Theor Appl Genet (1986) 72:353–358

© Springer-Verlag 1986

Genetic analyses of *Oryza* species by molecular markers for chloroplast genomes

H. Ichikawa *, A. Hirai and T. Katayama [1]

Graduate Division of Biochemical Regulation, Faculty of Agriculture, Nagoya University, Nagoya 464, Japan
[1] Faculty of Engineering, Kyushu Kyoritsu University, Kita-Kyushu 807, Japan

Accepted January 27, 1986
Communicated by R. Riley

Summary. Relationships in a wide range of *Oryza* species (13 species) were analyzed using the large subunits (LS) of Fraction I protein (Rubisco) and the Bam HI restriction patterns of chloroplast DNA (ctDNA) as molecular markers. Four types of LS were detected by isoelectrofocusing with and without S-carboxymethylation. The close relation between AA and CCDD genome species was suggested by analyses of LS and ctDNA. Intraspecific variation in *O. latifolia* was detected at the levels of both LS and ctDNA. The LS of the *BB, BBCC,* and *CC* genomes and *FF* (*O. brachyantha*) were not distinguishable, although the native Rubisco of the latter was slightly different from those of the first three. It was also shown that *O. australiensis,* the only *EE* genome species, might have evolved differently than the other *Oryza* species.

Key words: Rice – Chloroplast DNA – Fraction I protein – Rubisco – *Oryza* species

Introduction

There are about 20 species in the genus *Oryza*, most of which are located in Asia and Africa (Tateoka 1963) although some *Oryza* species can be found in South America and Australia.

The African cultivar *O. glaberrima* appears to have originated from an African wild species *O. breviligulata* while the Asian form of *O. perennis* (= *O. rufipogon*) is considered to be the progenitor of cultivated rice, *O. sativa.* However, there is still considerable disagreement about the ancestor of cultivated rice.

In order to have a broadly based breeding program of cultivated rice species, it is important to have information not only about their ancestor but also about the phylogenetic and evolutionary relationships of cultivated rice among all of the *Oryza* species. Many workers have studied the nuclear genomes of *Oryza* species for this purpose (Chang 1976; Morishima 1984), however there is little information about the relationship of chloroplast genomes in the genus. The chloroplast genome is maternally inherited and therefore gives information about the maternal parent in phylogenetic lineages. Since the main part of the genetic information responsible for photosynthesis is coded by chloroplast DNA, information concerning the maternal parent should be important for plant breeding.

Fraction I protein (ribulose-1,5-bisphosphate carboxylase/oxygenase; Rubisco) and chloroplast DNA (ctDNA) are useful molecular markers of chloroplast genomes. Fraction I protein is a key enzyme in photosynthesis and a major soluble protein found in chloroplasts. It consists of two subunits, namely the large (LS) and small (SS) subunits, coded by chloroplast and nuclear DNA, respectively (Wildman 1979). Recently, Pental and Barnes used Fraction I protein to study the phylogeny of *Oryza* species having AA and EE genomes (1985). Furthermore, the evolutionary relationships among many plant species have been evaluated by comparing the restriction endonuclease patterns of ctDNA (Kung et al. 1982; Ogihara and Tsunewaki 1982; Palmer and Zamir 1982; Ichikawa and Hirai 1983; Salts et al. 1984; Palmer et al. 1985).

We have analyzed the chloroplast genome in a wide range of *Oryza* species by using LS and ctDNA as molecular markers, and based on these chloroplast genomes, have attempted to gain information on relationships within the genus *Oryza.*

* Present address: Plantech Research Institute, 1000 Kamoshidacho, Midori-ku, Yokohama 227, Japan

Table 1. Classification of electrophoretic patterns of Fraction I protein form *Oryza* species

Species	(Strain)	2n	Genome	Geographical distribution	Type of LS			Type of native FIP
					Antiserum[c]		Page[d]	
					(– CM)[e]	(+CM)	(– CM)	
sativa	(Sasanishiki)	24	*AA*	Worldwide	M	M	M	II
perennis	(W630)[a]	24	*AA*	Asia	M	M	M	II
glaberrima		24	A^gA^g	Africa	M	M	M	II
breviligulata	(W828)	24	A^gA^g	Africa	M	M	M	II
punctata	(W1515)	24	*BB*	Africa	L	L	L	IV
punctata	(W1145)	48	*BBCC*	Africa	L	L	L	IV
malampuzhaensis	(W021)	48	*BBCC*	Africa	L	L	L	IV
minuta	(Y22)	48	*BBCC*	Asia	L	L	L	IV
collina	(W0006)[b]	24	*CC*	Asia	L	L	L	IV
officinalis	(W0065)	24	*CC*	Asia	L	L	L	IV
latifolia	(W1162)	48	*CCDD*	America	M	M	M	III
latifolia	(W1168)	48	*CCDD*	America	M	H	M	III
latifolia	(W1176)	48	*CCDD*	America	M	H	M	III
latifolia	(W1197)	48	*CCDD*	America	M	M	M	III
alta	(para-1)	48	*CCDD*	America	M	M	M	III
alta	(para-3)	48	*CCDD*	America	M	M	M	III
alta	(W1144)	48	*CCDD*	America	M	M	M	III
australiensis	(W0008)	24	*EE*	Australia	H	H	H	I
brachyantha	(W0237	24	*FF*	Africa	L	L	L	III

[a] Asian form (= *O. rufipogon*); [b] Ceylonese race of *O. officinalis;* [c] FIP was isolated by the use of anti-FIP antiserum; [d] FIP was isolated by polyacrylamide gel electrophoresis; [e] CM: S-carboxymethylation

Materials and methods

Plant materials

A total of 13 species (19 strains), collected from all over the world were used (Table 1). Seeds of *O. sativa* and *O. perennis* were gifts of Dr. T. Abe. Plants were grown individually in a greenhouse.

Preparation of Fraction I protein using antiserum

Leaf tissues were blended in liquid nitrogen followed by homogenization with buffer B (25 mM Tris-HCl, pH 7.4; 0.2 M NaCl; 0.5 mM EDTA) containing 1% (v/v) 2-mercaptoethanol, 12.5% glycerol and 0.2 g/g-tissue of Polyclar AT. Fraction I protein was precipitated from the extracts of *Oryza* species by using anti-tobacco Fraction I protein antiserum and analyzed by isoelectrofocusing according to the method described previously (Hirai 1982).

Preparation of Fraction I protein using polyacrylamide gel electrophoresis

Fraction I protein was isolated according to the method described by Cammaerts and Jacobs (1980) with some modifications. One gram of leaf tissue was frozen by liquid nitrogen and ground in a mortar with 1.5 ml of sample buffer (62.5 mM Tris-HCl, pH 6.8; 20% glycerol, 1% 2-mercaptoethanol) containing 0.2 g/g tissue Polyclar AT. The leaf extract was centrifuged for 15 min at 12,000×g and Fraction I protein in the supernatant was separated by polyacrylamide gel electrophoresis as described by Cammaerts and Jacobs (1980) but modified by adding 12.5% glycerol to the gel. The Fraction I protein band was stained in a solution containing 0.25% Coomassie blue, 50% methanol and 10% acetic acid cut from the gel, and shaken at 25 °C for 2 h in a small amount of urea buffer (0.5 M Tris-HCl, pH 8.5; 8 M urea; 0.5 mM EDTA) containing 5% (w/v) sucrose and 1% (w/v) dithiothreitol. The gel fragments were washed with some changes of urea buffer containing 1.5% Ampholine (LKB), pH 5–8, to remove dithiothreitol and then inserted into slots in the isoelectrofocusing gel. Isoelectrofocusing was performed as previously described (Hirai 1982). After the isoelectrofocusing, Fraction I protein in the slab gel was fixed in a solution containing 6% (w/v) trichloroacetic acid, 4% (w/v) sulfosalicylic acid and 25% (v/v) methanol for at least 3 h. The gel was subsequently washed in a solution containing 25% (v/v) ethanol and 10% (v/v) acetic acid for at least 2 h, and then stained in Coomassie blue solution.

Preparation of total DNA

Twenty-five grams of rice leaves were briefly blended in liquid nitrogen. After the liquid nitrogen was completely evaporated, the leaf powder was suspended in 125 ml of Kool's buffer A (50 mM Tris-HCl, pH 8.0; 0.35 M sucrose; 7 mM EDTA; 5 mM 2-mercaptoethanol: Bovenberg et al. 1981) containing 0.1% bovine serum albumin and 30ml swelled Sephadex G-25. The suspension was filtered through two layers of Micracloth, and centrifuged for 10 min at 1,000×g. The green pellet, containing chloroplasts and nuclei, was washed once with buffer A and DNA was then extracted (Ichikawa and Hirai 1983).

Preparation of ctDNA from Oryza sativa

Chloroplast DNA was extracted from leaf tissue of *O. sativa* as described by Hirai et al. (1985).

Restriction endonuclease analysis

Restriction endonucleases were purchased from Takara Shuzo Co. Ltd. DNA was digested by the enzymes according to the

supplier's instructions. Total DNA was digested with Bam HI and fragments were fractionated by agarose gel electrophoresis according to the method of Sugiura and Kusuda (1979) and Southern blotted to nitrocellulose filters. Nick translated *O. sativa* ctDNA fragments generated by double digestion with Bam HI and Eco RI were used for hybridization with total DNA fragments following the method of Shinozaki and Sugiura (1982).

Results

1 Analysis of Fraction I protein by isoelectrofocusing

a) Using anti-Fraction I protein antiserum. By using anti-Fraction I protein antiserum, Fraction I proteins were precipitated from *Oryza* leaves of 13 species (19 strains) and analyzed by isoelectrofusing in the presence of 8 M urea. Part of the electrophoretic patterns are shown in Fig. 1. With S-carboxymethylation, LS focused at approximately pH 6.5; without carboxymethylation, at pH 7.0. As shown in Fig. 1, three kinds of LS were detected. We designated these LS types H, M, and L in the order of descending pI. The type of LS in each strain is presented in Table 1.

Oryza sativa, a cultivated rice, and *O. perennis,* both having the *AA* genome, and *O. glaberrima,* an African cultivated rice, and *O. breviligulata,* both having the A^gA^g genome, had LS of the middle isoelectric point (M) type. All species having the *CCDD* genome – 4 strains of *O. latifolia* and 3 strains of *O. alta* (= *O. paraguaensis*) – possessed LS of the M type in the non-carboxymethylated protein analysis. However, an intra-specific variation in *O. latifolia* was observed in the case of S-carboxymethylated LS. Although carboxymethylated LS in two strains, W1162 and W1197, of *O. latifolia* were M type as well as *O. alta,* S-carboxymethylated LS in the other strains of *O. latifolia,* W1168 and W1176, were H type LS, which was the same type as *O. australiensis* (*EE* genome). Diploid *O. punctata* (*BB*), *O. collina* and *O. officinalis* (*CC*), *O. brachyantha* (*FF*), tetraploid *O. punctata, O. malampuzhaensis,* and *O. minuta* (*BBCC*) had L type LS, the lowest pI in the genus *Oryza.*

Although we could detect that the bands of SS focused at approximately pH 6.0, the patterns were not clear (Fig. 1). Therefore, the other method to isolate Fraction I protein from *Oryza* was attempted.

b) Using polyacrylamide gel electrophoresis. The band containing Fraction I protein separated by native polyacrylamide gel electrophoresis was excised, treated with urea and SH reagent without carboxymethylation, and applied to isoelectrofocusing. Typical results are presented in Fig. 2.

The various types of LS are summarized in Table 1. The classification of LS was completely identical with the results obtained using antiserum, however the bands of SS were better resolved by electrophoresis. We found that there are no less than two different SS polypeptides in the genus *Oryza.* While the SS of *O. australiensis* is a more basic polypeptide, *AA* and A^gA^g genome species and *O. alta* possess a more acidic form (Fig. 2).

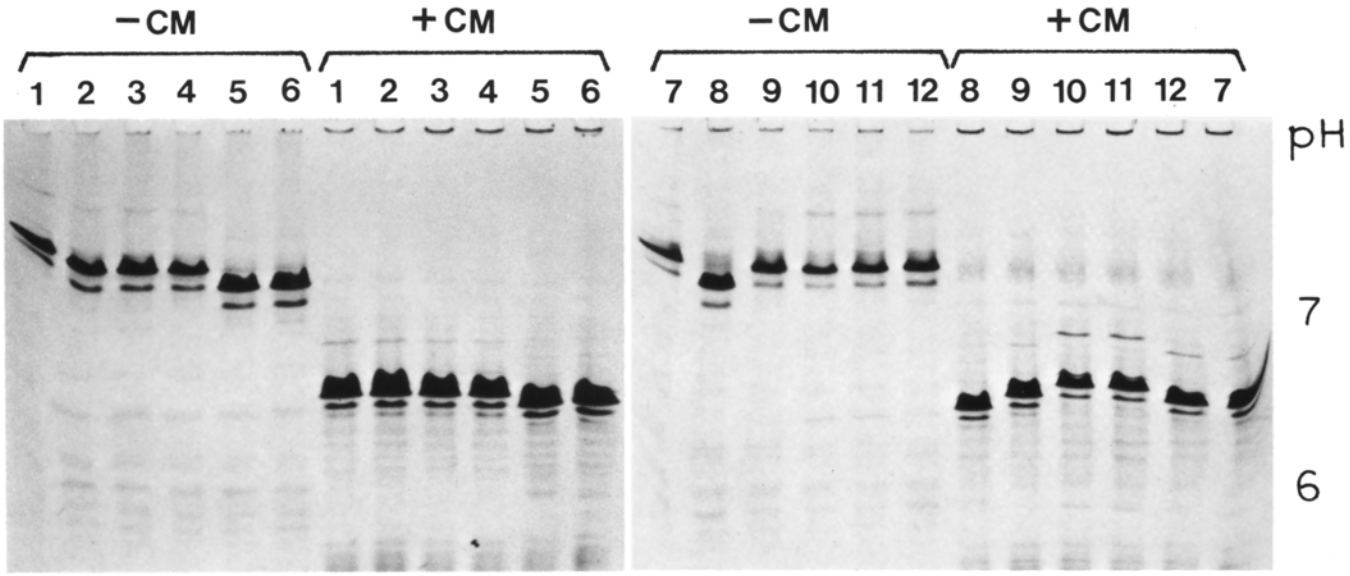

Fig. 1. Isoelectrofocus patterns of Fraction I protein from *Oryza* species. Fraction I protein was precipitated by antiserum and analyzed by isoelectrofocusing in the presence of 8M urea. +CM: with S-carboxymethylation. – CM: without S-carboxymethylation. 1: *O. sativa,* 2: *O. perennis,* 3: *O. glaberrima,* 4: *O. breviligulata,* 5: *O. punctata* (W1515), 6: *O. punctata* (W1145), 7: *O. alta* (para-3), 8: *O. collina,* 9: *O. latifolia* (W1162), 10: *O. latifolia* (W1168), 11: *O. latifolia* (W1176), 12: *O. latifolia* (W1197)

2 Classification of Oryza Fraction I proteins by native polyacrylamide gel electrophoresis

Oryza Fraction I proteins were classified into four different types by analyses with native polyacrylamide gel electrophoresis, as shown in Fig. 3. The four types were named I, II, III and IV in the order of their migration rate. The classification is summarized in Table 1.

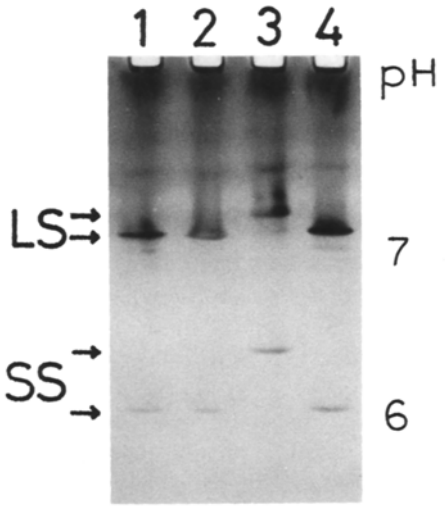

Fig. 2. Isoelectrofocusing of Fraction I proteins from *Oryza* species. Fraction I protein was isolated by polyacrylamide gel electrophoresis and analyzed by isoelectrofocusing as described in "Materials and methods". LS: the large subunits of Fraction I protein. SS: the small subunits of Fraction I protein. 1: *O. alta* (para-1), 2: *O. alta* (W1144), 3: *O. australiensis*, 4: *O. sativa*

Although LS of the *CCDD* genome species had the same type as those of the *AA* and A^gA^g genome species, native Fraction I proteins are in different classes. Species having *CCDD* genomes are type III, while *AA* and A^gA^g genome species are type II native Fraction I proteins.

Similarly, *O. brachyantha* (*FF*), possessing LS indistinguishable from *BB*, *BBCC* and *CC* genome species, was distinguished from these species by the different electrophoretic mobility of native Fraction I protein.

3 Analysis of chloroplast DNAs

Because of the limitation of sample size, we developed the following strategy to analyse chloroplast DNA. Total leaf DNA was extracted, digested by a restriction endonuclease, and fragments were fractionated by agarose electrophoresis and Southern blotted. Chloroplast DNA fragments were visualized after autoradiography using ^{32}P-labelled ctDNA from cultivated rice as probes.

The Southern hybridization patterns of Bam HI fragments from wild rice ctDNAs are shown in Fig. 4. Different fragments are not detected between *O. punctata* (*BB*) and *O. collina* (*CC*). Similarly, two *BBCC* species, *O. punctata* and *O. malapuzhaensis*, have the same restriction patterns. However, the ctDNAs in these two *BBCC* species are distinct from both *BB* and *CC* species (Fig. 4A). Any restriction pattern of ctDNA found in *BB*, *CC*, *BBCC* or *EE* genome species are distinguishable from that of *O. sativa* (*AA*).

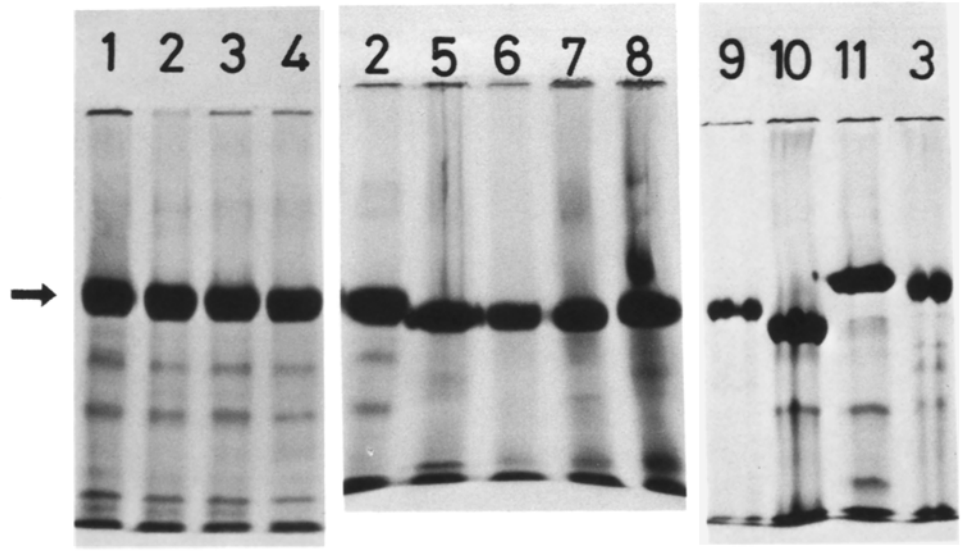

Fig. 3. Analysis of native *Oryza* Fraction I proteins by polyacrylamide gel electrophoresis. Leaf extracts containing Fraction I protein of *Oryza* species were separated by electrophoresis as described in "Materials and methods". *Arrow* indicates the position of native Fraction I proteins. 1: *O. glaberrima*, 2: *O. breviligulata*, 3: *O. sativa*, 4: *O. perennis*, 5: *O. latifolia* (W1162), 6: *O. latifolia* (W1168), 7: *O. latifolia* (W1176), 8: *O. latifolia* (W1197), 9: *O. alta* (para-3), 10: *O. collina*, 11: *O. australiensis*

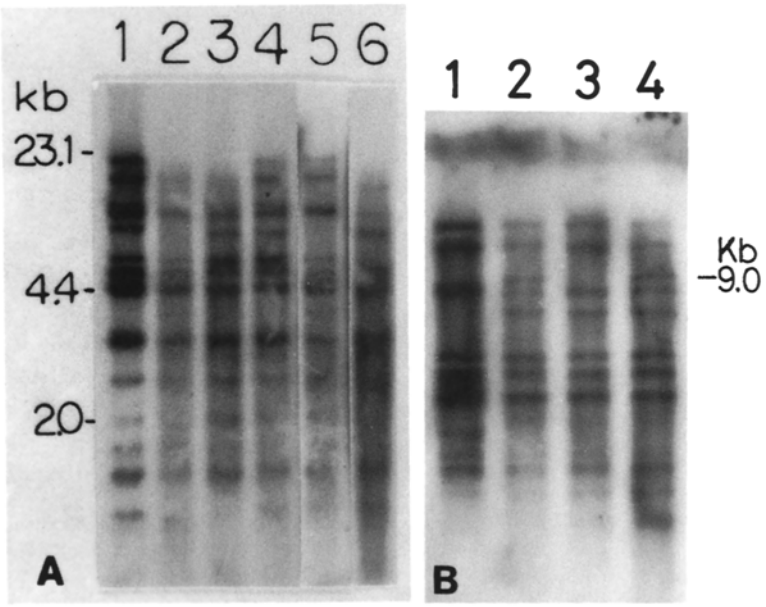

Fig. 4 A, B. Southern blot hybridization of Bam HI fragments from *Oryza* ctDNAs with Bam HI + Eco RI digests of *O. sativa* ctDNA. **A** 1: *O. sativa,* 2: *O. punctata* (W1515), 3: *O. punctata* (W1145), 4: *O. malampuzhaensis,* 5: *O. collina,* 6: *O. australiensis.* **B** 1: *O. collina,* 2: *O. latifolia* (W1162), 3: *O. latifolia* (W1168), 4: *O. alta* (*W1144*)

We have also compared the ctDNA restriction patterns between two *CCDD* species, *O. latifolia* (W1162 and W1168) and *O. alta* (W1144). All species possessing the *CCDD* genome have M type non-carboxymethylated LS. With carboxymethylation, W1168 and W1176 have H type LS while others have M type LS. Therefore, it is interesting to compare ctDNAs from two kinds of rice plants having *CCDD* genomes. As shown in Fig. 4 B, the third largest fragment (9 kb) in W1162 and *O. alta* ctDNA is missing in W1168 ctDNA, and the third fragment of W1168, which was the same size as the fourth fragment in W1162 and *O. alta*, appeared more strongly. Therefore, the intraspecific variation was detected not only at the level of LS but also at the level of ctDNA.

Bam HI fragments of ctDNAs from W1162 and *O. alta* possess considerably homology to that of *O. sativa,* at least in the sizes of five large fragments. CtDNA of *O. collina* (*CC*) was not only different from those of the *AA* and *EE* genome species, but also those of the *BBCC* and *CCDD* genome species.

Discussion

The aim of this research was to investigate the relationships between *Oryza* species by using chloroplast markers. The isoelectrofocusing patterns of LS and the electrophoretic patterns of native Fraction I protein from *AA* genome species were indistinguishable. These results support the findings of Pental and Barnes (1985) who found that LS from *AA* genome species, except for two *O. perennis* accessions from South America (which we did not analyse) were similar.

BB, BBCC, and *CC* genome species, which are to be found throughout Africa and Asia, contain L type LS. Similarly, the African native species, *O. brachyantha,* has the *FF* genome as well as L type LS. Although the electrophoretic mobilities of native Fraction I protein from *BB, BBCC* and *CC* species are different from those of *FF* species, the results might suggest that *BB, BBCC,* and *CC* species have a close relation with *O. brachyantha* (*FF*).

Since the LS from *BB, BBCC,* and *CC* genome species could not be distinguished by isoelectrofocusing, we have tried to compare Bam HI fragments of ctDNAs from these species to determine the female parent of *BBCC* species. However, Bam HI fragment patterns of ctDNAs from *BB* and *CC* species are the same and they are different from that from *BBCC* species. Therefore *BBCC* genome species may have arisen in comparatively ancient times.

O. australiensis (*EE*), which is distributed only in Australia, contains the most basic LS and the slowest mobility of native Fraction I protein in the electrophoresis of *Oryza* species. It also contains rather different ctDNA compared with the other species. *O. australiensis* would have evolved from its ancestor after the fracture of the Gondwanaland continents in which the genus *Oryza* established itself (Chang 1976).

Although the native Fraction I proteins between *CCDD* species and *AA* species are slightly different (Fig. 3; Table 1), we could not detect any differences in subunits of Fraction I proteins between *CCDD* and *AA* species by isoelectrofocusing patterns. Furthermore, the Bam HI fragment patterns of ctDNAs from *AA* species show a stronger resemblance to that found in *CCDD* species than in other species. A close relationship be-

tween *CCDD* species and *AA* species is suggested by these results.

Large subunits of Fraction I proteins from two strains (W1168, W1176) of *O. latifolia* (*CCDD*) have different pI (from LS of other strains of same species) when LS are carboxymethylated, while the LS of both groups have same pI when they are not carboxymethylated (Fig. 1). Bam HI fragment patterns of ctDNAs from two groups were also compared, and showed that ctDNAs from the two groups are different. Therefore, heterogeneity in LS is not an artifact of analysis, but results from different chloroplast genomes. Similar results were also seen in the isoelectrofocusing pattern of Fraction I protein from *Nicotiana* species (Akada and Hirai, unpublished data).

Acknowledgements. We thank Dr. T. Abe for gifts of rice seeds and for discussions, and Dr. H. Uchimiya for discussions. We are also grateful to Prof. S. D. Kung for his critical reading of this manuscript. This work was supported in part by a Grant-in-Aid for Scientific Research from Ministry of Education, Science and Culture.

References

Bovenberg WA, Kool AJ, Nijkamp HJJ (1981) Isolation, characterization and restriction endonuclease mapping of the *Petunia hybrida* chloroplast DNA. Nucl Acids Res 9:503–517

Cammaerts D, Jocobs M (1980) A simple electrophoretic procedure for the determination of the polypeptide composition of the subunits of Fraction I protein. Anal Biochem 109:317–320

Chang TT (1976) The origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. Euphytica 25:425–441

Hirai A (1982) Isoelectrofocusing of non-carboxymethylated fraction I protein from green callus. Plant Sci Lett 25: 37–41

Hirai A, Ishibashi T, Morikami A, Iwatsuki N, Shinozaki K, Sugiura M (1985) Rice chloroplast DNA: a physical map and the location of the genes for the large subunit of ribulose 1,5-bisphosphate carboxylase and the 32KD photosystem II reaction center protein. Theor Appl Genet 70:117–122

Ichikawa H, Hirai A (1983) Search for the female parent in the genesis of *Brassica napus* by chloroplast DNA restriction patterns Jpn J Genet 58:419–424

Kung SD, Zhu YS, Shen GF (1982) Nicotiana chloroplast genome. 3. Chloroplast DNA evolution. Theor Appl Genet 61:73–79

Morishima H (1984) Wild plants and domestification. In: Tsunoda S, Takahashi N (eds) Biology of rice. Japan Sci Soc Press, Tokyo/Elsevier, Amsterdam, pp 3–30

Ogihara Y, Tsunewaki K (1982) Molecular basis of the genetic diversity of the cytoplasm in *Triticum* and *Aegilops*. 1. Diversity of the chloroplast genome and its lineage revealed by the restriction pattern of ct-DNAs. Jpn J Genet 57:371–396

Palmer JD, Jorgensen RA, Thompson WF (1985) Chloroplast DNA variation and evolution in *Pisum:* patterns of change and phylogenetic analysis. Genetics 109:195–213

Palmer JD, Zamir D (1982) Chloroplast DNA evolution and phylogenetic relationship in *Lycopersicon*. Proc Natl Sci Acad USA 79:5006–5010

Pental D, Barnes SR (1985) Interrelationship of cultivated rices *Oryza sativa* and *O. glaberrima* with wild *O. perennis* complex. Analysis of Fraction 1 protein and some repeated DNA sequences. Theor Appl Genet 70:185–191

Salts Y, Herrmann RG, Peleg N, Lavi U, Izhar S, Frankel R, Beckmann JS (1984) Physical mapping of plastid DNA variation among eleven *Nicotiana* species. Theor Appl Genet 69:1–14

Shinozaki K, Sugiura M (1982) The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Gene 20:91–102

Sugiura M, Kusuda, J (1979) Molecular cloning of tobacco chloroplast ribosomal RNA genes. Mol Gen Genet 172: 137–141

Tateoka T (1963) Taxonomic studies of *Oryza*. 3. Key to the species and their enumeration. Bot Mag 76:165–173

Wildman SG (1979) Aspect of Fraction I protein evolution. Arch Biochem Biophys 196:598–610