Loss of chloroplast *trn*L_{UAA} intron in two species of *Hedysarum* (Fabaceae): evolutionary implications

Atefeh Amirahmadi[1], Shahrokh Kazempour Osaloo[1*], Ali Asghar Maassoumi[2]

[1]Department of Plant Biology, Faculty of Biological Sciences, Tarbiat Modares University, P.O. Box 14115-175, Tehran, I.R. Iran [2]Department of Botany, Research Institute of Forests and Rangelands, P.O. Box 13185-116, Tehran, I.R. Iran

Abstract
Previous studies have indicated that in all land plants examined to date, the chloroplast gene *trn*L_{UAA} is interrupted by a single group I intron ranging from 250 to over 1400 bp. The parasitic *Epifagus virginiana* has lost, however, the entire gene. We report that the intron is missing from the chloroplast genome of two arctic species of the legume genus *Hedysarum* (*H. alpinum*, *H. boreale*). DNA sequencing of the *trn*L gene and *trn*L-*trn*F intergenic spacer (*trn*L-F), as well as portion of *trn*F exon in these species confirms the absence of *trn*L intron and shows that it has been deleted from the gene precisely along established exon/intron splicing sites. Phylogenetic analysis of *trn*L-F sequence data revealed that they are closely related species. This indicates that the intron was lost from the chloroplast genome before the divergence of the two *Hedysarum* species. It is concluded that this rare genomic structural mutation may have occurred once during the evolution of land plants.
***Keywords*:** Chloroplast DNA; Fabaceae; *Hedysarum*; Structural mutation; *trn*L_{UAA} intron loss

INTRODUCTION

Based on the conserved characteristics of the primary sequence and the predicted secondary structure, chloroplast introns are classified as either group I, group II, or group III introns. A basic set of chloroplast introns, consisting of one group I intron and ca 20 group II introns, is a common feature of embryophytes (land plants) and their closest algal relatives-members of the Charophyta (Schmitz-Linneweber and Barkan, 2007). Almost all land plants and charophyte algae have a single group I intron in the *trn*L_{UAA} gene (Shaw *et al*., 2005). This intron is the most ancient intron that is considered to have been present in the common ancestor of cyanobacteria and acquired by plastids via vertical transmission (Asakura and Barkan 2007; Simon *et al*., 2003; Xu *et al*., 1990). The frequent loss of the intron in red algae and some of their secondary plastid derivatives as well as among green algae was already documented (Simon *et al*., 2003). The *trn*L_{UAA} gene is located between the tandemly arranged tRNA genes *trn*T_{UGU} and *trn*F_{GAA} in the large single copy region of the chloroplast genome in land plants (Taberlet *et al*., 2007; Sugiura, 1996; Taberlet *et al*., 1991). The *trn*L intron ranges from ca 250 to over 1400 bp in land plants (Shaw *et al*., 2005). Sequences of this *trn*L intron are usually coamplified with the *trn*L-*trn*F intergenic spacer and together these two fragments (hereafter *trn*L-F) have become the most popular and widely used noncoding cpDNA markers (Kazemi *et al*., 2009; Shaw *et al*., 2005). This intron shows sequence conservation in the regions flanking the *trn*L exons, while the central part is variable (Hao *et al*., 2009; Quandt and Stech, 2005; Bakker *et al*., 2000). It has a specific secondary structure and several highly conserved motifs that are found among all group I introns (Won and Renner, 2005; Shaw *et al*., 2005 and references therein).

Here, we report loss of the intron in two arctic species of the legume genus *Hedysarum*. Prior to this

*Correspondence to: **Shahrokh Kazempour Osaloo,** Ph.D.
Tel: + 98 21 82884404; Fax: +98 21 82884717
E-mail: skosaloo@modares.ac.ir

Table 1. List of the examined species.

Species	Voucher accession	GenBank Accession No.
Astragalus curvipes Trautv.	Iran: Gorgan, Golestan National Park, Maassoumi 47553 (TARI)	AB485938[a]
Alhagi persarum Boiss. & Buhse	Iran: Rudbar, Kazempour Osaloo 2008-1 (TMUPC)	AB558510
Ebenus stellata Boiss.	Iran: Hormozgan, Hajiabad, Ahangarian & Kazempour Osaloo 2006-1 (TMUPC)	AB558511
Hedysarum aculeolatum Boiss.	Morocco: Nador, Segangane, Podlech 122215 (MSB)	AB558512
Hedysarum alpinum L.	USSR: Arctic Siberia, Upianova s.n. (MSB)	AB558513
Hedysarum alpinum L. var. *americanum*	Canada: Yukon Territory, Whitehorse Rapids, Mitchell 135 (MSB)	AB558514
Hedysarum boreale Nutt. var. *boreale*	Canada: Saskatchewan, Prince Albert, Doppelbaur 15788 (MSB)	AB558515
Hedysarum boreale var. *mackenzii* (Richards.) Hitchc.	Canada: Manitoba, Churchill, Doppelbaur s.n. (MSB)	AB558516
Hedysarum hedysaroides (L.) Schinz & Thell.	USSR: Magadan, Bilibinsky, Zaslavskaja 6184 (TARI)	AB558517
Sulla pallida (Desf.) Choi & Ohashi (=*Hedysarum pallidum* Desf.)	Morocco: Oudja, Taourirt, Podlech 97026 (MSB)	AB558518
Onobrychis pulchella Miller	Iran: Khorasan, Kalate-Naderi, Ghahraman *et al*. 27318 (TUH)	AB558519

Abbreviations used in voucher information:
TARI, Herbarium of the Research Institute of Forests and Rangelands, Tehran, Iran; TMUPC, Tarbiat Modares University Plant Collection, Tehran, Iran; TUH, Tehran University Herbarium, Tehran, Iran; MSB, Herbarium of Ludwig-Maximilians-Universität, München, Germany. With the exception of TMUPC, herbarium acronyms are according to Holmgren and Holmgren (2008).
[a]*trn*L-F sequence for *Astragalus curvipes* was retrieved from GenBank.

report, the loss of this intron (actually whole *trn*L gene) in chloroplast of land plants has only been reported in the parasitic plant *Epifagus virginiana* (Wolfe *et al*., 1992).

MATERIALS AND METHODS

Taxon sampling: In the framework of a phylogenetic analysis of the tribe Hedysareae and allies using both nrDNA ITS and *trn*L-F sequences for over 70 taxa (Amirahmadi *et al*. unpubl. data), a subset of eight species including, *Alhagi persarum*, *Ebenus stellata*, *Hedysarum alpinum* (two accessions), *H. boreale* (two accessions), *H. aculeolatum*, *H. hedysaroides*, *Sulla pallida* and *Onobrychis pulchella* were chosen in the present work. *Astragalus curvipes* was included here as a positive control, containing the intron previously reported (Kazemi *et al*., 2009; Table 1).

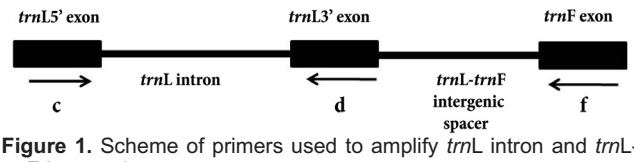

Figure 1. Scheme of primers used to amplify *trn*L intron and *trn*L-*trn*F intergenic spacer.

DNA isolation, amplification and sequencing: Total genomic DNA was isolated from fresh or herbarium materials using the modified cetyl trimethylammonium bromide (CTAB) procedure of Doyle and Doyle (1987). For each genomic DNA, the *trn*L intron (if present) and portions of its flanking exons were ampli-

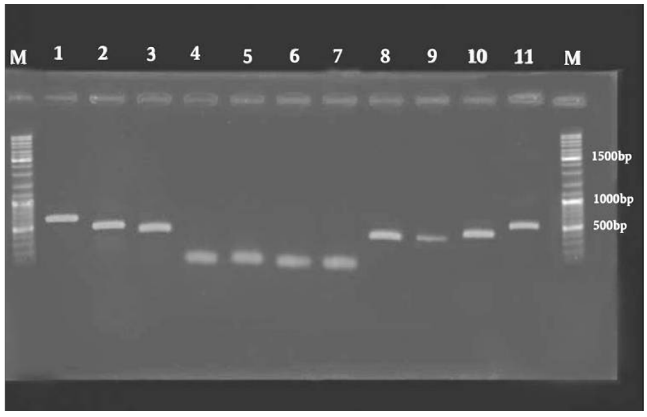

Figure 2. Ethidium bromide stained 1% agarose gel showing PCR amplified products for selected legume taxa screened for presence/absence of the *trn*L $_{UAA}$ intron. Lanes 1-11: (1) *Alhagi persarum*, present (625 bp); (2) *Onobrychis pulchella*, present (491bp); (3) *Ebenus stellata*, present (499 bp); (4,5) *Hedysarum alpinum*, absent (74 bp); (6,7) *Hedysarum boreale*, absent (74 bp); (8) *Hedysarum aculeolatum*, present (423 bp); (9) *Hedysarum hedysaroides*, present (432 bp); (10) *Sulla pallida*, present (428 bp); (11) *Astragalus curvipes*, present (615 bp). M= 100bp ladder.

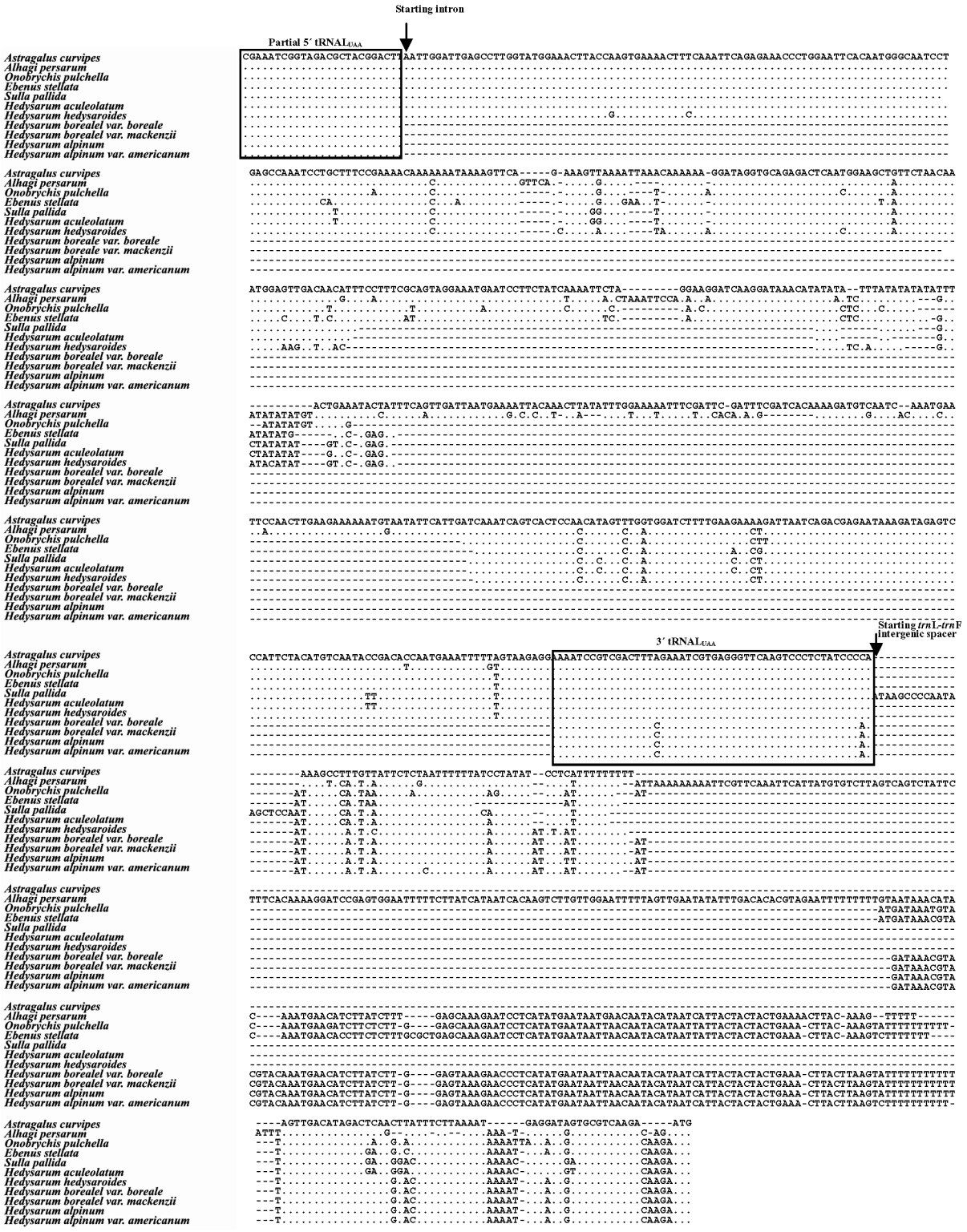

Figure 3. A sequence alignment of the chloroplast *trn*L_{UAA} gene and adjacent *trn*L-*trn*F intergenic spacer. Boxes indicate the *trn*L_{UAA} 5´ and *trn*L_{UAA} 3´ exons, respectively. Gaps (deleted nuclotides) are represented with a dash (-).

fied using the universal "c" and "d" primers of Taberlet *et al*. (1991). In order to sequencing, the whole *trn*L-*trn*F region (*trn*L gene, *trn*L-*trn*F intergenic spacer, and a portion of *trn*F exon) was then amplified using the "c" and "f" primers of Taberlet *et al*. (1991). Position of all primers used for the amplifi-

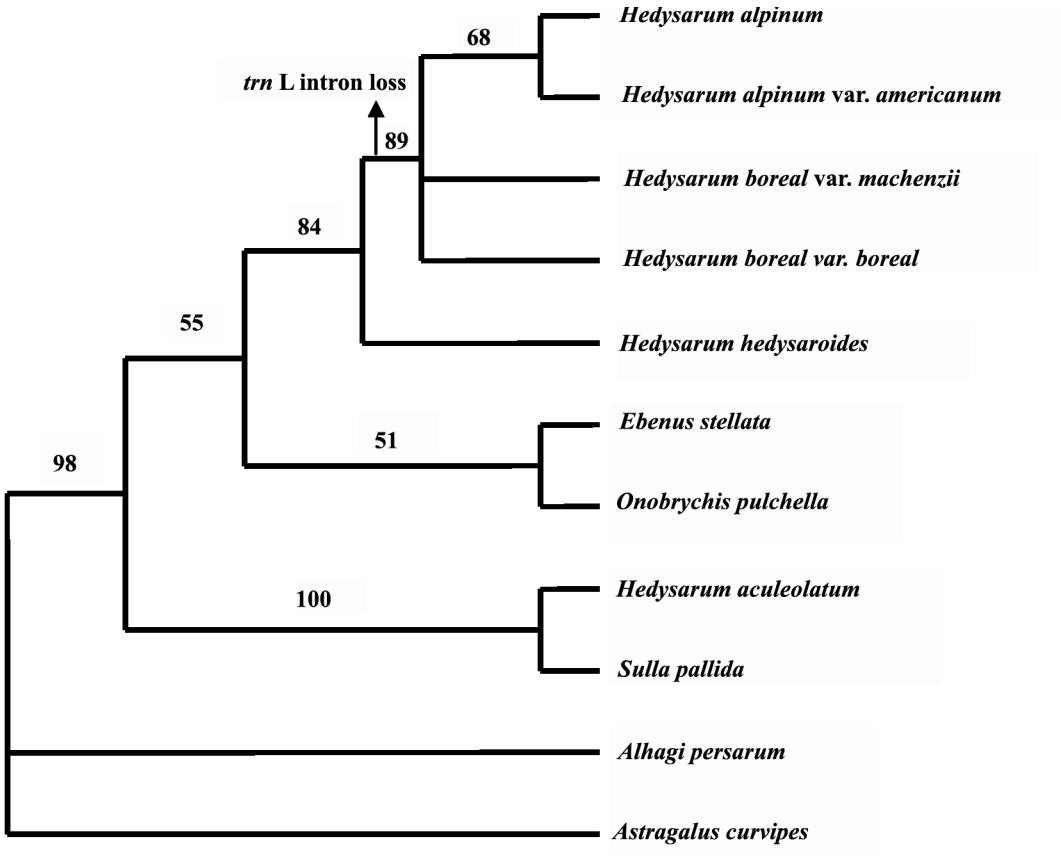

Figure 4. A single most parsimonious tree resulting from phylogenetic anaysis of *trn*L-F sequence data. Length of tree= 56 steps, Consistency index= 0.821 and Retention index= 0.839 (with the exclusion of uninformative characters). Numbers above the branches are bootstrap values.

cation is presented in Figure 1. Polymerase chain reaction (PCR) was carried out in 20 µl final volume of mixture containing 1.0 µl of template DNA (5 ng/µl), 0.5 µl of each primer (5 pmol/µl), 10 µl of the 2X *Taq* DNA polymerase master mix Red (Amplicon, Cat. No. 180301, Germany) and 8.0 µl sterile water. PCR cycles consisted of 31 cycles of 50 s at 94ºC for template denaturation, 40 s at 58ºC for primer annealing, and 55 s at 72ºC for primer extension, followed by 7 min at 72ºC for completion of primer extension. PCR products were separated by electrophoresis in 1% agarose gel stained with ethidium bromide and were photographed with a UVI gel documentation system (UVItec, Cambridge, UK). The *trn*L-F region was then sequenced using the 'Big dye terminator cycle sequencing ready reaction kit' (Applied Biosystems, USA) with the same "c" and "f" primers in an ABI Prism 3730xl DNA Analyzer (Applied Biosystems, USA).

Phylogenetic analysis: *trn*L-F sequences for the abovementioned taxa were edited using BioEdit ver. 7.0.9.0 (Hall, 1999) and aligned using ClustalX (Larkin *et al*., 2007), followed by manual adjustment. Alignment of the dataset required the introduction of numerous single and multiple base indels (insertions/deletions). Positions of indels were treated as missing data. Phylogenetic analysis of the *trn*L-F dataset was performed by the maximum parsimony method using PAUP* program version 4.0b10 (Swofford, 2002), employing the same search strategy as described previously (Tavakkoli *et al*., 2010). *Astragalus curvipes* was chosen as an outgroup.

RESULTS

According to our PCR experiments, the *trn*L_{UAA} intron was inferred to be absent only in four accessions of the two species of *Hedysarum*, *H. alpinum* and *H. boreale* (Fig. 2). The length of *trn*L gene ranged from 74 bp in *H. alpinum* and *H. boreale* to 625 bp in *Alhagi persarum*. For those species having the intron, size variation differed 116 to 192 bp from the control species

(*Astragalus curvipes*), except for *Alhagi persarum*, where it was 10 bp longer than in *A. curvipes*. DNA sequencing of the *trn*L-F region confirmed that the intron was completely missing from *H. alpinum* and *H. boreale*, and that it has led to a junction of 5´*trn*L exon and 3´*trn*L exon to form an uninterrupted gene (Fig. 3).

DISCUSSION

As noted above, DNA sequencing showed that the *trn*L intron was completely lost in both *H. alpinum* and *H. boreale*. Parsimony analysis of *trn*L-F sequences revealed that these two species with four sampled accessions formed a well supported clade which was sister to *H. hedysaroides* (Fig. 4). It is fully consistent with our comprehensive phylogenetic analyses of nrDNA ITS, *trn*L-F and the combined nrDNAITS-*trn*L-F sequence data for Hedysareae and related taxa (Amirahmadi *et al*. unpubl. data). This indicates that the intron loss has occurred in the chloroplast genome before the divergence of the two *Hedysarum* species (Fig. 4). The precise loss of an intron has been reported from other chloroplast genes (e.g., Jansen *et al*., 2008, 2007; Campagna and Downie, 1998; Wallace and Cota, 1996; Downie and Palmer, 1992; Downie *et al*., 1991; Hiratsuka *et al*., 1989). The legumes are one of a few angiosperm families that have experienced multiple losses of introns such as *rpl*2, *rps*12 and *clp*P (Jansen *et al*., 2008 and references therein). The loss of the *trn*L intron in the two species of *Hedysarum* is another exemplar that might be *per se* a rare evolutionary event among photosynthetic land plants.

Earlier studies (Simon *et al*., 2003; Xu *et al*., 1990) hypothesized that *trn*L intron in the chloroplast of the land plants has completely lost the self-splicing ability and thus has become dependent on a host factor to facilitate excision. Furthermore, Asakura and Barkan (2007) revealed that Crm Family Member 2 (CFM2), a nucleus-encoded protein, is bound to and enhances the splicing of the *trn*L intron. Nakamura *et al*. (1999) reported, however, that chloroplast ribonucleoproteins (cpRNPs) interact with unspliced $trnL_{UAA}$ RNA from tobacco chloroplast extract. One possible mechanism for the loss of *trn*L intron might be reverse transcription of the spliced RNA, which does not contain intron, followed by homologous recombination of the cDNA with the chloroplast genome at precisely the same position in the common ancestor of the two *Hedysarum* species. Such process has been already reported for the loss of both group I and group II introns in plants, fungi and animals (e.g., Jeffares *et al*., 2006; Hu 2006; Campagna and Downie, 1998; Downie *et al*., 1991; Hiratsuka *et al*., 1989; Dujon, 1989). Other mechanisms of intron loss could be simple genomic deletion and in-frame intron deletion (Niu *et al*., 2005 and references therein).

Acknowledgements

We thank the staff of the Herbarium of Ludwig-Maximilians-Universität, München, Germany, for providing the leaf materials of *Hedysarum* and *Sulla*. Support was provided by a research grant from Tarbiat Modares University. This work represents partial fulfillment of the requirements from the University for obtaining MSc degree by the first author.

References

Asakura Y, Barkan A (2007). A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *Plant Cell*. 19: 3864-3875.

Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, Compton J, Dawtrea R, Gibby M (2000). Patterns of nucleotide substitutions in angiosperm cpDNA *trn*L (UAA)-*trn*F (GAA) regions. *Mol Biol Evol*. 17: 1146-1155.

Campagna ML, Downie SR (1998). The Intron in chloroplast gene *rpl*16 is missing from the flowering plant families Geraniaceae, Goodeniaceae, and Plumbaginaceae. *Transact Illinois State Acad Sci*. 91: 1-11.

Dujon B (1989). Group I introns as mobile genetic elements: Facts and mechanistic speculations-a review. *Gene*. 82: 91-114.

Downie SR, Olmstead RG, Zurawski G, Soltis DE, Soltis PS, Watson JC, Palmer JD (1991). Six independent losses of the chloroplast DNA *rpl*2 intron in dicotyledons: molecular and phylogenetic implications. *Evolution*. 45: 1245-1259.

Downie SR, Palmer JD (1992). Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Molecular Systematics of Plants. Soltis, Soltis and Doyle ed., Chapman and Hall, New York. PP. 14-35.

Doyle JJ, Doyle JJ (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19: 11-15.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*. 41: 95-98.

Hao DC, Huang BL, Chen SL, Mu J (2009). Evolution of the Chloroplast *trn*L-*trn*F region in the Gymnosperm lineages Taxaceae and Cephalotaxaceae. *Biochem Genet*. 47: 351-369.

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distant tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol Gen Genet*. 217: 185-194.

Hu K (2006). Intron exclusion and the mystery of intron loss. *FEBS Lett*. 580: 6361-6365.

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal J, Kuehl JV, Boore JL (2007). Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci USA*. 104: 19369-19374.

Jansen RK, Wojciechowski MF, Sanniyasi E, Lee SB, Daniell H (2008). Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps*12 and *clp*P intron losses among legumes (Leguminosae). *Mol Phylogenet Evol*. 48:1204-1217.

Jeffares DC, Mourier T, Penny D (2006). The biology of intron gain and loss. *Trends Genet*. 22: 16-22.

Kazemi M, Kazempour Osaloo SH, Maassoumi AA, Rastegar Pouyani E (2009). Molecular Phylogeny of selected Old World *Astragalus* (Fabaceae): Incongruence among chloroplast *trn*L-F, *ndh*F and nuclear ribosomal DNA ITS sequences. *Nord J Bot*. 27: 425-436.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

Nakamura T, Ohta M, Sugiura M, Sugita M (1999). Chloroplast ribonucleoproteins are associated with both mRNAs and intron-containing precursor tRNAs. *FEBS Lett*. 460: 437-441.

Niu DK, Hou WR, Li SW (2005). mRNA-mediated intron losses: Evidence from extraordinarily large exons. *Mol Biol Evol*. 22: 1475-1481.

Quandt D, Stech M (2005). Molecular evolution of the *trn*L UAA intron in bryophytes. *Mol Phylogenet Evol*. 36: 429-443.

Schmitz-Linneweber S, Barkan A (2007). RNA splicing and RNA editing in Chloroplasts. In: Cell and Molecular Biology of Plastids, *Currt Topics Genet*, Vol.19, Bock ed., Springer, Heidelberg. PP. 213-248.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Ame J Bot*. 92: 142-166.

Simon D, Fewer D, Friedl T, Bhattacharya D (2003). Phylogeny and self splicing ability of the plastid tRNA-Leu group I intron. *J Mol Evol*. 57: 710-720.

Sugiura M. (1996). Structure and replication of chloroplast DNA. In: Molecular Genetics of Photosynthesis, Andersson, Hugh Salter and Barber ed., IRL Press, Oxford. PP. 58-74.

Swofford DL (2002). Phylogenetic analysis using parsimony (*and other methods). Version. 4. Sinauer Associates, Sunderland, Massachusetts, USA.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991). Universal primers for amplification of three non coding regions of chloroplast DNA. *Plant Mol Biol*. 17: 1105-1109.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, VermatT, Corthier G, Brochmann C, Willerslev E (2007). Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. *Nucl Acids Res*. 35: e14.

Tavakkoli S, Kazempour Osaloo S, Maassoumi AA (2010). The phylogeny of *Calligonum* and *Pteropyrum* (Polygonaceae) based on nuclear ribosomal DNA ITS and chloroplast *trn*L-F sequences. *Iran J Biotech*. 8: 7-15.

Wallace RS, Cota JH (1996). An intron loss in the chloroplast gene *rpo*C1 supports a monophyletic origin for the subfamily Cactoideae of the Cactaceae. *Curr Genet*. 29: 275-281.

Wolfe KH, Modern CW, Palmer JD (1992). Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci USA*. 89: 10648-10652.

Won H, Runner SS (2005). The Chloroplast *trn*L-*trn*F Region in the Seed Plant Lineage Gnetales. *J Mol Evol*, 61:425-436.

Xu MQ, Kaathe S, Goodrich Blair H, Nierzwicki Bauer S, Shub D (1990). Bacterial origin of a chloroplast intron: Conserved self-splicing group I introns in cyanobacteria. *Science* 250: 1566-1569.