Zea Systematics: Ribosomal ITS Evidence

Edward S. Buckler IV and Timothy P. Holtsford
Division of Biological Sciences, University of Missouri

Ribosomal internal transcribed spacer (ITS) sequences were used to evaluate the phylogenetics of *Zea* and *Tripsacum.* Maximum likelihood and polymorphism parsimony were used for phylogenetic reconstructions. *Zea* ITS nucleotide diversity was high compared to other plant species, but approximately equivalent to other maize loci. Coalescence of ITS alleles was rapid relative to other nuclear loci; however, there was still much diversity within populations. *Zea* and *Tripsacum* form a clade clearly differentiated from all other Poaceae. Four *Zea* ITS pseudogenes were identified by phylogenetic position and nucleotide composition. The phylogenetic position of *Z. mays* ssp. *huehuetenangensis* was clearly established as basal to the other *Z. mays.* The ITS phylogeny disfavored a *Z. luxurians* and *Z. diploperennis* clade, which conflicted with some previous studies. The introgression of *Z. mays* alleles into *Z. perennis* and *Z. diploperennis* was also established. The ITS data indicated a near contemporary divergence of domesticated maize and its two closest wild relatives.

Introduction

Maize domestication is inexorably linked to the development of New World cultures, and maize remains a dominant food source. Maize's agricultural preeminence has led to *Zea*'s use as a model system for genetics, molecular biology, and systematics (review of *Zea* systematics in Doebley 1990*a*; Kellogg and Birchler 1993). Despite our wealth of information on *Zea,* the phylogenetics of the genus *Zea* and the subspecies of *Z. mays* are not entirely clear. We used nuclear ribosomal internal transcribed spacer (ITS) sequences to elucidate *Zea*'s phylogenetics and *Zea*'s position within the Poaceae. Systematic information from other *Zea* loci and organellar genomes was compared to provide a more complete and synthetic phylogenetic reconstruction.

Chloroplast restriction site, isozyme, and cytogenetic analyses (Kato Y. 1976; Doebley and Goodman 1984; Doebley, Renfroe, and Blanton 1987) established that maize (*Z. m.* ssp. *mays*) was domesticated from *Z. mays* populations in Central Mexico (Doebley 1990*a*), while other *Zea* and *Tripsacum* are more distantly related to maize. These molecular studies further refined the *Zea* taxonomy (Doebley 1990*b*), but only the chloroplast phylogeny was rooted, and it did not resolve the subspecies of *Z. mays.* Hence, our rooted nuclear ITS phylogeny was undertaken to provide a nuclear evaluation of the genus and to refine the infraspecific *Z. mays* relations.

ITS regions have rates of substitution that are useful for evaluating the generic and species level relationships in plants (Baldwin et al. 1995; Hsiao et al. 1995). Substitution rates are accelerated in Poaceae probably due to short generation times (Gaut et al. 1992), hence *Zea* ribosomal ITS sequences could elucidate species and even subspecies divergences. But several molecular evolution obstacles must be considered. First, markers must coalesce (alleles must share a common ancester) faster than speciation to provide a strong phylogenetic signal. Nuclear ribosomal (nrDNA) genes exist in large arrays of tandem repeats, which evolve together through gene conversion, unequal crossing over, and perhaps repeat amplification (Baldwin et al. 1995). Most plant ITS surveys suggest little variation within species (Baldwin et al. 1995), which suggests that the repeats are coalescing quickly, but this needs to be empirically tested.

A second potential problem is polymorphism within the arrays. Homogenization of nrDNA genes is not instantaneous, and individual plants may contain a mixture of older and more-derived alleles (for an extreme example see Ritland, Ritland, and Straus 1993). Recombination can also result in individual alleles with multiple lineages. We model this polymorphism at the infraspecific level with polymorphism parsimony (Felsenstein 1979), which accounts for a high probability of polymorphism persistence.

Third, ribosomal ITS substitution rates vary several-fold between the various taxa of *Zea* (Buckler and Holtsford 1996). The maximum-parsimony algorithm is biased with unequal rates of substitution, while maximum-likelihood (ML) methods are most successful under these conditions (Kuhner and Felsenstein 1994). Parsimony is also inconsistent in resolving multiple short interior branches (Takezaki and Nei 1994), as would be the case in this *Zea* phylogeny with many closely related alleles. When both rate differences and short interior branches are simulated, ML performs best and is robust to substitution model violations (Huelsenbeck 1995).

Finally, information from a single locus only represents the evolution and coalescence of that locus and not the whole genome. Coalescence theory shows that an individual locus can be positively misleading for recent divergences (Wu 1991; Avise 1994), especially for large populations such as maize (Gaut and Clegg 1993*a*). Hence the ribosomal ITS region results will be compared with other studies.

Here we estimated the phylogenetic position of *Zea* and *Tripsacum* among the Poaceae with an ITS phylogeny, while considering the effects of ribosomal poly-

Abbreviations: ITS, internal transcribed spacer; ML, maximum likelihood; nrDNA, nuclear ribosomal DNA.

Key words: nuclear rDNA internal transcribed spacer (ITS), pseudogenes, coalescence, introgression, nucleotide diversity, *Zea* and teosinte, maize domestication, phylogeny reconstruction.

Address for correspondence and reprints: Edward S. Buckler IV, Division of Biological Sciences, 105 Tucker Hall, University of Missouri, Columbia, Missouri 65211. E-mail: buckler@biosci.mbp.missouri.edu.

Mol. Biol. Evol. 13(4):612–622. 1996

morphism and substitution rate variation. We then compared the results of the ITS phylogeny with previous studies, and evaluated the validity of the current taxonomy and signs of introgression. Putative ITS pseudogenes were identified, and their phylogenetic positions were determined. ITS coalescence rate was compared to other *Zea* loci. In the accompanying paper (Buckler and Holtsford 1996), we characterized the pseudogenes, and compared the rates and the patterns of ITS substitution.

Materials and Methods

Sampling and DNA Manipulation

Samples were chosen to represent all species and subspecies of *Zea* and four representative taxa of *Tripsacum* (table 1).

DNA was extracted from leaf and kernel tissue. Leaf tissue was extracted using a CTAB method (Bult, Källersjö, and Suh 1992) without the agarose purification. CTAB extraction of kernels produced unamplifiable DNA, therefore the following SDS procedure was developed. Kernels were smashed with a sledge hammer under sterile conditions and extracted with 800 μL of buffer (1.4 M NaCl, 20 mM Na_2EDTA, 100 mM Tris [pH 8.0], 2% SDS [w/v], and 0.2% B-mercaptoethanol [v/v]) for 1 hour at 70°C and centrifuged to pellet particulates. The supernatant was then precipitated with one tenth volume of 3 M NaOAc and two volumes of ethanol, washed with 80% ethanol, and dried. The pellet was resuspended in TE (10 mM Tris [pH 7.5] and 1 mM EDTA), RNAsed, and precipitated with one fifth volume of 2 M NaCl and two volumes of ethanol for 1 hour at −20°C, washed with 80% ethanol, dried, and resuspended in TE. The resulting genomic DNA from the CTAB leaf extractions and SDS kernel extractions was further purified by the desalting procedure of the QIAEX gel extraction kit (a silicagel extraction method from Qiagen, Chatsworth, Calif.).

A lambda clone of the maize ribosomal repeat was provided by Dr. B. Burr (Brookhaven Nat. Lab.). The ITS region was subcloned into pUC13 and sequenced. This first sequence allowed us to refine amplification and sequencing conditions for the ITS region of *Zea.*

The high GC content and secondary structure in the ITS region necessitated strong denaturing conditions for PCR amplification. We used two methods to circumvent these PCR problems. Both methods used 50 μL reactions with 1 ng of genomic DNA, 0.4 μM of each primer, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 1.0% Triton X-100, 2 mM $MgCl_2$, and 0.5 U of *Taq* polymerase. Amplification profiles always included a 5-minute 94°C initial denaturation. In the first method, 7-deaza-2′-deoxyguanosine triphoshate (c^7dGTP) PCR was used to reduce base pair stacking energy (Innis 1990). Primers ITS5 and 26sr (table 2), 37.5 μM of c^7dGTP, 12.5 μM of GTP, and 50 μM of dATP, dCTP, and dTTP were used in a 35-cycle reaction with a PCR profile of 15 sec at 98°C (denaturation), 60 sec at 56°C (annealing), and 90 sec at 72°C (extension). Since the cloning and restriction of c^7dGTP PCR fragments is inefficient, 3 μL of PCR product was reamplified for 20 cycles with 200 μM of dATP, dCTP, dGTP, and dTTP. Primers Fred and Barney, which are internal to ITS5 and 26sr, have *Bam*HI and *Eco*RI restriction sites, respectively. The PCR reamplification profile consisted of 15 sec at 98°C, 60 sec at 61°C, and 90 sec at 72°C.

The second method was a much more efficient single amplification using dimethylsulfoxide (DMSO). DMSO facilitates amplification by reducing strand reannealing (Varadaraj and Skinner 1994). The PCR reactions included 10% DMSO, 200 μM of each dNTP, and primers Fred and Barney; the 35 cycle profile was 30 sec at 94°C, 30 sec at 61°C, and 90 sec at 75°C.

All PCR products were purified by using the GENECLEAN II kit (BIO 101, La Jolla, Calif.). The purified PCR products were restricted using *Bam*HI and *Eco*RI, repurified with GENECLEAN II, and ligated into a pUC13 vector. Clones were sequenced using chain-termination sequencing with a Sequenase c^7dGTP sequencing kit (USB, Cleveland, Ohio) with the addition of 15% DMSO to the annealing mixture (Sun, Hegamyer, and Colburn 1993). Complete bidirectional sequences were produced with Tom, Jerry, and pUC13 primers.

Sequences were aligned using Lasergene's (DNASTAR Inc., Madison, Wis.) clustal alignment method and refined by eye.

Tree Reconstruction

To accommodate base composition bias, unequal rates of substitution, and ITS polymorphism, we used several tree reconstruction methods. Nucleotide composition bias was removed by calculating the distance between two taxa with the LogDet transformation (natural logarithm of the determinant of the divergence matrix, Lockhart et al. 1994). A LogDet-transformed distance matrix was produced for Poaceae, and a tree was constructed by neighbor-joining (Saitou and Nei 1987).

We estimated *Zea* and Poaceae phylogenies using ML (DNAML of Felsenstein 1981), as ML deals well with unequal rates of substitution between taxa. A computationally faster program was also used for most of the searches (fastDNAml of Olsen et al. 1994). The best tree was searched for using global branch swapping and 10 random additions of the taxa. The four ITS regions with informative insertions and deletions (indels) were coded as bases in some searches; because DNAML and fastDNAml only model changes between base states, we converted indels to bases while maintaining the transition/transversion ratio. A 74-taxa run took approximately 60 CPU hours on an IBM RS/6000. Statistical significance of branches and topologies was evaluated with the robust Kishino and Hasegawa (1989) test, as bootstrapping was not computationally possible. Recoding indels partially compromises the ML model so that likelihood scores will not be exact, but recoding should not greatly affect the relative likelihood tests of the alternate hypotheses (table 5).

Polymorphism parsimony was used to make a tree which accounts for the polymorphism among ribosomal alleles among the *Z. mays* subspecies. A polymorphism

Table 1
Accession List for Samples Used in This Study

Species (Race)	Locality	Source	Accession	Clones #[a]
Zea perennis	Jalisco, Mexico	Birchler	Guzman 1524	1–4
	Jalisco, Mexico	USDA	Ames 21881	5, 6
Zea diploperennis	Las Joyas, Jalisco, Mexico	Doebley	M001	7–11
Zea luxurians	Chiquimula, Guatemala	Doebley	M018	12, 13
	Chinandega, Nicaragua	Doebley	M111	14–17
Zea mays ssp. *mays* (*Pepitilla*)	Guerrero, Mexico	USDA	Ames 8212	18
(Maize Ancho)	Guerrero, Mexico	USDA	Ames 15820	19, 20
(Gaspe Flint)	North Dakota, USA	USDA	PI 213803	21–23
(Longfellow)	Ontario, Canada	USDA	PI 214195	24–26
(Argentine Pop)	Argentina	USDA	PI 217404	27, 28
(Nal Tel)	Yucatan, Mexico	USDA	PI 479091	29, 30
(Cónico)	Puebla, Mexico	USDA	PI 515436	31–33†
(Hopi)	Arizona, USA	USDA	PI 213733	34†
(Lambda Clone)		Burr, BNL	Lambda clone	35
Zea mays ssp. *parviglumis*	El Rodeo, Jalisco, Mexico	Doebley	M046	36†, 37†, 38, 39
	Teloloapan Hwy., Guer., Mexico	Doebley	M106	40, 41
	Iguala-Telo. reg., Guer., Mexico	USDA	PI 331783	42–44†
	Mazatlán, Guerrero, Mexico	USDA	PI 384061	45, 46
	Tingambato, Michoacan, Mexico	USDA	PI 331788	47, 48
Zea mays ssp. *mexicana*	Chalco, Mexico DF, Mexico	USDA	Ames 8083	49, 50
	Texcoco, Mexico DF, Mexico	Doebley	M092	51, 52
	Nobogame, Chihuahua, Mexico	Doebley	M075	53†, 54†, 55–57
	Oaxaca, Mexico	USDA	PI 384060	58†, 59†, 60, 61
Zea mays ssp. *huehuetenangensis*	San Antonio, Huehuetenango, Guatemala	Doebley	M031	62, 63
	Santa Ana, Huehuetenango, Guatemala	Doebley	M033	64–66
Tripsacum dactyloides	USA	Blakey	WW1582	82–85
Tripsacum australe	Matto Grosso, Brazil	Doebley	Timothy 68-67-1	86–88
Tripsacum maizar	Jalisco, Mexico	Doebley	de Wet 3721	89, 90
Tripsacum laxum	Veracruz, Mexico	Doebley	de Wet 3766	91, 92
Coix sp.				93

NOTE.—GenBank accession numbers are: U46583–U46660

[a] Clones amplified with c^{7}dGTP are indicated by †, all other clones were amplified with DMSO.

OTHER SEQUENCES USED IN THIS STUDY:

Species	GenBank # or Source	Sequence #
Clestachne sorghoides	CSU04790	67
Sorghum bicolor	SBU04789	68
Sorghum matarankense	SMU04792	69
Sorghum puppureo-sericeum	SPU04793	70
Zea mays?	ZMU04796	71
Avena longiglumis	AL58SRDNX	72
Hordeum vulgare	HV58SRDNX	73
Bothriochloa insculpta	Spies and Kellogg	74
Cymbopogon plurinodis	Spies and Kellogg	75
Elionurus mutica	Spies and Kellogg	76
Hyparrhenia anamesa	Spies and Kellogg	77
Oryza sativa	RICRGSBHA	78
Triticum aestivum	TA58SRDNX	79
Triticum speltoides	TS58SRDNX	80
Secale montanum	SM58SRDNX	81

Table 2
Primers Used for PCR and Sequencing

Primer	Sequence (5′ to 3′)	Position
ITS5	GGAAGGAGAAGTCGTAACAAGG	16S
26sr	CCGGTTCGCTCGCCGTTACT	26S
Fred	GTAGGGGATCCTGCGGAAGGATCA	16S
Barney	GCGAATTCAACTCAGCGGGTAGTCC	26S
Tom	CTTGCGTTCAAAGACTCGATGGTTC	5.8S
Jerry	GAACCATCGAGTCTTTGAACGCAAG	5.8S

ML model would be optimal (Felsenstein 1979) but computationally impossible for this data set. Therefore we developed a polymorphism parsimony step matrix for maximum parsimony. We allowed each polymorphic site to have up to two character states at any one time. Substitutions to a polymorphic state were weighted 8, maintenance of polymorphism was weighted 1, and loss of polymorphism was weighted 1. The polymorphism phylogeny was estimated with PAUP 3.1.1 (Swofford 1993) using steepest descent without MULPARS for 100 random additions.

DNA distances between taxa were calculated with the ML nucleotide substitution model of DNADIST

Table 3
Nucleotide Diversity (×100) and its Standard Deviation over the Stochastic Process (in Parentheses)

Taxon	ITS1	5.8S	ITS2	*Adh1*[a]	*Adh2*[a]
Zea[b]	2.64 (1.42)	1.26 (0.79)	2.26 (1.23)	2.28 (1.27)	2.64 (1.51)
Z. m. ssp. *mays*	2.23 (1.29)	1.37 (0.89)	1.97 (1.14)	2.14 (1.26)	3.26 (2.06)
Z. m. ssp. *parviglumis*	2.50 (1.46)	1.12 (0.78)	1.80 (1.08)		4.19 (4.41)
Z. m. ssp. *mexicana*	1.99 (1.20)	1.00 (0.72)	1.57 (0.97)		0.98 (1.19)
Z. m. ssp. *huehuetenangensis*	1.80 (1.27)	1.46 (1.11)	2.15 (1.47)		
Z. luxurians	2.35 (1.53)	0.41 (0.41)	1.97 (1.30)		2.53 (2.75)
Z. diploperennis and *Z. perennis*	2.16 (1.35)	1.83 (1.21)	1.80 (1.14)		1.52 (1.73)
Z. diploperennis and *Z. perennis*[c]	2.49 (1.47)	1.89 (1.20)	2.17 (1.29)		
Z. mays pseudogenes	6.98 (4.75)	3.27 (2.38)	11.31 (7.56)		

[a] Only silent sites were included for *Adh1* (Gaut and Clegg 1993*a*) and *Adh2* (Goloubinoff, Pääbo, and Wilson 1993).
[b] Comparisons involving all *Zea* are not necessarily equivalent between loci, as sampling intensity among the various *Zea* taxa was not equivalent.
[c] The introgressed alleles were included in these values.

with the deletions recoded as above (Felsenstein 1989). Average substitution distances were calculated for the alleles between taxonomic groups.

Nucleotide Diversity Estimates and Coalescence

Nucleotide diversity ($\hat{\pi}$) was estimated separately for ITS1, 5.8S, and ITS2 regions (Nei 1987) (table 3). The within- and between-population components of nucleotide diversity were estimated by the N_{st} statistic for the entire ITS region (Lynch and Crease 1990) (table 4). We used Monte Carlo simulations of the K_s statistic to estimate the significance of the differentiation among populations (Hudson, Boos, and Kaplan 1992). K_s is a weighted measure of differentiation, which is designed to maximize the power of the test. Significant differentiation between populations indicated the ITS alleles were coalescing at that taxonomic level. For some diversity analyses the races of maize were grouped into geographical complexes as follows: Northern Flints (clones #21–26, table 1), Mexico (#18–20, 29–33), and Argentine Popcorn (#27, 28). Nucleotide diversity was also compared with silent sites at *Adh1* and *Adh2* (Gaut and Clegg 1993*a*; Goloubinoff, Pääbo, and Wilson 1993).

Table 4
ITS Genetic Differentiation of *Z. mays* Subspecies

	N_{st}	P[a]
Among subspecies:		
ssp. *parviglumis, mexicana, mays*	0.055	0.0000
ssp. *parviglumis, mexicana, mays*[b]	0.082	0.0000
ssp. *parviglumis, mays*	0.034	0.0010
ssp. *mays, mexicana*[b]	0.151	0.0000
ssp. *mexicana, parviglumis*[b]	0.057	0.0052
Within subspecies:		
ssp. *mays* races	0.058	0.0186
ssp. *mays* racial complex	0.032	0.0084
ssp. *parviglumis* populations	0.025	0.2872
ssp. *mexicana* populations	0.181	0.0046
ssp. *mexicana* populations[b]	0.152	0.0396

[a] P value for test of differentiation as measured by the K_s statistic (Hudson, Boos, and Kaplan 1992).
[b] Nobogame population excluded from *Z. mays* ssp. *mexicana*.

Results

Sequence Analyses

Seventy-eight ribosomal clones were sequenced from *Zea* and *Tripsacum* (table 1). *Zea* ITS1, 5.8S, and ITS2 regions have average GC contents of 70.4%, 56.3%, and 73.3% respectively. The two denaturing amplification methods seemed to have different thermal stability preferences, although both produced full-length product efficiently. Thirteen clones were sequenced from PCR amplifications using c^7dGTP. Four of the 13 clones were low (62% to 65%) GC content pseudogenes (Buckler and Holtsford 1996), while the remaining clones showed a high GC base composition similar to those recovered from DMSO amplification. Since c^7dGTP only decreases steric hindrances in amplification once incorporated, in some reactions PCR selection (Wagner et al. 1994) probably favored the amplification of low-GC-content alleles during the initial cycles. DMSO amplification did not yield low GC content pseudogenes, since DMSO decreases the energy of steric hindrances throughout the amplification. DMSO amplification of the accessions producing pseudogenes and restriction with diagnostic enzymes suggested the pseudogenes were a very rare component of the genome.

A published maize ITS sequence from a *Sorghum* study (ZMU04796 of Sun et al. 1994) is highly (12% to 13%) and significantly (table 5) diverged from our *Zea* and *Tripsacum* sequences, including the lambda clone (fig. 1). We think that the maize sequence from the *Sorghum* study is actually a *Sorghum* contaminant. Considering the difficulty of amplifying *Zea* ITS, this result is not surprising.

We found length variation among the *Zea* alleles. The ITS1 varied from 211 to 217 bp, the 5.8S was 164 bp (one allele had a one base deletion), and the ITS2 varied from 216 to 220 bp except for one clone (#32), which had a 36-bp deletion. All four regions of length variation shared by multiple alleles were repeat sequences. Three follow strings of C's and G's, and the fourth region involves a GTT repeat (most *Zea mays* alleles have three repeats while other *Zea* have two repeats).

Table 5
Summary of Alternate Hypotheses Tested with the Kishino and Hasegawa Method

	Alternate Hypotheses	Indels Recoded[a]			Indels Missing		
		ΔLnL[b]	SD[c]	*P*[d]	ΔLnL[b]	SD[c]	*P*[d]
Zea and *Tripsacum* . . .	ML tree (figs. 1 and 2)	−4353.88[e]			−3919.34[e]		
Zea topology	Zl and Zd form a monophyletic clade	−18.28	9.37	0.05	−9.25	6.50	0.15
	Closest Zd not basal to Zm and Zl	−5.33	5.95	0.36	*		
	Zl not diverged from Zmh	−15.85	9.05	0.08	−19.72	10.07	0.05
	Zmmx-Nobogame not diverged from the Zm polytomy	−16.44	9.37	0.08	*		
	Zmh not diverged from closest Zm	−7.42	7.15	0.30	*		
	Zmh not diverged from Zm polytomy	−36.15	17.45	0.04	−8.87	8.02	0.27
Pseudogenes	All pseudogenes are not basal to Zea	−16.29	8.42	0.05	−13.69	7.91	0.08
	Pseudogenes are not basal to Central Mexican Zm	−49.71	14.21	0.00	−39.85	12.96	0.00
	Tripsacum is not basal to pseudogenes	−101.81	18.98	0.00	−80.90	19.96	0.00
Introgression	Introgressed Zd&p cluster with other Zd&p	−106.09	22.82	0.00	−72.66	19.81	0.00
Poaceae ML tree	(fig. 1)				−6139.10[e]		
	The branch to *Zea* and *Tripsacum* has length 0				−19.85	8.33	0.02
	Coix is basal to *Zea* and *Tripsacum*				−34.89	14.18	0.01
	ZMU04796 is basal to Zm				−189.90	27.94	0.00

Note.—Zl = *Zea luxurians*, Zd = *Z. diploperennis*, Zm = *Z. mays*, Zmh = *Z. mays* ssp. *heuheutenangensis*, Zmmx = *Z. m.* ssp. *mexicana*, Zd&p = *Z. diploperennis* and *Z. perennis*.

[a] Recording of indels compromises the likelihood model, therefore the *P* values are approximations.

[b] ΔLnL is the difference in natural log likelihoods scores between the best tree and the alternate hypothesis.

[c] The standard deviation of the ΔLnL.

[d] The probability of the alternate hypothesis.

[e] The observed LnL for the best tree.

* ML topology excluding indels agrees with the alternate hypothesis.

Nucleotide Diversity and Coalescence

The nucleotide diversity ($\hat{\pi}$) varied among subunits of the ITS region (highest to lowest diversity): ITS1 > ITS2 > 5.8S (table 3). The nucleotide diversity of ITS1 and ITS2 were comparable to the synonymous levels of maize nuclear *Adh1* and *Adh2* loci. The ITS nucleotide diversity was also similar to polymorphism levels at four maize anonymous single-copy nuclear loci ($0.040 \geq \Theta \geq 0.006$, Shattuck-Eidens et al. 1990). Among the *Z. mays* subspecies, the vast majority of the diversity was between individual alleles, but there was significant differentiation between subspecies, populations, and races (table 4). This genetic subdivision indicates that the ribosomal alleles have partially coalesced even at the level of populations.

Poaceae Trees

For the Poaceae, large divergences made alignment of indels less accurate; therefore, indels were excluded from the ML tree. The ML tree indicated that *Tripsacum*

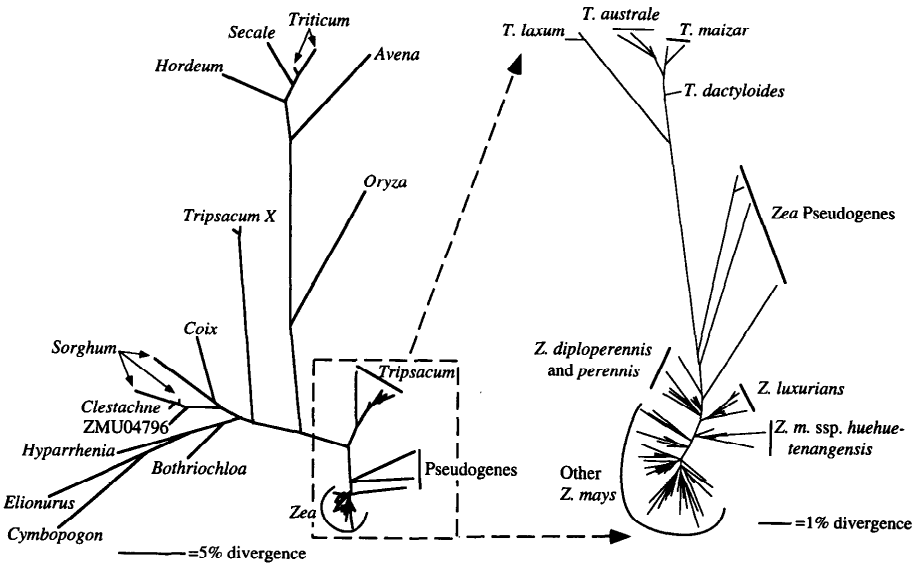

Fig. 1.—Tree on the left is the ITS maximum-likelihood tree for Poaceae with indels counted as missing. The tree on the right shows the maximum-likelihood tree for *Tripsacum* and *Zea* with the inclusion of indels.

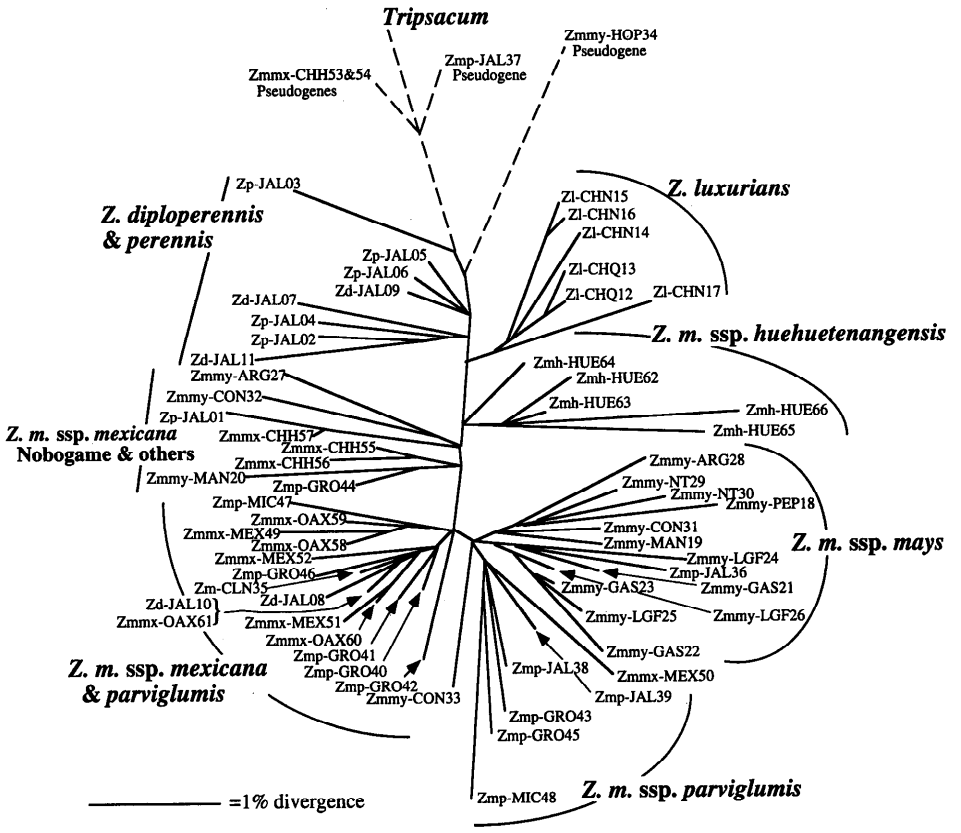

FIG. 2.—Maximum-likelihood tree with indels included. Dashed lines are unscaled branches. Zd—*Z. diploperennis,* Zp—*Z. perennis,* Zl—*Z. luxurians,* Zmh—*Z. m.* ssp. *huehuetenangensis,* Zmmx—*Z. m.* ssp. *mexicana,* Zmp—*Z. m.* ssp. *parviglumis,* Zmmy—*Z. m.* ssp. *mays.* Allele names: species abbreviation, state or race, clone number; e.g., Zmp-GRO45 = Z. *m.* ssp. *parviglumis,* Guerrero, clone 45.

was the closest relative to *Zea* (fig. 1). *Coix,* an Old World genus, is often thought to be closely related to *Zea* (Kellogg and Birchler 1993), but the tree significantly indicated affinities with *Sorghum* and the awned Andropogoneae and not *Zea* nor *Tripsacum* (table 5). Two *Tripsacum dactyloides* alleles were very different from other *Tripsacum* alleles, which may reflect an allotetraploid origin for *Tripsacum* (Galinat, Chaganti, and Hager 1964). This *Tripsacum* question will be pursued in future studies.

Zea Trees

LogDet, ML, DNA distance, and polymorphism parsimony were used to determine a phylogeny for the alleles of *Zea.* Each method had strengths and weaknesses in modeling the biology of ribosomal evolution. Trees from all reconstruction methods were consistent with the following: When *Tripsacum* was used as an outgroup, the pseudogenes were basal to the other *Zea* alleles. *Zea m.* ssp. *huehuetenangensis* was basal to the Central Mexican *Z. mays* clade. Two *Z. diploperennis* (#8 and 10) and one *Z. perennis* (#1) alleles were always associated with the *Z. mays* clade. The other *Z. perennis* and *Z. diploperennis* alleles were always together, although they did not always form a monophyletic clade. *Zea luxurians* alleles were always grouped. *Zea luxurians, Z. diploperennis,* and *Z. perennis* were always basal and clearly differentiated from the monophyletic *Z. mays* clade.

The LogDet neighbor-joining reconstruction indicated that the pseudogenes were basal to the other *Zea* alleles but more closely related to *Zea* than *Tripsacum.* This congruence between the ML and LogDet trees shows that variation in nucleotide composition (most apparent between pseudogenes and normal alleles), did not affect branch placement.

Since indels might be less homoplasious than other types of substitutions (Baldwin et al. 1995; Buckler and Holtsford 1996), we performed the ML analyses of *Zea* both with and without indels (figs. 1 and 2). Species level structure was compatible but not identical for both analyses. Figure 2 describes the relationships found when indels were included, while differences with indel exclusion were as follows: The *Z. diploperennis/perennis* clade, the *Z. luxurians* clade and the *Z. mays* clade formed a basal polytomy. The *Z. mays* were well defined, and the subspecies diverged in a pattern very similar to that of indel-excluded polymorphism parsimony (fig. 3). The most basal alleles of *Z. mays* included all of *Z. m.* ssp. *huehuetenangensis,* four *Z. m.* ssp. *parviglumis* alleles, and one *Z. m.* ssp. *mexicana.* The rest of the *Z. mays* alleles branch from one derived node. There were two *Z. m.* ssp. *mays* clades, Northern Flint and Mexican, branching from this node. Another monophyletic clade contains most of the *Z. m.* ssp. *mexicana* and *Z. m.* ssp. *parviglumis* alleles. This suggested that *Z. m.* ssp. *mays* divergence was basal to *Z. m.* ssp. *parviglumis* and *Z. m.* ssp. *mexicana* divergence.

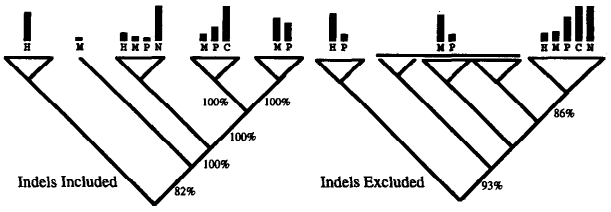

FIG. 3.—Polymorphism parsimony majority consensus trees. Histograms represent the frequency of a taxon's alleles within each clade. Percentages are the frequency of equally parsimonious trees supporting that clade, not bootstrapping frequencies. When indels were excluded, the *Z. m.*-ssp.-*mays*-dominated clades either formed the grouping indicated or formed a single monphyletic clade. H = *Z. m.* ssp. *huehuetenangensis,* M = *Z. m.* ssp. *mays,* P = *Z. m.* ssp. *parviglumis,* N = *Z. m.* ssp. *mexicana* Nobogame, C = *Z. m.* ssp. *mexicana* (excluding Nobogame).

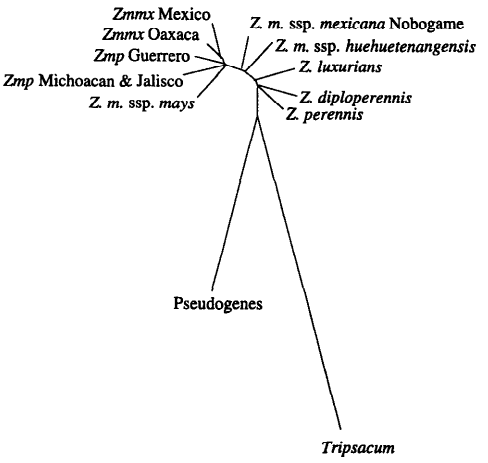

FIG. 4.—Fitch tree of average DNA substitution distance between taxonomic groups (indels included). Zmp = *Z. m.* ssp. *parviglumis,* Zmmx = *Z. m.* ssp. *mexicana.*

Polymorphism parsimony was used to model the maintenance of polymorphism of ribosomal repeats within *Z. mays* (fig. 3). Polymorphism parsimony with indels excluded suggested that *Z. m.* ssp. *mays* was basal to *Z. m.* ssp. *parviglumis* and *mexicana,* while the inclusion of indels suggested a polytomy. Within *Z. m.* ssp. *mays,* the Central Mexican races were separate from the Northern Flints (not shown), as was the case in our ML trees.

The Fitch tree of average substitution distances between groups of taxa provided a good summary of the ribosomal data (fig. 4). *Zea m.* ssp. *mays* was slightly basal to all the Central Mexican *Z. mays* except Nobogame, while *Z. luxurians* and *Z. mays* formed a clade.

Determining significance by bootstrapping was computationally impossible for this analysis (we estimated it would take 5,800 hours on an IBM RS/6000 computer for the ML tree). Branches discussed in the ML tree were significant according to the likelihood-ratio test ($P < 0.01$), but these tests err toward oversignificance (Felsenstein 1989). The Kishino and Hasegawa (1989) log-likelihood test is a more robust test and indicated several significant conclusions (table 5). The closest *Z. diploperennis* and *Z. perennis* alleles were not significantly basal to *Z. luxurians. Zea luxurians* did not form a clade with *Z. diploperennis* when indels were included; the rejection of this hypothesis could have resulted from this topology being rejected or from a failure of alleles to coalesce. *Zea luxurians* was basal to *Z. m.* ssp. *huehuetenangensis. Zea m.* ssp. *mexicana* Nobogame had a tendency to be basal to Central Mexican *Z. mays. Zea m.* ssp. *huehuetenangensis* was basal to the majority of Central Mexican *Z. mays* alleles. Three of the four pseudogenes were basal to all of *Zea,* while all were basal to Central Mexican *Z. mays.* The *Z. diploperennis* and *Z. perennis* alleles (#2–7, 9, 11) were significantly basal to three other *Z. diploperennis* and *Z. perennis* alleles (#1, 8, 10). There is little significant divergence within the Central Mexican *Z. mays.*

Discussion

Nucleotide Diversity and Coalescence

How much of the intraspecific diversity is a PCR artifact? Considering the conditions used here, the error rate should be less than 0.2 bases per ITS clone or 27-fold smaller than the observed diversity (Gelfand and White 1990; Kwiatowski et al. 1991). Although the Chester and Marshak (1993) study did not directly address *Taq* fidelity in the presence of DMSO, there is no suggestion that DMSO elevates errors.

Most plants appear to have complete homogenization of repeats within an individual and even within species (Baldwin et al. 1995) which is suggestive of rapid ITS coalescence. However some plants exhibit individual and intraspecific diversity (Ritland, Ritland, and Straus 1993; Sun et al. 1994; Baldwin et al. 1995). Since nucleotide diversity is directly related to effective population size and mutation rates ($\pi = 4N\mu$, Nei 1987), which are lineage-specific parameters, comparison of ITS nucleotide diversity must be relative to other *Zea* loci. Maize is known to have a much larger nucleotide (Shattuck-Eidens et al. 1990; Gaut and Clegg 1993*b*) and isozyme (Doebley and Goodman 1984) diversity than other plants, which probably reflects a large effective population size and an elevated substitution rate. Maize ITS nucleotide diversity was roughly equivalent to other maize nuclear loci (table 3).

Despite high nucleotide diversity, the genetic subdivision of *Z. mays* subspecies, populations, and races suggested that a few sites in the ITS were coalescing rapidly. Since ribosomal regions probably recombine frequently (Sanderson and Doyle 1992), coalescence can be described for individual sites rather than for an entire region. The domestication of maize cannot be older than the significant human migrations to the New World (roughly 15000 BP), while the origins of maize racial complexes are probably not older than 5500 BP (Weaver 1993, pp.1–24). While the maize *Adh1* and *Adh2* alleles have suggested a very old coalescence (*Adh1*: average = 1.3 MYA, shortest = 340,000 years ago) and there is little suggestion of coalescence even among the species of *Zea* ($P > 0.05$ for genetic subdivision [K_s] of species in *Adh1* and *Adh2,* Buckler, unpublished data; Gaut and Clegg 1993*a*; Goloubinoff, Pääbo, and Wilson 1993). This argues that some of the ITS sites coalesced

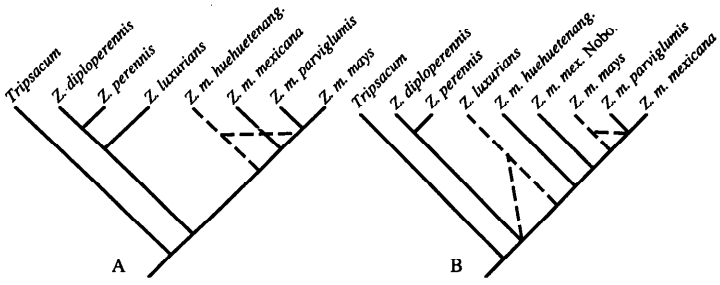

FIG. 5.—Tree A represents the reconstruction of Doebley (Doebley and Goodman 1984; Doebley, Renfroe, and Blanton 1987; Doebley 1990*b*), while tree B represents a summary of the ITS data. Dashed lines indicate the alternate placements of branches. Branch lengths are not scaled.

on the order of $1.3 \times 10^6/5{,}500 \approx 200$ times faster than the entire *Adh1* locus, suggesting that ribosomal concerted evolution and selection reduced the effective population size for a few ITS sites 200-fold. The entire ITS coalesces in a longer time, but how much longer is not apparent without a time-calibrated ITS mutation rate.

Poaceae Phylogeny

Zea and *Tripsacum* formed a monophyletic group which was distinct from other analyzed Andropogoneae, although members of the Rottboelliinae were not examined. The species of *Zea* have evolved very recently in comparison to *Zea*'s divergence from *Tripsacum.* The subtribe Maydeae, which are the monoecious species of Poaceae, normally include *Coix, Zea, Tripsacum,* and *Polytoca.* However, the ITS data suggested that *Coix* was no more closely related to *Zea* than it is to *Sorghum.* Supporting this view, Southern hybridization with *Zea* knob sequences shows a close relationship between *Zea* and *Tripsacum,* while homology with *Coix* and *Sorghum* sequences was not detectable even under low-stringency conditions (Dennis and Peacock 1984). The ITS and knob data argue for multiple origins of monoecy in Poaceae in contrast with morphological data (Kellogg and Birchler 1993). A molecular analysis of all the Maydeae species and many of the New and Old World Andropogoneae should clarify this issue.

Zea Phylogeny

Four putative pseudogenes, which came from three different *Z. mays* accessions, were clearly basal to most *Zea* spp. and to all *Z. mays.* The two pseudogenes from the same accession were closely related, but the other pseudogenes were not monophyletic. They appear to be evolutionary relicts, which diverged paralogously from the active rDNA genes before the divergence of the modern *Zea.* Normal *Zea* alleles were also amplified from these accessions; therefore, we believe that PCR amplification methods enriched for these rare pseudogenes. The *Zea* pseudogenes have undergone many substitutions relative to normal alleles, were heavily deaminated at methylated cytosines, and do not appear to have undergone recombination with other alleles (Buckler and Holtsford 1996). A ribosomal pseudogene has also recently been encountered in *Drosophila* promoter sequences (Linares, Bowen, and Dover 1994). The *Zea* pseudogenes are unlike the multiple ITS types found in *Mimulus* and Winteraceae, which were prevalent and closely related to each other within type (Ritland, Ritland, and Straus 1993; Suh et al. 1993). These *Mimulus* ITS types might represent alleles within different arrays, while the pseudogenes are probably inactive repeats, perhaps in the terminal regions of the *Zea* rDNA array (Linares, Bowen, and Dover 1994; Buckler and Holtsford 1996).

The ribosomal ITS data had a couple advantages for reconstructing the *Zea* phylogeny. First, the pseudogenes provided an excellent outgroup for *Zea,* which could not be equaled by extant taxa. Second, there was a high level of diversity even among the subspecies, and some of this diversity appears to be coalescing rapidly. The ML tree including indels forms the basis of this discussion, since this tree was statistically tested and partially modeled the polymorphism (summary tree B, fig. 5). In the ITS region, indels might be less likely than other substitutions to be homoplasious (Baldwin et al. 1995; Buckler and Holtsford 1996), because indels are likely to result in a series of ITS structural changes and hence are probably more constrained. There were a couple of points of disagreement between the ITS trees and previous phylogenies based on chloroplast and isozyme evidence (tree A, fig. 5). Rate variability among lineages, the use of plesiomorphic characters, and rooting problems may account for some of the discrepancies, while different coalescence patterns might explain other differences.

The ITS data indicate that *Zea perennis* and *Z. diploperennis* were essentially indistinguishable, probably reflecting a recent divergence (see below). They also indicated that the geographically distant *Z. luxurians* populations were monophyletic.

Ribosomal ITS data favored *Z. diploperennis* as basal to or in a polytomy with *Z. luxurians* and *Z. mays* clades. Chloroplast restriction site data significantly supported a *Z. luxurians, Z. diploperennis,* and *Z. perennis* clade (Doebley, Renfroe, and Blanton 1987), while the ITS data significantly rejected this clade (table 5). A likely explanation is a difference in the coalescence of ancestral polymorphism between these two loci. A *Zea* and *Tripsacum* restriction site study of the entire rDNA repeat found variability among a small sample of *Zea* in the IGS region (Zimmer, Jupe, and Walbot 1988); one site favors a *Z. diploperennis, Z. perennis,* and *Z. lux-*

urians clade, while the other favors a *Z. mays* and *Z. luxurians* clade. Restriction analysis of a small sample of 5S DNA on chromosome 2 found three sites supporting a *Z. mays* and *Z. luxurians* clade and none supporting a *Z. luxurians* and *Z. diploperennis* clade (Zimmer, Jupe, and Walbot 1988). Morphological data favors a *Z. luxurians* and *Z. diploperennis* clade, but only based on several plesiomorphic characters (Doebley 1983), which might have been retained in the more slowly evolving *Z. luxurians* and *Z. diploperennis* (Buckler and Holtsford 1996). Mitochondrial interaction studies suggest more compatibility between *Z. mays* and *Z. luxurians* than between *Z. mays* and *Z. diploperennis* (Allen, Emenhiser, and Kermicle 1989). *Magellan* retroposon elements have been sequenced from various *Zea,* and the average distances suggest a polytomy between *Z. mays, Z. luxurians,* and *Z. diploperennis* (Purugganan and Wessler 1994). However, the placement of *Z. luxurians* can only be rigorously evaluated with more loci (with outgroups), such as the ITS and the chloroplast data sets. If the divergence between the *Z. luxurians, Z. diploperennis,* and *Z. mays* lineages were nearly concurrent, then data from many loci will be needed to evaluate this question with a statistically sound analysis (Kishino and Hasegawa 1989; Wu 1991). The large effective population size of *Zea* will retard coalescence and complicate the resolution (Gaut and Clegg 1993*a*).

Z. mays Subspecies

For the first time, the phylogenetic position of *Z. m.* ssp. *huehuetenangensis* was clearly defined as being the basal taxon of *Z. mays*. This phylogenetic position is compatible with unrooted isozyme and knob studies (Kato Y. 1976; Doebley and Goodman 1984), and the unresolved polytomy of the chloroplast study (Doebley, Renfroe, and Blanton 1987).

The ribosomal ITS relationships between *Z. m.* ssp. *parviglumis, Z. mexicana,* and *Z. mays* were complicated by the recent divergence of these taxa. The Nobogame race of *Z. m.* ssp. *mexicana* appeared to diverge before the other Central Mexican *Z. mays,* which argued against its inclusion with other *Z. m.* ssp. *mexicana.* The early divergence of Nobogame is also suggested by knob data (Kato Y. 1976), but contradicted by isozyme data (Doebley and Goodman 1984). Genetic diversity estimates suggest that the Nobogame population might have undergone a bottleneck, which could complicate a determination of its phylogenetic position with distance measures. The ITS data set suggested that *Z. m.* ssp. *mays* diverged before or at the same time as the *Z. m.* ssp. *parviglumis* and *Z. m.* ssp. *mexicana* divergence, which conflicts with the isozyme data but is compatible with the distribution of chloroplast and mitochondrial alleles (Doebley, Goodman, and Stuber 1987; Doebley 1990*a*; Allen 1992). We speculate that this early divergence suggests a terminal Pleistocene/early Holocene domestication of maize. This level of phylogeny should be reevaluated with a synthesis of many loci. Eubanks (1995) recently suggested that a *Z. diploperennis* and *Tripsacum* hybrid was responsible for maize domestication. This theory is refuted by all available molecular and knob evidence including this ribosomal study.

Introgression of *Z. diploperennis* and *Z. perennis*

Three of the 11 *Z. diploperennis* and *Z. perennis* ITS clones were in the *Z. mays* clade, which suggested introgression. The introgressed *Z. perennis* ribosomal allele was almost identical to the *Z. m.* ssp. *mexicana* Nobogame (#57) allele, while the introgressed *Z. diploperennis* alleles were allied with *Z. m.* ssp. *mexicana* and *Z. parviglumis* alleles. Three points support introgression rather than differential segregation of ancestral polymorphism. First, *Z. mays* did not preserve any of the *Z. diploperennis*-like polymorphism, which is significantly different from an expectation of equal segregation of ancestral alleles into either clade ($G = 10.51$, $df = 1$, $P = 0.001$). The introgressed alleles had significantly higher substitution rates than other *Z. diploperennis* and *Z. perennis,* suggesting they may have spent time in the more rapidly evolving *Z. mays* genome ($P = 0.05$, for methods see Buckler and Holtsford 1996). Third, the two introgressed *Z. diploperennis* alleles were placed in clades that were shown to have significant differentiation between recently diverged populations.

Fertile hybrids between *Z. perennis,* the sole *Zea* tetraploid, and diploid *Zea* are very rare (Doebley 1989). However, hybrids between *Z. diploperennis* and *Z. mays* are fertile and meiosis is usually regular (Pasupuleti and Galinat 1982). Therefore, *Z. mays* alleles in *Z. diploperennis* probably came from simple hybridization, while the *Z. mays* allele in *Z. perennis* must have either crossed the ploidy barrier or been in the diploid ancestral pool. *Zea perennis* is generally considered an autotetraploid of some ancestral population of *Z. diploperennis* (Kato Y. and Lopez R. 1990); but several lines of evidence, including this ITS data, suggest substantial *Z. mays* germplasm in *Z. perennis.*

The ribosomal ITS suggested no differentiation between *Z. perennis* and *Z. diploperennis,* while isozyme evidence indicates a substantial divergence (Rogers' $D = 0.337$, Doebley and Goodman 1984). *Zea perennis* had a smaller isozyme distance to *Z. mays* than *Z. diploperennis* and *Z. luxurians,* which could reflect the introgression of *Z. mays* alleles into *Z. perennis.* For example, 11 of the 12 *Z. perennis* alleles of *Acp1* and *Pgm2* are shared with either *Z. diploperennis* or *Z. mays.* The large number of shared alleles argues for multiple origins of *Z. perennis* or multiple introgressions across the ploidy barrier. *Zea mays* introgression into the Piedra Ancha population of *Z. perennis* is also apparent from examination of mitochondria and chloroplast haplotypes (Doebley 1989; Allen 1992).

Chromosomal morphology data is ambiguous as to whether introgression has occurred. Chromosomal knobs do not indicate introgression of *Z. mays* into *Z. diploperennis* (Kato Y. and Lopez R. 1990). *Z. diploperennis* and *Z. perennis* have inversions on chromosomes 5 and 9, and these are shared with some *Z. m.* ssp. *mexicana,* especially the Nobogame race (Kato Y. 1976; Pasupuleti and Galinat 1982; Kato Y. and Lopez

R. 1990). *Zea perennis* has a higher level of bivalents (9.6 per cell) and a lower frequency of quadrivalents (0.499) than synthetic tetraploid *Z. m.* ssp. *mays* (Shaver 1962; Kato Y. and Lopez R. 1990), which suggests that *Z. perennis* is an old autotetraploid (partially diploidized), or that high levels of introgression have led to differentiation of the chromosomes.

The diploid *Z. diploperennis* and the tetraploid *Z. perennis* both seem to contain considerable *Z. mays* germplasm. It is difficult to discern whether introgression occurred before or after the tetraploid origin of *Z. perennis.* Multiple autotetraploid or allotetraploid origins for *Z. perennis* should be considered in future discussions of this taxon.

Conclusions

The ITS genetic diversity within the various *Zea* taxa was high compared to other genera, but it was not high compared to other *Zea* loci. Rapid ITS coalescence was suggested by analysis of maize alleles. *Zea* and *Tripsacum* were diverged from other examined Andropogoneae and do not form a monophyletic clade with *Coix.* The ITS pseudogenes were relicts of ribosomal evolution before the divergence of *Zea. Zea m.* ssp. *huehuetenangensis* was definitely basal to other *Z. mays. Z. luxurians* did not form a clade with *Z. diploperennis/perennis,* in conflict with chloroplast data. Future research with many nuclear loci and outgroups could elucidate this problem. The *Z. m.* ssp. *mexicana* Nobogame population may have diverged before the Central Mexican *Z. m.* ssp. *mexicana.* Evidence was provided for *Z. mays* introgression into *Z. diploperennis* and *Z. perennis.* The ITS data suggest an early divergence of ssp. *mays* from ssp. *mexicana* and *parviglumis,* but this should be reevaluated with many loci.

Acknowledgments

We would like to thank John Doebley for critical review of an earlier version of this manuscript and for seed samples. We also thank J. Birchler, K. Newton, and H. Iltis for useful discussions and seed samples, and two anonymous reviewers for comments on this manuscript. Thanks also to the USDA Plant Introduction Station (Ames, Iowa), B. Burr, and A. Blakey for samples. J. Spies and E. Kellogg graciously provided prepublication sequences for several of the Andropogoneae. This work was supported by an NSF Predoctoral Fellowship to E.S.B and UM Research Board award 93-060.

LITERATURE CITED

ALLEN, J. O. 1992. Teosinte cytoplasmic genomes: interaction with maize nuclear genomes and molecular genetic characterization of the mitochondria. PhD. dissertation, University of Wisconsin, Madison.

ALLEN, J. O., G. K. EMENHISER, and J. L. KERMICLE. 1989. Miniature kernel and plant: interaction between teosinte cytoplasmic genomes and maize nuclear genomes. Maydica **34**:277–290.

AVISE, J. C. 1994. Interpretative tools. Pp. 92–137 *in* Molecular markers, natural history and evolution. Chapman Hall, New York.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, and M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. **82**:247–277.

BUCKLER, E. S. IV, and T. P. HOLTSFORD. 1996. *Zea* ribosomal repeat evolution and mutation patterns. Mol. Biol. Evol. **13**: 623–632.

BULT, C., M. KÄLLERSJÖ, and Y. SUH. 1992. Amplification and sequencing of 16/18S rDNA from gel-purified total plant DNA. Plant Mol. Biol. Rep. **10**:273–284.

CHESTER, N., and D. R. MARSHAK. 1993. Dimethyl sulfoxide-mediated primer T_m reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. Anal. Biochem. **209**:284–290.

DENNIS, E. S., and W. J. PEACOCK. 1984. Knob heterochromatin homology in maize and its relatives. J. Mol. Evol. **20**:341–350.

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

DOEBLEY, J. 1989. Molecular evidence for a missing wild relative of maize and the introgression of its chloroplast genome into *Zea perennis.* Evolution **43**:1555–1559.

DOEBLEY, J. 1990*a*. Molecular evidence and the evolution of maize. Econ. Bot. **44**:6–27.

DOEBLEY, J. 1990*b*. Molecular systematics of *Zea* (Gramineae). Maydica **35**:143–150.

DOEBLEY, J. F., and M. M. GOODMAN. 1984. Isoenzymatic variation in *Zea* (Gramineae). Syst. Bot. **9**:203–218.

DOEBLEY, J. F., M. M. GOODMAN, and C. W. STUBER. 1987. Patterns of isozyme variation between maize and Mexican annual teosinte. Econ. Bot. **41**:234–246.

DOEBLEY, J., W. RENFROE, and A. BLANTON. 1987. Restriction site variation in the *Zea* chloroplast genome. Genetics **117**: 139–147.

EUBANKS, M. 1995. A cross between two maize relatives: *Tripsacum dactyloides* and *Zea diploperennis.* Econ. Bot. **49**: 172–182.

FELSENSTEIN, J. 1979. Alternative methods of phylogenetic inference and their interrelationships. Syst. Zool. **28**:49–62.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

FELSENSTEIN, J. 1989. PHYLIP: phylogeny inference package (version 3.2). Cladistics **5**:164–166.

GALINAT, W. C., R. S. K. CHAGANTI, and F. D. HAGER. 1964. *Tripsacum* as a possible amphidiploid of wild maize and *Manisuris.* Botanical Museum Leaflets, Harvard University **20**:289–316.

GAUT, B. S., and M. T. CLEGG. 1993*a*. Molecular evolution of the Adh1 locus in genus *Zea.* Proc. Natl. Acad. Sci. USA **90**:5095–5099.

GAUT, B. S., and M. T. CLEGG. 1993*b*. Nucleotide polymorphism in the Adh1 locus of Pearl Millet (*Pennisetum glaucum*) (Poaceae). Genetics **135**:1091–1097.

GAUT, B. S., S. V. MUSE, W. D. CLARK, and M. T. CLEGG. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. J. Mol. Evol. **35**:292–303.

GELFAND, D. H., and T. J. WHITE. 1990. Thermostable DNA polymerases. Pp. 129–141 *in* M. A. INNIS, D. H. GELFAND, J. J. SNINSKY, and T. J. WHITE, eds. PCR protocols: a guide to methods and applications. Academic Press, San Diego, Calif.

GOLOUBINOFF, P., S. PÄÄBO, and A. C. WILSON. 1993. Evolution of maize inferred from sequence diversity of an *Adh2*

gene segment from archaeological specimens. Proc. Natl. Acad. Sci. USA **90**:1997–2001.

HSIAO, C., N. J. CHATTERTON, K. H. ASAY, and K. B. JENSEN. 1995. Molecular phylogeny of the Pooideae (Poaceae) based on nuclear rDNA (ITS) sequences. Theor. Appl. Genet. **90**:389–398.

HUDSON, R. R., D. D. BOOS, and N. L. KAPLAN. 1992. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. **9**:138–151.

HUELSENBECK, J. P. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. **44**:17–48.

INNIS, M. A. 1990. PCR with 7-deaza-2′-deoxyguanosine triphosphate. Pp. 54–59 *in* M. A. INNIS, D. H. GELFAND, J. J. SNINSKY, and T. J. WHITE, eds. PCR protocols: a guide to methods and applications. Academic Press, San Diego, Calif.

KATO Y., T. A. 1976. Cytological studies of maize (*Zea mays* L.) and teosinte (*Zea mexicana* Schrader Kuntze) in relation to their origin and evolution. Mass. Agric. Expt. Sta. Bull. **635**:1–185.

KATO Y., T. A., and A. LOPEZ R. 1990. Chromosome knobs of the perennial teosintes. Maydica **35**:125–141.

KELLOGG, E. A., and J. A. BIRCHLER. 1993. Linking phylogeny and genetics: *Zea mays* as a tool for phylogenetic studies. Syst. Biol. **42**:415–439.

KISHINO, H., and M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. **29**:170–179.

KUHNER, M. K., and J. FELSENSTEIN. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. **11**:459–468.

KWIATOWSKI, J., D. SKARECKY, S. HERNANDEZ, D. PHAM, F. QUIJAS, and F. J. AYALA. 1991. High fidelity of the polymerase chain reaction. Mol. Biol. Evol. **8**:884–887.

LINARES, A. R., T. BOWEN, and G. A. DOVER. 1994. Aspects of nonrandom turnover involved in the concerted evolution of intergenic spacers within the ribosomal DNA of *Drosophila melanogaster*. J. Mol. Evol. **39**:151–159.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

LYNCH, M., and T. J. CREASE. 1990. The analysis of population survey data on DNA sequence variation. Mol. Biol. Evol. **7**:377–394.

NEI, M. 1987. DNA polymorphism within and between populations. Pp. 254–286 *in* Molecular evolutionary genetics. Columbia University Press, New York.

OLSEN, G. J., H. MATSUDA, R. HAGSTROM, and R. OVERBEEK. 1994. fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. **10**:41–48.

PASUPULETI, C. V., and W. C. GALINAT. 1982. *Zea diploperennis*: I. Its chromosomes and comparative cytology. J. Hered. **73**:168–170.

PURUGGANAN, M. D., and S. R. WESSLER. 1994. Molecular evolution of *magellan,* and maize Ty3/*gypsy*-like retroposon. Proc. Natl. Acad. Sci. USA **91**:11674–11678.

RITLAND, C. E., K. RITLAND, and N. A. STRAUS. 1993. Variation in the ribosomal internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttatus* species complex. Mol. Biol. Evol. **10**:1273–1288.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing trees. Mol. Biol. Evol. **4**: 406–425.

SANDERSON, M. J., and J. J. DOYLE. 1992. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. Syst. Biol. **41**:4–17.

SHATTUCK-EIDENS, D. M., R. N. BELL, S. L. NEUHAUSEN, and T. HELENTJARIS. 1990. DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing. Genetics **126**: 207–217.

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

SUH, Y., L. B. THIEN, H. E. REEVE, and E. A. ZIMMER. 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. Am. J. Bot. **80**:1042–1055.

SUN, Y., G. HEGAMYER, and N. H. COLBURN. 1993. PCR-direct sequencing of a GC-rich region by inclusion of 10% DMSO: application to mouse c-*jun*. Biotechniques **15**:372–374.

SUN, Y., D. Z. SKINNER, G. H. LIANG, and S. H. HULBERT. 1994. Phylogenetic analysis of sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. **89**:26–32.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign.

TAKEZAKI, N., and M. NEI. 1994. Inconsistency of the maximum parsimony method when the rate of nucleotide substitution is constant. J. Mol. Evol. **39**:210–218.

VARADARAJ, K., and D. M. SKINNER. 1994. Denaturants or cosolvents improve the specificity of PCR amplification of a G+C-rich DNA using genetically-engineered DNA-polymerases. Gene **140**:1–5.

WAGNER, A., N. BLACKSTONE, P. CARTWRIGHT, M. DICK, B. MISOF, P. SNOW, G. P. WAGNER, J. BARTELS, M. MURTHA, and J. PENDLETON. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Syst. Biol. **43**:250–261.

WEAVER, M. P. 1993. The Aztecs, Maya, and their predecessors: archaeology of Mesoamerica. Academic Press, San Diego, Calif.

WU, C. I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. Genetics **127**:429–435.

ZIMMER, E. A., E. R. JUPE, and V. WALBOT. 1988. Ribosomal gene structure, variation and inheritance in maize and its ancestors. Genetics **120**:1125–1136.

BARBARA A. SCHAAL, reviewing editor

Accepted January 8, 1996