Sequence of the chloroplast-encoded *atpB*-*atpE*-*trnM* gene clusters from rice

Eunpyo Moon, Teh-hui Kao and Ray Wu

Section of Biochemistry, Cornell University, Ithaca, NY 14853, USA
Submitted March 3, 1987

Accession no. Y00323

Previously, we described the isolation and characterization of two rice chloroplast *Hind*III fragments, Ct-1 and Ct-3. Each of these contains a gene cluster coding for the large subunit of ribulose-1,5-bisphosphate carboxylase (*rbcL*), β and ϵ subunits of ATPase (*atpB* and *atpE*), *tRNA*^{met} (*trnM*), and *tRNA*^{val} (*trnV*) (Moon *et al.*, Nucleic Acids Res. 15, 611-630, 1987). We also isolated and characterized a rice mitochondrial *Hind*III fragment (Mt-0) which contains an almost complete transferred copy of this chloroplast gene cluster (Moon *et al.*, in preparation). In the above reports, the *atpB*-*atpE*-*trnM* gene cluster was not included. Here, the *atpB*-*atpE*-*trnM* sequence of the functional copy in Ct-3 is shown and compared to the defective copy in Ct-1 and the transferred copy of Mt-0.

atpB, *atpE* and *trnM* Sequence

Ct-3	5' - CCCCCTTTTC TTATTTTGAG TCCAAATACC TAAATACTAC GAAAATTCTC TGTTCACAGC AATCTATGCT TCACAGTAGT ATATATTTTG	-308
Ct-1		
Mt-0	-- T G A	
Ct-3, 1	TATATCGAAG TCTAGATAA GAAAGTAGAG TAGGCACAAA TCGTTTACAA AAGGCCAAAT GTATATGAAA AAAAGATTGA TTGAACCTTC	-218
Mt-0	- G G C C G	
Ct-3, 1	CGACGGGCTC ATTCCATGAG TAAAGGATTG AATGGGATTC GTTTGGGCCAA CGAAATCAAG TCGTGGTCCC CTTTTCTCTC TTATTGAAAT	-128
Mt-0	A GGA	
Ct-3, 1	AACATAATCA TTTCCTTTTG ACTTTTC--G GATTTTGGGA TATTTTTTTG GTGTGATTT GGCATTATTC AACAAAGAAA AAATCAAAAT	-40
Mt-0	TT TT T	
Ct-3, 1	TTGATAAAT TCGTTTTTTT T-GAAAAATTA TGTGATAATT ATG AGA ACC AAT CCT ACT ACT TCT CGT CCC GGG GTT TCT ACA	42
Mt-0	T -T-----	
Ct-3, 1	ATT GAA GAA AAA AGT ACA GGG CGT ATC GAT CAA ATT ATT GGA CCC GTG CTG GAT GTC ACT TTT CCC CCG GGC AAG	117
Mt-0	C	
Ct-3, 1	TTA CCT TAT ATT TAT AAT GCT TTG GTA GTC AAG ACT CGA GAC ACT GAG GGT AAG CAA ATT AAT GTA ACT TGT GAG	192
Mt-0	C - - - - -	
Ct-3, 1	GTA CAA CAA TTA TTA GGA AAT AAT CGA GTT AGA GCT GTA GCT ATG AGT GCT ACA GAT GGG TTG ATG AGA GGA ATG	267
Mt-0	Set I	
Ct-3, 1	GAA GTG ATT GAC ACG GGA GCT CCT CTC AGT GTT CCT GTC GGT GGA GCT ACT CTT GGA CGA ATT TTC AAC GTT CTT	342
Mt-0	-----	
Ct-3, 1	GGG GAG CCT GTT GAC AAT TTG GGT CCT GTA GAT ACT AGT GCA ACA TTC CCT ATT CAT AGA TCC GCG CCC GCC TTT	417
Mt-0	-----	
Ct-3, 1	ATC GAG TTA GAT ACG AAA TTA TCC ATC TTT GAA ACT GGT ATT AAG GTG GTC GAT CTT TTA GCT CCT TAT CCG CGT	492
Mt-0	-----	
Ct-3, 1	GGA GGA AAA ATC GGA CTA TTT GGG GGA GCT GGA GTA GGT AAA ACA GTA CTC ATC ATG GAA TTA ATC AAC AAT ATT	567
Mt-0	A	
Ct-3, 1	GCT AAA GCT CAC GGG GGC GTA TCC GTA TTT GGC GGA GTA GGG GAA CGG ACT GGT GAA GGA AAT GAT CTT TAT ATG	642
Mt-0	T A A	
Ct-3, 1	GAA ATG AAG GAA TCT GGA GTA ATT AAT GAA AAA AAT-CTT GAG GAA TCA AAG GTA GCT CTA GTC TAT GGC CAA ATG	717
Mt-0	A - - - - - C	
Ct-3, 1	AAT GAA CCG CCA GGA GCT GCT ATG AGA GTT GGT TTG ACT GCC CTA ACT ATG GCA GAA TAT TTC CGA GAT GTT ATT	792
Mt-0	Set I	
Ct-1	-	

Ct-3	AAG CAA GAC GTG CTT CTA TTC	<u>ATC GAT</u>	AAT ATC TTT CGT TTT GTT CAA GCA GGA TCG GAG GTA TCT GCC TTA TTA	867
Ct-1	-			
Mt-0			G T	
Ct-3, 1, Mt-0	GGG AGA ATG CCC	<u>TCT GCA</u>	GTG GGT TAT CAA CCT ACT CTT AGT ACA GAA ATG GGT TCT TTG CAA GAA AGA ATT ACT	942
Ct-3, 1, Mt-0	TCT ACT AAA AAG GGA TCT ATA ACT TCG ATC CAA GCG GTT TAT GTA CCT GCG GAC GAT TTG ACC GAC CCT GCT CCT			1017
Ct-3	GCT ACA ACA TTT GCA CAT TTG GAT GCT ACT ACC GTA CTT TCC AGA GGA TTA GCT TCC AAA GGG ATT TAT	<u>CCT GCA</u>		1092
Ct-1				
Mt-0	C		G T	
Ct-3, 1	GTA GAT CCT TTA GAT TCA ACC TCA ACT ATG TTA CAA CCT CGG ATC GTT GGC AAC GAA CAT TAT GAA ACT GCG CAA			1167
Mt-0	-			
Ct-3, 1, Mt-0	AGA GTT AAG CAA ACT TTA CAA CGT TAC AAA GAA CTT CAG GAC ATT ATC GCA ATT CTT GCG TTG GAT GAA TTA TCG			1242
Ct-3, 1	GAG GAG GAT CCT TTA ACT GTA GCA AGA GCA CGA AAA ATT GAG CGC TTC TTA TCA CAA CCG TTT TTT GTG GCA GAA			1317
Mt-0			C C	
Ct-3	GTT TTT ACC GGT TCT CCA GGA AAG TAT GTT GGT CTT GCA GAA ACT ATT AGG GGA TTT CAA CTA ATC CTT TCC GGA			1392
Ct-1	-			
Mt-0		A C		
Ct-3, 1, Mt-0	GAA TTA GAC GGC CTA CCC GAA CAG GCT TTT TAT TTG GTG GGT	<u>ATC GAT</u>	GAA GCT AGC ACG AAA GCT ATA AAC	1467
Ct-3, 1	TTA GAA GAG GAG AAC AAC TTG AAG AAATGAAA TTA AAT CTT TAT GTA CTG ACT CCT AAG CGA ATT ATT TGG GAT TGT			1544
Mt-0	A			
Ct-3, 1	GAA GTG AAA GAA ATC ATT TTA TCT ACT AAT AGT GGC CAA ATT GGC GTA TTA CCA AAC CAC GCC CCC ATT AAC ACA			1619
Mt-0	T			
Ct-3, 1, Mt-0	GCT GTA GAT ATG GGT CCC TTG AGA ATA GCG CTC CTC AAC GAT CAA TGG TTA ACG GCG GTT CTG TGG AGC GGT TTT			1694
Ct-3, 1, Mt-0	GCC AGA ATA CTT AAT AAT GAG ATC ATC ATT TTA GGA AAT GAT GCG GAA CTG GGT AGT GAC ATT GAT CCG GAA GAA			1769
Ct-3, 1, Mt-0	GCT CAA CAG GCA CTT GAA ATA GCC GAA GCT AAC GTG ACT AGA GCT GAG GGT ACG AAA GAA TTG GTT GAA GCG			1841
Ct-3, 1	AAG GTA GCT CTC AGA GCA GCT AGG ATA CGA GTC GAG GCT GTT AAT TGG ATT CCC CCA TCT AAT TGA AGACAACCCA			1917
Mt-0	C	G	T	
Ct-3, 1	ACGGTTTAGT TGATACAAG AAAAAGGGAA GAGGGGTAGA AAAAATTATT AGATAACGAA GCGAAGTAGG GCCAATGCTA TCTAGTAATT			2007
Mt-0		-- ----- CT G A T		
Ct-3, 1, Mt-0	atggatga taacctaac ttggttactg agggcgcat actttcgtta tgagattggt gactcaattc atccg			2097
Ct-3, 1, Mt-0	TTTCTACCTA CTACCTACT ATTGGATTG AACCAATGAC TCCCGCCGTA TGAAAGCAAT ACTCTAACCA CTGAGTTAAG TAGGCAATT			2129
Ct-3, 1, Mt-0	ATCACCACAA AGGAAGACCC TTTACTTCGA TC			

Figure Legend. Sequences were aligned to maximize homology. Only those nucleotides which differ from the Ct-3 sequence are shown for Ct-1 and Mt-0. Dashes indicate gaps that are necessary to align sequences. The number at the right of each row indicates the position of the last base in that row relative to the first base of the initiation codon of atpB (nucleotide #1). The first nucleotide of the initiation codon for atpE is #1494. Restriction sites are shown above the nucleotide sequence. Nucleotides 1494-1497 represent the sequence overlap between atpB and atpE. Since trnM is transcribed in the opposite direction to atpB and atpE, this sequence on the complementary strand is represented in lower case letters.

Acknowledgements

This work was supported by research grants RF84066, Allocation No. 3, from the Rockefeller Foundation, and GM29179 from NIH.