

Vann *et al.*: Variation in teosinte at the *tb1* locus

Natural variation in teosinte at the domestication locus *teosinte branched1* (*tb1*)¹

Laura Vann², Thomas Kono^{2,3}, Tanja Pyhäjärvi^{2,4}, Matthew B. Hufford^{2,5,7}, and Jeffrey
Ross-Ibarra^{2,6,7}

²Department of Plant Sciences, University of California Davis

³Department of Agronomy and Plant Genetics, University of Minnesota Twin Cities

⁴Department of Biology, University of Oulu

⁵Department of Ecology, Evolution, and Organismal Biology, Iowa State University

⁶Center for Population Biology and Genome Center, University of California Davis

⁷Author for Correspondence

Manuscript received _____; revision accepted _____.

Acknowledgements

We would like to thank the Department of Plant Sciences at UC Davis for Graduate Student Research funding to LEV, *Matt UC Mexus* , and G. Coop for helpful discussion.

Grants as far as I know were Matt's UC Mexus money, whatever money paid for the 55K data since we used that, and I am not sure what paid for greenhouse space 55K data was public, no prob. acknowledge support from the Dept. of Plant Sciences for LV's GSR and for "research funds" or something vague, as this came from startup. acknowledge graham coop for helpful discussion

1 Abstract

2 *Premise of the study:* The *teosinte branched1* (*tb1*) gene is a major QTL controlling
3 branching differences between maize and its wild progenitor, teosinte. Previous work has
4 shown that the insertion of a transposable element (*Hopscotch*) upstream of *tb1* enhances
5 the gene's expression, causing much of the reduction in tillering observed in maize.
6 Observations of the maize *tb1* allele in teosinte and estimates of an age of insertion of the
7 *Hopscotch* element that predates domestication led us to investigate its prevalence and
8 potential role in teosinte.

9 *Methods:* Prevalence of the *Hopscotch* element was assessed across an Americas-wide
10 sample of 1110 maize and teosinte individuals using a co-dominant PCR assay.
11 Population genetic summaries were calculated for a subset of individuals from four
12 teosinte populations in central Mexico. Phenotypic data were also collected from a single
13 teosinte population where *Hopscotch* was found segregating.

14 *Key results:* Genotyping results suggest the *Hopscotch* element is at higher than expected
15 frequency in teosinte. Analysis of linkage disequilibrium near *tb1* does not support recent
16 introgression of the *Hopscotch* allele from maize into teosinte. Population genetic
17 signatures are consistent with selection on this locus revealing a potential ecological role
18 for *Hopscotch* in teosinte. Finally, two greenhouse experiments with teosinte do not
19 suggest *tb1* controls tillering in natural populations.

20 *Conclusions:* Our findings suggest the role of *Hopscotch* differs between maize and
21 teosinte. Future work should assess *tb1* expression levels in teosinte with and without the
22 *Hopscotch* and more comprehensively phenotype teosinte to assess the ecological
23 significance of the *Hopscotch* insertion and, more broadly, the *tb1* locus in teosinte.

24 **Key words:** domestication; maize; teosinte; *teosinte branched1*; transposable element

INTRODUCTION

Domesticated crops and their wild progenitors provide an excellent system in which to study adaptation and genomic changes associated with human-mediated selection (Ross-Ibarra et al., 2007). Perhaps the central focus of the study of domestication has been the identification of genetic variation underlying agronomically important traits such as fruit size and plant architecture (Olsen and Gross, 2010). 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). But 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).

Maize (*Zea mays* ssp. *mays*) was domesticated from the teosinte *Zea mays* ssp. *parviglumis* (hereafter, *parviglumis*) roughly 9,000 B.P. in southwest Mexico (Piperno et al., 2009; Matsuoka et al., 2002). Domesticated maize and the teosintes are an attractive system in which to study domestication due to the abundance of genetic tools developed for maize and well-characterized domestication loci (Hufford et al., 2012a; Doebley, 2004; Hufford et al., 2012b). Additionally, large naturally occurring populations of both *Zea mays* ssp. *parviglumis* (the wild progenitor of maize) and *Zea mays* ssp. *mexicana* (highland teosinte; hereafter *mexicana*) can be found throughout Mexico (Wilkes, 1977; Hufford et al., 2013), and genetic diversity of these taxa is estimated to be high (Ross-Ibarra et al., 2009).

Many morphological changes are associated with maize domestication, and understanding the genetic basis of these changes has been a focus of maize research for a number of years (Doebley, 2004). One of the most dramatic changes is found in plant architecture: domesticated maize is characterized by a central stalk with few tillers and

1 lateral branches terminating in a female inflorescence, while teosinte is highly tillered and
2 bears tassels (male inflorescences) at the end of its lateral branches. The *teosinte*
3 *branched1* (*tb1*) gene, a repressor of organ growth, was identified as a major QTL
4 involved in branching (Doebley et al., 1995) and tillering (Doebley and Stec, 1991)
5 differences between maize and teosinte. A 4.9 kb retrotransposon (*Hopscotch*) insertion
6 into the upstream control region of *tb1* in maize acts to enhance expression of *tb1*, thus
7 repressing lateral organ growth (Doebley et al., 1997; Studer et al., 2011). Dating of the
8 *Hopscotch* retrotransposon suggests that its insertion predates the domestication of
9 maize, leading to the hypothesis that it was segregating as standing variation in ancient
10 populations of teosinte and increased to high frequency in maize due to selection during
11 domestication (Studer et al., 2011). The effects of the *Hopscotch* insertion have been
12 studied in maize (Studer et al., 2011), and analysis of teosinte alleles at *tb1* has identified
13 functionally distinct allelic classes (Studer and Doebley, 2012), but little is known about
14 the role of *tb1* or the *Hopscotch* insertion in natural populations of teosinte.

15 In teosinte and other plants that grow at high population density, individuals detect
16 competition from neighbors via the ratio of red to far-red light. An increase in far-red
17 relative to red light accompanies shading and triggers the shade avoidance syndrome: a
18 suite of physiological and morphological changes such as reduced tillering, increased plant
19 height and early flowering (Kebrom and Brutnell, 2007). The *tb1* locus appears to play
20 an important role in the shade avoidance pathway in *Zea mays* and other grasses and
21 may therefore be crucial to the ecology of teosinte (Kebrom and Brutnell, 2007; Lukens
22 and Doebley, 1999). In this study we aim to characterize the distribution of the
23 *Hopscotch* insertion in *parviglumis*, *mexicana*, and landrace maize, and to examine the
24 phenotypic effects of the insertion in *parviglumis*. We use a combination of PCR
25 genotyping for the *Hopscotch* element in our full panel and sequencing of two small
26 regions upstream of *tb1* in a subset of teosinte populations to explore patterns of genetic
27 variation at this locus. Finally, we test for an association between the *Hopscotch* element

1 and tillering phenotypes in a population of *parviglumis*.

2 MATERIALS AND METHODS

3 **Sampling and genotyping**—We sampled 1,110 individuals from 350 accessions
4 (247 maize landraces, 17 *mexicana* populations, and 86 *parviglumis* populations) and
5 assessed the presence or absence of the *Hopscotch* insertion (Table S1 and Table S2).
6 DNA was extracted from leaf tissue using a modified CTAB approach (Doyle and Doyle,
7 1990; Maloof et al., 1984). We designed primers using PRIMER3 (Rozen and Skaletsky,
8 2000) implemented in Geneious (Kearse et al., 2012) to amplify the entire *Hopscotch*
9 element, as well as an internal primer allowing us to simultaneously check for possible
10 PCR bias between presence and absence of the *Hopscotch* insertion. Two PCRs were
11 performed for each individual, one with primers flanking the *Hopscotch* (HopF/HopR)
12 and one with a flanking primer and an internal primer (HopF/HopIntR). Primer
13 sequences are HopF, 5'-TCGTTGATGCTTTGATGGATGG-3'; HopR,
14 5'-AACAGTATGATTTTCATGGGACCG-3'; and HopIntR,
15 5'-CCTCCACCTCTCATGAGATCC-3' (Fig. S1, Fig. S2) *Primers in Fig. S1 should be labeled* .
16 Homozygotes show a single band for absence of the element (~300bp) and two bands for
17 presence of the element (~5kb and ~1.1kb), whereas heterozygotes are three-banded
18 (Fig. S2). When only one PCR resolved well, we scored one allele for the individual. We
19 used Phusion High Fidelity Enzyme (Thermo Fisher Scientific Inc., Waltham,
20 Massachusetts, USA) and the following conditions for amplifications: 98°C for 3 min, 30
21 cycles of 98°C for 15 s, 65°C for 30 s, and 72°C for 3 min 30 s, with a final extension of
22 72°C for 10 min. PCR products were visualized on a 1% agarose gel and scored for
23 presence/absence of the *Hopscotch* based on band size.

24 **Sequencing**—In addition to genotyping, we chose a subset of *parviglumis*
25 individuals for sequencing. We chose twelve individuals from each of four populations
26 from Jalisco state, Mexico (San Lorenzo, La Mesa, Ejutla A, and Ejutla B). For

1 amplification and sequencing, we selected two regions approximately 600bp in size from
2 within the 5' UTR of *tb1* (Region 1) and from 1,235bp upstream of the start of the
3 *Hopscotch* (66,169bp upstream from the start of the *tb1* ORF; Region 2). We designed
4 the following primers using PRIMER3 (Rozen and Skaletsky, 2000): for the 5' UTR,
5 5'-GGATAATGTGCACCAGGTGT-3' and 5'-GCGTGCTAGAGACACYTGTTGCT-3'; for the
6 66kb upstream region, 5'-TGTCCTCGCCGCAACTC-3' and
7 5'-TGTACGCCCCGCCCTCATCA-3' (Fig. S1). We used Taq polymerase (New England
8 Biolabs Inc., Ipswich, Massachusetts, USA) and the following thermal cycler conditions to
9 amplify fragments: 94°C for 3 min, 30 cycles of 92°C for 40 s, annealing for 1 min, 72°C
10 for 40 s, and a final 10 min extension at 72°C. Annealing temperatures for Region 1 and
11 Region 2 were 59.7°C and 58.8°C, respectively. To clean excess primer and dNTPs we
12 added two units of Exonuclease1 and 2.5 units of Antarctic Phosphatase to 8.0 μ L of
13 amplification product. This mix was placed on a thermal cycler with the following
14 program: 37°C for 30 min, 80°C for 15 min, and a final cool-down step to 4°C.

15 We cloned cleaned fragments into a TOPO-TA vector (Life Technologies, Grand
16 Island, New York, USA) using OneShot TOP10 chemically competent *E. coli* cells, with
17 an extended ligation time of 30 min for a complex target fragment. We plated cells on LB
18 agar plates containing kanamycin, and screened colonies using vector primers M13
19 Forward and M13 Reverse under the following conditions: 96°C for 5 min; then 35 cycles
20 at 96°C for 30 s, 53°C for 30 s, 72°C for 2 min; and a final extension at 72°C for 4 min.
21 We visualized amplification products for incorporation of our insert on a 1% agarose TAE
22 gel.

23 Amplification products with successful incorporation of our insert were cleaned using
24 Exonuclease 1 and Antarctic Phosphatase following the procedures detailed above, and
25 sequenced with vector primers M13 Forward and M13 Reverse using Sanger sequencing at
26 the College of Agriculture and Environmental Sciences (CAES) sequencing center at UC
27 Davis. We aligned and trimmed primer sequences from resulting sequences using the

1 software Geneious (Kearse et al., 2012). Following alignment, we verified singleton SNPs
2 by sequencing an additional one to four colonies from each clone. If the singleton was not
3 present in these additional sequences it was considered an amplification or cloning error,
4 and we replaced the base with the base of the additional sequences. If the singleton
5 appeared in at least one of the additional sequences we considered it a real variant and
6 kept it for further analyses.

7 **Genotyping analysis**—We examined discrepancies between observed and expected
8 genotype frequencies by calculating Hardy-Weinberg Equilibrium (HWE). To calculate
9 differentiation between populations (F_{ST}) and subspecies (F_{CT}) we used HierFstat
10 (Goudet, 2005). These analyses only included populations in which 8 or more individuals
11 were sampled. To test the hypothesis that the *Hopscotch* insertion may be adaptive
12 under certain environmental conditions, we looked for significant associations between the
13 *Hopscotch* frequency and environmental variables using BayEnv (Coop et al., 2010).
14 BayEnv creates a covariance matrix of relatedness between populations and then tests a
15 null model that allele frequencies in populations are determined by the covariance matrix
16 of relatedness alone against the alternative model that allele frequencies are determined
17 by a combination of the covariance matrix and an environmental variable, producing a
18 posterior probability (*i.e.*, Bayes Factor; Coop et al. 2010). We used genotyping and
19 covariance data from Pyhäjärvi et al. (2013) for BayEnv, with the *Hopscotch* insertion
20 coded as an additional SNP (Table S3). Environmental data were obtained from
21 www.worldclim.org, the Harmonized World Soil Database
22 (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and www.harvestchoice.org and summarized
23 by principle component analysis following Pyhäjärvi et al. (2013).

24 **Sequence analysis**—For population genetic analyses of sequenced Region 1 and
25 sequenced Region 2 we used the Libsequence package (Thornton, 2003) to calculate
26 pairwise F_{ST} between populations and to calculate standard diversity statistics (number
27 of haplotypes, haplotype diversity, Watterson’s estimator $\hat{\theta}_W$, pairwise nucleotide

1 diversity $\hat{\theta}_\pi$, and Tajima's D). To produce a visual representation of differentiation
2 between sequences and examine patterns in sequence clustering by *Hopscotch* genotype
3 we used Phylip (<http://evolution.genetics.washington.edu/phylip.html>), creating
4 neighbor-joining trees with bootstrap-supported nodes (100 repetitions). For creation of
5 trees we also included homologous sequence data from Maize HapMapV2 (Chia et al.,
6 2012) for teosinte inbred lines (TILs), some of which are known to be homozygous for the
7 *Hopscotch* insertion (TIL03, TIL17, TIL09), as well as 59 lines of domesticated maize.

8 **Introgression analysis**—In order to assess patterns of linkage disequilibrium (LD)
9 around the *Hopscotch* element in the context of chromosomal patterns of LD we used
10 Tassel (Bradbury et al., 2007) and calculated LD between SNPs across chromosome 1
11 using previously published data from twelve plants each of the Ejutla A (EjuA), Ejutla B
12 (EjuB), San Lorenzo (SLO), and La Mesa (MSA) populations (Pyhäjärvi et al., 2013).
13 We chose these populations because we had both genotyping data for the *Hopscotch* as
14 well as chromosome-wide SNP data for chromosome 1. For each population we filtered
15 the initial set of 5,897 SNPs on chromosome 1 to accept only SNPs with a minor allele
16 frequency of at least 0.1, resulting in 1,671, 3,023, 3,122, and 2,167 SNPs for SLO, EjuB,
17 EjuA, and MSA, respectively. We then used Tassel (Bradbury et al., 2007) to calculate
18 linkage disequilibrium (r^2) across chromosome 1 for each population.

19 We examined evidence of introgression on chromosome 1 in these same four
20 populations (EjuA, EjuB, MSA, SLO) using STRUCTURE (Falush et al., 2003) and
21 phased data from Pyhäjärvi et al. (2013), combined with the corresponding SNP data
22 from a diverse panel of 282 maize lines (Cook et al., 2012). SNPs were anchored in a
23 modified version of the IBM genetic map (Gerke et al., 2013). We created haplotype
24 blocks using a custom Perl script that grouped SNPs separated by less than 5kb into
25 haplotypes. We ran STRUCTURE at K=2 under the linkage model, performing 3
26 replicates with an MCMC burn-in of 10,000 steps and 50,000 steps post burn-in.

27 **Phenotyping of *parviglumis***—To investigate the phenotypic effects of the

1 *Hopscotch* insertion in teosinte, we conducted an initial phenotyping trial (Phenotyping
2 1). We germinated 250 seeds of *parviglumis* collected in Jalisco state, Mexico (population
3 San Lorenzo) (Hufford, 2010) where the *Hopscotch* is segregating at highest frequency
4 (0.44) in our initial genotyping sample set. In order to maximize the likelihood of finding
5 the *Hopscotch* in our association population we selected seeds from sites where genotyped
6 individuals were homozygous or heterozygous for the insertion. We chose between 10-13
7 seeds from each of 23 sampling sites. We treated seeds with Captan fungicide (Southern
8 Agricultural Insecticides Inc., Palmetto, Florida, USA) and germinated them in petri
9 dishes with filter paper. Following germination, 206 successful germinations were planted
10 into one-gallon pots with potting soil and randomly spaced one foot apart on greenhouse
11 benches. Plants were watered three times a day by hand and with an automatic drip
12 containing 10-20-10 fertilizer.

13 Starting on day 15, we measured tillering index as the ratio of the sum of tiller
14 lengths to the height of the plant (Briggs et al., 2007). Following initial measurements,
15 we phenotyped plants for tillering index every 5 days through day 40, and then on day 50
16 and day 60. On day 65 we measured culm diameter between the third and fourth nodes
17 of each plant. Culm diameter is not believed to be correlated with tillering index or
18 variation at *tb1*. Following phenotyping we extracted DNA from all plants using a
19 modified SDS extraction protocol. We genotyped individuals for the *Hopscotch* insertion
20 following the protocols listed above. Based on these initial data, we conducted a *post hoc*
21 power analysis using data from day 40 of Phenotyping 1, indicating that a minimum of 71
22 individuals in each genotypic class would be needed to detect the observed effect of the
23 *Hopscotch* on tillering index.

24 We performed a second phenotyping experiment (Phenotyping 2) in which we
25 germinated 372 seeds of *parviglumis*, choosing equally between sites previously
26 determined to have or not have the *Hopscotch* insertion. Seeds were germinated and
27 planted on day 7 post fruit-case removal into two gallon pots. Plants were watered twice

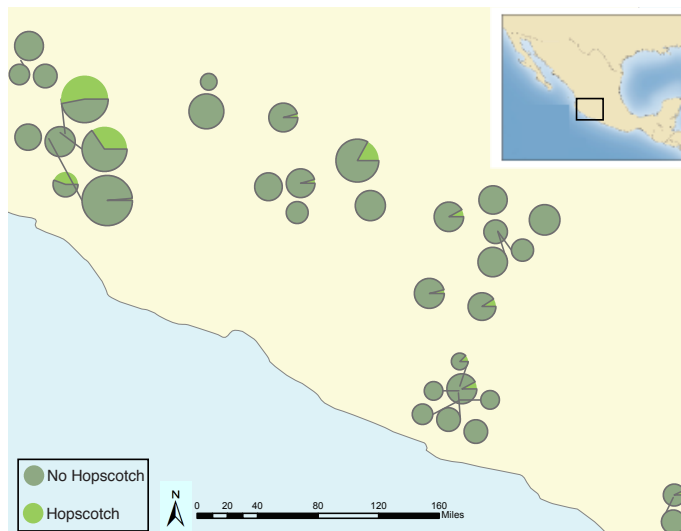
1 daily, alternating between fertilized and non-fertilized water. We began phenotyping
2 successful germinations (302 plants) for tillering index on day 15 post fruit-case removal,
3 and phenotyped every five days until day 50. At day 50 we measured culm diameter
4 between the third and fourth nodes. We extracted DNA and genotyped plants following
5 the same guidelines as in Phenotyping 1.

6 Tillering index data for each genotypic class did not meet the criteria for a repeated
7 measures ANOVA, so we transformed the data using a Box-Cox transformation ($\lambda = 0$)
8 Car Package for R, Fox and Weisberg 2011) to improve the normality and homogeneity of
9 variance among genotype classes. We analyzed relationships between genotype and
10 tillering index and tiller number using a repeated measures ANOVA through a general
11 linear model function implemented in SAS v.9.3 (SAS Institute Inc., Cary, NC, USA).
12 Additionally, in order to compare any association between *Hopscotch* genotype and
13 tillering and associations at other presumably unrelated traits, we performed an ANOVA
14 between culm diameter and genotype using the same general linear model in SAS.

15 RESULTS

16 **Genotyping**—Genotype of the *Hopscotch* insertion was confirmed with two PCRs
17 for 837 individuals. Among the 247 maize landrace accessions genotyped, all but eight
18 were homozygous for the presence of the insertion (Table S1 and Table S2). Within our
19 *parviglumis* and *mexicana* samples we found the *Hopscotch* insertion segregating in 37
20 and 4 populations, respectively, and at highest frequency in the states of Jalisco, Colima,
21 and Michoacán in central-western Mexico (Fig. 1). Using our *Hopscotch* genotyping, we
22 calculated differentiation between populations (F_{ST}) and subspecies (F_{CT}) for
23 populations in which we sampled 8 or more alleles. We found that $F_{CT} = 0$, and levels of
24 F_{ST} among populations within each subspecies (0.22) and among all populations (0.23)
25 are similar to those reported genome-wide in previous studies (Pyhäjärvi et al. 2013;
26 Table 1). Although we found large variation in *Hopscotch* allele frequency among our

Figure 1: Map showing the frequency of the *Hopscotch* allele in populations of *parviglumis* where we sampled more than 6 individuals. Size of circles reflects number of alleles sampled.



- 1 populations, BayEnv analysis did not indicate a correlation between the *Hopscotch*
- 2 insertion and environmental variables (all Bayes Factors < 1; Table S3).

Table 1: Pairwise F_{ST} values from sequence and *Hopscotch* genotyping data

Comparison	Region 1	Region 2	<i>Hopscotch</i>
EjuA & EjuB	0	0	0
EjuA & MSA	0.326	0.328	0.186
EjuA & SLO	0.416	0.258	0.280
EjuB & MSA	0.397	0.365	0.188
EjuB & SLO	0.512	0.290	0.280
MSA & SLO	0.007	0	0.016

- 3 **Sequencing**—To investigate patterns of sequence diversity and linkage
- 4 disequilibrium (LD) in the *tb1* region, we sequenced two small (<1kb) regions upstream
- 5 of the *tb1* ORF in four populations. After alignment and singleton checking we recovered

1 48 and 40 segregating sites for the 5' UTR region (Region 1) and the 66kb upstream
 2 region (Region 2), respectively. For Region 1, Ejutla A has the highest values of
 3 haplotype diversity, and θ_π , while Ejutla B and La Mesa have comparable values of these
 4 summary statistics, and San Lorenzo has much lower values. Additionally, Tajima's D is
 5 strongly negative in the two Ejutla populations and La Mesa, but is less negative in San
 6 Lorenzo (Table 2, Table S2). For Region 2, haplotype diversity and θ_π , are similar for
 7 Ejutla A and Ejutla B, while La Mesa and San Lorenzo have slightly lower values for
 8 these statistics (Table 2). *Do you want me to edit table here to have Hop freq as a column for the 4 pops?.*
 9 *I didn't know if it would be confusing because the frequencing is not representative of the frequencing of Hop in the*
 10 *sequenced sample set* Tajima's D is positive in all populations except La Mesa, indicating an
 11 excess of low frequency variants in this population (Table 2). Pairwise values of F_{ST}
 12 within population pairs Ejutla A/Ejutla B and San Lorenzo/La Mesa are 0 for both
 13 sequenced regions as well as for the *Hopscotch* *table 1 shows 0.016 for hopscotch, not 0. which is*
 14 *right?* , while they are high for other population pairs (Table 1). Neighbor joining trees of
 15 our sequence data and data from the teosinte inbred lines (TILs; data from Maize
 16 HapMapV2, Chia et al. 2012) do not reveal any clear clustering pattern with respect to
 17 population or *Hopscotch* genotype (Figure S3); individuals within our sample that have
 18 the *Hopscotch* insertion do not group with the teosinte inbred lines or domesticated maize
 19 that have the *Hopscotch* insertion.

20 **Evidence of introgression**—The highest frequency of the *Hopscotch* insertion in
 21 teosinte was found in *parviglumis* sympatric with cultivated maize. Our initial hypothesis
 22 was that the high frequency of the *Hopscotch* element in these populations could be
 23 attributed to introgression from maize into teosinte. To investigate this possibility we
 24 examined overall patterns of linkage disequilibrium across chromosome one and
 25 specifically in the *tb1* region. If the *Hopscotch* is found in these populations due to recent
 26 introgression we would expect to find large blocks of linked markers near this element.
 27 We find no evidence of elevated linkage disequilibrium between the *Hopscotch* and SNPs

Table 2: Population genetic statistics from resequenced regions near the *tb1* locus

Population	# Haplotypes	Hap. Diversity	$\hat{\theta}_\pi$	Tajima's D
<i>Region 1(5' UTR)</i>				
EJUA	8	0.859	0.005	-1.650
EJUB	5	0.709	0.004	-1.831
MSA	6	0.682	0.004	-1.755
SLO	3	0.318	0.001	-0.729
<i>Region 2 (66kb upstream)</i>				
EJUA	8	0.894	0.018	0.623
EJUB	8	0.894	0.016	0.295
MSA	3	0.682	0.011	-0.222
SLO	4	0.742	0.014	0.932

1 surrounding the *tb1* region in our resequenced populations (Fig. 2), and r^2 in the *tb1*
2 region does not differ significantly between populations with (average r^2 of 0.085) and
3 without (average $r^2 = 0.082$) the *Hopscotch* insertion. In fact, average r^2 is lower in the
4 *tb1* region ($r^2 = 0.056$) than across the rest of chromosome 1 ($r^2 = 0.083$; Table 3).

Table 3: mean r^2 values between SNPs on chromosome 1, in the broad *tb1* region, within the 5' UTR of *tb1* (Region 1), and 66kb upstream of *tb1* (Region 2).

Population	Chr. 1	<i>tb1</i> region	Region 1	Region 2
Ejutla A	0.095	0.050	0.747	0.215
Ejutla B	0.069	0.051	0.660	0.186
La Mesa	0.070	0.053	0.914	0.766
San Lorenzo	0.101	0.067	0.912	0.636

5 The lack of clustering of *Hopscotch* genotypes in our NJ tree as well as the lack of LD
6 around *tb1* do not support the hypothesis that the *Hopscotch* insertion in these
7 populations of *parviglumis* is the result of recent introgression. However, to further
8 explore this hypothesis we performed a STRUCTURE analysis using Illumina
9 MaizeSNP50 data from four of our *parviglumis* populations (EjuA, EjuB, MSA, and
10 SLO) and the maize 282 diversity panel (Cook et al., 2012; Pyhäjärvi et al., 2013). The
11 linkage model implemented in STRUCTURE can be used to identify ancestry of blocks of
12 linked variants which would arise as the result of recent admixture between populations.
13 If the *Hopscotch* insertion is present in populations of *parviglumis* as a result of recent
14 admixture with domesticated maize, we would expect the insertion and linked variants in
15 surrounding sites to be assigned to the "maize" cluster in our STRUCTURE runs, not
16 the "teosinte" cluster. In all runs, assignment to maize in the *tb1* region across all four
17 *parviglumis* populations is low (average 0.017) and much below the chromosome-wide

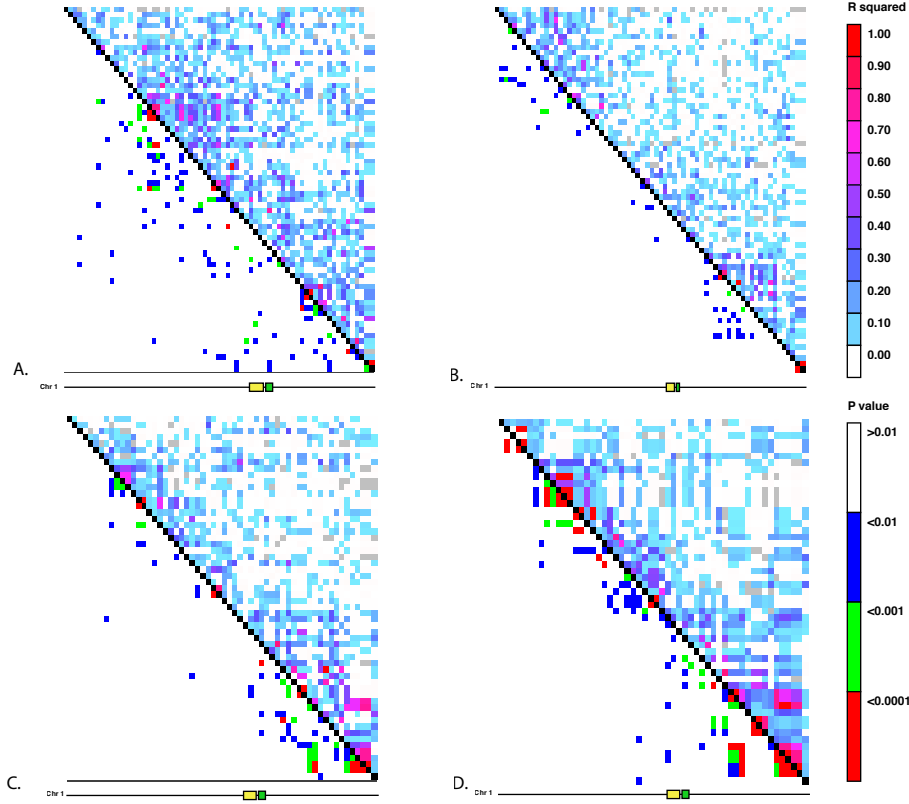


Figure 2: Linkage disequilibrium for SNPs in Mb 261-268 on chromosome 1. The yellow rectangle indicates the location of the *Hopscotch* insertion and the green represents the *tb1* ORF. A) Ejutla A; B) Ejutla B; C) La Mesa; D) San Lorenzo. The upper triangle above the black diagonal is colored based on the r^2 value between SNPs while the bottom triangle is colored based on p-value for the corresponding r^2 value.

- 1 average (0.20; Table 4; Fig. 3).
- 2 **Phenotyping**—To assess the contribution of *tb1* to phenotypic variation in tillering
- 3 in a natural population, we grew plants from seed sampled from the San Lorenzo
- 4 population of *parviglumis*, which had a high mean frequency (0.44) of the *Hopscotch*
- 5 insertion based on our initial genotyping. We measured tillering index (TI), the ratio of

Table 4: Assignments to maize and teosinte in the *tb1* and chromosome 1 regions from STRUCTURE

	<i>tb1</i> region		Chr 1	
Population	Maize	Teosinte	Maize	Teosinte
Ejutla A	0.022	0.978	0.203	0.797
Ejutla B	0.019	0.981	0.187	0.813
La Mesa	0.012	0.988	0.193	0.807
San Lorenzo	0.016	0.984	0.205	0.795

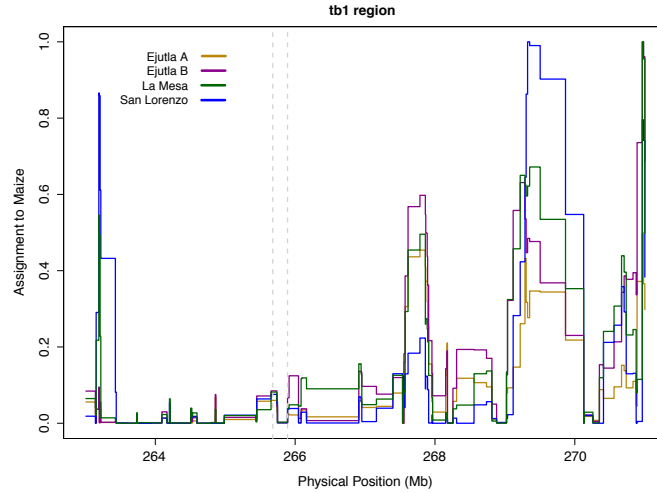


Figure 3: STRUCTURE assignment to maize across a section of chromosome 1. The dotted lines mark the beginning of the sequenced region 66kb upstream (Region 2) and the end of the *tb1* ORF.

- 1 the sum of tiller lengths to plant height, for 216 plants (Phenotyping 1) from within the
- 2 San Lorenzo population, and genotyped plants for the *Hopscotch* insertion. We found the
- 3 *Hopscotch* segregating at a frequency of 0.65 with no significant deviations from expected

1 frequencies under Hardy-Weinberg equilibrium. After performing a repeated measures
2 ANOVA between our transformed tillering index data and *Hopscotch* genotype we find no
3 correlation between genotype at the *Hopscotch* insertion and tillering index (Fig. 4), tiller
4 number, or culm diameter.

5 We performed a second grow-out of *parviglumis* from San Lorenzo (Phenotyping 2)
6 to assess whether lighting conditions or sample size may have affected our ability to
7 detect an effect of *tb1*. For the second grow-out we measured tillering index every five
8 days through day 50 for 302 plants. We found the *Hopscotch* allele segregating at a
9 frequency of 0.69, *is it in HWE in this pop? my guess is no!* with a 0.6 frequency of *Hopscotch*
10 homozygotes, and a 0.2 frequency of both heterozygotes and homozygotes for the teosinte
11 allele. Results were similar to Phenotyping 1, with no significant correlation between
12 *Hopscotch* and any of the three phenotypes measured.

13 DISCUSSION

14 Adaptation occurs due to selection on standing variation or *de novo* mutations.
15 Adaptation from standing variation has been well-described in a number of systems; for
16 example, selection for lactose tolerance in humans (Plantinga et al., 2012; Tishkoff et al.,
17 2007), variation at the *Eda* locus in three-spined stickleback (Kitano et al., 2008;
18 Colosimo et al., 2005), and pupal diapause in the Apple Maggot fly (Feder et al., 2003).
19 Although the adaptive role of standing variation has been described in many systems, its
20 importance in domestication is not as well studied.

21 In maize, alleles at domestication loci (*RAMOSA1*, Sigmon and Vollbrecht 2010;
22 *barren stalk1*, Gallavotti et al. 2004; and *grassy tillers1*, Whipple et al. 2011) are thought
23 to have been selected from standing variation, suggesting that diversity already present in
24 teosinte may have played an important role in maize domestication. The *teosinte*
25 *branched1* gene is one of the best characterized domestication loci, and, while previous
26 studies have suggested that differences in plant architecture between maize and teosinte

1 are a result of selection on standing variation at this locus (Clark et al., 2006; Studer
2 et al., 2011), much remains to be discovered regarding natural variation at this locus and
3 its ecological role in teosinte.

4 Studer et al. (2011) genotyped 90 accessions of teosinte (inbred and outbred),
5 providing the first evidence that the *Hopscotch* insertion is segregating in teosinte (Studer
6 et al., 2011). Given that the *Hopscotch* insertion has been estimated to predate the
7 domestication of maize, it is not surprising that it can be found segregating in
8 populations of teosinte. However, by widely sampling across teosinte populations our
9 study provides greater insight into the distribution and prevalence of the *Hopscotch* in
10 teosinte. While our findings are consistent with Studer et al. (2011) in that we identify
11 the *Hopscotch* allele segregating in teosinte, we find it at higher frequency than
12 previously suggested. Many of our populations with a high frequency of the *Hopscotch*
13 allele fall in the Jalisco cluster identified by Fukunaga et al. (2005), perhaps suggesting a
14 different history of the *tb1* locus in this region than in the Balsas River Basin where
15 maize was domesticated (Matsuoka et al., 2002). Potential explanations for the high
16 frequency of the *Hopscotch* element in *parviglumis* from the Jalisco cluster include gene
17 flow from maize, genetic drift, and natural selection.

18 While gene flow from crops into their wild relatives is well-known, (Ellstrand et al.,
19 1999; Zhang et al., 2009; Thurber et al., 2010; Baack et al., 2008; Hubner et al., 2012;
20 Wilkes, 1977; van Heerwaarden et al., 2011; Barrett, 1983), our results are more
21 consistent with Hufford et al. (2013) who found resistance to introgression from maize
22 into teosinte around domestication loci. We find no evidence of recent introgression in
23 our analyses. Clustering in our NJ trees do not reflect the pattern expected if maize
24 alleles at the *tb1* locus had introgressed into populations of teosinte. Moreover, there is
25 no signature of elevated LD in the *tb1* region relative to the rest of chromosome 1, and
26 Bayesian assignment to a maize cluster in this region is both low and below the
27 chromosome-wide average (Fig. 3, Table 4). Together, these data point to an explanation

1 other than recent introgression for the high observed frequency of *Hopscotch* in a subset
2 of our *parviglumis* populations.

3 Although recent introgression seems unlikely, we cannot rule out ancient introgression
4 as an explanation for the presence of the *Hopscotch* in these populations. If the
5 *Hopscotch* allele was introgressed in the distant past, recombination may have broken up
6 LD, a process that would be consistent with our data. We find this scenario less
7 plausible, however, as there is no reason why gene flow should have been high in the past
8 but absent in present-day sympatric populations. In fact, early generation maize-teosinte
9 hybrids are common in these populations today (MB Hufford, pers. observation), and
10 genetic data support ongoing gene flow between domesticated maize and both *mexicana*
11 and *parviglumis* in a number of sympatric populations (Hufford et al., 2013; Ellstrand
12 et al., 2007; van Heerwaarden et al., 2011).

13 Remaining explanations for differential frequencies of the *Hopscotch* among teosinte
14 populations include both genetic drift and natural selection. Previous studies using both
15 SSRs and genome-wide SNP data have found evidence for a population bottleneck in the
16 San Lorenzo population (Hufford, 2010; Pyhäjärvi et al., 2013), and the lower levels of
17 sequence diversity in this population in the 5' UTR (Region 1) coupled with more
18 positive values of Tajima's D are consistent with these earlier findings. Such population
19 bottlenecks can exaggerate the effects of genetic drift through which the *Hopscotch* allele
20 may have risen to high frequency entirely by chance. A bottleneck in San Lorenzo,
21 however, does not explain the high frequency of the *Hopscotch* in multiple populations in
22 the Jalisco cluster. Moreover, available information on diversity and population structure
23 among Jaliscan populations (Hufford, 2010; Pyhäjärvi et al., 2013) is not suggestive of
24 recent colonization or other demographic events that would predict a high frequency of
25 the allele across populations. Finally, diversity values in the 5' UTR of *tb1* are suggestive
26 of natural selection acting upon the gene in natural populations of *parviglumis*. Overall
27 nucleotide diversity is 76% less than seen in the sequences from the 66kb upstream

1 region, and Tajima's D is considerably lower and consistently negative. In fact, values of
2 Tajima's D in the 5' UTR are toward the extreme negative end of the distribution of this
3 statistic previously calculated across loci sequenced in *parviglumis* (Wright et al., 2005;
4 Moeller et al., 2007). Though not definitive, these results are consistent with the action
5 of selection on the upstream region of *tb1*, perhaps suggesting an ecological role for the
6 gene in *parviglumis*.

7 Significant effects of the *Hopscotch* insertion on lateral branch length, number of
8 cupules, and tillering index in domesticated maize have been well documented (Studer
9 et al., 2011). Weber et al. (2007) described significant phenotypic associations between
10 markers in and around *tb1* and lateral branch length and female ear length in a sample
11 from 74 natural populations of *parviglumis* (Weber et al., 2007); however, these data did
12 not include markers from the *Hopscotch* region 66kb upstream of *tb1*. Our study is the
13 first to explicitly examine the phenotypic effects of the *Hopscotch* insertion across a wide
14 collection of individuals sampled from natural populations of teosinte. We have found no
15 significant effect of the *Hopscotch* insertion on tillering index or tiller number, a result
16 that is discordant with its clear phenotypic effects in maize. One interpretation of this
17 result would be that the *Hopscotch* controls tillering in maize (Studer et al., 2011), but
18 tillering in teosinte is affected by variation at other loci. Consistent with this
19 interpretation, *tb1* is thought to be part of a complex pathway controlling branching,
20 tillering and other phenotypic traits (Kebrom and Brutnell, 2007; Clark et al., 2006). A
21 recent study by Studer and Doebley (2012) examined variation across traits in a
22 three-taxa allelic series at the *tb1* locus. Studer and Doebley (2012) introgressed nine
23 unique teosinte *tb1* segments (one from *Zea diploperennis*, and four each from *mexicana*
24 and *parviglumis*) into an inbred maize background and investigated their phenotypic
25 effects. Phenotypes were shown to cluster by taxon, indicating *tb1* may underlie
26 morphological diversification of *Zea*. Additional analysis in Studer and Doebley (2012)
27 suggested tillering index was controlled both by *tb1* and loci elsewhere in the genome.

1 Clues to the identity of these loci may be found in QTL studies that have identified loci
2 controlling branching architecture (*e.g.*, Doebley and Stec 1991, 1993). Many of these loci
3 (*grassy tillers*, *gt1*; *tassel-replaces-upper-ears1*, *tru1*; *terminal ear1*, *ter1*) have been
4 shown to interact with *tb1* (Whipple et al., 2011; Li, 2012), and both *tru1* and *ter1* affect
5 the same phenotypic traits as *tb1* (Doebley et al., 1995). *tru1*, for example, has been
6 shown to act either epistatically or downstream of *tb1*, affecting both branching
7 architecture (decreased apical dominance) and tassel phenotypes (shortened tassel and
8 shank length and reduced tassel number; Li 2012). Variation in these additional loci may
9 have affected tillering in our collections and contributed to the lack of correlation we see
10 between *Hopscotch* genotype and tillering. Finally, although photoperiod for Phenotyping
11 2 reasonably approximated that of the normal *parviglumis* growing season,
12 greenhouse-specific environmental conditions (plant density, light regime, etc...) may have
13 contributed to tillering responses different from those found in nature, obscuring the
14 effect of the *Hopscotch* insertion on variation.

15 In conclusion, our findings demonstrate that the *Hopscotch* allele is more widespread
16 in populations of *parviglumis* and *mexicana* than previously thought. Analysis of linkage
17 using SNPs from across chromosome 1 does not suggest that the *Hopscotch* allele is
18 present in these populations due to recent introgression; however, it seems unlikely that
19 the insertion would have drifted to high frequency in multiple populations. We do,
20 however, find preliminary evidence of selection on the *tb1* locus in *parviglumis*; this
21 coupled with our observation of high frequency of the *Hopscotch* insertion in a number of
22 populations suggests that the locus may play an ecological role in teosinte. In contrast to
23 domesticated maize, the *Hopscotch* insertion does not appear to have a large effect on
24 tillering in *parviglumis*. Future studies should examine expression levels of *tb1* in teosinte
25 with and without the *Hopscotch* insertion and further characterize the effects of
26 additional loci involved in branching architecture (*e.g.* *gt1*, *tru1*, and *ter1*). These data,
27 in conjunction with more exhaustive phenotyping, should help reveal the ecological

- 1 significance of the domesticated *tb1* allele in natural populations of teosinte.

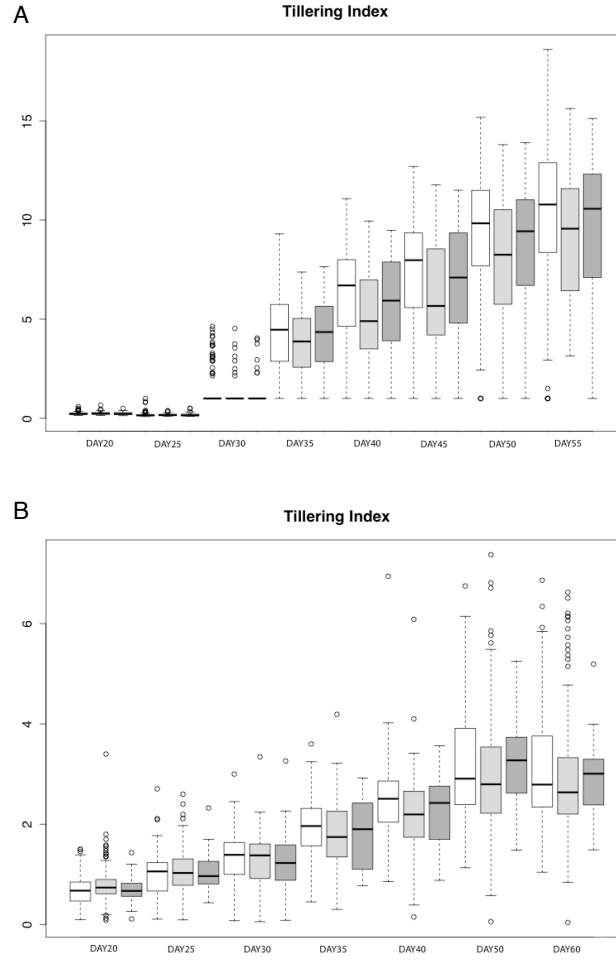


Figure 4: Box-plots showing tillering index in greenhouse grow-outs for Phenotyping 1 (A) and Phenotyping 2 (B). White indicates individuals homozygous for the *Hopscotch*, light grey represents heterozygotes, and dark grey represents homozygotes for the *teosinte* (No *Hopscotch*) allele. Within boxes, dark black lines represent the median, and the edges of the boxes are the first and third quartiles. Outliers are displayed as dots, the maximum value excluding outliers is shown with the top whisker, while the minimum excluding outliers is shown with the bottom whisker.

LITERATURE CITED

- 1
2 BAACK, E., Y. SAPIR, M. CHAPMAN, J. BURKE, AND L. RIESEBERG. 2008. Selection
3 on domestication traits and quantitative trait loci in crop-wild sunflower hybrids. *Mol*
4 *Ecol* 17: 666–677.
- 5 BAI, Y. AND P. LINDHOUT. 2007. Domestication and breeding of tomatoes: What have
6 we gained and what can we gain in the future? *Annals of Botany* 100: 1085–1094.
- 7 BARRETT, S. 1983. Crop mimicry in weeds. *Econ Bot* 37: 255–282.
- 8 BRADBURY, P., Z. ZHANG, D. KROON, T. CASSTEVENS, Y. RAMDOSS, AND
9 E. BUCKLER. 2007. Tassel: software for association mapping of complex traits in
10 diverse samples. *Bioinformatics* 23: 2633–2635.
- 11 BRIGGS, W., M. McMULLEN, B. GAUT, AND J. DOEBLEY. 2007. Linkage mapping of
12 domestication loci in a large maize-teosinte backcross resource. *Genetics* 177:
13 1915–1928.
- 14 CHIA, J., C. SONG, P. BRADBURY, D. COSTICH, N. DE, LEON, J. DOEBLEY,
15 R. ELSHIRE, B. GAUT, L. GELLER, J. GLAUBITZ, M. GORE, K. GUILL,
16 J. HOLLAND, M. HUFFORD, J. LAI, M. LI, X. LIU, Y. LU, R. McCOMBIE,
17 R. NELSON, J. POLAND, B. PRASANNA, T. PYHÄJÄRVI, T. RONG, R. SEKHON,
18 Q. SUN, M. TENAILLON, F. TIAN, J. WANG, X. XU, Z. ZHANG, S. KAEPLER,
19 J. ROSS-IBARRA, M. McMULLEN, E. BUCKLER, G. ZHANG, Y. XU, AND D. WARE.
20 2012. Maize hapmap2 identifies extant variation from a genome in flux. *Nat Genet* 44:
21 803–U238.
- 22 CLARK, R., T. WAGLER, P. QUIJADA, AND J. DOEBLEY. 2006. A distant upstream
23 enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and
24 inflorescent architecture. *Nat Genet* 38: 594–597.

- 1 COLOSIMO, P., K. HOSEMAN, S. BALABHADRA, G. VILLARREAL, M. DICKSON,
2 J. GRIMWOOD, J. SCHMUTZ, R. MYERS, D. SCHLUTER, AND D. KINGSLEY. 2005.
3 Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin
4 alleles. *Science* 307: 1928–1933.
- 5 COOK, J., M. McMULLEN, J. HOLLAND, F. TIAN, P. BRADBURY, J. ROSS-IBARRA,
6 E. BUCKLER, AND S. FLINT-GARCIA. 2012. Genetic architecture of maize kernel
7 composition in the nested association mapping and inbred association panels. *Plant*
8 *Physiol* 158: 824–834.
- 9 COOP, G., D. WITONSKY, A. DI, RIENZO, AND J. PRITCHARD. 2010. Using
10 environmental correlations to identify loci underlying local adaptation. *Genetics* 185:
11 1411–1423.
- 12 DOEBLEY, J. 2004. The genetics of maize evolution. *Annu Rev Genet* 38: 37–59.
- 13 DOEBLEY, J. AND A. STEC. 1991. Genetic-analysis of the morphological differences
14 between maize and teosinte. *Genetics* 129: 285–295.
- 15 DOEBLEY, J. AND A. STEC. 1993. Inheritance of the morphological differences between
16 maize and teosinte: Comparison of results for two F₂ populations. *Genetics* 134:
17 559–570.
- 18 DOEBLEY, J., A. STEC, AND C. GUSTUS. 1995. *teosinte branched1* and the origin of
19 maize: Evidence for epistasis and the evolution of dominance. *Genetics* 141: 333–346.
- 20 DOEBLEY, J., A. STEC, AND L. HUBBARD. 1997. The evolution of apical dominance in
21 maize. *Nature* 386: 485–488.
- 22 DOYLE, J. AND J. DOYLE. 1990. A rapid total dna preparation procedure for small
23 quantities of fresh tissue. *Phytochemical Bulletin* 19: 11–15.

- 1 ELLSTRAND, N., L. GARNER, S. HEGDE, R. GUADAGNUOLO, AND L. BLANCAS. 2007.
2 Spontaneous hybridization between maize and teosinte. *Journal of Heredity* 98:
3 183–187.
- 4 ELLSTRAND, N., H. PRENTICE, AND J. HANCOCK. 1999. Gene flow and introgression
5 from domesticated plants into their wild relatives. *Annu Rev Ecol Syst* 30: 539–563.
- 6 FALUSH, D., M. STEPHENS, AND J. PRITCHARD. 2003. Inference of population structure
7 using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*
8 164: 1567–1587.
- 9 FAO/IIASA/ISRIC/ISSCAS/JRC. 2012. Harmonized World Soil Database, version
10 1.2. FAO, Rome, Italy and IIASA, Laxenburg, Austria.
- 11 FEDER, J., S. BERLOCHER, J. ROETHELE, H. DAMBROSKI, J. SMITH, W. PERRY,
12 V. GAVRILOVIC, K. FILCHAK, J. RULL, AND M. ALUJA. 2003. Allopatric genetic
13 origins for sympatric host-plant shifts and race formation in rhagoletis. *P Natl Acad*
14 *Sci Usa* 100: 10314–10319.
- 15 FOX, J. AND S. WEISBERG. 2011. An R Companion to Applied Regression, vol. Second
16 Edition. Sage, Thousand Oaks, CA.
- 17 FUKUNAGA, K., T. NUSSBAUM-WAGLER, B. LI, Q. ZHAO, Y. VIGOUROUX,
18 M. FALLER, K. BOMBLIES, L. LUKENS, AND J. DOEBLEY. 2005. Genetic diversity
19 and population structure of teosinte. *Genetics* 169: 2241–2254.
- 20 GALLAVOTTI, A., Q. ZHAO, J. KYOZUKA, R. MEELEY, M. RITTER, J. DOEBLEY,
21 M. PE, AND R. SCHMIDT. 2004. The role of barren stalk1 in the architecture of maize.
22 *Nature* 432: 630–635.
- 23 GERKE, J., J. EDWARDS, G. KE, J. ROSS-IBARRA, AND M. McMULLEN. 2013. The

1 genomic impacts of drift and selection for hybrid performance in maize. *arXiv*
2 1307.7313.

3 GOUDET, J. 2005. Hierfstat, a package for r to compute and test hierarchical f-statistics.
4 *Mol Ecol Notes* 5: 184–186.

5 HUBNER, S., T. GUNTHER, A. FLAVELL, E. FRIDMAN, A. GRANER, A. KOROL, AND
6 K. SCHMID. 2012. Islands and streams: clusters and gene flow in wild barley
7 populations from the levant. *Mol Ecol* 21: 1115–1129.

8 HUFFORD, M. 2010. Genetic and ecological approaches to guide conservation of teosinte
9 (*zea mays* ssp. *parviglumis*), the wild progenitor of maize. *PhD Dissertation* : 130pp.

10 HUFFORD, M., P. BILINSKI, T. PYHÄJÄRVI, AND J. ROSS-IBARRA. 2012a. Teosinte as a
11 model system for population and ecological genomics. *Trends in Genetics* 12: 606–615.

12 HUFFORD, M., P. LUBINSKY, T. PYHÄJÄRVI, M. DEVENGENZO, N. ELLSTRAND, AND
13 J. ROSS-IBARRA. 2013. The genomic signature of crop-wild introgression in maize.
14 *PLoS Genetics* 9: e1003477.

15 HUFFORD, M., X. XU, J. VAN, HEERWAARDEN, T. PYHÄJÄRVI, J. CHIA,
16 R. CARTWRIGHT, R. ELSHIRE, J. GLAUBITZ, K. GUILL, S. KAEPLER, J. LAI,
17 P. MORRELL, L. SHANNON, C. SONG, N. SPRINGER, R. SWANSON-WAGNER,
18 P. TIFFIN, J. WANG, G. ZHANG, J. DOEBLEY, M. McMULLEN, D. WARE,
19 E. BUCKLER, S. YANG, AND J. ROSS-IBARRA. 2012b. Comparative population
20 genomics of maize domestication and improvement. *Nat Genet* 44: 808–U118.

21 KEARSE, M., R. MOIR, A. WILSON, S. STONES-HAVAS, M. CHEUNG, S. STURROCK,
22 S. BUXTON, A. COOPER, S. MARKOWITZ, C. DURAN, T. THIERER, B. ASHTON,
23 P. MEINTJES, AND A. DRUMMOND. 2012. Geneious basic: An integrated and
24 extendable desktop software platform for the organization and analysis of sequence
25 data. *Bioinformatics* 28: 1647–1649.

- 1 KEBROM, T. AND T. BRUTNELL. 2007. The molecular analysis of the shade avoidance
2 syndrome in the grasses has begun. *Journal of Experimental Botany* 58: 3079–3089.
- 3 KITANO, J., D. BOLNICK, D. BEAUCHAMP, M. MAZUR, S. MORI, T. NAKANO, AND
4 C. PEICHEL. 2008. Reverse evolution of armor plates in the threespine stickleback.
5 *Curr Biol* 18: 769–774.
- 6 KOVACH, M. AND S. MCCOUCH. 2008. Leveraging natural diversity: back through the
7 bottleneck. *Genome studies and Molecular Genetics* 11: 193–200.
- 8 LI, W. 2012. Tassels replace upper ears1 encodes a putative transcription factor that
9 regulates maize shoot architecture by multiple pathways. *PhD Dissertation* : 122.
- 10 LUKENS, L. AND J. DOEBLEY. 1999. Epistatic and environmental interactions for
11 quantitative trait loci involved in maize evolution. *Genet Res* 74: 291–302.
- 12 MALOOF, M., K. SOLIMAN, R. JORGENSEN, AND R. ALLARD. 1984. Ribosomal dna
13 spacer length polymorphisms in barley - mendelian inheritance, chromosomal location,
14 and population dynamics. *P Natl Acad Sci Usa* 81: 8014–8018.
- 15 MATSUOKA, Y., Y. VIGOUROUX, M. GOODMAN, G. SANCHEZ, E. BUCKLER, AND
16 J. DOEBLEY. 2002. A single domestication for maize shown by multilocus
17 microsatellite genotyping. *P Natl Acad Sci Usa* 99: 6080–6084.
- 18 MOELLER, D. A., M. I. TENAILLON, AND P. TIFFIN. 2007. Population structure and its
19 effects on patterns of nucleotide polymorphism in teosinte (*zea mays* ssp. *parviglumis*).
20 *Genetics* 176: 1799–1809.
- 21 OLSEN, K. AND B. GROSS. 2010. Genetic perspectives on crop domestication. *Trends in*
22 *Plant Science* 15: 529–537.
- 23 PIPERNO, D., A. RANERE, I. HOLST, J. IRIARTE, AND R. DICKAU. 2009. Starch grain

1 and phytolith evidence for early ninth millennium bp maize from the central balsas
2 river valley, mexico. *P Natl Acad Sci Usa* 106: 5019–5024.

3 PLANTINGA, T., S. ALONSO, N. IZAGIRRE, M. HERVELLA, R. FREGEL, J. VAN DER
4 MEER, M. NETEA, AND C. DE LA RUA. 2012. Low prevalence of lactase persistence in
5 neolithic south-west europe. *Eur J Hum Genet* 20: 778–782.

6 PYHÄJÄRVI, T., M. HUFFORD, AND J. ROSS-IBARRA. 2013. Complex patterns of local
7 adaptation in the wild relatives of maize. *Genome Biology and Evolution* 5: 1594–1609.

8 ROSS-IBARRA, J., P. MORRELL, AND B. GAUT. 2007. Plant domestication, a unique
9 opportunity to identify the genetic basis of adaptation. *P Natl Acad Sci Usa* 104:
10 8641–8648.

11 ROSS-IBARRA, J., M. TENAILLON, AND B. GAUT. 2009. Historical divergence and gene
12 flow in the genus *zea*. *Genetics* 181: 1399–1413.

13 ROZEN, S. AND H. SKALETSKY. 2000. Primer3 on the www for general users and for
14 biologist programmers. *Methods in Molecular Biology* : 365–386.

15 SIGMON, B. AND E. VOLLBRECHT. 2010. Evidence of selection at the *ramosa1* locus
16 during maize domestication. *Mol Ecol* 19: 1296–1311.

17 STUDER, A. AND J. DOEBLEY. 2012. Evidence for a natural allelic series at the maize
18 domestication locus *teosinte branched1*. *Genetics* 19: 951–958.

19 STUDER, A., Q. ZHAO, J. ROSS-IBARRA, AND J. DOEBLEY. 2011. Identification of a
20 functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43:
21 1160–U164.

22 THORNTON, K. 2003. libsequence: a c++ class library for evolutionary genetic analysis.
23 *Bioinformatics* 19: 2325–2327.

- 1 THURBER, C., M. REAGON, B. GROSS, K. OLSEN, Y. JIA, AND A. CAICEDO. 2010.
2 Molecular evolution of shattering loci in us weedy rice. *Mol Ecol* 19: 3271–3284.
- 3 TISHKOFF, S., F. REED, A. RANCIARO, B. VOIGHT, C. BABBITT, J. SILVERMAN,
4 K. POWELL, H. MORTENSEN, J. HIRBO, M. OSMAN, M. IBRAHIM, S. OMAR,
5 G. LEMA, T. NYAMBO, J. GHORI, S. BUMPSTEAD, J. PRITCHARD, G. WRAY, AND
6 P. DELOUKAS. 2007. Convergent adaptation of human lactase persistence in africa and
7 europe. *Nat Genet* 39: 31–40.
- 8 VAN HEERWAARDEN, J., J. DOEBLEY, W. BRIGGS, J. GLAUBITZ, M. GOODMAN,
9 J. GONZALEZ, AND J. ROSS-IBARRA. 2011. Genetic signals of origin, spread, and
10 introgression in a large sample of maize landraces. *P Natl Acad Sci Usa* 108:
11 1088–1092.
- 12 WEBER, A., R. CLARK, L. VAUGHN, J. SANCHEZ-GONZALEZ, J. YU, B. YANDELL,
13 P. BRADBURY, AND J. DOEBLEY. 2007. Major regulatory genes in maize contribute to
14 standing variation in teosinte (*zea mays* ssp *parviglumis*). *Genetics* 177: 2349–2359.
- 15 WHIPPLE, C., T. KEBROM, A. WEBER, F. YANG, D. HALL, R. MEELEY,
16 R. SCHMIDT, J. DOEBLEY, T. BRUTNELL, AND D. JACKSON. 2011. grassy tillers1
17 promotes apical dominance in maize and responds to shade signals in the grasses. *P*
18 *Natl Acad Sci Usa* 108: E506–E512.
- 19 WHITTON, J., D. WOLF, D. ARIAS, A. SNOW, AND L. RIESBERG. 1997. The
20 persistence of cultivar alleles in wild populations of sunflowers five generations after
21 hybridization. *Theoretical and Applied Genetics* 95: 33–40.
- 22 WILKES, H. 1977. Hybridization of maize and teosinte, in mexico and guatemala and the
23 improvement of maize. *Economic Botany* 31: 254–293.
- 24 WRIGHT, S. I., I. V. BI, S. G. SCHROEDER, M. YAMASAKI, J. F. DOEBLEY, M. D.

- 1 McMULLEN, AND B. S. GAUT. 2005. The effects of artificial selection on the maize
2 genome. *Science* 308: 1310–1314.
- 3 ZHANG, L., Q. ZHU, Z. WU, J. ROSS-IBARRA, B. GAUT, S. GE, AND T. SANG. 2009.
4 Selection on grain shattering genes and rates of rice domestication. *New Phytol* 184:
5 708–720.