Curr Genet (1993) 24:171–176

Current Genetics

A large deletion in the plastid DNA of the holoparasitic flowering plant *Cuscuta reflexa* concerning two ribosomal proteins (*rpl2*, *rpl23*), one transfer RNA (*trnI*) and an ORF 2280 homologue

Dagmar Bömmer, Gerd Haberhausen, Klaus Zetsche

Institut für Pflanzenphysiologie, Justus-Liebig-Universität, Heinrich-Buff-Ring 58, 35392 Gießen, Germany

Received: 15 October 1992 / Accepted: 16 December 1992

Abstract. We have determined the nucleotide sequence of a 5.3-kb region of the plastid DNA (ptDNA) from the heterotrophic holoparasitic plant *Cuscuta reflexa*. The cloned area contains genes for the D1-protein (32-kDa protein; *psbA*), $tRNA^{His}$ (*trnH*), ORF 740 (homologous to ORF 2280 from *Nicotiana tabacum*), ORF 77 (homologous to ORF 70), $tRNA^{Leu}$ (*trnL*) and a hypothetical ORF 55 which has no homology to any known gene among higher plants. This 5.3-kb area is colinear with a 12.4-kb region of tobacco ptDNA and has therefore undergone several deletions totalling 7.1 kb. Most of the missing nucleotides belong to one large deletion in the ptDNA of *C. reflexa* of approximately 6.5 kb. This deletion involves two ribosomal protein genes, *rpl2* and *rpl23*, as well as the transfer RNA for Isoleucin (*trnI*) and a region encoding 1540 amino-acid residues of an ORF 2280 homologue, as compared to tobacco chloroplast DNA. This is remarkable since the remaining genes, especially the *psbA* gene, are highly conserved in *C. reflexa*. Furthermore, we found that the expression of the *psbA* gene is in the same range as in the autotrophic *Ipomoea purpurea* which belongs to the same family as *Cuscuta* (Convolvulaceae). Here we hypothesize a total loss of *rpl2* and *rpl23* in the entire genome of *C. reflexa*. The phylogenetic position of, and the evolutionary change of ptDNA from, *Cuscuta* are discussed.

Key words: *Cuscuta* – Parasitic plant – Chloroplast genome – Plastid evolution – *psbA* – Gene expression

Introduction

The genus *Cuscuta* exclusively consists of holoparasitic flowering plants wich have lost their ability to grow photoautotrophically. In this genus, different degrees of adaptation to parasitic life can be observed. We have demonstrated this at the cytological as well as the biochemical level. While *Cuscuta europaea* is totally achlorophyllous, *Cuscuta reflexa* contains small amounts of chlorophyll, but in both species no Rubisco activity is detectable in crude extracts (Machado and Zetsche 1990; Haberhausen et al. 1992). At the cytological level, the plastids of *C. reflexa* show only a small number of thylakoid membranes, while in *C. europaea* no thylakoid membranes can be observed (Machado and Zetsche 1990).

We now have to elucidate whether this different adaptation to parasitic life could also be seen at the molecular level, and especially in the plastid DNA (ptDNA) of *Cuscuta*. We believe that *Cuscuta* is an excellent tool to answer questions regarding the stability and the evolution of ptDNA.

The chloroplast genome has been extensively examined among photoautotrophic angiosperms (for reviews see: Whitfeld and Bottomley 1983; Palmer 1985; Sugiura 1992) and the complete sequence is known from two higher land plants, tobacco and rice (Shinozaki et al. 1986; Hiratsuka et al. 1989). Comparison of these sequences reveals a great similarity in both the gene arrangement and the coding capacity of chloroplast DNA. Therefore, the ptDNA from higher plants has been described as a very stable genetic system during land plant evolution with a relatively low mutation rate as compared to that of the nucleus (Wolfe et al. 1987; Zurawski and Clegg 1987; Palmer 1990).

Chloroplasts of higher plants encode for about 30 tRNAs which are probably sufficient to recognize all 64 codons (for review see Sugiura 1992). As in the case of the ribosomal protein genes, chloroplasts code for one-third of the 60 different protein components found in chloroplast ribosomes (Sugiura 1992). In tobacco, several ribosomal protein genes are clustered near the junction between the inverted repeat (IR) and the large single copy region (LSC). The genes *rpl23* and *rpl2* are duplicated in *Nicotiana tabacum* by being part of both segments of the IR (Shinozaki et al. 1986).

Little is known about the arrangement and development of ptDNA in parasitic plants during the course of

Correspondence to: K. Zetsche

evolution, and where presumably no selective constraint effects are any longer operative. There are only three non-photosynthetic species reported so far. The root-parasitic *Epifagus virginiana* (dePamphilis and Palmer 1990) lacks most of the photosynthetic and putative chlororespiratory (*ndh*) genes. This results in a chloroplast genome of only 71 kb in size. A similar situation has been found in a related member of the Orobanchaceae, *Conopholis americana* (Wimpee et al. 1991, 1992). In *Conopholis*, some regions in the ptDNA are highly conserved, whereas other regions are either absent or highly divergent (Wimpee et al. 1991). In a previous paper (Haberhausen et al. 1992), we have shown that in the stem parasite *C. reflexa* most of the photosynthesis-related genes are present. The nucleotide sequences of the *rbcL* and *atpB* genes are highly conserved as compared to those of tobacco but the regulatory sequences (promoter, termination sites) show *Cuscuta*-specific deletions. In the present study, we describe an area of ptDNA from *C. reflexa* which shows more differences to the corresponding region in tobacco and where a large deletion has taken place. Two ribosomal protein genes, *rpl2* and *rpl23*, are affected by this deletion. For this reason, the whole translation apparatus might be non-functional in *C. reflexa*.

Materials and methods

Nucleic acid isolation. C. reflexa Roxb. was grown in a greenhouse on *Coleus sp.* host plants. Total cellular DNA was isolated from *C. reflexa* and *N. tabacum* as described (Haberhausen et al. 1992). For the isolation of total RNA from *C. reflexa* and *Ipomoea purpurea* (L.), small stem pieces of both plants were ground in liquid nitrogen and RNA was isolated according to the LiCl method (Ausubel et al. 1988).

Molecular cloning of the psbA gene and its flanking regions. Total cellular DNA from *C. reflexa* was digested with *Hin*dIII and the *psbA* gene was located on a 6.1-kb fragment by heterologous hybridization (*psbA* gene probe kindly provided by P. Westhoff and R. G. Herrmann). This fragment was cloned into pUC18 using standard methods (Maniatis et al. 1982).

Sequence analysis. Parts of the 6.1-kb clone were subcloned into pUC18 and sequenced using a T7 sequencing system (Pharmacia). The sequence of both strands was determined and all restriction sites were confirmed. Computer analysis was performed using DNASIS (HIBIOsystems) and 'Kröger-menu' (Kröger and Kröger-Block 1984). The sequences reported in this paper have been deposited in the EMBL Data Library under the accession number X67512.

Southern analysis. To check the deletion in *C. reflexa* ptDNA, two gene probes from *N. tabacum* containing parts of ORF2280 and the *rpl2* and *rpl23* genes were hybridized to *N. tabacum* and *C. reflexa* total cellular DNA (1 µg/lane) using the ECL direct labelling and detection system (Amersham). Gene probes were kindly provided by M. Sugiura.

Northern analysis. Total cellular RNA isolated from the stems of *C. reflexa* and *I. purpurea* was electrophoresed on 1.5% formaldehyd-agarose gels (15 µg/lane) and transferred onto nitrocellulose as described by Davies et al. (1986). Filters were hybridized to direct-labelled probes (ECL, Amersham).

Fig. 1.

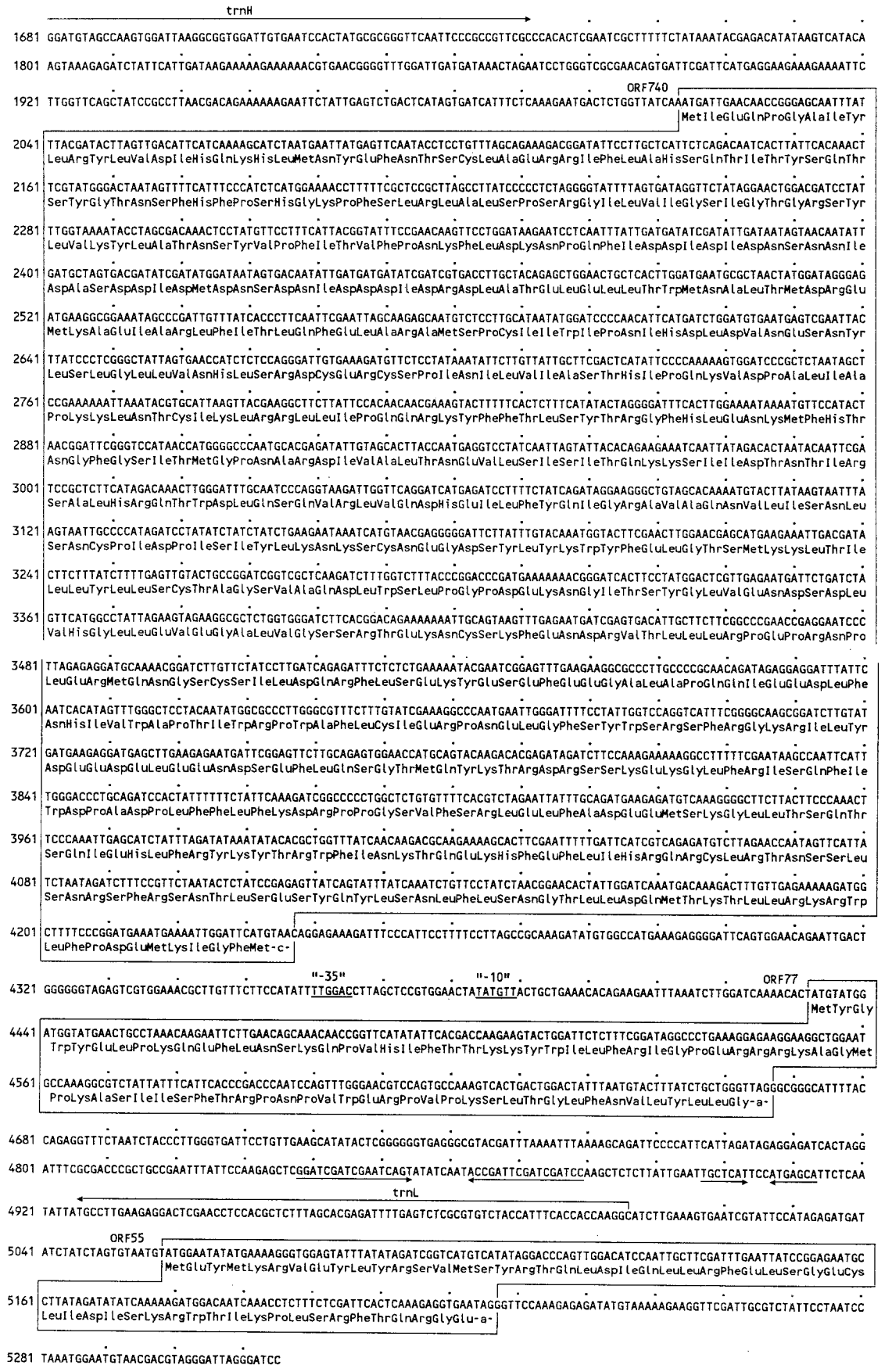

Fig. 1. Nucleotide and deduced amino-acid sequences of a 5.3-kb region of the plastid DNA from *C. reflexa*. Putative promoters are *underlined* (−35 region and −10 region). *Opposite arrows* indicate potential termination sequences

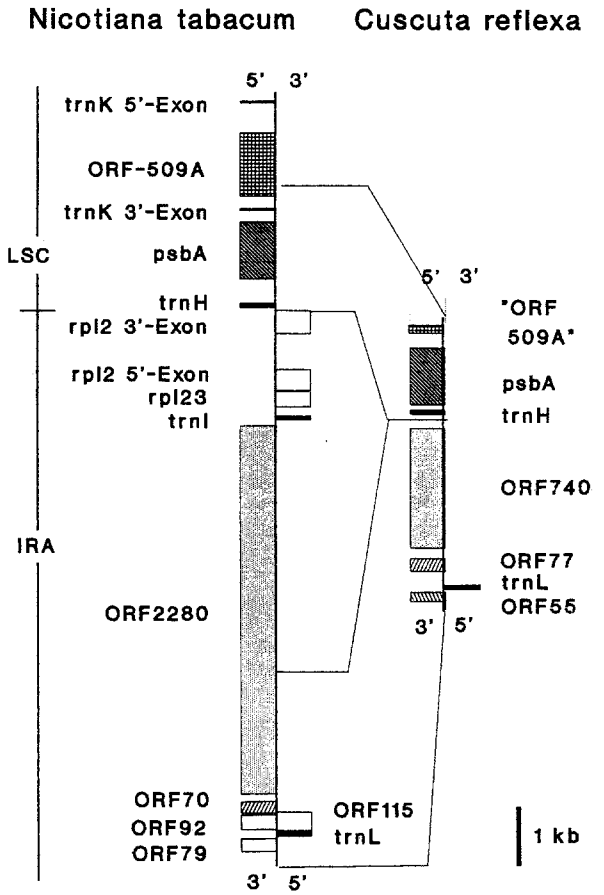

Fig. 2. Comparative gene arrangement of *N. tabacum* (*NT*) and *C. reflexa* (*CR*). A 12.4-kb area from tobacco corresponds with a colinear 5.3-kb region in the ptDNA of *C. reflexa*. Data for the tobacco gene map is derived from Shinozaki et al. (1986) and Sugiura (1992)

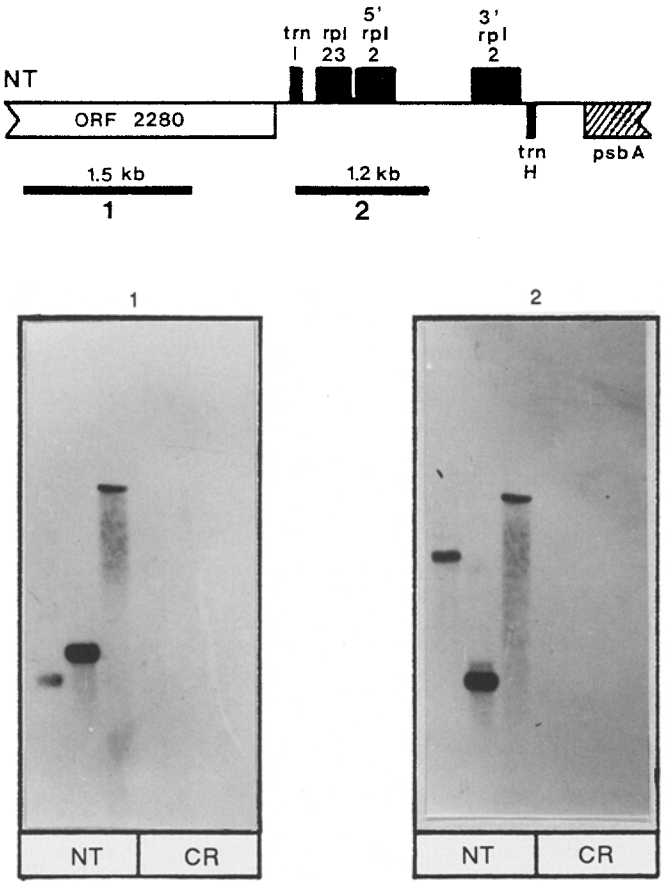

Fig. 3. Southern blots of *Nicotiana* (*NT*) and *Cuscuta* (*CR*) total cellular DNA (digested with *Bam*H1, *Eco*RI, *Hin*dIII) hybridized against two gene probes derived from a tobacco-homologous subclone (pTB28). Blots illustrate the absence of the 5'-ORF 2280 (*probe 1*) and *rpl2*/*rpl23* genes (*probe 2*) from *Cuscuta* total cellular DNA. Stringency conditions: 2 × 20 min 1 × SSC, 0.5% SDS at 25 °C

Results

Gene arrangement in Cuscuta reflexa

The nucleotide sequence of a 5.3-kb region of the *C. reflexa* plastid genome was determined (Fig. 1). The gene arrangement, presented in Fig. 2, is the same as in the autotrophic higher plant *N. tabacum*, a closely related angiosperm (*Nicotiana* and *Cuscuta* both belong to the order Solanales). The cloned area contains genes for the D1 protein (32-kDA protein; *psbA*), $tRNA^{His}$ (*trnH*), ORF740, ORF77, $tRNA^{Leu}$ (*trnL*), and ORF55, and is colinear with a 12.4-kb region of tobacco ptDNA (Fig. 2). The difference in size is due to several deletions, totalling approximately 7.1 kb, in the ptDNA of *C. reflexa*.

Deletions upstream of the psbA gene affect trnK

The *psbA* gene in *C. reflexa* ptDNA is highly conserved as compared to other higher plants. The homology to corresponding genes is in the range of 94% (e.g., *N. tabacum*, Sugita and Sugiura 1984). Neither the coding region nor the regulating sequences (promoter, terminator) are remarkably different. Upstream of the *psbA* gene the 3'-end of an ORF 509A homologue is located (Fig. 2). In *N. tabacum*, this ORF 509A is present in an intron between the two exons of $tRNA^{Lys}$ (*trnK*). The 3'-exon of $tRNA^{Lys}$ (UUU) seems to be lost in *C. reflexa* due to a deletion of about 180 bp. No statement about the 5'-exon is possible as yet (Fig. 2).

rpl2 and rpl23 have been deleted in the plastid DNA of C. reflexa

Downstream from *psbA*, a large deletion of about 6.5-kb can be observed. The deleted area starts with the first base of the large inverted repeat A (IRA) when compared to tobacco. This deletion comprises two ribosomal protein genes, *rpl2* and *rpl23*, the gene for $tRNA^{Ile}$ (*trnI*), as well as 1540 amino-acid residues of an ORF 2280 homologue. In order to find out whether these genes have been translocated in *C. reflexa*, heterologous hybridization experiments were carried out. Neither gene probes for *rpl2* and *rpl23* nor one for the 5'-end of ORF 2280 hybridized to total DNA from *C. reflexa* (Fig. 3).

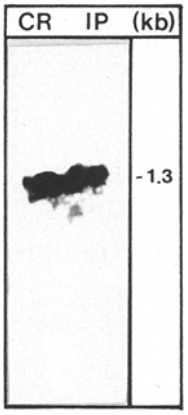

Fig. 4. Transcript of the *psbA* gene from *I. purpurea* (*IP*) and *C. reflexa* (*CR*). In Northern blots a *Cuscuta*-specific gene probe hybridized with total cellular RNA of approximately 1.3 kb in size. Stringency conditions: 2 × 20 min 6 M Urea, 1 × SSC, 0.4% SDS at 42 °C

ORF 740, ORF 77 and ORF 55 downstream from the psbA gene

As a result of this large deletion the open reading frame in *C. reflexa* homologous to ORF 2280 (tobacco, Sugiura 1992) and ORF 2131 (spinach, Zhou et al. 1988) is reduced to 740 amino acids. The overall degree of amino-acid identity is in the range of 90%.

There are two other ORFs between the *psbA* gene and *trnL* in the ptDNA of *C. reflexa* (Fig. 1). ORF 77 shows homology (approximately 87% identical residues) to ORF 70 of tobacco in a 54 amino-acid region. ORF 55 shows no homology to any known gene and may have been formed spontaneously due to several base substitutions. Two other ORFs (ORF 92, ORF 115) found in the chloroplast genome of tobacco between ORF 2280 and *trnL* are lost in *C. reflexa* ptDNA.

tRNAs downstream from the psbA gene

In the chloroplast genome of *C. reflexa* only two tRNAs, $tRNA^{His}$ (GUG) and $tRNA^{Leu}$ (CAA); are located downstream from the *psbA* gene; $tRNA^{Ile}$ (CAU) is deleted. The *trnH* gene shows identical nucleotides at 94.7% of positions compared to the *trnH* from *N. tabacum* while *trnL* is identical (100%). The *trnI* (CAU) gene located between *rpl23* and ORF 2280 of *N. tabacum* ptDNA is lost in the corresponding region of the *Cuscuta* chloroplast genome but another *trnI* (GAU) gene has been detected in the 16S–23S spacer where it is 100% homologous to tobacco (Gerd Haberhausen, unpublished results).

Transcript level of the psbA gene

We used a *psbA*-specific gene probe from *C. reflexa* to investigate the transcript level in Northern experiments using total cellular RNA from *C. reflexa* and *I. purpurea*. For this purpose, the same organs (stems) were chosen to obtain more reliable results. Surprisingly similar amounts of mRNA specific for *psbA* could be detected in both *C. reflexa* and *I. purpurea* (Fig. 4). No hybridization signals for ORF740 were found in *C. reflexa* (data not shown).

Discussion

Sequence analysis of a 5.3-kb region of the chloroplast genome from the holoparasitic flowering plant *C. reflexa* reveals a genomic organization that differs to a large extent from the corresponding region in the non-parasitic plant *N. tabacum*. This is surprising since examination of another region in the ptDNA of *C. reflexa* revealed that neither the gene content nor the nucleotide sequence of six plastid genes are significantly different from those of *N. tabacum* (Haberhausen et al. 1992). The ptDNA organization discussed here includes one gene (*psbA*) that shows no remarkable changes at the nucleotide sequence level in either its coding region or its the regulatory sequences. Also the transcription of the *psbA* gene in *C. reflexa* is comparable to that of an autotrophic plant (*I. purpurea*). In contrast to this, we found open reading frames (ORF 740, ORF 77) that could still be functional but with the grade of homology to corresponding genes decreased due to several deletions/insertions. Furthermore, some genes in *C. reflexa* ptDNA seem to be completely lost (*rpl2, rpl23, trnI, trnK*). Astonishingly, all lost genes are part of the translation apparatus. This raises the question of whether plastid-encoded genes in *Cuscuta* can be translated into proteins. In relation to this, two question scenarios are conceivable: (1) the missing genes have migrated into the nucleus and the gene products must be imported into the plastids from the cytoplasm, or (2) the genes are lost completely, which implies that the translation apparatus in plastids of *C. reflexa* is non-functional. The first of these two possibilities has been described for another ribosomal protein (*rpl22*) in some legumes by Gantt et al. (1991) and is the most recent case of gene transfer reported up to now. A number of similar events are known, due to a massive evolutionary transfer of genes from former endosymbionts to the nucleus of the host cell (Gray 1983; Palmer 1990).

In spinach, the *rpl23* gene seems to have become a pseudogene (Thomas et al. 1988) by splitting into two overlapping reading frames. As a consequence of this no *rpl23* gene product is detectable. Results from several other plants examined indicate either a loss of requirement for the CL23 protein in chloroplasts or a transfer of the *rpl23* gene to nuclear genome (Zurawski and Clegg 1987). On the other hand, ribosomal protein CL23 is functional in *Marchantia polymorpha* (Ohyama et al. 1986) and tobacco (Yokoi et al. 1991), which is more closely related to *Cuscuta* than is spinach. Here we hypothesize for a total loss of *rpl2* and *rpl23*, based on the fact that no Rubisco protein (encoded by the plastid *rbcL* gene) is detectable in *C. reflexa* by immunological methods (Haberhausen et al. 1992). If one assumes total loss of the *rpl2* and *rpl23* genes then we are confronted with the interesting phenomenon that, on the one hand, sever-

al photosynthesis-related genes are present in *C. reflexa* in a relatively unaltered and hence functional form. On the other hand, the translational machinery of the plastids appears to be non-functional. Clear evidence regarding this question cannot be given at present. However, the homology of the area examined, especially the flanking sequences in non-coding regions close to the deletion, points to the fact that the deletion (transfer or loss) is a relatively recent event in the ptDNA of *C. reflexa*. Interestingly, the deletion starts exactly with the beginning of the large inverted repeat (IRA, IRB) as compared to tobacco, probably indicating a 'hot spot region' in plastid DNA.

It seems that the ptDNA in *C. reflexa* is not reduced to the same extent as in other parasitic plants (e.g., *Epifagus* or *Conopholis*). On the genetic level this species may be regarded as a transition state between photosynthetic plants and those which are fully adapted to a parasitic life. Therefore, we believe that *C. reflexa* is an excellent model for studying the evolutionary and molecular development of the chloroplast genome.

Acknowledgements. We are grateful to Prof. Dr. P. Westhoff, Prof. Dr. R. G. Herrmann and Prof. Dr. M. Sugiura for providing several plastid gene probes. The authors like to thank Dr. M. Kröger for DNA software and J. Hustedt for reading the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (Ze 71/22-1)

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1988) Current protocols in molecular biology. Greene Publishing associated and Wiley-Interscience, Wiley and Sons, New York

Davies LG, Dibner MD, Battey JF (1986) Basic methods in molecular biology Elsevier Scientific Publishing, New York dePamphilis CW, Palmer JD (1990) Nature 348:337–339

Gantt JS, Baldauf SL, Calie PJ, Weeden NF, Palmer JD (1991) EMBO J 10:3073–3078

Gray MW (1983) Bioscience 33:693–699

Haberhausen G, Valentin K, Zetsche K (1992) Mol Gen Genet 232:154–161

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamota M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) Mol Gen Genet 217:185–194

Kröger M, Kröger-Block A (1984) Nucleic Acids Res 12:113–123

Machado MA, Zetsche K (1990) Planta 181:91–96

Maniatis T, Fritsch EF, Sambrook J (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Nature 322:572–574

Palmer JD (1985) Annu Rev Genet 19:325–354

Palmer JD (1990) Trends Genet 6:115–120

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) EMBO J 5:2043–2049

Sugita M, Sugiura M (1984) Mol Gen Genet 195:308–313

Sugiura M (1992) Plant Mol Biol 19:149–168

Thomas F, Massenet O, Dorne AM, Briat JF, Mache R (1988) Nucleic Acids Res 16:2461–2472

Whitfeld PR, Bottomley W (1983) Annu Rev Plant Physiol 34:279–310

Wimpee CF, Wrobel RL, Garvin DK (1991) Plant Mol Biol 17:161–166

Wimpee CF, Morgan R, Wrobel R (1992) Plant Mol Biol 18:275–285

Wolfe K, Li WH, Sharp PM (1987) Proc Natl Acad Sci USA 84:9054–9058

Yokoi F, Tanaka M, Wakasugi T, Sugiura M (1991) FEBS Lett 281:64–66

Zhou DX, Massenet O, Quigley F, Marion MJ, Moneger F, Huber P, Mache R (1988) Curr Genet 13:433–439

Zurawski G, Clegg MT (1987) Annu Rev Plant Physiol 38:391–418

Communicated by H. Kössel