2782–2790 *Nucleic Acids Research, 2006, Vol. 34, No. 9*
doi:10.1093/nar/gkl328

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

Jennifer Li-Pook-Than and Linda Bonen*

Biology Department, University of Ottawa, Ottawa, Canada K1N 6N5

Received March 11, 2006; Revised April 3, 2006; Accepted April 13, 2006

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*.

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 is 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

*To whom correspondence should be addressed. Tel: +1 613 562 5800 ext. 6356; Fax: +1 613 562 5486; Email: lbonen@science.uottawa.ca

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.

MATERIALS AND METHODS

RT–PCR and cRT–PCR analysis of intron RNAs

Mitochondrial RNA was isolated from 24 h 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 (~5 μg) was denatured at 70°C for 10 min and cDNA was synthesized using 10 U/μl reverse transcriptase, 0.2 vol 5× First Strand buffer (Invitrogen), 10 mM DTT and 0.5 mM dNTPs incubated at 45°C for 2 h. 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 (~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 10× 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.

RNase H analysis of excised introns

Wheat mitochondrial RNA (~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).

RESULTS

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 1A, boxed) and the two unpaired terminal nucleotides (5′-AY-3′) are involved in long-range interactions including γ−γ′ 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 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 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 1C), adjacent A:A and G:G mispairs in *nad5* intron 2 (Figure 1D), A:A mispairs in *nad4* intron 2 (Figure 1F) and to a much more pronounced extent in *nad1* intron 1 [Figure 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 1 fall in that category (Figure 1D and 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 1D–E, G–H). The *cox2* intron contains an A:G 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 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 1H) and considerable variation in sequence is seen among plants. In the case of *nad2* intron 1 and *nad4* intron 2 (Figure 1E–F), although a bulged nucleotide is present at the expected location (albeit flanked by pairs of varying strength among plants)

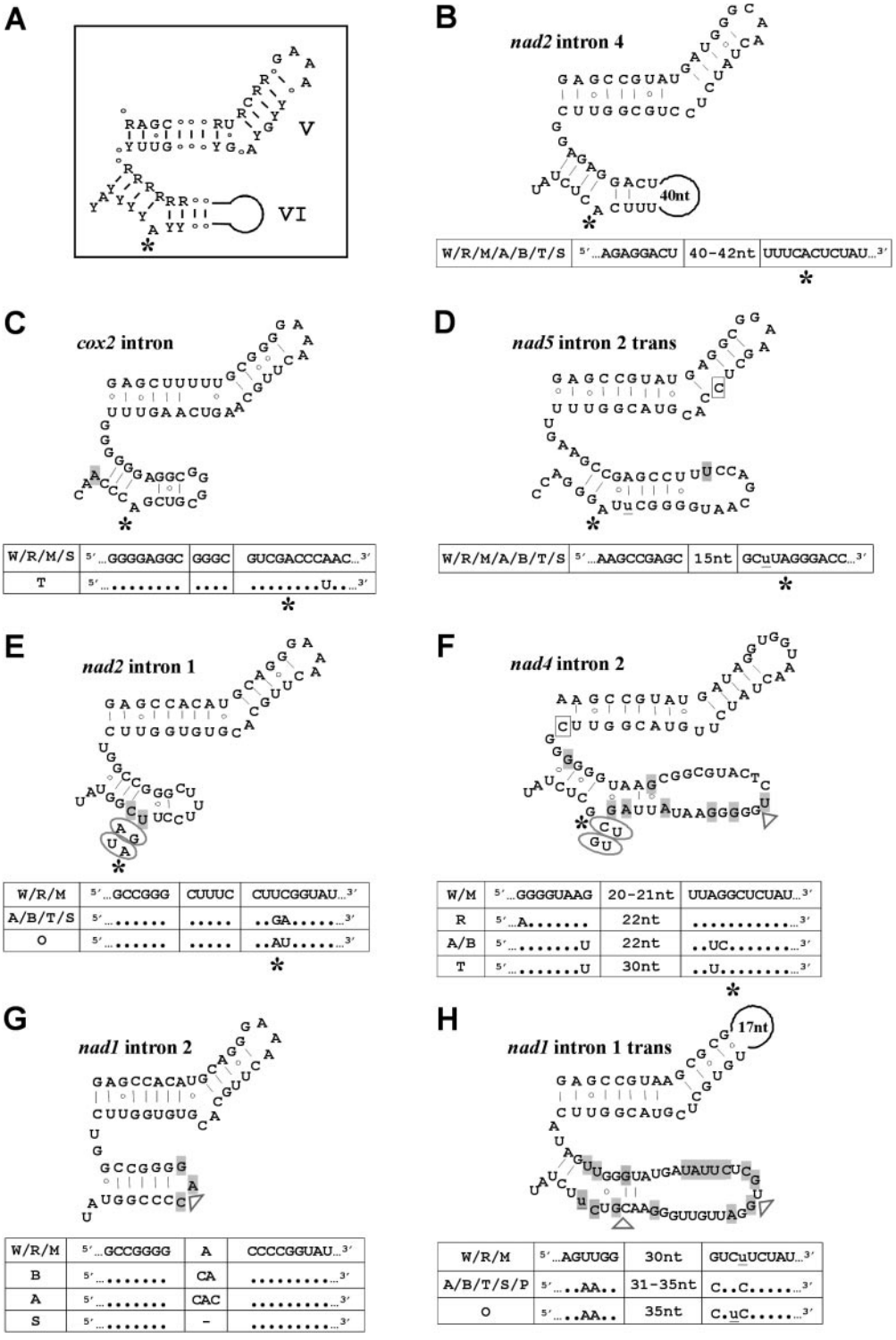

Figure 1. Secondary structure models of domains V and VI of wheat mitochondrial group II introns. (**A**) Classical group IIA intron (boxed), (**B**) *nad2* intron 4 (Y14434), (**C**) *cox2* intron (X01108), (**D**) *trans*-splicing *nad5* intron 2 (M74158), (**E**) *nad2* intron 1 (Y14433), (**F**) *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. Refs (2,15)]. 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 [Refs (22,27,38,39) and J. Li-Pook-Than and L. Bonen, unpublished data). Boxed upper case Cs indicate A–C mispairs which are not edited (J. Li-Pook-Than and L. Bonen, 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.Refs (27,40) and J. Li-Pook-Than and L. Bonen unpublished data] and differ from Y14434 and X57164, respectively.

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 1E, F and 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 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 1D and F, boxed Cs 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 (Figure1F and 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 (Figure 1B), *nad1* intron 2, which has a short tight dVI helix (Figure 1G), and the *cox2* intron, which has an A:G mispair adjacent to the branchpoint and mispairing in dV (Figure 1C).

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 whether linear intron molecules might also be present in the population (Figure 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 2A, lane 1 versus 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 2A, lanes 4–9, arrowheads). Several larger products were also observed (Figure 2A, lanes 4, 6, 7); however only *nad2* intron 4 (Figure 2A, lane 4) exhibited discrete hybridization signals with intron-specific probes in Southern blot analysis (Figure 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 2A, lane 4 versus 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 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 2B, lanes 8, 9), and this is consistent with the absence of Southern hybridization signal because of the intron oligomer probe maps outside that region. Comparable results with those shown in Figure 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 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 six out of six RT–PCR clones and four out of four 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 (Table 1). This was unexpected as this intron contains a bulged A in dVI (Figure 1B). The other four 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 ~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 six out of eight clones had short non-encoded polyA tracts (1–5 nt 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 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 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 2E, boxed nts). Of the remaining four clones,

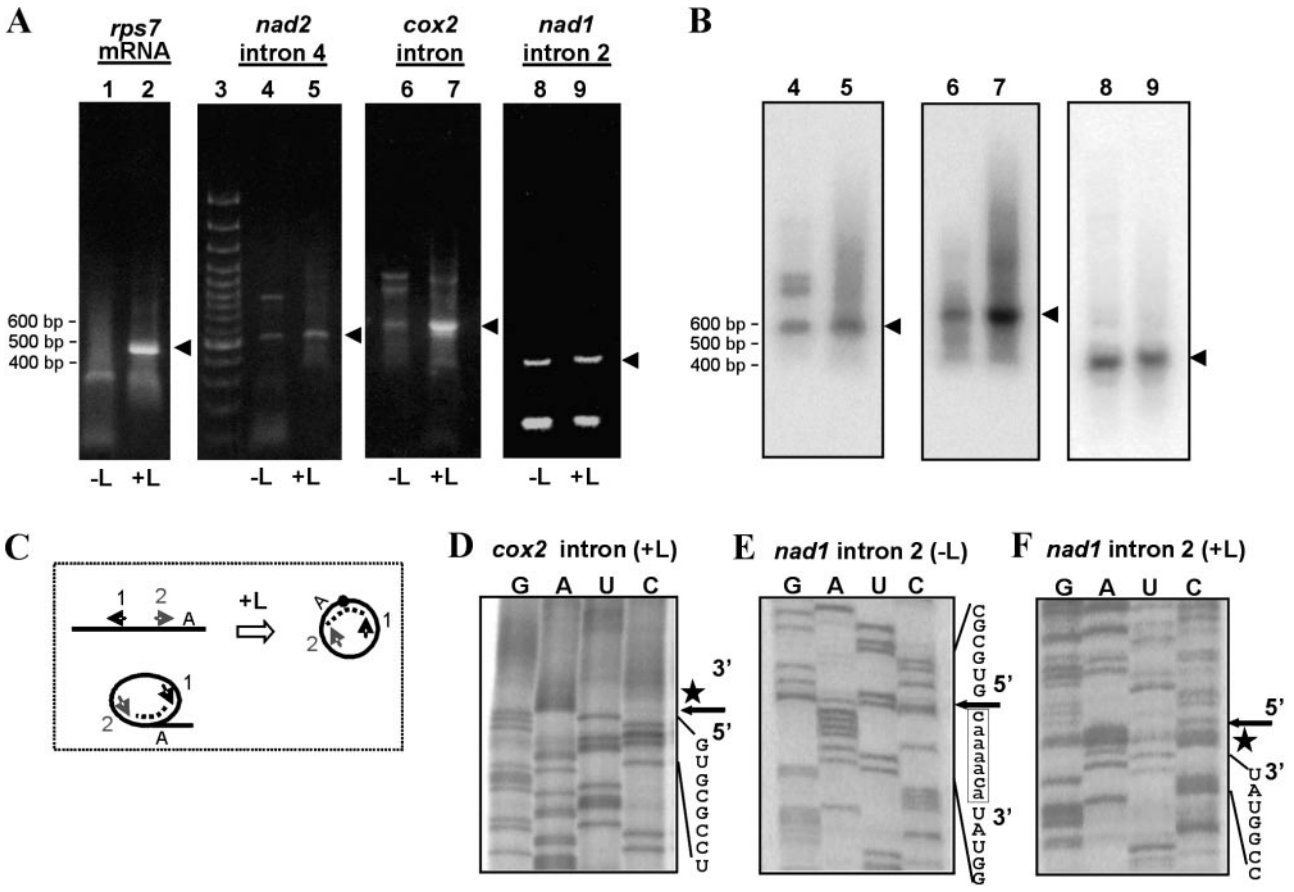

Figure 2. RT–PCR and cRT–PCR detection of lariat (and/or circularized) versus linear wheat mitochondrial intron molecules. (**A**) RT–PCR across the excised intron junction using 24 h wheat mitochondrial RNA as template, either pre-treated with ligase (+L; lanes 2, 5, 7 and 9) or untreated (−L; lanes 1, 4, 6 and 8) generated products of sizes 450, 520, 630 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 nucleotide of the introns, stars highlight short non-encoded A-rich stretches and the discrete non-encoded insert sequence is boxed in (E).

one represented a full-length circularized intron, while two clones also contained 5′-ACTGT-3′ which matches the first 5 nt of *nad1* exon 2, and one clone lacked the 3′ terminal 32 nt of the intron (Table 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 2F) appears to be somewhat more heterogeneous than what was seen for the *in vivo* population (Figure 2E).

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 with the 3′–5′ bond generated by ligation of the ends of a previously linear molecule (cf. Figure 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 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 3B, black arrowhead denoting 900 nt product detected by probe), when compared with the native 1.4 kb intron (Figure 3B, black dot). A minor signal was detected at the position corresponding to a linear form (∼400 nt, Figure 3B, open arrowhead). Also, major products of ∼4.8 kb derived from the linear precursor of ∼6 kb can be seen (Figure 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

Table 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

Intron	Clone	Ligase	3′ termini	non-encoded nucleotides	5′ termini
nad2 intron 4			3′ termini of *nad2* intron 4		5′ termini of *nad2* intron 4
			...ACCCTTTCACTCTAT	.....	GGGCGGCCGGAA...
	6/6	−L	...ACCCTTTCT.....	.....	GGGCGGCCGGAA...
	4/4	+L	...ACCCTTTCT.....	.....	GGGCGGCCGGAA...
cox2 intron			3′ termini of *cox2* intron		5′ termini of *cox2* intron
			...GTCAAGTTTGGGGGGAGGCGGGCGTCGACCCAAC	.....	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	.....	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...

Note that 5′ terminal nucleotides are shown on the right, 3′ terminal nucleotides 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 h wheat embryos.

intron (Figure 3C, black dot) to products consistent with circular (or lariat) molecules of ∼600 nt (Figure 3C, black arrowhead) and linear ones of ∼300 nt (Figure 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 3C, open versus 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 2 and Table 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, and this results in RNA species of ∼6, 4.6 and 3.6 kb (Figure 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 3D, black dot versus open arrowhead). This outcome was also seen when oligomer h was used, the shift being to 960 nt (Figure 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 3D, black arrowhead and black dot). The high molecular weight precursors shift to the expected sizes of 2.6 and 2.3 kb. It is worth noting that the RNA ligase (cRT–PCR) experiments (cf. Figure 2A, lane 9) might be influenced by protective structures at the intron termini, such as the 3′ terminal hairpin of *nad1* intron 2.

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 and L. Bonen 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). 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 (i) polyadenylation at the 3′ terminus [perhaps mediated by the bacterial-type tagging of transcripts for degradation observed in plant organelles, Ref. (28)] or (ii) endogenous circularization after 3′ exonucleolytic attack (Figure 4B). Linear molecules have also previously been

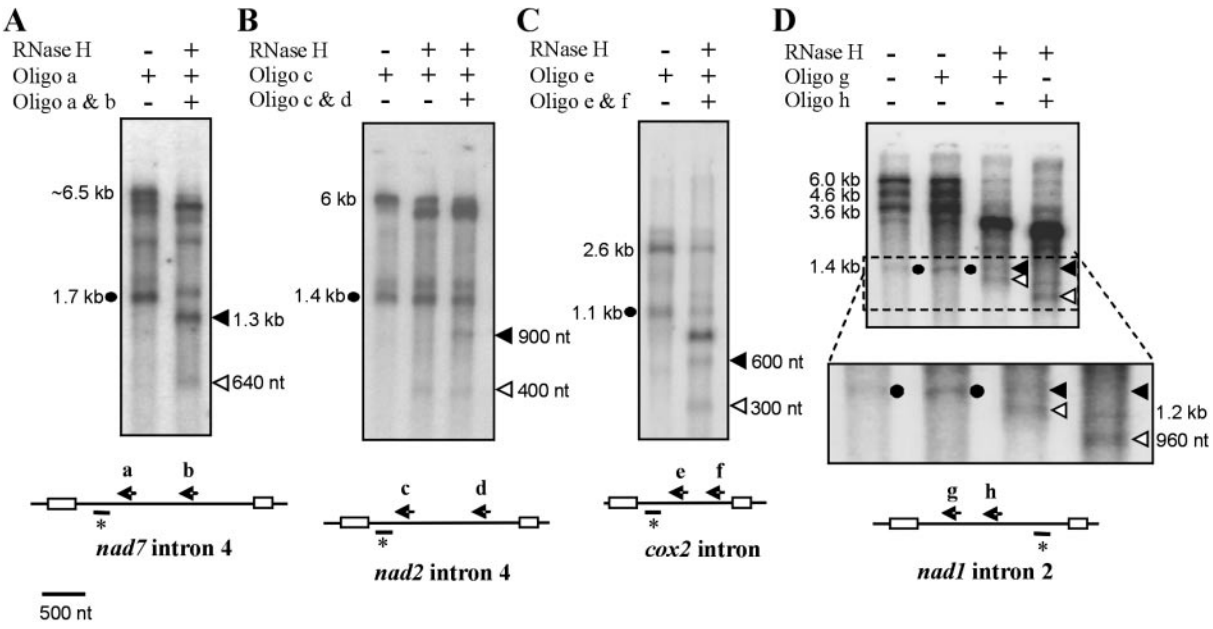

Figure 3. Determination of *in vivo* abundance of lariat (and/or circularized) versus linear excised introns by RNase H analysis. Wheat mitochondrial RNA (24 h) was annealed with oligomers as shown in schematics (arrows a–h) for (**A**) *nad7* intron 4, (**B**) *nad2* intron 4, (**C**) *cox2* intron and (**D**) *nad1* intron 2, prior to RNase H treatment. Positions of oligomer probes used in subsequent northern blot analysis are shown by overlined 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.

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 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 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 with the *cox2* intron (Figure 4B and C). Notably, the excised introns fall into three categories: (i) precisely full-length molecules, (ii) those which in addition have a discrete non-encoded ACAAAAC block and (iii) ones which have the full-length intron plus ACTGT (which corresponds to the first 5 nt 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 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 4B and 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 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.

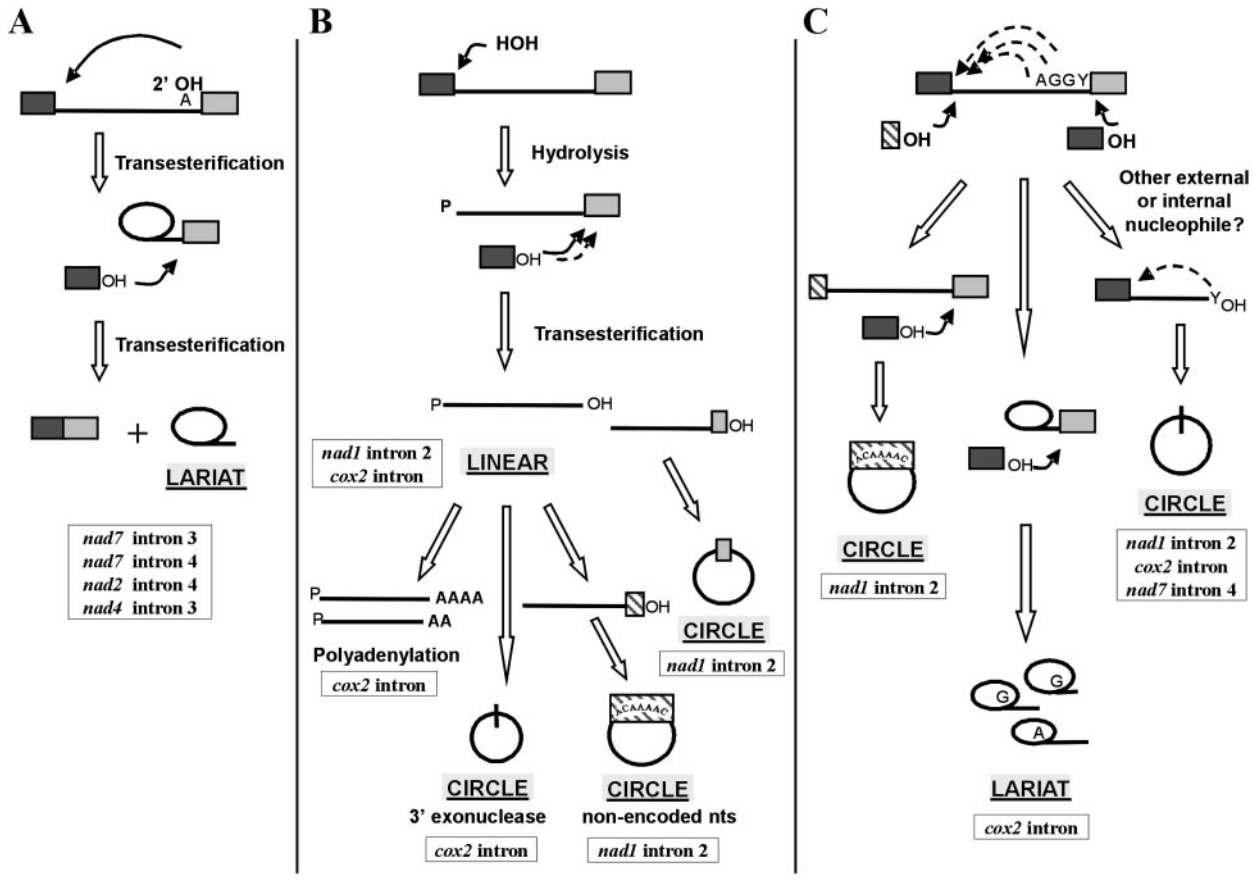

Figure 4. Models for plant mitochondrial group II intron splicing. (**A**) Conventional two-step transesterification yielding intron lariat. Upstream and downstream exons are denoted by dark and light shaded boxes, respectively. (**B**) Hydrolytic pathway generating linear excised intron which can further undergo (i) 3′ polyadenylation, (ii) 3′ exonuclease activity prior to endogenous circularization of heterogeneous molecules, [cf. models in Refs (8,35)], (iii) addition of discrete non-encoded block (hatched box) prior to endogenous circularization or (iv) 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 (i) external nucleophilic attack by a discrete non-encoded block (ACAAAAC, hatched box) followed by circularization of the full-length intron, (ii) an internal intronic nucleophile for first transesterification step (dashed arrows) or (iii) 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, Refs (8,35)].

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 (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.

SUPPLEMENTARY DATA

Supplementary data are available at *NAR* Online.

ACKNOWLEDGEMENTS

The authors thank C. Carrillo, N. Niknejad, S. Subramanian and J. Crosthwait for their earlier work on wheat and rice mitochondrial introns in our laboratory. The authors also thank

Dr R. Pandeya (Agriculture and Agri-Food Canada, Ottawa) for kindly providing the wheat seeds used in this study. Financial support in the form of a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (L.B.) and an Ontario Graduate Scholarship (J.L.P.T.) is gratefully acknowledged. Funding to pay the Open Access publication charges for this article was provided by the Natural Sciences and Engineering Research Council of Canada.

Conflict of interest statement. None declared

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. and 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. 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.
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 edn*. 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. *Nature 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 fersn, 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. and 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.