Dear Dr. Schnabel and editorial staff at the *American Journal of Botany*,

We would like to thank you and two anonymous reviewers for your helpful comments and suggestions regarding our manuscript. In the attached revision, we have addressed the majority of issues raised, but, while we feel that the new experiments suggested by Reviewer 1 would provide interesting results, they would require considerable extra time (many months) and are outside of the scope of this current work. In our submitted manuscript, we report results of a genotyping assay surveying several hundred wild and domesticated maize individuals for presence of the domesticated maize allele at *tb1*, Sanger sequencing of a targeted subset of these samples to assess evidence of selection on the domesticated allele in wild populations, a phenotyping experiment to gauge the effects of the domesticated allele in a wild population, and reanalysis of existing genome-wide SNP data to place the signature of selection at this locus in a wider context. This manuscript reports a substantial amount of novel work and explores, in a meaningful way, the concept of domestication from standing variation and the potential ecological relevance of domesticated alleles in wild populations. In addition to our revised manuscript, we include a detailed response to reviewers below.

Sincerely,

Matthew Hufford and Jeffrey Ross-Ibarra

Response to Associate Editor:

Comments from the Associate Editor and Reviewers are in black text and our responses are in blue text

I thought that the neighbor-joining analysis was of little value, because it was not clear to me just how the result could be expected to provide information about Hopscotch genotype. Why would you expect individuals with similar Hopscotch genotypes to cluster together? Also, although you don't provide any information on the distribution of variable sites among accessions and taxa, your data set of only 48 and 40 segregating sites is almost certainly not sufficiently variable to resolve relationships among the very large number of accessions you were analyzing. As a result, you produce mostly a huge polytomy, with a small amount of structure that appears to reflect the

genetic similarities and differences already evident in Table 1. The trees themselves are difficult for a reader to interpret, because none of the accession names at the branch tips (other than TIL) are defined in the text.

We agree entirely with this assessment of the Neighbor Joining trees. Our intention with the NJ trees was to show that there is no clear signal of selection on the *Hopscotch* genotype across sequence in the *tb1* genomic region. Presumably, if the *Hopscotch* element were under strong selection in teosinte populations we would expect to see *tb1* haplotype groups in an NJ tree due to linkage disequilibrium and Hill-Robertson effects. In fact, the original discovery of selection on the *Hopscotch* insertion during maize domestication was based on such a signal from linked sequence 60kb distant from the *Hopscotch* (Studer et al., 2011). We agree with the reviewer that the trees could be confusing and difficult to interpret, and the large polytomies do not provide information beyond our assertion that there is not strong haplotype structure broadly in the *tb1* genomic region that correlates with the *Hopscotch* genotype. We have therefore opted to keep the NJ trees in the supplement where interested readers can view them but where they don't unnecessarily complicate our main story. In addition, we have added explanation of the accession names in the supplemental figure legend.

Phenotyping experiments - These are curious experiments. First, why did you not generate a population of individuals with known genotypes and then test for the effect of genotype on phenotype? Why take a random sample and just hope that you will get the necessary number of individuals of each genotype? How can you randomize the growing conditions with respect to genotype, if you don't know genotypes ahead of time but instead determine them after all the growing measurements are taken?

We agree that a planned grow-out with previously determined genotypes could be useful. However, prior work has already shown that *Hopscotch* has a phenotypic effect in isogenic and inbred backgrounds (Studer et al., 2011), and our interests lie in the effects of the *Hopscotch* in genetic backgrounds occurring in nature. Moreover, our sampling of individuals from this population for our grow-out was not random. We initially genotyped a single individual from each sampling site within the population and then targeted our sample to include individuals from sampling sites where individuals were homozygous or heterozygous for the *Hopscotch* allele. Our reasoning in using this

approach was that we wanted a high proportion of both *Hopscotch* positive and *Hopscotch* negative individuals for phenotypic comparisons and, given that the *Hopscotch* allele is typically rare in teosinte, we wanted to enrich our sample for this allele. We were overly successful in this approach and had a higher frequency of individuals with the *Hopscotch* positive allele. We have clarified our sampling methodology on pages 8-9, lines 174-185. Finally, plants were situated randomly in the greenhouse prior to genotyping, which effectively represents a randomized experimental design.

Why are sample sizes not reported? You provide total samples sizes (eg, 206 or 216 for Phenotyping 1, depending on whether you read the Methods or the Results), but not sample sizes for each genotypic class.

We have reported sample size information for all three genotypes on page 11 lines 266-267.

It was not clear to me what "to detect the observed effect" means in lines 12-13 of p. 9. Your Phenotyping 1 results show no effect of genotype on tillering. Thus, it is not clear what effect you are referring to. What was the purpose of the Phenotyping 2 experiment? You already showed no effect of genotype on phenotype using populations where you are most likely to see segregation for the insertion, so it was not clear to me why sampling more broadly, especially from populations with no evidence of the Hopscotch insertion, would be expected to improve your understanding of this relationship.

On Day 40 of the Phenotyping 1 experiment we saw a weak positive correlation between the *Hopscotch* insertion and tillering index (p=0.0848). Using data from Phenotyping 1, we conducted a power analysis to determine the number of individuals we would need in both homozygous genotypic classes to observe a significant result. This analysis indicated we would need 71 individuals in both classes to detect a significant difference in tillering index. For the Phenotyping 2 experiment, we attempted to sample enough individuals from the San Lorenzo population (the same population sampled for Phenotyping 1) in order to detect a significant difference between homozygous genotypes. Our Phenotyping 2 experiment detects the same trend reported in Phenotyping 1, but the trend is still not significant. We agree with the editor that this second, non-significant result adds little to the manuscript and we have thus removed Phenotyping 2 and refer to a single grow-out in our revised version of the manuscript.

Note that sampling for both phenotyping experiments was limited to a single population from which we had a large number of seed collected. Seed was collected from various "sampling sites" within the contiguous population at which we had estimated *Hopscotch* allele frequency. For our broader genotyping of landrace maize, ssp. *parviglumis*, and ssp. *mexicana* we included samples from all populations to which we had access. For most of these populations the prevalence of the *Hopscotch* was previously unknown, and we were interested in having an idea of the overall distribution and frequency of the *Hopscotch* throughout the natural ranges of these taxa.

The discussion of introgression, genetic drift, and selection in your Discussion section seems to lack coherence. It sometimes focuses on explaining the unexpectedly high frequency of Hopscotch in present-day populations, whereas other times it seems to be addressing the presence or absence of Hopscotch in teosinte in general, and it also sometimes seems to be addressing selection on the tb1 locus.

We have edited this section for clarity to better emphasize our main point that the *Hopscotch* allele, which is known to produce a domesticated phenotype of reduced tiller number, is at higher than expected frequency in a subset of teosinte populations and may play an adaptive role in certain environments encountered by teosinte. Surprisingly, based on our grow-out of individuals from the San Lorenzo population, tiller number is not a major phenotypic difference between teosinte plants with and without the *Hopscotch* insertion, perhaps indicating the insertion affects multiple phenotypes in teosinte

Response to Reviewer 1

Reviewer #1: In their paper 'Natural variation in teosinte at the domestication locus teosinte branched1 (tb1)', Vann et al survey a large sample of maize and teosinte individuals for the presence of the Hopscotch transposable element upstream of tb1. The authors find that Hopscotch is more widespread, across a large sampled area in Mexico, than previously thought, particularly in parviglumis. The authors go on to sequence regions up and downstream of Hopscotch in subset of their initial sample. In analyzing their sequence data, they find no evidence for recent introgression

of the maize Hopscotch locus into wild teosintes. They also find evidence for selection acting on tb1 in parviglumis. In greenhouse experiments, no difference in tiller index or tiller number was observed between teosinte lines with or without Hopscotch.

They present an interesting result, their conclusions are supported by their data, and the methods they use are appropriate. However, I feel like they need to make some further effort to understand what is happening mechanistically. What is Hopscotch doing in maize vs these sampled teosintes? I feel like the expression analyses and phenotyping experiments they suggest should be part of a later study should be part of this study. At a minimum, qRT-PCR experiments should be performed assessing tb1 expression in their with/without Hopscotch lines. Expression analyses of gt1, te1, and tru1 would be great, since they invoke variation at these loci to explain their results.

Extensive expression analysis of individuals with and without the *Hopscotch* insertion at *tb1* and a number of other genes is a good suggestion for follow-up work but would be a substantial undertaking and indeed more appropriate for another paper.

Another potential avenue, that the authors also suggest, lies in more directed and exhaustive greenhouse experiments. Variation in red and far-red light has been shown to be important in regulating both tb1 (in Sorghum) and gt1 in maize. Growth chamber experiments (with or without supplemental far red light, with or without Hopscotch), coupled to expression analyses in seedlings, seem like they should be possible. In the greenhouse experiments that were done, phenotyping was incomplete. Apart from tillering, Hopscotch and natural variation at the tb1 locus appears to affect ear architecture and internode length. Was there any evidence for variation in these other traits? In the longer term, it would be nice to see what happens when you introgress these particular parviglumis loci into maize.

We did not measure ear architecture or internode length. While *tb1* does impact these traits, our primary interest was in the effect of the *Hopscotch* insertion at *tb1* on tillering given that this is an ecologically relevant phenotype that is known to be substantially affected by *Hopscotch* in a maize background. Both additional greenhouse and expression analyses would indeed be useful avenues of further investigation. Nonetheless, these are extensive additional experiments which we are unable to include at this time. We have, however, added text to the conclusion (Page 15, lines 372-377) of this paper discussing their utility and the need for further work to elucidate the role of

tb1 in branching architecture in teosinte: "Future studies should examine expression levels of *tb1* in teosinte with and without the *Hopscotch* insertion and further characterize the effects of additional loci involved in branching architecture (e.g. *gt1*, *tru1*, and *te1*). These data, in conjunction with more exhaustive phenotyping, should help to further clarify the ecological significance of the domesticated *tb1* allele in natural populations of teosinte."

One minor issue: in Fig. 1, it would be nice to see where the Balsas River Basin is in relation to the sampling shown.

We have added the Balsas River in a revised version of Figure 1.

Response to Reviewer 2

Although this paper brings up an interesting question regarding the role of tb1 standing variation in teosinte, it is not able to answer it completely. It presents results that have been published previously about the presence of hopscotch in teosinte, maize, and mexicana.

Reviewer 2 is correct in stating that it was known that the Hopscotch was present in some teosinte individuals (Studer et al., 2011); however, previous studies sampled a small fraction of the populations and individuals per population that we present here and thus did not provide comprehensive estimates. Our study is not only concerned with the presence or absence of the *Hopscotch* allele in teosinte but also the frequency at which it segregates in a range-wide sample of teosinte. This finer level of detail sheds light on the evolutionary history of the *Hopscotch* insertion (*i.e.*, the domestication allele at *tb1*) in teosinte. We have added clarifying text to the introduction (Page 5, lines 55-61): "The effects of the *Hopscotch* insertion have been studied in maize (Studer et al., 2011), and analysis of teosinte alleles at *tb1* has identified functionally distinct allelic classes (Studer and Doebley, 2012), but little is known about the role of *tb1* or the *Hopscotch* insertion in natural populations of teosinte. Previous studies have confirmed the presence of the *Hopscotch* in samples of ssp. *parviglumis*, ssp. *mexicana*, and landrace maize; yet little is known about the frequency at which the *Hopscotch* is segregating in natural populations."

Nevertheless, the manuscript is worth publishing as it does present data that may help rule out some possible explanations as to why tb1 is so common in parviglumus.

Thank you; we hope that our paper in conjunction with future research and experiments can help shed light on the ecological significance of both *tb1* and the *Hopscotch* insertion in teosinte.

This paper is fairly maize-centric. A more general message is included in the discussion, not as much in the introduction, and not at all in the abstract. This could be easily remedied.

need to add something here. I had added something I thought in the text...

Methods and statistics were appropriately used as far as my expertise allows me to judge. Specific suggestions follow:

Introduction: Page 4, Lines 7 - 9: You say "Additionally, many domesticates show reduced genetic diversity when compared to their wild progenitors, and an understanding of the distribution of diversity in the wild and its phenotypic effects has become increasingly useful to crop improvement (Kovach and McCouch, 2008).". Can you tell us why an understanding is useful? Can you give some examples?

We have expanded on this concept in the introduction using *Oryza* as an example on page 4, lines 31-43: "Through the process of domestication and improvement, genetic diversity in many domesticates was reduced relative to their wild progenitors (Kovach and McCouch, 2008). Many of these wild alleles no longer found in their domestic counterparts often harbor beneficial traits such as disease resistance or drought tolerance, and an understanding of the distribution of this diversity in the wild and its phenotypic effects has become increasingly useful to crop improvement (Kovach and McCouch, 2008). For example, *Oryza rufipogon*, the wild progenitor of domesticated rice, has proven useful for the integration of a number of beneficial QTL controlling traits such as grain size and yield into domesticated rice (**?**). Furthermore, in teosinte, many of the alleles underlying domestication phenotypes have been shown to be segregating in populations of teosinte, suggesting that natural variation in teosinte has played an important role in the domestication of maize (Gallavotti et al., 2004; **?**). While some effort has been invested into understanding how wild alleles behave in their domesticated relatives (e.g. Bai and Lindhout, 2007), very little is known

about the role that alleles found most commonly in domesticates play in natural populations of their wild progenitors (Whitton et al., 1997)."

8) Page 5, line 12 - 14: You say "The tb1 locus appears to play an important role in the shade avoidance pathway in Zea mays and other grasses and may therefore be crucial to the ecology of teosinte (Kebrom and Brutnell, 2007; Lukens and Doebley, 1999)." Please clarify; do the references belong to the first half of the sentence, supporting tb1 playing a role in Zea mays, or the second half of the sentence, supporting tb1 possibly being crucial to the ecology of teosinte? Which of the two distinct things have actually been studied? If the first, could you tell us how tb1 has been shown to work in maize under shade conditions? If the second, was it actually studied, or merely hypothesized in teosinte?

We have moved the references cited here to more clearly reflect the papers' findings. We have also added text describing the role of *tb1* in the shade avoidance pathway and clarifying the manner in which this could affect fitness in teosinte. Changes can be seen on page 5, lines 71-77: "The *tb1* locus appears to play an important role in the shade avoidance pathway in *Zea mays* (Lukens and Doebley, 1999) and other grasses (Kebrom and Brutnell, 2007), exhibiting changes in expression levels in response to shading. Lukens and Doebley (1999) introgressed the teosinte *tb1* allele into a maize inbred background and noted that under low density conditions plants were highly tillered, whereas under high density, plants showed significantly reduced tillers and grew taller (Lukens and Doebley, 1999). Based on these results we hypothesized that *tb1* may play a role in the ecology of teosinte."

9) Page 5, Lines 15 - 16: you say you "aim to characterize the distribution of the Hopscotch insertion in parviglumis, mexicana, and landrace maize, and to examine the phenotypic effects of the insertion in parviglumis." The mexicana seem a bit tacked on and not much discussed. Did you phenotype mexicana? Can you present the data here, if so? Does the analysis of this group of teosintes add anything to your conclusions?

Our goal in screening for the *Hopscotch* insertion was to assess its taxonomic and geographic distribution to the extent possible given the samples to which we had access. Reviewer 2 is correct in noting that ssp. *mexicana* is not as prevalent in our paper as ssp. *parviglumis*. Though our

sample of *mexicana* is range-wide, we only had access to seeds from 17 populations. Frequency of the *Hopscotch* insertion was low in these populations relative to *parviglumis*, an interesting point that we have discussed further on page X
, lines XX
of the text. The low frequency of *Hopscotch* in *mexicana* precluded a phenotyping experiment with this subspecies since a very high number of seeds would have to be germinated in order to obtain enough individuals with the *Hopscotch* allele.

10) Materials and Methods: Page 5, lines 22 - 23: how many individuals per accession? A fixed number, or variable, and depending on what? This can be mentioned so the reader does not have to go to the appendix or supplements.

We have added a range of values (1-38 individuals) per population in the methods, pages 5-6, lines 85-90: "We sampled 1,110 individuals from 350 accessions (247 maize landraces, 17 *mexicana* populations, and 86 *parviglumis* populations; ranging from 1-38 individuals per population, with an average of 11 individuals per population for *parviglumis* and *mexicana* and 2 individuals per landrace accession) and assessed the presence or absence of the *Hopscotch* insertion (Appendix 1 and Appendix 2, See Supplemental Materials with the online version of this article)."

11) Page 5, Line 28: mention how big the hopscotch element is. Can it be amplified in one amplicon? Again, this would be useful to know without having to look in appendix or supplements.

The sizes of amplicons are now mentioned explicitly in the Materials and Methods (page 6, lines 99-104) following primer sequences: "Homozygotes show a single band for absence of the *Hopscotch* element (∼300bp) and two bands for presence of the element (∼5kb, amplification of the entire element, and ∼1.1kb, amplification of part of the element), whereas heterozygotes are three-banded (Appendix 2, See Supplemental Materials with the online version of this article)." The 5kb amplicon including the entire *Hopscotch* could be amplified in a single reaction.

12) Page 6 line 6: "When only one PCR resolved well, we scored one allele for the individual". Sorry, what does this mean? That you have one band (one assumes the short one?). If you are talking presence/absence, how do you know the allele?

We developed primers for two PCR reactions that would amplify products with size polymorphism characteristic of the two different alleles (*i.e.*, *Hopscotch* vs. no *Hopscotch*). If PCR products were clearly resolved for only one of these reactions, we included one allele for that individual in our data set rather than infer the diploid genotype. We have clarified this in the text on page 6, lines 92-104. : "Two PCRs were performed for each individual, one with primers flanking the *Hopscotch* (HopF/HopR) and one with a flanking primer and an internal primer (HopF/HopIntR). Primer sequences are HopF, 5'-TCGTTGATGCTTTGATGGATGG-3'; HopR, 5'-AACAGTATGATTTCATGGGACCG-3'; and HopIntR, 5'-CCTCCACCTCTCATGAGATCC-3' (Appendix 3 and Appendix 4, See Supplemental Materials with the online version of this article). Homozygotes show a single band for absence of the element (∼300bp) and two bands for presence of the element (∼5kb, amplification of the entire element, and ∼1.1kb, amplification of part of the element), whereas heterozygotes show all three bands (Appendix 2, See Supplemental Materials with the online version of this article). Since we developed a PCR protocol for each allele, if only one PCR resolved well, we scored one allele for that individual rather than infer the diploid genotype."

13) Page 7 lines 14 - 15 "These analyses only included populations in which eight or more individuals were sampled." How many populations did that include?

Thirty-two populations were included in these analyses, and we have edited the text to reflect this number on page 7, lines 143-144: "These analyses only included populations ($n = 32$) in which eight or more individuals were sampled."

14) Page 7 lines 17 - 26: how many year's worth of environmental data went into your analysis? Can you hypothesize that tb1 differences can be correlated with current climactic data if the changes and subsequent selections must have happened many generations ago?

Our climatic data consisted of the 19 bioclim layers commonly used in ecological niche analyses. Full methods for generation of these data can be found at the following url: `http://www.worldclim.org/methods`. Briefly, temperature and precipitation were averaged for each global location over a minimum of ten years during the period from 1950-2000. Soil data were downloaded from

the Harmonized World Soil Database (FAOHWSD; http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HWSD_Documentation.pdf). This recently published data set drew together global soil maps generated by a world-wide network of scientists over the last several decades and summarized them in GIS data layers that are available as polygon coverages. Data for our sampling sites were extracted from these data sets based on latitude and longitude using the DIVA-GIS platform. Since data for our analyses were either summarized over the last several decades (climatic data) or are likely to be stable over time (soil data), they have potentially influenced current frequencies of the *Hopscotch* insertion in our sampled populations. We have further clarified this on page 8, lines 154-157 of the text : "Environmental data were obtained from www.worldclim.org and soil data were downloaded from the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) at www.harvestchoice.org. These data represent average values for the last several decades (climatic data) or are likely stable over time (soil data) and therefore represent conditions important for local adaptation of our samples. Information from these data sets was summarized by principle component analysis following Pyhäjärvi et al. (2013)."

15) Page 7 lines 22 - 23: "We used genotyping and covariance data from Pyhajarvi et al. (2013) for BayEnv" what kind of dataset is this? Teosinte?

We have clarified the dataset in the paper on page 8, lines 151-154: "We used teosinte (ssp. *parviglumis* and ssp. *mexicana*) SNP genotyping and covariance data from Pyhäjärvi et al. (2013) for BayEnv, with the *Hopscotch* insertion coded as an additional SNP. SNP data from Pyhäjärvi et al. (2013) were obtained using the MaizeSNP50 BeadChip and Infinium HD Assay (Illumina, San Diego, CA, USA)."

16) Page 8 lines 20 - 21:" We created haplotype blocks using a custom Perl script that grouped SNPs separated by less than 5kb into haplotypes." Does LD data in your parviglumus support 5Kb?

While the recombination landscape will vary to some extent across individuals of teosinte and maize, broad-scale patterns are conserved (Bauer et al., 2013). We therefore relied on previous estimates of the decay of linkage disequilibrium based on data from HapMap2 of maize that suggest LD decays over an average distance of 5500bp (**?**); because LD decay is even more rapid in teosinte (**?**) we used a conservative haplotype block size of 5kb as an estimate of the extent

of linkage disequilibrium in our sample and the physical distance over which SNPs would have to be separated in order to be in weak linkage and not violate assumptions of the STRUCTURE model.

17) Page 9 lines 2 - 3: "Plants were watered three times a day by hand and with an automatic drip containing 10-20-10 fertilizer." This sounds odd; by hand or automatically? Or you turned on the drip by hand? In which case, it was not automatic

We have edited the text to clarify, page 9, lines 199-200 : "Plants were watered three times a day with an automatic drip containing 10-20-10 fertilizer, which was supplemented with hand watering on extremely hot and dry days."

18) Page 9 lines 7 - 8: "Culm diameter is not believed to be correlated with tillering index or variation at tb1." So, why did you do it? Because of the last sentence in the materials and methods? Perhaps do not mention here that is not correlated; wait until the end of the materials and methods where you seem to explain it.

Thank you for the suggestion, we have edited the text accordingly and have added our explanation of the inclusion of culm diameter at the end of the Materials and Methods section: "Additionally, in order to compare any association between *Hopscotch* genotype and tillering and associations at other presumably unrelated traits, we performed an ANOVA between culm diameter and genotype using the same general linear model in SAS. Culm diameter is not believed to be correlated with tillering index or variation at *tb1* and is used as our independent trait for phenotyping analyses."

19) Page 9 lines 9 - 10: "We genotyped individuals for the Hopscotch insertion following the protocols listed above." The PCR protocols or the sequencing protocols?

We have clarified in the manuscript that genotypes were obtained using our PCR protocol.

20) Results: Page 10 lines 5 - 6: "Within our parviglumis and mexicana samples we found the Hopscotch insertion segregating in 37 and four populations, respectively"; remind us how many parviglumis and Mexicana samples you had total? So we can see if this comes out to a very different ratio than in maize."

We have added ratios to the text: "Within our *parviglumis* and *mexicana* samples we found the *Hopscotch* insertion segregating in 37 ($n = 86$) and four ($n = 17$) populations, respectively, and at highest frequency in the states of Jalisco, Colima, and Michoacán in central-western Mexico."

21) Page 10 lines 7 - 9: "Using our Hopscotch genotyping, we calculated differentiation between populations (FST) and subspecies (FCT) for populations in which we sampled eight or more alleles." What do you mean, 8 or more alleles? I thought you only had presence or absence; how are alleles differentiated with your PCR?

Please see our response to point 12. We developed a co-dominant assay that was capable of detecting whether an individual was homozygous or heterozygous for the *Hopscotch* insertion or lacked the insertion entirely. We have also edited this to read "eight or more chromosomes", which is a more precise description of our methodology.

22) Page 13 lines 25 - 27: "and genetic data support ongoing gene flow between domesticated maize and both mexicana and parviglumis in a number of sympatric populations (Hufford et al., 2013; Ellstrand et al., 2007; van Heerwaarden et al., 2011). Add Warburton, et al., 2011. Gene flow among different teosinte taxa and into the domesticated maize gene pool. Genet Resour Crop Evol 58:1243-1261.

We have added this reference as suggested.

23) Page 14 line 15 to page 15 line 14: too long a paragraph. Break into two, perhaps at page 14 line 32.

We have split the paragraph into two.

References

Bai, Y. and P. Lindhout. 2007. Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Annals of Botany* 100: 1085–1094.

Bauer, E., M. Falque, H. Walter, C. Bauland, C. Camisan, L. Campo, N. Meyer, N. Ranc, R. Rincent, W. Schipprack, T. Altmann, P. Flament, A. E. Melchinger, M. Menz, J. Moreno-Gonzalez, M. Ouzunova, P. Revilla, A. Charcosset, O. C. Martin, and C.-C. Schoen. 2013. Intraspecific variation of recombination rate in maize. *GENOME BIOLOGY* 14.

FAO/IIASA/ISRIC/ISSCAS/JRC. 2012. Harmonized World Soil Database, version 1.2. FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Gallavotti, A., Q. Zhao, J. Kyozuka, R. Meeley, M. Ritter, J. Doebley, M. Pe, and R. Schmidt. 2004. The role of barren stalk1 in the architecture of maize. *Nature* 432: 630–635.

Kebrom, T. and T. Brutnell. 2007. The molecular analysis of the shade avoidance syndrome in the grasses has begun. *Journal of Experimental Botany* 58: 3079–3089.

Kovach, M. and S. McCouch. 2008. Leveraging natural diversity: back through the bottleneck. *Genome studies and Molecular Genetics* 11: 193–200.

Lukens, L. and J. Doebley. 1999. Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. *Genet Res* 74: 291–302.

Pyhäjärvi, T., M. Hufford, and J. Ross-Ibarra. 2013. Complex patterns of local adaptation in the wild relatives of maize. *Genome Biology and Evolution* 5: 1594–1609.

Remington, D. L., J. M. Thornsberry, Y. Matsuoka, L. M. Wilson, S. R. Whitt, J. Doebley, S. Kresovich, M. M. Goodman, and E. S. Buckler. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences* 98: 11479–11484.

Sigmon, B. and E. Vollbrecht. 2010. Evidence of selection at the ramosa1 locus during maize domestication. *Mol Ecol* 19: 1296–1311.

Studer, A. and J. Doebley. 2012. Evidence for a natural allelic series at the maize domestication locus teosinte branched1. *Genetics* 19: 951–958.

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–U164.

Whitton, J., D. Wolf, D. Arias, A. Snow, and L. Riesberg. 1997. The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. *Theoretical and Applied Genetcs* 95: 33–40.