

Fragments of plastid DNA in the nuclear genome of tomato: prevalence, chromosomal location, and possible mechanism of integration

Eran Pichersky, John M. Logsdon Jr*, J. Mitchell McGrath and Robert A. Stasys

Department of Biology, University of Michigan, Ann Arbor, MI 48109, USA

Received August 6, 1990

Summary. We have undertaken a systematic search for plastid DNA sequences integrated in the tomato nuclear genome, using heterologous probes taken from intervals of a plastid DNA region spanning 58 kb. A total of two short integrates (202 and 141 nucleotides) were isolated and mapped to chromosomes 9 and 5, respectively. The nucleotide sequence of the integrates and that of the flanking regions were determined. The integration sites contain direct repeat elements similar in position (but not in length or sequence) to the direct repeats previously observed with another plastid integrate in the tomato nuclear genome. Based on these results, a model for the process of movement and integration of plastid sequences into the nuclear genome is discussed.

Key words: *Lycopersicon esculentum* – DNA transfer – Plastid DNA – Restriction fragment length polymorphism

Introduction

According to the endosymbiotic hypothesis, a massive movement of plastid DNA to the nucleus must have occurred (Weeden 1981). The majority of genes originally residing on the plastid genome are now found in the nucleus, and not in the plastids, of present-day plants, with the gene products being imported back into these organelles. This gene transfer phenomenon, whether in single or multiple events, must have mostly occurred early in the evolution of photosynthetic eukaryotes, since the gene complement of plastid genomes is remarkably similar among land plants and some algae (Palmer 1985). However, occasionally a gene found in the plastid genome of some species is found in the nucleus in others, suggesting more recent transfer events (Baldauf and

Palmer 1990; Palmer 1990a). Theoretical considerations suggest that for the transfer to be successful, intermediate steps are highly advantageous (Palmer 1990b). The first step would be the transfer of a plastid sequence into the nucleus without the concomitant loss of the same sequence in the plastid (in essence, a gene duplication). The transferred copy is unlikely to be immediately functional in the nuclear genome due to lack of eukaryotic regulatory sequences. The gene product, even if synthesized, is unlikely to be imported back into the plastid; therefore, additional genetic changes are required to produce a functional gene and a functional gene product. Later, the active nuclear gene may totally supplant the plastid copy, and the latter may be eliminated without adverse effects on the organism. Nevertheless, it is likely that in most cases of plastid DNA transfer to the nucleus, the transferred sequence would not have given rise to active genes but instead would have eventually been mutated and lost.

It has been shown by the Southern blotting technique that many regions of the plastid genome are also present in the nuclear genome of spinach and related species (Timmis and Scott 1983; Scott and Timmis 1984; Ayliffe et al. 1988). Since such hybridization experiments rely on a relatively high level of DNA sequence similarity (>80%), it is likely that only sequences which were recently transferred (and have not had time to diverge) are detected. Alternatively, these sequences may have been transferred to the nucleus long ago but have maintained a function which has been conserved by selection. Without primary sequence information, these hypotheses could not be tested.

However, the physical isolation of nuclear fragments containing plastid “integrates” has proven technically difficult. Plant cells in all tissues contain plastids and hundreds of copies of plastid DNA molecules, and thus plant nuclear DNA isolated from any tissue by the usual procedures is highly contaminated with plastid DNA. Thus, when a plastid DNA segment is used as a probe to screen “genomic” libraries, most of the recombinant phages identified contain *bona fide* plastid DNA inserts,

* Present address: Department of Biology, Indiana University, Bloomington, IN 47405, USA

Offprint requests to: E. Pichersky

and it is an extremely time-consuming process to purify each clone and characterize it in order to identify the few clones which contain nuclear DNA with plastid integrates. The first nuclear fragment with a plastid integrate was cloned independently of the plastid sequence (Pichersky et al. 1988), and it was only when sequence comparisons were carried out that it was determined that a plastid integrate was present on that fragment (Pichersky and Tanksley 1988). Cheung and Scott (1989) attempted to construct libraries from purified nuclear DNA, but although the DNA was substantially clean of plastid contamination, it was not clonable. Consequently, they resorted to first digesting the DNA sample with the frequently cutting, methylation-sensitive restriction enzyme *HpaII* (reasoning that the entire plastid genome was not extensively methylated and therefore should be digested into very small fragments), and then cloning the unrestricted DNA into a phage vector. In this way, they were able to isolate a single recombinant phage with a nuclear fragment containing a plastid integrate.

We report here a systematic search for plastid DNA integrates in the tomato nuclear genome, using a novel screening procedure and utilizing probes taken from intervals of the tobacco plastid DNA region spanning 58 kb. The two new integrates isolated are short (<250 nucleotides), as was the first tomato integrate previously analyzed, and integration sites also show direct repeats previously observed. Based on these results, a model for the process of movement and integration of plastid sequences into the nuclear genome is discussed.

Materials and methods

Probes. We have used 11 probes derived from the Large Single Copy (LSC) region of *Nicotiana tabacum* plastid genome (Fig. 1). These fragments were obtained from the appropriate plasmids of a *N. tabacum* chloroplast DNA library constructed by Dr. M. Sugiura and provided to us by Dr. J.D. Palmer. The fragments to be used as probes were subcloned into plasmid pUC18 or

pBluescript (SK⁺). Plasmid DNAs from these subclones were then digested with the appropriate restriction enzymes, run in a 1% agarose gel, and the fragments cut out and electroeluted. Fragments were labelled with ³²P by the random-primer method according to standard procedures (Pichersky and Tanksley 1988).

Library screen, cloning, and nucleotide sequence determination. Procedures for library screening, cloning, and Southern blotting were as described in Pichersky et al. (1985) and references therein. Hybridization was carried out with 6 × SSC solution at 60° C. Under these conditions, sequences with >70% homology can be detected (Pichersky et al. 1987). Two tomato libraries, in the phage vectors EMBL3 and Charon 35, which together cover >95% of the nuclear genome (Sugita et al. 1987; unpublished data) were used in the screen. Nucleotide sequence determination was by both enzymatic (Sanger et al. 1977) and chemical (Maxam and Gilbert 1980) methods.

Genetic mapping. Linkage analysis of DNA markers was as previously described (Pichersky et al. 1985). The F₂s used in the mapping and the databank associated with them were provided by Dr. S.D. Tanksley.

Results

Isolation of tomato nuclear fragments containing sequences homologous to plastid DNA

We have undertaken a systematic search for sequences in the nuclear genome of tomato which share high (>70%) sequence identity with plastid DNA. Probes were selected from the Large Single Copy (LSC) region of the tobacco plastid genome. Tobacco, whose plastid genome has been sequenced in full (Shinozaki et al. 1986), is phylogenetically closely related to tomato (Pichersky et al. 1986), and the overall sequence similarity of the tobacco plastid genome to that of tomato is approximately 98% (J.D. Palmer, personal communication; Pi-

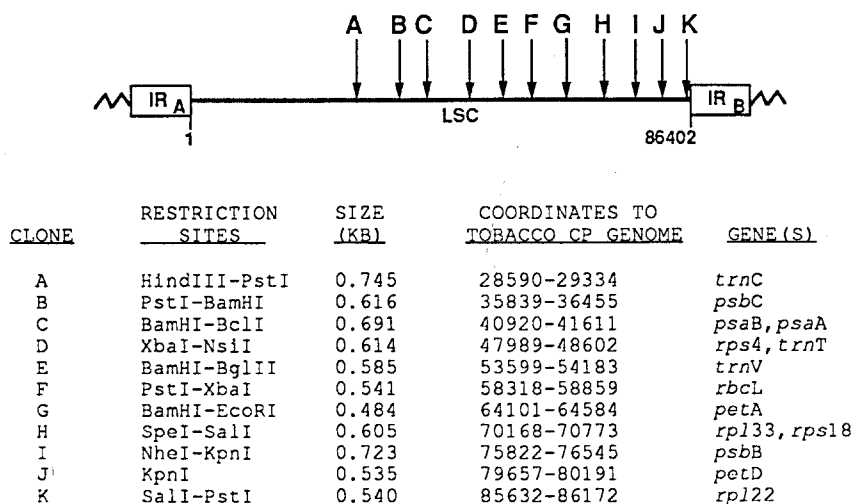


Fig. 1. Location of tobacco plastid DNA clones used for screening of tomato DNA library. Upper portion: schematic representation of a portion of the tobacco chloroplast genome. Position of clones A–K are shown to scale. LSC, Large Single Copy; IR, Inverted Repeat; numbers below represent nucleotide sequence coordinates. Lower portion: detailed description of each clone, including genes or portions of genes contained in each

chersky and Tanksley 1988). Short probes (0.5–0.75 kb) located at 3.1–6.5 kb intervals from each other were chosen (Fig. 1). These probes were used to screen DNA libraries constructed from tomato leaf DNA. To distinguish the few positive recombinant phage containing nuclear fragments with plastid integrates from the majority of positive phage simply containing plastid DNA, we used the following strategy. From each of the plates containing the tomato library phages to be screened, we obtained three nitrocellulose filter “lifts”. While one filter was hybridized with a given probe (probes A–K in Fig. 1), the second filter was hybridized with a sequence upstream of the first probe, and the third filter was hybridized with a sequence downstream of the first probe. For example, to isolate nuclear sequences with homology to probe B in Fig. 1, one filter was hybridized with probe B, the second filter was hybridized with probe A, and the third filter was hybridized with probe C. In phages containing plastid DNA, the entire insert consists of contiguous plastid DNA, but in inserts of nuclear DNA with plastid integrates, only a portion of the insert consists of sequences with homology to plastid sequences. Since each recombinant phage in our genomic libraries contains 12–20 kb of insert DNA, *bona fide* plastid clones hybridizing to a given probe will *always* hybridize either to one or both flanking probes, while a phage containing nuclear DNA with a plastid integrate will only hybridize to the flanking probes if the homology extends to these sequences. In practice, with our choice of probes (Fig. 1) this screening method meant that all plastid integrates shorter than 3 kb (the shortest distance between adjacent probes in Fig. 1) would not hybridize to any of the outside markers and could thus be identified. However, most probe combinations would have allowed the identification of considerably longer integrates (up to 13 kb).

Employing this screening method and the probes described in Fig. 1, we isolated 11 recombinant phages. Using total mitochondrial DNA as a probe, it was determined that nine of these contained mostly mitochondrial DNA (2 isolated with probe C, 1 with probe D, 2 with probe G, 1 with probe H, 1 with probe I, and 2 with probe J; since probes A and K were used only as outside markers, these represent a total of 9 for 9 probes). The length of the plastid integrates in these clones ranged from 3 to 10 kb. These nine clones were not analyzed further in the study reported here.

Characterization of nuclear DNA clones with plastid “integrates”

Two of the recombinant phage did not contain mitochondrial DNA, and in each of the two only a short fragment hybridized to the original probes used in their isolation, and to total plastid DNA. In each case, the region containing the sequence with homology to the plastid DNA was subcloned into plasmid vectors. One clone, designated 2D3, contained a 2.5 kb *Eco*RI fragment with a region hybridizing to probe H (Fig. 2). Another clone, designated 2B2, contained a 3.0 kb *Pst*I fragment with a region hybridizing to probe J (Fig. 2).

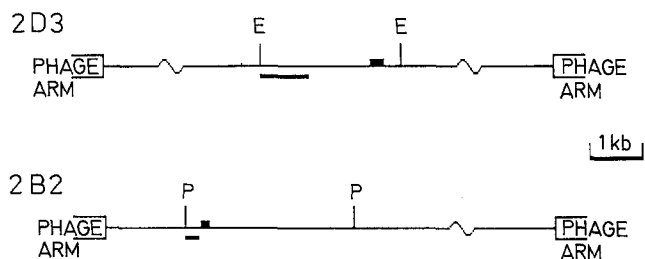


Fig. 2. Partial restriction map of the cloned tomato nuclear DNA fragments with plastid integrates. Only restriction sites flanking the integrates are shown. E, *Eco*RI; P, *Pst*I. An integrate is shown as a bar above the horizontal line. The bar below each drawing indicates the probe used in Southern blots shown in Fig. 3. Wavy line indicates additional DNA not shown. The total length of the insert in both 2D3 and 2B2 is 18 kb

To test whether the cloned DNA was of nuclear origin, and to determine its position on the genetic map, a segment of DNA not including the sequence with homology to plastid DNA was used from each of the two clones (Fig. 2) as a probe in Southern blots of F_2 individuals of the *Lycopersicon esculentum* \times *L. pennellii* cross. These F_2 s have been previously used for mapping numerous restriction fragment length polymorphism (RFLP) markers (Tanksley et al. 1987). Both probes revealed RFLPs segregating in Mendelian ratios, consistent with nuclear locations (data not shown). Analysis of co-segregation with previously mapped markers indicated that DNA from clone 2D3 mapped to the middle of chromosome 9, within 2 cM of the marker TG150, and that DNA from clone 2B2 mapped to the end of chromosome 5, 16.7 cM from the marker CD41.

To determine whether the cloned fragments faithfully represented fragments present in the tomato genome, and to rule out the possibility that the cloned DNAs were the product of rearrangement(s) occurring during the cloning process (such as recombination between different phages), we performed Southern blots of the restricted fragments derived from the cloned DNA and from DNA isolated from leaf tissue (Fig. 3A, B). The same length of fragment was observed for both sources of DNA in each case, demonstrating that no rearrangements took place.

We next determined the nucleotide sequences of the regions in 2D3 and 2B2 with sequence similarity to plastid DNA, and compared these sequences with the corresponding tobacco plastid sequences (Fig. 4A). The integrate in clone 2D3, a total of 202 nucleotides (nt), is derived from the *rps18* gene. The tobacco *rps18* consists of 306 nt; thus, the 2D3 integrate comprises two thirds of the *rps18* gene. Within this region, the 2D3 integrate has a 5 nt insertion and 16 mismatches compared with the tobacco sequence. Some of the mismatches are probably due to the difference between the plastid genomes of tomato and tobacco. However, one nucleotide substitution (at position 109 in Fig. 4A) creates a stop codon, and most likely originated after the transfer of the sequence into the nucleus. Taken together, it is clear that the 2D3 integrate is unlikely to encode a functional protein in the nucleus.

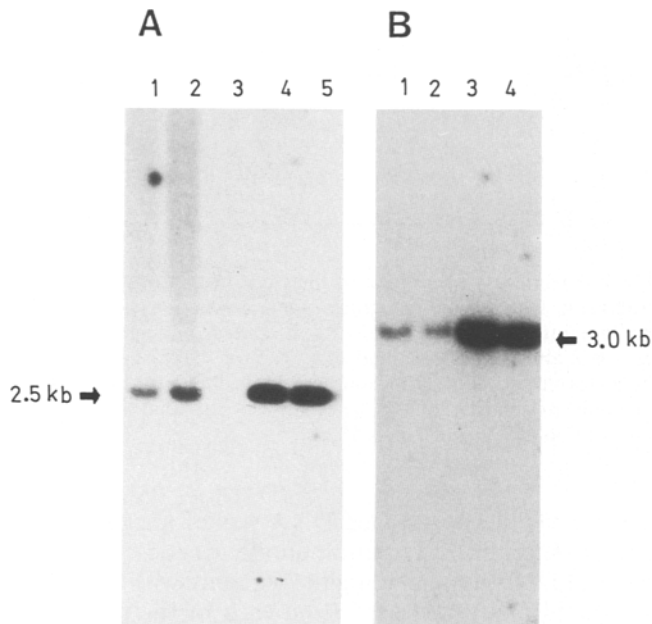


Fig. 3 A and B. Southern blots demonstrating identical size of restriction fragments of cloned and genomic DNA. **A** Lanes 1 and 2, tomato nuclear DNA; lane 3, empty; lanes 4 and 5, DNA from clone 2D3. DNA samples in all lanes were digested with *Eco*RI. The probe was derived from clone 2D3 (see Fig. 2). **B** Lanes 1 and 2, tomato nuclear DNA; lanes 3 and 4, DNA from clone 2B2. DNA samples in all lanes were digested with *Pst*I. The probe was derived from clone 2B2 (see Fig. 2)

The 2B2 clone contained an integrate of 141 nt, with homology to the *petD* gene (Fig. 4B). The tobacco *petD* gene consists of 483 nt (excluding an intron); the 2B2 integrate contains less than 30% of the *petD* sequence. Compared with the corresponding tobacco *petD* sequence, the 2B2 integrate has two insertions of 1 nt each (at positions 65 and 131 in Fig. 4B), a deletion of 4 nt (between positions 82 and 83), and a total of 14 mismatches, not including the deletions and insertions. As with the 2D3 integrate, it is unlikely that the 2B2 integrate encodes a functional protein in the nucleus.

Inspection of the sequences of the integrates and the integration sites reveals few features of note. The two integrates and their immediate surroundings (the sequences shown in Fig. 4) are approximately 70% A + T. In the 2D3 integrate, the 10 nt sequence flanking the 5' end of the integrate has a counterpart (9 matches out of 10) on the same strand inside the integrate near its 3' end (Fig. 4A). Similarly, the 5 nt sequence ATTTC, which flanks the 5' end of the 2B2 integrate, is also found inside the 2B2 integrate close to its 3' end. However, this short sequence is also found twice elsewhere in the 2B2 integrate (Fig. 4B). In addition, other short sequences with some possible homology to each other may be present, such as an inverted repeat of 4 nt also flanking the 2D3 integrate (Fig. 4A). However, the significance of such perceived relationships is not clear.

Discussion

Prevalence, size, and location of integrates

The two integrates whose isolation and characterization we report here are rather short (141 and 202 nt), as was

the first integrate (133 nt) previously identified in the intron of the *Cab7* gene (Pichersky and Tanksley 1988). Thus, our results are in general agreement with those obtained by Scott and Timmis (1984), who estimated from Southern blot analyses that the majority of integrates in the nucleus of spinach were smaller than 2 kb, although the level of resolution of their analysis could not be further improved. Our method of isolating plastid integrates was designed so that we would be able to detect any integrate under 3.1 kb, and in most cases much larger than that (up to 13 kb, which is the distance between some pairs of flanking probes, see Fig. 1), depending on the probes used. Although in this scheme very large integrates could not be detected at all, our failure to isolate any plastid integrate in the nucleus longer than 202 nt suggests that such integrates are rare or non-existent in tomato. In contrast, we did isolate larger (up to 10 kb) plastid integrates in the mitochondrial genome, indicating no technical limitations in our screening procedure to the isolation of integrates of such size.

Our results do conflict with the results obtained with spinach integrates in one respect. Scott and Timmis (1984) estimated from both renaturation experiments and Southern blots (using purified nuclear DNA without prior digestion with *Hpa*II) that each plastid sequence is represented on the average 5 times in the nuclear genome of spinach. In our study, only two of the nine probes we have employed (excluding the two outside markers used only to distinguish integrates from *bona fide* plastid sequences) detected corresponding sequences in the tomato nuclear genome, and these two integrates were each smaller than the probes used to isolate them. Thus, a total of approximately 5 kb of probe (the sum length of the nine probes) identified a total integrate length of approximately 0.4 kb. Although the sample size of the integrates we have obtained, two, is too small to draw any firm conclusion concerning their frequency, our results suggest that only a portion, perhaps even less than a half, of the plastid genome is represented in the nucleus of tomato. One plausible explanation for this discrepancy may be that the nuclear genomes of tomato and spinach are indeed different in regard to the frequency of plastid integrates. Another possible explanation is that although our screening procedure for integrates covered more than 95% of the nuclear genome, the tomato nuclear genome may contain many integrates in "unclonable" regions not present in our tomato DNA libraries.

Recently, du Jardin (1990) has used the Southern blotting technique to demonstrate the existence of both short (<1.0 kb) and long (several kb) plastid integrates in the genome of potato, a close relative of tomato. This researcher employed the methylation-sensitive enzyme *Hpa*II to degrade plastid DNA as the method for purifying nuclear DNA. Some evidence was presented that little nuclear DNA was digested by *Hpa*II. However, if short integrates are found in undermethylated, transcriptionally active nuclear regions, such integrates would have been missed. The only integrate physically isolated and characterized from a species other than tomato has been obtained by Cheung and Scott (1989) from spinach. It is at least 3.4 kb long, but since only part of it is present on the recombinant phage, its full size cannot

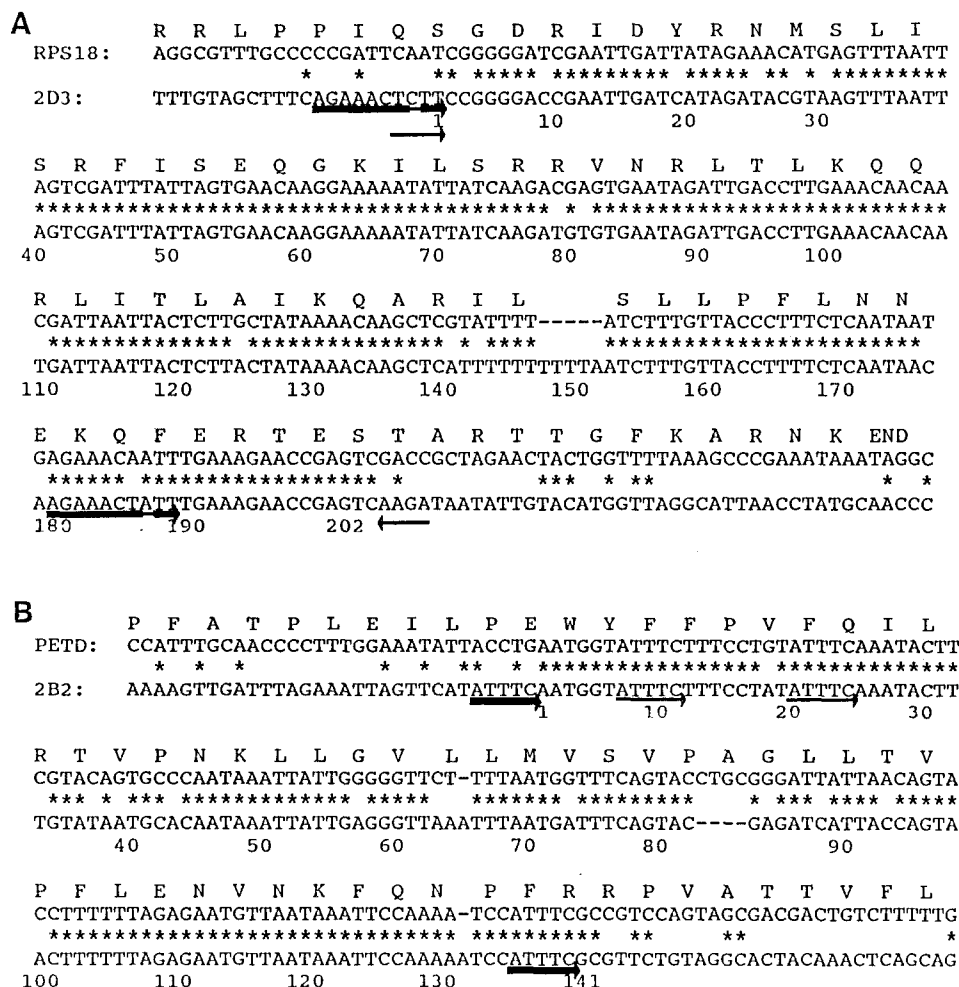


Fig. 4A and B. Sequence of the plastid integrates found in clones 2D3 (**A**) and 2B2 (**B**) and comparisons with the homologous regions from the tobacco plastid DNA. In each panel, the top DNA sequence is that of the tobacco chloroplast (with the sequence of the protein it encodes above it), and the bottom DNA sequence is that of the tomato integrate. An *asterisk* indicates nucleotide identity, a *dash* indicates a deletion/insertion introduced to maximize sequence similarity. Numbering begins with the first perceived nucleotide of the integrate. The tobacco *rps18* sequence begins at position 70617 and ends at position 70817 in the sequence of Shinozaki et al. (1986). The tobacco sequence of the *petD* gene is from 79982 to 80183. *Heavy arrows* indicate direct repeats at or close to the ends of the integrates. *Light arrows* in **A** indicate inverted repeats flanking the integrate. *Light arrows* in **B** indicate additional sequences identical with the direct repeats identified by the *heavy arrows*

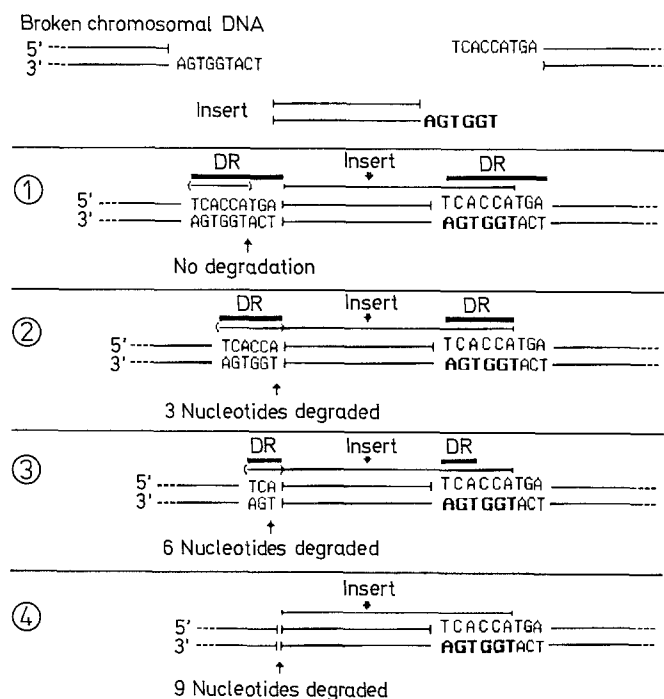


Fig. 5. A model for the integration of plastid DNA into double-stranded breaks of nuclear DNA, and the generation of direct repeats. The nucleotide sequence at the double-stranded break is purely arbitrary, as is its length

be determined. It is possible that the integrate Cheung and Scott isolated comes from a heterochromatin region which is highly methylated. Of the three tomato integrates so far analyzed, two (2B2 and the integrate in the *Cab7* gene) map to ends of chromosomes, and the third, 2D3, maps to the middle of a chromosome. However, the *Cab7* integrate is clearly found in a transcribed region, and both the 2D3 and the 2B2 integrates are tightly linked to several genes (unpublished data). Perhaps the tentative conclusion could be drawn that undermethylated, possibly transcriptionally active, nuclear regions contain short plastid integrates, while hypermethylated regions contain the relatively larger plastid integrates.

Possible mechanism of integration

Although nucleotide sequence determination of plastid sequences integrated into the mitochondrial genome and of the integration sites has not revealed general features such as the direct repeats observed after integration of transposable elements, nevertheless some direct repeats of a different kind were occasionally found (Fejes et al. 1988). Similarly, it was found in the case of the integrate in the *Cab7* gene, that the exact sequence of the 11 nucleotides at its 3' end was directly repeated at its 5' border (Pichersky and Tanksley 1988). Such direct repeats are

different from those generated by transposable elements because the sequence of at least one unit of the repeat is derived from the insertion sequence, whereas direct repeats flanking an integrated transposable element are generated by the duplication of the target DNA (Döring and Starlinger 1986). Based on these observations, it was suggested that homologous recombination is involved in the integration process of plastid DNA into the mitochondrial genome (Fejes et al. 1988), and a general mechanism for the integration of linear DNA fragments into double-stranded breaks of chromosomal DNA has been elaborated (Pichersky and Tanksley 1988; Pichersky 1990). In this model, insert DNA fits into the staggered part of the double-stranded break because it forms a heteroduplex at one end due to fortuitous sequence similarities. It was hypothesized that the other end is blunt-end ligated because it is less likely that the insert will have sequence similarity to the target DNA at both ends. Because a heteroduplex is formed only on one end, while the other end is presumably filled-in and then ligated, a direct repeat, whose basic unit is originally present in both the insert and target DNA, is created (Fig. 5). The observed direct repeat in fact flanks the insert only on one end, while at the other end this sequence is actually part of the insert itself.

Similar direct repeats are observed also in the 2D3 and 2B2 integrates (Fig. 4A, B). In the proposed model of integration, depending on the extent of the single-stranded regions of the chromosomal break and the extent of DNA degradation, if any, of the non-heteroduplexed single-stranded protruding end of the target DNA, the direct repeat units will seem to coincide either exactly or inexactly with the borders of the insert (Fig. 5, panels 1–4). In the 2B2 integrate, direct repeats of 5 nt are observed (Fig. 4B) which fit the example of panel 3 in Fig. 5. In the case of the 2D3 integrate, imperfect direct repeats (Fig. 4A), with 9 out of 10 matches, are observed which also fit the situation illustrated in panel 3 of Fig. 5. Additional evidence for the model may come from the examination of other recent insertions in nuclear DNA and from the demonstration of the activity of enzymes (Pfeiffer and Vielmetter 1988) which catalyze the insertion process in plants.

Acknowledgments. We wish to thank Dr. J.D. Palmer for providing the tobacco chloroplast clones and plastid and mitochondrial DNAs, Dr. R. Fisher for providing a tomato genomic library in the Charon 35 vector, Dr. S.D. Tanksley for providing the tomato F_2 plants and associated database, and Dr. S.L. Baldauf for critical reading of the manuscript. This work was supported by a grant from NSF (DMB-8819210) to EP, and by the NIH National Research Service Award (No. 5-732-GM07544-12) from the National Institute of General Medicine to JML.

References

- Ayliffe MA, Timmis JN, Scott NS (1988) Homologies to the chloroplast DNA in the nuclear DNA of a number of Chenopod species. *Theor Appl Genet* 75:282–285
- Baldauf SL, Palmer JD (1990) Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature* 344:262–265
- Cheung WY, Scott NS (1989) A contiguous sequence in spinach nuclear DNA is homologous to three separated sequences in chloroplast DNA. *Theor Appl Genet* 77:625–633

- Döring HP, Starlinger P (1986) Molecular genetics of transposable elements in plants. *Annu Rev Genet* 20:175–200
- Fejes E, Masters BS, McCarthy DM, Hauswirth WW (1988) Sequence and transcriptional analysis of a chloroplast insert in the mitochondrial genome of *Zea mays*. *Curr Genet* 13:509–515
- du Jardin P (1990) Homologies to plastid DNA in the nuclear and mitochondrial genomes of potato. *Theor Appl Genet* 79:807–812
- Maxam AM, Gilbert W (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol* 65:499–560
- Palmer JD (1985) Comparative organization of chloroplast genomes. *Annu Rev Genet* 19:325–354
- Palmer JD (1990a) Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet* 6:115–120
- Palmer JD (1990b) Plastid chromosomes: structure and evolution. In: Bogorad L, Vasil IK (eds) *The molecular biology of plants. (Cell culture and somatic cell genetics in plants, vol 7)*, Academic Press, Orlando, Florida
- Pfeiffer P, Vielmetter W (1988) Joining of nonhomologous DNA double strand breaks *in vitro*. *Nucleic Acids Res* 16:907–924
- Pichersky E (1990) Nomad DNA – a model for movement and duplication of DNA sequences in plant genomes. *Plant Mol Biol* 15:431–448
- Pichersky E, Tanksley SD (1988) Chloroplast DNA sequences integrated into an intron of a tomato nuclear gene. *Mol Gen Genet* 215:65–68
- Pichersky E, Bernatzky R, Tanksley SD, Breidenbach W, Kausch A, Cashmore AR (1985) Isolation, molecular characterization, and genetic mapping of two clusters of genes encoding chlorophyll a/b binding proteins in *Lycopersicon esculentum* (tomato). *Gene* 40:247–258
- Pichersky E, Bernatzky R, Tanksley SD, Cashmore AR (1986) Evidence for selection as a mechanism in the concerted evolution of the *Lycopersicon esculentum* (tomato) nuclear genes encoding the small subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase. *Proc Natl Acad Sci USA* 83:3880–3884
- Pichersky E, Hoffman NE, Malik VS, Bernatzky R, Tanksley SD, Szabo L, Cashmore AR (1987) The tomato *Cab-4* and *Cab-5* genes encode a second type of CAB polypeptides localized in PSII. *Plant Mol Biol* 9:109–120
- Pichersky E, Tanksley SD, Piechulla B, Stayton MM, Dunsmuir P (1988) Nucleotide sequence and chromosomal location of *Cab-7*, the tomato gene encoding the Type I chlorophyll a/b-binding polypeptide of photosystem I. *Plant Mol Biol* 11:69–71
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467
- Scott NS, Timmis JS (1984) Homologies between nuclear and plastid DNA in spinach. *Theor Appl Genet* 67:279–288
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Cunwongse 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, Tohdon N, Shimoda H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J* 5:2043–2049
- Sugita M, Manzara T, Pichersky E, Cashmore AR, Gruissem W (1987) Genome organization, sequence analysis and expression of all five genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from tomato. *Mol Gen Genet* 209:247–256
- Tanksley D, Miller J, Paterson A, Bernatzky R (1987) Molecular mapping of plant chromosomes. In: Gustafson JP, Appels RA (eds) *Chromosome structure and function*. Plenum Press, New York, pp 157–173
- Timmis JN, Scott NS (1983) Spinach nuclear and chloroplast DNAs have homologous sequences. *Nature* 305:65–67
- Weeden NF (1981) Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J Mol Evol* 17:133–139

Communicated by E. Meyerowitz