Molecular Ecology Notes (2002) **2**, 428–430 doi: 10.1046/j.1471-8278 .2002.00263.x

PRIMER NOTE

A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA

J. DUMINIL, M.-H. PEMONGE and R. J. PETIT
INRA, Laboratoire de génétique et amélioration des arbres forestiers, BP 45, F-33611 Gazinet Cedex, France

Abstract

Being generally uniparentally inherited, the plant mitochondrial genome is a source of original markers potentially useful for studies of phylogeny and population genetics. We designed 24 new pairs of plant mitochondrial DNA primers that allow amplification of either introns, intergenic regions or genes. They have been defined for consensus over angiosperms and were tested along with 11 previously described mitochondrial primer pairs on 28 plant species representing 19 families of higher plants. The total set allows amplification of 40 kb (~11%) of the mtDNA genome of *Arabidopsis thaliana*. The amplification rate ranged between 76% and 100% depending on the species.

Keywords: *Arabidopsis*, *Beta*, forest tree, multiple alignment, polymorphism, primer design

Received 17 April 2002; revision received 20 May 2002; accepted 20 May 2002

Plants have two cytoplasmic genomes, chloroplast (cp) DNA and mitochondrial (mt) DNA, which are generally uniparentally inherited. They are therefore a source of potentially useful markers for evolutionary and ecological studies. However, higher plant mitochondrial (mt) DNA is very conserved in sequence but not in structure. PCR-based markers useful at low taxonomic levels are therefore difficult to obtain, despite the large size of plant mtDNA: from 200 to 2400 kb, compared to 14–42 kb for animals and 18–176 kb for fungi (Backert *et al*. 1997). This is due to the presence, in higher plant mtDNA, of introns, intergenic sequences, duplicate sequences and sequences of plastid and nuclear origin (Marienfeld *et al*. 1999). The gene arrangement of mtDNA in higher plants varies enormously due to the presence of repeated regions, source of recombination within and between mtDNA genomes (Schuster & Brennicke 1994). However, exonic sequences of mtDNA are very well conserved, facilitating the identification of consensus regions within coding sequences.

We compared the completely sequenced mitochondrial genomes of *Arabidopsis thaliana* (Y08501 and Y08502) and *Beta vulgaris* (AP000397) to identify target DNA regions for amplification. We focused on noncoding sequences (introns or intergenic spacers) common to both species. Primers were also defined that amplify only coding sequences of genes, but excluding those already well-studied and constant in size (such as *atp*1, *atp*6–1, *rps*7 and *tat*C), too small (*nad*4L) or absent from one of the two genomes (e.g. *atp*8, present in *Beta* but absent in *Arabidopsis*).

For the intergenic sequences of ribosomal region *rrn*5/*rrn*18 two different combinations have been defined, one that amplifies the intergenic region only (using the reverse primer *rrn*18–1), and the second that amplifies the intergenic region and the gene *rrn*18 (using the reverse primer *rrn*18–2). For the intergenic region *rps*12/*nad*3 two different combinations have also been defined: the first one, *rps*12–1/*nad*3–1, amplifies the two genes *nad*3 and *rps*12, and the intergenic region between them; the second one, *rps*12–2/*nad*3–2, amplifies almost exclusively the intergenic sequence.

The conservation of candidate sequences in higher plants was tested and PCR primers were defined according to thermodynamic parameters necessary for their use in PCR reactions. The sequences of interest in the *Arabidopsis* genome were aligned with all Viridiplantae sequences available in databases using the computer software BLAST of NCBI (Altschul *et al*. 1997). Primers could then be chosen in regions consensus across angiosperms. The exact positions and lengths of the primers were chosen according to their parameters using the OLIGO Primer Analysis software version 5.0 (Rychlik *et al*. 1990), with *A. thaliana* sequence as reference.

Correspondence: R. J. Petit. Fax: + 33557979088; E-mail: *remy@pierroton.inra.fr*

Table 1 Description, amplification conditions and degree of conservation of the 35 mt consensus primer pairs*

No.	Name	Primer forward	Sequence (5′–3′)	Primer reverse	Sequence (5′–3′)	Position in *Arabidopsis*	Size (bp)	Ta† (°C)	Elongation time (min)	Reference‡	Score§
1	***atp*6–2****	***atp*6–2**	**GCATCATTCAAGTAAATACA**	***atp*6–2**	**GTGAAGCTGTCTGGAGGG**	**96836–98081**	**1264**	**50**	**2**	**this study**	23
2	***atp*9**	***atp*9**	**CCAAGTGAGATGTCCAAGAT**	***atp*9**	**CTTCGGTTAGAGCAAAGCC**	**78821–79081**	**279**	**50**	**1**	**this study**	24
3	***ccb*203**	***ccb*203**	**ASGTTCTACGGACCGATGCC**	***ccb*203**	**CACGGGGAGGGAGCRGGCGA**	**34274–34766**	**511**	**56**	**1**	**this study**	28
4	***ccb*206****	***ccb*206**	**TCAATCTTGTRAACTAATCG**	***ccb*206**	**CYYCTCCACACCAATCACGA**	**31081–30507**	**593**	**62**	**1**	**this study**	24
5	***ccb*256****	***ccb*256**	**GGAAGTTAGCAAAGTTAGAC**	***ccb*256**	**TTGTTCTTAACAGCGATGGC**	**40152–40661**	**528**	**56**	**1**	**this study**	28
6	***ccb*452**	***ccb*452**	**TATTACCAGACATAATTGGG**	***ccb*452**	**TGAACTGTATCTTCTTTGTG**	**53342–51849**	**1512**	**50**	**2**	**this study**	26
7	***cox*1**	***cox*1**	**TTGTTACGACCACGAAGA**	***cox*1**	**TCGGTGCCATTGCTGGAG**	**149988–151329**	**1360**	**58**	**2**	**this study**	26
8	*cox*2/1–2¶,**	*cox*2/1	ttttcttcctcattctkattt	*cox*2/2	ccactctattgtccacttcta	42124–42481	376	50	1	a	25
9	***cox*2/2–3¶**	***cox*2/2**	**TAGRAACAGCTTCTACGACG**	***cox*2/3**	**GRGTTTACTATGGTCAGTGC**	**40561–41963**	**1421**	**52**	**2**	**this study**	28
10	***cox*3**	***cox*3**	**CCGTAGGAGGTGTGATGT**	***cox*3**	**CTCCCCACCAATAGATAGAG**	**18365–19050**	**704**	**58**	**1**	**this study**	28
11	*nad*1/2–3	*nad*1/2	gcattacgatctgcagctca	*nad*1/3	ggaagccgattagtttctgc	87978–89043	1084	57.5	2	b	28
12	*nad*1/4–5	*nad*1/4	gccaatatgatcttaatgag	*nad*1/5	tcfaccttgatactaaaccag	143379–146990	3630	47	4	a	27
13	***nad*2/1–2**	***nad*2/1**	**AATGTGGGTTGGCTTGGWTT**	***nad*2/2**	**AATATGTAAAATTGTCCCTC**	**81121–80035**	**1105**	**62**	**2**	**this study**	21
14	***nad*2/3–4**	***nad*2/3**	**AGAAARGAATGCTGTAACCG**	***nad*2/4**	**ATGGGGATTKTYARTATCGC**	**130248–132983**	**2715**	**58**	**3**	**this study**	21
15	***nad*2/4–5****	***nad*2/4**	**TTCATATAGAATCCATGTCC**	***nad*2/5**	**CTATTTGTTCTTCGCCGCTT**	**129794–128013**	**1800**	**62**	**2**	**this study**	21
16	*nad*4/1–2	*nad*4/1	cagtgggttggtctggtaatg	*nad*4/2	tcatatgggctactgaggag	162009–164093	2103	57.5	3	b	27
17	*nad*4/2–3	*nad*4/2	ctcctcagtagcccatatga	*nad*4/3	aaccagtccatgacttaaca	164093–167257	3184	55	4	a	24
18	*nad*4/3–4	*nad*4/3	ggagctttccaaagaaatag	*nad*4/4	gccatgttgcactaagttac	167460–169643	2203	57	3	a	20
19	***nad*4L/*orf*25**	***orf*25**	**CTGTYTTTTCGCACTTAGGC**	***nad*4L**	**GTCCGRGGTACTATTGCTGT**	**188305–188957**	**671**	**52**	**1**	**this study**	27
20	*nad*5/1–2	*nad*5/1	ttttttcggacgttttctag	*nad*5/2	tttggccaagtatcctacaa	140724–142929	2225	57	3	a	28
21	*nad*5/4–5	*nad*5/4	ccaatttttgggccaattcc	*nad*5/5	cattgcaaaggcataatgat	22067–20697	1390	47	2	a	28
22	***nad*6****	***nad*6**	**TGAGTGGGTCWGTCGTCCTC**	***nad*6**	**TGATACTTTCTGTTTTGTCG**	**76653–77239**	**605**	**58**	**1**	**this study**	21
23	*nad*7/1–2	*nad*7/1	acctcaacatcctgctgctc	*nad*7/2	cgatcagaataaggtaaagc	132118–133320	1222	47	2	a	27
24	*nad*7/2–3	*nad*7/2	gctttaccttattctgatcg	*nad*7/3	tgttcttgggccatcataga	133220–134335	1135	57	2	a	28
25	*nad*7/3–4	*nad*7/3	tctatgatggcccaagaaca	*nad*7/4	acaccaaattctccctttagg	134316–137968	3652	47	4	a	28
26	***nad*7/4–5**	***nad*7/4**	**TGTCCTCCATCACGATVTCG**	***nad*7/5**	**CCAAATTCTCCTTTAGGTGC**	**136033–137965**	**1952**	**58**	**2**	this study	26
27	***nad*9**	***nad*9**	**GGTCATCTCAATGGGYTCAG**	***nad*9**	**TATAGTTGGGAGACTTTACC**	**23720–24192**	**491**	**58**	**1**	**this study**	25
28	***orf*25**	***orf*25**	**AAGACCRCCAAGCYYTCTCG**	***orf*25**	**TTGCTGCTATTCTATCTATT**	**188115–188600**	**504**	**50**	**1**	**this study**	28
29	***rpl*5**	***rpl*5**	**AGTGGTAAAGTCTCATCT**	***rpl*5**	**ATYGTGTGAAATAAGAGTAG**	**58256–57842**	**432**	**50**	**1**	**this study**	26
30	***rps*3****	***rps*3**	**GGCGTATTTCGGATGCTT**	***rps*3**	**TCAAGTYGGTTCAGTGAG**	**26598–28658**	**2076**	**50**	**3**	**this study**	27
31	***rps*4**	***rps*4**	**CSTTTCYGCTCCGAAGAG**	***rps*4**	**TCTCCGAAGATTGAGG**	**82048–82979**	**950**	**58**	**1**	**this study**	21
32	***rps*12–1/*nad*3–2**	***rps*12**	**TTTCTTCTCTACCATGACGA**	***nad*3**	**TGATCCYACTCGGTSTTCCT**	**60248–60929**	**700**	**50**	**1**	**this study**	22
33	***rps*12–2/*nad*3–1**	***rps*12**	**ACCATATTTDGATCTGCCDC**	***nad*3**	**YACGATHGGATTTCTMTATG**	**60559–60679**	**139**	**50**	**1**	**this study**	25
34	***rrn*5/*rrn*18–1**	***rrn*5**	**GAGGTCGGAATGGGATCGGG**	***rrn*18–1**	**GGGTGAAGTCGTAACAAGGT**	**161129–161383**	**273**	**58**	**1**	**this study**	27
35	***rrn*5/*rrn*18–2**	***rrn*5**	***Idem***	***rrn*18–2**	**CGTAASGCGTGGGAATCTGC**	**161129–163086**	**1976**	**58**	**2**	**this study**	26

*The 24 new consensus primer pairs (in bold type) and the 11 previously designed ones (in normal type).

†Annealing temperature.

‡a, Dumolin-Lapègue *et al*. (1997); b, Demesure *et al*. (1995).

§Number of species amplified (over the 28 tested) for each primer pair.

¶Different introns are present in *Arabidopsis* (*cox*2/1–2) and in *Beta* (*cox*2/2–3).

**With these primers, the presence of faint bands in addition to the major one was observed in some species.

The annealing temperatures of the primers were optimized by testing them on a Thermocycler Gradient 96 (Stratagene, La Jolla, CA, USA). The PCR mix solution (25 μL) contained 10 ng of template genomic DNA, 75 mM of Tris HCl, 1.8 mM of $MgCl_2$, 0.2 ng of BSA, 20 mM of ammonium sulphate, 200 μM of each of the four dNTP, 0.2 μM of each primer (or twice the quantity in the case of degenerated primers) and 0.2 units of silvestar DNA *Taq* polymerase (Eurogentec, Liège, Belgium). The amplifications were carried out using 1 cycle of 4 min at 94 °C, 35 cycles of 45 s at 94 °C, 45 s at 44–66 °C, 1–3 min (depending on the length of the fragment to be amplified) at 72 °C and one cycle at 72 °C. The list of the new primer pairs, their optimal annealing temperature and time of elongation are given in Table 1.

The efficiency of the primers was tested on 18 families (28 species) of angiosperms including both monocotyledons and dicotyledons (principally shrubs and trees): Aceraceae (*Acer campestre*), Aquifoliaceae (*Ilex aquifolium*), Araliaceae (*Hedera helix*), Betulaceae (*Alnus glutinosa*, *Betula pendula*), Chenopodiaceae (*Beta vulgaris*, *Spinacia oleracea*) Corylaceae (*Carpinus betulus*, *Corylus avellana*), Cruciferae (*Arabidopsis thaliana*), Ericaceae (*Calluna vulgaris*), Fabaceae (*Cytisus scoparius*), Fagaceae (*Fagus sylvatica*, *Quercus robur*), Poaceae (*Oryza sativa*, *Zea mays*), Oleaceae (*Fraxinus excelsior*), Onagraceae (*Oenothera speciosa*), Rosaceae (*Crataegus monogyna*, *Prunus avium*, *Rubus fructicosa*, *Sorbus torminalis*), Salicaceae (*Populus tremula*, *Salix caprea*), Solanaceae (*Nicotiana tabacum*), Tiliaceae (*Tilia cordata*) and Ulmaceae (*Ulmus minor*). One gymnosperm (family Pinaceae, *Pinus pinaster*) was also included. Most of these species have been used by Grivet *et al*. (2001) to check a large set of cpDNA primers. The score provided in Table 1 is the number of species yielding the expected PCR product for a given primer pair.

The score of the primer pairs ranged from 20 to 28 (71% to 100%). The lowest efficiency was found for *rps*4, *nad*6, *nad*4/3–4, *nad*2/4–5 and *nad*2/1–2. Generally these primers do not amplify the two monocots (*Zea mays* and *Oryza sativa*) and the gymnosperm (*Pinus pinaster*); the corresponding gene sequences are therefore not as well conserved among all higher plants. On the contrary, several primers (30%) give an amplification efficiency of 100% (*ccb*203, *ccb*256, *cox*2/1–2, *cox*3, *nad*1/2–3, 5/1–2, 5/4–5, 7/2–3, 7/3–4 and *orf*25).

This set of 35 consensus primer pairs allows the amplification of ~11% of the mtDNA genome of *A. thaliana* (approximately 23kb of noncoding sequences and 17 kb of coding sequences; total size of *Arabidopsis* mtDNA is 366924 nucleotides, Unseld *et al*. 1997). For all of the species tested, except for *Pinus pinaster*, the primer set allowed the amplification of more than 30 kb of the mitochondrial genome. The rate of amplification varies across the 28 species from 76% (in *P. pinaster*) to 100% (for *A. thaliana*, *Fagus sylvatica*, *Quercus robur*, *Prunus avium*, *Crataegus monogyna* and *Acer campestre*).

Some primer pairs produced multiband patterns in a few of the species analysed. This lack of specificity may be due to the presence of one or more copy of the gene in the mtDNA genome, or to the transfer to the nucleus of one copy of the gene (presence of nuclear pseudogenes or 'Numnts', Bensasson *et al*. 2001), and care should be taken when interpreting such results.

We hope that this set of primers, designed specifically to be robust and as consensus as possible, will be useful to detect mtDNA diversity at low taxonomic levels (including for intraspecific diversity studies).

Acknowledgements

We gratefully acknowledge the help of Delphine Grivet and Céline Latouche-Hallé for initiating JD to primer design and of Kate Porter for checking the English. The study has been carried out with financial support from the European Communities, Agriculture and Fisheries (FAIR) specific programme, CT97-3795, 'CYTOFOR' and the Bureau des Ressources Génétiques (France).

References

Altschul SF, Madden TL, Schäffer A *et al*. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Research*, **25**, 3389–3402.

Backert S, Nielsen BL, Börner T (1997) The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends in Plant Science*, **2**, 477–483.

Bensasson D, Zhang D-X, Hartl DL, Hewitt M (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**, 314–321.

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129–131.

Dumolin-Lapègue S, Pemonge M-H, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.

Grivet D, Heinze B, Vendramin GG, Petit RJ (2001) Genome walking with consensus primers: application to the large single copy of chloroplast DNA. *Molecular Ecology Notes*, **1**, 345–349.

Marienfeld J, Unseld M, Brennicke A (1999) The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. *Trends in Plant Science*, **4–12**, 495–502.

Rychlik W, Spencer WJ, Rhoades RE (1990) Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Research*, **18**, 6409–6412.

Schuster W, Brennicke A (1994) The plant mitochondrial genome: physical structure, information content, RNA editing, and gene migration to the nucleus. *Annual Review of Plant Physiology and Plant Molecular Biology, Plant Molecular Biology*, **45**, 61–78.

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366 924 nucleotides. *Nature Genetics*, **15**, 57–61.