Copyright © 1987 by the Genetics Society of America

Restriction Site Variation in the *Zea* Chloroplast Genome

J. Doebley,[1] W. Renfroe[2] and A. Blanton

Department of Biology. Texas A&M University, College Station, Texas 77843

Manuscript received March 16, 1987
Revised copy accepted May 23, 1987

ABSTRACT

Nineteen accessions selected from the four species and three subspecies of the genus *Zea* and one accession from the related genus *Tripsacum* were surveyed for variation with 21 restriction endonucleases. In all, 580 restriction sites were assayed in each chloroplast (cp)DNA, this representing 2.2% of the genome. Twenty-four of the 580 sites were variable in one or more of the cpDNAs. The number of nucleotide substitutions per site (p) between *Zea* and *Tripsacum* (0.0056) approximates that between other closely related angiosperm genera. The range in values of p among *Zea* species (0.0003–0.0024) is on the lower end of the range reported for other angiosperm genera. Analysis of the distribution of restriction site mutations throughout the genome indicated that the inverted repeat evolves more slowly than either the small or large unique sequence regions. Parsimony phylogenetic analysis of the restriction site data produced a tree consistent with isoenzymatic and morphological measures of affinity among the species. Chloroplast DNA analysis was not useful in discriminating the subspecies within *Zea mays*. The lack of any detectable differences between the cpDNA of maize (*Z. mays* subsp. *mays*) and some teosintes (*Z. mays* subsps. *mexicana* and *parviglumis*) is consistent with the hypothesis that maize is a domesticated form of teosinte. Comparison of the degree of sequence divergence for *Z. mays* cpDNA and the *Adh1* locus suggests the latter may be evolving at 10 times the rate of the former. Comparison of rates of sequence evolution for the mitochondrial and chloroplast genomes was inconclusive and could not clarify whether these two genomes have dissimilar rates of sequence evolution.

THE chloroplast (cp) genomes of flowering plants consist of circular DNA molecules between 120 and 217 kb in size. In most species, the genome contains small and large unique sequence regions which are separated by a pair of inverted repeats. This arrangement of the genome is conserved throughout the angiosperms with few exceptions (WHITFELD and BOTTOMLEY 1983). Inversions within the large unique sequence region have altered gene order among some groups; however, these events are rare (PALMER 1985). The low frequency of structural changes in angiosperm chloroplast genomes is complemented by a conservative rate of sequence evolution. By some estimates, the synonymous substitution rate is 0.1% per million years (ZURAWSKI, CLEGG and BROWN 1984). Together, the structural and sequence stability of the angiosperm chloroplast genome makes it suitable for restriction fragment analysis among closely related species and genera.

The ability of restriction endonucleases to cleave DNA at specific four, five or six nucleotide recognition sites provides a means for estimating the percent sequence divergence among the chloroplast genomes of angiosperms. When individual base substitutions create or destroy a restriction site, the numbers and sizes of DNA fragments produced by the restriction digest are altered. These alterations can be detected by agarose gel electrophoresis and Southern blotting, and the proportion of altered to conserved sites determined. If one makes the assumption that all site alterations are caused by single nucleotide substitutions, then the number of base substitutions per base pair (p) can be estimated between two plants (UPHOLT 1977; BROWN, GEORGE and WILSON 1979). These estimates may then be used to propose phylogenies (PALMER and ZAMIR 1982) and assess rates of sequence evolution among organisms or among different regions of the genome in a single organism (CLEGG, RAWSON and THOMAS 1984).

In this paper, we examine restriction site variation in the *Zea* chloroplast genome. We compare the proportion of conserved to altered sites in the different regions of the genome and among the species in the genus. Further, we assess whether the degree of cpDNA sequence divergence is similar to what has been reported for some other plant genera, and compare cpDNA evolution in *Zea* with the evolution of the *Zea* nuclear and mitochondrial genomes. Presences/absences of specific restriction sites in the chloroplast genome are used to propose a phylogeny for *Zea*.

[1] Present address: Botany Department, University of Minnesota, St. Paul, MN 55108.
[2] Present address: Biology Department, James Madison University, Harrisonburg, Virginia.

TABLE 1

Accessions analyzed

Taxa	Locality	Source[a]	Collection
Zea Section *Luxuriantes*			
1. *Z. perennis*	Piedra Ancha, Jalisco, Mex.	ILTIS	1050
2. *Z. perennis*	Ciudad Guzman, Jalisco, Mex.	COLLINS	s.n.[b]
3. *Z. diploperennis*	La Ventana, Jalisco, Mex.	GUZMAN	777
4. *Z. diploperennis*	Las Joyas, Jalisco, Mex.	ILTIS (KEMICLE)	1250 (G119)
5. *Z. luxurians*	Ipala, Chiquimula, Gua.	ILTIS	G-42
6. *Z. luxurians*	El Progreso, Jutiapa, Gua.	ILTIS	G-5
Zea Section *Zea*			
Z. mays			
7. subsp. *parviglumis* var. *huehuetenangensis*	San Antonio Huista, Huehuetenango, Gua.	ILTIS	G-120
8. var. *huehuetenangensis*	Santa Ana, Huehuetenango, Gua.	USDA	89360
9. subsp. *parviglumis* var. *parviglumis*	Teloloapan, Guerrero, Mex.	ILTIS	80
10. subsp. *parviglumis* var. *parviglumis*	Teloloapan, Guerrero, Mex.	WILKES (KERMICLE)	47890 (P447)
11. subsp. *parviglumis* var. *parviglumis*	El Salado, Guerrero, Mex.	BEADLE	s.n.
12. subsp. *parviglumis* var. *parviglumis*	La Huertita, Jalisco, Mex.	GUZMAN	s.n.
13. subsp. *mexicana* Race Central Plateau	Durango, Mex.	DOEBLEY	625
14. subsp. *mexicana* Race Central Plateau	Copandiro, Michoacan, Mex.	WILKES (KERMICLE)	48703 (P446)
15. subsp. *mexicana* Race Nobogame	Nobogame, Chihuahua, Mex.	BEADLE	2050
16. subsp. *mexicana* Race Chalco	Texcoco, Mexico, Mex.	DOEBLEY	479
17. subsp. *mexicana* Race Chalco	Tlalmanalco, Mexico, Mex.	DOEBLEY	642
18. subsp. *mays* Race Northern Flint	USA	USDA	214195
19. subsp. *mays* Race Southern Dent	USA	USDA	9432
Tripsacum			
20. *T. dactyloides*	Waldo, Florida	DOEBLEY	644

[a] Some lines were created by J. KERMICLE of the University of Wisconsin by taking a single teosinte plant and pollinating it and its offspring with pollen from a maize inbred line. The process was continued for several generations until a stock with teosinte cytoplasm and the maize nuclear genome was ultimately created.

[b] s.n. = sine numero.

MATERIALS AND METHODS

The genus *Zea* includes maize and its wild relatives, the teosintes. As treated by ILTIS and DOEBLEY (1980), *Zea* contains four species divided between two sections. Section *Zea* contains one species, *Z. mays*. This species is highly polymorphic containing three subspecies: subsp. *mays* (corn or maize), subsp. *mexicana* (Chalco, Central Plateau and Nobogame teosintes) and subsp. *parviglumis* (Balsas and Huehuetenango teosintes). Section *Luxuriantes* contains three wild species (teosintes), *Z. perennis*, *Z. diploperennis* and *Z. luxurians* (WILKES 1967; ILTIS and DOEBLEY 1980). For the present study, 19 accessions of *Zea* were chosen to represent the array of taxonomic variation in *Zea* (Table 1). For most accessions, seed from a single plant was used for chloroplast isolation. The only exceptions are samples 9, 11, and 20 (Table 1) for which chloroplasts were isolated from the pooled leaves of several plants. For comparative purposes, cpDNA was isolated from *Tripsacum dactyloides*, which belongs to the same tribe (Andropogoneae) of the grass family as *Zea* (Table 1).

Chloroplasts were isolated from 2–3-week-old seedlings by the sucrose gradient method (PALMER 1986). After isolation, the chloroplasts were resuspended in 50 mM Tris-25 mM EDTA (pH 8.0) and frozen (−20°) for later DNA extractions. DNA was isolated from the chloroplasts by phenol and chloroform extractions (ZIMMER and NEWTON 1982).

Restriction enzymes were purchased from Bethesda Research Laboratories or New England Biolabs and the restrictions carried out according to manufacturer's recommendations. Each of the 20 cpDNAs (Table 1) was restricted with 21 restriction endonucleases (Table 2). Restricted DNAs were electrophoresed in 0.8% agarose gels with a running buffer of 100 mM Tris-acetate, 1 mM EDTA (pH 8.1). DNA fragment sizes were estimated using the maximum likelihood method with *Hin*dIII digested bacteriophage λDNA and/or *Hae*III digested φX174 DNA as standards (SCHAFFER and SEDEROFF 1981). DNA fragments in the gels were denatured and transferred to a nylon hybridization membrane (Gene Screen Plus) according to manufacturer's (New England Nuclear) recommendations. Nylon filters were prehybridized overnight in a hybridization buffer of 10% dextran sulfate, 1 M NaCl and 1% SDS. Cloned DNA fragments were labeled with ^{32}P-dATP by nick translation as described by MANIATIS, FRITSCH and SAMBROOK (1982). Nick translated probes were separated from unincorporated ^{32}P-dATP on spun columns (MANIATIS, FRITSCH and SAMBROOK 1982), denatured, and then added to the hybridization buffer. Hybridizations were carried out at 65° overnight. Hybridization membranes were washed according to manufacturer's instructions and exposed to X-ray film (Kodak XAR-5) at −80° for 6 hr to 1 week using Dupont Cronex intensifier screens.

Cloned portions of the chloroplast genome were used to probe the nylon filters. These include (1) plasmid clones of the sorghum chloroplast genome—pLD 3, 9 and 24 (DANG and PRING 1986); (2) charon 4A clones of maize cpDNA—λ-9, λ-11 and λ-12 (LAURRINUA *et al.* 1983); (3) a single

TABLE 2

Number of bands scored (and the number of variable sites observed, if any) with each probe for each restriction enzyme

	Probe						
Enzyme	cB9	λ-9	λ-11	λ-12	pS11	pLD3	pLD9 + 24
*Bam*HI	10 (1)	7 (2)	5	5	—[a]	2	2
*Bcl*I	13	4	6 (1)	6	—	3	2
*Bgl*I	4	1	1	4	—	3	1
*Bgl*II	8	5	7	3	—	3	5
*Bst*EII	8	1	1	1	—	6	2
*Cfo*I	14 (1)	4	10	6 (1)	—	5	3
*Cla*I	6	7	6	6 (1)	—	3	6
*Dra*I	14 (3)	12 (2)	7	5	—	4	1
*Eco*RI	10	10 (2)	12 (2)	5	—	1	5
*Eco*RV	6	5	4	3	—	2	4
*Hae*III	—	—	—	—	17 (3)	7	14
*Hin*dIII	9	3	6	2	—	6	3
*Kpn*I	4	2	4	2	—	4	2
*Nco*I	7 (1)	2	3	4	—	2	2
*Nru*I	4	2	3	4 (1)	—	3	2
*Nsi*I	12 (1)	8	4	2	—	2	1
*Pst*I	7	4	3	3	—	3	3
*Sac*I	7	2	5	1	—	5	2
*Ssp*I	10	7	7	12 (2)	—	5	3
*Stu*I	5	1	2	2	—	3	5
*Xba*I	3	5	2	6	—	2	1
Total	161 (7)	92 (6)	98 (3)	82 (5)	17 (3)	74	69

[a] Dash indicates that no hybridization was performed.

cosmid clone of maize cpDNA—cB9 (D. M. LONSDALE, unpublished data); and (4) a 12.5-kb *Sma*I fragment cloned from *Z. perennis*—pS11. The positions of all clones are indicated on the maize chloroplast genome map (Figure 1). If blots probed with large λ and cosmid clones produced too many bands for easy interpretation, then they were reprobed with the smaller sorghum plasmid clones (DANG and PRING 1986), or *Pst*I or *Sma*I clones of the *Zea* chloroplast genome made by the authors.

Estimates of the number of nucleotide substitutions per site (p) between each of the cpDNAs were calculated by the maximum likelihood method. Equations 19, 21 and 23 from NEI and TAJIMA (1983) were executed by the program MAXLIKE provided to the authors by NEI. A Wagner parsimony phylogenetic tree for the *Zea* taxa was constructed with the Mix program of the Phylogenetic Inference Package version 2.9 created by J. FELSENSTEIN.

RESULTS AND DISCUSSION

Restriction site variation: The number of fragments visualized for each enzyme with each probe is given (Table 2). After correcting for the fact that adjacent probes hybridize to many of the same fragments, it was estimated that a total of 580 sites had been assayed, representing 2.2% of the genome. Of the 580, 24 were variable in one or several of the cpDNAs.

For each of the 24 variant sites, the expected additive pattern of fragment sizes was observed (Table 3). Thus, when *Bam*HI digests were probed with cB9, cpDNAs 1 to 6 differed from all other cpDNAs by the addition of a new restriction site which caused the loss of a 12.82-kb fragment and the gain of 8.35- and 4.16-kb fragments (Table 3; Figure 2). While all mutations gave straightforward patterns such as this, often it was necessary to reprobe the filters with subclones of the large λ and cosmid clones. Southern blots for 8 of the 24 mutations are shown (Figure 2).

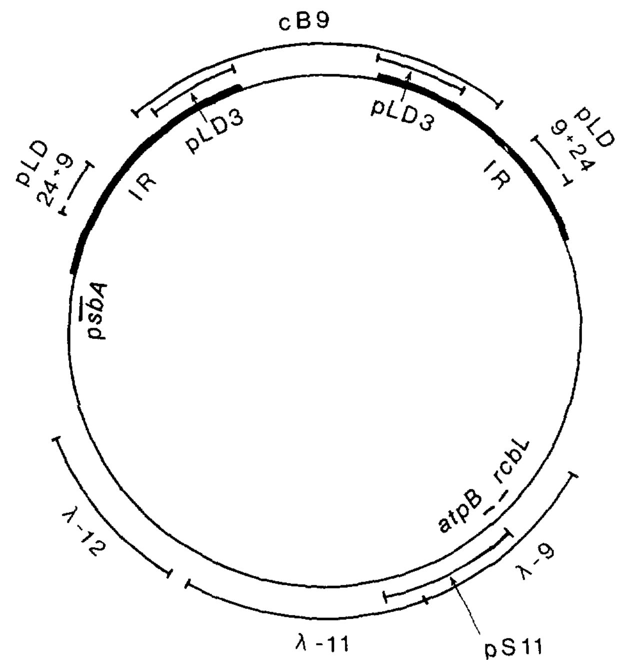

FIGURE 1.—Schematic diagram of the *Zea* chloroplast genome. The locations of cloned cpDNA sequences used as probes are shown. Locations of the inverted repeats (*heavy line*), ATP synthase subunit β (*atpB*), large subunit of ribulose-1,5-bisphosphate carboxylase (*rbcL*) and the 32-kd thylakoid membrane protein (*psbA*) are indicated.

Several of the Southern blots depicted in Figure 2 deserve further comment. First, mutations 1 and 8 (Figure 2; Table 3), both involving *Bam*HI, account for the differences in the *Bam*HI cpDNA banding patterns previously shown by TIMOTHY *et al.* (1979). Although the methods employed by these authors did not allow determination of the genetic basis for the differences in banding patterns, our data demonstrate that the two restriction site changes are responsible. Similarly, mutations 12 and 13 (Figure 2; Table 3) account for the differences in *Eco*RI banding patterns shown by TIMOTHY *et al.* (1979). These two mutations are somewhat more difficult to interpret because they occur in adjacent fragments (*Eco*RI-d and e) of similar size (*cf.* LARRINUA *et al.* 1983). Mutation 12 splits the 3.9-kb *Eco*RI-d into 2.94- and 0.84-kb fragments, while mutation 13 splits the 3.74-kb *Eco*RI-e into 2.91- and 0.78-kb fragments. Our data suggest that the differences shown among *Hin*dIII cpDNA banding patterns for *Zea* by TIMOTHY *et al.* (1979) result from several small insertion/deletion mutations (DOEBLEY, MA and RENFROE 1987).

Sequence divergence: An estimate of the number

TABLE 3

Changes in the numbers and sizes of cpDNA fragments caused by loss/gain of a restriction site[a]

Mutation	Enzyme	Probe	Region[b]	Losses (kb)	Gains (kb)	Mutated Samples
1	*Bam*HI	cB9	SUS	12.82	8.35 + 4.16	1–6
2	*Cfo*I			5.11	2.90 + 2.24	10, 12
3	*Dra*I			5.98	5.32 + 0.62	20
4	*Dra*I			6.10 + 0.62	6.67	1, 2, 20
5	*Dra*I			0.86 + 0.72	1.66	1–6
6	*Nco*I			7.00	6.40 + 0.63	20
7	*Nsi*I			2.54	2.05 + 0.38	20
8	*Bam*HI	λ-9	LUS	6.65	4.45 + 2.09	7–19
9	*Bam*HI			6.23	5.80 + 0.43	20
10	*Dra*I			3.56 + 0.45	4.22	20
11	*Dra*I			1.86	1.47 + 0.38	20
12	*Eco*RI			3.90	2.95 + 0.84	9, 11, 15, 17
13	*Eco*RI			3.74	2.91 + 0.78	1–6
14	*Bcl*I	λ-11	LUS	6.77 + 2.43	9.49	1–4
15	*Eco*RI			1.00 + 0.91	1.89	20
16	*Eco*RI			0.95	0.78 + 0.21	7–19
17	*Cfo*I	λ-12	LUS	5.01	2.45 + 2.45	20
18	*Cla*I			3.83	3.40 + 0.39	20
19	*Nru*I			12.64 + 2.43	15.59	20
20	*Ssp*I			1.15 + 1.00	2.25	20
21	*Ssp*I			2.08 + 0.34	2.46	20
22	*Hae*III	pS11	LUS	3.15	2.68 + 0.51	5, 6
23	*Hae*III			2.48	1.37 + 1.17	20
24	*Hae*III			1.86	1.45 + 0.44	20

[a] Restriction site changes were polarized using *Tripsacum* as the outgroup. For mutations that distinguish *Zea* from *Tripsacum*, no outgroup is available, and assignment of fragments to loss or gain categories is arbitrary.
[b] SUS = small unique sequence region and LUS = large unique sequence region.

of nucleotide substitutions per site (p) between the cpDNAs (Table 4) was calculated by the maximum likelihood method. Values for p between *Tripsacum* and *Zea* species ranges from 0.0052 to 0.0058. Similar values were reported by Clegg, Rawson and Thomas (1984) between *Cenchus* and *Pennisetum*, two other grass genera. Likewise, Palmer and Zamir (1982) reported percent sequence divergence of 0.6–0.7% between *Lycopersicon* and closely related species of *Solanum*. Thus, all three studies show relatively little sequence divergence between the chloroplast genomes of closely related angiosperm genera.

It is important to note that our values for p between *Zea* and *Tripsacum* may not be comparable to those reported by some other authors because we did not employ probes covering the portions of the large single copy region bordering the inverted repeats. These regions have been shown to be the most variable segments of the genome in the angiosperm genera *Brassica* (Palmer *et al.* 1983) and *Clarkia* (Sytsma and Gottlieb 1986). Thus, our values for p may underestimate the actual values for the entire genome.

The comparative restriction maps of *Zea* and *Sorghum* (Dang and Pring 1986) enable one to calculate that the number of nucleotide substitutions per site is 0.0155 ± 0.0043 between these two genera, both of which belong to the tribe Andropogoneae. If the evolution of the chloroplast genome among these grass genera has been clocklike, then these values suggest that the divergence between *Zea* and *Sorghum* occurred 2–4 times as long ago as did the divergence between *Zea* and *Tripsacum* (0.0056). This multiple (2–4×) seems small because *Zea* and *Tripsacum* are closely allied and can be hybridized (albeit with difficulty), while *Sorghum* belongs to a wholely separate portion of the Andropogoneae than *Zea* and *Tripsacum*.

Within *Zea*, values for p among the four species range from 0.0003 to 0.0024, representing 1–8 restriction site changes. These values are similar to those reported among species of *Lycopersicon* (Palmer and Zamir 1982) and *Pennisetum* (Clegg, Rawson and Thomas 1984). They are somewhat smaller than sequence divergences reported among species of *Brassica* (Palmer *et al.* 1983) and *Pisum* species (Palmer, Jorgensen and Thompson 1985), and an order of magnitude smaller than those reported for *Clarkia* (Sytsma and Gottlieb 1986). The small cpDNA sequence divergence values for *Zea* are consistent with the hypothesis that it is a relatively young genus (Doebley 1983).

No variation in cpDNA was observed within *Z. perennis*, *Z. diploperennis* and *Z. luxurians*; however, only two samples (from separate populations) were studied for each species. There is relatively little morphological variation within these three species and

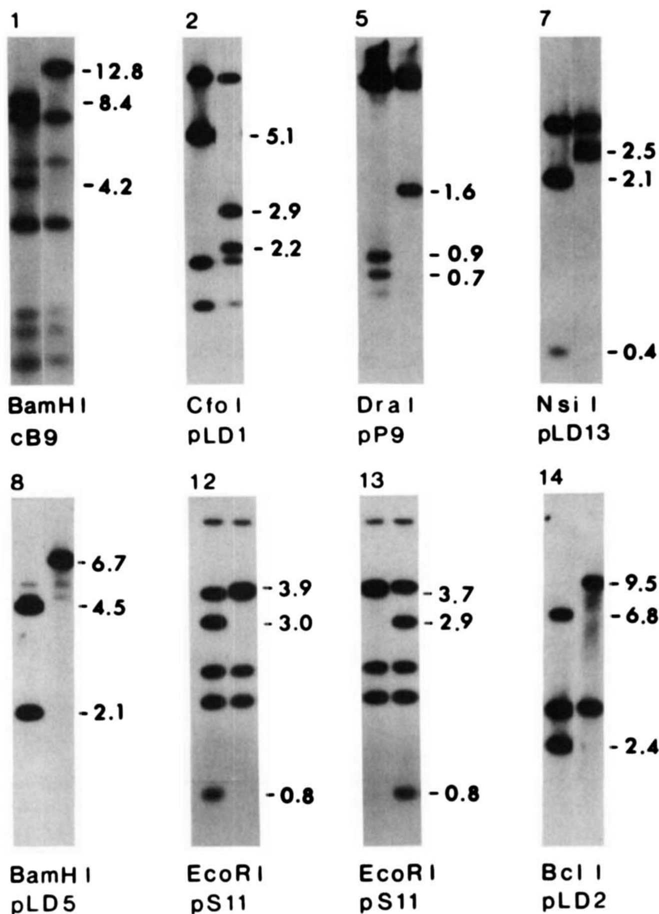

FIGURE 2.—Southern blots showing restriction site mutations in the *Zea* chloroplast genome. The identifying number of the mutation (Table 3) appears above each blot and the restriction enzyme and name of the cloned probe below. Fragment sizes are indicated in kb. Cloned probes labelled pLD are described by DANG and PRING (1986), cB9 and pS11 are shown on Figure 1, and pP9 is a pUC clone of *Pst*I-9 fragment of *Z. mays*.

they have restricted geographic ranges. Among the 13 samples of *Z. mays* (*sensu lato*), three distinct cpDNA types were defined by two variable restriction sites. Values for p varied among populations of *Z. mays* (*sensu lato*) from 0.00 to 0.0006, representing 0–2-site differences. Chloroplast DNA restriction site variation within species has been observed in several other genera (PALMER and ZAMIR 1982; PALMER, JORGENSEN and THOMPSON 1985; BANKS and BIRKY 1985). *Z. mays* as defined by ILTIS and DOEBLEY (1980) is a large polymorphic species including several wild subspecific taxa as well as the cultigen, *Z. mays* subsp. *mays* (maize). This species encompasses considerable morphological variation and is geographically wide spread. Thus, it is not surprising to find some cpDNA sequence divergence within it.

Distribution of variation within the genome: For the present data set, no variable sites were observed among the 164 sites within the inverted repeat. Seven variable sites were found among the 101 sites (6.9%) in the small unique sequence region, and 17 variable

TABLE 4

Estimates of the number of nucleotide substitutions per site, p (given as 100p), and their standard errors in parentheses (upper triangle) and number of restriction site mutations among the cpDNAs (lower triangle)[a]

	Cytoplasms[b]						
	T	D	P	L	M1	M2	M3
T	—	0.581	0.550	0.581	0.550	0.550	0.519
	—	(0.096)	(0.093)	(0.096)	(0.093)	(0.093)	(0.090)
D	19	—	0.030	0.060	0.212	0.212	0.181
		—	(0.021)	(0.030)	(0.057)	(0.057)	(0.053)
P	18	1	—	0.090	0.242	0.242	0.212
			—	(0.037)	(0.061)	(0.061)	(0.057)
L	19	2	3	—	0.212	0.212	0.181
				—	(0.057)	(0.057)	(0.053)
M1	18	7	8	7	—	0.060	0.030
					—	(0.030)	(0.021)
M2	18	7	8	7	2	—	0.030
						—	(0.021)
M3	17	6	7	6	1	1	—
							—

[a] Rows and columns are labeled with the first letter of the species name except T for *Tripsacum dactyloides*.
[b] For *Zea mays*, M1 includes samples 10 and 12, M2 includes samples 9, 11, 15 and 17, and M3 includes samples 7, 8, 13, 14, 16, 18 and 19 (Table 1). The accessions within the other species were invariant.

sites among the 315 sites (5.4%) in the large unique sequence region. This pattern is similar to what has been observed in *Lycoperiscon* (PALMER and ZAMIR 1982), *Pennisetum* (CLEGG, RAWSON and THOMAS 1984) and *Clarkia* (SYTSMA and GOTTLIEB 1986). All studies suggest that the inverted repeats evolve more slowly than the unique sequence regions, and the latter study suggests that different portions of the large unique sequence region have dissimilar rates of sequence evolution. TIMOTHY, HU and LEVINGS (1981) had previously predicted that the inverted repeats are conserved relative to the unique sequence regions in *Zea* based on the comparison of restriction fragment patterns. Our data confirm their suspicions.

Phylogenetic analysis: A Wagner parsimony phylogenetic tree for the *Zea* taxa was constructed using *Tripsacum dactyloides* as the outgroup to root the tree (Fig. 3). In addition to the 24 restriction site mutations described herein, five length (insertion/deletion) mutations, four (LM-1–4) of which have been previously described (DOEBLEY, MA and RENFROE 1987), were included. The fifth insertion/deletion mutation (LM-5) occurs in the region of the inverted repeat covered by pLD9 and pLD24. It is approximately 50 bp in size and by outgroup analysis is judged to be a deletion. This mutation was found only in the two samples of *Z. luxurians*.

The parsimony tree (Fig. 3) requires 30 steps to account for the 29 mutations. Only restriction site mutation number 4 (Table 2) undergoes a reversal, which is the convergent or parallel loss of a *Dra*I site

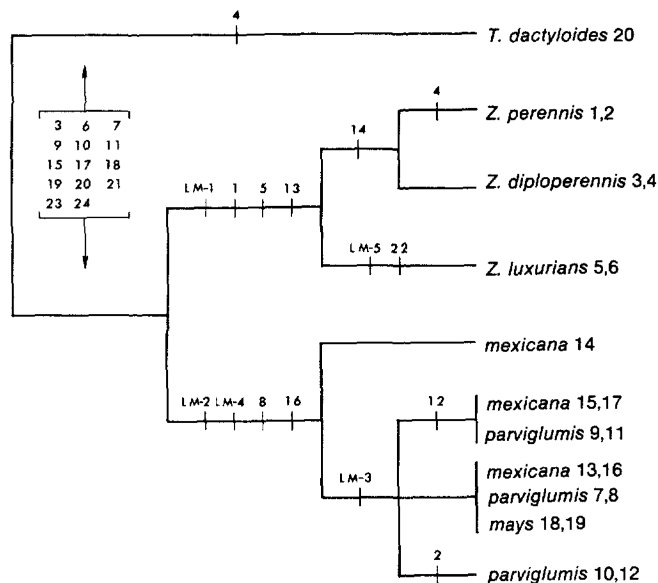

FIGURE 3.—Wagner parsimony tree for *Zea* based on restriction site loss/gain and insertion/deletion mutations in the chloroplast genome. Sample numbers from Table 1 and specific or subspecific names appear at the branch ends. The identification numbers of restriction site mutations (Table 3) and of length mutations (LM) described by DOEBLEY, MA and RENFROE (1987) or in the text appear along the branch segments.

in *T. dactyloides* and *Z. perennis.* This rate of convergence is similar to that reported for *Lycopersicon-Solanum* (1 of 39), *Brassica* (2 of 50) and *Clarkia* (6 of 119) (PALMER and ZAMIR 1982; PALMER *et al.* 1983; SYTSMA and GOTTLIEB 1986). Wilcoxon matched-pairs signed ranks tests (TEMPLETON 1983) for the phylogenetic tree (Fig. 3) did not allow the rejection of the null hypothesis that the *Zea* cpDNAs are evolving in a clock-like fashion.

Figure 3 agrees in several aspects with previous morphological and isozyme studies (DOEBLEY 1983; DOEBLEY, GOODMAN and STUBER 1984). First, *Zea* is split into two sections with precisely the same species compositions as in the classification proposed by DOEBLEY and ILTIS (1980). Second, all lines of evidence suggest *Z. luxurians* has a close relationship to *Z. perennis* and *Z. diploperennis.* Third, *Z. perennis* and *Z. diploperennis* are distinguished by only a single restriction site mutation which is congruent with the hypothesis that the former is a relative recent autopolyploid derivative of the latter (SHAVER 1962). Fourth, *Z. mays* subsp. *mays* is closely allied to subsps. *parviglumis* and *mexicana,* and it could not be distinguished from some populations of these two teosintes.

The principal discordance between cpDNA data and previous isozyme and morphological studies is the failure of the cpDNA restriction analysis to resolve *Z. mays* subsp. *mexicana* and *parviglumis.* This is not completely unexpected as cpDNA analyses have not been shown taxonomically informative below the species level. Curiously, what heterogeneity of cpDNA does exist in *Z. mays* (*sensu lato*) is not organized along taxonomic lines. This suggests either that some chloroplast restriction sites are convergent (or parallel) or that the pattern of evolution among *Z. mays* subspecies is reticulate. Lack of concordance between cpDNA data and other measures of genetic relatedness have been encountered within other species (PALMER and ZAMIR 1982; PALMER, JORGENSEN and THOMPSON 1985).

The morphological differences between maize and teosinte are of such magnitude that until rather recently the two were placed by most authors in separate genera (REEVES and MANGELSDORF 1942). However, this is incongruent with much genetic and biosystematic evidence which suggests that maize and some types of teosinte are essentially conspecific (BEADLE 1939, 1972; GALINAT 1971; ILTIS and DOEBLEY 1980; DOEBLEY, GOODMAN and STUBER 1984). Most authorities have given priority to the genetic-biosystematic evidence and supported the hypothesis that maize is a domesticated form of teosinte that has undergone a dramatic morphological change under domestication (BEADLE 1972; GALINAT 1983; ILTIS 1983; DOEBLEY 1983). The cpDNA data are consistent with this hypothesis in that they indicate maize occupies a recent branchlet in the phylogeny of *Zea* (Fig. 3). The cpDNA genotype of maize is also found in *Z. mays* subsps. *parviglumis* and *mexicana,* suggesting one of these two taxa (or their ancestral forms) was the progenitor of maize.

Evolution of chloroplast and other *Zea* genomes: A conservative rate of evolution for the chloroplast genome of angiosperms has been suggested, in part, because populations with identical chloroplast genomes can exhibit high levels of polymorphism for their nuclear genomes. This is also the case with *Zea.* For example, the cpDNAs of *Z. luxurians* and *Z. diploperennis* are distinguished by only two restriction site mutations and one insertion/deletion event. Yet, for isozymes these two species have different alleles or frequencies of alleles at 17 of 21 isozyme loci (DOEBLEY, GOODMAN and STUBER 1984). Similarly, we have examined 12 accessions of *Z. mays* subsp. *mays* from throughout Latin America, and all had the same cpDNA genotype. Yet, isozyme diversity among Latin American maize races is extremely high (GOODMAN and STUBER 1983; DOEBLEY, GOODMAN and STUBER 1984).

Comparative rates of sequence evolution between *Zea* chloroplast and nuclear genomes can be compared more directly using the DNA sequences of *Adh1-S* and *F* (SACHS *et al.* 1986). These two alleles, which are also designated *Adh1-6* and *4,* have known distributions in *Zea* (DOEBLEY, GOODMAN and STUBER 1984). *Adh1-F* (*4*) is found throughout the genus, while *Adh1-S* (*6*) is restricted to the three subspecies

of *Z. mays*. This distribution suggests (although it does not prove) that *Adh1-S* arose by mutation after the divergence of *Z. mays* (*sensu lato*) from the other species of *Zea*. The translated regions of *Adh1-S* and *F* have 1.23% sequence divergence, their introns 2.01%, and the matched-up 3′ flanking regions 4.5%. These values compare with between 0.03 and 0.06% during the same evolutionary period for the portions of the chloroplast genome we assayed. Thus, if our interpretation that *Adh1-S* arose from *Adh1-F* after the origin of the *Z. mays* is correct, then *Adh1* is evolving at over 10 times the rate of the *Zea* chloroplast genome.

These data suggest a much greater rate of sequence evolution for *Adh1* as compared to the chloroplast genome. Yet, *Adh1* is among the more conservative isozyme loci in *Zea*, with several other loci, most notably *Glu1*, *Idh2*, *Mdh2* and *Pgm2*, showing much higher rates of polymorphism (DOEBLEY, GOODMAN and STUBER 1984, 1985). Of 23 loci surveyed in Mexican maize races, 17 had a greater number of alleles than *Adh1*. Thus, the *Adh1* sequences may underestimate the rate of evolution for isozyme genes encoded in the nucleus. On the other hand, restriction fragment analysis, because it can not discriminate between base substitutions and small insertion/deletion events, is apt to overestimate the rate of sequence evolution in the chloroplast genome. Thus, the disparity in rate of sequence evolution for the portions of the chloroplast we assayed and nuclear genomes of *Zea* may be even greater than our estimate indicates. Several other authors have suggested that nuclear genes evolve more rapidly than cpDNA in *Lycopersicon* (PALMER and ZAMIR 1982), *Pennisetum* (CLEGG, RAWSON and THOMAS 1984) and *Lisianthius* (SYTSMA and SCHAAL 1985). This is also the case in *Zea* and accounts for the fact that subspecies, races and populations which have the same cpDNA genotypes may have a considerable degree of isozyme variation (DOEBLEY, GOODMAN and STUBER 1984).

Limited comparison of the rates of sequence evolution in cpDNA and mtDNA is possible from published reports. However, because angiosperm mitochondrial genomes undergo relatively frequent structural rearrangements (SEDEROFF *et al.* 1981; SCHARDL *et al.* 1984), accurate estimates of percent sequence divergence may not be obtained from restriction fragment analyses. For this reason, we compare our estimates for cpDNA sequence evolution to evidence from DNA sequence analysis of the *Z. mays* and *Z. diploperennis* 18S-5S ribosomal gene region and cytochrome oxidase II gene (GWYNN *et al.* 1987). These mitochondrial genes (and flanking regions) show 0.67% sequence divergence between these two species, whereas the cpDNA analysis reveals only 0.18%. Because these two estimates were obtained by wholely different means, this comparison should be viewed cautiously. Nevertheless, they apparently indicate a more rapid rate of sequence evolution for the mitochondrial genome. This interpretation is contradicted if one compares the rates of sequence evolution between mitochondrial ribosomal genes and cytochrome oxidase II, and the chloroplast *rbcL* (GWYNN *et al.* 1987). The former shown an overall substitution rate of 0.019% per million years and the latter 0.12% per million years (GWYNN *et al.* 1987), indicating more rapid sequence evolution for the chloroplast gene. The question of whether these two genomes have dissimilar rates of sequence evolution should remain open.

CONCLUDING REMARKS

Variation within the chloroplast genome has now been examined in a small number of angiosperm genera. In studies where authors have determined the genetic nature of variation, both loss/gain of restriction sites and small (50–500 bp) insertion/deletion events have generally been observed. Inversions occur much less frequently and are all but unknown among species within genera (PALMER, JORGENSEN and THOMPSON 1985). *Zea* conforms to this pattern precisely, showing 24 restriction site mutations, 5 insertion/deletion events, and no inversions among its 4 species and the related genus *Tripsacum*.

Studies of the distribution of variation within the chloroplast genome have uniformly shown that the vast majority of restriction site variation occurs in the unique sequence regions, whereas the inverted repeats are highly conserved. This same distribution of variation appears in *Zea-Tripsacum*, where we observed zero, seven and 17 restriction site mutations in the inverted repeat and small and large unique sequence regions, respectively. Two studies (PALMER *et al.* 1983; SYTSMA and GOTTLIEB 1986) indicate that different provinces within the large unique sequence region have disparate rates of sequence evolution. Because we have not studied the entire chloroplast genome and because the amount of variation encountered in *Zea* was limited, this possibility could not be addressed. Insertion/deletion mutations have been found largely in the large unique sequence region; however, one study (SYTSMA and GOTTLIEB 1986) found several insertion/deletions in the portion of the inverted repeat proximal to the large unique sequence region. Four of five insertion/deletions in *Zea* occur in the large unique sequence region (DOEBLEY, MA and RENFROE 1987) and the fifth one in the portion of the inverted repeats adjacent to the large unique sequence region.

A phylogeny for *Zea* constructed with cpDNA data agrees precisely with assessments on relationships among the species based on morphology (DOEBLEY

1983) and isozymes (Doebley, Goodman and Stuber 1984). However, the pattern of variation among the subspecies of *Z. mays* did not conform to isozyme and morphological studies. This result is not unexpected as there was very little cpDNA variation within *Z. mays* and cpDNA analyses have not been shown taxonomically useful below the species level. The chloroplast genome analyses presented here are consistent with the hypothesis that maize is a domesticated form of teosinte, either *Z. mays* subsps. *parviglumis* or *mexicana*. This is an important result for these data are independent of previous isozyme and morphological assessments of the phylogenetic relationship between maize and its wild relatives.

Previous studies indicate that the chloroplast genome has a very conservative rate of sequence evolution (Palmer 1985). Other analyses suggest that *Z. mays* has a very active nuclear genome (Johns, Strommer and Freeling 1983) with degrees of isozyme variation in considerable excess of that typically found within angiosperm species (Doebley, Goodman and Stuber 1985). Comparison of the rate of sequence evolution for *Zea* cpDNA and the *Adh1* locus suggests that the latter evolves at over 10× the rate of the former. This and other results reported herein underscore the greater utility of isozyme studies for assaying variation among populations within species and of cpDNA analyses for analyzing variation among species and closely related genera.

McClean and Hansen (1986) have suggested that mtDNA and cpDNA have dissimilar rates of sequence evolution in *Solanum-Lycopersicon*. However, because their estimates of sequence evolution of mtDNA probably violate the assumption that all fragment differences result from base substitutions, this result should be viewed askance. Our estimate for sequence divergence between *Z. mays* and *Z. diploperennis* cpDNA (0.18%) is also less than the 0.67% estimate for mitochondrial genes and flanking regions for the same species pair. However, when mitochondrial and chloroplast gene sequence substitution rates, rather than estimates from restriction site data, are compared the opposite result is obtained (Gwynn *et al.* 1987). Whether overall rates of sequence evolution in plant mitochondrial and chloroplast genomes are dissimilar remains unclear.

We thank Peter Bretting and David Timothy for critical review of the manuscript, and Lawrence Bogarod, D. M. Lonsdale and Daryl Pring for providing clones. This research was supported in part by grants from the National Science Foundation (BSR-8508490) and Pioneer Hi-Bred, International of Johnston, Iowa.

LITERATURE CITED

Banks, J. A. and W. Birky, Jr., 1985 Chloroplast DNA diversity is low in a wild plant, *Lupinus texensis*. Proc. Natl. Acad. Sci. USA **82:** 6950–6954.

Beadle, G. W., 1939 Teosinte and the origin of maize. J. Hered. **30:** 245–247.

Beadle, G. W., 1972 The mystery of maize. Field Mus. Nat. Hist. Bull. **43:** 2–11.

Brown, W. M., M. George, Jr. and A. C. Wilson, 1979 Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA **76:** 1967–1971.

Clegg, M. T., J. R. Rawson and K. Thomas, 1984 Chloroplast DNA variation in pearl millet and related species. Genetics **106:** 449–461.

Dang, L. H. and D. R. Pring, 1986 A physical map of the sorghum chloroplast genome. Plant Mol. Biol. **6:** 119–123.

Doebley, J., 1983 The maize and teosinte male inflorescence: a numerical taxonomic study. Ann. Mo. Bot. Gard. **70:** 32–70.

Doebley, J. and H. H. Iltis, 1980. Taxonomy of *Zea*. I. Subgeneric classification with key to taxa. Am. J. Bot. **67:** 982–993.

Doebley, J., M. M. Goodman and C. W. Stuber, 1984 Isozyme variation in *Zea* (Gramineae). Syst. Bot. **9:** 203–218.

Doebley, J., M. M. Goodman and C. W. Stuber, 1985 Isozyme variation in the races of maize from Mexico. Am. J. Bot. **72:** 629–639.

Doebley, J., D. P. Ma and W. T. Renfroe, 1987 Insertion/deletion mutations in the *Zea* chloroplast genome. Curr. Genet. **11:** 617–624.

Galinat, W. C., 1971 The origin of maize. Annu. Rev. Genet. **5:** 447–478.

Galinat, W. C., 1983 The origin of maize as shown by key morphological traits of its ancestor, teosinte. Maydica **28:** 121–138.

Goodman, M. M. and C. W. Stuber, 1983 Races of maize. VI. Isozyme variation among races of maize in Bolivia. Maydica **28:** 169–187.

Gwynn, B. F., R. E. Dewey, R. R. Sederoff, D. H. Timothy, and C. S. Levings III, 1987 Sequence of the 18S-5S ribosomal gene region and the cytochrome oxidase II gene from mtDNA of *Zea diploperennis*. Theor. Appl. Genet. In press.

Iltis, H. H., 1983 From teosinte to maize: the catastrophic sexual transformation. Science **222:** 886–894.

Iltis, H. H. and J. F. Doebley, 1980 Taxonomy of *Zea* (Gramineae). II. Subspecific categories in the *Zea mays* complex and a generic synopsis. Am. J. Bot. **67:** 982–993.

Johns, M. A., J. N. Strommer and M. Freeling, 1983 Exceptionally high levels of restriction site polymorphism in DNA near the maize *Adh1* gene. Genetics **105:** 733–743.

Larrinua, I. M., K. M. T. Muskavitch, E. J. Gubbins and L. Bogorad, 1983 A detailed restriction endonuclease site map of the *Zea mays* plastid genome. Plant Mol. Biol. **2:** 129–140.

Maniatis, T., E. F. Fritsch and J. Sambrook, 1982 *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

McClean, P. E. and M. R. Hanson, 1986 Mitochondrial DNA sequence divergence among *Lycopersicon* and related *Solanum* species. Genetics **112:** 649–667.

Nei, M. and F. Tajima, 1983 Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics **105:** 207–217.

Palmer, J. D., 1985 Evolution of chloroplast and mitochondrial DNA in plants and algae. pp. 131–240. In: *Molecular Evolutionary Genetics*, Edited by R. J. MacIntyre. Plenum, New York.

Palmer, J. D., 1986 Isolation and structural analysis of chloroplast DNA. Methods Enzymol. **118:** 167–186.

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

Palmer, J. D., R. A. Jorgensen and W. F. Thompson, 1985 Chloroplast DNA variation and evolution in *Pisum*:

patterns of change and phylogenetic analysis. Genetics **109:** 195–213.

PALMER, J. D., C. R. SHIELDS, D. B. COHEN and T. J. ORTON, 1983 Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. Theor. Appl. Genet. **65:** 181–189.

REEVES, R. G. and P. C. MANGELSDORF, 1942 A proposed taxonomic change in the tribe Maydeae (family Gramineae). Am. J. Bot. **29:** 815–817.

SACHS, M. M., E. S. DENNIS, W. L. GERLACH and W. J. PEACOCK, 1986 Two alleles of maize *alcohol dehydrogenase 1* have 3′ structural and poly(A) addition polymorphisms. Genetics **113:** 449–467.

SCHAFFER, H. E. and R. R. SEDEROFF, 1981 Improved estimation of DNA fragment lengths from agarose gels. Anal. Biochem. **115:** 113–122.

SCHARDL, C. L., D. M. LONSDALE, D. R. PRING and K. R. ROSE, 1984 Linearization of maize mitochondrial chromosomes by recombination with linear episomes. Nature **310:** 292–296.

SEDEROFF, R. R., C. S. LEVINGS III, D. H. TIMOTHY and W. W. L. HU, 1981 Evolution of DNA sequence organization in mitochondrial genomes of *Zea*. Proc. Natl. Acad. Sci. USA **78:** 5953–5957.

SHAVER, D. L., 1962 A study of meiosis in perennial teosinte, in tetraploid maize and in their tetraploid hybrids. Caryologia **15:** 43–57.

SYTSMA, K. J. and B. A. SCHAAL, 1985 Phylogenetics of the *Lisianthus skinneri* (Gentianaceae) species complex in Panama utilizing DNA restriction fragment analysis. Evolution **39:** 594–608.

SYTSMA, K. J. and L. D. GOTTLIEB, 1986 Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). Evolution **40:** 1248–1262.

TEMPLETON, A. R., 1983 Phylogenetic inference from restriction endonuclease site maps with particular reference to the evolution of humans and apes. Evolution **37:** 221–244.

TIMOTHY, D. H., W. W. L. HU and C. S. LEVINGS III, 1981 Conservation of inverted repeats in chloroplast DNA from the genus *Zea*: teosinte. Crop Sci. **21:** 327–330.

TIMOTHY, D. H., C. S. LEVINGS III, D. R. PRING, M. F. CONDE and J. L. KERMICLE, 1979 Organelle DNA variation and systematic relationships in the genus *Zea*: teosinte. Proc. Natl. Acad. Sci. USA **76:** 4220–4224.

UPHOLT, W. B., 1977 Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res. **4:** 1257–1265.

WHITFELD, P. R. and W. BOTTOMLEY, 1983 Organization and structure of chloroplast genes. Annu. Rev. Plant. Physiol. **34:** 279–310.

WILKES, H. G., 1967 *Teosinte: The Closest Relative of Maize.* Bussey Institute, Harvard University, Cambridge, Mass.

ZIMMER, E. A. and K. J. NEWTON, 1982 A simple method for the isolation of high molecular weight DNA from individual maize seedlings and tissues. pp. 165–168. In: *Maize for Biological Research*, Edited by W. F. SHERIDAN, Plant Molecular Biology Assoc., Charlottesville, Va.

ZURAWSKI, G., M. T. CLEGG and A. H. D. BROWN, 1984 The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. Genetics **106:** 735–749.

Communicating editor: M. T. CLEGG