of a typical secondary structure of a conventional group II intron and a plant mitochondrial intronic open reading frame (ORF).

[A] Group II introns contain six helical domains (I-VI) surrounding a central hub. Exon binding sites (EBS1 and EBS2) within dI interact with intron binding sites (IBS1 and IBS2) found at the 3' end of the upstream exon. Other long range interactions are also shown by dotted lines and Greek symbols α–α', β–β', ε–ε', ζ–ζ', λ–λ', κ–κ', η–η', γ–γ' and θ–θ'. These pairings form the catalytic core for splicing (reviewed in Michel et al., 1989; Qin and Pyle, 1997; Lehmann and Schmidt, 2003). dV and dVI are in bold, while the detailed secondary structure models for these regions are shown in Figure 1.4 for plant mitochondrial introns. Typically invariant sequences are shown with uppercase letters representing the nucleotides (Y = pyrimidine), including an AGC triad in dV and the unpaired adenosine in dVI. The dashed region in dIV represents the position of the intron encoded protein. **[B]** Features of the plant mitochondrial intronic ORF (*mat-R*) include a reverse transcriptase (RT) domain, a maturase (X) and an endonuclease domain (En, formerly associated to a Zn finger motif) which are showed in warm colored boxes. Exons are shown as light blue boxes. Figure 1.2 was adapted from Michel and Ferat, 1995 and Bonen and Vogel, 2001.


A
EBS1
β'
α'
II
III
IV
EBS2
η
θ'
γ
ζ'
V
κ
ζ
AGC
λ'
I
κ'
β
GYGUG
ε
γ'
YA
YY
A
IBS1
IBS2
η'
VI
ε'
λ
α
θ


B
intronic ORF
RT2-7
X
En
Exon 1
A
Exon 2
dI dII dIII
dV dVI
dIV

the creation of the catalytic core for splicing (reviewed in Michel et al., 1989; Qin and Pyle, 1997; Lehmann and Schmidt, 2003). The catalytic core consists mainly of the interaction between dI and dV. The most phylogenetically conserved domain, dV contains a total of 34 nucleotides of 9+5 or 8+6 conserved base pairs and is capped by a tetraloop (Michel and Ferat, 1995). An AGC triad at the base of the dV stem is highly conserved, and any substitution of these nucleotides results in loss of splicing (Boulanger et al., 1995). Termed the 'catalytic effector', the presence of dIII in association with dI and dV, improves the rate of splicing (Xiang et al., 1998) and is suggested to interact with dV (Jestin et al., 1997), as well as with divalent metal ions (Sigel et al., 2000). Both dII and dIV (excluding the intron encoded protein within dIV) are variable in both sequence and length, and when deleted only slightly reduced splicing specificity (Koch et al., 1992). However, the region within the stem of dII has been linked to tertiary interactions with both dI (θ–θ'; Costa et al., 1997) and the hairpin loop of dVI (η–η'; Chanfreau and Jacquier, 1996). A site between dII and dIII (γ–γ') also interacts closely with the catalytic core and the last nucleotide at the 3' end of the intron (Fedorova et al., 2002). The unpaired adenosine involved in the first step of splicing is located in dVI, 6-8 nt upstream of the 3' splice site. The linker between dV and dVI is typically conserved 2-3 nt in length. Based on these secondary and tertiary structure associations, as well as an intronic encoded protein located in dIV, the group II introns have been further subdivided into three classes; group IIA-C (rev. in Toor et al., 2001). The mitochondrial group II introns, including those in the plant lineages, mainly fall into the group IIA category while chloroplast-type and bacterial-type group II introns are of the group IIB and C classes, respectively (Figure 1.4A for dV/dVI structure).

Self-splicing *in vitro* for group II introns was first observed in *cox1* group II intron (*al5γ*) in yeast mitochondria (van der Veen et al., 1986; Peebles et al., 1987). To date, no examples of self-splicing of group II introns in plant mitochondria have been demonstrated. To better study the key components of splicing within these introns, mutational analysis of transcripts that are reinserted into mitochondria via electroporation was performed by Farré and Araya (2001). Using this transformation method, splicing was found to be inhibited when nucleotides in EBS sites and the stem-loop structure of dVI were modified (Farré and Araya, 2002).

1.1.2 The intronic encoded protein

Splicing *in vivo* requires association with proteins that can be host-encoded or intron-encoded. In the latter case, the dIV intronic ORF was found to interact with the excised lariat structure of *aI2* in yeast mitochondria forming a ribonucleoprotein complex (Huang et al., 2005). The intronic encoded protein (IEP) contains a single reading frame that typically encodes one long protein with three domains: 1) a reverse transcriptase (RT) domain, 2) a maturase (X) containing an RNA binding domain for splicing, and 3) an endonuclease domain (Zn, Zn finger motif) involved in intron mobility (Figure 1.2B, adapted from Zimmerly et al., 2001). The mode of action for splicing involves binding to a high affinity site in dIV of the intron, as well as dI, dII and dVI, forming a ribonucleoprotein structure (RNP) and changing the structural conformation of the intron (Wank et al., 1999). The IEP is also involved in group II intron retrotransposition, which is a postulated mechanism for their proliferation throughout fungal, protozoan and plant organelles, as well as in the bacterial and archaebacterial kingdoms (reviewed in Lambowitz and Zimmerly, 2004).

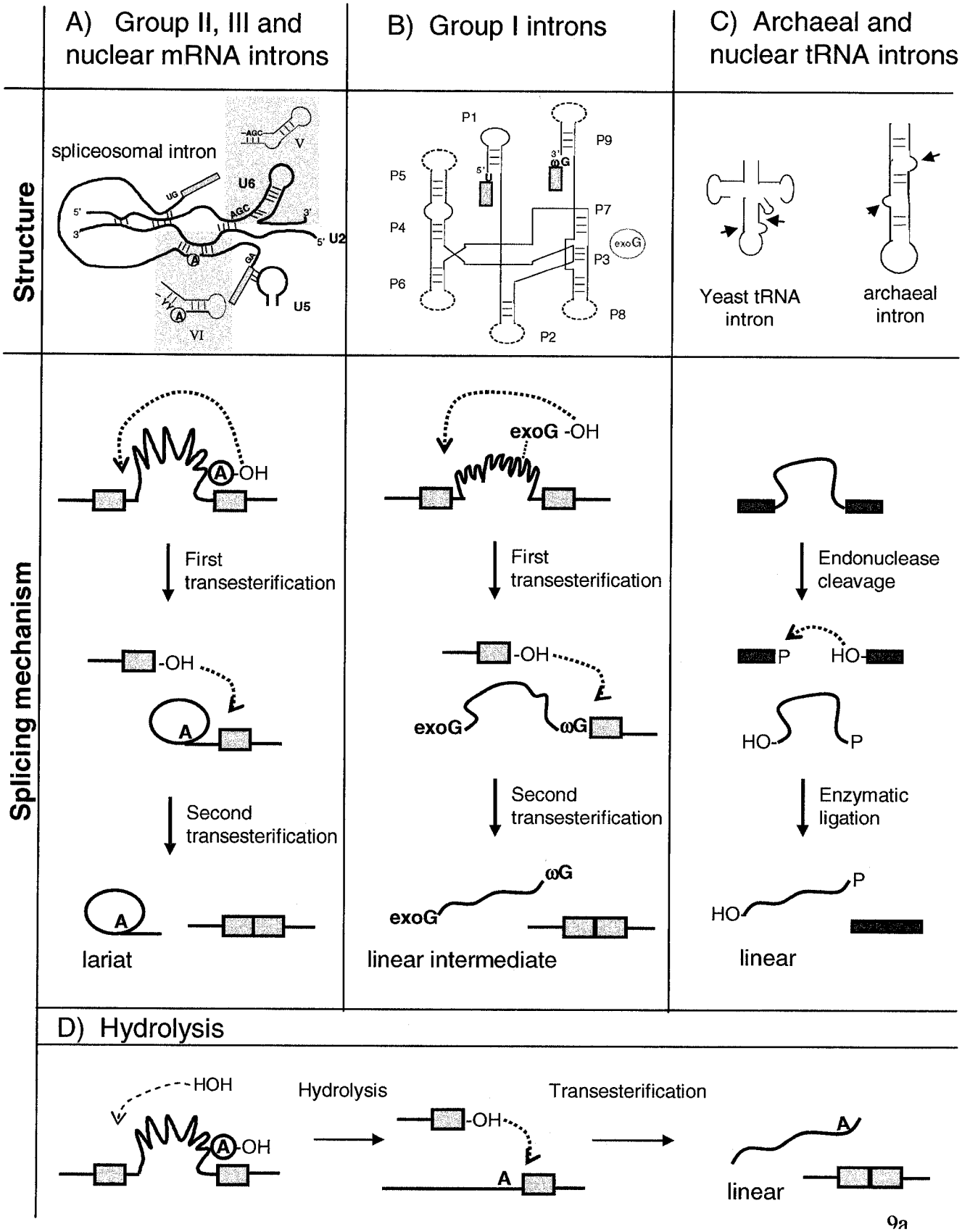
In the case of flowering plants, the mitochondrial intronic ORF or *mat-R* is located in *nad1* intron 4, containing only the RNA maturase (X) and a shortened reverse transcriptase domain (Figure 1.2B). Interestingly, remnants of a *cox2* intron *mat-R* were found in the sole species of the ancient gymnosperm *Gingko biloba*, supporting the idea that ancestral plant mitochondrial group II introns each originally encoded the protein needed for their own splicing and mobility (Ahlert et al., 2006). Notably, mitochondrial *mat-R* genes were also found in the nuclear genomes of both Arabidopsis and rice, suggesting that their RNA splicing function may be maintained by import back into the organelle (Mohr and Lambowitz, 2003). A more recent study by Nakagawa and Sakurai (2006) showed that in Arabidopsis, a mutation in this nuclear encoded *mat-R* resulted in reduced splicing efficiency of the mitochondrial *nad4* transcript.

1.1.3 Non-conventional splicing of group II introns

There are several examples of non-conventional splicing (Figure 1.3 A-D), some that result in linear rather than lariat forms of excised group II introns observed when *in vitro* mutational analyses of core ribozymic sites have been performed (van der Veen et al., 1987; Jarrell et al., 1988). These linear forms are generated via a hydrolytic pathway (Figure 1.3D) when water acts as the first-step attacking nucleophile (Daniels et al., 1996). This mechanism is

Figure 1.3 Comparison of structure and splicing mechanisms of the four major classes of introns.

[A] Spliceosomal introns, group II-type and group III-type introns use the two step transesterification pathway of splicing where the 2' OH of a branchpoint adenosine near the 3' intron end (circled A) acts as the nucleophile in first-step splicing. The resulting excised intron is in a lariat form as shown. Similarities between the spliceosomal intron and certain interacting snRNAs (U2 and U6) with dV and dVI of group II introns (highlighted in pink) have strengthened the hypothesis that they share a common evolutionary ancestor (adapted from Valadkhan, 2005). **[B]** The secondary structural model of the group I introns contains nine helical domains (P1-P9). Group I intron removal also occurs via a two step transesterification pathway, however an exogenous guanosine (exoG) acts as the nucleophile in first step splicing. A 3' terminal guanosine (ωG) replaces the exoG at the G-binding site, allowing for the second splicing step (reviewed in Fedor and Williamson, 2005) generating an intron excised in linear form. **[C]** Archaeal and yeast nuclear tRNA introns are removed through an endonucleolytic cleavage event at specific sites shown by arrows in the schematic (reviewed in Lykke-Andersen et al., 1997). The tRNA pieces are joined together via an unknown RNA ligation, possibly a tRNA ligase (Englert and Beier, 2005). **[D]** An additional mechanism of splicing observed for group II introns, based on *in vitro* studies (reviewed in Lehmann and Schmidt, 2003), is the hydrolytic pathway of splicing. Water acts as the first-step attacking nucleophile, while the second transesterification step results in the excised linearized form of intron. The exons are shown as light blue boxes, while regions coding for tRNAs are represented as grey boxes.


A) Group II, III and nuclear mRNA introns
B) Group I introns
C) Archaeal and nuclear tRNA introns
Structure
Splicing mechanism
spliceosomal intron
U6
U2
U5
V
VI
P1
P2
P3
P4
P5
P6
P7
P8
P9
exoG
Yeast tRNA intron
archaeal intron
-OH
First transesterification
Second transesterification
lariat
exoG -OH
ωG
linear intermediate
Endonuclease cleavage
Enzymatic ligation
HO-
P
linear
D) Hydrolysis
HOH
Hydrolysis
Transesterification
linear

proposed to substitute for lariat splicing in *in vitro* mutational analysis of dV and dVI in yeast mitochondria (Podar et al., 1998). Another study conducted by Grandlund et al., (2001) under high salt conditions showed the bacteria *Streptococcus* group II intron *GBSi1* to be excised in linear form despite containing the conventional bulged adenosine. In this case however, while many classical group II features are present, the GBSi1 intron lacks the EBS2-IBS2 long-range pairing.

To date there are only a few reported *in vivo* examples of the group II introns being excised in 'non-lariat' physical forms, for example by a hydrolytic pathway or other proposed non-traditional mechanisms. An interesting case is the barley chloroplast *trnV* group II intron, which lacks the dVI bulged A and was found excised in linear form (Vogel and Borner, 2002). Here, a hydrolytic pathway was proposed to generate these linear molecules, which were also susceptible to polyadenylation. Hydrolysis could also play an intermediary role for yeast mitochondrial aI5γ, in which full-length circularized excised introns were found (Murray et al., 2001). An additional proposed mechanism of splicing, that would generate such circularized introns, is the spliced exon reopening (SER) model, where the unattached upstream exon would act as nucleophile in first-step transesterification (Murray et al., 2001; Lehmann and Schmidt, 2003). The recent study by Molino-Sanchez et al. (2006) revealed both fully-circularized and lariat populations of excised group II intron in *Sinorhizobium meliloti*, a nitrogen-fixing bacterium. Furthermore, they suggested that the balance between these two forms was controlled by its intron-encoded *mat-R* protein. Interestingly, wheat mitochondrial cis- *nad7* intron 4 was found to excise both in lariat and circular physical forms consistent with the presence of more than one splicing pathway in plant mitochondria (C. Carrillo, Ph.D. thesis, University of Ottawa, 2001). Under the conventional splicing biochemistry, those plant mitochondrial trans-splicing introns that contain the bulging A in dVI would most likely excise in 'Y' shaped form, such as *nad1* intron 4, *nad5* intron 2 and *nad5* intron 3 (Figure 1.4). This raises the possibility of multiple splicing pathways being available for the removal of plant mitochondrial introns.

Interestingly a sub-class of group II introns is found within euglenoids, namely the group III introns. They are a streamlined version of group II lacking the dIII-dV, but still retaining a dVI branched adenosine for lariat splicing (Copertino et al., 1994). Another subcategory of

introns in euglenoids are the twintrons, which are group II introns that are found within other introns (Copertino and Hallick, 1993).

1.1.4 Plant mitochondrial group II introns

Plant mitochondrial group II introns notoriously deviate from conventional secondary structure. The dV and dVI secondary structural models of all 22 wheat mitochondrial group II introns are shown in Figure 1.4A, while the remaining 4 of the 26 plant mitochondrial group II introns found in plant species other than wheat are represented in Figure 1.4B. Aberrancies in the core dV features are observed in wheat mitochondrial introns which lack Watson-Crick base pairing (*cox2* intron 1 and *nad2* intron 3) and contain large end loops (*nad1* intron 1, *nad4* intron 2 and the *ccmFC* intron). Furthermore, these features are not conserved in different species of plants. For example, rice *nad2* intron 1 contains a dVI bulged A, while counterparts in wheat and Arabidopsis have instead a U residue and G residue, respectively. In some cases, such as *nad1* intron 2 of various plant species, the dVI branchpoint region is altogether missing. The lack of a branchpoint A would exclude first-step transesterification as the mechanism for splicing *in vivo*, however no alternate biochemistry of splicing has yet been elucidated for these plant mitochondrial introns.

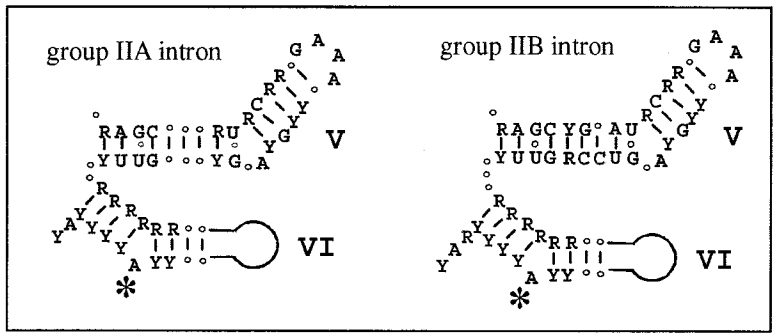
1.2 Splicing mechanisms of other classes of introns

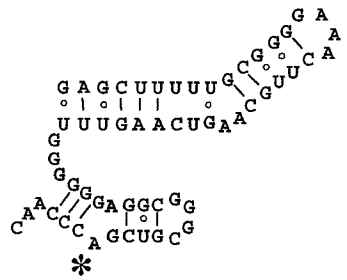
Apart from the group II type, the other three major classes of introns are 1) nuclear spliceosomal introns, 2) group I ribozymic type introns and 3) yeast nuclear and archaebacterial tRNA introns. The splicing mechanisms of these introns are summarized in Figure 1.3. Certain similarities in their biochemistries of splicing have led to the postulation that all four classes of introns stem from an ancient viral-like ancestor (Cavalier-Smith, 1991; Lehmann and Schmidt, 2005). Particularly, spliceosomal introns and group II introns share several common features, including the use of a downstream intronic bulged adenosine as the nucleophile in first step transesterification. The structure and positioning of the U2 and U6 snRNAs with the spliceosomal intron is similar to the group II intron catalytic core site (Figure 1.3A, pink boxes), (reviewed in Valadkhan, 2005). U6 and dV of group II introns also show remarkable similarities, including the invariable AGC triad (Sashital et al., 2004). The parallels between these two intron classes have been linked to a common evolutionary ribozymic RNA ancestor (Valadkhan, 2005). However, it cannot be excluded that these striking similarities may have

Figure 1.4 Secondary structure models of domains V and VI of 26 known plant mitochondrial group II introns.

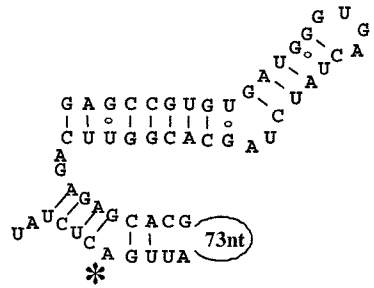
Known plant mitochondrial group II introns are represented by those in **[A]** wheat mitochondria (as this was the species used for this research) and **[B]** those that are not found in wheat but in other plant species. Branchpoint adenosines are shown with asterisks. Regions within introns that deviate from the classical structure by lacking a conventional dVI branchpoint are highlighted in green while those that contain unpaired bases within helical regions are highlighted in yellow. Certain lineage specific differences at the branchpoint regions are shown with ovals. Boxed uppercase Cs indicate A-C mispairs that are potentially corrected by editing (untested), while uppercase underlined Cs (red) are tested not to be corrected by editing and lower case underlined uridines (red) show positions that are edited thereby correcting the mispair (Bonen lab, unpublished data). Other deviancies from the conventional group II structure, such as longer or shorter stems or large dV tetraloops, are shown in parenthesis. Plant mitochondrial introns mostly fall into class A type group II intron, though some share some features with the class B type. The dV and dVI secondary structure of both classes are shown boxed [A]. Sequence data was obtained from AP008982, NC_00128, BA000029, NC_006581, X63625, Choury et al., 2005 and Regina et al., 2005 for wheat, Arabidopsis, rice, tobacco, carrot, potato and Gymnosperms, respectively.

A) Wheat mitochondrial group II introns

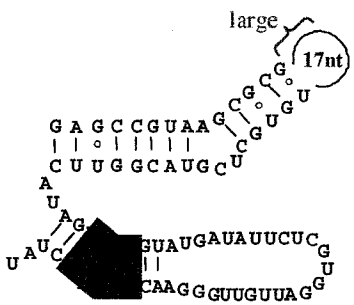




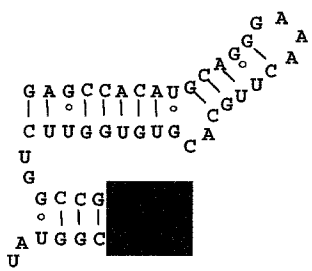
cox2 intron



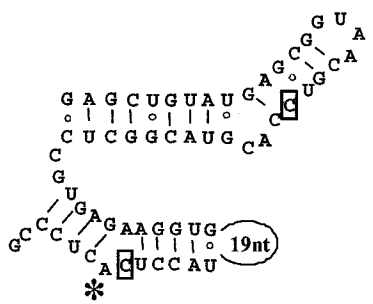
rps3 intron



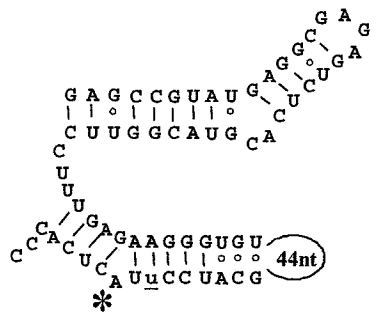
nad1 intron 1 trans



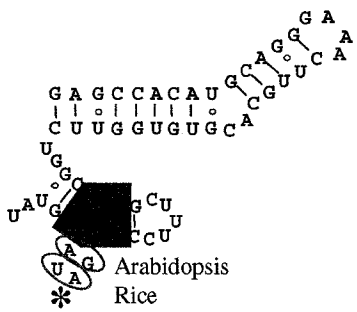
nad1 intron 2



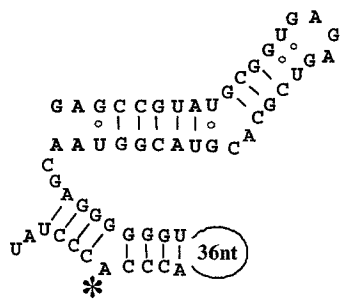
nad1 intron 3 trans



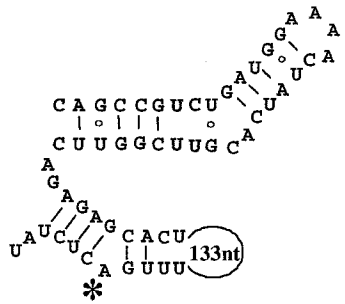
nad1 intron 4 trans



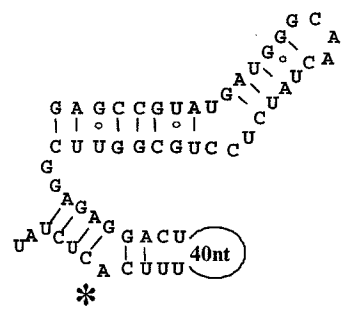
nad2 intron 1



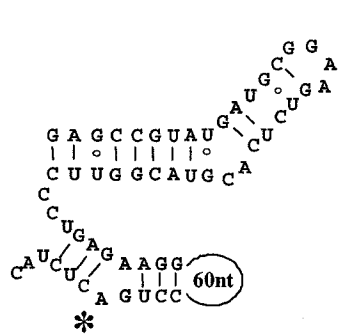
nad2 intron 2 trans



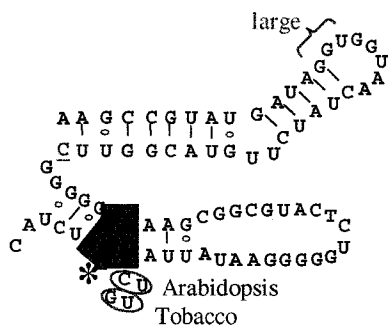
nad2 intron 3



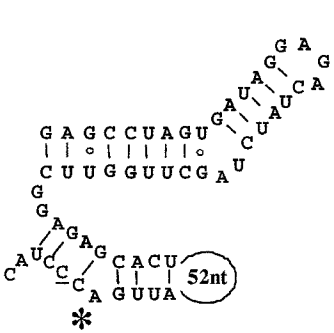
nad2 intron 4



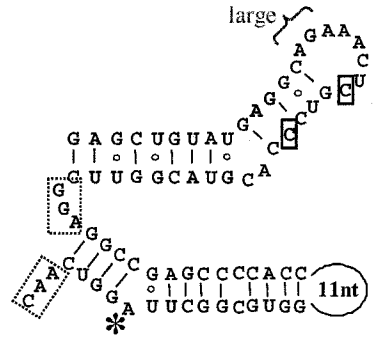
nad4 **intron 1**



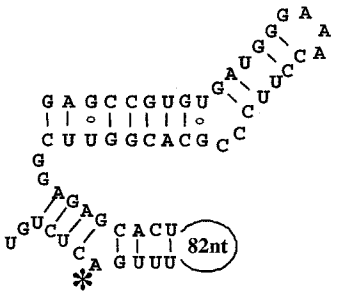
nad4 **intron 2**



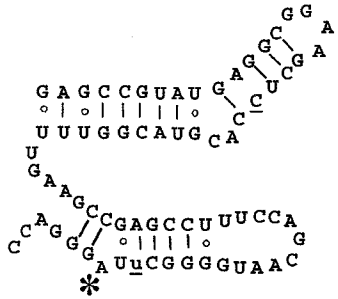
nad4 **intron 3**



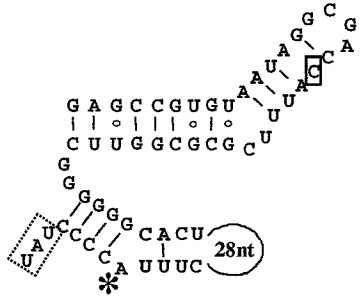
ccmFC **intron**



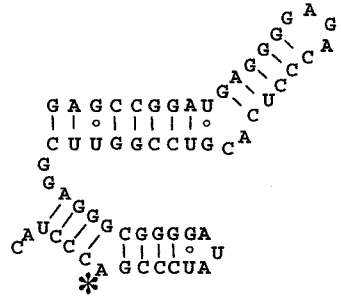
nad5 **intron 1**



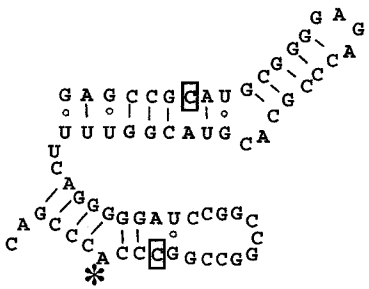
nad5 **intron 2 trans**



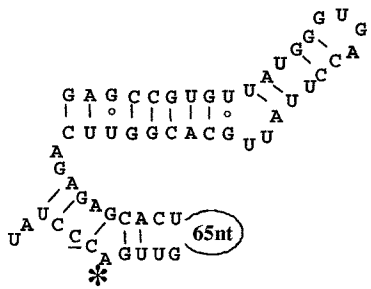
nad5 **intron 3 trans**



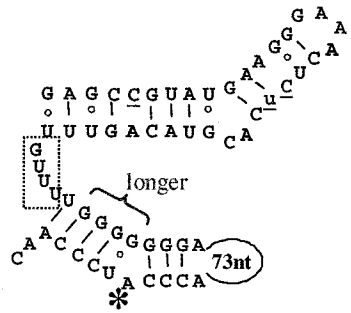
nad5 **intron 4**



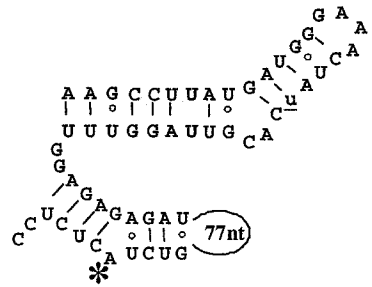
nad7 **intron 1**



nad7 **intron 2**

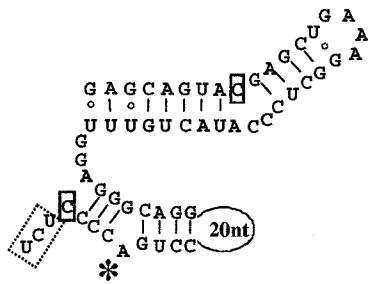


nad7 **intron 3**

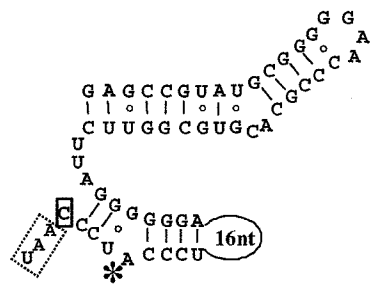


nad7 **intron 4**

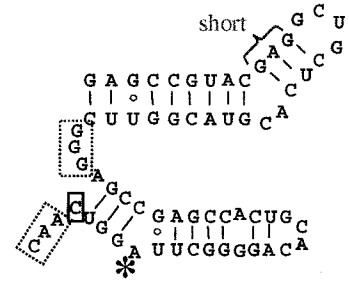
B) Group II introns not found in wheat mitochondria



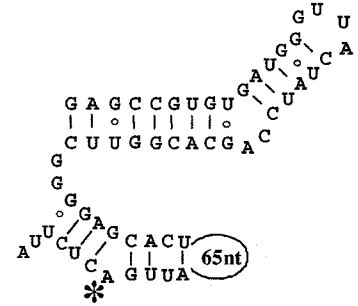
cox2 **intron 2 carrot**



rps10 **intron potato, Arabidopsis**



rpl2 **intron rice**



rps3 **intron 2 Gymnosperms**

instead arisen through convergent evolution as a product of chemical determinism (Weiner, 1993).

1.2.1 Nuclear spliceosomal introns

Nuclear spliceosomal introns are excised via a large ribonucleoprotein complex composed of over 200 proteins and 5 small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) (reviewed in Jurica and Moore, 2003; Roy and Gilbert, 2006). In spliceosomal intron excision, U1 and U2 snRNAs first recognize the 5' splice site and 3' end spliceosomal intron adenosine branch site (Figure 1.3A). This is followed by two transesterification reactions, releasing the excised spliceosomal intron as a lariat, with a longer tail (~100 nt). The nuclear spliceosomal intron also interacts with a divalent metal (Mg^{2+}) essential for proper intron core folding (Huppler et al., 2002). Another conserved region (ACAGAGA) found in U6 is positioned in a bulged region and interacts with the branchpoint nucleophile adenosine of spliceosomal introns.

1.2.2 Group I introns

The group I introns are similar to group II introns in that they are ribozymic, contain an intron encoded protein (IEP) and are located in bacteria as well as the organelles of plants, fungi and protists. An isolated case of a group I intron in animal mitochondria (sea anemone) was also found (Beagley et al., 1996). Group I introns, like group II introns and nuclear spliceosomal introns, undergo two transesterification steps; however an external guanosine acts as the nucleophile in the first splicing step (Figure 1.3B). Group I introns contain nine conserved paired regions (P1-P9) in which P7 contains a G site pocket where the exogenous G (exoG) binds and the first step of splicing is initiated, creating a 2'-5' OH bond with the 5' end of the intron (Figure 1.3B, reviewed in Huang et al., 2005). A 3' terminal guanosine (ωG) replaces the exoG at the G-binding site, allowing for the second splicing step (reviewed in Fedor and Williamson, 2005). A third transesterification step can also occur when ωG attacks a nucleotide close to the 5' end (P1) resulting in a shortened 5' end of the intron that is circularized (Nielsen et al., 2003; Lehmann and Schmidt, 2003). Another circularization method is also proposed by Nielsen et al. (2003), such that the 3' terminal ωG is released via hydrolysis, which becomes available for nucleophilic attack on the 5' end of the intron resulting in full-length circles, as well

as unligated exons. Like group II and spliceosomal introns, group I intron excision is also dependent on divalent cations in a three-metal model of catalysis (Stahley and Strobel, 2006).

1.2.3 Nuclear tRNA introns and archaebacterial introns

The splicing pathway for yeast nuclear tRNA and archaebacterial introns is distinct from that of the group II intron (Figure 1.3C). They instead use an endonuclease, to recognize and cleave their respective introns, and a RNA ligase for the covalent linkage of the RNA pieces (reviewed in Lykke-Andersen, 1997). A homologous endonuclease domain of 130 amino acids was found in both yeast nuclear tRNA and archaebacterial introns suggesting that they share a common evolutionary origin (Trotta et al., 1997).

1.3 Plant mitochondrial genetic systems and their expression

Based on the endosymbiont theory, mitochondria (as well as chloroplasts) are thought to be descendants of free-living eubacteria-like organisms, with the closest living relative of mitochondria being a α-proteobacteria (reviewed in Gray, 1999 and Hoffmann et al., 2001). Like their animal and fungal counterparts, plant mitochondrial genomes contain genes mainly involved in energy production via the electron transport chain pathway. However their genomes can range in size from ~200 kb in *Brassica* species to ~2400 kb in the muskmelon, in contrast to the mammalian mitochondrial genomes which are small (only 16 kb in humans) (reviewed in Gray, 1999). In contrast, the mutation rates of plant mitochondria are remarkably low, approximately 0.34×10^{-9} base substitutions per site per year, 100 times less than in mammalian mitochondria (Lynch et al., 2006). Within the last decade, eight plant mitochondrial genomes namely, *Marchantia polymorpha* (Oda et al., 1992), *Beta vulgaris* (Kubo et al., 2000), *Arabidopsis thaliana* (Unseld et al., 1997), *Brassica napus* (Handa, 2003), *Nicotiana tabacum* (Sugiyama et al., 2005), *Oryza sativa* (Notsu et al., 2002), *Zea mays* (Clifton et al., 2004) and *Triticum aestivum* (Ogihara et al., 2005) have been sequenced. An overview of their genome content, known protein coding genes (subdivided into number of respiratory-type, ribosomal and other types), as well as other RNA features (group II intron and number of edit sites), are summarized in Table 1.2.

Table 1.2 Summary overview of size and content of eight completed plant mitochondria genomes.

These are *Marchantia polymorpha* (Oda et al., 1992), *Beta vulgaris* (Kubo et al., 2000), *Arabidopsis thaliana* (Unseld et al., 1997), *Brassica napus* (Handa, 2003), *Nicotiana tabacum* (Sugiyama et al., 2005), *Oryza sativa* (Notsu et al., 2002), *Zea mays* (Clifton et al., 2004) and *Triticum aestivum* (Ogihara et al., 2005). Proteins encoded within plant mitochondrial genomes are ribosomal proteins (rp), electron transport chain pathway proteins (etc) and other (*) proteins including *mttB* (*orfX*) homologous to a transporter protein and the intron encoded protein (*mat-R*).

Plant	Genome size	Genes			RNA features		Reference
		rRNA	tRNA	Coding (rp, etc, other*)	Grp II intron (cis, trans)	Edits	
Marchantia polymorpha	186.6 kb	3	29	30 (16, 14, 3)	25 [6 group I]	no	Oda et al., 1992
Beta vulgaris	368.8 kb	5	25	31 (6, 22, 3)	20 in 7 (14,6)	370	Kubo et al., 2000; Mower & Palmer, 2006
Arabidopsis thaliana	366.9 kb	3	22	32 (7, 23,2)	24 (19, 5)	441	Unseld et al., 1997
Brassica napus	221.5 kb	3	17	34 (8, 23,3)	24 (19, 5)	427	Handa, 2003
Nicotiana tabacum	430.6 kb	3	21	36 (10, 23, 3)	23 in 10 (17, 6)	n/a	Sugiyama et al., 2005
Oryza sativa	490.5 kb	3	17	35 (11, 20, 2)	23	491	Notsu et al., 2002
Zea mays	569.6 kb	3	21	33 (8, 23, 2)	22 (15, 7)	n/a	Clifton et al., 2004
Triticum aestivum	452.5 kb	3	17	35 (9, 24, 2)	21 in 7 (15, 6)	n/a	Ogihara et al., 2005

1.3.1 Plant mitochondrial genome content

The inner membrane of mitochondria contains the components for ATP energy production via the electron transport pathway. These five components are composed of complex I - NADH dehydrogenase (*nad*), complex II - succinate:CoQ oxidoreductase (*sdh*), complex III - Cytochrome bc1 oxidoreductase (*cob*), complex IV - cytochrome oxidase (*cox*), and complex V - ATP synthase (*atp*) (Figure 1.1). A subset of proteins making up these complexes are from genes encoded within plant mitochondrial genomes typically *nad1-nad7, nad4L, nad9, sdh3, cob, cox1-3, atp1, atp4, atp6, atp8* and *atp9*. Notably, the remaining proteins that compose these complexes are nuclear-encoded and require import from the cytosol. Other genes located within plant mitochondria are those involved in translation, including the small and large subunit ribosomal protein genes (*rps* and *rpl*), tRNAs and rRNAs. Over evolutionary time, some of the the 14 ribosomal protein genes presently observed in the plant mitochondrial lineages have undergone or are in the process of undergoing transfer to the nucleus, often leaving behind lineage-specific pseudogenes (Ψ) in the mitochondrial genome (Adams and Palmer, 2003). Other genes found in plant mitochondria are the cytochrome c biogenesis genes (*ccm*) and the membrane targeting and translocating gene (*mttB*). In the seeds plants, an additional gene encoding for a reverse transcriptase (*mat-R)* is found within *nad1* intron 4, while in the bryophyte *Marchantia*, eight of its introns that are found in different genes (*18S rRNA* intron, *cox1* intron 1 and 2, *cox2* intron 2, *atp9* intron 2, *atpA* intron 1 and 2 and *cob* intron 3) contain copies of the *mat-R* gene (Oda et al., 1992; Toor et al., 2001).

1.3.2 Transcription of plant mitochondrial genes

In a few cases, plant mitochondrial genes are organized as clusters. An example that is present in several plant mitochondria is the ribosomal gene cluster *rps10-rpl2-rps19-rps3-rpl16-rpl5-rps14* found in *Marchantia* as well as partly intact in some flowering plants (Sugiyama et al., 2005). This impacts on gene expression, in that those genes that are clustered together are often co-transcribed. In tobacco mitochondria, eighteen gene clusters were suggested to be co-transcribed by Sugiyama et al., (2005) which include a mixture of respiratory type, ribosomal protein and tRNA genes (ex. *atp6-rps13-nad1b/c*), while the remaining 23 genes, mostly tRNAs, appear to be monocistronic. In rice, *rps3-rpl16* appeared to be co-transcribed with *nad3-rps12* from two alternative promoters (Nakazono et al., 1995). No strict consensus for plant

mitochondrial promoters has yet been elucidated, though general characteristics include an AT-rich region and a CRTA motif (reviewed in Brennicke et al., 1999). Interestingly, run-on transcription studies revealed that transcription rates differ among gene clusters in Arabidopsis, and are counterbalanced by RNA processing (Giegé et al., 2000; Kuhn et al., 2001). A recent study in Arabidopsis demonstrated that mature mRNA can originate from multiple promoters and their abundance can be regulated by stability events via polyadenylation and phosphorylation (Holec et al., 2006).

1.3.3 Transcript stability and turnover

RNA splicing and editing as well as endonucleolytic cleavages affect both co-transcribed and monocistronic genes. These post-transcriptional processes can influence the stability of the transcript (reviewed in Binder and Brennicke, 2003). For pea mitochondrial mRNAs, inverted repeats that form into stem loops at the 3' end of the untranslated regions were found to be involved in stability (Dombrowski et al., 1997). Polyadenylation tail tagging has been associated with degradation of mRNA in plant mitochondria. This tagging for decay has been observed in the 3' ends of plant mitochondria mRNA, chloroplast transcripts, yeast mitochondria RNAs and even human mitochondria mRNAs (Slomovic et al., 2005), and is residual from bacterial ancestry contrasting to the polyadenylation for stability seen on nuclear mRNAs (reviewed in Galgliardi et al., 2004).

1.3.4 RNA editing

In addition to splicing, virtually all plant mitochondrial transcripts undergo RNA editing whereby specific nucleotides are altered from their encoded genomic sequence often changing the final amino acid composition. This nucleotide modification from cytidine-to-uridine (C-to-U; Figure 1.1A) has been observed in the organellar mRNAs of land plants, and to a much lesser extent in tRNAs and rRNAs (reviewed in Gray, 2003). In Arabidopsis mitochondria, 441 C-to-U ORF edits were identified (Giegé and Brennicke, 1999), while in sugarbeet mitochondria, 357 C-to-U edits were found in protein coding transcripts (Mower and Palmer, 2006). Examples of C-to-U modifications were also found in nuclear encoded apoplipoprotein B transcripts in vertebrates, while adenosine to inosine or A-to-I RNA editing were detected mostly in neuronal tissues (reviewed in Maas et al., 2003). A deamination mechanism would be required for this

RNA editing process, and indeed both C-to-U and A-to-I deaminases have been isolated for vertebrates, termed Apobec and ADAR (adenosine deaminase acting on RNA), respectively (reviewed in Navaratnam and Sarwar, 2006). Additionally, U-to-C edits have been observed in 'lower' plant mitochondria and the plastids of hornwort (Kugita et al., 2003) and viral RNAs.

The mitochondria of higher plants predominantly undergo C-to-U editing, which in some cases have been found to create a start codon (ACG → AUG). Silent edits can also occur when the nucleotide being altered is at the third-codon postion. Partial editing has been found to mainly affect silent edit sites, as observed in sugar beet mitochondrial transcripts (Bonen et al., 1994; Mower and Palmer, 2006). Though infrequent, RNA editing can also take place in intronic regions affecting secondary structure and thus splicing efficiency (Carrillo et al., 2001 and Farre and Araya, 2001). Another type of RNA editing is the insertion of nucleotides, usually U's and C's observed in the mitochondrial mRNAs of *Trypanosomes* and *Physarum*. This insertion/deletion editing ultimately changes the reading frame of the mRNA and can also increase the transcript size up to a remarkable 50% as observed in *cox3* of the *Trypanosome*. Furthermore, this mechanism includes the use of small antisense guide RNAs as templates for insertion (reviewed in Simpson et al., 2003), adding another layer of complexity to this post-transcriptional modification.

1.3.5 Potential mechanisms of RNA editing in plant mitochondria

The mechanism or consensus nucleotide sequences associated with plant mitochondrial RNA editing have not yet been elucidated. Studies to characterize cis-elements required for editing in chloroplasts and mitochondria show those regions ~20 nt upstream, and less than 10 nt downstream of the edit site are important for editing (Bock et al., 1997; Chateigner-Boutin and Hanson, 2003; Mulligan et al., 1999). To gain insight about cis-elements required for editing, novel electroporation experiments were used to introduce the Sorghum *atp6* gene and an Arabidopsis *cox*2 gene into maize mitochondria resulting in *cox2* being fully spliced and edited, while the *atp6* remained unprocessed (Staudinger and Kempken, 2003). In tobacco chloroplasts editing site recognition appears to be mediated via site-specific trans-acting factors (Bock and Koop, 1997; Shikanai, 2006). Recently, members of the large family pentatricopeptides (PPR) have been proposed as part of the editosome interacting with cis-elements in chloroplasts (Kotera et al., 2005). In Petunia mitochondria, PPR proteins have been found to both restore fertility in

cytoplasmic male sterile (CMS) plants and promote the editing of *atp6* mRNAs (Wang et al., 2006). It has been suggested that C-to-U editing in plants is a deaminase reaction, rather than a transglycosylation reaction (reviewed in Mulligan et al., 1999). An example of a cytidine deaminase of prokaryotic origin was found in Arabidopsis suggesting possible association with C-to-U editing in the plant mitochondria, however, no organellar-specific targeting sequence has yet been observed (Faivre-Nitschke et al., 1999).

1.4 Plant mitochondrial gene expression during embryo-to-seedling development

An interesting period to investigate RNA processing events in plant mitochondria is during early seedling germination, when a transition from low to high energy demand is in progress. Dry seed mitochondria are poorly differentiated and lack cristae, but just hours after imbibition, ATP production by oxidative phosphorylation is evident (Bewley and Black, 1994). This suggests that mitochondrial development occurs rapidly after imbibition, and complexes assemble in the inner mitochondrial membrane, required for the oxidative pathway. In response to these developmental cues, mitochondria undergo active transcription and translation of respiratory genes, as well as protein synthesis type ribosomal genes. The relative steady state levels of respiratory chain pathway transcripts (*atp* and *cox)* from both mitochondria and cytosol increased during early maize germination (Logan et al., 2001). In contrast, relative levels of excised group II introns decreased from germinating embryos to 6 day seedlings (Subramanian, S., Ph.D. thesis, University of Ottawa, 1999; Carrillo, C., Ph.D. thesis, University of Ottawa, 2001). Also, non-respiratory genes, such as one encoding the ribosomal proteins in plant mitochondria, are particularly interesting to examine during the shift from low to high translational activity during early plant development. Apart from their essential role in translation, certain ribosomal proteins may also have a dual function, possibly containing RNA or protein binding/recognition sites associating with post-transcriptional regulatory gene products. This is seen with *E.coli rps12*, which was found to also enhance T4 bacteriophage group I intron splicing *in vivo* (Wool, 1996).

1.4.1 Alternative respiratory pathways of plant mitochondria

Apart from aerobic respiration, plants can have other sources for ATP production that include an anaerobic pathway or an alternative respiratory pathway that bypasses certain

respiratory chain complexes. The rotenone-insensitive pathway uses NAD(P)H proteins found on the outer surface of the inner mitochondrial membrane that can oxidize cytosolic NAD(P)H (Elhafez et al., 2006). Another alternative respiratory route is through the cyanide-insensitive pathway, which bypasses complex IV (*cox*) via an alternate oxidase (*aox*) complex (Moller, 2001) as shown in Figure 1.1B. Studies on germinating Arabidopsis seeds showed the presence of stored nuclear encoded *aox2* mRNAs in dry seeds and relatively low levels of *aox1a* mRNAs which increased after 2 days (Saisho et al., 2001). During early maize development, respiration via complex IV was found to be essential, while complex I could be bypassed through an exogenous NADH complex (Logan et al., 2001). Both these alternative electron transport pathways have also been associated with the removal of ROS (reactive oxygen species), reducing damage to the mitochondria caused by oxidative stress (Moller, 2000). The accumulation of group II introns (19/26) in the complex I genes, and the coordination between conventional and alternative respiratory pathways in plants, raises questions about a potential role of these introns in *nad* gene regulation.

1.4.2 Nuclear encoded proteins targeted to plant mitochondria

A critical parameter for proper mitochondrial function is the coordination of mitochondrial gene expression with that of the nucleus. Of the estimated 800-1000 proteins present in the mitochondria, only 30-40 of these are encoded within its genome, while the remaining are found in the nucleus (Hoffman et al., 2001). Thus, active transcription in the nucleus, translation in the cytosol, as well as protein import into the mitochondria is crucial. The well-studied mechanism of recognition and import of cytosolic proteins into the correct organelle usually occurs via signal peptides that are typically composed of hydrophobic and positively charged residues, and are located at the N-terminus of the protein (reviewed in Zhang and Glaser, 2002). In the case of those targeted to the mitochondria, proteins are recognized by TOM and TIM (translocases of the outer and inner mitochondrial membrane), and aided by chaperones (reviewed in Truscott et al., 2001) in passing through the mitochondrial membranes. Thus, during the period when plant seeds are leaving dormancy, nuclear encoded gene products may be a limiting factor in early post-transcriptional processing in plant mitochondria.

1.4.3 Plant mitochondrial system of this research: *Triticum aestivum*

The major Canadian crop plant *Triticum aestivum* or common bread wheat has been selected as the model species for this study on group II intron splicing during embryo-to-seedling development. Recent work has shown that hexaploid *Triticum aestivum* (AABBDD) was derived from *Aegilops tauschii* or goat grass (DD), *Triticum urartu* (AA) and *Aegilops speltoides* (BB) (Petersen et al., 2006). Wheat is thought to originate from the Fertile Crescent in the Middle East, but crossbreeding with *Aegilops* has allowed the modern day bread wheat to be cultivated in more temperate regions such as Canada.

1.5 Objectives

My research objectives focus on gene expression patterns, at the RNA level, of those wheat mitochondrial genes that contain group II introns. This study was performed during the embryo-to-seedling stages of development when there is a shift of low to high mitochondrial biogenesis which is an interesting period to study RNA processing events such as RNA splicing and editing. Furthermore, the physical excised form of wheat mitochondrial group II introns that lack conventional features was also examined to gain a better understanding of their splicing pathways. In this light, some the questions proposed are:

What are the expression profiles of mitochondrially-encoded genes during the shift from seed dormancy to seedling development? The expression patterns of certain oxidative phosphorylation genes, as well as ribosomal protein genes encoded within wheat mitochondrial genomes were studied using northern analysis of transcript profiles during the period when seeds leave dormancy and begin developing into seedlings. Several of these genes contain one or multiple group II introns, and their splicing profiles, as well as the relationship between *de novo* transcription and splicing, were also examined using this method.

How are the RNA processing events of editing and group II intron splicing coordinated in wheat mitochondria? RNA editing is typically an early event in the process between transcription and mRNA maturation (Mulligan et al., 1999). Wheat mitochondrial transcripts, including *nad2, nad4, nad5, nad7* and *cox2* transcripts contain edit sites that are close to exon/intron splice junctions. Using RT-PCR and direct sequencing analysis, the editing status of

partially-spliced transcripts were verified at these sites to examine the relationship between splicing and RNA editing. The editing status of *cox2* precursors during embryo-to-seedling development were also studied to gain some insight on the inter-relationship between editing and *de novo* transcription.

What is the biochemistry of splicing of group II introns in wheat mitochondria? The presence of the conventional dVI bulging adenosine would predict splicing that would generate lariat physical forms. However, certain plant mitochondrial group II introns lack the classical secondary structural features and thus may require the use of different splicing mechanisms. I examined the physical forms of candidate wheat mitochondrial introns that contain the branchpoint A (*nad2* intron 4 and *cox2* intron), as well as those that lack this feature (*nad1* intron 2).

CHAPTER 2: MATERIALS AND METHODS

2.1 Isolation of wheat mitochondrial RNA

Wheat seeds (*Triticum aestivum* cv. Frederick) used in this study were kindly provided by Dr. R. Pandeya (Agriculture and Agri-Food Canada, Ottawa). Methods for the isolation of wheat mitochondrial RNA from dry embryos (0 hr), imbibed 6 hr, 12 hr 18 hr, 24 hr, 2 day germinating seeds on filter paper in petri dishes, and 4 day and 6 day seedlings in vermiculite are described in Li-Pook-Than et al., 2004 (Chapter 3).

2.1.1 Dissection experiments on wheat seeds

Several parameters were tested to assess if the manipulation of the wheat seeds contributed to changes in gene expression:

1) The endosperm of wheat seeds were excised pre- or post- imbibition

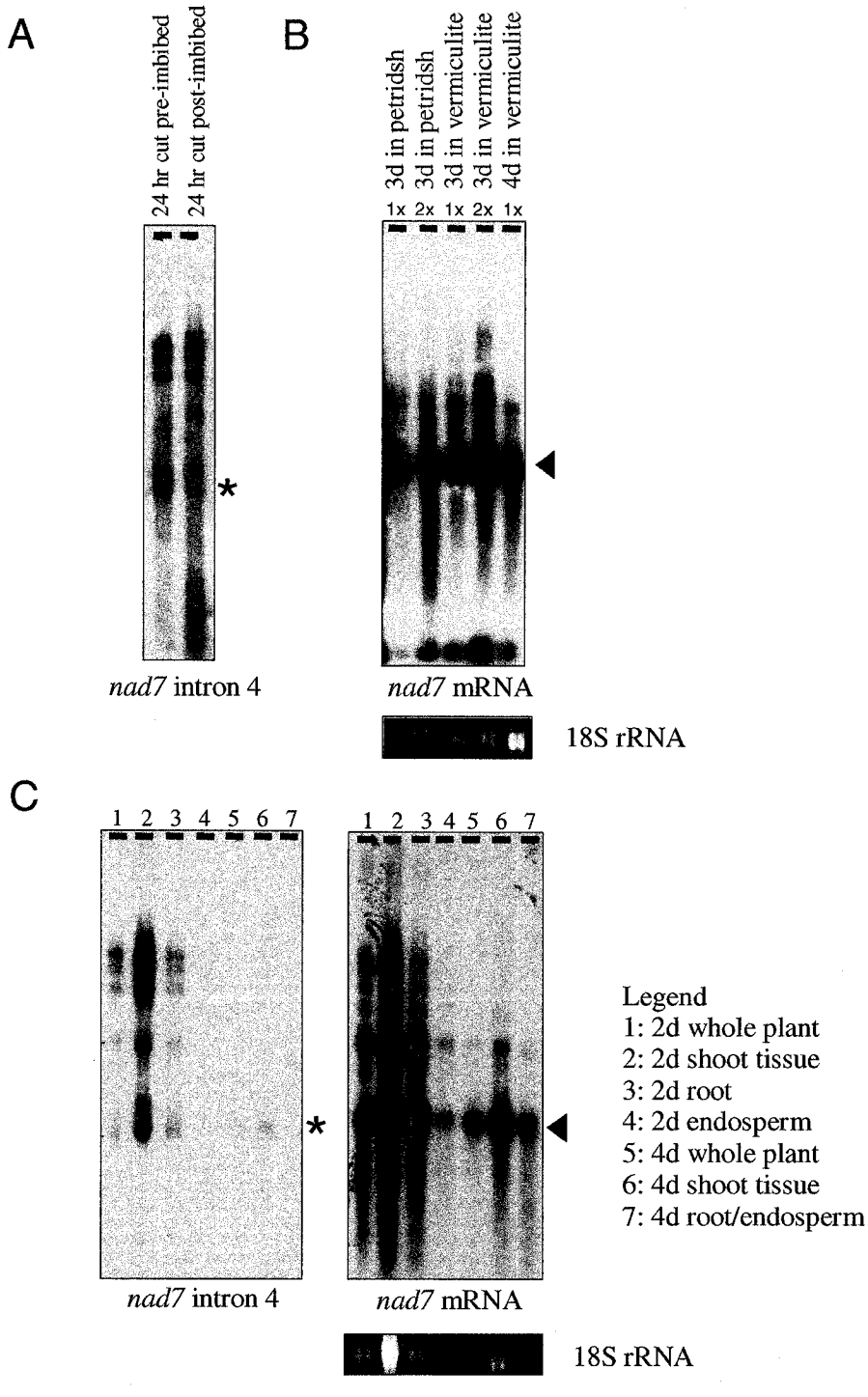
In previous experiments of our lab, part of the endosperm was dissected from the embryo pre-imbibition and subsequently the embryo-containing halves was grown on filter paper in petri dishes (C. Carrillo, Ph.D thesis, University of Ottawa, 2001). This dissection ensures minimal polysaccharide contamination during the RNA extraction procedure. Logan et al. (2001) described removal of the endosperm from maize seedlings post-imbibition, while the seeds were imbibed in between filter papers. To evaluate these two procedures and their effect (if any) on gene expression, I excised the endosperm using a sterile scalpel from both pre- and post-hydrated seeds. The embryo-containing halves were imbibed in between filter papers (2 beneath the seeds and 1 covering the seeds). Based on northern analysis using probes specific to the *nad7* gene, I observed no differences in the gene expression profiles of *nad7* intron 4 when 24 hr wheat germinating seeds were dissected prior to or after imbibition (Figure 2.1A). For the experiments reported in this thesis, the endosperm of wheat seeds were dissected using the traditional method, which is pre-imbibition.

2) Wheat seeds were grown on filter paper in petri dishes versus vermiculite

Based on northerns probed with primers specific to *nad7* mRNA (Figure 2.1B), no significant differences in gene expression patterns were observed when 3 day old seedlings were

Figure 2.1 Northern blot analysis following wheat seed dissection experiments of wheat mitochondrial gene expression.

[A] Embryos of wheat seeds were excised pre- or post- 24 hr imbibition, **[B]** wheat seeds were grown on filter paper for 3 days in pertri dishes or vermiculite and **[C]** wheat mitochondrial RNA was extracted from whole plants, shoots, endosperm and roots, as listed in figure legend to the right. Approximately equal amounts of wet weight tissue was used in [A] and [B], while for [C] 5X more wet weight tissue, and 2X more root tissue, was used than in shoots. Black arrowheads and asterisks show positions of *nad7* mRNA at 1.6 kb and excised *nad7* intron 4 at 1.7 kb, respectively. Exon-specific (LB92) and intron-specific (LB133) oligomer probes are as indicated below each panel. **[D]** Photographs of unexcised wheat seeds at the dormant stage of development (0 hr), 24 hr – 2 day imbibed and excised wheat seeds and a 4-6 day grown etiolated wheat seedling.



Embryo-to-seedling stages of wheat growth

grown in petri dishes versus vermiculite. For this thesis 6 hr, 12 hr, 18 hr, 24 hr and 2 day germinating seeds were grown on filter paper in petri dishes, while 4 day and 6 day seedlings were grown in vermiculite.

3) Wheat mitochondrial RNA was extracted from shoot, endosperm or root tissue

To examine if gene expression patterns differed in diverse tissues of germinating wheat embryos, RNA was extracted from shoot, endosperm and root tissue and northern analysis was performed using *nad7* mRNA and *nad7* intron 4 specific probes. Similar RNA profiles were obtained when mitochondrial RNA was extracted from shoot, root or endosperm tissue in both 2 day old etiolated germinating seeds (Figure 2.1C, lane 1 versus lanes 2-4) and 4 day old seedlings (Figure 2.1C, lane 5 versus lanes 6-7) However, levels of RNA/g wet weight tissue varied significantly. Despite containing 5X more wet weight tissue in endosperm than in shoots (Figure 2.1C, lane 3), and 2X more in roots (Figure 2.1C, lane 4), RNA was the most concentrated in 2 day germinating shoot tissue (approximately 10-12 μg); Figure 2.1C, lane 2) relative to endosperm and roots. In 4 day and 6 day old vermiculite-grown seedlings, only shoot tissue is used for this study.

2.2 Oligomers utilized in experiments

Oligomers used for each analysis, including northern and Southern blot hybridization, Rnase H analysis, sequencing, RT-PCR analysis and primer extension analysis are given in Table 2.1 and their positions within each respective gene are summarized in respective Figure 2.2A-G.

2.3 Northern and Southern blot hybridization

Northern and Southern blot analysis, utilizing oligomers end-labelled with γ-^{32}P-ATP as a probe for hybridization, is described in Li-Pook-Than et al., 2004 (Chapter 3) and Li-Pook-Than and Bonen, 2006 (Chapter 4). In some cases, (such as Figure 3.1, 3.2 and 3.3) the same blot was consecutively re-hybridized with different probes after allowing for adequate radioactive decay period between each usage.

Table 2.1 Sequence and position of oligomers specific to wheat mitochondrial genes used for RT-PCR, cRT-PCR, RNase H analysis, primer extension and sequencing analysis.

For oligomer positions on schematics see Figure 2.2.

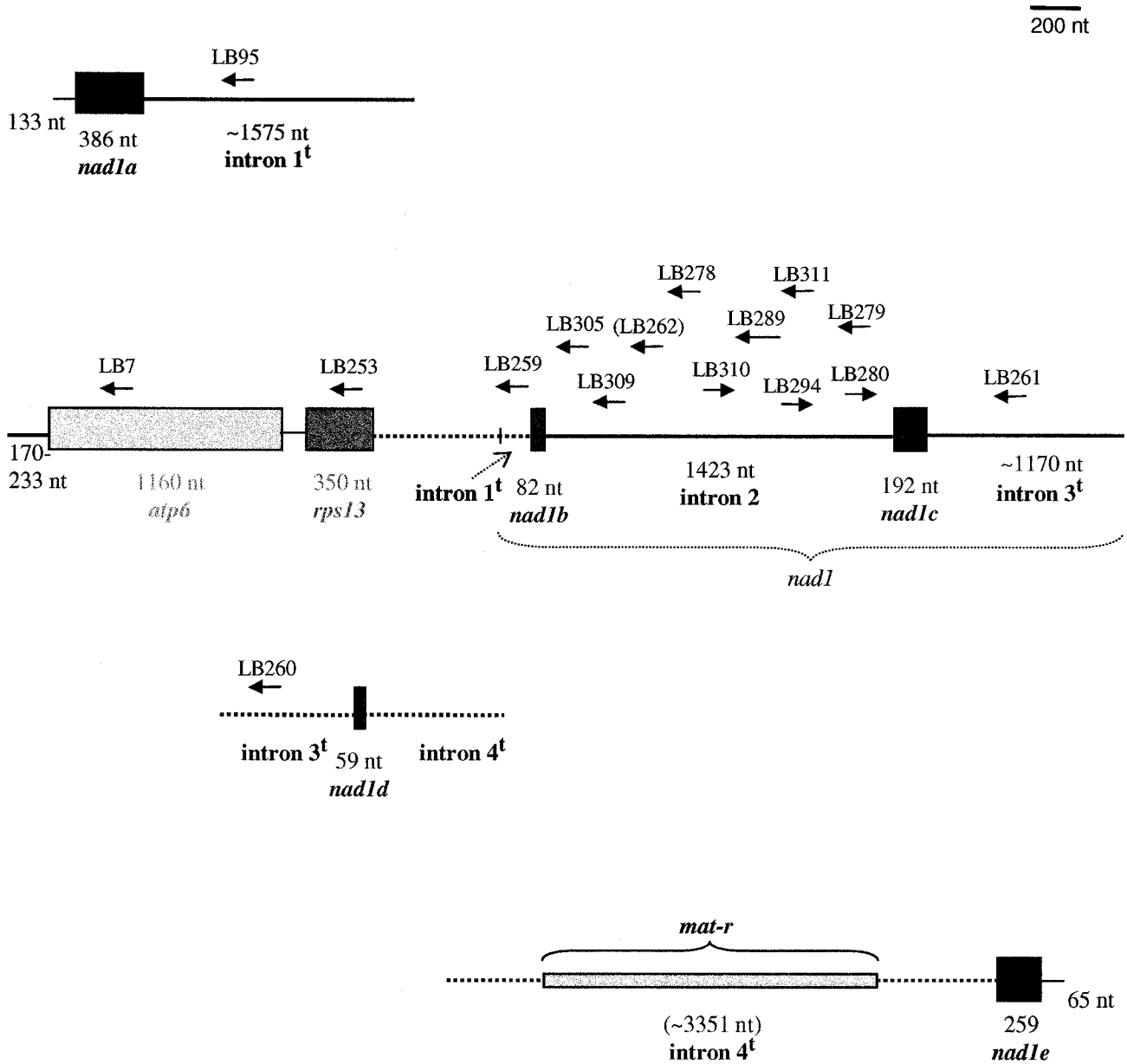
	Location	Position	Sequence (5′ to 3′)	Direction	Comments (Accession no.)
A) *nad1a, atp6/rps13/nad1bc,nad1d and nad1e*					
LB7	*atp6*		AGTTGCCCACATGAAGATCA	antisense	M24084
LB95	*nad1* intron 1	687-706	GACTTCCACTTATCGATCCC	antisense	X57968
LB253	*rps13*	n/a	AATCGTTCGATGTCTGCTCG	antisense	n/a
LB259	*nad1* intron 1	573-594	CCACGAGAATATCATACCCA	antisense	X57967
LB260	*nad1* intron 3	67-86	GGAAGTGTGTATAAATCGAC	antisense	X57967
LB261	*nad1* intron 3	2441-2460	GAAGAATGAAGAGTGCCCTG	antisense	X57966
LB262	*nad1* intron 2	903-922	TTGCTCATTGCAGTCGCCTC	antisense	X57967
LB278	*nad1* intron 2	1141-1160	CCGATCTGTCACCCATACAT	antisense	X57967
LB279	*nad1* intron 2	1843-1862	AAGGAAAGTCTAGGTTGGCG	antisense	X57967
LB280	*nad1* intron 2	1843-1862	CGCCAACCTAGACTTTCCTT	sense	X57967
LB289	*nad1* intron 2	1649-1688	CATCAAATGATGCATGTGGGCTG AGCCATTCCCATCCCAA	antisense	40mer, X57967
LB294	*nad1* intron 2	1953-1972	CAGCTTACTCACCCTACTCC	sense	X57967
LB305	*nad1* intron 2	728-747	AGATCCGAGCGATTGCCTTC	antisense	X57967
LB309	*nad1* intron 2	810-829	CTCAAAATGAGCCTTGCGAC	antisense	X57967
LB310	*nad1* intron 2	1474-1493	GGCTGGCTACAAGTATAGCC	sense	X57967
LB311	*nad1* intron 2	1953-1972	GGAGTAGGGTGAGTAAGCTG	antisense	X57967
B) *nad2cde*					
LB235	*nad2c*	928-947	GCTACCCACTTCGATCAATT	sense	Y14434
LB236	*nad2* intron 4	4247-4266	CCTAAGGCCAACTTCCATTC	antisense	Y14434
LB237	*nad2* intron 4	5441-5460	GAGAGGACTCAGCTGTTAGT	sense	Y14434
LB390	*nad2* intron 4	4485-4503	TTCTTCATTGGAGGGTTCGG	antisense	Y14434
LB391	*nad2* intron 4	5058-5077	AGAGACCCCTGTAGAGAATG	antisense	Y14434
LB400	*nad2* intron 4	4940-4959	AACGAGTCGACGATTGAAGC	antisense	Y14434
LB425	*nad2e*	5512-5531	GTAACGACTTGTCACGATCC	antisense	Y14434
C) *nad4*					
LB90	*nad4d*	7905-7924	TTTCTTTGCTGATTCCTCTC	sense	X57164
LB93	*nad4* intron 2	5470-5489	GTTCCGGTTCACTGATAAGG	antisense	X57164
LB203	*nad4* intron 3	6447-6464	CGTTCGAGTTATTGGCTCGT	antisense	X57164
LB221	*nad4* intron 4	7687-7702	GAATCAAAGTCCAGGTTGGT	antisense	X57164
LB229	*nad4* intron 4	7611-7632	AGGCTTATGTCATATGGGAA	sense	X57164
LB230	*nad4* intron 4	6527-6548	GAGCTCTGCCATAAGAACTTGG	antisense	X57164
LB389	*nad4* intron 4	7220-7249	AGACGCACGGAAGAACTATG	antisense	X57164
LB417	*nad4c*	6007-6026	TAGTAGCCACATTACGTGCG	sense	X57164
LB420	*nad4b*	2176-2195	ATACCCATGTTTCCCGAGGC	sense	X57164
nad5					
LB357	*nad5* intron 2	172-191	ACAACCTTAACCGCACAAGC	sense	M74157
LB410	*nad5* intron 4	1006-1025	AAAGCTCCGGTCTTAACGCC	antisense	M74159
D) *nad7*					
LB39	3′ UTR	6639-6620	TAGGATCCTGATCGAGCAAG	antisense	X75036
LB40	*nad7* intron 4	6493-6512	GCAGTCGACTGAGTTCTGAA	sense	X75036
LB41	*nad7* intron 3	4226-4245	CCACCAGATCTTAAGGAAAG	sense	X75036
LB42	*nad7* intron 4	5640-5621?	GCCGGTGCGTGGTGAACGGA	antisense	X75036
LB43	5′ UTR	552-571	CTGGACAAGCTTTAGGGGAA	sense	X75036
LB46	*nad7c*	1594-1623	ACTTATCTTCAAGCTTTACC	sense	X75036
LB47	*nad7c*	3361-3380	GCACAGCAAGCAAAGGATTG	sense	X75036
LB53	*nad7* intron 1	1408-1427	ACGCGAATTCGCTTCCGAGG	sense	X75036
LB60	*nad7* intron 3	4390-4371	CCACCACTTCACTTTTGCAC	antisense	X75036
LB64	*nad7* intron 3	3493-3474	ATGCATGCTTTTGTAGGGTC	antisense	X75036

LB87	*nad7* intron 3	4164-4183	TCTGACTCTATTGGTCATAG	sense	X75036
LB91	*nad7c*	3214-3233	GGAGTGGCACAAGATCTGCC	sense	X75036
LB92	*nad7d*	4596-4577	ATCGGCTTTGATCATGCCAC	antisense	X75036
LB94	*nad7* intron 3	3973-3954	CACTTTCATTCGACGTTCCG	antisense	X75036
LB133	*nad7* intron 4	4977-4958	CGTGTCAGCTTAGTTATC	antisense	X75036
LB159	*nad7* intron 4	4752-4733	AACCTCGGGATGCTTTACTC	antisense	X75036
LB198	*nad7* intron 2	2781-2802	CATGGAATAGCATGGCCCGG	sense	X75036
LB226	*nad7* intron 2	1780-1799	GATACGTGAGTCTTAGTGGG	antisense	X75036
LB335	*nad7* intron 4	5249-5268	TCCCGTACATCGTCTTTCTC	antisense	X75036
E) *cox2*					
LB13	*cox2a*	190-209	CAGCATCACAAAGAGCGATT	antisense	X01108
LB240	*cox2* intron 1	751-770	CCGAGAAGAGGTATGTGTAC	antisense	X01108
LB242	*cox2b*	2113-2132	TGATTGGATACCCAATCCGC	antisense	X01108
LB256	*cox2a*	189-210	CAATCGCTCTTTGTGATGCT	sense	X01108
LB287	*cox2* intron 1	837-857	CTGTGGTGCCGATAGATTCA	antisense	X01108
LB288	*cox2* intron 1	1683-1703	ATAAGAGTAGGCGTGGAGAG	sense	X01108
LB347	*cox2* intron 1	1703-1683	CTCTCCACGCCTACTCTTAT	antisense	X01108
LB348	*cox2* intron 1	1460-1441	AGTCGGCTTTGGCGAAGTTG	antisense	X01108
LB383	*cox2* intron 1	1441-1460	CAACTTCGCCAAAGCCGACT	sense	X01108
F) rps3/rpl16					
LB215	*rps3a*	170056-170877	CAGATCAGGTCTTACCGAAA	antisense	BA000029
LB217	*rps3* intron	171138-171157	TCCTTATCCTCGCGTTCATG	antisense	BA000029
LB290	*rps3* intron	171061-171079	TCTTTCACGACATGCTCTGG	antisense	BA000029
Lb291	*rps3* intron	172430-172449	GTTAGCGGCTTAAGCTAAGC	sense	BA000029
LB293	*rps3b*	173383-173402	GTGCCACGAAATGATTGAGA	antisense	BA000029
LB297	*rpl16*	174528-174547	GGCTTGTCGAGCATTTGACA	antisense	BA000029
LB319	*rps3* intron	170923-170944	GGACCGAACCCATTCGGTGA	antisense	BA000029
LB330	*rps3* intron	172275-172294	GGAATGGACAGCTGAGGTCA	antisense	BA000029
LB334	*rps3* intron	172430-172449	GCTTAGCTTAAGCCGCTAAC	antisense	BA000029
G) *rps7*					
LB27	*rps7*	442-461	ATCAATTTATCGGCCTCGTA	sense	X67242
LB37	*rps7*	610-631	GTTAGTTCGAGCTAGGCGGT	antisense	X67242
LB274	*rps7*	912-931	CACCAATCGGAGTTTCGCGC	sense	X67242

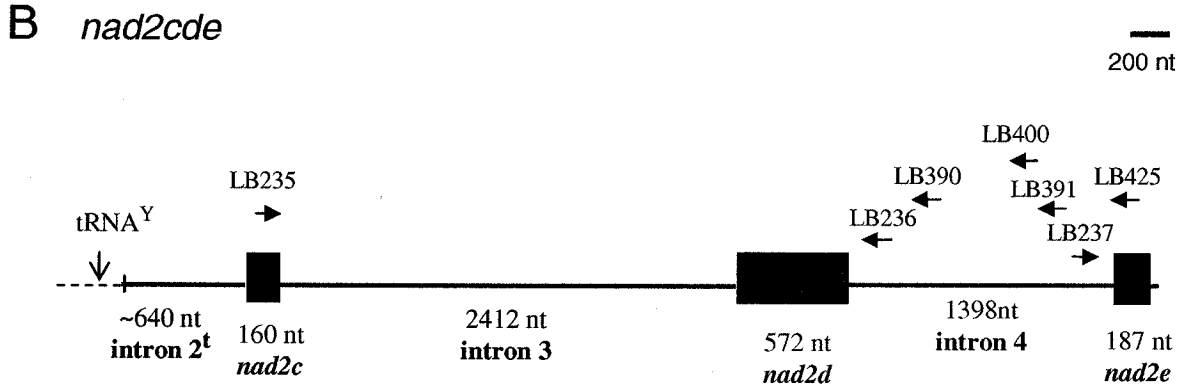
Figure 2.2 Schematics showing the positions of oligomers based on selected wheat mitochondrial genes.

[A] *nad1a, atp6/rps13/nad1bc, nad1d* and *nad1e*, **[B]** *nad2cde*, **[C]** *nad4*, **[D]** *nad7*, **[E]** *cox2*, **[F]** *rps3/rpl16* and **[G]** *rps7*. Oligomers (arrows) shown are used in multiple experiments including northern analysis, RT-PCR, cRT-PCR, the RNase H method, primer extension and sequencing analysis. Sizes of exons (colored boxes), introns (colored lines) and UTRs (black lines) shown in each schematic are from GenBank annotations where associated accession numbers for wheat mitochondrial genes are given in Table 2.1. The lengths of *nad1* trans-introns (') shown here are from S1 nuclease analysis mapping of the termini of *nad1* trans-introns (Chapedelaine, Y., Ph.D. thesis, University of Ottawa, 1992), while unknown UTR sizes are represented by dotted lines.

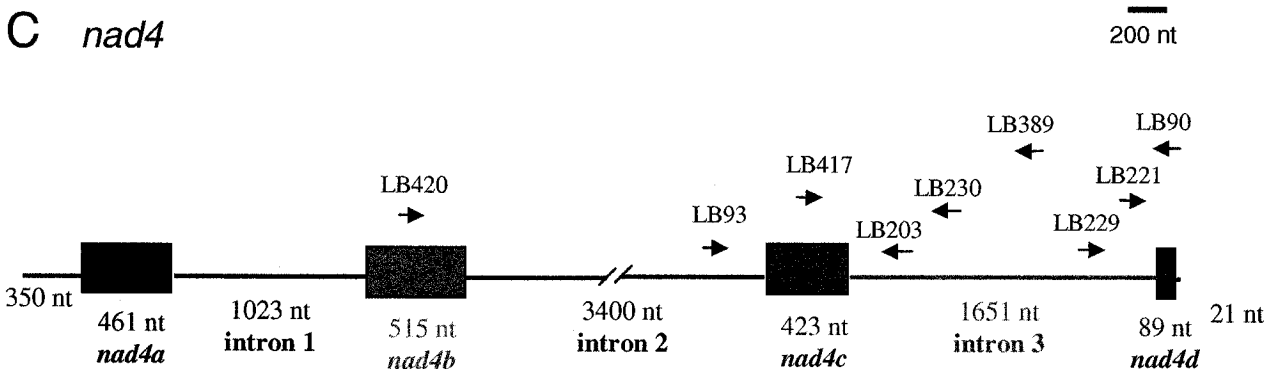
A *nad1a, nad1b/c, nad1d* and *nad1e*



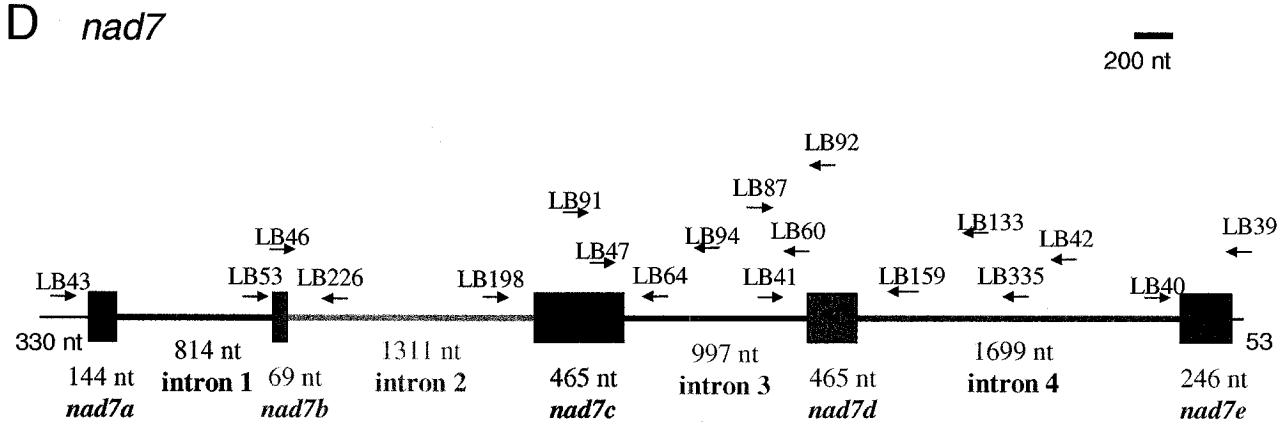
B *nad2cde*



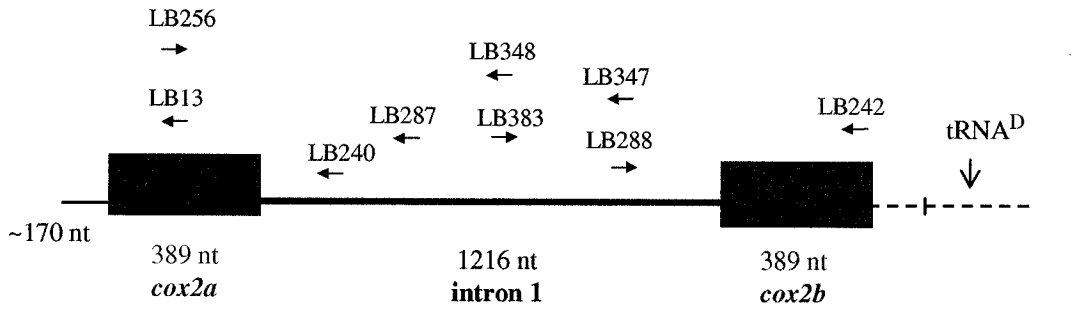
C *nad4*



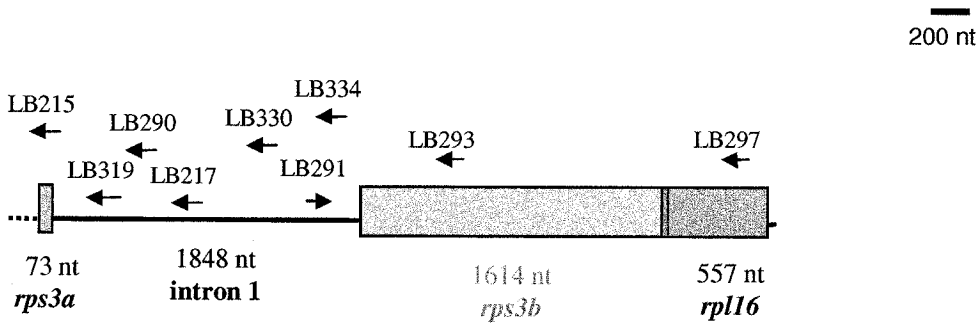
D *nad7*



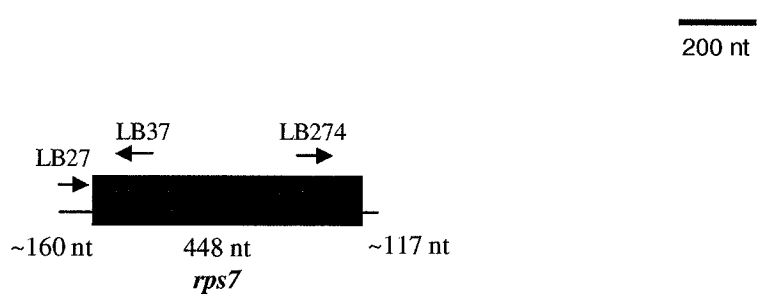
E *cox2*



F *rps3/rpl16*



G *rps7*



2.4 RNase H analysis

RNase H analysis was used for the study of physical forms of excised introns *in vivo*, and the method is described in Li-Pook-Than and Bonen, 2006 (Chapter 4). The integrity of the RNA was greatly improved with the addition of an RNAase inhibitor (RNasin, Promega) and dithiothretol (DTT) neither of which inhibits the activity of RNase H (Sambrook et al., 1989). Similarly, the omission of the RNA buffer which contained $MgCl_2$ (and subsequent replacement with ddH_2O) from the original RNase H protocol (as described in Carrillo, C., Ph.D. thesis, University of Ottawa, 2001) reduced RNA degradation.

2.5 Sequencing of clones and direct sequencing of RT-PCR products

The sequencing of clones and direct sequencing of RT-PCR and cRT-PCR (RNA pre-treated with RNA ligase) products are described in Li-Pook-Than and Bonen, 2006 (Chapter 4). C-to-U editing at exon sites close to unspliced introns were also examined using direct sequencing of RT-PCR products of wheat mitochondrial genes (Chapter 5, Li-Pook-Than et al., in press *Physiologia Plantarum*). In the latter case, to minimize contribution from contaminating DNA and to examine editing in transcripts containing total number of intron(s), mitochondrial RNA samples were successively treated three times with 2 U/μL DNase I (Amersham) in 32 mM Tris pH 7.5, 5 mM $MgCl_2$, 0.1 mM DTT and then phenol-extracted (3X) prior to RNA recovery by ethanol precipitation (Sambrook et al, 1989). Automated sequencing was performed by the Ottawa Health Research Institute DNA sequencing facility (OHRI) in Ottawa and the McGill University/Genome Québec Centre in Montréal.

2.6 Primer extension analysis

Primer extension is typically used to map the 5' ends of transcripts. Here, it was utilized to examine the presence of linear excised introns. Approximately 10 ug of 24 hr wheat mitochondrial RNA (24 hr) was extracted, ethanol precipitated, denatured at 70°C for 10 min, quickly cooled in ice prior to the addition of 1X First Strand Buffer (Invitrogen), 10 μM DTT, 0.1 mM dNTPs, and 100 U of reverse transcriptase. RNA was annealed with ^{32}P-end labeled antisense oligomers (as described in Li-Pook-Than et al., 2004) and primer extension occurred for 2 hrs at 45°C. Three reverse transcriptase enzymes were tested separately to examine the efficiency of extension, namely Superscript II reverse transcriptase (SSII – Invitrogen), MMLV

reverse transcriptase (Invitrogen) and Omniscript reverse transcriptase (Qiagen). Omniscript was found to be the most efficient at extending over RNA secondary structures and was used for the remaining primer extension analyses as shown in Figure 4.7. The samples were run on a 7% denaturing polyacrylamide gel adjacent to a DNA sequencing ladder usually of the corresponding DNA template (of the same oligomer).

2.7 RNA electrophoresis on a denaturing polyacrylamide gel

Examination of linear versus lariat (circularized) physical forms of excised introns also included the extraction of wheat mitochondrial RNA from 6 hr, 24 hr germinating embryos and 6 day seedlings (as previously described in Section 2.1) and running the RNA instead on a denaturing polyacrylamide gels (4%-6%). Lariat or circularized molecules would migrate slowly at positions above high molecular weight precursors, in contrast to linear molecules of the same size that would migrate with markers . Mitochondrial RNA (~10 μg) was denatured at 90°C for 5 min, cooled on ice and was run on a 6% polyacrylamide - urea gel (Sambrook et al, 1989) for 7 hrs in a vertical electrophoresis system (25 cm x 16 cm, Owl Scientific) at 500 V- 600 V (12 mAmps) in 0.5X TBE. The gel was removed from the glass plates using a dry filter paper of the same size and placed in an electroblotting support apparatus in the following order: Scotch-Brite pad, 3mm filter paper, gel, nylon membrane, pre-soaked 3 mm filter paper and Scotch-Brite pad. The nylon membrane (Hybond), pre-soaked in 0.5 X TBE, was placed on the gel and all air bubbles were removed. The gel was transferred onto the nylon membrane (faced towards the anode, +) using an electroblotting apparatus EC-420 (E-C Apparatus Corporation, Florida) for no more than 1 hr in 1X TBE at 15V (0.2 Amps). Membrane hybridizations were performed with ^{32}P-end-labeled oligomer probes (Li-Pook-Than et al., 2004) specific to wheat mitochondrial genes (*rps7* and *cox2*) as described in Chapter 4.7.4.

CHAPTER 3: RNA PROFILES OF WHEAT MITOCHONDRIAL GENES DURING EARLY PLANT DEVELOPMENT

3.0 Rationale

Plant mitochondria undergo a shift from low to high respiratory activity during the transition from dormant seeds to germinating embryos upon imbibition. Little is known about the regulation of expression of mitochondrially-encoded genes during this period of high energy demand and rapid mitochondrial biogenesis. I have used northern blot analysis to examine the steady state RNA profiles of several group II intron-containing respiratory chain complex genes (*nad1, nad7* and *cox2*) as well as selected ribosomal protein genes (*rps3/rpl16*). This was done to address if transcription and RNA splicing events are tightly coordinated during this developmental transition. The relative abundance of steady state mRNAs, precursors and excised group II introns of both the respiratory pathway and ribosomal protein genes were compared during early wheat development to determine if genes of different functions demonstrate gene-specific regulation patterns. Furthermore, to better understand the particular order of group II intron excision (if any), the comparison between steady state levels of precursor transcripts versus splicing intermediates were also examined in *nad7* which contains four cis-group II introns.

This section (3.1-3.5) was published as: "Li-Pook-Than, J., Carrillo, C. and Bonen, L. (2004) Variation in mitochondrial transcript profiles of protein-coding genes during early germination and seedling development in wheat. *Current Genetics*, **46**(6), 374-80." Experiments represented by the one shown in Figure 4 in Li-Pook-Than et al., 2004 were performed by C. Carrillo. This section is followed by supplementary data (Section 3.6) that was not included in this paper.

Chapter 3: RNA profiles of wheat mitochondrial genes during early plant development

Jennifer Li-Pook-Than, Catherine Carrillo, and Linda Bonen

3.1 Abstract

We have examined RNA profiles of wheat mitochondrial genes during the developmental period when seeds leave dormancy, germinate and develop into seedlings. Mitochondrial RNAs isolated from 0 hr to 6 days post-imbibition were subjected to northern analysis using coding-specific and intron-specific probes. Stable, edited mRNAs were observed in dormant seeds, and precursor RNAs were subsequently detected early in embryo germination. The respiratory chain genes *(nad7, cox1, cox2, atp6)* showed mRNA profiles which paralleled those of the ribosomal RNAs, whereas ribosomal protein genes (*rps2, rps3, rps7)* had proportionately lower steady-state mRNA levels in later stages of seedling development. The relative levels of precursors compared to respective mRNAs shifted down during development, consistent with transcription outpacing RNA processing in early stages but co-ordination being more effective several days after imbibition. In the case of multiply-split genes containing group II introns, complex patterns of splicing intermediates were observed, suggesting a lack of strict polarity of intron removal, although splicing efficiency appears to differ among introns. Excised intron RNAs typically are relatively more abundant in embryos than seedlings. These observations are consistent with a transient imbalance of RNA processing machinery at the onset of seed germination, which is a period of rapid mitochondrial biogenesis.

Key words: Mitochondria, RNA processing, group II intron, plant, development

3.2 Introduction

The production of mature messenger RNAs within plant mitochondria involves complex RNA processing events, including C-to-U type RNA editing, cis-/trans-splicing, and the maturation of transcript termini, in some cases with polycistronic transcripts being cleaved to monocistronic mRNAs. Plant mitochondrial mRNAs are typically quite stable compared to bacterial mRNAs even though they lack 5' caps or stabilizing polyA tails (as in eukaryotic cytosol mRNAs) nor do they universally have a 3' stem-loop structure to which RNA stability proteins bind (as is common in plant chloroplasts) (reviewed in Hoffmann et al., 2001). Because the mitochondrial respiratory chain complexes and ribosomes are comprised of both nuclear-encoded proteins as well as ones synthesized inside the mitochondrion, these processes must be tightly coordinated with the expression of nuclear genes, and virtually all of the RNA processing machinery is expected to be imported from the cytosol.

The mitochondrial genomes of flowering plants have approximately 35 protein coding genes, primarily ones for respiratory chain proteins in complex I-V of the electron transport chain, as well as ribosomal proteins. Eleven of these genes contain introns, although not all are necessarily present in a given plant, and all but one (namely, a group I intron in *Peperomia cox1*, Cho et al., 1998) have been classified as group II based on distinctive structural features (Michel and Ferat, 1995). Some group II introns (although none in plant mitochondria) have ribozymic activity *in vitro*, and these introns have also attracted interest because of their mobile (retroviral-like) activity and possible evolutionary link to nuclear spliceosomal introns. Both classes are spliced via two-step transesterification with the release of the excised intron as a lariat (reviewed in Michel and Ferat, 1995; Bonen and Vogel, 2001). Of the 25 introns identified as group II in plant mitochondria, 19 are located in *nad* genes encoding subunits of NADH dehydrogenase (complex I in the inner membrane), and several undergo trans-splicing because gene segments have been dispersed by mtDNA rearrangements. This bias of intron distribution raises the possibility of a regulatory role, and it is notable that in plants, complex I of the inner mitochondrial membrane is not essential for respiration. The alternative rotenone-insensitive pathway utilizes NAD(P)H dehydrogenases located on the cytosolic side of the membrane (reviewed in Rasmusson et al., 2004) and external NADH oxidation is believed to play a fundamental role in mitochondrial assembly during germination (cf. Logan et al., 2001).

We are interested in the temporal relationships among these post-transcriptional events of RNA maturation/stability, particularly for intron-containing genes. In our present study of mitochondrial RNA profiles during wheat germination and seedling development, we have selected respiratory and ribosomal protein genes, focusing on ones with varying intron content, from intronless (*atp6, cox1, rps2, rps7*), to those with a single intron (*cox1, rps3*) or multiple introns (*nad7*). Seed germination is a period of high energy demand with rapid mitochondrial biogenesis, and the hydration and activation of electron transport chain components during early stages (0 hr - 12 hrs after imbibition) result in a sharp increase in oxygen consumption. This is followed by a lag in respiration perhaps related to the accumulation of immature electron transport chain components, and then at the time of radicle emergence (typically ~18 hr post-imbibition) there is a second respiratory burst when the mitochondrial machinery is fully functional (reviewed in Bewley and Black, 1994). The respiratory response during germination is independent of mitochondrial genome copy number and the abundance of transcripts for mitochondrial proteins appears low until radicle emergence, based on studies in maize (Logan et al., 2001); however, very little is known about mitochondrial RNA processing events during this period. To this end, we have examined wheat mitochondrial transcript profiles of precursor RNAs relative to mature mRNAs during a 6-day period when seeds leave dormancy, germinate and then develop into young seedlings.

3.3 Materials and Methods

3.3.1 Isolation of wheat mitochondrial RNA

Wheat seeds (*Triticum aestivum* var. Frederick), which were kindly provided by Dr. R. Pandeya (Agriculture and Agri-Food Canada), were surface-sterilized prior to dissection or imbibition. Embryos were dissected from dry seeds (0 hr) or at various times (6 hr, 12 hr, 18 hr, 24 hr and 2 days) after imbibition with sterile water on filter paper in Petri dishes maintained in the dark at room temperature (Logan et al., 2001). The emerging radicles were typically visible after approximately 18 hrs. Seeds were also grown in vermiculite for 4 days or 6 days before the etiolated shoot tissue was harvested. After subcellular fractionation by centrifugation to recover crude mitochondria, RNA was isolated as previously described (Subramanian et al., 2001). No differences in RNA profiles were seen if embryo dissection was performed before or after

imbibition. In addition, RNA from 2-day seedlings grown in parallel on either Petri dishes or in vermiculite showed similar mitochondrial RNA profiles.

3.3.2 Northern blot analysis

Wheat mitochondrial RNA (approximately 5 μg) was electrophoresed on 1.2% agarose/formaldehyde gels and after membrane transfer, hybridizations were performed with ^{32}P-end-labelled oligomer probes specific to wheat mitochondrial genes. The hybridizations were carried out overnight at 42°C in 5% deionized formamide, 5 x SSC, 0.1% SDS and 50 μg/ml yeast tRNA, and followed by stringent washes. Gel loadings were standardized relative to approximately equivalent amounts of mitochondrial 18S rRNA for the various time points, and the same blots (as well as sister blots) were successively hybridized after radiodecay. Northern blot autoradiographs of different exposures as well as from independent experiments were subjected to densitometry using AlphaEase software (Alpha Innotech, California, USA) and signal intensities were standardized relative to mitochondrial 18S rRNA.

The synthetic oligonucleotides used as hybridization probes are given in Table 3.1. The *nad7* intron 3 lariat-junction oligomer contains 7 nt complementary to the 5' end of intron 3 and 10 nt annealing upstream of the branchpoint adenosine near the 3' end of the intron. Stringent hybridization conditions (51°C) were used to preclude formation of short 7 nt or 10 nt hybrids, in order to discriminate lariat forms from precursor RNAs or linear excised intron forms.

3.4 Results

3.4.1 Gene-specific differences in relative abundances of mitochondrial mRNAs during seedling development

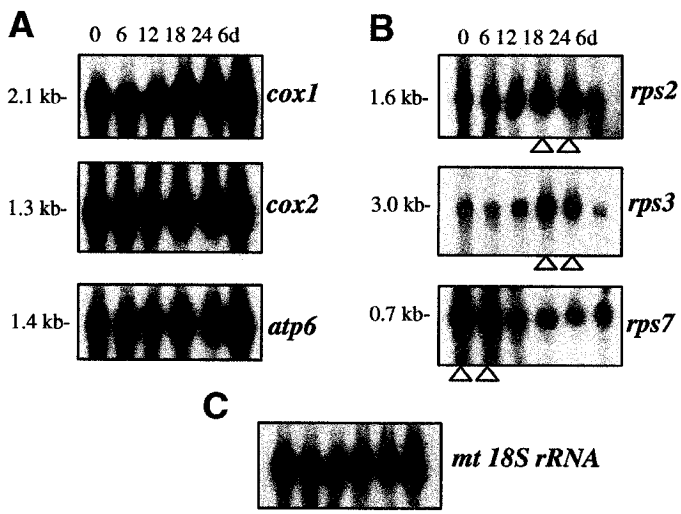
Wheat mitochondrial RNAs isolated at different stages of embryo-through-seedling development, namely 0 hr, 6 hr, 12 hr, 18 hr, 24 hr and 6 days after imbibition of dormant seeds, were subjected to northern hybridization analysis using oligomer probes specific to respiratory chain genes (*cox1, cox2, atp6)* (Figure 3.1A) and ribosomal protein genes *(rps2, rps3, rps7)* (Figure 3.1B). RNA blots were standardized for loading relative to wheat mitochondrial 18S rRNA (Figure 3.1C) and the representative data shown in Figure 3.1 are from successive re-probings of the same blot to further ensure standardization. The sizes of mature mRNAs are in agreement with those previously reported for *cox1* (Bonen et al., 1987), *cox2* (Bonen et al.,

Table 3.1 Oligomer probes used in hybridization analysis (Table adapted from Li-Pook-Than et al., 2004).

Gene	Position	Sequence	Name
cox1	exon	5'TGTAGAGCAATGTCTAGCCC'3	LB300
cox2	exon 1	5'CAGCATCACAAAGAGCGATT'3	LB13
cox2	intron 1	5'CTGTGGTGCCGATAGATTCA'3	LB287
atp6	exon	5'AGTTGCCCACATGAAGATCA'3	LB7
rps2	exon	5'TGTATAGGATCATTCGCTGG'3	LB205
rps3	exon 1	5'CAGATCAAGTCTTACCGAAA'3	LB215
rps3	intron 1	5'TCCTTATCCTCGCGTTCATG'3	LB217
rps7	exon	5'GTTCAGTTCGAGCTAGGCGGTG'3	LB37
nad7	exon 4	5'ATCGGCTTTGATCATGCCAC'3	LB92
nad7	intron 2	5'GATACGTGAGTCTTAGTGGG'3	LB226
nad7	intron 4	5'CGTGTCAGCTTAGTTATC'3	LB133
nad7	intron 3	5'ATGCATGCTTTTGTAGGGTC'3	LB64
nad7	intron 3 (lariat-junction)	5'CATGTCGCACTGGGTAG'3	LB151
18S wheat mitochondrial rRNA		5'GTGATCATTGGTCCGATGCT'3	LB211

Figure 3.1 Northern blot analysis of wheat mitochondrial RNA isolated at different stages of germination and seedling development (0 hr, 6 hr, 12 hr, 18 hr, 24 hr and 6 days after imbibition).

The same blot was consecutively hybridized with oligomer probes specific to various mitochondrial respiratory chain genes **[A]** and ribosomal protein genes **[B]**. The sizes of mRNAs are given on the left. Arrowheads in [B] show the developmental stages with most abundant RNA species; and these exhibit at least twofold higher levels in the stage with the lowest abundance. RNA loading was standardized relative to wheat mitochondrial 18S ribosomal RNA **[C]**.


A
0 6 12 18 24 6d
2.1 kb-
cox1
1.3 kb-
cox2
1.4 kb-
atp6
B
0 6 12 18 24 6d
1.6 kb-
rps2
3.0 kb-
rps3
0.7 kb-
rps7
C
mt 18S rRNA

1984), *atp6* (Bonen and Bird, 1988), *rps2* (Vaitilingom et al., 1998) and *rps7* (Zhuo and Bonen, 1993). The wheat *rps3* mRNA of ~3.0 kb is a co-transcript including *rpl16* which is located immediately downstream, as in rice (Nakazano et al., 1995), and this was confirmed using *rpl16*-specific probes (data not shown). All the other genes shown in Figure 3.1 have monocistronic mRNAs, and only one of these genes (*atp6*) is present in two genomic copies in wheat. Both *atp6* copies are actively transcribed, and their mRNA lengths differ by only 63 nt (at the 5' end) so are unresolved under these electrophoretic conditions. Both *cox2* and *rps3* each contain a single group II-type intron and their RNA maturation pathways are discussed further below.

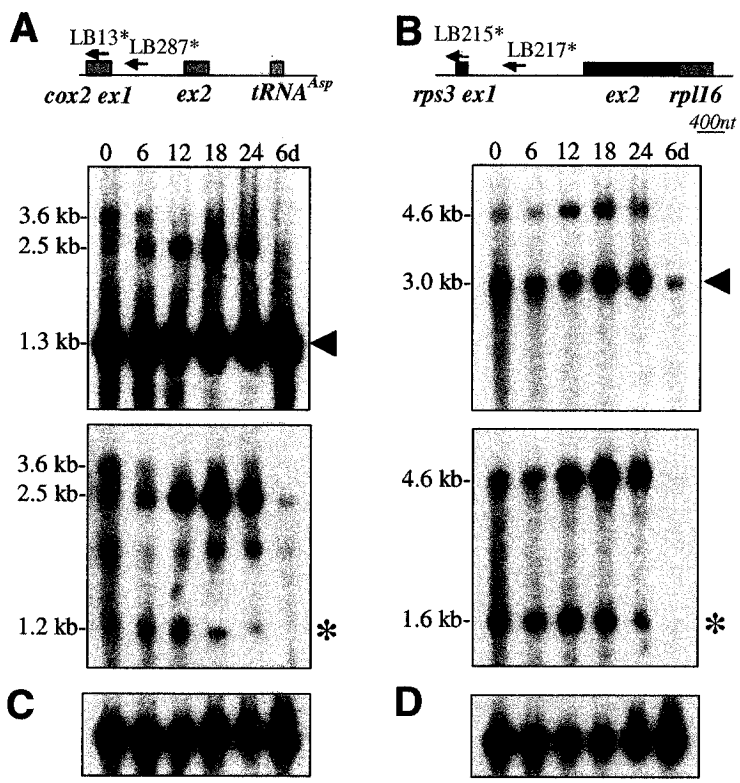
The mRNA levels for the respiratory chain genes, *cox1, cox2* and *atp6,* remain relatively constant compared to mitochondrial 18S ribosomal RNA throughout the developmental stages examined (Figure 3.1A). In contrast, *rps2* and *rps3* mRNAs show relatively higher levels at the 18 hr and 24 hr stages (Figure 3.1B, arrowheads), with a subsequent decrease in abundance in 6 day seedlings and *rps7* mRNA levels are highest in dormant seeds and at 6 hr and then drop to a rather constant level thereafter. Unlike the respiratory chain genes, all three of these ribosomal protein genes have relatively low levels of mRNAs in the 6 day seedling stage. Based on densitometric analysis of northern blots, the ribosomal protein mRNAs showed approximately 2-3 fold differences between their highest and lowest levels during development (Figure 3.1, arrowheads). The mature mRNAs which are stored in the dormant seeds appear to be fully edited at predicted sites, based on direct sequencing of RT-PCR products for *cox2* and *nad7* (data not shown), so these transcripts are potentially competent for translation, consistent with a role in "maintenance" respiration in pro-mitochondria, as well as an immediate energy source upon hydration of the seed.

3.4.2 Elevated levels of intron-containing precursors relative to respective mRNAs in early seedling development

To examine RNA splicing status during these stages, northern analysis was carried out using intron-specific probes for *cox2* and *rps3*, each of which contains a single intron of 1.2 kb and 1.6 kb, respectively (Figure 3.2). Intron-containing precursors with the expected sizes of 2.5 kb and 4.6 kb were most prominent in the 12 hr to 24 hr stages, and this is most clearly seen with intron-specific probes, where peaking occurs around 18 hr (Figure 3.2). In addition, these precursors could be detected in dry embryos and 6 day seedlings, albeit at low levels, with longer

Figure 3.2 Northern blots of wheat mitochondrial genes *cox2* and *rps3,* each containing a single intron.

Exon-specific oligomers (upper panels) and intron-specific oligomers (lower panels) for **[A]** *cox2* and **[B]** *rps3*, respectively, were hybridized consecutively on the same blot and their positions are shown by arrows in the schematics. Exons of *cox2*, *rps3* and their respective downstream genes (*tRNA*^{Asp} and *rpl16)* are denoted as colored blocks. The positions of mRNAs are shown by black arrowheads and excised introns by asterisks. RNA loading of the gels was standardized relative to wheat mitochondrial 18S ribosomal RNA **[C, D]**. The blot in panel A is the same as the one in Figure 1.


A
LB13*
LB287*
cox2 ex1
ex2
tRNA^Asp
0 6 12 18 24 6d
3.6 kb-
2.5 kb-
1.3 kb-
3.6 kb-
2.5 kb-
1.2 kb-
*
B
LB215*
LB217*
rps3 ex1
ex2
rpl16
400nt
0 6 12 18 24 6d
4.6 kb-
3.0 kb-
4.6 kb-
1.6 kb-
*
C
D

exposures of the autoradiographs. In the case of *cox2*, an additional ~3.6 kb precursor can be seen in the 0 hr and 6 hr lanes (Figure 3.2A) and its size is consistent with being a co-transcript with $tRNA^{Asp}$, which is located 1120 bp downstream of *cox2*. The *rps3* intron-containing transcript of 4.6 kb was confirmed to include the downstream *rpl16* using *rpl16*-specific oligomer probes. It should be noted that even at early developmental stages editing seems to be an early event in the RNA processing pathway, in that the spliced mRNAs for *cox2* and *nad7* were fully edited in 0 hr, 12 hr and 24 hr (based on RT-PCR sequencing data not shown) at the sites previously reported for 6 day seedlings (Bonen et al., 1994). This is similar to the editing status of maize *cox2* and *atp9* transcripts in 3 day and 7 day seedlings, although *nad3* mRNAs showed less complete editing in the 3 day stage (Grosskopf and Mulligan, 1996).

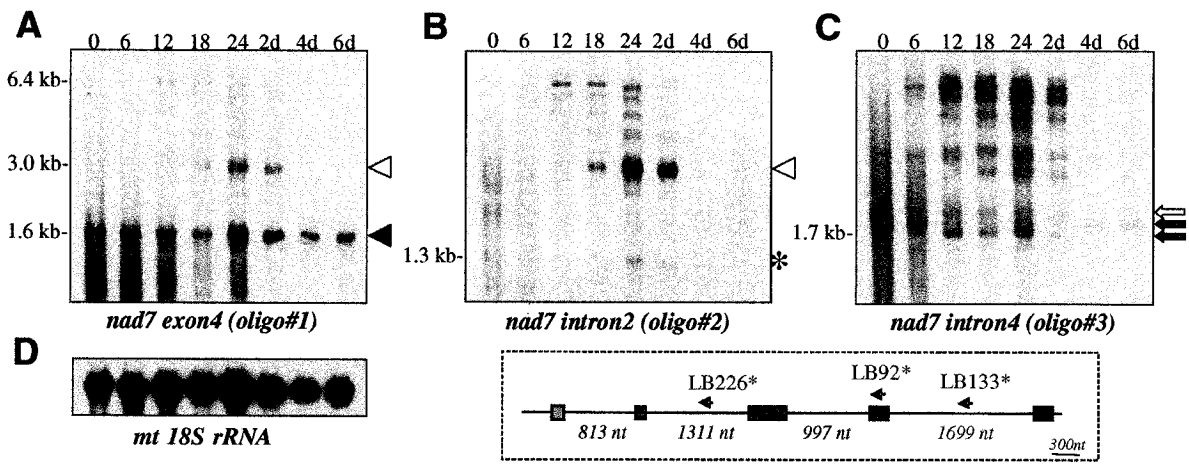
3.4.3 Partially-spliced precursors show a shift in relative abundance during development

The wheat *nad7* gene contains four group II-type introns of the sizes indicated in Figure 3.3, and this genomic region gives rise to complex RNA profiles which include unspliced precursors, partially-spliced transcripts, as well as the 1.6 kb mature mRNA (Figure 3.3A, black arrowhead) and excised intron RNA species, as shown for intron 2 (Figure 3.3B, asterisk) and intron 4 (Figure 3.3C, arrows). Mitochondrial RNAs isolated 2 days and 4 days after imbibition were included to better follow changes in stoichiometries of RNA splicing intermediates over time. The RNA profiles for *nad7* introns 1 and 3 were also examined (data not shown) and this extends our earlier observations for *nad7* expression at the 24 hr and 6 day stages (Carrillo and Bonen, 1997).

The steady-state levels of mature *nad7* mRNA roughly parallel the profiles of the ribosomal RNAs throughout the developmental stages examined (Figure 3.3A), including its presence in dormant seeds, as was the case for the other respiratory chain genes shown in Figure 1. The large *nad7* precursor transcripts (> 6 kb) can be seen at 6 – 12 hrs (cf. Figure 3.3C) and could be detected with this technique as early as 2 hrs post-imbibition (data not shown). The *nad7* precursors and partially-spliced transcripts, ranging in size from approximately 2.4 kb to 6.4 kb, are most abundant relative to mature mRNA levels during the 12 - 24 hr stages, and in this regard exhibit a temporal profile similar to that of *cox2* and *rps3* (Figure 3.2). In addition, the *nad7* profile shows a shift in size and abundance of transcripts, until precursors are very low level in the 4 - 6 day seedlings.

Figure 3.3 Northern blot analysis of wheat mitochondrial *nad7* at various developmental stages.

Blots hybridized with *nad7* exon-specific oligomer **[A]** and intron-specific oligomers for intron 2 **[B]** and intron 4 **[C].** The same blot was consecutively probed and their positions (LB92, LB133 and LB226) are shown by arrows in the schematic. The five exons are shown by boxes and intron lengths are given in nt (Bonen et al. 1994). The black arrowhead in [A] denotes the mature *nad7* mRNA and the open arrowheads [B,C] indicate the position of an abundant partially-spliced transcript. The asterisk in [B] shows the position of excised intron 2, the two orange arrows in [C] indicate excised intron 4, and the upper white arrow in [C] denotes the putative uncoupled splicing intermediate (see text). Gel loading of RNA was standardized with wheat mitochondrial 18S ribosomal RNA **[D]**.


A
0 6 12 18 24 2d 4d 6d
6.4 kb-
3.0 kb-
1.6 kb-
nad7 exon4 (oligo#1)
B
0 6 12 18 24 2d 4d 6d
1.3 kb-
nad7 intron2 (oligo#2)
C
0 6 12 18 24 2d 4d 6d
1.7 kb-
nad7 intron4 (oligo#3)
D
mt 18S rRNA
LB226*
LB92*
LB133*
813 nt
1311 nt
997 nt
1699 nt
300nt

The presence of four introns in *nad7* could theoretically lead to a large number of different partially-spliced transcripts if intron excision did not proceed in a strict order, and indeed a complex and diverse array of RNA species was seen with intron-specific probes (Figure 3.3B,C). This has also been verified by RT-PCR experiments using combinations of different intron and exon primers with RNAs from various developmental stages including 6 day seedlings (data not shown). However it does appear that certain introns are removed less efficiently than others. Most notable is the very abundant ~3.0 kb species detected with intron 2 (Figure 3.3B, open arrowhead), whose level in the 24 hr – 2 day period is even approaching that of mature *nad7* mRNA (Figure 3.3A). This RNA species corresponds to a precursor containing only intron 2, designated both on its size and lack of hybridization with any of the other three *nad7* intron probes. It is possible that this intron requires specific splicing factors that are rate-limiting in the early stages.

The *nad7* transcript profiles have a further degree of complexity due to the presence of splicing intermediates which result from the uncoupling of the two transesterification steps of splicing, that is, transcripts containing the lariat-intron linked to downstream exon(s) but lacking the upstream exon. Such species were detected by hybridization using a lariat-junction specific oligomer probe for *nad7* intron 3 (Figure 3.4B, LB151), an intron which had previously been shown to be excised as a lariat form (Carrillo et al., 2001). Based on size, there are several putative intermediate splicing by-products (Figure 3.4B, schematics at right), and as expected the lariat-junction specific probe does not hybridize to unspliced/partially-spliced higher molecular weight RNA species seen with the *nad7* intron 3 LB64 probe (Figure 3.4A). The relative abundances of these RNA species appear to shift down during development, suggesting differences in splicing kinetics or in the stabilities of splicing dead-end products. A minor *cox2* intron-containing species of ~ 2 kb (Figure 3.2A) may also fall in this category. Although such uncoupled splicing intermediates have rarely been reported for *in vivo* studies of group II intron splicing, they have been observed in spinach chloroplast and showed differences in steady state levels among genes (Kim and Hollingsworth, 1993). We also have observed a putative one for *rpl2* in germinating embryos and seedlings of rice (Subramanian et al., 2001).

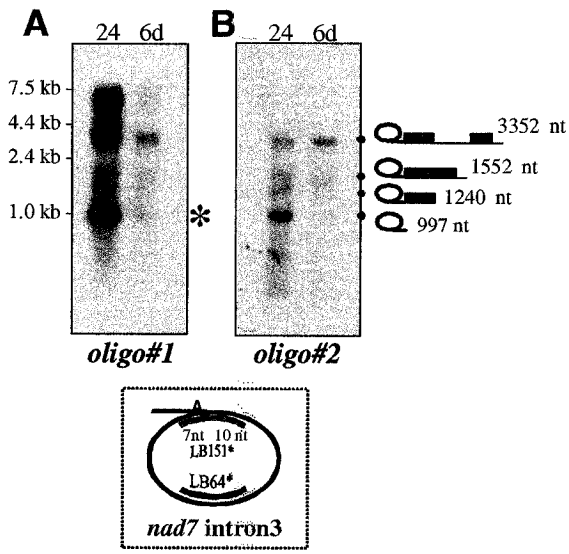
3.4.4 Variation in steady state levels of excised intron RNAs during development

Excised intron RNAs were detected for *cox2, rps3* (Figure 3.2, asterisks) and *nad7* (Figures 3.3 and 3.4, asterisks and grey arrows) during the early stages of development, including the dry embryo, although the *nad7* intron 2 and *cox2* ones appear very minor. It is worth noting that the data shown in Figures 3.2 and 3.3, were from long electrophoretic runs to better resolve high molecular weight precursor RNAs, and the excised intron species are rather diffuse compared to shorter runs (cf. Figure 3.4 and Carrillo et al., 2001). In all cases (except for *nad7* intron 4, see below), there are markedly lower levels of excised intron RNA in the 6 day seedlings compared to earlier stages, and this is also the case for the rice *rpl2* intron (Subramanian et al., 2001). This suggests that the excised introns have different degradation properties during development and may reflect availability of machinery needed for intron turnover, analogous to the debranching enzyme for nuclear spliceosomal introns. The tertiary structure of group II introns can also have a major impact on their stability and accumulation *in vivo*, as was observed by mutational analysis of a green algal mitochondrial intron using a *Chlamydomonas* chloroplast transformation system (Hollander and Kuck, 1999).

Although the relative levels of the excised *nad7* intron RNAs are typically much lower in later seedling stages, one exception is *nad7* intron 4. It is the most prominent RNA species detected with the intron 4 probe in 4 - 6 day seedlings (Figure 3.3C, middle grey arrow), a time when precursors are present at much reduced levels. Interestingly, this species appears to migrate more slowly than the excised intron in the 12 hr - 2 day stage (Figure 3.3C, lower grey arrow), but it is rather similar in mobility to faint signals in the 0 hr and 6 hr lanes. The 2 day RNA sample appears to be intermediate with both species present. A third slightly slower species visible in the 12 hr – 2 day lanes (Figure 3.3C, open arrow) is believed to be an uncoupled transesterification splicing by-product (as discussed above) which contains only the intron and downstream exon. We are further examining the physical nature of the *nad7* excised intron 4 and RT-PCR sequencing of the lariat junction region has revealed not only the expected lariat form in germinating embryos and seedlings, but also fully circularized molecules, analogous to those reported for yeast mitochondrial *cox1* intron 2 (Murray et al., 2001). We are also particularly interested in deviant plant mitochondrial introns which lack a conventional bulging adenosine (cf. Bonen and Vogel, 2001). Although group II introns are typically excised as lariats, there

Figure 3.4 Northen blot detection of lariat-type transcripts for *nad7* intron 3 in 24hr germinating embryos and 6-day seedlings.

The same blot was consecutively probed with *nad7* intron 3 LB64 **[A]** and lariat-junction specific LB151 **[B]**, as shown in the schematic below. The asterisk shows the position of the excised intron and the schematics in [B] show putative RNA species containing lariat intron 3 and various downstream sequences, with their predicted sizes in nt (see Figure 3.3 for intron lengths). This Figure is from Li-Pook-Than et al., 2004 and was performed by C. Carrillo.


A
24 6d
B
24 6d
7.5 kb
4.4 kb
2.4 kb
1.0 kb
3352 nt
1552 nt
1240 nt
997 nt
oligo#1
oligo#2
7nt 10 nt
LB151*
LB64*
nad7 intron3

have been several reports of linear forms generated *in vivo* by a hydrolytic pathway in plant chloroplasts (cf. Vogel and Borner, 2002) and bacteria (cf. Granlund et al., 2001).

3.5 Discussion

Plant mitochondrial intron-containing genes such as *nad7, cox2* and *rps3* show elevated levels of precursors relative to their respective mRNAs during wheat embryo germination, peaking about 18 - 24 hr after imbibition of dormant seeds, and then decreasing to very low steady state levels in 2 day to 6 day seedlings. This shift in relative stoichiometries suggests a lag in the efficiency of coupling between transcription and RNA processing. During early development (12-24 hr), *de novo* transcription appears to outpace RNA splicing, in contrast to 6 day seedlings where the low level of precursors relative to mature mRNA suggests a tighter co-ordination between transcription and RNA processing events.

Little is known about the regulation of mitochondrial transcription or RNA processing during early seedling development, but in maize, mRNA abundances for cytochrome oxidase and ATP synthase genes were seen to increase after 24 hr of imbibition (Logan et al., 2001). This contrasts with earlier studies where abundant *atp9* transcripts were observed in dormant maize embryos and an increase in *cox2* mRNA levels by 12 hrs (Ehrenshaft and Brambl, 1990). It is worth noting that those northern blots were normalized relative to total RNA rather than mitochondrial RNA as in the present study, which focuses on mitochondrial RNA processing pathways during early development. Analysis by real-time PCR to examine *cox1*, *atpA* and *rps10* expression during soybean cotyledon development (2 - 20 days after planting) suggested that transcripts for these mitochondrial genes peak around day 16 (Daley et al., 2003). Of particular pertinence to our findings is the observation made during the characterization of a putative wheat mitochondrial transcription factor which is nuclear-encoded (Ikeda and Gray, 1999). Its mRNA abundance increased during early germination and paralleled the profile for newly-synthesized *cox2* transcripts being detectable in the 1 hr to 15 hr period but decreasing by 25 hr, in agreement with our present observations. The *cox2* mature mRNAs maintained rather constant levels from 0 hr to 25 hr post-imbibition, similar to our data in Figure 3.1A. Transcripts observed during very early stages (<12 hr) may reflect a mixture of *de novo* transcription products and mitochondrial RNAs stored in the dormant seed.

The apparent outpacing of transcription relative to RNA splicing may reflect that RNA processing machinery is rate-limiting during early development. The nature of splicing or editing machinery in plant mitochondria is as yet unknown, but components are expected to be encoded by nuclear genes, synthesized in the cytoplasm and imported into the organelle. It is anticipated that there will be specialized as well as general splicing factors, based on studies of fungal mitochondrial and plant chloroplast group II introns (reviewed in Bonen and Vogel, 2001). This may explain the apparent differences in efficiency of intron excision (cf. Figure 3.3A and B, open arrowheads) and lack of strict polarity of intron excision in genes with multiple introns. Interestingly, a tobacco nuclear mutant (NMS1) is defective in the splicing of one particular intron, *nad4* intron1 (Brangeon et al., 2000), and in wheat this intron appears inefficiently excised compared to the two other *nad4* introns (N. Niknejad and L. Bonen, unpublished observations). In other systems, RNA binding proteins from families with pre-existing cellular functions have been recruited as group II splicing factors (cf. Ostheimer et al., 2003), and particularly interesting candidates in plant mitochondria are the RRM (RNA recognition motif) (Vermel et al., 2002) and the PPR (pentatricopeptide repeat) (Small and Peeters, 2000) classes of nucleic acid binding proteins. One of the latter has been implicated in the restoration of expression in cytoplasmic male sterile mutants in petunia (Bentolila et al., 2002). Other potential candidates are the nuclear-encoded genes, nMat-1 and nMat-2, which are homologues of the group II intronic ORF, *mat-R* located within *nad1* intron 4 in flowering plant mitochondria (Mohr and Lambowitz, 2003).

The lower steady state levels of ribosomal protein mRNAs compared to respiratory chain ones at later (post-24 hr) developmental stages raise interesting questions both with respect to expression mechanisms, as well as potential biological implications. It may reflect differences in transcription or RNA turnover, and the latter could be related to RNA structure, availability of RNA stability proteins, or tagging for degradation by short polyA tail addition (reviewed in Hoffmann et al., 2001; Gagliardi et al., 2004). A complex interplay between these events is expected, and in *Arabidopsis* mitochondria, differences in transcription rates appear counterbalanced by post-transcriptional processes controlling RNA stability (Giegé et al., 2000). It will be important to determine whether there are specific triggers controlling early, enhanced production of ribosomal machinery which is needed in order to synthesize the mitochondrial-encoded respiratory chain components for seed germination. It should be noted that the genomic

context of plant mitochondrial genes may also impact on their RNA profiles. For example, ribosomal protein genes are often co-transcribed with other genes, and in wheat mitochondria, the temporal profile of the *nad3-rps12* bicistronic mRNA during development resembles that of the respiratory chain genes shown in Figure 3.1, rather than ribosomal protein ones (data not shown).

Our observations illustrate the complexities of post-transcriptional events in wheat mitochondria during embryo germination and early seedling development. It will be of interest to dissect the processes that control mitochondrial RNA maturation, and to explore relationships between steady state mRNA levels and protein levels which lead to the functional respiratory chain complexes for energy production in plant cells.

3.5.1 Acknowledgements

We thank Dr. R. Pandeya (Agriculture and Agri-Food Canada, Ottawa) for kindly providing the wheat seeds used in this study, and Sophie Calixte for excellent technical assistance. Financial support for this research by the Natural Sciences and Engineering Research Council of Canada (L.B.) and Ontario Graduate Scholarship (J.L.P.T.) is gratefully acknowledged.

3.6 Additional information on expression profiles of wheat mitochondrial genes

3.6.1 Expression profiles of mitochondrial-encoded *cob* and *rps3-rpl16* during early wheat development

To further study the expression of other wheat mitochondrial genes, the RNA profiles of cytochrome b *(cob)*, a component of respiratory pathway complex III (Figure 1.1) were examined using northern analysis. The steady state mRNA levels (2.3 kb) of *cob* remained abundant throughout development, including at the 6d seedling stage (Figure 3.5 A). This is consistent with the mRNA patterns observed for the respiratory chain components *cox1, cox2* and *atp6* (Figure 3.1) and *nad7* (Figure 3.3). A 2.9 kb transcript, in addition to the 2.3 kb species, was also observed when a *cob* specific probe was used on northerns during the 6 hr to 2 day stages of development suggesting that primary transcript undergoes 5' or 3' termini cleavage events (Figure 3.5A, schematic).

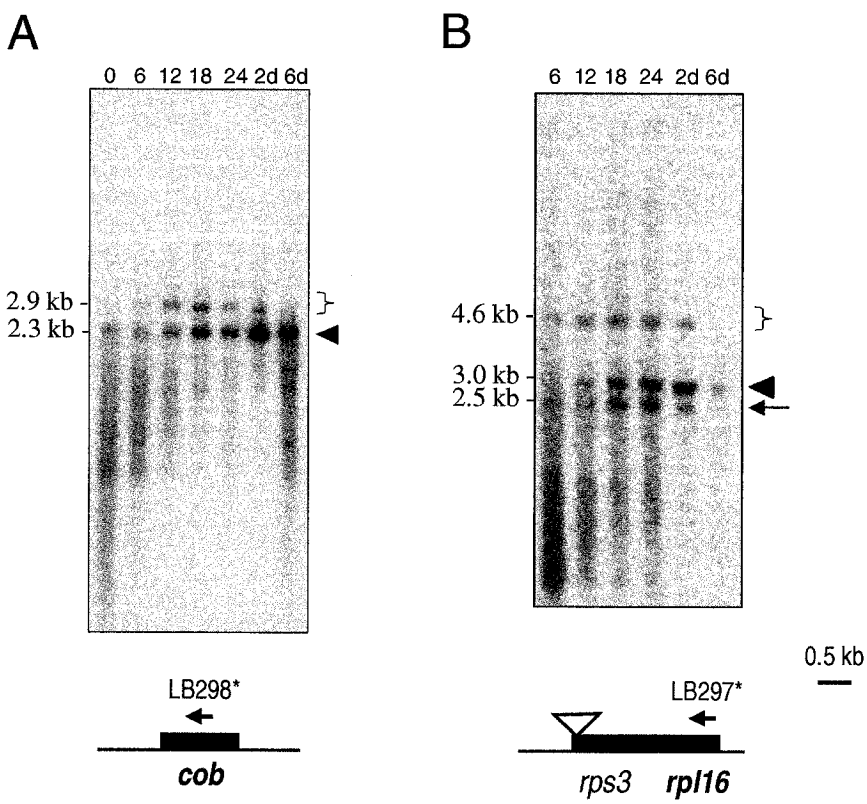
The expression profile of specifically *rpl16* (of the *rps3/rpl16*) was also included to determine if this result paralleled those of its *rps3* co-transcript (as seen in Figure 3.2). In Figure 3.2B, transcript profiles consistent with a 4.6 kb precursor and a 3.0 kb mature mRNA were revealed when an *rps3* exon 1 specific probe was utilized. In wheat mitochondria, *rps3* and *rpl16* slightly overlap by 3 nts with the putative GUG start codon of *rpl16* (Ogihara et al., 2005; Sakamuto et al., 1997). When an oligomer-probe specific to only *rpl16* (or within the 3' UTR *rps3*) was used, an additional 2.5 kb species was also detected (Figure 3.5 B, red arrow), which was not present when an *rps3* exon 1 specific oligo probe was used (Figure 3.2B). One possibility that would generate this *rpl16*-specific transcript is a RNA cleavage event. The lower steady state abundance of *rps3/rpl16* mRNAs in 6 day seedlings, in contrast to the abundant transcripts observed for *cob,* paralleled ribosomal protein expression profiles as discussed in Section 3.5 (Figure 3.1).

3.6.1 Expression profiles of mitochondrial-encoded *atp6-rps13-nad1bc* during embryo-to-seedling development of wheat

An interesting example showing the complexity of RNA processing events is the wheat mitochondrial poly-cistronic transcript *atp6-rps13-nad1b/c*, containing *atp6*, *nad1b/c* and *rps13*, as well as cis- (*nad1* intron 2) and trans-splicing introns (*nad1* intron 1 and intron 3). In wheat,

Figure 3.5 Northern blot analysis of wheat mitochondria *cob* and *rpl16* genes during embryo-to-seedling development.

Blots hybridized with **[A]** *cob* exon-specific probe, and **[B]** *rpl16* exon- specific probe. Precursor positions are indicated with parenthesis, while expected sizes and positions of mRNAs are shown with black arrowheads, respectively. The red arrow in [B] shows position of additional species at 2.5 kb. Sizes of RNA species are given at the left while schematics below each blot show exons as colored boxes, an intron as a triangle and the position and name of the oligo probe used. RNA loading was standardized relative to wheat mitochondrial 18S rRNA (data not shown).


A
0 6 12 18 24 2d 6d
2.9 kb
2.3 kb
LB298*
cob
B
6 12 18 24 2d 6d
4.6 kb
3.0 kb
2.5 kb
0.5 kb
LB297*
rps3
rpl16

atp6 and *rps13* were previously found to be co-transcribed (Takvorian et al., 1997). In my northern analysis, I observed a complex expression profile of co-transcribed *atp6-rps13-nad1b/c* as a long co-transcript of ~6.5 kb, and RNA processing intermediates, 4.6 kb and 3.6 kb (Figure 3.6 A-C). When using a *nad1* intron 3 -specific primer, multiple high molecular weight transcripts were also observed including one that matched the 4.6 kb species (data not shown). Notably, the 4.6 kb species is lower in intensity relative to the 6.5 kb band in the 2 day lane of Figure 3.6 C, while these two bands are abundant in the corresponding 2 day lanes of Figure 3.6 A-B. This is consistent with the 4.6 kb species containing a mixture of two processed *atp6-rps13-nad1b/c* transcripts, namely one with trans-spliced *nad1* intron 3 and one with only cis-spliced *nad1* intron 2 (Figure 3.6, right schematic). The 3.6 kb species is observed when a *nad1* intron 2 specific probe was used, suggesting that a cleavage event or a discrete transcription event via an alternate promoter would result in two distinct *atp6-rps13* and *nad1b/c* transcripts. Relatively abundant monocistronic 1.5 kb *atp6* mRNA is seen in the sequentially probed northern (Figure 3.6A, arrowhead), while no monocistronic transcript is observed for *rps13* (Figure 3.6B, arrowhead). The latter has a coding sequence of ~350 nts, and was formerly thought to be a silent gene in seedlings (Bonen et al., 1990). Using the *rps13* specific probe, no transcripts were detected in 4 day and 6 days seedlings (Figure 3.6B), although the products were detected in seedling using the more sensitive RT-PCR method (Takvorian et al., 1997).

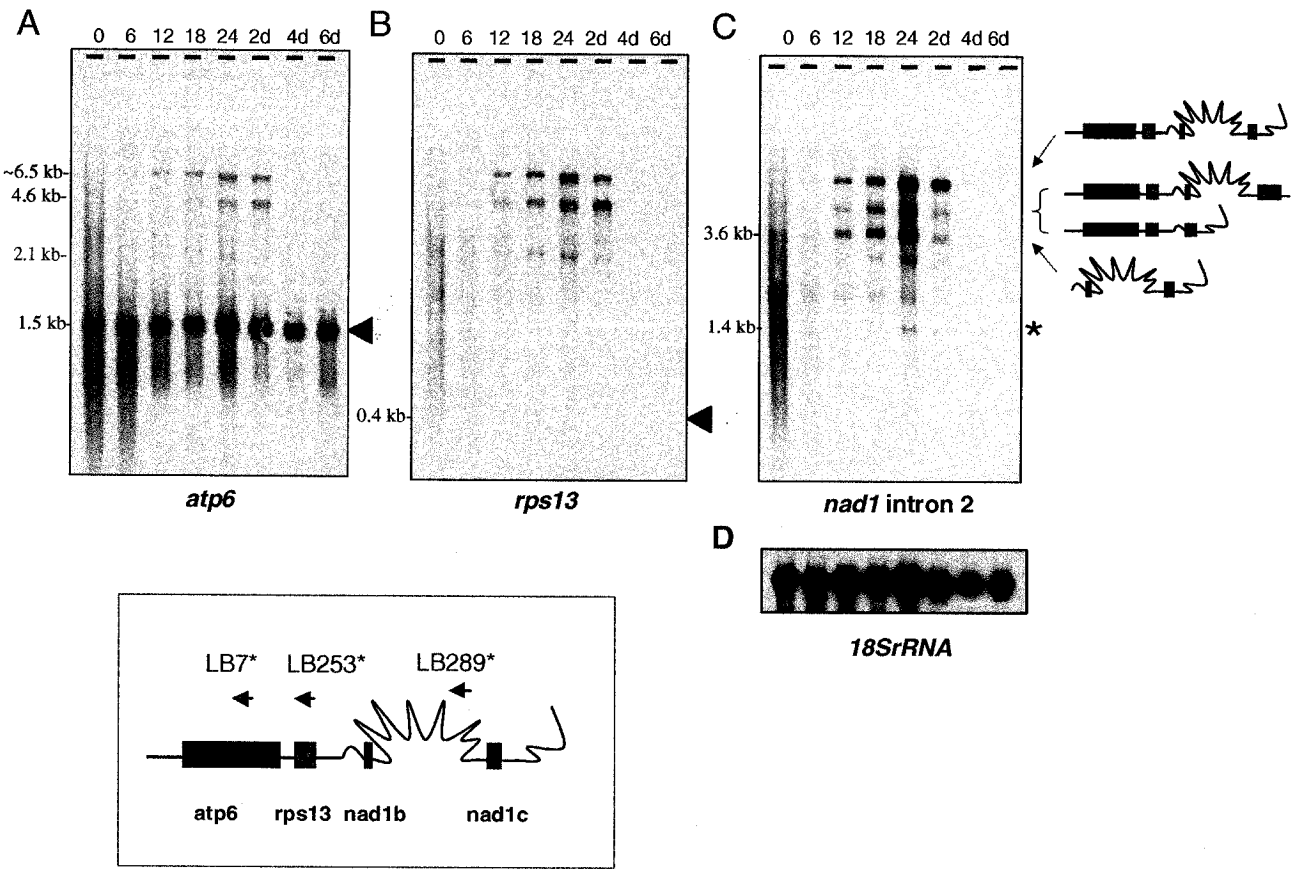
Northern analysis of the *atp6-rps13-nad1b/c* co-transcripts also revealed low levels of intron-containing precursors in both dormant seeds and 4-6 day seedlings relative to abundant precursor levels between the 12 hr to 2 day stages of development (Figure 3.6A-C). Elevated levels of *cob* precursor (2.9 kb) were also observed during the germinating seed stages from 6 hr to 2 days (Figure 3.5A), and less abundant in dry embryo and 6 day seedlings. These examples are similar to the northern analysis of *cox2*, *rps3/rpl16* and *nad7* (Li-Pook-Than et al., 2004; Section 3.3-3.4) where high molecular weight transcripts were abundant in germinating embryos, but at low levels in dormant seeds and 6 day seedlings, relative to respective mRNAs. The appearance of relatively abundant precursor transcripts at the early stages of development is consistent with *de novo* transcription of mitochondrially encoded genes. This is consistent with the observation that transcription outpaces the RNA processing events during the early stages of development, while in seedlings these events are more efficiently coordinated (Li-Pook-Than et al., 2004).

3.6.2 Excised intron profiles during early plant development

Lower steady-state levels of excised introns RNAs in seedling stages of development relative to germinating embryos were observed for rice *rpl2* intron (Subramanian et al., 2001) as well as for wheat *nad7* introns 2-4 (Carrillo and Bonen, 2001), *rps3* intron and *cox2* intron (Li-Pook-Than et al., 2004). Similarly, this pattern was also observed for *nad1* intron 2 (Figure 3.6, asterisk) and suggests a possible variation in physical form of excised intron during early germination which may be more accessible to intron turnover machinery at later stages of development than at the early developmental stages. Notably, sequencing of cloned RT-PCR products containing excised wheat mitochondrial *nad7* intron 4, revealed 12 lariats as well as 1 fully-circularized species in 24 hr germinating seeds, while in 6 day seedlings, 15 lariats and 4 circular species were found (C.Carrillo, Ph.D. thesis, University of Ottawa, 2001). In the case of the *nad1* intron 2 northern analysis shown here, it is worth noting that this gel was from a long electrophoretic run to resolve the high molecular weight precursors, but resulted in diffused excised intron species. A shorter run, as shown in the next Chapter (Figure 4.4D), revealed better resolved species corresponding to excised *nad1* intron 2.

Figure 3.6 Northern blot analysis of wheat mitochondria *atp6*, *rps13* and *nad1* intron 2 co-transcript during embryo-to-seedling development.

Blots hybridized with probes specific to **[A]** *atp6,* **[B]** *rps13* and **[C]** *nad1* intron 2. The position of the expected mRNA for *atp6* and *rps13* are shown with arrowheads and excised *nad1* intron 2 is indicated with an asterisk. RNA loading was standardized relative to wheat mitochondrial 18S rRNA **[D]**. Schematics depicting the various RNA species are shown to the right of each blot while positions of oligo probes are in the boxed schematic. Exons and introns are represented by colored boxes and wavy lines, respectively.


A
0 6 12 18 24 2d 4d 6d
~6.5 kb
4.6 kb
2.1 kb
1.5 kb
atp6
B
0 6 12 18 24 2d 4d 6d
0.4 kb
rps13
C
0 6 12 18 24 2d 4d 6d
3.6 kb
1.4 kb
*
nad1 intron 2
D
18SrRNA
LB7*
LB253*
LB289*
atp6
rps13
nad1b
nad1c

CHAPTER 4: MULTIPLE PHYSICAL FORMS OF EXCISED GROUP II INTRONS IN WHEAT MITOCHONDRIA

4.0 Rationale

Certain plant mitochondrial group II introns lack distinctive ribozymic secondary structural features involved in two-step transesterification splicing and subsequent release of the intron as a conventional lariat. What is the biochemistry of splicing of these introns that are lacking the bulged adenosine? Are there additional splicing mechanisms, such as hydrolysis, available *in vivo* in plant mitochondria? To investigate these potential splicing pathways, I examined the physical form(s) of excised introns in germinating wheat embryos, with particular interest in *nad1* intron 2, which contains a short and tight dVI helix and lacks a branchpoint adenosine. Furthermore, *nad7* intron 4, which contains a conventional dVI, was found to give rise to a minor population of fully-circularized as well as lariat molecules (Carrillo, C., Ph.D. thesis, University of Ottawa, 2001). The group II intron that contains a branchpoint adenosine, namely *cox2* intron, was first selected for this study and was expected to excise as majority lariats. However, my preliminary results on *cox2* intron surprisingly revealed a mixture of different excised forms. Therefore, *nad2* intron 4 was also included in the study and was found to excise as conventional lariats. To distinguish between lariats (or endogenously circularized) and linear forms of excised introns, RT-PCR amplification and sequencing analysis across the intronic junctions were performed and compared to similar experiments where RNA populations were pre-treated with RNA ligase (to circularize linear intron-molecules *in vitro*). In parallel, RNase H analysis was additionally used to examine the *in vivo* amounts of linear and circular RNA species.

The results reported in this section (4.1-4.6) were published as: "Li-Pook-Than, J. and Bonen, L. (2006) Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acid Research.* **34**(9), 2782-2790." Supplementary data for this section are found in Section 4.7 including northerns in which RNA was run on polyacrylamide gel in an attempt to distinguish between circular (lariats) and non-circular (linear) physical forms of excised introns.

Chapter 4: Multiple physical forms of excised group II intron RNAs in wheat mitochondria

Jennifer Li-Pook-Than and Linda Bonen

4.1 Abstract

Plant mitochondrial group II introns do not all possess hallmark ribozymic features such as the bulged adenosine involved in lariat formation. To gain insight into their splicing pathways, we have examined the physical form of excised introns in germinating wheat embryos. Using RT-PCR and cRT-PCR, we observed conventional lariats consistent with a two-step transesterification pathway for introns such as *nad2* intron *4*, but this was not the case for the *cox2* intron or *nad1* intron 2. For *cox2*, we detected full-length linear introns which possess non-encoded 3' terminal adenosines, as well as heterogeneous circular introns which lack 3' nucleotide stretches. These observations are consistent with hydrolytic splicing followed by polyadenylation as well as an *in vivo* circularization pathway, respectively. The presence of both linear and circular species *in vivo* is supported by RNase H analysis. Furthermore, the *nad1* intron 2, which lacks a bulged nucleotide at the branchpoint position, comprised a mixed population of precisely full-length molecules and circular ones which also include a short, discrete block of non-encoded nucleotides. The presence of these various linear and circular forms of excised intron molecules in plant mitochondria points to multiple novel group II splicing mechanisms *in vivo*.

4.2 Introduction

Almost all plant mitochondrial introns belong to the highly-structured ribozymic group II class and the majority are located within NADH dehydrogenase (*nad*) genes encoding complex I components of the respiratory chain. To date, a total of 25 have been identified among various flowering plants, and that includes six of the trans-splicing type (reviewed in 1,2). Group II introns are also found in organellar genes of other eukaryotes (excluding animals), as well as in a few bacterial genes (reviewed in 1,3,4). In some cases, these introns have autocatalytic activity *in vitro*, even though intron-encoded RNA maturases and/or host-encoded protein machinery are required for splicing *in vivo*. Only a few examples of self-splicing introns have been reported for

plant organelles, namely within algal chloroplast genes (5,6), but notably none in plant mitochondria.

Conventional group II introns are excised through two transesterification steps, with the 2'-OH of an unpaired adenosine 7-8 nt from the 3'end of the intron acting as the first attacking nucleophile. This is the same biochemical mechanism as for nuclear spliceosomal introns and such similarities suggest that they share a common evolutionary origin (cf. ref. 7). For both classes, the intron is liberated as a lariat, although group II ones have a much shorter tail and are highly structured with six distinctive domains (dI-dVI). From detailed biochemical and mutational analysis (primarily of fungal mitochondrial and bacterial ribozymic introns), key interactions required for proper folding and autocatalytic activity *in vitro* have been elucidated, and they include dV as part of the catalytic core and dVI which provides the first attacking nucleophile (reviewed in 3,8). In particular, branching requires an adenosine in a bulged context within the dVI helix and at a fixed distance from the 3'splice site, namely 7 nt for group IIA introns and 8 nt for group IIB (3,8-10). Mutations within the dV/dVI region, particularly at the branchpoint, can shift the first splicing step from transesterification to hydrolysis, and in some cases to incorrect selection of the 3' splice site (11,12).

In contrast to most other group II introns examined to date, those in plant mitochondria do not always conform to the conventional group II secondary structural model (reviewed in 1). The core helices can have weak base-pairing and certain key features sometimes appear missing. This raises questions about the mechanistic and evolutionary pathways taken by such "degenerate" introns. Interestingly, a plant chloroplast intron, *trnV*, which lacks a dVI bulged adenosine, is excised as a linear molecule which is generated by a hydrolytic pathway of splicing (13). The importance of the dVI structure for efficient splicing in plant mitochondria is also supported by mutational analysis of the wheat *cox2* intron, in that disruption of base-pairing resulted in loss of splicing as monitored by *in organello* splicing assays (14). In the present study, we have examined the excised physical forms of the wheat mitochondrial introns *nad2* intron 4, *nad1* intron 2 and *cox2* intron 1, designated as *nad2*i1282, *nad1*i477 and *cox2*i373, respectively in the nomenclature of Dombrovska and Qiu (cf. ref. 15). We provide evidence for multiple splicing pathways in wheat mitochondria.

4.3 Materials and Methods

4.3.1 RT-PCR and cRT-PCR analysis of intron RNAs

Mitochondrial RNA was isolated from 24 hr germinating wheat embryos (*Triticum aestivum* cv. Frederick) as previously described (16). To examine the sequences of lariat (or circularized) excised intron molecules, a synthetic oligomer mapping close to the 5' end of the intron was used as primer in cDNA synthesis with Superscript II reverse transcriptase (Invitrogen), which is able to cross 2'-5' bonds (17). Wheat mitochondrial RNA (approximately 5 μg) was denatured at 70°C for 10 min and cDNA was synthesized using 10 U/μL reverse transcriptase, 0.2 vol 5X First Strand buffer (Invitrogen), 10 mM DTT and 0.5 mM dNTPs incubated at 45°C for 2 hr. Subsequently, this primer in combination with one mapping to the 3' end of the same intron was used in PCR amplification with *Taq* DNA polymerase (Invitrogen). In parallel, cRT-PCR experiments (to circularize any linear intron molecules which contain a 5' monophosphate) were performed (18) with wheat mitochondrial RNA (approximately 5 μg) being incubated with 0.6 U/μL RNA ligase (New England Biolabs), 0.5 U/μL RNasin (Promega), 50 μg/mL BSA (Promega) and 0.1 vol 10X RNA ligase buffer (New England Biolabs) overnight at 14°C, and then recovered by ethanol precipitation prior to PCR amplification. The resulting cRT-PCR products will also include any endogenously circularized molecules already present in the RNA population. Prior treatment of RNA by heating at 70°C for 10 min or boiling for 5 min did not affect the nature or yield of products. In some experiments, Superscript III reverse transcriptase (Invitrogen) was used, although it appeared less efficient in traversing 2'-5' bonds. In the RT-PCR experiments, no product was seen in the absence of RT or when DNA was used as template with the cRT-PCR primer pairs (because of their inverted orientation). The sole exception was an abundant low molecular weight product for *nad1* intron 2, which was determined to be co-linear with DNA by sequence analysis.

RT-PCR or cRT-PCR products were subsequently gel-purified using UltraClean 15 (MoBio Laboratories Inc.) and either sequenced directly using Sequenase version 2.0 (US Biochemicals) and electrophoresed on urea-polyacrylamide sequencing gels or after being cloned into pGemT-Easy (Promega) plasmid vectors (19). Automated sequencing was performed by the Ottawa Health Research Institute DNA sequencing facility (OHRI) and the McGill University / Genome Québec Centre.

4.3.2 RNase H analysis of excised introns

Wheat mitochondrial RNA (approximately 5 μg) was incubated with intron-specific oligomers at 37°C for 30 min prior to treatment with 0.03 U/μL RNase H (Invitrogen), 1 U/μL RNasin (Promega), 0.27 μg/μL BSA (Promega) and 10 mM DTT at 37°C for 30 min (20). RNase H-treated samples were phenol-extracted and ethanol-precipitated prior to northern blot analysis. Hybridizations were performed with ^{32}P-end-labelled (T4 polynucleotide kinase) oligomer probes designed based on wheat mitochondrial sequences. The synthetic oligonucleotides used for RT-PCR, cRT-PCR, RNase H and as hybridization probes are shown in Supplementary data (online).

4.4 Results

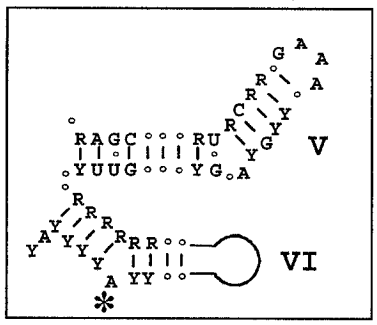
4.4.1 Nature of dVI branchpoint region of plant mitochondrial group II introns

Most plant mitochondrial introns fall in the group IIA category, in which the bulged adenosine of dVI is located 7 nt from the 3'end of the intron (Figure 4.1A, boxed) and the two unpaired terminal nucleotides (5'-AY-3') are involved in long range interactions including $\gamma - \gamma'$ with the dII/III linker (reviewed in 3,8,10,21). Figure 1B shows such an example in *nad2* intron 4, which conforms to this conventional structure and its sequence is highly conserved among plants (Figure 4.1B, inset box). In contrast, a subset of plant mitochondrial introns, of both cis- and trans-splicing types, deviate from the classical group II structure to varying extents (Figure 4.1C-H). The dVI helices of some have a lower predicted thermodynamic stability than for classical group II introns, and this can be seen by A:G mispairs in the *cox2* intron (Figure 4.1C), adjacent A:A and G:G mispairs in *nad5* intron 2 (Figure 4.1D), A:A mispairs in *nad4* intron 2 (Figure 4.1F) and to a much more pronounced extent in *nad1* intron 1 (Figure 4.1H, ref.22). It is notable that plant mitochondrial intron folding can be improved by RNA editing, which converts A-C mis-pairs to A-U pairs, although not all such mispairs in dV/dVI regions are seen to undergo editing in the introns of mitochondria (23) or chloroplasts (13), and very few of the dVI mispairs shown in Figure 4.1 fall in that category (Figure 4.1D & H, but cf. underlined lower case U in dVI of *nad5* intron 2). Even those introns possessing strongly paired dVI helices, in several cases deviate from the conventional group II format of having a purine stretch opposite a string of downstream pyrimidines (Figure 4.1D-E, G-H). The *cox2* intron contains an A:G

Figure 4.1 Secondary structure models of domains V and VI of wheat mitochondrial group II introns.

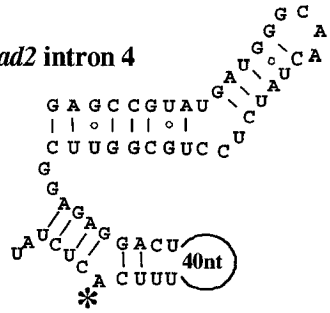
[A] Classical group IIA intron (boxed), **[B]** *nad2* intron 4 (Y14434), **[C]** *cox2* intron (X01108), **[D]** trans-splicing *nad5* intron 2 (M74158), **[E]** *nad2* intron 1 (Y14433), **[F]** *nad4* intron 2 (X57164), **[G]** *nad1* intron 2 (X57967) and **[H]** trans-splicing *nad1* intron 1 (X57967). The introns shown in (B-H) are designated as *nad2*i1282, *cox2*i373, *nad5*i1455, *nad2*i156, *nad4*i976, *nad1*i477 and *nad1*i394, respectively in the nomenclature of Dombrovska and Qiu (cf. ref. 22,15). Expected position of dVI bulged adenosine is indicated by an asterisk. Non-conserved positions are shaded and boxes below each wheat intron (AP008982) show homologous regions from the sequenced mitochondrial genomes of rice [R] (*Oryza sativa* BA000029), maize [M] (*Zea mays* AY506529), *Arabidopsis thaliana* [A] (NC_001284), *Brassica napus* [B] (AP006444), tobacco [T] (*Nicotiana tabacum* NC_006581), and sugar beet [S] (*Beta vulgaris* NC_002511), and selected introns from *Oenothera berteriana* [O] (M81725, M63033) and *Pisum sativum* [P] (22). Dots represent identical nucleotides and lower case underlined nucleotides show C-to-U editing (22,27,38,39, and Li-Pook-Than and Bonen, unpublished data). Boxed upper case C's indicate A-C mispairs which are not edited (our unpublished data). Lineage-specific differences are shown by triangles for indels and ovals for the branchpoint region. For wheat *nad2* intron 4 and *nad4* intron 2, the dV/dVI sequences are from AP008982 (cf. ref. 27, 40, and Li-Pook-Than. & Bonen, unpublished data) and differ from Y14434 and X57164, respectively.

A



B

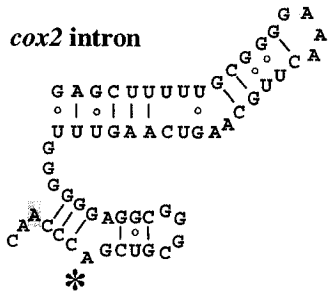
nad2 intron 4



W/R/M/A/B/T/S	5′...AGAGGACU	40-42nt	UUUCACUCUAU...3′

C

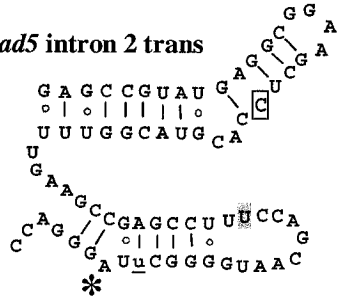
cox2 intron



W/R/M/S	5′...GGGGAGGC	GGGC	GUCGACCCAAC...3′
T	5′.....U.....3′	....	.....U.....3′

D

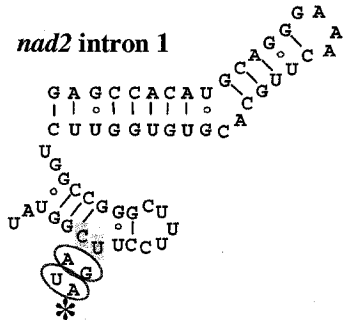
nad5 intron 2 trans



W/R/M/A/B/T/S	5′...AAGCCGAGC	15nt	GCuUAGGGACC...3′

E

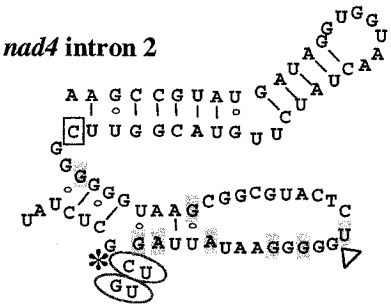
nad2 intron 1



W/R/M	5′...GCCGGG	CUUUC	CUUCGGUAU...3′
A/B/T/S	5′.....	.....	..GA.....3′
O	5′.....	.....	..AU.....3′

F

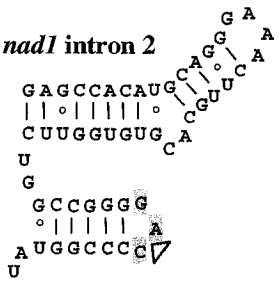
nad4 intron 2



W/M	5′...GGGGUAAG	20-21nt	UUAGGCUCUAU...3′
R	5′...A.....	22nt	.....3′
A/B	5′.....U	22nt	..UC.....3′
T	5′.....U	30nt	..U.....3′

G

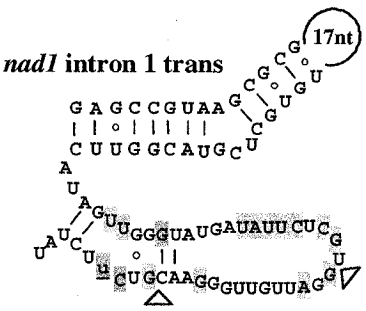
nad1 intron 2



W/R/M	5′...GCCGGGG	A	CCCCGGUAU...3′
B	5′.....	CA	.....3′
A	5′.....	CAC	.....3′
S	5′.....	-	.....3′

H

nad1 intron 1 trans



W/R/M	5′...AGUUGG	30nt	GUCuUCUAU...3′
A/B/T/S/P	5′.....AA..	31-35nt	C..C.....3′
O	5′.....AA..	35nt	C.uC.....3′

mispair immediately distal to the dVI branchpoint, as do a number of barley chloroplast introns, yet all of those were found to be excised as conventional lariats (13).

With respect to the lariat-generating branchpoint, the majority of plant mitochondrial introns possess a bulged adenosine at the expected position, however there are several interesting exceptions (Figure 4.1E-H). For example, a bulged nucleotide is completely absent from the dVI helix of *nad1* intron 2 and its short tight structure is conserved among all flowering plants examined to date. In contrast, in the very weakly-structured dVI of *nad1* intron 1, the first candidate adenosine is located at least 11 nt from the 3' end (Figure 4.1H) and considerable variation in sequence is seen among plants. In the case of *nad2* intron 1 and *nad4* intron 2 (Figure 4.1E-F), although a bulged nucleotide is present at the expected location (albeit flanked by pairs of varying strength among plants) it is not adenosine (cf. *nad4* intron 2) or adenosine in only some plants (cf. *nad2* intron 1).

Indeed the extent of sequence variation among orthologous introns from various plants in the vicinity of the branchpoint and its impact on potential base-pairing (Figure 4.1E, F & H, shaded residues and insets) is surprising given the crucial role of dVI in lariat formation, as well as the low rate of nucleotide substitution in plant mitochondrial DNA (24). This is not the behaviour expected of sequences under strong functional constraint. In this regard, we had previously noted that dV helices also tend to show weaker pairing than do classical group II introns (23). In some cases, dV structure is improved by editing, but in other instances the mispairs either remain unedited or they are not editing candidates (cf. Figure 4.1D & F, boxed C's and *cox2* intron U:U). Incidentally, *nad4* intron 2 and *nad1* intron 1 also have aberrantly long dV loops, and this is correlated with weak dVI helices and high sequence variation among plants (Figure 4.1F & H). Based on such examination of dVI features, we selected three cis-splicing introns from wheat for this study, namely, *nad2* intron 4 which has a conventional dVI (Fig.1B), *nad1* intron 2, which has a short tight dVI helix (Figure 4.1G), and the *cox2* intron, which has an A:G mispair adjacent to the branchpoint and mispairing in dV (Figure 4.1C).

4.4.2 Atypical circularized and linear introns detected by RT-PCR and cRT-PCR

To assess the presence of lariat (and/or circularized) excised introns in wheat embryo RNA populations *in vivo*, we used an RT-PCR strategy to amplify and then subsequently sequence the junction regions of *nad2* intron 4, *nad1* intron 2 and the *cox2* intron. These experiments were accompanied in parallel with RNA ligase-treated samples (cRT-PCR) to

determine if linear intron molecules might also be present in the population (Figure 4.2). We included *rps7* mRNA as a control, as it is known to be a 5' processed transcript (25) and thus can be circularized with RNA ligase. As expected, the *rps7* mRNA template resulted in an RT-PCR product of 450 bp only when pre-treated with RNA ligase (Figure 4.2A, lane 1 vs. lane 2, arrowhead).

For *nad2* intron 4, *nad1* intron 2 and the *cox2* intron, we observed RT-PCR products for both the RNA ligase-treated and untreated templates, and their respective sizes were as expected for excised introns (Figure 4.2A, lanes 4-9, arrowheads). Several larger products were also observed (Figure 4.2A, lanes 4,6,7); however only *nad2* intron 4 (Figure 4.2A, lane 4) exhibited discrete hybridization signals with intron-specific probes in Southern blot analysis (Figure 4.2B, lane 4). The latter may reflect amplification products which include neighbouring *nad2* sequences, whereas those generated with the *cox2* intron primers appear to be due to mis-priming from an unrelated genomic region. Similarly, higher molecular weight products were reported in the analysis of barley chloroplast introns (13). The apparent shift from discrete to heterogeneous high molecular weight products when RNA ligase-treated template was used with *nad2* intron 4 primers (Figure 4.2A, lane 4 vs. lane 5) may relate to contribution from circularized *nad2* trans-splicing precursor molecules. In the case of *nad1* intron 2, the abundant low molecular weight species (Figure 4.2A lanes 8,9) were confirmed by sequence analysis to be mis-primed products containing an internal section of *nad1* intron 2 that is co-linear with mitochondrial DNA (Figure 4.2B, lanes 8,9), and this is consistent with the absence of Southern hybridization signal because the intron oligomer probe maps outside that region. Comparable results to those shown in Figure 4.2 were reproducibly obtained with templates from different mitochondrial RNA preparations, different cDNA preparations and when alternative primers were used for cDNA synthesis.

To determine the precise nature of the excised intron junction, these RT-PCR products were cloned and sequenced. Such data are shown in Table 4.1. As expected, the *nad2* intron 4 junction sequence was consistent with the excised intron being a lariat (with a 6 nt tail) based on 6 out of 6 RT-PCR clones and 4 out of 4 cRT-PCR clones. At the branchpoint position, we observed mis-incorporation of adenosine instead of thymidine, as has been previously reported for templates with 2'-5' bonds (13,17,22).

In contrast, for the *cox2* intron only one out of five RT-PCR clones was consistent with a lariat model of splicing (Figure 1.3). This was unexpected as this intron contains a bulged A in

Figure 4.2 RT-PCR and cRT-PCR detection of lariat (and/or circularized) versus linear wheat mitochondrial intron molecules.

[A] RT-PCR across the excised intron junction using 24 hr wheat mitochondrial RNA as template, either pre-treated with ligase (+L; lanes 2, 5, 7 and 9) or untreated (-L; lanes 1, 4, 6 and 8) generated products of sizes 450 bp, 520 bp, 630 bp and 400 bp (denoted by arrowheads) for *rps7* mRNA, *nad2* intron 4, *cox2* intron and *nad1* intron 2, respectively. Size markers are shown in lane 3. **[B]** Southern blot of gels (shown in A, lanes 4-9) using intron-specific oligomer probes for *nad2* intron 4, *cox2* intron and *nad1* intron 2. **[C]** Schematic showing positions of primers (arrows 1 and 2) used for RT-PCR and cRT-PCR in panels [A] and [B]. Branchpoint is shown by A. Dotted lines indicate regions of amplification. **[D-F]** Direct sequencing of RT-PCR products for [D] *cox2* intron (+L), [E] *nad1* intron 2 (-L) and [F] *nad1* intron 2 (+L) using oligomers 6, 11 and 16, respectively (Supplementary data). Note that an inverted orientation of the sequencing gel is shown for [E]. Arrows show positions of the 5' terminal nt of the introns, stars highlight short non-encoded A-rich stretches, and the discrete non-encoded insert sequence is boxed in [E].

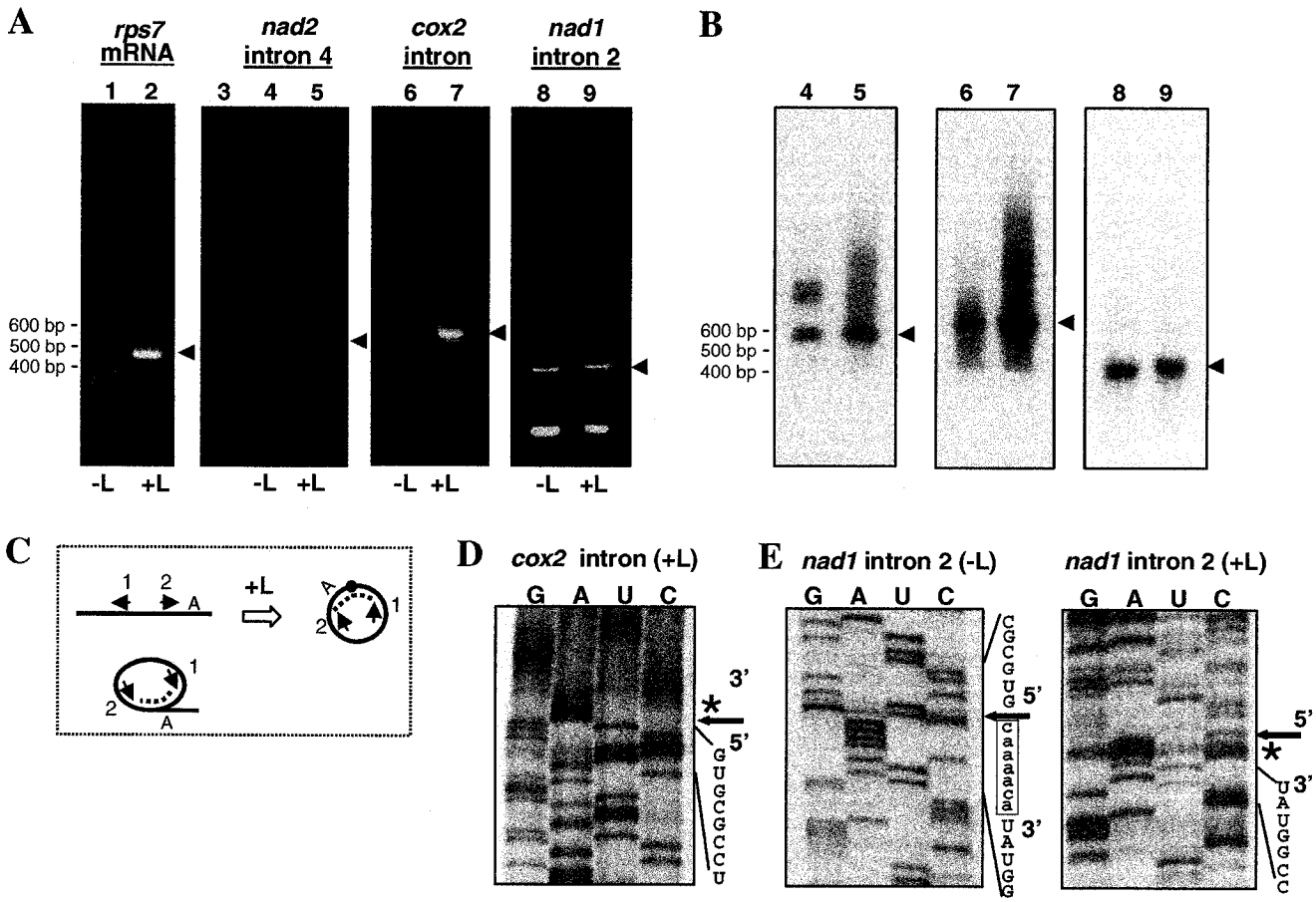

A
rps7 mRNA
nad2 intron 4
cox2 intron
nad1 intron 2
1 2
3 4 5
6 7
8 9
600 bp -
500 bp -
400 bp -
-L +L
-L +L
-L +L
-L +L
B
4 5
6 7
8 9
600 bp -
500 bp -
400 bp -
C
1 2
A
+L
1
2
A
1
2
A
D
cox2 intron (+L)
G A U C
3'
5'
E
nad1 intron 2 (-L)
G A U C
5'
3'
nad1 intron 2 (+L)
G A U C
5'
3'

Table 4.1 Sequence analysis of RNA-ligase treated (+L) and untreated (-L) RT-PCR clones of circularized or lariat intron junction regions for wheat mitochondrial *nad2* intron 4, the *cox2* intron and *nad1* intron 2.

Note that 5' terminal nts are shown on the right, 3' terminal nts on the left and non-encoded ones in the middle. Bold A or T nucleotide represents position of branchpoint A (cf. Figure 1, asterisk) and underlined nucleotides correspond to those which would be in the tail of lariat molecules. Lower case nucleotides (actgt) match the 5' terminus of *nad1* exon 3. The *nad1* intron 2 data includes the analysis of mitochondrial RNA from 2-day as well as 24-hr wheat embryos.

Clone	Ligase	3' termini	non-encoded nts	5' termini
***nad2* intron 4**		**3' termini of *nad2* intron 4**		**5' termini of *nad2* intron 4**
		...ACCCTTTCACTCTAT$^{3'}$	.....	$^{5'}$GGGCGGCCGGAA...
6/6	–L	...ACCCTTTCT.....	.....	GGGCGGCCGGAA...
4/4	+L	...ACCCTTTCT.....	.....	GGGCGGCCGGAA...
cox2 intron		**3' termini of *cox2* intron**		**5' termini of *cox2* intron**
		...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC$^{3'}$	.....	$^{5'}$GTGCGCCTCTT...
1/5	–L	...GTCAAGTTTGGGGGGAGGCGGGCG.....	.....	GTGCGCCTCTT...
1/5	–L	...GTCAA.....	.....	GTGCGCCTCTT...
2/5	–L	...GTCAAGTTTGGGGGGA.....	.....	GTGCGCCTCTT...
1/5	–L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGT.....	.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCG.....	.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGG.....	A.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGG.....	AAAA.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	A.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	AAA.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	AAAA.....	GTGCGCCTCTT...
1/8	+L	...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	AAAAA.....	GTGCGCCTCTT...
nad1* intron *2		**3' termini of *nad1* intron *2***		**5' termini of *nad1* intron 2**
		...GGACCCCGGTAT$^{3'}$	.....	$^{5'}$GTGCGCCTT...
1/9	–L	...GGACCCCGGTAT	.....	GTGCGCCTT...
5/9	–L	...GGACCCCGGTAT	ACAAAAC....	GTGCGCCTT...
1/9	–L	...AAACTTGCΔ32nt	.....	GTGCGCCTT...
2/9	–L	...GGACCCCGGTAT	actgt.....	GTGCGCCTT...
3/10	+L	...GGACCCCGGTAT	.....	GTGCGCCTT...
4/10	+L	...GGACCCCGGTAT	ACAAAAC....	GTGCGCCTT...
3/10	+L	...GGACCCCGGTAT	actgt.....	GTGCGCCTT...

dVI (Figure 4.1B). The other 4 clones displayed heterogeneous circles each missing different lengths (from 10-29 nts) from the 3' terminal region of the intron. It should be noted that approximately 50% of the clones obtained from multiple experiments had the correct insert size but sequencing revealed heterogeneous stretches of upstream and/or downstream exon sequences in addition to all or part of the intron (data not shown). This was not seen in our examination of any of the other introns. Such species might be derived from degraded RNA molecules, yet it is curious that they appear to be endogenously circularized. In this regard, evidence has been presented for a potential cryptic 5' splice site within *cox2* exon 1 of petunia (26), but that appeared to be discrete rather than heterogeneous in nature.

When the *cox2* intron cRT-PCR (RNA ligase-treated) products were cloned and sequenced to assess the presence of any linear intron RNA species in the population, we found that 6 out of 8 clones had short non-encoded polyA tracts (1 to 5 nucleotides in length) and four of these were full-length introns whereas the other two lacked 3'nucleotide stretches. In addition, one clone represented a full-length circularized intron species (end-to-end) and one simply lacked 3' terminal nucleotides. The latter presumably reflects contribution from the pre-existing non-RNA-ligase treated population of endogenously circularized forms. Notably in all cases, the precise 5' end of the *cox2* intron was observed. The presence of a mixed population of excised *cox2* intron species was corroborated by direct sequencing of the cRT-PCR products (Figure 4.2D). It can be seen that the sequence is homogeneous until the 5' end of the intron, but immediately past the junction there is a mixture of A-rich sequences (Figure 4.2D, asterisk).

The *nad1* intron 2, which contains a short tight dVI helix, showed yet a different set of features in its population of excised intron molecules and none were lariats. Rather 5/9 clones showed a discrete 5'-ACAAAAC-3' insert at the junction of full-length introns, and its relative abundance is evident in the directly-sequenced population of RT-PCR products (Figure 4.2E, boxed nts). Of the remaining 4 clones, one represented a full-length circularized intron, while two clones also contained 5'-ACTGT-3' which matches the first 5 nucleotides of *nad1* exon 2, and one clone lacked the 3' terminal 32 nt of the intron (Table 4.1). Similar results were observed for the cRT-PCR analysis, in that 3/10 clones showed precise end-to-end circularized species, and the others also contained either the non-encoded ACAAAAC block (4/10 clones) or ACTGT (3/10 clones). When the cRT-PCR products were examined directly, the sequence after

the junction (Figure 4.2F) appears to be somewhat more heterogeneous than what was seen for the *in vivo* population (Figure 4.2E).

4.4.3 Linear and circular excised introns *in vivo* supported by RNase H analysis

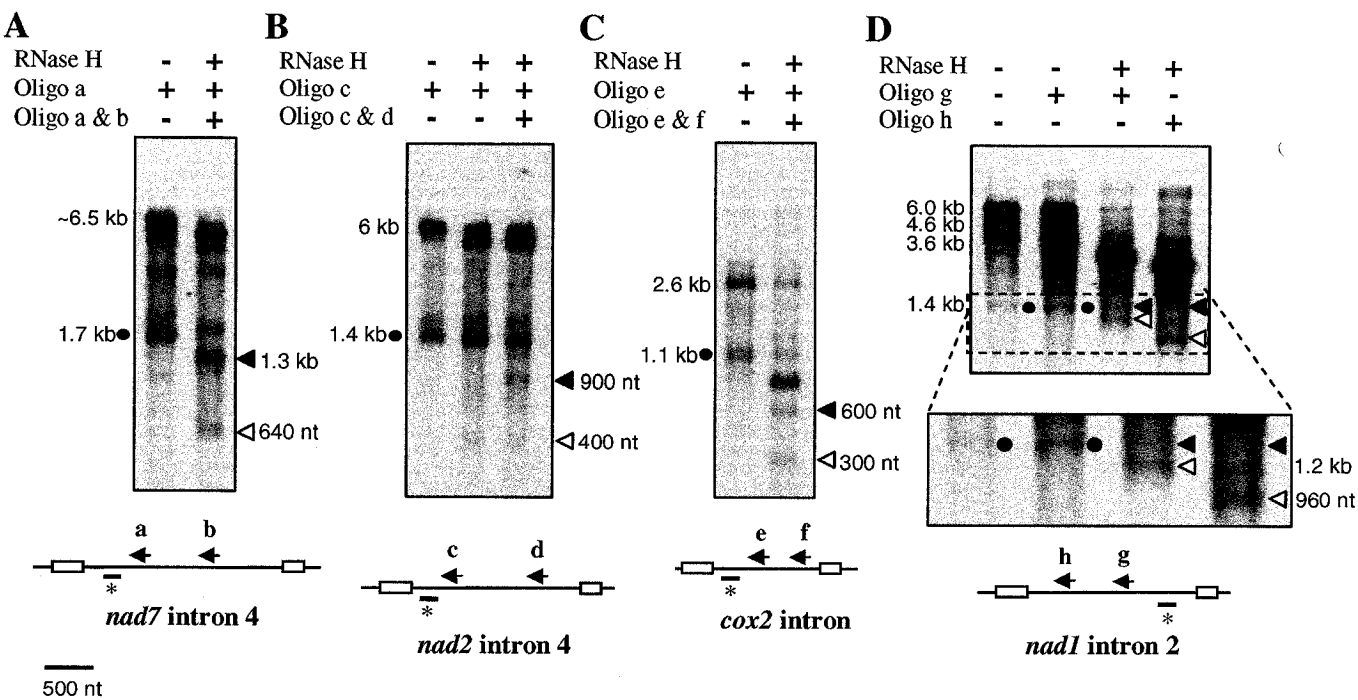
Although the above strategy provides information about the physical form of excised introns, it does not directly address their relative abundance, especially since the yields of RT-PCR products may be influenced by reverse transcriptase being less efficient at extending over the 2'-5' bond of a lariat species compared to the 3'-5' bond generated by ligation of the ends of a previously-linear molecule (cf. Figure 4.2C). To assess the relative levels of spliced lariat (and/or circularized) versus linear physical forms of introns *in vivo*, wheat mitochondrial RNA was annealed with intron-specific oligomers prior to incubation with RNase H and subsequent northern blot hybridization. In our analysis, we included *nad7* intron 4 which we had previously shown to be predominantly lariat, but which also exhibited a minor population of fully-circularized (end-to-end) intron molecules (16,27), and as expected, the vast majority appear to be lariat rather than linear (Figure 4.3A, black vs. open arrowhead). In the case of *nad2* intron 4, when we used two annealing oligomers prior to RNase H treatment, again the shift in size was consistent with a lariat (or circularized) excised intron (Figure 4.3B, black arrowhead denoting 900 nt product detected by probe), when compared to the native 1.4 kb intron (Figure 4.3B, black dot). A minor signal was detected at the position corresponding to a linear form (~ 400 nt, Figure 4.3B, open arrowhead). Also, major products of ~ 4.8 kb derived from the linear precursor of ~ 6 kb can be seen (Figure 4.3B) in addition to the precursor and intron species which were not completely digested by RNase H.

RNase H analysis of the *cox2* intron using two intron-specific primers revealed conversion of the 1.1 kb intron (Figure 4.3C, black dot) to products consistent with circular (or lariat) molecules of ~600 nt (Figure 4.3C, black arrowhead) and linear ones of ~300 nt (Figure 4.3C, open arrowhead), as well as a shift in the 2.6 kb precursor to a prominent species of ~800 nt. The relative strengths of the hybridization signals for the linear and circular forms (Figure 4.3C, open vs. black arrowheads) suggest that both species are rather equally represented in the population, and this corroborates the results obtained with RT-PCR and cRT-PCR (Figure 4.2 and Table 3.1).

In the case of wheat *nad1* intron 2, more complex northern profiles are observed because there are multiple precursors due to its co-transcription with upstream *atp6* and *rps13* genes,

Figure 4.3 Determination of *in vivo* abundance of lariat (and/or circularized) versus linear excised introns by RNase H analysis.

Wheat mitochondrial RNA (24 hr) was annealed with oligomers as shown in schematics (arrows a-h) for **[A]** *nad7* intron 4, **[B]** *nad2* intron 4, **[C]** *cox2* intron and **[D]** *nad1* intron 2, prior to RNase H treatment. Positions of oligo-probes used in subsequent northern blot analysis are shown by asterisks. Size markers are shown, black dots represent the positions of native excised introns, black arrowheads indicate positions expected for hybridizing products from lariat (or circularized) intron forms and white arrowheads for linear intron forms.


A
RNase H - +
Oligo a + +
Oligo a & b - +
~6.5 kb
1.7 kb
1.3 kb
640 nt
a b
nad7 intron 4
B
RNase H - + +
Oligo c + + +
Oligo c & d - - +
6 kb
1.4 kb
900 nt
400 nt
c d
nad2 intron 4
C
RNase H - +
Oligo e + +
Oligo e & f - +
2.6 kb
1.1 kb
600 nt
300 nt
e f
cox2 intron
D
RNase H - - + +
Oligo g - + + -
Oligo h - - - +
6.0 kb
4.6 kb
3.6 kb
1.4 kb
1.2 kb
960 nt
h g
nad1 intron 2
500 nt

resulting in RNA species of approximately 6 kb, 4.6 kb and 3.6 kb (Figure 4.3D). After RNase H treatment, when only one intron-specific primer was used (oligomer g), the signal for the excised intron of 1.4 kb shifts to a 1.2 kb product consistent with the presence of a linear species (Figure 4.3D, black dot vs. open arrowhead). This outcome was also seen when oligomer h was used, the shift being to 960 nt (Figure 4.3D, open arrowhead). In both cases, the linear species appear relatively abundant. Lariat (or circular) species when converted to Y-shaped structures by RNase H, would be expected to co-migrate with the native form under these electrophoretic conditions, but there is very little signal at that position (Figure 4.3D, black arrowhead and black dot). The high molecular weight precursors shift to the expected sizes of 2.6 kb and 2.3 kb. It is worth noting that the RNA ligase (cRT-PCR) experiments (cf. Figure 4.2A, lane 9) might be influenced by protective structures at the intron termini, such as the 3' terminal hairpin of *nad1* intron 2.

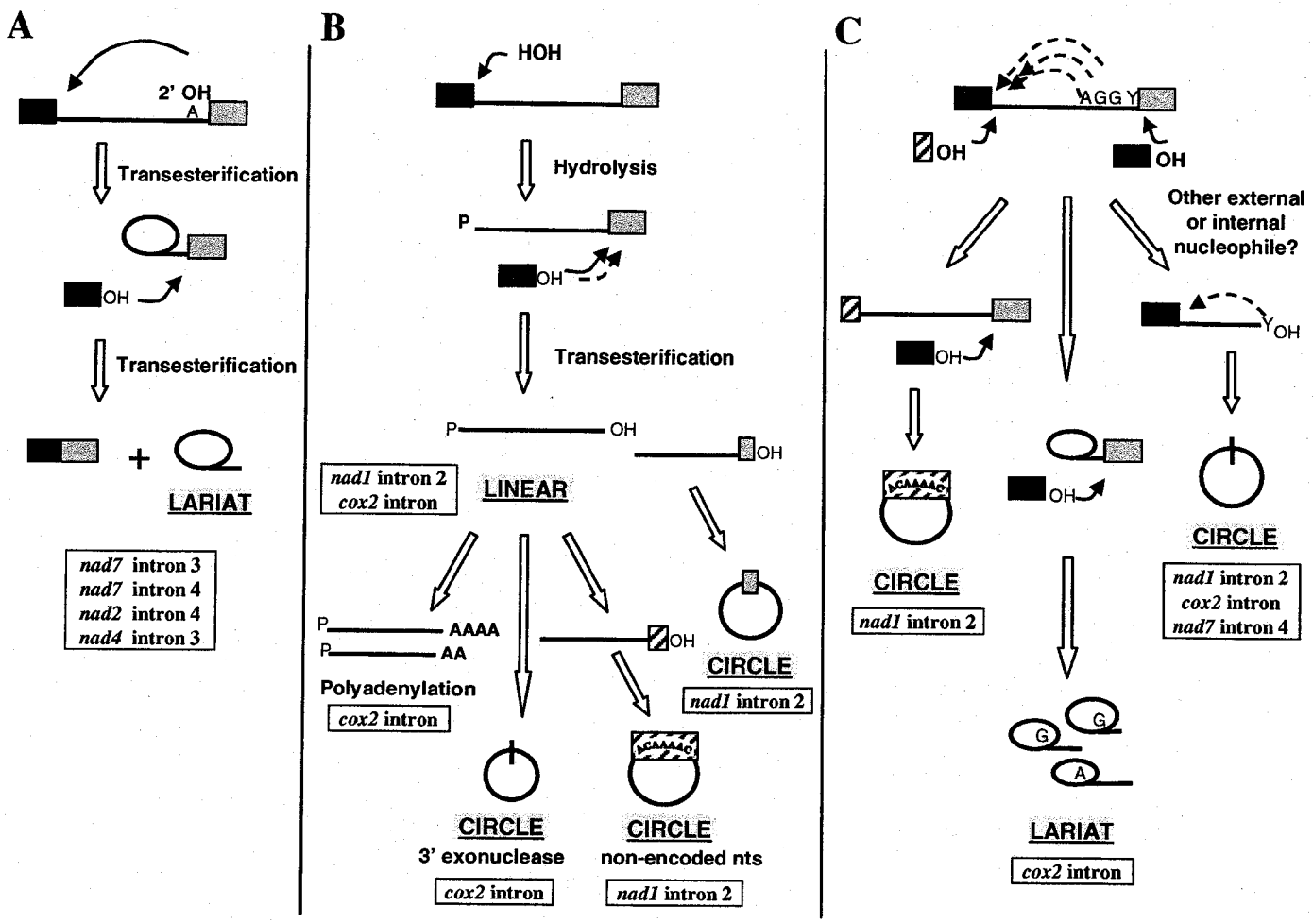
4.5 Discussion

Our analysis of the *cox2* and *nad1* excised introns in RNA isolated from germinating wheat embryos has uncovered novel physical forms which deviate from the classical lariat structure of group II introns. In contrast to introns such as *nad2* intron 4 (this work), *nad7* intron 3 (21,27) and *nad4* intron 3 (N. Niknejad & L.B. unpublished data) which possess a bulged adenosine branchpoint in dVI and are excised as lariats, our observations in the case of *nad1* intron 2 and the *cox2* intron support the presence of both circularized and linear molecules *in vivo* (Figure 4.4). Because the *cox2* intron has an apparently conventional branchpoint adenosine, it was unexpected that the lariat form was not predominant, but instead we observed heterogeneous circular excised introns which lack various lengths at the 3' terminus. Moreover, the population included full-length linear intron molecules with short, non-encoded polyA stretches.

Taken together, our data for the *cox2* intron support a hydrolytic pathway of splicing generating linear molecules which undergo either [1] polyadenylation at the 3' terminus (perhaps mediated by the bacterial-type tagging of transcripts for degradation observed in plant organelles, ref. 28) or [2] endogenous circularization after 3' exonucleolytic attack (Figure 4.4B). Linear molecules have also previously been reported for the *Streptococcus* GBS intron (which possesses a conventional branchpoint adenosine (29) and for the barley chloroplast *trnV* intron, which in addition is polyadenylated (13); however, the wheat *cox2* intron population includes

Figure 4.4 Models for plant mitochondrial group II intron splicing.

[A] Conventional two-step transesterification yielding intron lariat. Upstream and downstream exons are denoted by dark and light shaded boxes, respectively. **[B]** Hydrolytic pathway generating linear excised intron which can further undergo [1] 3' polyadenylation, [2] 3' exonuclease activity prior to endogenous circularization of heterogeneous molecules, (cf. models in ref. 8,35), [3] addition of discrete non-encoded block (hatched box) prior to endogenous circularization, or [4] incorrect 3'splice site selection (dashed arrow) resulting in a discrete exon fragment (ACTGT for *nad1* intron 2, light shaded box) being incorporated within the circular molecule. Note that the spliced exon-exon product is only shown in panel (**A**). **[C]** Hypothetical pathways of splicing via [1] external nucleophilic attack by a discrete non-encoded block (ACAAAAC, hatched box) followed by circularization of the full-length intron, [2] an internal intronic nucleophile for first transesterification step (dashed arrows), or [3] the hydroxyl group of the free upstream exon (dark shaded box) providing the first attacking nucleophile, followed by the terminal pyrimidine (Y) of the intron acting as the circularizing nucleophile (cf. "spliced exon reopening" model, ref. 8,35).


A
2' OH
A
Transesterification
OH
Transesterification
+
LARIAT
nad7 intron 3
nad7 intron 4
nad2 intron 4
nad4 intron 3
B
HOH
Hydrolysis
P
OH
Transesterification
P
OH
OH
nad1 intron 2
cox2 intron
LINEAR
P
AAAA
P
AA
OH
Polyadenylation
cox2 intron
CIRCLE
nad1 intron 2
CIRCLE
3' exonuclease
cox2 intron
CIRCLE
non-encoded nts
nad1 intron 2
C
AGG Y
OH
OH
Other external
or internal
nucleophile?
OH
OH
Y
OH
CIRCLE
nad1 intron 2
OH
CIRCLE
nad1 intron 2
cox2 intron
nad7 intron 4
G
G
A
LARIAT
cox2 intron

endogenously circularized molecules as well. This might occur through nucleophilic attack by 2'OH groups at other positions within dV or dVI to generate aberrant lariats with longer tails of varying lengths (Figure 4.4C). However, the cloned sequences correspond identically to the genomic sequence at such putative novel branchpoints, whereas typically an adenosine branchpoint is mis-read as T by reverse transcriptase (13, 17, 22). Alternatively, circularization might occur via an endogenous RNA ligase. It is also not excluded that the linear molecules which we observed are derived from the debranching of lariats with subsequent protection from exonuclease attack.

The *nad1* intron 2 appears to be present as linear and circular molecules *in vivo*, but they have quite different physical features compared to the *cox2* intron (Figure 4.4B & C). Notably, the excised introns fall in three categories: [1] precisely full-length molecules, [2] those which in addition have a discrete non-encoded ACAAAAC block and [3] ones which have the full-length intron plus ACTGT (which corresponds to the first 5 nucleotides of the downstream exon). The latter could be derived from incorrect 3' splice site selection (cf. *Bacillus cereus* ref. 30 and *Bacillus anthracis*, ref. 31) and such alternative splicing would lead to defective mRNA molecules. Because exactly the same non-encoded nucleotide blocks were consistently seen (and with or without RNA ligase treatment), these observations differ from those for the *cox2* intron and do not easily conform to a model of being generated by a polyA ragged tailing mechanism (Figure 4.4B). The endogenous linear molecules, which comprise the majority based on RNase H analysis, appear resistant to 3' exonuclease attack (unlike the *cox2* intron) and this may be related to the short tight dVI helix (reminiscent of the 3' termini of certain plant mitochondrial mRNAs, ref. 32). In all cases, the correct 5' splice site was seen both for *nad1* intron 2 as well as the heterogeneous *cox2* intron molecules. The apparent absence of exonuclease attack at the 5' end of putative linear molecules might relate to protective RNA structures or close association with proteins.

Our *nad1* intron 2 data suggest a combination of splicing mechanisms which exploit various internal or external nucleophiles to explain the discrete non-encoded block (Figure 4.4B & C, hatched box). The latter would be analogous to group I intron splicing where an external guanosine acts as the attacking nucleophile (reviewed in 33). The origin of the non-encoded ACAAAAC block is unknown. If it is added to the 3' end of the intron, it seems unlikely to be mediated by tRNA nucleotidyl transferase or an RNA-directed RNA polymerase, as was

suggested for tobacco chloroplast *ndhD* mRNA processing (34), since there is no obvious template. Other models consistent with a precisely full-length circularized intron (which we observed at low levels for *nad1* intron 2, *nad7* intron 4 and the *cox2* intron) include reverse splicing or a "spliced exon reopening" pathway (cf. yeast mitochondrial aI2 intron circles *in vivo*, ref. 35) whereby the 3' OH of the liberated 5'-exon acts as first attacking nucleophile *in trans* and then the hydroxyl group of the intron terminal pyrimidine attacks the 5' splice site (Figure 4.4C). Interestingly, in plant mitochondria, uncoupled splicing intermediates have been observed to be relatively abundant in some cases (16) and thus might contribute to such a pathway.

The apparent loss of classical group II features of mitochondrial introns in the flowering plant lineage over evolutionary time is supported by comparisons with homologues from early-diverging lineages of land plants. For example, *nad1* intron 1 in *Isoetes lacustris* (22,36) and the *cox2* intron in *Acorus calamus* (37) have conventional dVI structures. Moreover, the dynamic nature of plant mitochondrial genomes (reviewed in ref. 2) can have impact on intron sequences, such as the genomic rearrangements which have led to trans-splicing. It also appears that gene conversion (or swapping) events can occur between group II core structures. For example in Figure 1, it can be seen that dV and part of dVI of *nad1* intron 2 and *nad2* intron 1 are identical. This also underscores the plasticity and tolerance that the splicing machinery must have in light of such deviation from conventional structure.

The mixture of physical forms of excised introns observed in this study suggests that multiple biochemical mechanisms of splicing occur in plant mitochondria, such as a hydrolytic pathway as well as unusual circularization events. Their relative contributions appear to vary among introns and it will be interesting to learn if there are developmental effects, perhaps due to differences in the availability of nuclear-encoded machinery especially early in germination when seeds leave dormancy (16). In this regard, it is interesting that the steady state levels of excised introns show marked differences during germination, typically being present at higher levels in embryos than in seedlings. The apparently "degenerate" nature of the folding of these group II introns raises the possibility that additional proteins and/or small RNAs have been recruited for the splicing process in plant mitochondria.

4.5.1 Acknowledgements

We thank C. Carrillo, N. Niknejad, S. Subramanian and J. Crosthwait for their earlier work on wheat and rice mitochondrial introns in our laboratory. We also thank Dr. R. Pandeya (Agriculture and Agri-Food Canada, Ottawa) for kindly providing the wheat seeds used in this study. The financial support of the Natural Sciences and Engineering Research Council of Canada (L.B.) and Ontario Graduate Scholarship (J.L.P.T.) is gratefully acknowledged.

4.6 References

1. Bonen,L. and Vogel,J. (2001) The ins and outs of group II introns. *Trends Genet.*, **17**, 322-331.
2. Knoop,V. (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.*, **46**, 123-139.
3. Michel,F. and Ferat,J.L. (1995) Structure and activities of group II intron. *Annu. Rev. Biochem.*, **64**, 435-461.
4. Lambowitz,A.M. and Zimmerly,S. (2004) Mobile group II introns. *Annu. Rev. Genet.*, **38**, 1-35.
5. Sheveleva,E.V. and Hallick,R.B. (2004) Recent horizontal intron transfer to a chloroplast genome. *Nucleic Acids Res.*, **32**, 803-810.
6. Odom,O.W., Shenkenberg,D.L., Garcia,J.A., Herrin,D.L. (2004) A horizontally acquired group II intron in the chloroplast psbA gene of a psychrophilic *Chlamydomonas*: in vitro self-splicing and genetic evidence for maturase activity. *RNA*, **10**, 1097-1107.
7. Seetharaman,M., Eldho,N.V., Padgett,R.A. and Dayie,K.T. (2006) Structure of a self-splicing group II intron catalytic effector domain 5: Parallels with spliceosomal U6 RNA. *RNA*, **12**, 235-247.
8. Lehmann,K. and Schmidt,U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol.*, **38**, 249-303.
9. Chu,V.T., Adamidi,C., Liu,Q., Perlman,P.S. and Pyle,A.M. (2001) Control of branch-site choice by a group II intron. *EMBO J.*, **20**, 6866-6876.
10. Toor, N., Hausner, G. and Zimmerly, S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA*, **7**, 1142-1152.
11. Podar,M., Chu,V.T., Pyle,A.M. and Perlman,P.S. (1998) Group II intron splicing in vivo by first-step hydrolysis. *Nature*, **391**, 915-918.
12. Boulanger,S.C., Faix,P.H., Yang,H., Zhuo,J., Franzen,J.S. Peebles,C.L. and Perlman,P.S. (1996) Length changes in the joining segment between domains 5 and 6 of a group II intron inhibit self-splicing and alter 3' splice site selection. *Mol. Cell. Biol.*, **16**, 5896-5904.
13. Vogel,J. and Borner,T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.*, **21**, 3794-3803.
14. Farre,J.C. and Araya,A. (2002) RNA splicing in higher plant mitochondria: determination of functional elements in group II intron from chimeric *cox* II gene in electroporated wheat. *Plant J.*, **29**, 203-213.
15. Dombrovska,O., Qiu,Y.L. (2004) Distribution of introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implications. *Mol. Phylogenet. Evol.*, **32**, 246-263.
16. Li-Pook-Than,J., Carrillo,C. and Bonen,L. (2004) Variation in mitochondrial transcript profiles of protein-coding genes during early germination and seedling development in wheat. *Curr. Genet.*, **46**, 374-380.
17. Vogel,J., Hess,W.R. and Borner,T. (1997) Precise branch point mapping and quantification of splicing intermediates. *Nucleic Acids Res.*, **25**, 2030-2031.

18. Vogel,J. and Hess,W.R. (2001) Complete 5' and 3' end maturation of group II intron-containing tRNA precursors. *RNA*, **7**, 285-292.
19. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. Zangar,R.C., Hernandez,M., Kocarek,T.A. and Novak,R.F. (1995) Determination of the poly(A) tail lengths of a single mRNA species in total hepatic RNA. *Biotechniques*, **18**, 465-469.
21. Hollander,V. and Kuck,U. (1999) Group II intron splicing in chloroplasts: identification of mutations determining intron stability and fate of exon RNA. *Nucleic Acids Res.*, 27, 2345-2353.
22. Carrillo,C., Chapedelaine,Y. and Bonen,L. (2001) Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol. Gen. Genet.*, **264**, 595-603.
23. Carrillo,C. and Bonen,L. (1997) RNA editing status of *nad7* intron domains in wheat mitochondria. *Nucleic Acids Res.*, **25**, 403-409.
24. Wolfe,K.H., Li,W.H. and Sharp,P.M. (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl Acad. Sci. USA*, **84**, 9054-9058.
25. Zhuo,D. and Bonen,L. (1993) Characterization of the S7 ribosomal protein gene in wheat mitochondria. *Mol. Gen. Genet.*, **236**, 395-401.
26. Pruitt,K.D. and Hanson,M.R. (1991) Splicing of the *Petunia* cytochrome oxidase subunit II intron. *Curr. Genet.*, **19**, 191-197.
27. Carrillo,C. (2001) RNA splicing and editing of group II introns in flowering plant mitochondria. *Ph.D. thesis*, University of Ottawa, Ottawa, Canada.
28. Galgliardi,D., Stepien,P.P, Temperley,R.J., Lightowlers,R.N. and Chrzanowska-Lightowlers,Z.M.A. (2004) Messenger RNA stability in mitochondria: different means to an end. *Trends Genet.*, **20**, 260-267.
29. Grandlund,M., Michel,F. and Norgren,M. (2001) Mutually exclusive distribution of IS1548 and GBSi1, an active group II intron identified in human isolates of group B *Streptococci. J. Bacteriol.*, **183**, 2560-2569.
30. Tourasse,N.J., Stabell,F.B., Reiter,L. and Kolsto,A.B. (2005) Unusual group II introns in bacteria of the *Bacillus cereus* group. *J. Bacteriol.* 187, 5437-5541.
31. Robart,A.R., Montgomery,N.K., Smith,K.L. and Zimmerly,S. (2004) Principles of 3' splice site selection and alternative splicing for an unusual group II intron from *Bacillus anthracis*. *RNA*, **10**, 854-862.
32. Dombrowski,S., Brennicke,A. and Binder,S. (1997) 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.*, **16**, 5069-5076.
33. Fedor,M.J. and Williamson,J.R. (2005) The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell Biol.*, **6**, 399-412.
34. Zandueta-Criado,A. and Bock,R. (2004) Surprising features of plastid ndhD transcripts: addition of non-encoded nucleotides and polysome association of mRNAs with an unedited start codon. *Nucleic Acids Res.*, **32**, 542-550.
35. Murray,H.L., Mikheeva,S., Coljee,V.W., Turczyk,B.M., Donahue,W.F., Bar-Shalom,A. and Jarrell,K.A. (2001) Excision of group II introns as circles. *Mol. Cell*, **8**, 210-211.
36. Malek,O. and Knoop,V. (1998) Trans-splicing group II introns in plant mitochondria: the complete set of cis-arranged homologs in ferns, fern allies, and a hornwort. *RNA* **4**, 1599-1609.
37. Albertazzi,F.J., Kudla,J. and Bock,R. (1998) The *cox2* locus of the primitive angiosperm plant *Acorus calamus*: molecular structure, transcript processing and RNA editing. *Mol. Gen. Genet.*, **259**, 591-600.
38. Farré,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**, 959-967.
39. Wissinger,B., Schuster,W. and Brennicke,A. (1991) Trans splicing in *Oenothera* mitochondria: *nad1* mNAs are edited in exon and trans-splicing group II intron sequences. *Cell* **65**: 473-482.

40. Ogihara,Y., Yamazaki,Y., Murai,K., Kanno,A., Terachi,T., Shiina,T., Miyashita,N., Nasuda,S., Nakamura,C., Mori,N., Takumi,S., Murata,M., Futo,S. and Tsunewaki,K. (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res.*, **33**, 6235-6250.

4.7 Additional data on the physical forms of excised introns in wheat mitochondria

4.7.1 Preparatory experiments using RNA ligase for circularized RT-PCR (cRT-PCR) on rps7

The RNA ligase method was first applied to examine the termini of wheat mitochondrial *rps7* transcripts. This *rps7* mRNA was subsequently used as a positive control for circularized RT-PCR analysis when examining excised introns (as observed in Section 4.3). Zhuo and Bonen (1993) showed the presence of minor high molecular weight species and several abundant smaller molecular weight mRNAs were consistent with RNA processing events generating 5' mono-phosphates. Consequently, RNA ligase is effective for circularization of the *rps7* molecules (Figure 4.5A-C). The ligated circularized *rps7* mRNA termini were examined by direct sequencing of RT-PCR products at 6 hr and 4 day stages of development. Based on northern analysis, *rps7* mRNAs levels decreased relative to *18S rRNA* during embryo-to-seedling development (Figure 3.1, Li-Pook-Than et al., 2004). These two stages were selected based on the hypothesis that populations of 4 day *rps7* transcripts (which were less abundant in the seedling stages, Figure 3.1) would have modified termini at different stages of development (tagging for turnover, Gagliardi et al., 2004). Similar sequence profiles were observed at the 3'-5' ligated junctions in 6 hr or 4 day developmental stages. Instead, the sequences are relatively cleanly read through the 3' terminus and past the 5' end (Figure 4.5B, red arrow), revealing that the majority of cleaved transcripts ended ~116 nt downstream of the stop codon and began ~165 nt upstream of the start codon (Figure 4.5C, black arrows). However, the slightly mixed appearance (overlapping of potentially 2-5 different leader sequences) past the 5' UTR region corroborates the mixed population of mapped 5' ends reported for *rps7* (Zhuo and Bonen, 1993). These results also suggest that a major cleavage event would generate the 165 nt 5' UTR of the *rps7* transcript.

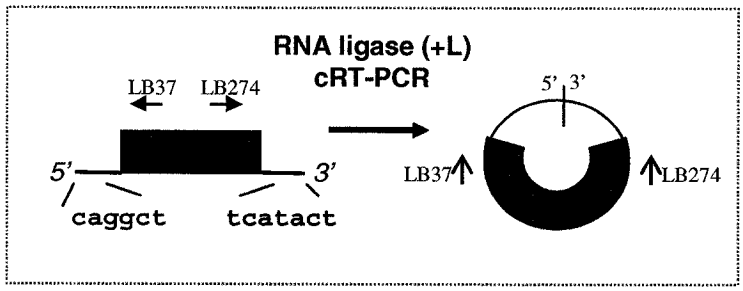
4.7.2 RNase H analysis suggests a subpopulation of linear excised wheat *rps3* intron and *nad4* intron 3

To examine the excised physical forms of a broader range of wheat mitochondrial group II introns, *rps3* intron and *nad4* intron 3 were included in the RNase H study (at the 24 hr stage of development). Similar to the RNase H results of *cox2* intron and *nad1* intron 2 (Figure 4.3), a mixed *in vivo* population of both lariat (or circularized) and linear molecules was observed for

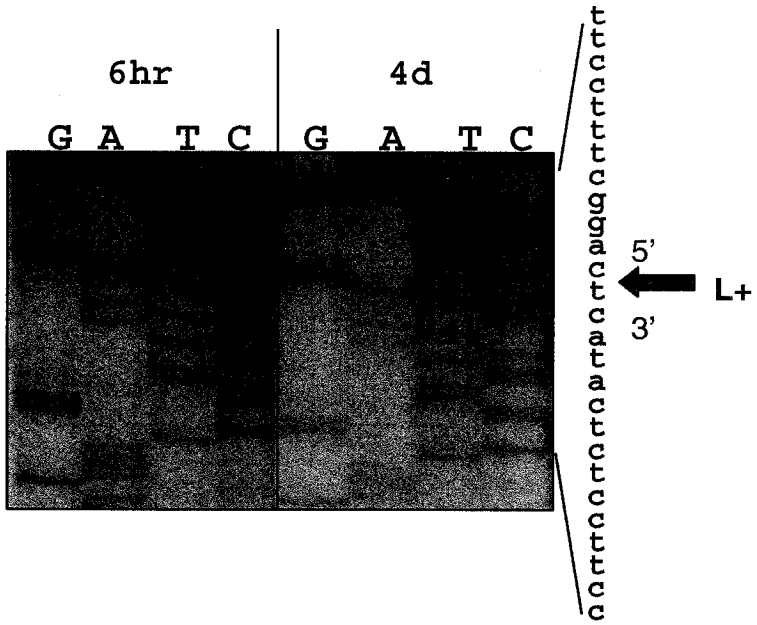
Figure 4.5 Circularization of *rps7* processed transcripts using RNA ligase and cRT-PCR analysis.

[A] Schematic depicting position of *rps7* primers used in cRT-PCR analysis and for sequencing. **[B]** Direct sequencing of cRT-PCR products from 6 hr (left panel) and 4 day (right panel) wheat plants. Arrow to right shows junction position of ligated (+L) 3' and 5' *rps7* ends. **[C]** Complete sequence of wheat mitochondrial *rps7* (X67242) with open arrowheads showing positions of 5' and 3' ends based on S1 nuclease mapping (Zhuo and Bonen, 1993). The black arrowheads show processed *rps7* transcript positions, that are corroborated by this experiment, and boxed sequence represent the coding sequence of *rps7*. Blue and green sequences represent 5' and 3' UTR regions, respectively, while the 't' nucleotide is in aqua to denote that its source may be from either the 3' or 5' end of the *rps7* transcript.

A



B



C

agatcttggc tcgccctaca aaaaaacaac tatatgccct ataaacttgc ttcttcgaat ctcgaaataa
catatagaaa gtgtttctgt gatgagacca ttctcgatcg ataaatgcga taggagccta tgtgtttcac
gggcagccgg ccaatagggg aaggttgtga atccggcgaa ccgaccgctt taagaacgtg attgtcttct
ctcactcagt tatcaattga caaagatgac ccattctttt tcgccctaaa aacgacaatg gtatggtact
ttttcttcaa atcgagattg tgtgggtgtc cgctcatgtt cacgttacat gctaa**atcag gcttccttg**
gaaaaaccaa ggacaacccc tatctcagtc tcctttcctt tctcttttcg ggagcagagc tgaaaaagat
ggacagtaac gatcgcgtaa tatcaattta tcggcctcgt catcgaaagc ggcttccaat tgctcggaaa
ttctcagcta tatggggggac tttgatggtg agcaaaaaga attgatcaag aaattggtaa actttcgcat
gatcgatggt aaaagaacga gagttcgtgc_tattgtttat aaaacttttc accgcctagc tcgaactgaa
cgcgatgtaa taaaacttat ggttgacgcc gtagataata taaagccaat atgcgaagtg gtcaaagtag
gagtcgcagg tactatttat gatgttcctg ggattgtagc cagggatcgt caacaaacct tagctattcg
ttggatcctt ggagcagctt tcaaacgacg tataagctac aggataagct tagagaaatg ttcatttgct
gagatactgg atgcttaccg aaagagggga atttcacgta agagaaggga gaatcttcat ggactggctt
ccaccaatcg gagtttcgcg catttcagat ggtggtaaag **tgataccaca taaggagctc ttcctcattc**
agtcatactc aacaaaagta agaaatgttt gacccggatc ctttttttcat cttcatatag aaagaaaatc
ggccttcctc atactccccc ttcattcata gagttggag gaatccacaa gaggcctgcc cgttcataat
tgcataaaag aaccattctt

wheat mitochondrial *rps3* intron (Figure 4.6A, black vs white arrowhead; in collaboration with O. Antoniuk, NSERC summer student, 2003). These results paralleled those of *nad7* intron 4 (Figure 4.3) in that the wheat mitochondrial *rps3* intron also contains the classical branchpoint region for the first step of splicing and lariat formation (Figure 4.6C) while the unexpected detection of linear excised molecules suggest the occurrence of supplemental splicing mechanisms *in vivo* (Figure 4.5). For rice mitochondrial *rps3* intron, one sequenced RT-PCR cloned product revealed a lariat excised species (O. Antoniuk, NSERC summer student, 2003).

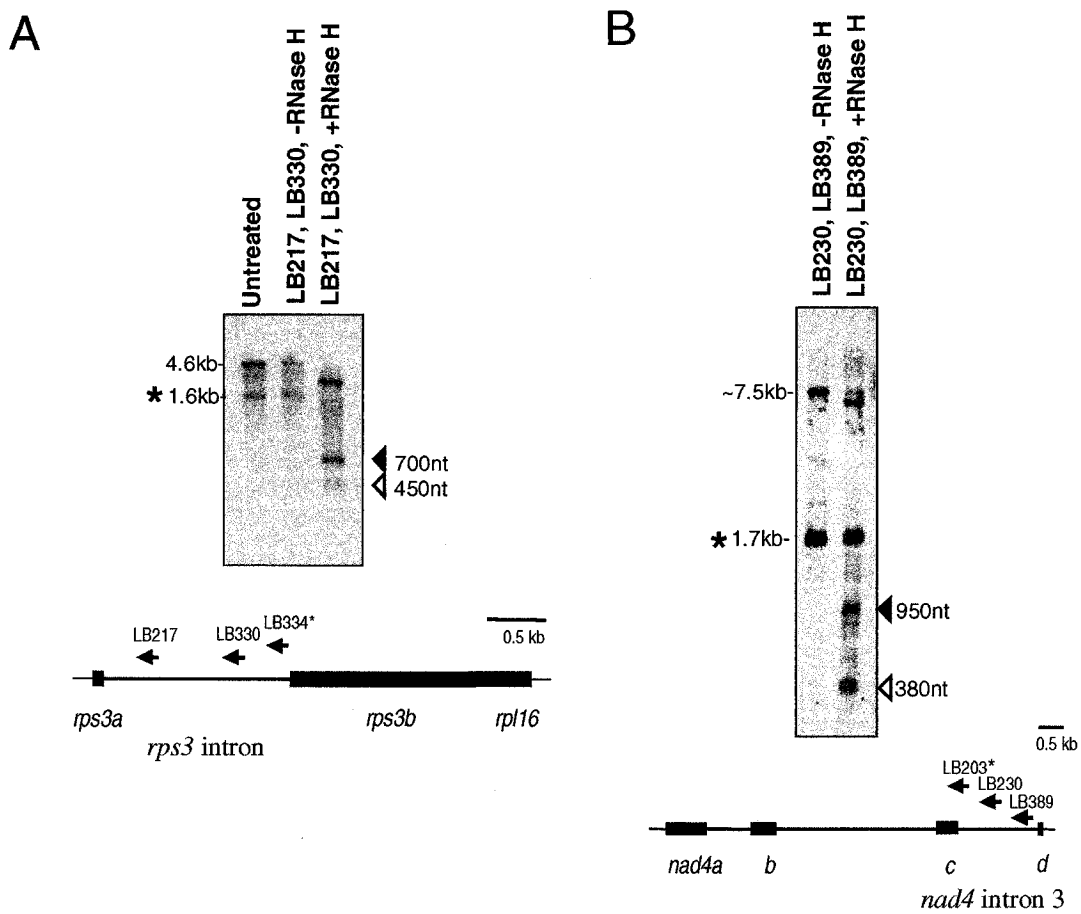
In the case of *nad4* intron 3, it had previously been observed via RT-PCR sequencing across the junction, that 3/5 clones corresponded to lariat excised introns, while the remaining 2 clones lacked one nucleotide at the lariat junction (N.Niknejad, M.Sc. thesis, University of Ottawa, 2003). The latter phenomenon is consistent with the pausing and misincorporation of a nucleotide by the reverse transcriptase enzyme due to 2'-5' nature of the lariat bond as previously reported in Vogel et al. (1997). RNase H analysis of excised *nad4* intron 3, revealed ~50% of the population as linear molecules of 370 nts (Figure 4.6B). The transcripts and excised intron were incompletely digested (+ RNase H lane), suggesting that the secondary and tertiary conformation of *nad4* intron 3 may limit RNase H analysis. Notably, *nad4* intron 3 as well as *rps3* intron, each contain a classical branch point region (Figure 4.6C), but like the *cox2* intron, the presence of abundant linear molecules suggests that an additional form of splicing, such as first-step hydrolysis, occurs commonly in germinating wheat embryos. A schematic showing RNase H treated lariats or linear excised introns when two primers are used is shown in Figure 4.6D.

4.7.3 Primer extension analysis for the study of excised *cox2* intron and *nad1* intron 2

An additional technique to provide information about physical forms of excised introns is primer extension analysis. This method is typically used to map the 5' termini of RNAs, and in this case, the 5' ends of linear excised group II introns. A caveat to this analysis however is that the reverse transcriptases (used in this method) are inefficient at extending over branchpoint regions, and thus the deficient extension over lariat 2'-5' bond could be mistaken for a linear species. However, primer extension analysis can be utilized to distinguish between *in vivo* circularized from linear (or lariat) excised introns, if the circularized molecules contain 3'-5' bonds at their junctions.

Figure 4.6 Determination of *in vivo* abundance of lariat (or circularized) versus linear excised introns (of wheat mitochondrial *rps3* intron and *nad4* intron 3) by RNase H analysis.

Wheat mitochondrial RNA (24 hr) was annealed with gene-specific oligomers as shown below each blot for **[A]** *rps3* intron and **[B]** *nad4* intron 3 prior to RNase H treatment. Positions of oligo-probes used in subsequent northern blot analysis are shown in red in schematic below each blot. Sizes of RNA species are adjacent to blots, asterisks represent the positions of native excised introns, black arrowheads indicate positions expected for hybridizing products from lariat (or circularized) intron forms and white arrowheads for linear intron forms. Models of dV and dVI secondary structure of *rps3* intron (above) and *nad4* intron 3 (below) with branchpoint A (asterisk) are shown in **[C]**. **[D]** Schematic depicting RNase H treated lariat or linear excised introns when two primers are used.


A
Untreated
LB217, LB330, -RNase H
LB217, LB330, +RNase H
4.6kb
* 1.6kb
700nt
450nt
LB217
LB330
LB334*
0.5 kb
rps3a
rps3b
rpl16
rps3 intron
B
LB230, LB389, -RNase H
LB230, LB389, +RNase H
~7.5kb
* 1.7kb
950nt
380nt
0.5 kb
LB203*
LB230
LB389
nad4a
b
c
d
nad4 intron 3
C
rps3 intron
73nt
nad4 intron 3
52nt
D
1.7kb
in vivo splicing
and/or
RNase H
950nt
380nt
LARIAT
(or fully-circularized)
LINEAR

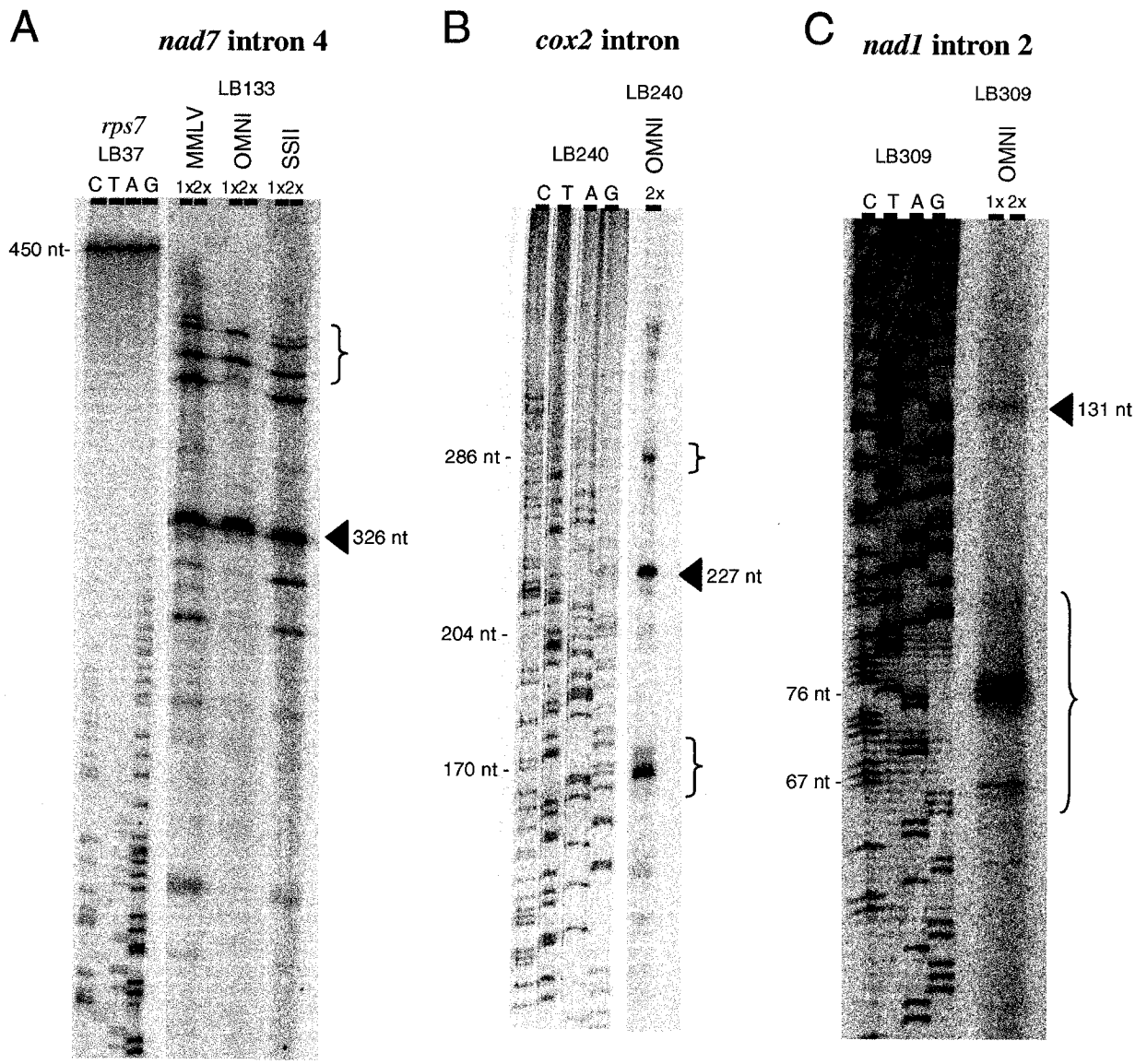
Previous analysis showed that wheat *nad7* intron 4 (24 hr) was excised as lariat as well as minor fully-circularized molecules (C.Carrillo, Ph.D. thesis, University of Ottawa, 2001). My studies corroborated these results and also suggest the presence of linear excised molecules (Figure 4.3). Primer extension was performed on this intron to first test the efficiency of three different RT's, namely MMLV (Invitrogen) and Omniscript (Qiagen) and SuperScript II (Invitrogen). I observed the expected extension product of 326 nt (Figure 4.7A, arrowhead), supportive of the major form being lariat (incomplete extension over the 2'-5' branchpoint) and linear excised *nad7* intron 4, when all three RT's were used. Omniscript RT was found to be the most efficient at traversing through RNA secondary structures as represented by the fewest premature stops in the OMNI versus the MMLV and SSII lanes in Figure 4.7A. Similar results were obtained when using a different stock of 24 hr wheat mitochondrial RNA. Higher molecular weight species were also seen (Figure 4.7A, brackets), between 326 nt – 450 nt in length, that are consistent with premature primer extension stops within the 3' region of *nad7* exon 4. Consequently, Omniscript reverse transcriptase was selected for primer extension of *cox2* intron and *nad1* intron 2 and the expected sizes for lariat or linear excised introns were observed, 227 nt and 131 nt denoted by arrowheads in Figure 4.7B-C, respectively. However, premature primer extension stops were also observed suggesting the presence of secondary structures at the 5' region of the introns (Figure 4.7, brackets). The 131 nt extension product of *nad1* intron 2 would point towards linear excised molecules, as no lariat molecules were detected via cloning of RT-PCR products (Li-Pook-Than and Bonen, 2006) and circular excised introns would not be detected using method. In the latter case, the reverse transcriptase would extend through the 3'-5' bond at the junction of the circularized introns (Figure 4.7D).

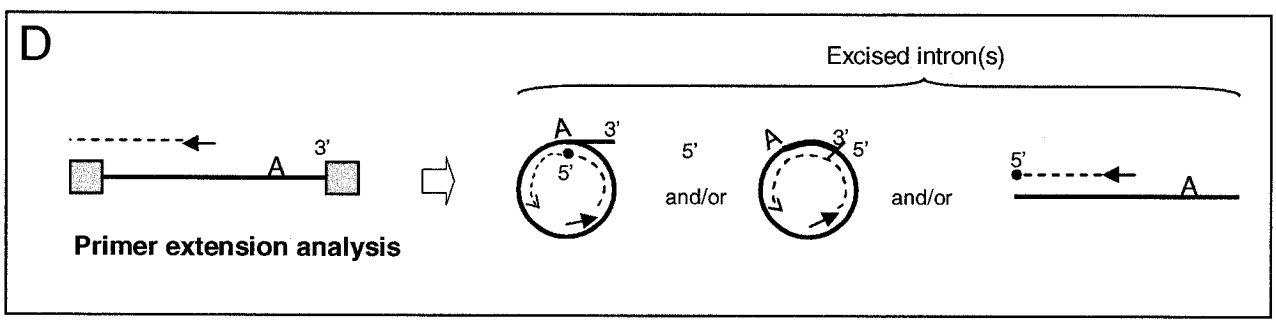
4.7.4 Northerns based on polyacrylamide gels for the study of excised *cox2* intron

Northern blot analysis from denaturing urea/polyacrylamide gels were also tested to distinguish between *in vivo* lariat and linear excised intron populations in wheat mitochondria during early seedling development. This type of gel is used to analyze *in vitro* self-splicing reactions of group II introns such that the excised lariat molecules would migrate very slowly, at positions above the high molecular weight linear precursors (Figure 4.8A, schematic). RNA was transferred from the gel to the membrane using an electroblotting technique. Preliminary studies using oligo-probes specific to *rps7* showed both mRNA and precursors profiles

Figure 4.7 Primer extension analysis of *nad7* intron 4, *cox2* intron and *nad1* intron 2.

[A] Wheat mitochondrial RNA was treated with an oligo-probe specific to *nad7* intron 4 and three different types of reverse transcriptase enzymes (MMLV, Omniscript - OMNI, Superscript II - SSII). Omniscript RT was selected for the primer extension of **[B]** *cox2* intron and **[C]** *nad1* intron 2. The sizes of expected extension products for linear and/or lariat molecules are 326 nt, 227 nt 131 nt, respectively, and are shown by arrowheads based on the DNA sequencing ladder on the left of each extension lane. Parentheses indicate positions of premature stops during primer extension due to inefficiencies of the reverse transcriptases to traverse RNA secondary structures. The extension lanes 1 and 2 contain extension products based from approximately 3μg (1x) and 6μg (2x) wheat mitochondrial RNA. The oligomers used for sequencing and primer extensions are indicated above each figure and their positions are shown in Figures 2.1D, E and A, respectively. A schematic showing the three possible types of extension products if the excised intron is lariat, circular or linear is shown in **[D]**.

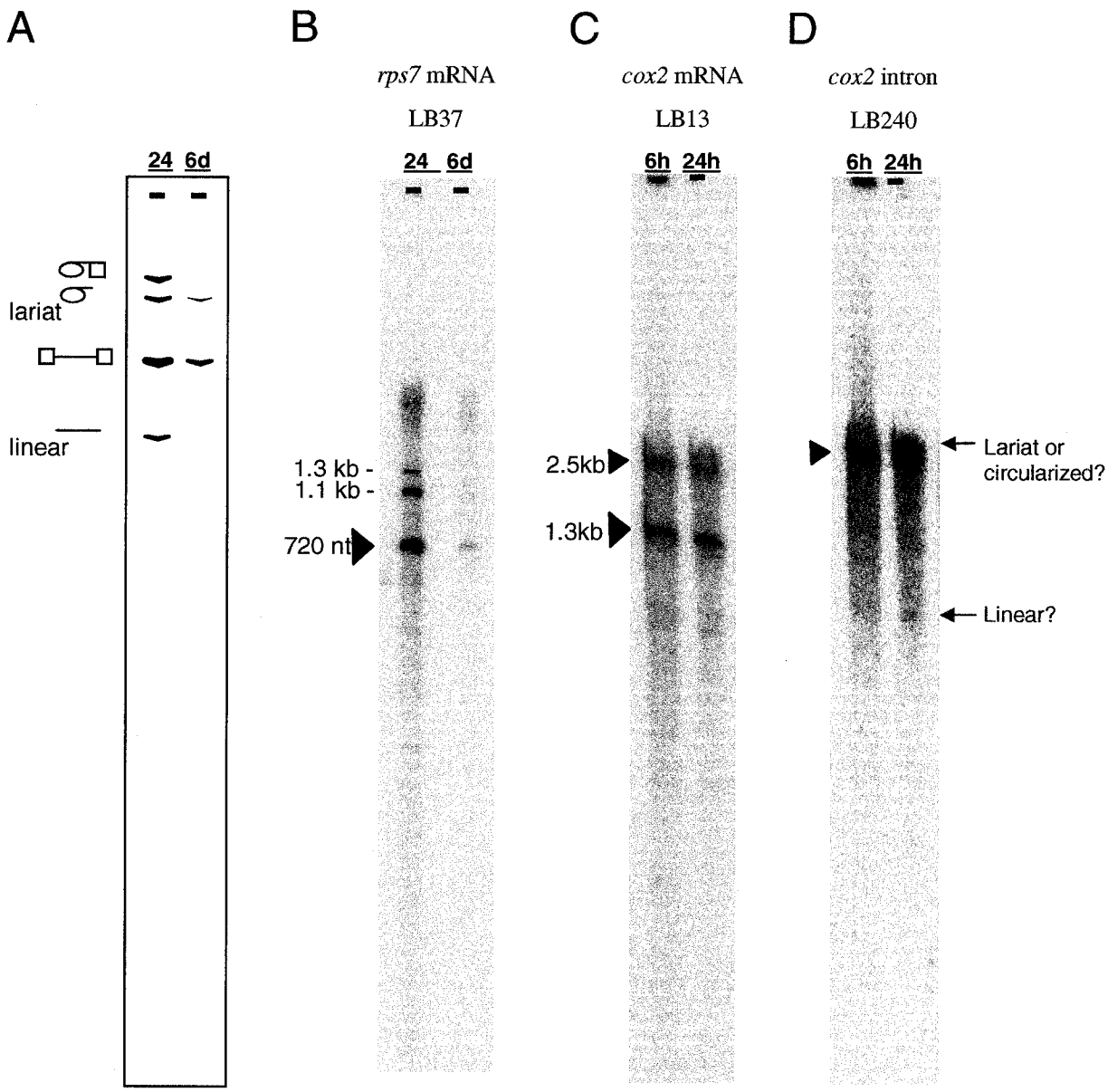

A
nad7 intron 4
LB133
rps7
LB37
MMLV
OMNI
SSII
C T A G
1x2x 1x2x 1x2x
450 nt-
326 nt
B
cox2 intron
LB240
LB240
OMNI
C T A G
2x
286 nt -
227 nt
204 nt -
170 nt -
C
nad1 intron 2
LB309
LB309
OMNI
C T A G
1x 2x
131 nt
76 nt -
67 nt -


D
Excised intron(s)
A
3'
5'
Primer extension analysis
5'
and/or
and/or
A

comparable to those from northerns obtained from agarose/formaldehyde gels (Figure 4.8B versus Figure 3.1B, respectively). Similarly, when a *cox2* exon-specific probe was used, both precursor species and mRNAs were observed (Figure 4.8 C-D, blue arrowhead versus black arrowhead, respectively). When a *cox2* intron-specific probe was used, a high molecular weight doublet band was observed only in the 24 hr lane of Figure 4.8D, as well as a faint low molecular weight band at the approximate expected size of excised linear introns (1.1 kb). These results are consistent with a population of both *in vivo* circularized molecules, as well as linear excised molecules, respectively, and corroborate results observed using RNase H analysis (Figure 4.3C, Section 4.3). The relatively low levels of excised introns make it difficult to use such methods for analysis, however abundant mRNAs can be clearly observed. Based on this preliminary analysis, this method has potential for distinguishing between lariat and linear forms of excised introns. Improving the sensitivity of this technique include adjusting gel concentration and electrophoretic conditions.

Figure 4.8 Analysis of physical forms of excised introns using northern blots based on polyacrylamide gels.

[A] Schematic showing the hypothetical positions of excised lariat or linear introns (in red) as well as partially spliced and precursor species if the blot is probed with an intron specific probe. Northern blots based on 4%-6% polyacrylamide gels were hybridized with oligomers specific to wheat mitochondrial **[B]** *rps7* mRNA, **[C]** *cox2* mRNA and **[D]** *cox2* intron. Development stage of extracted wheat mitochondrial RNA and oligo-probes used are indicated above each blot. Positions of mRNAs, precursors and excised introns are shown with large black arrowheads, small blue arrowheads and red arrows, respectively. Sizes of RNA species are given on the left of each blot. The same blot was consecutively used for [C] and [D].


A
B
C
D
rps7 mRNA
LB37
cox2 mRNA
LB13
cox2 intron
LB240
24 6d
6h 24h
lariat
linear
1.3 kb -
1.1 kb -
720 nt
2.5kb
1.3kb
Lariat or circularized?
Linear?

CHAPTER 5: RELATIONSHIP BETWEEN RNA SPLICING AND EXON EDITING NEAR INTRON JUNCTIONS IN WHEAT MITOCHONDRIA

5.0 Rationale

C-to-U editing has been observed in precursor and partially-spliced transcripts in plant mitochondria, and thus is considered an early RNA processing event, preceding group II intron splicing (Mulligan et al., 1999). Of particular interest are the RNA edits that are at exon sites close to intron junctions. In this section, these sites are studied in wheat mitochondria during early germination to better understand the relationship between editing and splicing. These junction edit sites also alter the corresponding codon and subsequently change the amino acid from their respective coding DNA. Notably, in wheat *nad7* (Carrillo and Bonen, 1997) and in *cox2* of cold stressed wheat (Kurihara-Yonemoto and Handa, 2001) these junction sites remained unedited when the adjacent intron was unspliced however edit sites more internal to the exon showed editing. This work investigated selected edit sites, close to splice junctions that are found in the respiratory chain pathway genes *nad2*, *nad5*, *nad7* and *cox2*, via sequence analysis of respective RT-PCR products. This Chapter also focuses on examining if RNA editing is exclusively an early post-transcriptional event, if editing and splicing are independent of one another, and how well is editing synchronized with transcription during early seedling development.

Sections of this chapter (**5.1-5.2**) are in press as: "Li-Pook-Than, J., Carrillo, C., Niknejad, N., Calixte, S., Crosthwait, J. and Bonen, L. Relationship between RNA splicing and exon editing near intron junctions in wheat mitochondria during germination and seedling development. Invited paper to appear in a special issue on plant mitochondria in *Physiologia Plantarum* (accepted May 15, 2006)". The full manuscript is given in Appendix I.

5.1 RT-PCR sequencing analysis of exon editing sites near intron junctions

To address if editing and splicing are independent of each other, I examined four of the seven exon edit sites in wheat mitochondria that are close (no more than 6 nt away) to unspliced cis-introns. These were *nad2d* at the -4 position (4 nt upstream of intron 4), *nad5d* at the -5 position (5 nt upstream of intron 4), *nad7c* at the -1 position (1 nt upstream of intron 3) and *cox2a* at the -4 position (4nt upstream of its intron). Their positions within their respective genes and their subsequent amino acid conversions are shown in Figure 5.1 (asterisks and hatched boxes, respectively). The remaining exon junction edit sites *nad4c* at the +1 position, *nad4d* at the +6 position and *nad7e* at the +3 position, were also examined by members of our lab and the complete data set of all seven exon edit sites near exon/intron junctions for wheat mitochondria will be found in Li-Pook-Than et al., *Physiologia Plantarum* (in press). For multiply-split genes, such as *nad2*, *nad5* and *nad7*, I focused on partially-spliced transcripts to ensure that there was no contribution from residual contaminating DNA. Various intron/exon primer combinations were used to generate RT-PCR products for direct sequencing at these four positions in both 24-hr germinating embryos and 6-day etiolated seedlings to also assess developmental differences (Figure 5.2-5.4, schematics).

In the case of the *nad2d* (-4) site, virtually complete editing was observed in partially-spliced transcripts for both 24 hr embryos and 6 day seedlings (Figure 5.2A, black arrowheads), with only faint signals for the unedited state (cf. C lane in sequencing gels). For 24 hr RT-PCR sequencing data of *nad2d* (-4), my results confirmed those obtained by J. Crosthwait (laboratory project). These results also showed full editing at the edit site close to the junction, as well as other sites in *nad2d* that are further away from the intron (Figure 5.2B, three black arrowheads, star on schematic). Interestingly, two side-by-side *nad2d* sites which are located 33-34 nt upstream of the *nad2* intron 4 junction, appear about 50% edited in these precursors (Figure 5.2A, grey arrowheads), raising the possibility of a steric exclusion effect at such adjacent sites. Both these edited positions are part of the same codon whereby the edits observed alters the amino acid from that of the coding DNA (P→L). All these exon sites, including sites close to the exon/exon junction, exhibited complete editing in 24 hr fully-spliced *nad2* mRNAs, as expected (Figure 5.2C).

Figure 5.1 Distribution of exon editing sites in intron-containing wheat mitochondrial genes, *nad2, nad4, nad5, nad7* and *cox2*.

Coding regions are shown by open boxes, cis-introns are shown as triangles and Y-shaped symbols indicate trans-introns. Edit sites are shown by vertical lines, and those with black dots are located within 15 nt from a splice junction, with the seven positions either upstream (negative symbol) or downstream of an intron (plus symbol) shown in boxes above. The positions of the four wheat edit sites examined in this section of this study are indicated with asterisks, while C-to-U changes within codons and subsequent amino acid changes are boxed. The examination of the remaining three exon/intron junction edit sites in wheat mitochondrial genes can be found in Appendix I. Editing data were compiled for wheat *nad2* (Morawala-Patell et al., 1998), *nad4* (Lamattina et al., 1991), *nad5* (Pereira de Souza et al., 1991), *nad7* (Bonen et al., 1994) and *cox2* (Covello and Gray, 1990). The schematics in Figure 2.2 show the exon sizes to scale.

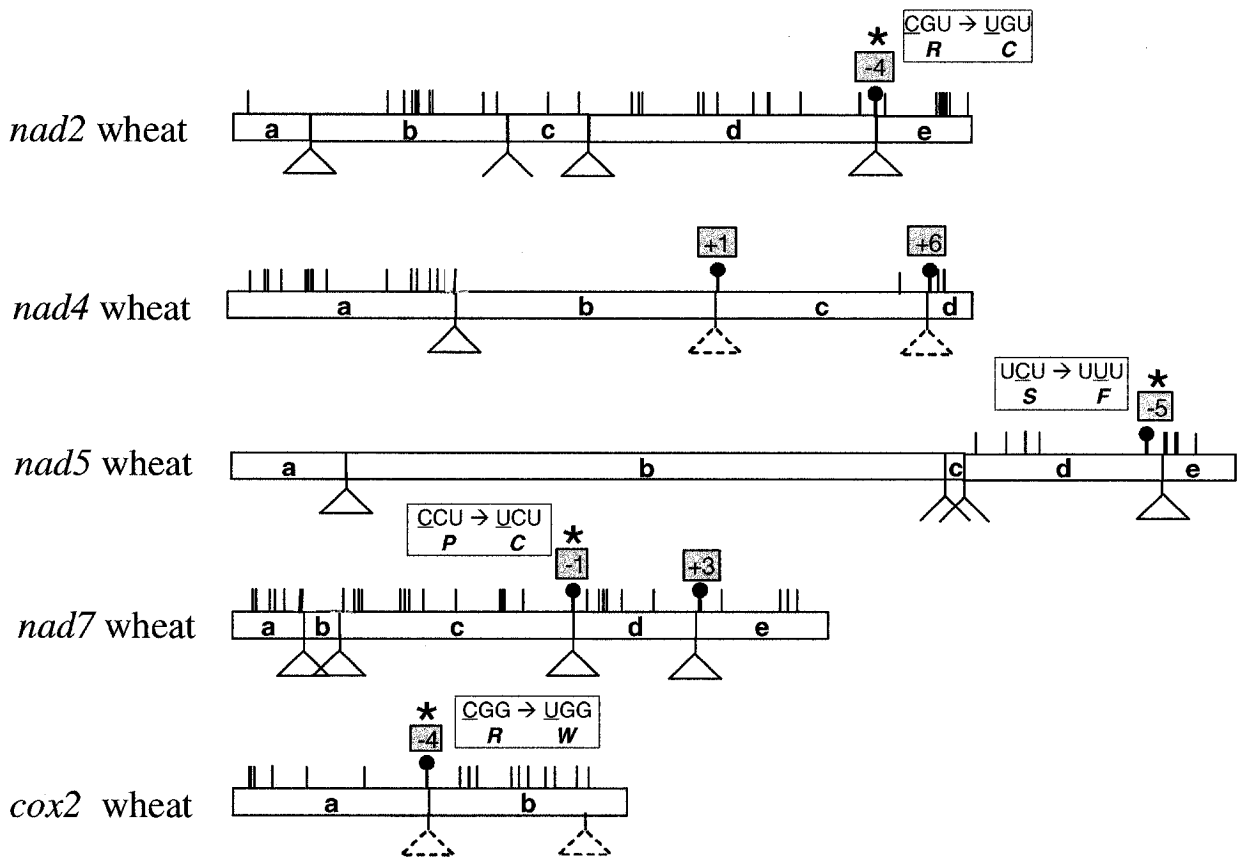
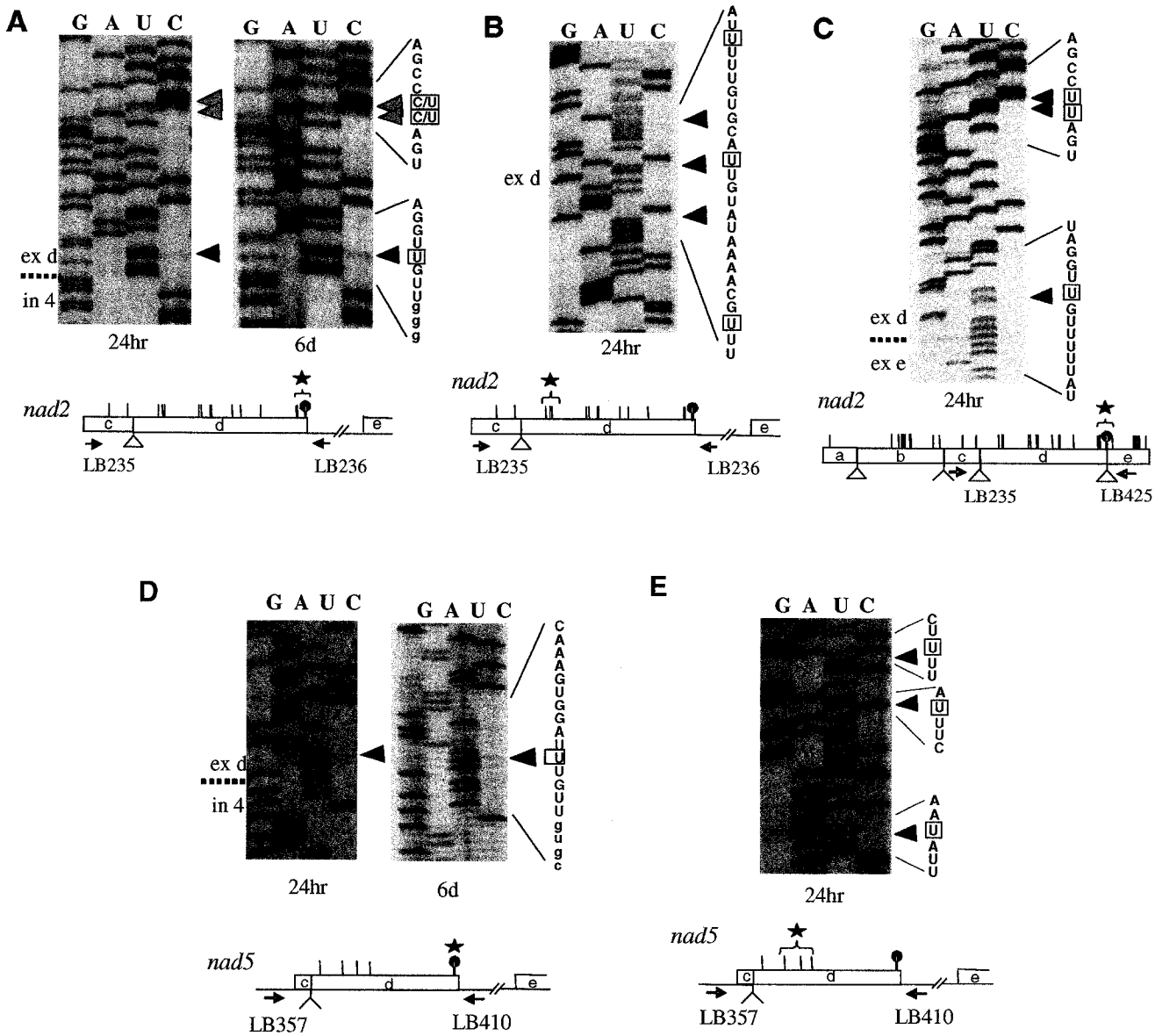

nad2 wheat
CGU → UGU
R C
-4
a
b
c
d
e
nad4 wheat
+1
+6
a
b
c
d
nad5 wheat
UCU → UUU
S F
-5
a
b
c
d
e
nad7 wheat
CCU → UCU
P C
-1
+3
a
b
c
d
e
cox2 wheat
CGG → UGG
R W
-4
a
b

Figure 5.2 Exon editing status near splice junction of *nad2d* (-4 position) and *nad5d* (-5 position) for wheat mitochondrial precursor transcripts and mRNAs from 24 hr germinating embryos and 6 day etiolated seedlings.

[A-C] Direct sequencing of RT-PCR products from partially-spliced *nad2d*-intron 4 template showing [A] the -4 edit region or [B] 3 distal editing sites, and [C] from *nad2* mRNA template showing the -4 edit region. **[D-E]** Direct sequencing of RT-PCR products from partially-spliced *nad5d*-intron 4 template showing [D] the -5 edit region and [E] 3 distal editing sites. Sequences are shown at the right of sequencing gels, with exon (upper case), intron (lower case) and candidate edit sites (boxed). Black, grey or white arrowheads indicate observed C-to-U edits, incomplete editing of the population or unedited sites, respectively. The schematics below (as described in Figure 1 legend) show positions of primers (with numbered arrows, see Table 1) and positions of edits under examination (stars).


A
G A U C
G A U C
ex d
in 4
24hr
6d
nad2
c
d
e
LB235
LB236
B
G A U C
ex d
24hr
nad2
c
d
e
LB235
LB236
C
G A U C
ex d
ex e
nad2
24hr
a
b
c
d
e
LB235
LB425
D
G A U C
G A U C
ex d
in 4
24hr
6d
nad5
c
d
e
LB357
LB410
E
G A U C
24hr
nad5
c
d
e
LB357
LB410

For the *nad5d* (-5), similar results to *nad2d* (-4) were obtained whereby the -5 site was edited at the junction site in partially-spliced transcripts (of both 24 hr and 6 day plants) when the adjacent intron was still present (Figure 5.2D). Full editing was also observed at internal positions in *nad5* exon d in both 24 hr (Figure 5.2E) as well as 6 day (data not shown) RNA populations. Analogous observations were made in the case of the wheat *nad7c* (+12) site (Carrillo and Bonen, 1997) and for potato *rps10* (-13) site based on 7/8 cDNA clones (Zanlungo et al., 1995). In contrast, wheat *nad4*d (+6) was found to be unedited in partially-spliced transcripts as shown in Li-Pook-Than et al., *Physiologia Plantarum* (in press).

When exon editing sites located immediately beside intron junctions were studied, interestingly in the case of *nad7*c (-1), like *nad4*c (+1), these sites remained unedited when the adjacent intron was unspliced (Li-Pook-Than et al., *Physiologia Plantarum*, in press). In Figure 5.3A (white arrowheads), this can be seen for the *nad7*c (-1) site at the position one nucleotide upstream of intron 3 confirming previous analysis (C. Carrillo, Ph.D. thesis, University of Ottawa, 2001). Here, I also show that this site underwent virtually complete editing in partially-spliced transcripts when intron 2 is still present but intron 3 has been excised (Figure 5.3B, black arrowhead). With respect to the *nad7e* (+3) exon editing site, the lack of editing was also observed for a *nad7e* (+3) site in RNA precursor populations which still contain intron 4 (Carrillo and Bonen, 1997). None of 22 clones from 24 hr or 6 day seedlings showed editing at that site in contrast to other *nad7d* and *nad7e* sites (which showed at least 50% editing in the population). Fully-spliced mRNA including dormant seeds (0 hr) and germinating embryos (12 hr and 24 hr stages of development) were also found to be completely edited (Figure 5.3C). This suggests that the editing factors that are associated to positions close to the site of editing are hindered or unavailable when the adjacent, and not distal, intron is still present. This study also included RT-PCR sequencing of *nad4* mRNAs to examine the editing status of *nad4c* (+1) and *nad4d* (+6) in 24 hr germinating embryos. Here I show that *nad4* mRNAs are completely edited (Figure 5.3D), analogous to fully edited mRNA observations of *nad2* and *nad7*.

5.2 Developmental differences in editing status of intron-containing wheat mitochondrial *cox2* transcripts

At this point, when both 24 hr and 6 day partial-transcripts were examined in *nad2*, *nad5* and *nad7*, each with one or more introns removed, no changes in editing were observed between

these two developmental stages at edit sites close to splice-junctions or at sites more internal to the exon. However, changes in editing status at different developmental stages were observed when examining precursor transcripts of *cox2* (its single intron unspliced). The editing status of spliced and unspliced wheat *cox2* transcripts during the developmental period when seeds leave dormancy, germinate and develop into etiolated seedlings (0 hr to 6 days post-imbibition) were examined in collaboration with S. Calixte (current M.Sc. student in lab). These stages were selected because I had previously observed that intron-containing *cox2* precursors appeared relatively more abundant during the 12 hr – 24 hr period than they do in either earlier (cf. 6 hr) or later (cf. 4d and 6d) stages where only very minor amounts of unspliced forms or excised intron were detected, even though spliced mRNA levels are high (Figure 5.4A, top panel vs. bottom panel). Because *cox2* contains a single intron in wheat, our editing analysis RNA samples were repeatedly (3X) DNase-treated prior to RT-PCR (Figure 5.4B, inset) to minimize as much as possible any contribution from residual contaminating mitochondrial DNA. No amplification products are observed when RT-PCR was performed after reverse transcriptase was omitted from the reaction (Figure 5.4B, inset, -RT lane) suggesting a minimal contribution of contaminating mitochondrial DNA.

For the unspliced *cox2* transcripts in dormant seeds, the splice junction (-4) exon site shows only about 50% editing (Figure 5.4B, upper panel grey arrowheads) and interestingly, a more distal *cox2* exon 1 site which is 130 nt further away from the intron exhibits about only 10% editing (Figure 5.4B, lower panel, white arrowhead). In the case of 24 hr embryos, the *cox2* (-4) editing status resembles that of the 0 hr RNA population, whereas the distal exon site shows greater than 50% editing (Figure 5.4B, grey arrowheads) and in 6-day seedlings more than 90% editing (or virtually complete if the minor signal in the C lane is due to contaminating DNA) (Figure 5.4B, black arrowheads). This variation in degree of editing at various sites in unspliced molecules and lack of correlation with proximity to the splice junction suggests independence in the nature of editing determinants. In contrast, the spliced wheat *cox2* mRNA, as expected, showed virtually complete editing at all sites examined (including the -4 position) in all these developmental stages as well as in dormant seeds (Figure 5.4C, black arrowheads). This corroborates observations that mRNAs are typically fully edited, as previously found with *nad2, nad4* and *nad7* mRNAs (Figure 5.2C and 5.3C-D, respectively). The finding that *cox2* unspliced precursors, show different degrees of editing between the embryo and seedling stages not only at

splice junctions but also at other exon sites, raises questions about the efficiency of editing in the early stages of development and the availability of nuclear-encoded RNA processing machinery. It will be of interest to further elucidate the inter-relationships between RNA editing and splicing events especially during periods of high respiratory demand, as in the early stages of germination.

Figure 5.3 Editing status for wheat mitochondrial *nad7c*/intron 3 and *nad4* mRNA at junction sites during early seedling development.

Direct sequencing of RT-PCR products for **[A]** partially spliced *nad7c*/intron 3 at the (-1) junction region in 12 hr and 24 hr germinated embryos, **[B]** *nad7*c/d with a distal intron still present at the 24 hr stage of development, **[C]** n*ad7* mRNA at the exon c/d (-1) site in dormant seeds, 12 hr and 24 hr germinated embryos, and **[D]** n*ad4* mRNA at the exon b/c (+1) and exon c/d (+6) in 24 hr germinated embryos. Designations are as described in Figures 5.1 and 5.2.

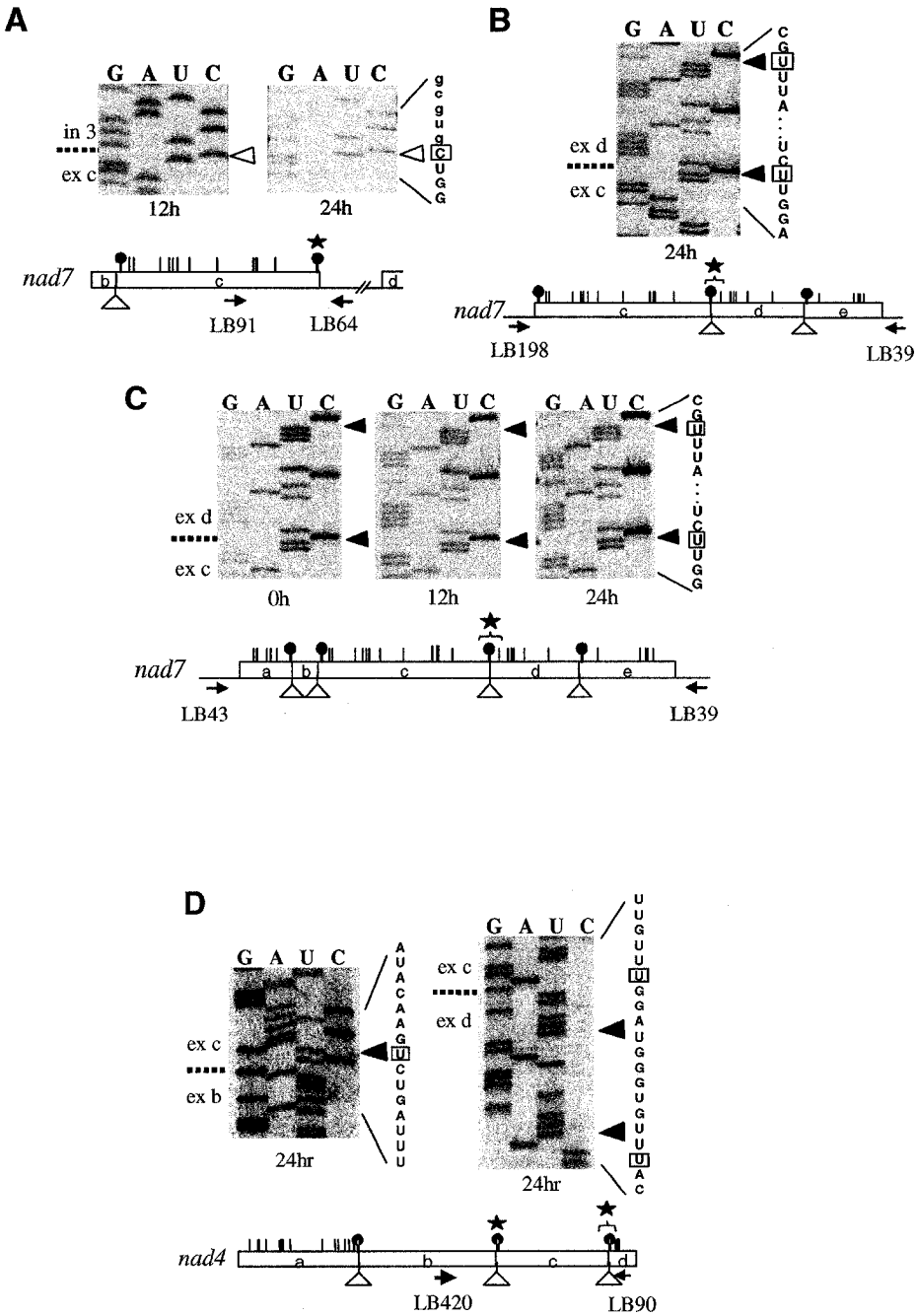
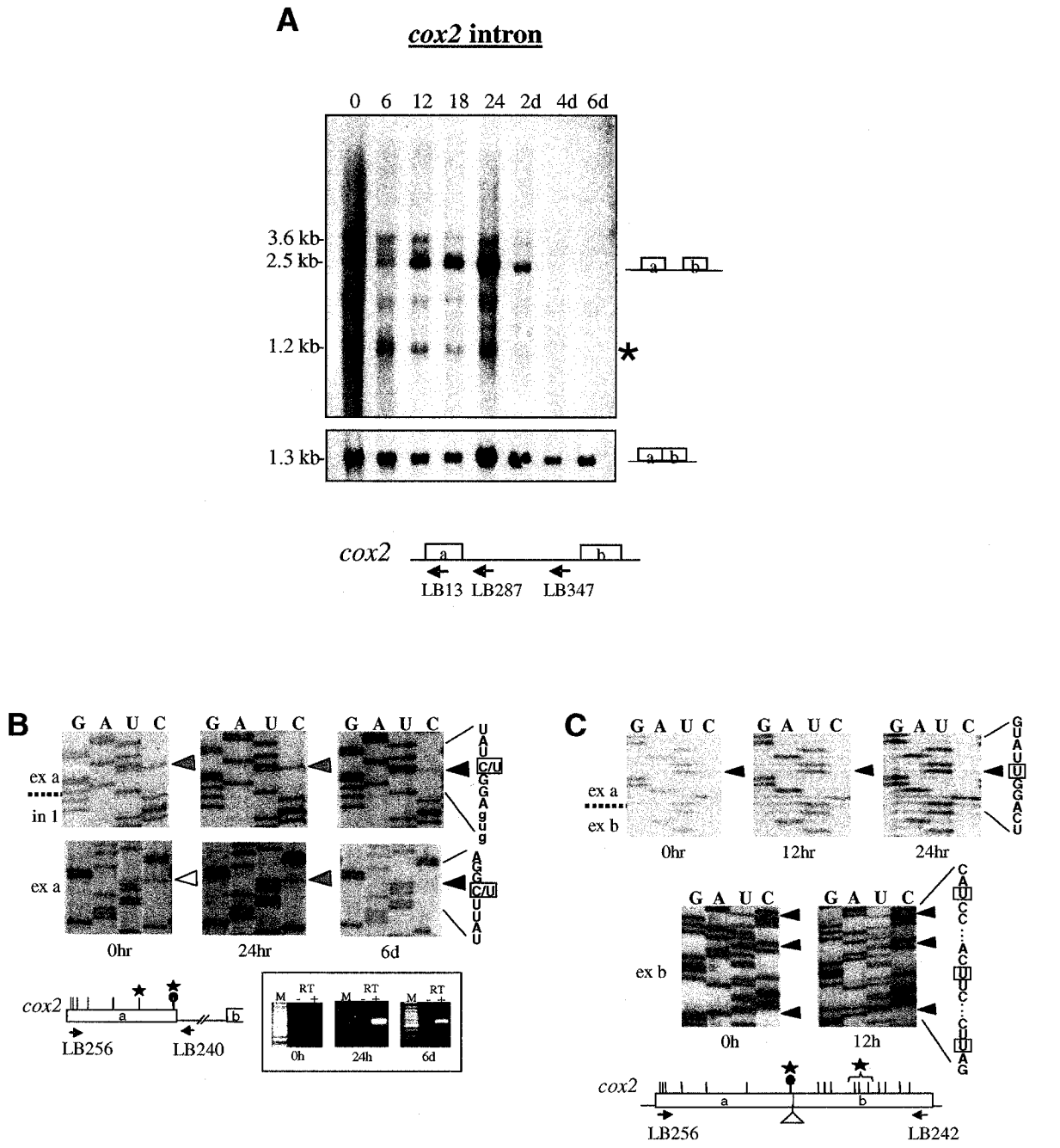

A
G A U C
G A U C
in 3
ex c
12h
24h
nad7
b
c
d
LB91
LB64
B
G A U C
ex d
ex c
24h
nad7
c
d
e
LB198
LB39
C
G A U C
G A U C
G A U C
ex d
ex c
0h
12h
24h
nad7
a
b
c
d
e
LB43
LB39
D
G A U C
ex c
ex b
24hr
G A U C
ex c
ex d
24hr
nad4
a
b
c
d
LB420
LB90

Figure 5.4 Northern blot analysis and editing status of wheat mitochondrial *cox2* spliced and unspliced transcripts during embryo-to-seedling development (0 hr – 6 days post-imbibition).

[A] Northern blot analysis of wheat mitochondrial RNA isolated from various developmental stages (0 hr – 6 days post-imbibition) and hybridized first with *cox2* intron-specific oligomers (upper) and subsequently with a *cox2* exon 1 specific oligomer (lower). Sizes are indicated on the left while schematics of *cox2* precursors, excised intron (asterisk) and mRNA are shown on the right. Lower schematic shows positions of hybridization probes. **[B]** Direct sequencing of RT-PCR products for *cox2* precursors at the exon (-4) junction site (above) and at editing sites distal from the junction (below). Inset show gels of RT-PCR products (of 600 bp) representing *cox2* precursor transcripts when RNA (which had been pre-treated with DNase I) was incubated either with (+RT) or without (-RT) reverse transcriptase. **[C]** Direct sequencing of RT-PCR products for *cox2* mRNA at -4 position (upper panel) and at distal editing sites (lower panel) at 0 hr, 12 hr and 24 hr post-imbibition. Designations are as described in Figures 5.1 and 5.2.


A
cox2 intron
0 6 12 18 24 2d 4d 6d
3.6 kb
2.5 kb
1.2 kb
1.3 kb
a
b
*
cox2
LB13 LB287 LB347
B
G A U C
ex a
in 1
ex a
0hr
24hr
6d
cox2
LB256
LB240
M RT - +
0h
24h
6d
C
ex a
ex b
0hr
12hr
24hr
0h
12h
LB256
LB242

6.0 GENERAL DISCUSSION

6.1 Coordination of splicing with transcription during wheat embryo-to-seedling development

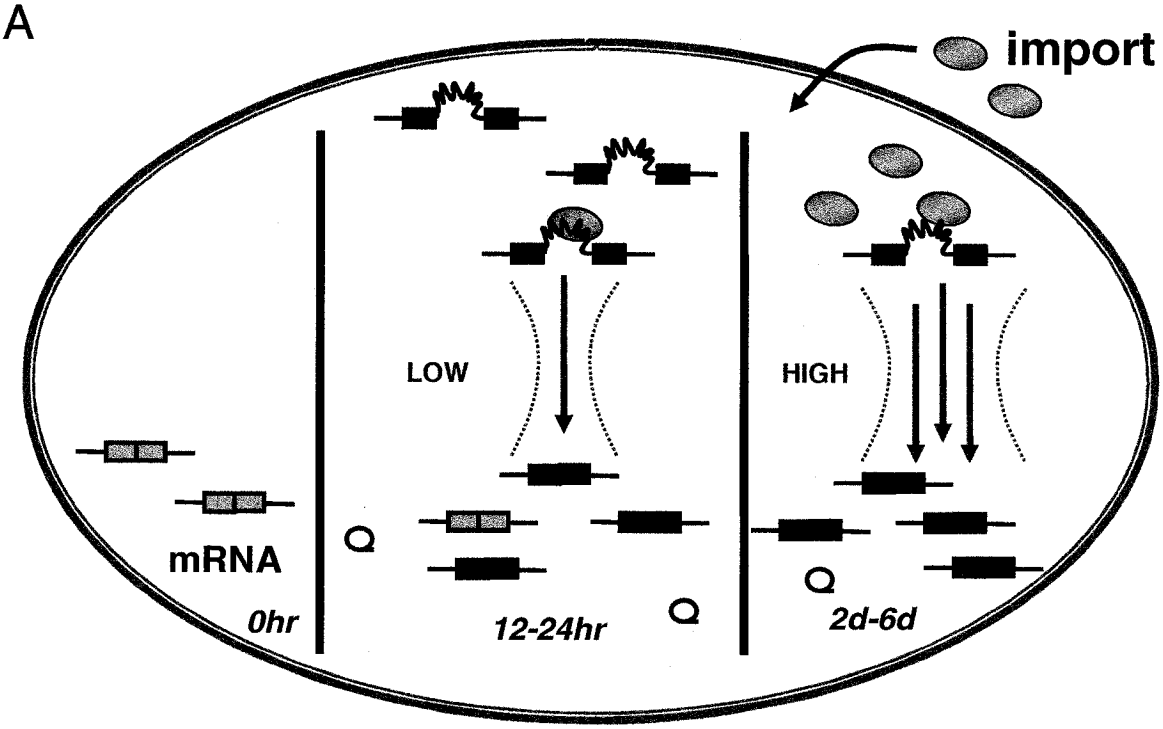
For wheat mitochondrial intron-containing genes such as *cox2, rps3, nad7* and *nad1* (Chapter 3), the transient accumulation of high molecular weight precursors relative to mature mRNAs was observed during the 6 hr-to-2 day stages of development. This is consistent with a transition from poor synchronization between transcription and RNA processing in germinating embryos, to more efficient coordination at the seedling stages of development. This apparent lag in coordination between transcription and RNA splicing events occurs at approximately the same period as Phase II of respiration when the newly synthesized electron chain components are thought to be accumulating and assembling in plant mitochondria (Bewley and Black, 1994). Howell et al. (2006) suggest a succession of transcription events for a subset of mitochondrially-encoded and nuclear-encoded rice genes that they studied. In their work, the earliest transcripts to reach maximal abundance (after 2 hrs of imbibition) were those for mitochondrial import (translocases of the inner and outer mitochondrial membrane, such as nuclear-encoded *TIM22* and *TOM20*), followed by those involved in transcription and translation (nuclear-encoded RNA polymerase (*RpoTm*) and several ribosomal proteins, mt-encoded *rps4* and nuclear-encoded *rps14*). The majority to reach maximum abundance the latest, after 24 hrs, were the election transport chain genes (nuclear-encoded *sdh2* and *cytc*-1, as well as mt-encoded *atp1*and *cox2*). Based on their *in vitro* import studies, protein import machinery was found to be abundant in dry seeds and operational after 2 hr of imbibition. My northern results for wheat mitochondrial intron-containing genes show an increase of precursor RNAs at the 6 hr stage of development that reached relative higher abundances between 12-24 hrs and is in overall agreement with Howell's study (2006) concerning the respiratory chain genes. In this light, it may be expected that the RNA splicing machinery is unavailable in 6 hr imbibed embryos but fully functional after the 24 hr stage of development, after the translocases are established and *de novo* transcription is occurring.

A model depicting this apparent transition from lower efficiency of RNA splicing relative to mitochondrial transcription during early germination to a more efficient coordination in the seedling stages is shown in Figure 6.1A. Mitochondrial mRNAs are stored in dormant seeds, but

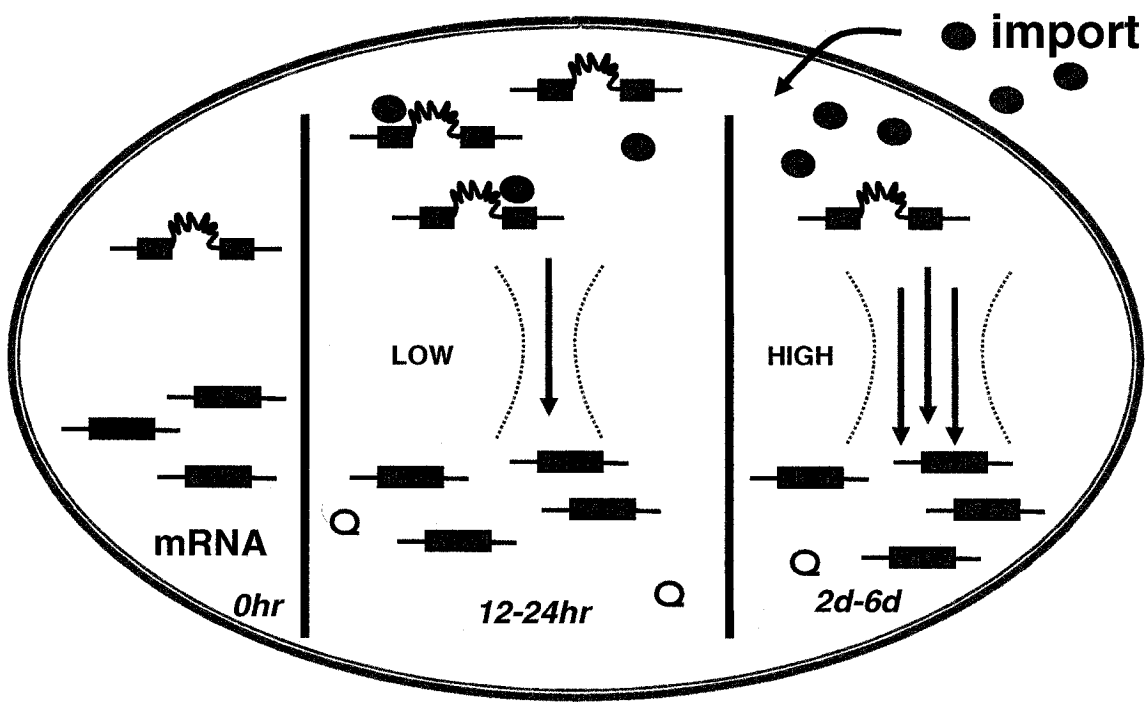
Figure 6.1 Schematics depicting the transition from inefficient coordination between transcription and RNA processing events, splicing and editing, in 24 hr germinating seeds to a more synchronized coupling in 6 day seedlings.

[A] Group II intron splicing is inefficiently coordinated with transcription in 6-24 hr germinating seeds, but more efficiently coordinated in 4-6 day seedlings. This is based on high steady state levels of wheat mitochondrial *nad1*, *nad7*, *cox2* and *rps3/rpl16* intron-containing transcripts relative to respective mRNAs observed during the germinating stages of development (12-24 hrs); as described in Chapter 3. **[B]** C-to-U editing appears inefficiently coordinated with transcription in 24 hr germinating seeds. This is based on analysis of *cox2* transcript C-to-U edit sites, both at internal exon and exon/intron junction positions, which were incompletely edited in embryos and 24 hr germinating wheat seeds, while in 6 day seedlings sites were fully edited (as shown in Figure 5.4). The shift from low to high efficiency of splicing [A] and editing [B] is depicted with downward arrows. Exons, introns, UTRs and excised introns are represented by colored boxes, waved lines, horizontal lines and conventional lariat forms, respectively. Stored mRNAs are shown by yellow in [A] and pink colored boxes in [B]. Imported splicing machinery (*s*) and editing machinery (*e*) are shown as purple and orange ovals, respectively.

A


import
LOW
HIGH
mRNA
0hr
12-24hr
2d-6d
Q
Q
Q

B


import
LOW
HIGH
mRNA
0hr
12-24hr
2d-6d
Q
Q
Q

after hydration, *de novo* transcription is activated and detected with the accumulation of precursors in germinating embryos (6-24 hrs). Transcription and intron excision only appear better synchronized after nuclear-encoded splicing machinery is more efficiently imported into the mitochondria, assembled and fully functional at the late seedling stages of development in wheat (2-6 days). Another example of a nuclear-encoded protein gene, *hsp70* (heat shock protein), is expressed more abundantly at seedling stages of development, relative to the mitochondrial genes, and is targeted to the mitochondria and is required for mitochondrial biogenesis (Logan et al., 2001). This 'bottleneck' effect for the availability of nuclear-encoded proteins is also observed during pollen formation (Smart et al., 1994). Here, a significant increase in transcripts of mitochondrial *atpA, atp9, cob* and *26S rRNA* was observed relative to nuclear encoded transcripts (*atpB* and adenine nucleotide translocator- *ant*) in sunflower. Notably, during the period of rapid mitochondrial generation, plant mitochondria were found to closely surround the nuclear envelope (reviewed in Binder et al., 1996). Further analysis such as quantitative PCR can be used to more specifically quantify the relative abundance of transcripts in the population. Inter-organellar crosstalk plays an important role during plant development. This phenomenon is observed in barley mutants that compensated for a chloroplast deficiency with increased levels of mitochondrially encoded cytochrome oxidase and ATP synthase transcripts in leaves, but not in roots (Hedtke et al., 1999).

6.1.1 mRNA abundance of wheat mitochondrial genes during wheat embryo-to-seedling development

I observed a variation in ribosomal protein gene expression profiles (*rps2, rps3/rpl16* and *rps7)* during early plant development, relative to the steady state levels of their respiratory chain pathway counterparts (*nad7, cox1, cox2, cob and atp6)*. This is consistent with the requirement of operational ribosomal proteins during the translation of the respiratory chain transcripts. Analogously, large amounts of cytosolic ribosomes and ribosomal subunits are stored in dry maize embryos and become fully functional at the onset of hydration (Beltra-Pena et al., 1995). These patterns also suggest regulatory differences between functionally different genes during early plant development, as described in Section 6.1. Comparable results were observed in *cox2, atp1* and *atp9* of maize, where low to undetectable levels of transcripts were observed in dormant

embryos, relative to more abundant RNAs at the 12 hr stages (Ehrenshaft and Brambl, 1990). Similarly, *rps10, atpα* and *cox1* transcript abundance increased from 2-5 days and slightly decreased at 7 days in soybean cotyledons (Daley et al., 2003).

In Chapter 3 of this work, abundant mRNAs were revealed for both the respiratory chain and ribosomal protein genes relative to precursor transcripts in dormant seeds (Figure 3.1). The two possible classes of mRNA present in the dry embryo are 1) residual mRNAs that are marked for degradation during early seedling development, and 2) conserved mRNAs that are immediately hydrating on imbibition and are integral for germination as enzymes for primary metabolism or fermentation (Bewley and Black, 1994). This suggests that dry embryos contain stored mRNAs for latent respiration of the viable seed, which may then become activated upon hydration. On the onset of *de novo* transcription, a mixture of stored and newly synthesized mRNAs would thus be present during the early stages of development (Figure 6.1A, pale green vs. bright green). Distinction between these stored and newly synthesized mRNAs using *in organelle* transcription run-on assays may help to address questions regarding the transcription rate versus RNA turnover for any particular developmental stage. The steady state amount of mitochondrial RNA is controlled by the transcription rate and transcript stability (Hoffmann et al., 2001). As observed in Arabidopsis mitochondria, variation in transcription rates for different genes may be counterbalanced by post-transcriptional processes, namely RNA degradation (Giegé et al., 2000; Holec, 2006). Interestingly, promoter strength was shown to be uncorrelated with steady-state abundance of mRNA. For example, maize *rps12* was found to have a strong promoter, but lower steady state levels of RNA (Muise and Hauswirth, 1992), whereas *atp9* had a higher transcript abundance relative to other protein coding genes, but had a weaker promoter than *atp1* (Mulligan et al., 1991).

6.2 Temporal relationship between RNA editing and group II intron splicing at editing sites near exon/intron junctions

C-to-U editing can occur in both exons as well as sites within introns. In the latter case, editing can improve RNA secondary structure (Carrillo et al., 2001, Li-Pook-Than and Bonen, 2006), and potentially improve long-range RNA:RNA interactions or association with trans-acting factors. This would suggest that editing would have to be an early event that would precede splicing, though the editing would not be efficiently coordinated with splicing and

transcription during early stages of development (reviewed in Mulligan et al., 1999). In 7 day etiolated maize seedlings, mitochondrial precursor *cox2* RNAs showed less editing at exon sites close to introns, and at sites more internal in the exon (Yang and Mulligan, 1991).

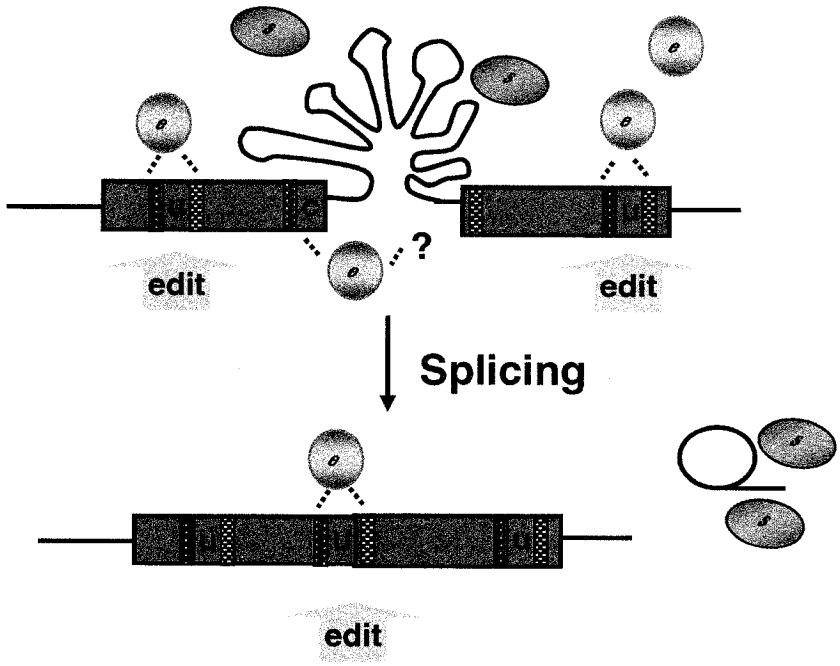
We have gained some insight on the nature of editing regions by looking at the editing of exon sites adjacent to unspliced introns (summarized in Table 6.1). C-to-U editing did not occur at the junctions sites of *nad4c* (+1), *nad4d* (+6), *nad7c* (-1) and *nad7d* (+3) precursors, while in contrast, virtually complete editing was seen for *nad2d* (-4) and *nad5d* (-5) (Carrillo, C., Ph.D. thesis, University of Ottawa, 2001; Li-Pook-Than et al., *Physiologia Plantarum*, in press; Chapter 5). When sequences before and after splicing were compared, no generalized editing consensus motifs were observed in these examples (Table 6.1). These results, however, remain in keeping with previous analyses suggesting that flanking nucleotides around edit sites (for stretches of approximately 10-20 nt) provide key information for editing recognition, with elements upstream of the candidate site being more important than downstream ones (reviewed in Shikanai, 2006). It is possible that editing at such sites might be a rate limiting step that triggers splicing, but low levels of such intermediate molecules would be difficult to detect *in vivo*. Our results suggest that splicing is the rate limiting step for editing at the exon/intron junction site.

The relationship between splicing with editing can be described through two models:
1) In the 'exon-exon motif' model, the proper motif can be recognized by editing machinery which is available only after the exons are spliced together (Figure 6.2A). The presence of the adjacent exon may also provide a secondary structural motif that could act as a cis-editing element. Such cis-acting anti-sense RNA has been proposed to work with RNA adenosine deaminases for A-to-I editing (Higuchi et al., 1993). Another recent study by Bolle and Kempken (2006) showed correct editing in *cox2* transcripts in both mono- and eudicot plant-specific *in organello* systems. They proposed a self-guiding transcript model, where a complex secondary or tertiary structured element contained within the transcript itself could direct editing specificity (Bolle and Kempken, 2006). This is compatible with the 'exon-exon motif' model proposed here, where this RNA structure would be available as a cis-element once the exons are spliced together. A variation in editing status is observed in precursor populations for sites slightly away from intron borders (Table 6.1). This is consistent with the view that there is variation among sites as well as between orthologous sites in different plants based on *in vitro*

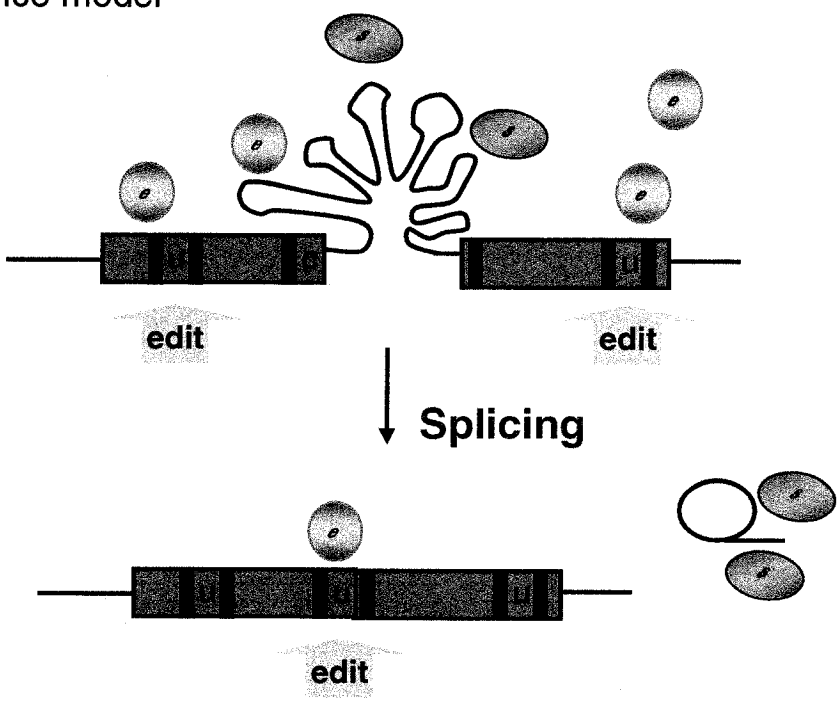
Figure 6.2 Two models depicting coordination between editing and splicing.

[A] The 'exon-exon' motif model whereby the correct motif for recognition by editing machinery is only available when the two exons are spliced together. This hypothetical motif is represented by adjacent red and blue stippled boxes. **[B]** The 'structural inhibition' model suggests that the secondary (or teriary) structure of the unspliced group II intron may sterically inhibit access by the editing machinery. Once the intron is removed, the specific site can be edited. Exons, introns, UTRs and excised introns are represented by colored boxes, waved lines, horizontal lines and conventional lariat forms, respectively. Imported splicing machinery (***s***) is shown purple ovals and editing machinery (***e***) are shown as purple and orange ovals, respectively.

A) Exon-exon motif model



B) Steric hindrance model



editing studies (Takenaka et al., 2004; Neuwirt et al., 2005) and *in organelle* studies (Farré et al., 2001; Staudinger and Kempken, 2003).

2) The 'structural inhibition' model suggests that the presence of the unspliced group II intron may sterically inhibit access by the editing machinery (Figure 6.2B). The nature of this 'physical blocking' might well be due to interactions between intron and exon sequences. Notably for group II introns, sequences in the upstream exon undergo base-pairing with complementary intronic sequences (EBS1 and EBS2) located within domain I of the core structure (reviewed in Michel and Ferat, 1995; Bonen and Vogel, 2001; Lehmann and Schmidt, 2003). There is evidence from mutational electroporation studies of *cox2* splicing in isolated wheat mitochondria that supports EBS-IBS interaction (Farré and Araya, 2002), although it is worth noting that some plant mitochondrial introns appear degenerate and lack certain conventional group II features (Bonen and Vogel, 2001; in wheat see Figure 1.4). An EBS-IBS interaction might be expected to block entry of editing machinery, and indeed if editing were to convert G-C pairs to G-U pairs one might expect a bias against editing preceding splicing. Alternatively, if A-C mispairs were edited to A-U, the intron-exon interaction would be strengthened, and hence editing might improve the efficiency of splicing, as proposed in the case of *Isoetes lacustris nad1* (Malek and Knoop, 1998). For the *cox2* (-4 position) edit, there are examples in both directions. In *Acorus calamus*, an early diverging angiosperm lineage, EBS1-IBS1 pairing is improved (Albertazzi et al., 1998), whereas in wheat, editing would weaken the base-pairing (Li-Pook-Than et al., *Physiologia Plantarum*, in press). In addition to interactions with upstream exon sequences, group II introns also have an important long-range interaction, designated δ - δ', between the first nucleotide of the downstream exon and an intronic site (Michel and Ferat, 1995; Bonen and Vogel, 2001). Interestingly, a chloroplast *ndhA* (+13) exon was found to be unedited in intron-containing transcripts in barley leaves, and this observation prompted the suggestion that editing status at this site might influence intron folding within domains V and VI (which are crucial for the ribozymic splicing reaction) (del Campo et al., 2000).

6.2.1 Coordination of editing and transcription during embryo-to-seedling development

Based our studies, a lag in effective editing during early seedling development relative to plant mitochondrial transcriptional activity is also proposed (summarized in Figure 6.1B). This limited RNA editing was observed in precursor transcripts of *cox2* in dormant and germinating

wheat seeds, while the respective sites were fully edited in 6 day seedlings (Chapter 5). Incomplete editing throughout precursor transcripts, including at positions close to the exon/intron junction, was also observed in suspension cells of Petunia *cox2* and trans-spliced *nad1* (Sutton et al., 1991). Interestingly, heat stress was found to be correlated to an increase in steady state levels of *rps14* and *rpl20* transcripts in maize chloroplasts, while their relative editing efficiencies were found to decrease (Nakajima and Mulligan, 2001). Analogous to splicing and editing, which appear inefficiently coordinated with transcription during early seedling germination, RNA processing in general, as well as RNA turnover events, may be delayed during this period. Choury et al. (2005) used the electroporation in organelle transformation technique to introduce potato *rps10* group II intron into wheat mitochondria. They found that *rps10* genes containing wheat promoters, and not potato ones, were transcribed but not spliced or edited. This study proposed that transcription works in coordination with RNA processing events, such as editing and splicing, as part of a multiprotein super-complex.

6.3 Multiple splicing mechanisms in wheat mitochondria

The conventional model of group II intron removal describes the use of a bulged adenosine located at the 3' region of the intron as the nucleophile attacking the 5' end of the intron and subsequently resulting in the formation of a lariat structure. This was the case for excised *nad2* intron 4 molecules as observed in this work. However, the most unexpected finding in this present research was the identification of an *in vivo* mixture of atypical excised physical forms for the *cox2* intron that were distinct from those of *nad1* intron 2, and neither of which were majority lariats. This raises questions about the potential variability of excised physical forms for other introns found in plant mitochondria, particularly those that appear degenerate in higher order structure (Figure 1.4). Are the aberrant plant mitochondrial introns more 'flexible' in their biochemistry of splicing? We propose several models for splicing *in vivo* that include the conventional two-step transesterification pathway, a hydrolytic pathway, as well as proposed alternate nucleophile mechanism(s) internal or external to the intron (Figure 4.4). As more *in vivo* studies on the excision of group II introns are being performed, a few instances of unusual excised forms are being found, namely linear introns $tRNA^{val}$ in barley chloroplasts (Vogel and Borner, 2002), as well as circularized introns in both aI5γ in yeast mitochondria (Murray et al., 2001) and nitrogen-fixing bacteria (Molino-Sanchez et al., 2006). It is becoming

Table 6.1: Sequence context of the seven wheat mitochondrial exon editing sites located within 6 nucleotides of splice junctions and their editing status in precursor RNAs of 24 hr germinating wheat embryos.

[A] Exons (upper case) and introns (lower case) are shown with the candidate edit sites being boxed and their codon context denoted by underlining. The vertical line represents the splice junction position. The gene name (with underlining to denote exon containing the edit site of interest) is followed by the distance (in parentheses) of the edit site downstream (+ nt) or upstream (- nt) of the splice junction. Their editing status in wheat 24 hr precursor RNA is summarized by plus and minus symbols on the right. **[B]** Summary of physical form (lariat, circular or linear) of excised intron adjacent to respective exon editing junction site.

Edit site	Splice junction	Editing status of pre-mRNA	Physical form of adjacent group II intron
nad2de (-4)	GCGTTATAGGTCGTT ↓ gggcggccggaagg	+	Lariats
nad4bc (+1)	Aatattaggctctat ↓ CGAACATACAGGGAA	-	n/a (odd branchpoint)
nad4cd (+6)	aactattgaccctac ↓ TTGTTCGGATGGGTG	-	Lariats (preliminary linear)
nad5de (-5)	CAAAGTGGATCTGTT ↓ gtgcgaagagtgct	+	n/a (dVI bulged A)
nad7cd (-1)	TATTGTTAAGAGGTC ↓ gtgcgacatgaagac	-	Lariats (minor linear)
nad7de (+3)	atcggtctactctac ↓ CTCAATTCACCATTT	-	Lariats (circularized, minor linear)
cox2ab (-4)	ATCAATGGTATCGGA ↓ gtgcgcctcttaacg	+/-	Mixed (circular, linear)

more evident that additional splicing pathways, such as hydrolysis, are playing a more dominant role than previously thought.

It will also be of interest to examine the physical forms of the trans-spliced introns. Following conventional splicing the excised trans-introns would be "Y" shaped but predicted to be unstable. The 'Y' form could be tested by using inverse RT-PCR over the junction. In the case of *nad1* intron1, a branchpoint A is not present in dVI of this intron (Figure 4.1) and it is possible that another nucleophile is available for the first transesterification step to occur, analogous to one of our proposed splicing models for *nad1* intron 2 observations (Chapter 4). Another potential avenue of study would be to determine whether the physical forms of excised introns in wheat mitochondria are observed in the eudicot plants such as in the family Fabaceae.

An additional intriguing finding is the observation of the non-intron encoded heptamer (ACAAAAC) in the circularized forms of excised *nad1* intron 2, lacking the dVI branchpoint adenosine. One of the splicing pathways proposed suggests that ACAAAAC acts as an external nucleophile, analogous to the external guanosine for group I introns (Figure 4.4). A potential source of this heptamer would be in investigating the non-coding regions of the Arabidopsis mitochondrial genome which were found to be transcribed and processed (Holec et al., 2006). The ACAAAAC could be acting as a 'guide RNA', similar to the editing mechanism in trypanosomes. Here, in the case of intron removal, these small RNAs would recognize and interact with splice sites or motifs specific to each intron.

6.3.1 Evolutionary history of plant mitochondrial group II introns

The parallels between the splicing of organellar group II introns and nuclear spliceosomal introns (Figure 1.3) have led to the belief that they share a common ribozymic ancestor (Valadkhan, 2005). Little is known about the machinery involved in group II intron splicing, but the use of spliceosomal-like snRNAs cannot be ruled out (Mattick, 1994). Group II and group I-type introns also share similarities, having ribozymic and retrotransposable activity, encoding a RNA maturase, and splicing via two transesterification steps, although group I introns use an external guanosine (reviewed in Lehmann and Schmidt, 2003). With the discovery of group II introns in bacterial genomes (reviewed in Ferat and Michel, 1995), it has been speculated that mitochondrial group II introns were first introduced as early as pro-mitochondria of the endosymbiont (~1.5-1.8 billion years ago) (reviewed in Burger et al., 2003). Another hypothesis for the invasion of group II introns is through genetic recombination or retro-transposition from a

bacterial or even fungal source as proposed for the group I intron of *cox1* in *Peperomia* mitochondria (Cho et al., 1998).

Notably, 19 of the 26 plant mitochondrial group II introns are found in the NADH dehydrogenase genes (*nad*). This predisposition of these introns to the complex I genes suggest a regulatory role, that is notable in plants since alternative respiratory pathways are available that can bypass complex I of the electron transport chain (via a rotenone-insensitive pathway) (Moller, 2001). An interesting example is the tobacco CMSII mutant, which entirely lacks complex I and instead uses the rotenone-insensitive pathway for respiration, albeit resulting in diminished physiological traits (Pineau et al., 2005). Previous observations have shown that ATP synthesis, during the early stages of development, is dependent on an external NADH dehydrogenase, and not complex I (*nad* genes) (Logan et al., 2001). This was similarly shown in the mitochondria of rice embryos 2 hr post-imbibition, where the NADH dependent complex I pathway is utilized (Howell et al., 2006).

6.4 The stability and turnover of excised wheat mitochondrial group II introns

In plant mitochondria, transcript stability is believed to be linked to RNA secondary structure elements that are able to bind to protein complexes for the maturation and degradation of the mRNA (Hoffmann et al., 2001). The stable stem loop of dVI in *nad1* intron 2 may confer such stability when excised as linear molecules. The endogenous circularization step proposed in our splicing models, for both *cox2* intron and *nad1* intron 2 spliced molecules (Figure 4.4), would also promote stability. Although no endogenous RNA ligases have yet been identified in plant mitochondria, multifunctional genes similar in sequence to that of plant tRNA ligases have been found in both Arabidopsis and rice genomes, and have been suggested to have other essential functions apart from pre-tRNA splicing (Englert and Beier, 2005). Three isoforms of wheat cytosolic RNA ligases were found to circularize linear uncapped RNAs (Makino et al., 2006). Using the mitochondria targeting programs, MitoProtII 1.0a4 (Claros and Vincens, 1996), TargetP 1.1 (Emanuelsson et al., 2000) and Predotar v. 1.03 (Small et al., 2004), I found that in two out of three cases, isoform 2 and isoform 3 of the wheat RNA ligases were predicted to be targeted to the mitochondrion.

The examination of excised wheat *cox2* intron molecules revealed a mixed population of physical forms including linearized molecules that appear to have undergone 3' exonuclease

activity, and in some cases, endogenous circularization (Chapter 4). This 3' exonuclease activity is consistent with mRNA turnover events observed for pea *atp9* and *atp1* mitochondrial transcripts (Dombrowski et al., 1997; Kuhn et al., 2001). The 3' end polyadenylation promotes nucleolytic activity, typically observed with bacterial RNAs, in contrast to the 3' poly A tracts rendering stability in cytosolic RNAs (Gagliardi et al., 2001). Indeed, for excised *cox2* introns, I observed polyadenylation for both full length linear and 3' degraded linear ends (1-5 As), while the 5' ends remained unaltered (Table 4.1). Through northern analysis, I also observed more abundant excised introns in germinating seeds than in 6 day seedlings (Chapter 3). Similar to the lag experienced in splicing machinery during early developmental stages (Figure 6.1), the RNA turnover machinery, such as a degradosome, may be more efficient at the later seedling stages of development. Debranching enzymes are possible candidates for the turnover machinery of lariat introns as observed with yeast counterparts that were able to digest yeast mitochondrial group II lariats (Nam et al., 1994). When I performed a BLAST search, a rice lariat debranching enzyme was found, although it was not predicted to be targeted to plant mitochondria using the three targeting programs (as above: MitoProtII 1.0a4; Clarens and Vincens, 1996, TargetP 1.1; Emanuelsson et al., 2000 and Predotar v. 1.03; Small et al., 2004).

6.5 Potential trans-factors involved in group II introns splicing in plant mitochondria

My observations of multiple physical forms of excised group II introns *in vivo* within wheat mitochondria suggests the presence of novel proteins that are involved in these variously proposed splicing pathways. This addresses the questions proposed for this research such that aberrant plant mitochondrial introns can utilize non-convential pathways of splicing. In fact, little is known about the machinery that is involved in classical lariat splicing in plant mitochondria. Presently, only one tobacco mutant (*nms1)* has been found to be involved with the splicing machinery of only group II *nad4* intron 1 (Brangeon et al., 2000). This defective gene was recently identified to encode a nMat-R protein in Arabidopsis by Nakagawa and Sakurai (2006). Moreover, several copies of the group II intron related maturase gene were found in both rice and *Arabidopsis* nuclear genomes, and may potentially be imported back into the organelle for splicing activity (Mohr and Lambowitz, 2003). The intronic-encoded maturase RIR-3 was recently demonstrated in the cyanobacteria *Trichodemium* to aid in the general splicing of other maturase-less group II introns containing genes (Meng et al., 2005).

The defining characteristics of RNA splicing proteins may involve RNA recognition and/or RNA binding motifs. Such proteins may include the pentatricopeptide family (PPR) of proteins which have recently been found to interact closely with plant mitochondrial transcripts, particularly in the RNA editing in Arabidopsis plastids (Small and Peeters 2000; Kotera et al., 2005). Of the ~200 PPR proteins encoded in the Arabidopsis nucleus, two thirds have been predicted to be targeted to the mitochondria or chloroplast (Bentolila et al., 2002). Analysis of rice mitochondrial proteome revealed two PPR proteins of unknown function (Heazlewood et al, 2004). Another set of binding proteins associated with post-transcriptional events in higher plant mitochondria is the RRM type (N-terminal RNA-recognition motif) RNA binding proteins (Vermel et al., 2002). DEAD box RNA helicases were also found to bind to RNA and have been associated with group I and group II intron splicing in yeast (Huang et al, 2005). Similarly, an Arabidopsis Mg^{2+} transporter was shown to complement a yeast mitochondrial group II intron splicing mutant (Schock et al., 2000). Other potential types of group II intron splicing machinery in plant mitochondria could be related to counterparts found in plastids. These include CRS1 and CRS2 proteins which were found to be involved in the group II intron splicing in maize chloroplasts (Jenkins et al., 1997; Till et al., 2001). They are associated with the splicing of nine group II introns, and interact in a complex involving the CRS2 associated factors (CAF1 and CAF2) (Ostheimer et al., 2005).

Experimental approaches that can be used to obtain clues about the proteins involved in post-transcriptional processing (spliceosome-like, editosome-like, or degradosome-like machinery) include gel shift assays, which can be performed to observe the proteins that interact with different types of RNA (specifically dV/dVI intronic sequences). Another feature of proteins being imported into mitochondria is that their mRNAs are often first targeted to the surface of the organelles, and then translated in polysomes attached to the outer mitochondrial surface. Long mRNAs of nuclear encoded mitochondrial transcripts, linked to prokaryotic origin, have been found localized close to yeast mitochondria (Sylvestre et al., 2003). It would be interesting to examine whether long mRNAs also occur in plant mitochondria and if their abundance shifts during early seedling development.

6.6 Concluding remarks

RNA processing events in plant mitochondria, such as group II intron splicing and C-to-U editing are intricate and in some cases can be co-dependent on one another, as observed in this work. The variability in secondary structure of plant mitochondrial group II introns, including those that lack the dVI branchpoint A, add another level of complexity to gene expression. For certain group II introns, conventional excision resulting in the formation of lariats was not predominant. Several new models for the biochemistry of splicing are proposed to explain the diversity in excised physical forms found among populations *in vivo*, including the use of potential alternative nucleophiles, as in the case of *nad1* intron 2. Within the same population, wheat mitochondrial introns were also found to excise via multiple pathways. This is seen with excised *cox2* introns, which are postulated to undergo 3' exonuclease activity and endogenous circularization, as well as full-length linearization and polyadenylation. Splicing factors that may be involved in these multiple pathways include the recruitment of small RNAs (snRNA-like) and/or novel proteins. These cis- or trans-acting factors may replace the features and functions that are found to be 'degenerative' in plant mitochondrial group II introns (Figure 4.1). The availability of this excision machinery appear to be inefficiently coordinated with transcription during the period when embryos leave dormancy and begin germinating, while in later stages of development they are more effectively coupled. Identification and isolation of the proteins involved in splicing and editing, may serve as useful molecular tools for genetic research. The study of RNA processing events affecting wheat mitochondrial genes helps provide a better understanding of basic molecular processes that affect flowering plants, and may be useful for agricultural and bioenergetic research involving crop plants for bio fuels.

REFERENCES

Adams, K.L. and Palmer, J.D. (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* 29, 380-395.

Ahlert, D., Piepenburg, K., Kudla, J., and Bock, R. (2006) Evolutionary origin of a plant mitochondrial group II intron from a reverse transcriptase/maturase-encoding ancestor. *J Plant Res* 119, 363-371.

Albertazzi, F.J., Kudla, J., and Bock, R. (1998) The *cox2* locus of the primitive angiosperm plant Acorus calamus: molecular structure, transcript processing and RNA editing. *Mol Gen Genet* 259, 591-600.

Beagley, C.T., Okada, N.A., and Wolstenholme, D.R. (1996) Two mitochondrial group I introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. *PNAS* 93, 5619-5623.

Beltrán-Peña, E., Ortíz-López, A., Sánchez de Jiménez, E., and Abbas, G.M. (1995) Synthesis of ribosomal proteins from stored mRNAs early in seed germination. *Plant Mol Biol* 28, 327-336.

Bentolila, S., Alfonso, A.A., and Hanson, M.R. (2002) A pentatricopeptide repeat-containing gene restore fertility in cytoplasmic male-sterility plants. *Proc Natl Acad Sci U S A* 99, 10887-10892.

Bewley, J.D. and Black, M. (1994) Seeds: physiology of development and germination, 2nd ed. Plenum. *New York.*

Binder, S., Marchfelder, A., and Brennicke, A. (1996) Regulation of gene expression in plant mitochondria. *Plant Mol Biol* 32, 303-314.

Binder, S. and Brennicke, A. (2003) Gene expression in plant mitochondria: transcriptional and post-transcriptional control. *Philos Trans R Soc Lond B Biol Sci* 358, 181-189.

Bock, R., Hermann, M., and Fuchs, M. (1997) Identification of critical nucleotide positions for plastid RNA editing site recognition. *RNA* 3, 1194-1200.

Bock, R. and Koop, H.U. (1997) Extraplastidic site-specific factors mediate RNA editing in chloroplasts. *EMBO J* 16, 3282-3288.

Bolle, N. and Kempken, F. (2006) Mono- and dicotyledonous plant-specific RNA editing sites are correctly edited in both *in organello* systems. *FEBS Lett* 580, 4443-4448.

Bonen, L., Boer, P.H., and Gray, M.W. (1984) The wheat cytochrome oxidase subunit II gene has an intron insert and three radical amino acid changes relative to maize. *EMBO J* 3, 2531-2536.

Bonen, L., Boer, P.H., McIntosh, J.E., and Gray, M.W. (1987) Nucleotide sequence of the wheat mitochondrial gene for subunit I of cytochrome oxidase. *Nucleic Acids Res* 15, 6734-6734.

Bonen, L. and Bird, S. (1988) Sequence analysis of the wheat mitochondrial *atp6* gene reveals a fused upstream reading frame and markedly divergent N termini among plant ATP6 proteins. *Gene* 73, 47-56.

Bonen, L., Bird, S., and Belanger, L. (1990) Characterization of the wheat mitochondrial orf25 gene. *Plant Mol Biol* 15, 793-795.

Bonen, L. and Brown, G.G. (1993) Genetic plasticity and its consequences: perspectives on gene organization and expression in plant mitochondria. *Can J Bot* 71, 645-660.

Bonen, L. (1993) Trans-splicing of pre-mRNA in plants, animals, and protists. *FASEB J* 7, 40-46.

Bonen, L., Williams, K., Bird, S., and Wood, C. (1994) The NADH dehydrogenase subunit 7 gene is interrupted by four group II introns in the wheat mitochondrial genome. *Mol Gen Genet* 244, 81-89.

Bonen, L. and Vogel, J. (2001) The ins and outs of group II introns. *Trends Genet* 17, 322-331.

Boulanger, S.C., Belcher, S.M., Schmidt, U., Dib-Hajj, S.D., Schmidt, T., and Perlman, P.S. (1995) Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol Cell Biol* 15, 4479-4488.

Brangeon, J., Sabar, M., Gutierres, S., Combettes, B., Bove, J., Gendy, C., Chetrit, P., Colas des Francs-Small, C., Pla, M., Vedel, F., and De Paepe, R. (2000) Defective splicing of the first nad4 intron is associated with lack of several complex I subunits in the Nicotiana sylvestris NMS1 nuclear mutant. *Plant J* 21, 269-280.

Brennicke, A., Zablen, L., Dombrowski, S., Hoffman, M., and Binder, S. (1999) Transcription signal of mitochondrial and nuclear genes for mitochondrial proteins in dicot plants. *J Hered* 90, 345-350.

Burger, G., Gray, M.W., and Lang, B.F. (2003) Mitochondrial genomes: anything goes. *Trends Genet* 19, 709-716.

Carrillo, C. (2001) RNA splicing and editing of group II introns in flowering plant mitochondria. *Ph.D.Thesis* University of Ottawa, Ontario, Canada.

Carrillo, C. and Bonen, L. (1997) RNA editing status of nad7 intron domains in wheat mitochondria. *Nucleic Acids Res* 25, 403-409.

Carrillo, C., Chapdelaine, Y., and Bonen, L. (2001) Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol Gen Genet* 264, 595-603.

Cavalier-Smith, T. (1991) Intron phylogeny: a new hypothesis. *Trends Genet* 7, 145-148.

Cech, T.R. (1993) The efficiency and versatility of catalytic RNA: implications for an RNA world. *Gene* 135, 33-36.

Chanfreau, G. and Jacquier, A. (1996) An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J* 15, 3466-3476.

Chapedelaine, Y. (1992) Structure and expression of the trans-split gene for NADH dehydrogenase subunit 1 in wheat mitochondria. *Ph.D.Thesis* University of Ottawa, Ontario, Canada.

Chateigner-Boutin, A. and Hanson, M.R. (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar *cis*-element. *Nucleic Acids Res* 31, 2586-2594.

Cho, Y., Qiu, Y.L., Kuhlman, P., and Palmer, J.D. (1998) Explosive invasion of plant mitochondria by a group I intron [see comments]. *Proc Natl Acad Sci U S A* 95, 14244-14249.

Choury, D., Farré, J.C., Jordana, X., and Araya, A. (2005) Gene expression studies in isolated mitochondria: *Solanum tuberosum rps10* is recognized by cognate potato but not by the transcription, splicing and editing machinery of wheat mitochondri. *Nucleic Acids Res* 33, 7058-7065.

Claros, M.G. and Vincens, P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur.J.Biochem.* 241, 779-786.

Clifton, S.W., Minx, P., Fauron, C.M.R., Gibson, M., Allen, J.O., Sun, H., Thompson, M., Barbazuk, W.B., Kanuganti, S., Tayloe, C., Meyer, L., Wilson, R.K., and Newton, K.J. (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol* 136, 3486-3503.

Copertino, D.W. and Hallick, R.B. (1993) Group II and group III introns of twintrons: potential relationships with nuclear pre-mRNA introns. *Trends Biochem Sci* 18, 467-471.

Copertino, D.W., Hall, E.T., Van Hook, F.W., Jenkins, K.P., and Hallick, R.B. (1994) A group III twintron encoding a maturase-like gene excises through lariat intermediates. *Nucleic Acids Res* 22, 1029-1036.

Costa, M., Dème, E., Jacquier, A., and Michel, F. (1997) Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol* 267, 520-536.

Dai, L. and Zimmerly, S. (2002) Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res* 30, 1091-1102.

Daley, D.O., Considine, M.J., Howell, K.A., Millar, A.H., Day, D.A., and Whelan, J. (2003) Respiratory gene expression in soybean cotyledons duting post-germinative development. *Plant Mol Biol* 51, 745-755.

Daniels, D.L., Michels, W.J., and Pyle, A.M. (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. *J Mol Biol* 256, 31-49.

del Campo, E.M., Sabater, B., and Martin, M. (2000) Transcripts of the *ndh-D* operon of barley plastids: possible role of unedited site II in splicing of the *ndhA* intron. *Nucleic Acids Res* 28, 1092-1098.

Dombrovska, O. and Qiu, Y.L. (2004) Distribution of introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implications. *Mol Phylogenet Evol* 32, 246-263.

Dombrowski, S., Brennicke, A., and Binder, S. (1997) 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J* 16, 5069-5076.

Ehrenshaft, M. and Brambl, R. (1990) Respiration and mitochondrial biogenesis in germinating embryos of Maize. *Plant Physiol* 93, 295-304.

Elhafez, D., Murcha, M.W., Clifton, R., Soole, K.L., Day, D.A., and Whelan, J. (2006) Characterization of mitochondrial alternative NAD(P)H dehydrogenases in Arabidopsis: intraorganelle location and expression. *Plant Cell Physiol* 47, 43-54.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300, 1005-1016.

Englert, M. and Beier, H. (2005) Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins phylogenetic origins. *Nucleic Acids Res* 33, 388-399.

Faivre-Nitschke, S.E., Grienenberger, J., and Gualberto, J. (1999) Looking for an editing emzyme: study of plant cytidine deaminase genes and activities. *Eur J Biochem* 263, 896-903.

Farré, J.C., Leon, G., Jordana, X., and Araya, A. (2001) *cis* recognition elements in plant mitochondrion RNA editing. *Molecular and cellular Biolog* 21, 6731-6737.

Farré, J.C. and Araya, A. (2001) Gene expression in isolated plant mitochondria: high fidelity of transcription, splicing and editing of a transgene product in electroporated organelles. *Nucleic Acids Res* 29, 2484-2491.

Farré, J.C. and Araya, A. (2002) RNA splicing in higher plant mitochondria: determination of functional elements in group II intron from a chimeric *cox II* gene in electroorated wheat mitochondria. *The plant journal*

Fedor, M.J. and Williamson, J.R. (2005) The catalytic diversity of RNAs. *Nat Rev Mol Cell Biol* 6, 399-412.

Fedorova, O., Mitros, T., and Pyle, A.M. (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site. *J Mol Biol* 330, 197-209.

Ferat, J.L. and Michel, F. (1993) Group II self-splicing introns in bacteria [see comments]. *Nature* 364, 358-361.

Gagliardi, D., Perrin, R., Marechal-Drouard, L., Grienenberger, J. , and Leaver, C.J. (2001) Plant mitochondrial polyadenylated mRNAs are degraded by a 3'- to 5'-exoribonuclease activity, which proceeds unimpeded by stable secondary structures. *J Biol Chem* 276, 43541-43547.

Gagliardi, D., Stepien, P.P., Temperley, R.J., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M. (2004) Messenger RNA stability in mitochondria: different means to an end. *Trends Genet* 20, 260-267.

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

Giegé, P., Hoffman, M., Binder, S., and Brennicke, A. (2000) RNA degradation buffers asymmetries of transcription in *Arabidopsis* mitochondria. *EMBO reports* 1, 164-170.

Granlund, M., Michel, F., and Norgren, M. (2001) Mutually exclusive distibution of IS*1548* and GBSi1, an active group II intron identified in human isolates of group B streptococci. *J Bacteriol* 183, 2560-2569.

Gray, M.W. (1999) Evolution of organellar genomes. *Curr Opin Genet Dev* 9, 678-687.

Gray, M.W. (2003) Diversity and evolution of mitochondrial RNA editing systems. *IUBMB Life* 55, 227-233.

Grosskopf, D. and Mulligan, R.M. (1996) Developmental- and tissue-specificity of RNA editing in mitochondria of suspension-cultured maize cells and seedlings. *Curr Genet* 29, 556-563.

Groth-Malonek, M., Pruchner, D., Grewe, F., and Knoop, V. (2005) Ancestors of trans-splicing mitochondrial introns support serial sister group relationships of hornworts and mosses with vascular plants. *Mol Biol Evol* 1, 117-225.

Gualberto, J.M., Weil, J.H., and Grienenberger, J.M. (1990) Editing of the wheat coxIII transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. *Nucleic Acids Res* 18, 3771-3776.

Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res* 31, 5907-5916.

Haugen, P., Simon, D.M., and Bhattacharya, D. (2005) The natural history of group I introns. *Trends Genet* 21, 111-119.

Heazlewood, J.L., Tonti-Filippini, J.S., Gout, M., Day, D.A., Whelan, J., and Millar, A.H. (2004) Experimental analysis if the Arabidopsis mitochondrial proteome highlights and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *The Plant Cell* 16, 241-256.

Hedtke, B., Wagner, I., Borner, T., and Hess, W.R. (1999) Inter-organellar crosstalk in higher plants: impaired chloroplast development affects mitochondrial gene and transcript levels. *Plant J* 19, 635-643.

Hiesel, R., Wissinger, B., and Brennicke, A. (1990) Cytochrome oxidase subunit II mRNAs in Oenothera mitochondria are edited at 24 sites. *Curr Genet* 18, 371-375.

Higuchi, M., Single, F.N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P.H. (1993) RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* 75, 1361-1370.

Hoffman, M., Kuhn, J., Daschner, K., and Binder, S. (2001) The RNA world of mitochondria. *Progress in Nucleic acid research and molecular biology* 70 , 119-154.

Holec, S., Lange, H., Kuhn, K., Alioua, M., Borner, T., and Gagliardi, D. (2006) Relaxed transcription in Arabidopsis mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Mol Cell Biol* 26, 2869-2876.

Holländer, V. and Kück, U. (1999) Group II intron splicing in chloroplasts: identification of mutations determining intron stability and fate of exon RNA. *Nucleic Acids Res* 27, 2345-2353.

Howell, K.A., Millar, A.H., and Whelan, J. (2006) Ordered assembly of mitochondria during rice germination begins with pro-mitochondrial structures rich in components of the protein import apparatus. *Plant Mol Biol* 60, 201-223.

Huang, H.R., Rowe, C.E., Mohr, S., Jiang, Y., Lambowitz, A.M., and Perlman, P.S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc Natl Acad Sci U S A* 102, 163-168.

Huppler, A., Nikstad, L.J., Allmann, A.M., Brow, D.A., and Butcher, S.E. (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol* 9, 431-435.

Ikeda, T.M. and Gray, M.W. (1999) Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Mol Cell Biol* 19, 8113-8122.

Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L., and Perlman, P.S. (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products. *J Biol Chem* 263, 3432-3439.

Jenkins, B.D., Kulhanek, D.J., and Barkan, A. (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell* 9, 283-296.

Jestin, J.L., Dème, E., and Jacquier, A. (1997) Identification of structural elements critical for inter-domain interactions in a group II self-splicing intron. *EMBO J* 16, 2945-2954.

Jurica, M.S. and Moore, M.J. (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell* 12, 5-14.

Kim, J.K. and Hollingsworth, M.J. (1993) Splicing of group II introns in spinach chloroplasts (in vivo): analysis of lariat formation. *Curr Genet* 23, 175-180.

Klein, R.R., Klein, P.E., Mullet, J.E., Minx, P., Rooney, W.L., and Schertz, K.F. (2005) Fertility restorer locus Rf1 of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome *12*. *Theor Appl Genet* 111, 994-1012.

Knoop, V., Altwasser, M., and Brennicke, A. (1997) A tripartite group II intron in mitochondria of an angiosperm plant. *Mol Gen Genet* 255, 269-276.

Knoop, V. (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr Genet* 46, 123-139.

Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K., and Perlman, P.S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity. *Mol Cell Biol* 12, 1950-1958.

Kotera, E., Tasaka, M., and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* 433, 326-330.

Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A., and Mikami, T. (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for $tRNA^{Cys(GCA)}$. *Nucleic Acids Res* 28, 2571-2576.

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

Kuhn, J., Tengler, U., and Binder, S. (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Molecular and cellular biology* 21, 731-742.

Kurihara-Yonemoto, S. and Handa, H. (2001) Low temperature affects the processing pattern and RNA editing status of the mitochondrial *cox2* transcripts in wheat. *Curr Genet*

Lambowitz, A.M., Caprara, M.G., Zimmerly, S., and Perlman, P. (1999) Group I and group II ribozymes as RNPs: clues to the past and guides to the future (book). *the RNA world* chap 18, 451-485.

Lambowitz, A.M. and Zimmerly, S. (2004) Mobile group II introns. *Annu Rev Genet* 38, 1-35.

Lehmann, K. and Schmidt, U. (2003) Group II introns: Structure and catalytic versatility of large natural ribozymes. *Crit.Rev.Biochem.Mol.Biol.* 38, 249-303.

Li-Pook-Than, J., Carrillo, C., and Bonen, L. (2004) Variation in mitochondrial transcript profiles of protein-coding genes during early germination and seedling development in wheat. *Curr Genet* 46, 374-380.

Li-Pook-Than, J. and Bonen, L. (2006) Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res* 34, 2782-2790.

Liu, X.Q., Yang, J., and Meng, Q. (2003) Four inteins and three group II introns encoded in a bacterial ribonucleotide gene. *J Biol Chem* 278, 46826-46831.

Logan, D.C., Millar, A.H., Sweetlove, L.J., and Leaver, C.J. (2001) Mitochondrial biogenesis during germination in Maize embryos. *Plant Physiol* 125, 662-672.

Lykke-Andersen, J., Aagaard, C., Semionenkov, M., and Garrett, R.A. (1997) Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem Sci* 22, 326-331.

Lynch, M., Koskella, B., and Schaack, S. (2006) Mutation pressure and evolution of organelle genomic architecture. *Science* 311, 1727-1730.

Maas, S., Rich, A., and Nishikura, K. (2003) A-to-I RNA editing: recent news and residual mysteries. *J Biol Chem* 278, 1391-1394.

Makino, S., Sawasaki, T., Tozawa, Y., Endo, Y., and Takai, K. (2006) Covalent circularization of exogenous RNA during incubation with a wheat embryo cell extract. *Biochem Biophys Res Commun* 347, 1080-1087.

Malek, O. and Knoop, V. (1998) Trans-splicing group II introns in plant mitochondria: the complete set of cis-arranged homologs in ferns, fern allies, and a hornwort. *RNA* 4, 1599-1609.

Martinez-Abarca, F. and Toro, N. (2000) Group II introns in the bacterial world. *Mol Microbiol* 38, 917-926.

Mattick, J.S. (1994) Introns: evolution and function. *Curr Opin Genet Dev* 4, 823-831.

Meng, Q., Wang, Y., and iu, X.Q. (2005) An intron-encoded protein assists RNA splicing of multiple similar introns of different bacterial genes . *J Biol Chem* 280, 35085-35088.

Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.D., and Goldschmidt-Clermont, M. (2006) A novel multifunctional factor involved in trans-splicing of chloroplast introns in Chlamydomonas. *Nucleic Acids Res.* 34, 262-274.

Michel, F., Umesono, K., and Ozeki, H. (1989) Comparative and functional anatomy of group II catalytic introns--a review. *Gene* 82, 5-30.

Michel, F. and Ferat, J.L. (1995) Structure and activities of group II introns. *Annu Rev Biochem* 64, 435-461.

Mohr, G. and Lambowitz, A.M. (2003) Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res* 32, 647-652.

Molina-Sanchez, M.D., Martinez-Abarca, F., and Toro, N. (2006) Excision of the *Sinorhizobium meliloti* group II intron RmInt1 as circles *in vivo*. *J Biol Chem* 281, 28737-28744.

Moller, I.M. (2001) Plant mitochondria and oxidative stress: Electron Transport, NADPH turnover, and metabolism of reactive oxygen species. *Annu Rev Plant Physiol Plant Mol Biol* 52, 561-591.

Mower, J.P. and Palmer, J.D. (2006) Patterns of partial RNA editing in mitochondrial genes of *Beta vulgaris*. *Mol Genet Genomics* 276, 285-293.

Muise, R.C. and Hauswirth, W.W. (1992) Transcription in maize mitochondria: effects of tissue and mitochondrial genotype. *Curr Genet* 22, 235-242.

Mulligan, R.M., Leon, P., and Walbot, V. (1991) Transcriptional and posttranscriptional regulation of maize mitochondrial gene expression. *Mol Cell Biol* 11, 533-543.

Mulligan, R.M., Williams, M.A., and Shanahan, M.T. (1999) RNA editing site recognition in higher plant mitochondria. *J Hered* 90, 338-344.

Murray, H.L., Mikheeva, S., Coljee, V.W., Turczyk, B.M., Donahue, W.F., Bar-Shalom, A., and Jarell, K.A. (2001) Excision of group II introns as circles. *Mol Cell* 8, 201-211.

Nakagawa, N. and Sakurai, N. (2006) A mutation in At-nMat1a, which encodes a nuclear gene having high similarity to group II intron maturase, causes impaired splicing of mitochondrial NAD4 transcript and altered carbon metabolism in *Arabidopsis thaliana*. *Plant Cell Physiol* 47, 772-783.

Nakajima, K. and Mulligan, R.M. (2001) Heat stress results in incomplete C-to-U editing of maize chloroplast mRNAs and correlates with changes in chloroplast transcription rate. *Curr Genet* 40, 209-213.

Nakazono, M., Itadani, H., Wakasugi, T., Tsutsumi, N., Sugiura, M., and Hirai, A. (1995) The *rps3-rpl16-nad3-rps12* gene cluster in rice mitochondrial DNA is transcribed from alternative promoters. *Curr Genet* 27, 184-189.

Nam, K., Hudson, R.H., Chapman, K.B., Ganeshan, K., Damha, M.J., and Boeke, J.D. (1994) Yeast lariat debranching enzyme. Substrate and sequence specificity. *J Biol Chem* 269, 20613-20621.

Navartnam, N. and Sarwar, R. (2006) An overview of cytidine deaminases. *Int J Hematol* 83, 195-200.

Neuwirt, J., Takenaka, M., Van der Merwe, JA., and Brennicke, A. (2005) An *in vitro* RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. *RNA* 11, 1563-1570.

Nielsen, H., Fiskaa, T., Birgisdottir, A.B., Haugen, P., Einvik, C., and Johansen, S. (2003) The ability to form full-length intron RNA circles is a general property of nuclear group I introns. *RNA* 9, 1464-1475.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A., and 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 plant. *Mol Genet Genomics* 268, 434-445.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., and et.al. (1992) Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. *J Mol Biol* 223, 1-7.

Ogihara, Y., Yamazaki, Y., Murai, K., Kanno, A., Terachi, T., Shiina, T., Miyashita, N., Nasuda, S., Nakamura, C., Mori, N., Takumi, S., Murata, M., Futo, S., and Tsutsumi, N. (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res* 33, 6235-6280.

Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. (2003) Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J* 22, 3912-3929.

Ostheimer, G.J., Hadjivassiliou, H., Kloer, D.P., Barkan, A., and Matthews, B.W. (2005) Structural analysis of the group II intron splicing factor CRS2 yields insights into its protein and RNA interaction surface. *J Mol Biol* 345, 51-68.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., and Cheng, H.L. (1987) A self-splicing RNA excises an intron lariat. *Cell* 44, 213-223.

Petersen, G., Seberg, O., Yde, M., and Berthelsen, K. (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the **A**, **B**, and **D** genomes of common wheat (*Triticum aestivum*). *Mol Phylogenet Evol* 39, 70-82.

Pineau, B., Mathieu, C., Gerard-Hirne, C., De Paepe, R., and Chetrit, P. (2005) Targeting the NAD7 subunit to mitochondria restores a functional complex I and a wild type phenotype in the *Nicotiana sylvestris* CMS II mutant lacking *nad7*. *J Biol Chem* 280, 25994-26001.

Podar, M., Chu, V.T., Pyle, A.M., and Perlman, P.S. (1998) Group II intron splicing in vivo by first-step hydrolysis. *Nature* 391, 915-918.

Pyle, A.M. (1993) Ribozymes: a distinct class of metalloenzymes. *Science* 261, 699-700.

Qin, P.Z. and Pyle, A.M. (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr Opin Struct Biol* 8, 301-308.

Qiu, Y.L., Cho, Y., Cox, J.C., and Palmer, J.D. (1998) The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* 394, 671-674.

Qiu, Y.L. and Palmer, J.D. (2004) Many independent origins of trans splicing of a plant mitochondrial group II intron. *J Mol Evol* 59, 722-724.

Rasmusson, A.G., Soole, K.L., and Elthon, T.E. (2004) Alternative NAD(P)H dehydrogenases of plant mitochondria. *Annu Rev Plant Biol* 55, 23-39.

Regina, T.M., Picardi, E., Lopez, L., Pesole, G., and Quagliariello, C. (2005) A novel additional group II intron distinguishes the mitochondrial *rps3* gene in gymnosperms. *J Mol Evol* 60, 196-206.

Roy, S.W. and Gilbert, W. (2006) The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat Rev Genet* 7, 211-221.

Saisho, D., Nakazono, M., Lee, K.-H., Tsutsumi, N., Akita, S., and Hirai, A. (2001) The gene for alternative oxidase-2 (AOX2) form *Arabidopsis thaliana* consists of five exons unlike other AOX genes and is transcribed at an early stage during germination. *Genes Genet Syst* 76, 89-97.

Sakamoto, W., Tan, S.H., Murata, M., and Motoyoshi, F. (1997) An unusual mitochondrial *atp9-rpl16* cotranscript found in the maternal distorted leaf mutant of *Arabidopsis thaliana*: implication of GUG as an initiation codon in plant mitochondria. *Plant Cell Physiol* 38, 975-979.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning. *A Laboratory Manual, 2nd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sashital, D.G., Cornilescu, G., McManus, C.J., Brow, D.A., and Butcher, S.E. (2004) U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* 11, 1237-1242.

Schock, I., Gregan, J., Steinhauser, S., Schweyen, R.J., Brennicke, A., and Knoop, V. (2000) A member of a novel *Arabidopsis thaliana* gene family of candidate Mg^{+2} ion transporters complements a yeast mitochondrial group II intron-splicing mutant. *Plant J* 24, 489-501.

Schuster, W., Hiesel, R., Wissinger, B., and Brennicke, A. (1990) RNA editing in the cytochrome b locus of the higher plant Oenothera berteriana includes a U-to-C transition. *Mol Cell Biol* 10, 2428-2431.

Schuster, W. and Brennicke, A. (1990) RNA editing of ATPase subunit 9 transcripts in *Oenothera* mitochondria. *FEBS Lett* 268, 252-256.

Shikanai, T. (2006) RNA editing in plant organelles: machinery, physiological function and evolution. *Cell Mol Life Sci* 63, 698-708.

Sigel, R.K., Vaidya, A., and Pyle, A.M. (2000) Metal ion binding sites in a group II intron core. *Nat Struct Biol* 7, 1111-1116.

Simpson, L., Sbicego, S., and Aphasizhev, R. (2003) Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* 9, 265-276.

Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2005) Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. *Mol Cell Biol* 25, 6427-6435.

Small, I. and Peeters, N. (2000) The PPR motif- a TPR-related motif prevalent in plant organellar proteins. *TIBS* 25, 47

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4, 1581-1590.

Smart, C.J., Moneger, F., and Leaver, C.J. (1994) Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. *Plant Cell* 6, 811-825.

Stahley, M.R. and Strobel, S.A. (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr Opin Struct Biol* 16, 319-326.

Staudinger, M. and Kempken, F. (2003) Electroporation of isolated higher-plant mitochondria: transcripts of an introduced *cox2* gene, but not an *atp6* gene, are edited *in organello*. *Mol Gen Genomics* 269, 553-561.

Subramanian, S. (1999) Plant mitochondrial L2 ribosomal protein genes and their expression. *Ph.D.Thesis* University of Ottawa, Ontario, Canada.

Subramanian, S., Fallahi, M., and Bonen, L. (2001) Truncated and dispersed *rpl2* and *rps19* pseudogenes are co-transcribed with neightouring downstream genes in wheat mitochondria. *Curr Genet* 39, 264-272.

Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A., and 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, 603-615.

Sutton, C.A., Conklin, P.L., Pruitt, K.D., and Hanson, M.R. (1991) Editing of pre-mRNAs can occur before cis- and trans-splicing in Petunia mitochondria. *Mol Cell Biol* 11, 4274-4277.

Sylvestre, J., Vialette, S., Corral Debrinski, M., and Jacq, C. (2003) Long mRNAs coding for yeast mitochondrial proteins of prokaryotic origin preferentially localize to the vicinity of mitochondria. *Genome Biol* 4, 44

Takenaka, M., Neuwirt, J., and Brennicke, A. (2004) Complex *cis*-elements determine an RNA editing site in pea mitochondria. *Nucleic Acids Res* 32, 4137-4144.

Takvorian, A., Coville, J.L., Haouazine-Takvorian, N., Rode, A., and Hartmann, C. (1997) The wheat mitochondrial *rps13* gene: RNA editing and co-transcription with the *atp6* gene. *Curr Genet* 31, 497-502.

Till, B., Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2001) CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA* 7, 1227-1238.

Toor, N., Hausner, G., and Zimmerly, S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* 7, 1142-1152.

Trotta, C.R., Miao, F., Arn, E.A., Stevens, S.W., Ho, C.K., Rauhut, R., and Abelson, J.N. (1997) The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. *Cell* 89, 849-858.

Truscott, K.N., Pfanner, N., and Voos, W. (2001) Transport of proteins into mitochondria. *Rev Physiol Biochem Pharmacol.2001;143:81-136* 143, 81-136.

Turmel, M., Otis, C., and Lemieux, C. (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. *Proc Natl Acad Sci U S A* 99, 11275-11280.

Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997) The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. *Nat Genet* 15, 57-61.

Vaitilingom, M., Stupar, M., Grienenberger, J.M., and Gualberto, J.M. (1998) A gene coding for an *rps2* protein is present in the mitochondrial genome of several cereals, but not in dicotyledons. *Mol Gen Genet* 258, 530-537.

Valadkhan, S. (2005) snRNAs as the catalysts of pre-mRNA splicing. *Curr Opin Chem Biol* 9, 603-608.

van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986) Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* 44, 225-234.

van der Veen, R., Kwakman, J.H., and Grivell, L.A. (1987) Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J* 6, 3827-3831.

Vermel, M., Delage, L., Grienenberger, J.M., Marechal-Drouard, L., Gualberto, J.M., and Guermann, B. (2002) A family of RRM-type RNA-binding proteins specific to plant mitochondria. *Proc Natl Acad Sci U S A* 99, 5866-5871.

Vogel, J., Hess, W.R., and Börner, T. (1997) Precise branch point mapping and quantification of splicing intermediates. *Nucleic Acids Res* 25, 2030-2031.

Vogel, J. and Borner, T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J* 21, 3794-3802.

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., Su, D., Chen, Y., Guo, J., Luo, D., Long, Y., Zhong, Y., and 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, 676-687.

Wank, H., SanFilippo, J., Singh, R.N., Matsuura, M., and Lambowitz, A.M. (1999) A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. *Mol Cell* 4, 239-250.

Weiner, A.M. (1993) mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? *Cell* 72, 161-164.

Wolf, J., Gerber, A.P., and Keller, W. (2006) *tadA*, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J* 21, 3841-3852.

Wool, I.G. (1996) Extraribosomal functions of ribosomal proteins. *Trends Biochem Sci* 21, 164-165.

Xiang, Q., Qin, P.Z., Michels, W.J., Freeland, K., and Pyle, A.M. (1998) Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* 37, 3839-3849.

Yang, A.J. and Mulligan, R.M. (1991) RNA editing intermediates of cox2 transcripts in maize mitochondria. *Mol Cell Biol* 11, 4278-4281.

Zangar, R.C., Hernandez, M., Kocarek, T.A., and Novak, R.F. (1995) Determination of the polyA tail lengths of a single mRNA species in total hepatic RNA. *BioTechniques* 18, 465-469.

Zhang, X.P. and Glaser, E. (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the *hsp70* molecular chaperone. *Trends Plant Sci* 7, 14-21.

Zhuo, D. and Bonen, L. (1993) Characterization of the S7 ribosomal protein gene in wheat mitochondria. *Mol Gen Genet* 236, 395-401.

Zimmerly, S., Hausner, G., and Wu, X. (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* 29, 1238-1250.

APPENDIX

	P	P	L	0	7	7	0		B	Dispatch: 24.7.06	Journal: PPL	CE: Suchitra
	Journal Name				Manuscript No.					Author Received:	No. of pages: 11	ME: Varma

Physiologia Plantarum 0: 1–11. 2006

Relationship between RNA splicing and exon editing near intron junctions in wheat mitochondria

Jennifer Li-Pook-Than, Catherine Carrillo, Nima Niknejad, Sophie Calixte, Jennifer Crosthwait and Linda Bonen*

Biology Department, University of Ottawa, 30 Marie Curie, Ottawa, Canada K1N 6N5

Correspondence
*Corresponding author,
e-mail: lbonen@science.uottawa.ca

Received 31 March 2006; revised 25 April 2006

doi: 10.1111/j.1399-3054.2006.00770.x

We have compared the RNA editing status of wheat mitochondrial unspliced precursor transcripts with spliced RNAs, focusing in particular on exon editing sites located very close to intron/exon junctions. Using direct sequencing of reverse transcriptase–PCR products to assess C-to-U editing in various RNA populations for the respiratory chain genes *nad2*, *nad4*, *nad5*, *nad7*, and *cox2*, we observed that candidate exon sites immediately upstream or downstream of introns remained unedited in the presence of the adjacent intron, whereas sites further away were typically partially (or completely) edited in precursor molecules. The 'late' editing of exon sites adjacent to splice junctions is consistent with access of the editing machinery being sterically hindered by the intron or alternatively by an editing recognition element being created by the exon/exon structure. When we examined RNA isolated from the embryo-to-seedling stages of wheat development, we found that fully spliced messenger RNAs showed editing at the expected sites in all stages, including dormant seeds. In contrast, we observed that *cox2* precursors were less completely edited in RNA populations from 0 to 24 h postimbibition embryos than in 6-day-old seedlings not only at splice junctions but also at other exon sites, consistent with more efficient coupling between editing and transcription during later stages of development.

Introduction

RNA splicing and editing are crucial steps in the conversion of precursors to mature messenger RNAs (mRNAs) in plant mitochondria. Virtually all protein-coding sequences undergo C-to-U type editing at some sites, and such changes at the RNA level usually improve amino acid sequence similarity with counterparts from other organisms (reviewed in Maier et al. 1996, Shikanai 2006). Variation, however, in distribution of editing sites is observed among mitochondrial genes, as well as among homologues from different flowering plants. In addition to editing, about one-third of the protein-coding genes also require splicing for transcript maturation. Plant mitochondrial introns fall almost exclusively in the group II category, and a subset is discontinuous in the genome so that expression requires *trans*-splicing (reviewed in Bonen and Vogel 2001). Not all introns are present in all flowering plants, and a correlation between intron absence and a lack of need for editing at positions near former splice sites supports the idea that intron loss is RNA mediated, with spliced complementary DNAs (cDNAs) being integrated into the mitochondrial genome by homologous recombination (Geiss et al. 1994). Although editing occurs primarily within coding sequences, it has also been observed within introns, and in some cases converts A–C mispairs to A–U pairs within core intron helices, and so may be important for proper intron folding, and thus would obligatorily precede splicing.

Abbreviations – cDNA, complementary DNAs; mRNA, messenger RNA; RT-PCR, reverse transcriptase–PCR.

However, editing is not observed in all expected cases and variation in editing status of intronic A–C mispairs has been seen among plants (Carrillo et al. 2001).

Mitochondrial RNA processing events appear coordinated with transcription and although the expression machinery (for either editing or splicing) is just beginning to be elucidated (reviewed in Lehmann and Schmidt 2003, Shikanai 2006), it undoubtedly includes nuclear-encoded components, which are imported into the mitochondrion. Recent wheat in organelle gene expression studies suggest that editing may only occur in vivo when transcripts are associated with an active 'RNA maturation machinery complex' (Choury et al. 2005), and studies of yeast and animal mitochondria point to such expression complexes being membrane associated (reviewed in Shadel 2004). The efficiency of coupling of these expression events appears to shift during plant development, with transcription outpacing splicing during early germination of seeds compared with more efficient coordination at later stages of development based on studies in wheat (Li-Pook-Than et al. 2004). Seed germination is a period of rapid mitochondrial biogenesis and the roles of nuclear and mitochondrial-encoded components in this process are under investigation (cf. Logan et al. 2001). However, very little is known about the editing status of transcripts during development. A comparison of mitochondrial RNA from 3- to 7-day maize seedlings revealed a shift from 50 to 75% in the extent of editing of *nad3* transcripts; however, no difference was observed in the case of spliced *cox2* or *atp6* mRNAs (Grosskopf and Mulligan 1996). Interestingly, chloroplast studies have demonstrated that editing efficiency differs among genes and among given sites within a particular gene (ranging from less than 10 to 100%) between roots and green leaf tissue in maize (Chateigner-Boutin and Hanson 2003, Peeters and Hanson 2002).

Such observations raise interesting questions about the temporal relationship between splicing and editing. The view that editing is usually a relatively 'early' post-transcriptional event is well supported by editing data from precursor transcripts, for both intron-containing genes (cf. maize *cox2*, Yang and Mulligan 1991; petunia *cox2*, Sutton et al. 1991, potato *rps10*, Zanlungo et al. 1995) and ones which are cotranscribed but subsequently cleaved to monocistronic mRNAs (cf. wheat *nad3-rps12* Gualberto et al. 1991). Such precursor RNA populations have been shown to exhibit editing, although it may be incomplete and there is no polarity (from one end or the other), although certain sites may appear more 'readily' edited than others. We had observed an extreme example of this in the case of a *nad7* exon site located three nucleotides downstream of intron 4 (Carrillo and Bonen 1997). It exhibited no editing in precursor molecules (unlike other exon editing sites), yet was fully edited in spliced RNAs. This suggested that such relatively 'late' editing events may be related to the context of sequences around the junction. To extend this study, we have now examined additional exon editing sites, which are close to exon/intron junctions and compared their status in unspliced transcripts with those in spliced RNAs.

Materials and methods

Reverse transcriptase–PCR analysis of wheat mitochondrial transcripts

Mitochondrial RNA was isolated from dry wheat seeds (*Triticum aestivum* cv. Frederick), or from dissected embryos before imbibition with sterile water for 12 or 24 h at room temperature in the dark, or from etiolated shoot tissue harvested after growth for 6 days in vermiculite as previously described (Li-Pook-Than et al. 2004). To minimize contribution from contaminating DNA, mitochondrial RNA samples were successively treated three times with 2 U μl^{-1} DNase I (Amersham) in 32 m*M* Tris, pH 7.5, 5 m*M* $MgCl_2$, 0.1 m*M* DTT and then phenol extracted prior to RNA recovery by ethanol precipitation. [1]

Wheat mitochondrial RNA (approximately 5 μg) was denatured at 70°C for 10 min prior to cDNA synthesis using an intron-specific or terminal-exon–specific synthetic primer (oligomers shown in Table 1) and incubation with 10 U μl^{-1} Superscript II reverse transcriptase

Table 1. Synthetic oligomers used for reverse transcriptase–PCR (RT-PCR) and direct sequencing. Oligomers that were used solely for RT-PCR analysis are shown by [R].

Oligomer	5′–3′ sequence	Position
nad2 (Y14434)		
Oligo #1	cctaaggccaacttccattc	5′ end of intron 4
Oligo #2	gctacccacttcgatcaatt	5′ end of exon 3
Oligo #3 [R]	cttcatggtgaatggcccac	3′ UTR
Oligo #4	gtaacgacttgtcacgatcc	5′ end of exon 5
nad4 (X57164)		
Oligo #5	gttccggttcactgataagg	3′ end of intron 2
Oligo #6 [R]	tagtagccacattacgtgcg	Mid-region of exon 3
Oligo #7 [R]	atacccatgtttcccgaggc	Mid-region of exon 2
Oligo #8	tttctttgctgattcctctc	3′ end of exon 4
Oligo #9	gaatcaaagtccaggttggt	3′ end of intron 3
cox2 (X01108)		
Oligo #10	ccgagaagaggtatgtgtac	5′ end of intron
Oligo #11	caatcgctctttgtgatgct	5′ end of exon 1
Oligo #12	tgattggatacccaatccgc	3′ end of exon 2

[9]

(Invitrogen), 0.2 vol of 5× first strand buffer (Invitrogen), 10 m*M* DTT and 0.5 m*M* dNTPs at 45°C for 2 h. To examine exon editing sites of partially spliced (or precursor) molecules, the cDNAs were subsequently subjected to PCR amplification using *Taq* DNA polymerase (Invitrogen), and the cDNA synthesis primer was typically used in combination with one mapping to a distal exon (or intron) location (Table 1). When alternative primers were used for cDNA synthesis, similar results were obtained. These experiments were also repeated using different cDNA preparations and multiple mitochondrial RNA preparations.

Direct sequencing of reverse transcriptase–PCR products

Reverse transcriptase–PCR (RT-PCR) products were subsequently gel purified using UltraClean 15 (MoBio Laboratories, Inc.) and then sequenced directly using Sequenase version 2.0 (US Biochemicals) and electrophoresed on urea–polyacrylamide sequencing gels (Sambrook et al.1989). In the case of a *nad4d* edit within the terminal exon, RT-PCR products derived from precursors (which are colinear with DNA) were selected by their resistance to cleavage by *Msp1* restriction enzyme, because editing causes the loss of a *MspI* site (CCGG to TCGG). A similar strategy was used to select for RT-PCR products derived from edited *nad7*e-containing transcripts, where a *NheI* restriction site is generated by editing (Carrillo and Bonen 1997).

Results

Distribution of exon editing sites within intron-containing plant mitochondrial genes

To further explore the timing between editing and splicing events, we have focused on exon editing sites located close to intron borders and compared their status in RNAs in which the nearby intron is still present with ones in which it has been excised. We also examined more distal sites in such transcripts, the premise being that because editing is typically an 'early' RNA processing event, transcripts in which at least one intron has been excised might be expected to be edited. To facilitate our choice of sites, we surveyed the positions of known editing sites in coding sequences of wheat mitochondrial intron-containing genes (Fig. 1). A total of 26 group II–type introns have been identified within 10 protein-coding genes in the mitochondria of various seed plants, although not all are present within a given plant (reviewed in Knoop 2004). Fig. 1 shows the positions of introns (triangles for *cis*-splicing and Y-shaped symbols for *trans*-splicing) and editing sites (vertical lines) within the wheat mitochondrial genes encoding NADH dehydrogenase subunits (*nad1*, *nad2*, *nad4*, *nad5* and *nad7*) and cytochrome oxidase subunit 2 (*cox2*). Such data extend those previously compiled for wheat *nad* genes by Maier et al. (1996) in that we have included *cox2* (Covello and Gray 1990) and additional *nad1e* data (Y. Chapdelaine, PhD thesis, University of Ottawa). Because editing data were unavailable for wheat *rps3* and *ccmFc*, we have shown the counterparts from rice (Notsu et al. 2002), as well as for *rpl2* (Kubo et al. 1996), which is a truncated pseudogene in wheat (Subramanian et al. 2001). The *rps10* gene is nuclear-encoded in cereals such as wheat (Adams et al. 2000); however, in plants such as potato the mitochondrial-located gene has a single intron (Zanlungo et al. 1995) (Fig. 1). Other differences in intron composition have been observed among flowering plants for *nad4* (Gass et al. 1992), *nad7* (Sugiyama et al. 2005), *cox2* (reviewed in Bonen and Brown 1993) and *rps3* (Kubo et al. 2000, Regina et al. 2005), and introns, which are 'optionally' present among plants, are shown by dotted triangles in Fig. 1. In addition, the *cox1* gene in certain plants (such as *Peperomia*) contains a group I intron, which is believed to be of fungal origin and recently acquired through horizontal transfer (Adams et al. 1998).

In Fig. 1, it can be seen that editing sites are scattered along the coding sequences and they show a marked difference in density both among genes (cf. *nad5* vs *nad7*) and within exons of a given gene (cf. *nad4a* vs *nad4b*). The uneven distribution of editing sites has led to the suggestion that genes sporadically undergo retro-copy correction with edited mRNA acting as template in gene conversion (Geiss et al. 1994). Interestingly, the two intron-containing ribosomal protein genes in rice (*rps3* and *rpl2*) have very few editing sites, whereas the closely linked downstream unsplit *rpl16* and *rps19* have a much higher density (Fig. 1, stippled blocks). Notably the three-nucleotide spacer between *rpl2* and *rps19* also contains an edit site in rice (Kubo et al. 1996).

There does not appear to be any strong bias either for or against editing sites being close to exon/intron junctions, and in Fig. 1 those which are within 15 nucleotides of either an upstream or downstream intron are shown with a black dot. Of these 14 sites, nine are located upstream of a splice site and five are downstream. We selected a subset of these editing junction sites, namely, the seven in wheat, which were no more than six nucleotides away from an intron (Fig. 1, shaded boxes), to experimentally examine their editing status in populations of intron-containing precursor molecules compared with that of spliced transcripts. The sequences of these regions within wheat *nad2*, *nad4* (two sites), *nad5*, *nad7* (two sites) and *cox2*-coding sequences are shown in Fig. 2, with the

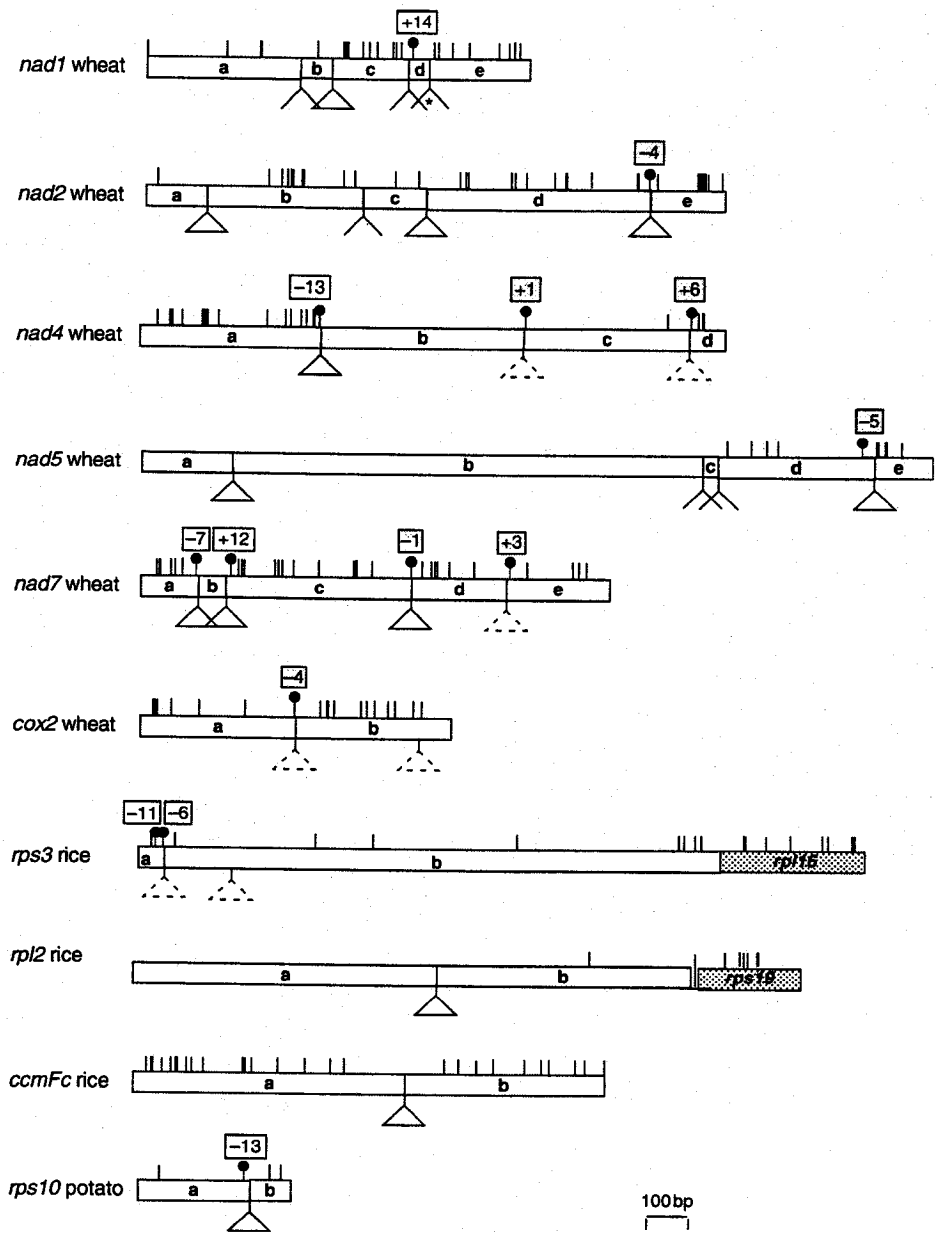


Fig. 1. Distribution of exon editing sites in intron-containing mitochondrial genes from wheat, as well as several from rice and potato. Coding regions are shown by open boxes, and adjacent genes by stippled boxes. *cis*-introns are shown as triangles, Y-shaped symbols indicate *trans*-introns, asterisk indicates *cis*-/*trans*-variation among plants and introns that are not widely present in seed plants are denoted by dotted triangles (Knoop 2004). Edit sites are shown by vertical lines, and those with black dots are located within 15 nucleotides from a splice junction, with precise positions either upstream (negative symbol) or downstream of an intron (plus symbol) shown in boxes above. The positions of the seven wheat edit sites examined in this study are highlighted in grey. Editing data were compiled for wheat *nad1* (Chapdelaine and Bonen 1991, Chapdelaine, PhD Thesis, University of Ottawa 1992), *nad2* (Morawala-Patell et al. 1998), *nad4* (Lamattina and Grienenberger 1991), *nad5* (Pereira de Souza et al. 1991), *nad7* (Bonen et al. 1994) and *cox2* (Covello and Gray 1990), as well as rice *rps3*, *rpl16*, *rpl2*, *rps19* and *ccmFc* (Notsu et al. 2002; BA000029) and potato *rps10* (Zanlungo et al. 1995). The three-nucleotide spacer between *rpl2* and *rps19* also contains an edit site (Kubo et al. 1996).

editing sites of interest being boxed. Four of these sites are located one to five nucleotides upstream of an intron, three are one to six nucleotides downstream of an intron, and all lead to amino acid changes (Fig. 2, codons underlined). The flanking contexts for both precursor and spliced molecules are shown. It can be seen that the neighbouring sequences before and after splicing are markedly different and the conserved nature of the 5′

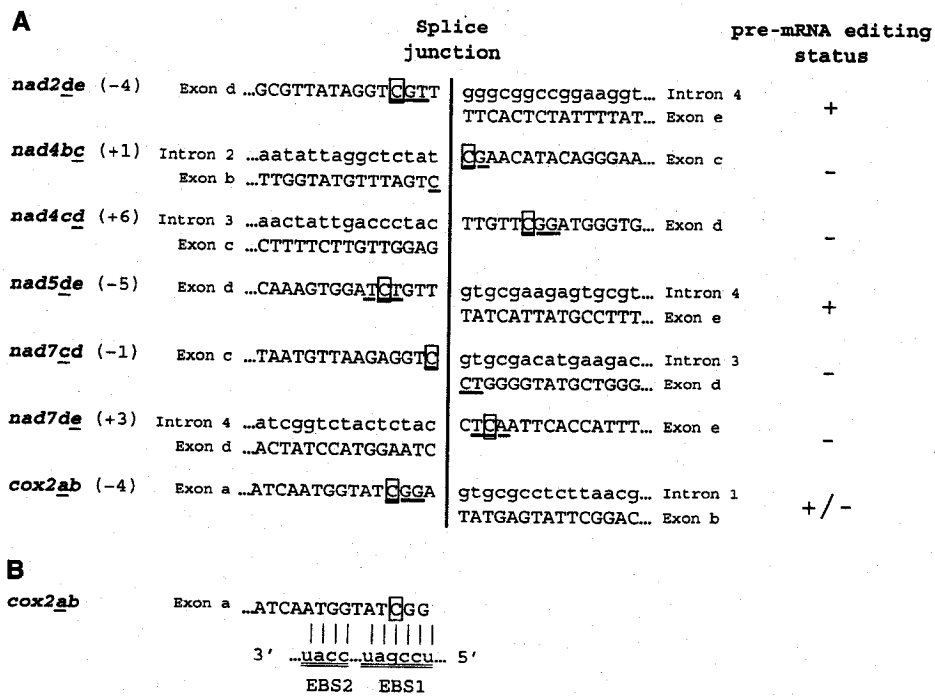


Fig. 2. Sequence context of the seven wheat mitochondrial exon editing sites located within six nucleotides of splice junctions and their editing status in precursor RNAs of 24-h germinating wheat embryos. (A) Exons (upper case) and introns (lower case) are shown with the candidate edit sites being boxed and their codon context denoted by underlining. The vertical line represents the splice junction position. The gene name (with underlining to denote exon containing the edit site of interest) is followed by the distance (in parentheses) of the edit site downstream (+ nucleotide) or upstream (− nucleotide) of the splice junction. Their editing status in wheat 24-h precursor RNA is summarized by plus and minus symbols on the right. (B) Exon–intron base-pairing [8] interactions (EBS–IBS) are shown for *cox2* (cf. Farré et al. 2002, with modification).

ends of group II introns is also evident. As expected, these editing sites show the typical bias against purine in the −1 position (cf. Giegé and Brennicke 1999).

RT-PCR sequencing analysis of exon editing sites near intron junctions

To examine the status of these seven candidate editing sites in the RNA population at different stages of RNA processing, we used various intron/exon primer combinations to generate RT-PCR products for direct sequencing. For multiple split genes, we focused on partially spliced transcripts to ensure that there was no residual contribution from contaminating DNA. We included analysis of mitochondrial RNA from 24-h germinating embryos and 6-day etiolated seedlings of wheat to assess developmental differences. In the case of the *nad2d* (−4) site, which is located four nucleotides upstream of intron 4, we observed virtually complete editing in both embryos and seedlings (Fig. 3A, black arrowheads), with only faint signals for the unedited state (cf. C lane in sequencing gels). Other sites within *nad2d*, which are further away from the intron, are also edited in precursor transcripts (Fig. 3B, three black arrowheads, star on schematic) although, interestingly two side-by-side *nad2d* sites, which are located 33–34 nucleotides upstream of the intron 4 junction, appear about 50% edited in these precursors (Fig. 3A, grey arrowheads), raising the possibility of a steric exclusion effect at such adjacent sites. All these exon sites (as well as others) exhibited complete editing in fully spliced *nad2* mRNAs (Fig. 3C). Similarly, a *nad5c* (−5) site is edited in both the precursor and spliced RNA populations (Supplementary data, Fig. S1). This also is the case for a wheat *nad7c* (+12) site (Carrillo and Bonen 1997) and was observed for a potato *rps10* (−13) site (Fig. 1) based on 7/8 cDNA clones (Zanlungo et al. 1995).

In contrast, when we examined exon editing sites, which are located immediately beside intron junctions, a different outcome was observed, that is these sites remained unedited when the adjacent intron was present. In Fig. 4A (white arrowheads), this can be seen for the *nad4c* (+1) site located one nucleotide downstream of intron 2. Similarly, a *nad7c* (−1) site did not show editing in unspliced molecules (Supplementary data, Fig. S2).

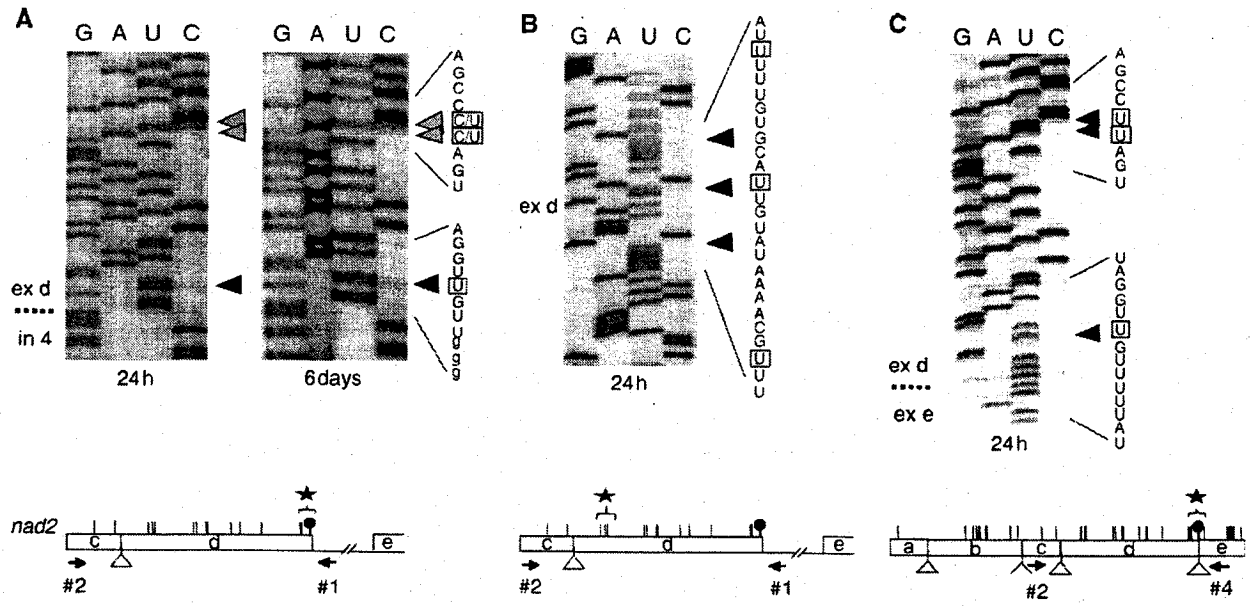


Fig. 3. Exon editing status near splice junction of *nad2d* (−4 position) for wheat mitochondrial precursor transcripts and messenger RNAs (mRNAs) from 24-h germinating embryos and 6-day etiolated seedlings. (A–C) Direct sequencing of reverse transcriptase–PCR products from partially spliced *nad2d*-intron 4 template showing (A) the −4 edit region or (B) three distal editing sites, and (C) from *nad2* mRNA template showing the −4 edit region. Sequences are shown at the right of sequencing gels, with exon (upper case), intron (lower case) and candidate edit sites (boxed). Black, grey or white arrowheads indicate observed C-to-U edits, incomplete editing of the population or unedited sites, respectively. The schematics below (as described in Fig. 1 legend) show positions of primers (with numbered arrows, see Table 1) and positions of edits under examination (stars).

Although most editing sites, which are not extremely close to intron junctions, showed editing in the presence of the intron, we observed an exception in *nad4d* where a site six nucleotides downstream of intron 3 appeared not to be edited at all (Fig. 4B, white arrowhead). This is in contrast to a more distal exon site (18 nucleotides away), which showed approximately equal signal in the C and U lanes in these precursors, as well as an intronic editing site (also with approximately 50% editing) located in *nad4* intron 3 within the domain V helix of the canonical secondary structure exhibited by group II introns. The latter edit converts an A–C mispair to A–U and thus would improve intron folding (Fig. 4B, upper grey arrowhead). Notably, the *nad4d* (+6) site shows virtually complete editing in partially spliced transcripts when intron 2 is still present but intron 3 has been excised (Fig. 4C, black arrowhead) as well as in fully spliced mRNA (data not shown). Thus, the presence of a distal intron does not appear to impair editing at other intron junction sites. Similar results were seen with the *nad7c* (−1) edit site at 0-, 12- and 24-h stages (Supplementary data, Fig. S2). With respect to *nad7* exon editing sites, we had also previously observed a complete lack of editing for an *nad7e* (+3) site in RNA precursor populations, which still contain intron 4 (Carrillo and Bonen 1997). None of 22 clones from 24-h or 6-day-old seedlings showed editing at that site in contrast to other *nad7d* and *nad7e* sites (which showed at least 50% editing in the population).

Developmental differences in editing status of intron-containing wheat mitochondrial *cox2* transcripts

We have examined the editing status of spliced and unspliced *cox2* transcripts during the developmental period when seeds leave dormancy, germinate and develop into etiolated seedlings (0 h to 6 days post-imbibition). These stages were selected because we had previously observed that intron-containing *cox2* precursors appear relatively more abundant during the 12- to 24-h period than they do in either earlier (cf. 6 h) or later (cf. 6 days) stages where only very minor amounts of unspliced forms or excised intron were detected, even though spliced mRNA levels are high (Li-Pook-Than et al. 2004). Because *cox2* contains a single intron in wheat, in our editing analysis RNA samples were repeatedly DNase treated prior to RT-PCR (Fig. 5A inset) although it cannot be excluded that there is still minor contribution from residual contaminating mitochondrial DNA.

For the unspliced *cox2* transcripts in dormant seeds, the splice junction (−4) exon site shows only about 50% editing (Fig. 5A, upper panel grey arrowheads) and

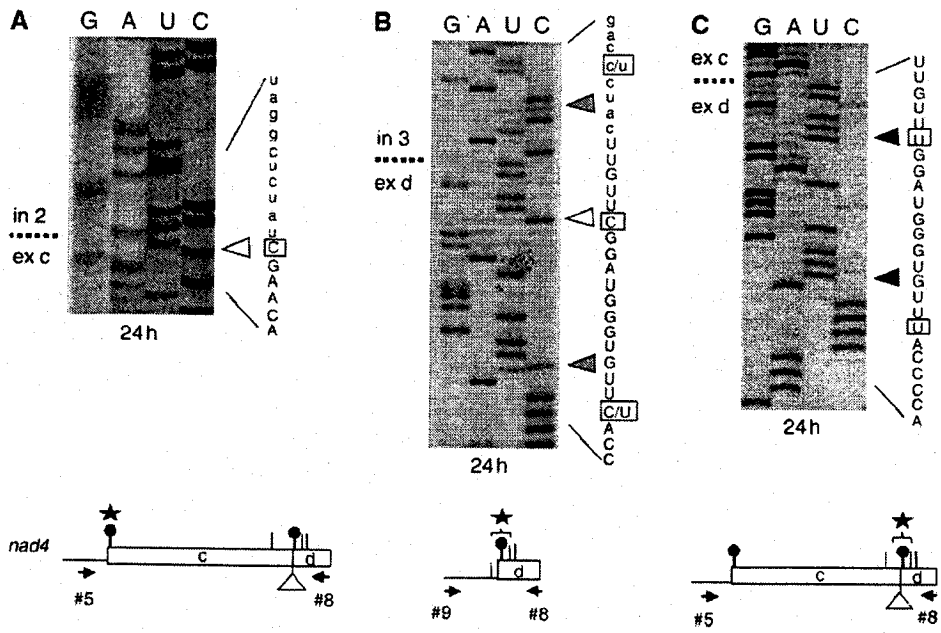


Fig. 4. Editing status near splice junction of *nad4c* (+1 position) and *nad4d* (+6 position) for wheat mitochondrial transcripts from 24-h germinating embryos. Direct sequencing of reverse transcriptase–PCR products for (A) *nad4c* (+1) junction region in partially spliced intron 2/*nad4cd* transcripts, (B) *nad4d* (+6) junction region in *nad4* intron 3/*nad4d* transcripts and (C) *nad4d* (+6) junction region in partially spliced intron 2/*nad4cd* transcripts in which the distal intron is still present. Designations are as described in Figs 1 and 3.

interestingly, a more distal *cox2* exon 1 site, which is 130 nucleotides further away from the intron, exhibits about only 10% editing (Fig. 5A, lower panel, grey arrowheads). In the case of 24-h embryos, the *cox2* (−4) editing status resembles that of the 0-h RNA population, whereas the distal exon site shows greater than 50% editing (Fig. 5A) and in 6-day-old seedlings we saw more than 90% editing (or virtually complete if the minor signal in the C lane is because of contaminating DNA) (Fig. 5A, black arrowheads). This variation in degree of editing at various sites in unspliced molecules and lack of correlation with proximity to the splice junction suggests an independence in the nature of editing determinants. In contrast, the spliced wheat *cox2* mRNA, as expected, showed virtually complete editing at all sites examined (including the −4 position) in all these developmental stages as well as in dormant seeds (Fig. 5B, black arrowheads).

Discussion

We have shown that wheat mitochondrial editing sites, which are located either immediately upstream or immediately downstream of splice junctions, remain unedited in unspliced transcripts, whereas ones further away show complete (or partial) editing in the intron-containing precursor RNA population. This contrasts with their fully edited status in spliced mRNAs as well as in splicing intermediates where only distal introns are still present. Our findings are consistent with models of (1) steric hindrance, which prevents the editing machinery from interacting with preexisting recognition sites or (2) the creation of such editing *cis*-elements by splicing, i.e. the required primary or secondary structure for editing is provided by exon/exon (but not exon/intron or intron/exon) sequences (Fig. 2A). It is also possible that editing at such sites might be a rate-limiting step, which triggers splicing, so that there would be low levels of such intermediate molecules detected in vivo.

With respect to the 'physical blocking' model, although the determinants that provide editing specificity are not yet well understood in plant mitochondria, it appears that flanking nucleotides (for stretches of approximately10–20 nucleotides) provide key information, with elements upstream of the candidate site being more important than downstream ones (reviewed in Shikanai 2006). The variation in editing status that we observed in precursors populations for sites slightly away from intron borders, namely, virtually complete editing for *nad2* (−4) and *nad5* (−5), about 50% editing for *cox2* (−4), and no editing *nad4* (+6) (Fig. 2A) is in keeping with the view that there is variation among sites as well as between orthologous sites in different plants based on in vitro editing studies (cf. Neuwirt et al. 2005, Takenaka et al.

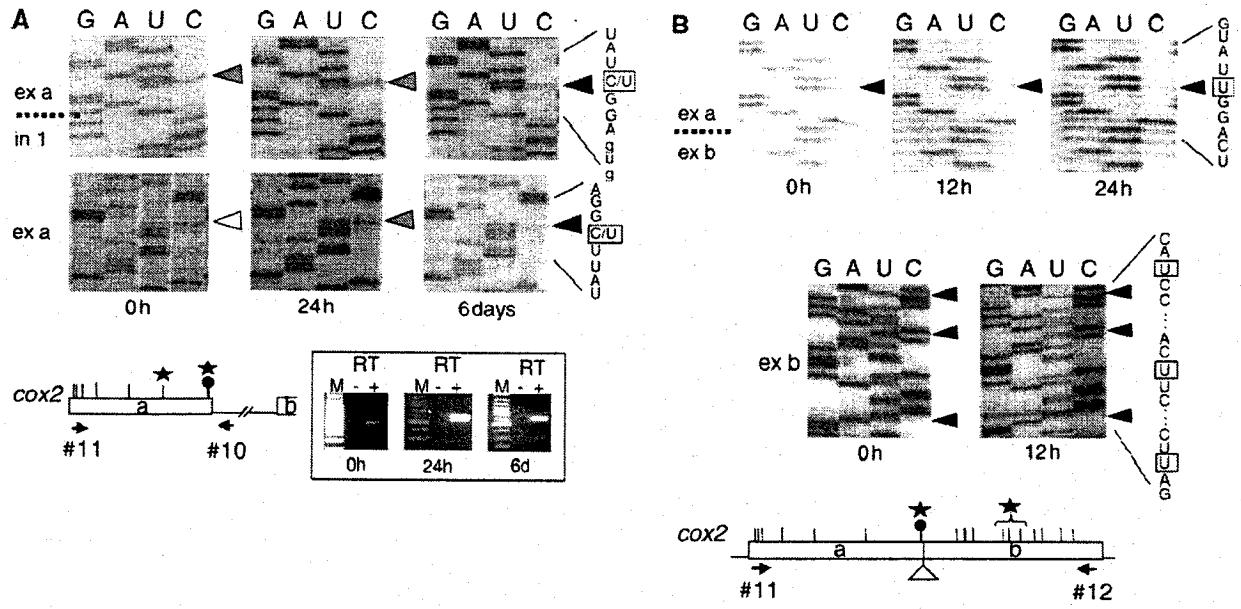


Fig. 5. Editing status of wheat mitochondrial *cox2* spliced and unspliced transcripts during embryo-to-seedling development (0 h to 6 days postimbibition). (A) Direct sequencing of reverse transcriptase–PCR (RT-PCR) products for *cox2* precursors at the exon (−4) junction site (above) and at editing sites distal from the junction (below). Inset shows gels of RT-PCR products (of 600 bp) representing *cox2* precursor transcripts when RNA (which had been pretreated with DNase I) was incubated either with (+RT) or without reverse transcriptase (−RT). (B) Direct sequencing of RT-PCR products for *cox2* mRNA at −4 position (upper panel) and at distal editing sites (lower panel) at 0, 12 and 24 h postimbibition. Designations are as described in Figs 1 and 3.

2004) and in organelle studies (cf. Farré et al. 2001, Staudinger and Kempken 2003) and the absence of any generalized consensus motifs.

The nature of such steric hindrance might well be because of interactions between intron and exon sequences. It is well known for group II introns that exon sequences immediately preceding the intron undergo base pairing with complementary intronic sequences (EBS1 and EBS2) located within domain I of the core structure (reviewed in Bonen and Vogel 2001, Lehmann and Schmidt 2003). Indeed, there is evidence from mutational studies of *cox2* splicing in isolated wheat mitochondria in support of such an EBS–IBS interaction (Farré and Araya 2002), although it is worth noting that some plant mitochondrial introns appear degenerate and lack certain conventional group II features (Bonen and Vogel 2001). An EBS–IBS interaction might be expected to block entry of editing machinery, and indeed if editing were to convert G–C pairs to G–U one might expect a bias against editing preceding splicing. Alternatively, if A–C mispairs were edited to A–U, the intron–exon interaction would be strengthened, and hence editing might improve the efficiency of splicing, as proposed in the case of *Isoetes lacustris nad1* (Malek and Knoop 1998). For the *cox2* (−4 position) edit, there are examples of both directions. In *Acorus calamus*, from an early diverging angiosperm lineage, IBS1–EBS1 pairing is improved (Albertazzi et al. 1998), whereas in wheat editing would weaken the base pairing (Fig. 2B). In this regard, the *cox2* (−4) position was reported to show partial editing in unspliced transcripts from 7-day etiolated maize seedlings (Yang and Mulligan 1991) and petunia suspension cells (Sutton et al. 1991), based on 4/7 and 3/7 cDNA clones, respectively (a level similar to that reported for more distal editing sites). Incidentally, a relatively lower level of editing at the *cox2* (−4) site was observed in cold-treated wheat seedlings (Kurihara-Yonemoto and Handa 2001). In addition to interactions with upstream exon sequences, group II introns also have an important long-range interaction (designated δ–δ′) between the first nucleotide of the downstream exon and an intronic site (Bonen and Vogel 2001). Interestingly, a chloroplast *ndhA* (+13) exon was found to be unedited in intron-containing transcripts in barley greening leaves, and this observation prompted the suggestion that editing status at this site might influence intron folding within domains V and VI (which are crucial for the ribozymic splicing reaction) (del Campo et al. 2000).

Our observation that editing is a relatively 'late' RNA processing event at exon sites very close to splice junctions raises the possibility that such sites might fulfil a regulatory role in gene expression. If so, it might be

expected that this requirement for editing would be conserved among different plant lineages. When we surveyed the published editing data for the seven sites shown in Fig. 2 (boxed), we found variation among plants. For example, the *nad4c* (+1) site remains as C in *nad4* mRNA in *Arabidopsis thaliana* (Giegé and Brennicke 1999) and *Brassica napus* (Handa 2003), whereas the *nad7c* (−1) and *nad7e* (+3) sites are edited to U in *Brassica* but not *Arabidopsis* mRNAs, and both the *nad2d* (−4) and *nad4d* (+6) sites are edited to U in these two plants. In the case of the *nad5d* (−5) site, there is a genomically encoded T in both *Arabidopsis* and *Brassica* so that editing is not required, and this is also the case for the *cox2* (−4) site in *Brassica*. These latter observations therefore preclude a regulatory role conserved among all plants.

The finding that wheat *cox2* unspliced precursors show different degrees of editing between the embryo and seedling stages not only at splice junctions but also at other exon sites, raises questions about the efficiency of editing in the early stages of development and the availability of nuclear-encoded RNA processing machinery. It will be of interest to further elucidate the interrelationships between RNA editing and splicing events especially during periods of high respiratory demand, as in the early stages of germination.

Supplementary material

The following material is available to download from www.blackwell-synergy.com/loi/ppl

Fig. S1. Exon editing status near splice junction of *nad5d* (−5 position) for wheat mitochondrial precursor transcripts and messenger RNAs from 24-h germinating embryos and 6-day etiolated seedlings.

Fig. S2. Editing status near splice junction of *nad7c* (−1 position) for wheat mitochondrial transcripts from dormant seeds (0 h), 12- (postimbibition) and 24-h germinating embryos.

Table S1.

Acknowledgements – We thank Dr R. Pandeya (Agriculture and Agri-Food Canada, Ottawa) for kindly providing the wheat seeds used in this study and the Natural Sciences and Engineering Council of Canada for continued financial support.

References

Adams KL, Clements MJ, Vaughn JC (1998) The *Peperomia* mitochondrial *coxI* group I intron: timing of horizontal transfer and subsequent evolution of the intron. J Mol Evol 46: 689–696

Adams KL, Daley DO, Qiu YL, Whelan J, Palmer JD (2000) Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature 408: 354–357

Albertazzi FJ, Kudla J, Bock R (1998) The *cox2* locus of the primitive angiosperm plant *Acorus calamus*: molecular structure, transcript processing and RNA editing. Mol Gen Genet 259: 591–600

Bonen L, Brown GG (1993) Genetic plasticity and its consequences: perspectives on gene organization and expression in plant mitochondria. Can J Bot 71: 645–660

Bonen L, Vogel J (2001) The ins and outs of group II introns. Trends Genet 17: 322–331

Bonen L, Williams K, Bird S, Wood C (1994) The NADH dehydrogenase subunit 7 gene is interrupted by four group II introns in the wheat mitochondrial genome. Mol Gen Genet 244: 81–89

Carrillo C, Bonen L (1997) RNA editing status of *nad7* intron domains in wheat mitochondria. Nucleic Acids Res 25: 403–409

Carrillo C, Chapdelaine Y, Bonen L (2001) Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. Mol Gen Genet 264: 595–603

Chapdelaine Y, Bonen L (1991) The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a *trans*-splicing model for this gene-in-pieces. Cell 65: 465–472

Chateigner-Boutin A, Hanson MR (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar *cis*-element. Nucleic Acids Res 31: 2586–2594

Choury D, Farré JC, Jordana X, Araya A (2005) Gene expression studies in isolated mitochondria: Solanum tuberosum *rps10* is recognized by cognate potato but not by the transcription, splicing and editing machinery of wheat mitochondria. Nucleic Acids Res 33: 7058–7065

Covello PS, Gray MW (1990) Differences in editing at homologous sites in messenger RNAs from angiosperm mitochondria. Nucleic Acids Res 18: 5189–5196

del Campo EM, Sabater B, Martin M (2000) Transcripts of the *ndh-D* operon of barley plastids: possible role of unedited site III in splicing of the *ndhA* intron. Nucleic Acids Res 28: 1092–1098

Farré JC, Araya A (2002) RNA splicing in higher plant mitochondria: determination of functional elements in group II intron from a chimeric *cox II* gene in electroporated wheat mitochondria. Plant J 29: 203–213

Farré JC, Leon G, Jordana X, Araya A (2001) *cis* recognition elements in plant mitochondrion RNA editing. Mol Cell Biol 21: 6731–6737

Gass DA, Makaroff CA, Palmer JD (1992) Variable intron content of the NADH dehydrogenase subunit 4 gene of plant mitochondria. Curr Genet 21: 423–430

Geiss KT, Abbas GM, Makaroff CA (1994) Intron loss from the NADH dehydrogenase subunit 4 gene of lettuce mitochondrial DNA: evidence for homologous recombination of a cDNA intermediate. Mol Gen Genet 243: 97–105

Giegé P, Brennicke A (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci USA 96: 15324–15329

Grosskopf D, Mulligan RM (1996) Developmental- and tissue-specificity of RNA editing in mitochondria of suspension-cultured maize cells and seedlings. Curr Genet 29: 556–563

Gualberto JM, Bonnard G, Lamattina L, Grienenberger JM (1991) Expression of the wheat mitochondrial *nad3-rps12* transcription unit: correlation between editing and mRNA maturation. Plant Cell 3: 1109–1120

Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica Napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucleic Acids Res 31: 5907–5916

Knoop V (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46: 123–139

Kubo N, Ozawa K, Hino T, Kadowaki K (1996) A ribosomal protein L2 gene is transcribed, spliced, and edited at one site in rice mitochondria. Plant Mol Biol 31: 853–862

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys) (GCA). Nucleic Acids Res 28: 2571–2576

Kurihara-Yonemoto S, Handa H (2001) Low temperature affects the processing pattern and RNA editing status of the mitochondrial *cox2* transcripts in wheat. Curr Genet 40: 203–208

Lamattina L, Grienenberger JM (1991) RNA editing of the transcript coding for subunit 4 of NADH dehydrogenase in wheat mitochondria: uneven distribution of the editing sites among the four exons. Nucleic Acids Res 19: 3275–3282

Lehmann K, Schmidt U (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. Crit Rev Biochem Mol Biol 38: 249–303

Li-Pook-Than J, Carrillo C, Bonen L (2004) Variation in mitochondrial transcript profiles of protein-coding genes during early germination and seedling development in wheat. Curr Genet 46: 374–380

Logan DC, Millar AH, Sweetlove LJ, Hill SA, Leaver CJ (2001) Mitochondrial biogenesis during germination in maize embryos. Plant Physiol 25: 662–672

Maier RM, Zeltz P, Kössel H, Bonnard G, Gualberto JM, Grienenberger JM (1996) RNA editing in plant mitochondria and chloroplasts. Plant Mol Biol 32: 343–365

Malek O, Knoop V (1998) Trans-splicing group II introns in plant mitochondria: the complete set of *cis*-arranged homologs in ferns, fern allies, and a hornwort. RNA 4: 1599–1609

Morawala-Patell V, Gualberto JM, Lamattina L, Grienenberger JM, Bonnard G (1998) Cis- and trans-splicing and RNA editing are required for the expression of *nad2* in wheat mitochondria. Mol Gen Genet 258: 503–511

Neuwirt J, Takenaka M, Van der Merwe JA, Brennicke A (2005) An in vitro RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. RNA 11: 1563–1570

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, 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: 434–445

Peeters NM, Hanson MR (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA 8: 497–511

Pereira de Souza A, Jubier MF, Delcher E, Lancelin D, Lejeune B (1991) A *trans*-splicing model for the expression of the tripartite *nad5* gene in wheat and maize mitochondria. Plant Cell 3: 1363–1378

Regina TMR, Picardi E, Lopez L, Pesole G, Quagliariello C (2005) A novel additional group II intron distinguishes the mitochondrial *rps3* gene in gymnosperms. J Mol Evol 60: 196–206

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Shadel GS (2004) Coupling the mitochondrial transcription machinery to human disease. Trends Genet 20: 513–519

Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 63: 698–708

Staudinger M, Kempken F (2003) Electroporation of isolated higher-plant mitochondria: transcripts of an introduced *cox2* gene, but not an *atp6* gene, are edited in organelle. Mol Genet Genomics 269: 553–561

Subramanian S, Fallahi M, Bonen L (2001) Truncated and dispersed *rpl2* and *rps19* pseudogenes are co-transcribed with neighbouring downstream genes in wheat mitochondria. Curr Genet 39: 264–272

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: 603–615

Sutton CA, Conklin PL, Pruitt KD, Hanson MR (1991) Editing of pre-mRNAs can occur before *cis*- and *trans*-splicing in Petunia mitochondria. Mol Cell Biol 11: 4274–4277

Takenaka M, Neuwirt J, Brennicke A (2004) Complex *cis*-elements determine an RNA editing site in pea mitochondria. Nucleic Acids Res 32: 4137–4144

Yang AJ, Mulligan RM (1991) RNA editing intermediates of *cox2* transcripts in maize mitochondria. Mol Cell Biol 11: 4278–4281

Zanlungo S, Quiñones V, Moenne A, Holuigue L, Jordana X (1995) Splicing and editing of *rps10* transcripts in potato mitochondria. Curr Genet 27: 565–571

Edited by P. Gardeström