

## CHARACTERIZATION OF A MAJOR MAIZE DOMESTICATION QTL ON THE SHORT ARM OF CHROMOSOME 1

P. Quijada<sup>1</sup>, L.M. Shannon, J.C. Glaubitz, A.J. Studer, J. Doebley\*

*Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin, USA*

Received September 15, 2009

**ABSTRACT** - Inheritance of the basic morphological differences between primitive maize and teosinte is mainly controlled by genes falling within five or six regions of the maize genome. Herein, we focus on one of these regions, the short arm of chromosome 1, for which we created a nearly-isogenic line (NIL) consisting of a teosinte chromosome segment (T1S) introgressed into maize inbred line (W22). By crossing this NIL with W22 and then selfing for 6 generations, 135 nearly isogenic recombinant inbred lines (NIRILs) were recovered for the T1S chromosomal region. We mapped the cross-overs within the T1S segment for each NIRIL and collected phenotypic data for domestication and other traits on each NIRIL. Using these data, we mapped 15 quantitative trait loci (QTLs) controlling several domestication, developmental, and seed-related traits. The proportion of the phenotypic variance explained by the QTLs for a trait ranged from 7% to 65% and the direction of the effects of most QTLs agreed with the expectation that teosinte alleles should be associated with teosinte-like phenotypes. Most QTL for domestication-related traits co-localized to the same 15 cM region near the center of the introgressed segment. These QTL could represent either a single major gene with pleiotropic effects or several tightly linked genes. Positional cloning studies are now underway to distinguish these two hypotheses.

**KEY WORDS:** Domestication; QTL, Quantitative Trait Loci; NIRIL, Nearly isogenic recombinant inbred line; Teosinte; Maize.

### INTRODUCTION

Despite the fact that maize (*Zea mays* ssp. *mays*) and its closest relatives, the annual teosintes (*Z.*

*mays* ssp. *parviglumis* and ssp. *mexicana*), belong to the same biological species, they differ strikingly in their morphology, particularly of their female inflorescences or ears (DOEBLEY, 2004). Since maize exists only as a cultivated plant and teosinte is only known as a wild plant, BEADLE (1972, 1980) proposed that maize is a domesticated form of teosinte and that their morphological differences resulted from human selection during domestication. BEADLE (1972, 1980), using an F<sub>2</sub> population of 50,000 plants from a cross between *Z. mays* ssp. *mexicana* (race Chalco) and the primitive maize race Chapalote, reported that maize-like and teosinte-like segregants were recovered at a frequency of 1/500. From this, he deduced that five major and independently inherited genes distinguish maize and teosinte and thus he viewed the origin of maize as the result of a small number of mutations each of major phenotypic effect.

Over the past 15 years, our research group has carried out a series of quantitative trait locus (QTL) mapping experiments to study the inheritance of the key traits distinguishing maize and teosinte (DOEBLEY, 2004). Following Beadle's logic, we initially used primitive maize races (Chapalote and Reventador) as the maize parents of our maize-teosinte populations, allowing us to map genes involved in maize domestication instead of those involved in maize improvement (DOEBLEY *et al.*, 1990, 1994, 1995; DOEBLEY and STEC, 1991, 1993). For the key traits that distinguish the inflorescences of maize and teosinte, we found a total of 50 and 64 significant QTLs in the Reventador x teosinte (ssp. *parviglumis*) and Chapalote x teosinte (ssp. *mexicana*) populations, respectively. In these two populations, large effect QTLs were found in only six chromosomal regions (chromosomes 1S, 1L, 2S, 3L, 4S, and 5S) which control the key differences between maize and teosinte. These results are remark-

<sup>1</sup> Current address: Seminis Vegetable Seeds, Inc., Woodland, California, USA.

\* For correspondence (fax: +1 608-262-2976; e.mail: jdoebly@wisc.edu).

ably concordant with Beadle's hypothesis that the major differences between maize and teosinte result from the actions of about five major genes. However, they do not rule out the possibility that the differences could result from the effects of a small number of blocks of multiple linked QTLs.

Formal proof that a single, major gene underlies each strong domestication QTL peak is provided if the causative factor can be positionally cloned. To date, the genes underlying two of the major domestication QTL have been successfully cloned, *tb1* (DOEBLEY *et al.*, 1997) and *tga1* (WANG *et al.*, 2005). These two genes are currently under active investigation to shed light upon their modes of molecular action and their roles in the evolution and domestication of maize. To determine if the remaining four major domestication QTL represent single major genes or tightly linked groups of genes each with small effects, we initiated fine-mapping studies by developing four near-isogenic lines (NILs). Each NIL has one of the teosinte genomic regions of interest isolated in a common maize (inbred line W22) genetic background. Each NIL was crossed again to W22 and a population of nearly isogenic recombinant inbred lines (NIRILs) was developed with recombination events distributed throughout the introgressed region of interest.

In this study, we focus on the NIRIL population constructed to allow mapping within the teosinte introgression on the short arm of chromosome one (T1S). In previous studies (DOEBLEY *et al.*, 1990; DOEBLEY and STEC, 1991, 1993; BRIGGS *et al.*, 2007), this chromosomal region has been repeatedly associated with the multiple domestication traits. The newly constructed NIRIL population was phenotyped for domestication, developmental, and kernel-related traits and genotyped with a set of PCR-based markers (SSRs and indels) in order to confirm previously detected QTLs, refine their locations and facilitate their future positional cloning.

## MATERIALS AND METHODS

### Plant materials

A maize near isogenic line (NIL) was developed by introgressing a teosinte *Zea mays* L. ssp. *parviglumis* (Iltis and Cochrane collection 81) chromosome 1 short arm (T1S) segment into the maize inbred line W22 via six generations of backcrossing, followed by one generation of selfing to recover a T1S homozygote. During the creation of this NIL, molecular markers were used both to follow the target chromosomal segment of the chromosome 1 short arm as well as to eliminate teosinte chromosome segments at other major domestication QTL identified by

DOEBLEY and STEC (1993). This NIL (W22-T1S) was then crossed to W22 to develop a population of nearly isogenic recombinant inbred lines (NIRIL). From a single F<sub>1</sub> plant of this cross, 135 F<sub>2</sub> plants were selfed for five additional generations to create a set of 135 highly, but not completely, homozygous NIRILs.

### Molecular markers and linkage map construction

To guide the introgression of the T1S genomic segment into the W22 background, we employed six RFLP loci from the T1S region (*php20640*, *ts2*, *umc11a*, *umc13*, *umc76a*, *umc157a*) that were previously shown by DOEBLEY and STEC (1993) to associate with several domestication related traits. To genotype the 135 NIRILs, we assembled a set of 72 PCR-based markers (56 SSRs and 16 indels) using the positions of the six RFLP loci on the IBM2 (intermated B73 x Mo17 population) map available at the Maize Genetics and Genomics Database as a guide (LAWRENCE *et al.*, 2007). Of these 72 markers, 25 (24 SSRs and 1 indel; see Results for the marker names) were polymorphic and used to genotype the 135 NIRILs. The resulting data were used to construct a genetic map with the aid of the software JoinMap v 3.0 (VAN OORJEN and VOORRIJS, 2001).

For the 3375 genotypes (25 markers x 135 lines), 5.2% of the genotypes were heterozygous and another 1.9% had missing data. Excluding cross-overs involving the heterozygous markers, there were a total 139 cross-overs in the data set or slightly more than one per line. The distribution of cross-overs among lines was as follows: 0 (15 lines), 1 (57 lines), 2 (27 lines), 3 (8 lines) and 4 (1 line).

### Phenotypic data collection

The NIRILs, along with their parents (W22 and W22-T1S), were each grown in single rows containing 10 plants per row in a completely randomized block design with three replicates at the University of Wisconsin West Madison Agricultural Research Station, Madison, WI, USA during summer 2005.

The following ten traits were evaluated: days to pollen (DTP; number of days after planting when 50% of the plants in a row were shedding pollen), plant height (PLHT; the distance, in cm, from the ground to the tip of the tassel), tillering (TILL; the ratio of the sum of tiller heights/plant height), ear number (EN; number of ears showing silk on a plant), ear length (EL; distance, in cm, from the base to the tip of the ear), ear diameter (ED; diameter, in cm, of the midsection of each ear), 10-kernel length (10KL; length, in cm, of 10 consecutive kernels in a single rank along the ear), cupules per rank (CUPR; number of cupules in a single rank from base to the tip of the ear), kernel weight (KW; kernel weight, in mg, derived from the average weight of 50 kernels), and kernel type (KT; 0 = non-dent, 1 = semi-dent, 2 = dent). The latter six traits were all measured on the uppermost, well-formed ear of each plant. In contrast, DTP was evaluated for an entire row and each of the remaining three traits, PLHT, TILL, and EN, were averages from approximately five plants per row. We consider TILL, EN, EL, ED, CUPR, 10KL, KT and KW to be domestication and improvement traits, and DTP and PLHT to be diversification traits (under selection during the diversification of maize).

### Data analysis

To estimate least-square means for each NIRIL, we used the MIXED procedure of SAS (LITTELL *et al.*, 1996). NIRIL (or parental) lines were considered as fixed effects while replicates (rows; also referred to as plots) and samples (plants) within NIRIL or parental lines were treated as random effects. The linear model used was

$$Y_{ijk} = m + a_i + b_j + e_{ij} + d_{ijk}$$

where  $Y_{ijk}$  is the trait value for the  $k$ th plant in the  $j$ th block on the  $i$ th NIRIL,  $m$  is the overall mean of the experiment,  $a_i$  is the NIRIL (or parental) line effect,  $b_j$  is the block effect,  $e_{ij}$  is the experimental error (random variation among plots), and  $d_{ijk}$  is the (within-plot) sampling error. The least-square means estimates for each trait were used to calculate phenotypic correlations using the SAS CORR procedure (SAS INSTITUTE, 2000) and as phenotypes in the QTL mapping analyses. The significance of each phenotypic correlation was determined via a t-test of the coefficient of correlation (EDWARDS, 1976), using the Bonferroni-Holm sequential method (RICE, 1989) to adjust the significance levels and thereby account for multiple testing. We calculated broad-sense heritabilities ( $H^2$ ) on a plot-basis as

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ge}^2 + \sigma_e^2}$$

where  $\sigma_g^2$  is the genotypic variance,  $\sigma_{ge}^2$  is the genotype x environment interaction variance, and  $\sigma_e^2$  is the experimental error variance. We used the MIXED procedure of SAS to fit a linear random-effect model for the estimation of the variance components (LITTELL *et al.*, 1996).

QTL mapping was conducted in the R/qtl module of the R statistical computing package (BROMAN *et al.*, 2003). For each trait, an initial QTL scan was performed using simple interval mapping with a 1 cM step (LANDER and BOTSTEIN, 1989) and the position of the highest LOD score recorded. Statistical significance of the peak LOD score was assessed using 10,000 permutations of the data (DOERGE and CHURCHILL, 1996). Then, the position and effect of the QTL was refined using the multiple-imputation method (SEN and CHURCHILL, 2001) by executing the “sim.geno” command (0.5 cM steps, 7000 joint genotype distribution imputations, and an assumed genotyping error rate of 0.001), followed by the “fitqtl” command. To search for additional QTL, the “addqtl” command was used. If a second QTL was detected, then “fitqtl” was used to test a model containing both

QTL and their interaction effect. If both QTL remained significant, the “refineqtl” command was used to re-estimate the QTL positions based on the full model including both QTL. Interaction effects were never significant and thus not included in the full model. Finally, each QTL was removed from the model and then added back using the “addqtl” command to re-confirm its significance and position. Approximate confidence intervals for the locations of the identified QTL were obtained via 1.5 LOD support intervals to each side of the position of the LOD maximum. Since R/qtl does not allow heterozygous genotypes for RILs, heterozygous genotypes were converted to missing data. The phenotypic and genotypic data from this paper are available on-line at [www.panzea.org](http://www.panzea.org) (ZHAO *et al.*, 2006) or from the corresponding author.

## RESULTS

### Quantitative trait variation

Significant phenotypic variation among NIRILs was detected for all the traits reported here; additional traits (disarticulation of the ears, pedicellate spikelet, and rank) were also measured but did not display significant variation (data not shown). Most of the reported traits appeared to be normally distributed, with the exception of days to pollen and number of ears, which were slightly skewed toward maize phenotypes (Fig. 1). For all the traits, the parental lines had phenotypic values close to the average value of the NIRILs. The highest phenotypic correlations ( $r_p$ ) were found among some of the domestication and improvement traits such as KW, CUPR, ED, and EL; longer ears tended to have more cupules in a single rank and heavier kernels tended to be found on ears of greater diameter (Table 1).

TABLE 1 - Heritabilities and phenotypic correlation coefficients among traits.

	Days to Pollen	Plant height	Tillering	Number of ears	Ear length	Ear diameter	Cupules per rank	10-kernel length	Kernel type	Kernel weight
Heritability <sup>a</sup>	0.52	0.42	0.34	0.24	0.68	0.52	0.66	0.41	0.55	0.72
Days to pollen		-0.20	0.16	0.15	0.21	-0.49 ***	0.19	0.21	-0.47 ***	-0.62 ***
Plant height			-0.03	-0.07	0.10	0.22	0.07	-0.17	0.10	0.17
Tillering				0.13	-0.11	-0.12	-0.10	0.03	-0.21	-0.20
Number of ears					-0.19	0.00	-0.06	-0.06	-0.10	-0.03
Ear length						-0.46 ***	0.87 ***	0.03	-0.46 ***	-0.48 ***
Ear diameter							-0.49 ***	-0.02	0.55 ***	0.74 ***
Cupules per rank								-0.24	-0.39 ***	-0.58 ***
10-kernel length									-0.15	0.10
Kernel type										0.55 ***

<sup>a</sup> Heritabilities were calculated on a plot-basis.

\*, \*\*, \*\*\* Significant at the 0.05, 0.01, and 0.001 probability levels, respectively. Significance levels were corrected for multiple testing according to the Bonferroni-Holm sequential method.

Late-flowering lines tended to have slender ears, less dented and smaller kernels as reflected by the phenotypic correlations between these traits and DTP. In contrast, plant height, tillering, ear number, and 10-kernel length did not show any significant correlations among themselves or with the other traits (Table 1). Heritabilities of the traits are low to moderate, ranging from 0.24 to 0.72 (Table 1).

### QTL analysis

The genetic map of our 25 PCR-based markers spanned 68.4 cM across the introgressed T1S genomic region (Fig. 2). In total, 15 significant QTLs were detected within this region for the ten traits examined in this study (Table 2). For five of the 10 traits, we detected two QTL per trait which in some cases were surprising close together. For example, *kt1.1* and *kt1.2* are only 14 cM apart and *dtp1.1* and *dtp1.2* are only 25 cM apart. The percentage of the phenotypic variation ( $R^2$ ) explained by an individual QTL ranged from 6.6 (*10kl1.1*) to 52.3 (*kw1.1*). The directions of allelic effects at most of the QTLs agreed with the expectation that the teosinte (T) alleles should be associated with teosinte-like phenotypes and the maize (M) alleles with maize-like phenotypes. The exceptions to this expectation were the effects of the T alleles for *el1.1*, *cupr1.1* and *cupr1.2*, all traits that measure aspects of ear length (Table 2). A comparison of heritabilities to the  $R^2$  values for the QTL shows that the QTL usually explain most but not all of the genetic variation (Tables 1 and 2). For example, *dtp1.1* and *dtp1.2* explain 46% of the variation for flowering time and this trait has a heritability of 52%. However, *cupr1.1* and *cupr1.2* explain only 33% of the variation for cupules per rank and this trait has a heritability of 66%.

For the eight domestication and improvement traits, we detected 12 QTL. Since the LOD threshold ( $p = 0.05$ ) was between 1.66 and 1.78 for all traits and all of the LOD scores except one are greater than 4.0, these QTL have strong statistical support (Table 2). Nevertheless, many of these QTL have relatively small effects in terms of the difference between the maize and teosinte homozygous classes, i.e. twice the additive effect. For example, *till1.1* adds only 0.3 tillers, *ed1.1* decreases ear diameter by only 2 mm, and *el1.1* increases ear length by only 1 cm. A few QTL have relatively large effects. For example, *kw1.1* produces a difference of 26 mg in kernels that weight about 180 mg or a 14% change. *cupr1.1* adds about 2.3 cupules per rank. Since there are two kernels per cupule and 7 ranks of

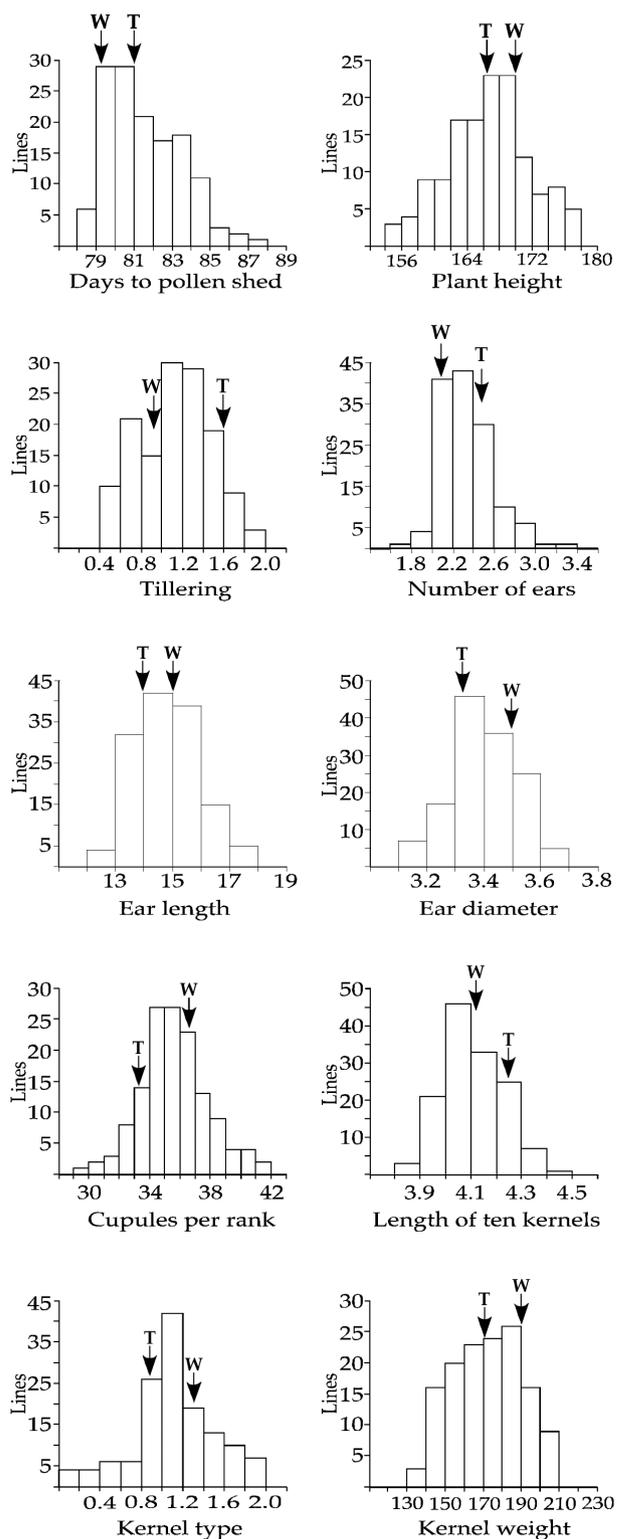


FIGURE 1 - Frequency distributions of the nearly isogenic recombinant inbred lines (NIRILs) least-square means of the 10 traits reported in this study. The arrows indicate the positions of inbred W22 (W) and the W22-T1S introgression line (T).

TABLE 2 - *Quantitative trait loci detected.*

Trait	QTL	Position (cM)	LOD <sup>a</sup>	Marker <sup>a</sup>	C.I. (cM) <sup>a</sup>	Additive effect <sup>b</sup>	R <sup>2c</sup>
Cupules per rank	<i>cupr1.1</i>	40.5	9.71	bnlg1083	32.6 - 45.6	1.14 cupules	27.9
Cupules per rank	<i>cupr1.2</i>	68.4	6.03	umc1598	51.5 - end	0.93 cupules	18.4
Cupules per rank Model	–	–	12.19			2.07	33.6
Days to pollen	<i>dtp1.1</i>	10.9	13.65	umc2225	6.8 - 20.7	1.21 days	36.8
Days to pollen	<i>dtp1.2</i>	35.4	11.51	umc2204	32.0 - 41.0	1.11 days	32.1
Days to pollen Model	–	–	18.50			2.32	46.3
Ear diameter	<i>ed1.1</i>	45.6	16.50	bnlg1803	42.8 - 49.1	-0.08 cm	42.8
Ear length	<i>el1.1</i>	42.8	7.81	bnlg1803	33.1 - 45.6	0.54 cm	23.1
Ear number	<i>en1.1</i>	18.1	4.83	umc1166	7.8 - 24.5	0.10 ears	15.0
Kernel type	<i>kt1.1</i>	22.8	17.11	umc1070	17.6 - 28.4	-0.27	43.7
Kernel type	<i>kt1.2</i>	36.4	17.22	umc2204	32.0 - 47.3	-0.26	43.9
Kernel type Model	–	–	21.67			-0.53	51.7
Kernel weight	<i>kw1.1</i>	36.9	22.03	umc2204	32.6 - 42.2	-13.04 mg	52.3
Kernel weight	<i>kw1.2</i>	59.9	17.03	AY106592	57.8 - 66.1	-12.16 mg	43.6
Kernel weight Model	–	–	30.84			-25.20	64.5
Plant height	<i>plbt1.1</i>	5.8	6.14	umc2224	0.6 - 10.3	-2.28 cm	18.7
Tillering	<i>till1.1</i>	42.2	5.11	bnlg1803	33.7 - 59.3	0.15 tillers	15.9
10-kernel length	<i>10kl1.1</i>	8.8	2.00	umc2224	0.6 - 19.1	0.03 cm	6.6
10-kernel length	<i>10kl1.2</i>	68.4	5.05	umc1598	65.0 - end	-0.05 cm	15.7
10-kernel length Model	–	–	9.68			-0.02	27.9

<sup>a</sup> Peak LOD, closest marker to the peak LOD, and the 1.5-LOD confidence interval (C.I.) for significant QTL (significance based on 10,000-permutations).

<sup>b</sup> Additive effect was estimated as  $(Q^TQ^T - Q^MQ^M)/2$ , where  $Q^TQ^T$  and  $Q^MQ^M$  represent the mean phenotypes of NIRILs for teosinte and maize genotypes at a QTL position. The sign of the additive effects corresponds to the direction of effect of the teosinte allele on the phenotype.

<sup>c</sup>  $R^2$  is the percentage of phenotypic variation explained by the QTL.

cupules on the ear, this translates to an increase of 32 kernels per ear.

Inspection of Fig. 2 suggests that the domestication-improvement QTL are somewhat clustered near the center of the introgressed teosinte segment. First, for three of the domestication and improvement traits (EL, ED and TILL), only a single QTL was detected and these three QTL all map to the same region near the center of the T1S introgression between positions 35 - 50 cM (Fig. 2; Table 2). Second, for three other domestication-improvement traits (CUPR, KT, and KW), two QTL each were detected and the QTL of largest effect of each pair was located in the same chromosomal segment between positions 35 and 50 cM (*cupr1.1*, *kw1.1* and *kt1.2*; Fig 2, Table 2). Thus, this 15 cM segment harbors six domestication related QTL including the largest effect QTL. At the top of Fig. 2, the results of a sliding window analysis of the concentration of domestication-improvement QTL is shown and it indicates that there is a high density of such QTL be-

tween positions 35 - 50 cM as indicated by the intensity of the gray-scale bar.

Outside of the region between positions 35 - 50 cM, there are six additional domestication-improvement QTL that are scattered across the introgressed teosinte chromosome segment. These QTL tend to have more modest effects with smaller LOD scores than the domestication-improvement QTL within the 15 cM region. *en1.1* is the only QTL for ear number, accounts for only 15% of the variation, and the teosinte allele increases ear number as expected. *kt1.1* accounts for 44% of the variation in kernel denting with the teosinte allele reducing the degree of denting as expected. *cupr1.2* accounts of 18% of the variation with the teosinte allele adding extra cupules to the ear which is counter to the expected direction of the effect. *kw1.2* accounts for 42% of the variation in kernel weight with the teosinte allele reducing kernel weight. *10kl1.1* and *10kl1.2* account for 7 and 16% of the variation respectively and they have effects in opposite directions.

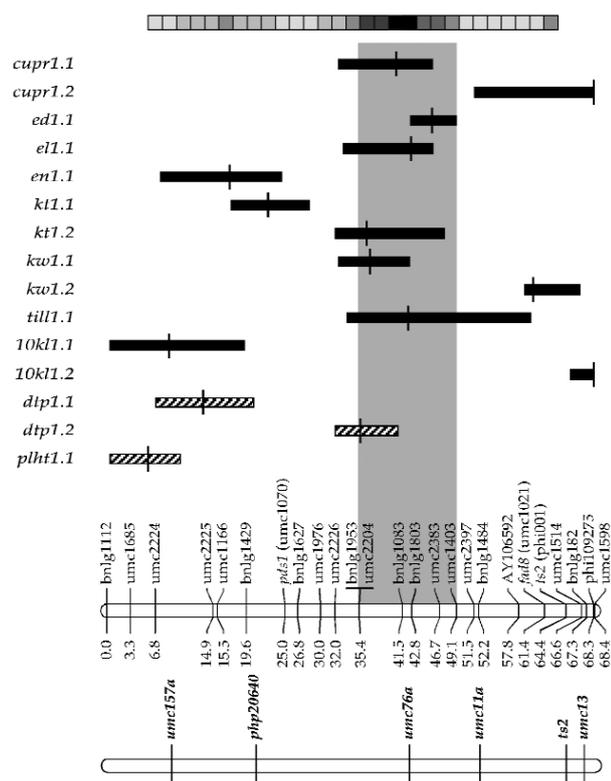


FIGURE 2 - Map of the 15 QTL detected in this study on chromosome arm 1S. Horizontal bars for each QTL represent the 1.5 LOD support interval and the narrow vertical line the position of the peak LOD score: black-bars for domestication-improvement QTL and hatched-bars for diversification QTL. The gray rectangle outlines a concentration of domestication-improvement QTL between position 35 and 50 cM. At the top of the figure, the shaded squares depict the relative-concentration of domestication-improvement QTL based on a 14 cM window sliding analysis with a 2 cM step-size. QTL names are based upon the trait name abbreviations (see Materials and Methods) followed by the chromosome number; the numbers after the period enumerate the QTLs detected for each trait. Two genetic maps are provided below the QTL plot: the SSR map built with the NIRIL genotypes (top) and the RFLP map used during the development of the NIRILs (bottom). Position of each marker locus in cM is indicated.

We also mapped QTL for plant height and days to pollen shed which are not domestication traits but are expected to differ between W22 which is adapted to a more northern (long-day) environment than teosinte. The single plant height QTL, *plht1.1*, explained 19% of the variation with plants homozygous for the teosinte allele being 4.5 cm shorter than maize homozygotes. Two QTL for days to pollen shed were detected. *dtp1.1* explains 37% of the variance with plants homozygous for the

teosinte allele being 2.4 days later than maize homozygotes. *dtp1.2* explains 32% of the variance with plants homozygous for the teosinte allele being 2.2 days later than maize homozygotes. The combined effect of both QTL gives a difference of 4.6 days in flowering time between the maize and teosinte homozygous classes.

## DISCUSSION

Maize and its progenitor teosinte, provide a good system for the study of domestication QTL since there are dramatic morphological differences between them. Herein we focused on the role of the short arm of chromosome 1 by constructing a NIL containing a teosinte chromosomal segment on this chromosome arm (T1S) introgressed into modern maize line W22 and then performing QTL mapping in a NIRIL population derived from a cross between this NIL (W22-T1S) and W22. This chromosome segment was previously identified as containing one or more QTL affecting the differences in plant and ear morphology between maize and teosinte (DOEBLEY *et al.*, 1990; DOEBLEY and STEC, 1991, 1993; BRIGGS *et al.*, 2007).

We estimated both the heritabilities ( $H^2$ ) of the traits and the amount of variance explained by the QTL ( $R^2$ ; Tables 1 and 2). If the QTL explained all the genetic variance, then the  $R^2$  and  $H^2$  values would be equal for each trait. For a few traits, these values are close. For example, for kernel type  $R^2$  is 52% and  $H^2$  is 55%. Similarly, for days to pollen shed,  $R^2$  is 46% and  $H^2$  is 52%. However, for other traits, the values are quite different. For example, for the number of cupules per rank,  $R^2$  is 34% and  $H^2$  is 66%. Multiple factors could contribute to this type of difference. First, QTL elsewhere in the genome would increase  $H^2$  but not  $R^2$ . These could be either QTL inherited from the teosinte parent or new mutations affecting the traits accumulated among the NIRILs during their creation. Second, the seed of each line came from a single ear. If the ears for some lines had poor seed quality, then  $H^2$  would be inflated relative to  $R^2$ . Third, epigenetic differences among the NIRILs could be a factor.

These three factors might also explain two anomalous features of the phenotypic data. First, the trait distributions are approximately normal (Fig. 1) rather than bimodal as one might expect for a set of NIRILs segregating for a single QTL affecting a trait with a high heritability. Low heritability and

variance in seed quality for the NIRILs could cause this outcome. Second, as shown in Fig. 1, the mean values for the W22 and W22-T1S parents are not strikingly different relative to the amount of variation among the NIRILs for many traits. Low heritability and variance in seed quality would tend to increase the variance relative to the difference between the parent lines.

For most traits, the directions of the QTL effects are in the expected direction with the teosinte allele giving a more teosinte-like phenotype. Thus, the teosinte alleles cause later flowering (*dtp1.1* and *dtp1.2*), a narrower ear (*ed1.1*), more ears per plant (*en1.1*), less dented kernels (*kt1.1* and *kt1.2*), smaller kernels (*kw1.1* and *kw1.2*), and more tillers (*till1.1*). For plant height, there is no expectation for the direction of the effect. For 10-kernel length, the direction is also not strictly predictable since two factors contribute in opposite directions to the difference between maize and teosinte for this trait. First, maize has a general enlargement (gigantism) of the ear as compared to teosinte, which suggests that the maize allele should be associated with larger values for 10-kernel length as seen with *10kl1.2*. However, maize also has tighter packing of the kernels due to less elongation of the internodes (cupules) to which the kernels are attached, which suggests that the maize allele should be associated with smaller values for 10-kernel length as seen for *10kl1.1*.

There are two traits for which the directions of the QTL effects are clearly in the wrong direction with the teosinte allele producing a more maize-like phenotype. First, the teosinte alleles of *cupr1.1* and *cupr1.2* both add more cupules per rank, although teosinte has fewer cupules (or kernels) along a rank in its ears than does maize. Curiously, DOEBLEY and STEC (1993) detected QTL for this trait in this same chromosomal region in two different maize-teosinte F<sub>2</sub> populations, however, their QTL acted in the expected direction. The unexpected direction of the effects for *cupr1.1* and *cupr1.2* may be the result of an epistatic interaction between the teosinte allele(s) and the maize (W22) genetic background. Second, the teosinte allele of *el1.1* contributes to longer ears, although teosinte has shorter ears than maize. Since *el1.1* and *cupr1.1* are in the same location and both act in the wrong direction, they may represent a single pleiotropic QTL with the additional cupules it produces adding to the length of the ear. Cupules per rank and ear length are highly correlated ( $r_p = 0.9$ ).

It may seem surprising that for five of the 10

traits we detected two QTL within the 68 cM introgressed segment. To assess whether this result was an artifact of the multiple-imputation method of QTL mapping, we repeated the analysis using composite interval mapping (CIM). CIM identified two QTL for four of these same five traits (data not shown). The LOD score for a second QTL at the fifth trait fell just below the threshold for significance with CIM. Thus, finding two QTL for some traits does not appear to be an artifact of the method of analysis. Indeed, *10kl1.1* and *10kl1.2* are 60 cM apart and act in opposite directions and thus they should be easy to detect. *dtp1.1* and *dtp1.2* are only 25 cM apart, however there are 25 NIRILs with cross-overs between these two QTL, which may explain why these two QTL could be separated. *kt1.1* and *kt1.2* are the closest pair, being only 14 cM apart with at least 13 cross-overs between them. We noticed that *kt1.2* is located very close to *kw1.1*. Perhaps, *kt1.1* is the kernel type (or denting) QTL and *kt1.2* corresponds to *kw1.1* and influences the degree of denting through an effect on kernel weight.

Our principal goals in this research were to confirm the QTL on chromosome arm 1S that were identified by DOEBLEY and STEC (1993) and to map these QTL to a narrow chromosomal segment. We have achieved the first goal, but we have made only modest progress on the second goal. At the start of this experiment, we had envisioned that the data would allow us to map the QTL to an interval of 5 cM or less flanked by two adjacent markers. This would have been possible if the QTL had Mendelized such that each of the 135 NIRILs could be classified as "+" or "-" for the QTL, i.e. if the distribution of phenotypes were bimodal. Contrary to this expectation, the trait distributions were all approximately normal and the QTL could not be Mendelized (Fig. 1).

Some of the possible reasons for the failure of the QTL to Mendelize and how these factors could be overcome include the following. (1) For five or the 10 traits, we detected two QTL, which indicates complex and not Mendelian inheritance. This problem could be mitigated by creating lines that segregate for only one of the two QTLs. (2) The traits have low heritabilities as reported in Table 1. This problem could be overcome by growing the lines at multiple locations over multiple years and measuring a larger number of plants per plot. (3) The 135 NIRILs may differ for QTL affecting the traits in regions of the genome other than chromosome arm 1S. This problem could be overcome by using

genome-wide markers and incorporating any other QTL into the statistical model. (4) The seed for each line used in this experiment all came from a single ear and thus a single mother plant (ear parent). If the ears or ear parents for different lines differed for random factors such as low-level fungal infections, degree of seed maturation, or residual seed moisture content after drying, then the trait mean of one line that is "+" for the QTL could differ considerably from another line that is "+" for the same QTL. We tried to control for this source of variation by using seed that was produced in the same year and location; however, one might additionally bulk seed of several independent ears for each line. In summary, factors such as these likely explain why the trait distributions were normal rather than bimodal and why the QTL did not Mendelize.

The majority of the domestication-related QTL mapped within a common chromosomal segment of approximately 15 cM in length (from *umc2204* to *umc1403*). The corresponding traits for most these QTL (CUPR, EL, ED, and KW) were highly correlated. One explanation for this finding is that the different traits are controlled by the same gene within this segment, which has pleiotropic effects; the alternative explanation is that these domestication-related traits are each controlled by distinct, but tightly linked genes upon which selection, during domestication, acted in concert. To further dissect this 15 cM genomic segment, additional lines with recombination events within the region will be phenotyped in replicated field experiments and screened with more PCR-based markers. By saturating this region with recombination events and genetic markers, we hope to pin down and eventually clone the gene(s) from this chromosomal segment that substantially contribute to the differences between maize and teosinte.

ACKNOWLEDGEMENTS - We thank David Wills for comments on an earlier version of this manuscript, and Karl Broman and Natalia de Leon for advice on the statistical analyses. This work was supported by the USDA Hatch grant MSN101593 and the National Science Foundation grants DBI0321467 and DBI0820619.

## REFERENCES

- BEADLE G.W., 1972 The mystery of maize. *Field Mus. Nat. Hist. Bull.* **43**: 2-11.
- BEADLE G.W., 1980 The ancestry of corn. *Scientific American* **242**: 112-119.
- BRIGGS W.H., M.D. McMULLEN, B.S. GAUT, J. DOEBLEY, 2007 Linkage mapping of domestication loci in a large maize-teosinte backcross resource. *Genetics* **177**: 1915-1928.
- BROMAN K.W., H. WU, S. SEN, G.A. CHURCHILL, 2003 R/ql: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889-890.
- DOEBLEY J.F., 2004 The genetics of maize evolution. *Ann. Rev. Genet.* **38**: 37-59.
- DOEBLEY J.F., A. STEC, 1991 Genetic analysis of the morphological differences between maize and teosinte: comparison of results of two F<sub>2</sub> populations. *Genetics* **134**: 559-570.
- DOEBLEY J.F., A. STEC, 1993 Inheritance of the morphological differences between maize and teosinte. *Genetics* **129**: 285-295.
- DOEBLEY J.F., A. BACIGALUPO, A. STEC, 1994 Inheritance of kernel weight in two maize-teosinte hybrid populations - Implications for crop evolution. *J. Hered.* **85**: 191-195.
- DOEBLEY J.F., A. STEC, C. GUSTUS, 1995 *Teosinte branched1* and the origin of maize: evidence of epistasis and the evolution of dominance. *Genetics* **141**: 333-346.
- DOEBLEY J.F., A. STEC, L. HUBBARD, 1997 The evolution of apical dominance in maize. *Nature* **386**: 485-488.
- DOEBLEY J.F., A. STEC, J. WENDEL, M. EDWARDS, 1990 Genetic and morphological analysis of a maize teosinte F<sub>2</sub> population - implications for the origin of maize. *Proc. Natl. Acad. Sci. USA* **87**: 9888-9892.
- DOERGE R.W., G.A. CHURCHILL, 1996 Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**: 285-294.
- EDWARDS A.L., 1976 An introduction to linear regression and correlation. WH Freeman and Company, San Francisco, CA.
- LANDER E.S., D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.
- LAWRENCE C.J., M.L. SCHAEFFER, T.E. SEIGFRIED, D.A. CAMPBELL, L.C. HARPER, 2007 MaizeGDB's new data types, resources and activities. *Nucl. Acids Res.* **35**: D895-D900.
- LITTELL R.C., G.A. MILLIKEN, W.W. STROUP, R.D. WOLFINGER, 1996 SAS system for mixed models. SAS Institute, Cary, NC.
- RICE W., 1989 Analyzing tables of statistical tests. *Evolution* **43**: 223-225.
- SAS INSTITUTE INC., 2000 SAS/STAT® User's guide, version 8. SAS Inst, Cary, NC.
- SEN S., G.A. CHURCHILL, 2001 A statistical framework for quantitative trait mapping. *Genetics* **159**: 371-387.
- VAN OOIEN J.W., R.E. VOORRIJS, 2001 JoinMap Version 3.0: Software for the calculation of Genetic Linkage Maps. *Plant Research Intl., Wageningen, The Netherlands.*
- WANG H., T. NUSSBAUM-WAGLER, B. LI, Q. ZHAO, Y. VIGOUROUX, M. FALLER, K. BOMBLIES, L. LUKENS, J.F. DOEBLEY, 2005 The origin of the naked grains of maize. *Nature* **436**: 714-719.
- ZHAO W., P. CANARAN, R. JURKUTA, T. FULTON, J. GLAUBITZ, E. BUCKLER, J. DOEBLEY, B. GAUT, M. GOODMAN, J. HOLLAND, S. KRESOVICH, M. McMULLEN, L. STEIN, D. WARE, 2006 Panzea: a database and resource for molecular and functional diversity in the maize genome. *Nucl. Acids Res.* **34**: D752-D757.