<https://helda.helsinki.fi>

Identification and characterization of plastid trnF((GAA)) pseudogenes in four species of Solanum (Solanaceae)

Poczai, Peter

2011

Poczai , P & Hyvönen , J 2011 , ' Identification and characterization of plastid trnF((GAA))
pseudogenes in four species of Solanum (Solanaceae) ' , Biotechnology Letters , vol. 33 ,
no. 11 , pp. 2317-2323 . <https://doi.org/10.1007/s10529-011-0701-x>

<http://hdl.handle.net/10138/42435>

<https://doi.org/10.1007/s10529-011-0701-x>

submittedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Identification and characterization of plastid *trnF*_(GAA) pseudogenes in four species of *Solanum* (Solanaceae)

Péter Poczai · Jaakko Hyvönen

Received: 28 June 2011 / Accepted: 4 July 2011 / Published online: 16 July 2011
© Springer Science+Business Media B.V. 2011

Abstract Non-functional *trnF* pseudogenes that rarely occur in embryophytes have been found in Solanaceae. We have sequenced the *trnL*-F intergenic spacer of four species of *Solanum*, and found duplicated regions of the original *trnF* gene. These repeats were 94–260 bp long causing large length variation in the *trnL*-F intergenic spacer resulting from differences in pseudogene copy number (2–4). The duplicated *trnF* regions are comprised of several highly structured motifs, which were partial residues, or entire parts of the Anticodon, T- and D-domains of the original gene, but all lacked the acceptor stems at the 5'- or 3'-end. Pseudogenes included several transitions and transversions in their sequences compared to the original *trnF* gene. Among pseudogene copies, T-domains were more frequent and fragmented than D-domain elements. Our results demonstrate that although chloroplast evolution is uniform such structural duplications in the sequences used for phylogenetic reconstructions should be treated with great caution.

Keywords Chloroplast DNA · Gene duplication · Pseudogenes · Structural mutations · *TrnL*-F intergenic spacer

Introduction

The genus *Solanum* L. (Solanaceae) contains about 1,400 species and is one of the largest genera of flowering plants (Bohs 2005; Stern et al. 2010) with many economically important crops such as potato, tomato and eggplant. Recent phylogenetic studies have identified the major groups in the genus (Weese and Bohs 2007) and now these results provide a solid framework for investigating processes of molecular evolution, and provide species level taxonomic keys (Knapp et al. 2004; <http://www.solanaceaesource.org>). The plastid *trnT*-F region is widely applied to resolve deep level phylogeny of land plants (Quandt and Stech 2004; Gobert et al. 2006; Zhao et al. 2011). It has also been used in studies of *Solanum* (Stern et al. 2010; Tepe and Bohs 2010). The *trnT*-F region is located in the large single copy region of the chloroplast genome and it contains three highly conserved transfer RNA genes, namely the tRNA genes for threonine (UGU), leucine (UAA) and phenylalanine (GAA). These tRNA genes are separated by two intergenic spacers and the *trnL* intron intercalated within the first and second exon of the *trnL*_(UAA) gene. The *trnL*-F intergenic spacer

P. Poczai (✉) · J. Hyvönen
Department of Plant Biology, University of Helsinki,
PO Box 65, 00014 Helsinki, Finland
e-mail: peter.poczai@gmail.com

J. Hyvönen
e-mail: jaakko.hyvonen@helsinki.fi

P. Poczai
Department of Plant Science and Biotechnology,
Georgikon Faculty, University of Pannonia, Festetics 7,
8360 Keszthely, Hungary

separating the second exon of the *trnL*_(UAA) gene and the exon of the *trnF*_(GAA) gene (Koch et al. 2007) has remarkably high length variation in angiosperms (Borsch et al. 2003). Such microstructural changes can provide phylogenetically informative characters, if properly treated (e.g. Borsch et al. 2003); otherwise conflicting phylogenies could be inferred (e.g. Vijverberg and Bachmann 1999). These inconsistencies may be due to specific recombination mechanisms which are accompanied by high mutational rates and trigger the parallel generation of particular molecular changes, like the integration of tandem repeats or the formation of pseudogenic copies (Dobeš et al. 2007). The occurrence of plastid *trnF* pseudogenes in seed plants is extremely rare (Schmickl et al. 2009). They have been only described from a few plant genera in Asteraceae (Vijverberg and Bachmann 1999), Annonaceae (Pirie et al. 2007), Brassicaceae (Koch et al. 2007) and Juncaceae (Drábková et al. 2004), but they have not been reported in Solanaceae.

As part of our ongoing research we utilize plastid *trnT*-*F* sequences to study phylogeny and evolution of different clades and species groups of the genus *Solanum*. We report here the occurrence and sequence motif characterization of *trnF* pseudogenes found in the *trnL*-*F* intergenic spacer (IGS) region of four *Solanum* species.

Materials and methods

Sampling, PCR amplification and sequencing

We used one herbarium and two living accessions of *Solanum physalifolium* Rusby, and *Solanum sisymbriifolium* Lam., while only one herbarium sample was used from *Solanum montanum* L. and *Solanum trisectum* Dunal. These taxa represent a smaller portion of dataset used for phylogenetic reconstruction in the genus *Solanum*. All specimens are included here where suspicious pseudogenic repeats have been found. Voucher specimens were deposited previously in the Georgikon Herbarium of the University of Pannonia. Further details are given in Table 1. DNA was purified from 10 to 25 mg dried herbarium, or fresh plant leaves with the DNeasy Plant Mini Kit (Qiagen, Germany). The *trnL*-*F* region was PCR amplified with the *E* forward primer (5'-GGT TCA AGT CCC TCT ATCC C-3') described by Taberlet et al. (1991), and

with a modified reverse primer *F'* (5'-TCC AGG AAC CAG ATT TGAA C-3') designed to anneal to the end of the 3'-acceptor stem of the *trnF*_(GAA) exon. Amplification reactions were performed in 50 µl containing: 25 µl nuclease free water, approx. 20 ng template DNA, 50 µg bovine serum albumin BSA, 0.5 µM each primer, 0.2 mM dNTP, 5 µl 10× PCR buffer (1 mM Tris-HCl, pH 8.8 at 25°C, 1.5 mM MgCl₂, 50 mM KCl and 0.1% Triton X-100) and 0.5 U of Phusion High-Fidelity DNA Polymerase (Finnzymes, Finland). All reactions were performed in a MasterCycler ep96 (Eppendorf, Germany): 2 min at 94°C for initial denaturation, 35 cycles of 30 s denaturation at 94°C, 1 min annealing at 50°C, and 2 min extension at 72°C, followed by a final extension for 5 min at 72°C. Amplification products were separated on 1.5% agarose gels in 0.5× TBE buffer (220 V, 0.5 h) and stained with ethidium-bromide. PCR products were purified with the QIAquick PCR purification kit (Qiagen) according to the protocol of the manufacturer. Sequencing was performed on an ABI 3730XL automated sequencer in both directions using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit v.3.0. All sequences were annotated and deposited in NCBI GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>); accession numbers are available in Table 1.

*TrnF*_(GAA) pseudogene recognition and copy number assessment

Putative pseudogene repeats were identified with screening using Repbase (<http://www.girinst.org/>) with the “mask pseudogenes” and “report simple repeats” options of the online tool CENSOR (Kohany et al. 2006). This was done to identify repetitive elements by comparing our sequences to known eukaryotic repeats and prototypic sequences stored in Repbase utilizing WU-BLAST. A second search was conducted with FastPCR (Kalendar et al. 2009) using the repeat search option of the program. Under “type of repeats” we checked for simple, direct, inverted, direct antisense, and direct reverse repeats, respectively. Default values were used under a kMers repeat screening. After each search, repetitive motifs and sequences were recorded and compared with results obtained from the Repbase search. After repeats were identified in the *trnL*-*F* IGS sequences, further structural *trnF*_(GAA) gene elements (see Fig. 1) or

Table 1 Information on accessions and herbarium specimens used in this study

Taxa	Ψ copy number	Major clade within <i>Solanum</i> ^a	Collection locality	Herbarium code	Voucher	GenBank accession number
<i>Solanum montanum</i>	2	Regmandra	Peru	H1620485	P. Poczai (H) 2011	JN130369
<i>Solanum physalifolium</i>	4	Morelloid	Spain	H1701364	X. Giraldez 1986	JN130371
		Morelloid	Germany	HG09015	P. Poczai (H) 2011	JN130372
		Morelloid	Hungary	HG09016	P. Poczai (H) 2011	JN130373
<i>Solanum sisymbriifolium</i>	4	Leptostemonum	Argentina	H1548976	P. Poczai (H) 2011	JN130366
		Leptostemonum	Cultivated	HG10012	P. Poczai (H) 2011	JN130367
		Leptostemonum	Cultivated	HG10013	P. Poczai (H) 2011	JN130368
<i>Solanum trisetum</i>	2	Normania	France	HG10011	P. Poczai (H) 2011	JN130370

^a Major clades are indexed according to Weese and Bohs (2007) currently utilized by the PBI *Solanum* worldwide treatment (<http://www.solanaceaesource.org>)

Fig. 1 Nucleotide sequence of the *Solanum trnF*_(GAA) gene. Structural elements are represented by boxes, while acceptor stems with connecting line at the 5'- and 3'-end. Secondary

structure is indicated by arrow symbols a–d and a'–d', while the exact structure edited and visualized by 4SALE (Seibel et al. 2006) is shown under the sequence panel

residues were annotated via manual or ClustalW (Larkin et al. 2007) alignments with BIOEDIT (Hall 1999). For all annotations the complete chloroplast genome of *Solanum bulbocastanum* Dunal (DQ347958) was used. The annotated sequence alignment is shown in Fig. 2.

Results and discussion

We characterized eight multiple sequences from four *Solanum* species belonging to different clades in the infrageneric phylogenetic scheme of the genus (Bohs 2005; Weese and Bohs 2007). The length of the *trnL*-F

<i>trnF</i> _(GAA) gene		<Acc. stem> <----D-domain-----> <-----Anticodon----->	<-----T-domain-----><Acc. stem>
		GCCGGGATA GCTCAGTTGGTAGAGCA GAGGACTGAAAATCCTCGTGTC	ACCAGTTCAAATCTGGT TTCTGGCA
5'copies			
<i>S. physalifolium</i>	Ψ ₁	----- -AGGACTGAAAATCCTCATGTT-----	AGCGGTTCCAA-----
<i>S. montanum</i>	Ψ ₁	----- -AGGACTGAAAACCTCATGTT-----	AGCGGTTCCAA-----
<i>S. trisectum</i>	Ψ ₁	----- -AGAAGTAAAATCTCATGTT-----	AGCGGTTCCAA-----
<i>S. sisymbriifolium</i>	Ψ ₁	----- -AGGACTCGAAAATCCTCATGTT-----	AACGGTTCCAAT TTCCAAT-----
3'copies			
<i>S. physalifolium</i>	Ψ ₄	P ₁ ---P ₂ ---CAGTTGGTAGAGCA GAGGACTGAAAATCCTCATGTT-----	AGCGGTTCCAA-----
<i>S. montanum</i>	Ψ ₂	P ₁ ---P ₃ ---GTGGTAGAGTA GAGGACTGTAAATCCTCGTGCC-----	ACCTGTTCCAATCTGGA-----
<i>S. trisectum</i>	Ψ ₂	P ₁ ---P ₃ ---GTGGTAGAGTA GAGGACTGTAAATCCTCGTGCC-----	ACCTGTTCCAATCTGGA-----
<i>S. sisymbriifolium</i>	Ψ ₄	----- -GATTGGTAGGTA GAGGACTGAAAATCCTC-----	AG-----CCAATCTAAT --Bf ₁ ---A _t
Middle copies			
<i>S. physalifolium</i>	Ψ _{2,3}	P ₁ ---P ₂ ---CAGTTGGTAGAGCA GAGGACTGAAAATCCTCATGTT CATGTT	AACGGTTCCAA
<i>S. sisymbriifolium</i>	Ψ ₂	---CCAGATTGGTAGGTA GAGGACTGAAAATCCTCATGTT	AACGGTTCCAATCTAAT --Bf ₂ ----
<i>S. sisymbriifolium</i>	Ψ ₃	----- -GCTCGATTGGTAGGTA GAGGACTGAAAATCCTC--GTT	--CCAATCA-----

Fig. 2 Manual alignment of *Solanum* pseudogene (Ψ_n) copies demonstrating polymorphisms and structural rearrangements. Sequence repeats are represented as boxes, while shaded bases in the alignment positions show nucleotide polymorphisms observed across taxa or within each copy type. At the top of the alignment the consensus and conserved sequence of the *Solanum trnF*_(GAA) is shown, with annotating lines and arrows

intergenic spacers varied between 420 and 624 bp. Duplicated elements were found in the 5'-region upstream of the *trnF*_(GAA) (Fig. 3). These duplicated sequences have been classified as 5'-copy, 3'-copy and middle copy pseudogenes (Ψ), based on their positions in the IGS region, and shared synteny. We identified four pseudogenic repeats in *S. physalifolium* and *S. sisymbriifolium*, and two copies in *S. montanum* and *S. trisectum*. The *trnF*_(GAA) pseudogenes varied in length and structure in different species (Fig. 4). All repeats contained functional elements from the *trnF*_(GAA) exon, but their structure, sequence and distribution was variable in each taxon. The 5'-copies showed high structural similarities. They contained only the sequence of the anticodon and a fragmented T-domain, but completely lacked other gene elements. Both the anticodon and the partial T-domain contained few substitutions

indicating structural elements of the phenylalanine transfer RNA gene. Letters in the alignment are unaligned putative promoters abbreviated with P₁ (TCGAGA), P₂ (TTGGGA-ATT) and P₃ (TTGGG), respectively, while Bf₁ (CTGGGT), Bf₂ (CTGGGTTG) are bordering motifs and A_t (TGGAAATC) represents a fragment of an additional Anticodon loop sequence

compared to the *trnF*_(GAA) gene (see Fig. 2). Transitions were more common in the Anticodon region, then in the T-domain fragment of 5' copy pseudogenes. Overall four transitions (C → T; G → A) and two transversions (T → G; G → C) were found in the Anticodon; from these two transitions were commonly preserved in all copies. In the partial T-domain only three transversions can be observed (C → G; C → A). Furthermore, we observed an additional inserted seven base pair long repeat (TTCCAAT) in the partial T-domain sequence of *S. sisymbriifolium*. The 3'-copies had 13 polymorphic sites in their sequence, from these three transitions (G ↔ A; C → T) occurred in the D-domain, while four polymorphisms were found to be transitions in the Anticodon domain (G ↔ A; C ↔ T), except in *S. montanum* and *S. trisectum* where an A → G transition was replaced by an A → T transversion.

Fig. 3 Schematic representation of the *trnL*-F intergenic spacer region, showing pseudogene repeats

Fig. 4 Shared syntenic presentation of pseudogene copies in four different *Solanum* species. Structural elements are indicated by coloured boxes, while black lines indicate bordering motifs. Arrow signs depict sequence repeats. Colour

code for boxes follows: red putative promoters, grey D-domain, blue Anticodon domain, purple T-domain, orange acceptor stem

The T-domain of 3'-pseudogene copy had a partial five base pair deletion (CAGTT) in *S. sisymbriifolium* and had three transversions (C → G; T ↔ A) and two transitions (G ↔ A) in its sequence. Middle copies were the most variable regarding structural fragmentation and single nucleotide polymorphisms (SNPs), since we observed 16 SNPs in these copies and a six base pair repeat at the 3' end of *S. physalifolium* Anticodon regions. Twelve SNPs were found to be transitions (T → C; G ↔ A), while four were transversions (C ↔ A; G ↔ C). Detailed distribution of SNPs is shown in Fig. 2.

Of particular interest is that in all pseudogenes both acceptor stems were missing but anticodon domains were relatively conserved and found in highest number across all copies. Instead of acceptor stems, bordering motifs were found in some pseudogenes. Some of these were identified as putative promoter regions in *S. physalifolium*, *S. montanum* and *S. trisetum*. These putative promoters, defined here as region P, were grouped into three different types as P₁ (TCGAGA), P₂

(TTGGGAATT) and P₃ (TTGGG), respectively. Region P₁ showed high similarity to a putative sigma70-type bacterial promoter motif (−35 TTGACA; Koch et al. 2007), while P₂ and P₃ to other bacterial promoter elements (−10 GAGAT/−35 TGGGA; Azevedo et al. 1993; Poirel and Nordmann 2006). Such promoter motifs have also been found together with pseudogenes, or separately in the *trnL*-F IGS region (Quandt et al. 2004). In a comprehensive study of embryophytes, the −35 TTGACA motif has been found consistently; it has been speculated that it represents the original promoter of the *trnF*_(GAA) gene (Quandt et al. 2004). Previously, Kanno and Harai (1993) concluded that the *trnF*_(GAA) is co-transcribed with the *trnL*_(UAA) gene, indicating that our promoter like motifs should be non-functional. These findings are in accordance with previous conclusions reported by Koch et al. (2007) for *Arabidopsis* Heynh. in Holl and Heynh. (Brassicaceae). They reported that such pseudogenic promoter motifs are hardly functional, since all promoter like elements were inserted between

*trnL*_(UAA) and *trnF*_(GAA) genes. In *S. sisymbriifolium* two other elements (CTGGGT/CTGGGTTG) were observed bordering the 3'-end of Ψ copies. While also in this taxa, Ψ_4 was followed by a highly fragmented Anticodon part, merely containing the sequence of just the Anticodon loop.

In *Solanum* pseudogenes, T-domains were more common than D-domains which seem to be lost more easily. However, T-domain sequences were more fragmented and variable than D-domains where the structure is obviously relatively conserved. Considerably more nucleotide polymorphism was accumulated in middle copy pseudogenes compared to 5'- or 3'-copies. Koch et al. (2007) reported similar features of *Arabidopsis* pseudogenes, where the mutation rate for some copies exceeded even the rate calculated for the entire *trnL*-intron-*trnL*-F IGS region. This might be due to the fact that recombination mediated intraspecific variation is usually more active in middle copies, which participate more actively in complex molecular mechanisms leading to structural variation than the 5'- or 3'-copies which are not being prone to such recombination. Recombination is not thought to play a major role in the evolution of plastid genomes, due to predominance of uniparental (maternal or even paternal) inheritance and haploidy (Birky 2001). Therefore, the overall structure of the chloroplast genome, at least in embryophytes, is considered to be generally conserved (Ansell et al. 2007), and interspersed by short structural changes (<10 bp) that are frequently detected. However, it seems that cpDNA pseudogenes behave differently since all the reported pseudogenes from different plant families were structurally variable (Vijverberg and Bachmann 1999; Drábková et al. 2004; Koch et al. 2007). This leads to the assumption of complex or at least different molecular read mechanism driven development and evolution. It has been previously concluded that tandemly repeated *trnF*_(GAA) pseudogenes are undergoing parallel independent changes which are mediated by complex processes of illegitimate recombination (see Dobeš et al. 2007). Ansell et al. (2007) studying *trnF*_(GAA) pseudogenes in *Arabidopsis lyrata* (L.) O'Kane and Al-Shehbaz populations found that disperse repeats which share similar secondary structure patterns interact and facilitate reciprocal exchange of structural motifs between copies through intra- and intermolecular recombinations. These in the end form chimeric sequences and result in iterative expansion/contraction

in pseudogene copy numbers (Ansell et al. 2007). Such recombination processes might also be responsible for structural and sequence variability found in *Solanum* pseudogenes. The widely accepted assumption that chloroplast sequence evolution is simple and structural changes are informative is violated in the case of pseudogenes; we suggest handling such data in phylogenetic reconstruction with great caution. Inferring phylogenies from characters where nucleotide substitutions have no longer unique histories can mislead phylogenetic reconstructions when base to base homology is assumed. Until the evolutionary mechanisms of *Solanum* pseudogene diversification are known more in detail we suggest using approaches like fixed state optimization (Wheeler 1999), or some other similar form of coding, where only fragment based homology is assumed a priori, and these assumptions are not made at the level of individual nucleotides. Equally important is to analyze numerous sources of data simultaneously (Kluge 1989; Nixon and Carpenter 1996). Only this will show congruence among different gene-regions and reveal potentially homoplastic and phylogenetically misleading changes.

Acknowledgments Thanks are due to the Botanical Museum (H) of University of Helsinki, Finland for providing leaf segments for our analysis. This study was supported by a Hungarian State Eötvös Research Grant (MÖB/90-1/2011) provided for P. P.

References

- Ansell SW, Schneider H, Pedersen N, Grundmann M, Russell SJ, Vogel J (2007) Recombination diversifies chloroplast *trnF* pseudogenes in *Arabidopsis lyrata*. *J Evol Biol* 20:2400–2411
- Azevedo V, Sorokin A, Ehrlich D, Serrón P (1993) The transcriptional organization of the *Bacillus subtilis* 168 chromosome region between the *spoVAF* and *serA* genetic loci. *Mol Microbiol* 10:397–405
- Birky CW (2001) The inheritance of genes in mitochondria and chloroplasts, mechanisms and models. *Annu Rev Genet* 35:125–148
- Bohs L (2005) Major clades in *Solanum* based on *ndhF* sequence data. In: Keating RC, Hollowell VC, Croat TB (eds) A festschrift for William G. D'Arcy: the legacy of a taxonomist. Missouri Botanical Garden Press, St. Louis, pp 27–49
- Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C, Barthlott W (2003) Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J Evol Biol* 16:558–576

- Dobeš C, Kiefer C, Kiefer M, Koch MA (2007) Plastidic *trnF_{UUC}* pseudogenes in North American genus *Boechera* (Brassicaceae): mechanistic aspects of evolution. *Plant Biol* 9:502–515
- Drábková L, Kirschner J, Vlček Č, Paček V (2004) *TrnL-trnF* intergenic spacer and *trnL* intron define major clades within *Luzula* and *Juncus* (Juncaceae): importance of structural mutations. *J Mol Evol* 59:1–10
- Gobert V, Moja S, Taberlet P, Wink M (2006) Heterogeneity of three molecular data partition phylogenies of mints related to *M. × piperita* (*Mentha*; Lamiaceae). *Plant Biol* 8:470–485
- 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
- Kalendar R, Lee D, Schulman AH (2009) FastPCR software for PCR primer and probe design and repeat search. *Genes, Genomes, Genomics* 3:1–14
- Kanno A, Harai A (1993) A transcription map of the chloroplast genome from rice (*Oryza sativa*). *Curr Genet* 23:166–174
- Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst Zool* 38:7–25
- Knapp S, Bohs L, Nee M, Spooner DM (2004) Solanaceae—a model for linking genomics with biodiversity. *Comp Funct Genomics* 5:285–291
- Koch MA, Dobeš C, Kiefer C, Schmickl R, Klimeš L, Lysak MA (2007) Supernetwork identifies multiple events of plastid *trnF_(GAA)* pseudogene evolution in the Brassicaceae. *Mol Biol Evol* 24:63–73
- Kohany O, Gentles AJ, Hankus L, Jurka I (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* 7:474
- 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
- Nixon KC, Carpenter JM (1996) On simultaneous analysis. *Cladistics* 12:221–242
- Pirie MD, Vargas MPB, Botermans M, Bakker FT, Chatrou LW (2007) Ancient paralogy in the cpDNA *trnL-F* region in Annonaceae: implications for plant molecular systematics. *Am J Bot* 94:1003–1016
- Poirol L, Nordmann P (2006) Genetic structures at the origin of acquisition and expression of the carbapenem-hydrolyzing oxacillinase gene *bla_{OXA-58}* in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 50:1442–1448
- Quandt D, Stech M (2004) Molecular evolution of the *trnT_{UGU}-trnF_{GAA}* region in bryophytes. *Plant Biol* 6:545–554
- Quandt D, Müller K, Stech M, Frahm JP, Hilu KW, Borsch T (2004) Molecular evolution of the chloroplast *trnL-F* region in land plants. *Monogr Syst Bot Mo Bot Gard* 98:13–37
- Schmickl R, Keifer C, Dobeš C, Koch MA (2009) Evolution of *trnF(GAA)* pseudogenes in cruciferous plants. *Plant Syst Evol* 282:229–240
- Seibel PN, Müller T, Dandekar T, Schultz J, Wolf M (2006) 4SALE—a tool for synchronous RNA sequence and secondary structure alignment and editing. *BMC Bioinformatics* 7:498
- Stern SR, Weese T, Bohs L (2010) Phylogenetic relationships in *Solanum* section *Androceras* (Solanaceae). *Syst Bot* 35:885–893
- 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
- Tepe EJ, Bohs L (2010) A molecular phylogeny of *Solanum* sect. *Pterioidea* (Solanaceae) and the utility of COSII markers in resolving relationships among closely related species. *Taxon* 59:733–743
- Vijverberg K, Bachmann K (1999) Molecular evolution of tandemly repeated *trnF(GAA)* gene in the chloroplast genomes of *Microseris* (Asteraceae) and the use of structural mutations in phylogenetic analysis. *Mol Biol Evol* 16:1329–1340
- Weese TL, Bohs L (2007) A three gene phylogeny of the genus *Solanum* (Solanaceae). *Syst Bot* 32:445–463
- Wheeler WC (1999) Fixed character states and the optimization of molecular sequence data. *Cladistics* 15:379–385
- Zhao T, Wang Z-T, Branford-White CJ, Xu H, Wang C-H (2011) Classification and differentiation of the genus *Peganum* indigenous to China based on chloroplast *trnL-F* and *psbA-trnH* sequences and seed coat morphology. *Plant Biol*. doi:10.1111/j.1438-8677.2011.00455