ΦYTON
REVISTA INTERNACIONAL DE BOTÁNICA EXPERIMENTAL
INTERNATIONAL JOURNAL OF EXPERIMENTAL BOTANY
FUNDACION ROMULO RAGGIO
Gaspar Campos 861, 1638 Vicente López (BA), Argentina
www.revistaphyton.fund-romuloraggio.org.ar

Evolutionarily conserved untranslated regions facilitate the cloning of complete coding sequences of chondriogenes encoding NADH dehydrogenase subunits in higher plants

Regiones conservadas no traducidas facilitan el clonado de los genes mitocondriales que codifican las subunidades de NADH dehidrogenasa en plantas

Jin G[1,2], XM Tang[3], Y Niu[4], XY Huang[2], T Chen[2], Q Huang[2], J Zhang[2], RY Zhou[1]

Abstract. In plants, the mitochondrial NADH dehydrogenase (complex I) is a large protein complex transferring electrons to ubiquinone. For the nine chondriogenes encoding complex I subunits (*nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7,* and *nad9*), an efficient strategy for the cloning of complete coding sequences (CDSs) is important. Specific orthologous portions of untranslated regions (UTRs) were found based on multiple sequence alignments of chondriogene orthologues encoding complex I subunits in plant species. Based on the conservation of partial UTRs, a one-step PCR strategy was conceived for the cloning of CDSs of the nine chondriogene orthologues. Using this strategy, the five complete mitochondrial open reading frames (ORFs) , which encode mitochondrial NADH dehydrogenase subunits, *nad1, nad2, nad6, nad7* and *nad9* respectively, were cloned in three angiosperm species: kenaf (*Hibiscus cannabinus*), camphor tree (*Cinnamomum camphora*), and ramie (*Boehmeria nivea*). The fifteen cloned PCR products also included 5' and 3'-UTR partial sequences. Moreover, a potential C-U RNA editing site was identified in the start codon of kenaf *nad9*. In conclusion, the simple and efficient strategy avoids the use of time-consuming rapid amplification of cDNA ends (RACE) process, and facilitates the cloning mitochondrial complete ORFs whose 5' and 3' flanking UTR contain an orthologous region with some degeneracy in higher plant species.

Keywords: NADH dehydrogenase subunit; RNA edit, Cloning; Degenerate primer; Untranslated region.

Abbreviations: CDS: coding sequence; UTR: untranslated region; CMS: cytoplasmic male sterility; mtDNA: mitochondrial DNA; ORF: open reading frame; RACE: rapid amplification of cDNA ends.

Resumen. En plantas, la deshidrogenasa NADH de la mitocondria (complejo I) es un gran complejo de proteínas que transfiere electrones a la ubiquinona. Para los nueve genes mitocondriales que codifican las subunidades del complejo I (*nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7,* y *nad9*), es importante una estrategia eficiente para el clonado de secuencias completas de codificación (CDSs). Se observaron porciones ortólogas específicas de regiones no traducidas (UTRs) basado en alineamientos de secuencias múltiples de ortólogos de genes mitocondriales que codifican subunidades del complejo I en las especies vegetales. Basado en la conservación de UTRs parciales, se determinó una estrategia PCR de un paso para la clonación de CDSs de los nueve ortólogos de genes mitocondriales. Usando esta estrategia, las cinco estructuras de lectura abierta mitocondriales (ORFs) , que codifican subunidades de la NADH deshidrogenasa mitochondrial (i.e., *nad1*, *nad2*, *nad6*, *nad7*, y *nad9*), se clonaron en tres especies de angiospermas: *Hibiscus cannabinus*, *Cinnamomum camphora*, y *Boehmeria nivea*. Los quince productos PCR clonados también incluyeron las secuencias parciales 5' y 3'-UTR. Más aún, se identificó un sitio de edición potencial C-U RNA en el codón de iniciación del nad9 de *Hibiscus cannabinus*. En conclusión, la estrategia simple y eficiente evita el uso del proceso de amplificación rápida de las terminaciones de cDNA (RACE) que consume mucho tiempo, y facilita la clonación del ORFs mitochondrial completo cuyos lados 5' y 3' del UTR contienen una region ortóloga con alguna degeneración en las especies de plantas superiores.

Palabras clave: Subunidad de la NADH deshidrogenasa; Edición, clonación del RNA; Iniciador degenerado; Región no traducida.

Abreviaturas. CDS: secuencia de codificación; UTR: región no traducida; CMS: esterilidad citoplásmica en el macho; mtDNA: DNA mitocondrial; ORF: estructuras de lectura abierta; RACE: rapida amplificación de terminaciones cDNAs.

[1] College of Agriculture, Guangxi University, Nanning, Guangxi 530004, China
[2] Guangxi Subtropical Crops Research Institute, Nanning, Guangxi 530001, China
[3] Cash Crops Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, Guangxi 530007, China
[4] Guangxi Academy of Specialty Crops, Guilin, Guangxi 541004, China
Address correspondence to: Prof. Ruiyang Zhou, ruiyangzhou@163.com. Dr. Xiangmin Tang, txm906@163.com. Mr. Qiang Huang, gxscrijg@163.com. Mr Tao Chen, 499481815@qq.com. Jin G and Tang X-M contributed equally to this work.
Received 15.XII.2016. Accepted 20.V.2017.

INTRODUCTION

Mitochondrion play crucial role in bioenergy production (Schwarzländer & Finkemeier, 2013), cellular development and metabolism (Mackenzie & McIntosh, 1999), signaling (Millar et al., 2008; Chandel, 2014), programmed cell death (Yao et al., 2004; Kim et al., 2006), and biotic stress responses (Amirsadeghi et al., 2006). In higher plants, mitochondrial genes are closely related to several phenotypes such as cytoplasmic male sterility (CMS) (Wang et al., 2006) and stress resistance (Maxwell et al., 2002). The nine subunits of mitochondrial complex I, i.e., *nad9*, *nad7*, *nad6*, *nad5*, *nad4L*, *nad4*, *nad3*, *nad2*, and *nad1*, are encoded by mitochondrial DNA. Functional deficiency in complex I has been associated with various deficient phenotypes, including delay in flowering time and abnormal development.

RNA editing at the transcripts of *nad1*, *nad2*, *nad5* and *nad7* is essential for mitochondrion biogenesis, cell division, and plant growth (Kühn et al., 2011; Sosso et al., 2012; Li et al., 2014). To identify RNA editing sites in coding regions of the chondriogenes, suitable and efficient strategies are necessary to clone complete coding sequences (CDSs). To acquire complete open reading frames including start and stop codens, chondriogenes can be traditionally cloned by homologous amplification for internal fragments and rapid amplification of cDNA ends (RACE) technology or Genome Walking for 3' and 5' end unknown sequences (Liu et al., 2012). However, these procedures are complicated and laborious. In addition, the cloned sequences of chondriogenes did not convey complete CDS information when upstream and downstream primers cross at the start and stop codons, respectively. Therefore, the potential RNA editing sites could not be identified in primer regions.

Recently, high-throughput sequencing technology provides an efficient method to obtain the complete sequence of plant mitochondrial DNA. Several studies have shown that higher plant mtDNA is highly conserved in the coding sequences. Non-coding regions of plant mitochondrial genomes are extraordinarily dynamic with respect to DNA recombination (Handa, 2003). In the present study, orthologous UTR portions were found based on multiple sequence alignments of nine orthologous chondriogenes encoding complex I subunit. According to these results, a universal and efficient strategy based on the conserved UTR portion was presented to clone the complete coding sequences of chondriogenes encoding NADH dehydrogenase subunits in plant species by one-step cloning.

MATERIALS AND METHODS

Plants material. The kenaf cultivar 'P3' was grown in an experimental field at Guangxi Academy of Agricultural Sciences, Nanning, China (108° 14' E, 22° 50' N). The mature camphortree and ramie plants were maintained under natural conditions at the Guangxi Subtropical Crops Research Institute, Nanning, China (108° 20' E, 22° 53' N). According to the APGIII system, a distance tree showed distant phylogenetic relationships between the three species and other plant species with sequenced mitochondrial genome.

Homology analysis of untranslated regions. The mitochondrial genome sequences of 20 species (Table 1) were obtained from the GenBank database. The twenty angiosperm species included both monocots and eudicots referencing the APGIII system. Based on the annotations of their mitochondrial genomes, the 3' and 5'-UTR in orthologous chondriogenes were analyzed by multiple alignment. A 200-bp sequence upstream of the start codon and downstream of the stop codon was referred as the 'putative UTR portion'. DNAMAN software (Lynnon Corporation, USA) was used to obtain the results of multiple alignment of the nine orthologous chondriogenes encoding complex I subunits.

Table 1. Plant species analyzed in this study and accession data.
Tabla 1. Especies vegetales analizadas en este estudio y datos de las variedades.

Plant species	mtDNA GenBank no.
Arabidopsis thaliana	NC_001284.2
Beta vulgaris subsp. vulgaris	NC_002511.2
Butomus umbellatus	NC_021399.1
Carica papaya	NC_012116.1
Gossypium hirsutum	NC_027406.1
Oryza sativa	NC_001751.1
Populus tremula	NC_028096.1
Raphanus sativus	NC_018551.1
Cucumis sativus	NC_016004.1
Glycine max	NC_020455.1
Liriodendron tulipifera	NC_021152.1
Citrullus lanatus	NC_014043.1
Ajuga reptans	NC_023103.1
Silene conica	NC_016219.1
Vitis vinifera	NC_012119.1
Triticum aestivum	NC_007579.1
Sorghum bicolor	NC_008360.1
Nicotiana tabacum	NC_006581.1
Daucus carota subsp. sativus	NC_017855.1
Lotus japonicus	NC_016743.2

Primer design for target genes. Based on multiple alignment, degenerate primer pairs were designed in conserved UTR regions of the five orthologous gene (*nad1, nad2, nad5, nad7,* and *nad9*) for the cloning of complete CDSs (Table 2). Primers of the target genes for RT-PCR were designed and analyzed using Primer Premier V5.0.

Table 2. List of designed degenerate primers.
Tabla 2. Lista de iniciadores degenerados.

Gene names	Primers	Primer sequences	Tm/°C
nad1	nad1F	CGAGAAAAGGTCCCMTCC	51
	nad1R	TAGCCCGCTTCCCATTN	
nad2	nad2F	ATACGGGGGAAATGAAAGC	53
	nad2R	GAAYCTTSTCAATGATCGAAC	
nad6	nad6F	TCACGAWTTGGATTYGAAC	51
	nad6R	GCATGAAAGTTCCATTTCAG	
nad7	nad7F	CGAAGGAGATGCATTTCTG	53
	nad7R	CAATAAAGCTARGGTCCTGATC	
nad9	nad9F	TAGCATTTCCTATTGMTTTGTCC	50
	nad9R	GCCTTTCTTGGTYGGACCA	

Total RNA extraction and molecular cloning. Total RNA was extracted from the young leaf tissues of three plant species metioned above (Chang et al., 1993; Shi et al., 2016) and reverse transcribed with random hexamers using M-MLV transcriptase (Promega, USA). Taq DNA polymerase (Biomiga, USA), which is capable of adding a single 3'-A overhang to PCR products, was used for PCR amplification with cDNA templates and degenerate primer pairs specific to the target chondriogenes (*nad1*, *nad2*, *nad5*, *nad7*, and *nad9*) in a Tprofessional-96 thermal cycler (Biometra, Germany). PCR products were analysed on 1.2% agarose gels and the single target fragment was recovered and purified using a gel extraction kit (Biomiga, USA). The purified PCR products were subcloned into pUCm-T vector, sequenced, and assembled using DNAstar SeqMan software (Madison, WI, USA).

RESULTS

Sequence analysis of chondriogene orthologs. UTR orthologues of key chondriogenes in higher plants were analysed. According to the results of multiple sequence alignments (data not shown), orthologues of several chondriogenes in angiosperms showed evolutionary homology in partial 5' and 3'-UTRs. In addition to chondriogenes encoding complex I, some other genes showed evolutionarily conserved 5' or 3'-UTRs (Table 3).

The application of cloning strategy. The overall cloning strategy was illustrated with kenaf nad9 as an example (Fig. 1 and 2). The five mitochondrial genes (*nad1*, *nad2*, *nad6*, *nad7*, and nad9) of kenaf, camphor tree, and ramie, which encode corresponding NADH dehydrogenase subunits, were successfully cloned by using this strategy. Both *nad1* and *nad7* genes included introns, suggesting that this strategy could clone chondriogenes with introns. In higher plants, the sequences encoding complex I subunits show evolutionary conservation

Table 3. Orthology of several chondriogenes in angiosperm plant.
Tabla 3. Ortología de varios genes mitocondriales en angiospermas.

Gene names	5'-UTR	3'-UTR
nad1	√	√
nad2	√	√
nad3	√	√
cob	√	√
rps3	×	√
rps4	×	√
rps12	√	×
matR	√	×
mttB(orfx)	×	√
coxIII	√	×
nad4	√	√
nad5	√	√
nad6	√	√
nad7	√	√
nad4L	√	√
nad9	√	√

Note: The table includes the chondriogene of orthologous UTR ≥ 30 bp, with the sequence identity greater than or equal to 90%. √ indicate orthology, × indicate non-orthology.
Nota: La Tabla incluye los ortólogos UTR de longitud mayor ó igual a 30 pb e identidad mayor o igual a 90%. √ indica ortología, × indica ausencia de ortología.

Table 4. GenBank accession number of the fifteen cloned chondriogenes.
Tabla 4. Número de accesión en el Banco de Genes de los 15 genes mitocondriales clonados.

Plant species	Gene name	GenBank no.
Hibiscus cannabinus	*nad1*	KU755467
	nad2	KU755453
	nad6	KU755461
	nad7	KU755458
	nad9	KU755464
Cinnamomum camphora	*nad1*	KU755456
	nad2	KU755454
	nad6	KU755462
	nad7	KU755459
	nad9	KU755465
Boehmeria nivea	*nad1*	KU755457
	nad2	KU755455
	nad6	KU755463
	nad7	KU755460
	nad9	KU755466

in partial UTR, suggesting a crucial role in transcription. All of the 15 cloned genes were submitted to NCBI (Table 4). Notably, a potential C-U RNA editing site was found in the start codon of kenaf *nad9* (Fig. 1 and 2).

The nucleotide sequences exhibited a high degree of similarity to the corresponding orthologues (data not shown). According to the APGIII system, ramie is a member of the Rosales order, and kenaf is a member of the Malvales order. Both Rosales and Malvales are ascribed to the eurosids I branch. However, the camphor tree is a member of the magnoliids, which have a distant phylogenetic relationship with Rosales and Malvales. These results suggest that the conservation of partial UTR in the cloned chondriogenes is a common phenomenon in angiosperm species, and the designed degenerate primers could be used for cloning complete ORFs of the corresponding genes in angiosperms.

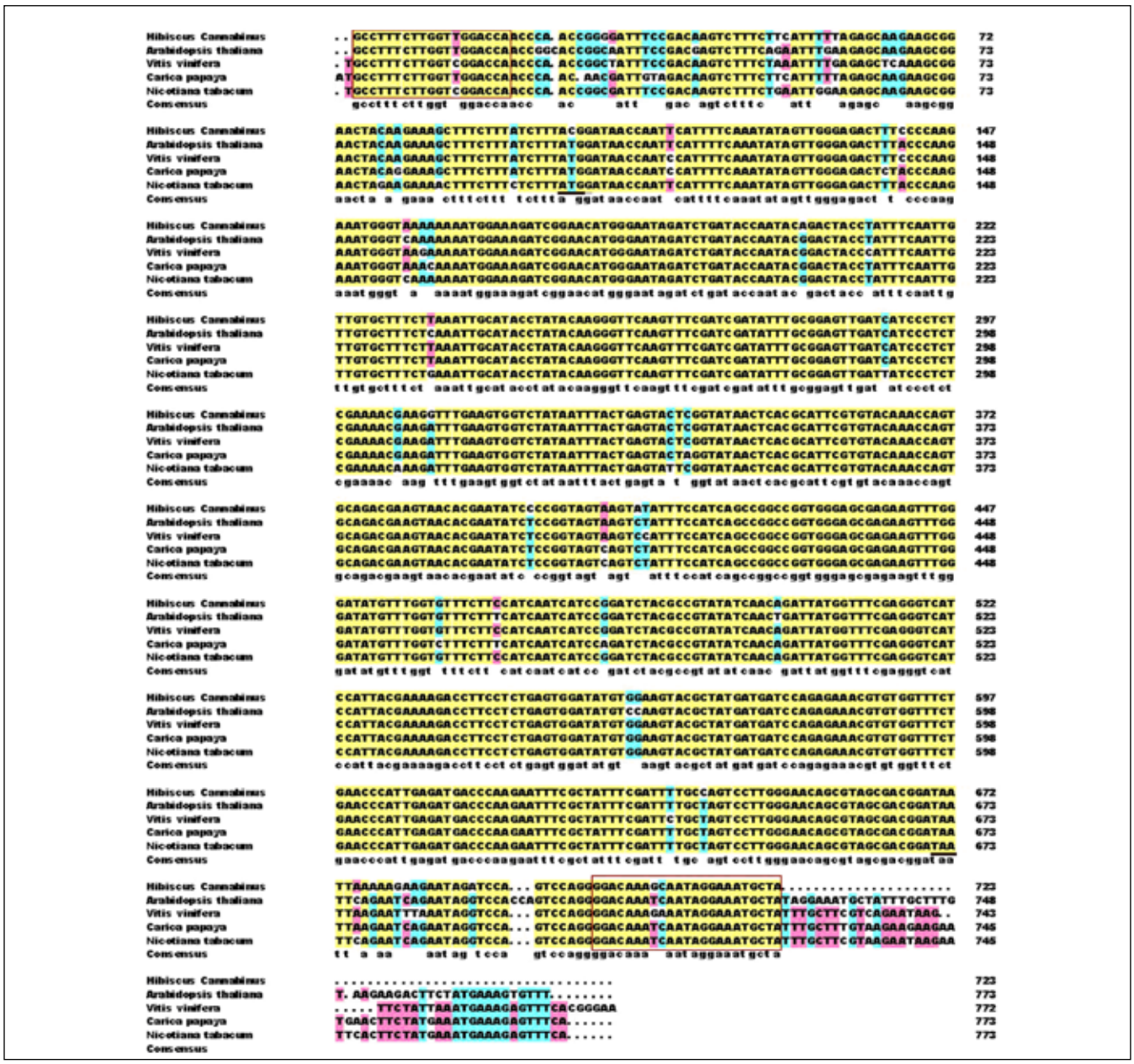

Fig. 1. Illustrated map of cloning chondriogenes encoding complex I subunits, using *nad9* from *Hibiscus cannabinus* as an example. Note: The two boxes indicate the primer sequences; ATG and TAA are underlined. A potential C-U RNA editing site was identified in the start codon.

Fig. 1. Mapa de los genes mitocondriales que codifican las subunidades del complejo I usando "*nad9*" de *Hibiscus cannabinus* como un ejemplo. Nota: los dos recuadros señalan la posición de los iniciadores; codones ATG y TAA se marcan con subrayado. Un potencial sitio de editado de ARN se identificó en el codón de iniciación.

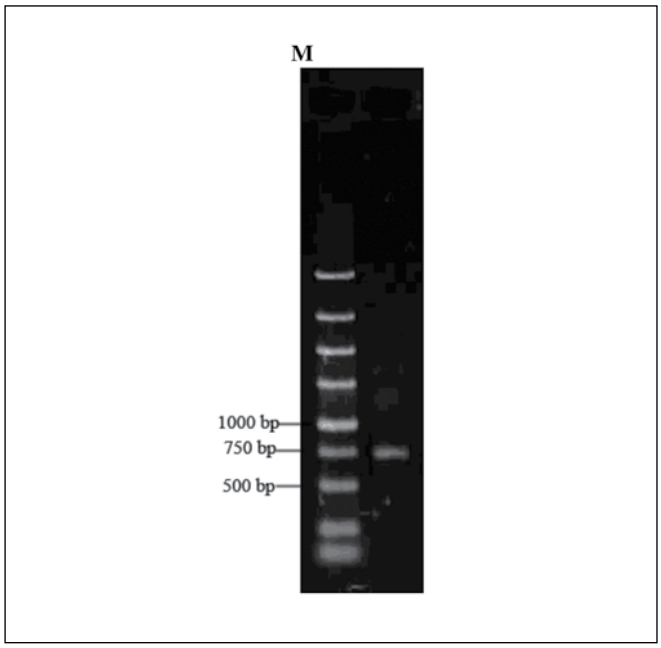

Fig. 2. PCR amplification of the complete CDS encoding nad9 from *Hibiscus cannabinus*. M: DNA Marker 2000.
Fig. 2. Amplificación de la secuencia codificante completa de nad9 de *Hibiscus cannabinus*. Marcador de peso 2000.

DISCUSSION

Selective evolutionary pressure could result in the conservation of certain nucleotide sequences that may affect gene expression (Hardison, 2000; Mrowka et al., 2003). It was found that the 5'-UTRs of the *orf25*, *atp6*, and *coxII* mRNAs contain three conserved sequence blocks compared to plant mitochondrial mRNAs from both monocot and dicot species (Pring et al., 1992). The conservation of these non-coding regions may be under evolutionary selective pressure to remain relatively unchanged because it binds to transcription factor or miRNA (Binder et al., 1996; Toscano-Garibay & Aquino-Jarquin, 2012).

The results from the present study showed conserved 3'and 5'-UTR regions of the cloned chondriogenes in the molecular evolution of angiosperms. Based on the conservation of untranslated regions in chondriogenes, it is concluded that the orthologous UTR regions could play an essential role in the regulation of the expression of chondriogenes encoding complex I. Furthermore, a universal and efficient strategy was developed to clone complete coding sequences of chondriogenes encoding complex I by one-step homologous cloning. The cloning strategy will be useful for cloning mitochondrial complete ORFs whose 5' and 3' flanking UTR contain an orthologous region with some degeneracy among angiosperms.

The feasibility of the strategy presented in the present study was verified by cloning of *nad1*, *nad2*, *nad6*, *nad7*, and *nad9* genes in kenaf, ramie, and camphor trees. Orthologous UTRs in higher plant chondriogenes could be utilized as a new molecular marker system for the analysis of genetic diversity of maternal inheritance.

However, not all UTRs of chondriogenes remain conserved. For example, the UTR portion of *orfB* and cox1 in rice (*Oryza sativa*) show faster evolutionary speeds (data not shown), perhaps due to less evolutionary pressure. Despite some insufficiencies, the cloning strategy reported here could increase the efficiency of cloning chondriogenes encoding complex I subunits in plants, and help identify RNA editing sites in coding regions of mitochondrial complex I subunit genes. The strategy could also be applied to cloning chondriogenes encoding complex I in animals.

CONCLUSION

In angiosperm species, comparison of partial 3' and 5'-UTR regions in orthologous chondriogenes reveals some conserved aspects of the structure of UTRs. The conservative property facilitates the cloning of complete coding sequences of chondriogenes encoding NADH dehydrogenase subunits.

ACKNOWLEDGEMENTS

This work was supported by the Natural Science Foundation of Guangxi (No. 2016GXNSFAA380093, 2016GXNSFBA380127, 2014GXNSFAA118130) and the Earmarked Fund for China Agriculture Research System (CARS-19-S15).

REFERENCES

Amirsadeghi, S., C.A. Robson & G.C. Vanlerberghe (2006). The role of the mitochondrion in plant responses to biotic stress. *Physiology Plantarum* 129: 253-266.

Binder, S., A. Marchfelder & A. Brennicke (1996). Regulation of gene expression in plant mitochondria. *Plant Molecular Biology* 32: 303-314.

Chandel, N.S. (2014). Mitochondria as signaling organelles. *BMC Biology* 12: 34.

Chang, S.J., J. Puryear & J. Cairney (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11: 113-116.

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 Research* 31: 5907-5916.

Hardison, R.C. (2000). Conserved noncoding sequences are reliable guides to regulatory elements. *Trends in Genetics* 16: 369-372.

Kim, M., J.H. Lim, C.S. Ahn, K. Park, G.T. Kim, W.T. Kim & H.S. Pai (2006). Mitochondria-associated hexokinases play a role in the control of programmed cell death in *Nicotiana benthamiana*. *Plant Cell* 18: 2341–2355.

Kühn, K., C. Carrie, E. Giraud, Y. Wang, E.H. Meyer, R. Narsai, C.C. des Francs-Small, B. Zhang, M.W. Murcha & J. Whelan (2011). The RCC1 family protein RUG3 is required for splicing of nad2 and complex I biogenesis in mitochondria of *Arabidopsis thaliana*. *Plant Journal* 67: 1067-1080.

Li, X.J., Y.F. Zhang, M. Hou, F. Sun, Y. Shen, Z.H. Xiu, X. Wang, Z.L. Chen, S.S. Sun, I. Small & B.C. Tan (2014). Small kernel 1 encodes a pentatricopeptide repeat protein required for mitochondrial nad7 transcript editing and seed development in maize (*Zea mays*) and rice (*Oryza sativa*). *Plant Journal* 9: 797-809.

Liu, X.L., S.W. Zhang, J.Q. Duan, G.H. Du & F.H. Liu (2012). Mitochondrial genes *atp6* and *atp9* cloned and characterized from ramie [*Boehmeria nivea* (L.) Gaud.] and their relationship with cytoplasmic male sterility. *Molecular Breeding* 30: 23-32.

Mackenzie, S. & L. McIntosh (1999). Higher plant mitochondria. *Plant Cell* 11: 571-586.

Maxwell, D.P., R. Nickels & L. McIntosh (2002). Evidence of mitochondrial involvement in the transduction of signals required for the induction of genes associated with pathogen attack and senescence. *Plant Journal* 29: 269-279.

Millar, A.H., I.D. Small, D.A. Day & J. Whelan (2008). Mitochondrial biogenesis and function in *Arabidopsis*. *Arabidopsis Book/ American Society of Plant Biologists* 6: e0111.

Mrowka, R., K. Steinhage, A. Patzak & P.B. Persson (2003). An evolutionary approach for identifying potential transcription factor binding sites: the renin gene as an example. *American Journal Physiology Regulatory Integrative Comparative Physiology* 284: R1147-1150.

Pring, D.R., J.A. Mullen & F. Kempken (1992). Conserved sequence blocks 5' to start codons of plant mitochondrial genes. *Plant Molecular Biology* 19: 313-317.

Shi, X.P., C.J. Zhang, Q.H. Liu, Z. Zhang, B. Zheng & M.Z. Bao (2016). De novo comparative transcriptome analysis provides new insights into sucrose induced somatic embryogenesis in camphor tree (*Cinnamomum camphora* L.). *BMC Genomics* 17: 26.

Schwarzländer, M. & I. Finkemeier (2013). Mitochondrial energy and redox signaling in plants. *Antioxid Redox Signal* 8: 2122-2144.

Sosso, D., S. Mbelo, V. Vernoud, G. Gendrot, A. Dedieu, P. Chambrier, M. Dauzat, L. Heurtevin, V. Guyon, M. Takenaka & P.M. Rogowsky (2012). PPR2263, a DYW-Subgroup Pentatricopeptide repeat protein, is required for mitochondrial nad5 and cob transcript editing, mitochondrion biogenesis, and maize growth. *Plant Cell* 24: 676-691.

Toda, T., S. Fujii, K. Noguchi, T. Kazama & K. Toriyama (2012). Rice MPR25 encodes a pentatricopeptide repeat protein and is essential for RNA editing of *nad5* transcripts in mitochondria. *Plant Journal* 72: 450-460.

Toscano-Garibay, J.D. & G. Aquino-Jarquin (2012). Regulation exerted by miRNAs in the promoter and UTR sequences: MDR1/ P-gp expression as a particular case. *DNA Cell Biology* 31: 1358-1364.

Wang, Z.H., Y.J. Zou, X.Y. Li, Q.Y. Zhang, L.T. Chen, H. Wu, D.H. Su, Y.L. Chen, J.X. Guo & D. Luo (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.

Yao, N., B.J. Eisfelder, J. Marvin & J.T. Greenberg (2004). The mitochondrion – an organelle commonly involved in programmed cell death in *Arabidopsis thaliana*. *Plant Journal* 40: 596-610.