Biochemical and Biophysical Research Communications 376 (2008) 1–4

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Mini Review

Gene duplication and transfer events in plant mitochondria genome

Ai-Sheng Xiong, Ri-He Peng, Jing Zhuang, Feng Gao, Bo Zhu, Xiao-Yan Fu, Yong Xue, Xiao-Fen Jin, Yong-Sheng Tian, Wei Zhao, Quan-Hong Yao *

Biotechnology Research institute, Shanghai Academy of Agricultural Sciences, Shanghai 201106, China

ARTICLE INFO

Article history:
Received 19 August 2008
Available online 31 August 2008

Keywords:
Gene duplication
Transfer
Evolution
Plant
Mitochondria

ABSTRACT

Gene or genome duplication events increase the amount of genetic material available to increase the genomic, and thereby phenotypic, complexity of organisms during evolution. Gene duplication and transfer events have been important to molecular evolution in all three domains of life, and may be the first step in the emergence of new gene functions. Gene transfer events have been proposed as another accelerator of evolution. The duplicated gene or genome, mainly nuclear, has been the subject of several recent reviews. In addition to the nuclear genome, organisms have organelle genomes, including mitochondrial genome. In this review, we briefly summarize gene duplication and transfer events in the plant mitochondrial genome.

Mitochondria

Mitochondria are membrane-enclosed organelles that occur in most eukaryotic cells. Mitochondria have inner and outer membranes composed of phospholipid bilayers and proteins [1]. Mitochondria, which have been called "cellular power plants", produce most of the cell's supply of adenosine triphosphate (ATP), which is used as a source of energy. In addition to supplying energy to the cell, mitochondria are involved in a range of processes, such as signaling, cellular differentiation, and cell death, as well as control of the cell cycle and cell growth [2]. Mitochondria are not necessarily inherited solely through the maternal line; they can be inherited from both parents [3].

Mitochondrial genome

The non-Mendelian genetics of extracellular genomes were first reported a century ago. Mitochondria are believed to be the products of endosymbiotic events. Most of the DNA present in eukaryotic organisms is in the cell nucleus, but they also have independent mitochondrial genomes [4–5]. Mitochondrial DNA is maternally inherited in most multicellular organisms. Coding regions in the mitochondrial genome accumulate sequence changes very slowly, but the linear arrangement of genes changes quite quickly [6]. Mitochondrial DNA has direct repeats spread throughout the genome. More than 1000 complete mitochondrial DNA sequences have been published for organisms including mammals, protists, ascomycete fungi and plants [7–10].

Gene duplication

Gene duplication and subsequent divergence are important in the evolution of genes by providing genetic material from which novel functions can arise [11–13]. The function of genes after duplication can be categorized as neofunctionalization, subfunctionalization, or nonfunctionalization [14]. Careful analysis of duplicated regions shows that the majority of duplicated genes disappear during evolution [15]. Some duplicated genes may be lost by accumulation of deleterious mutations (nonfunctionalization). Paralogous genes (newly duplicated genes) that are not silenced may be maintained by subfunctionalization (partitioning ancestral functions, with the duplicated genes performing different aspects of the original gene's function) and/or neofunctionalization (one of the genes acquires a novel function) [14,16–18].

There are several pathways by which genes can duplicate [19–23]. Gene or genome duplication events can involve a single gene, a segment of the genome, a single chromosome or even the whole genome. Genome duplications differ from single gene duplications in that whole chromosomes are simultaneously doubled, and the overall number of genes is thereby increased [19,21,22,24]. It has long been known that new genes frequently emerge through gene and genome duplication. Spontaneous duplication of large chromosomal segments has been demonstrated experimentally in yeast [25–26]. Polyploidy, which results in duplication of the whole gene complement of an organism, is widespread in eukaryotes, and

* Corresponding author. Fax: +86 21 62205704.
E-mail addresses: aishengxiongcn@yahoo.com.cn (A.-S. Xiong), yaoquanhong@saas.sh.cn (Q.-H. Yao).

0006-291X/$ - see front matter Crown Copyright © 2008 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.08.116

is probably one of the main mechanisms underlying evolutionary divergence [27–28].

The mitochondrion is a nearly autonomous organelle that contains the biochemical machinery necessary to replicate and transcribe its own genome and to synthesize protein. In addition to the nuclear genome, organisms have mitochondrial genomes, most of which undergo intra- and intergenomic recombination and rearrangement [29]. Mitochondrial genomes are simplified relics of the much larger cellular genomes of their cyanobacterial and proteobacterial ancestors [30]. Gene duplication is an important process in mitochondrial evolution, but duplication of large fragments of mitochondrial genomes is infrequent. Knowledge of the sequences of the mitochondrial genome of a number of organisms has made it possible to conduct comprehensive searches for duplicated genes, enabling informative study of their evolution [31]. The availability of complete mitochondrial and nuclear genome sequences has made it possible to study the extent of gene duplication events and gene transfer events.

Gene duplication events in plant mitochondria

The plant mitochondrial genome is a circular double-stranded DNA molecule that encodes tRNAs, rRNAs, ribosomal proteins, and a portion of the enzymes used in respiration [32–33]. Plant mitochondrial genomes range in size from 200 kb to 2400 kb and are at least 10–100 times larger than animal mitochondrial genomes [34–35]. At 208 kb, *Brassica hirta* is one of the smallest and *Cucumis melo* at 2300 kb is the largest known mitochondrial genome in higher plants [6,34,36]. The plant mitochondrial genome has an unusually dynamic structure due to recombination between repeated sequences, which generates a population of molecules of different sizes and molecular configurations. The plant mitochondrial genome has a very low substitution rate, and its evolution is characterized by frequent structural rearrangements [36–37]. Several descriptive models have been proposed to explain the occurrence of deletion–duplication events in the plant mitochondrial genome [37–39].

The entire mitochondrial genome of rapeseed (*Brassica napus* L.), which contains a 2427 bp sequence as a direct repeat, was sequenced by Handa [40]. This sequence includes the first exon, the intron, and part of the second exon of the *cox2* gene. Due to duplication, two copies of *cox2* genes exist in rapeseed, although these copies diverge from each other 55 bp upstream of the stop codon. One copy (*cox2-1*) is homologous to other plant mitochondrial *cox2* genes, but the other copy (*cox2-2*) has an extension that shows no homology to any other sequence examined to date [40].

The cucumber (*Cucumis sativus*) has some unique attributes that make it a potential model system for mitochondrial transformation of higher plants. Microspores have relatively few, huge mitochondria, which have paternal transmission. The cucumber has unique mitochondrial mutations that result in strongly mosaic phenotypes [33]. Lilly and Havey investigated mitochondrial genome expansion within the cucurbits using hybridization to select mitochondrial sequences present in high copy numbers in cucumber and at low levels in watermelon (*Citrullus lanatus*). Lilly and Havey sequenced 15 clones to identify sequences repeated throughout the cucumber mitochondrial genome. On the basis of dot-blot hybridizations, seven repetitive DNA motifs account for over 13% of the cucumber mitochondrial genome, equaling over 50% of the size of the Arabidopsis mitochondrial genome. Sequence analysis of 136 kb of cucumber mitochondrial DNA revealed only 11.2% with significant homology to previously characterized mitochondrial sequences, 2.4% to chloroplast DNA, and 15% to the seven repetitive DNA motifs. The remaining 71.4% of the sequence was unique to the cucumber mitochondrial genome. These results demonstrate that the expanded cucumber mitochondrial genome is due, in part, to extensive duplication of short repetitive sequences, possibly by recombination and/or replication slippage [35].

There are two divergent copies of the chloroplast origin *rps13* gene, which encodes ribosomal protein S13, are found in the nucleus of the rosids Arabidopsis, Gossypium, and Glycine. One is *nucp-rps13*, which encodes chloroplast-imported RPS13; the other is *numit-rps13*, which encodes mitochondria-imported RPS13 [41]. The function of *numit-rps13* has been modified after gene duplication, and one could argue that *numit-rps13* has gained a new function. Subsequently *mt-rps13* was lost from mitochondrial DNA many times during the evolutionary history of rosids. Those organellar *rps13* genes in rosids provide a distinctive case of gene duplication involving the co-evolution of the nuclear and cytoplasmic genomes [41–42].

Gene transfer events between mitochondrial and nuclear genomes

In addition to the nuclear genome, plants have chloroplast and mitochondrial genomes, all of which undergo intra- and intergenomic recombination and rearrangement. The nuclear, chloroplast and mitochondrial genomes have been sequenced in some plants, such as Arabidopsis [32,43,44] and rice [9,45,46]. With the completion of those plant genome sequencing projects, it was possible to identify transfer events from the organelles to the nucleus on a whole-genome scale.

It is well known that mitochondria donated many genes to nuclear chromosomes during evolution. Rujan and Martin compared 3961 Arabidopsis nuclear protein-coding genes with the complete set of proteins from yeast and 17 reference prokaryotic genomes [47]. The degree of conservation in protein sequences in addition to lateral gene transfer between free-living prokaryotes pose substantial challenges to genome phylogenetics.

To date, several genes transferred from the mitochondrial genome to the nuclear genome have been identified in flowering plants. However, the mechanism of gene transfer events is only poorly understood [48]. Gene transfer events between the mitochondrial genome and the nuclear genome in plants are believed to often occur as single gene transfers through an RNA intermediate. The majority of mitochondria gene transfer events have been reported to range from hundreds to thousands of base pairs [49–51]. The sequence analysis of *Arabidopsis thaliana* chromosome 2 revealed a mitochondrial-to-nuclear DNA transfer of nearly the entire mitochondrial genome into the pericentric region on the short arm. The size of the mitochondrial DNA insert was about 270 kb [52]. However, DNA fiber-based fluorescence in situ hybridization analysis revealed that the mitochondrial DNA insert is about 620 kb, which is near the centromere on chromosome 2 [51].

A promiscuous nuclear sequence containing a mitochondrial DNA fragment was isolated from rice by Kubo and his colleagues [53], who found that the integration of the mitochondrial sequence into the nuclear genome was mediated by a DNA fragment, and that the nuclear sequence was transcribed and spliced, but it appeared to be a pseudogene. The mitochondrial sequence was integrated in an antisense orientation into the pre-existing *V-ATPase B* pseudogene, which can be transcribed and spliced. They suggested that the DNA transfer event may be a case of unsuccessful gene transfer from mitochondrion to nucleus (Fig. 1).

The *rps13* gene, encoding mitochondrial ribosomal protein S13, is normally present in the mitochondrial genome of higher plants, but is lacking from the *A. thaliana* mitochondrial genome [54–55]. Mollier et al. demonstrated that the nuclear gene encoding the plastid S13 has been partially duplicated in *A. thaliana*, such that

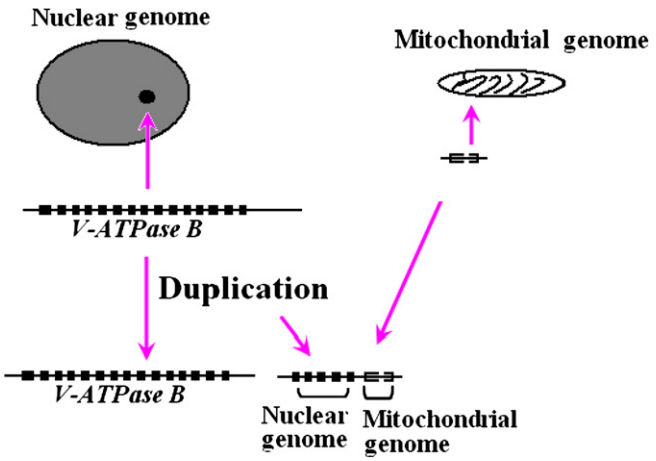

Fig. 1. The transfer of the *V-ATPase B* gene between the mitochondrial genome and the nuclear genome (modified from Kubo et al. [[53]]).

the copy has lost the exon encoding the plastid transit peptide and has acquired a sequence capable of encoding a mitochondrial targeting sequence. The mitochondrial S13 ribosomal protein has probably been replaced by its homologue from plastids in *A. thaliana*. The differences between the S13 gene sequence of the mitochondrion and the chloroplast suggest that the gene duplication occurred after the Brassicaceae arose but before the divergence of Arabidopsis and Brassica [55].

The complete mitochondrial genome of rice has been sequenced. Notsu et al. found that 6.3% and 13.4% of the mitochondrial genome sequence is derived from the plastid and the nuclear genome, respectively [9]. They demonstrated frequent and independent DNA sequence flow among the mitochondrial, plastid and nuclear genomes during the evolution of flowering plants, and this may account for the range of genetic variation observed between the mitochondrial genomes of higher plants [9].

Conclusion and perspective

The complete genome sequences, including the mitochondrial and nuclear genomes, of more and more organisms are becoming available, and this can be considered a major step forward toward exploiting the usefulness of mitochondrial genetic engineering technology. Earlier work showed that many gene transfers to the nuclear occurred during mitochondrial evolution, but there is no reliable estimate of the total number of genes that have been transferred.

Gene duplication is a fundamental process in the evolution of eukaryotic genomes. After duplication, one copy of a gene may undergo divergence in sequence, expression pattern, and function, and the rate of gene duplication is an important parameter in the study of evolution. During the past decade, the amount of sequence data (primarily DNA sequence data) has increased nearly 100-fold, and will continue to increase rapidly as the result of immense technical progress in DNA sequencing. Completely sequenced mitochondrial genomes are a valuable source of data for determining the evolutionary history of the organelle. Many mitochondrial enzymatic subunits are nuclear-encoded, cytoplasmically translated, and imported into the mitochondria. Gene duplication events in the mitochondrial genome and gene transfer events between the mitochondrial genome and the nuclear genome are important processes in the evolution of the eukaryotic cell.

Acknowledgments

Authors acknowledge the funding from Project for International Scientific and Technological Cooperation (Shanghai China-Alberta Canada); Shanghai Rising-Star Program (08QH14021); Hi-tech research and development program of China (2006AA10Z117; 2008AA10Z401). We are grateful to Prof. Max Cheng (Department of Plant Sciences, University of Tennessee, USA) for his useful comments on the manuscript.

References

[1] K. Henze, W. Martin, Evolutionary biology: essence of mitochondria, Nature 426 (2003) 127–128.
[2] H.M. McBride, M. Neuspiel, S. Wasiak, Mitochondria: more than just a powerhouse, Curr. Biol. 16 (2006) R551–R560.
[3] M. Schwartz, V.J. John, Paternal Inheritance of Mitochondrial DNA, N. Engl. J. Med. 347 (2002) 576–580.
[4] M.W. Gray, The endosymbiont hypothesis revisited, Int. Rev. Cytol. 141 (1992) 233–357.
[5] J. Maple, S.G. Møller, Plastid division: evolution, mechanism and complexity., Ann. Bot. 99 (2007) 565–579.
[6] J.D. Palmer, L.A. Herbon, Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence, J. Mol. Evol. 28 (1989) 87–97.
[7] G. Burger, D.V. Lavrov, L. Forget, B.F. Lang, Sequencing complete mitochondrial and plastid genomes, Nat. Protoc. 2 (2007) 603–614.
[8] P. Mereu, M. Palici di Suni, L. Manca, B. Masala, Complete nucleotide mtDNA sequence of Barbary sheep (*Ammotragus lervia*), DNA Seq. 19 (3) (2008) 241–245. Jun.
[9] Y. Notsu, S. Masood, T. Nishikawa, N. Kubo, G. Akiduki, M. Nakazono, A. Hirai, K. Kadowaki, The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants, Mol. Genet. Genomics 268 (2002) 434–445.
[10] Y. Sugiyama, Y. Watase, M. Nagase, N. Makita, S. Yagura, A. Hirai, M. Sugiura, The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants, Mol. Genet. Genomics 272 (2005) 603–615.
[11] S. Ohno, Evolution by Gene Duplication, Springer, New York, 1970.
[12] A. Hoshino, H.Y. Johzuka, S. Iida, Gene duplication and mobile genetic elements in the morning glories, Gene 265 (2001) 1–10.
[13] J.S. Taylor, Y. Van de Peer, A. Meyer, Genome duplication, divergent resolution and speciation, Trends Genet. 17 (2001) 299–301.
[14] M. Lynch, J.S. Conery, The evolutionary fate and consequences of duplicate genes, Science 290 (2000) 1151–1155.
[15] C. Simillion, K. Vandepoele, M.C. Van Montagu, M. Zabeau, Y. Van de Peer, The hidden duplication past of *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. USA 99 (2002) 13627–13632.
[16] X. He, J. Zhang, Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution, Genetics 169 (2005) 1157–1164.
[17] M. Lynch, A. Force, The probability of duplicate gene preservation by subfunctionalization, Genetics 154 (2000) 459–473.
[18] B. Causier, R. Castillo, J. Zhou, R. Ingram, Y. Xue, S.Z. Schwarz, B. Davies, Evolution in action: following function in duplicated floral homeotic genes, Curr. Biol. 15 (2005) 1508–1512.
[19] S.M. Shimeld, Gene function, gene networks and the fate of duplicated genes, Semin. Cell Dev. Biol. 10 (1999) 549–553.
[20] J.S. Taylor, J. Raes, Duplication and divergence: the evolution of new genes and old ideas, Annu. Rev. Genet. 38 (2004) 615–643.
[21] R.V. Samonte, E.E. Eichler, Segmental duplications and the evolution of the primate genome, Nat. Rev. Genet. 3 (2002) 65–72.
[22] K.H. Wolfe, W.H. Li, Molecular evolution meets the genomics revolution, Nat. Genet. 33 (2003) 255–265.
[23] I. Hurley, M.E. Hale, V.E. Prince, Duplication events and the evolution of segmental identity, Evol. Dev. 7 (2005) 556–567.
[24] M. Sémon, K.H. Wolfe, Consequences of genome duplication, Curr. Opin. Genet. Dev. 17 (2007) 505–512.
[25] R. Koszul, S. Caburet, B. Dujon, G. Fischer, Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments, EMBO J. 23 (2004) 234–243.
[26] R. Koszul, B. Dujon, G. Fischer, Stability of large segmental duplications in the yeast genome, Genetics 172 (2006) 2211–2222.
[27] J.F. Wendel, Genome evolution in polyploids, Plant Mol. Biol. 42 (2000) 225–249.
[28] G. Blanc, K.H. Wolfe, Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes, Plant Cell 16 (2004) 1667–1678.
[29] Q. Yuan, J. Hill, J. Hsiao, K. Moffat, S. Ouyang, Z. Cheng, J. Jiang, C.R. Buell, Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion, Mol. Genet. Genomics 267 (2002) 713–720.
[30] B.F. Lang, M.W. Gray, G. Burger, Mitochondrial genome evolution and the origin of eukaryotes, Annu. Rev. Genet. 33 (1999) 351–397.

[31] W. Martin, O. Deusch, N. Stawski, N. Grünheit, V. Goremykin, Chloroplast genome phylogenetics: why we need independent approaches to plant molecular evolution, Trends Plant Sci. 10 (2005) 203–209.
[32] M. Unseld, J.R. Marienfeld, P. Brandt, A. Brennicke, The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides, Nat. Genet. 15 (1997) 57–61.
[33] M.J. Havey, J.W. Lilly, B. Bohanec, G. Bartoszewski, S. Malepszy, Cucumber: a model angiosperm for mitochondrial transformation?, J Appl. Genet. 43 (2002) 1–17.
[34] B.L. Ward, R.S. Anderson, A.J. Bendich, The mitochondrial genome is large and variable in a family of plants (cucurbitaceae), Cell 25 (1981) 793–803.
[35] J.W. Lilly, M.J. Havey, Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber, Genetics 159 (2001) 317–328.
[36] J.D. Palmer, L.A. Herbon, Unicircular structure of the *Brassica hirta* mitochondrial genome, Curr. Genet. 11 (1987) 565–570.
[37] B. Albert, B. Godelle, P. Gouyon, Evolution of the plant mitochondrial genome: dynamics of duplication and deletion of sequences, J. Mol. Evol. 46 (1998) 155–158.
[38] I.D. Small, R. Suffolk, C.J. Leaver, Evolution of plant mitochondrial genomes via substoichiometric intermediates, Cell 58 (1989) 69–76.
[39] C.M.R. Fauron, B. Moore, M. Casper, Maize as a model of higher plant mitochondrial genome plasticity, Plant Sci. 112 (1995) 11–32.
[40] H. Handa, The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (Brassica napus L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*, Nucleic Acids Res. 31 (2003) 5907–5916.
[41] S.L. Liu, K. Adams, Molecular adaptation and expression evolution following duplication of genes for organellar ribosomal protein S13 in rosids, BMC Evol. Biol. 8 (2008) 25.
[42] K.L. Adams, Y.L. Qiu, M. Stoutemyer, J.D. Palmer, Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution, Proc. Natl. Acad. Sci. USA 99 (2002) 9905–9912.
[43] S. Sato, Y. Nakamura, T. Kaneko, E. Asamizu, S. Tabata, Complete structure of the chloroplast genome of *Arabidopsis thaliana*, DNA Res. 6 (1999) 283–290.
[44] The Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, Nature 408 (2000) 796-815.
[45] J. Hiratsuka, H. Shimada, R. Whittier, T. Ishibashi, M. Sakamoto, M. Mori, The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals, Mol. Gen. Genet. 217 (1989) 185–194.
[46] J. Yu, S. Hu, J. Wang, G.K. Wong, S. Li, B. Liu, A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica), Science 296 (2002) 79–92.
[47] T. Rujan, W. Martin, How many genes in Arabidopsis come from cyanobacteria? An estimate from 386 protein phylogenies, Trends Genet. 17 (2001) 113–120.
[48] U. Bergthorsson, K.L. Adams, B. Thomason, J.D. Palmer, Widespread horizontal transfer of mitochondrial genes in flowering plants, Nature 424 (2003) 197–201.
[49] C.W. Sun, J. Callis, Recent stable insertion of mitochondrial DNA into an Arabidopsis polyubiquitin gene by nonhomologous recombination, Plant Cell 5 (1993) 97–107.
[50] J.L. Blanchard, G.W. Schmidt, Pervasive migration of organellar DNA to the nucleus in plants, J. Mol. Evol. 41 (1995) 397–406.
[51] R.M. Stupar, J.W. Lilly, C.D. Town, Z. Cheng, S. Kaul, C.R. Buell, J. Jiang, Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats, Proc. Natl. Acad. Sci. USA 98 (2001) 5099–5103.
[52] X. Lin, S. Kaul, S. Rounsley, T.P. Shea, M.I. Benito, C.D. Town, C.Y. Fujii, T. Mason, Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*, Nature 402 (1999) 761–768.
[53] N. Kubo, M. Takano, M. Nishiguchi, K. Kadowaki, Mitochondrial sequence migrated downstream to a nuclear V-ATPase B gene is transcribed but non-functional, Gene 271 (2001) 193–201.
[54] R. Kumar, J. Drouaud, M. Raynal, I. Small, Characterization of the nuclear gene encoding chloroplast ribosomal protein S13 from *Arabidopsis thaliana*, Curr. Genet. 28 (1995) 346–3520.
[55] P. Mollier, B. Hoffmann, C. Debast, I. Small, The gene encoding *Arabidopsis thaliana* mitochondrial ribosomal protein S13 is a recent duplication of the gene encoding plastid S13, Curr. Genet. 40 (2002) 405–409.