Theor Appl Genet (2009) 119:1129–1142
DOI 10.1007/s00122-009-1115-1

ORIGINAL PAPER

Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte

Sherry A. Flint-Garcia · Anastasia L. Bodnar · M. Paul Scott

Received: 17 October 2008 / Accepted: 20 July 2009 / Published online: 22 August 2009

Abstract All crop species have been domesticated from their wild relatives, and geneticists are just now beginning to understand the consequences of artificial (human) selection on agronomic traits that are relevant today. The primary consequence is a basal loss of diversity across the genome, and an additional reduction in diversity for genes underlying traits targeted by selection. An understanding of attributes of the wild relatives may provide insight into target traits and valuable allelic variants for modern agriculture. This is especially true for maize (*Zea mays* ssp. *mays*), where its wild ancestor, teosinte (*Z. mays* ssp. *parviglumis*), is so strikingly different than modern maize. One obvious target of selection is the size and composition of the kernel. We evaluated kernel characteristics, kernel composition, and zein profiles for a diverse set of modern inbred lines, teosinte accessions, and landraces, the intermediate between inbreds and teosinte. We found that teosinte has very small seeds, but twice the protein content of landraces and inbred lines. Teosinte has a higher average alpha zein content (nearly 89% of total zeins as compared to 72% for inbred lines and 76% for landraces), and there are many novel alcohol-soluble proteins in teosinte relative to the other two germplasm groups. Nearly every zein protein varied in abundance among the germplasm groups, especially the methionine-rich delta zein protein, and the gamma zeins. Teosinte and landraces harbor phenotypic variation that will facilitate genetic dissection of kernel traits and grain quality, ultimately leading to improvement via traditional plant breeding and/or genetic engineering.

Communicated by A. Bervillé.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1115-1) contains supplementary material, which is available to authorized users.

S. A. Flint-Garcia
USDA-ARS, Plant Genetics Research Unit,
Columbia, MO 65211, USA

S. A. Flint-Garcia
Division of Plant Sciences, University of Missouri,
Columbia, MO 65211, USA

A. L. Bodnar · M. P. Scott
Department of Agronomy, Iowa State University,
Ames, IA 50011, USA

M. P. Scott
USDA-ARS Corn Insects and Crop Genetics Research Unit,
Ames, IA 50011, USA

S. A. Flint-Garcia (✉)
USDA-ARS, University of Missouri,
301 Curtis Hall, Columbia, MO 65211, USA
e-mail: Sherry.Flint-Garcia@ars.usda.gov

Introduction

Maize was domesticated in a single event from teosinte about 7,500 years ago in central Mexico (Matsuoka et al. 2002). It is believed that a founding population of teosinte individuals was isolated from the progenitor population via human selection to form an ancestral maize population. By the time Columbus discovered the Americas, artificial selection and the steady accumulation of mutations allowed the range of maize to expand from Mexico to Canada (Ducrocq et al. 2008). During this expansion, different maize lineages adapted to local growing conditions (soil type, temperature, altitude, and biotic and abiotic stresses) and desired human uses. While the resulting heterogeneous, open-pollinated landraces resemble modern maize more than teosinte, they can be considered an intermediate

between teosinte and modern inbreds (Yamasaki et al. 2005). Landraces have not undergone inbreeding and have not been selected to perform under highly intensive agriculture practices. Focused maize breeding efforts beginning in the early 1900s resulted in inbred lines that, when crossed, produce hybrids with increased vigor and substantially higher yields.

The *mays* and *parviglumis* subspecies differ substantially in plant, ear, and seed morphologies (reviewed in Doebley 2004). The most striking examples include differences in plant and inflorescence architecture conferred by the *teosinte branched1* locus (Doebley et al. 1995) and the hardened glume structure (fruitcase) surrounding the teosinte kernel conferred by *teosinte glume architecture1* (Dorweiler et al. 1993). Large scale sequencing studies comparing teosinte to modern inbred lines have indicated 2–4% of the maize genome has experienced artificial selection throughout its history, i.e., during domestication and/or plant breeding (Wright et al. 2005; Yamasaki et al. 2005). These selected genes can be roughly divided into two classes: "domestication genes" where diversity is greatly reduced in *landraces and inbreds*, and "improvement genes" where diversity is severely reduced in *only inbreds*. For either class of selected genes, there is little or no genetic variation remaining in inbred lines to contribute to crop improvement by traditional breeding or gene discovery by genetic analysis.

Maize kernel composition is important in terms of human and animal nutrition. Typical kernel composition values for the commodity yellow dent corn on a dry matter basis are 71.7% starch, 9.5% protein, 4.3% oil, 1.4% ash, and 2.6% sugar (Watson 2003); 80% of the protein is stored in the endosperm, the nutritive tissue of the seed. The essential amino acids lysine, tryptophan, and methionine are limited because they are lacking or present at low levels in zeins, the major class of storage proteins. Zeins constitute about half of the endosperm protein, and thus nearly half of the total seed protein (Paulis and Wall 1977). Genetic and genomic studies have revealed great complexity of zein gene families (Song et al. 2001; Song and Messing 2002; Wilson and Larkins 1984; Woo et al. 2001) and support earlier observations that different classes of zeins have varying amino acid compositions (Melcher and Fraij 1980; Sodek and Wilson 1971).

Because the zeins are so abundant, they have a large impact on the amino acid composition of the kernel. Most attempts to improve the nutritional quality of maize protein involve altering zein content. One example of this is the development of Quality Protein Maize, where the *opaque2* mutation alters zein levels in the kernels (Prasanna et al. 2001). A second example is the use of *dzr1,* a mutation that results in overproduction of a methionine-rich zein, to develop high-methionine inbred lines (Phillips et al. 2008). A number of transgenic approaches involving modification of zein levels have been explored as well (Huang et al. 2004, 2005; Lai and Messing 2002).

Teosinte and landrace accessions may be sources of genetic variation for maize improvement, especially for genes that have limited or no variation remaining in modern inbred lines due to initial domestication events and plant breeding (Wright et al. 2005; Yamasaki et al. 2005). An example of the utility of wild species in genetic studies and crop improvement is using *Oryza rufipogon* grain yield in rice (Xiao et al. 1998). This is likely the case for the maize starch pathway, where three of six genes have experienced selection during domestication (Whitt et al. 2002). Several studies suggest that progenitors of modern maize contain a diversity of zein genes that is lacking in modern inbreds (Swarup et al. 1995; Wilson and Larkins 1984). For example, Swarup et al. (1995) found that exotic maize and wild members of the genus *Zea* exhibited higher levels of methionine-rich delta zeins than maize inbreds, leading the authors to hypothesize that the high methionine trait was lost in the course of domestication. Whether loss of the high methionine trait was a result of artificial selection or random genetic drift is unclear. Introgression of *Z. mays* ssp. *mexicana*, a teosinte more distantly related to maize than ssp. *parviglumis*, into maize resulted in lines with significantly higher protein content, as well as higher lysine, methionine, and/or phenylalanine content on a kernel weight basis (Wang et al. 2008).

Geneticists are just now beginning to understand the consequences of domestication and breeding history of a crop species (Hamblin et al. 2006; Hyten et al. 2006; Tang et al. 2006). One such consequence is that useful genetic variation may have been lost during the domestication process. Knowledge of the growth, physiology, and other various attributes of wild relatives may provide insights into key traits and allelic variants that are useful in modern agriculture. The objective of this study was to test the hypothesis that teosintes carry variation in seed traits that exceeds the variation found in domesticated maize, and to determine if this variation could be useful for improving maize germplasm. To do this, we examined seed characteristics and the zein seed storage proteins in a panel of diverse germplasm that includes modern inbred lines, landraces, and teosinte.

Materials and methods

Plant materials and experimental design

The germplasm used in this study were selected to represent a broad diversity within the teosintes, maize landraces, and maize inbred lines. We obtained the following 11 geographically diverse teosinte (ssp. *parviglumis*) accessions

from the North Central Regional Plant Introduction Station (NCRPIS): PI 384063, PI 384065, PI 384066, PI 384071, Ames 21889, Ames 21785, Ames 21786, Ames 21789, Ames 21809, Ames 21812, and Ames 21814. We obtained the following 17 landrace (ssp. *mays*) accessions from NCRPIS or M. Goodman at North Carolina State University: Assiniboine (PI213793), Bolita (OAX68), Cateto Sulino (URG II), Chalqueno (MEX48), Chapalote (SIN2), Conico (PUE32), Costeno (VEN453), Cristalino Norteno (CHI349), Dzit Bacal (GUA131), Gordo (CHH160), Guirua (MAG450), Nal-tel (YUC7), Pisccorunto (APC13), Sabanero (SAN329), Serrano (GUA14), Tuson (CUB57), and Zapalote Chico (OAX70). The inbred lines included in this study are the 27 parental lines of the Nested Association Mapping (NAM) population (McMullen et al. 2009; Yu et al. 2008): B73, B97, CML103, CML228, CML247, CML277, CML322, CML333, CML52, CML69, Hp301, IL14H, Ki11, Ki3, Ky21, M162W, M37W, Mo17, Mo18W, MS71, NC350, NC358, Oh43, Oh7B, P39, Tx303, and Tzi8.

The 55 entries were planted in two replicates at the Illinois Crop Improvement Association winter nursery site near Ponce, Puerto Rico in winter 2005–2006. A day-neutral site was required as teosinte will not flower and set seed in long day (>12 h/day) environments due to photoperiod sensitivity. The 55 entries were randomized within groups (inbred lines, landraces, and teosintes) and groups were randomized within replicates. All entries were allowed to open pollinate in order to obtain an adequate amount of seed for analyses. Ears were harvested, and balanced bulks of seed were created for each plot.

Kernel composition and seed characteristics

The stony fruitcases of the teosinte seeds were removed and discarded prior to analysis. Kernel weights of all entries were determined by weighing 100 kernels from each replicate. Kernels were ground into a fine meal and submitted to the University of Missouri Experiment Station Chemical Laboratories for proximate analysis following the Official Methods of AOAC International (2006). Crude fat, moisture, ash, and crude fiber were determined on a per tissue mass basis for each sample by methods 920.39 (A), 934.01, 942.05, and 978.10, respectively. Crude protein was determined by combustion analysis (LECO; method 990.03). Total carbohydrate was calculated by subtraction.

Percent endosperm (wt/wt) was determined for the inbreds and landraces from both experimental replicates. Inadequate seed quantities of the teosinte entries prevented determining percent endosperm from the same field experiment as other traits. However, the same 11 teosinte accessions used in this project had been seed increased individually via open pollination in isolation under growth chamber and greenhouse conditions. From these stocks, a sample of three teosinte accessions was chosen to span the range of seed size (small, medium, and large) present among the 11 teosinte accessions. Percent endosperm was determined for these three teosinte accessions for comparison. The removal of the fruitcase is extremely labor intensive and prevented analyzing the percent endosperm for all accessions.

HPLC analysis of the zein storage proteins

A single replicate of the field experiment was used for HPLC analysis of the alcohol-soluble proteins. For each entry, a bulk of kernels was ground into fine flour with a coffee grinder. The number of kernels ground per sample was determined by kernel size: approximately 20 kernels for the teosintes, and five for the landraces and inbred lines. In addition, three whole individual kernels from each of two teosintes were ground with a handheld drill. Zeins were quantified on a "per tissue mass" basis for each sample. Alcohol-soluble proteins were extracted from 10 mg of flour using 100 µL extraction buffer consisting of 70% EtOH, 61 mM NaOAc, and 5% β-mercaptoethanol. The mixture was vortexed briefly, horizontally shaken for 1 h at 37°C, then centrifuged for 10 min at 12,000 rpm. The supernatant was diluted 1:4 with extraction buffer. An aliquot of 25 µL of each extract was injected into a C18 protein and peptide column in a Waters 2695 Separation Module, and absorbance at 200 nm was measured with a Waters 2487 Dual Absorbance Detector. Separation of distinct proteins based on hydrophobicity was achieved with a gradient of ultrapure water and acetonitrile, both containing 0.01% trifluoroacetic acid. The gradient ranged from 65 to 25% acetonitrile for a total of 40 min of elution at a flow rate of 2 mL/min, excluding equilibration steps before and after elution. The entire set of 55 samples was extracted and injected in seven separate sets, each including a B73 entry before and after the samples as a control.

Statistical analysis

For the HPLC of the zein proteins, the total area under the curve (excluding the injection peak prior to 7 min and the wash peak after 35 min elution time) was calculated for each entry using integration via the Empower software (Waters) with a minimum peak width of 30 and threshold of 800. Specific zein peaks from 7 to 35 min elution time were identified by comparison to known inbred HPLC profiles (Wilson 1991). Area under the peak was estimated for a subset of peaks and converted to percent of total area. Alpha peak areas (which eluted between 15 and 28 min) were added together to form a total alpha zein value. Within each germplasm group, an average absorbance and

standard deviation were calculated for each point in the HPLC traces. Analysis of variance was conducted to determine whether there was significant variation "among groups" using SAS PROC MIXED (SAS Institute Inc. 1999–2001). Where significant variation among groups existed, SAS PROC TTEST was used to compare mean values of teosinte versus landraces, teosinte versus inbred, and landraces versus inbreds.

For the kernel composition and seed characteristic traits, analysis of variance was conducted using SAS PROC MIXED with entries fixed and replicates random. The phenotypic variance of all entries was partitioned into "among groups", and specific contrasts of teosinte versus landraces, teosinte versus inbreds, and landraces versus inbreds were tested. Least squares means for entries were calculated and least significant differences (LSD) were obtained for $P = 0.05$. Pearson's correlation coefficients were calculated using SAS PROC CORR. Principal component analysis was conducted using the PROC PRINCOMP procedure of SAS, and the number of "meaningful" components was determined using the eigenvalue-one criterion (Kaiser 1960).

Results

Phenotypic traits were organized into three major groupings for this study: kernel composition, seed characteristics, and zein profiles.

Kernel composition

Analysis of variance revealed that replicates were not significantly different ($P > 0.05$) for any of the kernel composition traits: moisture, protein, fat, fiber, ash, and carbohydrate (data not shown). Significant differences existed among entries, and among the three germplasm groups for each trait ($P < 0.05$). The teosintes had less carbohydrate and more protein than either landraces or inbred lines ($P < 0.001$), and landraces had less carbohydrate and more protein than the inbreds ($P < 0.01$; Fig. 1; Table 1). These results demonstrate a dramatic shift in seed nutrient storage during domestication and/or plant breeding, especially between teosinte and the other two germplasm groups.

Teosinte also had higher fat and ash content, and lower moisture and fiber content than either the landraces or the inbred lines ($P < 0.01$; Fig. 1; Table 1). While these comparisons are statistically significant, the differences are far less marked than those for protein and carbohydrate.

Seed characteristics

Replicates were not a significant source of variation for either seed weight or percent endosperm ($P > 0.05$; data not shown). Landrace and inbred kernels weighed eight to nine times more than teosinte kernels after the stony fruitcases were removed ($P < 0.001$; Fig. 2; Table 1). Inbred and landrace kernel weights were not significantly different at $P = 0.05$. Percent endosperm did not differ between teosinte and landraces ($P > 0.05$), while inbred percent endosperm was significantly higher than the other two germplasm groups ($P < 0.05$).

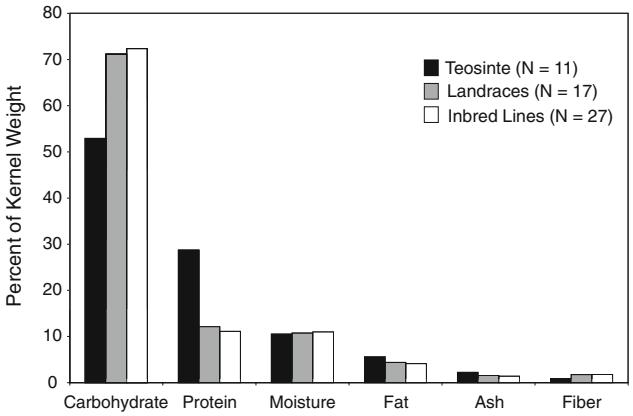

Fig. 1 Kernel composition of teosinte (*black*), landrace (*grey*), and inbred lines (*white*)

Zein profile

A graph of HPLC traces for all individuals can be found in the online supplemental materials (Supplemental Fig. 1). A simple comparison of the inbred line B73 and teosinte accession Ames 21785 demonstrates our peak naming convention (Fig. 3) based on the peak assignments of Wilson (1991) and the genetic class system of Thompson and Larkins (1989). Variation among the groups can be clearly seen in a qualitative analysis. For this purpose, we defined three regions of the chromatograms for detailed examination. The gamma region contains the 16 and 27 kDa gamma zein peaks and the beta zein peaks, the alpha region contains the alpha zein peaks, and the delta region contains the delta zeins. In order to simplify the complex HPLC trace data and to show differences in zein profiles among the groups, an average zein profile was calculated for each germplasm group (Fig. 4). The zein profiles among the three groups differ considerably, especially the teosintes as compared to the landraces and inbreds.

In order to express these data quantitatively, we identified six prominent peaks and integrated them separately but simultaneously in all samples (Table 2). Because the alpha zein region (which eluted between 15 and 28 min) was so complex and variable among entries, we integrated each of the alpha zein peaks separately and then summed them together to obtain a total alpha zein value for each entry.

Table 1 Least squares entry means and analysis of variance (ANOVA) for kernel composition and seed characteristic traits for a panel of teosinte accessions, landraces, and inbred lines

Entry	Group	Moisture	Protein	Fat	Fiber	Ash	Carbohydrate	Seed weight	Percent endosperm
B73	Inbred	10.65	11.06	3.75	1.41	1.24	73.31	0.28	92.66
B97	Inbred	11.21	9.08	3.61	1.42	1.23	74.89	0.27	92.98
CML103	Inbred	11.37	9.14	3.54	1.62	1.39	74.58	0.32	91.45
CML228	Inbred	10.50	12.48	5.53	1.43	1.54	69.97	0.33	91.64
CML247	Inbred	11.23	12.45	6.95	2.72	1.59	67.80	0.22	91.56
CML277	Inbred	10.93	11.63	3.48	1.65	1.29	72.68	0.25	91.79
CML322	Inbred	10.74	11.20	4.46	2.44	1.55	72.06	0.23	90.96
CML333	Inbred	11.11	11.88	4.69	2.11	1.53	70.79	0.24	90.25
CML52	Inbred	10.78	12.63	3.22	2.10	1.35	72.03	0.24	91.71
CML69	Inbred	10.91	12.70	3.34	1.97	1.43	71.63	0.27	92.22
Hp301	Inbred	10.31	12.15	3.92	2.91	1.37	72.26	0.11	94.54
Il14H	Inbred	10.90	12.38	4.68	1.57	1.64	70.41	0.17	91.82
Ki11	Inbred	10.48	11.29	4.17	1.74	1.35	72.72	0.32	91.32
Ki3	Inbred	11.38	11.12	3.87	1.99	1.48	72.16	0.27	91.45
Ky21	Inbred	11.26	11.42	3.43	1.90	1.49	72.42	0.29	91.91
M162W	Inbred	11.71	10.66	4.10	1.80	1.31	72.32	0.35	92.20
M37W	Inbred	11.23	10.56	3.57	1.91	1.45	73.20	0.29	93.21
Mo17	Inbred	11.03	11.90	3.80	1.77	1.34	71.94	0.30	91.79
Mo18W	Inbred	11.37	9.05	4.03	1.66	1.37	74.20	0.23	92.45
MS71	Inbred	10.74	11.09	4.23	2.07	1.43	72.52	0.29	91.62
NC350	Inbred	10.89	12.95	3.67	1.39	1.35	71.15	0.21	92.21
NC358	Inbred	11.12	10.63	3.37	1.50	1.38	73.51	0.25	92.99
Oh43	Inbred	11.34	8.70	3.78	1.51	1.19	75.00	0.27	90.71
Oh7B	Inbred	10.71	8.97	4.51	1.38	1.37	74.46	0.28	91.13
P39	Inbred	10.12	12.15	6.71	1.42	1.55	69.48	0.20	90.72
Tx303	Inbred	11.61	8.11	3.59	1.68	1.33	75.37	0.31	91.43
Tzi8	Inbred	11.55	12.70	3.27	1.46	1.41	71.17	0.29	91.35
Assiniboine	Landrace	10.63	12.91	4.30	1.99	1.60	70.58	0.30	91.20
Bolita	Landrace	10.82	11.50	4.63	1.50	1.58	71.49	0.35	89.11
Cateto Sulino	Landrace	10.65	13.12	5.11	1.88	1.54	69.60	0.28	87.58
Chalqueno	Landrace	10.44	12.89	4.80	2.09	1.60	70.37	0.28	91.28
Chapalote	Landrace	10.56	13.60	4.81	1.95	1.56	69.48	0.21	90.15
Conico	Landrace	10.53	12.52	4.71	1.67	1.57	70.69	0.20	86.96
Costeno	Landrace	10.76	11.11	3.93	1.70	1.42	72.79	0.29	90.97
Cristalino Norteno	Landrace	11.13	11.13	4.25	1.30	1.46	72.04	0.30	91.64
Dzit Bacal	Landrace	11.06	12.25	4.35	1.80	1.60	70.76	0.34	90.80
Gordo	Landrace	10.83	11.17	4.74	1.69	1.51	71.77	0.36	91.11
Guirua	Landrace	10.85	13.35	3.83	1.78	1.69	70.30	0.22	91.22
Nal-tel	Landrace	10.50	12.46	5.01	1.98	1.54	70.51	0.16	89.05
Pisccorunto	Landrace	11.30	10.09	3.47	1.74	1.65	73.50	0.32	88.91
Sabanero	Landrace	10.92	11.54	4.38	1.31	1.46	71.71	0.31	90.95
Serrano	Landrace	10.94	13.34	4.01	1.66	1.59	70.12	0.22	90.68
Tuson	Landrace	10.64	10.90	4.11	1.83	1.48	72.87	0.33	90.60
Zapalote Chico	Landrace	10.58	12.33	4.44	1.92	1.54	71.12	0.33	90.09
Ames 21785	Teosinte	10.45	30.45	5.68	1.01	2.27	51.16	0.02	n.d.
Ames 21786	Teosinte	10.46	29.63	6.15	0.95	2.27	51.50	0.02	n.d.

Table 1 continued

Entry	Group	Moisture	Protein	Fat	Fiber	Ash	Carbohydrate	Seed weight	Percent endosperm
Ames 21789	Teosinte	10.53	30.72	5.00	0.92	2.16	51.60	0.02	n.d.
Ames 21809	Teosinte	10.59	29.82	5.33	0.98	2.08	52.20	0.02	91.31
Ames 21812	Teosinte	10.57	26.49	5.43	0.87	2.24	55.28	0.03	n.d.
Ames 21814	Teosinte	10.66	27.83	6.56	0.93	2.39	52.57	0.03	n.d.
Ames 21889	Teosinte	10.54	26.63	5.92	0.87	2.28	54.64	0.04	n.d.
PI 384063	Teosinte	10.47	28.97	4.96	0.85	2.07	53.54	0.03	90.05
PI 384065	Teosinte	10.53	27.59	5.45	0.90	2.38	54.06	0.03	n.d.
PI 384066	Teosinte	10.43	27.59	5.96	0.94	2.44	53.60	0.03	n.d.
PI 384071	Teosinte	10.50	30.15	5.27	0.82	2.12	51.98	0.03	89.19
LSD ($P = 0.05$)		0.74	1.83	1.86	0.33	0.33	2.70	0.05	3.04
Teosinte mean		10.52	28.71	5.61	0.91	2.24	52.92	0.03	90.18
Landrace mean		10.77	12.13	4.40	1.75	1.55	71.16	0.28	90.13
Inbred mean		11.00	11.11	4.12	1.80	1.40	72.37	0.26	91.85
ANOVA									
Among Groups		***	***	***	***	***	***	***	***
Teo. versus LR		***	***	***	***	***	***	***	ns
Teo. versus Inbreds		***	***	***	***	***	***	***	*
LR versus Inbreds		*	**	ns	ns	**	**	ns	***

n.d. not determined, *LSD* least significant difference for comparing entry means, *ns* not significantly different at $P = 0.05$

*, **, *** Significantly different at $P < 0.05$, 0.01, and 0.001, respectively

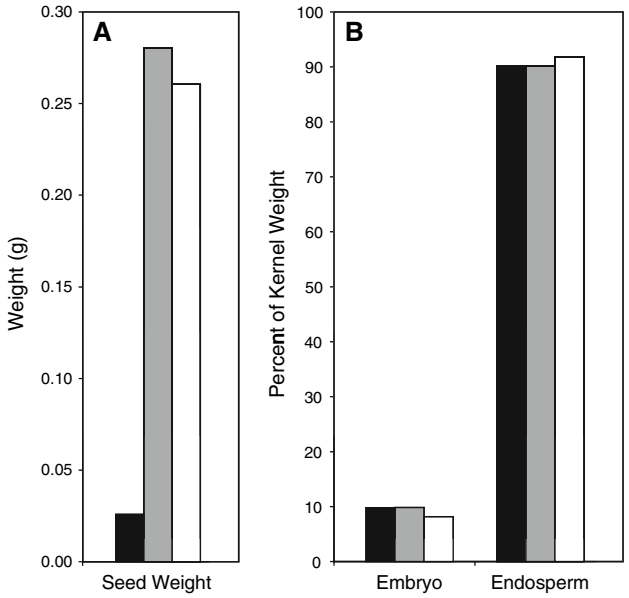

Fig. 2 Seed weight (**a**) and percent endosperm and embryo (**b**) of teosinte (*black*), landrace (*grey*), and inbred lines (*white*)

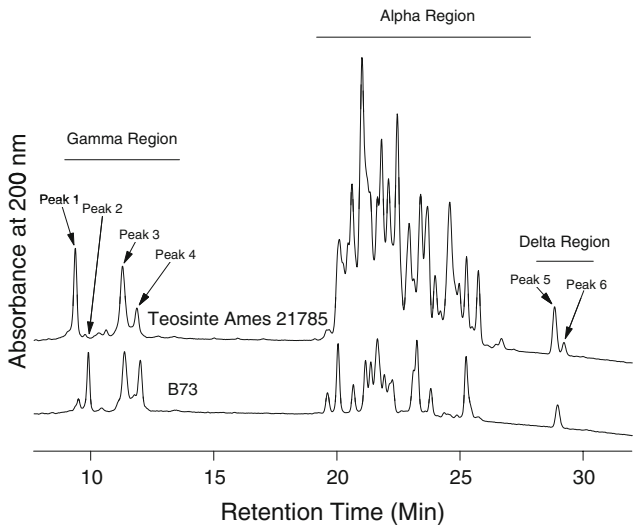

Fig. 3 Comparison of inbred line B73 and teosinte accession Ames 21785 demonstrating typical zein profiles and the naming of specific peaks

Examination of the region of the chromatogram where the alpha zein proteins elute (Fig. 4; Supplemental Fig. 2) shows the first striking difference between teosintes, and landraces and inbred lines. The teosintes have very high levels of alpha zeins relative to the landraces and inbreds, both in terms of total area under the curve and in number of peaks. While landraces and inbred lines tend to have five to seven distinct alpha peaks, no discrete peaks are evident in the teosintes (Fig. 4; Supplemental Fig. 2). Three prominent peaks are present in both inbred lines and landraces, while other peaks are prominent in either inbred lines or landraces (Supplemental Fig. 2). The teosintes have rela-

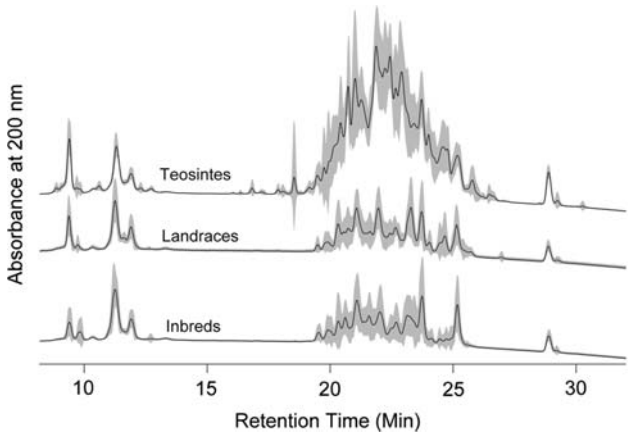

Fig. 4 Average zein profiles for comparison of teosinte, landrace, and inbred line germplasm groups. At each point along the *x*-axis, the *black line* represents the average value and the *grey* region represents the standard deviation

tively prominent peaks outside of the region containing the landrace and inbred line peaks, for example, between 17 and 19 min and between 26 and 27 min, presumably representing novel alpha zein proteins. It is clear that the alpha zein protein family in teosintes as a group is much more complex than those of either landraces or inbred lines.

One potential explanation for the increased complexity, in terms of number of zein peaks, of the teosinte alpha zein protein family is that there may be more heterogeneity within the individual teosinte accessions than among the inbred or landrace accessions. Each teosinte accession is a population of related, segregating individuals. To test this hypothesis, we analyzed the zein content of three individual kernels from two teosinte accessions. There was variability between individual kernels within each accession in terms of relative amount of each peak (data not shown). However, there were still more alpha zein peaks in single teosinte kernels than in the landrace and inbred groups. It is also worth noting that each landrace is also a population of related, segregating individuals, yet they do not display the increased number of peaks of the teosintes. A close examination of the alpha zein region demonstrates noticeably more variability in the teosintes than the landraces despite the fact that there are more landraces ($N = 17$) sampled than teosintes ($N = 11$) (Supplemental Fig. 2).

To account for differences in total alcohol-soluble protein, the area under each peak was converted to percent of total area for specific peaks (Table 2). In comparing the percent alpha zeins among the groups, teosintes had a significantly higher ($P < 0.001$) alpha zein content (average of 26 peaks representing 90.4% of total alcohol-soluble proteins) than the landraces (17 peaks representing 76.5% of the total) or the inbred lines (15 peaks representing 72.6% of the total).

A second striking difference among the groups was in the gamma zein region of the chromatograms. Peak 1, the 14 kDa beta zein, appeared to be more abundant in teosinte than the other groups (Fig. 4; Supplemental Fig. 3). However, after peak areas were adjusted for total alcohol-soluble proteins, peak 1 was significantly less abundant in teosinte than landraces ($P < 0.01$), but was equally abundant in teosinte and inbred lines (Table 2). This apparent discrepancy is a reflection of the total alcohol-soluble proteins, and therefore total protein content. Because the alcohol-soluble proteins are very abundant but generally have low nutritional value, the high level of these proteins may mask the presence of other proteins with better nutritional properties. The percentages of peaks 3 and 4, the 27 and 16 kDa gamma zeins, respectively, were significantly higher in inbreds and landraces than teosintes ($P < 0.001$). The variation in gamma zein content among the germplasm groups suggests that there may be differences in the processes that lead to zein deposition in the protein body as proposed by Coleman et al. (1996).

There was also variation in the delta zein region (Table 2; Fig. 4; Supplemental Fig. 4). Teosintes had lower levels of peak 5, compared to landraces and inbred lines. Additionally, there are many novel (i.e., unknown) peaks that exist in one group or another (Table 2; Supplemental Fig. 1). In general, teosinte has a significantly greater number of novel peaks than other groups ($P < 0.001$), especially in the elution period of 16–19 min. The peaks in the 16- to 19-min elution period likely represent alpha zeins due to the hydrophobicity and charge of these proteins as reflected by HPLC. Other potentially novel peaks exist in the beta and gamma zein region (elution period of 9–13 min).

Correlated traits

We measured the seed characteristics seed weight and percent endosperm as it is unclear how these general seed characteristics relate to kernel composition (Vasal 2000).

When using the entire dataset consisting of entry means, the correlations among the kernel composition data (moisture, protein, fat, fiber, ash, and carbohydrate) were all highly significant ($P < 0.01$; Table 3) for all but one comparison. These results were anticipated as composition values must sum to 100%, and are consistent with prior studies (Hopkins 1899; Pollmer et al. 1978; Rossi et al. 2001). All composition traits were highly correlated ($P < 0.001$) with seed weight, presumably because of the extreme differences in kernel composition and kernel weight between teosinte and the other two groups. The fact that only crude fat and ash were significantly correlated ($P < 0.05$) with percent endosperm suggests that modifications in carbohydrate and protein can be achieved without affecting the endosperm to embryo ratio. There were many significant correlations between the kernel composition traits and aspects of the zein profiles. Of interest is the consistently significant

Table 2 Quantification and analysis of variance (ANOVA) of zein protein peaks for a panel of diverse teosinte accessions, landraces, and inbred lines

Entry	Group	Alpha (no. of peaks)	Alpha (% area)	Peak 1[a] C1[b] 14 kDa-β[c] (% area)	Peak 2 C2 15 kDa-β (% area)	Peak 3 E 27 kDa-γ (% area)	Peak 4 F 16 kDa-γ (% area)	Peak 5 (% area)	Peak 6 (% area)	Unknown (no. of peaks)
Ames 21785	Teosinte	22	90.31	2.56	0.08	3.27	1.22	1.54	0.37	3
Ames 21786	Teosinte	29	90.77	2.87	0.11	2.36	1.08	1.50	0.08	5
Ames 21789	Teosinte	29	90.10	2.34	0.00	2.72	1.17	1.63	0.80	5
Ames 21809	Teosinte	27	90.14	2.66	0.00	3.10	1.23	1.40	0.31	7
Ames 21812	Teosinte	26	89.91	1.53	0.72	2.77	1.26	1.26	0.00	7
Ames 21814	Teosinte	29	90.56	1.93	0.15	2.72	1.61	1.33	0.00	8
Ames 21889	Teosinte	23	91.12	1.07	1.10	3.14	0.28	1.27	0.14	8
PI 384063	Teosinte	30	92.24	2.19	0.00	2.67	1.03	1.20	0.00	4
PI 384065	Teosinte	25	89.35	2.47	0.00	3.03	1.39	1.42	0.17	6
PI 384066	Teosinte	24	89.26	3.77	0.00	2.81	0.80	1.40	0.00	9
PI 384071	Teosinte	25	90.62	1.63	0.00	3.44	1.53	1.50	0.00	6
Assiniboine	Landrace	16	83.11	6.08	0.00	4.01	1.26	2.02	0.29	3
Bolita	Landrace	17	77.75	4.22	0.00	10.73	4.48	2.30	0.00	2
Cateto Sulino	Landrace	19	74.72	4.29	0.61	14.71	0.98	2.82	0.00	4
Chalqueno	Landrace	15	79.26	3.88	0.43	8.10	2.63	2.23	0.00	4
Chapalote	Landrace	20	78.31	3.64	0.25	10.91	4.15	1.49	0.00	4
Conico	Landrace	19	82.49	1.00	2.42	8.05	1.71	1.89	0.64	2
Costeno	Landrace	20	71.94	6.24	0.00	13.66	4.52	2.50	0.00	3
Cristalino Norteno	Landrace	15	62.91	10.29	0.00	13.85	7.84	2.32	0.00	4
Dzit Bacal	Landrace	19	75.86	5.03	0.68	9.51	2.89	2.28	0.88	4
Gordo	Landrace	16	75.84	2.59	2.22	8.59	0.00	2.84	1.23	2
Guirua	Landrace	17	79.52	4.48	0.00	9.05	3.71	2.33	0.42	2
Nal-tel	Landrace	16	72.86	5.95	0.00	9.00	5.69	3.31	0.00	5
Pisccorunto	Landrace	19	76.32	5.16	0.90	4.28	8.19	3.07	1.74	1
Sabanero	Landrace	17	76.33	6.34	0.00	10.53	3.07	2.71	0.00	2
Serrano	Landrace	15	80.78	4.90	0.00	7.00	1.56	1.88	0.00	4
Tuson	Landrace	17	72.22	0.30	3.72	13.31	3.00	2.95	1.47	3
Zapalote Chico	Landrace	19	80.56	3.00	0.81	8.99	3.99	2.19	0.00	2
B73	Inbred	14	70.40	1.46	5.53	10.13	6.99	3.05	0.00	2
B97	Inbred	16	65.39	4.85	0.00	16.97	5.70	3.37	0.00	3
CML103	Inbred	15	66.92	0.00	8.41	11.72	7.12	2.88	0.00	3
CML228	Inbred	17	80.95	0.32	4.31	7.90	3.88	2.19	0.00	1
CML247	Inbred	17	78.90	2.85	0.21	11.55	3.06	2.97	0.00	1
CML277	Inbred	18	83.64	1.37	0.17	10.49	0.42	2.86	0.67	1
CML322	Inbred	18	67.45	4.81	0.00	17.94	4.58	2.29	2.00	2
CML333	Inbred	19	76.54	3.49	0.69	12.99	3.68	1.87	0.00	3
CML52	Inbred	14	78.24	2.14	0.00	13.54	2.69	2.73	0.00	2
CML69	Inbred	17	76.42	2.03	0.96	13.35	4.16	2.05	0.00	2
Hp301	Inbred	17	79.75	1.20	4.67	8.46	0.85	1.55	0.00	3
Ill4H	Inbred	15	73.64	5.13	0.00	14.54	3.49	1.83	0.00	3
Ki11	Inbred	14	72.57	1.72	5.82	11.25	4.62	2.31	0.15	2
Ki3	Inbred	14	77.87	3.23	0.00	11.44	3.75	3.20	0.00	1
Ky21	Inbred	15	61.45	1.13	3.45	24.01	4.76	3.76	0.00	2
M162W	Inbred	16	73.30	4.13	0.00	13.46	4.81	3.29	0.00	3

Table 2 continued

Entry	Group	Alpha (no. of peaks)	Alpha (% area)	Peak 1[a] C1[b] 14 kDa-β[c] (% area)	Peak 2 C2 15 kDa-β (% area)	Peak 3 E 27 kDa-γ (% area)	Peak 4 F 16 kDa-γ (% area)	Peak 5 (% area)	Peak 6 (% area)	Unknown (no. of peaks)
M37W	Inbred	14	69.16	0.96	2.93	18.23	4.65	2.95	0.00	2
Mo17	Inbred	15	71.92	2.32	0.22	15.67	3.86	1.91	0.00	5
Mo18W	Inbred	17	60.88	4.89	0.00	22.67	4.76	3.97	1.80	2
MS71	Inbred	16	73.69	1.69	0.19	16.46	2.56	2.31	2.22	3
NC350	Inbred	16	81.52	3.79	0.41	8.20	3.75	1.89	0.00	2
NC358	Inbred	15	76.76	4.56	1.00	9.75	4.86	2.25	0.00	3
Oh43	Inbred	15	60.64	2.56	5.94	14.22	8.44	3.86	0.00	2
Oh7B	Inbred	13	64.85	6.75	0.94	15.28	6.85	3.70	1.00	1
P39	Inbred	16	70.53	7.52	0.00	8.84	5.07	6.93	0.00	3
Tx303	Inbred	11	63.08	7.07	0.00	18.33	2.53	7.35	0.00	3
Tzi8	Inbred	10	83.82	1.17	0.00	10.38	1.82	2.25	0.00	2
Teosinte Mean		26	90.40	2.27	0.20	2.91	1.15	1.40	0.17	6
Landrace Mean		17	76.52	4.55	0.71	9.66	3.51	2.42	0.39	3
Inbred Mean		15	72.60	3.08	1.70	13.62	4.21	3.02	0.29	2
ANOVA										
Among groups		***	***	**	ns	***	***	***	ns	***
Teo. versus LR		***	***	**	ns	***	***	***	ns	***
Teo. versus inbreds		***	***	ns	ns	***	***	***	ns	***
LR versus inbreds		**	ns	*	ns	**	ns	*	ns	*

ns not significantly different at $P = 0.05$

*, **, *** Significantly different at $P < 0.05$, 0.01, and 0.001, respectively

[a] Peak designation corresponds to labeling of Figs. 3 and 4, and Supplemental Figs. 1, 3, and 4

[b] Letter–number designation refers to the nomenclature proposed by Wilson (1991)

[c] Molecular mass and Greek letter designation refers to the genetic classification proposed by Thompson and Larkins (1989)

correlation between the kernel composition traits and the alpha zein and total zein fractions. Generally, the lower the seed weight, the higher the protein content and more alpha zein.

Percent endosperm (as a measure of the endosperm to embryo ratio) was significantly correlated with crude fat, ash, number of alpha zein peaks, and percentage peak 3 (Table 3). Seed weight was significantly ($P < 0.01$) correlated with alpha zein content and peaks 2–5. A number of significant correlations were detected among the zein profile traits (Table 3).

Principal component analysis revealed five components that explained 77.2% of the variation (Fig. 5). The first component captures the trend of low seed weight, low carbohydrate content, high protein content, and high alpha zein content, as determined above. Again, this likely reflects the striking differences observed for these traits in teosinte. The second principal component, while much less concise, encompasses the relationship between low values for percent endosperm, percent of alpha peaks, and percent of peak 2, and high values for fat and percent peaks 1, 4, and 5. Components 3 and 4 describe relationships between the various zein peaks as revealed by the correlations, and component 6 relates percent endosperm with fat and seed weight (data not shown).

Discussion

There is increased demand placed on maize today compared to a few decades ago when maize was primarily used as a feed source in the US. Modern technologies allow the creation of a wide array of food, feed, fuel, and industrial products from maize, and improvements in maize must meet these new and increased demands. Researchers have the potential to create designer maize, as each end use calls for different kernel qualities. Redesigning maize to meet these challenges may require the introduction of novel alleles not presently found in today's maize.

The events that led to the domestication of maize from teosinte involved artificial selection for traits such as inflorescence and plant architecture, ear architecture, and kernel architecture (reviewed in Doebley 2004; Doebley et al. 1995; Dorweiler et al. 1993). The genes underlying traits

Table 3 Significance of Pearson correlation coefficients between kernel quality, seed characteristics, and zein profiles for a panel of diverse teosinte, landrace, and inbred lines

	Protein	Fat	Fiber	Ash	Carb.	Seed weight	Endosperm (%)	No. of alpha peaks	Alpha (%)	Peak 1 (%)	Peak 2 (%)	Peak 3 (%)	Peak 4 (%)	Peak 5 (%)	Peak 6 (%)	Unknown
Moisture	***	***	ns	***	***	***	ns	***	***	ns	ns	***	**	**	ns	**
Protein		***	***	***	***	***	ns	***	***	*	*	***	***	***	ns	***
Fat			**	***	***	***	*	***	***	ns	ns	***	**	ns	ns	***
Fiber				***	***	***	ns	***	***	ns	ns	***	ns	ns	ns	***
Ash					***	***	**	***	***	ns	*	***	***	***	ns	***
Carbohydrate						***	ns	***	***	ns	*	***	***	***	ns	***
Seed weight							ns	***	***	ns	*	***	***	***	ns	***
Endosperm (%)								**	ns	ns	ns	*	ns	ns	ns	ns
No. of alpha peaks									***	ns	ns	***	***	***	ns	***
Alpha (%)										**	*	***	***	***	ns	***
Peak 1 (%)											***	ns	*	**	ns	ns
Peak 2 (%)												ns	**	ns	ns	ns
Peak 3 (%)													***	***	ns	***
Peak 4 (%)														**	ns	***
Peak 5 (%)															ns	**
Peak 6 (%)																ns

Grey and white cells represent negative and positive correlation coefficients, respectively

ns not significant at $P = 0.05$

*, **, *** Significant at $P < 0.05$, 0.01, and 0.001, respectively

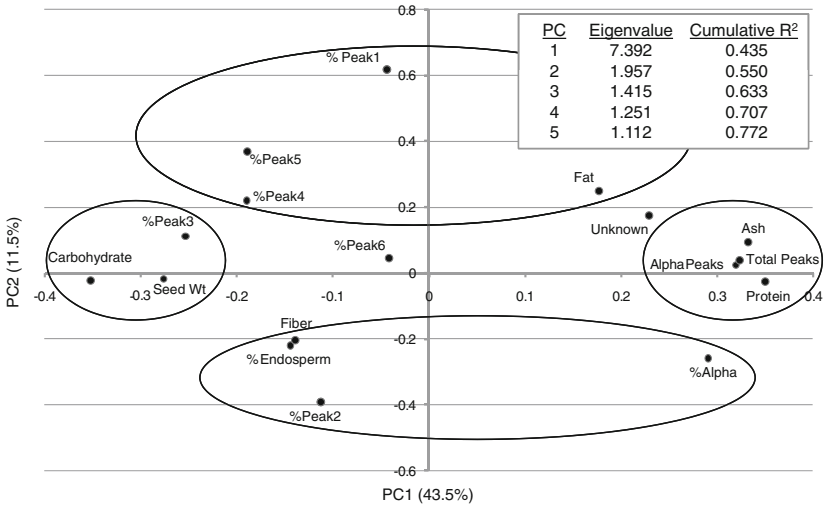

Fig. 5 Principal component analysis of 17 kernel composition, seed weight, percent endosperm, and zein profile traits. Eigen values for five principal components and the cumulative variation explained are displayed (*inset*). Plot of the first two components and the trait patterns they describe are enclosed in *ovals*

targeted during artificial selection have significantly reduced genetic variation in landraces and/or modern inbred lines compared to teosinte and relative to the rest of the maize genome (Clark et al. 2004; Wang et al. 1999; Wright et al. 2005). Large-scale sequencing studies indicate that 2–4% of the maize genome, corresponding to approximately 1,000–1,200 genes, has experienced selection during domestication or plant breeding (Wright et al. 2005). Conversely, 96–98% of maize genes are neutral genes that retain high levels of diversity in modern inbred lines as compared to teosinte. Recapturing variability in selected genes from teosinte and/or landraces would enhance the variability present in neutral genes in modern maize, and together could be exploited for continued maize improvement. Maize geneticists and breeders must work together to distinguish between selected and neutral genes, assay the allelic variation present in diverse inbreds, landraces, and teosinte, and reintroduce these alleles into breeding programs in a manner that increases the efficiency of germplasm use. It is also necessary to characterize the phenotypic diversity in agronomically relevant traits for the various gene pools, namely landraces and teosinte, in order to gain insight into which germplasm pools harbor valuable phenotypic variation.

In the current study, we chose to evaluate kernel quality and seed characteristics, with an emphasis on protein content and quality. It is apparent that kernel traits (e.g., seed size and starch production) were targets of selection during domestication and/or plant breeding (Whitt et al. 2002). Of primary concern regarding protein quality is the poor amino acid balance, due to low abundance of tryptophan, lysine, and methionine in zein proteins. Other seed characteristics

such as seed size are relevant as they too were likely targets of selection.

Generally, we found wide variability for all traits measured that often correlated with germplasm group. The fact that plants grown in this experiment were allowed to open pollinate may have contributed to variability within the sample. The existence of xenia effects, i.e., the direct effect of pollen source on in the character of the resulting kernel, was described over a century ago (Focke 1881; Webber 1900). While early reports of xenia focused on qualitative traits such as kernel color (yellow vs. white, or purple vs. white) and kernel type (normal vs. sugary), later reports of xenia included quantitative traits such as oil content (Curtis et al. 1956). Xenia effects of 8–13% were reported for kernel weight in normal endosperm materials (Bulant and Gallais 1998; Bulant et al. 2000). However, the existence and extent of xenia effects remain largely unknown for many traits including zein profiles. For this experiment, there was no feasible way to make controlled pollinations so as to eliminate xenia effects. Teosinte and several of the landraces are photoperiod sensitive and must be grown in a day-neutral setting such as a growth chamber or short-day winter nursery site. Manual pollinations cannot be made effectively on teosinte, leaving open pollination the only means to obtain adequate seed quantities of all materials growing in the same environment.

Kernel composition and seed characteristics

The striking differences in kernel protein and carbohydrate between teosinte and the landrace and inbred groups are highly significant, both statistically and agronomically (Fig. 1). Our results are consistent with those of Paulis and Wall (1977), where they found protein levels of 28.7% in their *parviglumis* teosinte accession. The protein content of teosinte was almost as high as protein levels obtained by the Illinois Long-Term Selection Experiment (Dudley 2007) without a corresponding change in the embryo to endosperm ratio (Bjarnason and Pollmer 1972).

The marked increase of carbohydrate during the progression from teosinte to landraces is consistent with evidence of artificial selection in the starch pathway (Whitt et al. 2002), though crude carbohydrate content obtained by proximate analysis is not perfectly correlated to starch ($r = 0.82$; Flint-Garcia, unpublished data). This dramatic change in kernel composition may have occurred in concert with the loss of the hard fruitcase that restricts the growth of the seed. A theoretical scenario is as follows: mutations in the *teosinte glume architecture1* gene opened the stony fruitcase slightly (Dorweiler et al. 1993). Increased access to the teosinte seed would have been desirable by humans, and an initial population was likely isolated from the progenitor population. Artificial selection for increased seed size occurred, likely through increased starch production. Furthermore, modern maize breeding has selected primarily on increased yield, which translates into higher starch content at the expense of protein. Though teosinte per se contains less starch than modern maize (53% in teosinte vs. 73% in inbred lines), teosinte still harbors more genetic variability in the genes that underlie expression of starch traits (Whitt et al. 2002). This variability may be valuable for maize improvement for human and animal nutrition, as well as industrial application including the production of biofuels.

Zein profiles

Given that various zein proteins have different amino acid compositions, and our observation that several novel zeins accumulate in teosinte and not in landraces or inbred lines (Fig. 4; Supplemental Fig. 1), this study suggests that teosinte may potentially contribute genes for improvement of amino acid content as previously suggested (Swarup et al. 1995; Wang et al. 2008) or demonstrated with ssp. *mexicana* (Swarup et al. 1995; Wang et al. 2008). Many novel peaks were observed in the gamma region of the teosinte chromatographs (Supplemental Fig. 3). Further studies are needed to determine the amino acid sequence and nutritional value of the novel zein proteins. If these novel peaks are found to have potential nutritional value (i.e., increased methionine or cysteine content), genetic studies will be required to elucidate their gene sequence and regulation. In addition, there were differences in abundance of known zeins between teosinte and the other germplasm groups. Therefore, teosinte may be valuable in genetic studies attempting to define the regulation of the beta or gamma zein proteins.

The alpha zein gene family is the result of a complex series of duplication events (Song and Messing 2003) resulting in very large multigene families in maize. However, the alpha zeins appear to be even more complex in teosinte than in maize (Supplemental Fig. 2). It has been shown that many alpha zein genes in modern maize are inactive (Thompson et al. 1992), so it may be that some of these genes are active in teosintes. Alternatively, teosinte may contain duplicated chromosomal segments that have been lost or diverged in function in the course of development of modern maize (Thompson et al. 1992). It is possible that some of the novel teosinte peaks in the alpha region are mere variants of known alpha zein proteins. Additions and deletions of amino acids change the hydrophobicity of the protein (e.g., methionine and tryptophan are hydrophobic while lysine is hydrophilic), and thus result in altered elution times when compared to known alpha zein peaks in inbreds. Some of the shifts in the zein profiles observed in this study may indeed reflect important amino acid substitutions. It is also possible that these novel peaks are unrelated or more distantly related proteins. Further characterization

of interesting zein proteins is therefore required before any conclusion can be drawn regarding these novel peaks in the alpha region.

Implications for maize breeding and genetics

For all kernel traits analyzed, there was substantial variation both *between* and *within* germplasm groups. This phenotypic variation may be useful for both improving modern maize and for genetic studies to investigate the genetic architecture of these agronomically relevant traits. The inbred lines used in this study are the parental lines of the maize NAM population. NAM was designed to enable high power and resolution QTL mapping through joint linkage-association analysis (Buckler et al. submitted; McMullen et al. 2009; Yu et al. 2005). By integrating genetic design, natural diversity, and genomics technologies, NAM analysis of the kernel traits surveyed in the current study will provide tremendous opportunities to link molecular variation with phenotypic variation and mine maize variation for kernel improvement.

The germplasm enhancement of maize (GEM) project is a collaboration of maize breeders and scientists from other disciplines from public research institutions and the seed industry with the objective of incorporating novel germplasm, primarily from landraces, into elite germplasm (Pollak 2003). In the last 8 years, 65 GEM releases have improved amino acid profiles (index of lysine, methionine, and tryptophan), oil content greater than 4.5%, protein content greater than 13%, and unique starch thermal properties (http://www.public.iastate.edu/~usda-gem). Likewise, phenotypic selection in Burr's White population for high and low protein and oil in the Illinois Long-Term Selection Experiments (Dudley 2007) clearly demonstrate the power of phenotypic selection. In these cases, the original landraces contained variation for the genes under selection.

However, for "domestication genes," diversity is severely reduced both inbreds and landraces (Yamasaki et al. 2005), leaving teosinte as the only source for variability. Maize and teosinte can be crossed readily, and populations have been derived for QTL mapping and cloning experiments (Briggs et al. 2007). Since the selection status of all structural and regulatory genes involved in kernel composition (starch, protein, and oil) and the zein protein accumulation is currently unknown, it is possible that teosinte contains diversity for genes controlling these traits that is not present in landraces or inbred lines. This diversity may be valuable for breeding efforts.

Breeding projects can immediately utilize teosinte germplasm to introduce novel zein proteins. Indeed, promising results have been reported by Wang et al. (2008) for increased protein content and amino acid composition using *Z. mays* ssp. *mexicana* germplasm in a maize breeding program. As discussed earlier, however, teosinte has many undesirable attributes, both in the plant and the seed, that must be selected against. The use of teosinte introgression materials will greatly assist in this background selection and speed the process of deriving agronomically acceptable lines. However, for genetic and biochemical studies, substantial research must be conducted to identify the novel zein peaks, characterize them in terms of their amino acid content, and to determine the regulatory mechanisms governing them. Additional efforts are underway to create near isogenic line introgression libraries of ten teosinte accessions in a maize background (Flint-Garcia, unpublished data) for bridging teosinte and elite germplasm in breeding programs, and for the genetic analysis of seed protein content traits.

Conclusion

When comparing kernel traits between teosinte, landraces, and inbred lines, there is variation in (1) the restrictive nature of the stony fruitcase that surrounds the teosinte kernel which is absent in maize, (2) kernel size, (3) starch and protein content, (4) zein profiles, and (5) amino acid profiles. Teosinte offers a unique combination of these traits that result in higher protein and lower carbohydrate than maize, and these traits may be exploited for maize improvement. However, major obstacles exist concerning the use of teosinte in breeding programs: the hard restrictive fruitcase, seed shattering, small seed size, and photoperiod sensitivity. Can these traits be separated to yield a large seeded, high protein, nutritionally balanced line? It is up to the creativity of geneticists and breeders to explore these options in an effort to meet the demands of today and tomorrow.

Acknowledgments We would like to thank Susan Melia-Hancock (USDA-ARS in Columbia, MO) for collecting data regarding seed weight and embryo-endosperm ratio; these are especially challenging traits to evaluate in teosinte. We would also like to thank Hari Krishnan (USDA-ARS in Columbia, MO) for conducting preliminary experiments leading to this larger study, and Adrienne Moran Lauter (USDA-ARS in Ames, IA) for technical assistance. This research was supported by a National Science Foundation Plant Genome Award (DBI0321467), the USDA-CSREES special appropriation "Plant Biotechnology Iowa," and by research funds provided by USDA-ARS. Names of products are necessary to report factually on available data; however, neither the USDA nor any other participating institution guarantees or warrants the standard of the product, and the use of the name does not imply approval of the product to the exclusion of others that may also be suitable.

References

AOAC (2006) Official methods of analysis of AOAC INTERNATIONAL, 18th edn. AOAC INTERNATIONAL, Gaithersburg

Bjarnason M, Pollmer WG (1972) The maize germ: its role as a contributing factor to protein quantity and quality. Z Pflanzenzuchtg 68:83–89

Briggs WH, McMullen MD, Gaut BS, Doebley J (2007) Linkage mapping of domestication loci in a large maize teosinte backcross resource. Genetics 177:1915–1928

Bulant C, Gallais A (1998) Xenia effects in maize with normal endosperm: I. Importance and stability. Crop Sci 38:1517–1525

Bulant C, Gallais A, Matthys-Rochon E, Prioul JL (2000) Xenia effects in maize with normal endosperm: II. Kernel growth and enzyme activities during grain filling. Crop Sci 40:182–189

Clark R, Linton E, Messing J, Doebley J (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc Natl Acad Sci USA 101:700–707

Coleman CE, Herman EM, Takasaki K, Larkins BA (1996) The maize gamma-zein sequesters alpha-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. Plant Cell 8:2335–2345

Curtis JJ, Brunson AM, Hubbard JE, Earle FR (1956) Effect of the pollen parent on oil content of the corn kernel. Agron J 48:551–555

Doebley J (2004) The genetics of maize evolution. Annu Rev Genet 38:37–59

Doebley J, Stec A, Gustus C (1995) *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141:333–346

Dorweiler J, Stec A, Kermicle J, Doebley J (1993) *Teosinte glume architecture 1*: a genetic locus controlling a key step in maize evolution. Science 262:233–235

Ducrocq S, Madur D, Veyrieras J-B, Camus-Kulandaivelu L, Kloiber-Maitz M, Presterl T, Ouzunova M, Manicacci D, Charcosset A (2008) Key impact of *Vgt1* on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. Genetics 178:2433–2437

Dudley JW (2007) From means to QTL: the Illinois long-term selection experiment as a case study in quantitative genetics. Crop Sci 47:S20–S31

Focke WO (1881) Die Pflanzen-Mischlinge. Ein Beitrag zur Biologie der Gewächse

Hamblin MT, Casa AM, Sun H, Murray SC, Paterson AH, Aquadro CF, Kresovich S (2006) Challenges of detecting directional selection after a bottleneck: lessons from *Sorghum bicolor*. Science 173:953–964

Hopkins CG (1899) Improvement in the chemical composition of the corn kernel. Ill Agric Expt Sta Bul 55:205–240

Huang S, Adams WR, Zhou Q, Malloy KP, Voyles DA, Anthony J, Kriz AL, Luethy MH (2004) Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. J Agric Food Chem 52:1958–1964

Huang S, Kruger DE, Frizzi A, D'Ordine RL, Florida CA, Adams WR, Brown WE, Luethy MH (2005) High-lysine corn produced by the combination of enhanced lysine biosynthesis and reduced zein accumulation. Plant Biotechnol J 3:555–569

Hyten DL, Song Q, Zhu Y, Choi I-Y, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB (2006) Impacts of genetic bottlenecks on soybean genome diversity. Proc Natl Acad Sci USA 103:16666–16671

Kaiser HF (1960) The application of electronic computers to factor analysis. Educ Psychol Meas 20:141–151

Lai J, Messing J (2002) Increasing maize seed methionine by mRNA stability. Plant J 30:395–402

Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci USA 99:6080–6084

McMullen MM, Kresovich S, Buckler ES, Holland JB, Sanchez Villeda H, Bradbury P, Li H, Sun Q, Bottoms C, Flint-Garcia S, Hanson M, Acharya C, Yates H, Mitchell SE, Browne C, Eller M, Guill K, Kroon D, Lepak N, Romero S, Salvo S, Peterson B, Jones E, Smith S, Brown P, Pressoir G, Oropeza Rosas M, Harjes C, Glaubitz JC, Goodman M, Ware D (2009) Genetic properties of the maize nested association mapping population (accepted)

Melcher U, Fraij B (1980) Methionine-rich protein fraction prepared by cryoprecipitation from extracts of corn meal. J Agric Food Chem 28:1334–1336

Paulis JW, Wall JS (1977) Comparison of the protein compositions of selected corns and their wild relatives, teosinte and *Tripsacum*. J Agric Food Chem 25:265–270

Phillips RL, Suresh J, Olsen M, Krone T (2008) Registration of high-methionine versions of maize inbreds A632, B73, and Mo17. J Plant Regist 2:243–245

Pollak LM (2003) The history and success of the public-private project on germplasm enhancement of maize (GEM). Adv Agron 78:45–87

Pollmer WG, Eberhard D, Klein D (1978) Inheritance of protein and yield of grain and stover in maize. Crop Sci 18:757–759

Prasanna BM, Vasal SK, Kassahun B, SN N (2001) Quality protein maize. Curr Sci 81:1308–1319

Rossi V, Hartings H, Thompson RD, Motto M (2001) Genetic and molecular approaches for upgrading starch and protein fractions in maize kernels. Maydica 46:147–158

SAS Institute Inc. (1999–2001) SAS® Proprietary Software Release 8.2, Cary

Sodek L, Wilson CM (1971) Amino acid compositions of proteins isolated from normal, opaque-2, and floury-2 corn endosperms by a modified Osborne procedure. J Agric Food Chem 19:1144–1150

Song R, Messing J (2002) Contiguous genomic DNA sequence comprising the 19-kDa zein gene family from maize. Plant Physiol 130:1626–1635

Song R, Messing J (2003) Gene expression of a gene family in maize based on noncollinear haplotypes. Proc Natl Acad Sci USA 100:9055–9060

Song R, Llaca V, Linton E, Messing J (2001) Sequence, regulation, and evolution of the maize 22-kDa alpha zein gene family. Genome Res 11:1817–1825

Swarup S, Timmermans MC, Chaudhuri S, Messing J (1995) Determinants of the high-methionine trait in wild and exotic germplasm may have escaped selection during early cultivation of maize. Plant J 8:359–368

Tang T, Lu J, Huang J, He J, McCouch SR, Shen Y, Kai Z, Purugganan MD, Shi S, Wu C-I (2006) Genomic variation in rice: genesis of highly polymorphic linkage blocks during domestication. PLoS Genet 2:e199

Thompson GA, Larkins BA (1989) Structural elements regulating zein gene expression. BioEssays 10:108–113

Thompson GA, Siemieniak DR, Sieu LC, Slightom JL, Larkins BA (1992) Sequence analysis of linked maize 22 kDa alpha-zein genes. Plant Mol Biol 18:827–833

Vasal SK (2000) High quality protein corn. In: Hallauer AR (ed) Specialty corns, 2nd edn. CRC Press, Boca Raton

Wang RL, Stec A, Hey J, Lukens L, Doebley J (1999) The limits of selection during maize domestication. Nature 398:236–239

Wang L, Xu C, Qu M, Zhang J (2008) Kernel amino acid composition and protein content of introgression lines from *Zea mays* ssp. *mexicana* into cultivated maize. J Cereal Sci 48:387–393

Watson SA (2003) Description, development, structure, and composition of the corn kernel. In: White PJ, Johnson LA (eds) Corn: chemistry and technology, 2nd edn. American Association of Cereal Chemists, St. Paul

Webber HJ (1900) Xenia, or the immediate effect of pollen, in maize. USDA Bull 22:1–44

Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler ES (2002) Genetic diversity and selection in the maize starch pathway. Proc Natl Acad Sci USA 20:12959–12962

Wilson C (1991) Multiple zeins from maize endosperms characterized by reversed-phase high performance liquid chromatography. Plant Physiol 95:777–786

Wilson DR, Larkins BA (1984) Zein gene organization in maize and related grasses. J Mol Evol 20:330–340

Woo Y-M, Hu DW-N, Larkins BA, Jung R (2001) Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. Plant Cell 13:2297–2317

Wright SI, Vroh Bi I, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS (2005) The effects of artificial selection on the maize genome. Science 308:1310–1314

Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. Genetics 150:899–909

Buckler ES, Holland JB, Bradbury P, Acharya C, Brown P, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman M, Harjes C, Guill K, Kroon D, Larsson S, Lepak N, Li H, Mitchell SE, Pressoir G, Peiffer J, Oropeza Rosas M, Rocheford T, Romay C, Romero S, Salvo S, Sanchez Villeda H, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MM The genetic architecture of Maize flowering time (submitted)

Yamasaki M, Tenaillon MI, Vroh Bi I, Schroeder SG, Sanchez-Villeda H, Doebley JF, Gaut BS, McMullen MD (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. Plant Cell 17:2859–2872

Yu J, Arbelbide M, Bernardo R (2005) Power of in silico QTL mapping from phenotypic, pedigree, and marker data in a hybrid breeding program. Theor Appl Genet 110:1061–1067

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551