Evidence That the Origin of Naked Kernels During Maize Domestication Was Caused by a Single Amino Acid Substitution in *tga1*

Huai Wang,*,[1] Anthony J. Studer,*,[2] Qiong Zhao,*,[3] Robert Meeley,[†] and John F. Doebley*,[4]
*Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53076, and [†]DuPont Pioneer, Johnston, Iowa 50131

ABSTRACT *teosinte glume architecture1* (*tga1*), a member of the SBP-box gene family of transcriptional regulators, has been identified as the gene conferring naked kernels in maize *vs.* encased kernels in its wild progenitor, teosinte. However, the identity of the causative polymorphism within *tga1* that produces these different phenotypes has remained unknown. Using nucleotide diversity data, we show that there is a single fixed nucleotide difference between maize and teosinte in *tga1*, and this difference confers a Lys (teosinte allele) to Asn (maize allele) substitution. This substitution transforms TGA1 into a transcriptional repressor. While both alleles of TGA1 can bind a GTAC motif, maize-TGA1 forms more stable dimers than teosinte-TGA1. Since it is the only fixed difference between maize and teosinte, this alteration in protein function likely underlies the differences in maize and teosinte glume architecture. We previously reported a difference in TGA1 protein abundance between maize and teosinte based on relative signal intensity of a Western blot. Here, we show that this signal difference is not due to *tga1* but to a second gene, *neighbor of tga1 (not1). Not1* encodes a protein that has 92% amino acid similarity to TGA1 and that is recognized by the TGA1 antibody. Genetic mapping and phenotypic data show that *tga1*, without a contribution from *not1*, controls the difference in covered *vs.* naked kernels. No trait differences could be associated with the maize *vs.* teosinte alleles of *not1*. Our results document how morphological evolution can be driven by a simple nucleotide change that alters protein function.

KEYWORDS *tga1*; glume architecture; teosinte; maize; domestication

ALTHOUGH the study of adaptive evolution through natural or artificial selection dates back to Darwin, the genetic mechanisms that drive changes in morphology remain strongly debated (Stern and Orgogozo 2008). Questions surrounding the number of genes (few or many) and types of mutations (regulatory or coding) in these genes have been a focus in this debate (Carroll 2005, 2008; Hoekstra and Coyne 2007). These questions can only be answered through the genetic and molecular dissection of genes that have undergone selective pressure between lineages or within populations. Crop species offer a powerful system for investigating these questions since crops are the products of continuous directional selection to adapt them to the human controlled environment and human needs (Meyer and Purugganan 2013). Research on crop models is further facilitated by the extensive genetic and genomic resources available for them. Moreover, because of the recent divergence of crop species from their wild progenitors, crop-progenitor pairs remain cross-compatible and amenable to genetic analysis.

Maize was domesticated in the central Balsas valley of Mexico ~9000 years ago from a wild relative called teosinte (Piperno *et al.* 2001; Matsuoka *et al.* 2002). Because the morphology of modern maize is drastically different from teosinte species, the progenitor of maize was highly debated until molecular evidence proved the ancestral link (Mangelsdorf and Reeves 1938; Beadle 1939; Doebley 2001). Remarkably, genes controlling much of the morphological difference between maize and teosinte were shown to map to just six regions of the genome or major domestication loci (Doebley *et al.* 1990). Subsequent studies have sought to elucidate the genetic nature of these loci. Fine mapping of two of these regions, located on opposite arms of chromosome 1, identified causative polymorphisms in the regulatory region of *teosinte*

doi: 10.1534/genetics.115.175752
Manuscript received February 23, 2015; accepted for publication April 27, 2015; published Early Online May 4, 2015.
Supporting information is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.175752/-/DC1.
[1]Present address: Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, MO 63141.
[2]Present address: Department of Crop Sciences, 1201 W. Gregory Ave., University of Illinois, Urbana, IL 61801
[3]Present address: Otsuka America Pharmaceutical, Inc., 2440 Research Blvd., Rockville, MD 20850.
[4]Corresponding author: Laboratory of Genetics, University of Wisconsin, 425 Henry Mall, Madison, WI 53706. E-mail: jdoebley@wisc.edu

branched1 and *grassy tillers1* (Studer *et al.* 2011; Wills *et al.* 2013). These regulatory changes lead to altered plant architecture, and population genetic data indicate that the regulatory regions of these two genes were targets of selection during domestication. Recently, the domestication locus on chromosome 5 was shown to contain multiple factors (genes) rather than a single large effect gene (Lemmon and Doebley 2014).

Another major maize domestication locus, which is located on chromosome 4, controls whether the grains are enclosed in a "fruitcase" as in teosinte or uncovered as in maize (Dorweiler *et al.* 1993). The fruitcase that encapsulates the teosinte grain is formed from a hardened cup-shaped stem segment (cupule) in which the grain is located and a hardened bract (glume) that seals the grain in the cupule. In maize, the grains are borne naked on the exterior of the ear, and the organs that form the fruitcase in teosinte are redeployed to form the internal central axis of the ear (the cob). The transition from encased to exposed grain greatly facilitated the use of the grain as food. The locus that largely controls this difference has been resolved to a single gene called *teosinte glume architecture1* (*tga1*) (Wang *et al.* 2005). The maize allele of *tga1* disrupts the normal development of the cupulate fruitcase, exposing the grain on the surface of the ear. *tga1* encodes a squamosa-promoter binding protein (SBP), a transcription factor family that has been shown to regulate floral development (Klein *et al.* 1996; Wang *et al.* 2005). Although *tga1* has been identified as the major gene controlling changes in fruitcase development during domestication, the causative polymorphism in *tga1* and how this polymorphism affects the phenotype has not been resolved.

In this article, we show that a single fixed nucleotide polymorphism in the coding sequence of *tga1* distinguishes the maize and teosinte alleles. This difference creates an amino acid substitution that changes TGA1 protein dimerization and alters how TGA1 regulates its targets, with the maize allele acting more as a repressor relative to the teosinte allele. We also describe the pleiotropic effects of RNAi lines for *tga1* on multiple traits, indicating that *tga1* may play a broad role in development, having effects on kernel size and shape as well as plant architecture. Finally, we describe another gene, *neighbor of tga1* (*not1*), that is tightly linked to *tga1* and has a high sequence homology with *tga1*, but for which we observed no differences in phenotypic effect between the maize and teosinte alleles. Our molecular and genetic analyses of *tga1* show how a simple amino acid change can alter protein function and thereby drive the evolution of a new phenotypic state.

Materials and Methods

Detailed materials and methods can be found in Supporting Information.

Plant materials

Maize inbred W22 and W22:*tga1*, an introgression line that contains a teosinte chromosomal segment surrounding *tga1* in a W22 background, were used in most experiments (Dorweiler and Doebley 1997). A set of recombinant lines (T249, T1214, T1464, and T2956) derived from a W22 × W22:*tga1* F_2 fine mapping population that was previously described (Wang *et al.* 2005), were also utilized. The W22:*tga1-ems* line, previously reported in Wang *et al.* (2005), contains the amino acid substitution Leu5 to Phe5, which is immediately upstream of the single amino acid difference between maize- and teosinte-TGA1. The *tga1-ems* allele was recovered from an ethyl methanesulfonate mutagenesis and displays ear phenotypes similar to maize lines containing the teosinte allele of *tga1*. Additional plant materials include the *not1-Mu2* stock from the Trait Utilization System for Corn (TUSC) collection and the transgenic *tga1-RNAi* lines made at the Plant Transformation Facility at Iowa State University.

Protoplast transient assays

Dual luciferase reporter assays were used to determine the repressor function of *tga1* in transient protoplast expression experiments. N-terminal sequences of *tga1* were fused to a GAL4-DNA binding domain (DBD) and cotransformed with a firefly luciferase gene downstream of two GAL4 binding sites. Firefly luciferase expression was measured and then normalized to a *Renilla* luciferase internal control. Maize mesophyll protoplasts used for the transient expression experiments were extracted and transformed using a protocol developed by the Jen Sheen laboratory (see Supporting Information).

Protein purification from plant tissue

Protein was extracted from ear tissue using a plant total protein extraction kit (Sigma, St. Louis, MO) following the manufacturer's instructions. To show that TGA1 forms a dimer *in vivo*, formaldehyde was used to fix immature ear tissue prior to protein extraction.

Binding assays

Electrophoretic mobility shift assays (EMSAs) for PCR-assisted binding site selection experiments were performed as described previously (Tang and Perry 2003), and for experiments testing the *in vitro* binding of maize-TGA1 to the *not1* promoter, performed as described previously with some modifications (Wang *et al.* 2004). Chromatin immunoprecipitation (ChIP) experiments with an anti-TGA1 polyclonal antibody were used in to verify that TGA1 bound the *not1* promoter *in vivo*. ChIP assays were performed as described previously (Gendrel *et al.* 2002; Wang *et al.* 2002).

Results

Only one fixed nucleotide difference exists between maize and teosinte alleles of tga1

Wang *et al.* (2005) demonstrated by fine mapping that the causal polymorphism that distinguishes maize and teosinte lies within a 1042-bp segment of *tga1* (GenBank: AY883436–AY883460), which includes the first 18 bp of the ORF and 1024 bp upstream of the ATG. With a small sample of maize and teosinte alleles, these authors observed seven nucleotide

differences between maize and teosinte: six that are upstream of the start codon and one that is at position 18 of the ORF. To determine if these seven candidate polymorphisms could be narrowed to a smaller number, we assayed a larger sample (20) of teosinte alleles (Figure S1, GenBank: KR261098–KR261108). These data show that the six upstream polymorphisms no longer represent fixed differences between maize and teosinte but rather that some teosintes possess the same nucleotide as maize at each of these six sites. However, a nucleotide difference at position 18 of the ORF (C for maize and G for teosinte), which encodes a Lys6-to-Asn6 substitution from teosinte-TGA1 to maize-TGA1, still remains a fixed difference. Thus, this is the only fixed difference in the causative region that defines the glume architecture difference between maize and teosinte.

neighbor of tga1 is a tightly linked paralog of tga1

Examination of the genomic region near *tga1* revealed a gene (AC233751.1_FG002) that shares high nucleotide similarity to *tga1*. This gene is located only ~270 kb away from *tga1* and thus it was named *neighbor of tga1* (*not1*) (Preston *et al.* 2012). Comparing the TGA1 and NOT1 proteins from maize inbred W22, they have 92% identity in sequence. *not1* also exists in teosinte and a sequence alignment of different alleles of *tga1* and *not1* is shown in Figure S2.

To investigate whether *not1* contributes to the glume architecture difference between teosinte and maize, we investigated the glume phenotype of recombinant inbred lines (RILs) carrying different combinations of the *tga1* and *not1* alleles. RILs that carry *tga1-teosinte* alleles, such as T249 and W22:*tga1*, all show teosinte glume architecture regardless of whether they have the maize or teosinte allele at *not1* (Table 1). Similarly, RILs that possess the *tga1-maize* allele, such as W22, T1464, and T2956, all show maize glume architecture regardless of whether they have the maize or teosinte allele at *not1*. Thus, the teosinte glume architecture phenotype is associated with the *tga1-teosinte* allele but not related to the *not1-teosinte* allele.

We investigated the *tga1* and *not1* gene expression across different lines with RT-qPCR. As shown in Figure 1A, there is no statistical difference in message accumulation between the maize and teosinte alleles of *tga1* (ANOVA: $P = 0.9467$), and *tga1* message level is not affected by the genotype at *not1*. However, the *not1* message level is lowest when there are maize alleles at both *not1* and *tga1*, highest when there are teosinte alleles at both genes, and intermediate for the heteroallelic genotypes. These results indicate that the maize allele of *not1* accumulates less message than the teosinte allele, and they suggest that maize-TGA1 represses *not1* expression compared to teosinte-TGA1.

An antibody was generated using the TGA1 C-terminal protein and was used for a Western blot as part of the initial characterization of *tga1* (Wang *et al.* 2005). On the Western blot, the signal associated with protein samples for genetic stocks with the *tga1-teosinte* allele were stronger than for stocks with the *tga1-maize* allele (Wang *et al.* 2005). However, these stocks not only differ for their *tga1* allele but they also differ for their alleles at *not1*. This situation raises the possibility that the difference in signal strength on the Western blot was due to *not1* rather than *tga1*. If the anti-TGA1 antibody cross-reacts with NOT1, then the Western signal might be a combination of both TGA1 and NOT1.

Table 1 Glume architecture phenotypes of isogenic lines with different genotypes at *tga1* and *not1*

Genotype	Allele at *tga1*	Allele at *not1*	Glume architecture trait
W22	*maize*	*maize*	Maize like
T249	*teosinte*	*maize*	Teosinte like
T1214	*maize*	*teosinte*	Maize like
T1464	*maize*	*teosinte*	Maize like
T2956	*maize*	*teosinte*	Maize like
W22: *tga1*	*teosinte*	*teosinte*	Teosinte like
W22:*tga1-ems*	*ems*	*maize*	Teosinte like

To investigate this possibility, we performed a Western blot with proteins from comparable developmentally staged ears of our different RILs. The strongest signals were detected from the ears of the RILs that carry a *not1-teosinte* allele including W22:*tga1* and T1464 (Figure 1B). By comparison, line T249, which contains a teosinte allele at *tga1* but a maize allele at *not1*, has a signal intensity on the Western blot that is much less. In addition, the Western blot signals from W22:*tga1-ems*, which carries a *tga1-ems* allele and a *not1-maize allele*, do not show a dramatic difference in signal from W22. These observations suggest that strong signals in Western blots are associated with *not1-teosinte* instead of *tga1-teosinte*. Thus, the observation made by Wang *et al.* (2005) that the *tga1-teosinte* allele confers greater protein accumulation than the *tga1-maize* allele does not appear to be correct. Rather, the level of protein accumulation associated with these two alleles appears roughly equivalent.

We further investigated the TGA1 and NOT1 proteins by performing gel electrophoresis for an extended period of time to see if the TGA1 and NOT1 proteins could be separated (Figure 1C). In this analysis, we included an additional *not1* allele with a *Mu* element insertion (*not1-Mu2*) for which RT-qPCR shows no evidence of a transcript (Figure S3). The RILs with *not1-maize* all show two protein bands corresponding to TGA1 and NOT1. The *not1-Mu2* line is missing the lower of these two bands. Thus, the lower band appears to be NOT1. The RILs with *not1-teosinte* show a single thick band, which could be comigrating TGA1 and NOT1 proteins. Importantly, a comparison of RIL T249 and W22 shows that the band corresponding to NOT1 is stronger when *tga1-teosinte* is present than when *tga1-maize* is present. This observation is consistent with our interpretation of Figure 1B that the stronger protein signal for the W22:*tga1* line represents NOT1 rather than TGA1. The combined RT-qPCR and Western data suggest that *not1-teosinte* is expressed higher than *not1-maize* and that *tga1-maize* represses *not1* such that the greatest difference in *not1* expression is seen between lines that are *tga1-maize*; *not1-maize vs. tga1-teosinte*; *not1-teosinte*

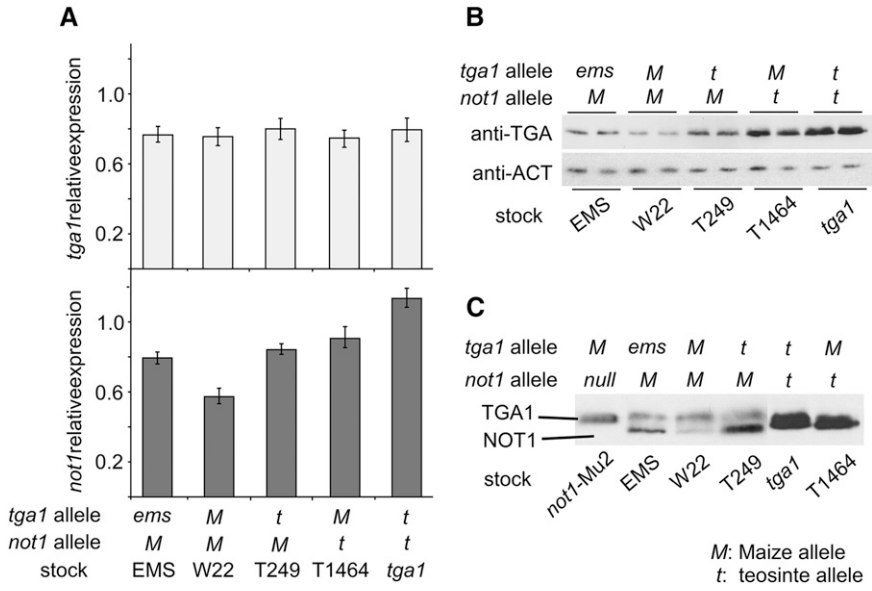

Figure 1 *tga1* and *not1* gene expression and protein accumulation. (A) RT-qPCR results showing *tga1* (top panel) and *not1* (bottom panel) expression. *tga1* expression was not statistically different between genetic stocks (ANOVA: $P = 0.9467$, $n = 10$). *not1* expresses less when *tga1* is a maize allele. Maize-*not1* expresses less in W22 than in T249 and EMS. Teosinte-*not1* expresses less in T1464 than in tga1. (B) Anti-TGA1 recognizes both TGA1 and NOT1; however, TGA1 and NOT1 cannot be separated with a short gel run performed in duplicate. (C) Teosinte-NOT1 cannot be separated from TGA1, but maize-NOT1 can be resolved under extended gel running conditions. The allelic compliment of each genetic stock is indicated *M*, maize allele; *t*, teosinte allele; *ems*, ems allele.

(Figure 1, A–C). Again, the statement by Wang *et al.* (2005) that the *tga1-teosinte* allele confers greater protein accumulation than the *tga1-maize* allele does not appear to be correct.

Lys6-to-Asn6 substitution converted TGA1 N-terminal into a repression domain

Based on the maize-teosinte sequence comparison and lack of an expression difference between the maize and teosinte alleles of *tga1*, our working hypothesis is that the single amino acid substitution (Lys to Asn) at the sixth position of *tga1* is the causative site for the loss of teosinte glume architecture during maize domestication. To test if the Lys6-to-Asn6 substitution controls the functional difference between maize- and teosinte-TGA1, we employed a protoplast transient assay system.

We constructed six effectors and two reporter constructs for the transient assays (Figure 2, A and B). We mixed plasmids with different effector and reporter combinations (Figure 2C) and introduced the mixture into maize mesophyll protoplasts by electroporation. After incubating the transfected protoplasts for 16–18 hr, the firefly luciferase activity was normalized to the *Renilla* luciferase activity for each assay. Transfection of the LD-VP16 effector along with the reporters gave strong firefly gene expression (Figure 2C); however, all other effectors (maize-GD, teosinte-GD, and ems-GD) along with the reporters in absence of LD-VP16 showed no activity in reporter gene expression relative to the negative control effector (GD) (Figure 2C, lightly shaded columns). These results suggest the TGA1 N-terminal domain is not an activation domain for any of the three alleles.

Cotransfection of LD-VP16 with an effector encoding the GAL4-DBD fused to the IAA17(I) (positive control for repression) resulted in repressed firefly expression (Figure 2C). Cotransfection of LD-VP16 with the effector carrying the maize allele TGA1 N-terminal fusion also showed repressed firefly luciferase activity. Interestingly, cotransfection with the TGA1 N-terminal domain from teosinte or ems allele effectors showed similar reporter activities to the GD effector (negative control). These results indicate that the N-terminal domain of maize-TGA1 is an active repression domain, but the N-terminal of teosinte-TGA1 or ems-TGA1 are not active repressors. The results also suggest that the repression activity of TGA1 N-terminal domain is established by just a single naturally occurring amino acid substitution (Lys6 to Asn6) and it was reversed by the ems-induced substitution (Leu5 to Phe5).

PCR-assisted binding site selection

To assay whether the Lys6-to-Asn6 substitution affects the DNA binding activity of the protein and to better understanding TGA1 function, we performed PCR-assisted binding site selection to determine the binding site of TGA1. Using EMSAs, we observed that both maize- and teosinte-TGA1 can shift up two bands but their patterns are opposite (Figure 3). Maize-TGA1 shifted up a stronger upper band while teosinte-TGA1 shifted up a stronger lower band. However, sequence consensus of DNA fragments derived from these shifted bands showed that they all contain a GTAC motif. This is consistent with previous reports that SBP-domain proteins bind to GTAC (Birkenbihl *et al.* 2005; Kropat *et al.* 2005). These results indicate that the Lys6-to-Asn6 substitution did not affect the binding site specificity of TGA1 for the GTAC motif, but likely affected the configuration of TGA1 binding to DNA.

TGA1 binds to the promoter of not1 in vitro and in vivo

We isolated the promoter segments of *not1-maize* and *not1-teosinte* alleles (Figure S4, GenBank: KR261109 and KR261110). Interestingly, both sequences share a conservative region that contains the GTAC motif. We took partial sequence surrounding this GTAC motif to make a probe for EMSA (Figure 4A). Our results show this sequence from the *not1* promoter can bind maize-TGA1 *in vitro* (Figure 4B). Furthermore, the binding activity of maize-TGA1 to this DNA fragment was

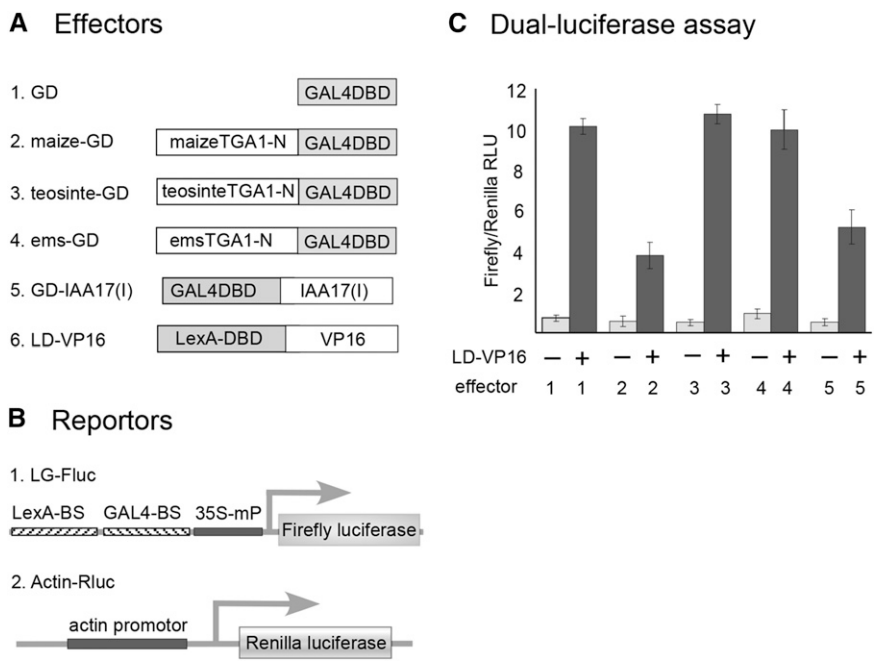

Figure 2 Protoplast transient dual-luciferase assays show maize-TGA is a transcriptional repressor. (A) Diagram showing effector construct components. (B) Diagram showing reporter construct components. (C) Effectors 1–5 were cotransformed with both reporters into maize protoplasts with and without the activation control (effector 6). Dual-luciferase assay results are shown in relative light units (RLU), *y*-axis, (±SE; $n = 6$). Maize-TGA1 (effector 2) shows stronger repression than the repression control (effector 5).

abolished totally with the GTAC motif mutated to CTAC (Figure 4B). These results suggest that TGA1 binds to the *not1* promoter at the GTAC motif.

A ChIP assay using an anti-TGA1 antibody was performed to confirm the *in vivo* interaction between TGA1 and the *not1* promoter. Because anti-TGA1 can also cross-react with NOT1, we used tissue from plants with the *not1-Mu2* allele, which is a null mutant for *not1*. The ChIP-PCR results showed that the DNA fragment of the *not1* promoter region was enriched in the anti-TGA1 ChIP population as compared to the input control (Figure 4C). Furthermore, ChIP populations generated using non- or preimmune serum did not show specific selection of the *not1* fragment. To quantify the degree of enrichment, we also performed qPCR, which shows the *not1* promoter region was enriched ~4.8-fold in the anti-TGA1 population (ChIP = 4.68 ± 0.27; input control = 0.97 ± 0.11) (Figure 4D).

Maize-TGA1 forms more stable dimers than teosinte-TGA1

The DNA probes used for EMSA in Figure 3 and Figure 4 are different, but they both showed doublet bands. Thus, the double banding pattern is not likely related to the DNA probe sequences. To determine if the doublet was caused by protein dimerization, we performed additional EMSA using tagged and tag-free versions of TGA1 that are expected to migrate differentially during electrophoresis and thus be distinguishable. In addition to these two full-length versions of TGA1, we also used a shortened version of TGA1 with the N terminus removed.

As shown in Figure 5A, all the full-length TGA1 proteins, including tag-maize-TGA1, tag-teosinte-TGA1, tag-ems-TGA1, and maize-TGA1, shifted up two bands. The upper band is the major band when using tag-maize-TGA1 or maize-TGA1 while the lower band is the major band when using tag-teosinte-TGA1 or tag-ems-TGA1. These results suggest that TGA1 exists as dimers and monomers dynamically, and that maize-TGA1 tends to form more stable dimers.

In Figure 5A, one can also see that the DNA–protein complexes with tag-maize-TGA1 (version 1) run slower in the gel than the complex from tag-free maize-TGA1 (version 2). When we mixed tag-maize-TGA1 and tag-free maize-TGA1 together in the same binding reactions, we found that a novel intermediated band was detected. We interpret this new band as a heterodimer of tag-maize-TGA1/tag-free maize-TGA1 proteins. Intriguingly, a truncated TGA1 with 103 amino acids removed from the N terminus only shifts up a single band. These results indicate that TGA1 can form dimers and that the N terminus of TGA1 is necessary for dimerization. We conclude that the Lys6-to-Asn6 mutation, which is located in the TGA1 N terminus, can affect the dynamic ratio between dimers and monomers.

To investigate whether TGA1 forms dimers *in vivo*, we performed a Western blot assay using formaldehyde cross-linked ear tissue from *not1-Mu2* plants. Formaldehyde can covalently preserve protein dimers *in vivo*, thus prevent the dimer disassociation during protein preparation and electrophoresis (Sang *et al.* 2005). As shown in Figure 5B, only a single band was detected at ~50 kDa from nonfixed tissue, and this band is approximately the size of TGA1 as monomer. However, when using cross-linked tissue, an additional band, which is approximately twice the weight of the TGA1 monomer, was recognized by the anti-TGA1 antibody. These results suggest that maize-TGA1 forms homodimers *in vivo*, which is consistent with our *in vitro* assay data.

tga1/not1 loss-of-function plants via RNAi and their phenotype

To understand the function of *tga1*, we screened the available maize mutant collections for *tga1* loss-of-function

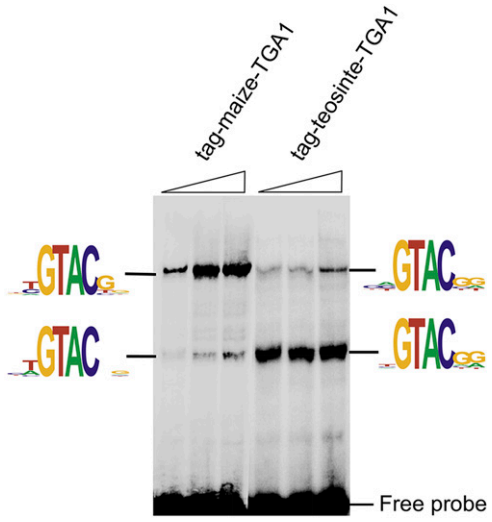

Figure 3 Binding site selection for maize-TGA1 and teosinte-TGA1. α-^{32}P-labeled DNA fragments that contain 26 randomized nucleotides were incubated with maize or teostine-TGA1. After five rounds of EMSAs, the shifted DNA bands were excised, and the consensus binding motif was obtained. Both maize-TGA1 and teosinte-TGA1 proteins bind to DNA fragments with a GTAC core motif. For each allele, 50 ng, 100 ng, and 200 ng of TGA1 protein are shown (left to right).

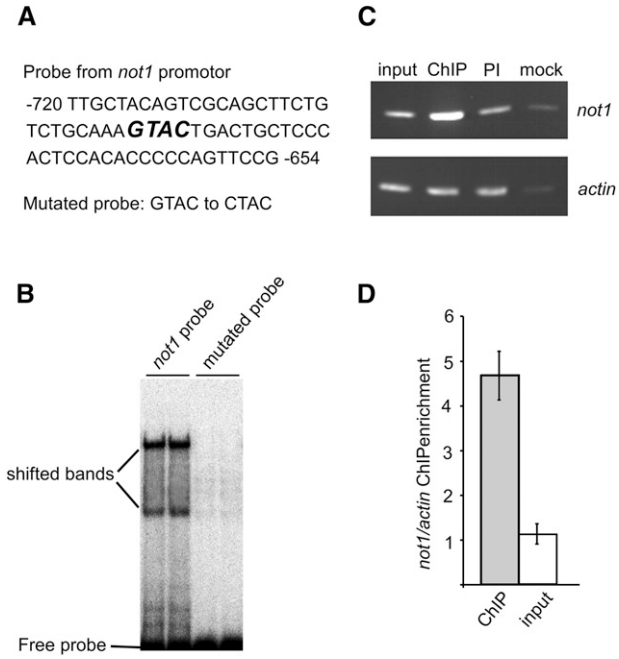

Figure 4 TGA1 binds the *not1* promoter *in vitro* and *in vivo* through the GTAC core motif. (A) The *not1* promoter DNA probe sequence containing the GTAC motif. (B) EMSA showing that maize-TGA1 binds to the *not1* promoter *in vitro* through the GTAC motif. The *not1* probe shifts up two bands with maize-TGA1 protein, but the band shifting is abolished when the GTAC core is mutated to CTAC. (C and D) Chromatin immunoprecipitation (ChIP) confirms that TGA1 binds the *not1* promoter *in vivo*. (C) RT-PCR amplification of the *not1* promoter region containing the GTAC core motif. Input, total input chromatin DNA before precipitation; ChIP, chromatin DNA precipitated with anti-TGA1 antibody; PI, DNA precipitated with preimmune serum; mock, no antibody or serum added. (D) RT-qPCR showing the relative enrichment of the *not1* promoter in the anti-TGA1 ChIPs (±SE; n = 7).

alleles without success (see *Materials and Methods*). We did not recover any mutant alleles for *tga1*; however, we found 5 *Mu* insertions in *not1*. The *not1-Mu2* allele is a null mutation lacking detectable expression (Figure S3), although we did not observe any phenotypic difference between homozygous *not1-Mu2* plants and wild-type sibs in a segregating population.

We attempted to generate transgenic *tga1* overexpression plants by putting *tga1-maize* or *tga1-teosinte* under a constitutive rice *Actin1* promoter. Surprisingly, no *tga1* overexpression transgenic plants could be generated with either the maize or teosinte allele. *tga1* contains a miRNA156 target site, and thus miRNA156 could inhibit the overexpression. To rule out this possibility, we built two new constructs with the miRNA156 site in *tga1* mutated synonymously and placed them under a maize *Ubiqitin* promoter. However, we still could not produce transgenic plants and there was difficulty even obtaining transgenic callus with these vectors. Other control constructs that were transformed side by side worked fine as reported to us by Dr. Kan Wang at the Plant Transformation Facility, Iowa State University. These results suggest that the overexpression of *tga1* inhibits some aspect of plant development and thus plant regeneration during the transformation process.

We next used an RNAi-based approach to generate *tga1* loss-of-function plants. Because RNAi is achieved by overexpression of a hairpin structure (Kusaba 2004), it is likely not efficient enough to generate a complete knockout of *tga1* mRNA, especially during plant regeneration. Using this approach, we were able to recover *tga1-RNAi* transgenic plants. Five of the *tga1-RNAi* events were grown to maturity and crossed to maize inbred W22. Then, 60 progeny plants from each cross were grown and the segregation of transgene was identified by a Basta leaf painting assay. Progeny groups for four events showed an approximate 1:1 ratio for Basta resistance:susceptible plants, suggesting that these four events have single insertions.

As shown in Figure 6, *tga1-RNAi* plants present some interesting phenotypes. The *tga1/not1* loss-of-function plants have much longer lateral branches (Figure 6, A–C). A few lateral branches even have secondary ears developed along them (Figure 6C). The glumes of *tga1-RNAi* plants are enlarged, which is a characterization of W22:*tga1* glumes (Figure 6, F, G, J, and K). However, these glumes are relatively paperish and less thick, less hard, and less polished. The kernels from the *tga1-RNAi* plants are narrower (Figure 6I) compared to control kernels (Figure 6H). Furthermore, they have a pointed tip, which makes the ear "prickly" (Figure 6M), while the control ear is smooth (Figure 6L). Interestingly, the *tga1-RNAi* plants have prop roots developed at the first four to six nodes of the stalk, while the control plants only have two nodes with prop roots (Figure 6, D and E). This collection of phenotypes suggests that *tga1/not1* not only have function in glume architecture, but also function in juvenile growth, lateral branch formation, and ear

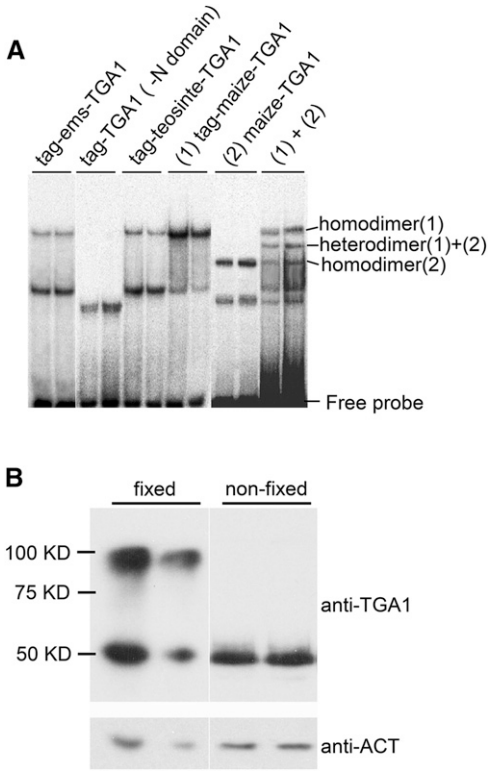

Figure 5 TGA1 forms a dimer *in vitro* and *in vivo*. (A) An EMSA showing that maize-TGA1 forms more stable dimers than teosinte- or ems-TGA1. Removing the N-terminal domain of TGA1 abolished the dimerization of TGA1, shown in lane3 and 4. A mixture of maize-TGA1 proteins with/without tags for EMSA confirmed TGA1 forms homo and heterodimers (lane11 and 12). (B) Protein from formaldehyde cross-linked ears produce two bands on a Western blot at ~50 kDa and 100 kDa, while protein from unfixed ears only have the lower band.

development. The fact that the *tga1-RNAi* plants display some characteristics of the *tga1-teosinte* glume phenotype suggests that the Lys6 to Asn6 from teosinte to maize is a gain-of-function mutation.

Quantitative traits associated with tga1-RNAi

Statistically significant associations were identified between the *tga1-RNAi* transgene and some quantitative traits related to plant and ear architecture. We compared plants with and without the *tga1-RNAi* transgene as determined by a Basta painting assay. We analyzed 30 resistance (TG = transgene present) and 30 susceptible (NTG = transgene absent) plants each from segregating F_2 families for four events. Associations between the phenotypes and *tga1-RNAi* transgene genotype were tested using *T*-tests.

As shown in Figure 7, the *tga1-RNAi* transgene is associated with statistically significant effects on lateral branch number, length of the uppermost lateral branch, the length of the blade of the first husk leaf, and the number of nodes with prop roots. The results suggest that TGA1/NOT1 represses the growth of lateral branches in length and numbers. The association of some ear traits with *tga1-RNAi* was also observed (Figure 7). Glume length was significantly increased when TGA1/NOT1 were knocked down/out. In contrast, ear diameter and the weight of 50 kernels decreased in *tga1-RNAi* plants. We also measured some additional ear traits, such as ear length and kernel row number; however, there were no statistically significant associations between these traits and the *tga1-RNAi* transgene.

In summary, these results suggest that TGA1/NOT1 have broader functions beyond controlling glume architecture in maize and teosinte. Specifically, they may also affect traits such as seed shape, seed weight, ear and lateral branch morphology, as well as the juvenile seedling phase of plant growth.

Discussion

A previous population genetic analysis of *tga1* identified seven fixed nucleotide differences between maize and teosinte using a small sample of accessions (Wang *et al.* 2005). Here, we used a larger sample size of maize and teosinte, which revealed that six of these seven nucleotide sites are polymorphic within either maize or teosinte (Figure S1). With this larger sample, only a single fixed difference was identified between maize and teosinte within *tga1*. This single site encodes an Lys6-to-Asn6 substitution between teosinte and maize at the sixth amino acid position from the N terminus of the protein. This observation makes the Lys6-to-Asn6 substitution the most likely candidate for the causative difference between maize and teosinte in *tga1*.

We tested the functional consequences of the Lys6-to-Asn6 substitution using protoplast transient assays (Tiwari *et al.* 2004). These assays revealed that this single amino acid substitution alters the function of the TGA1 protein. The Lys6-to-Asn6 substitution gives the maize-TGA1 a strong repressor function as seen by reporter expression levels that are comparable to the repressor control (Figure 2). Furthermore, this amino acid substitution also increases dimerization of TGA1, as shown by gel shift assays (Figure 5). The increased dimer formation and repressor function of the maize-TGA1 are likely concomitants. Arguably, the most well-characterized repressor, the *lac* repressor (LacR) in *Escherichia coli*, forms a stable dimer of dimers, which then produces a functional tetramer through weak associations (reviewed in Lewis 2005). Amino acid substitutions that weaken dimer formation of LacR eliminate repressor function. Furthermore, a single amino acid substitution has been shown to block dimerization (Dong *et al.* 1999; Spott and Dong 2000). Thus, the observed correlation of dimerization and repressor strength reported here is consistent with functional studies in other systems.

The binding site of TGA1 was determined using EMSA and shown to be GTAC (Figure 3). Both the maize- and teosinte-TGA1 were used; however, no difference in binding site preference was observed. Thus, the amino acid substitution that leads to repressor function is likely the mechanism by which plant morphology is altered, not differences in binding site specificity. The *tga1* binding motif was found

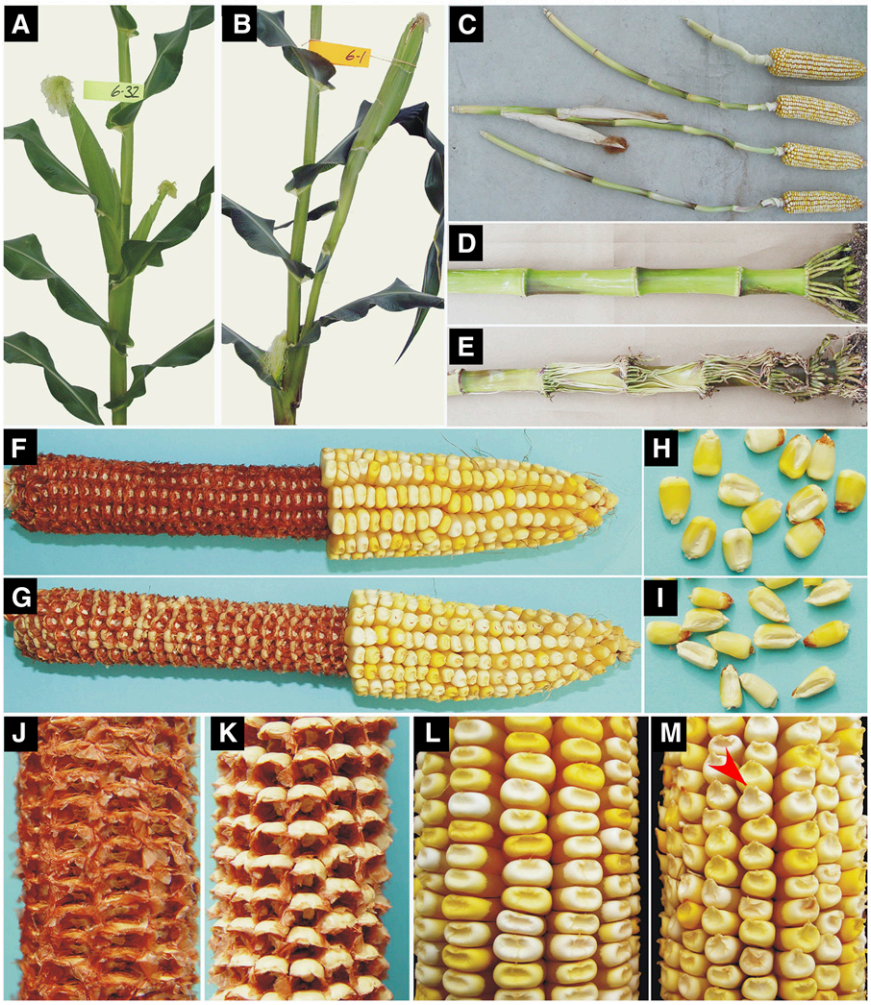

Figure 6 Phenotypes of *tga-RNAi* plants. (A–C) *tga1-RNAi* plants have longer lateral branches. (A) A nontransgenic control plant with a short lateral branch. (B) A typical *tga1-RNAi* plant from event 4 with a long lateral branch. (C) Lateral branches with leaves removed; top one is from nontransgenic control; the others are from *tga1-RNAi* plants, which have long shanks and some even have secondary ears. (D and E) A typical *tga1-RNAi* plant (E) has prop roots on more nodes than a control plant (D). (F–M) The ear and kernel phenotypes. G, I, K, and M are from *tga1-RNAi*. F, H, J, and L are controls. Compared to the control, *tga1-RNAi* ears have enlarged glumes (G and K) and narrow shaped seeds (I). Furthermore, each *tga1-RNAi* kernel is pointed (M) on the surface, which results in a spiky ear.

in the promoter of *not1* and TGA1 binds this site both *in vitro* and *in vivo*, suggesting that *not1* and *tga1* may function in the same pathway. This result begins to unravel a potential cascade of gene expression changes that accompanies the alteration of a major domestication gene.

tga1 was discovered based on the specific effects of the maize *vs.* teosinte allele on the development of the teosinte fruitcase and maize cob (Dorweiler *et al.* 1993). Although these alleles differ for the Lys6-to-Asn6 amino acid substitution, both alleles are expressed at comparable levels (Wang *et al.* 2005). We assayed the broader effects of *tga1* by using an RNAi construct to reduce or eliminate *tga1* gene expression. Maize lines expressing an RNAi construct targeting *tga1* displayed pleiotropic morphological effects on several branching and kernel traits (Figure 6 and Figure 7), which had not previously been associated with this domestication gene (most recently Brown *et al.* 2011). With regard to branching, these RNAi lines likely remove the repressive function of TGA1/NOT1, allowing the outgrowth of axillary branches. The effects on kernel shape and size may be related to the fact that the kernel resides within the fruitcase in teosinte, and thus, fruitcase and kernel development are coordinately regulated by *tga1*. Whether the maize *vs.* teosinte allele of *tga1* affects any of these additional traits is unknown, but such effects have not previously been reported (Dorweiler and Doebley 1997).

An open question is whether the effects of the RNAi construct on kernel and branching traits result from a knockdown of *tga1* or *not1*. The RNAi construct was generated using *tga1* sequence, but given the sequence similarity between *tga1* and *not1*, it is likely that the RNAi construct targets both genes. However, given that the *Mu* insertions in *not1* did not produce the morphological phenotypes seen in the RNAi lines, we infer that the phenotypes observed in the RNAi lines are either attributable to *tga1* or a redundant function of *tga1* and *not1* and thus can only be observed when both are knocked down. However, without a null *tga1* allele, we cannot show conclusively that these phenotypes are specific to *tga1*. These inferences about *tga1 vs. not1* function are further complicated by the binding of TGA1 to the *not1* promoter. While multiple morphological phenotypes are observed in the RNAi lines, the single amino acid substitution fixed during domestication seems to be specific to the ear traits reported previously (Wang *et al.* 2005).

Both *tga1* and *not1* belong to the SBP family of transcription factors. Members of this family have been shown to

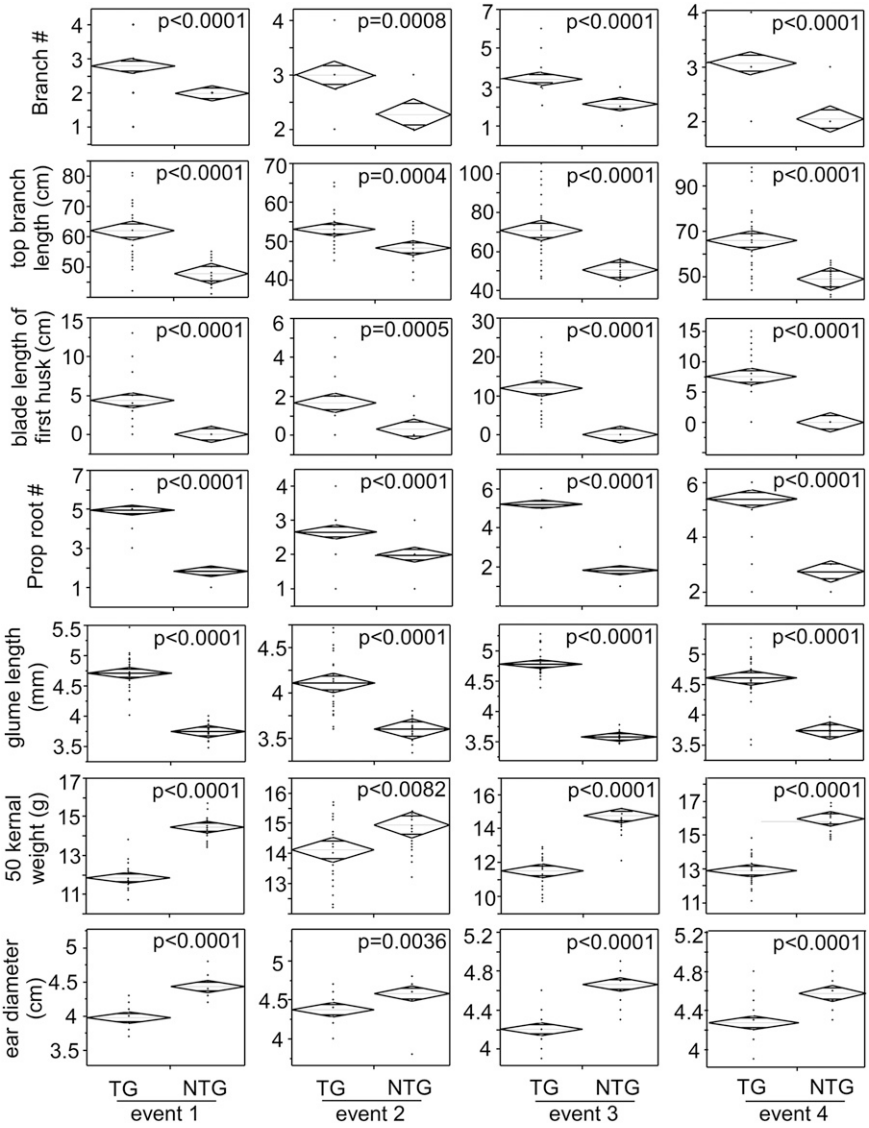

Figure 7 Quantitative effects associated with the *tga1-RNAi* lines. Trait values were plotted on the *y*-axis and dots represent the value range for each phenotype. Diamonds are centered on the mean values for the traits. Data from four events are presented with TG signifying plants that contain the *tga1-RNAi* transgene and NTG for nontransgenic plants. The *P*-value for each *T*-test is displayed in the upper right corner of each graph.

regulate meristem development, and manipulations of these regulators have produced both plant architecture and ear phenotypes (Chuck *et al.* 2010, 2014). While there is a clear homolog for *tga1* in other grass species, the *tga1/not1* duplication occurred at the base of the *Zea* genus, evident by the absence of this duplication in other lineages (Preston *et al.* 2012). The presence of the *tga1/not1* duplication in maize may have facilitated the subfunctionalization of *tga1/not1* such that *tga1* alone controls the fruitcase/cob in teosinte/maize while *tga1* functions in a redundant manner with *not1* to regulate plant architecture traits. This would explain the phenotypes observed in the RNAi lines, which likely target both *tga1/not1*, when no phenotype was present in *not1* mutant plants. Furthermore, various *tga1* alleles have only been reported to display ear phenotypes (Wang *et al.* 2005; Brown *et al.* 2011). This hypothesis is further supported by work on the ortholog of *tga1* in rice (LOC_Os08g41940; *OsSPL16*), which has pleiotropic plant and inflorescence phenotypes (Wang *et al.* 2011, 2012).

In this article, we have shown that an amino acid substitution in *tga1* is the causal variant that underlies the origin of the naked grains of maize as compared to the covered grains of teosinte. Although the predominant mechanism for morphological evolution may be alterations in gene expression (Carroll 2008), changes in protein function are also involved as shown here. Investigation of how protein evolution contributes to the evolution of new morphological forms enhance our understanding of how adaptions arise.

Acknowledgments

We thank Bao Kim Nugyen, Jesse Rucker, Tina Nussbaum Wagler, and Lisa Kursel for technical assistance. This work was supported by National Science Foundation grants IOS1025869 and IOS1238014 and US Department of Agriculture–Hatch grant MSN169062.

Literature Cited

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

Birkenbihl, R. P., G. Jach, H. Saedler, and P. Huijser, 2005 Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. J. Mol. Biol. 352: 585–596.

Brown, P. J., N. Upadyayula, G. S. Mahone, F. Tian, P. J. Bradbury *et al.*, 2011 Distinct genetic architectures for male and female inflorescence traits of maize. PLoS Genet. 7: e1002383.

Carroll, S. B., 2005 Evolution at two levels: on genes and form. PLoS Biol. 3: 1159–1166.

Carroll, S. B., 2008 Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134: 25–36.

Chuck, G., C. Whipple, D. Jackson, and S. Hake, 2010 The maize SBP-box transcription factor encoded by tasselsheath4 regulates bract development and the establishment of meristem boundaries. Development 137: 1243–1250.

Chuck, G. S., P. J. Brown, R. Meeley, and S. Hake, 2014 Maize SBP-box transcription factors unbranched2 and unbranched3 affect yield traits by regulating the rate of lateral primordia initiation. Proc. Natl. Acad. Sci. USA 111: 18775–18780.

Doebley, J., 2001 George Beadle's other hypothesis: one-gene, one-trait. Genetics 158: 487–493.

Doebley, J., A. Stec, J. Wendelt, and M. Edwardst, 1990 Genetic and morphological analysis of a maize-teosinte F2 population: implications for the origin of maize. Proc. Natl. Acad. Sci. USA 87: 9888–9892.

Dong, F., S. Spott, O. Zimmermann, B. Kisters-woike, B. Mu *et al.*, 1999 Dimerisation mutants of Lac repressor. I. A monomeric mutant, L251A, that binds Lac operator DNA as a dimer. J. Mol. Biol. 290: 653–666.

Dorweiler, J., and J. Doebley, 1997 Developmental analysis of teosinte glume architecture1: a key locus in the evolution of maize (Poaceae). Am. J. Bot. 84: 1313–1322.

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

Gendrel, A., Z. Lippman, and C. Yordan, 2002 Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science 297: 1871–1873.

Hoekstra, H. E., and J. Coyne, 2007 The locus of evolution: evo devo and the genetics of adaptation. Evolution 61: 995–1016.

Klein, J., H. Saedler, and P. Huijser, 1996 A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. Mol. Gen. Genet. 250: 7–16.

Kropat, J., S. Tottey, R. P. Birkenbihl, N. Depège, P. Huijser *et al.*, 2005 A regulator of nutritional copper signaling in Chlamydomonas is an SBP domain protein that recognizes the GTAC core of copper response element. Proc. Natl. Acad. Sci. USA 102: 18730–18735.

Kusaba, M., 2004 RNA interference in crop plants. Curr. Opin. Biotechnol. 15: 139–143.

Lemmon, Z. H., and J. F. Doebley, 2014 Genetic dissection of a genomic region with pleiotropic effects on domestication traits in maize reveals multiple linked QTL. Genetics 198: 345–353.

Lewis, M., 2005 The lac repressor. C. R. Biol. 328: 521–548.

Mangelsdorf, P. C., and R. G. Reeves, 1938 The origin of maize. Proc. Natl. Acad. Sci. USA 24: 303–312.

Matsuoka, Y., Y. Vigouroux, M. M. Goodman, G. J. Sanchez, E. Buckler *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite genotyping. Proc. Natl. Acad. Sci. USA 99: 6080–6084.

Meyer, R. S., and M. D. Purugganan, 2013 Evolution of crop species: genetics of domestication and diversification. Nat. Rev. Genet. 14: 840–852.

Piperno, D. R., K. V. Flannery, and N. Cave, 2001 The earliest archaeological maize (*Zea mays L.*) from highland Mexico: new accelerator mass spectrometry. Proc. Natl. Acad. Sci. USA 98: 2101–2103.

Preston, J. C., H. Wang, L. Kursel, J. Doebley, and E. A. Kellogg, 2012 The role of teosinte glume architecture (tga1) in coordinated regulation and evolution of grass glumes and inflorescence axes. New Phytol. 193: 204–215.

Sang, Y., Q. Li, V. Rubio, Y. Zhang, J. Mao *et al.*, 2005 N-terminal domain– mediated homodimerization is required for photoreceptor activity of Arabidopsis. Plant Cell 17: 1569–1584.

Spott, S., and F. Dong, 2000 Dimerisation mutants of Lac repressor. II. A single amino acid substitution, D278L, changes the specificity of dimerisation. J. Mol. Biol. 296: 673–684.

Stern, D. L., and V. Orgogozo, 2008 The loci of evolution: How predictable is genetic evolution? Evolution 62: 2155–2177.

Studer, A., Q. Zhao, J. Ross-Ibarra, and J. Doebley, 2011 Identification of a functional transposon insertion in the maize domestication gene tb1. Nat. Genet. 43: 1160–1163.

Tang, W., and S. E. Perry, 2003 Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. J. Biol. Chem. 278: 28154–28159.

Tiwari, S. B., G. Hagen, and T. J. Guilfoyle, 2004 Aux/IAA proteins contain a potent transcriptional repression domain. Plant Cell 16: 533–543.

Wang, H., W. Tang, C. Zhu, and S. Perry, 2002 A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. Plant J. 32: 831–843.

Wang, H., L. V. Caruso, A. B. Downie, and S. E. Perry, 2004 The embryo MADS domain protein AGAMOUS-like 15 directly regulates expression of a gene Encoding an enzyme involved in gibberellin metabolism. Plant Cell 16: 1206–1219.

Wang, H., T. Nussbaum-Wagler, B. Li, Q. Zhao, Y. Vigouroux *et al.*, 2005 The origin of the naked grains of maize. Nature 436: 714–719.

Wang, S. S., C. S. Wang, T. H. Tseng, Y. L. Hou, and K. Y. Chen, 2011 High-resolution genetic mapping and candidate gene identification of the SLP1 locus that controls glume development in rice. Theor. Appl. Genet. 122: 1489–1496.

Wang, S., K. Wu, Q. Yuan, X. Liu, Z. Liu *et al.*, 2012 Control of grain size, shape and quality by OsSPL16 in rice. Nat. Genet. 44: 950–954.

Wills, D. M., C. J. Whipple, S. Takuno, L. E. Kursel, L. M. Shannon *et al.*, 2013 From many, one: genetic control of prolificacy during maize domestication. PLoS Genet. 9: e1003604.

Communicating editor: J. A. Birchler

GENETICS

Supporting Information
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.175752/-/DC1

Evidence That the Origin of Naked Kernels During Maize Domestication Was Caused by a Single Amino Acid Substitution in *tga1*

Huai Wang, Anthony J. Studer, Qiong Zhao, Robert Meeley, and John F. Doebley

DOI: 10.1534/genetics.115.175752

Supporting Information

File S1

Material and Methods

Nucleotide Variation in *tga1*. We augmented the sample of teosinte individuals used by Wang *et al.* (2005) for nucleotide diversity analysis by sequencing an additional 11 teosinte plants (Table S1) for the 1042 bp control region defined by Wang *et al.* (2005) which includes the first 18 bp of the ORF and 1024 bp upstream of the ORF. The primer sequences, PCR programs, and sequencing protocols are the same as those used by Wang *et al.* (2005). For heterozygous DNA samples, PCR fragments were cloned into TOPO®-TA vectors (Life Technologies) and multiple clones were sequenced to acquire the individual allele sequences without potential Taq polymerase incorporated errors. To determine the putative causative site underlying the glume architecture difference between maize and teosinte, we aligned the complete set of 16 maize and 20 teosinte sequences and visually inspected the alignment for fixed differences between maize and teosinte.

Screening for *tga1* Mutants and Identification of *neighbor of tga1* (*not1*). Efforts to obtain a loss-of-function allele of *tga1* were made by screening several maize mutant collections, including maize targeted mutagenesis (May et al. 2003), maize TILLING project (Till et al. 2004) and Pioneer Hi-Bred International's Trait Utility System for Corn (TUSC). We also performed informatics searchers of the databases for the RescueMu (Lunde et al. 2003) and UniformMu (Settles et al. 2007) projects. No *tga1* mutants were found in any of these collections, however, five *Mu* insertion events were identified in *neighbor of tga1* (*not1*), a gene closely related to *tga1* (Preston et al. 2012), from TUSC using a *Mu* terminal repeat primer (5'-agagaagccaacgccawcgcctcyatttcgtc-3') with a primer designed from *tga1* (5'agaaagcgtctggacgggcacaatc3').

Quantitative PCR. RNA extractions were performed, as previously described (Wang et al. 2005), on immature top ears that were 25mm in length. All gene expression comparisons were based on plants from the same nursery location and year. cDNA was produced as previously reported (Wang et al. 2005) and then used for quantitative PCR (RT-qPCR). RT-qPCR was performed on ABI Prism 7000 sequence detection system (Applied Biosystems) with Power SYBR® Green PCR Master Mix (Applied Biosystems) according to the manufacturer's instructions. Specific forward primers were used for *tga1* (5'-cagtgcagcaggttccatc-3') and *not1* (5'-cgcatcactcaccaaatcca-3') along with a common reverse primer (5'gaagcttatctcgccctcct3'). We sequenced PCR products from 20 independent clones for each primer set to confirm that they were specific to either *tga1* or *not1*. A maize actin gene (AY104722, GRMZM2G126010) was selected as a control, and amplified using the primers 5'- ccaaggccaacagagagaaa

-3’ and 5’- ccaaacggagaatagcatgag -3’ (Bomblies 2003). Gene expression quantification was performed using 10 biological replicates for each genetic stock.

Protoplast Transient Assays. The TGA1 N-terminal coding sequences (amino acids 1-103) from the maize, teosinte and ems alleles were fused to a GAL4-DNA Binding Domain (DBD) followed by the *Nopaline synthase (Nos) terminator* (maize-GD, teosinte-GD and ems-GD respectively). These fusion gene effectors were expressed using the rice *Actin1* promoter. An effector with the GAL4-DBD alone (GD) served as the negative control. The effector GD-IAA17(I), encodes a GAL4-DBD fused to the IAA17 protein domain I, which is an known repressor (Tiwari et al. 2004). This effector served as a positive control for trans-repression. The effector LD-VP16 encodes an *E. coli* LexA-DBD fused to the *Herpes simplex virus* VP16 activation domain (LaMarco and McKnight 1989). This effector serves as a positive control for activation. Both the GD-IAA17(I) and LD-VP16 effectors are under the control of the cauliflower mosaic virus 35S minimal promoter (-46bp). These two effectors were generous gifts from Dr. Tom J. Guilfoyle at the University of Missouri-Columbia.

Two reporters were constructed for the transient assays. The LG-Fluc reporter contains a firefly luciferase gene under the control of the cauliflower mosaic virus 35S minimal promoter (-46bp) with two LexA and two GAL4 binding sites upstream. The LexA and GAL4 binding sites upstream of the minimal promoter allowed the fusion gene effectors containing Gal4-DBD or LexA-DBD to bind and drive the expression of the firefly luciferase gene. The Actin-Rluc reporter contains a *Renilla* luciferase gene under rice *Actin1* promoter. This reporter was co-transformed into the protoplasts with the LG-luc reporter and served as an internal control. Using both firefly and *Renilla* luciferase enabled us to take advantage of using the dual-luciferase reporter assay system (Promega Corp., Madison, WI) to take two readings from the same samples.

Transient expression assays using maize mesophyll protoplasts were performed following a detailed protocol from the Sheen’s lab (http://molbio.mgh.harvard.edu/sheenweb/protocols_reg.html). Briefly, 2-3 x 10^5 freshly isolated protoplasts in 400 □l electroporation buffer were mixed with 50 □l of plasmids which include different combination of effectors and reporters. All the plasmids were prepared with EndoFree Plasmid Mega Kit (Qiagen) and each effector : reporter LG-Fluc : reporter Actin-Rluc mixture contained 25 □g : 15 □g : 5 □g, respectively. The protoplast-plasmid mixes were transferred in 0.5 ml cuvettes and electroporated with a Gene Pulser II Electroporation System (Bio-rad) set to 250 Volts for 1.5 msec. Three pulses spaced 20 sec apart were used for each sample. After electroporation, protoplasts were incubated for 18 hrs at 25 °C before harvesting for quantification. The harvested protoplasts were then lysed with CCLR (Cell Culture Lysis Reagent, Promega) and subjected to dual-luciferase assay using a Dual-Luciferase Reporter Assay System (Promega) following the manufacturer’s instruction. Six biological replicates were measured per effector/reporter mixture.

Heterologous Protein Expression and Western Blot Assays. Tagged TGA1 proteins were generated by cloning various full-length *tga1* cDNA alleles, and a *tga1* maize allele with an N-terminal coding region deletion, into the pET151/D-TOPO® vector (Invitrogen) and transforming into BL21 codonPlus (DE3)-RIRL cells (Stratagene). All of the proteins cloned into the pET151/D-TOPO® vector are expressed with a 6xHis-V5-tag derived from the vector. Overexpressed proteins were extracted using BugBuster® Master Mix (Navagen) and then solubilized from inclusion bodies and purified under denaturing condition (with 6M urea) using a His-bind purification Kit (Navagen). The purified proteins were then re-natured by dialyzing through serial refolding solution with decreasing urea concentration (4 M, 2 M, 1 M to 0 M). The refolding solution consists of 0.1 M Tris, pH 8.0, 0.4M L-Arginine, 0.1 mM ZnCl2, 5 mM DTT, 0,2 mM PMSF and plant protease inhibitor cocktail (Sigma). Dialysis was performed in 2 liter volume at 4 °C with at least 2 hours in between solution exchanges. The purity and integrity of the proteins were assessed by SDS-PAGE and western analysis.

A soluble tag-free maize-TGA1 was generated, but could only be retained at low concentrations (≤ 5 ng/ul). The maize full-length *tga1* cDNA was cloned into a pVP-GW vector (Singh et al, 2005), which can produce proteins with S-His-tag and enhanced solubility. Proteins were overexpressed in *E. coli*, and then extracted and purified using Ni-NTA Superflow Columns (Qiagen) following the manufacturer's instructions. After purification, the S-His-tag was cleaved off using AcTEV protease (Invitrogen). All proteins were diluted to low concentration (< 40 ng/µl) prior to the TEV protease reaction to prevent precipitation of TGA1 upon cleavage of the tag. Final purification of the tag-free TGA1 proteins was performed by passing them through Ni-NTA Superflow Columns.

Western blot assays were performed on immature ear tissue 25mm in length using the anti-TGA1 antibody as previously described (Wang et al. 2005). The mini gel for the western blot in Figure 1B was run at 150 Volts for 80 min. The blot was probed with anti-TGA1 first, and then stripped and re-probed with anti-ACTIN to confirm equal loading. A large gel that was run at 150 Volts for 18 hours, was used for the western blot in Figure 1C to resolve TGA1 and NOT1. Immature *not1-Mu2* ears were used for the western blot in Figure 5B.

PCR-Assisted Binding Site Selection. PCR-assisted binding site selections were performed via Electrophoretic Mobility Shift Assays (EMSAs) as described previously (Tang and Perry 2003). A 76-mer oligo (5'-actcgaggaattcggcaccccgggt$(n)_{26}$tggatccggagagctcccaacgcgt -3') containing a core of 26 randomized nucleotides was synthesized and then converted to double-stranded DNA by PCR using two primers that corresponding to the non-degenerate ends of the

76-mer (5'-actcgaggaattcgctaccccgggt -3' and 5'- acgcgttgggagctctccggatcca -3'). The double-stranded 76-mer population was labeled with [α-^{32}P]dATP and purified on a 8% polyacrylamide gel for use as an EMSA probe.

EMSA were performed as reported previously (Wang et al. 2004) with slight modifications. Approximately 100 ng of maize-TGA1 or teosinte-TGA1 was pre-incubated in 1× binding buffer (50 mm Tris–HCl, pH 7.5, 50 mm NaCl, 1 mM ZnCl2, 1 mM DTT, 10% glycerol, 0.1 μg μl^{-1} poly(dI-dC), 0.5 μg μl^{-1} BSA) at room temperature for 10 min before adding 10^4-10^5 cpm of the radiolabeled DNA fragments. Then, the entire reaction mix (20 μl) was incubated for an additional 10 min and analyzed on a 5% PAGE gel containing 0.5 x TBE and 2.5% glycerol. The gel was dried and exposed to a PhosphorImager screen (Amersham Biosciences). The regions corresponding to shifted bands were excised and electroeluted. Then, an aliquot from each elutant was used for PCR amplification and radiolabeling to generate new probes. The new probes were used for the next round of binding and EMSA and this procedure was repeated for 5 rounds. Finally, the enriched TGA1 bound oligonucleotides were cloned into pCR2.1-TOPO® cloning vector (Invitrogen) and at least 40 clones from each shifted band were sequenced and aligned to obtain the consensus sequence.

EMSA Testing *in vitro* Interaction Between maize-TGA1 and *not1* Promoter. An 825 bp of promoter segment of *not1-maize* as well as 480 bp of 5' end sequence of *not1-teosinte* were isolated using the GenomeWalker Universal Kit (Clontech, Mountain View, CA) following the manufacture's instructions. A DNA fragment corresponding to the *not1* regulatory region (from –389 to –343) and containing a GTAC motif (–372 to –361) was generated using two partially complementary oligonucleotides that were extend by PCR. The two oligonucleotides were as follows: oligonucleotide 1, 5'- ttgctacagtcgcaactgtctgtgtctgcaaaGTACgactgct -3', and oligonucleotide 2, 5'- ccggaactgggggtggagtgggagcagtcaGTACtttgcagacaca -3'. A mutated version of the DNA fragment was created using oligonucleotide 3 (same as oligonucleotide 1 except that the **G**TAC was changed to **C**TAC) and oligonucleotide 4 (same as oligonucleotide 2 except that except that the **G**TAC was changed to **C**TAC). The wild-type and mutant DNA fragments were labeled with [α-^{32}P]dATP and EMSA performed as described above.

Chromatin Immunoprecipitation (ChIP) Assays. ChIP assays were performed to verify *in vivo* binding of TGA1 to DNA fragments containing the GTAC motif from the *not1* regulatory region. Immature ear primordia (1-5 cm in length) from *not1-Mu2* plants were isolated and sliced to pieces, with at least one dimension less than 5 mm, and then fixed in 1% formaldehyde. The experimental procedure was done as described previously (Gendrel et al. 2002; Wang et al. 2002). An anti-TGA1 polyclonal antibody was used for chromatin immunoprecipitation. DNA populations recovered from ChIP were then used as templates for

qPCR with a maize actin gene as the control. Primers specific to *not1* promoter region are 5'- acaggtgcacagcacaacat -3' and 5'- agcagcagccaacaaagatt -3'. qPCR was performed as described above using 7 biological replicates.

***tga1-RNAi* Plants Generation and Phenotyping.** An RNAi vector that targets *tga1* was constructed by generating a hairpin loop with an inverted exon3 sequence (amplified using the primers 5'-GAGAGTCCATATCACATCACTCACC-3' and 5'-CACTAAAGCCAGTATCTCCCTACCAGATCGTC-3') link to intron2-exon3 of *tga1* (5'-AGTCTGTACTTGGTTCGAATACATCTAC-3' and 5'-CACTAAAGCCAGTATCTCCCT-3'). The hairpin structure was driven with a maize ubiquitin promoter (Christensen and Quail 1996) and transformed into Hi-II maize using the Agrobacterium-mediated method. Maize transformation was done at the Plant Transformation Facility at Iowa State University (http://agron-www.agron.iastate.edu/ptf/index.aspx).

Different transgenic events of *tga1-RNAi* were recovered and crossed to W22. Then, 80 progeny from each cross were grown out and subjected to a BASTA painting assay. Briefly, a 1% solution of glufosinate (BASTA) solution with 0.01% Tween was painted on the tip of an expanded leaf when plants were at 8-10 leaf stage. Four days after the BASTA application, the plants were scored for resistance or susceptibility to the herbicide. The correlation between RNAi transgene presence and BASTA resistance in plants were further validated by western blot (Figure S4). Western blots using anti-TGA1 were performed with four BASTA resistant and four susceptible plants for each event using young ear tissue. Consistent with expectation, TGA1 was not detected, or had a weak signal, in all ear samples from BASTA resistant plants, while samples from BASTA susceptible plants produced a strong signal with western blot analysis.

Thirty BASTA resistant and 30 susceptible plants, from segregating F_2 families derived from a cross between BATSA resistant T_0 plants with W22, were phenotyped for plant and ear architecture traits. These include lateral branch number, length of the uppermost primary lateral branch, blade length of the first husk leaf of the top ear, number of nodes with prop roots, glume length, the weight of 50 kernels, ear diameter, and ear length. To measure glume length, the kernels were removed from the middle of the ears and then the cobs were broken to expose the glumes for measurement. Eight glumes were measured from each ear and the average value calculated. T-tests were used to evaluate whether the trait values were significantly different between plants with/without the *tga1-RNAi* transgene. Segregating families from four independent transgenic events were assayed.

Literature Cited

Bomblies, K. 2003 Duplicate FLORICAULA/LEAFY homologs zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. Development 130: 2385–2395.

Christensen, A. H., and P. H. Quail. 1996 Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res. 5: 213–218.

Gendrel A., Z. Lippman, and C. Yordan, 2002 Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science. 297: 1871–1873.

LaMarco, K. L., and S. L. McKnight. 1989 Purification of a set of cellular polypeptides that bind to the purine-rich cis-regulatory element of herpes simplex virus immediate early genes. Gene. Dev. 3: 1372–1383.

Lunde, C. F., D. J. Morrow, L. M. Roy, and V. Walbot. 2003 Progress in maize gene discovery: a project update. Func. Integr. Genomics 3: 25–32.

May, B. P., H. Liu, E. Vollbrecht, L. Senior, P. D. Rabinowicz, et al. 2003 Maize-targeted mutagenesis: A knockout resource for maize. Proc. Natl. Acad. Sci. USA 100: 11541–11546.

Preston, J. C., H. Wang, L. Kursel, J. Doebley, and E. A. Kellogg. 2012 The role of teosinte glume architecture (tga1) in coordinated regulation and evolution of grass glumes and inflorescence axes. New Phytol. 193: 204–215.

Settles, M., D. R. Holding, B. C. Tan, S. P. Latshaw, J. Liu, et al. 2007 Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. BMC Genomics 8: 116.

Singh, S., C.C. Cornilescu, R. C. Tyler, G. Cornilescu, M. Tonelli, et al. 2005 Solution structure of a late embryogenesis abundant protein (LEA14) from Arabidopsis thaliana, a cellular stress-related protein. Protein Sci. 14: 2601-2609.

Tang, W., and S. E. Perry. 2003 Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. J. Biol. Chem. 278: 28154–28159.

Till, B. J., S. H. Reynolds, C. Weil, N. Springer, C. Burtner, et al. 2004 Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol. 4: 12.

Tiwari, S. B., G. Hagen, and T. J. Guilfoyle. 2004 Aux/IAA Proteins Contain a Potent Transcriptional Repression Domain. Plant Cell. 16: 533–543.

Wang H., Tang W., Zhu C., Perry S., 2002 A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. Plant J. 32: 831–843.

Wang, H., L. V. Caruso, A. B. Downie, and S. E. Perry. 2004 The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism. Plant Cell 16: 1206–1219.

Wang, H., T. Nussbaum-Wagler, B. Li, Q. Zhao, Y. Vigouroux, et al. 2005 The origin of the naked grains of maize. Nature 436: 714–719.

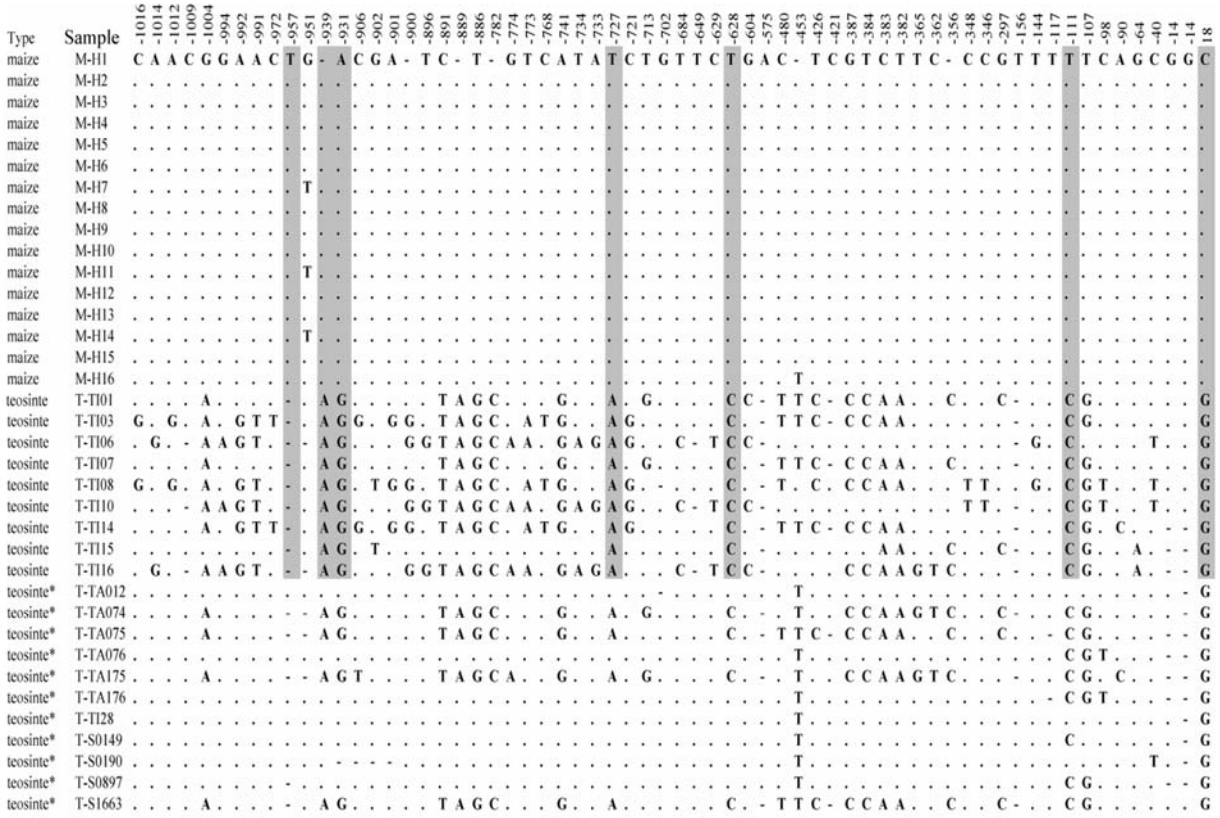

Figure S1 Nucleotide sequence alignment (polymorphic sites only) of the 1042-bp region of *tga1* identified as containing the causative site (WANG *et al.* 2005). A sample of maize and teosinte genotypes showing that the only fixed difference between maize and teosinte is in the 18th nucleotide of the open reading frame. Numbers in the top row indicate the nucleotide positions corresponding to the W22:*tga1* sequence. Indel polymorphisms were shown as single mutations. Dots indicate identity to the sequence in the top row. The seven "fixed" differences initially identified by Wang *et al.* (2005) are shown by shading. Eleven teosinte individuals not sampled by Wang *et al*. (2005) are indicated by asterisks.

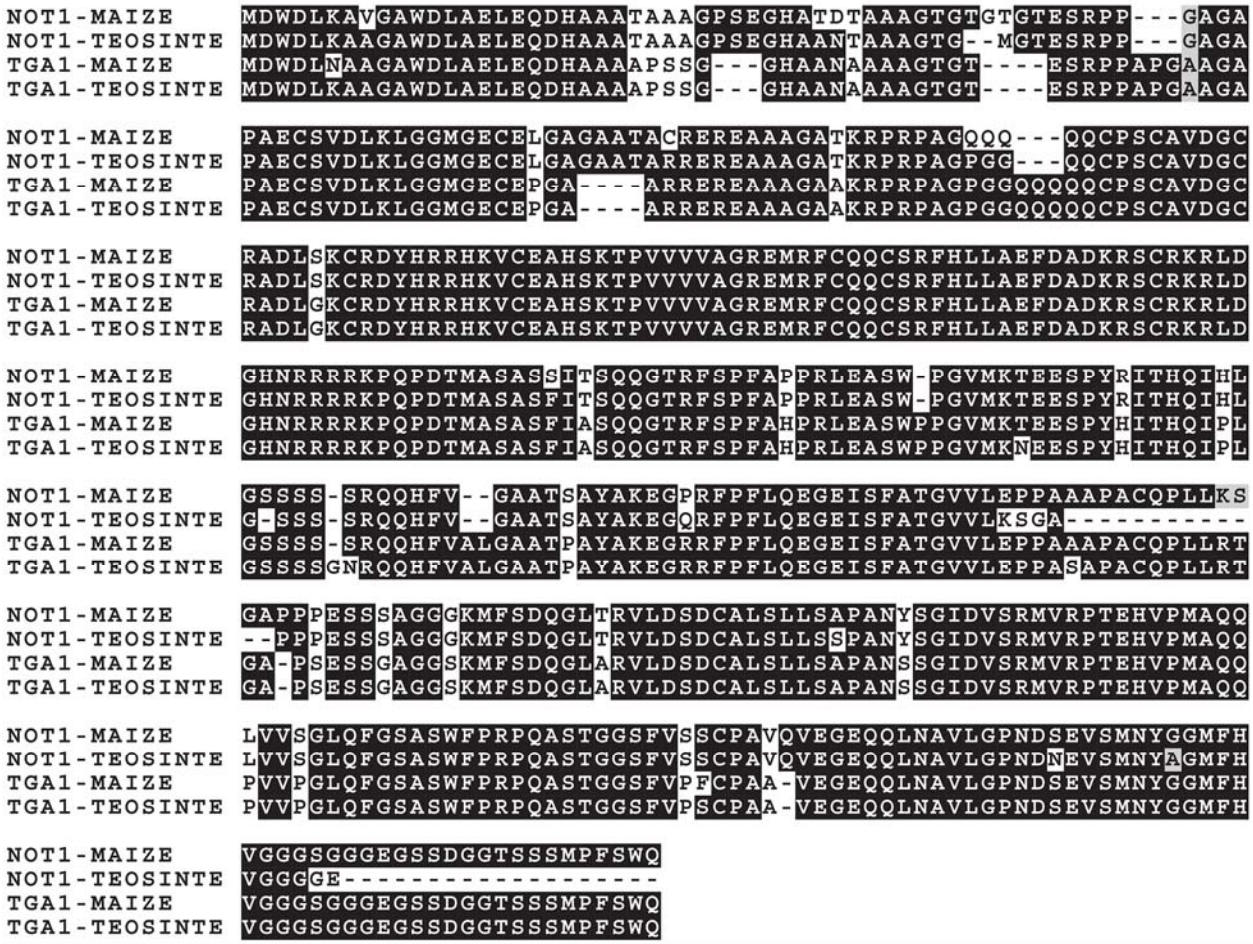

Figure S2 Amino acid alignment of TGA1 and NOT1 from the maize inbred W22 and the teosinte parent of the mapping population used by Wang *et al.* (2005).

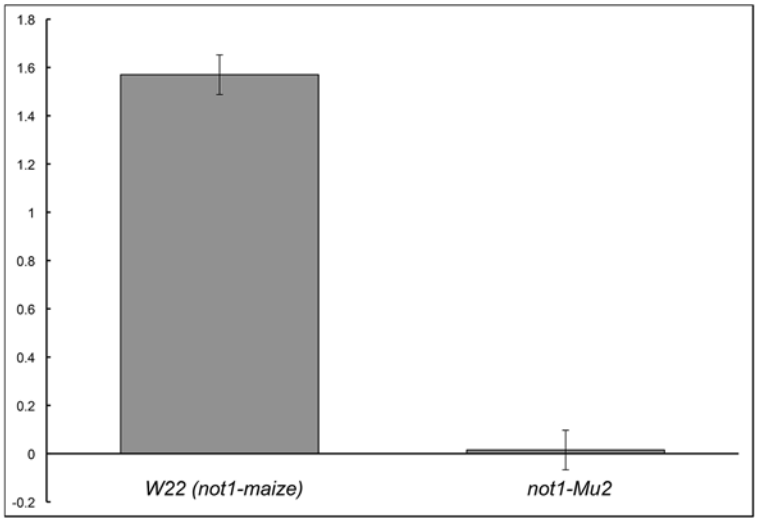

Figure S3 Column graph showing results of RT-qPCR using *not1* specific primers. Expression of *not1* was normalized to a maize actin gene. *not1-Mu* expression (0.015 ± 0.082) was 100-fold less than *not1-maize* (1.570 ± 0.082) in W22 and not statistically different from 0 (no expression).

>*not1-maize* promotor (from W22)

ATCATTTCCACTTCTTGAGTTACCGTTTCCAATTTCATTTTTTTAAAAAATATA
AAAATAAAAATACTTTTAATATTTATCACTGTTTCCGTCCTGCCTCCCTTGCTA
CAGTCGCAACTGTCTGTGTCTGCAAAGTACTGACTGCTCCCACTCCACCCCCA
GTTCCGGGCCTCTTGGGCATCGCTTCTGTTGCTGCAAATCTTTGGGCCTGTTTG
GTTCAGCTTTTTTCTGACCAGCTTTTCTGAGAATCTGGCTGTAGGAAGAATCT
GGCTGTGTAGAGAATCTGTGTATCATTAGGATTACGTGCGGAGGAAGATAAA
GGTGTCCATAGGACTCAGGATATAGAAAGTGACGGATTCCTACTATTGCAAC
GACTCAACCGATTATGTGTTTATGTTGATTTTGAATGATTTTTACTCAAACGA
ATTTTATAGAAGCTGACTGAAAAGCTGAGCGTTTGGCAGTCCACAACAGCTT
TTGGTGGCGAGAAGCTCAAAAAAGCTGAAACAAACAGTGGCTTTGTTGCCTG
CTGCTGCCTTGCTATCTATCTATACACCGCCCGCACCTTCCATTCCTCCTCCTC
CGAAGCAGCAGGCAGCAGCTGCATCGCACCTCACACCTCTCGTGTCCATCGA
TCCAGCCGCCCGCCGCAGCTGCAGCTCTCACTGTTGCTGTGCCACCTCCTCGT
CGCCTGTAGTGTCGACAGAGAAACGCCCGCTGAATGAGAGGGAAGGAACGG
AAGCTGCAGCGGGCGCGCGAGTGCAAGGACTAGGACTAGCGGTTGCAACGT
CGGCGCGCGCGCGCGTGCGGCGTACGTCGGGC**ATG**

>*not1-teosinte* promotor (from W22:*tga1*)

AAATACGTATATATTATTACTAATTTTTTTGTGGATGTACAACTTGTGTGAATT
TTTTAAGTCAATTGAGGGTGGATGTTATATTTTCACCAATTGTATTTTAGATTC
CATACGTTTATTTTTAGGGAAATTTGGATCCATACCATTAAAAGATCACCACT
TTGGATCCATACCATTACTATCTCACTTACATGTGGGTCCACATGAGTCAATG
ACATGTGGGGTCCATGGTATATATCTAAAGTTTGGATCTTTTAATGGTATAGA
TCCAATTGTTCCTTATTTTTATACCAAAGTTTTGTTTTTCATGTACACGAATTA
CTGATATCATTTTCACTTCTGGAGTTACCGTTTCTAATTTCATTTCCTGAAAAA
TGTAAAAATAAAAATACTTTTAATATTTATCGCTGTTTCCGTCCTGCCTCCCTT
GCTACAGTCGCAGCTGTCTGTGTCTGCAAAGTACTGACTGCTCCCACTCCACT
CCACCCCCAGTTCCGGGCCACTCGGGCATCGTTTCTGTTGCTGCAAATCTTTG
TTGGCTGCTGCTGCCTTGCTATCTATCTATACACCGCCCGCACCTTCCATTCCT
CCTCCTCCGAAGCAGCAGGCAGCAGCTGCATCGCACCTCACACCTCTCGTGT
CCATCGATCCAGCCGCCGCCGCAGCTCTCACTTCACTGTTGCTGTGCCACCTC
CTCGTCGCCTGTAGTGTCGACAGAGAAACGCCCGCGGAATGAGAGGGAAGG
AACGGAAGCTGCAGCGGGCGCGCGAGTGCAAGGACTAGGACTAGCGGTTGC
AACGTCGGGC**ATG**

Figure S4 Nucleotide sequences of the maize and teosinte *not1* promoter sequences. The GTAC motif bound by tga1 is highlighted in yellow. The start codon (ATG) for the ORFs are shown in bold.

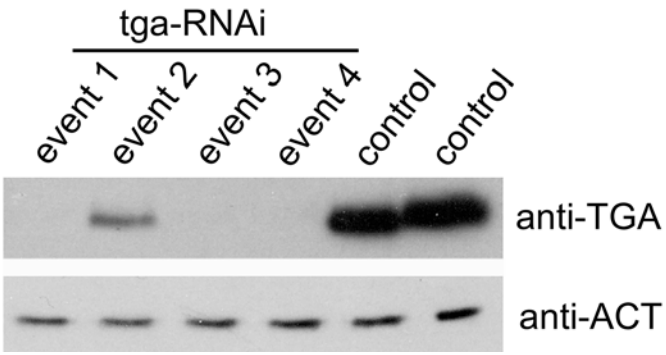

Figure S5 Western blot showing the TGA1 and NOT1 protein accumulation were knocked down or out in 4 different *tga1-RNAi* events. Blot was probed with anti-TGA1 first (top panel), then it was stripped and probed with anti-ACTIN show equal loading. Controls is from plants with out transgene.

Table S1 List of additional teosinte accessions for diversity analysis

Sample	Genbank ID	Accession	Source	Taxon	Country	State/Province	Collector	Collection	Altitude	Latitude	Longitude
S1663	KR261108	G-120	HHI	Z. mays ssp. huehuetenangensis	Guatemala	Huehuetenango	H. Iltis	G-120	1350	N15:39	W91:46
S0149	KR261105	11369	CIMMYT	Z. mays ssp. mexicana	Mexico	Guanajuato	T. A. Kato	K-69-7	1925	N20:10	W101:5
S0190	KR261106	11387	CIMMYT	Z. mays ssp. mexicana	Mexico	Chihuahua	Wilkes, Sanchez, Taba	WST-85-2	1850	N26:14	W106:58
S0897	KR261107	11372	CIMMYT	Z. mays ssp. mexicana	Mexico	Guanajuato	T. A. Kato	K-69-10	2100	N20:5	W101:13
TA175	KR261102	8765	CIMMYT	Z. mays ssp.	Mexico	Michoacan	T. A. Kato	K-67-22	950	N19:5	W100:28

				parviglumis							
TA176	KR261103	8765	CIMMYT	Z. mays ssp. parviglumis	Mexico	Michoacan	T. A. Kato	K-67-22	950	N19:5	W100:28
TA074	KR261099	69	INIFAP	Z. mays ssp. parviglumis	Mexico	Mexico	Sanchez & Ordaz	JSG Y LOS - 172	1620	N19:10	W100:12
TA075	KR261100	69	INIFAP	Z. mays ssp. parviglumis	Mexico	Mexico	Sanchez & Ordaz	JSG Y LOS - 172	1620	N19:10	W100:12
TA076	KR261101	69	INIFAP	Z. mays ssp. parviglumis	Mexico	Mexico	Sanchez & Ordaz	JSG Y LOS - 172	1620	N19:10	W100:12
TA012	KR261098	18	INIFAP	Z. mays ssp. parviglumis	Mexico	Jalisco	Sanchez & Ordaz	JSG Y LOS - 3	1520	N20:25	W102:21
TI28	KR261104	G-347-3	Kermicle	Z. mays ssp. mexicana	Mexico	Chihuahua	Kermicle	G-347-3	1850	N26:14	W106:58