

2014

# Genes and Junk in Plant Mitochondria—Repair Mechanisms and Selection

Alan C. Christensen

*University of Nebraska - Lincoln*, [achristensen2@unl.edu](mailto:achristensen2@unl.edu)

Follow this and additional works at: <http://digitalcommons.unl.edu/bioscifacpub>

---

Christensen, Alan C., "Genes and Junk in Plant Mitochondria—Repair Mechanisms and Selection" (2014). *Faculty Publications in the Biological Sciences*. 352.

<http://digitalcommons.unl.edu/bioscifacpub/352>

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in the Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Genes and Junk in Plant Mitochondria—Repair Mechanisms and Selection

Alan C. Christensen\*

School of Biological Sciences, University of Nebraska–Lincoln

\*Corresponding author: E-mail: achristensen2@unl.edu.

Accepted: May 27, 2014

## Abstract

Plant mitochondrial genomes have very low mutation rates. In contrast, they also rearrange and expand frequently. This is easily understood if DNA repair in genes is accomplished by accurate mechanisms, whereas less accurate mechanisms including nonhomologous end joining or break-induced replication are used in nongenes. An important question is how different mechanisms of repair predominate in coding and noncoding DNA, although one possible mechanism is transcription-coupled repair (TCR). This work tests the predictions of TCR and finds no support for it. Examination of the mutation spectra and rates in genes and junk reveals what DNA repair mechanisms are available to plant mitochondria, and what selective forces act on the repair products. A model is proposed that mismatches and other DNA damages are repaired by converting them into double-strand breaks (DSBs). These can then be repaired by any of the DSB repair mechanisms, both accurate and inaccurate. Natural selection will eliminate coding regions repaired by inaccurate mechanisms, accounting for the low mutation rates in genes, whereas mutations, rearrangements, and expansions generated by inaccurate repair in noncoding regions will persist. Support for this model includes the structure of the mitochondrial *mutS* homolog in plants, which is fused to a double-strand endonuclease. The model proposes that plant mitochondria do not distinguish a damaged or mismatched DNA strand from the undamaged strand, they simply cut both strands and perform homology-based DSB repair. This plant-specific strategy for protecting future generations from mitochondrial DNA damage has the side effect of genome expansions and rearrangements.

**Key words:** mitochondrial genome, mutation rate, DNA repair, junk DNA.

## Introduction

Plant mitochondrial genomes have followed different evolutionary trajectories from their counterparts in animals and fungi. The genomes are very large (up to 11 Mb) but still have only 30–60 genes, thus most of the DNA is noncoding. The mutation rate measured in protein-coding regions and rRNA regions is very low, but the genomes are subject to major rearrangements and expansions (Palmer and Herbon 1988). The mutational burden hypothesis was proposed as an explanation for the paradox of low mutation rates and high expansion rates (Lynch et al. 2006; Lynch 2007), but exceptional species with both high mutation rates and high expansion rates have been found that defy this explanation (Cho et al. 2004; Parkinson et al. 2005; Sloan, Muller, et al. 2012; Sloan et al. 2012). After comparing the mitochondrial noncoding sequences of two *Arabidopsis thaliana* ecotypes that had been diverged for approximately 200,000 years, I proposed that coding and noncoding DNAs are repaired by

different mechanisms and thus have different mutation rates and spectra (Christensen 2013). Although coding regions are highly conserved, noncoding DNA has diverged so rapidly that over 200 kb of the *A. thaliana* mitochondrial genome is not alignable with any sequences outside the Brassicales family of plants, suggesting that it is nonfunctional junk (Brenner 1998; Christensen 2013). This also explains why noncoding DNA has not previously been used in mutational or phylogenetic studies—it evolves too quickly to be useful over evolutionary time scales. The model proposes that coding regions are repaired very accurately, likely by homologous recombination or gene conversion. Noncoding regions are repaired by inaccurate mechanisms of double-strand break (DSB) repair that produce rearrangements, chimeric genes, and genome expansion (Davila et al. 2011). Because there is no mechanism available for precisely removing junk DNA, it accumulates by Muller's ratchet (Muller 1964). The common feature in both coding

and noncoding DNA is DSB repair, leading either to homology-based accurate repair or to inaccurate repair with duplications expanding the genome.

Although this model explains the observed features of mitochondrial genomes, how the coding and noncoding DNA have such distinctly different mutation rates and spectra is still a mystery. One possible explanation is that the primary mechanisms of DNA repair are different in genes and in junk, and the only plausible mechanism for this is transcription-coupled repair (TCR) (Ganesan et al. 2012; Vermeulen and Fouteri 2013; Howan et al. 2014). The existence of cotranscribed genes in plant mitochondria provides an opportunity to test this hypothesis. In this work, I find the hypothesis of TCR to be unlikely and suggest a model for how mitochondrial genomes are repaired differently in genes and in junk.

## Results

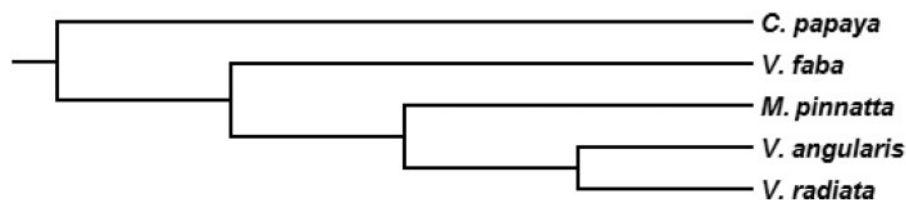
The hypothesis of TCR can be tested by examining both coding and noncoding transcribed regions, for example, the protein-coding regions and intergenic regions of cotranscribed genes. The model predicts that the mutation rate in the coding regions should be equal to the mutation rate in the intergenic regions. In *A. thaliana*, there are four gene clusters shown to be cotranscribed: *nad4L-atp4*, *rpl5-cob*, *nad3-rps12*, and *rps3-rpl16* (Hoffmann et al. 1999; Forner et al. 2007) and these same clusters are observed in a wide variety of angiosperms (Richardson et al. 2013). The lengths of the intergenic regions in these transcripts in *A. thaliana* are 266 bp, 1.9 kbp, 45 bp, and 0 bp (*rps3* and *rpl16* overlap by 134 bp), respectively (Davila et al. 2011). Because selection might be acting near the translation start and stop sites, the two larger intergenic regions are most suitable as a test of the hypothesis. In several species including *A. thaliana*, there is an *rps14* pseudogene between *rpl5* and *cob* (Aubert et al. 1992; Quinones et al. 1996; Figueroa et al. 1999; Ong and Palmer 2006). Because in some species *rps14* is a functional gene and in others it is a pseudogene in the intergenic region, several species were chosen for analysis all of which have a functional *rps14* gene. The *rpl5* and *rps14* genes are just a few nucleotides apart, so only the *rps14-cob* intergenic region was used. The species chosen were all legumes with completely sequenced mitochondrial genomes containing single copies of the *nad4L-atp4* and *rpl5-rps14-cob* clusters. Four legumes were chosen: The mung bean (*Vigna radiata*), the azuki bean (*Vigna angularis*), the pongam tree (*Millettia pinnata*), and the fava bean (*Vicia faba*). *Carica papaya* was chosen as outgroup (fig. 1).

The five coding regions, *nad4L*, *atp4*, *rpl5*, *rps14*, and *cob*, were aligned (supplementary fig. S1, Supplementary Material online), and the synonymous substitutions per synonymous site were measured using the concatenation of all five. The genes of plant mitochondria also show extensive RNA editing (Barkan and Small 2014). The edited sites were confirmed and

annotated in the *M. pinnata* genome (Kazakoff et al. 2012). All of these are C to U edits in the mRNA and most change the amino acid encoded. This alters the definitions of synonymous and nonsynonymous sites for two reasons. If an edit of a C to a U in the mRNA changes the amino acid codon, then a mutation in the genome at that site from a C to a T will be a synonymous change, but standard methods will count that position as a nonsynonymous site. Several examples are in this data set. Of the 48 edited cytosines in these 5 genes, all are conserved within the legumes, but 12 of those edited sites have mutated to T in *C. papaya* (see supplementary fig. S1, Supplementary Material online). Of those sites, 11 would be classified as nonsynonymous substitutions, but the editing in the legumes means that the differences in *C. papaya* are actually synonymous substitutions. Furthermore, the pentatricopeptide repeat (PPR) proteins that mediate editing recognize the RNA sequence upstream of the edit (Barkan and Small 2014), so changes in these positions will all be nonsynonymous if they affect editing efficiency, even if the amino acid sequence at the site of the mutation does not change. For this reason, the analysis was done twice: Once using the entire coding regions and again with any edited codons and the six preceding codons removed from the alignment. The intergenic regions between *nad4L* and *atp4* and between *rps14* and *cob* were also aligned (supplementary fig. S2, Supplementary Material online), and the mutation rate was determined using the concatenation of both alignments, including both transitions and transversions, but not indels. The rates are shown in table 1 and graphed in figure 2. The substitution rate in the intergenic region is higher than the synonymous substitution rate in the complete coding sequences. When edited sites are removed, the substitution rate in the intergenic sequence is still higher than in coding sequences but not statistically significant in most cases.

However, the substitution rate is only a small part of the story. The alignments also reveal frequent nucleotide losses and gains in the intergenic regions (particularly just upstream of *cob*). The intergenic regions have mutated much more extensively than the coding regions when indels are taken into account. As shown previously, most of the intergenic regions in plant mitochondria cannot even be aligned except between very closely related species (Christensen 2013). Without the flanking coding regions of *rps14* and *cob*, the intergenic region between them cannot be accurately aligned using these five species.

If TCR is the mechanism of repair in plant mitochondria, then the mutation rate in a transcribed intergenic region should be the same as the neutral mutation rate measured by synonymous substitutions in the coding regions of the same transcripts, and the frequency of indels in the intergenic regions should be low, as in the coding regions. Indels in the coding regions are rare and are always in-frame, whereas in the intergenic regions, there are more indels per nucleotide; therefore, the hypothesis of TCR is most likely incorrect.

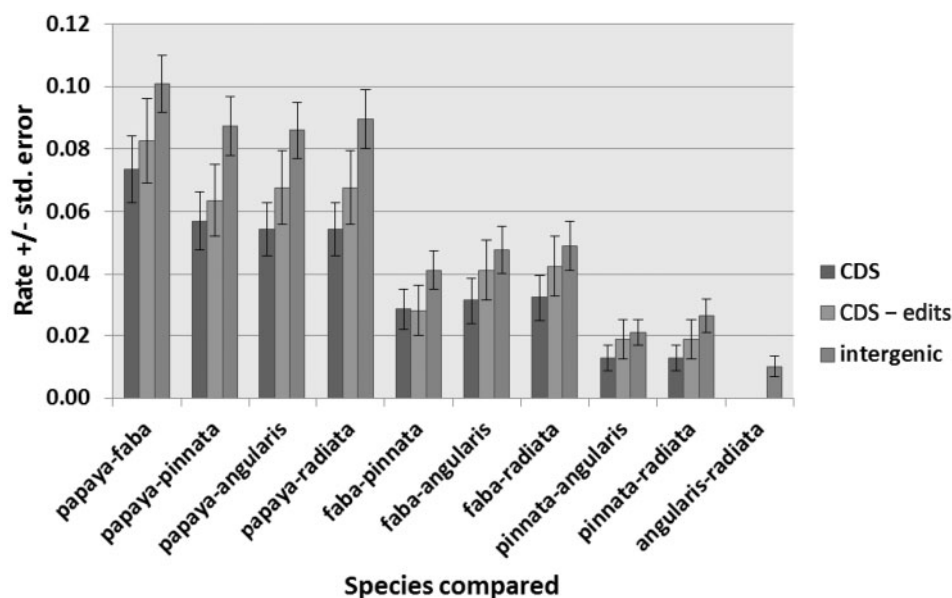


**Fig. 1.**—Phylogenetic relationships of the species studied. Tree showing the relationships between the four legumes used in this study and the outgroup *Carica papaya*. Based on Soltis et al. (2011).

**Table 1**  
Mutation Rates in Coding and Intergenic Regions

Species 1	Species 2	CDS	CDS – edits	Intergenic
<i>Carica papaya</i>	<i>Vicia faba</i>	0.0735 ± 0.0107	0.0826 ± 0.0135	0.1009 ± 0.0092
<i>C. papaya</i>	<i>Millettia pinnata</i>	0.0570 ± 0.0093	0.0636 ± 0.0114	0.0873 ± 0.0095
<i>C. papaya</i>	<i>Vigna angularis</i>	0.0543 ± 0.0084	0.0677 ± 0.0117	0.0860 ± 0.0091
<i>C. papaya</i>	<i>V. radiata</i>	0.0543 ± 0.0085	0.0677 ± 0.0117	0.0896 ± 0.0094
<i>Vic. faba</i>	<i>M. pinnata</i>	0.0285 ± 0.0064	0.0280 ± 0.0080	0.0411 ± 0.0064
<i>Vic. faba</i>	<i>V. angularis</i>	0.0313 ± 0.0074	0.0411 ± 0.0097	0.0477 ± 0.0075
<i>Vic. faba</i>	<i>V. radiata</i>	0.0323 ± 0.0074	0.0424 ± 0.0097	0.0491 ± 0.0078
<i>M. pinnata</i>	<i>V. angularis</i>	0.0128 ± 0.0041	0.0189 ± 0.0063	0.0209 ± 0.0040
<i>M. pinnata</i>	<i>V. radiata</i>	0.0128 ± 0.0041	0.0189 ± 0.0063	0.0264 ± 0.0054
<i>V. angularis</i>	<i>V. radiata</i>	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0102 ± 0.0032

NOTE.—Synonymous substitution rates in the coding sequences (CDS), coding sequences with edited regions removed (CDS – edits), and intergenic regions are shown ( $\pm$  standard errors). Analyses were conducted using the Kumar model (Nei and Kumar 2000). The analysis involved 5 nt sequences. All positions containing gaps and missing data were eliminated. There were a total of 967 positions in the CDS data set, 712 positions in the CDS – edits data set, and 1,620 positions in the intergenic data set. Of these positions in the CDS data set, there were 51 variants within the 4 legumes, including 20 synonymous substitutions, 26 nonsynonymous substitutions, and 5 in-frame indels.



**Fig. 2.**—Mutation rates in coding regions (CDS) and noncoding regions. Synonymous substitution rates in the CDS of *nad4L*, *atp4*, *rpl5*, *rps14*, and *cob* and the coding regions without the edited regions (CDS – edits) were calculated as described in the text. Substitution rates in the intergenic regions between *nad4L* and *atp4* and between *rps14* and *cob* were also calculated as described in the text. Standard errors are shown.

## Discussion

How do plant mitochondrial genomes and their repair systems produce genes with very low synonymous substitution rates, but intergenic regions with high substitution, indel, genome expansion, and rearrangement rates? One possibility was a different DNA repair pathway in genes and in junk, but the only plausible mechanism is TCR, which can be ruled out. The explanation must therefore be a combination of the available DNA repair pathways and selection on the DNA postrepair. Plant mitochondria have a short-patch base-excision repair system, at least for removal of uracil (Boesch et al. 2009), but there is no evidence for long-patch base-excision repair or nucleotide-excision repair (Gualberto et al. 2013). Genome evolution and the rearrangements seen in mutants suggest that DSB repair is an important process in plant mitochondria (Shedge et al. 2007; Arrieta-Montiel et al. 2009; Davila et al. 2011; Janicka et al. 2012; Miller-Messmer et al. 2012; Christensen 2013). DSB repair has multiple modalities that can produce either very accurate or inaccurate repair. One pathway, break-induced replication (BIR), can also result in large duplications, particularly if the break invades another DNA molecule at a homeologous site (Llorente et al. 2008; Cappadocia et al. 2010).

Other than short-patch base-excision repair, little is known about DNA repair proteins in mitochondria, except for the MSH1 protein, a mitochondrially targeted homolog of mismatch repair proteins. It has been suggested that the MSH1 protein plays a role in homology surveillance during DSB repair (Abdelnoor et al. 2003; Shedge et al. 2007; Arrieta-Montiel et al. 2009; Davila et al. 2011). Nuclear and bacterial mismatch repair systems include a strand-discrimination mechanism that directs endonuclease cleavage and repair to the newly synthesized DNA strand (Kunkel and Erie 2005; Ghodgaonkar et al. 2013). Homologs of the strand-discrimination components have not been identified in plant organelles; however, the MSH1 protein of higher plants is fused directly to an endonuclease domain (Abdelnoor et al. 2006). Sequence comparisons and modeling showed that the endonuclease domain is similar to the GIY-YIG homing endonuclease I-TevI, which makes DSBs as a monomer (Mueller et al. 1995; Kleinstiver et al. 2013). This suggests a model for DNA repair in plant mitochondria of lesion recognition followed by double-strand breakage, catalyzed by MSH1 and other unknown nucleases. A DSB eliminates the need for a strand-discrimination system but requires a template.

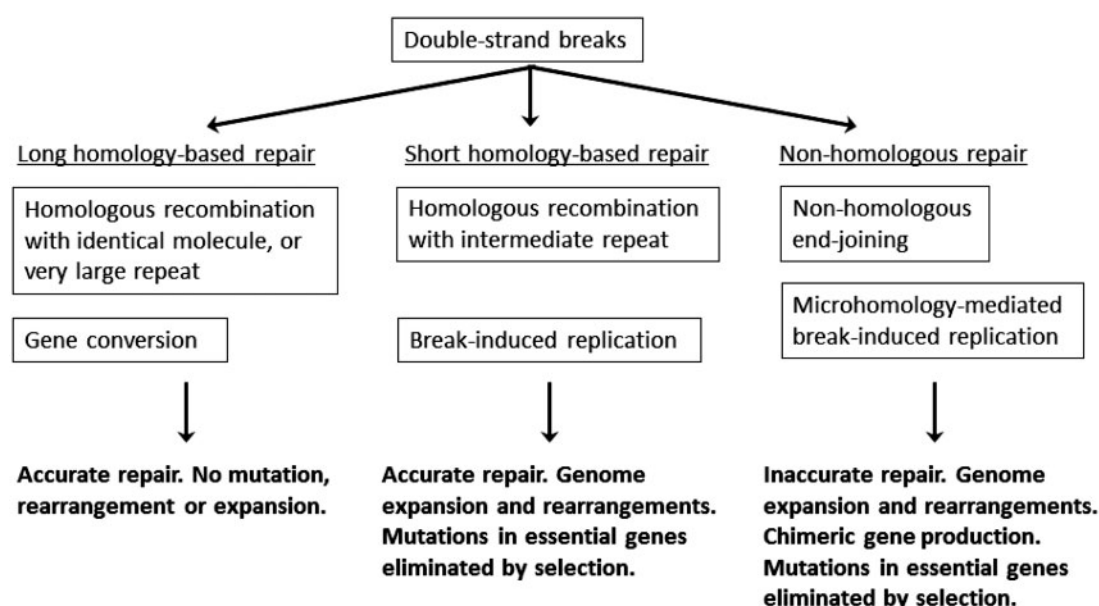
If DNA damage (other than what can be repaired by short-patch base-excision repair, such as deaminated cytosine) is converted into DSBs, and these breaks are then processed by DSB repair mechanisms, there are a number of possible outcomes. Alternative pathways for processing the DSB will depend on whether a template molecule is available and whether the second broken end is captured by the repair event. If the two DNA ends are coordinated, nonhomologous

end joining can be very accurate, but otherwise it can lead to chimeric gene formation and duplications. BIR at a homologous region may lead to large duplications and can also shift the stoichiometry of different parts of the genome. BIR at a short region of homology (such as the 50–500 bp repeats) will lead to rearrangements and genome expansion; BIR at micro-homologies of a few nucleotides can also produce chimeric genes. Homologous recombination or gene conversion will accurately repair the DSBs. The question still remains of how coding sequences are repaired so accurately while the noncoding regions experience rapid change.

The most likely explanation is that both types of DSB repair occur in all parts of the genome, but selection determines which outcomes we can observe (fig. 3). DSB repair can occur in either coding or noncoding DNA and can either be accurate or inaccurate. In noncoding DNA, accurate repair presumably occurs but is impossible to observe in alignments. Inaccurate repair leads to expansions, mutations, and rearrangements, which are observed. In coding DNA, mitochondria with inaccurately repaired essential genes may be eliminated from the cell, or not inherited, thus what we observe in coding DNA is repair that maintains gene function, explaining the low synonymous substitution and indel rate. Accurate, homology-based repair such as gene conversion can explain the observations in coding sequences. If a template is not available within a mitochondrion, mitochondrial fusion could occur to make a template DNA molecule available. This model, that most DNA repair is mediated via generating DSBs followed by the DSB repair pathways and selection for functional mitochondria within a cell, can explain the evolution of plant mitochondrial genomes.

An interesting additional question is why natural selection has favored this mechanism of DNA repair in plant mitochondria but not in animal mitochondria or the nucleus. Recent work showed that in animals the female germline sequesters a subset of mitochondria that are relatively inactive in producing reactive oxygen species and other DNA damaging agents, to minimize transmission of mitochondrial mutations (de Paula et al. 2013). Both plants and animals need to avoid the inheritance of accumulated mitochondrial mutations and appear to use different mechanisms to accomplish that. Plants do not have the luxury of specifying a germline, so converting damage into DSBs followed by accurate template-directed repair ensures that the genes will be faithfully inherited. The side effect of using DSB repair for nearly every type of damage is genome expansion and accumulation of chimeric genes, but the benefit of accurate transmission of mitochondrial genes to the next generation must outweigh the relatively minor cost of replicating a large mitochondrial genome. Finally, the mutational burden hypothesis does not appear to apply to plant mitochondria. In addition to mutations in the junk DNA apparently being mostly neutral, the specific repair mechanisms available do not lead to an inverse correlation between mutation rate and genome size. This model further





**FIG. 3.**—Model for mitochondrial DNA repair explaining differences between genes and junk. The diagram shows the fate of DSBs. These can be repaired by nonhomologous or template-based repair, and a template can either be a sister DNA molecule or be a short stretch of identity in a different context in the same or a different DNA molecule.

predicts that if mechanisms such as base-excision repair or mismatch repair are less effective or transiently lost in a lineage, DSB repair will produce genome expansions at the same time as base substitution rates increase. This also predicts a loss of editing sites and can explain the counterintuitive positive correlation between mutation rates and genome expansions in plant mitochondria.

## Materials and Methods

Complete mitochondrial genome sequences used were accessions KC189947 for *V. faba* (Negruk 2013), JN872550 for *M. pinnata* (Kazakoff et al. 2012), AP012599 for *V. angularis* (Naito et al. 2013), HM367685 for *V. radiata* (Alverson et al. 2011), and EU431224 for *C. papaya* (Ming et al. 2008). Sequence manipulation to extract the specific genes and intergenic regions studied was done using the VectorNTI 11.5.0 package from Invitrogen.

Alignments were done using MUSCLE (Edgar 2004) as implemented in MEGA6 (Tamura et al. 2013). Alignments were prepared for figures using Jalview (Waterhouse et al. 2009). Synonymous substitution rates and standard error estimates were calculated by MEGA6, using the Kumar model (Nei and Kumar 2000) with all ambiguous positions removed for each sequence pair. Substitution rates in noncoding regions were calculated by MEGA6, using Kimura's two-parameter model (Kimura 1980), including both transitions and transversions, with all ambiguous positions removed for each sequence pair.

## Supplementary Material

Supplementary figures S1 and S2 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org>).

## Acknowledgments

The author is grateful to Matt Wynn, Nnamdi Uche, Rachael Schulte, and Brent Nielsen for helpful comments on the article. John Osterman recommended a nice Chianti to go with the fava bean sequence. This work was supported in part by an EAGER award from the National Science Foundation (MCB-1104677).

## Literature Cited

- Abdelnoor RV, et al. 2003. Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to *MutS*. *Proc Natl Acad Sci U S A*. 100:5968–5973.
- Abdelnoor RV, et al. 2006. Mitochondrial genome dynamics in plants and animals: convergent gene fusions of a *MutS* homolog. *J Mol Evol*. 63: 165–173.
- Alverson AJ, Zhuo S, Rice DW, Sloan DB, Palmer JD. 2011. The mitochondrial genome of the legume *Vigna radiata* and the analysis of recombination across short mitochondrial repeats. *PLoS One* 6:e16404.
- Arrieta-Montiel MP, Shedge V, Davila J, Christensen AC, Mackenzie SA. 2009. Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* 183:1261–1268.
- Aubert D, Bisanz-Seyer C, Herzog M. 1992. Mitochondrial rps14 is a transcribed and edited pseudogene in *Arabidopsis thaliana*. *Plant Mol Biol*. 20:1169–1174.

- Barkan A, Small I. 2014. Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol.* 65:415–442.
- Boesch P, et al. 2009. Plant mitochondria possess a short-patch base excision DNA repair pathway. *Nucleic Acids Res.* 37:5690–5700.
- Brenner S. 1998. Refuge of spandrels. *Curr Biol.* 8:R669.
- Cappadocia L, et al. 2010. Crystal structures of DNA-Whirly complexes and their role in *Arabidopsis* organelle genome repair. *Plant Cell* 22: 1849–1867.
- Cho Y, Mower JP, Qiu YL, Palmer JD. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc Natl Acad Sci U S A.* 101:17741–17746.
- Christensen AC. 2013. Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol.* 5: 1079–1086.
- Davila JL, et al. 2011. Double-strand break repair processes drive evolution of the mitochondrial genome in *Arabidopsis*. *BMC Biol.* 9:64.
- de Paula WB, et al. 2013. Female and male gamete mitochondria are distinct and complementary in transcription, structure, and genome function. *Genome Biol Evol.* 5:1969–1977.
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.
- Figuroa P, et al. 1999. The gene for mitochondrial ribosomal protein S14 has been transferred to the nucleus in *Arabidopsis thaliana*. *Mol Gen Genet.* 262:139–144.
- Forner J, Weber B, Thuss S, Wildum S, Binder S. 2007. Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Res.* 35:3676–3692.
- Ganesan A, Spivak G, Hanawalt PC. 2012. Transcription-coupled DNA repair in prokaryotes. *Prog Mol Biol Transl Sci.* 110:25–40.
- Ghodgaonkar MM, et al. 2013. Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell.* 50:323–332.
- Gualberto JM, et al. 2013. The plant mitochondrial genome: dynamics and maintenance. *Biochimie* 100:107–120.
- Hoffmann M, Dombrowski S, Guha C, Binder S. 1999. Cotranscription of the rpl5-rps14-cob gene cluster in pea mitochondria. *Mol Gen Genet.* 261:537–545.
- Howan K, Monnet J, Fan J, Strick TR. 2014. Stopped in its tracks: the RNA polymerase molecular motor as a robust sensor of DNA damage. *DNA Repair (Amst)* pii:S1568–7864(14)00059-7.
- Janicka S, et al. 2012. A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *Plant J.* 72: 423–435.
- Kazakoff SH, et al. 2012. Capturing the biofuel wellhead and powerhouse: the chloroplast and mitochondrial genomes of the leguminous feedstock tree *Pongamia pinnata*. *PLoS One* 7:e51687.
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol.* 16:111–120.
- Kleinstiver BP, Wolfs JM, Edgell DR. 2013. The monomeric GIY-YIG homing endonuclease I-Bmol uses a molecular anchor and a flexible tether to sequentially nick DNA. *Nucleic Acids Res.* 41:5413–5427.
- Kunkel T, Erie D. 2005. DNA mismatch repair. *Annu Rev Biochem.* 74: 681–710.
- Llorente B, Smith CE, Symington LS. 2008. Break-induced replication: what is it and what is it for? *Cell Cycle* 7:859–864.
- Lynch M. 2007. The origins of genome architecture. Sunderland (MA): Sinauer Associates, Inc.
- Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. *Science* 311:1727–1730.
- Miller-Messmer M, et al. 2012. RecA-dependent DNA repair results in increased heteroplasmy of the *Arabidopsis* mitochondrial genome. *Plant Physiol.* 159:211–226.
- Ming R, et al. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452:991–996.
- Mueller JE, Smith D, Bryk M, Belfort M. 1995. Intron-encoded endonuclease I-TevI binds as a monomer to effect sequential cleavage via conformational changes in the td homing site. *EMBO J.* 14: 5724–5735.
- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat Res.* 106:2–9.
- Naito K, Kaga A, Tomooka N, Kawase M. 2013. De novo assembly of the complete organelle genome sequences of azuki bean (*Vigna angularis*) using next-generation sequencers. *Breed Sci.* 63:176–182.
- Negrak V. 2013. Mitochondrial genome sequence of the legume *Vicia faba*. *Front Plant Sci.* 4:128.
- Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. Oxford: Oxford University Press.
- Ong HC, Palmer JD. 2006. Pervasive survival of expressed mitochondrial rps14 pseudogenes in grasses and their relatives for 80 million years following three functional transfers to the nucleus. *BMC Evol Biol.* 6:55.
- Palmer JD, Herbon LA. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol.* 28:87–97.
- Parkinson CL, et al. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evol Biol.* 5:73.
- Quinones V, et al. 1996. The rpl5-rps14-cob gene arrangement in *Solanum tuberosum*: rps14 is a transcribed and unedited pseudogene. *Plant Mol Biol.* 31:937–943.
- Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD. 2013. The “fossilized” mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol.* 11:29.
- Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA. 2007. Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* 19:1251–1264.
- Sloan DB, Muller K, McCauley DE, Taylor DR, Storchova H. 2012. Intraspecific variation in mitochondrial genome sequence, structure, and gene content in *Silene vulgaris*, an angiosperm with pervasive cytoplasmic male sterility. *New Phytol.* 196:1228–1239.
- Sloan DB, et al. 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 10:e1001241.
- Soltis DE, et al. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *Am J Bot.* 98:704–730.
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Vermeulen W, Foustier M. 2013. Mammalian transcription-coupled excision repair. *Cold Spring Harb Perspect Biol.* 5:a012625.
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191.

Associate editor: Bill Martin.

Supplemental Data Figure 1. Alignment of coding regions.  
Alignments of the coding regions for *nad4L*, *atp4*, *rpl5*, *rps14* and *cob* are shown. Alignments were done using Muscle, implemented in Mega6, as described in Methods. Alignments were then prepared for presentation using Jalview, as described in Methods. Cytosines that are edited in *M. pinnatta* are indicated with stars.

A) Alignment of *nad4L*

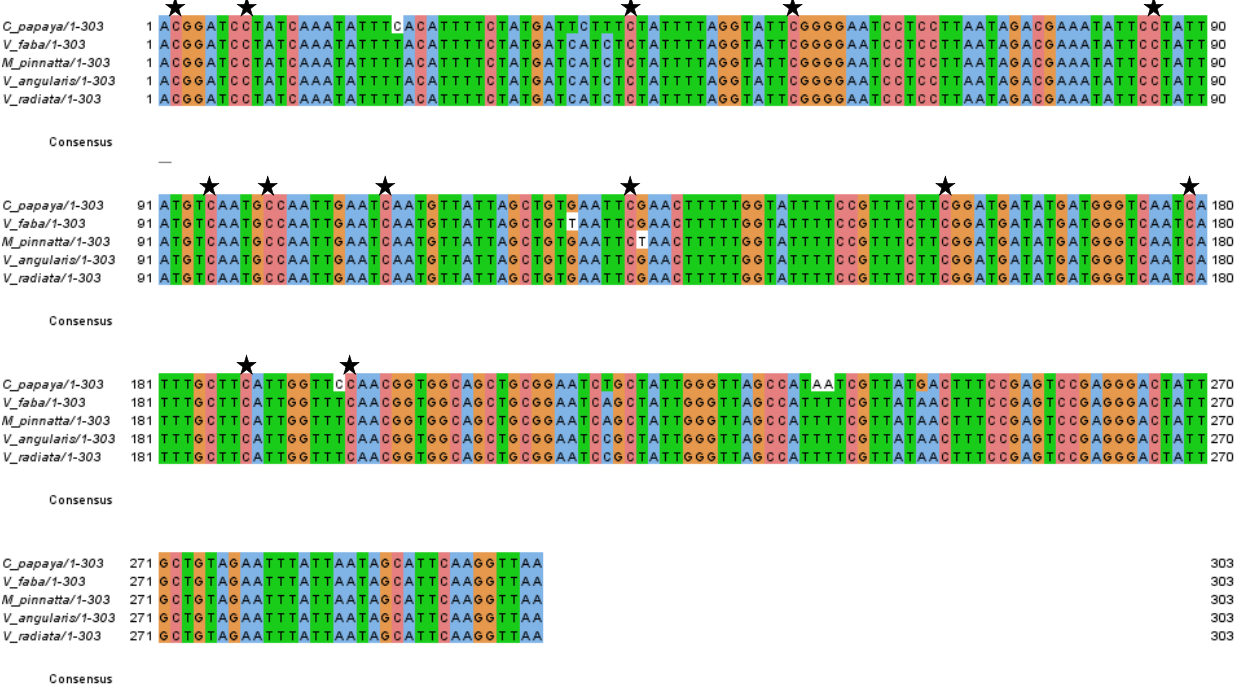




Figure 2 displays multiple sequence alignments of the 5' region of the 18S rDNA gene for five species: *C. papaya*, *V. faba*, *M. pinnatta*, *V. angularis*, and *V. radiata*. The alignments are presented in blocks, with the consensus sequence shown below each block. The sequences are color-coded to highlight conserved regions (green) and variable regions (red, blue, orange, yellow). Star symbols (\*) are placed above specific positions in the alignments to indicate sites of interest.

**Alignment 1 (Positions 1-90):**

- C. papaya*/1-579: 1 ATGAGATTGAGTTCCACGAATATGGAATGCTAGCAAGATGTTATTTGCTGCTATTCTATCTATTGTCGATTAAGTTCCGAAGAAAGATCTCA
- V. faba*/1-588: 1 ATGAGATTGAGTTCCACGCATATGCAAGATGCTATTGCTGCTATTCTCATCTATTGTCGATTAAGTTCCGAATTAAGATATCTCA
- M. pinnatta*/1-588: 1 ATGAGATTGAGTTCCACGAATATGCAAGCTAGAAAAGATGCTATTGCTGCTATTCTCATCTATTGTCGATTAAGTTCCGAAGAAAGATCTCA
- V. angularis*/1-585: 1 ATGAGATTGAGTTCCACGAATATGCT...CTAGAAAAGATGCTATTGCTGCTATTCTCATCTATTGTCGATTAAGTTCCGAAGAAAGATCTCA
- V. radiata*/1-594: 1 ATGAGATTGAGTTCCACGAATATGCT...CTAGAAAAGATGCTATTGCTGCTATTCTCATCTATTGTCGATTAAGTTCCGAAGAAAGATCTCA

**Consensus:** —

**Alignment 2 (Positions 91-177):**

- C. papaya*/1-579: 91 ATCTATAATGAAGAAATGATAGTAGCTCTTGTTTATAGGCTTTCTCATATTAGTCGTTCAGAGTTTAGGTAAGACTTTCAAAGTGACT
- V. faba*/1-588: 91 ATCTATAATGAAGAAATGATAGTAGCTCGTTGTTTATAGGCTTTCTCATATTAGTCGGAAGAGTTTAGGTAATACTTTCAAAGTTAAGT
- M. pinnatta*/1-588: 91 ATCTATAATGAAGAAATGATAGTAGCTCGTTGTTTATAGGCTTTCTCATATTAGTCGGAAGAGTTTAGGTAATACTTTCAAAGTTAAGT
- V. angularis*/1-585: 88 ATCTATAATGAAGAAATGATAGTAGCTCGTTGTTTATAGGCTTTCTCATATTAGTCGGAAGAGTTTAGGTAATACTTTCAAAGTTAAGT
- V. radiata*/1-594: 88 ATCTATAATGAAGAAATGATAGTAGCTCGTTGTTTATAGGCTTTCTCATATTAGTCGGAAGAGTTTAGGTAATACTTTCAAAGTTAAGT

**Consensus:** —

**Alignment 3 (Positions 181-270):**

- C. papaya*/1-579: 181 CTCGACGGGAGAAATCCAGGCTATTCCAGGAAGAAATCGCAGCAATTTCCCAATCTTAACGAAAGTAGTTCCCTCCGGAATCTAATGAACAACAA
- V. faba*/1-588: 181 CTCGACGGGAGAAATCCAGGCTATTCCAGGAAGAAATCGCAGCAATTTCCCAATCTTAACGAAAGTAGTTCCCTCCGGAATCTAATGAACAACAA
- M. pinnatta*/1-588: 181 CTCGACGGGAGAAATCCAGGCTATTCCAGGAAGAAATCGCAGCAATTTCCCAATCTTAACGAAAGTAGTTCCCTCCGGAATCTAATGAACAACAA
- V. angularis*/1-585: 178 CTCGACGGGAGAAATCCAGGCTATTCCAGGAAGAAATCGCAGCAATTTCCCAATCTTAACGAAAGTAGTTCCCTCCGGAATCTAATGAACAACAA
- V. radiata*/1-594: 178 CTCGACGGGAGAAATCCAGGCTATTCCAGGAAGAAATCGCAGCAATTTCCCAATCTTAACGAAAGTAGTTCCCTCCGGAATCTAATGAACAACAA

**Consensus:** —

**Alignment 4 (Positions 271-357):**

- C. papaya*/1-579: 271 CGATTACTTTAGGAATCCGCTTTGCGAAATTTGTGGCACCGTAGTAGAATCATTACCAATGGCAGCGTGTGCGCCTAAGTGCAGAAAAGACAGTGT
- V. faba*/1-588: 271 CGATTACTTTAGGAATCAGTTTTCGAAATTTGTGGCACCGTAGTAGAATCATTACCAATGGCAGCGTGTGCGCCTAAGTGCAGAAAAGACAGTGT
- M. pinnatta*/1-588: 271 CGATTACTTTAGGAATCAGTTTTCGAAATTTGTGGCACCGTAGTAGAATCATTACCAATGGCAGCGTGTGCGCCTAAGTGCAGAAAAGACAGTGT
- V. angularis*/1-585: 268 CGATTACTTTAGGAATCAGTTTTCGAAATTTGTGGCACCGTAGTAGAATCATTACCAATGGCAGCGTGTGCGCCTAAGTGCAGAAAAGACAGTGT
- V. radiata*/1-594: 268 CGATTACTTTAGGAATCAGTTTTCGAAATTTGTGGCACCGTAGTAGAATCATTACCAATGGCAGCGTGTGCGCCTAAGTGCAGAAAAGACAGTGT

**Consensus:** —

**Alignment 5 (Positions 361-447):**

- C. papaya*/1-579: 361 CAAGCTTTGTTATGCGCAAACTTAAATGTTAAGTCAGCAACACTTCCAAATGCCACTTCTTCCGTCGCATCCGTCCTTCAGGACGATCTTA
- V. faba*/1-588: 361 CAAGCTTTGTTATGCGCAAACTTAAATGTTAAGTCAGCAACACTTCCAAATGCCACTTCTTCCGTCGCATCCGTCCTTCAGGACGATCTTA
- M. pinnatta*/1-588: 361 CAAGCTTTGTTATGCGCAAACTTAAATGTTAAGTCAGCAACACTTCCAAATGCCACTTCTTCCGTCGCATCCGTCCTTCAGGACGATCTTA
- V. angularis*/1-585: 358 CAAGCTTTGTTATGCGCAAACTTAAATGTTAAGTCAGCAACACTTCCAAATGCCACTTCTTCCGTCGCATCCGTCCTTCAGGACGATCTTA
- V. radiata*/1-594: 358 CAAGCTTTGTTATGCGCAAACTTAAATGTTAAGTCAGCAACACTTCCAAATGCCACTTCTTCCGTCGCATCCGTCCTTCAGGACGATCTTA

**Consensus:** —

**Alignment 6 (Positions 451-537):**

- C. papaya*/1-579: 451 GTCCACAGGTTTTCACCTTCTCAGTGAGCGAAAGATTTTCCCCGGGTGACGTTTGAAGACTTCTATAGTAGAATCTATTCCGAGAGGGCTTG
- V. faba*/1-588: 451 GGCACAAAGTTTTCACCTTATTAGTTAGGAGGAGATTTATCCCCCAGTGATCTTCGAAAAGCAGAAAAAATAGAACTCATTTCCGAGAGAGCTTG
- M. pinnatta*/1-588: 451 GGCACAAAGTTTTCACCTTATTAGTTAGGAGGAGATTTTCCCCCAGTCTATCTCGAAAAGCAGAAAAAATAGAACTCATTTCCGAGAGAGCTTG
- V. angularis*/1-585: 448 GGCACAAAGTTTTCACCTTATTAGTTAGGAGGAGATTTTCCCCCAGTGATCTTCGAAAAGCAGAAAAAATAGAACTCATTTCCGAGAGAGCTTG
- V. radiata*/1-594: 448 GGCACAAAGTTTTCACCTTATTAGTTAGGAGGAGATTTTCCCCCAGTGATCTTCGAAAAGCAGAAAAAATAGAACTCATTTCCGAGAGAGCTTG

**Consensus:** —

**Alignment 7 (Positions 541-594):**

- C. papaya*/1-579: 541 GTGGTCTTAAGAAATGGTTCGGGCGGGGGGTTCTCT...TAA
- V. faba*/1-588: 541 GTGGTCTTAAGAAATGGTTCGGGCGGGGGGTTCTCTTAAGAAATAAA...TAA
- M. pinnatta*/1-588: 541 GTGGTCTTAAGAAATGGTTCGGGCGGGGGGTTCTCTTAAGAAATAAA...TAA
- V. angularis*/1-585: 538 GTGGTCTTAAGAAATGGTTCGGGCGGGGGGTTCTCTTAAGAAATAAA...TAA
- V. radiata*/1-594: 538 GTGGTCTTAAGAAATGGTTCGGGCGGGGGGTTCTCTTAAGAAATAAGAAACGAAATAG

**Consensus:** —

<i>V_papaya1</i> -558	1	A	T	G	T	T	C	C	A	C	T	C	T	T	T	T	C	A	T	C	G	A	A	G	A	T	G	A	T	C	A	G	T	C	A	G	T	C	G	T	G	T	C	A	A	A	C	G	A	A	T	A	C	G	C	C	A	A	G	T	T	A	G	G	A	A	G	T	T	C	T	G	G	A	G	
<i>V_faba1</i> -1558	1	A	T	G	T	T	C	C	A	C	T	C	A	T	T	T	C	A	T	C	G	A	A	G	A	T	G	A	T	C	A	G	T	C	A	G	T	C	G	T	G	T	C	A	A	A	C	G	A	A	T	A	C	G	C	C	A	A	G	T	T	A	G	G	A	A	G	T	T	C	T	G	G	A	G	
<i>M_pinnata1</i> -561	1	A	T	G	T	T	C	C	A	C	T	C	A	T	T	T	C	A	T	C	G	A	A	G	A	T	G	A	T	C	A	G	T	C	A	G	T	C	G	T	G	T	C	A	A	A	C	G	A	A	T	A	C	G	C	C	A	A	G	T	T	A	G	G	A	A	G	T	T	C	T	G	G	A	G	
<i>V_angustifolia</i> -1558	1	A	T	G	T	T	C	C	A	C	T	C	A	T	T	T	C	A	T	C	G	A	A	G	A	T	G	A	T	C	A	G	T	C	A	G	T	C	G	T	G	T	C	A	A	A	C	G	A	A	T	A	C	G	C	C	A	A	G	T	T	A	G	G	A	A	G	T	T	C	T	G	G	A	G	
<i>V_radiata1</i> -1558	1	A	T	G	T	T	C	C	A	C	T	C	A	T	T	T	C	A	T	C	G	A	A	G	A	T	G	A	T	C	A	G	T	C	A	G	T	C	G	T	G	T	C	A	A	A	C	G	A	A	T	A	C	G	C	C	A	A	G	T	T	A	G	G	A	A	G	T	T	C	T	G	G	A	G	

## Consensus

[illegible]

## Consensus

*V\_papaya1*-1598 178 **AAAT****T****C****A****T****C****A****G****A****C****A****A****A****G****G****G****T****T****C****G****A****C****A****G****A****A****A****G****T****C****G****T****T****C****G****A****T****C****A****A****T****C****C****A****T****T****C****T****T****T****G****G****G****T****C****A****A****A****A****A****A****A****A****C****A****A****A****G****G****A****T****A****T****G****T****C****A****G** 1267  
*V\_faba1*-1598 178 **AAAT****T****C****A****T****C****A****G****A****C****A****A****A****G****G****G****T****T****C****G****A****C****A****G****A****A****A****G****T****C****G****T****T****C****G****A****T****C****C****A****A****T****C****C****A****T****T****C****T****T****T****G****G****G****T****C****A****A****A****A****A****A****A****C****A****A****A****G****G****A****T****A****T****G****T****C****A****G** 1267  
*M\_pinnata1*-1561 181 **AAAT****T****C****A****T****C****A****G****A****C****A****A****A****G****G****G****T****T****C****G****A****C****A****G****A****A****A****G****T****C****G****T****T****C****G****A****T****C****C****A****A****T****C****C****A****T****T****C****T****T****T****G****G****G****T****C****A****A****A****A****A****A****A****C****A****A****A****G****G****A****T****A****T****G****T****C****A****G** 1270  
*V\_angustifolia1*-1558 178 **AAAT****T****C****A****T****C****A****G****A****C****A****A****A****G****G****G****T****T****C****G****A****C****A****G****A****A****A****G****T****C****G****T****T****C****G****A****T****C****C****A****A****T****C****C****A****T****T****C****T****T****T****G****G****G****T****C****A****A****A****A****A****A****A****A****C****A****A****A****G****G****A****T****A****T****G****T****C****A****G** 1267  
*V\_radiata1*-1598 178 **AAAT****T****C****A****T****C****A****G****A****C****A****A****A****G****G****G****T****T****C****G****A****C****A****G****A****A****A****G****T****C****G****T****T****C****G****A****T****C****C****A****A****T****C****C****A****T****T****C****T****T****T****G****G****G****T****C****A****A****A****A****A****A****A****C****A****A****A****G****G****A****T****A****T****G****T****C****A****G** 1267

## Consensus

*C. papaya*/1-558 268 G A C C T A G C A C G A C A A A G C A C T C T C C G A G G G C A T G G A A T G T C T C A T T T T T C G G T C A G A A T C T C G A C A G T A A T G T C T C T G T T A G A T T C T C C G 357  
*V. faba*/1-558 268 G A C C T A G C A C G A C A A A G C A C T C T C C G A G G G C A T G G A A T G T C T A A T T T T T C G G T C A G A A T C T C G A C A G T A A T G T C T C T G T T A G A T T C T C C G 357  
*M. pinnata*/1-561 271 G A C C T A G C A C G A C A A A G C A C T C T C C G A G G G C A T G G A A T G T C T A A T T T T T C G G T C A G A A T C T C G A C A G T A A T G T C T C T G T T A G A T T C T C C G 360  
*V. angularis*/1-558 268 G A C C T A G C A C G A C A A A G C A C T C T C C G A G G G C A T G G A A T G T C T A A T T T T T C G G T C A G A A T C T C G A C A G T A A T G T C T C T G T T A G A T T C T C C G 357  
*V. radiata*/1-558 268 G A C C T A G C A C G A C A A A G C A C T C T C C G A G G G C A T G G A A T G T C T C A T T T T T C G G T C A G A A T C T C G A C A G T A A T G T C T C T G T T A G A T T C T C C G 357

## Consensus

[illegible]

## Consensus

*C. papaya*/1-558 448 G A A C A T A T T C G A G G G T T C A A T G T G A C T A T T G T C A C T T C G G C C A A C A C A A G A T G A G A C T T T A C C A C C G T G G A G C G G C T T T T T G C A A A A A 537  
*V. faba*/1-558 448 G A A C A T A T T C G A G G G T T C A A T G T G A C T A T T G T C A C T T C G G C C A A C A C A A G A T G A G A C T T T A C C A C C G T G G A G C G G C T T T T T G C A A A A A 537  
*M. pinnata*/1-561 451 G A A C A T A T T C G A G G G T T C A A T G T G A C T A T T G T C A C T T C G G C C A A C A C A A G A T G A G A C T T T A C C A C C G T G G A G C G G C T T T T T G C A A A A A 540  
*V. angularis*/1-558 448 G A A C A T A T T C G A G G G T T C A A T G T G A C T A T T G T C A C T T C G G C C A A C A C A A G A T G A G A C T T T A C C A C C G T G G A G C G G C T T T T T G C A A A A A 537  
*V. radiata*/1-558 448 G A A C A T A T T C G A G G G T T C A A T G T G A C T A T T G T C A C T T C G G C C A A C A C A A G A T G A G A C T T T A C C A C C G T G G A G C G G C T T T T T G C A A A A A 537

## Consensus

<i>C_papaya</i> /1-558	538	G	A	T	G	A	G	G	G	G	A	A	C	T	C	A	G	T	A	558
<i>V_faba</i> /1-558	538	G	A	T	G	A	G	G	G	A	A	A	C	T	C	A	G	T	A	558
<i>M_pinnatta</i> /1-561	541	G	A	T	G	A	G	G	A	A	A	C	T	C	A	G	T	A	561	
<i>V_angulatis</i> /1-558	538	G	A	T	G	A	G	G	G	A	A	A	C	T	C	A	G	T	A	558
<i>V_radiata</i> /1-558	538	G	A	T	G	A	G	G	G	A	A	A	C	T	C	A	G	T	A	558

### Consensus

D) Alignment of *rps14*

*C\_papaya*/1-303 1 ATGTCGGAGAAAGCGAAATATACGAGATCACAAACGTAGATTGCTCGCGGCTAAATATGAATTGAGACGAAAGCTTTATAAAGCCTTTGT 90  
*V\_faba*/1-303 1 ATGTCGGAGAAAGCGAAATATACGAGATCACAAACGCAGATTGCTCGCGGCATAATATGAATTGAGACGAAAGCTTTATAAAGCCTTTGT 90  
*M\_pinnatta*/1-303 1 ATGTCGGAGAAAGCGAAATATACGAGATCACAAACGCAGATTGCTCGCGGCTAAATATGAATTGAGACGAAAGCTTTATAAAGCCTTTGT 90  
*V\_angulata*/1-303 1 ATGTCGGAGAAAGCGAAATATACGAGATCACAAACGCAGATTGCTCGCGGCTAAATATGAATTGAGACGAAAGCTTTATAAAGCCTTTGT 90  
*V\_radiata*/1-303 1 ATGTCGGAGAAAGCGAAATATACGAGATCACAAACGCAGATTGCTCGCGGCTAAATATGAATTGAGACGAAAGCTTTATAAAGCCTTTGT 90

Consensus

—

*C\_papaya*/1-303 91 AAAGATTC CGATCTTCTAGTGATATGCGAGACAAACATCGTTATAAGTTGTC CAAGTTGCCAAGAAAGAGTTCC TTTGCACGAGTAAGA 180  
*V\_faba*/1-303 91 AAAGATTC TGATCTTCTAGTGATATGCGGACAAAC TTCGTTATAAGTTGTC CAAGTTGCCAAGAAATAGTTCC TTTGCACGAGTAAGA 180  
*M\_pinnatta*/1-303 91 AAAGATTC CGATCTTCTAGTGATATGCGGACAAAC TTCGTTATAAGTTGTC CAAGTTGCCAAGAAAGAGTTCC TTTGCACGAGTAAGA 180  
*V\_angulata*/1-303 91 AAAGATTC CGATCTTCTAGTGATATGCGGACAAAC TTCGTTATAAGTTGTC CAAGTTGCCAAGAAATAGTTCC TTTGCACGAGTAAGA 180  
*V\_radiata*/1-303 91 AAAGATTC CGATCTTCTAGTGATATGCGGACAAAC TTCGTTATAAGTTGTC CAAGTTGCCAAGAAATAGTTCC TTTGCACGAGTAAGA 180

Consensus

*C\_papaya*/1-303 181 AACCGATG TATTTCCACGGGT CGCCCTCGTTCCGTATATGAGTTCTTTGCAATTTCTCGTATCG TTTTTCGTGGATTAGCATCTCGAGGT 270  
*V\_faba*/1-303 181 AACCGATG TATTTCCACGGGT CGCCCTCGTTCCGTATATGAGTTCTTTGCAATTTCTCGTATCG TTTTTCGTGGATTAGCATCTCGAGGT 270  
*M\_pinnatta*/1-303 181 AACCGATG TATTTCCACGGGT CGCCCTCGTTCCGTATATGAGTTCTTTGCAATTTCTCGTATCG TTTTTCGTGGATTAGCATCTCGAGGT 270  
*V\_angulata*/1-303 181 AACCGATG TATTTCCACGGGT CGCCCTCGTTCCGTATATGAGTTCTTTGCAATTTCTCGTATCG TTTTTCGTGGATTAGCATCTCGAGGT 270  
*V\_radiata*/1-303 181 AACCGATG TATTTCCACGGGT CGCCCTCGTTCCGTATATGAGTTCTTTGCAATTTCTCGTATCG TTTTTCGTGGATTAGCATCTCGAGGT 270

Consensus



*C\_papaya*/1-303 271 CCTTTGATGGGCATAAAGAAATCGTCTTGGTAG 303  
*V\_faba*/1-303 271 CCTTTGATGGGCATAAAGAAATCGTCTTGGTAG 303  
*M\_pinnatta*/1-303 271 CCTTTGATGGGCATAAAGAAATCGTCTTGGTAG 303  
*V\_angulata*/1-303 271 CCTTTGATGGGCATAAAGAAATCGTCTTGGTAG 303  
*V\_radiata*/1-303 271 CCTTTGATGGGCATAAAGAAATCGTCTTGGTAG 303

Consensus

<i>C. papaya</i> /1-1182	1	A	T	G	A	C	T	A	A	G	A	C	C	A	C	G	A	T	T	C	T	C	T	T	C	T	A	A	A	G	A	C	C	T	A	T	C	C	T	C	A	C	A	C	T	T	A	A	T	C	A	G	A	C	T	T	T	G	A	T	A	T	C	C	A	C	C	C	G	A	G	C	90
<i>V. faba</i> /1-1179	1	A	T	G	A	C	T	A	A	G	A	C	C	A	C	A	T	T	C	T	C	T	T	C	T	T	C	T	A	A	A	C	A	C	C	T	A	T	C	C	T	C	A	C	A	C	T	T	A	A	T	C	A	C	A	C	C	C	G	A	G	C	90										
<i>M. pininnata</i> /1-1179	1	A	T	G	A	C	T	A	A	G	A	C	C	A	C	A	T	T	C	T	C	T	T	C	T	T	C	T	A	A	A	C	A	C	C	T	A	T	C	C	T	C	A	C	A	C	T	T	A	A	T	C	A	C	A	C	C	C	G	A	G	C	90										
<i>V. angustifolius</i> /1-1182	1	A	T	G	A	C	T	A	A	G	A	C	C	A	C	A	T	T	C	T	C	T	T	C	T	T	C	T	A	A	A	C	A	C	C	T	A	T	C	C	T	C	A	C	A	C	T	T	A	A	T	C	A	C	A	C	C	C	G	A	G	C	90										
<i>V. radiata</i> /1-1182	1	A	T	G	A	C	T	A	A	G	A	C	C	A	C	A	T	T	C	T	C	T	T	C	T	T	C	T	A	A	A	C	A	C	C	T	A	T	C	C	T	C	A	C	A	C	T	T	A	A	T	C	A	C	A	C	C	C	G	A	G	C	90										

### Consensus

<i>C_papaya</i> /1-1182	91	AAT	CTT	AGT	TAT	TGG	GGGGG	TT	CGG	TCG	CTT	AGC	TGG	CTT	TGT	TAG	CAT	CAAA	TAG	GCAT	TGGC	TTTTTTT	TAG	CTAT	GCATT	AC	180
<i>V_faba</i> /1-1179	91	AAT	CTT	AGT	TAT	TGG	GGGGC	TT	CGG	TT	CGT	TAGC	TGG	ATT	TGT	TTAG	CAT	CAGAT	TAG	GCAT	TGGC	TTTTTTT	TAG	CTAT	GCATT	AC	180
<i>M_pinnatta</i> /1-1179	91	AAT	CTT	AGT	TAT	TGG	GGGGC	TT	CGG	TT	CGT	TAGC	TGG	ATT	TGT	TTAG	CAT	CAGAT	TAG	GCAT	TGGC	TTTTTTT	TAG	CTAT	GCATT	AC	180
<i>v_angularis</i> /1-1182	91	AAT	CTT	AGT	TAT	TGG	GGGGC	TT	CGG	TT	CGT	TAGC	TGG	ATT	TGT	TTAG	CAT	CAGAT	TAG	GCAT	TGGC	TTTTTTT	TAG	CTAT	GCATT	AC	180
<i>V_radiata</i> /1-1182	91	AAT	CTT	AGT	TAT	TGG	GGGGC	TT	CGG	TT	CGT	TAGC	TGG	ATT	TGT	TTAG	CAT	CAGAT	TAG	GCAT	TGGC	TTTTTTT	TAG	CTAT	GCATT	AC	180

## Consensus

<i>C_papaya</i> /1-1182	181	A	C	A	C	C	T	A	T	G	T	A	A	C	T	T	T	C	A	A	C	A	C	G	T	A	G	A	C	A	T	T	A	T	G	A	G	A	G	A	T	G	T	G	A	A	G	G	G	G	G	G	T	G	T	T	G	C	C	G	T	T	A	T	G	C	A	T	G	C	T	A	A	T		2700
<i>V_faba</i> /1-1179	181	A	C	A	C	C	T	A	T	G	T	A	A	C	T	T	T	C	A	A	C	A	C	G	T	A	G	A	C	A	T	T	A	T	G	A	G	A	G	A	T	G	T	G	A	A	G	G	G	G	G	T	G	T	T	G	C	C	G	T	T	A	T	G	C	A	T	G	C	T	A	A	T		2700	
<i>M_pinnatifida</i> /1-1179	181	A	C	A	C	C	T	A	T	G	T	A	A	C	T	T	T	C	A	A	C	A	C	G	T	A	G	A	C	A	T	T	A	T	G	A	G	A	G	A	T	G	T	G	A	A	G	G	G	G	T	G	T	T	G	C	C	G	T	T	A	T	G	C	A	T	G	C	T	A	A	T		2700		
<i>V_angularis</i> /1-1182	181	A	C	A	C	C	T	A	T	G	T	A	A	C	T	T	T	C	A	A	C	A	C	G	T	A	G	A	C	A	T	T	A	T	G	A	G	A	G	A	T	G	T	G	A	A	G	G	G	G	T	G	T	T	G	C	C	G	T	T	A	T	G	C	A	T	G	C	T	A	A	T		2700		
<i>V_radiata</i> /1-1182	181	A	C	A	C	C	T	A	T	G	T	A	A	C	T	T	T	C	A	A	C	A	C	G	T	A	G	A	C	A	T	T	A	T	G	A	G	A	G	A	T	G	T	G	A	A	G	G	G	G	T	G	T	T	G	C	C	G	T	T	A	T	G	C	A	T	G	C	T	A	A	T		2700		

## Consensus

C.*papaya*/1-1182 271 GGGGCAAGATGTTTCTCATTTGGTGTACCTTCATCTCTTGGTGTCTATATCATGCGAGTTATAGCAGCTTAGAGGAAATTTGTTGG 3600  
V.*faba*/1-1179 271 GGGGCAAGATGTTTCTCATTTGGTGTACCTTCATATTTTTCGGTGTCTATATCATGCGAGTTATAGCAGCTCTAGAGGAAATTTGTTGG 3800  
M.*pinната*/1-1179 271 GGGGCAAGATGTTTCTCATTTGGTGTACCTTCATATTTTTCGGTGTCTATATCATGCGAGTTATAGCAGCTCTAGAGGAAATTTGTTGG 3800  
V.*angularis*/1-1182 271 GGGGCAAGATGTTTCTCATTTGGTGTACCTTCATATTTTTCGGTGTCTATATCATGCGAGTTATAGCAGCTCTAGAGGAAATTTGTTGG 3800  
V.*radicata*/1-1182 271 GGGGCAAGATGTTTCTCATTTGGTGTACCTTCATATTTTTCGGTGTCTATATCATGCGAGTTATAGCAGCTCTAGAGGAAATTTGTTGG 3800

## Consensus

[illegible]

## Consensus

<i>V_papaya1/1-1182</i>	451	A	C	A	G	T	A	A	T	T	A	C	A	G	C	T	T	A	G	C	G	C	A	T	A	C	C	T	G	T	A	G	A	G	A	T	A	C	C	A	T	A	G	T	G	A	C	T	T	G	G	G	G	T	T	G	G	G	G	T	T	C	C	G	T	G	A	C	A	A	T	G	C	C	
<i>V_faba1/1-1179</i>	451	A	C	A	G	T	A	A	T	T	A	C	A	G	C	T	T	A	G	C	G	C	A	T	A	C	C	T	G	T	A	G	A	G	A	T	A	C	C	A	T	A	G	T	G	A	C	T	T	G	G	G	G	T	T	C	C	G	T	G	A	C	A	A	T	G	C	C							
<i>M_pinnata1/1-1179</i>	451	A	C	A	G	T	A	A	T	T	A	C	A	G	C	T	T	A	G	C	G	C	A	T	A	C	C	T	G	T	A	G	A	G	A	T	A	C	C	A	T	A	G	T	G	A	C	T	T	G	G	G	G	T	T	C	C	G	T	G	A	C	A	A	T	G	C	C							
<i>V_rufus1/1-1182</i>	451	A	C	A	G	T	A	A	T	T	A	C	A	G	C	T	T	A	G	C	G	C	A	T	A	C	C	T	G	T	A	G	A	G	A	T	A	C	C	A	T	A	G	T	G	A	C	T	T	G	G	G	G	T	T	C	C	G	T	G	A	C	A	A	T	G	C	C							
<i>V_andalut1/1-1182</i>	451	A	C	A	G	T	A	A	T	T	A	C	A	G	C	T	T	A	G	C	G	C	A	T	A	C	C	T	G	T	A	G	A	G	A	T	A	C	C	A	T	A	G	T	G	A	C	T	T	G	G	G	G	T	T	C	C	G	T	G	A	C	A	A	T	G	C	C							

## Consensus

<i>p_paypal/1-1182</i>	541	A	C	C	T	T	A	A	A	C	G	C	G	T	T	T	T	T	A	G	T	C	T	C	A	T	C	A	T	T	A	C	C	C	C	T	T	T	A	G	G	G	C	C	A	G	T	C	T	T	C	A	T	C	T	G	G	C	G	C	A	T	T	G	C	A	T	T	G	C	A	T	A	A	T	A		630
<i>V_faba/1-1179</i>	541	A	C	C	T	T	A	A	A	C	G	C	G	T	T	T	T	T	A	G	T	C	T	C	A	T	C	A	T	T	A	C	C	C	C	T	T	T	A	G	G	G	C	C	A	G	T	C	T	T	C	A	T	C	T	G	G	C	G	C	A	T	T	G	C	A	T	A	A	T	A		630					
<i>M_pinnata/1-1179</i>	541	A	C	C	T	T	A	A	A	C	G	C	G	T	T	T	T	T	A	G	T	C	T	C	A	T	C	A	T	T	A	C	C	C	C	T	T	T	A	G	G	G	C	C	A	G	T	C	T	T	C	A	T	C	T	G	G	C	G	C	A	T	T	G	C	A	T	A	A	T	A		630					
<i>V_angularis/1-1182</i>	541	A	C	C	T	T	A	A	A	C	G	C	G	T	T	T	T	T	A	G	T	C	T	C	A	T	C	A	T	T	A	C	C	C	C	T	T	T	A	G	G	G	C	C	A	G	T	C	T	T	C	A	T	C	T	G	G	C	G	C	A	T	T	G	C	A	T	A	A	T	A		630					
<i>V_radiata/1-1182</i>	541	A	C	C	T	T	A	A	A	C	G	C	G	T	T	T	T	T	A	G	T	C	T	C	A	T	C	A	T	T	A	C	C	C	C	T	T	T	A	G	G	G	C	C	A	G	T	C	T	T	C	A	T	C	T	G	G	C	G	C	A	T	T	G	C	A	T	A	A	T	A		630					

## Consensus

<i>C_papaya1</i> -1192	631	GGATCAAAATAACCATTTGGGTGTACATTTCAGAGAAATGGATAAAA	TTCTTTT	TTCCTTTT	ACCCCTT	ATTTTTAT	GTAAAGGAATCTAG	TAAGGTTGGGTA	720
<i>V_faba1</i> -1-1179	631	GGATCAAAATAACCATTTGGGTGTACATTTCAGAGAAATGGATAAAA	TTCTTTT	TTCCTTTT	ACCCCTT	ATTTTTAT	GTAAAGGAATCTAG	TAAGGTTGGGTA	720
<i>M_pinnata1</i> -1-1179	631	GGATCAAAATAACCATTTGGGTGTACATTTCAGAGAAATGGATAAAA	TTCTTTT	TTCCTTTT	ACCCCTT	ATTTTTAT	GTAAAGGAATCTAG	TAAGGTTGGGTA	720
<i>V_anguai1</i> -1-1182	631	GGATCAAAATAACCATTTGGGTGTACATTTCAGAGAAATGGATAAAA	TTCTTTT	TTCCTTTT	ACCCCTT	ATTTTTAT	GTAAAGGAATCTAG	TAAGGTTGGGTA	720
<i>V_radiata1</i> -1192	631	GGATCAAAATAACCATTTGGGTGTACATTTCAGAGAAATGGATAAAA	TTCTTTT	TTCCTTTT	ACCCCTT	ATTTTTAT	GTAAAGGAATCTAG	TAAGGTTGGGTA	720

## Consensus

<i>C_papaya1</i> -1182	721	60	TTTTTGGCTATCTTTTCCATTTGGATTATTTTATGCTCCAAATGTTTTGGGCGATCCCGACAAATATATACCTGCTAATCCGATGCC	810
<i>V_faba1</i> -1-1179	721	60	TTTTTGGCTATCTTTTCCATTTGGATTATTTTATGCTCCAAATGTTTTGGGCGATCCCGACAAATATATACCTGCTAATCCGATGCC	810
<i>M_pinnata1</i> -1-1179	721	60	TTTTTGGCTATCTTTTCCATTTGGATTATTTTATGCTCCAAATGTTTTGGGCGATCCCGACAAATATATACCTGCTAATCCGATGCC	810
<i>V_angularis1</i> -1182	721	60	TTTTTGGCTATCTTTTCCATTTGGATTATTTTATGCTCCAAATGTTTTGGGCGATCCCGACAAATATATACCTGCTAATCCGATGCC	810
<i>V_radiata1</i> -1182	721	60	TTTTTGGCTATCTTTTCCATTTGGATTATTTTATGCTCCAAATGTTTTGGGCGATCCCGACAAATATATACCTGCTAATCCGATGCC	810

## Consensus

<i>C_papaya1-1192</i>	811	A	C	C	C	C	C	C	C	T	C	A	T	T	G	T	G	C	C	G	G	A	A	T	G	G	T	A	T	T	C	C	A	C	C	A	T	C	C	A	T	C	C	A	T	T	C	T	T	C	G	T	A	G	T	A	A	C	C	T	G	A	C	A	A	A	T	C	G	G	A	A	G	T	G	T	A	G	C	C	G	C	A		9000
<i>V_faba1-1179</i>	811	A	C	C	C	C	C	C	C	T	C	A	T	T	G	T	G	C	C	G	A	A	T	G	G	T	A	T	T	C	C	A	C	C	A	T	C	C	A	T	C	C	A	T	T	C	T	T	C	G	T	A	G	T	A	A	C	C	T	G	A	C	A	A	A	T	C	G	G	A	A	G	T	G	T	A	G	C	C	G	C	A		9000	
<i>M_pinnata1-1179</i>	811	A	C	C	C	C	C	C	C	T	C	A	T	T	G	T	G	C	C	G	A	A	T	G	G	T	A	T	T	C	C	A	C	C	A	T	C	C	A	T	C	C	A	T	T	C	T	T	C	G	T	A	G	T	A	A	C	C	T	G	A	C	A	A	A	T	C	G	G	A	A	G	T	G	T	A	G	C	C	G	C	A		9000	
<i>V_angustata1-1182</i>	811	A	C	C	C	C	C	C	C	T	C	A	T	T	G	T	G	C	C	G	A	A	T	G	G	T	A	T	T	C	C	A	C	C	A	T	C	C	A	T	C	C	A	T	T	C	T	T	C	G	T	A	G	T	A	A	C	C	T	G	A	C	A	A	A	T	C	G	G	A	A	G	T	G	T	A	G	C	C	G	C	A		9000	
<i>V_radiata1-1182</i>	811	A	C	C	C	C	C	C	C	T	C	A	T	T	G	T	G	C	C	G	A	A	T	G	G	T	A	T	T	C	C	A	C	C	A	T	C	C	A	T	C	C	A	T	T	C	T	T	C	G	T	A	G	T	A	A	C	C	T	G	A	C	A	A	A	T	C	G	G	A	A	G	T	G	T	A	G	C	C	G	C	A		9000	

*C\_papaya*/1-1182 901 A T A G C A C C A G T T T T T A T A T G T C T T T T G G C T T T A C C T T T T T T T A A A A G T A T G T A T G T G C G T A G T T C A A G T T T T C G C C C G A T T C A C C A A G G A 990  
*V\_faba*/1-1179 901 A T A G C A C C T G T T T T T A T A T G T C T G T T G G C T T T A C C T T T T T T T A A A A G T A T G T A C G T G C G T A G T T C A A G T T T T C G C C C T A T T C A C C A A G G A 990  
*M\_pinnatta*/1-1179 901 A T A G C A C C A G T T T T T A T A T G T C T G T T G G C T T T A C C T T T T T T T A A A A G T A T G T A C G T G C G T A G T T C A A G T T T T C G C C C T A T T C A C C A A G G A 990  
*V\_angulata*/1-1182 901 A T A G C A C C A G T T T T T A T A T G T C T G T T G G C T T T A C C T T T T T T T A A A A G T A T G T A C G T G C G T A G T T C A A G T T T T C G C C C T A T T C A C C A A G G A 990  
*V\_radiata*/1-1182 901 A T A G C A C C A G T T T T T A T A T G T C T G T T G G C T T T A C C T T T T T T T A A A A G T A T G T A C G T G C G T A G T T C A A G T T T T C G C C C T A T T C A C C A A G G A 990

Consensus

*C\_papaya*/1-1182 991 A T A T T T T G G T T G C T T T T G G C G G A T T G C T T A C T A C T A G G T T G G A T C G G A T G T C A A C C T G T G G A G G C A C C A T T T G T G A C T A T T G G A C A A A T T 1080  
*V\_faba*/1-1179 991 A T A T T T T G G T T G C T T T T G G C G G A T C G C T T A C T A C T A G G T T G G A T C G G A T G T C A A C C T G T G G A G G C A C C A T T T G T T A C T A T T G G A C A A A T T 1080  
*M\_pinnatta*/1-1179 991 A T A T T T T G G T T G C T T T T G G C A G A T C G C T T A C T A C T A G G T T G G A T C G G A T G T C A A C C T G T G G A G G C A C C A T T T G T T A C T A T T G G A C A A A T T 1080  
*V\_angulata*/1-1182 991 A T A T T T T G G T T G C T T T T G G C G G A T C G C T T A C T A C T A G G T T G G A T C G G A T G T C A A C C T G T G G A G G C A C C C T T T G T T A C T A T T G G A C A A A T T 1080  
*V\_radiata*/1-1182 991 A T A T T T T G G T T G C T T T T G G C G G A T C G C T T A C T A C T A G G T T G G A T C G G A T G T C A A C C T G T G G A G G C A C C C T T T G T T A C T A T T G G A C A A A T T 1080

Consensus

*C\_papaya*/1-1182 1081 C C T C C T T T T G T T T T C T T T T G T T C T T T G C A T A A C G C C C A T T C T G G G A C G A G T T G G A A G A G G A A T T C C T A A T T C T T A T . . . A C G G A T G A A 1187  
*V\_faba*/1-1179 1081 C C T C C T T T T G T T T T C T T T G T T C T T T G C C A T A A C G C C C A T T C C G G G A C G A G T T G G A A G A G G A A T T C C T A A T T C T T A C . . . A C G G A T G A G 1187  
*M\_pinnatta*/1-1179 1081 C C T C C T T T T G T T T T C T T T G T T C T T T G C C A T A A C G C C C A T T C C G G G A C G A G T T G G A A G A G G A A T T C C G A A T T C T T A C . . . A C G G A T G A G 1187  
*V\_angulata*/1-1182 1081 C C T C C T T T T G T T T T C T T T G T T C T T T G C C A T A A C G C C C A T T C C G G G A C G A G T T G G A A G A G G A A T T C C T A A T T C T T A C A C T A C G G A T G A G 1170  
*V\_radiata*/1-1182 1081 C C T C C T T T T G T T T T C T T T G T T C T T T G C C A T A A C G C C C A T T C C G G G A C G A G T T G G A A G A G G A A T T C C T A A T T C T T A C A C T A C G G A T G A G 1170

Consensus

*C\_papaya*/1-1182 1188 A C T G A T C A C A C C T G A 1182  
*V\_faba*/1-1179 1188 A C T G A T C A G . . . T G A 1179  
*M\_pinnatta*/1-1179 1188 A C T G A T C A G . . . T G A 1179  
*V\_angulata*/1-1182 1171 A C T G A G A T G . . . T G A 1182  
*V\_radiata*/1-1182 1171 A C T G A G A T G . . . T G A 1182

Consensus



# Supplemental Data Figure 2. Alignment of intergenic regions.

Alignments of the intergenic regions were done as described in Methods, and in the legend to Supplemental Data Figure 1.

## A) Alignment of intergenic region between *nad4L* and *atp4*

<i>C_papaya</i> /1-182	1	CCATTCCTCCTATAGAAATTAGGAAAAACGAAAGAACCACTTGATTCTTTTCTG	61
<i>V_faba</i> /1-200	1	ACATGACTCCTAGAGAAATTACCAAAAATACGAAG--TTATCTTTTCTCCTTTCTG	62
<i>M_pinnatta</i> /1-203	1	AGATGACTCCTAGAGAAATTACCAAAAATACGAAG--TTCTCTTTTCTCCTTTCTGTTCTCAATTTCTCTTC	67
<i>V_angulatis</i> /1-195	1	ACATGACTCCTAGAGAAATCCCAAAAATACGAAG--TTCTCTTTTCTCCTTTCTG	59
<i>V_radiata</i> /1-195	1	ACATGACTCCTAGAGAAATTACCAAAAATACGAAG--TTCTCTTTTCTCCTTTCTG	59
<i>C_papaya</i> /1-182	62	TTTCTTTTGTGTTGGTT--GGCAGGGTCAGGGCCTTTCTCGCTTGGCTAACGAATCCGATT	120
<i>V_faba</i> /1-200	63	ATTTATTTTATTTTACTTTGGTTGGCAGGGCAGGGTCGGGGCCTTTATCGCTGGGCGAGCGCATCCGATT	134
<i>M_pinnatta</i> /1-203	68	TTTCTTTTGTATTTTGTGTTGGTT--GGCAGGGTCAGGGCCTTTCTCGCTGGGCGAGCGCATCCGATT	133
<i>V_angulatis</i> /1-195	60	TTTCTTTTGTATTTTGTGTTGGTT--GGCAGGGTCAGGGCCTTTCTCGCTGGGCGAGCGCATCCGATT	125
<i>V_radiata</i> /1-195	60	TTTCTTTTGTATTTTGTGTTGGTT--GGCAGGGTCAGGGCCTTTCTCGCTGGGCGAGCGCATCCGATT	125
<i>C_papaya</i> /1-182	121	CTTC--AGTCTCTTTCTTAAACCACCTTCCCGTTTCAGTTGCTGAAAGATAGAGAGAAGCTTTCTAA	182
<i>V_faba</i> /1-200	135	TATTCTAAAGTCTTTCTTAAACCACCTTCCCGTTTGTGCTGAAAGATAGATAGAACCCTTTCTAA	200
<i>M_pinnatta</i> /1-203	134	GATTCTAAAGTCTTTCTTAAACCACCTTCCCGTTTCAGTTGCTGAAAGATAGATAGAACCCTTTCTAACTAA	203
<i>V_angulatis</i> /1-195	126	GATTCTAAAGTCTTTCTTAAACCACCTTCCCGTTTCAGTTGCTGAAAGATAGATAGAACCCTTTCTAACTAA	195
<i>V_radiata</i> /1-195	126	GATTCTAAAGTCTTTCTTAAACCACCTTCCCGTTTCAGTTGCTGAAAGATAGATAGAACCCTTTCTCACTAA	195

## B) Alignment of intergenic region between *rps14* and *cob*

C_papaya/1-1341	1	CAACCA	CCAAACCAAT	AGAAC	AAGGGTTAGCT	CTGCAGCT	TGGTCC	CAAGCAA	.....	GGTAAGTAGGTT	66								
V_faba/1-1232	1	CAACCG	CCAAACCCAT	AAACCA	AAGGGTTAGCT	CCGCAGCT	TGGTCC	CAAGCAA	GGCTT	GGTAAGTAGGTT	71								
M_pinnatta/1-863	1	CAACCG	CCAAACCAAT	AGAAC	AAGGGTTAGCT	CTGCAGCT	TGGTCC	CAAGCAA	.....	GGTAAGTAGGTT	66								
V_angularis/1-897	1	CAACCG	CCAAACCAAT	AGAAC	AAGGGTTAGCT	CTGCAGCT	TGGTCC	CAAGCAA	.....	GGTAAGTAGGTT	66								
V_radiata/1-894	1	CAACCG	CCAAACCAAT	AGAAC	AAGGGTTAGCT	CTGCAGCT	TGGTCC	CAAGCAA	.....	GGTAAGTAGGTT	66								
C_papaya/1-1341	67	ATTAC	AGCCGGCT	CCGGACCG	AAAAGACCT	AACGGAGT	AATCCCT	TATCT	TTGGAT	CGGAG	.....	ATGCT	132						
V_faba/1-1232	72	ATTAC	GGCCGGCT	CCGGACCG	AAAAGACCT	AACGGAGT	AATCCCT	TATCT	CGGAT	CGGAG	.....	GTGCT	137						
M_pinnatta/1-863	67	ATTAC	GGCCGGCT	CCGGACCG	AAAAGACCT	AACGGAGT	AATCCCT	TATCT	CGGAT	CGGAG	.....	ATGCT	132						
V_angularis/1-897	67	ATTAC	GGCCGGCT	CCGGACCG	AAAAGACCT	AACGGAGT	AATCCCT	TATCT	CGGAT	CGGAG	ATGCT	ATGCT	137						
V_radiata/1-894	67	ATTAC	GGCCGGCT	CCGGACCG	AAAAGACCT	AACGGAGT	AATCCCT	TATCT	CGGAT	CGGAG	ATGCT	ATGCT	137						
C_papaya/1-1341	133	AACGGGG	CGGGAAT	CGAAGT	TGGGGG	ACCTCT	CTACCG	CGCTGT	TGTCT	TATCT	CTCTGT	CAAGT	ATGCT	CCCCAG	203				
V_faba/1-1232	138	AACGGGG	CGGGAAT	CGAAGT	TGGGGG	ACCTAT	CTACCG	CGCTGT	TGTCT	TATCT	CTCTGT	CAAGT	ATGCT	CCCCAG	208				
M_pinnatta/1-863	133	AACGGGG	CTGGAAT	CGAAGT	TGGGGG	ACCTAT	CTACCG	CGCTGT	TGTCT	TATCT	CTCTGT	CAAGT	ATGCT	CCCCAG	203				
V_angularis/1-897	138	AACGGGG	CTGGAAT	CGAAGT	TGGGGG	ACCTAT	CTACCG	CGCTGT	TGTCT	TATCT	CTCTGT	CAAGT	ATGCT	CCCCAG	208				
V_radiata/1-894	138	AACGGGG	CTGGAAT	CGAAGT	TGGGGG	ACCTAT	CTACCG	CGCTGT	TGTCT	TATCT	CTCTGT	CAAGT	ATGCT	CCCCAG	208				
C_papaya/1-1341	204	ACATAG	ACTAGGT	TACA	GATAGT	ACTCTT	TGGAAAG	ATAGA	ATAT	CACCG	.....	CGTGAA	CATAA	CATT	267				
V_faba/1-1232	209	ACATAG	ACTACGT	TACAGGG	TAGT	ACTCTT	TG	.....	GATATA	ATAT	CACCG	CGTGAA	CGTGAA	CATAA	CATT	275			
M_pinnatta/1-863	204	ACATAG	ACTACGT	TACAGGG	TAGT	ACTCTT	TG	.....	GATATA	ATAT	CACCG	CGTGAA	CGTGAA	CATAA	CATT	270			
V_angularis/1-897	209	ACATAG	ACTACGT	TACAGGG	TAGT	ACTCTT	TG	.....	GATAGA	ATAT	CACCG	.....	CGTGAA	CATAA	CATT	269			
V_radiata/1-894	209	ACATAG	ACTACGT	TACAGGG	TAGT	ACTCTT	TG	.....	GATAGA	ATAT	CACCG	.....	CGTGAA	CATAA	CATT	269			
C_papaya/1-1341	268	TGTTAC	GGAATGT	CACT	CC	.....	CGACC	ACTATT	CTAA	ATAT	AGTA	AGGCGG	GAGAACT	CTTGTT	CATT	330			
V_faba/1-1232	276	AGTTAC	GGAATGT	CACT	CC	CGAGG	ATCG	ACC	ACTATT	CTAA	AGAA	AGTA	AGGCGG	GAGAACT	CTTGTT	CATT	346		
M_pinnatta/1-863	271	AGTTAC	GGAATGT	CACT	CC	CGAGG	ATCG	ACC	ACTATT	CTAA	AGAA	AGTA	AGGCGG	GAGAACT	CTTGTT	CATT	341		
V_angularis/1-897	270	AGTTAC	GGAATGT	CACT	CC	CGAGG	ATCG	ACC	ACTATT	CTAA	AGAA	AGTA	AGGCGG	GAGAACT	CTTGTT	CATT	340		
V_radiata/1-894	270	AGTTAC	GGAATGT	CACT	CC	CGAGG	ATCG	ACC	ACTATT	CTAA	AGAA	AGTA	AGGCGG	GAGAACT	CTTGTT	CATT	340		
C_papaya/1-1341	331	GGAGCG	CCGGAGT	CGCGAGG	TTCT	TCCCAT	CATTG	AAGT	CAGAGT	GTGGG	ACTG	AGCCTT	CCGAAT	GAGAA	401				
V_faba/1-1232	347	GGAGCG	CCGGAGT	CGCGGGG	TTGTT	TCCCAT	CATTG	AAGT	CAGAGT	GTGGG	ACTG	AGCCTT	CCGAAT	GATAA	417				
M_pinnatta/1-863	342	GGAGCG	CCGGAGT	CGCGAGG	TTGTT	TCCCAT	CATTG	AAGT	CAGAGT	GTGGG	ACTG	AGCCTT	CCGAAT	GATAA	412				
V_angularis/1-897	341	GGAGCG	CCGGAGT	CGCGAGG	TTGTT	TCCCAT	CATTG	AAGT	CAGAGT	GTGGG	ACTG	AGCCTT	CCGA	.....	404				
V_radiata/1-894	341	GGAGCG	CCGGAGT	CGCGAGG	TTGTT	TCCCAT	CATTG	AAGT	.....	ATAGT	GTGGG	ACTG	AGCCTT	CCGA	.....	400			
C_papaya/1-1341	402	A	GACAAAAA	AGT	CTT	AGTTT	CGTTG	GGAAAA	ACC	AACG	CA	.....	AATAT	CATATT	GACTTT	CTCT	CGCCC	467	
V_faba/1-1232	418	A	TAAAAAAA	AGT	CTT	AGTTT	CGTTG	GGAAAA	ACC	AACG	CA	.....	AATAT	CATATT	GACTTT	CTCT	CGCCC	484	
M_pinnatta/1-863	413	A	GAAAAAAA	AGT	CTT	AGTTT	CGTTG	GGAAAA	ACC	AACG	CA	.....	AATAT	CATATT	GACTTT	CTCT	CGCCC	478	
V_angularis/1-897	405	-	GAAAAAAA	AGT	CTT	AGTTT	CGTTG	GGAAAA	ACC	AACG	CA	AAAT	CATATT	GACTTT	CTCT	CGCCC	473		
V_radiata/1-894	401	-	GAAAAAAA	AGT	CTT	AGTTT	CGTTG	GGAAAA	ACC	AACG	CA	AAAT	CATATT	GACTTT	CTCT	CGCCC	469		
C_papaya/1-1341	468	TACTT	CTAAGG	ATAG	ATAG	.....	AAAAGG	TTGG	AGAG	AGT	GACTTT	ATGAA	ATCT	CTCT	TTTCT	TAAAG	CTGCG	537	
V_faba/1-1232	485	AAC	TTCTAAGG	ATAG	ATAG	.....	AAAAGG	TTGG	AGAG	AGT	TACTTT	ATGAA	ATCT	CTCT	TTTCT	TAAAG	CTGCG	554	
M_pinnatta/1-863	479	AAC	TTCTAAGG	ATAG	ATAG	.....	AAAAGG	TTGG	AGAG	AGT	GACTTT	CTGAA	ATCT	CTCT	TTTCT	TAAAG	CTGCG	548	
V_angularis/1-897	474	AAC	TTCTAAGG	ATAG	ATAG	.....	AAAAGG	TTGG	AGAG	AGT	GACTTT	CTTAA	ATCT	CTCT	TTTCT	TAAAG	CTGCG	543	
V_radiata/1-894	470	AAC	TTCTAAGG	ATAG	ATAG	.....	AAAAGG	TTGG	AGAG	AGT	GACTTT	CTTAA	ATCT	CTCT	TTTCT	TAAAG	CTGCG	540	
C_papaya/1-1341	538	AAAGG	CGCCCC	CCCA	AAAG	.....	CCCC	ACCC	CTTTTT	.....	CCCC	TTTAA	TGAA	AGAC	CCCT	.....	CGCC	TCG	603
V_faba/1-1232	555	CAAGG	CG	.....	.....	.....	GCCCC	ACCC	CTTTTT	.....	CCCC	TTTAA	TGAA	AGAC	CCCT	.....	CGCC	TC	605
M_pinnatta/1-863	549	CAAGG	CG	.....	.....	.....	GCCCC	ACCC	CTTTTT	.....	CCCC	TTTAA	TGAA	AGAC	CCCT	CGCC	CGCC	CG	604
V_angularis/1-897	544	CAAGG	CG	.....	.....	.....	GCCCC	ACCC	CTTTTT	.....	CCCC	TTTAA	TGAA	AGAC	CCCT	CGCC	CGCC	CG	599
V_radiata/1-894	541	CAAGG	CG	.....	.....	.....	GCCCC	ACCC	CTTTTT	.....	CCCC	TTTAA	TGAA	AGAC	CCCT	CGCC	CGCC	CG	596
C_papaya/1-1341	604	CTTCC	TATT	CATTTT	TGGG	TCGGG	ACCCG	AGGGCC	TT	TTTTTT	CT	CAAG	AGG	AGATT	TGAG	AGAT	CAACC	674	
V_faba/1-1232	805	CTTCC	TATT	CATTTT	TGGG	TCGGG	ACCCG	AGGGCC	TT	TTTTTT	CT	CAAG	AGG	AGATT	TGAG	AGAT	CAACC	674	
M_pinnatta/1-863	805	CTTCC	TATT	CATTTT	TGGG	TCGGG	ACCCG	AGGGCC	TT	TTTTTT	CT	CAAG	AGG	AGATT	TGAG	AGAT	CAACC	674	
V_angularis/1-897	800	CTTCC	TATT	CATTTT	TGGG	TCGGG	ACCCG	AGGGCC	TT	TTTTTT	CT	CAAG	AGG	AGATT	TGAG	AGAT	CAACC	668	
V_radiata/1-894	597	CTTCC	TATT	CATTTT	TGGG	TCGGG	ACCCG	AGGGCC	TT	TTTTTT	CT	CAAG	AGG	AGATT	TGAG	AGAT	CAACC	665	

*C\_papaya/1-1341* 675 AACCACAAATCCGTAGCCCGGGTGACTCACAGCCCTCCCTCTCCCAATGAAATGGGATGAATCAAAATC 745  
*V\_faba/1-1232* 606 .....AAAAAAATGGGATTAATTAAAAA 628  
*M\_pinnatta/1-863* 675 AACCACAAATCCGTAGCCCAAGG..... 697  
*V\_angulatis/1-897* 669 AACCACAAATCCGTAGCCCAAGGTTGACTCACAGCCCTCCCTCTCCCAAGAAATGGGATGAATTAAAAA 739  
*V\_radiata/1-894* 666 AACCACAAATCCGTAGCCCAAGGTTGACTCACAGCCCTCCCTCTCCCAAGAAATGGGATGAATTAAAAA 736

*C\_papaya/1-1341* 746 AAT..AAGCTTTTGTATTGATTGAGGGCGCAGCGAAGCCAAATTTCAATCAAGGCAAAAGGGGG..... 806  
*V\_faba/1-1232* 629 AA...AAGCTTTTGTATTGATTGAGGGCGCAGCGAAGCCAAATTTATTTCAAGGCAAAAGGGGGACTTACTT 696  
*M\_pinnatta/1-863* 740 AAAAAAGCTTTTGTATTGATTGAGGGCGCAGCGAAGCCAAATTTCTTTCAAGGCAAAAGGGGG..... 802  
*V\_angulatis/1-897* 737 AAAAAAGCTTTTGTATTGATTGAGGGCGCAGCGAAGCCAAATTTCTTTCAAGGCAAAAGGGGG..... 799  
*V\_radiata/1-894*

*C\_papaya/1-1341* 807 ..CTTACTTTTCTGTACGCTGAGTCATCCTCTTCAAATTTAGCTATGCTAATGGAAGAGGAAAAGTTTGA 875  
*V\_faba/1-1232* 697 ACTTTACTTTTCTGTACATC...CTATTCAAATCAAATTTAGCTATGTTAATCGAACAGGAAAAGTTTTAA 764  
*M\_pinnatta/1-863* .....  
*V\_angulatis/1-897* .....  
*V\_radiata/1-894* .....

*C\_papaya/1-1341* 876 GATGATATGGACCCCAAGAGATGAGCGAGAACCTCCAATTGCTTAGGGGTCTGCACTCTCTCCCGCTTGGC 946  
*V\_faba/1-1232* 765 GATGATATGGACCC...TGGACGAGAACCTCCAATTGCTTAGGGGTCTGCACTCTCTCCCGCTTGGT 827  
*M\_pinnatta/1-863* .....  
*V\_angulatis/1-897* .....  
*V\_radiata/1-894* .....

*C\_papaya/1-1341* 947 GGACGAGATCTTCTCTCCTTTTATCATTCTTTCTCCACACACCTTTTCCCTCTTTTACCAGAGAGGAAAG 1017  
*V\_faba/1-1232* 828 GGACGAAATCTTCTCTCCTTTTAGAATGATTCTCTCCACACACCTTTTCCCTCTTTTACCAGAGAGGAAAG 898  
*M\_pinnatta/1-863* .....  
*V\_angulatis/1-897* .....  
*V\_radiata/1-894* .....

*C\_papaya/1-1341* 1018 AATAATCTTCCAAAGCGACAGAGACCTAAATTTCCATTAGATTCAATTCCTAAGCTTGCTTTTGTTCAGCAA 1088  
*V\_faba/1-1232* 899 AAAAAATCTTCCAAAGCGACAGAGACCTAAATTTCCATTAGATTCAATTCCTAAGCTTGCTTTTGTTCAGTAA 969  
*M\_pinnatta/1-863* .....  
*V\_angulatis/1-897* .....  
*V\_radiata/1-894* .....

*C\_papaya/1-1341* 1089 GATGATCAGTCCAAAGAGTCTCTGGAGAGAGAGAAAGCGGTAAAAACCTCTCTGATTCTGGTCACCGAGCGGT 1159  
*V\_faba/1-1232* 970 GATGATCAGTCCGAGAGTCTCTGGGAGAGAGAGAAAGCGGTAAAAACCTCTCTGATTCTGGTCACCGAGCAGT 1040  
*M\_pinnatta/1-863* .....  
*V\_angulatis/1-897* .....  
*V\_radiata/1-894* .....

*C\_papaya/1-1341* 1160 GGGACGACCTTTTCAATAACCCAGGATGACCCATTGACGCTTCTTTT...GCTGTATACCCCTTCCATCC 1226  
*V\_faba/1-1232* 1041 CGTCCGACTCTTAAGTAACCCAGGATGACCCCTTTCAGCGCTCTCTTTTCTGTATACCCCTTCCATCC 1111  
*M\_pinnatta/1-863* 698 .....CCCTTTCAGCGCTCTCTTTTCTGTATACCCCTTCAATCC 739  
*V\_angulatis/1-897* 803 .....GCTTACT..... 809  
*V\_radiata/1-894* 800 .....GCTTACT..... 806

*C\_papaya/1-1341* 1227 TTCAA...AAGT...CAAAGAAAGGGTACTATATATATATATATATATTATATATATATAGTAGGGC 1288  
*V\_faba/1-1232* 1112 TTCGG...AGGTGGAGAGAGAGAAAGGTACTATAAAAAATGT-TGTATTTATCTTTTTA...AGTAGGGC 1175  
*M\_pinnatta/1-863* 740 TTCGGAGGTAGGTGGAGAGAGAGAAAGGTACTATAAATAATAAATAATGGATTCTTTTTTT...TAGGGC 806  
*V\_angulatis/1-897* 810 .....TACAATA.....ATGGATTCTTTTTTTAAGTAGGGC 840  
*V\_radiata/1-894* 807 .....TACAAGA.....ATGGATTCTTTTTTTAAGTAGGGC 837

*C\_papaya/1-1341* 1289 CCCAGAACGCTCAAAAGGTGGGGGAACCAAGAGTAGTCACGATAGGAAAGAG...AA 1341  
*V\_faba/1-1232* 1176 CCCAGAACGAAAAAAGGTGGGGGAACCTGAGTTGTACGATAGGAAAGAGCAAAAA 1232  
*M\_pinnatta/1-863* 807 CCCAGAACGAAAAAAGGTGGGGGAACCAAGAGTTGTACGATAGGAAAGAGAAAAA 863  
*V\_angulatis/1-897* 841 CCCAGAACGAAAAAAGGTGGGGGAACCAAGAGTTGTACGATAGGAAAGATCCAAAA 897  
*V\_radiata/1-894* 838 CCCAGAACGAAAAAAGGTGGGGGAACCAAGAGTTGTACGATAGGAAAGATCCAAAA 894