Plant Molecular Biology (2006) 61:1–12
DOI 10.1007/s11103-005-5773-3

Copy number of bean mitochondrial genes estimated by real-time PCR does not correlate with the number of gene loci and transcript levels

Magdalena Woloszynska[1], Beata Kmiec[1], Pawel Mackiewicz[2] and Hanna Janska[1,*]
[1]*Department of Cell Molecular Biology, Institute of Biochemistry and Molecular Biology, University of Wroclaw, ul. Przybyszewskiego 63/77, Wroclaw 51-148, Poland (*author for correspondence; e-mail Hanna.Janska@ibmb.uni.wroc.pl);* [2]*Department of Genomics, Institute of Genetics and Microbiology, University of Wroclaw, ul. Przybyszewskiego 63/77, Wroclaw 51-148, Poland*

Received 18 October 2005; accepted in revised form 8 December 2005

Key words: gene copy number, *Phaseolus vulgaris*, plant mitochondria, real-time PCR, steady-state transcript level, transcription regulation

Abstract

Structural rearrangements characteristic for plant mitochondrial DNA often result in the appearance of genes in new genomic environments. The determination of the real number of gene copies is difficult since the *in vivo* structure of plant mitochondrial genomes is questionable. It is still uncertain whether the gene copy number regulates transcription in plant mitochondria. Using the real-time PCR technique we have quantified the copies of mitochondrial genes and their transcripts in four related *Phaseolus vulgaris* lines. We found low intergenomic variation both in the copy number of particular genes and the abundance of their transcripts, while the intragenomic differences between copy numbers and transcripts levels of various genes were much higher. Furthermore, we found that the appearance of a gene in a new location is not correlated with a proportional increase in its copy number estimated by real-time PCR. This observation seems to result from gene dosage compensation which is probably associated with the multimolecular plant mitochondrial genome structure and particularly with the recombinogenic activity of large repeats. Based on the relative gene copy numbers we propose the existence of two types of *Phaseolus* mitochondrial genomes: one associated with fertility and the other inducing cytoplasmic male sterility. We also show that there is no correlation between the observed number of copies of the analyzed genes and the steady-state level of their transcripts.

Abbreviations: atp9, gene encoding subunit 9 of ATP synthase; *atpA*, gene encoding alpha-subunit of ATP synthase; CMS, cytoplasmic male sterility; *cob*, gene encoding cytochrome *b*; *coxIII*, gene encoding subunit 3 of cytochrome oxidase; mtDNA, mitochondrial DNA; pvs, *Phaseolus vulgaris* sterility sequence; *rrn18*, gene encoding 18S rRNA; r_s, The Spearman's rank correlation coefficient; *trnfM*, gene encoding initiator methionine tRNA

Introduction

Plant mitochondrial genomes are well known for their structural variability observed even between closely related species. This variability is predominantly driven by recombinationally active repeated sequences (Small *et al.*, 1987; Fauron *et al.*, 1995; Mackenzie and McIntosh, 1999). Selective amplification or suppression of preexisting sequences can also affect genome structure (Small *et al.*, 1987; Janska *et al.*, 1998; Woloszynska *et al.*, 2001). Rearrangements like

inversions, duplications or deletions have often been identified by physical mapping and/or sequencing of plant mitochondrial genomes (Clifton *et al*., 2004; Sugiyama *et al*., 2005). If the duplicated or deleted DNA fragment contains a coding sequence, then the copy number of respective gene/genes may be affected. In the light of the high polyploidy of plant mitochondrial genomes it is intriguing if the appearance of a gene in a new genomic environment correlates with a proportional increase in its actual copy number or if mechanisms exist that can buffer the resulting variations. Since our knowledge on the *in vivo* structure of plant mtDNA is still poor, it is difficult to predict the actual copy number of a certain gene based on the number of genomic environments inferred from a physical map. Consequently, quantitative proportions among various genes in plant mitochondria can not be precisely predicted.

Accurate determination of plant mitochondrial gene copy number seems to be an important issue considering that it could influence gene expression at least in some systems. Specific examples of altering gene expression by gene amplification or suppression are connected with stoichiometric shifts in heteroplasmic population of common bean mitrochondrial genomes (Janska *et al*., 1998). In this species, the appearance of cytoplasmic male sterile line (CMS-Sprite) was caused by amplification of the male sterility inducing sequence *pvs* and reversion to fertility was achieved by *pvs* suppression. The amplification and suppression events were followed by expression of a *pvs*-encoded protein or its silencing, respectively. Hedtke *et al*. (1999) reported that the lack of photosynthetically active chloroplasts in a barley mutant is correlated with a global 3-fold increase of mitochondrial gene copy number and increased mtRNA levels. A direct relationship between maize and *Brassica hirta* mitochondrial gene copy numbers and their transcriptional rates was described by Muise and Hauswirth (1995). Contrary to those reports, Mulligan and co-authors (1991) found that the steady-state abundance of transcripts of nine maize genes did not correlate with either the transcriptional activity or the copy number of those genes. Consequently, the dependence between plant mitochondrial gene copy number and expression level is still poorly understood. All reports mentioned above employed hybridization techniques for quantification of nucleic acids. In this study we applied the most accurate method available today, real-time PCR, to verify how the copy numbers and steady-state transcript levels of mitochondrial genes coding for proteins or RNAs differ within and between genomes of four *Phaseolus vulgaris* lines (POP, GO8063, CMS-Sprite and WPR−3). We chose these lines because they are related to each other and their mitochondrial genomes are mapped (Janska and Mackenzie, 1993; Janska *et al*., 1998). The GO8063 line is the fertile progenitor of the sterile CMS-Sprite line which combines the cytoplasm of GO8063 and nuclear genome of Sprite. GO8063 originated from a cross of the fertile POP and NEP-2 lines. The WPR-3 line was derived from the CMS-Sprite line via spontaneous reversion to fertility. As we already mentioned, the reversion was associated with a stoichiometric suppression of mtDNA molecules containg *pvs* (Janska *et al*., 1998). It is believed that the GO8063 and CMS-Sprite lines have the same mtDNAs but different nuclear background, while CMS-Sprite and WPR-3 possess identical nuclear genomes but vary with respect to mitochondrial genetic information.

In this work we show that despite differences in physical maps mitochondrial genomes may comprise similar copy numbers of individual genes. Basing on the relative copy numbers of six mitochondrial genes, two types of *Phaseolus* mitochondrial genomes can be distinguished – one associated with fertility (POP, WPR-3) and the other with cytoplasmic male sterility (GO8063, CMS-Sprite). Furthermore, we found no significant correlation between the copy number of mitochondrial genes and the steady-state abundance of respective transcripts.

Methods

Plant materials and isolation of nucleic acids

Common bean (*Phaseolus vulgaris*) accession lines POP, GO8063, CMS-Sprite and WPR−3 were as described by Janska *et al*. (1998). Mitochondrial DNA and RNA were purified from 7-day-old etiolated seedlings. Mitochondrial DNA and mitochondria for RNA preparation were obtained according to the procedure described by

Mackenzie *et al*. (1988). Mitochondrial RNA was extracted from mitochondria using RNeasy Plant mini Kit (Qiagen) with on-column DNase digestion. Total RNA was isolated from leaves using the same procedure. Total genomic DNA was extracted from etiolated seedlings, leaves or stalks of growth chamber-grown plants with DNeasy Plant Mini Kit (Qiagen). At least three independent DNA or RNA preparations were performed for each *Phaseolus* line. From each preparation the gene copy number or transcript levels were analyzed at least three times giving a minimum of nine replicates.

Primers and hybridization probes for real-time PCR

Primers and hybridization probes listed in Table 1 were designed basing on the *atpA* sequence of *P. vulgaris* available in the BLAST database (GeneBank Accession Number M64246) or sequences of gene fragments obtained in our laboratory (*atp9*, *cob*, *coxIII*, *rrn18*, *trnfM*). To select the sequences of all oligonucleotides the Accelrys DS Gene 1.5 program was used. The melting temperatures of primers were in the range of 51.4–52.8 °C, the melting temperatures of probes were appropriately higher (58.1–61.8 °C). Each gene-specific pair of probes consisted of an upstream probe labelled with fluorescein at the 3′ end and a downstream probe labelled with LC Red 640 at the 5′ terminus. To allow the fluorescence resonance energy transfer between fluorescein and LC Red 640, the probes were designed to hybridize in tandem orientation to the same strand of the PCR products with a 1- or 2-nucleotide gap between the dyes. The PCR products varied in length in the

Table 1. Sequences of primers and hybridization probes used in real-time PCR.

Gene	Primers	Hybridization probes
atpA	Forward GGAATTGTGGTCTTTGGTAGTG	Probe 1 CGCACTGGATCTATTGTGGATGTTCCTGC-Fluorescein
	Reverse CCCCTCTTCCATCAATAGGTAC	Probe 2 LC Red 640-GAAAGGCTATGCTAGGGCGTGTGGTC-Phosphate
*atpA**	Forward AGGTCGTGGAGTGTATTC	Probe 1 TTCTTTAAAAAAAGATAGGAATCAGAGATGGCAT-Fluorescein
	Reverse AGTTGGAAGAAGTCAAATTTGG	Probe 2 LC Red 640-ATAAGATCTTGATATGGCTGGCTGCGGTG-Phosphate
atp9	Forward TACAATTGCTTCAGCGGGA	Probe 1 TTCATTCCGTGGCAAGAAATCCATCATTGGC-Fluorescein
	Reverse TTAATGCGAACAAGGCAATAG	Probe 2 LC Red 640-AACAGTTATTCGGATATGCAATCCTGGGC-Phosphate
cob	Forward CGGTTCGTTAGCTGGTATTTG	Probe 1 TGCATTACACACCTCATGTGGATCTAGCTT-Fluorescein
	Reverse CATGCATATAACGGAGCAACC	Probe 2 LC Red 640-CAACAGCGTAGAACACGTTATGAGAGATG-Phosphate
coxIII	Forward GGAGCAGCCGTAACTTGG	Probe 1 GGGAAGGAAAAACGAGCAGTTTACGCTTTA-Fluorescein
	Reverse CGAAATAGTGAAGGGTGCT	Probe 2 LC Red 640-AGCTACCGTTTCACTGGCTCTAGTATTCA-Phosphate
rrn18	Forward GCGCTGTTTGATGAGCCT	Probe 1 GTAAAGGCTGACCAAGCCAATGATGCTTA-Fluorescein
	Reverse AAGATTCCCCACTGCTGC	Probe 2 LC Red 640-TGGTCAAAACGGATGATCAGCCACACTG-Phosphate
trnfM	Forward AGAAAGAGTAGGTGAAGAAGCG	Probe 1 CGTCAGGCTCATAATATGAAGATTGCAGGT-Fluorescein
	Reverse GGTTCAAGTTCAGACTTGGAGA	Probe 2 LC Red 640-GAATCCTGCCCCCGCCATGTCTTTC-Phosphate

The products of PCR and RT-PCR were gel-analyzed yielding a single band of expected size for each pair of primers. The DNA fragments represented by these bands were excised and sequenced to confirm their identity.
atpA – both variants of the *atpA* gene.
*atpA** – only this variant of the gene which is common to all *Phaseolus* lines analyzed.

range of 145–155 bp, the only exception was the *trnfM* gene which is very short and thus the amplicon size was only 124 bp.

Real-time PCR analysis of gene copy number and transcript levels

Gene copy numbers and transcript levels were assayed with the LightCycler 2.0 instrument using respectively the FastStart DNA Master Hybridization Probes kit or RNA Master Hybridization Probes kit for one-step RT-PCR (Roche Diagnostics). Reactions were carried out in a total volume of 10 μl with a final concentration of 2.5 mM $MgCl_2$ (FastStart DNA Master Hybridization Probes kit) or 3.25 mM $Mn(OAc)_2$ (RNA Master Hybridization Probes kit). The protocol for gene copy number estimation consisted of three programs: denaturation, 95 °C for 10 min; amplification, 40 cycles at 95 °C for 10 s, 52 °C for 10 s with single data acquisition, 72 °C for 6 s; cooling, 40 °C for 30 s. For transcript level quantification the following four programs were used: reverse transcription, 61 °C for 20 min; denaturation, 95 °C for 30 s; amplification, 95 °C for 1 s, 52 °C for 15 s with single data acquisition, 72 °C for 10 s with a transition rate of 2 °C/s; cooling as previously.

The intragenomic variation of gene copy numbers and steady-state transcript levels were estimated basing on the following equation: $N = N_0 \times E^n$, where N is the copy number after n cycles, N_0 is the initial copy number and E is the amplification efficiency. N_0 can be calculated from this equation assuming that at the cycle number equal to Cp (crossing point estimated by real-time PCR) the N value is identical for all samples in the experiment. The values of amplification efficiency were calculated for each gene by the LightCycler v 4.0 software basing on the standard curve. The standard curves were generated using six serial 2-fold dilutions of the sample that reached exponential amplification at the earliest cycle. The copy numbers and steady-state transcript levels of the analysed genes (N_{0g}) were expressed relative to the *cob* gene/transcript copy number which was set to 1 ($N_{0cob} = 1$). Consequently, $N_{0g} = E_{cob}^{Cpcob} / E_g^{Cpg}$.

An alternative quantification approach used in this study was applied to estimate the relative gene copy or transcript abundances for chosen mitochondrial genes in different *Phaseolus* lines. For this purpose Relative Quantification analysis using the second derivative maximum method of the LightCycler v 4.0 software was used. One of the POP line samples served as the calibrator and the *cob* gene was used as the reference.

The analysed sets fulfilled the assumption about normal distribution and homogeneity of variance so appropriate analysis of variance (ANOVA) was carried out to define the significance of differences in the normalized number of copies of particular genes and their transcripts between *Phaseolus* lines (intergenomic variation) and within them (intragenomic variation). The Scheffe's and Tukey's (HSD) tests were applied in *post hoc* analysis. Moreover, Student's *t*-test was used for the comparison of observed and predicted copy numbers for particular genes.

The dendrogram classifying *Phaseolus* lines according to the numbers of copies of all genes was constructed by the Unweighted Pair-Group Method Arithmetic Averages UPGMA (Sneath and Sokal, 1973) with the assumption of Euclidean distance between objects.

Results

Intragenomic variation in mitochondrial gene copy number established by real-time PCR

We applied real-time PCR to estimate relative copy numbers of six genes within mitochondrial genomes of four related *Phaseolus vulgaris* lines. Mitochondrial DNA isolated from 7-day etiolated seedlings was used as a template. We analyzed four genes coding for subunits of the respiratory chain complexes: *cob*, *coxIII*, *atpA*, *atp9*, and two RNA-encoding genes: *rrn18* and *trnfM*.

In the simplest approach the number of gene copies can be considered as the number of different gene locations on physical maps (Table 2, ''R1 and R2 activity not considered''). However, given that bean genomes contain two major families of recombinationally active large repeats, R1 and R2 (Janska and Mackenzie, 1993; Janska *et al*., 1998), this not necessarily reflects the real number of DNA molecules on which a certain gene resides. To include the R1 and R2 recombination activity in our considerations we performed two types of theoretical predictions of gene copy numbers. First, we have calculated the number of gene copies that could be predicted if the recombinational activity

Table 2. Predicted and real-time PCR-estimated values of *Phaseolus vulgaris* mitochondrial gene copy numbers in lines POP, WPR−3, GO8063 and CMS-Sprite.

Number of gene copies												
Predicted										Real-time PCR-estimated values		
Gene	R1 and R2 activity not considered (number of different gene locations on physical maps)			R1 and R2 activity considered								
				Flanking repeats			Circular chromosomes					
	POP	WPR-3	GO8063/ CMS-Sprite	POP	WPR-3	GO8063/ CMS-Sprite	POP	WPR−3	GO8063/ CMS-Sprite	POP	WPR-3	GO8063/ CMS-Sprite
cob	1	1	1	1	1	1	1	1	1	1	1	1 / 1
atpA	1	1	2	1	0.4	1.1	1	0.5	0.6	0.45	0.52	0.61 / 0.74
*atpA**	1	1	1	1	0.4	0.4	1	0.5	0.5	0.45	0.41	0.42 / 0.35
atp9	2	3	4	2	1.3	1.8	2	1.5	1.6	1.55	1.92	2.12 / 2.33
coxIII	1	1	1	1	0.7	0.9	1	0.5	0.5	0.66	0.61	0.70 / 0.67
rrn18	1	1	2	1	1	1.7	1	1	1.1	0.71	0.73	0.83 / 0.94
trnfM	2	3	3	2	2	2.7	2	2	2.1	2.64	3.08	3.94 / 3.93

R1 and R2 – recombinatinally active large repeats (Janska and Mackenzie, 1993; Janska *et al*., 1998). All values were normalized relative to the copy number of the *cob* gene that was set to 1.
atpA – both variants of the *atpA* gene.
*atpA** – only this variant of the gene which is common to all *Phaseolus* lines analyzed.

of two repeats, flanking each of the analyzed genes, was taken into account (Table 2, ''Flanking repeats''). However, this type of prediction misses the influence of other recombinationally active repeats on the copy number of genes. As already mentioned, the structure and organization of plant mitochondrial genomes are still uncertain but models of circular chromosomes have been proposed for all four analyzed *Phaseolus vulgaris* genomes (Janska and Mackenzie, 1993; Janska *et al*., 1998). Based on those models we have determined the numbers of the gene copies that could be generated via intra- and intermolecular recombinations of R1 and R2 repeats located within the chromosomes (Table 2, ''Circular chromosomes''). The three sets of values obtained basing on the theoretical predictions described above were compared with the values resulting from real-time PCR (Table 2, ''Real-time PCR estimated values'').

In the intragenomic analysis the copy numbers of the investigated genes were estimated in each genome by real-time PCR and normalized by the *cob* copy number that was set to 1 (Figure 1). We chose the *cob* gene for this purpose because this gene was present in only one location on the physical maps constructed for the mitochondrial genomes of all analyzed *Phaseolus vulgaris* lines (Janska and Mackenzie, 1993; Janska *et al*., 1998). In all lines the patterns of the gene copy numbers were similar: the *atpA*, *coxIII* and *rrn18* genes were less numerous than *cob*, while *atp9* and particularly *trnfM* were significantly more abundant. The low abundance of *atpA* and *coxIII* was unexpected considering the number of copies of these genes calculated basing on their locations on physical maps (Figure 1 and Table 2). The *coxIII* gene, similarly to *cob*, is always present in one copy whereas the *atpA* gene is present in one (POP, WPR-3) or two (GO8063 and CMS-Sprite) copies on the maps (Janska and Mackenzie, 1993; Janska *et al*., 1998). However, when the R1 and R2 recombinational activity was taken into account in the predictions, the relatively low representation of the *coxIII* and *atpA* genes relative to *cob* became easier to explain (Table 1). Unlike the *atpA* and *coxIII* genes, the low observed copy number of *rrn18* was not expected from any of the predictions. The higher *atp9* and *trnfM* abundances compared to *cob* was predicted by all three types of theoretical calculations. However, it should be noted that the ratios of the real-time PCR estimated copy numbers of *atp9* and *trnfM* to the *cob* gene remain usually significantly different from the values suggested by the predictions (Table 2). For example, according to all theoret-

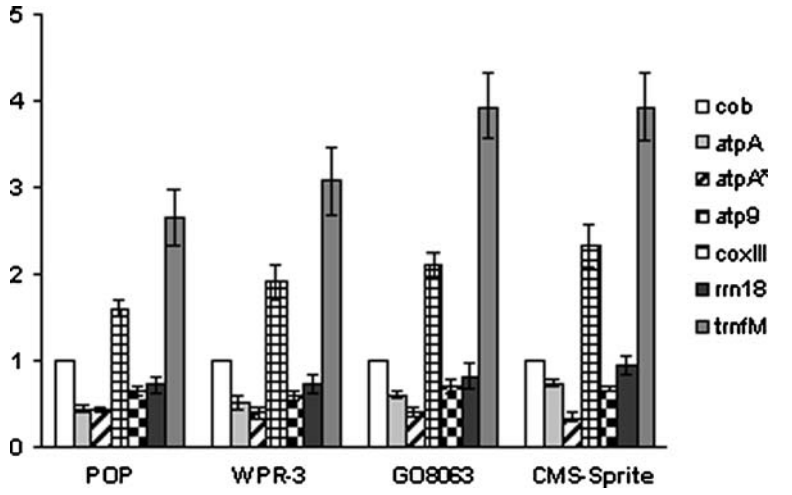

Figure 1. Intragenomic analysis of gene copy numbers. The *atpA*, *atp9*, *cob*, *coxIII*, *rrn18* and *trnfM* gene copies were quantified in mitochondrial genomes of four *Phaseolus vulgaris* lines (POP, WPR−3, GO8063 and CMS-Sprite) by real-time PCR using fluorescently labeled probes for product detection. The results are normalized to the copy number of *cob* that was set to 1. The *atpA** abbreviation indicates only this variant of the *atpA* gene which is common to all *Phaseolus* lines analyzed.

ical estimates, the *atp9* gene is present in two copies in the POP line (Janska and Mackenzie, 1993; Janska *et al*., 1998), but the relative ratio of this gene to the *cob* gene was established experimentally as 1.55. In the case of *trnfM* almost all predictions significantly underestimated the observed values. The differences between the predicted and real-time PCR estimated gene copy numbers were statistically insignificant ($P > 0.48$) only in few cases: for *trnfM* in WPR-3 when R1 and R2 activity was not considered, and for the *atpA* gene in the WPR-3 and GO8063 genomes when the proposed circular chromosomes were considered in the theoretical predictions. From our results we conclude that finding a gene in more than one copy by mapping (sequencing) is not related to a proportional increase of the real copy number of the gene. Even if the activity of large recombinogenic repeats is taken into consideration, the actual gene copy numbers cannot be predicted precisely.

Despite the lack of a statistically significant correspondence between real-time PCR estimated and predicted copy numbers for particular genes we performed correlation analysis to establish which of the theoretical predictions presents the best approximation of the overall observed gene copy numbers. This time, however, we did not consider particular genes but rather the global relations between the genes copy numbers expressed by the Spearman's rank correlation coefficient (r_S). We found that all three theoretical predictions are the same for the POP genome and approximate the observed genes copy numbers rather poorly ($r_S = 0.79$; $P = 0.03$). However, for the remaining genomes the correlation coefficients between the theoretical and experimentally estimated values increase in the following order: ''R1 and R2 activity not considered'' (WPR-3: $r_S = 0.79$, $P = 0.03$; GO8063: $r_S = 0.64$, $P = 0.12$; CMS-Spirite: $r_S = 0.73$, $P = 0.06$), ''Flanking repeats'' (WPR-3: $r_S = 0.98$, $P = 8.3 \cdot 10^{-5}$; GO8063: $r_S = 0.82$, $P = 0.02$; CMS-Spirite: $r_S = 0.89$, $P = 0.007$) and ''Circular chromosomes'' (WPR-3: $r_S = 0.95$, $P = 8.4 \cdot 10^{-4}$; GO8063: $r_S = 0.90$, $P = 0.006$; CMS-Spirite: $r_S = 0.95$, $P = 8.1 \cdot 10^{-4}$). The last type of theoretical prediction presents the best match with the observed gene copy numbers for most *Phaseolus vulgaris* genomes (except POP).

As we mentioned, the *atpA* gene is present in one or two genomic environments (locations on the physical map) depending on the *Phaseolus vulgaris* genome. One environment is common to all investigated bean lines, while the second version of *atpA* is linked to the sterility-inducing *pvs* sequence and resides only in the genomes of GO8063 and CMS-Sprite (Janska and Mackenzie, 1993). To evaluate the contribution of each *atpA* environment to the total copy number of this gene, two pairs of primers and hybridization probes

were used. One set of oligonucleotides allowed us to amplify and quantify only this version of the gene which is common to all *Phaseolus vulgaris* lines analyzed (*atpA**) while the second set of oligos was designed to recognize both variants of the *atpA* gene (*atpA*). As expected, the results of *atpA** and *atpA* quantification were identical in POP or nearly identical in WPR-3 (the observed difference was not statistically significant, $P > 0.08$). The differences were, however, statistically significant ($P < 0.0002$) for CMS-Sprite but, surprisingly, not for GO8063 ($P > 0.05$). Consequently, the contribution of the additional *atpA* gene copy (*atpA* minus *atpA**) seems to be higher in CMS-Sprite than in GO8063. As described below, this finding is further supported by intergenomic analysis. Finally, the greatest difference in the gene copy number (6.5-fold) in the intragenomic analysis was observed in GO8063 between *atpA* and *trnfM*.

Despite the mentioned similarities in the gene copy number patterns in the investigated genomes, clustering analysis of the results presented in Figure 1 allowed us to distinguish two types of genomes (Figure 2). Two fertility-associated genomes (POP and WPR-3) formed one set, and the two cytoplasmic male sterility-inducing genomes (GO8063 and CMS-Sprite) formed the second one characterized by relatively higher *atpA*, *atp9* and *trnfM* copy numbers. The method groups genomes according to their similarity based on the overall patterns in the numbers of all gene copies. To better understand this observation we compared the gene copy numbers among the four *Phaseolus* genomes using Relative Quantification analysis available in the LightCycler software.

Quantification of intergenomic differences in relative gene copy number

This analysis requires a calibrator genome (one of defined or arbitrarily assumed gene copy numbers) and a reference gene (whose copy number is assumed to be constant in all analyzed genomes). The reference gene is included in the experiment to eliminate the PCR amplification differences resulting from unequal DNA template content. The mitochondrial genome of the POP line was arbitrarily chosen as the calibrator for our study, meaning that the copy numbers of all analyzed genes were set to one in the POP line. As the reference we used the *cob* gene. Consequently, the resultant gene copy numbers were expressed relative to that in the POP line and normalized relative to the content of the *cob* gene in the analyzed DNA samples.

This new type of analysis (shown in Figure 3) supports the observation from the intragenomic analysis that the POP genome is similar to WPR-3

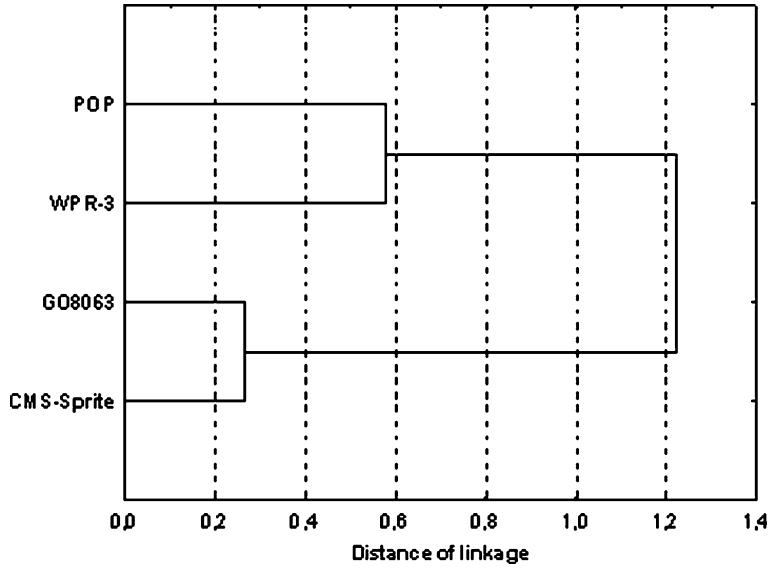

Figure 2. Dendrogram classifying *Phaseolus vulgaris* lines according to the numbers of copies of all genes, constructed by UPGMA with the assumption of Euclidean distance between objects.

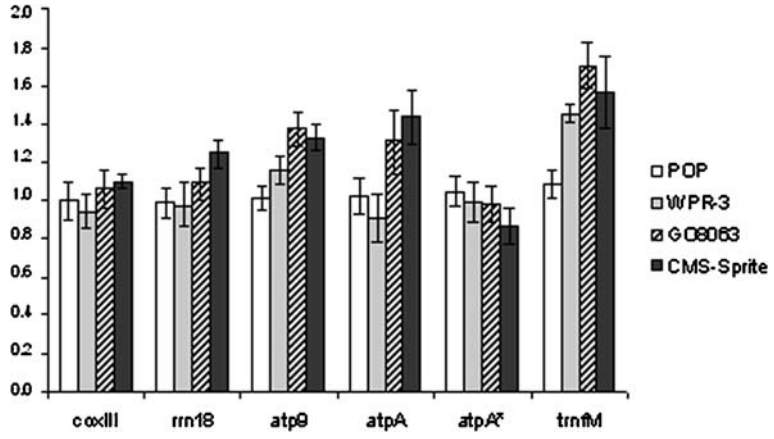

Figure 3. Intergenomic analysis of gene copy numbers. The copy numbers of the *atpA*, *atp9*, *cob*, *coxIII*, *rrn18* and *trnfM* genes were quantified, normalized to the copy number of *cob* (set to 1) and compared between genomes of four *Phaseolus vulgaris* lines (POP, WPR−3, GO8063 and CMS-Sprite). The genome of the POP line was arbitrarily chosen as the calibrator, e.g. the copy numbers of all genes in this line were set to 1. The comparison was performed using Relative Quantification analysis with the second derivative maximum method of the LightCycler v 4.0 software.

but differs from GO8063 and CMS-Sprite. The differences in the copy number between the fertility associated genome (POP, WPR-3) and the cytoplasmic male sterility-inducing genomes (GO8063, CMS-Sprite) are not the same for all genes. The copy numbers of *coxIII* and *atpA** compared between genomes are quite similar, while the copy numbers of the *rrn18*, *atp9* and especially *atpA* genes differ in individual *Phaseolus vulgaris* lines (Figure 3). An unexpected observation from this analysis was a small difference between the genomes of CMS-Sprite and GO8063 which previously were assumed identical (Janska and Mackenzie, 1993). The mitochondrial genome of CMS-Sprite contains a statistically significant higher copy number of *rrn18* ($P < 0.009$) compared to GO8063. It also has a higher total copy number of *atpA* ($P > 0.18$) and a lower one of the *atpA** gene. However, in the latter case the difference in copy number is at the limit of statistical significance ($P = 0.05$).

Intergenomic differences in atpA copy numbers do not depend on plant organ and are not influenced by light

All the data presented in the two former chapters were obtained using mtDNA isolated from etiolated seedlings. We wanted to check if the results would remain the same if total genomic DNA was used instead, and also if the results would vary depending on the plant organ or light conditions. For this purpose we reproduced the *atpA* quantification in the four *Phaseolus* lines but this time we used as templates samples of total genomic DNA obtained from etiolated stalks as well as from green stalks or leaves. Regardless of the type of the DNA (mitochondrial or genomic) used as the template or plant organs from which the DNA was obtained, the differences in the *atpA* copy number observed for the analyzed *Phaseolus* lines were virtually identical (data not shown).

Steady-state levels of transcripts are not related to gene copy number

The next aim of our study was to verify whether the observed variation of gene copy numbers is correlated with the steady-state levels of respective transcripts. We chose *cob*, *rrn18*, *atpA*, and *atp9* genes for this analysis.

In an intragenomic analysis we observed great differences in the accumulation of transcripts of the respective genes (Figure 4). The most plentiful were the *rrn18* transcripts – 60–90 times more abundant than the *cob* mRNA, even though in all *Phaseolus vulgaris* lines the gene copy number of *rrn18* was lower than the *cob* gene copy number (Figures 1 and 4). There was also no statistically significant correlation between the steady-state

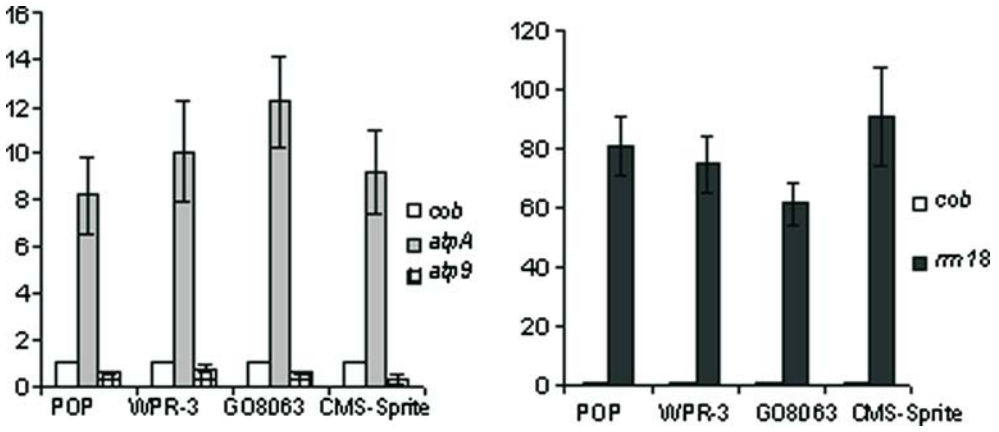

Figure 4. Intragenomic analysis of transcript levels. The steady-state levels of transcripts of the *atpA*, *atp9*, *cob* and *rrn18* mitochondrial genes were quantified by one-step real-time RT-PCR in four *Phaseolus vulgaris* lines: POP, WPR−3, GO8063 and CMS-Sprite. The results are normalized to the transcript level of *cob* that was set to 1.

level of transcripts of *atpA*, *atp9* and *rrn18* genes and their copy numbers ($r = -0.44$, $P = 0.16$). The *atpA* gene copy number was always below that of *cob* in the analyzed *Phaseolus vulgaris* genomes but the level of the *atpA* transcripts was on average 10 times higher than in the case of *cob*. In contrast to *atpA*, the *atp9* gene was more numerous than *cob* but its transcript abundance was below that observed for the *cob* gene.

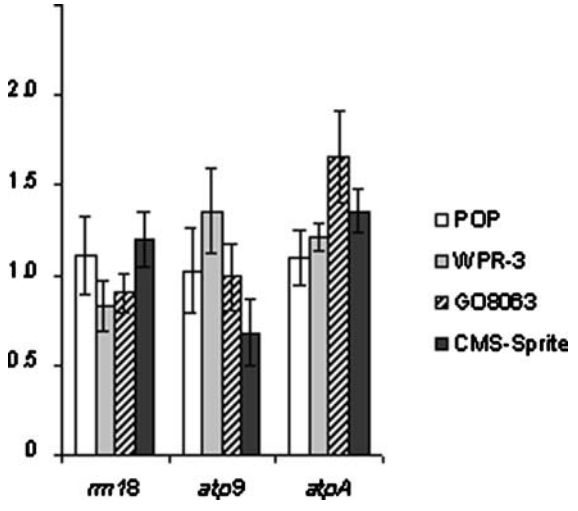

Figure 5. Intergenomic analysis of transcript levels. Transcripts of the *atpA*, *atp9*, *cob* and *rrn18* genes were quantified in four *P. vulgaris* lines by one-step real-time RT-PCR and normalized to the copy number of *cob* (set to 1). The obtained data were then compared between *P. vulgaris* lines with the POP line taken as the calibrator (transcript levels of all genes in this line were set to 1).

We observed some positive correlation between gene copy numbers and transcript levels in the intergenomic analysis (Figure 5) but it was not statistically significant ($r = 0.39$, $P = 0.14$). In specific cases even a negative relationship was observed. It was especially pronounced for the *atp9* transcripts in CMS-Sprite. A 1.3-fold increase in gene copy number was associated with a 0.7-fold decrease in transcript level between CMS-Sprite and POP (Figures 3 and 5). The levels of the *rrn18* ($P < 0.0005$), *atpA* ($P < 0.007$) and *atp9* ($P < 0.029$) transcript steady-states differed between CMS-Sprite and GO8063 with a statistical significance. The differences detected in *rrn18* transcript levels were much more evident than those observed for the copy numbers of this gene. Despite the very similar numbers of *atp9* gene copies in CMS-Sprite and GO8063, the abundances of respective mRNAs differed significantly. The transcripts of *atpA* were more numerous in GO8063 than in CMS-Sprite, while at the DNA level the situation was opposite – the mitochondrial genome of CMS-Sprite contained a higher *atpA* copy number than GO8063. Finally, in contrast to the gene copy number profiles, we could not distinguish between the sterility-inducing genomes and the genomes associated with fertility based on transcript levels.

Discussion

Using real-time PCR we have performed quantitative analysis of mitochondrial genes and their transcripts in four lines of the *Phaseolus* genus. These closely related lines contained mitochondrial genomes varying in the number of copies of specific genes. With this experimental model we could compare the gene and transcript copy numbers in lines believed to contain identical mitochondrial but different nuclear genomes (GO8063 and CMS-Sprite) and the opposite, identical nuclear but different mitochondrial DNAs (CMS-Sprite and WPR-3). We also included the POP line which differs from the remaining lines with respect to both the mitochondrial and nuclear genomes. We found that the variability in the gene copy numbers among the four *Phaseolus vulgaris* lines was low for any particular gene, the detected differences being not higher than 1.7-fold (Figure 3), whereas the copy number differences among individual genes observed within the same genome were more variable, up to 6.5-fold (Figure 1). The copy numbers of genes estimated by intragenomic or intergenomic real-time PCR analyses were different than those calculated basing on three types of theoretical predictions. Interestingly, none of these predictions appeared to be even close to the gene copy numbers estimated experimentally for the POP genome. However, for the remaining *Phaseolus vulgaris* lines the best approximation of the observed values was given by the prediction based on the model organizations proposed for these mitochondrial genomes (Janska and Mackenzie, 1993; Janska *et al*., 1998). The models assume that mitochondrial DNA is organized as two (WPR-3) or three (GO8063 and CMS-Sprite) circular chromosomes that may undergo intra- and intermolecular recombinations via two sets of large repeats R1 and R2. We assumed in our calculations that the amplification rates and consequently the numbers of each molecule were the same and that the frequency of each recombination event was equal regardless of the type of the repeat or its flanking sequences. If any of these two conditions was not fulfilled, then the discrepancy between the experimentally estimated and predicted values would be at least partly explained. Another possible explanation requires the assumption that molecules different than circular chromosomes and their derivatives generated by recombinations exist in *Phaseolus* mitochondria.

The discrepancy observed between the predicted and observed values as well as the low intergenomic variability in gene stoichiometry suggest that gene dosage compensation may play a role in controlling gene copy numbers between related mitochondrial genomes. The prediction that gives the best match to the experimentally estimated values is based on the previously proposed multimolecular organization of the analyzed genomes and recombinations between large repeated sequences. Consequently, it appears that these factors are involved in the gene dosage compensation.

Basing on the physical mapping data and predicted organizations it was believed that the four *Phaseolus vulgaris* lines represent three types of mitochondrial DNA: POP-like, cytoplasmic male sterility-inducing (GO8063 and CMS-Sprite), and that characteristic for the WPR-3 revertant line (Janska and Mackenzie, 1993; Janska *et al*., 1998). Instead, the comparison of the gene copy numbers allowed us to distinguish only two types of *Phaseolus vulgaris* mtDNAs: a fertility-associated one and a cytoplasmic male sterility-inducing genome. The former is represented by the POP and WPR-3 lines showing relatively similar patterns of gene copy numbers. This is quite surprising considering that the predicted organizations of POP and WPR-3 are considerably different. On the other hand, our quantitative analysis supports the view about the similarity between the CMS-Sprite and GO8063 mitochondrial genomes. However, our results do not confirm a complete identity that was previously claimed (Janska and Mackenzie, 1993). The genome of CMS-Sprite contains more copies of the *rrn18* gene and of the *atpA* version which resides near the *pvs* sequence. The differences observed are not striking but they can be crucial considering that the intergenomic variability in the gene copy number is generally low and that these two genes flank the sterility associated *pvs* sequence.

The low intergenomic variability in gene stoichiometry and the observation that it does not determine the transcription level may suggest that the copy numbers of plant mitochondrial genes are not crucial and not strictly controlled. This, however, does not seem to be the case. Each gene in each *Phaseolus vulgaris* line was quantified at

least in three independent DNA samples and at least in triplicate. Considering the small standard deviation obtained, the copy numbers seem to be well controlled. Moreover, the abundance of the *atpA* gene was established in etiolated stalks, green leaves and green stalks and the resulting values were consistent, indicating that the mitochondrial DNA stoichiometry is maintained independently of plant organ.

One of the most important conclusions from our results is that the steady-state levels of mitochondrial transcripts are not regulated by the respective genes' copy numbers. The previously published data concerning this issue is rather scarce and contradictory (Mulligan *et al*., 1991; Muise and Hauswirth, 1995; Janska *et al*., 1998; Hedtke *et al*., 1999). While some authors describe a direct relationship between the number of gene copies and transcript abundance (Janska *et al*., 1998; Hedtke *et al*., 1999) or transcription activity (Muise and Hauswirth, 1995), our results and those reported by Mulligan *et al*. (1991) deny such correlation. However, it has to be underlined that the data presented in this paper concerns only those genes that are maintained in the main mitochondrial genome. We have already shown that an increase in the copy number of a given gene occurring via a genomic shift from a substoichiometric level to the main genome level, results in the gene's expression (Janska *et al*., 1998). Consequently, a correlation between the gene copy number and transcription exists when the gene is shifted from sublimon to normal abundance but it does not operate within the main genome.

The lack of correlation between the gene copy number and transcript steady-state level that we observed in *Phaseolus vulgaris* mitochondria brings to mind the results of a similar analysis performed in chloroplasts of *Chlamydomonas reinhardtii* (Eberhard *et al*., 2002). The authors found that a decreased number of chloroplast genome copies had limited effects on the abundance of transcripts and that chloroplast protein synthesis was relatively insensitive to changes in the gene copy number or transcript level.

We found that in the intragenomic analysis the abundances of transcripts varied greatly (up to a few dozen-fold) among individual genes. Rather unexpectedly, a negative correlation was observed between the copy number of a gene and the steady-state level of its transcript. In other words, transcription of relatively scarce genes (*rrn18*, *atpA*) resulted in high steady-state levels of transcripts while transcripts of an abundant gene (*atp9*) were relatively few. Like for the gene copy numbers also in the case of transcript levels the intergenomic variability was low. A maximal 2-fold difference was observed between CMS-Sprite and WPR-3 for *atp9* mRNA.

The steady-state level of transcripts measured by real-time RT-PCR reflects not only changes in transcription activity, but are also affected by post-transcriptional regulation, e.g. RNA processing and degradation. Therefore, our results do not allow one to conclude which level of regulation (transcriptional or post-transcriptional) is more responsible for the observed differences in transcript levels. A comparative survey of transcriptional activities and steady-state levels of various mRNAs in *A. thaliana* mitochondria revealed that variations in transcriptional activities of mitochondrial genes are higher than the steady-state RNA level variations (Giege *et al*., 2000). That observation suggested that high variations of transcription rates are buffered by post-transcriptional processes in *Arabidopsis*. However, this does not seem to be happening to the same extent in the case of beans because the steady-state RNA levels of the investigated transcripts differ significantly. Moreover, studies in *Arabidopsis* have shown that the steady-state levels of transcripts for different subunits of a given protein complex correlate with the protein complex subunit stoichiometries. In contrast, in our results the proportions of *atpA* and *atp9* transcripts in *P. vulgaris* (1:0.03–1:0.07 depending on the bean line) do not correspond to the stoichiometry of the respective protein subunits in complex V (1:4).

The discrepancy between the levels of transcripts and respective proteins involved in functional complexes could be explained if translational or posttranslational control of protein expression operated in *Phaseolus* mitochondria. This conclusion is consistent with a recent view that the biogenesis of the inner mitochondrial membrane complexes is coordinated post-translationally at the level of assembly of the protein-complexes (Giege *et al*., 2005).

Within the last decade we have witnessed a quick accumulation of physical maps and complete genomic nucleotide sequences. This enormous experimental effort has brought

unquestionable benefits but also disappointments. Among others, plant mitochondrial genomes of several species have been mapped and mtDNAs of six species sequenced (Unseld *et al*., 1997; Kubo *et al*., 2000; Notsu *et al*., 2002; Handa *et al*., 2003; Clifton *et al*., 2004; Sugiyama *et al*., 2005). Our results indicate that predictions about the relative gene copy numbers cannot be based on the number of gene locations found during mapping and/or sequencing.

Acknowledgements

This work was supported by the State Committee for Scientific Research (Grant No. 3 P06A 007 25). Purchase of the LightCycler 2.0 instrument was supported by the State Committee for Scientific Research (Grant No. 4947/IA/105/2004). M. Woloszynska is grateful to Prof W.H. Trzeciak (Department of Biochemistry and Molecular Biology, University of Medical Sciences, Poznan, Poland) who enabled her to learn the real-time PCR method in his laboratory and to Dr. S. Wisniewski for introducing her to this technique. M. Woloszynska is a recipient of a fellowship from the Foundation for Polish Science.

Reference

Clifton, S.W., Minx, P., Fauron, C.M.R., Gibson, M., Allen, J.O., Sun, H., Thompson, M., Barbazuk, W.B., Kanuganti, S., Tayloe, C., Meyer, L., Wilson, R.K. and Newton, K.J. 2004. Sequence and comparative analysis of the maize NB mitochondrial genome. Plant Physiol. 136: 3486–3503.

Eberhard, S., Drapier, D. and Wollman, F.A. 2002. Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J. 31: 149–160.

Fauron, C.M.R., Moore, B. and Casper, M. 1995. Maize as a model of higher plant mitochondrial genome plasticity. Plant Sci. 112: 11–32.

Giege, P., Hoffman, M., Binder, S. and Brennicke, A. 2000. RNA degradation buffers asymetries of transcription in *Arabidopsis* mitochondria. EMBO Rep. 1: 164–170.

Giege, P., Sweetlove, L.J., Cognat, V. and Leaver, C.J. 2005. Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in *Arabidopsis*. Plant Cell 17: 1497–1512.

Handa, H. 2003. The complete nucleotide and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucl. Acids Res. 31: 5907–5916.

Hedtke, B., Wagner, I., Borner, T. and Hess, W.R. 1999. Inter-organellar crosstalk in higher plants: impaired chloroplast development affects mitochondrial gene and transcript levels. Plant J. 19: 635–643.

Janska, H. and Mackenzie, S.A. 1993. Unusual mitochondrial genome organization in cytoplasmic male sterile common bean and the nature of cytoplasmic reversion to fertility. Genetics 135: 869–879.

Janska, H., Sarria, R., Woloszynska, M., Arrieta-Montiel, M. and Mackenzie, S.A. 1998. Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. Plant Cell 10: 1163–1180.

Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A. and Mikami, T. 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for $tRNA^{Cys}$(GCA). Nucl. Acids Res. 28: 2571–2576.

Mackenzie, S.A., Pring, D.R., Basset, M.J. and Chase, C.D. 1988. Mitochondrial DNA rearrangement associated with fertility restoration and cytoplasmic reversion to fertility in cytoplasmic male sterile *Phaseolus vulgaris* L. Proc. Natl. Acad. Sci. USA 85: 2714–2717.

Mackenzie, S. and McIntosh, L. 1999. Higher plant mitochondria. Plant Cell 11: 571–585.

Muise, R.C. and Hauswirth, W.W. 1995. Selective DNA amplification regulates transcript levels in plant mitochondria. Curr. Genet. 28: 113–121.

Mulligan, R.M., Leon, P. and Walbot, V. 1991. Transcriptional and posttranscriptional regulation of maze mitochondrial gene expression. Mol. Cell. Biol. 11: 533–543.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A. and Kadowaki, K. 2002. 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: 434–445.

Small, I.D., Isaac, P.G. and Leaver, C.J. 1987. Stoichiometric differences in DNA molecules containing the *atpA* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. EMBO J. 6: 865–869.

Sneath, P.H.A. and Sokal, R.R. 1973. Numerical Taxonomy. W. H. Freeman & Co, San Francisco.

Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A. and Sugiura, M. 2005. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol. Gen. Genomics 272: 603–615.

Unseld, M., Marienfeld, J.R., Brandt, P. and Brennike, A. 1997. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat. Genet. 15: 57–61.

Woloszynska, M., Kieleczawa, J., Ornatowska, M., Wozniak, M. and Janska, H. 2001. The origin and maintenance of the small repeat in the bean mitochondrial genome. Mol. Genet. Genomics 265: 865–872.