Cloning and sequence analysis of tomato cpDNA fragments: towards developing homologous chloroplast transformation vectors

Irving Joseph Berger[1,2], Dirce Maria Carraro[2], Ralph Bock[3], Ricardo Antunes Azevedo[1] and Helaine Carrer[2,*]

[1]*Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias 11, CEP 13418-900, Piracicaba, SP, Brasil;* [2]*Laboratório de Biologia Molecular e Genômica, Departamento de Ciências Biológicas, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias 11, CEP 13418-900, Piracicaba, SP, Brasil;* [3]*Max Plant Institut for Molecular Plant Physiology, Am Muehlenberg 1, D-14476, Potsdam-Golm, Germany. *Corresponding author: hecarrer@esalq.usp.br*

With the view of developing chloroplast transformation vectors based on homologous targeting regions for tomato (*Lycopersicon esculentum* L.), plastid DNA fragments of tomato cv. IAC-Santa Clara were cloned and analyzed. Isolation and cloning of *Pst*I/*Sal*I chloroplast fragments into the pBlueScript vectors yielded the plasmids pIJB1 and pIJB2 with cpDNA fragments of 8.6 kb and 6.4 kb, respectively. DNA sequencing and computer analyses revealed that the tomato sequences cloned display from 93 to 100 % of identity to the respective fragments in tobacco, which is more pronounced in coding regions. The intergenic spacers are somewhat less conserved suggesting that evolutionary divergences occurred mainly in these putative non-coding regions. The most remarkable difference found is a 437 bp sequence present in tobacco chloroplast genome in the intergenic region of the genes *trn*E and *trn*T but completely absent in the tomato chloroplast DNA. Analysis of restriction enzyme digestion patterns revealed several unique restriction sites in the intergenic spacer regions suggesting potential utility of these sequences in species-specific vector construction for tomato chloroplast transformation.

Key words: *Lycopersicon esculentum*, *Nicotiana tabacum*, chloroplast DNA, plastid transformation, plastome, tobacco, tomato.

Clonagem e análise do seqüenciamento de fragmentos de cpDNA de tomate: desenvolvimento de vetores espécie-específicos de transformação de cloroplastos: Visando o desenvolvimento de vetores homólogos de transformação de cloroplastos para tomate (*Lycopersicon esculentum* L.), fragmentos cloroplastidiais de tomate cv. IAC-Santa Clara foram clonados e analisados. Isolamento e clonagem de fragmentos *Pst*I/*Sal*I de cloroplastos de tomate deram origem aos plasmídios pIJB1 e pIJB2, contendo fragmentos de 8,6 pb e 6,4 pb respectivamente. Seqüenciamento e análise dos fragmentos revelaram que as seqüências apresentam de 93 a 100 % de identidade com os respectivos fragmentos em tabaco nas regiões codificadoras. Espaços intergênicos mostraram-se menos conservados, sugerindo que divergências evolutivas ocorreram principalmente em regiões não diretamente envolvidas na tradução. A diferença mais marcante encontrada está em um fragmento de 437 pb presente em tabaco na região intergênica entre os genes *trn*E e *trn*T, mas ausente no DNA cloroplastidial de tomate. Análise do padrão de restrição das seqüências cloroplastidiais apontou sítios únicos de restrição, sugerindo potencial uso dessas seqüências para a construção de vetores de transformação de cloroplastos. O recente sucesso na expressão de genes exógenos de interesse agronômico em cloroplastos de plantas sugere que essa tecnologia promete revolucionar o conceito e a aplicação da engenharia genética no melhoramento de plantas.

Palavras-chave: *Lycopersicon esculentum*, *Nicotiana tabacum*, DNA cloroplastidial, fumo, tomate, transformação plastidial.

INTRODUCTION

Besides the genetic information localized in the nucleus, plant cells contain DNA in chloroplasts and mitochondria. These organelles are semi-autonomous and have their own machinery for transcription, protein synthesis and DNA replication (Herrmann et al., 1992). The expression of the three different genetic systems within a plant cell is tightly coordinated during plant development.

Organellar genomes provide attractive alternative targets for plant genetic engineering. Recent methodological progress has made possible the development of transformation technologies for chloroplast genomes (plastomes; Maliga, 2004). Transplastomic plants, i.e. plants harboring transformed chloroplast genomes, have several advantages as compared to transgenic plants generated by transformation of the nuclear genome. Most notably, they enhance environmental safety concerning transgene distribution with the pollen since chloroplasts are uniparentally maternally inherited in most crop plants. Moreover, transplastomic plants also show extraordinarily high expression levels of the introduced transgene(s) (Carrer et al., 1993, De Cosa et al., 2001). A chloroplast transformation technology for higher plants was first reported for tobacco in 1990 by Maliga's group (Svab et al., 1990) and took almost one decade of efforts for *Arabidopsis thaliana* (Sikdar et al., 1998), potato (Sidorov et al., 1999) and tomato (Ruf et al., 2001) chloroplasts to be successfully transformed. Due to this success and specially because *Arabidopsis*, potato and tomato were transformed with homeologous vectors (not species-specific vectors), the plastid genomes of important crop species have received increasing attention with the expectation to develop chloroplast transformation systems and to introduce agronomically relevant traits for plant improvement into the chloroplast genome (revised in Daniell et al., 2005). Following these tendencies, the crop species carrot (Kumar et al., 2004a), soybean (Dufourmantel et al., 2004) and cotton (Kumar et al., 2004b) also have had their chloroplast genomes transformed recently.

As chloroplast transformation occurs by homologous recombination events (Bock and Hagemann, 2000), it is very important to know the sequence of the chloroplast DNA (cpDNA) region to which the gene of interest should be specifically targeted. During the past 20 years, 19 higher plant plastid genomes have been completely sequenced, including *Nicotiana tabacum* (Shinozaki et al., 1986), *A. thaliana* (Sato et al., 1999), *Spinacea oleracea* (Schmitz-Linneweber et al., unpublished - direct submission to NCBI,

accession NC 002202), *Oryza sativa* (Hiratsuka et al., 1989), *Zea mays* (Maier et al., 1995), *Pinus thunbergii* (Wakasugi et al., 1994) and *Saccharum officinarum* (Calsa Junior et al., 2004). From the comparative analyses of these genomes, the concept has emerged that higher plant plastomes show a remarkably conserved structure and coding capacity across large phylogenetic distances. Also, these sequence data showed that related species display highly homologous DNA sequences and a high level of conservation concerning both gene content and gene order (Calsa et al., 2004).

Since both tobacco and tomato belong to the Solanacean family, it can be expected that they share high chloroplast genome homology, a characteristic also indicated by the recent tomato chloroplast transformation with a vector developed for tobacco (Ruf et al., 2001). Given the great success of tobacco chloroplast transformation showing several different target sites for transgene insertion (Maliga, 2003), our goal has been to find alternative regions in the tomato chloroplast genome suitable for efficient uptake of transgenes, and hence the importance of constructing transformation vectors based on homologous targeting regions. Since tomato is a horticulture species grown worldwide and an excellent plant system for agricultural biotechnology, the development of a chloroplast transformation technology for tomato would provide a highly efficient and versatile tool for plant breeders and biotechnologists.

In order to select homologous targeting regions for constructing tomato chloroplast transformation vectors, we have isolated, cloned and sequenced two *Pst*I/*Sal*I tomato chloroplast DNA fragments of 8,660 bp and 6,366 bp, respectively, in size. Our data show that the gene organization in tobacco plastomes is also well conserved in tomato. The sequenced tomato cpDNA fragments were found to contain several non-coding regions suitable for the uptake of transgenes. Consequently, these fragments will serve as the basis for constructing plasmid vectors species-specific for tomato chloroplast transformation. The vectors would be important for tomato plastome engineering for both applied and basic research. Several physiological and biochemical processes related to cellular mechanisms occur inside plastids and the putative *orf*s related to these processes could be potential targets for gene function studies with possibilities for the elucidation of these physiological mechanisms. Some attempts have been made to produce engineered Rubisco proteins to improve photosynthesis and yield in plants (Whitney and Andrews, 2001; Whitney and Andrews, 2003). In addition, since sugar and starch synthesis are located in the plastids,

genes related to carbohydrate metabolism are interesting genes to be expressed in this organelle. An encouraging example of successful carbohydrate engineering using the transplastomic technology is the recent production of trehalose in tobacco chloroplasts giving rise to drought resistant plants (Lee et al., 2003). Furthermore, trehalose has been used as a valuable substance for food preservation, pharmaceutical and medical purposes (Schiraldi et al., 2002).

MATERIAL AND METHODS

Chloroplast isolation and DNA extraction: Intact chloroplasts were isolated from leaves of tomato cv. IAC-Santa Clara (*Lycopersicon esculentum* Mild.) in Percoll gradients as described by Bock (1998). Extracted cpDNA was further purified by cesium chloride gradient ultracentrifugation (Romano, 1998).

Cloning and sequencing of tomato cpDNA fragments: Based on a previously published *Sal*I/*Pst*I/*Hpa*I/*Kpn*I tomato chloroplast DNA restriction map (Phillips, 1985), *Sal*I/*Pst*I fragments were cloned by shotgun cloning into pBlueScript KS(-) vectors. Two plasmid clones containing cpDNA fragments of 8,660 bp (subsequently referred to as pIJB1) and 6,366 bp (named pIJB2), respectively, were selected for further analysis.

For the complete sequencing of the inserts, sub-cloning strategies were combined with primer walking. The pIJB1 plasmid was sub-cloned as pIJB3, pIJB4, pIJB5, pIJB6 and pIJB8 as *Pst*I/*Cla*I (2.3Kb), *Cla*I/*Sma*I (1.6Kb), *Sma*I/*Bam*HI (1.6Kb), *Bam*HI/*Sal*I (3.1Kb) and *Xho*I/*Sal*I (partial from pIJB6, 2.3Kb) fragments respectively. The pIJB2 plasmid was fragmented by nebulization for a shotgun library construction into pUC18/*Sma*I (SureClone/Amersham Pharmacia). DNA sequencing was performed with universal forward and reverse primers and Big Dye terminators on an ABI Prism 377 DNA Sequencer (Applied Biosystems). The complete sequences were assembled using the Sequencher™ software version 3.1 (Gene Codes Corporation).

Computer analysis: The generated tomato cpDNA sequences were compared with the chloroplast genome sequence of tobacco (*Nicotiana tabacum*) using the BLAST 2.0 algorithm (NCBI - National Center for Biotechnology Information – http://www.ncbinlm.nih.gov/BLAST/). Tomato nucleotide sequences were aligned with the corresponding sequences from tobacco with the DIALIGN 2.0 software (Morgenstern, 1999). In order to compare the chloroplast coding sequences

from tomato (*L. esculentum*) with those from other plant species we have used multialignment analyses of nucleotide and amino acid sequences of *Z. mays* (GenBank/EMBL/DDBJ accession number - #X86563), *Oryza sativa* (#X15901), *P. thumbergii* and *Marchantia polymorpha* (#X04465). In addition, RNA editing (based on an earlier analysis in tobacco; Hirose et al., 1999), GC content, restriction patterns, occurrence of nucleotide substitutions, duplications and deletions were analyzed.

RESULTS AND DISCUSSION

Cloning and sequencing of tomato ptDNA fragments: Two tomato chloroplast DNA fragments of 8,660 bp and 6,366 bp were cloned into pBS KS (-) vectors as *Pst*I/*Sal*I fragments resulting in the plasmids pIJB1 and pIJB2, respectively. Sequencing of the fragment ends and comparison with the homologous tobacco cpDNA sequences (Shinozaki et al., 1986) revealed that the sequences correspond to the region between the *psb*C and *rpo*B genes (pIJB1) and the region between the *rpl*20 and *psb*N genes (pIJB2). The sizes of the corresponding regions in tobacco plastome are 9,127 bp and 6,375 bp as compared to 8,660 bp and 6,366 bp, respectively, in tomato.

Comparative analysis and identification of coding sequences and intergenic spacers: Comparative analyses of the nucleotide sequences of coding and intergenic regions (spacers) between tomato and tobacco are shown in tables 1 and 2. For the chloroplast fragment in plasmid pIJB1 (8,660 bp), the nucleotide sequences showed 93 to 100 % identity (table 1). The coding regions for *trn*C, *pet*N (*ycf*6), *trn*D, *trn*Y, *trn*E and *trn*T are most highly conserved (table 1). They exhibit 100 % identity to the corresponding tobacco sequences. The photosystem II gene *psb*M is also highly conserved with the exception of a single nucleotide substitution in the position +78, that in tobacco is a C, but a T nucleotide in the respective gene in tomato. However, this nucleotide substitution is silent and does not change the coding properties (TAC $\rightarrow$ TAT; both codons specify isoleucine). Interestingly, this TAT isoleucine codon found in tomato is also present in the corresponding genes of *Z. mays*, *O. sativa*, *S. officinarum* and *M. polymorpha* (alignments not shown) suggesting that a point mutation may have occurred at this position in tobacco. The *psb*D coding regions from tobacco and tomato share 99.2 % nucleotide sequence identity (table 1). Again, this does not result in any deviation at the amino acid sequence level as all nucleotide

substitutions are also silent. For the chloroplast fragment in plasmid pIJB2 (6,366 bp), the nucleotide sequences share 78.9 to 100 % identity with the corresponding tobacco sequences (tables 1 and 2).

Since high level of similarities has been shown for coding sequences in different plastomes (Sato et al., 1999, Kato et al., 2000, Calsa Junior et al., 2004), we performed alignments and comparative analyses of the sequence homology for the *trn*C, *pet*N (*ycf*6), *psb*M, *trn*D, *trn*Y, *trn*E, *trn*T and *psb*D (derived from plasmid pIJB1), *rpl*20, 5'*rps*12, *clp*P, *psb*B, *psb*T and *psb*N (derived from plasmid pIJB2) coding sequences in several plastomes, such as *Z. mays* (GenBank #X86563), *Oryza sativa* (GenBank #X15901), *N. tabacum* (GenBank #Z00044), *P. thunbergii* (GenBank #D17510), *Marchantia polymorpha* (GenBank #X04465) and *S. officinarum* (GenBank #NC005878). The alignments (not shown) clearly demonstrate that, among all genes analyzed, *clp*P is the least conserved gene. Remarkably, the *clp*P gene showed a great accumulation of point mutations in two of the three exons (table 1). The exons 1 and 3 showed 78.9 % and 81.7 % identity, respectively, revealing an interesting exception from the generally higher conservation in the coding regions as compared to non-coding intergenic spacers. Analysis of the corresponding amino acid sequence alignments also confirmed the high degree of interspecific conservation also for the level of the gene products (proteins). The tomato *psb*B coding region is an example (table 1). Although the tomato *psb*B carries 11 nucleotide substitutions as compared to the tobacco sequence (Shimada and Sugiura, 1991), the tomato and tobacco sequences specify fully identical amino acid sequences, since all of the substitutions occur at the third position in the codon and hence are silent (Rodolphe and Mathe, 2000). Again, the *clp*P exons and their products are exceptions. Changes in their nucleotide sequences result in both silent mutations and mutations with consequent changes in the protein sequence.

Analysis of RNA editing based on Hirose et al. (1999) showed that the pattern of RNA editing sites in tobacco is also conserved in the tomato chloroplast genes sequenced. Both *rpl*20 and *rpo*B encode potential RNA editing sites in tomato. Within the *rpl*20 coding region, there is an editing site generating a leucine codon (uUa) from an original serine codon (uCa) present in the DNA sequence. Similarly, editing sites at the codons 113 (serine (uCu) → phenylalanine (uUu)), 158 (serine (uCa) → leucine (uUa)) and 184 (serine (uCa) → leucine (uUa)) present in the tobacco *rpo*B gene were confirmed for tomato. Two other editing sites present within

Table 1. Sizes of coding regions and homology of the nucleotide sequences of the analyzed tomato chloroplast genome fragments with the corresponding regions in tobacco. Plasmids pIJB1 and pIJB2 carry cpDNA inserts of 8,660 bp and 6,366 bp, respectively.

Coding Region	Size (bp)		Identity (%)
	Tomato	Tobacco	
Transfer RNAs			
*trn*E-UUC (pIJB1)	73	73	100
*trn*Y-GUA (pIJB1)	84	84	100
*trn*D-GUC (pIJB1)	74	74	100
*trn*C-GCA (pIJB1)	72	72	100
*trn*T-GGU (pIJB1)	72	72	100
Photosystem II			
*psb*D (pIJB1)	1062	1062	99.2
*psb*M (pIJB1)	105	105	99.0
*pet*N/*ycf*6 (pIJB1)	90	90	100
*psb*B (pIJB2)	1527	1527	99.3
*psb*T (pIJB2)	104	104	97.1
*psb*N (pIJB2)	132	132	100
Other genes			
*pet*N (pIJB1)	90	90	100
*rpl*20 (pIJB2)	387	387	99.2
5' *rps*12 (pIJB2)	114	114	100
*clp*P (exon 1) (pIJB2)	71	71	81.7
*clp*P (exon 2) (pIJB2)	292	292	94.2
*clp*P (exon 3) (pIJB2)	228	228	78.9

Table 2. Size of intergenic spacers and homology of the nucleotide sequences of the tomato chloroplast DNA fragments with the corresponding regions in tobacco.

Intergenic Spacer	Size (bp)		Identity (%)
	Tomato	Tobacco	
pIJB1 (8.660 kb)			
*psb*D - *trn*T	1178	1218	96.5
*trn*T - *trn*E	411	848	93.4 (44.5[1])
*trn*E - *trn*Y	59	59	93.2
*trn*Y - *trn*D	108	108	97.2
*trn*D - *psb*M	1081	1072	93.3
*psb*M - *pet*N	1107	1132	95.3
*pet*N - *trn*C	648	670	94.0
*trn*C - *rpo*B	1328	1281	95.6
pIJB2 (6.366 kb)			
*psb*N - *psb*T	77	77	94.8
*psb*T - *psb*B	201	201	99.0
*psb*B - *clp*P(1)	444	449	98.9
*clp*P(1) - *clp*P(2)[2]	798	807	94.1
*clp*P(2) - *clp*P(3)[3]	617	637	95.6
*clp*P(3) - 5'*rps*12	138	138	99.3
5'*rps*14 - *rpl*20	811	811	95.8

[1] Homology when the 437 bp deletion in tomato is considered;
[2] Intron 1; [3] Intron 2.

Table 3. Alignment of the *trn*E-*trn*T intergenic spacer regions from tomato and tobacco showing the 437 bp deletion within the tomato sequence. Asterisks indicate homology; lowercase letters denote sequence divergence (Morgenstern, 1999).

*trn*E-*trn*T spacer: alignment and homology between tomato and tobacco

```
NT     1    CtGCTTGACC AACCGCCATC ATACTATGAT CATAGTATGA TCaGTTTTTT    50
LE     1    CcGCTTGACC AACGGCCATC ATACTATGAT GATAGTATGA TCcGTTTTTT    50
            * ******** ********** ********** ********** ** *******

NT    51    GAAATTGTCA ATATAATCGA ATGATTCTAt CCGAGGGATC TTTCCCCCTT   100
LE    51    AAAATTGTCA ATATAATCAA ATGATTCTAg CCGAGGGATC TTTCCCCCCT   100
            ********** ********** ********* ********** **********

NT   101    TCAGAATTGC ATAGAATT-- TTTTTATTCG TCATTGATGA ATTATTCATT   148
LE   101    TCAGAATTGC ATAGAATTgt TTTTTATTCG TCATTGACGA ATTATTCATT   149
            ********** ********   ********** ********** **********

NT   149    AGAATCGcCA TTAGAAATCT AGTAGTAGTA TTTTTTTTtt tttggAATTA   198
LE   151    AGAATCGaCA TTAGAAATCT AGTAGTAGTA TTTTTTTT-- -----AATTA   193
            ******* ** ********** ********** ********       *****

NT   199    TTTCAATTGA ATTTCTTTCg attattttag tttagattat ttagtattta   248
LE   194    TTTCAATTGA ATTTATTTCt ---------- ---------- ----------   214
            ********** ********

NT   249    gaattttctt tttttattat aaataaaaaa aaaattaata aatacaaaaa   298
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   299    atagaaataa taaggaagag taggattttt gcagggaatg attggtccgt   348
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   349    cagaaaagga aaaaggtgtg aaattctatt tctttcactt tcatttgatt   398
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   399    cattgttaag acgagatatc cttatctccc tcccaccaag acaggaaatt   448
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   449    aacaaacgag aaatctagta agcgggatca agaagaaaat tcttttttct   498
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   499    ccaagaattt agttcaggag acaagtagaa tctcttcatt ccatgattcg   548
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   549    atgaaatatc ttgaatttta tgttgaattg ctaggtgtat gtacatgtat
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   599    caatcaagtg aattttgttc tggtgggatc aattcaataa aagaaaaaaa   648
LE   214    ---------- ---------- ---------- ---------- ----------   214

NT   649    gcaATTCGAG TCGGTCTTGA AACAATTCAT TGCATTTTCT CCTAgACTTC   698
LE   214    ---ATTCGAG TCGGTCTTGA AACAATTCAT TGCATTTTCT CCTAtACTTC   260
               ******* ********** ********** ********** **** *****

NT   699    CTAGGTAAAT CCATTTTATT ATTCAACAAT GAGCCACTAG ACACTATGTA   748
LE   261    CTAGGTAAAT CCATTTTATT ATTCAACAAT GAGCCACTAG ACACTATGTA   310
            ********** ********** ********** ********** **********

NT   749    TCTACTGCAT GTACTTA-TG CATATATACT TATGTTTATA ATATATGTAC   797
LE   311    TCTACTGCAC GTACTTAaTG CATATATACT TATGTTTATA ATATAAGTAC   360
            ********** ******* ** ********** ********** **********

NT   798    CTATAGATAT TTtATCCACA TAGTGAATAA TTCCGGAATT AAATCAAAAA   848
LE   361    ATATAGATAT TTaATCCACA TAGTGAATAA TTCCGGAATT AAATAAAAAA   410
            ********** ** ******* ********** ********** **********

NT   849    G
LE   411    G
            *
```

the maize *rpo*B gene (codons 189 and 208) (Maier et al., 1995) are absent from both tobacco and tomato sequences.

As compared to the coding sequences, our data revealed significantly less homology in the intergenic regions due to the accumulation of different mutations including substitutions, insertions, duplications and deletions in tomato compared to tobacco. This indicates that divergences occur mainly in regions that are not directly involved in protein codification or tRNA structure (table 2). The most remarkable difference is a 437 bp fragment present in tobacco in between the *trn*E and *trn*T genes, but absent from the tomato chloroplast DNA. The alignment for this spacer is shown in table 3. One possible explanation is that, at some point in evolution, this 437 bp fragment was deleted from the tomato ptDNA by either recombination or strand slippage during DNA replication. Other noteworthy differences between the tomato and tobacco sequences are, for example, the sequences CTTTTTA (*psb*D-*trn*T spacer), TCAGT and GATATAT (*trn*C-*rpo*B spacer), TCTAGATTTGAGTGA (*rpl*20-*rps*18 spacer), ATAGATATCTTAC (*clp*P(1)-*clp*P(2) intron) and TTTT (*psb*T-*psb*N spacer) which are duplicated in tandem in tomato. In contrast, the sequences AGAA (*psb*M-*pet*N spacer), CCTCC, TTTTT and TTTTTT (*clp*P(2)-*clp*P(3) intron) and TTCTAATTGA (*clp*P(1)-*clp*P(2) intron) are duplicated in tobacco but occur as a single copy in tomato.

The GC content in the 8,660 bp and 6.366 bp fragments was calculated to be 35.13 % and 36.94 %, respectively, which is comparable to the GC content of the corresponding tobacco sequences (34.75 % and 37.00 %; Shimada and Sugiura, 1991; 37.9 % for the complete tobacco chloroplast genome) reflecting the high homology of the sequences. Also, the same pattern of codon usage was found in tomato and tobacco with the preferential occurrence of A or T in the third codon position (Shimada and Sugiura, 1991).

Since both tomato and tobacco belong to the *Solanaceae*, a large family of dicotyledonous plants, the high homology found between tobacco and tomato chloroplast sequences as well as the conserved gene order, the similar GC content and the high degree of conservation of the gene products is not surprising and confirms the close genetic relationship between members of this plant family.

Physical maps illustrating gene organization in the tomato DNA fragments cloned in pIJB1 and pIJB2 are shown in figures 1 and 2. Also, these figures show the unique sites for selected restriction endonucleases identified in the intergenic regions. Within the 8,660 bp fragment (pIJB1), *Sma*I, *Stu*I, *Bam*HI, *Xho*I and *Eco*RV are potential insertion sites options while *Ava*I and *Eco*NI are unique restriction sites within the 6,366 bp (pIJB2).

Existence of unique restriction sites identified in intergenic regions in both fragments will make possible the construction of homologous transformation vectors for tomato chloroplast transformation. Plastid transformation technologies promise to be very useful for the expression of agronomical relevant transgenes in commercial tomato lines for genetic breeding programs. Chloroplast transformation makes possible high levels of transgene expression and foreign protein accumulation and has additional attractive advantages, such as, absence of epigenetic effects, transgene insertion by homologous recombination, environmental safety and simultaneous expression of several foreign genes (as polycistronic mRNAs) by a single transformation event (Bock and Hagemann, 2000, Bock and Khan, 2004, Maliga, 2004). This technology allows the expression of a variety of genes to improve crops with new agronomical traits as well as to produce plants with higher nutritional contents by expressing higher levels of essential amino acids, fatty acids, vitamins or pharmaceutical compounds.
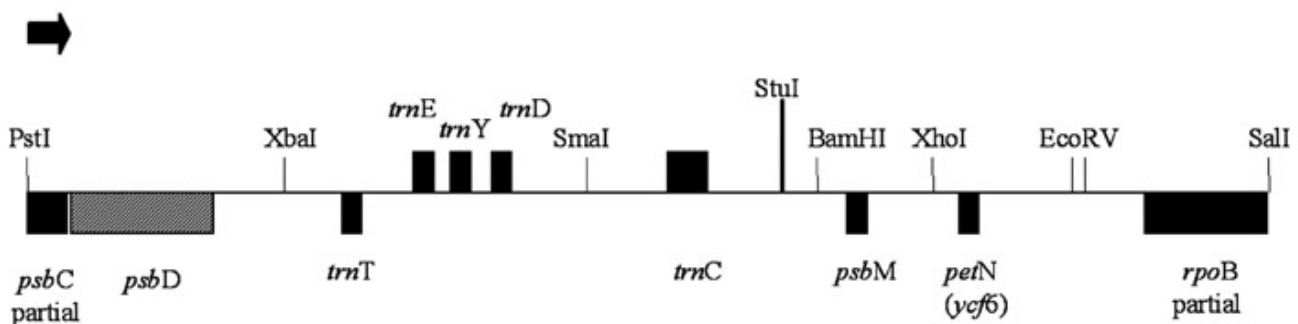
Figure 1. Physical and restriction maps of the 8,660 bp tomato chloroplast fragment as contained in plasmid pIJB1. The restriction map shows suitable enzymes having unique sites in the intergenic spacers. *Pst*I and *Sal*I are the cloning sites of the fragment. Black arrows indicate direction of gene transcription above and below the central black line.
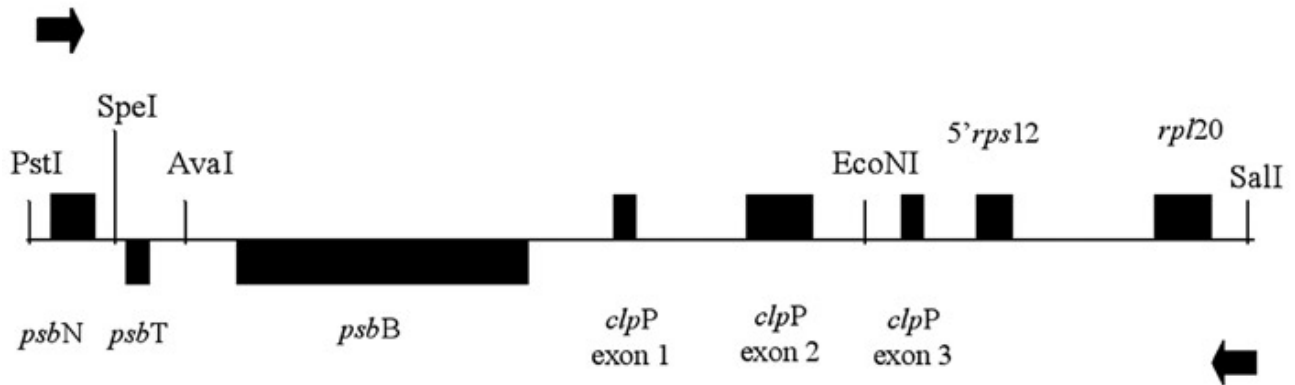
Figure 2. Physical and restriction maps of the 6,366 bp tomato chloroplast fragment cloned in plasmid pIJB2. The restriction map shows suitable enzymes having unique sites in the intergenic spacers. *Pst*I and *Sal*I are the cloning sites of the fragment. Black arrows indicate direction of gene transcription above and below the central black line.

REFERENCES

Bock R, (1998) Analysis of RNA editing in plastids. Methods - A Companion to Meth. Enzymol. 15:75-83.

Bock R, Hagemann R (2000) Plastid genetics: Manipulation of plastid genomes and biotechnological applications. Progr. Bot. 61:76-90.

Bock R, Khan MS (2004) Taming plastids for a green future. Trends Biotechnol. 22:311-318.

Calsa Junior T, Carraro DM, Benatti MR, Barbosa AC, Kitajima JP, Carrer H (2004) Structural features and transcript-editing analysis of sugarcane (*Saccharum officinarum* L.) chloroplast genome. Curr. Genet. 46:366-373.

Carrer H, Hockenberry TN, Svab Z, Maliga P (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. Mol. Gen. Genet. 241:49-56.

De Cosa B, Moar W, Lee SB, Miller M, Daniell H (2001) Overexpression of the Bt cry2Aa2 operon in chloroplasts leads the formation of insecticidal crystals. Nat. Biotechnol. 19:71-74.

Dufourmantel N, Pelissier B, Garçon F, Peltier G, Ferullo J-M, Tissot G (2004) Generation of fertile transplastomic soybean. Plant Mol. Biol. 55:479-489.

Hager M, Bock R (2000) Enslaved bacteria as new hope for plant biotechnologists. Applied Microbiol. Biotech. 54:302-310.

Hirose T, Kusumegi T, Tsudzuki T, Sugiura M (1999) RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. Mol. Gen. Genet. 262:462-467.

Kato T, Kaneko T, Sato S, Nakamura Y, Tabata S (2000) Complete structure of the chloroplast genome of a legume, *Lotus japonicus*. DNA Res. 31:323-330.

Kumar S, Dhingra A, Daniell H (2004a) Plastid-expressed *betaine aldehyde dehydrogenase* gene in carrot cultured cells, roots, and leaves confers enhanced salt tolerance. Plant Physiol. 136:2843-2854.

Kumar S, Dhingra A, Daniell H (2004b) Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. Plant Mol. Biol. 56:203-216.

Lee S-B, Kwon H-B, Kwon S-J, Park S-C, Jeong M-J, Han S-E, Byun M-O, Daniell H (2003) Accumulation of trehalose within transgenic chloroplasts confers drought tolerance. Mol. Breed. 11:1-13.

Maier RM, Neckermann K, Igloi GL, Kössel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J. Mol. Biol. 251:614-628.

Maliga P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotech. 21:20-28.

Maliga P (2004) Plastid transformation of higher plants. Annu. Rev. Plant Biol. 55:289-313.

Morgenstern B (1999) Dialign 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics 15:211-218.

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572-574.

Phillips AL (1985) Restriction map and clone bank of tomato plastid DNA. Curr. Genet. 10:147-152.

Rodolphe F, Mathe C (2000) Translational conditional models for protein coding sequences. J. Comput. Biol. 7:249-260.

Romano E (1998) Extração de DNA de Tecidos Vegetais. In: Brasileiro ACM, Carneiro VTC (eds), Manual de transformação genética de plantas, pp.163-175. Embrapa-SPI/Embrapa-Cenargen, Brasília.

Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nature Biotechnol. 19:870-875.

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res. 29:283-290.

Schiraldi C, Di Lernia I, De Rosa, M (2002) Trehalose production: exploiting novel approaches. Trends Biotechnol. 20:420-425.

Shimada H, Sugiura M (1991) Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. Nucl. Acid Res. 19:983-995.

Shinozaki K, Ohme M, TanakaM, 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) The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. EMBO J. 5:2043-2049.

Sidorov VA, Kasten D, Pang S-Z, Hajdukiewicz PTJ, Staub JM, Nehra NS (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J. 19:209-216.

Sikdar SR, Chaudhuri S, Maliga P (1998) Plastid transformation in *Arabidopsis thaliana*. Plant Cell Rep. 18:20-24.

Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc. Natl. Acad. Sci. U.S.A. 87:8526-8530.

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudsuki T, Sugiura M (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thumbergii*. Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798.

Whitney SM, Andrews TJ (2001) Plastome-encoded bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) supports photosynthesis and growth in tobacco. Proc. Natl. Acad. Sci. U.S.A. 98:14738–14743.

Whitney SM, Andrews TJ (2003) Photosynthesis and growth of tobacco with a substituted bacterial Rubisco mirror the properties of the introduced enzyme. Plant Physiol. 133:287–294.