Plant Molecular Biology Reporter **18:** 243–253, 2000

Genetic Resources

Chinese Spring Wheat (*Triticum aestivum* L.) Chloroplast Genome: Complete Sequence and Contig Clones

YASUNARI OGIHARA[1,*], KAZUHIRO ISONO[2], TOSHIO KOJIMA[1], AKIRA ENDO[1], MITSUMASA HANAOKA[3], TAKASHI SHIINA[3], TORU TERACHI[4], SHIGEKO UTSUGI[5], MINORU MURATA[5], NAOKI MORI[6], SHIGEO TAKUMI[6], KAZUHO IKEO[7], TAKASHI GOJOBORI[7], RIKA MURAI[8], KOJI MURAI[8], YOSHIHIRO MATSUOKA[8], YUKARI OHNISHI[8], HIKARU TAJIRI[8] and KOICHIRO TSUNEWAKI[8]
[1]*Kihara Institute for Biological Research and Graduate School of Integrated Science, Yokohama City University, Yokohama 244-0813, Japan;* [2]*PE Biosystems Japan Ltd., Chiba 279-0011, Japan;* [3]*Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-8501, Japan;* [4]*Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, Kyoto 603-8047, Japan;* [5]*Research Institute for Bioresources, Okayama University, Kurashiki 710-0046, Japan;* [6]*Laboratory of Plant Genetics, Department of Biological and Environmental Science, Faculty of Agriculture, Kobe University, Kobe 657-8501, Japan;* [7]*Center for Information Biology, National Institute of Genetics, Mishima 411-8540, Japan;* [8]*Department of Bioscience, Fukui Prefectural University, Fukui 910-1195, Japan*

Abstract. Libraries of plasmid clones covering the entire chloroplast (cp) genome of the common wheat, *Triticum aestivum* cv. Chinese Spring were constructed and assembled into contig-clones. From these, we obtained the complete nucleotide sequence of wheat chloroplast DNA – a 134,540 bp circular DNA (DDBJ accession no. AB042240) containing four species of ribosomal RNA, 30 genes for 20 species of transfer RNA, and 71 protein coding genes. Additionally, we detected five unidentified open reading frames conserved among grasses. Plasmid clones are available on request.

Key Words: Chinese Spring wheat, chloroplast DNA, clone bank, complete nucleotide sequence, plastome evolution

Introduction

The plastome is a circular DNA molecule containing long inverted repeats ranging from 10 to 85 kbp depending upon the species (Palmer, 1991). These occur at two locations that divide the genome into small and large single copy regions. It is well known that cpDNA genome structure and gene content are highly conserved. The complete cpDNA sequence is now available for tobacco (Shinozaki et

---

[*] Author for correspondence. Present address: Kihara Institute for Biological Research, Yokohama City University, Maioka-cho 641-12, Totsuka-ku, Yokohama 244-0813, Japan; e-mail: ogihara@yokohama-cu.ac.jp; fax: 81-45-820-1901; ph: 81-45-820-1903.

al., 1986), rice (Hiratsuka et al., 1989), maize (Maier et al., 1995), black pine (Wakasugi et al., 1994), liverwort (Ohyama et al., 1986), Chlorella (Wakasugi et al., 1997) and Euglena (Hallick et al., 1993). Although the overall structure of the chloroplast genome is believed to be conserved by the stabilizing action of its long inverted repeats, small structural changes including inversions (Howe et al., 1988; Hiratsuka et al., 1989), translocations (Ogihara et al., 1988), and deletions (Palmer, 1991) have been found among the angiosperms. Furthermore, comparative sequence analysis of cpDNAs in related plants has uncovered a hot spot for length mutations (Ogihara et al., 1988). This approach will be useful for investigating the mechanisms underlying chloroplast genome evolution. Wheat is one of the major crops in the Gramineae, and the most important member of Triticeae. Since only selected genes and/or regions in the wheat cpDNA were available, we initiated efforts to complete the nucleotide sequence of the wheat plastome as a third candidate from the gramineous plants.

Materials and Methods

Isolation of chloroplast DNA and construction of plasmid contigs

Chloroplast DNA was isolated from 10 to 14 day-old seedling of common wheat, *Triticum aestivum* cv. Chinese Spring by the method of Ogihara and Tsunewaki (1982). Contig clones covering the entire chloroplast genome of common wheat (Table 1; Figure 1) were constructed by cloning restriction fragments (Ogihara and Tsunewaki, 1988) in pBR322 and/or PCR products in pGEM-T plasmids. For determination of the complete nucleotide sequences, the cloned DNA fragments were subcloned in pBluescript or pUC (Table 1). The DNA sequence from both strands was obtained with restricted subclones, exonuclease III deletion subclones, or internal primers. PCR products were determined by sequencing at least three independent fragments. The sequences around the junctions of the clones were verified by the PCR fragments containing the region overlapping the clones.

Results and Discussion

Complete nucleotide sequence and gene organization of the wheat chloroplast genome

The circular wheat cpDNA is 134,540 bp long, of which 20,702 bp are contained in the inverted repeats (IR). The IR disrupts the remaining DNA into 80,347 bp of large single copy (LSC) region and 12,789 bp of small single copy (SSC) region. The nucleotide sequences have been deposited with the DDBJ nucleotide sequence database under the accession number, AB042240. The sequence number is given counterclockwise from the border between the LSC and IRA. The wheat plastome genes were identified by homology to their counterparts in rice (Hiratsuka et al., 1989) and maize (Maier et al., 1995). Identified genes as well as conserved open reading frames among grass plastomes are listed in Table 2. The arrangement and locations of these genes are given in Table 3.

Ribosomal RNA, 16S, 23S, 4.5S and 5S genes are localized in the IR region, as with other land plants (Palmer, 1991). Thirty tRNA genes that can

Table 1. Wheat chloroplast contig clones.

Fragment name	Clone name	Plasmid vector	Fragment size (bp)	Position		Restriction enzyme
				from	to	
Primary clones						
B2	pTacB2	pUC18	9523	50939	60462	*Bam*H I
P1-1	pTacP1-1	pBluescript II SK-	2060	60462	62522	PCR* (*Hin*d III - *Xba* I)
P1-2	pTacP1-2	pBluescript II SK-	2062	62491	64553	PCR* (*Hin*d III - *Xba* I)
P1-3	pTacP1-3	pBluescript II SK-	1993	64552	66545	PCR* (*Pst* I - *Xho* I)
P1-4	pTacP1-4	pBluescript II SK-	1947	66526	68473	PCR* (*Pst* I - *Xho* I)
P1-5	pTacP1-5	pBluescript II SK-	2112	68454	70566	PCR* (*Pst* I - *Xho* I)
P1-6	pTacP1-6	pBluescript II SK-	2283	70556	72839	PCR* (*Pst* I - *Xho* I)
P1-7	pTacP1-7	pGEM-T	2089	72821	74910	PCR*
P1-8	pTacP1-8	pBluescript II SK-	2008	74891	76899	PCR* (*Kpn* I - *Sac* I)
P1-9	pTacP1-9	pBluescript II SK-	2344	76881	79225	PCR* (*Bam*H I - *Sal* I)
P1-10	pTacP1-10	pBluescript II SK-	1969	79205	81174	PCR* (*Bam*H I - *Sal* I)
P2-1	pTacP2-1	pGEM-T	2869	-548	2321	PCR*
P2-2	pTacP2-2	pGEM-T	1432	2089	3521	PCR*
P2-3	pTacP2-3	pGEM-T	3593	3479	7072	PCR*
P2-4	pTacP2-4	pGEM-T	1505	6976	8481	PCR*
P2-5	pTacP2-5	pGEM-T	1545	8356	9901	PCR*
P2-6	pTacP2-6	pGEM-T	2194	9663	11857	PCR*
P3	pTacP3	pBR322	14475	18118	32593	*Pst* I
P4	pTacP4	pBR322	12387	32593	44980	*Pst* I
P5	pTacP5	pBR322	11059	107426	118485	*Pst* I
P6	pTacP6	pBR322	8244	88154	96398	*Pst* I
P7	pTacP7	pBR322	8153	44980	53133	*Pst* I
P8	pTacP8	pBR322	5622	96398	102020	*Pst* I
P9	pTacP9	pBR322	5406	102020	107426	*Pst* I
P10	pTacP10	pBR322	5283	12835	18118	*Pst* I
P12	pTacP12	pBR322	1481	11354	12835	*Pst* I
S3a	pTacS3a	pBR322	13132	75903	89035	*Sal* I
Sub-clones						
B2-1	pTacB2-1	pUC18	2194	50939	53133	*Bam*H I - *Pst* I
B2-2	pTacB2-2	pUC18	1939	53133	55072	*Pst* I
B2-3	pTacB2-3	pUC18	5390	55072	60462	*Pst* I - *Bam*H I
P3-1	pTacP3-1	pBluscript II SK+	2575	18118	20693	*Pst*I - *Sal* I
P3-2	pTacP3-2 SK+	pBluscript II	3699	20693	24392	*Sal* I - *Hin*d III
P3-3	pTacP3-3	pBluscript II SK+	2270	24392	26662	*Hin*d III

Table 1. Wheat chloroplast contig clones [concluded].

Fragment name	Clone name	Plasmid vector	Fragment size (bp)	Position from	to	Restriction enzyme
P3-4	pTacP3-4 SK+	pBluscript II	3708	26662	30370	*Hind* III
P3-5	pTacP3-5	pBluscript II	2223	30370	32593	*Hind* III - *Pst* I
P4-1	pTacP4-1	pUC18	2090	32593	34683	*Pst* I - *Sal* I
P4-2	pTacP4-2	pUC18	1225	34683	35908	*Sal* I
P4-3	pTacP4-3	pUC18	3388	35769	39157	*Bam*H I
P4-4	pTacP4-4	pUC18	3818	38700	42518	*Kpn* I - *Hind* III
P4-7	pTacP4-7	pUC18	2736	42244	44980	*Pst* I - *Xba* I
P5-1	pTacP5-1	pUC18	2982	107426	110408	*Pst* I - *Bgl* II
P5-2	pTacP5-2	pBluscript II SK+	3609	110408	114017	*Bgl* II
P5-3	pTacP5-3	pBluscript II SK+	1377	114017	115394	*Bgl* II
P5-4	pTacP5-4	pBluscript II SK+	141	115394	115535	*Bgl* II
P5-5	pTacP5-5	pUC18	2950	115535	118485	*Bgl* II - *Pst* I
P6-1	pTacP6-1	pBluscript II SK+	181	88154	88335	*Pst* I - *Eco*R I
P6-2	pTacP6-2	pBluscript II SK+	1504	88335	89839	*Eco*R I - *Xba* I
P6-3	pTacP6-3	pBluscript II SK+	2087	89839	91926	*Xba* I - *Apa* I
P6-4	pTacP6-4	pBluscript II SK+	1888	91926	93814	*Apa* I - *Sac* I
P6-5	pTacP6-5	pBluscript II SK+	592	93814	94406	*Sac* I
P6-6	pTacP6-6	pBluscript II SK+	950	94406	95356	*Sac* I
P6-7	pTacP6-7	pBluscript II SK+	1042	95356	96398	*Sac* I - *Pst* I
P7-1	pTacP7-1	pUC18	2577	44980	47557	*Pst* I - *Bgl* II
P7-2	pTacP7-2	pUC19	1489	47273	48760	*Eco*R I
P7-3	pTacP7-3	pUC19	2088	48760	50848	*Eco*R I
P7-4	pTacP7-4	pUC18	2649	50484	53133	*Bgl* II - *Pst* I
P8-1	pTacP8-1	pBluscript II SK-	837	96398	97235	*Pst* I - *Xho* I
P8-2	pTacP8-2	pBluscript II SK-	1206	97235	98441	*Xho* I - *Kpn* I
P8-3	pTacP8-3	pBluscript II SK-	1666	98441	100107	*Kpn* I - *Xho* I
P8-4	pTacP8-4	pBluscript II	1913	100107	102020	*Xho* I - *Pst* I
S3a-1	pTacS3a-1	pUC19	960	80461	81421	*Xba* I - *Hind* III
S3a-2	pTacS3a-2	pUC19	2594	81421	84015	*Hind* III
S3a-3	pTacS3a-3	pUC19	2654	84015	86669	*Hind* III
S3a-4	pTacS3a-4	pUC19	2366	86669	89035	*Hind* III - *Sal* I

* Indicates that PCR product was cloned into the plasmid vector. Inserts can be released by digestion of endonucleases given in parentheses.

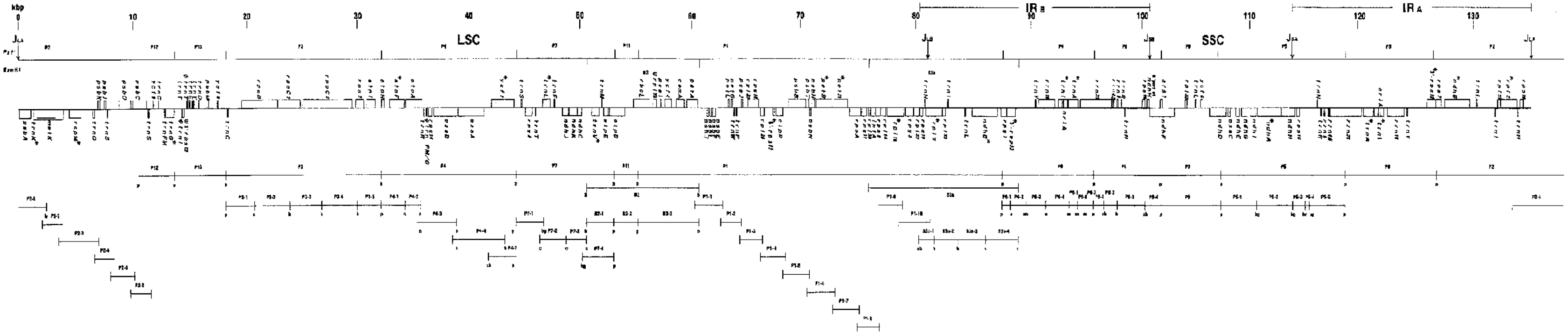

Figure 1. Locations of individual contig clones in relation to the physical and gene maps of the wheat chloroplast genome

Table 2. Genes contained in the wheat chloroplast DNA.

RNA genes		
Ribosomal RNAs		
	23S rDNA, 16S rDNA, 5S rDNA, and 4.5S rDNA	
Transfer RNAs		
	30 different species, corresponding to 20 amino acids, have been identified.	
	trnA	tRNA-Ala (UGC)
	trnC	tRNA-Cys (GCA)
	trnD	tRNA-Asp (GUC)
	trnE	tRNA-Glu (UUC)
	trnF	tRNA-Phe (GAA)
	trnG	tRNA-Gly (GCC), tRNA-Gly (UCC)
	trnH	tRNA-His (GUG)
	trnI	tRNA-Ile (CAU), tRNA-Ile (GAU)
	trnK	tRNA-Lys (UUU)
	trnL	tRNA-Leu (UAA), tRNA-Leu (CAA), tRNA-Leu (UAG)
	trnM	tRNA-Met (CAU)
	(trnfM	tRNA-fMet (CAU))
	trnN	tRNA-Asn (GUU)
	trnP	tRNA-Pro (UGG)
	trnQ	tRNA-Gln (UUG)
	trnR	tRNA-Arg (UCU), tRNA-Arg (ACG)
	trnS	tRNA-Ser (GCU), tRNA-Ser (UGA), tRNA-Ser (GGA)
	trnT	tRNA-Thr (GGU), tRNA-Thr (UGU)
	trnV	tRNA-Val (UAC), tRNA-Val (GAC)
	trnW	tRNA-Trp (CCA)
	trnY	tRNA-Tyr (GUA)
Protein genes		
Photosynthesis		
Photosystem I		
	psaA, psaB, psaC, psaI, psaJ	
Photosystem II		
	psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT	
Cytochrome		
	petA, petB, petD*, petG, petL*	
ATP synthase		
	atpA, atpB, atpE, atpF, atpH, atpI*	
Rubisco		
	rbcL	
Ribosomal proteins		
Large subunits		
	rpl2, rpl14, rpl16*, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36*	
Small subunits		
	rps2, rps3, rps4, rps7, rps8, rps11, rps12, rps14, rps15, rps16*, rps18, rps19*	
Transcription/translation		
RNA polymerase		
	rpoA, rpoB, rpoC1, rpoC2	
Translation factor		
	infA	
NADH-dehydrogenase		
	ndhA, ndhB*, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*	

Table 2. Genes contained in the wheat chloroplast DNA [concluded].

Proteinase	
	clpP
Envelope membrane protein	
	cemA
Maturase	
	matK
Conserved ORFs among plastome of grasses	
	*ycf3**, *ycf4*, *ycf5*, *ycf6*, *ycf9*

*Indicates intron-containing genes.

Table 3. Chloroplast genes in Chinese Spring wheat.

			Coding region	
Gene	Product	Strand	from	to
(LSC)	80,347bp		1	80,347
psbA	PSII 32kDa protein	B	1,152	91
trnK	tRNA-Lys (UUU) 3' exon	B	1,419	1,385
	5' exon	B	3,913	3,972
matK	maturase within trnK intron	B	3,308	1,680
rps16	ribosomal protein S16 3' exon	B	4,734	4,516
	5' exon	B	5,624	5,584
trnQ	tRNA-Glu (UUG)	B	6,803	6,685
psbK	PSII protein	A	7,104	7,289
psbI	PSII protein	A	7,684	7,794
trnS	tRNA-Ser (GCU)	B	7,908	7,995
psbD	PSII D2 protein	A	8,995	10,056
psbC	PSII 43kDa protein	A	10,004	11,424
trnS	tRNA-Ser (UGA)	B	11,659	11,572
ycf9		A	12,017	12,205
trnG	tRNA-Gly (GCC)	A	12,483	12,553
trnfM	tRNA-fMet (CAU)	B	13,074	13,002
trnG	tRNA-Gly (UCC) 3' exon	B	13,214	13,167
	5' exon	B	13,914	13,891
ψtrnI	homology to *trnI* (LAU)	B	14,789	14,698
ψ3'-rps12 a	homology to 3'-rps12	A	15,035	15,106
trnT	tRNA-Thr (GGU)	A	15,035	15,106
ψtrnT	homology to trn-Thr (GGU)	A	15,114	15,174
ψtrnE	homology to trnE (UUC)	A	15,461	15,479
trnE	tRNA-Glu (UUC)	A	15,480	15,552
trnY	tRNA-Tyr (GUA)	A	15,614	15,695
trnD	tRNA-Asp (GUC)	A	16,063	16,136
psbM	PSII low MW protein	A	16,910	17,014
trnC	tRNA-Cys (GCA)	B	18,750	18,680
rpoB	RNA polymerase beta subunit	A	19,915	23,145
rpoC1	RNA polymerase beta' subunit-1	A	23,183	25,234
rpoC2	RNA polymerase beta' subunit-2	A	25,439	29,878
rps2	ribosomal protein S2	A	30,204	30,194
atpI	ATPase a subunit	A	31,187	31,930

Table 3. Chloroplast genes in Chinese Spring wheat [continued].

			Coding region	
Gene	Product	Strand	from	to
atpH	ATPase III subunit	A	32,502	32,747
atpF	ATPase I subunit 5' exon	A	33,201	33,345
	3' exon	A	34,170	34,576
atpA	ATPase alpha subunit	A	34,668	36182
trnR	tRNA-rg (UCU)	B	36,389	36,318
ψtrnfM/G	chimeric pseudogene fusing trnfM (CAU) to trnG (UCC)	B	36,615	36,541
rps14	ribosomal protein S14	B	37,087	36,776
psaB	PSI p700 apoprotein A2	B	39,437	37,233
psaA	PSI p700 apoprotein A1	B	41,714	39,463
ycf3	intron-containing reading frame			
	3rd exon	B	42,566	42,408
	2nd exon	B	43,527	43,298
	1st exon	B	44,407	44,284
trnS	tRNA-Ser (GGA)	A	45,038	45,124
rps4	ribosomal protein S4	B	46,041	45,436
trnT	tRNA-Thr (UGU)	B	46,415	46,343
trnL	tRNA-Leu (UAA) 5' exon	A	47,028	47,062
	3' exon	A	47,651	47,700
trnF	tRNA-Phe (GAA)	A	48,055	48,127
ndhJ	NADH dehydrogenase 30kDa subunit	B	49,054	48,575
ndhK	NADH dehydrogenase 27kDa subunit	B	49,892	49,155
ndhC	NADH dehydrogenase ND3	B	50,245	49,833
trnV	tRNA-Val (UAC) 3' exon	B	51,190	51,156
	5' exon	B	51,827	51,789
trnM	tRNA-Met (CAU)	A	52,033	52,105
atpE	ATPase epsilon subunit	B	52,633	52,219
atpB	ATPase beta subunit	B	54,126	52,629
rbcL	RuBisCO	A	54,909	56,342
ψrpl23	homology to rpl23	A	55,635	56,918
psaI	PSI small peptide	A	57,222	57,332
ycf4		A	57,664	58,221
cemA	envelope membrane protein	A	58,696	59,388
petA	cytochrome f	A	59,620	60,582
psbJ	PSII J protein	B	61,525	61,403
psbL	PSII L protein	B	61,767	61,651
psbF	PSII cytochrome b559 4kDa subunit	B	61,909	61,790
psbE	PSII cytochrome b559 8kDa subunit	B	62,171	61,920
petL	cytochrome b/f complex 3.5kDa subunit	A	63,340	63,435
petG	cytochrome b/f complex subunit 5	A	63,615	63,728
trnW	tRNA-Trp (CCA)	B	63,922	63,849
trnP	tRNA-Pro (UGG)	B	64,136	64,062
psaJ	PSI smallpeptide	A	64,501	64,635
rpl33	ribosomal protein L33	A	65,072	65,272
rps18	ribosomal protein S18	A	65,542	66,054
rpl20	ribosomal protein L20	B	66,576	66,217
5'-rps12	ribosomal protein S12 exon-1	B	67,384	67,271
clpP	ATP-dependent protease proteolytic subunit	B	68,176	67,526
psbB	PSII 47kDa protein	A	68,671	70,197

Table 3. Chloroplast genes in Chinese Spring wheat [continued].

			Coding region	
Gene	Product	Strand	from	to
psbT	PSII T-protein	A	70,361	70,477
psbN	PSII N-protein	B	70,657	70,526
psbH	PSII 10kDa phosphoprotein	A	70,761	70,982
petB	cytochrome B6 5' exon	A	71,112	71,117
	3' exon	A	71,866	72,507
pet D	cytochrome b/f complex subunit 4 5' exon	A	72,694	72,701
	3' exon	A	73,451	73,925
rpoA	RNA polymerase alpha subunit	B	75,140	74,121
rps11	ribosomal protein S11	B	75,648	75,217
rpl36	ribosomal protein L36	B	75,948	75,835
infA	initiation factor 1	B	76,400	76,059
rps8	ribosomal protein S8	B	76,891	76,481
rpl14	ribosomal protein L14	B	77,404	77,033
rpl16	ribosomal protein L16 3' exon	B	77,918	77,517
	5' exon	B	78,970	78,962
rps3	ribosomal protein S3	B	79,831	79,112
rpl22	ribosomal protein L22	B	80,319	79,873
(JLB)	Junction LSC-IR		80,347	80,348
rps19	ribosomal protein S19	B	80,679	80,398
trnH	tRNA-His (GUG)	A	80,811	80,885
rpl2	ribosomal protein L2 3' exon	B	81,371	80,941
rpl2	ribosomal protein L2 5' exon	B	82,425	82,035
rpl23	ribosomal protein L23	B	82,725	82,444
trnI	tRNA-Ile (CAU)	B	82,973	82,900
trnL	tRNA-Leu (CAA)	B	84,551	84,471
ndhB	NADH dehydrogenaseND2 3' exon	B	85,910	85,155
	5' exon	B	87,400	86,624
rps7	ribosomal protein S7	B	88,174	87,705
3'-rps12	ribosomal protein s12 exon-3	B	88,256	88,228
	exon-2	B	89,018	88,797
trnV	tRNA-Val (GAC)	A	90,748	90,819
16S rDNA	16S rRNA	A	91,049	92,540
trnI	tRNA-Ile (GAU) 5' exon	A	92,848	92,884
	3' exon	A	93,691	93,725
trnA	tRNA-Ala (UGC) 5' exon	A	93,792	93,829
	3' exon	A	94,635	94,669
23S rDNA	23S rRNA	A	94,821	97,708
4.5S rDNA	4.5S rRNA	A	97,804	97898
5S rDNA	5S rRNA	A	98,126	98,246
trnR	tRNA-Arg (ACG)	A	98,475	98,548
trnN	tRNA-Asn (GUU)	B	98,893	98,822
rps15	ribosomal protein S15	A	100,425	100,697
5'-ndhH	5' portion of NADH dehydrogenase 49kDa subunit	A	100,843	101,049
(JSB)	Junction IR-SSC		101,049	101,050
ndhF	NADH dehydrogenase ND5	B	103,337	101,118
rpl32	ribosomal protein L32	A	104,256	104,447
trnL	tRNA-Leu (UAG)	A	105,137	105,236

Table 3. Chloroplast genes in Chinese Spring wheat [concluded].

			Coding region	
Gene	Product	Strand	from	to
ycf5		A	105,308	106,270
ndhD	NADH dehydrogenase ND4	B	107,942	106,440
psaC	PSI 9kDa protein	B	108,307	108,062
ndhE	NADH dehydrogenase ND4L	B	109,091	108,786
ndhG	NADH dehydrogenase ND6	B	109,834	109,304
ndhI	NADH dehydrogenase 18kDa subunit	B	110,640	110,098
ndhA	NADH dehydrogenase ND1 3' exon	B	111,276	110,738
	5' exon	B	112,862	112,313
ndhH	NADH dehydrogenase 49kDa subunit	B	114,045	112,864
(JSA)	Junction SSC-IR		113,838	113,839
rps15	ribosomal protein S15	B	114,463	114,191
trnN	tRNA-Asn (GUU)	A	115,995	116,066
trnR	tRNA-Arg (ACG)	B	116,415	116,340
5S rDNA	5S rRNA	B	116,762	116,642
4.5S rDNA	4.5S rRNA	B	117,084	116,990
23S rDNA	23S rRNA	B	120,067	117,180
trnA	tRNA-Ala (UGC) 3' exon	B	120,253	120,219
	5' exon	B	121,096	121,059
trnI	tRNA-Ile (GAU) 3' exon	B	121,197	121,163
	5' exon	B	122,040	122,004
16S rDNA	16S rRNA	B	123,838	122,348
trnV	tRNA-Val (GAC)	B	124,140	124,069
3'-rps12	ribosomal protein S12 exon-2	A	125,870	126,091
	exon-3	A	126,632	126,660
rps7	ribosomal protein S7	A	126,714	127,183
ndhB	NADH dehydrogenase ND2 5' exon	A	127,488	128,264
	3' exon	A	128,978	129,733
trnL	tRNA-Leu (CAA)	A	130,337	130,417
trnI	tRNA-Ile (CAU)	A	131,915	131,988
rpl23	ribosomal protein L23	A	132,163	132,444
rpl2	ribosomal protein L2 5' exon	A	132,463	132,853
	3' exon	A	133,517	133,947
trnH	tRNA-His (GUG)	B	134,077	134,000
rps19	ribosomal protein S19	A	134,209	134,490
(JLA)	Junction IRA-LSC		134,540	1

Accession No. = AB042240.
Total number of nucleotides = 134,540.
Base ratio: A = 41666 (31.0%), C = 25726 (19.1%), T = 41329 (30.7%), G = 25819 (19.2%).
Nucleotide numbers are given counter-clockwise from the start of large single copy region.

transfer 20 species of amino acids were identified. Structures of these genes encoding stable RNAs are similar to those of rice, maize and other angiosperms. A novel small RNA gene, *sprA* found in tobacco chloroplast DNA (Wakasugi et al., 1998) was not detected in wheat cpDNA.

76 protein-encoding genes were identified as follows: 31 for the photosynthetic apparatus, 26 for the transcription-translation system, 11 for NADH-dehydrogenase, 3 for miscellaneous proteins, and 5 for unidentified open

reading frames that are conserved among grass plastomes. The number and content of functional wheat chloroplast genes so far identified are identical to those in rice and maize (Hiratsuka et al., 1989; Maier et al., 1995). However, some wheat genes did exhibit structural differences when compared to other grass genes.

Acknowledgements

The work was supported in part by a Grant-in-Aid (No. 10309008) from the Ministry of Education, Science, Sports and Culture of Japan.

References

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A and Stutz E (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res 21: 3537-3544.

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K and Sugiura M (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.

Howe CJ, Barker RF, Bowman CM and Dyer TA (1988) Common features of three inversions in wheat chloroplast DNA. Curr Genet 13: 343-349.

Maier RM, Neckermann K, Igloi GL and Koessel 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.

Ogihara Y and Tsunewaki K (1982) Molecular basis of the genetic diversity of the cytoplasm in *Triticum* and *Aegilops*: I. Diversity of the chloroplast genome and its lineage revealed by the restriction pattern of ct-DNAs. Jpn J Genet 57: 371-396.

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.

Ogihara Y, Terachi T and Sasakuma T (1988) Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. Proc Natl Acad Sci USA 85: 8573-8577.

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H and Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322: 572-574.

Palmer J (1991) Plastid chromosomes: structure and evolution. In: Bogorad L and Vasil IK (eds), The Molecular Biology of Plastids, pp 5-53, Academic Press, San Diego.

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Simada H and Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. EMBO J 5: 2043-2049.

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudzuki T and Sugiura M (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc Natl Acad Sci USA 91: 9794-9798.

Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsuzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K and Sugiura M (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: The existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci USA 94: 5967-5972.

Wakasugi T, Sugita M, Tsudzuki T and Sugiura M (1998) Updated gene map of tobacco chloroplast DNA. Plant Mol Biol Reptr 16: 231-241.