
Gene conversion as a mechanism for divergence of a chloroplast tRNA gene inserted in the mitochondrial genome of *Brassica oleracea*

M.Dron, C.Hartmann, A.Rode and M.Sevignac

Laboratoire Biologie Moléculaire Végétale, Bâtiment 430, Université Paris Sud 91405, Orsay, France

Received 23 August 1985; Revised 15 October 1985; Accepted 25 October 1985

ABSTRACT

We have characterized a 1.7 kb sequence, containing a tRNA Leu2 gene shared by the ct and mt genomes of *Brassica oleracea*. The two sequences are completely homologous except in two short regions where two distinct gene conversion events have occurred between two sets of direct repeats leading to the insertion of 5 bp in the T loop of the mt copy of the ct gene. This is the first evidence that gene conversion represents the initial evolutionary step in inactivation of transferred ct genes in the mt genome. We also indicate that organelle DNA transfer by organelle fusion is an ongoing process which could be useful in genetic engineering.

INTRODUCTION

The presence of promiscuous DNA (1) in organelle genomes is now accepted as a general phenomenon (2). Most data (3-5) indicates that the DNA movement occurred long ago and, during evolution, the transferred sequences have diverged by point mutations and recombination processes.

In plants; DNA transfer has been observed between chloroplast and mitochondrial genomes of a wide variety of species (6,7). The movement has always been found unidirectionally from chloroplasts to mitochondria (2,6,7). Stern and Palmer (7) propose two types of mechanisms of interorganellar sequence transfer. The first one involves possible physical membrane interactions between the two organelles resulting in the transformation of the mitochondrion by the chloroplast, the mitochondrial genome would be modified by intermolecular recombination. The second one concerns the existence of unidirectional transfer vectors as transposable elements or transducing phages. In such a case, recombination will occur by site specific recombination between two highly homologous sequences. Stern and Palmer (7) feel it also unlikely that these integrated sequences play any biological role in the mitochondria because they are related to genes which are not known to function in the mitochondria rbcL,16S ct rRNA gene. These hypotheses have not been demonstrated yet.

Here we present the characterization of a 1.7 kb ctDNA fragment located in the mitochondrial genome of Brassica oleracea. Part of this fragment containing a tRNA Leu 2 gene has been sequenced. The two chloroplast and mitochondrial sequences share complete homology except in two very short regions inside and close to the tRNA gene. We discuss the mechanisms of integration and evolution of the chloroplast sequence in the mitochondrial genome. These modifications probably lead to the inactivation of the transferred tRNA gene.

METHODS

Ct DNA was extracted and the ct Sal I fragment 3 cloned as described (8, 9). Mt Sal I fragment 3 was subcloned from a cosmid containing mt Sal I fragments 3, 8, 12, 19 (8). Restriction analysis, gel electrophoresis, ligation, E. coli transformation, nick translation and DNA/DNA hybridizations were as described (9). Maps were constructed by ³²P end labelling (10) or by single and double restriction enzyme digestions. For Maxam-Gilbert sequencing of mt and ct homologous fragments, DNA was 3' end labeled by Klenow enzyme, the strands separated on 5% denaturing polyacrylamide gels, electroeluted and used for chemical cleavages (11). Sequences were analysed with a Hewlett-Packard computer.

RESULTS AND DISCUSSION

A 12 kb region located in the inverted repeat of the ct DNA molecule has been detected in the mt DNA of a wide variety of plants (6,7). However, only 1.7 kb of this ct DNA sequence is present in the mt genome of B. oleracea (Fig.1); this segment corresponds to the left part of the ct inverted repeat which, in maize, contains a tRNA Leu 2 gene (12). This is also the situation in B. oleracea because a clone of this ct gene from Vicia faba (gift of A. Steinmetz) strongly hybridized to a 500 bp Hae III/Ava II fragment from the 1.7 kb sequence common to mt and ct DNAs (Fig.1).

In Fig.2, the nucleotide sequence of the B. oleracea ct tRNA Leu2 gene (non coding strand) and flanking region are presented and compared to the homologous B. oleracea mt copy. The Brassica ct and mt sequences are highly homologous: 100% from bases 1-103 (except one substitution at position 66) and 100% from position 196 to the end. However, 13 bp modifications can be detected between bases 105-196; 7 between positions 105-113 and two insertions in the mt sequence at position 181 (4bp) and position 202 (1bp).

A mechanism can be put forward to explain these differences such as two gene conversion steps would have been necessary to obtain the mt

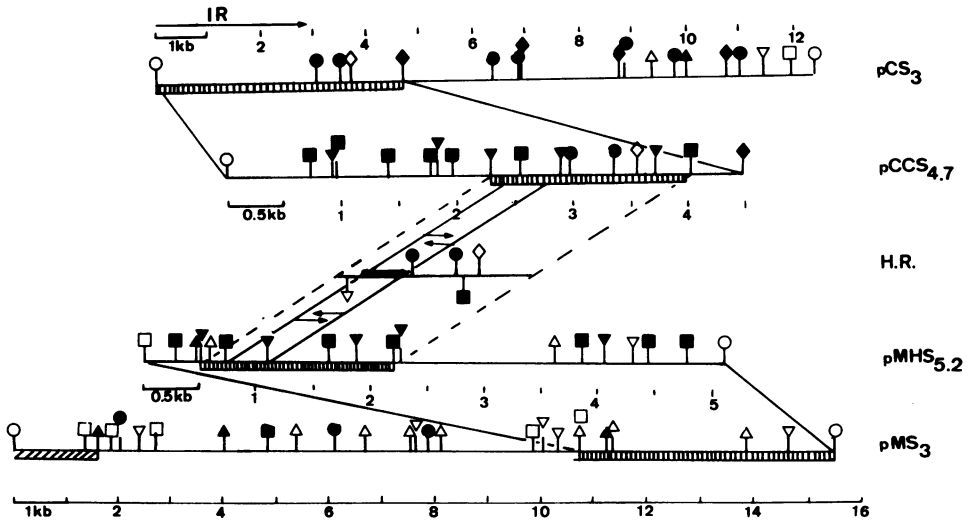


Fig.1 Physical maps and precise localization of the 1.7 kb common sequence on restriction fragment Sal 3 (30) from *B.oleracea* ct genome (pCS3) and Sal 3 (8) from *B.oleracea* mt genome (pMS3): pCS3 was cloned in pBR322 while pMS3 is a subclone from a larger recombinant in cosmid pHc 79 (Hartmann et al submitted). The homologous region (HR) has been determined by hybridization (pC against pM and reversibly) followed by subcloning of the Sal I/ Cla I fragment of pCS₃ (termed pCCS_{4.7}) and of the Sal I/ Hind III fragment of pMS₃ (termed pMHS_{5.2}). Restriction sites listed on the upper side of the HR line are present in pCCS_{4.7} but not in pMHS_{5.2} while the sites on the lower side are specific to pMHS_{5.2}. The thick region of the HR line corresponds to the common Hae III/ Ava II fragment which hybridized to the ct *Vicia faba* tRNA Leu2 gene. The ct and mt 500 bp Hae III/ Ava II fragments were subcloned in pHp 34 (31) and sequenced as shown by arrows. Scale is presented for each fragment. Homologous ct/mt region is represented by ||||| (HR). Mt 26S rRNA gene is represented by ////. IR represents the beginning and direction of the ct inverted repeat. Restriction enzymes are as follows: Sal I ○, Bam H I ●, Cla I ◆, Hind III □, Pst I ▽, Xho I △, Ava II ▼, Hae III ■, Sma I ▲, Bgl I ◇.

structure (mt) from the original ct arrangement (Fig. 3). They would involve two sets of direct repeats (A,B), each one containing sequences labeled as 1', 2, 4 and 5, and three individual sequences, labeled as 1, 3 and 6. The recombination mechanism should probably be intermolecular rather than intramolecular because the two sets of direct repeats (Fig.3,A and B) are too close to have enabled sufficient DNA folding. The first step would have been probably the insertion of the original 1.7kb ct sequence into the mt genome. We propose that divergence has occurred according to the following process: two mt DNA molecules would have been

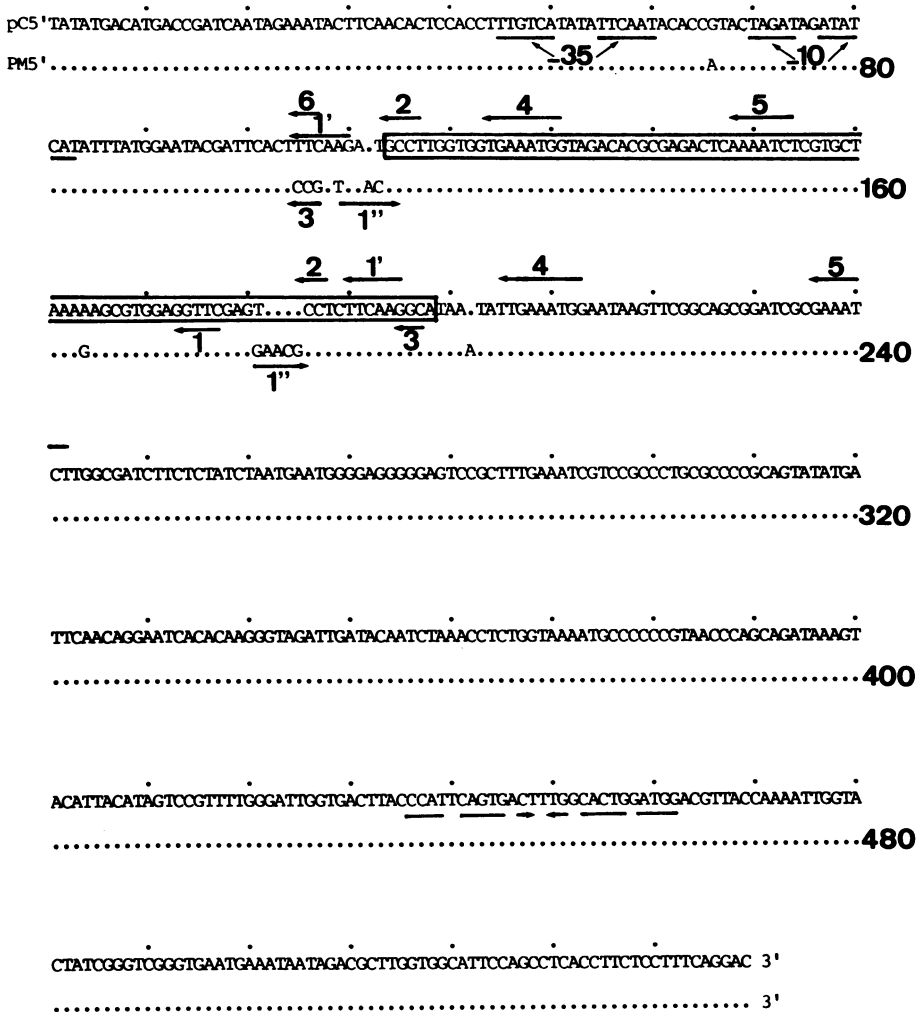


Fig.2 Nucleotide sequence of the *B. oleracea* ct (pC) 542 bp *Hae* III/ *Ava* II restriction fragment. The position of the tRNA *Leu2* is boxed. Known ct signals for transcription initiation are underlined. Divergences occurring in the *B. oleracea* mt (pM) sequence are reported in the lower line. The direct and inverted repeated sequences are numbered and indicated by arrows. Scale is labeled for the mt non coding strand.

mispaired so that the repeat B of one mt molecule can be paired with the repeat A of another mt molecule. This heteroduplex would have been submitted to gene conversion resulting in an hypothetical intermediate (int) arrangement (fig.3). Then, two int mt DNA molecules would have been

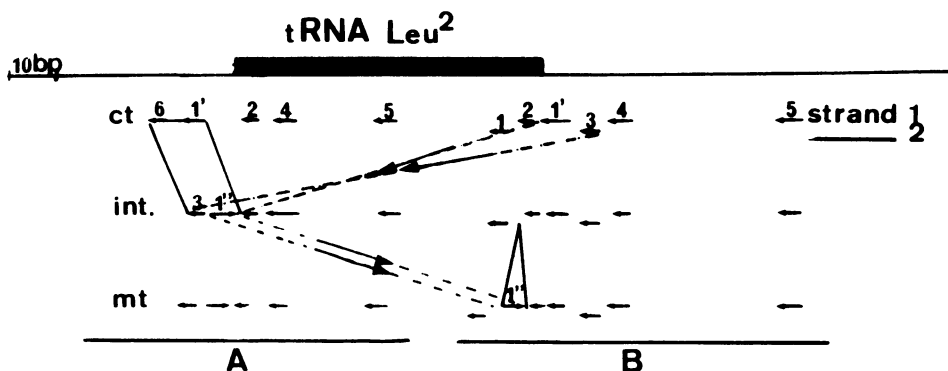


Fig.3 Arrangement of the two blocks of directly repeated sequences lying inside or close to the *B. oleracea* tRNA Leu2 gene. The black box represents the tRNA Leu2 gene. Ct represents the ct DNA arrangement, mt represents the mt arrangement and int represents an hypothetical intermediate structure. Each direct repeat (A and B) is composed of sequences labelled as 1, 2, 4 and 5. Sequences labelled 1, 3 and 6 are unique and involved in the gene conversion process. The two DNA strands (strand 1, strand 2) are indicated. Lines between 6, 1' (ct) and 3, 1' (int) or between int and 1' (mt) correspond to the target sites while the broken lines with arrows correspond to the donor sequences of the gene conversion process. The composition of sequences numbered 1 to 6 are given in Fig. 2. The mechanism of gene conversion between ct and int requires a double stranded ct break.

mispaired again; sequence 1" from repeat A was used as donor to target sequences 1 or 1' of repeat B leading in the present mt DNA structure (fig.3). The mechanism leading to apparent gene conversion may have been foldback during replication or something else.

This gene conversion mechanism (Fig. 3) creates inverted (1', 1'') and direct (1''; 1'') repeats which would be able to delete parts of the common 1.7 kb sequence (13). The insertion of 5 bp (GCAAG) in the mt copy of the ct gene tRNA Leu2 is the principal consequence of these two gene conversion steps. These 5 bp have been inserted at an invariable pyrimidine (T) (14) in the T loop which is invariable in size. We therefore consider tRNA Leu2 gene is now a pseudogene in *Brassica* mitochondria for the following reasons :

1) No tRNA gene introns resemble the one reported here. Would it be an intron, it would be the shortest so far reported and the only one to be located in the T loop. Two types of tRNA introns have been previously reported : nuclear tRNA introns (15) are short (14 - 65 bp) and always located in the anticodon loop sequence while cytoplasmic (ct or mt) tRNA

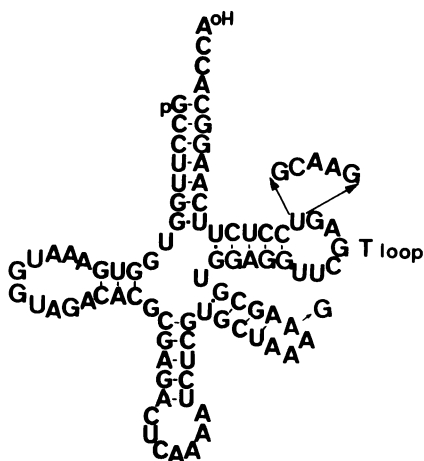


Fig.4 Clover leaf structure of the *B.oleracea* ct tRNA Leu2. The differences with the proposed structure of the homologous mt tRNA Leu2 are indicated by arrows outside the cloverleaf structure.

introns are longer (≥ 500 bp) and generally inserted in the anticodon sequence (16-18).

2) Except in human mitochondria (19), the T loop is invariable in size (7 bp) and always contains a pyrimidine at the last position (U in ct tRNA Leu2). Recently published plant mt tRNA sequences correspond to this universal scheme (20-22).

3) *Vicia faba* tRNA Leu 2 gene environment (23) is completely different from that of maize or *Brassica* (this paper); however the structure of the gene has been maintained, it is expressed in all these chloroplasts (24). In the contrary after its insertion in the mt genome, the first modification between the ct and mt sequences is an insertion in the tRNA gene: in mitochondria, there is no selective pressure to maintain the tRNA gene active.

Although these three arguments favour a pseudogene, the transcription initiation signals (-35 and -10) have not been modified in the mt copy of the ct gene (Fig.2) suggesting that it may still be transcribed. Suitable *in vivo* and *in vitro* experiments with, as yet uncharacterized, plant mt RNA polymerases will be necessary to determine if the mt copy of the gene is expressed but unable to function.

Our data, and the finding that maize mt and ct tRNA His genes are 100% homologous (25), suggest that DNA transfer between chloroplasts and

mitochondria are recent, ongoing events because insufficient sequence diversity has occurred to calculate an evolutionary divergence rate (3,5). Because the common ct and mt 1.7 kb sequence is located 11 kb from the ct 23S rRNA gene and 9 kb from the mt 26S rRNA gene, it is possible that, after organelle fusion, the two genomes were aligned by the highly homologous ($\geq 70\%$) rRNA genes (26,27) and the common sequence transferred by gene conversion. This 1.7 kb sequence then diverged by specific gene conversions in the mt copy of the ct sequence (the direct repeats indicated in Fig.3 are unmodified in the ct genome but modified in the mt genome). These common ct and mt sequences may account, in part, for the large variable size of the plant mt genome (28). There are no reports of mt sequences being transferred to ct genomes indicating that ct genomes have an invariable size which cannot accommodate foreign DNA insertions. It is likely that DNA modifying and repair mechanisms are different in mt and ct and that the enzymes involved in mt recombination are site specific unlike the long homologous pairing mechanism of the recA protein system (29), at least in this special case dealing with ct sequences inserted in the mt DNA molecule.

A collection of common mt and ct sequences, an understanding of organelle DNA transfers and a knowledge of mt DNA modifying and repair enzymes may enable the manipulation and expression of foreign genes in plant mitochondria.

Acknowledgements: This work was supported by a C.N.R.S ATP to M.D. Authors thank R.KEMBLE, J.L.ROSSIGNOL, M.SCHNEIDER, F.QUETIER for critical reading of the manuscript.

REFERENCES

1. Ellis, J., Nature (1982) 299, 678-679
2. Timmis, J.N and Scott, N.S (1984) Trends in Biochem. Sci., 271-273
3. Jacobs, H.T, Posakony, J.W., Grula, J.W., Roberts, J.W., Xin, J.H., Britten, R.J. and Davidson, E.H. (1983) J.Mol.Biol. 165, 609-632
4. Fareilly, F and Butow R.A (1983) Nature 301, 296-301
5. Nomiya, H., Fukuda, M., Wakasugi, S., Tsuzuki, T. and Shimada, K. (1985) Nucleic Acids Res. 13, 1649-1657
6. Stern, D.B and Lonsdale, D.M (1982) Nature 299, 698-702
7. Stern, D.B and Palmer, J.D (1984) Proc.Natl.Acad.Sci.USA 81, 1946-1950
8. Chetrit, P., Matthieu, C., Muller, J.P. and Vedel, F (1984) Curr.Genet. 8, 413-421
9. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Cold Spring Harbor Lab. Press (N.Y)
10. Smith, H.O and Birnstiel, M.L (1976) Nucleic Acids Res. 3, 2387-2398
11. Maxam, A. and Gilbert, W. (1980) Meth. in Enzymol. 65, 499-560
12. Steinmetz, A., Krebbers, E.T., Schwarz, Z., Gubbins, E.J. and

Nucleic Acids Research

- Bogorad, L. (1983) *J. Biol. Chem.* 258, 5503-5511
13. Kourilsky, P (1983) *Biochimie* 65, 85-93
 14. Rich, A and Kim, S.H (1978) *Sci. Am.* 238, 52-73
 15. Peebles, C.L, Gegenheimer, P and Abelson, J (1983) *Cell* 32, 525-536
 16. Koch, W., Edwards, K and Kossel, H. (1981) *Cell* 25, 203-213.
 17. Deno, H., Kato, A., Shinozaki, K and Suciura, M. (1982) *Nucleic Acids Res.* 10, 7511-7520.
 18. Deno, H and Suciura, M (1984) *Proc. Natl. Acad. Sci. USA.* 81, 405-408.
 19. Anderson, S., Bankier, A.T., Barrell, B.G., deBruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Shreier, P.H, Smith, A.J.W., Staden, R and Young, I.G (1981) *Nature*, 290, 457-465
 20. Parks, T.D., Dougherty, W.G., Levins III, C.S. and Timothy, D.H. (1984) *Plant Physiol.* 76, 1079-1082
 21. Gottschalk, M and Brennicke, A (1985) *Curr. Genet.* 9, 165-168
 22. Marechal, L., Guillemaut, P., Grienberger, J.M., Jeannin, G. and Weil, J.H. (1985) *Febs Lett.* 184, 289-294.
 23. Bonnard, G., Weil, J.H. and Steinmetz, A. (1985) *Curr. Genet.* 9, 417-422.
 24. Osorio-Almeida, M.L., Guillemaut, P., Keith, G., Canaday, J., Weil, J.H. (1980) *Biochem. Biophys. Res. Commun.* 92, 102-108.
 25. Iams, K.P, Heckman, J.E and Sinclair, J.H (1985) *Plant Mol. Biol.* 4, 225-233.
 26. Spencer, D.F, Schnare, M.W and Gray, M.W (1984) *Proc. Natl. Acad. Sci. USA.* 81, 493-497
 27. Stern, D.B, Hodge, T.P and Lonsdale, D.M (1984) *Plant Mol. Biol.* 3, 355-361.
 28. Quetier, F., Lejeune, B., Delorme, S. and Falconet, D. (1985) in *Encyclopedia of Plant Physiol. new series*, vol. 18, 25-36.
 29. Howard-Flanders, P., West, S.L. and Stasiak, A. (1984) *Nature* 309, 215-220.
 30. Vedel, F. and Matthieu, C. (1983) *Curr. Genet.* 7, 13-20.
 31. Prentki, P. and Krisch, H. (1982) *Gene* 17, 189-196.