Genomic Recombination of the Mitochondrial atp6 Gene in Arabidopsis thaliana at the Protein Processing Site Creates Two Different Presequences

Joachim Marienfeld, Michael Unsell, Petra Brandt, and Axel Brennicke*
Institut für Genbiologische Forschung, Ihnestraße 63, D-14195 Berlin and Allgemeine Botanik, Universität Ulm, Albert-Einstein-Allee, D-89069 Ulm, Germany

(Received 2 October 1996)

Abstract

In the mitochondrial genome of the flowering plant Arabidopsis thaliana the atp6 open reading frame is located on the border of one of the repeats resulting in two copies with different presequence extensions. The two presequences of 135 and 97 amino acids respectively show no similarity to each other, while the mature protein sequences are identical. Both preproteins are most likely synthesized in Arabidopsis mitochondria from promoter elements upstream of each copy. The presence of two arrangements in the mitochondrial genome of fertile Arabidopsis plants suggests this recombination to be unrelated to a cytoplasmic male sterile phenotype. This recombination precisely at the mature protein terminus is reminiscent of the domain shuffling model in protein evolution.

Key words: Arabidopsis; mitochondria; atp6; recombination

1. Introduction

Subunit 6 of the F1-F0 ATPase complex in mitochondria (ATP6) is in most eukaryotes encoded in the mitochondrial genome, but is nevertheless usually synthesized with a presequence which is cleaved off during protein maturation. While the mature part of the protein is highly conserved among plants, fungi and animals, the respective presequences show little similarity even between closely related plant species. The precise function of this presequence is at present still unclear, but it could be required for targeting the subunit to the mitochondrial inner membrane or it may be necessary at a later stage for assembly of the different subunits into the protein complex.

Another example of a usually mitochondrially encoded protein with a presequence is subunit 2 of complex IV (COX2), from which the presequence is likewise cleaved off during maturation of the preprotein. In this instance the presequence appears to be involved in the transport of the preprotein to the outside of the inner mitochondrial membrane. In analogy the ATP6 presequence may serve a similar transport function. The function of the ATP6 protein presequence, be it targeting, membrane insertion or assembly, appears to be supported by many different primary sequence structures, which meet one or more as yet unclear, apparently not very stringent, but necessary parameter(s).

In virtually all plant mitochondrial genomes sequence regions of several kilobases are found duplicated, the two copies usually being involved in homologous recombination events. These duplications occur apparently irrespective of coding regions and often duplicate genes or parts of them. In several plant species such as rice, sorghum and some Brassicaceae, the comparatively small atp6 gene is closely connected to or even resides in such repeated sequences. In the Arabidopsis mitochondrial genome the atp6 gene is located at the border of a repeated sequence region, the recombination occurring precisely at the protein processing site.

2. Materials and Methods

The nucleotide sequences of the atp6 genomic arrangements in Arabidopsis thaliana var. Columbia were determined as part of the Arabidopsis mitochondrial genome project by non-radioactive fluorescent technology on ordered cosmids and their subclones as to be detailed elsewhere (M.U., J.M., P.B.; A.B., in preparation). Active recombination across the repeated sequences was identified in the physical map of the mitochondrial DNA. Nucleic acid technology followed standard protocols.
Mitochondrial ATP6 Recombination in Arabidopsis

3. Results and Discussion

3.1. Two different presequences for the ATP6 protein

In the Arabidopsis mitochondrial genome the ATP6 gene is located at the border of one of the two recombination-ally active repeats (Fig. 1). The recombination occurs precisely at the border of the mature protein sequence and results in two genes with identical ATP6 mature proteins, but with completely different presequences. One of the two N-terminal extensions shows significant similarity to the presequences in other plant species, including also the monocot maize (At-1; Fig. 2), suggesting that this arrangement represents an ancestral gene organisation. The second presequence in At-2 shows no similarity to any other ATP6 presequence in the plants analysed to date (data not shown). In At-2 the only in frame ATG-methionine codon is found 96 amino acids upstream of the mature protein terminus SPL, the closest ATG is located only 7 codons before the processing site. In one of the two ATP6 genes in Brassica napus the only in frame ATG is found 8 triplets upstream of the processing site suggesting that such a short presequence could be sufficient for a functional ATP6 precursor protein.
3.2. RNA editing and transcription

In the Arabidopsis atp6 open reading frame one RNA editing site is predicted to alter codon 26 from specifying the amino acid proline to code for either leucine or serine as in other plants (Fig. 3). This editing event is also the only change observed in Brassica napus atp6 transcripts, while in some radish lines no RNA editing has been observed in atp6 transcripts. The atp6 transcripts in Brassicaceae thus represent the least edited conserved mRNAs in plant mitochondria observed to date.

The significance of the highly conserved atp6 core protein is further highlighted by the precise carboxy terminus introduced in many plant species by RNA editing and predicted in others such as the last four species of Observed in many plant species by RNA editing transcription initiation in dicot plants.

Figure 3. Alignment of the mature ATP6 protein sequences deduced from cDNA and genomic sequences. Genomic encoded amino acids identical to Arabidopsis are represented by dots. Amino acids identical as result of RNA editing are shown as exclamation marks (!). Silent RNA editing sites are underlined. Positions changed by RNA editing to amino acids not conserved in Arabidopsis are in bold face and the vertical arrow (?) shows the likely RNA editing site in Arabidopsis deduced from this alignment. Missing amino acids are indicated by dashes. Amino acid sequences are deduced from edited cDNA sequences of Brassica napus (B.n.; acc.no. X58276), Orothamnus beterina (O.b.; acc.no. Y00485), Sorghum bicolor (S.b.; acc.no. X52161) and Zea mays (Z.m.; acc.no. Z11843). The sequence of Oryza sativa (O.s.; acc.no. X58276) is only in the carboxy-terminal part based on cDNA. The ATP6 protein sequences of Arabidopsis thaliana (A.t.1 and A.t.2; acc.nos. Y08501 and Y08502 resp.), Raphanus sativus (R.s.; acc.no. J04945), Triticum aestivum (T.a.; acc.no. M24804), Vicia faba (V.f.; acc.no. X54285), Glycine max (G.m.) and Nicotiana tabacum (N.t.; acc.no. S00651) are based on genomic DNA. Numbering begins at the mature part of the ATP6 polypeptide. Lengths of the amino-terminal extensions are indicated in brackets. Termination codons are shown as asterisks.

3.3. Recombination involving atp6 genes in other plants

Genomic recombination involving the atp6 gene is also observed in a number of other plant species (Fig. 1B), suggesting the atp6 gene contains sequences preferentially recognized by the as yet unknown homologous recombination activities in plant mitochondria. Alternatively the relatively small size of the reading frame coding the mature protein and the deduced relaxed constraints on the presequence identity allow the atp6 genes to tolerate its involvement in recombination events, which would more often than not disrupt and inactivate larger genes.

In some Brassicaceae other than Arabidopsis likewise two atp6 gene arrangements have been observed. Between fertile and sterile radish (Raphanus sativus), recombination has occurred as in Arabidopsis at the border between presequence and mature protein and likewise results in presequences largely unrelated to each other. Recombination in Raphanus is however no more active in this genome, but one arrangement is dominant in the male sterile Ogura cytoplasm, while the other is the major gene in the fertile cytoplasm.

In some lines of sorghum two atp6 genes have been identified in mitochondrial genomes, while in other cytoplasms of this species only one or the other of the two genes is present. Like in Arabidopsis one of the two presequences is similar to those in other plant species with only one atp6 gene, whereas the other is of unknown origin. The divergence between the two genes is as in Arabidopsis at the mature protein terminus, but the coding
sequence in sorghum is not involved in an active homologous recombination.2

In sunflower, Helianthus annuus, recombination occurs close to the transition from presequence to mature protein, at or within the last codon of the presequence.9 Duplication of the atp6 gene in soybean alters the upstream 5’ nucleotide sequences, but maintains identical pre- and mature protein sequences.10 In rapeseed, Brassica napus, mitochondrial recombinations in the transition from fertile to a male sterile cytoplasms have altered also parts of the respective upstream sequences, but have left conserved the entire protein structure.7

In rice, particularly sequences downstream of the atp6 gene are affected by sequence rearrangements again between different sterile and fertile cytoplasms, while the atp6 coding region itself is not affected.11,12

3.4. Functional conclusions

The recombination site in Arabidopsis precisely at the protein processing site suggests that upstream amino acid identities have little if any influence on the processing event itself, but suggest that rather the conserved mature protein terminus SPL... may be sufficient to trigger the specific protease. Comparison of the different pre-sequences now available in different plant species (Fig. 2 and data not shown) does not reveal any crucial parameter that might specify the function of this presequence in e.g. targeting the protein to a specific mitochondrial location such as a certain orientation in or at the inner membrane.

The recombination event at the mature protein terminus in Arabidopsis is suggestive of the model of evolution by modules,13 in which certain unrelated protein domains are linked by genomic recombinations (often via introns) in different permutations and combinations to create novel gene functions much faster than by completely de novo mutational creation of functional genes from scratch.

The presence of two distinct atp6 genes in male fertile Arabidopsis with different presequences suggests that this gene and its various guises of upstream and downstream sequence organisation in different cytoplasms should not be crucially involved in the cytoplasmic male sterile phenotype.7,8,11,12,14 Rather, these atp6 recombinations are most likely byproducts of genomic recombinations, which occur as frequently between fertile and sterile genomes as they are found within fertile mitochondrial DNAs.

Acknowledgements: This work was supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie in the Genzentrum Berlin.

References

1. Bonen, L. and Bird, S. 1988, Sequence analysis of the wheat mitochondrial atp6 gene reveals a fused upstream reading frame and markedly divergent N termini among plant ATP6 proteins, Gene, 73, 47–56.
2. Mullen, J. A., Pring, D. R., Kempken, F., Ferguson, J., and Chase, C. D. 1992, Sorghum mitochondrial atp6: divergent amino extensions to a conserved core polypeptide, Plant Mol. Biol., 20, 71–79.
3. Pratje, E., Mannhaupt, G., Michaelis, G., and Beyreuther, K. 1983, A nuclear mutation prevents processing of a mitochondrially encoded membrane protein in Saccharomyces cerevisiae, EMBO J., 2, 1049–1054.
4. Palmer, J. D., and Shields, C. R. 1984, Tripartite structure of the Brassica campestris mitochondrial genome, Nature, 307, 437–440.
5. Klein, M., Eckert-Ossenkopp, U., Schmiedeberg, I., Brandt, P., Unseld, M., Bremnicke, A., and Schuster, W. 1994, Physical mapping of the mitochondrial genome of Arabidopsis thaliana by cosmid and YAC clones, Plant J., 6, 447–455.
6. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, Molecular cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
7. Handa, H. and Nakajima, K. 1992, Different organization and altered transcription of the mitochondrial atp6 gene in the male-sterile cytoplasm of rapeseed (Brassica napus L.), Curr. Genet., 21, 153–159.
8. Krishnasamy, S., Grant, R. A., and Makaroff, C. 1994, Subunit 6 of the F0-ATP synthase complex from cytoplasmic male-sterile radish: RNA editing and NH2-terminal protein sequencing, Plant Mol. Biol., 24, 129–141.
9. Spassova, M., Moneger, F., Leaver, C. J., Petrov, P., Atanassov, A., Nikjamp, H. J., and Hille, J. 1994, Characterisation and expression of the mitochondrial genome of a new type of cytoplasmic male-sterile sunflower, Plant Mol. Biol., 26, 1819–1831.
10. Grabau, E., Havlik, M., and Gesteland, R. 1988, Chimeric organization of two genes for the soybean mitochondrial ATPase subunit 6, Curr. Genet., 13, 83–89.
11. Kadowaki, K., Suzuki, T., and Kazama, S. 1990, A chimeric gene containing the 5’portion of atp6 is associated with cytoplasmic male-sterility of rice, Mol. Gen. Genet., 224, 10–16.
12. Iwabuchi, M., Kyoizuka, J., and Shimamoto, K. 1993, Processing followed by complete editing of an altered mitochondrial atp6 RNA restores fertility of cytoplasmic male sterile rice, EMBO J., 12, 1437–1446.
13. Long, M., De Souza, S. J., Rosenberg, C., and Gilbert, W. 1996, Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor, Proc. Natl. Acad. Sci. USA, 93, 7727–7731.
14. Dewey, R. E., Timothy, D. H., and Leavings III, C. S. 1991, Chimeric mitochondrial genes expressed in the C male-sterile cytoplasm of maize, Curr. Genet., 20, 475–482.