Sequencing of a Segment of a Monilophyte Species Mitochondrial Genome Reveals Features Highly Similar to those of Seed Plant mtDNAs

S. Panarese, G. Rainaldi, C. De Benedetto and R. Gallerani*

Department of Biochemistry and Molecular Biology, University of Bari, via Orabona 4, 70126 Bari, Italy

Abstract: A continuous sequence of 20,374 bp has been produced corresponding to an equivalent region of the mitochondrial genome of the fern Asplenium nidus. The information content of this sequence includes: several segments from chloroplast origin, three tRNA genes of probable native type, a couple of inverted repeats, three protein genes and a segment of a fourth.

Among the tRNA genes trnN, usually from chloroplast origin in the Spermatophyte mitochondrial genomes, shows the characteristics of a native gene.

Keywords: mtDNA, Plant, Monilophytes, Asplenium nidus.

INTRODUCTION

Mitochondrial genomes of land plants have been fully sequenced and characterized in several species belonging to the Briophytes (Marchantia polymorpha [1] and Physcomitrella patens [2] and Spermatophytes (Arabidopsis thaliana [3], Beta vulgaris [4], Oryza sativa [5], Brassica napus [6], Zea mays [7], Nicotiana tabacum [8] and Triticum aestivum [9]).

The comparison of organization, structure and expression between Spermatophyte mitochondrial genomes reveals several homogeneous features which can be summarized as follows: i) the presence of repeated sequences, ii) a heterogeneous structure, iii) the presence of DNA segments of extra mitochondrial origin (mainly chloroplastic) carrying in some cases active genes (usually for tRNAs) [10-12], iv) the editing of transcription products of structural genes v) an incomplete set of tRNA genes. On the contrary not all of these features can be considered peculiar properties of Briophyte mitochondrial genomes, in particular: RNA editing, active in some species but not in all [13], absence of homologous recombination events; absence of incorporation of foreign genetic information; absence of chloroplast DNA insertions [14,15].

As far as the mtDNA of Monilophyte plant species is concerned, very little is known about their structure and organization and the expression products of genes encoded on them. For these plants the sequences of only a few characteristic mitochondrial genes or of incomplete parts of them are available up to now in the data banks (nad1, nad2, nad5, atp1, coxII, coxIII, rnl18, rnr26) [13,16-22]. One of the reasons why no complete gene sequences are available is the difficulty to isolate pure mitochondrial fractions and therefore uncontaminated mitochondrial DNA, on the one hand. On the other hand this topic is of relevant interest to better understand the relative position of Monilophytes, compared to the more ancient Briophytes and the more recent Spermatophytes.

One of the most recent acquisitions on this topic is the results of the studies of Pryer [23], who proposed that the ferns belong to a monophyletic group which is the closest to seed plants, using combined data (morphological and derived from multigene sequence analysis).

To gain more knowledge on the mitochondrial biogenesis of Monilophytes, we chose plants of a filicale family, the fern Asplenium nidus, available at the Botanical Garden of the University of Bari.

Using an unusual procedure we were able to obtain a continuous sequence of DNA (20,374 bp) corresponding, bona fide, to an equivalent region of the A. nidus mitochondrial genome. The main results obtained from our investigation are the following: i) the detection of DNA segments of chloroplast origin; ii) the identification of complete genes for both proteins and tRNAs (nad4L, nad9, atp9, trnN, trnR, trnK); iii) a couple of inverted repeats although of small size.

SOURCES OF MITOCHONDRIAL DNA

Two alternative procedures for the isolation of organelles have been developed depending on the tissue used as starting material: roots or leaves. In the former the soil contained in the thick network of roots was removed by hand and washed in distilled water. This step was followed by drying the roots on filter paper and weighing and wrapping them with a double layer of sterile gauze. After washing several times with sterile water, the roots were suspended in sterile buffer (Mannitol 0.4 M, Mops 25 mM, EGTA 1 mM, PVP 1% pH 7.8), and homogenized in a blender with five hits every five seconds at medium speed. Fractionation, lysis of organelles and DNA extraction was as reported by Hanson [24].

The isolation of organelles and DNA from green leaves (leaf-procedure) was carried out using the same protocol with only a few small modifications [24].
The mitochondrial DNA isolated by the two alternative procedures showed significant differences in chloroplast contamination as judged from the amplification of highly conserved chloroplast regions [25]. Using as a template the mtDNA prepared from the roots, where the copy number of plastid DNA are reduced [26], no amplified products could be detected. The results of these experiments are reported in Fig. (1).

The root procedure was also used for establishing according to Kawata [27] a partial library of sonified DNA fragments and for the synthesis of two distinct amplification products (see below).

The leaf procedure was used mainly for the isolation of total RNA employed for reverse transcription and cDNA synthesis, an investigation concerning editing of transcripts of three specific protein genes (nad4L, nad9, atp9).

**DETECTION AND SEQUENCING OF THREE DISTINCT DNA SEGMENTS**

The fern mitochondrial DNA library was screened using a list of probes corresponding to genes usually encoded on plant mitochondrial genomes. Among them an atp9 probe led to the identification of a specific clone (10C9). The sequence analysis of its insert (of 2,574 bp) revealed the presence of the atp9 gene together with segments of genes from chloroplast origin (rps11 and psbA).

The sequence of the insert begins at position 17,800 and ends at position 20,374 of the continuous sequence (GenBank accession number AM600641) assembled as described below.

Using the Tryple Master PCR System (Eppendorf), particularly suitable for obtaining long PCR products with high fidelity, two large amplification products were obtained.

Unsuccessful preliminary experiments were carried out in the presence of direct and reverse primers already used in previous investigations having as a final goal the identification of tRNA genes or deduced from the sequence of 10C9 clone insert. In further attempts other pairs of oligos having different sequences were used. Among them the primer rtREV (5’– TTGCTTGCCCGTGTCTCTG – 3’) deduced from the sequence of the clone 10C9 was used in combination with several other different oligos. In all these cases amplification products of the same size (about 10,000 bp) were detected.

Indeed the sequencing of CLA amplification product (9,790 bp) confirmed the presence of two inverted repeats at its termini (from 8,555 to 8,692 and from 18,284 to 18,427 bp).

Another gene usually of mitochondrial origin (nad4L) and three further segments of chloroplast origin (see below) were also detected within this amplification product (CLC).

A second amplification product was generated using the pair of primers (CLA1R: 5’- TTCTGTAGGGGACCGAA ACC-3’ deduced from the 5’ terminal side of the CLA segment and the oligo ND1E4F: 5’- CTCACCTCTCTCTA GTCTGTG -3’ deduced from the 3’ terminus of the nad1 gene sequence available in the data bank (GenBank accession number AY353954)).

Further genes (or segments of them) of mitochondrial and chloroplast origins were detected on this second long amplified product:

i) in particular a region of 154 bp corresponding to the 3’ terminus of fourth exon of the nad1 gene.

ii) a segment of 132 bp of chloroplast origin.

Other than these DNA stretches three tRNA genes were detected; two of native type and one usually considered being from chloroplast origin.

Bioinformatic analysis of these three DNA sequences was carried out using FASTA and BLAST programs to search for structural genes. tRNA scan-SE for tRNA genes
and “Repeat search” derived from “Fast PCR” to search for direct or reverse repeats.

ASSEMBLY OF THREE SEQUENCES IN A CONTINUOUS FORM

The sequence of the continuous region described under the accession number AM600641 was obtained connecting that of 10C9 insert to CLA amplification product through a common overlapping region of 601 bp. At the same time the right border of CLC amplification product could be connected to the left border of CLA through a second common and overlapping region of 214 bp (Fig. (2)).

Despite the unusual procedure used for the production of the 20,374 bp DNA continuous sequence (part belonging to a clone and part to amplification products), we think that it can be considered, bona fide, as it corresponds to that of a region of A. nidus mitochondrial DNA for several different reasons: i) the 2,574 kbp insert of clone 10C9 is most probably a mitochondrial segment because the \( atp9 \) gene has always been detected among genes of that origin; ii) the amplified segment of 9,790 bp region contains a segment of 601 bp completely overlapping for the same extent to part of the 2,574 kbp insert of clone10C9 and two further genes of mitochondrial origin (\( nad4L \) and \( nad9 \)); iii) the amplified product of 8,794 bp overlaps perfectly the CLA product for 214 bp, contains tRNA genes highly similar to those detected on other plant mitochondrial genomes and a segment of another gene of mitochondrial origin: the \( nad1 \) gene (the fourth exon); iv) the results of the editing studies carried out on the transcripts of the genes detected on the continuous DNA region of 20,374 bp. These studies revealed in all cases high levels of both C-U and U-C substitutions (see below). This analysis strongly supports the hypothesis that they are of mitochondrial origin and therefore that the region of 20,374 bp is part of the \( A. \) nidus mitochondrial genome.

This conclusion can be also drawn in consideration that in the plant kingdom the transcript editing is active exclusively in the mitochondria.

The sequence produced has been confirmed by control experiments of amplification carried out using three different preparations of DNA isolated by the root-procedure and various pairs of primers. They were different from those used to generate CLC and CLA products; in particular some of them were selected to confirm the sequences of the overlapping regions. All the amplification products (about ten) gave sequences almost identical to those corresponding to the continuous sequence.

SEQUENCE ANALYSIS OF THE 20,374 BP DNA SEGMENT

Detection of Segments from Chloroplast Origin

About 20% of the 20,374 bp segment contains sequences from chloroplast origin. They account for regions of various sizes of chloroplast genes such as \( rps11 \) [28] and \( psbA \) and \( rpoB \) [29] encoded on the \( Adiantum \) \( capillus \) \( veneris \) and \( Goebelia \) \( cornigera \) chloroplast genomes respectively. Further short stretches of \( trnA \) 232 and of the second exon of \( trnA \) gene [29] and of part of \( Pteridium \) \( aquilinum \) \( rrn16 \) gene (GenBank accession number Z81323) were also identified (Table 1). For further details about these gene segments see also its legend.

Table 1. The accession numbers of \( A. \) \( capillus \) \( veneris \) and of partial sequences of \( G. \) \( cornigera \) and \( P. \) \( aquilinum \) chloroplast genome are AM600641, AM607944 and Z81323. The coordinates of the segments of the chloroplast genes are as follows: \( psbA \) : 19712-19735; \( rps11 \) : 18621-18772; \( rpoB \) (b): 15686-16240; \( rpoB \) (a): 16294-16683; \( trnA \) : 13863-13899; \( rrn16 \) : 432-564.

<table>
<thead>
<tr>
<th>Element Compared</th>
<th>Dimension of Segments Detected on 20374 BP</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Identity</td>
<td>Positivity</td>
</tr>
<tr>
<td>( psbA )</td>
<td>( G. ) cornigera</td>
<td>123</td>
</tr>
<tr>
<td>( rps11 )</td>
<td>( A. ) capillus veneris</td>
<td>153</td>
</tr>
<tr>
<td>( rpoB ) (b)</td>
<td>( A. ) capillus veneris</td>
<td>555</td>
</tr>
<tr>
<td>( rpoB ) (a)</td>
<td>( A. ) capillus veneris</td>
<td>389</td>
</tr>
<tr>
<td>fragment ( (trnA, \ rrn23) )</td>
<td>( A. ) capillus veneris</td>
<td>361</td>
</tr>
<tr>
<td>( rrn16 )</td>
<td>( P. ) aquilinum</td>
<td>132</td>
</tr>
</tbody>
</table>

The segments \( psbA \) and \( rps11 \) are oriented in opposite directions (Fig. (1)). The former starts at nucleotide 19,712 and ends at nucleotide 19,835. The coordinates of the latter are from 18,621 to 18,772.

The segments b and a of the \( rpoB \) gene coding for the beta subunit of chloroplast RNA polymerase are located at nucleotide number 15,686 to 16,240 and from 16,294 to 16,683, respectively.

The relative position of segments \( rpoB-b \) and \( rpoB-a \) suggests that their insertion could be related to two distinct transfer events. Indeed the \( rpoB-b \) segment corresponding to amino acid 887-1,071 precedes the second segment (\( rpoB-a \)) which codes for the first 145 amino acids of the corresponding protein and is located at a distance of only 53 nucleotides from sequence b.

The position on the 20,374 bp sequence of the other three short stretches similar to chloroplast sequences are: \( rrn16 \) from 432 to 564, \( rrn23 \) from 13,673 to 13,694, second exon of the \( trnA \) gene from 13,863 to 13,899.

The identity of nucleotide sequences between regions of the 20,374 segment and the various chloroplast DNA counterparts range from 85% for \( psbA \) segment to 100% (segment of the \( rrn16 \) gene) whereas amino acid similarities range from 67% for the segment “a” of RpoB protein to 95% for the segment of PsbA protein (Table 1).

The overall conclusion that can be drawn from the analysis of results described in this section is that the transfer of genetic material from chloroplast to mitochondria was already active in Monilophytes. Moreover this investigation shifts for several millions of years in the past the appearance of this event during plant mitochondrial evolution.

The detection of this feature which is also peculiar to mtDNAs of Spermatophytes plant species is in agreement
with the proposal of Pryer [23] concerning the position of the ferns being close to the Spermatophytes.

**tRNA Genes**

Three tRNA genes were detected on the 20,374 bp sequence: the first two (trnR and trnK) show high similarities with corresponding genes of the native kind encoded on other plant mitochondrial genomes [3-8]. This comparison has led to the conclusion that they both have the same origin.

The analysis of the trnN gene is of particular interest. Usually it is in most of the higher plant mitochondrial genomes studied so far from chloroplast origin. To speculate about its possible genetic origin the sequence was aligned with the corresponding gene encoded on the A. capillus veneris plastid genome (the only available among fern chloroplast genomes) and those detected on the N. tabacum cp genome [30] and the Z. mays mt genome [7] (Fig. (3)).

Multialignment reveals that the similarity between the A. nidus mitochondrial and A. capillus veneris chloroplast gene is very low (about 68%). In contrast similarity between A. capillus veneris chloroplast genes and the Z. mays mitochondrial genes is high (about 87%) as it is between N. tabacum chloroplast and Z. mays mitochondrial genes (about 96%). Therefore, this analysis suggests that the trnN gene is most probably of native origin. This conclusion is also supported by considering that A. nidus and A. capillus veneris are taxonomically closer than Z. mays and N. tabacum [30].

This is the first case showing the existence of a native trnN gene encoded by a plant mitochondrial genome.

**Detection of Protein Genes**

Sequence analyses of the 20,374 bp segment revealed the presence of three protein genes of mitochondrial origin: atp9, nad4L and nad9.

**Identification of Editing Sites on the Transcripts of three Mitochondrial Protein Genes**

ClustalW program was used to determine editing sites, aligning DNA and cDNA sequences.

The overall analysis of results obtained in this investigation shows several specific features for the genes coded on A. nidus mtDNA compared with the same genes studied on other angiosperm genomes: i) editing levels on A. nidus transcripts are always higher (quite often twice as much). See also legends of Fig. (3) and Fig. (4); ii) reverse editing events (U-C) are present only in transcripts of A. nidus genes; iii) in particularly critical positions editing events of both types generate transcripts which can be translated as functional proteins.

**Analysis of Editing Sites on atp9 transcripts**

The multi-alignment of cDNAs of five different plant species [3, 5-7] produced from atp9 transcripts and described in Fig. (3) reveals some relevant features mainly related to A. nidus transcripts. In particular, a direct editing event generates a stop codon as detected only for O. sativa cDNA (codon 75).

As previously anticipated, no reverse editing sites are detected on Spermatophyta cDNAs whereas three of them have been detected on A. nidus mitochondrial atp9 cDNA. Two of them in particular have critical consequences because they suppress two stop codons within the gene: codons 37 and 44.

As far as the reproducibility of cDNA analysis for the atp9 gene is concerned, it must be stressed that 100% of cDNA clones (fifty) showed the same sequences and therefore revealed the same editing pattern.

**Analysis of Editing Sites on nad4L Transcripts**

Fig. (4) describes the comparison of cDNAs from atp9 transcripts in five plant species. Direct and reverse editing sites are reported in small letters. The percentage of editing sites for the four angiosperm species ranges from 1.6 % (A. thaliana) to 3.6 % (O. sativa). The same value for A. nidus is 4.9 %. Data from other papers in which editing sites are reported as a whole have not been considered.
of the nad4L gene on the genome, ii) the kinetics of editing for sites of genes with different orientations are different; iii) finally, a combination of the events considered in the two hypotheses can generate the complex editing pattern described in this section.

---

**Fig. (5).** Comparison of cDNAs from nad4L transcripts in four plant species. Direct and reverse editing sites are indicated as reported in the legend of Fig. (4). The percentage of editing sites ranges from an almost similar value for angiosperm species (3%) to 8% for A. nidus.

### Analysis of Editing Sites on nad9 Transcripts

As well as atp9 and nad4L for nad9, editing events have a higher frequency on A. nidus transcripts. For the transcripts of nad9 gene the direct editing events (1.6%) and the reverse editing events (1.5%), are almost equivalent. Again they have a high relevance because suppress three stop codons within the nad9 gene (Fig. (6)). One interesting point has to be highlighted: the A.nidus nad9 gene is longer than what usually found in any other plant species. In particular it has a 3' extension of 105 nt (35 amino acids on carboxil-terminus as deduced protein) longer than that of B. napus, (accession number BAC98862) *O. sativa* (BAC19900), *Z. mais* (ABE98687) and almost similar to that of *M. polymorpha* (NP_054446): 87 bp (29 a.a.). This observation implies further studies because the extension has been detected in other plant mitochondrial genes. In particular in wild beet the NAD9 subunit has a C-terminal extension of 14 amino acids, which has been related to the male sterile G cytoplasmin [31].

### Repeated Sequences

Several direct and inverted repeats were detected within the 20,374 bp segment. Most of them are very small in size and located short distances apart.

As described above, the use of a single primer (rtREV) made possible the amplification of a large region (of 9,790 bp). Sequencing of this segment revealed at a distance of 9,592 bp, two inverted repeats of 137 bp and 143 bp.

The relevance of the presence of these repeats remains to be understood although, the possibility that the region included between them could be present in the genome in two different orientations, cannot be at the moment completely ruled out.
Unfortunately the unusual procedure used in this paper can not answer further questions concerned with the study of plant mitochondrial genomes. The most relevant are the following: the size and shape of the genome and the position of 24,374 bp segment related to the reminder of genome part. The reasons for why no suitable answers are available at the moment for the above mentioned questions depends on the incompleteness of library which makes not yet possible to draw a complete map of the mtDNA master chromosome.

Our next goal will be to demonstrate whether cp-like tRNA genes are on the A. nidus mitochondrial genome. The results of preliminary experiments (Panarese unpublished) demonstrate the existence of 24 tRNA genes on this genome, 5 of which are likely from chloroplast origin. However, no direct evidence is available at the moment, although the approach used in this paper should help in finding tRNA genes inserted in full mitochondrial contexts, as reported for other cp-like genes [10-12].

ACKNOWLEDGMENTS

This work was supported by “Ateneo funds” from University of Bari.

The authors thank Prof. Graziano Pesole for the fruitful discussion on experimental data and conclusions, Dr. Antonio Placido for the supply of some specific primers and templates and Dr. Richard Lusardi lecturer in scientific English at Bari University for the revision of language.

REFERENCES


