Angiosperm mitochondrial genomes and mutations

Tomohiko Kubo a, Kathleen J. Newton b,*

a Laboratory of Genetic Engineering, Research Faculty of Agriculture, Hokkaido University, N-9, W-9, Kita-ku, Sapporo 060-8589, Japan
b Division of Biological Sciences, University of Missouri, Columbia, MO 65211, USA

Received 8 August 2007; received in revised form 9 October 2007; accepted 18 October 2007
Available online 4 November 2007

Abstract

Flowering plants harbor the largest mitochondrial genomes reported so far. At present, the nucleotide sequences of 15 mitochondrial genomes from seven angiosperm species are available, making detailed comparative analysis feasible. The gene content is variable among the species, but the most striking feature is the fluidity of intergenic regions, where species-specific sequences predominate. Additionally, angiosperm mitochondrial genomes, even within a species, show a remarkable amount of rearrangement. We also review mitochondrial mutants in angiosperms from a genomic viewpoint, and discuss how they have arisen. The involvement of nuclear genes in mitochondrial genome stability and organization is currently being revealed through the analysis of mutants.

© 2007 Elsevier B.V. and Mitochondria Research Society. All rights reserved.

Keywords: Genome organization; Heteroplasmy; Homologous recombination; Horizontal gene transfer; Nuclear–mitochondrial interaction; Plant mitochondria

1. Introduction

The largest mitochondrial genomes reported so far are those of flowering plants, whose sizes are estimated to range from 200 to more than 2400 kbp (reviewed in Scheffler, 1999). From investigations of both plant and non-plant mitochondrial genomes (Knoop, 2004; Gray et al., 2004; Terasawa et al., 2007), it has been inferred that the sizes of angiosperm mitochondrial genomes have expanded since plants colonized the land although the genomes have lost some genes during this evolutionary period. This means that intergenic regions have expanded in angiosperm mitochondria. Therefore, the evolutionary trend of the angiosperm mitochondrial genome is counter to that of the mammalian mitochondrial genome, which has become smaller and more compact since the endosymbiotic origin of mitochondria (Scheffler, 1999). Another characteristic of angiosperm mitochondrial genomes is the high degree of variation in terms of the genomic organization, because frequent rearrangements are still ongoing in almost every angiosperm lineage. During this process, unique mitochondrial mutants have arisen as by-products.

Due to initial technical difficulties in sequencing the complex and large DNA molecules, the entire nucleotide sequence of an angiosperm mitochondrial genome was not known until 1997, when Brennicke’s group reported the 366,924 nucleotides of the model plant Arabidopsis thaliana (Unseld et al., 1997). Subsequently, the complete nucleotide sequences of 15 mitochondrial genomes from seven angiosperm species have been made available (Table 1) (Kubo et al., 2000; Notsu et al., 2002; Handa, 2003; Satoh et al., 2004; Clifton et al., 2004; Sugiyama et al., 2005; Oghara et al., 2005; Tian et al., 2006; Allen et al., 2007). In this review, we will first describe briefly the organization of angiosperm mitochondrial genomes. Then, we will give an overview of angiosperm mitochondrial mutants classified on the basis of their associated mutations. More detailed information has been presented in earlier reviews on the first topic (Kubo and Mikami, 2007; Kmiec et al., 2006; Knoop, 2004; Bullerwell and Gray, 2004; Adams and Palmer, 2003; Marienfeld et al., 1999). For details about cytoplasmic male sterility and nuclear–mitochondrial interaction, see Carlsson et al.
## Table 1
Comparison of gene content among angiosperm mitochondrial genome

<table>
<thead>
<tr>
<th>Plant:</th>
<th>Sugar beet</th>
<th>Tobacco</th>
<th>Arabidopsis</th>
<th>Rapseseed</th>
<th>Rice</th>
<th>Japonica, Nipponbare</th>
<th>Japonica, Nipponbare-S</th>
<th>Indica, 93-11</th>
<th>Maize</th>
<th>B37N</th>
<th>A188</th>
<th>B37S</th>
<th>B37T</th>
<th>B37C</th>
<th>Chinese spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivar/strain:</td>
<td>TK81-O</td>
<td>TK81-MS</td>
<td>Bright Yellow04</td>
<td>Col-0</td>
<td>Wester</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name of cytoplasm or genome type</td>
<td>Owen cytoplasm</td>
<td>nap cytoplasm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of master circle (nucleotides)</td>
<td>368,801</td>
<td>501,020</td>
<td>430,597</td>
<td>366,924</td>
<td>221,853</td>
<td>490,520</td>
<td>490,669</td>
<td>490,673</td>
<td>491,515</td>
<td>569,630</td>
<td>701,046</td>
<td>557,162</td>
<td>535,825</td>
<td>739,719</td>
<td>452,528</td>
</tr>
<tr>
<td>Accession number</td>
<td>BA000009</td>
<td>BA000024</td>
<td>BA000042</td>
<td>Y08501</td>
<td>AP006444</td>
<td>BA000029</td>
<td>DQ167400</td>
<td>DQ167807</td>
<td>DQ167399</td>
<td>AY506529</td>
<td>DQ490952</td>
<td>DQ490951</td>
<td>DQ490953</td>
<td>DQ645536</td>
<td>AP008982</td>
</tr>
<tr>
<td>Complex I</td>
<td>nad1, 2, 3, 4, 4L, 5, 6, 7, 9</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Complex II</td>
<td>sdh3</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Complex III</td>
<td>cob</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Complex IV</td>
<td>cox1, 2, 3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Complex V</td>
<td>atp1, 4, 6, 8, 9</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Cytochrome c biogenesis</td>
<td>ccmB</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ccmC</td>
<td>+*1</td>
<td>+*1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ccmFC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ccmFN</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+*2</td>
<td>+*2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ribosomal protein</td>
<td>rpl2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rpl8</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rpl16</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps1</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps10</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>rps11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>rps12</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rps19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Protein translocator</td>
<td>mttB</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>mat-R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Intronic ORF in the fourth intron of nad1</td>
<td>rRNA</td>
<td>rrm5, 18, 26</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Chloroplast-like</td>
<td>Native</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Origin unknown</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>CMS-associated ORF</td>
<td>preSatp6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>orf222</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>orf5531</td>
<td>orf77</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+, presence of the gene; -, absence of the gene; ψ, pseudo gene; *1, unusual N-terminal extension; *2, split into two genes.

2. Organization of angiosperm mitochondrial genomes

2.1. Mode of existence: does the 'master chromosome' really exist?

For historical reasons, the angiosperm mitochondrial genome is usually described as a single circular DNA molecule that houses a complete set of genes (Fig. 1), called the ‘master chromosome’ (reviewed in Scheffler, 1999). With one reported exception (Palmer and Herbon, 1987), each angiosperm master chromosome harbors one to several sets of repeated sequences where active recombination events occur between the repeat copies. Due to this property, isomeric forms of the master chromosome would be expected when the repeat copies of a set are present in inverted orientation to each other; alternatively, two subdivided molecules (subgenomes) would be expected when two copies of the same repeat are present in direct orientation.

Master circular chromosomes of angiosperm mtDNA have been generated by restriction mapping and then by shotgun or mapped-cosmid sequencing, although some groups reported difficulty in generating circular configurations (Kubo et al., 1999; Robinson and Wolyn, 2002). In one notable case, maize CMS-S, it is well-accepted that the mitochondrial genome exists mainly as multiple linear molecules, and the sequencing data was found to be consistent with this premise (discussed in Allen et al., 2007).

The multipartite structure of plant mitochondrial genomes generated by recombination within master circles was used to explain the heterogeneity that is observed when plant mitochondrial (mt) DNA is examined by electron microscopy and gel electrophoresis (Kmiec et al., 2006). However, the apparent morphology of carefully isolated mtDNA was not found to be consistent with the circular model (Oldenburg and Bendich, 1996). The observed molecules appeared to be linear and circular and of various sizes (including structures that exceed estimated genome sizes). In addition, Y-, H-, and theta-shaped branched forms were seen, which presumably represent recombination intermediates (Backert and Börner, 2000). Therefore, the entity of an angiosperm mitochondrial genome is likely to be a mixture of various DNA molecules. It has been suggested that plant mtDNA replicates via a recombination-dependent mechanism (Oldenburg and Bendich, 1996; Backert and Börner, 2000). It should be noted that the concept of the master chromosome remains because the question as to how multipartite and branched DNA molecules are transmitted properly to the next generation is unsolved (see below).

Fig. 1. Maize mitochondrial genomes differ in size mainly due to the presence of large duplications. Circular maps for CMS-C (740 kb) and NB (570 kb) were generated from mitochondrial genomic sequences (Allen et al., 2007). The positions of genes and gene fragments are shown around the periphery of each ring. Large repeats are color coded within the rings, with their orientations indicated by the directions of the arrowheads. NB and CMS-C share the 11 kb (yellow) and 17 kb (bright-red) duplications. The CMS-C genome has, in addition, three sets of very large repeats of 45 kb (olive-green), 50 kb (purple), and 100 kb (orange), which together account for the 30% larger size of the C genome relative to NB. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
2.2. Genetic information

The sizes of the master chromosomes of the sequenced mitochondrial genomes range from 221,853 (rapeseed) to 739,719 bp (maize CMS-C; Fig. 1) (Table 1). However, because these figures include repeated sequences or large duplications ranging from 0.2 to 120 kbp, the net genetic complexity (i.e., the size following removal of one copy of each large repeat) is much smaller. In addition to these long repeats, short repeated sequences can be found frequently. The difference between the large repeats and short repeats is not only the length but also the activity of recombination. Whereas many of the large repeats appear to actively and frequently recombine, the short repeats are usually inactive, but they appear to play a central role in the evolution of angiosperm mitochondrial genomes (Andre et al., 1992; Conklin and Hanson, 1994; see below).

The number of mitochondrial genes in angiosperms is 50–60 (not considering copy number), as summarized in Table 1. The differential number of genes is due to the differential gene content for the subunits of Complex II, and especially, ribosomal proteins and tRNAs. When the content of ribosomal protein genes is compared among angiosperms, one can realize how often genes have been lost from the mitochondrial genome during angiosperm evolution. Most of the genes that are lost from the mitochondrial appear to have been transferred to the nuclear genome (Adams and Palmer, 2003), but this is not always the case. For example, one of the duplicated rpl13 genes of plastid origin in the A. thaliana nuclear genome has been recruited to encode a mitochondrial targeted polypeptide (Mollier et al., 2002).

Plant mitochondrial genes are translated according to the universal genetic code. However, the composition of transfer RNA genes in angiosperm mitochondria is quite unique. Of the 15–21 tRNA genes encoded, 10–12 are orthologous to those residing in moss mitochondria, and thus are considered to be descendants of those harbored by the initial endosymbiont from the origin of mitochondria: this class of tRNA genes is called 'native' (Dietrich et al., 1992; Marienfeld et al., 1999; Kubo and Mikami, 2000). The remaining tRNA genes either have high homology to chloroplast DNA or their origin is unknown; i.e., their sequences do not match tRNAs from any known source. Differential allocation of tRNA genes to these classes is frequently found: for example, whereas tRNA^{Cys} (GCA) is encoded by a native class gene in many dicotyledonous plants, an origin-unknown gene is the only functional tRNA^{Cys} (GCA) gene in sugar beet, and grass family members use a chloroplast-like gene (Kubo et al., 2000). This suggests that recruitment of tRNA genes has occurred independently in several angiosperm lineages. Another example is tRNA^{His} (GUG); although most of the angiosperms have a corresponding chloroplast-like gene, no tRNA^{His} (GUG) gene is found in wheat mitochondria (Ogihara et al., 2005). Lack of this tRNA gene is compensated for by the import of cytosolic tRNA^{His} (GUG) molecules (Dietrich et al., 1992; Kumar et al., 1996).

In general, the import of tRNA molecules from the cytosol is common in angiosperm mitochondria. Sequence analysis of mitochondrial genomes revealed that tRNAs for alanine, arginine, leucine, threonine, and valine are commonly lost (Marienfeld et al., 1999; Kubo and Mikami, 2007), so the corresponding tRNAs would be expected to be imported. Additional tRNA molecules have to be imported in some plant lineages: for example, tRNAs for phenylalanine, methionine (elongator), and tryptophan are needed in A. thaliana mitochondria (Marienfeld et al., 1999; Duchene and Marechal-Drouard, 2001). tRNA molecules are presumed to pass through the mitochondrial membrane through voltage dependent anion channels (Salinas et al., 2006).

Some of the mitochondrial genes in angiosperms are interrupted by introns. In each of the sequenced genomes, the total number of the introns is 20–24, constituting 4–13% of the genome. All the introns in the sequenced mitochondrial genomes are classified as group II type; however, a horizontally transferred group I intron has also been documented (Cho et al., 1998). One of the characteristics of angiosperm mitochondrial introns is the presence of trans splicing, particularly for genes encoding Complex I subunits. This was first reported for nad1 (Chapdelaine and Bonen, 1991; Wissinger et al., 1991); then nad2 and nad5 were revealed to contain trans-splicing introns (Binder et al., 1992; Knoop et al., 1991; Pereira de Souza et al., 1991). That there was a transition from a cis to a trans arrangement of introns during their evolution was supported by the finding that all the trans introns in angiosperms are in cis arrangements in non-angiosperm plants (Malek and Knoop, 1998). Interestingly, sometimes the transition seems to have occurred independently in multiple lineages of angiosperms. For example, although sugar beet, petunia, and tobacco have a trans-splicing intron between the fourth and fifth exons of nad1, careful phylogenetic analysis concluded that they have independent origins (Qiu and Palmer, 2004). For details on plant introns and splicing, see Bonen (2007) in this issue.

RNA editing is a necessary step for angiosperm mitochondrial gene expression, which converts cytidine residues to uridine and, rarely, uridine to cytidine. Most of the RNA editing sites are found in the protein-coding genes, and a few are in tRNAs, untranslated regions, and introns (Shikanai, 2006). The numbers of editing sites have been determined for some of the sequenced angiosperm species, ranging from a total of 357 in sugar beet (Mower and Palmer, 2006) to 491 in rice (Notsu et al., 2007). Comparative analysis of editing sites revealed that there are a number of species-specific editing sites: for example, A. thaliana and rapeseed have species-specific editing at 83 and 69 sites, respectively (Handa, 2003). For details about RNA editing, see Takenaka et al. (2008) in this issue.
2.3. ‘Promiscuous’ DNA in intergenic regions

Because the gene-coding regions constitute only 7–17% of the mitochondrial genome in the angiosperms sequenced to date, the remaining regions (intergenic regions) were examined to discover how the genome sizes had expanded in angiosperms. However, no conclusive mechanism has been proposed yet. Database searching using intergenic regions as a query revealed that some intergenic regions contain chloroplast DNA sequences (1.6–6.2% of the genome) and nuclear DNA sequences (0.1–13.4% of the genome) (Kubo and Mikami, 2007; Allen et al., 2007). Nuclear or chloroplast DNA that has incorporated into the mitochondrial genome has been called ‘promiscuous’ DNA (see Timmis et al., 2004). Promiscuous DNA has never been reported in animal mitochondrial genomes (Scheffler, 1999). As was mentioned earlier, chloroplast DNA sequences in mtDNA can be sources of some functional tRNA genes, such as trnD-GAC, trnH-GUG, trnM-CAU, trnN-GUU, trnS-GGA and trnW-CCA (Marienfeld et al., 1999; Kubo and Mikami, 2007). It should be noted that chloroplast DNA sequences have also been reported to serve as the promoter for a mitochondrial gene (Nakazono et al., 1996); however, this is not commonly found. Other possible roles are currently unknown.

Nuclear DNA sequences have been found by database searches as well as by hybridization analyses, but there are no reports to identify any functions for these nuclear DNA sequences in mitochondria. Overall, the functional significance of promiscuous DNA does not seem great except for the above-mentioned examples. Promiscuous DNA is rarely conserved among angiosperm mitochondrial genomes.

2.4. Intergenic DNA ‘of unknown origin’ in plant mitochondrial genomes

Outside of genes, which themselves are highly conserved, and after accounting for chloroplast, nuclear and plasmid DNA insertions, the majority of the DNA in the sequenced plant mitochondrial genomes is of unrecognizable origin. Considering the compact and conserved nature of animal mitochondrial genomes, it was truly surprising to find that, in the first sequenced angiosperm mitochondrial genomes, over half of each genome had no homology to any sequences in the public databases (Unsold et al., 1997; Kubo et al., 2000; Notsu et al., 2002; Handa, 2003; Clifton et al., 2004).

An unusual case of large amounts of recognizable sequences in intergenic mtDNA has been discovered in the basal angiosperm, Amborella trichopoda (Bergthorsson et al., 2004). It contains over 250 mitochondrial gene sequences with high sequence similarity to other, usually eudicot, mtDNA. Palmer and colleagues have proposed that this situation results from relatively recent mitochondrion-to-mitochondrion gene transfer (Richardson and Palmer, 2007). Multiple and ongoing horizontal gene transfer events may be responsible for the expansion of the A. trichopoda mitochondrial genome to over 4000 kb in size (J. Palmer, personal communication). A number of other sporadic mitochondrial horizontal gene transfer (HGT) events have been recently reported in a variety of plant species (reviewed in Richardson and Palmer, 2007).

The ways that plant mitochondria from distant relatives could come in contact with one another to transmit information is currently a matter of speculation and could include illegitimate pollination, insect vectors, and parasitic and epiphytic growth on host plants (Richardson and Palmer, 2007). Thus, although A. trichopoda may represent the extreme case, rare HGT among mitochondria over evolutionary time spans could be one of the possible sources of the additional mtDNA present in angiosperm mitochondrial genomes. After this extra, unneeded DNA is incorporated, it would be subject to rearrangement and loss such that it becomes unrecognizable with the passage of time.

2.5. Genome rearrangements and acquisition/loss of nucleotide sequences

It is well known that the arrangement of mitochondrial genes can vary, even within a single angiosperm species (e.g., Satoh et al., 2004; see also Fig. 1). This is because angiosperm mitochondrial genomes have undergone frequent genome rearrangement over the course of time. In this process, DNA recombination leading to inversions or deletions has played a major role. For example, some but not all the rearrangements between two sugar beet genomes may involve homologous recombination via 9–376 bp repeated sequences (Satoh et al., 2006; see also below).

The distribution of rearrangement points was investigated through comparative analysis of five maize genomes (Allen et al., 2007). Interestingly, rearrangement points are unevenly distributed in the genome, but the reason for this is unclear. Most of the rearrangements do not affect gene expression; i.e., they just produce a polymorphism between the genomes. In rare cases they are critical for mitochondrial gene expression (see below).

During successive genome rearrangements, angiosperm mitochondrial genomes have lost and acquired some sequences. When the first two sequenced angiosperm genomes, A. thaliana and sugar beet, were compared, shared sequences between the two were found to comprise only 21% of each genome, consisting of gene-coding regions and their flanks (Kubo et al., 2000). It is also surprising that intra-specific comparison of angiosperm mitochondrial genomes revealed 0.48–13.6% of the genome to be specific to either of the genomes (Satoh et al., 2004, 2006; Allen et al., 2007). Much of the ‘mitotype-specific sequence’ has been determined to originate from the chloroplast, the nucleus or from plasmids resident in the mitochondria; however, DNA whose origin is unknown can also be specific to a single mitotype (Satoh et al., 2006; Allen et al., 2007).
3. Mitochondrial mutants in angiosperms

3.1. Caused by point mutation

Whereas a number of point mutations in mitochondrial genomes are reported to be associated with human diseases (Moraes, 1998), very few point mutations that impair mitochondrial function have been described in angiosperm mitochondria. A unique mitochondrial mutant has been found in wild beet populations in which the activity of cytochrome c oxidase is decreased by 50% compared to normal beet; its COXII polypeptide is missing the final eight amino acids due to truncation by a nonsense mutation in the cox2 ORF (Ducos et al., 2001). Interestingly, this mutated cox2 is present in a homoplasmic state, i.e., no normal copy of the cox2 gene is detected, and only a shorter, less stable COX2 protein is synthesized (Ducos et al., 2001). The phenotype is male sterile but otherwise normal. How does such a mitochondrial mutant survive in wild populations despite such a large decrease in mitochondrial electron transport activity? One answer is that plants have an alternative respiratory pathway (see Rasmusson et al. (2007) in this issue) that appears to be increased in this mutant, which may be sufficient to compensate for the 50% reduction in the cytochrome pathway.

Comparative analyses of nucleotide sequences between normal and mutant mitochondrial genomes that are responsible for maternally inherited male sterility phenotypes (called cytoplasmic male sterility or CMS) has been done in sugar beet (Satoh et al., 2004) and maize (Allen et al., 2007). These analyses revealed a total of 1–24 nucleotide substitutions in gene-coding regions; however, the functional significance for mitochondrial gene expression has not yet been experimentally elucidated. On the other hand, it should be noted that most CMS-related mutations, including the above-mentioned examples in sugar beet and maize, are presumed to be associated with the expression of aberrant ORFs (Yamamoto et al., 2005; Allen et al., 2007; see below).

3.2. Caused by deletion

Some angiosperm mitochondrial mutants are caused by deletions resulting from recombination via rather short (6 to less than 100 nucleotides) repeated sequences, leading to the loss of parts of specific mitochondrial genes. This class of mutants exhibits abnormal phenotypes such as retarded development and, in several cases, leaf variegation due to chlorophyll disorders (a pleiotropic effect of the mitochondrial deficiency). In maize, a series of spontaneous defective-growth mutants, known as non-chromosomal stripe (NCS), have been described that are heteroplasmic for deletions in mitochondrial genes such as cox2, nad4, and rps3 (Lauer et al., 1990; Marienfeld and Newton, 1994; Newton et al., 1996). In a tobacco-related species, Nicotiana sylvestris, apparently homoplasmic nad7-deficient mutants that exhibit male sterility and slow growth were obtained from protoplast culture (Pineau et al., 2005). These mtDNA deletion mutants have been utilized as a tool for investigating nuclear–mitochondrial interactions (Kuzmin et al., 2004; Vidal et al., 2007). A similar cucumber mutant called mosaic (MSC) was also reported, but it is still unclear which of the mitochondrial genes is deleted (Lilly et al., 2001).

In most of the defective-growth mitochondrial mutants, normal mitochondrial genes are also present in a heteroplastic state. Cells homoplastic for the maize NCS6 cox2 deletion can be obtained by tissue culture, but it is quite difficult to regenerate a plantlet (Gu et al., 1994), suggesting that mitochondrial activity is insufficient for development when the cox2 deletion mutation is homoplasmic. Homoplasmic NCS2 maize plants lacking functional nad4 have been found on rare occasions in the field, but such plants exhibit both male and female sterility and severely depressed growth (Yamato and Newton, 1999).

3.3. Caused by an aberrant ORF

In angiosperms, the most commonly seen class of mitochondrial mutant is cytoplasmic male sterility (CMS), in which male gametophytic development is impaired but the plant is otherwise normal (Chase, 2007; Hanson and Bentolila, 2004). Mitochondrial genomes of CMS plants have an identical set of genes to that of normal mitochondria (Satoh et al., 2004; Allen et al., 2007). Efforts to identify the responsible genes in mitochondrial genomes resulted in the discovery of aberrant open reading frames, consisting of fragments of mitochondrial genes and/or unknown sequences, usually with a chimeric structure.

In a few cases, aberrant ORFs are translated into unique polypeptides associated with mutant phenotypes. For example, pcf, which encodes a 25 kDa polypeptide associated with petunia CMS, consists of four parts: a fragment of atp9, two cox2 fragments, and a sequence of unknown origin (reviewed by Hanson and Bentolila, 2004). Aberrant ORFs can also exist within gene-coding regions such as the 5′ leader of the atp6 gene. In sugar beet, the 5′ leader of atp6 in a CMS variety is extended to 387 codons and is translated into a CMS-associated 35 kDa polypeptide, whereas the translation product of core atp6 is unaffected (Yamamoto et al., 2005).

To date, a number of aberrant ORFs, each producing unique polypeptides, have been reported to be associated with CMS (Chase, 2007; Hanson and Bentolila, 2004). Surprisingly, the aberrant ORFs have little in common, even when their phenotype is very similar. This suggests that the CMS-associated aberrant ORFs have arisen independently during the course of angiosperm evolution. In all the sequenced angiosperm mitochondrial genomes, ORFs of unknown function are found and some of them exhibit chimeric organization, although they are unlikely to be translated (Marienfeld et al., 1997; Yamamoto et al., 2005). Whether the ORFs of unknown function are starting to evolve into novel aberrant ORFs conferring some
specific phenotype or whether they are accumulating mutations and becoming degraded is not known.

Although most of the aberrant ORFs are associated only with male-sterile flowers, additional phenotypes can be observed. In maize CMS-T, male sterility is accompanied by sensitivity to specific fungal toxins and an insecticide, which results from pleiotropic effects of its aberrant ORF, T-urf13-T (Schnable and Wise, 1998). Another aberrant ORF in citrus, ACRS, also seems responsible for fungal-toxin sensitivity, but ACRS is not associated with CMS (Ohtani et al., 2002). Unlike the deletion-class mutants, retarded growth or chlorophyll disorders are not seen with the expression of the CMS-associated chimeric genes, even though the unique polypeptides are often accumulated in mitochondria throughout the whole plant. One exception is pvs-orf239 in common bean, which exhibits anther-specific accumulation of the translation products (Sarria et al., 1998).

Although the phenotypes of vegetative and female reproductive tissues of CMS plants show no obvious defects, physiological and gene expression differences in comparison with male-fertile (normal) plants have been reported for non-anther tissues. These include altered amounts of nucleus-coded mitochondrial proteins in ears of maize (Hochholdinger et al., 2004) and altered electron transport in cell cultures of petunia (Connett and Hanson, 1990). This suggests that accumulation of CMS-associated proteins may affect all the tissues but that there is a mechanism of compensation in non-anther tissues.

Most of the mitochondrially localized proteins that differ quantitatively in their expression between maize CMS and normal mitotypes are nucleus-coded, suggesting differential signaling from mitochondria to the nucleus (Hochholdinger et al., 2004). Differential signaling from mutant mitochondria to alter nuclear gene expression has also been reported for the deletion-class NCS mutants of maize (Kuzmin et al., 2005). The mutation-inducing nuclear genotype appears to involve one or more recessive alleles, which have been transferred into other maize lines through the male parent. In common bean, a dominant allele called Fr specifically reduces the copy number of a mtDNA molecule containing the CMS-associated pvs-orf239 sequence to a substoichiometric level, restoring male-fertility to the plant (Kmiec et al., 2006).

Clearly, there are nuclear genes that affect mitochondrial recombination and replication, and that stabilize mitochondrial genome organization in terms of transmission and/or segregation. Such nuclear genes have been isolated in A. thaliana. They include Msh1, the A. thaliana homolog of the Escherichia coli MutS mismatch repair component, RecA3, which is one of the three A. thaliana homologs of Escherichia coli RecA, and Osb, a member of a family of plant-specific DNA-binding proteins (Abdelnoor et al., 2003; Shedge et al., 2007; Zaegel et al., 2006). Homozygous mutations in any of the three genes results in the emergence of recombinant mtDNA that is not detectable in the wild type plant. The atp9 locus was chosen for a detailed investigation of the recombination events in the three mutants. Interestingly, the three mutants utilized the same short repeated sequences in the atp9 coding region for homologous recombination events, resulted in the emergence of the same new atp9 locus, although the entire genomic arrangement of the three mutants were different. Based
under some constraint, at least within a species. The nature of this constraint is unknown but nuclear genes must be involved, because the stability of genome organization is under control of nuclear genes, as exemplified by the several nuclear mutants mentioned above. If recombination via short repeated sequences is controlled by nuclear genes, the evolution of mitochondrial genomes and emergence of mitochondrial mutants in plants now comes under the purview of the nuclear genome.

Acknowledgements

We thank James Allen for providing Fig. 1. Parts of this work were supported by Grants in Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and a Grant from the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan (to T.K.), as well as by the U.S. National Science Foundation (Grants DBI-0110168 and DBI-0423898 to K.J.N.).

References


Satoh, M., Kubo, T., Mikami, T., 2006. The Owen mitochondrial genome
Satoh, M., Kubo, T., Nishizawa, S., Estiati, A., Itchoda, N., Mikami, T.,
Saber, M., Gagliardi, D., Balk, J., Leaver, C.J., 2003. ORFB is a subunit
dehydrogenases in the electron transport chain of plant mitochondria.
in plants. Mitochondrion 7, 177–194.
mitochondrial genome of petaloid CMS carrot. Mol. Genet. Genomics
268, 232–239.
Saber, M., Gagliardi, D., Balk, J., Leaver, C.J., 2003. ORFB is a subunit
of F1F0-ATPase synthase: insight into the basis of cytoplasmic male
sterility in sunflower. EMBO Rep. 4, 381–386.
Salinas, T., Duchene, A.-M., Delage, L., Nilson, S., Glaser, E., Zaepfel,
M., Marechal-Drouard, L., 2006. The voltage-dependent anion channel, a major component of the tRNA import machinery in plant
induction of mitochondrial rearrangements for cytoplasmic male
cytoplasmic male sterility-associated mitochondrial peptide in common
Satoh, M., Kubo, T., Nishizawa, S., Estiati, A., Itchoda, N., Mikami, T.,
2004. The cytoplasmic male-sterile type and normal type mitochondrial
genomes of sugar beet share the same complement of genes of
known function but differ in the content of expressed ORFs. Mol.
Genet. Genomics 272, 247–256.
Satoh, M., Kubo, T., Mikami, T., 2006. The Owen mitochondrial genome
in sugar beet (Beta vulgaris L): possible mechanisms of extensive
rearrangements and the origin of the mitotype-unique regions. Theor.
Appl. Genet. 113, 477–484.
Sedge, V., Arrieta-Montiel, M., Christensen, A.C., Mackenzie, S.A.,
2007. Plant mitochondrial recombination surveillance requires unusual
Shikanai, T., 2006. RNA editing in plant organelles: machinery, physi-