Evidence for a Natural Allelic Series at the Maize Domestication Locus *teosinte branched1*

Anthony J. Studer[1] and John F. Doebley
Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706

ABSTRACT Despite numerous quantitative trait loci and association mapping studies, our understanding of the extent to which natural allelic series contribute to the variation for complex traits is limited. In this study, we investigate the occurrence of a natural allelic series for complex traits at the *teosinte branched1* (*tb1*) gene in natural populations of teosinte (*Zea mays* ssp. *parviglumis*, *Z. mays* ssp. *mexicana*, and *Z. diploperennis*). Previously, *tb1* was shown to confer large effects on both plant architecture and ear morphology between domesticated maize and teosinte; however, the effect of *tb1* on trait variation in natural populations of teosinte has not been investigated. We compare the effects of nine teosinte alleles of *tb1* that were introgressed into an isogenic maize inbred background. Our results provide evidence for a natural allelic series at *tb1* for several complex morphological traits. The teosinte introgressions separate into three distinct phenotypic classes, which correspond to the taxonomic origin of the alleles. The effects of the three allelic classes also correspond to known morphological differences between the teosinte taxa. Our results suggest that *tb1* contributed to the morphological diversification of teosinte taxa as well as to the domestication of maize.

OVER the past several decades, there has been considerable interest in the genetic architecture of trait variation in natural populations as defined by number of genes involved and the effect sizes of these genes (Tanksley 1993; Mackay 2001). A key component of this issue is how variation is structured at individual genes. Are genes typically biallelic, like Mendel's classic loci, or do genes often harbor allelic series, *i.e.*, multiple alleles with measurably different effects on traits? While allelic series are known for pigmentation and other simple phenotypic traits, such as the *extension* allelic series controlling coat color in rabbits (Fontanesi *et al.* 2006), allelic series for complex morphological traits are not well documented. The unambiguous documentation of a natural allelic series for complex traits would further the understanding of the genetic architecture of variation in natural populations.

Maize and its wild relatives, the teosintes, are an attractive system for the study of natural variation and complex traits. Maize and the teosintes belong to the genus *Zea*, which has four species that are native to Mexico and Central America: *Z. perennis*, *Z. luxurians*, *Z. diploperennis*, and *Z. mays* (Doebley and Iltis 1980). The latter species includes four subspecies: one for domesticated maize (ssp. *mays*) plus three subspecies for teosinte (sspp. *parviglumis*, *mexicana*, and *huehuetenangensis*), each with a distinct ecogeographic distribution. Of these three wild subspecies, ssp. *parviglumis* has been identified as the wild progenitor of maize (Doebley 2004). Since these teosinte taxa are interfertile with maize, one can leverage the genetic tools of maize to study variation in teosinte. Some of these teosinte taxa are widespread and contain abundant natural genetic variation (Fukunaga *et al.* 2005). The teosintes are an appealing gene pool in which one could search for natural allelic series for complex traits.

Among the ~35,000 maize genes, an attractive candidate for the study of natural allelic series is *teosinte branched1* (*tb1*). This gene controls plant architecture (apical dominance) and ear morphology (Doebley *et al.* 1997). *tb1* is a member of the TCP family of transcription factors (Cubas *et al.* 1999), and it is one of the key genes involved in the domestication of maize (Doebley 2004). During maize domestication, ancient farmers selected an allele of *tb1*, which is expressed about twice as strongly as most teosinte alleles. The factor controlling this difference in gene expression has

doi: 10.1534/genetics.112.138479
Manuscript received January 7, 2012; accepted for publication April 3, 2012
Supporting information is available online at http://www.genetics.org/content/suppl/2012/04/13/genetics.112.138479.DC1.
Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. JQ900488–JQ900509.
[1]Corresponding author: Department of Genetics, 425 Henry Mall Genetics/Biotech Bldg. 5210, Madison, WI 53706. E-mail: tony.studer@gmail.com

been mapped to a regulatory region 58–69 kb upstream of the *tb1* ORF (Clark *et al.* 2006; Studer *et al.* 2011). Since teosinte possessed natural allelic variation at *tb1* upon which ancient farmers could apply selection, it seems plausible that teosinte might contain a natural allelic series at this gene for traits related to plant architecture and ear morphology.

In this article, we present evidence for a natural allelic series at *tb1*. We introgressed nine teosinte chromosomal segments encompassing *tb1* into the isogenic background of a maize inbred line. These *tb1* containing segments included four from *Z. mays* ssp. *mexicana*, four *Z. mays* ssp. *parviglumis*, and one *Z. diploperennis*. We compare the effects of these *tb1* introgressions to one another and to a maize reference allele for four morphological traits that are known to be controlled by this gene (Clark *et al.* 2006; Studer *et al.* 2011). We show that the introgressed teosinte *tb1* chromosomal segments separate into three distinct phenotypic classes and that these classes correspond to the taxonomic origin of the segments. Moreover, the effects of the *tb1* segments match the known morphological differences between these taxa. Our results suggest that *tb1*, which contributed to maize domestication, also played a role in the morphological divergence of teosinte taxa.

Materials and Methods

Plant materials

Segments of the long arm of chromosome 1 from nine different teosintes (IS1–IS9) (Supporting Information, Table S1) were introgressed into a maize inbred W22 background via six generations of backcrossing. The nine teosintes were chosen to represent multiple taxa over a wide geographic range without knowledge of the sequence makeup at *tb1*. During the backcrossing process, RFLP markers (NPI615, umc140, *tb1*, umc107, bnl15.18, *kn1*) flanking *tb1* were used to follow the target segment. After the sixth generation of backcrossing, the BC_6 plants were selfed and PCR-based markers were used to map each of the teosinte introgressed chromosomal segments (Figure 1, Table S2).

Phenotypic data collection and analysis

Plants were grown at the University of Wisconsin West Madison Agricultural Research Station, Madison, WI, during the summer of 2009. For each of the nine introgressed teosinte *tb1* chromosomal segment, a population of 140 BC_6S_1 plants was grown. All 1260 plants were grown together in a single fully randomized plot with 0.9-m spacing between plants in both dimensions. This spacing minimized the degree to which plants shaded their neighbors. Seedlings were genotyped using two PCR-based indel markers in *tb1* (Table S2). Phenotypic analysis was performed on all plants that were homozygous for either the maize or teosinte marker genotype. Using BC_6S_1 plants allowed us to compare individuals containing the introgressed teosinte chromosomal segments with individuals homozygous for the W22 segment. Seed for each of the nine populations was obtained from a single ear, thus eliminating any concern that differences among genotypic classes within a population are due to ear parent effects.

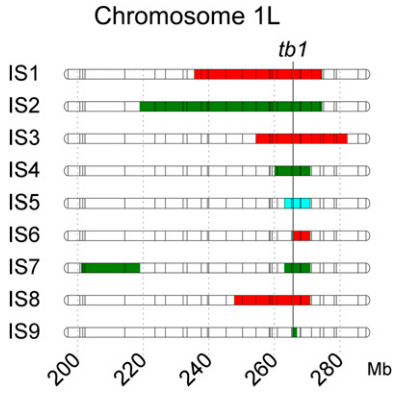

Figure 1 Physical map of the introgression lines. All introgressed segments are drawn to scale, and vertical dotted lines show AGPv2 reference position (Mb). Shaded areas indicate teosinte chromosome segments on the basis of taxonomic origin: (blue) *Zea diploperennis*, (red) *Z. mays* ssp. *parviglumis*, and (green) *Z. mays* ssp. *mexicana*; unshaded areas represent maize chromosome segments. Markers used for genotyping are shown along the chromosomes as solid black lines and listed in Table S2. The position of *tb1* is shown for reference.

The following four traits were phenotyped: cupules per rank (CUPR; number of cupules in a single rank from base to the tip of the ear), lateral branch internode length (LBIL; mean internode length, in centimeters, for the uppermost lateral branch), tillering (TILL; sum of tiller heights/plant height), and percentage of staminate spikelets (STAM; percentage of male spikelets in the inflorescence). CUPR and STAM were both measured on the uppermost lateral inflorescence (ear) of each plant.

The software package JMP IN 4.0.4 (SAS Institute, Cary, NC) was used for calculating means, standard errors, Levene's tests, and principal component analysis (PCA). The MIXED procedure of SAS (SAS Institute) was used to implement a mixed linear statistical model to test for genotypic, family, and allelic effects. Genotype (W22 control segment *vs.* teosinte introgression at *tb1*) was considered a fixed effect, while family (IS1–IS9) and a genotype by family interaction term were treated as random effects. The full linear model used was

$$Y_{cdf} = \mu + a_c + b_d + a_c \times b_d + e_{cdf},$$

where Y_{cdf} is the trait value for the *fth* plant with *cth* genotype from the *dth* family, μ is the overall mean of the experiment, a_c is the genotypic effect, b_d is the family effect, $a_c \times b_d$ is the genotype by family interaction, and e_{cdf} is the error term. Terms were added individually to the model and tested for significance using the likelihood ratio test which has a χ^2 distribution with one degree of freedom. Levene's test was used to assess the equality of the variance for plants containing the W22 control segments *vs.* plants containing teosinte

Table 1 *N* and additive effects for homozygous plants

Introgression	N_M/N_T	Additive effects TILL	Additive effects LBIL	Additive effects CUPR	Additive effects STAM
IS1	21/32	0.6993	0.8075	−3.2368	2.9186
IS2	34/31	0.5684	0.7766	−0.4941	0.3971
IS3	23/28	0.6558	0.5455	−3.3727	2.1882
IS4	33/41	0.5619	0.8682	−0.2175	−0.0095
IS5	33/23	0.4446	−0.0592	−0.6667	0.0000
IS6	31/34	0.3539	0.7284	N/A	N/A
IS7	32/29	0.6757	1.4879	−1.9267	0.0000
IS8	37/30	0.5191	0.9026	−3.0700	2.2961
IS9	24/32	0.6370	0.5928	−0.8644	0.0273

Subscript M, homozygous maize control plants; subscript T, homozygous teosinte introgression plants.

introgression segments. PCA was performed using the trait correlation matrix for the additive effects. Additive effects were calculated as half the difference between the mean of the homozygous teosinte introgression class and the mean of the corresponding maize class. The sample size for each introgression family and the additive effect estimates are listed in Table 1. Given that the difference between W22 and itself equals zero, zero was used for all W22 trait values in the PCA. All genotype and phenotype data are available at www.panzea.org.

Nucleotide sequence analysis

Polymerase chain reaction (PCR) was done using Qiagen Taq DNA Polymerase following the manufacture's instructions and standard methods. One primer set was used to amplify the *tb1* coding region (GGACATATGAGTAGGCCA CACTCCTCC, GATTTGCAGCTCATCAAGAAA) and two additional internal primers were used to sequence the PCR product (TCATGGACAACGATGAGTGG, CCAAGAAAATCGGC CAATAA). Two primer sets were used to amplify the *tb1* control region (CGGTCAAAGAGTAGGGCAAG, GCGTCTGTTCCG CATTCA and ACTCAACGGCAGCAGCTACCTA, CGTGTGTGTG ATCGAATGGT). Sequencing of PCR fragments was done using Applied Biosystems (ABI) BigDye and an ABI 3730xl DNA Analyzer at the University of Wisconsin Biotechnology Center DNA Sequencing Facility. Initial alignment of nucleotide sequences was performed using ClustalW (Thompson *et al.* 1994) and then finished by hand using MEGA version 5.03 (Tamura *et al.* 2011). Neighbor joining trees were constructed in PAUP 4.0b10 (Swofford 2003) using the absolute number of differences after gaps; missing and ambiguous bases were removed from the alignment.

Results

Genotypic, family, and allelic effects

To observe each of the *tb1* teosinte chromosomal segments' effects on phenotype relative to the control W22 chromosome segment, the mean for plants homozygous for each of the nine introgression segments and the mean for each of the control populations (plants homozygous for the maize segment) were plotted for the four phenotypes (Figure 2): CUPR, LBIL, TILL, and percentage of STAM. These traits represent ear morphology (CUPR and STAM) and plant architecture (LBIL and TILL), which are some of the major morphological differences between maize and teosinte. One of the teosinte introgressed segments (IS6) does not have data for ear morphology traits because all of the ears (from introgression and control plants) were sterile (without kernels).

To test for genotypic (W22 control segment *vs.* teosinte introgression), family (IS1–IS9), and allelic affects (genotype × family) a mixed linear statistical model was used. For all traits, the model indicates that plants containing a teosinte introgression segment are significantly different from plants with a W22 control segment (Table 2A, Figure 2). A significant family effect is also supported by the model for each trait, indicting a difference between family means (Table 2A). To explore further the differences between families, an interaction term was added to the model to ask whether the introgression family affects both genotypes equally. If significant, this interaction would suggest that not all of the teosinte introgression segments are equivalent, assuming that all of the W22 controls engender equivalent phenotypes given they carry the same allele at *tb1*. The interaction term is significant for LBIL, CUPR, and STAM but not for TILL (Table 2A). The insignificant interaction term for TILL indicates that all of the teosinte introgressed segments have equivalent effects and that the variance observed among families is due to factors other than an allelic series at *tb1*.

We observed a significant family effect for TILL but not a significant family by genotype interaction term. This result suggests that there are differences among the introgression families due to a factor other than the *tb1* introgression segment. Two possible explanations were considered for this result. First, significant phenotypic differences between families could be observed if additional genetic factors segregated between backcrossing populations at loci in the genome unlinked to the target segment encompassing *tb1*. Such factors would increase (or decrease) the trait mean for both plants with the introgressed teosinte chromosomal segment and the corresponding control plants, which would contribute to a significant family term in the model but not a significant interaction between genotype and family. Second, environmentally determined seed quality differences among the ear parents for different introgression families could be responsible. This is particularly possible since only a single ear parent was used for each introgression family. Ear parent effects such as seed weight, seed maturity, and speed of germination can influence adult phenotype. Thus, environmentally induced ear parent effects could account for the differences seen among the introgression families, which were derived from different ears.

For LBIL, the interaction term between genotype by family was found to be significant (Table 2A), suggesting that not all of the teosinte introgression segments are equivalent when

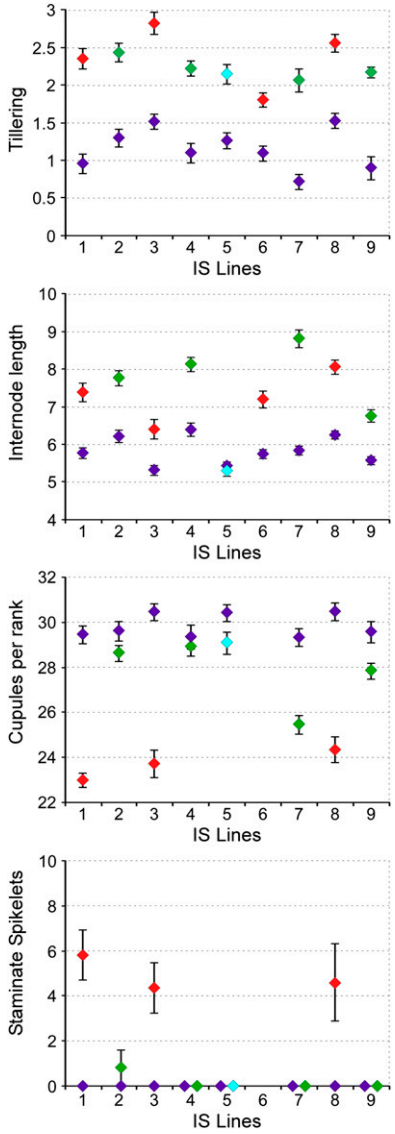

Figure 2 Phenotypic means. Points are shaded on the basis of taxonomic origin of the *tb1* introgressed segment: (purple) *Zea mays* ssp. *mays* control populations, (blue) *Z. diploperennis*, (red) *Z. mays* ssp. *parviglumis*, and (green) *Z. mays* ssp. *mexicana*. Error bars represent the standard error for each genotypic class. The *x*-axis shows the introgression segments; the *y*-axis shows trait means.

assuming all of the W22 control populations are equivalent. To investigate this result further, a model with just a family term was tested against a null model. This was performed separately for the W22 control and the teosinte introgression subsets of the data. The family term was significant for both subsets (Table 2B), indicating that while there are significant differences among the teosinte introgressions this could be the result of factors other than an allelic series at *tb1*, since there are also significant differences among the W22 controls.

Since there were significant differences among W22 controls for LBIL, we used Levene's test to ask whether the variance among teosinte *tb1* introgressions for LBIL is equivalent to the variance among the control populations as expected if there is not an allelic series at *tb1*. The results of this test indicate that there is greater variance among teosinte introgressions as compared to the control populations (Table 3), suggesting that the teosinte introgressions possess different allelic effects for LBIL. A graph of the additive effects for LBIL highlights the small effect of IS5 and the large effect IS7 has on LBIL, compared to the rest of the teosinte introgressions (Figure 3).

Our analyses using the mixed linear model for the ear morphology traits (CUPR and STAM) produced a different result than seen for the plant architecture traits (TILL and LBIL). For both CUPR and STAM, the interaction term for genotype by family was found to be significant (Table 2A), suggesting that not all of the teosinte introgression segments are equivalent. Furthermore, when the interaction was included in the model, the family term dropped out, indicating that all of the variance observed among families is due to differences between teosinte introgression segments rather

Table 2 Likelihood ratio test results comparing statistical models with genotype, family, and allelic effects (d.f. = 1)

Test	Genotype[a] *M vs. T*	Family[b] IS1–IS9	Genotype × Family[c] Allelic effects
A.			
TILL	<0.0001	<0.0001	0.1573
LBIL	<0.0001	<0.0001	<0.0001
CUPR	<0.0001	<0.0001	<0.0001
STAM	<0.0001	<0.0001	<0.0001
Test	**Family (*M*)[d]**	**Family (*T*)[d]**	
B.			
TILL	<0.0001	<0.0001	
LBIL	<0.0001	<0.0001	
CUPR	0.3711	<0.0001	
STAM	1	<0.0001	

[a] Tests whether the means of both genotypes are equal or whether the two genotypic means are different: $H_0 : Y = \mu + e$, $H_a : Y = \mu +$ genotype $+ e$.
[b] Tests whether the nine family means are equal or whether two or more family means are not equal: $H_0 : Y = \mu +$ genotype $+ e$, $H_a : Y = \mu +$ genotype $+$ family $+ e$.
[c] Tests whether the allelic effects between families are equal or whether two or more families have allelic effects that are not equal: $H_0 : Y = \mu +$ genotype $+$ family $+ e$, $H_a : Y = \mu +$ genotype $+$ family $+$ genotype $\times$ family $+ e$.
[d] Tests whether the nine family means are equal or whether two or more family means are not equal using maize (*M*) and teosinte introgression (*T*) subsets of the data: $H_0 : Y = \mu + e$, $H_a : Y = \mu +$ family $+ e$.

Table 3 Levene's test results for equal variance

Test M vs. T	Levene's DFNum[a]	Levene's DFDen[b]	Levene's *F*-ratio	Levene's *P*-value
TILL	1	16	0.0089	0.9262
IBIL	1	16	5.0864	0.0385
CUPR	1	14	40.6558	<0.0001
STAM	1	14	62.6364	<0.0001

[a] Numerator degrees of freedom.
[b] Denominator degrees of freedom.

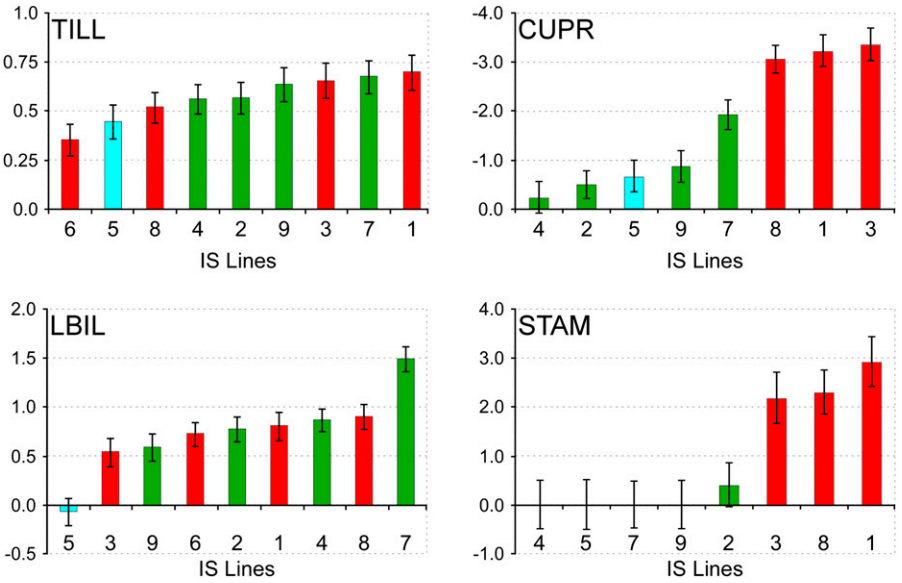

Figure 3 Additive effects. Traits are abbreviated as follows: cupules per rank (CUPR), lateral branch internode length (LBIL, in centimeters), staminate spikelets (STAM, percentage), and tillering (TILL). The *x*-axis shows the introgression segments; the *y*-axis shows additive effects. Error bars represent the standard error for each effect. Bars are shaded on the basis of taxonomic origin of the introgression segments: (blue) *Zea diploperennis*, (red) *Z. mays* ssp. *parviglumis*, and (green) *Z. mays* ssp. *mexicana*.

than unlinked genetic factors that differ among the introgression populations or ear parent effects as discussed above. A model with just a family term was used to test the assumption that the W22 control populations are equivalent. Significant family effects were found among teosinte introgression segments but not among W22 controls (Table 2B), further supporting the hypothesis that there is an allelic series for CUPR and STAM at *tb1*.

The tb1 introgressions form distinct classes

To assess whether the different teosinte *tb1* introgressions could be classified into groups on the basis of phenotype, a principal component analysis was performed using the additive effects of the four traits as input data. IS6 was not included in the analysis because it is missing data for two of the four traits. Two components were retained from the analysis, which explain 64 and 27% of the observed variance. The ear morphology traits, cupules per rank and staminate spikelets, load to component 1, which is represented by the *x*-axis in Figure 4. The plant architecture traits, tillering and lateral branch internode length, load to component 2, which is represented by the *y*-axis in Figure 4. The W22 control plots to the lower left quadrant of the graph with distance from this point corresponding to more teosinte-like phenotypes (Figure 4).

The principal component analysis suggests that there are three classes of teosinte *tb1* introgressions. The first class is composed of a single introgression (IS5), which plots away from the rest of the teosinte introgressions and is located in the quadrant containing the W22 control. This result suggests that IS5 is an allele that confers a phenotype that is only modestly different from the W22 control. This relationship can also be observed by looking at IS5 for each trait individually (Figures 2 and 3). The second class is composed of IS2, -4, -7, and -9, all of which plot to the upper left quadrant (Figure 4). This quadrant represents introgressions that produce teosinte-like plant architecture traits (long tillers and lateral branches), but maize-like ear morphology traits (more cupules per rank and few staminate spikelets). The final class is composed of IS1, -3, and -8 and occupies the right half of the graph along the *x*-axis (Figure 4). These introgressions produce both a more teosinte-like ear morphology and plant architecture. In particular, IS1, -3, and -8 have a high percentage of male spikelets in their ears (Figures 2 and 3).

Strikingly, the PCA reveals that the allelic classes correspond to the taxonomic origin of the teosinte *tb1* introgression (Figure 4). The allelic class with the most teosinte-like phenotypes corresponds to introgressions from *Z. mays* ssp. *parviglumis* (PAR). The allelic class with moderate teosinte-like phenotypes corresponds to introgressions from the *Z. mays* ssp. *mexicana* (MEX). Finally, the allelic class with the most maize-like phenotypes corresponds to the introgression from *Z. diploperennis* (DIP). Thus, the allelic series at *tb1* appears to have a taxonomic basis. Because of the isogenic nature of the introgression lines, the apparent allelic series cannot be the result of factors other than a difference at or near *tb1*.

Although the allele series shows a distinct taxonomic signature, we also asked whether the allele classes were correlated with the length of the introgressed segments (Figure 1). No obvious correlation between phenotype and introgression length is observed. For example, the largest introgression (IS2) does not have the most teosinte-like phenotypes, nor does the smallest introgression (IS9) have the most maize-like phenotypes (Figure 3). Moreover, different introgression lengths are represented in the different allelic classes defined in the PCA. This result supports the conclusion of an allelic series at *tb1*, as opposed to other linked genes in the introgressed segments causing the observed allelic differences.

To explore the possibility of a correlation between the nucleotide sequence of *tb1* and phenotype, we plotted the phenotypic classes defined by the PCA onto neighbor joining

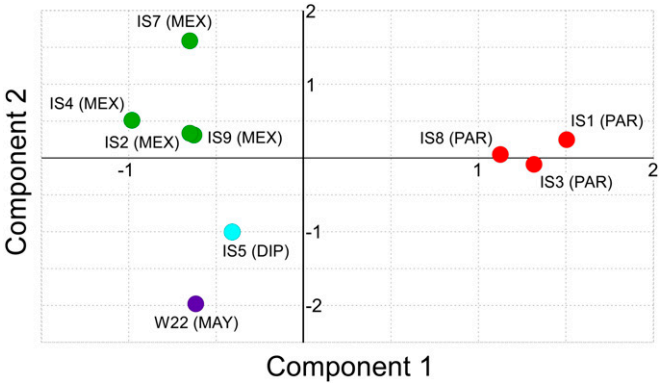

Figure 4 Principle components plot. The *x*-axis shows component 1, which represents ear morphology traits; the *y*-axis shows component 2, which represents plant morphology traits. Dots are shaded on the basis of taxonomic origin of the introgression segment: (purple) *Zea mays* ssp. *mays*, (blue) *Z. diploperennis*, (red) *Z. mays* ssp. *parviglumis*, and (green) *Z. mays* ssp. *mexicana*.

trees on the basis of two regions of the *tb1* nucleotide sequence (Figure 5). One portion is the protein coding region of the gene and 3′-UTR, and the other corresponds to a known upstream regulatory region of *tb1* (Clark *et al.* 2006). The teosinte introgressions representing any single allelic class defined by the PCA are scattered across both of the phylogenetic trees, and for the most part, no relationship between phylogeny and phenotype is apparent. For example, the class representing the most teosinte-like ear phenotypes (IS1, -3, and -8) does not cluster in either phylogeny. One striking feature of both phylogenies is that IS5, which was derived from a separate species (*Z. diploperennis*, DIP) and has unique phenotypic effects, stands apart from all other introgressions in both trees. This result suggests that the different phenotypes observed for IS5 when compared to the other introgressions could be due to sequence differences in the upstream control region and/or the coding sequence of *tb1*. Since the introgressions from neither *Z. mays* ssp. *parviglumis* (PAR) nor *Z. mays* ssp. *mexicana* (MEX) cluster on either of the phylogenetic trees, these trees do not enable the identification of sequence variants that control the putative allelic variation.

Discussion

Natural allelic series for simple phenotypic traits such as pigmentation are well documented in the literature. For example, five alleles have been described at the *R* locus in maize, which control plant and kernel pigmentation. Each of these five alleles produces a distinct phenotype on the basis of pigment quantity, spatial patterning in kernels, the timing of pigmentation onset during development, and which organs are pigmented (kernels, anthers, leaves, and/or roots) (Styles *et al.* 1973). A similar allelic series for pigmentation has been described for the *B* locus of maize (Styles *et al.* 1973; Radicella *et al.* 1992). Much like these examples from maize, an allelic series for coat color in mice has been described (Phillips 1966; Jackson 1994). Alleles of the *agouti* locus produce distinct coat colors and pattern differences due to factors in both the promoter and coding region of the gene. Allelic series have also been described for traits such as self-incompatibility in plants (Nasrallah *et al.* 1991; Takayama and Isogai 2005).

Evidence for natural allelic series for complex or morphological traits has come from association mapping and QTL studies (Purugganan and Suddith 1998; McKechnie *et al.* 2010; Todesco *et al.* 2010). For example, an allelic series for flowering time was reported among a diverse set of maize lines that display significant variation in flowering time (Buckler *et al.* 2009). In this example, statistical evidence for an allelic series is shown; however, there is no actual proof that a single locus with multiple alleles explains the observed phenotypic series, since the occurrence of several tightly linked genes each with two alleles cannot be excluded. Another concern with evidence for allelic series from QTL and association studies is that the alleles are each typically characterized in a different genetic background. Thus, it is possible that the QTL in question has only two alleles that form a number of apparent allelic classes on the basis of the background in which they were assayed. Using association mapping, Weber *et al.* (2007) assayed variation in a natural teosinte population and found multiple SNPs in and around *tb1* associated with small effects on plant and ear architecture. Their results are consistent with our data, which show relatively small amounts of natural variation within taxa. However, because Weber *et al.* (2007) only included a single taxon (*Z. mays* ssp. *parviglumis*), the among-taxa variation that we described may be distinct from the allelic effects they report.

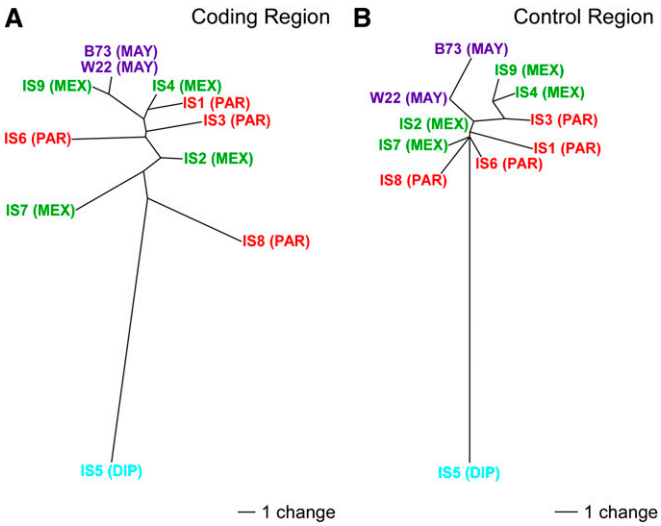

Figure 5 Phylogenetic trees. (A) Neighbor joining tree, based on sequence from the *tb1* coding region. (B) Neighbor joining tree, based on sequence from the *tb1* upstream control region. Text color is based on taxonomic origin of the introgression segment: (purple) *Zea mays* ssp. *mays*, (blue) *Z. diploperennis*, (red) *Z. mays* ssp. *parviglumis*, and (green) *Z. mays* ssp. *mexicana*.

In this article, we present evidence for a natural allelic series at *tb1* for three complex morphological traits: lateral branch internode length, the number of cupules per rank, and the number of staminate spikelets (Figures 2 and 3). Our evidence for allelic series at *tb1* largely eliminates concerns about the influence of genetic background by using isogenic lines. We also examined the role of linked genes on trait variation associated with *tb1* by considering the length of the introgressed chromosomal segment surrounding *tb1* for each of the teosinte introgressions. We saw no evidence that phenotype is correlated with the length of the introgression segment (Figures 1 and 4).

An argument could be made that *tb1* is contributing little or nothing to the observed phenotypic variation that we observed, and that the variation is caused by heterogeneity at linked genes, given that the introgressions contain between ~80 (IS9) and >800 (IS2) linked genes. However, IS9, which spans only 80 linked genes, has nearly identical phenotypic effects to IS2, which spans 800 linked genes, arguing against a role for linked genes contributing to the observed phenotypic variation. Ideally, teosinte introgression segments of a uniform length and that only contain the *tb1* gene itself would be compared. However, the creation of such lines would be a long and labor-intensive process.

A recent report (Studer and Doebley 2011) on the fractionation of the QTL effects at *tb1* sheds additional light on whether linked genes underlie the allelic effects that we observed. This report shows that the QTL at *tb1* for plant architecture traits (TILL and LBIL) does not fractionate, but rather maps narrowly to a 69-kb region upstream of the *tb1* coding sequence. This result is consistent with the hypothesis that the variation for LBIL near *tb1* is due sole to *tb1*, and by inference, that the variation observed in this study is attributable to allelic differences at *tb1* and not linked genes.

Studer and Doebley (2011) report that the QTL at *tb1* for CUPR does fractionate; however, presence/absence of the teosinte allele at this QTL does not correlate well with the phenotypic differences among the introgressions. For example, IS1–3 and -8 all carry the teosinte allele of the CUPR QTL identified upstream of *tb1*. While IS1, -3, and -8 all show large effects for CUPR, IS2 has only a small effect (Figure 3). Furthermore, *Z. mays* ssp. *mexicana* introgression IS7 has an intermediate effect on CUPR but does not have the teosinte allele for the CUPR QTL. The taxonomic origin of the introgression is a better predictor of variation for CUPR than the presence/absence of the teosinte allele for the CUPR QTL. While *Z. mays* ssp. *parviglumis* introgressions (IS1, -3, and -8) all have strong effects, most *Z. diploperennis* and *Z. mays* ssp. *mexicana* introgressions (IS2, -4, -5, and 9) have weak effects (Figure 3).

Studer and Doebley (2011) also report that the QTL at *tb1* for STAM fractionates, and in this case, there is a correlation, although imperfect, between presence/absence of the teosinte allele at this QTL and the phenotypic differences among the introgressions for STAM. Our introgressions IS1, -3, and -8 all carry the teosinte allele of this QTL for STAM and have the largest effects for STAM, while IS4, -5, -7, and -9 carry the maize allele and show no effect on STAM (Figure 3). The exception is IS2, which carries the teosinte allele at this linked STAM QTL but does not have a significant effect on STAM. Thus, taxonomy is still a better predictor of variation for STAM than the presence/absence of the teosinte allele for the STAM QTL, since *Z. mays* ssp. *parviglumis* introgressions (IS1, -3, and -8) all have strong effects and *Z. diploperennis* and *Z. mays* ssp. *mexicana* introgressions (IS2, -4, -5, -7, and -9) all have weak effects (Figure 3).

The feature that is best correlated with the phenotypic effects of the *tb1* alleles that we examined is the taxonomic origin of these alleles. In a principal components analysis on the basis of phenotype, the eight teosinte introgressions form three classes that correspond to *Z. mays* ssp. *parviglumis*, *Z. mays* ssp. *mexicana*, and *Z. diploperennis* (Figure 4). This result not only supports the existence of an allelic series at *tb1*, but it also implicates *tb1* in the morphological diversification of these taxa in addition to its role in maize domestication. There are several notable correspondences between known morphological differences between these taxa and the effects associated with the alleles of *tb1* we assayed. First, *Z. mays* ssp. *mexicana* has more fruitcases (a greater CUPR value) per ear than either *Z. mays* ssp. *parviglumis* or *Z. diploperennis* (Iltis and Doebley 1980), and our *Z. mays* ssp. *mexicana* alleles have greater CUPR values than our *Z. mays* ssp. *parviglumis* and *Z. diploperennis* alleles (Figure 3). Second, *Z. diploperennis* has shorter lateral branches that are tipped in a mixed male–female inflorescence unlike other teosintes that have longer lateral branches tipped by tassels (Iltis *et al.* 1979; Doebley and Iltis 1980). The one *Z. diploperennis* allele we assayed has the smallest value for LBIL (shorter branches) of all nine teosinte alleles assayed (Figure 3). While our observations suggest that *tb1* may partly control morphological differences among teosinte taxa, our study includes a limited sampling of each taxa and thus the data must be regarded as suggestive rather than conclusive.

Given the correlation between taxonomy and allelic effects (Figure 4), we examined phylogenetic trees on the basis of the nucleotide sequences of the control region and coding sequence of *tb1*, but we saw no relationship between phenotype and phylogeny. We also examined the sequence alignments for any fixed differences between the taxa that may not have been visible in the trees. No fixed differences were found between *Z. mays* ssp. *mexicana* and *Z. mays* ssp. *parviglumis* individuals for either sequenced region. The *Z. diploperennis* sequence is highly divergent from the other alleles with many sequence differences. With such a large number of differences and only a single *Z. diploperennis* sample, it is not possible to say which if any are potentially causative. However, there are two polymorphisms unique to the *Z. diploperennis* allele of *tb1* that cause radical amino acid changes in the helix II portion of the TCP domain, which is involved in DNA binding. An A > G substitution at AGP_v2 position 265,746,492 causes a T-to-A amino acid change, and

T > G substitution at AGP_v2 position 265,746,501 causes a S-to-A amino acid change. Both changes are from hydrophobic to hydrophilic amino acids, which could alter protein function. Further experimentation is needed to test whether these amino acid differences affect phenotype.

In summary, our experiments provide evidence for a natural allelic series at *tb1* with effects on complex morphological traits. It has been previously shown that *tb1* played a major role in the domestication of maize from its wild progenitor, teosinte (Doebley 2004). Since the allelic classes that we observed at *tb1* correspond with taxonomic origin, *tb1* may also have played a role in the morphological diversification of *Z. mays* ssp. *parviglumis*, *Z. mays* ssp. *mexicana*, and *Z. diploperennis*. To provide final proof of the allelic series at *tb1* and verify its role in the divergence of teosinte, the causal polymorphisms underlying the phenotypic differences need to be identified.

Acknowledgments

We thank Brian Yandell and Peter Bradbury for statistical advice concerning the data analyses in this manuscript. This work was supported by the Department of Agriculture Hatch grant MSN101593 and the National Science Foundation grants DBI0321467 and DBI0820619.

Literature Cited

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. Acharya, P. J. Brown *et al.*, 2009 The genetic architecture of maize flowering time. Science 325: 714–718.

Clark, R. M., T. Nussbaum Wagler, P. Quijada, and J. Doebley, 2006 A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescence architecture. Nat. Genet. 38: 594–597.

Cubas, P., N. Lauter, J. Doebley, and E. Coen, 1999 The TCP domain: a motif found in proteins regulating plant growth and development. Plant J. 18: 215–222.

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

Doebley, J., A. Stec, and L. Hubbard, 1997 The evolution of apical dominance in maize. Nature 386: 485–488.

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

Fontanesi, L., M. Tazzoli, F. Beretti, and V. Russo, 2006 Mutations in the *melanocortin 1 receptor* (*MC1R*) gene are associated with coat colours in the domestic rabbit (*Oryctolagus cuniculus*). Anim. Genet. 37: 489–493.

Fukunaga, K., J. Hill, Y. Vigouroux, Y. Matsuoka, J. Sanchez-Gonzalez *et al.*, 2005 Genetic diversity and population structure of teosinte. Genetics 169: 2241–2254.

Iltis, H. H., and J. F. Doebley, 1980 Taxonomy of *Zea* (Gramineae). II. Subspecific categories in the *Zea mays* complex with a generic synopsis. Am. J. Bot. 67: 994–1004.

Iltis, H. H., J. F. Doebley, R. M. Guzman, and B. Pazy, 1979 Zea diploperennis (Gramineae): a new teosinte from Mexico. Science 203: 186–188.

Jackson, I. J., 1994 Molecular and developmental genetics of mouse coat color. Annu. Rev. Genet. 28: 189–217.

Mackay, T. F. C., 2001 The genetic architecture of quantitative traits. Annu. Rev. Genet. 35: 303–339.

McKechnie, S. W., M. J. Blacket, S. V. Song, L. Rako, X. Carroll *et al.*, 2010 A clinally varying promoter polymorphism associated with adaptive variation in wing size in *Drosophila*. Mol. Ecol. 19: 775–784.

Nasrallah, J. B., T. Nishio, and M. E. Nasrallah, 1991 The self-incompatibility genes of *Brassica*: Expression and use in genetic ablation of floral tissues. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 393–422.

Phillips, R. J. S., 1966 A *cis-trans* position effect at the *A* locus of the house mouse. Genetics 54: 485–495.

Purugganan, M. D., and J. I. Suddith, 1998 Molecular population genetics of the *Arabidopsis CAULIFLOWER* regulatory gene: nonneutral evolution and naturally occurring variation in floral homeotic function. Proc. Natl. Acad. Sci. USA 95: 8130–8134.

Radicella, J. P., D. Brown, L. A. Tolar, and V. L. Chandler, 1992 Allelic diversity of the maize *B* regulatory gene: different leader and promoter sequences of two *B* alleles determine distinct tissue specificities of anthocyanin production. Genes Dev. 6: 2152–2164.

Studer, A., Q. Zhao, J. Ross-Ibarra, and J. Doebley, 2011 A transposon insertion was the causative mutation in the maize domestication gene *tb1*. Nat. Genet. 43: 1160–1163.

Studer, A. J., and J. F. Doebley, 2011 Do large effect QTL fractionate? A case study at the maize domestication QTL *teosinte branched1*. Genetics 188: 673–681.

Styles, E. D., O. Ceska, and K.-T. Seah, 1973 Developmental differences in action of *R* and *B* alleles in maize. Can. J. Genet. Cytol. 15: 59–72.

Swofford, D. L., 2003 *PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods)*, *Version 4.0b10*, Sinauer Associates, Sunderland, MA.

Takayama, S., and A. Isogai, 2005 Self-incompatibility in plants. Annu. Rev. Plant Biol. 56: 467–489.

Tamura, K., D. Peterson, N. Peterson, and G. Stecher, M. Nei *et al.*, 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.

Tanksley, S. D., 1993 Mapping polygenes. Annu. Rev. Genet. 27: 205–233.

Thompson, J. D., D. G. Higgins, and T. J. Gibson, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position- specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680.

Todesco, M., S. Balasubramanian, T. T. Hu, M. B. Traw, M. Horton *et al.*, 2010 Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. Nature 465: 632–636.

Weber, A., R. M. Clark, L. Vaughn, J. de Jesus Sanchez-Gonzalez, J. Yu *et al.*, 2007 Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). Genetics 177: 2349–2359.

Communicating editor: A. Charcosset

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2012/04/13/genetics.112.138479.DC1

Evidence for a Natural Allelic Series at the Maize Domestication Locus *teosinte branched1*

Anthony J. Studer and John F. Doebley

DOI: 10.1534/genetics.112.138479

Table S1 Teosinte alleles used for introgressed chromosomal segments

Line	Species	Subspecies	Country	State/ Province	Population	Collector	Collection	Lat Deg	Lat Min	Long Deg	Long Min
IS1	*mays*	*parviglumis*	Mexico	Guerrero	1 mile S of Palo Blanco	Beadle & Kato	Site 4	17	25	-99	30
IS2	*mays*	*mexicana*	Mexico	Mexico	km 43 on hwy from Chalco to Amecameca	Iltis *et al.*	28622	19	6	-98	42
IS3	*mays*	*parviglumis*	Mexico	Guerrero	30 km S of Chilpancingo	Beadle & Kato	Site 2-3	17	12	-99	30
IS4	*mays*	*mexicana*	Mexico	Jalisco	10 km S of Degollado	M. Puga	11066	20	22	-102	11
IS5	*diploperennis*		Mexico	Jalisco	Zarza Mora, 2 km E of Las Joyas	Iltis *et al.*	1250	19	35	-104	16
IS6	*mays*	*parviglumis*	Mexico	Guerrero	1 km N of Mazatlan	Beadle & Kato	Site 1	17	30	-99	30
IS7	*mays*	*mexicana*	Mexico	Chihuahua	Nobogame	Beadle	s.n.	26	6	-107	0
IS8	*mays*	*parviglumis*	Mexico	Guerrero	Sites 9-10, Teloloapan-Arcelia Hwy	Iltis & Cochrane	81	18	21	-100	12
IS9	*mays*	*mexicana*	Mexico	Mexico	km 1.8 WSW of Texcoco	H. Iltis	28620	19	30	-98	55

Table S2 Markers for genotyping

Markers	Forward Primer (5′ to 3′)	Reverse Primer (5′ to 3′)
umc2569 [a]	GTGACACCCTAGCCCTCTTAGACA	TAGCTGGAGTATGTCGTCTTGGTG
umc2237 [a]	CTCAGCTACAGGAGCGAAGAGG	GTCACTGCACGATCCATCACAT
umc1122 [a]	CACAACTCCATCAGAGGACAGAGA	CTGCTACGACATACGCAAGGC
umc2396 [a]	TGCATCTTTAGCTACGAGACAACCT	TGCATGCATTTTTAGGTTTGGAAT
bngl615 [a]	CTTCCCTCTCCCCATCTCCTTTCCAA	GCAACCTGTCCATTCTCACCAGAGGATT
bnlg1025 [a]	TGGTGAAGGGGAAGATGAAG	CCGAGACGTGACTCCTAAGC
bnlg1564 [a]	ACGGGAGAACAAAAGGAAGG	CTCTCCCTCACATCCGCC
bnlg1629 [a]	GTTGGATGGAAAATTCTAGATCG	TTGCGTCATTACAGCAGGAG
bnlg2228 [a]	GCAGCAATCGACACGAGATA	CTTGGATCGCACTCCGTC
umc2181 [a]	ATCGGGTCCGGATAGATTTTACAC	GTAGCTAGCTTAAGCAGTGCTCCG
mmc0041 [a]	AGGACTTAGAGAGGAAACGAA	TTTATCCTTACTTGCAGTTGC
umc1924 [a]	GGATGCGGTCGTACAGTACAAGTAT	CTACAACAACTGCTGCTCCCG
umc1991 [a]	GAAATTGATGCAATTCACCCTGAT	ATTGAATTGCGTGATGCAAGAGTA
umc1914 [a]	CAACATGAGCGTGCTAAATACTCG	ACAGGAACACATGAGGTCATCAAA
umc2047 [a]	GACAGACATTCCTCGCTACCTGAT	CTGCTAGCTACCAAACATTCCGAT
umc1298 [a]	AGCTGAACAAAATAAACGGAACGA	AGGACAAGAAAAAGAAGAAGCACG
PZD00117.indel1 [a]	CCCGCGGCCCGCCGTCAAGT	ATGCGCGGGCAAGCGCACCG
umc1306 [a]	CGAAACAAAACACCCAGCAGTAGT	CCAGGATGAATAAATCGTATTGCC
bnlg1502 [a]	AGGTCCTGGCACTAAGAGCA	AGAGGTGGTATGATCACCTGG
umc1082 [a]	CCGACCATGCATAAGGTCTAGG	GCCTGCATAGAGAGGTGGTATGAT
PZD00101.indel1 [a]	ATCGACCAACCAACTTCTCG	GCTTGGCAGTGGGTTAGTGT
umc1726 [a]	GATGAGGAAGAAAAGGGAAAAGGA	AGACTCAACCCTAACCCTAATGGG
bnlg1671 [a]	TCACGATCAGCAAGCAATTC	CCCCACCAACCTTAGAGTCA
umc1774 [a]	ATGGGACTATGCATGGTATTTTGG	TACACCATACGTCACCAGGTTCAC
umc2223 [a]	ACTTCTGCAGAGCGAGCAGG	TTTTGGGACTGAAGAAGAAGATCG
umc1500 [a]	TCTCTGACTATTCCACGAGCTCAA	CTGGTGCGTGCTACAACTGTG
umc1421 [a]	TGCTACGAACTGGGATACACTCAA	AGTGGTGAATGTGCCCTAGGAATA
GS1 [b]	ACACCGCCACCGACATCT	TTGTCCCTGAACGGCCAATA
CR Indel [c]	CGGTCAAAGAGTAGGGCAAG	GCGTCTGTTCCGCATTCA

[a] Markers used to map the introgressed teosinte segments

[b] Directly labeled FAM genescan marker used to genotype IS3 F_2 population.

[c] Agarose gel marker used to genotype all IS F_2 populations except IS3.