

Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome

Yasunari Ogiwara, Yukiko Yamazaki¹, Koji Murai², Akira Kanno³, Toru Terachi⁴, Takashi Shiina, Naohiko Miyashita⁵, Shuhei Nasuda⁵, Chiharu Nakamura⁶, Naoki Mori⁶, Shigeo Takumi⁶, Minoru Murata⁷, Satoshi Futo⁸ and Koichiro Tsunewaki^{5,*}

Faculty of Agriculture, Kyoto Prefecture University, Sakyo-ku, Kyoto 606-8522, Japan, ¹National Institute of Genetics, Misima 411-8540, Japan, ²Faculty of Bioscience and Biotechnology, Fukui Prefectural University, Matsuoka, Fukui 910-1195, Japan, ³Faculty of Life Sciences, Tohoku University, Aoba-ku, Sendai 980-8577, Japan, ⁴Faculty of Engineering, Kyoto Sangyo University, Kita-ku, Kyoto 603-8047, Japan, ⁵Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan, ⁶Faculty of Agriculture, Kobe University, Nada-ku, Kobe 675-0013, Japan, ⁷Research Institute for Bioresources, Okayama University, Kurashiki, Okayama 710-0046, Japan and ⁸FASMAC Co. Ltd, Atsugi, Kanagawa 243-0041, Japan

Received July 7, 2005; Revised September 5, 2005; Accepted October 7, 2005

DDBJ/EMBL/GenBank accession no. AP008982

ABSTRACT

The application of a new gene-based strategy for sequencing the wheat mitochondrial genome shows its structure to be a 452 528 bp circular molecule, and provides nucleotide-level evidence of intra-molecular recombination. Single, reciprocal and double recombinant products, and the nucleotide sequences of the repeats that mediate their formation have been identified. The genome has 55 genes with exons, including 35 protein-coding, 3 rRNA and 17 tRNA genes. Nucleotide sequences of seven wheat genes have been determined here for the first time. Nine genes have an exon–intron structure. Gene amplification responsible for the production of multicopy mitochondrial genes, in general, is species-specific, suggesting the recent origin of these genes. About 16, 17, 15, 3.0 and 0.2% of wheat mitochondrial DNA (mtDNA) may be of genic (including introns), open reading frame, repetitive sequence, chloroplast and retro-element origin, respectively. The gene order of the wheat mitochondrial gene map shows little synteny to the rice and maize maps, indicative that thorough gene shuffling occurred during speciation. Almost all unique mtDNA sequences of wheat, as compared with rice and maize mtDNAs, are redundant DNA. Features of the gene-based strategy are

discussed, and a mechanistic model of mitochondrial gene amplification is proposed.

INTRODUCTION

The mitochondrial genome is important in plant development, as well as in productivity (1–3), and extensive studies have been done on its functions (4). Although the complete nucleotide sequence has been determined for seven land plant species (5–11), the genomic makeup is not well understood (11–13) because of the multipartite structure of the genome (14–16). With a new gene-based strategy for sequencing the wheat mitochondrial genome, we obtained a number of recombinant molecules, analyses of which for the first time have provided proof, at the nucleotide sequence level, of the mechanism that produces multipartite molecules in the mitochondrial genome. Moreover, we demonstrate by gene map comparison that thorough gene shuffling occurred during the speciation of three cereals (wheat, rice and maize), leading to remarkable changes in their mitochondrial genome structures, as previously shown by the restriction fragment mapping of maize mitochondrial DNA (mtDNAs) (17) and by MultiPipMaker analysis of several sequenced plant mitochondrial genomes (10). Based on this information, we propose a new method for quantifying genome-wide molecular changes in mitochondrial genomes, which result in ontogenetic as well as phylogenetic variability of the cereal mitochondrial genomes.

*To whom correspondence should be addressed at 6-14-10 Kasugadai, Nishi-ku, Kobe, Hyogo 651-2276, Japan. Tel/Fax: +81 078 961 4085; Email: ktsunewa@kais.kyoto-u.ac.jp

In the wheat complex, *Triticum* (wheat) and *Aegilops* (goat grass), inter- as well as intra-specific molecular diversity of both the chloroplast and the mitochondrial genomes were studied in order to clarify the phylogenetic relationships of various taxa of the complex, including the origin of wheat (18,19). Diversity among plasmons of their phenotypic effects on various wheat characters also was investigated [for review see (20)]. However, we have not studied the functional relationships between molecular variation and differential phenotypic effects. We determined recently the complete nucleotide sequence and gene content of the wheat chloroplast genome (21). Here we report those of the mitochondrial genome. The information obtained provides a basis for future studies on the linkage of molecular diversity and phenotypic variability in the wheat complex.

MATERIALS AND METHODS

Plant material

The common wheat, *Triticum aestivum* cv. Chinese Spring, was the source of the mtDNA studied here, that was obtained from mitochondria of 14-day-old etiolated seedlings (22) and was purified before use (23).

MtDNA library construction and clone sequencing

An mtDNA library was constructed by the SuperCos1 *in vitro* packaging method (Stratagene, LaJolla) from partially digested wheat mtDNA with Sau3AI. From this library, 232 clones were randomly selected and dot-blotted with 32 mitochondrial genes as probes (24). All probe genes, except *rps13*, hybridized with 7 or more clones, from which 23 clones were

selected to cover 31 probe genes and were sequenced by the shotgun method. Sequenced fragments were aligned using BLASTn (25) to determine the entire sequence of each clone.

Sequence assembly and gene analysis

Alignment of the 23 clones gave two linear molecules of ~350 and 76 kb. Two additional clones, #194 and #204, whose ends hybridized to one end each of the two linear molecules, were selected and sequenced. Phrap (http://www.genome.washington.edu/UWGC/analysis_tools/phrap.htm), BLASTn and blast2sequences programs were used for the primary assembly of all the clones. Manual fine tuning was done to generate the final master circle (MC). Repeat sequences were analyzed by in-house script, window size 8 bp, and represented as a dot-plot image. Open reading frames (ORFs) were identified by a Genome Gambler (Xanagen Co.) and ORFfinder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). tRNA genes were searched for by tRNAscan-SE (26). The annotated rice and maize mitochondrial genes, BA000029 and AY506529, respectively, as well as individual wheat mitochondrial genes submitted to the DNA databank, were compared with our sequence data to annotate all the genes. Sequences homologous to known cereal transposable elements were searched for, referring to the TIGR grass transposable elements database after Clifton *et al.* (10).

Gene nomenclature and nucleotide position

The nomenclature of Clifton *et al.* (10) for maize mitochondrial genes was adopted, except for the designation of exons, for which *ex-1* to *ex-5* of a given gene are indicated by *a* to *e*, affixed to its gene symbol. Positions of a forward-strand

Table 1. Wheat mtDNA clones sequenced, showing their size, type, marker genes used and genes other than probe genes identified by sequencing

Clone	Size (bp)	Type ^a	Probe genes used ^b	Additional genes found by sequencing ^c
#1	37 129	R(S)	<i>nad1a</i> , <i>nad7</i> , <i>rrn5/18</i>	<i>mttB</i> , <i>trnfM</i> , <i>trnP</i> , <i>trnS</i>
#5	33 266	I	<i>ccmFC</i> , <i>rrn26(p)</i>	<i>trnK</i> , <i>trnQ</i>
#6	38 445	I	<i>nad4</i> , <i>nad5de</i>	<i>trnP</i>
#24	35 670	R(D)	<i>cob</i> , <i>rps7</i> , <i>rrn5/18</i> , <i>rrn26</i>	<i>trnF</i> , <i>trnfM</i> , <i>trnM</i> , <i>trnS</i>
#27	15 896	R(S)	<i>nad1a</i>	<i>ccmFC</i>
#31	35 843	R(S)	<i>cox2</i> , <i>nad3</i> , <i>nad9</i> , <i>nad2cde</i> , <i>rps12</i>	<i>orf173</i> , <i>orf349</i> , <i>rps2</i> , <i>trnD</i> , <i>trnS</i> , <i>trnY</i>
#39	16 661	I	<i>cox2</i> , <i>nad2cde</i> , <i>nad9</i> , <i>atp4</i>	<i>orf349</i> , <i>rps2</i> , <i>trnD</i> , <i>trnY</i>
#51	34 696	R(S)	<i>rps7</i> , <i>rrn5/18</i> , <i>rrn26</i>	<i>trnF</i> , <i>trnfM</i> , <i>trnS</i>
#63	35 769	I	<i>cox1</i> , <i>rrn26</i>	<i>trnK</i> , <i>trnQ</i>
#66	36 206	I	<i>nad7</i>	<i>nad4L</i> , <i>rps19(p)</i> , <i>trnD</i> , <i>trnfM</i> , <i>trnI</i> , <i>trnK</i> , <i>trnM</i> , <i>trnN</i> , <i>trnS</i>
#74	34 458	I	<i>nad1a</i> , <i>nad1d</i> , <i>nad5ab</i> , <i>nad6</i> , <i>rrn5/18</i>	<i>rpl2(p)</i> , <i>rps4</i> , <i>trnfM</i> , <i>trnP</i>
#75	35 217	R(S)	<i>atp6</i> , <i>nad1bc</i> , <i>nad5de</i>	<i>ccmFCa</i> , <i>orf194</i> , <i>orf359</i> , <i>rps13</i> , <i>rpl16</i> , <i>rps1</i> , <i>rps3</i> , <i>trnC</i>
#92	37 038	I	<i>atp6</i> , <i>cox3</i> , <i>matR</i> , <i>nad1bc</i> , <i>nad1e</i> , <i>nad5c</i> , <i>nad7</i>	<i>rpl5</i> , <i>rps13</i> , <i>trnE</i>
#94	27 319	R(S)	<i>nad1d</i> , <i>nad6</i> , <i>rrn26</i>	<i>rpl2(p)</i> , <i>rps4</i> , <i>trnK</i> , <i>trnQ</i>
#96	44 184	R(S)	<i>atp6</i> , <i>cob</i> , <i>cox3</i> , <i>rrn5/18</i>	<i>ccmFN</i> , <i>trnE</i> , <i>trnfM</i>
#102	37 709	I	<i>atp6</i> , <i>nad5de</i>	<i>ccmFCa</i> , <i>ccmFCb</i> , <i>orf194</i> , <i>orf359</i> , <i>rpl16</i> , <i>rps3</i> , <i>trnC</i> , <i>trnP</i>
#110	39 360	I	<i>atp6</i> , <i>cob</i> , <i>ccmFN</i> , <i>rrn5/18</i> , <i>rps1</i>	<i>trnfM</i>
#126	34 832	I	<i>cob</i> , <i>cox1</i> , <i>cox2</i> , <i>atp8</i>	<i>trnD</i>
#146	36 595	I	<i>cob</i> , <i>matR</i> , <i>nad1e</i> , <i>nad5c</i> , <i>ccmFN</i> , <i>rrn5/18</i> , <i>rps1</i>	<i>rpl5</i> , <i>trnfM</i>
#160	36 135	I	<i>nad2ab</i> , <i>nad2cde</i> , <i>nad9</i> , <i>atp4</i>	<i>ccmB</i> , <i>orf349</i> , <i>trnK</i> , <i>trnQ</i> , <i>trnY</i>
#162	38 872	R(S)	<i>atp1</i> , <i>atp6</i> , <i>atp9</i> , <i>nad1bc</i>	<i>ccmFCa</i> , <i>ccmFCb</i> , <i>orf194</i> , <i>orf359</i> , <i>rps13</i> , <i>rpl16</i> , <i>rps3</i> , <i>trnC</i>
#190	39 803	I	<i>cox2</i> , <i>nad3</i> , <i>atp8</i> , <i>rps12</i>	<i>mttB</i> , <i>trnS</i>
#194	38 416	I	(None)	<i>atp1</i> , <i>nad4L</i> , <i>rps19(p)</i> , <i>trnD</i> , <i>trnfM</i> , <i>trnI</i> , <i>trnK</i> , <i>trnM</i> , <i>trnN</i> , <i>trnW</i>
#204	34 585	I	<i>rps7</i> , <i>rrn26</i>	<i>trnF</i> , <i>trnS</i>
#224	38 351	R(D)	<i>cob</i> , <i>cox3</i> , <i>nad1a</i> , <i>rrn5/18</i>	<i>mttB</i> , <i>rpl5</i> , <i>trnfM</i> , <i>trnP</i>
Total	872 455		(Average size = 34 898 bp)	

^aI: intact clone; R(S): single-recombinant clone; S(D): double recombinant clone.

^bUnderlined: probe genes not detected by sequencing.

^c(p): partial gene.

nucleotide in the MC molecule and in a gene or repeat sequence, respectively, are shown as the 'MC coordinate' and 'gene or repeat coordinate'.

RESULTS

Sequencing of individual clones and their alignment

Twenty-five wheat mtDNA clones were sequenced (Table 1). Their sizes ranged from 27 to 44 kb, except for two (#27 and #39) ~16 kb in size. The average size was 34 898 bp, and the total size was 872 455 bp. Alignment showed a single 452 528 bp MC molecule (Figure 1). Fifteen clones occupied single locations in the genome ('intact clone'), while the remaining 10 were split into two or three segments located in different parts of the genome, tentatively called the 'recombinant clone'. Quetier *et al.* (15) estimated size of the wheat mitochondrial genome to be ~430 kb, based on its *SalI* restriction map. Their estimate is very close to the size, 452 528 bp, determined by the present sequencing work.

Intra-molecular recombination and site of recombination

Of the 10 recombinant clones, 8 were split into two segments. The other two (#24 and #224) were cleaved into three segments. Without exception, there was a pair of direct repeats

(DRs) or inverted repeats (IRs) at the split site (Figure 1). All the recombinant clones carried a completely or nearly identical copy of the same repeat at the recombination site (details in the next paragraph). DRs connected split fragments head-to-tail, whereas IRs connected them head-to-head or tail-to-tail. These facts indicate that the split clones were produced by intra-molecular recombination between the relevant repeats. In sum, nine repeat pairs, R1 to R9, were responsible for the production of all of the recombinant clones (Table 2). The production of clones #24 and #51 was mediated by the same R7 repeats, whereas #75 (#162 as well) and #96 were reciprocal products of recombination of R8 repeats (Figure 2A). Two clones, #24 and #224, were double recombinants (Figure 2B and C). The former was produced by recombination between two DR pairs, R3 and R7, and the latter recombination between two IR pairs, R2 and R6. Seven additional repeats, R10 to R16, larger than 100 bp were present in the genome (Table 2). Three repeats, R1, R7 and R10, shared a 1634 bp sequence in common, containing a part of *rrn26*. Similarly, three other repeats, R2, R3 and R4, shared a 4430 bp common sequence that carried *trnFM*, *rrn18* and *rrn5*. In addition, small repeats of 30–100 bp in size were detected in a dot-matrix image, of which 35 were the direct and 38 were the inverted types. All those repeats are shown in Figure 3, in which R1 to R16 are marked by arrows. We need to search for whether all of them serve as

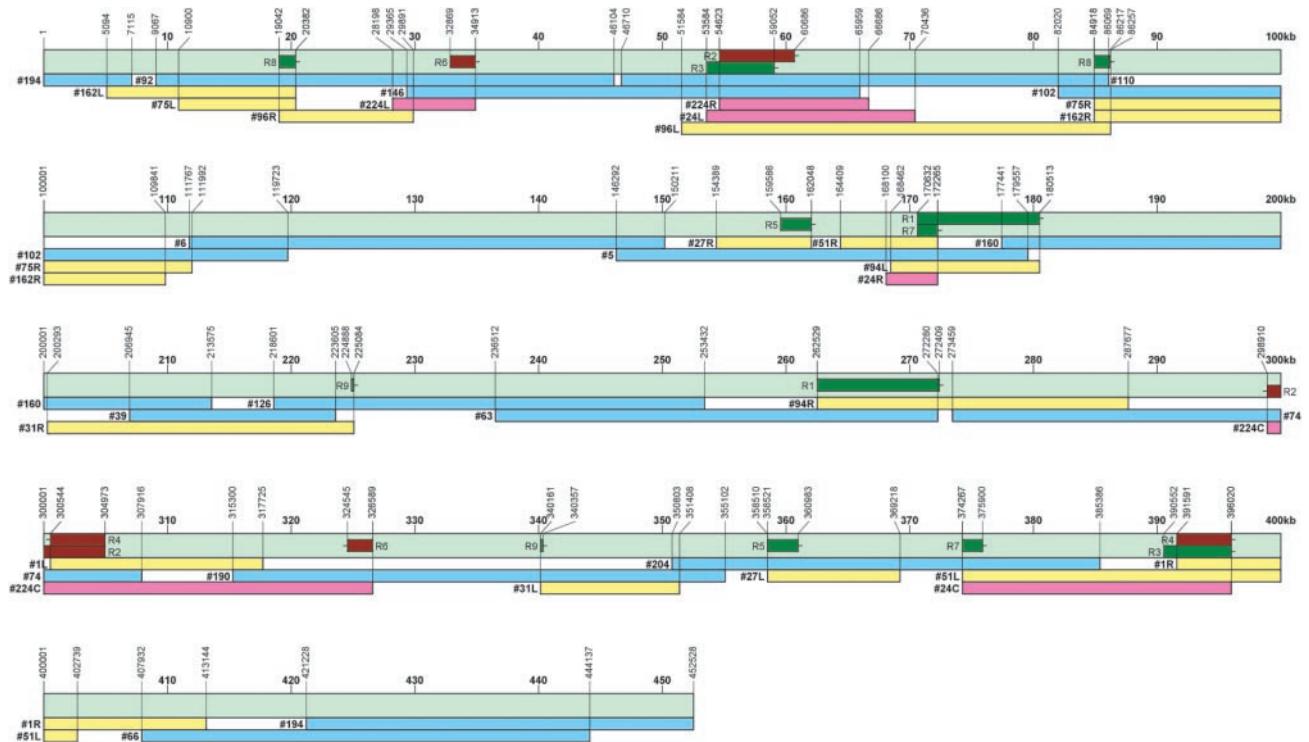


Figure 1. Alignment of 25 mtDNA clones in the 452 528 bp MC molecule of the wheat mitochondrial genome. Broad, light-green bar shows the MC molecule cleaved between MC coordinates 452 528 and 1. Numbers on the MC molecule show the MC coordinates of the ends of all the clones, their segments and repeat sequences. Rectangle with projection in the broad bar: R1–R9 repeat pairs involved in recombinant clone formation. DRs are dark green, IRs dark brown. The projection shows the direction of each repeat copy. Slender bar: individual clones; light blue, yellow and orange represent intact, single recombinant and double recombinant clones, respectively. L, C or R affixed to clone numbers: Left, central and right segments of a recombinant clone. Note that L and R segments of a single-recombinant clone have the same repeat copy at the end connecting two segments; head-to-tail for DRs, and head-to-head or tail-to-tail for IRs. The double-recombinant clone has a copy of one repeat pair at one end of its L and C segments and a copy of another repeat pair at the other end of C and at one end of R that, respectively, connect the L and C segments and the C and R segments by recombination.

Table 2. Repeats involved in intra-molecular recombination, and other repeats larger than 100 bp found in the wheat mitochondrial genome

No.	Type ^a	Size (bp)	Gene in repeats ^b	MC coordinates ^c			Difference between copies	Recombinant clone/fragment
				Copy-1	Copy-2	Copy-3		
R1	DR	9882	<i>rrn26(p)-trnQ-trnK</i>	170 632–180 513	262 529–272 409	—	copy-2 1bp def.	#94R/L
R2	IR	6064	<i>rrn5-rrn18-trnfM</i>	54 623–60 686	304 973–298 910	—	1bp mismatch	#224C/R
R3	DR	5469	<i>rrn5-rrn18-trnfM</i>	53 584–59 052	390 552–396 020	—	identical	#24C/L
R4	IR	4430	<i>rrn5-rrn18-trnfM</i>	304 973–300 544	391 591–396 020	—	1bp mismatch	#1L/R
R5	DR	2463	No gene	159 586–162 048	358 521–360 983	—	identical	#27L/R
R6	IR	2045	No gene	32 869–34 913	326 589–324 545	—	identical	#224C/L
R7	DR	1634	<i>rrn26(p)</i>	170 632–172 265	262 529–264 162	374 267–375 900	identical	#24C/R; #51L/R ^e
R8	DR	1341	<i>atp6</i>	19 042–20 382	84 918–86 257	—	copy-2 1bp def. +7 bp mismatch	#75L/R; #96L/R; #162L/R
R9	DR	197	No gene	224 888–225 084 ^d	340 161–340 357	—	1 bp mismatch	#31L/R
R10	DR	7035	<i>rrn26</i>	257 128–264 162	368 866–375 900	—	identical	None
R11	DR	493	<i>atp8</i>	233 393–233 885	338 884–339 376	—	5 bp mismatch	None
R12	DR	385	<i>trnK</i>	178 466–178 850	270 362–270 746	442 527–442 911	copy-3 1 bp dif.	None
R13	DR	207	No gene	63 737–63 943	233 154–233 360	—	1 bp mismatch	None
R14	DR	190	<i>trnP</i>	117 643–117 832	305 046–305 235	—	4 bp mismatch	None
R15	DR	186	<i>trnD</i>	222 071–222 256	429 332–429 517	—	identical	None
R16	DR	104	No gene	20 239–20 342	86 115–86 218	197 109–197 212	identical	None

^aDR and IR: direct and inverted repeats.^b*rrn26(p)*: partial 422 bp sequence of *rrn26*.^cBoldface: IR copy.^dIncluded in *cox2a*.^eBoth recombinations occurred between copy-1 and -3.

recombination sites or not, although our results showed that a repeat pair as small as 197 bp in size (= R9) mediated recombination.

We tried to identify the recombination site in each repeat pair. Four, R3, R5, R6 and R7, had identical copies. Four others, R1, R2, R4 and R9, had only 1 nt difference between the two copies, located at the extreme end of the repeat (Table 2). Identification of the recombination site therefore was informative only for repeat pair R8, which was involved in the production of three recombinant clones; #75, #96 and #162 (Figure 4). Two copies of this repeat, R8-1 and -2, which carried *atp6-1* and *atp6-2* at the same R8 coordinates, 91–1251, had 8 nt differences; one at R8 coordinate 6, the others between coordinates 1301 and 1316 (Figure 4A). Nucleotide sequences of the two R8 copies and their 5'- and 3'-flanking segments were compared with those of the three recombinants. As for the 5'-flanking sequence and sixth nucleotide of R8, #75 and #162 were the same as the R8-1 copy, whereas #96 was the same as R8-2. As for the 3'-flanking sequence and seven variable nucleotides at R8 coordinates 1301–1316, #75 and #162 were the same as the R8-2 copy, whereas #96 was identical to the R8-1 copy. These findings indicate that all three recombinant clones were produced by recombination in the same 1294 bp segment of the R8 repeat (Figure 4A). Previously, Bonen and Bird (27) sequenced wheat mtDNA segments flanking *atp6*, and found that there are two molecular forms at both the 5' ('downstream' in their designation) and 3' ('upstream') borders of the gene. Their nucleotide sequences were in complete agreement with ours, except for a 1 bp deletion in our R8-2 copy between MC coordinates 86 217 and 86 218. Their sequences 3 and 2 correspond to the 5' and 3' borders of the R8-1 copy, and the sequences 4 and 1 to the 5' and 3' borders of the R8-2 copy (Figure 4B). They located a 6 bp insertion in sequence 3, extending the homologous region between sequences 3 and 4 by 22 bp downstream (toward the 5' end), which was confirmed by our

findings. The sequence comparison (Figure 4B) indicated that the three recombinant clones were produced by recombination in the same 1291 bp segment (3 bp smaller, comparing with the alignment in Figure 4A). Because this segment occupies ~95% of the R8 repeat, it is not surprising that three independent recombination events occurred within this segment.

Bonen and Bird (27) also reported the presence of short DRs in three of the above four sequences, corresponding to the present 5' and 3' borders of the R8-1 copy and the 5' border of the R8-2 copy (Figure 4B), where 'border' means the boundary between a repeat end and its flanking sequence. We examined 60 bp sequences around the 5' and 3' borders (30 bp on both sides of each border) of all repeats shown in Table 2. The complete border sequences are given in Supplementary Table 1. Of 70 border sequences of the 35 repeat copies, 22 contained straight, DRs (no gap, no mismatch) of 3–7 bp while additional 24 possessed aberrant 4–10 bp DRs, having a mismatched nucleotide or a few nucleotides intervening between the repeats, and the remaining 24 did not have short DRs (Table 3). Fourteen repeats had short DRs at both ends, which did not show any sequence similarity, homologous or complementary, to each other. Thus, we conclude that the majority of the repeat ends are associated with short DRs, although their functional role is unknown.

Stern and Palmer (28) indicated that *rrn18* and *rrn26* often are contained in recombination sites of the wheat mitochondrial genome. Our results confirmed this because 6 of the 12 recombination events detected are mediated by repeat pairs containing those genes (Table 2).

Genes and the genetic map of the wheat mitochondrial genome

In all, 55 genes and their exons were identified (Table 4) and mapped on the MC molecule (Figure 5). All the protein-

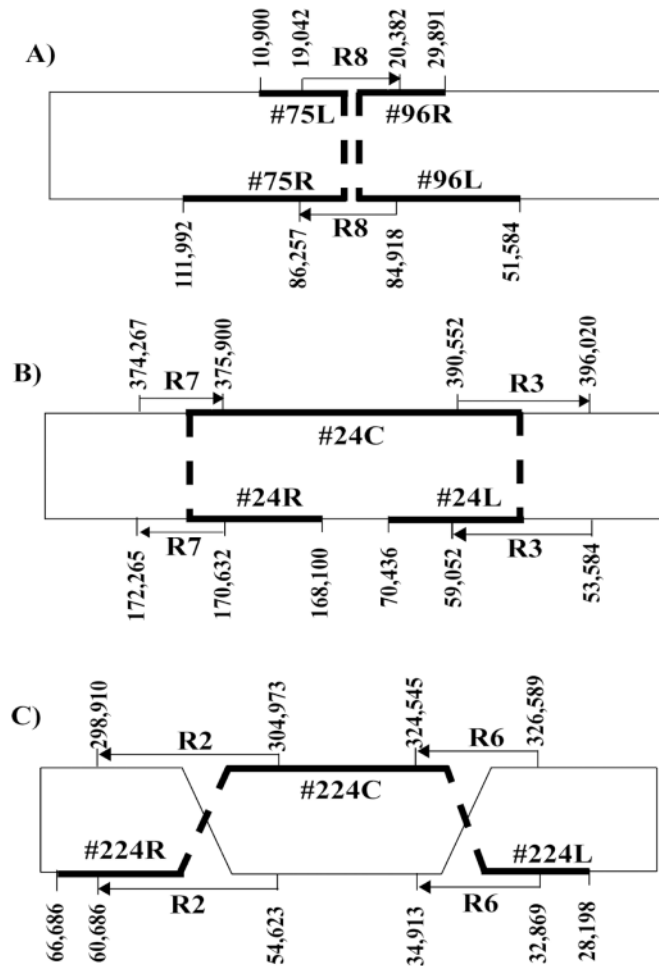


Figure 2. Origins of four recombinant clones obtained by recombination mediated by different repeat pairs. Rectangle: MC molecule. Arrows: DR or IR pairs. Broken line: fusion of separate segments by recombination. Thick and thin lines: cloned DNA segment and remaining part of the recombinant molecule not included in the clone. Numbers on MC molecules: MC coordinates at the ends of repeats and the cloned molecule. *Note:* DRs should be drawn in the same direction by folding the MC molecule with a 180° twist. This was omitted to simplify the figure. (A) Clones #75 and #96 as reciprocal products of R8-mediated recombination. They are part of two subgenomic molecules; (B) clone #24 is the product of double recombination at two DR pairs, R3 and R7; (C) clone #224 is the product of double recombination at two IR pairs, R2 and R6.

rRNA- and tRNA-coding genes known for wheat (24), rice (8) and maize (10) were present, i.e. 9 Complex I genes, 1 Complex III gene, 3 Complex IV genes, 5 Complex V genes, 4 cytochrome c biogenesis genes, 11 ribosomal protein genes, 2 other protein-coding genes, 3 rRNA genes and 17 tRNA genes. Nucleotide sequences of seven wheat genes, *rpl16*, *rps3*, *rps4*, *mttB*, *trnA*, *trnI* and *trnM*, were determined here for the first time. Three genes, *rpl2-p*, *rps19-p*, and *rrn26-p* (the third *rrn26* copy), were truncated. The first two are functional in rice but missing in maize (10). Nine genes, *nad1*, *nad2*, *nad4*, *nad5*, *nad7*, *cox2*, *ccmFC*, *rps3* and *trnA* (chloroplast origin), had the exon-intron structure. The chloroplast counterpart of *trnA* also has an intron (21). All exons of *nad4* (exons *a-d*), *nad7* (*a-e*), *cox2* (*a,b*), *ccmC* (*a,b*), *rps3* (*a,b*) and *trnA* (*5'*-,*3'-ex*) were *cis*-spliced, whereas some exons

of *nad1*, *nad2* and *nad5*, were *trans*-spliced (the slash indicating *trans*-spliced exons) as follows: *nad1a/nad1b,c*, *nad1d/nad1e*; *nad2a,b/nad2c-e*; and *nad5a,b/nad5c/nad5d,e*.

Ten genes were present in multi-copy: *atp6*, *atp8*, *rrn26*, *trnD* and *trnP* were duplicated and *rrn5*, *rrn18*, *trnFM*, *trnK* and *trnQ* triplicated. In addition, three *trnS* genes were found, but they greatly differed each other in nucleotide sequence and therefore were considered different genes, confirming the results of two previous works (29,30).

Restriction fragment analyses of wheat mtDNA revealed the presence of seven molecular forms of the *rrn18-rrn5* cluster (31,32). We identified three copies, Copy-1, -2 and -3, of a three-gene cluster, *trnFM-rrn18-rrn5*, in the MC molecule, all of which were included in three repeats, R2, R3 and R4 (Table 2). Figure 6 illustrates the production of two recombinant forms of this gene cluster from recombination between Copy-1 and -2 (pathway [A]) and Copy-2 and -3 (pathway [B]). Because recombination also occurs between Copy-1 and -3, six recombinants are expected altogether. We obtained three of them, which were produced by recombination between Copy-1 and -2 (#1L/R), Copy-1 and -3 (#224C/R) and Copy-2 and -3 (#24C/L) (Table 2). None of their reciprocal products was obtained, probably as a matter of chance owing to the small number of the clones examined, because the fourth recombinant molecule is reported by Lejeune *et al.* (32). As for *rrn26*, two molecular forms of its 5' end, and three forms of the 3' end had been predicted previously (15,32). This prediction was verified by the present findings confirming two complete and one partial copy (422 bp 3' end) of *rrn26*.

Two copies of *atp8* had five mismatched nucleotide pairs scattered within the 471 bp gene region. Sequence analyses of recombinant molecules supposedly produced by recombination between the R11 repeats containing this gene (Table 2) might be useful in specifying recombination site(s) within the repeat.

In addition to those genes, 179 ORFs larger than 300 bp were found (Supplementary Table 2). Their total size amounted to 75 465 bp, occupying ~16.7% of the entire genome. This number greatly exceeds the 121 ORFs of comparable size reported for maize (10), in spite of the fact that the wheat mitochondrial genome is much smaller than the maize genome. Functional analysis of those ORFs will be an important problem in the future mitochondrial genomics.

MtDNA sequences homologous to ctDNA

Homology search using the blast2sequence program revealed that the wheat mitochondrial genome has 55 sequences homologous (mostly with 80% or higher homology on the nucleotide basis) to the corresponding sequences of the wheat chloroplast genome (Table 5). Exceptions were nine sequences question-marked in the last column of Table 5, which were mosaic of highly conserved and variable sequences, showing segmental differentiation of the sequences. Sizes of individual sequences vary between 27 bp for the smallest and 4239 bp for the largest. The total size, 26 264 bp, corresponds to 5.80% of the entire genome.

Of the above 55 wheat mtDNA sequences, 8 carried native (not chloroplast-derived) mitochondrial genes, *atp1*, *rrn18-1*, -2, -3, *rrn26-1*, -2, -*p* and *trnM*, whose total size amounted to

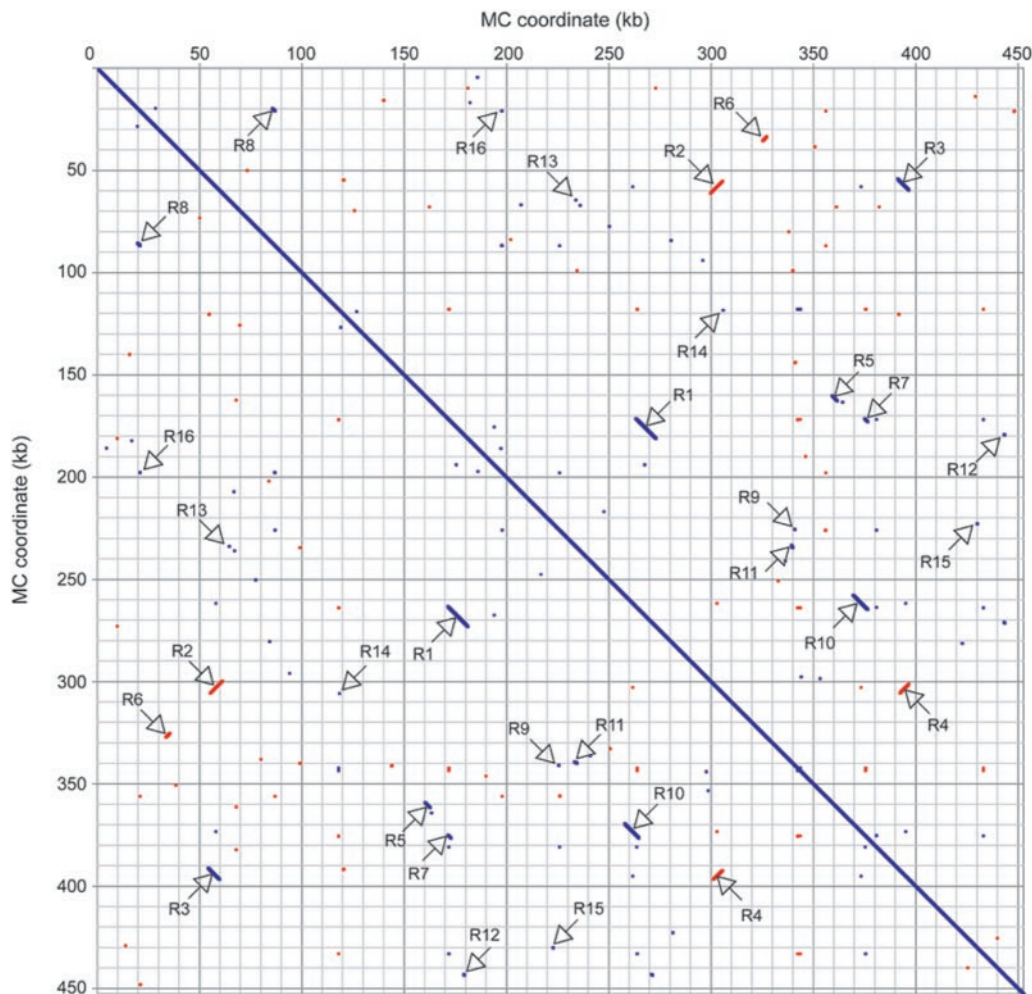


Figure 3. Dot matrix of the MC molecule of wheat mitochondrial genome, showing direct (blue) and inverted (orange) repeat pairs of larger than 30 bp. Sixteen repeat pairs, R1–R16, of larger than 100 bp are marked by arrows (Table 2).

12 809 bp. They showed homology to the ctDNA sequences carrying the corresponding chloroplast genes, *atpA*, *rrn16*, *rrn23* and *trnM* (marked by circles in Table 5). Each of the gene pairs, *atp1/atpA*, *rrn18/rrn16*, *rrn26/rrn23* and *mt-trnM/ct-trnM*, is assumed to have originated from a common prokaryotic gene, being homoeologous to each other (evidence will be reported elsewhere). The total size of the mtDNA sequences of real chloroplast origin therefore was estimated as 13 455 bp; 2.97% of the wheat mitochondrial genome, compared with 22 593 bp (6.3%) and 25 281 bp (4.4%) reported, respectively, for rice and maize (8,10). Thus, both the total size and proportion of the chloroplast-derived sequences relative to the entire genome were smallest in wheat, comparing with rice and maize.

Gene shuffling in the cereal mitochondrial genome

We compared mitochondrial gene maps of wheat, rice (8) and maize (10), excluding tRNA genes, pseudogenes and ORFs (Figure 7). Five exons of *nad7*, *nad7a* to *e*, showed a common arrangement in the three cereals. This gene was used to mark the common map origin, and the arrangement of *nad7a* to *nad7e* to mark the common map direction. A syntenic

gene/exon arrangement, then, should appear as a row of genes/exons parallel to either diagonal line. Only a few gene/exon clusters of the three cereals showed synteny. One 5-gene cluster, *ccmFN-rps1-matR-nad1e-nad5c*, and five 2-gene clusters, *rps13-nad1bc*, *rrn18-rrn5*, *rps3-rpl16*, *nad9-nad2cde* and *nad3-rps12*, showed synteny. The third and fourth ones are shown as 3-gene clusters in Figure 7, because maize has an extra copy of both *rps3a* and *nad2de* and, for this, *rps3a* and *rps3bcd*, and *nad2c* and *nad2de* were shown separately. Similarly, *nad4abc* and *nad4d* were shown as a 2-gene cluster because rice has two extra copies of *nad4d*. In addition, three 2-gene clusters, *rps19(p)-nad4L*, *ccmB-nad2ab* and *nad5ab-rpl2(p)*, of wheat and rice conserved synteny, and two 2-gene clusters (*cox1-rrn26* and *nad6-rps4*) of wheat and maize preserved synteny. No synteny was detected for any other gene combinations, indicative that frequent gene shuffling occurred during cereal speciation, resulting in remarkable structural differences in the cereal's mitochondrial genomes. Fauron *et al.* (17) showed by the physical map comparison that mitochondrial genome restructuring has taken place between three maize cytotypes, and Clifton *et al.* (10) demonstrated by MultiPipMaker analysis that little sequence similarity exists between mitochondrial genomes of six plant

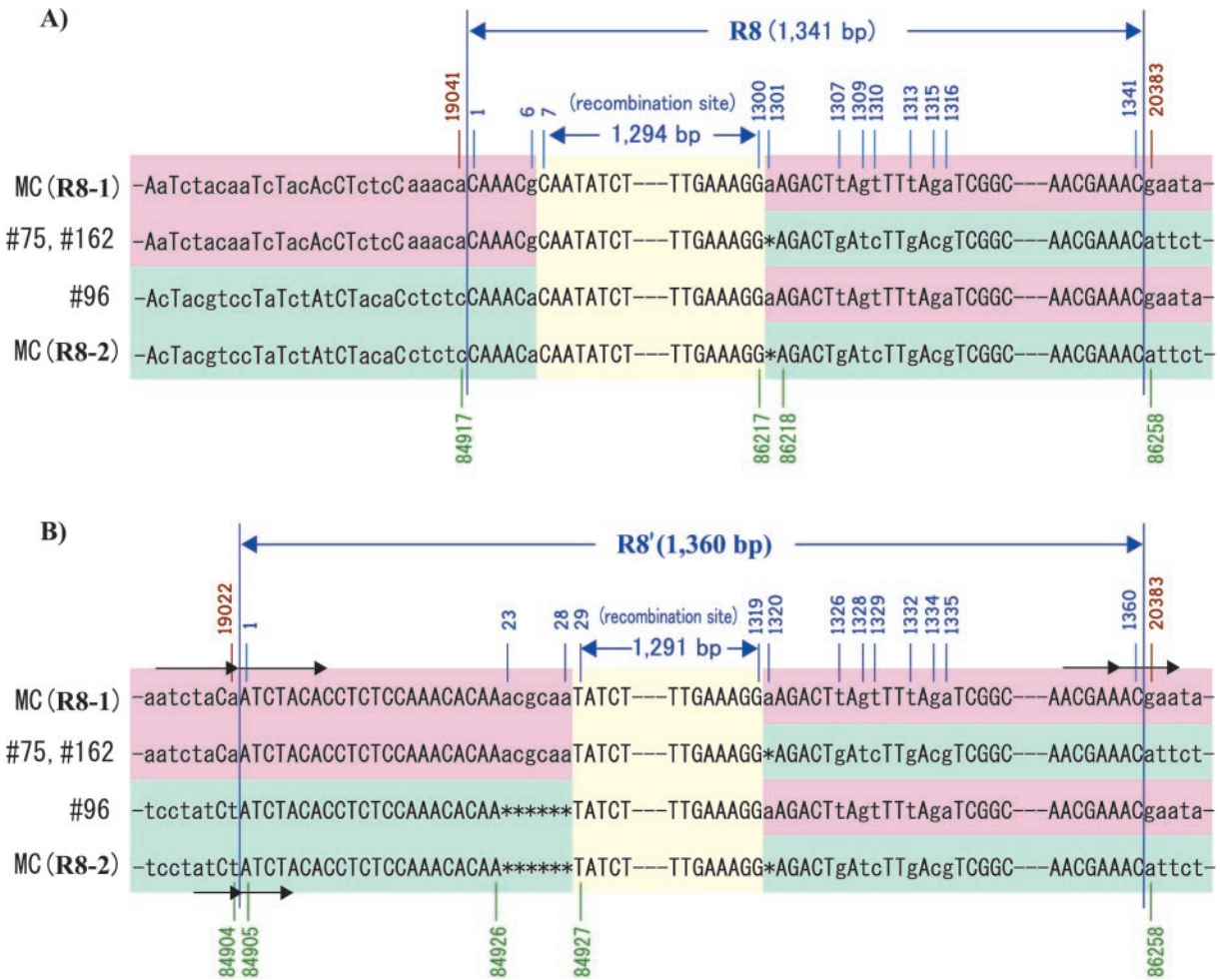


Figure 4. (A and B) Recombination site in R8 repeats which produced the three recombinant clones, #75, #96 and #162. Nucleotide sequences in pink, light green and yellow backgrounds, respectively, are sequences homologous to an R8 copy (R8-1) and its flanking regions, sequences homologous to the other R8 copy (R8-2) and its flanking regions, and the recombination site sequence. In this figure, the forward strands are shown, which are antisense relative to the *apt6*-coding sequence. Numbers outside and inside the R8 or R8' repeat: MC coordinates of the nucleotides flanked respectively by two R8 copies and the R8 or R8' coordinates of the variable nucleotides between them. Capital and lower case letter: Consensus and unique nucleotide between two R8 copies and their flanking regions. Asterisk: deficient nucleotide.

species. Those results agree with ours of the above cereal gene map comparison.

DISCUSSION

Features of the gene-based strategy for sequencing plant mitochondrial genomes

Two principal strategies have been used to sequence plant mitochondrial genomes; the physical map-based (5,7–9), and the genome shotgun strategies (6,10,11). We used a new gene-based strategy for wheat, facilitated by the fact that many wheat mitochondrial genes are available as probes (24) for selecting wheat mtDNA clones for sequencing. Use of this strategy gave a complete picture of the wheat mitochondrial genome by sequencing the 872.5 kb mtDNA, less than twice the genome size, 452 528 bp. Comparative values for the genome shotgun strategy are ~4, 8 and >20 times for *Arabidopsis*, tobacco and maize (6,11,10), indicating apparent high genome sequencing efficiency of the gene-based strategy.

However, application of this strategy requires construction of a cosmid mtDNA library and selection of mtDNA clones covering known mitochondrial genes by dot hybridization. The overall efficiency of the gene-based strategy, compared with that of the genome shotgun strategy, is not clear.

The advantage of the gene-based, compared with the physical map-based strategy, is that no physical map construction is required. This is difficult with some plants because of the multipartite structure of the mitochondrial genome. Based on the physical map of the mitochondrial genome of a common wheat cultivar, Capitole (15), Lejeune and Quetier [cited from (24)] constructed the first gene map of the wheat mitochondrial genome, to which 36 genes were allocated. Their map completely matches ours for five local gene maps: (i) *rrn18/rrn5-cob-atp6-nad5de-nad4abcd-nad2ab-orf25* (= *atp4*)-*nad2cde-nad9-cox2ab-cox1-rrn26*, (ii) *rps7-rrn18/rrn5-nad7abcde-atp1-atp9-nad1bc-rps13-atp6-cox3*, (iii) *nad1a-rrn18/rrn5-nad5ab-nad1d-nad6-rrn26*, (iv) *nad3-rps12-orf156* (= *atp8*) and (v) *matR-nad1e-nad5c*. The arrangement of these five gene groups within the genome,

Table 3. Short DRs found in the 5'- and 3'-borders of 16 repeats, R1–R16, in the wheat mitochondrial genome

Repeat	5'-Border	3'-Border
R1-1	—	TGGt/gg
R1-2	C/CCC+cccc	—
R2-1	—	AA/Aaaa
R2-2	AGGTagg/t	AA/Gaag
R3-1	TTT/TCT++++tttct	—
R3-2	CCC/T++ccct	—
R4-1	—	AA/Gaag
R4-2	ACATA+++acata/	—
R5-1	ACCTAa/ccta	—
R5-2	CCTA+/ccta	—
R6-1	GCAA/gcaa	ATTTC+++att/tc
R6-2	CAAc/aa	CTTGC/ATC++++ctgtctc
R7-1	AT/CCCaccucc	TATTTCA+/atttaa
R7-2	C/CCC+cccc	TATTTCA+/atttaa
R7-3	C/CCC+cccc	—
R8'-1	ATCTACA/atctaca	ACGAAac/gaa
R8'-2	ATCT/atct	—
R9-1	—	—
R9-2	—	—
R10-1	TGCTTTCCTC++++tctttcttc	TATTTCA+/atttaa
R10-2	/TCTTTCtctttct	—
R11-1	TCA/tca	AAG/Aaaga
R11-2	—	AAATAAG/+aaaaaag
R12-1	—	AGATCaga/tc
R12-2	—	AGATCaga/tc
R12-3	A/GTagt	—
R13-1	—	—
R13-2	—	TAT/TC+tattc
R14-1	GCGCTgc/ggt	—
R14-2	GGT/ggt	AGGCagg/c
R15-1	AAGA/aaga	T/ATAtata
R15-2	GAAGA+g/aaga	TTTCTT/ttctt
R16-1	/AATAGCA+++aatagca	GAAAGga/a*g
R16-2	/AATAGCA+++aatagca	/GA*CTgatct
R16-3	/AATAGCA+++aatagca	GAAAGga/a*g

Slash, border; plus, intervening nucleotide between short repeat sequences; asterisk, deficient nucleotide; capital and lower-case letters, short DR sequences; underlined, mismatched nucleotide; sequences of R8'-1 and -2, first identified by Bonen and Bird (27), and confirmed here.

however, differ both in order and direction. Their order in Lejeune and Quetier's map is that shown above, whereas in our map it is (i)–(iii, reverted)–(iv, reverted)–(ii)–(v, reverted). Whether this discrepancy is due to the different mtDNA sources, or to problems in the physical map they relied upon, needs to be clarified.

The gene-based strategy for complete mitochondrial genome sequencing can not achieve its goal by itself if the genome contains large gene-free region(s) of >35 kb (average insert size of the vector) when Cosmid mtDNA clones are used in sequencing. To cover such regions of the genome, we need to perform sequencing of some additional clones which do not carry any probe genes. In fact, it was necessary for us to sequence a probe gene-free clone, #194, to complete sequencing the wheat mitochondrial genome. The MC molecule obtained successfully integrated all sequenced mtDNAs in it without leaving any pieces out.

Blast search on the sequence homology between the present wheat and previously reported rice and maize mitochondrial genomes (8,10) provided supporting evidence that the present MC molecule represents the wheat mitochondrial genome. The rice and maize mitochondrial genomes were divided into successive 30 kb sections (sizes of the end sections

Table 4. Genes in the wheat mitochondrial genome

Gene ^a	Size (bp)	MC coordinates		Strand ^b	No. of amino acids	Previous accession no. ^c
		From	To			
I. Complex I genes						
<i>nad1a</i>	386	306 345	306 730	+	—	X57968
<i>nad1b</i>	82	17 602	17 683	—	—	X57967
<i>nad1c</i>	192	15 988	16 179	—	—	X57967
<i>nad1d</i>	59	282 283	282 341	—	—	X57966
<i>nad1e</i>	259	43 394	43 652	—	—	X57965
<i>nad1</i>	978	—	—	×	325	—
<i>nad2a</i>	153	182 361	182 513	—	—	Y14433
<i>nad2b</i>	392	181 155	181 546	—	—	Y14433
<i>nad2c</i>	161	210 083	210 243	—	—	Y14434
<i>nad2d</i>	573	207 093	207 665	—	—	Y14434
<i>nad2e</i>	188	205 502	205 689	—	—	Y14434
<i>nad2</i>	1467	—	—	—	488	—
<i>nad3</i>	357	341 099	341 455	—	118	X59153
<i>nad4a</i>	461	135 541	136 001	+	—	X57164
<i>nad4b</i>	515	137 026	137 540	+	—	X57164
<i>nad4c</i>	423	140 981	141 403	+	—	X57164
<i>nad4d</i>	89	143 059	143 147	+	—	X57164
<i>nad4</i>	1488	—	—	+	495	—
<i>nad4L</i>	303	421 629	421 931	—	100	AJ295996
<i>nad5a</i>	231	295 842	296 072	—	—	M74157
<i>nad5b</i>	1216	293 764	294 979	—	—	M74157
<i>nad5c</i>	21	43 010	43 030	—	—	M74158
<i>nad5d</i>	395	111 614	112 008	+	—	M74159
<i>nad5e</i>	150	112 942	113 091	+	—	M74159
<i>nad5</i>	2013	—	—	×	670	—
<i>nad6</i>	744	280 549	281 292	—	247	X62100
<i>nad7a</i>	143	409 840	409 982	+	—	X75036
<i>nad7b</i>	69	410 796	410 864	+	—	X75036
<i>nad7c</i>	467	412 177	412 643	+	—	X75036
<i>nad7d</i>	244	413 642	413 885	+	—	X75036
<i>nad7e</i>	262	415 585	415 846	+	—	X75036
<i>nad7</i>	1185	—	—	+	394	—
<i>nad9</i>	864	211 710	212 573	—	287	X69720
II. Complex III & IV genes						
<i>cob</i>	1197	63 122	64 318	—	398	X02352
<i>cox1</i>	1575	245 285	246 859	+	524	Y00417
<i>cox2a</i>	390	224 812	225 201	—	—	X01108
<i>cox2b</i>	393	223 200	223 592	—	—	X01108
<i>cox2</i>	783	—	—	—	260	—
<i>cox3</i>	798	28 053	28 850	—	265	X15944
III. Complex V genes						
<i>atp1</i>	1530	6832	8361	+	509	X15918
<i>atp4</i>	579	196 584	197 162	—	192	X54311
<i>atp6-1</i>	1161	19 132	20 292	—	386	M24084
<i>atp6-2</i>	1161	85 008	86 168	—	386	M24084
<i>atp8-1</i>	471	338 884	339 354	—	156	X59153
<i>atp8-2</i>	471	233 393	233 863	—	156	X59153
<i>atp9</i>	243	8824	9066	+	80	X15919
IV. Cytochrome c biogenesis genes						
<i>ccmB</i>	621	185 578	186 198	—	206	AF082025
<i>ccmC</i>	723	156 957	157 679	+	240	XY9609
<i>ccmFCa</i>	755	99 514	100 268	+	—	AY500223
<i>ccmFCb</i>	559	101 280	101 838	+	—	AY500223
<i>ccmFC</i>	1314	—	—	+	437	—
<i>ccmFN</i>	1770	50 112	51 881	—	589	X69205
V. Ribosomal protein genes						
<i>rpl2-p</i>	169	283 450	283 618	—	56+1/3	AJ295995
<i>rpl5</i>	570	29 778	30 347	—	189	AJ535507
<i>rpl16</i>	558	90 168	90 725	—	185	New
<i>rps1</i>	525	49 341	49 865	—	174	X69205
<i>rps2</i>	1083	215 091	216 173	+	360	Y13920
<i>rps3a</i>	74	93 925	93 998	—	—	New
<i>rps3b</i>	1612	90 574	92 185	—	—	New
<i>rps3</i>	1686	—	—	—	561	—
<i>rps4</i>	1074	273 586	274 659	—	357	New
<i>rps7</i>	447	379 580	380 026	—	148	X67242

Table 4. continued

Gene ^a	Size (bp)	MC coordinates		Strand ^b	No. of amino acids	Previous accession no. ^c
		From	To			
<i>rps12</i>	378	340 677	341 054	–	125	X59153
<i>rps13</i>	351	18 628	18 978	–	116	Y00520
<i>rps19-p</i>	198	422 830	423 027	–	66	AJ295996
VI. Other protein coding genes						
<i>matR</i>	2037	44 172	46 208	–	678	X57965
<i>mtbB</i>	816	314 992	315 807	+	271	New
VII. rRNA genes						
<i>rrn5-1</i>	122	302 949	303 070	+	–	Z14078
<i>rrn5-2</i>	122	393 494	393 615	–	–	Z14078
<i>rrn5-3</i>	122	56 526	56 647	–	–	Z14078
<i>rrn18-1</i>	1955	300 880	302 834	+	–	Z14078
<i>rrn18-2</i>	1955	393 730	395 684	–	–	Z14078
<i>rrn18-3</i>	1955	56 762	58 716	–	–	Z14078
<i>rrn26-1</i>	3467	371 222	374 688	–	–	Z11889
<i>rrn26-2</i>	3467	259 484	262 950	–	–	Z11889
<i>rrn26-p</i>	422	170 632	171 053	–	–	Z11889
VIII. tRNA genes						
<i>trnA 5'-ex*</i>	38	74 738	74 775	+	–	New
<i>trnA 3'-ex*</i>	35	75 581	75 615	+	–	New
<i>trnA*</i>	73	–	–	+	–	–
<i>trnC*</i>	71	97 420	97 490	+	–	X15119
<i>trnD-1</i>	74	429 341	429 414	–	–	X15379
<i>trnD-2</i>	74	222 080	222 153	–	–	X15379
<i>trnE</i>	72	27 050	27 121	–	–	X14698
<i>trnF*</i>	73	382 956	383 028	–	–	X15118
<i>trnfM-1</i>	74	300 805	300 878	+	–	Z14078
<i>trnfM-2</i>	74	395 686	395 759	–	–	Z14078
<i>trnfM-3</i>	74	58 718	58 791	–	–	Z14078
<i>trnI</i>	74	430 118	430 191	–	–	New
<i>trnK-1</i>	73	270 536	270 608	+	–	X15236
<i>trnK-2</i>	73	442 701	442 773	+	–	X15236
<i>trnK-3</i>	73	178 640	178 712	+	–	X15236
<i>trnM</i>	73	436 154	436 226	–	–	New
<i>trnN*</i>	72	428 634	428 705	–	–	X15379
<i>trnP-1</i>	75	305 095	305 169	+	–	Z14078
<i>trnP-2</i>	75	117 692	117 766	+	–	Z14078
<i>trnQ-1</i>	72	266 806	266 877	+	–	X15140
<i>trnQ-2</i>	72	174 909	174 980	+	–	X15140
<i>trnQ-3</i>	72	193 282	193 353	+	–	X06902
<i>trnS-1</i>	88	341 968	342 055	–	–	X13245
<i>trnS-2</i>	87	408 505	408 591	+	–	X15118
<i>trnS-3*</i>	87	383 444	383 530	–	–	X15118
<i>trnW*</i>	74	445 613	445 686	+	–	X05602
<i>trnY</i>	83	210 880	210 962	–	–	Y14434

^aBoldface, sum of all exons; lower-case letters, exons of a protein-coding gene; hyphenated, copies of the same gene; asterisk: probable chloroplast origin.

^bPlus and minus, coded by the forward and reverse strand; ×, *trans*-spliced gene.

^cNew, gene or exon whose nucleotide sequence is first reported for wheat.

were somewhat different), and sequences homologous to wheat mtDNA were investigated for each section (Table 6). All the sections contained homologous sequences of ~3 kb or larger (up to 15 kb) to wheat mtDNA. MtDNA sequences conserved between wheat and rice, and between wheat and maize were distributed all over the rice and maize genomes, with no large conserved sequence-free regions (larger than 10 kb; detailed data omitted) in the genomes. This fact indicates that rice and maize mitochondrial genomic sequences are well represented in the wheat MC molecule.

The most essential feature of the present gene-based strategy is that it facilitated the recovery of recombinant molecules. Restriction fragment mapping of plant mtDNA shows a multipartite structure of the mitochondrial genome, consisting

of isomeric as well as subgenomic molecules produced by intra-molecular recombination (12,14–17,33). None of the previous works on complete sequencing of flowering plant mitochondrial genomes, by use of either the genome shotgun or physical map-based strategies, has recovered recombinant molecules. This is why recombination events have not been analyzed at the nucleotide sequence level. By virtue of the gene-based strategy, we obtained 10 recombinant clones among 25 examined, determined their nucleotide sequences, and identified repeat sequences responsible for their formation.

Structural features of the wheat mitochondrial genome

The wheat mitochondrial genome was assumed to be a 452 528 bp MC molecule (Figure 1), that was ~92 and 79% the size of the rice and maize mitochondrial genomes, and possessed all the protein-, rRNA- and tRNA-coding genes known to be present in rice and maize (8,10). These facts indicate that wheat has the most compact mitochondrial genome among the three cereals.

Multicopy mitochondrial genes were compared between wheat, rice and maize (Table 7). Gene amplification in general was species-specific. All of the multicopy wheat genes were located in the repeated sequences (Table 2). With the exceptions of *atp8* and *trnQ*, multicopies of all the wheat genes had identical nucleotide sequences. As for *trnQ*, two copies were identical, whereas the third copy differed from them by a single nucleotide. These facts suggest their recent amplification, comparing with the divergence time of three cereals. One alternative possibility is copy correction through homologous recombination, which is known to occur in the case of chloroplast IRs (34).

To account for the observed species-specific gene amplification, a mechanistic model can be proposed. Recombination between the same repeat sequences in two subgenomic molecules produced by recombination between different repeat pairs will give rise to an aberrant MC molecule having a duplicate segment. Figure 6 illustrates an example, using a simplified MC molecule, in which only three copies (Copy-1, -2 and -3) of the *trnfM-rrn18-rrn5* cluster and two repeat pairs, R5 and R6, are shown. Recombination between the R6 sequences in two subgenomic molecules, II and III, which are produced in pathways [B] and [C], gives a new MC molecule with an extra copy of the *trnfM-rrn18-rrn5* cluster and R5 repeat together with their flanking regions. The size of the duplication corresponds to the sum of two segments, one between the recombination breakpoints in Copy-2 and one R5 copy, and the other between those in Copy-3 and the other R5 copy.

Search for transposable element sequences in the wheat mitochondrial genome revealed presence of five sequences, three of which were different partial sequences of the wheat Sabrina retrotransposon, and two others were a part of a rice Tos-14 retrotransposon and wheat Tar1 retrotransposon. Total size of the five sequences was 805 bp, being ~0.2% of the mitochondrial genome. Comparable figures for rice and maize were 20 sequences (total size 7003 bp, 14.3% of the genome) and 4 sequences (total size 641 bp, 0.1% of the genome), respectively (8,10). In this respect, wheat mitochondrial genome is similar to maize than to rice mitochondrial genome.

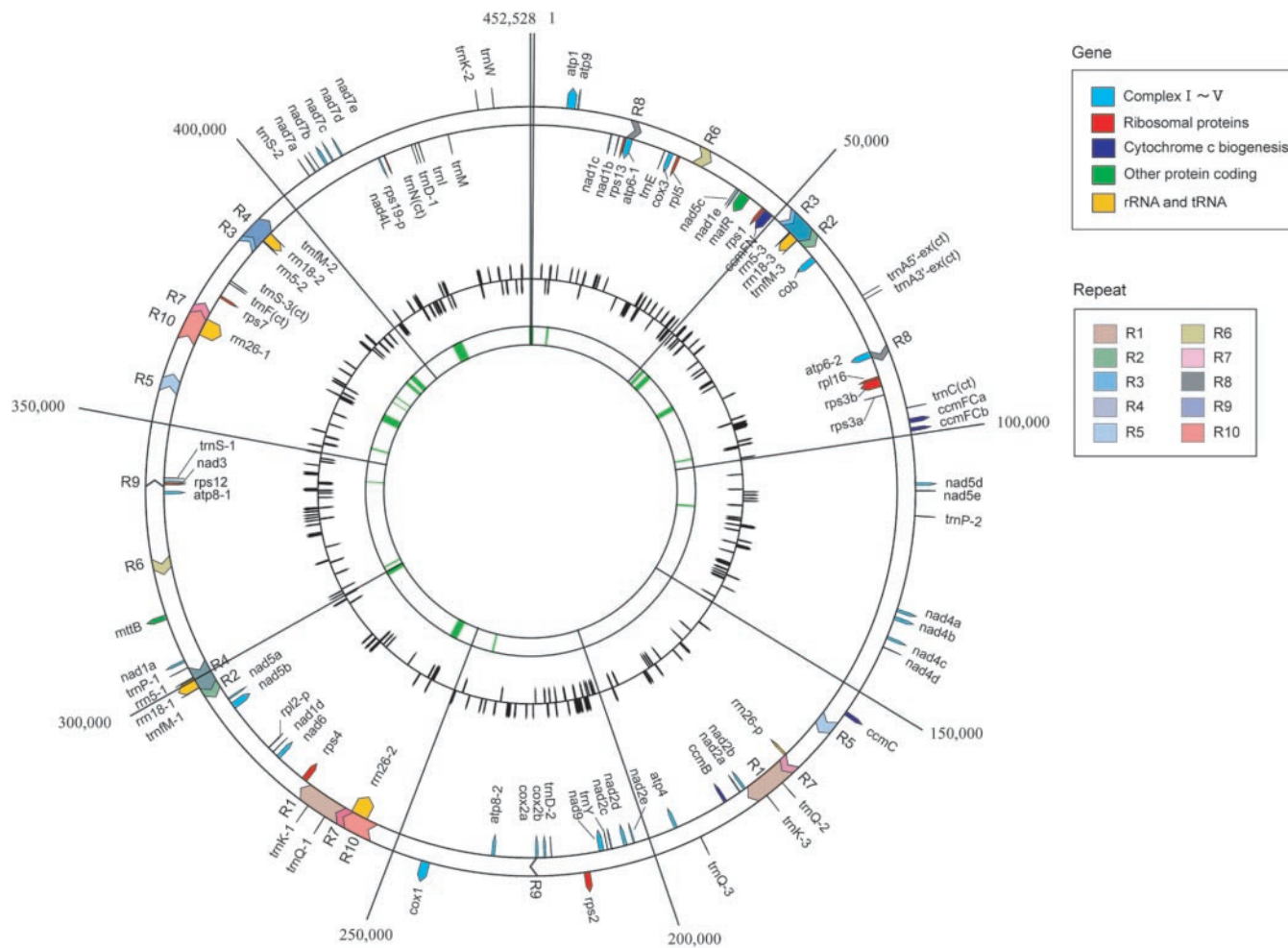


Figure 5. Genetic map of the wheat mitochondrial genome showing the location of all the genes and their exons in the outer-most circle, of ORFs larger than 300 bp in the central circle, and of chloroplast-derived DNA segments in the inner-most circle. The broad, outer-most circle represents the MC molecule, in which the nine repeat pairs, R1–R9, that mediate production of all the recombinant clones and an additional 7035 bp repeat pair, R10, are shown. Genes and exons coded by the forward and reverse DNA strands are shown outside and inside the MC molecule, respectively.

It is important to know what kinds of sequences were involved in the observed mitochondrial genome differentiation. For this purpose, the MC coordinates of all unique wheat sequences larger than 100 bp, comparing with both the rice and maize mtDNA sequences were enumerated (Supplementary Table 3). In total, 227 unique sequences distributed throughout the genome were identified. Comparison between their positions and those of all mitochondrial genes in the genome indicated that almost all unique sequences corresponded to intergenic spaces. The exceptions were nine sequences carrying partial sequence of a gene. Of those, six sequences carried 3–97 bp of the highly variable 3' end of the sense strand of *cob*, *nad6*, *rpl2-p*, *rrn5-1*, *rrn5-2* and *rrn5-3*. Two sequences contained a 324 bp segment of *atp6-1* and *-2*, that is located in the 3'-terminal region of these genes. The last sequence carried a 28 bp 5' end of *nad9*, that is variable among the three cereals. These facts taken together demonstrate that the mtDNA sequences diversified in the three cereals are mostly redundant DNAs.

In a summary, the wheat mtDNA sequences were partitioned into six categories, genic (including introns), ORF, repetitive, chloroplast-derived, retro-element and unique

sequences (Table 8). This partition was not orthogonal, because some sequences were enumerated in more than one category. Sizes of the genic, ORF, repetitive, chloroplast-derived and unique sequences were obtained from the data presented in Table 4, Supplementary Table 2, Table 2, Table 5 and Supplementary Table 3, respectively.

Structural dynamics of the mitochondrial genome in ontogeny

Arrieta-Montiel *et al.* (35) reported on the structural dynamics of the common bean mitochondrial genome, which was revealed by studying a single mtDNA segment carrying the *cms*-associated *pvs-orf239* sequence. Using the gene-based strategy, we isolated 10 recombinant mtDNA molecules, and determined the repeat sequences responsible for their production. Many other repeat pairs also were characterized (Table 2 and Figure 3), which are potential sites for additional recombination. Based on the entire wheat mitochondrial genome sequence (DNA Database accession no. AP008982) and the map positions of all repeat pairs larger than 100 bp (Table 2), we may prepare DNA primers for the sequences

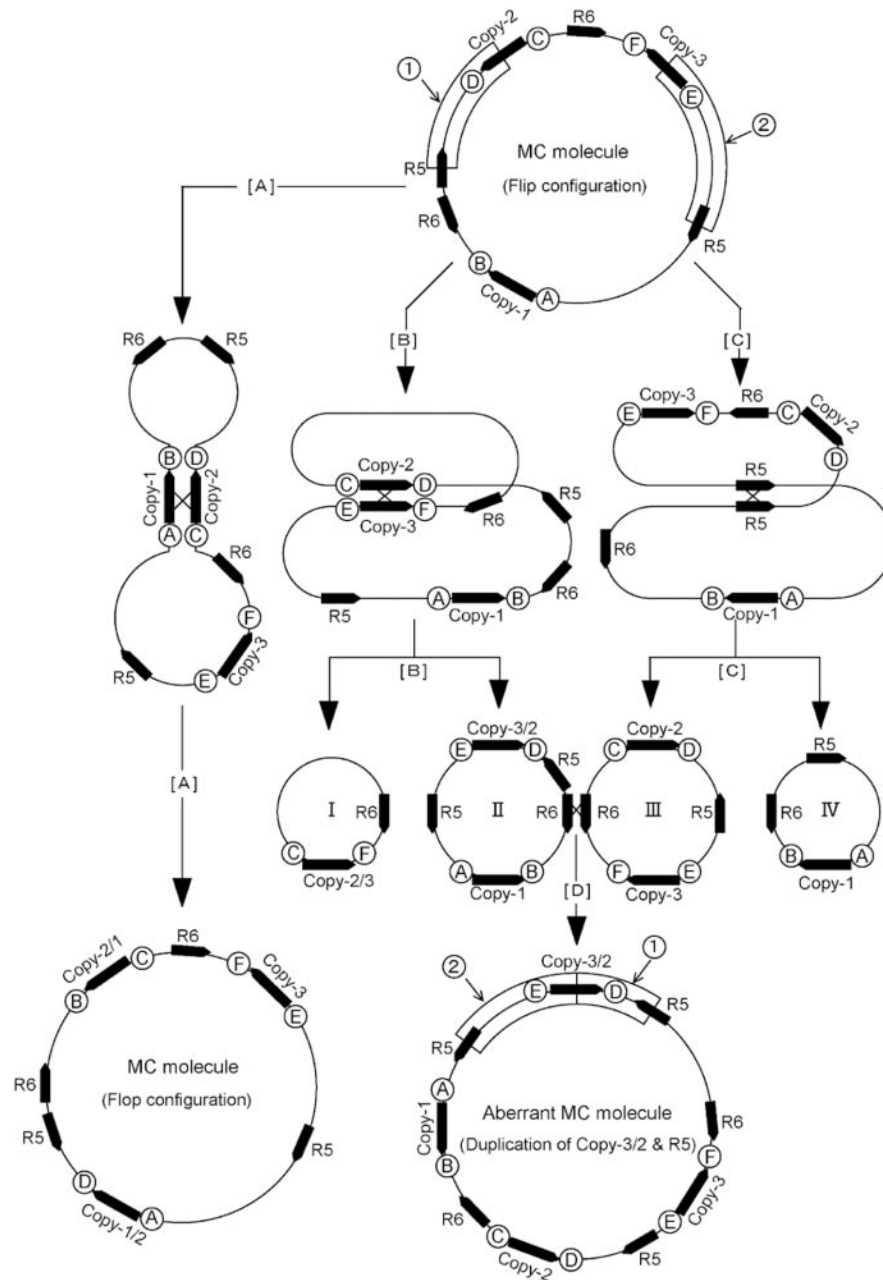


Figure 6. Production of various molecular forms from the MC molecule by intra-molecular recombination between different repeat pairs. Copy-1, -2 and -3 are three copies of the *trnM-rrn18-rrn5* gene cluster. Copy-2 and -3 are inverted relative to Copy-1. R5 and R6 represent a DR and an IR pair, respectively. A/B, C/D and E/F are PCR primer pairs to mark the 5'- and 3'-flanking regions of Copy-1, -2 and -3, respectively. [A]: production of an isomer (flop form) of the MC molecule (flip form) by recombination between an IR pair, Copy-1 and -2. [B] and [C]: production of two complementary subgenomic molecules by recombination between a DR pair, Copy-2 and -3, and two R5 copies, respectively. [D]: production of an aberrant MC molecule having extra copies of the three-gene cluster (Copy-3/2) and R5 repeat by recombination between R6 repeats in two subgenomic molecules, II and III.

flanking both ends of those repeats. Their use in long-range PCR will allow efficient screening of recombinant molecules produced by recombination between the marked repeat pairs and quantification of isomeric as well as subgenomic molecules, as proved by Sugiyama *et al.* (11) in tobacco. They also demonstrated that long-range PCR works for a distance as long as 23 kb between two primers, which is sufficient to cover all repeats present in the wheat mitochondrial genome (Table 2). The same method may also facilitate finding the difference in recombinational activity among various repeat pairs as well as

the equality or inequality of the reciprocal recombination products.

The methodological details for such studies are as follows: recombination between an IR pair will produce an isomer (flop form) of the MC molecule (flip form; Figure 6, pathway [A]). This event is detected by long-range PCR using four primer pairs, A/B, C/D, A/D and C/B. If either the A/B or C/D primer pair gives an amplified product in PCR, the template clone is regarded as the original MC molecule, whereas if either the A/D or C/B pair gives an amplified product, the template clone

Table 5. Wheat mtDNA sequences showing homology to ctDNA sequences

MtDNA sequence MC coordinates ^a	Size (bp)	Mt gene located ^b	Homologous ctDNA sequence		Ct gene located ^b	Nucleotide sequence homology (%) ^c
			CtDNA coordinates	Size (bp)		
	995–1157	No	62 057–62 218	162	<i>psbE</i> *	92.0
O	7309–7889	<i>atp1</i> *	35 143–35 696	554	<i>atpA</i> * (copy-3)	?
	53 841–54 950	No	36 037–34 918	1120	<i>atpA</i> # (copy-4)	97.2
	55 179–55 205	No	48 098–48 072	27	<i>trnF</i> *	100.0
O	56 769–58 700	<i>rrn18-3</i>	92 532–91 061	1472	<i>rrn16</i> (copy-1)	?
	74 171–76 003	No	93 226–95 059	1834	<i>trn1 3'-ex, trnA, rrn23</i> *	99.8
	79 301–79 405	No	63 441–63 336	106	<i>petL</i>	95.3
	97 417–97 542	<i>trnC</i>	18 754–18 628	127	<i>trnC</i>	91.3
	97 779–97 861	No	18 367–18 282	86	No	81.4
	98 764–99 133	No	77 395–77 024	372	<i>rpl14</i> #	81.6
	99 254–99 373	No	76 833–76 715	119	<i>rps8</i> *	86.7
	117 697–117 760	<i>trnP-2</i> #	64 131–64 069	63	<i>trnP</i> #	82.8
	119 524–120 020	No	35 542–36 072	531	<i>atpA</i> * (copy-5)	?
	146 989–147 055	No	109 596–109 526	71	<i>ndhG</i> *	84.5
	154 459–154 512	No	84 016–83 963	54	No	98.1
	157 714–157 745	<i>trnI-p</i>	82 976–82 945	32	<i>trnI</i> *	96.9
	162 458–162 515	No	34 309–34 366	58	<i>atpF 3'-ex</i> *	96.6
O	170 827–170 895	<i>rrn26-p</i> *	95 063–94 995	69	<i>rrn23</i> * (copy-3)	79.7
	174 918–174 971	<i>trnQ-2</i> #	6749–6696	54	<i>trnQ</i> *	83.3
	242 709–243 106	No	110 973–111 387	415	<i>ndhA 3'-ex</i> #	?
	249 838–249 890	No	51 175–51 225	51	<i>trnV 3'-ex</i> *	88.7
O	259 648–262 792	<i>rrn26-2</i>	97 615–94 995	2 621	<i>rrn23</i> (copy-1)	?
	266 815–266 868	<i>trnQ-1</i> #	6749–6696	54	<i>trnQ</i> *	83.3
	294 426–294 490	<i>nad5b</i> *	102 531–102 595	65	<i>ndhF</i> *	81.5
O	300 896–302 827	<i>rrn18-1</i>	91 061–92 532	1472	<i>rrn16</i> (copy-2)	?
	304 391–304 417	No	48 072–48 098	27	<i>trnF</i> *	100.0
	304 646–304 973	No	34 918–35 245	328	<i>atpA</i> * (copy-1)	95.7
	305 100–305 163	<i>trnP-1</i> #	64 131–64 069	63	<i>trnP</i> #	82.8
	316 034–316 102	No	21 268–21 343	76	<i>rpoB</i> *	86.8
	324 416–324 550	No	75 643–75 509	135	<i>rps11</i> *	94.1
	343 407–343 573	No	111 572–111 403	170	No	88.4
	349 029–349 064	No	33 284–33 319	36	<i>atpF 5'-ex</i> *	91.7
	358 076–358 521	No	41 098–40 653	446	<i>psaA</i> *	99.8
O	371 386–374 530	<i>rrn26-1</i>	97 615–94 995	2621	<i>rrn23</i> (copy-2)	?
	378 941–379 028	No	44 149–44 062	88	No	94.3
	379 044–379 150	No	44 039–43 932	108	No	88.0
	380 703–380 885	No	68 074–68 256	183	<i>clpP</i> *	95.6
	382 951–383 089	<i>trnF</i>	48 133–47 995	139	<i>trnF</i>	94.2
	383 203–383 278	No	47 851–47 776	76	No	82.9
	383 342–383 422	No	47 702–47 621	82	<i>trnL 3'-ex</i>	91.5
	383 409–383 603	<i>trnS</i>	45 160–44 967	194	<i>trnS</i>	90.8
	388 154–388 182	No	34 485–34 513	29	<i>atpF 3'-ex</i>	100.0
	390 809–391 918	No	36 037–34 918	1120	<i>atpA</i> * (copy-2)	97.2
	392 147–392 173	No	48 098–48 072	27	<i>trnF</i> *	100.0
O	393 737–395 668	<i>rrn18-2</i>	92 532–91 061	1472	<i>rrn16</i> (copy-3)	?
	400 525–400 556	No	49 041–49 072	32	<i>ndhJ</i> *	96.9
	408 510–408 585	<i>trnS-2</i> #	11 655–11 579	77	<i>trnS</i> #	80.5
	417 240–421 478	No	89 049–84 780	4270	<i>ndhB, rps7, rps12 ex-2, -3</i>	?
	421 513–421 558	No	84 756–84 711	46	No	97.8
	428 633–428 718	<i>trnN</i>	98 896–98 811	86	<i>trnN</i>	98.8
O	436 153–436 225	<i>trnM</i>	52 107–52 035	73	<i>trnM</i>	94.5
	445 372–445 416	No	64 290–64 246	45	No	88.9
	445 455–445 488	No	64 084–64 051	34	<i>trnP</i> *	100.0
	445 609–445 690	<i>trnW</i>	63 927–63 846	82	<i>trnW</i>	96.3
	452 168–452 356	No	61 844–62 032	189	<i>psbF</i> *, <i>psbE</i> *	88.4
	Total	26 264	—	—	—	—
	Total excluding O-marked sequences: 13 455 bp					

CtDNA sequences present in one IR_B, IR_A, are shown, omitting those in the other copy (IR_A), because of the same gene set present in two copies. Total size of 26 264 bp is 14 bp smaller than the sum of all the segments because a 14 bp sequence overlaps between two segments of the mtDNA coordinates 383 342–383 422 and 383 409–383 603.

^aO: native mtDNA sequence.

^b# and asterisk: genes, of which a large portion and only a small portion are located in the respective DNA sequences. Gene in boldface: complete or nearly complete gene sequence included in the respective DNA sequences. No: no gene present.

^c?: undetermined because of segmental differentiation of the sequence within the gene.

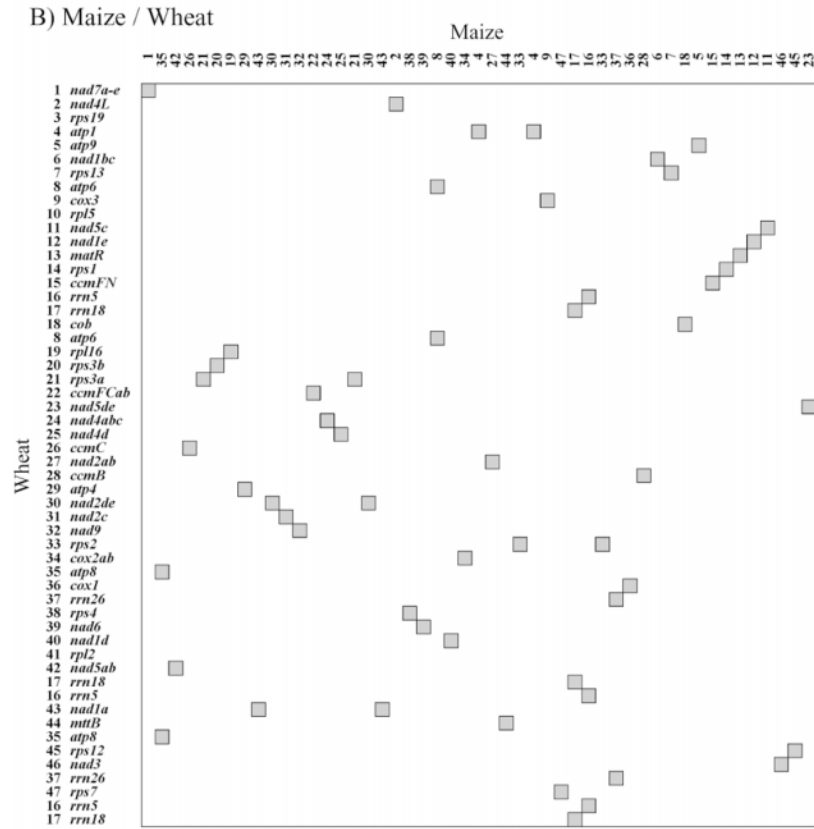
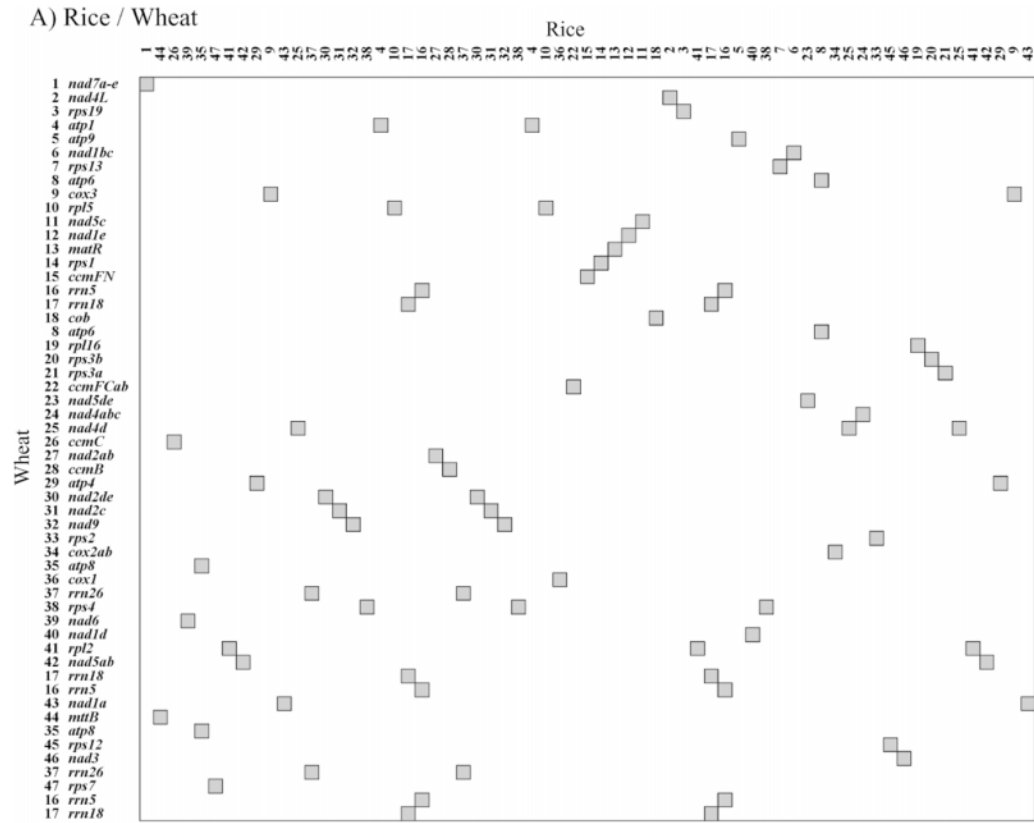


Figure 7. Correlation of gene order between the mitochondrial gene maps of wheat and rice (A) and wheat and maize (B). All the protein- and rRNA-coding genes and the former's *trans*-spliced exons are arranged from top to bottom for wheat, and from left to right for rice and maize, based on their order in the respective gene maps. Genes of rice and maize are indicated by code numbers given to the corresponding wheat genes in the left margin of figures. Duplicate genes carry the same number.

Table 6. Rice and maize mtDNA sequences homologous to wheat mtDNA in different sections of the genome

Rice mtDNA			Maize mtDNA		
Genome section ^a	Section size (bp)	Homologous section (bp) ^b	Genome section ^a	Section size (bp)	Homologous section (bp) ^b
1	30000	12778	1	30000	6518
2	30000	7676	2	30000	12814
3	30000	15569	3	30000	12460
4	30000	13001	4	30000	12185
5	30000	5465	5	30000	10010
6	30000	10057	6	30000	6241
7	30000	10989	7	30000	7782
8	30000	12804	8	30000	4308
9	30000	8527	9	30000	14730
10	30000	6881	10	30000	9739
11	30000	14898	11	30000	11500
12	30000	10985	12	30000	8014
13	30000	13182	13	30000	4667
14	30000	11009	14	30000	5216
15	30000	11748	15	30000	7642
16	40520	8112	16	30000	2842
Total	490520	173691	17	30000	5079
			18	30000	4233
			19	29630	6776
			Total	569630	152756

^aRice and maize mitochondrial genomes are divided into successive 30 kb sections, the last one being the remaining part of the respective genome.

^bTotal size of wheat mtDNA sequences of larger than 30 bp which are homologous to the rice or maize mtDNA sequences.

Table 7. Copy numbers of mitochondrial genes that differ in number in wheat, rice and maize: gene fragments, pseudogenes and chloroplast-derived genes are excluded

Gene	Wheat	Rice ^a	Maize ^a
(1) Protein-coding gene			
<i>atp1</i>	1	2	2
<i>atp4</i>	1	2	1
<i>atp6</i>	2	1	1
<i>atp8</i>	2	1	1
<i>cox3</i>	1	2	1
<i>nad1a</i>	1	2	2
<i>nad2c</i>	1	2	1
<i>nad2d, e</i>	1	2	2
<i>nad4d</i>	1	3	1
<i>nad5a,b</i>	1	2	1
<i>nad9</i>	1	2	1
<i>rpl2</i>	0	3	0
<i>rpl5</i>	1	2	0
<i>rps2</i>	1	1	2
<i>rps3a</i>	1	1	2
<i>rps7</i>	1	1	1
(2) RNA gene			
<i>rrn5</i>	3	2	1
<i>rrn18</i>	3	2	1
<i>rrn26</i>	2	2	1
<i>trnD</i>	2	1	2
<i>trnE</i>	1	1	2
<i>trnM</i>	3	1	1
<i>trnI</i>	1	1	2
<i>trnK</i>	3	1	1
<i>trnM</i>	1	1	0
<i>trnN</i>	0	1	1
<i>trnP</i>	2	1	2
<i>trnQ</i>	3	1	1

^aAfter Notsu *et al.* (8) for rice and Clifton *et al.* (10) for maize.

is regarded as the flop configuration of MC, so far as the marked IR is concerned. The ratio of the latter to the former clones in number gives the molar ratio of the recombinant to the non-recombinant clones. Similarly, recombination between a DR pair will produce two subgenomic molecules (Figure 6, pathway [B]), whose production is detected by successful DNA amplification by use of the C/F or E/D primer pair. Quantification of the subgenomic molecules over the non-recombinants is achieved in the same way as described above. Such studies targeted to different repeat pairs enumerated in Table 2, using Cosmid clones of wheat mtDNA extracted from different organs or different ages of the plant as the template, will disclose structural dynamics of the mitochondrial genome in plant development.

Evolutionary change in the mitochondrial genome structures of cereals

The chloroplast genomes of rice, maize and wheat have identical gene arrangements (36,37,21), evidence of the structure's evolutionary stability. In contrast, the mitochondrial genome structure differs markedly in the three cereals (Figure 7) although the kinds of genes present essentially are the same [(8,10), present findings]. We showed that a variety of mtDNA molecules are produced in somatic tissues by intra-molecular recombination mediated by different repeat pairs. The structural differences of several mitochondrial genes in wheat and rice are suspected to be caused by short repeat pairs (data to be published elsewhere). We postulate that the same mechanism operates in germ cell lines, creating structural diversity in the mitochondrial genomes of different plant phylogenies.

Another possible factor for high phylogenetic variability of the mitochondrial genome, compared with the chloroplast genome, is high DNA redundancy in the former than in the latter genome. The ratio of the genic sequences, including all exons and *cis*-introns, and excluding the sequences of chloroplast origin and pseudogenes, to the total mitochondrial genome size is 18.0% for rice (8), 11.7% for maize (10) and 15.9% for wheat (Table 8). Comparable values for the chloroplast genome are 58.8% for rice (36) and 60.4% for wheat (21), indicative of the presence of a much larger amount of redundant DNAs in the mitochondrial than in the chloroplast genome.

The MC molecule may represent the intact wheat mitochondrial genome

All previous works on complete sequencing of flowering plant mitochondrial genomes are based upon the MC molecule hypothesis (6–11). Because of the multipartite structure of the genome and the lack of direct electron-microscopic evidence, however, the existence of the MC molecule is still a matter of debate (11–13). After Andre *et al.* (12), we suspected reality of the MC molecule in wheat and upon this suspicion we adopted the gene-based sequencing strategy. It turns out, however, that analysis of the 10 recombinant clones obtained has given support to the existence of the MC molecule.

If we consider the MC molecule to be a flip configuration of the genome, then recombination between either of the three IR pairs (Table 2) will produce its flop (= isomeric) molecule, as shown in the pathway [A] of Figure 6, whereas recombination between either of the DR pairs produces two complementary,

Table 8. Classification of wheat mtDNA sequences into different categories

Category	No. sequences	Total size (bp)	Proportion (%)	Source
Entire genome	—	452 528	100.0	Figure 1
Genic, including introns	76	71 848	15.9	Table 4
ORFs (larger than 300 bp), including those in repeats	179	75 465	16.7	Supplementary Table 2
Repetitive (repeat sequences larger than 100 bp)	26	68 960	15.2	Table 2
Chloroplast origin	47	13 455	3.0	Table 5
Retro elements	5	805	0.2	Text
Unique (larger than 100 bp), comparing with rice and maize	277	257 762	57.0	Supplementary Table 3

Classification of the sequences is not orthogonal, because some sequences are enumerated in more than one category.

Table 9. Expected and actual products of recombination when the mitochondrial genome has alternative configurations

Alternative genome configuration	Case Configuration	Type of affected repeat ^b			Expected product of recombination at the affected repeats ^c	Actual recombinant produced (clone obtained) ^d	
		Recombination site ^a	Repeat Before recombination	Repeat After recombination			
1	Flop configuration at R2	R2 (IR)	R5	DR	IR	Double-flop configuration at R2 and R5	Subgenomic molecule (#27)
2	Flop configuration at R2	R2 (IR)	R7	DR	IR	Double-flop configuration at R2 and R7	Subgenomic molecule (#24*)
3	Flop configuration at R4	R4 (IR)	R6	IR	DR	Subgenomic molecules	Flop configuration at R6 (#224*)
4	Flop configuration at R6	R6 (IR)	R4	IR	DR	Subgenomic molecules	Flop configuration at R4 (#1)
5	Subgenomic molecules	R1 (DR)	R9	DR	Separated	Present MC configuration	Subgenomic molecule (#31)
6	Subgenomic molecules	R3 (DR)	R3	DR	Separated	Present MC configuration	Subgenomic molecule (#75, #96, #162)
7	Subgenomic molecules	R7 (DR)	R4	IR	Separated	Double-flop configuration at R7 and R4	Flop configuration at R4 (#1)
8	Subgenomic molecules	R5 (DR)	R2	IR	Separated	Double flop configuration at R2 and R5	Flop config. at R2 (#224*)

^aRepeat mediating recombination that results in respective genome configuration. DR and IR, direct and inverted repeats.

^bOne of several repeat pairs, of which type is changed by altered genome configuration. Separated, two repeat sequences are separated to different subgenomic molecules.

^cThe present MC molecule is considered as the flip configuration of the genome. Single- or double-flop configuration is caused by single or double recombination at the indicated repeats.

^dAsterisk: double recombinant clone.

subgenomic molecules (pathways [B] and [C] in Figure 6), where 'complementary' means that a complete gene set is shared by two or more molecules (15). The origin of eight recombinant clones can be explained by a single recombination event, while the remaining two double-recombination events occurred in the MC molecule. However, if the genome were in any other configuration, most of the recombinant clones obtained could not have been produced by simple recombination events (Table 9). Consider the following: if the genome existed in the flop configuration of MC (Table 9, Case 1–4), then recombination between any pair of the present DRs should produce double-flop configurations of the genome (Case 1 and 2), and recombination between IRs should produce two subgenomic molecules (Case 3 and 4). Similarly, if the genome consisted of two subgenomic molecules (Case 5–8), recombination between the repeat sequences in two separated molecules should produce the MC molecule (Case 5 and 6), or its double-flop configuration (Case 7 and 8). In all eight postulated cases, the expected recombination products do not match the ones we actually obtained. This fact supports the hypothesis that the MC molecule serves as the basic wheat mitochondrial genome structure.

A possible alternative is that the wheat mitochondrial genome contains all kinds of isomeric as well as subgenomic molecules (13). Lonsdale *et al.* (16) and Fauron *et al.* (17) showed that 5–14 subgenomic molecules are produced from the MC molecule of sugar beet and maize by intra-molecular recombination. In our study we prepared wheat mtDNA from 2-week-old seedlings. Now, if a seedling consists of

~10⁶ cells, it means that 19 successive cell divisions, on the average, occurred before DNA extraction. We do not know how many replication origins exist in the wheat mitochondrial MC molecule. An electron-microscopic study of mtDNA replication in *Chenopodium* indicates only a few, if not just one, origins in its mtDNA (38). Considering this fact, together with information on the single replication origin of bacterial chromosomes, it is hard to believe that all kinds of subgenomic molecules have replication origins necessary for their maintenance through many cell cycles. This is further support for the presence of the MC molecule.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank L. Bonen for her generous gift of the mitochondrial gene probes used in our clone selection, C.-H. Guo and T. Ochiai for technical help, Jeffrey D. Palmer for critical comments on our manuscript, and Val Woodward for valuable advice on improving the manuscript regarding its context and language. This work was supported by a Grant-in-Aid (No. 12309008, 2000–2002) from the Japan Society for the Promotion of Sciences, and research grants from the Iijima Memorial Foundation for the Promotion of Food Science and Technology (2003 and 2004) and the Nisshin Seifun Foundation (2001 and 2004). Funding to pay the Open Access

publication charges for this article was provided by the Graduate School of Agriculture, Kyoto University, Japan.

Conflict of interest statement. None declared.

REFERENCES

- Dewey, R.E., Timothy, D.H. and Levings, C.S., III (1987) A mitochondrial protein associated with cytoplasmic male sterility in the T cytoplasm of maize. *Proc. Natl Acad. Sci. USA*, **84**, 5374–5378.
- Miller, R.J. and Koeppe, D.E. (1971) Southern corn leaf blight: susceptible and resistant mitochondria. *Science*, **173**, 67–69.
- Siculella, L. and Palmer, J.D. (1988) Physical and gene organization of mitochondrial DNA in fertile and male sterile sunflower CMS-associated alterations in structure and transcription of the *atpA* gene. *Nucleic Acids Res.*, **16**, 3787–3799.
- Cooper, G.M. (2000) *The Cell—A Molecular Approach*, 2nd edn. Sinauer Assoc., Inc., Sunderland.
- Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T. *et al.* (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA, a primitive form of plant mitochondrial genome. *J. Mol. Biol.*, **223**, 1–7.
- Unsel, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366 924 nucleotides. *Nature Genet.*, **15**, 57–62.
- 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). *Nucleic Acids Res.*, **28**, 2571–2576.
- 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.
- Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.*, **31**, 5907–5916.
- Clifton, S.W., Minx, P., Fauron, C.M., Gibson, M., Allen, J.O., Sun, H., Thompson, M., Barbazuk, W.B., Kanuganti, S., Tayloe, C. *et al.* (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol.*, **136**, 3486–3503.
- 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. Genet. Genomics*, **272**, 603–615.
- Andre, C., Levy, A. and Walbot, V. (1992) Sam11 repeated sequences and the structure of plant mitochondrial genomes. *Trends Genet.*, **8**, 128–132.
- Bendich, A.J. (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr. Genet.*, **24**, 279–290.
- Palmer, J.D. and Shields, C.R. (1984) Tripartite structure of the *Brassica campestris* mitochondrial genome. *Nature*, **307**, 437–440.
- Quetier, F., Lejeune, B., Delorme, S., Falconet, D. and Jubier, M.F. (1985) Molecular form and function of the wheat mitochondrial genome. In Van Vloten-Doting, L., Groot, G.S.P. and Hall, T.C. (eds), *Molecular Form and Function of the Plant Genome*. Plenum Press, NY, pp. 413–420.
- Lonsdale, D.M., Brear, T., Hodge, T.P., Melville, S.E. and Rottmann, W.H. (1988) The plant mitochondrial genome: homologous recombination as a mechanism for generating heterogeneity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **319**, 149–163.
- Fauron, C., Casper, M., Gao, Y. and Moore, B. (1995) The maize mitochondrial genome: dynamic, yet functional. *Trends Genet.*, **11**, 228–235.
- Ogihara, Y. and Tsunewaki, K. (1988) Diversity and evolution of chloroplast DNA in *Triticum* and *Aegilops* as revealed by restriction fragment analysis. *Theor. Appl. Genet.*, **76**, 321–332.
- Wang, G.-Z., Matsuoka, Y. and Tsunewaki, K. (2000) Evolutionary features of chondriome divergence in *Triticum* (wheat) and *Aegilops* shown by RFLP analysis of mitochondrial DNAs. *Theor. Appl. Genet.*, **100**, 221–231.
- Tsunewaki, K. (1996) Plasmon analysis as the counterpart of genome analysis. In Jauhar, P.P. (ed.), *Methods of Genome Analysis in Plants*. CRC Press, NY, pp. 271–299.
- Ogihara, Y., Isono, K., Kojima, T., Endo, A., Hanaoka, M., Shiina, T., Terachi, T., Utsugi, S., Murata, M., Mori, N. *et al.* (2000) Chinese Spring wheat (*Triticum aestivum* L.) chloroplast genome: complete sequence and contig clones. *Plant Mol. Biol. Rep.*, **18**, 243–253.
- Bonen, L. and Gray, M.W. (1980) Organization and expression of the mitochondrial genome of plants. I. The genes for wheat mitochondrial ribosomal and transfer RNA: evidence for an unusual arrangement. *Nucleic Acids Res.*, **8**, 319–335.
- Kolodner, R. and Tewari, K.K. (1975) The molecular size and conformation of the chloroplast DNA from higher plants. *Biochim. Biophys. Acta.*, **402**, 372–390.
- Bonen, L. (1995) The wheat mitochondrial genome. In Levings, C.S. III and Vasil, I.K. (eds), *The Molecular Biology of Plant Mitochondria*. Kluwer Academic Publ., Netherlands, pp. 345–364.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
- Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **25**, 955–964.
- Bonen, L. and Bird, S. (1988) Sequence analysis of the wheat mitochondrial *atp6* gene reveals a fused upstream reading frame and markedly divergent N-termini among plant ATP6 proteins. *Gene*, **73**, 47–56.
- Stern, D.B. and Palmer, J.D. (1984) Recombination sequences in plant mitochondrial genomes: diversity and homologies to known mitochondrial genes. *Nucleic Acids Res.*, **12**, 6141–6157.
- Joyce, P.B.M., Spencer, D.F., Bonen, L. and Gray, M.W. (1988) Genes for tRNA-Asp, tRNA-Pro, tRNA-Tyr and two tRNAs-Ser in wheat mitochondrial DNA. *Plant Mol. Biol.*, **10**, 251–262.
- Joyce, P.B.M. and Gray, M.W. (1989) Chloroplast-like transfer RNA genes expressed in wheat mitochondria. *Nucleic Acids Res.*, **17**, 5461–5476.
- Falconet, D., Lejeune, B., Quetier, F. and Gray, M.W. (1984) Evidence for homologous recombination between repeated sequences containing 18S and 5S ribosomal RNA genes in wheat mitochondrial DNA. *EMBO J.*, **3**, 297–302.
- Lejeune, B., Delorme, S., Delcher, E. and Quetier, E. (1987) Recombination in wheat mitochondrial DNA: occurrence of nine different genomic contexts for the 18S–5S genes. *Plant Physiol. Biochem.*, **25**, 227–233.
- Fauron, C. and Havlik, M. (1989) The maize mitochondrial genome of the normal type and the cytoplasmic male sterile type T have very different organization. *Curr. Genet.*, **15**, 149–154.
- Goulding, S.E., Olmstead, R.G., Morden, C.W. and Wolfe, K.H. (1996) Ebb and flow of the chloroplast inverted repeat. *Mol. Gen. Genet.*, **252**, 195–206.
- Arrieta-Montiel, M., Lyznik, A., Woloszynska, M., Janska, H., Tohme, J. and Mackenzie, S. (2001) Tracing evolutionary and developmental implications of mitochondrial stoichiometric shifting in the common bean. *Genetics*, **158**, 851–864.
- Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y. *et al.* (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.*, **217**, 185–194.
- Maier, R.M., Neckermann, K., Igloi, G.L. and Kossel, H. (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.*, **251**, 614–628.
- Backert, S., Dorfel, P., Lurz, R. and Borner, T. (1996) Rolling-circle replication of mitochondrial DNA in the higher plant *Chenopodium album* (L.). *Mol. Cell. Biol.*, **16**, 6285–6294.