RNA EDITING IN *Calotropis procera* MITOCHONDRIAL NADH-DEHYDROGENASE SUBUNIT 3 GENE

A. M. RAMADAN[1,2]

1. *Department, Agricultural Genetic Engineering Research Institute (AGERI), Agriculture Research Center (ARC)*
2. *Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), PO Box 80141, Jeddah 21589, Saudi Arabia b Plant Molecular Biology, Giza, Egypt*

RNA editing refers to posttranscriptional alterations of RNA molecules through insertion, deletion, or modification of nucleotides, not including RNA splicing, capping, or polyadenylation (Nishikura, 2006; Farajollahi and Maas, 2010). RNA Editing was discovered for the first time in trypanosome mitochondria (Benne *et al*., 1986). RNA editing occurred as differences between genomic sequences and the corresponding RNA sequences. The predominant type of RNA editing in animals is the conversion of adenosine (A) to inosine (I), catalyzed by a family of adenosine deaminases that act on RNA (Nishikura, 2006). This editing is also known as A-to-G editing because inosine in RNA is read as guanosine (G) by the translational machinery (Nishikura, 2006). Another well-documented type of RNA editing in animals is cytidine-to-uridine (C-to-U) editing, catalyzed by the activation-induced cytidine deaminase/ apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like family of deaminase, but it is less frequent than A-to-G editing (Nishikura, 2006). In land plants, RNA editing highly specifically converts cytidine to uridine nucleotides in transcripts of both plastid and mitochondrial genes (Castandet and Araya, 2011); 34 cytidine residues in plastids and more than 500 residues in mitochondria have been reported to be editing target sites in *Arabidopsis thaliana* (Chateigner-Boutin and Small, 2007; Bentolila *et al*., 2008). Analysis of RNA editing in higher plant mitochondrial transcripts specifying the cytochrome b (*cytb*), subunit 1 of the NADH-dehydrogenase (*nadl*) and cytochrome oxidase subunits II and HI (*coxll* and *coxIll*) had revealed homogeneously edited cDNAs for these loci (Hiesel *et al*., 1989). RNA editing the *nad3* locus predominantly involves modification of cytidines to be recognized as uridines by the reverse transcriptase and presumably the ribosome (Schuster *et al*., 1990). One reverse alteration has been observed in the cytochrome b locus modifying a genomic encoded T to C in the cDNA sequence (Hiesel, *et al*., 1989).

A number of cytosines are altered to be recognized as uridines in transcripts of the nad3 locus in mitochondria of the higher plant Oenothera. Such nucleotide modifications can be found at 16 different sites within the *nad3* coding region of *Oenothera* mitochondria (Schuster *et al*.,

Egypt. J. Genet. Cytol., 43:353-364, July, 2014
Web Site (*www.esg.net.eg*)

1990) and 15 sites *Carthamus tinctorius* (Kalinati *et al*., 2008). The role of *nad3* editing in drought tolerance was investigated (Yuan and Liu, 2012).

Calotropis procera is flowering plant in the poison family, Apocynaceae, natively grown in North Africa, Tropical Africa, Western Asia, South Asia, and Indochina (Aiton, 2010). *Calitropis* species show high grown performance during the dry season, implying the occurrence of special strategies of drought tolerance (Colombo *et al*., 2007; Khan *et al*., 2007; Boutraa, 2010).

In our study, *nad3* gene was identified from genomic DNA (accession no. KP171516) and cDNA (accession no. KP171517) in desert plant *Calotropis procera*, and then RNA editing was investigated in 11 positions of this mitochondrial gene lead to change 11 amino acid in peptide sequence.

MATERIALS AND METHODS

Sample collection and isolation of total RNA and DNA

Three leaf discs of *C. procera* were collected from Jeddah region (KSA, latitude 21°26'6.00, longitude 39°28'3.00. Samples were frozen in liquid nitrogen (50 mg tissue each) and total RNA extraction was performed using RNeasy Plant Mini Kit (Qiagen, cat. no. 74903). To remove DNA contaminants, 3 µl of 10 mg/ml RNase A, DNase and protease-free (Thermo Scientific cat no. EN0531) were added to the RNA samples and tube was incubated at 30°C for 15 min. DNeasy Plant Mini Kit (Qiagen, cat. no. 69106) was used for DNA isolation Estimation of the DNA and RNA concentration in different samples was done by measuring optical density at 260 nm. DNA and RNA samples were sent to Beijing Genomics Institute (BGI), Shenzhen, China, for deep sequencing and dataset were provided for analysis.

Next-Generation Sequencing (NGS)

Whole-RNA-seq and DNA-seq, paired-end short-sequence reads of *C. procera* were generated using the Illumina Genome AnalyserIIx (GAIIx) according to manufacturer's instructions (Illumina, San Diego, CA).

Sequence filtering and bioinformatics analysis

The raw sequencing data were obtained using the Illumina python pipeline v. 1.3. For the obtained libraries, only high quality reads (quality >20) were retained. Then, reference assembly using *Rhazya stricta* mitochondrion DNA (accession No. KJ485850) as a reference DNA of the obtained short (paired-end) read dataset was performed using assembler CLC Genomics workbench 3.6.5.

Ten *nad3* sequences (Table 1) belonging to other plant species were obtained from GenBank and used as a reference for blasting (http://www.ncbi.nlm.nih.gov/BLAST). To produce *nad3* cDNA, genomic nad3 was used as a refer-

ence for raw RNA sequencing data (Illumina python pipeline v. 1.3).

Analysis of RNA editing and deduced amino acids

The genomic and cDNA sequences of nad3 transcripts of *Calotropis procera* obtained in the present study were analyzed for RNA editing status using multi sequence alignment using CLC genomic work bench 3.6.5 (http://www.clcbio.com/products/clc-genomics-workbench). Also, protein multi sequence alignment was achieved using the same program

Domain analysis

The functional domains were identified from the NCBI's conserved domain database (CDD) http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml)

Accession Numbers

Sequence data from this article have been submitted to GenBank data library under accession numbers; *C. procera* genomic nad3 gene (accession no. KP171516) and *C. procera* cDNA *nad3* gene (accession no. KP171517). Nad3 sequences of other plant species accession no. are*: Asclepias syriaca* (KF541337), *Rhazya stricta* (KJ485850), *Boea hygrometrica* (JN107812), *Salvia miltiorrhiza* (KF177345), *Petunia axillaris* (U61394), *Ajuga reptans* (KF709392), *Mimulus guttatus* (JN098455), *Vitis vinifera* (GQ220323), *Nicotiana tabacum* (BA000042), *Nicotiana sylvestris* (X96741).

Protein sequences accession no. are: *Allium cepa* (Q96007), *Helianthus annuus* (P60159), *Pinus sylvestris* (Q36664), *Panax ginseng* (P27062), *Solanum tuberosum* (O99869).

RESULTS AND DISCUSSION

Nad3 is a subunit of complex I of the electron transport chain in mitochondria. Interruption in *nad3* editing lead to accumulate large concentrations of ROS which leads to the deterioration afford to drought in *Arabidopsis* (Yuan and Liu, 2012). So we will try through this study to understand RNA editing of *nad3* in desert plant.

Characterization of C. procera nad3 gene

Through this study, *Nad3* gene was characterized in *C. procera* (accession no. KP171516) using DNAseq raw data. A total of 71,349,934 paird-end short DNA sequence reads was generated for *C. procera* using the HiSeq 2000 Illumina platform (Illumina, San Diego, CA). *Nad3* gene of *Rhazya stricta* (KJ485850) was used as reference in CLC genomic workbench. The best BLAST search hits were used to perform multi-sequence alignment (Table 1). This resulted in 10 *nad3* gene sequences from 10 different species, in addition to *C. procera*. A multiple sequence alignment of the 11 sequences was obtained (Fig. 1). Many investigators were used CLC genomic workbench to perform genome sequencing and characterize genes in different bio-systems (Christopher *et al*., 2011; Cerna *et al*., 2014; Courtney *et al*., 2014).

Characterization of C. procera nad3 mRNA

cDNA *nad3* gene in *C. procera* (accession no. KP171517) was characterized using RNAseq raw data. A total of 215, 841 and 902 paird-end short RNA sequence reads was generated for *C. Procera* using the HiSeq 2000 Illumina platform (Illumina, San Diego, CA). *Nad3* gene of *C. procera* (accession no. KP171516) was used as a reference in CLC genomic workbench program. Investigators used traditional methods to isolate and identify the cDNA, which depend on using 9 to 10 clones to confirm the right sequences (Hiesel *et al*., 1989; Schuster *et al*., 1990; Kalinati *et al*., 2008). On other hand, Anders and Huber (2010) used NGS data which contain millions of reads to confirm the right sequences depending on CLC genomic workbench program (http://www.clcbio.com/products/clc-genomics -workbench).

RNA editing in nad3 transcript

RNA editing is common in most organisms especially in mitochondria (Chateigner-Boutin and Small, 2007; Bentolila *et al*., 2008; Castandet and Araya, 2011). Editing in *nad3* gene was reporting (Schuster *et al*., 1990; Kalinati *et al*., 2008). RNA editing in *nad3* gene was detected in 14 sites in *Oenothera* mitochondria (Schuster *et al*., 1990), 19 sites in carrot mitochondria (Rurek *et al*., 2001) and 16 sites in safflower (Kalinati *et al*., 2008). In study, a comparison between *nad3* sequences of the genomic and cDNA (Fig. 2) revealed editing in the transcript. Editing is revealed in 11 sites (nucleotide no. 44, 62, 80, 209, 215,230, 247, 266, 275, 317 and 349). All of which were C to U conversion. Total of 11 amino acid substitution were detected due to editing, the most common being proline to leucine (P-L). Other changes were serine to leucine (S-L), serine to phenylalanine (S-F), proline to serine and arginine to tryptophan (R-W) (Table 2 and Fig. 2). Generally in *Arabidopsis* mitochondria, RNA editing is increase the proportion of hydrophobic amino acid codons (Giege´ and Brennicke, 1999). So it is suggested that increasing protein hydrophopicity is suitable to protein and enzyme function in mitochondrial membrane like *nad3* protein (Kalinati *et al*., 2008). The interruption of C250 editing (cytosine base No. 250 in nad3) lead to accumulate large concentrations of ROS Which leads to the deterioration of drought in Arabidopsis (Yuan and Liu, 2012). Although *Calotropis procera* is a desert plant, but there is no editing in C250. By check the edited amino acid in this site, proline edited to serine in *Arabidopsis*, *rice* and *sorghum* (Yuan and Liu, 2012), but in *Calotropis procera*, leucine is not edited. Several Investigators reported that it is normal and necessary the presence of serine or leucine in protein binding or recognition sites but proline is not normal in previous sites (Matthew *et al*., 2003). So we suggest that *C. procera nad3* does not need to be edited in this site but another *nad3* gene in other species which have proline in the same position may need to be edited in order for nad3 does not lose it′s activity. Partial RNA editing (some transcripts of the same gene

edit in certain sites and other not) was found in mitochondria of some plant species (Kalinaty *et al*., 2008), and other as well us not found this phenomena (Rurek *et al*., 2001). We suggest this heterogeneity is occurred according to RNA editing mechanism, which not exactly identified in plant till now (Aleel, 2011). Also, we excluded the effect of mtDNA copy on heterogeneous RNA editing because it is need different genomic nad3 sequences but investigators found that all clones of genomic nad3 gene of the same plant species have the same sequence, but the heterogenisity found in cDNA clones (Lu and Hanson, 1996; Kalinaty *et al*., 2008).

Analysis of the deduced protein sequence

Editing is only intermediate stage in the process of forming functional protein (Kalinati *et al*., 2008). The actual effect of editing needs to be assessed at the protein level. A comparison of amino acid sequences derived from genomic as well as cDNA of *C. procera* along with cDNA of derived amino acid profile of other species was achieved to clearly that editing in this gene of *C. procera* led to formation of conserved amino acid (Fig. 3).

Conserved domain analysis

Many investigators used to confirm the functionality of proteins (Copley *et al*., 2002; Ramadan *et al*., 2012; Shokry *et al*., 2014). Domain analysis indicated the presence of NADH-ubiquinone/ oxidoreductase, chain 3 (*nad3*). Conserved domain database accession number cl00535, and pfam accession number PF00507 (Fig. 4). Although conserved domain analysis of protein is classifying protein into families and predicting functional sites but this method cannot detect the activity difference between editing gene and it's original sequences because it depends on peptide sequence rather than amino acids properties. But the laboratory experiments proofed that interruption in nad3 editing results in the loss of its function (Yuan and Liu, 2012)

In conclusion, extensive editing takes place in transcript of *nad3* of *C. procera* and these edit sites are mostly conserved across plant species. This high degree of conservation in length and composition across plant species, as a result of *nad3* editing, indicates to the importance for editing. It seems that RNA editing minimizes the differences between sequences on protein level; in addition to maintain a conserved polypeptide sequence for this gene.

SUMMARY

Nad3 (NADH-dehydrogenase subunit 3) gene from genomic (accession no. KP171516) and cDNA (accession no. KP171517) was identified in desert plant *Calotropis procera* using RNA seq and DNA seq data. A number of cytosines are altered to be recognized as uridines in transcripts of the nad3 locus in mitochondria. The nucleotide modifications were found at 11 different nucleotide positions (nucleotide no. 44, 62, 80, 209, 215,230, 247, 266, 275, 317 and 349) within the *nad3* coding region. Heterogeneous RNA editing in *C. procera nad3* RNA was not

detected in this study. These alterations in the mRNA sequence change codon identities to specify 11 amino acids. The alteration in nucleotides leads to codons alteration specifying different amino acids, the common being proline to leuciene (P-L). Other changes were serine to leucine (S-L), serine to phenylalanine (S-F), proline to serine and arginine to tryptophan (R-W). These alterations are common in mitochondrial *nad3* gene of most plant species with few differences according to the properties of the amino acids involved.

REFERENCES

Aiton, W. T. (2010). Germplasm Resources Information Network. United States Department of Agriculture, 2001-10-19, Retrieved, 2010-06-26.

Aleel, K. G. (2011). To Thy Proteins Be True: RNA Editing in Plants. Plant Physiology, 156: 453-454.

Anders, S. and W. Huber (2010). Differential expression analysis for sequence count data. Genome Biol., 11: R106

Benne, R., J. Van den Burg, J. P. Brakenhoff, P. Sloof, J. H. Van Boom and M. C. Tromp (1986). Major transcript of the frame shifted *coxII* gene from *trypanosome* mitochondria contains four nucleotides that are not encoded in the DNA. Cell, 46: 819-826.

Bentolila, S., L. E. Elliott and M. R. Hanson (2008). Genetic architecture of mitochondrial editing in *Arabidopsis thal1iana*. Genetics, 178: 1693-1708.

Boutraa, T. (2010). Growth performance and biomass partitioning of the desert shrub *Calotropis procera* under water stress conditions. Research Journal of Agriculture and Biological Sciences, 6: 20-26.

Brown, C., T., L. K. Fishwick, B. M. Chokshi, M. A. Cuff, J. M. Jackson, T. Oglesby, A. T. Rioux, E. Rodriguez, G. S. Stupp, A. H. Trupp, J. S. Woollcombe-Clarke, T. N. Wright, W. J. Zaragoza, J. C. Drew, E. W. Triplett and L. Wayne (2011). Whole-genome sequencing and phenotypic analysis of *Bacillus subtilis* mutants following evolution under conditions of relaxed selection for sporulation. Applied and Environmental Microbiology, 77: 6867-6877.

Castandet, B. and A. Araya (2011). RNA editing in plant organelles. Why make it easy? Biochemistry (Mosc), 76: 924-931.

Chateigner-Boutin, A. L. and I. Small (2007). A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res., 35: e114.

Colombo, R., O. Marı´n, S. Iraza´bal and W. Tezara (2007). Relaciones hı´dricas, fotosı´ntesis y anatomı´a foliar de dos especies del ge´nero *Calotropis*. Interciencia, 32: 791-796.

Copley, R. R., T. Doerks, I. Letunic and P. Bork (2002). Protein domain analysis in the era of complete genomes. FEBS Lett., 513: 129-134.

Farajollahi, S. and S. Maas (2010). Molecular diversity through RNA editing: A balancing act. Trends Genet., 26: 221-230.

Giege´, P. and A. Brennicke (1999). RNA editing in *Arabidopsis* mitochondria affects 441 C to U changes in ORFs. Proc. Natl. Acad. Sci., USA, 96: 15324-15329.

Hiesel, R., B. Wissinger, W. Schuster and A. Brennicke (1989). RNA editing in plant mitochondria. Science, 246: 1632-1634.

Hollender, C. A., C. Kang, O. Darwish, A. Geretz, B. F. Matthews, J. Slovin, N. Alkharouf and Liu Zhongchi (2014). Floral trans-criptomes in woodland *Strawberry* uncover developing receptacle and anther gene networks. Plant Physiology, 165: 1062-1075.

Hui, Y. and L. Dong (2012). Functional disruption of the pentatrico-peptide protein SLG1 affects mitochondrial RNA editing, plant development, and responses to abiotic stresses in *Arabidopsis*. The Plant Journal, 70: 432-444.

Kalinati, Y. N., V. D. Kumar and S. S. Reddy (2008). RNA editing in NAD3 and ATP9 transcripts of saf-flower (*Carthamus tinctorius*). Int. J. Integr. Biol., 3: 143-149.

Katerina, C., J. Oppelt, L. Radova, K. Musilova, N. Tom, F. Pardy, J. Malcikova, K. Plevova, B. Tichy, Y. Brychtova, M. Doubek, M. Trbusek, J. Mayer, J. Koca, R. Calogero, S. Pospisilova and M. Mraz (2014). Abstract 5198: Identification of microRNAs involved in DNA damage response in malignant B cells and their biological and clinical relevance. Proceedings: AACR Annual Meeting 2014; April 5-9, 2014; San Diego, CA.

Khan, R., S. Shahzad, M. I. Choudhary, S. A. Khan and A. Ahmad (2007). Biodiversity of the endophytic fungi isolated from *Calotropis procera* (Ait.) R. Br. Pakistan J. Botany, 39: 2233-2239.

Lu, B. and M. R. Hanson (1996). Fully edited and partially edited *nad*9 transcripts differ in size and both are associated with polysomes in potato mitochondria. Nucleic Acids Res., 24: 1369-1374.

Matthew, J. B., R. B. Russell, M. R. Barnes and I. C. Gray (2003).

Amino Acid Properties and Consequences of Substitutions. Bioinformatics for Geneticists. John Wiley & Sons, Ltd. ISBNs: 0-470-84393-4 (HB); 0-470-84394-2 (PB).

Nishikura, K. (2006). Editor meets silencer: Crosstalk between RNA editing and RNA interference. Nat. Rev. Mol. Cell Biol., 7: 919-931.

Ramadan, A. M., A. M. Shokry, N. O. Gadalla, S. M. Hassan, S. Edris, M. A. Al-Kordy, O. A. Abuzinadah, J. S. M. Sabir, S. R. Al-Akilli, H. S. M. Al-Zahrani, R. M. Hussein, F. M. El-Domyati and A. Bahieldin (2012). Detection of a MAPK-Like Gene in *Calotropis procera* Plant from the De Novo Assembled Genome Contigs of the High Throughput Sequencing Dataset. Life Sci. J., 9: 157-166.

Rurek, M., S. Marek, A. Natalia, M. Barbara and A. Halina (2001). Differences in editing of mitochondrial *nad*3 transcripts from CMS and fertile carrots. Acta Biochimica Polonica, 48: 711-717.

Shokry, A., S. Al-Karim, A. Ramadan, N. Gadallah, Sanaa G. Al Attas, J. S. M. Sabir, Sabah M. Hassan, M. A. Madkour, R. Bressan, M. Mahfouz and A. Bahieldin (2014). Detection of a Usp-like gene in *Calotropis procera* plant from the de novo assembled genome contigs of the high-throughput sequencing dataset. C. R. Biologies, 337: 86-94.

Wolfgang, S., B. Wissinger, M. Unseld and A. Brennicke (1990). Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in Oenothera mitochondria. The EMBO Journal, 9: 263-269.

Table (1): Accession number for each DNA sequence, description, organism name and the calculated e-value of homologous sequence to *C. procera nad3* gene sequence identified using specialized BLAST search programs.

Accession	Description	T.S.	Ident. %	E-value
gb|KF541337.1|	Asclepiassyriaca mitochondrion, complete genome	654	99	0.0
gb|KJ485850.1|	Rhazya stricta mitochondrion, complete genome	627	98	4e-176
gb|JN107812.1|	Boea hygrometrica mitochondrion, complete genome	616	97	1e-172
gb|KF177345.1|	Salvia miltiorrhiza mitochondrion, complete ge...	610	89	4e-171
gb|U61394.1|	Petunia axillaris subsp. parodii atpA-2 ...	610	97	4e-171
gb|KF709392.1|	Ajuga reptans mitochondrion, complete genome	604	97	2e-169
gb|JN098455.1|	Mimulus guttatus mitochondrion, complete genome	604	97	2e-169
gb|GQ220323.1|	Vitis vinifera strain PN40024 mitochondrion, p...	604	97	2e-169
dbj|BA000042.1|	Nicotiana tabacum mitochondrial DNA, complete...	604	97	2e-169
emb|X96741.1|	Nicotiana sylvestris mitochondrial nad3, rps12 genes an...	604	97	2e-169

Table (2): Summary of RNA editing and encoded amino acids changes of *C. procera nad3* gene.

Edit site no.	Nucleotide position	Codon position	Amino acid change
1	44	15	TCG(S)-TTG(L)
2	62	21	CCA (P)-CTA(L)
3	80	27	CCA (P)-CTA(L)
4	209	70	CCT (P)-CTT(L)
5	215	72	CCG (P)-CTG(L)
6	230	77	TCC (S)-TTC (F)
7	248	83	CCT (P)-TCT (S)
8	266	90	CCG (P)-CTG(L)
9	275	92	TCT (S)-TTT (F)
10	317	106	TCT (S)-TTT (F)
11	349	117	CGG(R)-TGG (W)

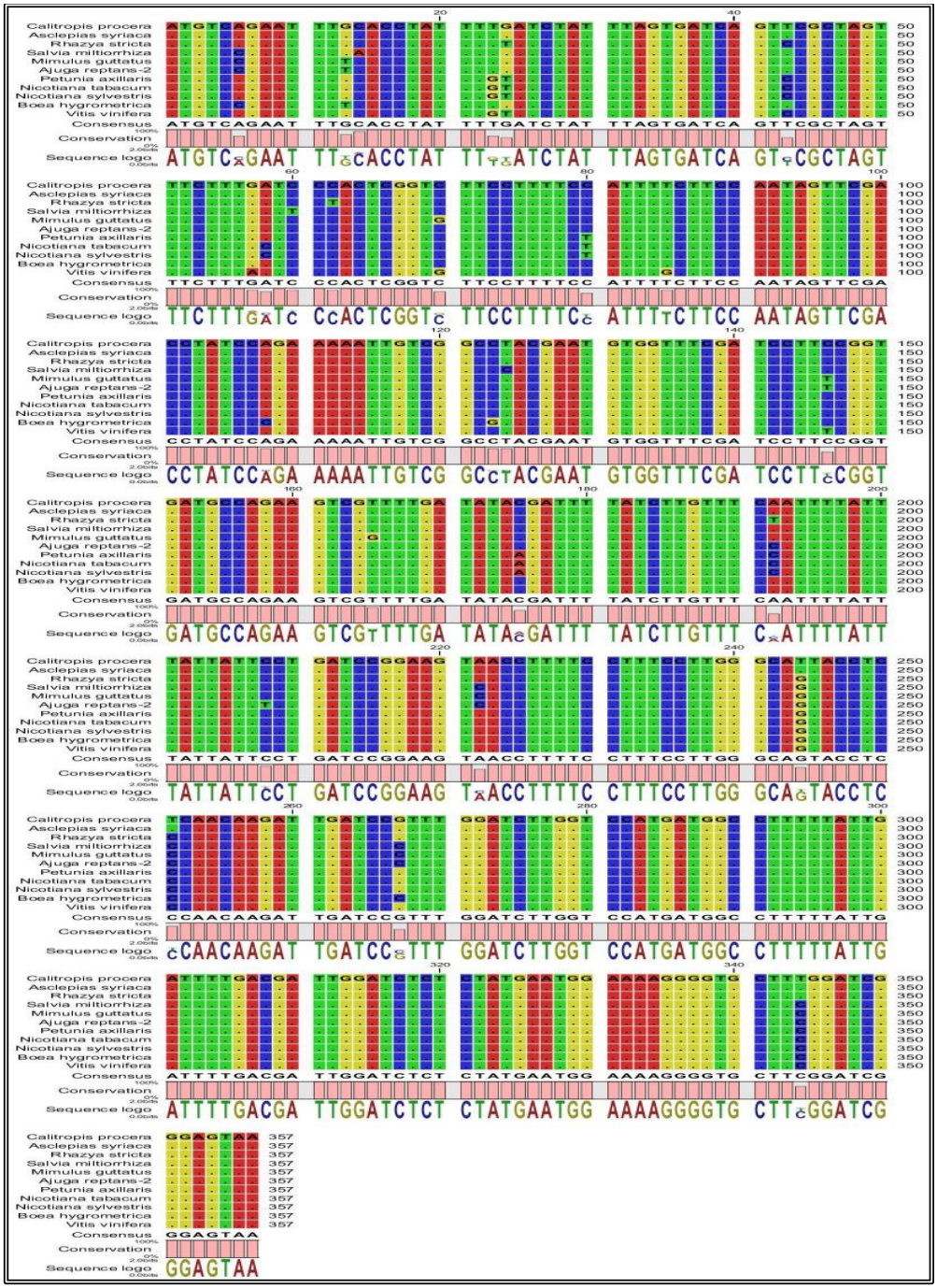

Fig. (1): A multiple sequence alignment of the 11 nad3 sequences included *C. procera* mt. genomic *nad3* sequence. Dots indicate to similarity to *C. procera* genomic *nad3* sequence.

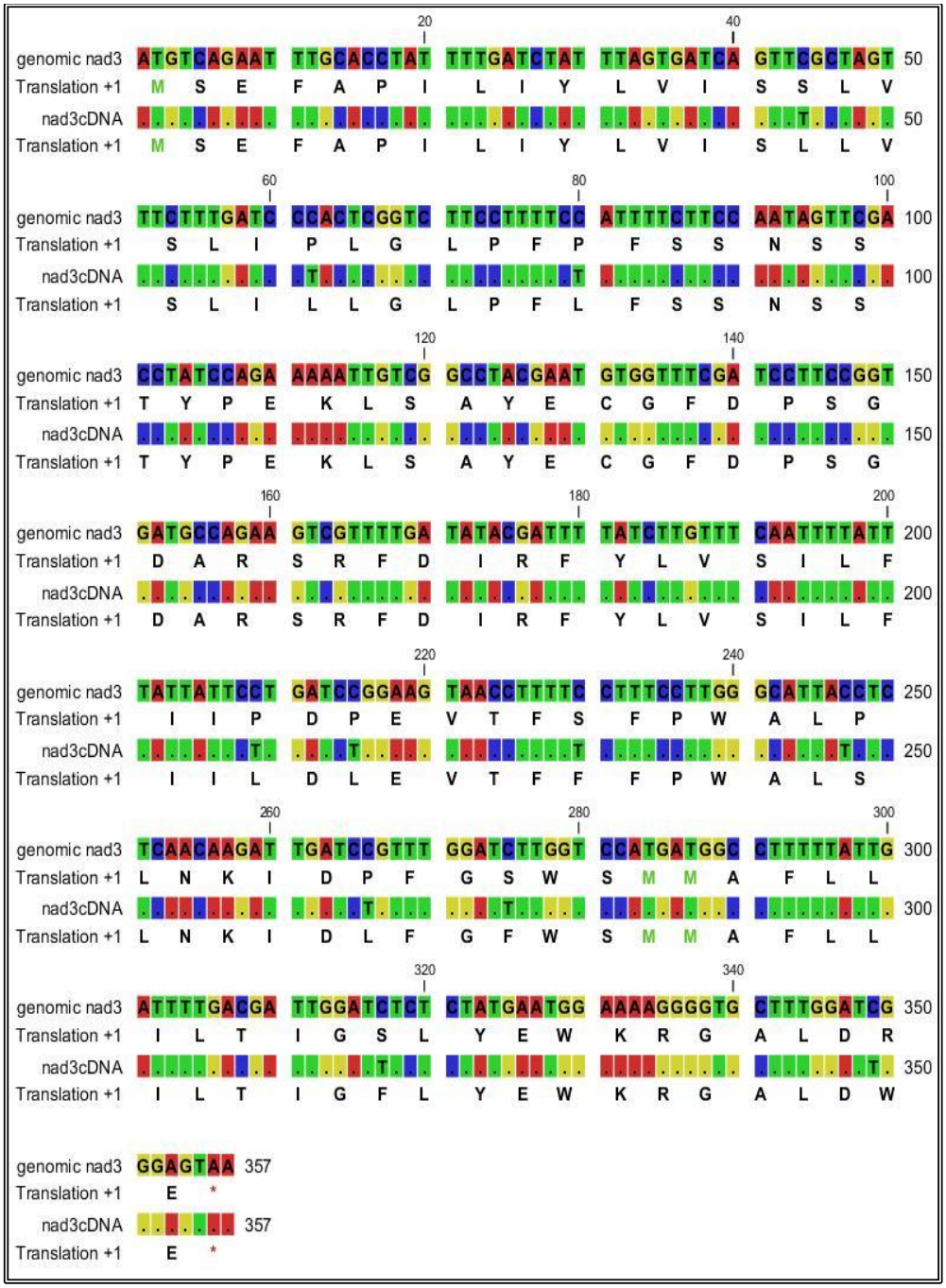

Fig. (2): A comparison between mt.genomic and cDNA sequences of *C. procera nad3*. The corresponding amino acids are given in the second and fourth lines respectively. Dots indicate to similarity between genomic and cDNA sequences *C. procera nad3* gene.

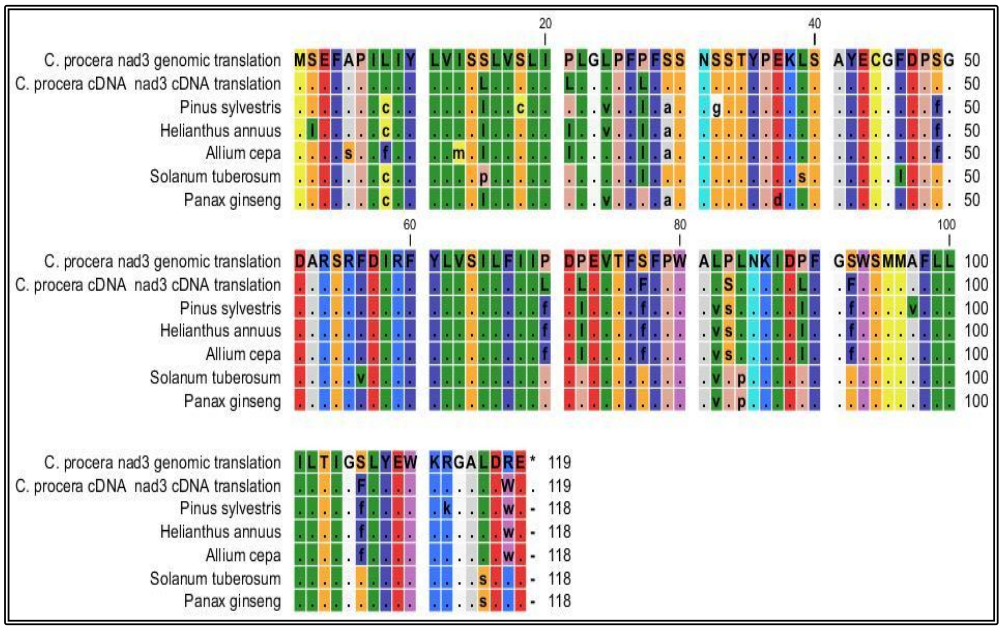

Fig. (3): Comparison of *nad3* protein deduced from mt. genomic and cDNA sequences of *C. procera* with that protein sequences from other species.

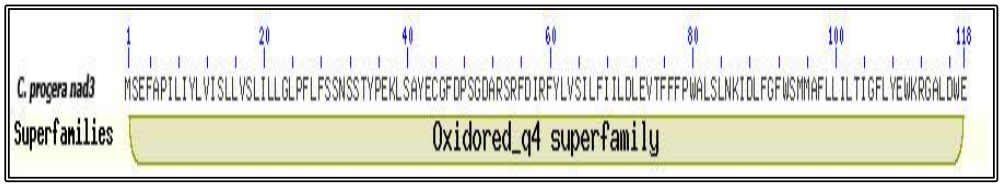

Fig. (4): Protein domains of the deduced amino acid sequence of the obtained *nad3* protein.