



# The genetic architecture of teosinte catalyzed and constrained maize domestication

Chin Jian Yang<sup>a</sup>, Luis Fernando Samayoa<sup>b,c</sup>, Peter J. Bradbury<sup>d</sup>, Bode A. Olukolu<sup>e</sup>, Wei Xue<sup>a</sup>, Alessandra M. York<sup>a</sup>, Michael R. Tuholski<sup>a</sup>, Weidong Wang<sup>a</sup>, Lora L. Daskalska<sup>a</sup>, Michael A. Neumeier<sup>a</sup>, Jose de Jesus Sanchez-Gonzalez<sup>f</sup>, Maria Cinta Romay<sup>g</sup>, Jeffrey C. Glaubitz<sup>g</sup>, Qi Sun<sup>g</sup>, Edward S. Buckler<sup>d</sup>, James B. Holland<sup>b,c</sup>, and John F. Doebley<sup>a,1</sup>

<sup>a</sup>Laboratory of Genetics, University of Wisconsin–Madison, Madison, WI 53706; <sup>b</sup>US Department of Agriculture–Agricultural Research Service Plant Science Research Unit, North Carolina State University, Raleigh, NC 27695; <sup>c</sup>Department of Crop Science, North Carolina State University, Raleigh, NC 27695; <sup>d</sup>US Department of Agriculture–Agricultural Research Service, Cornell University, Ithaca, NY 14853; <sup>e</sup>Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, TN 37996; <sup>f</sup>Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Zapopan, Jalisco CP45110, Mexico; and <sup>g</sup>Genomic Diversity Facility, Cornell University, Ithaca, NY 14853

Contributed by John F. Doebley, January 28, 2019 (sent for review December 14, 2018; reviewed by Loren H. Rieseberg and Bruce Walsh)

The process of evolution under domestication has been studied using phylogenetics, population genetics–genomics, quantitative trait locus (QTL) mapping, gene expression assays, and archaeology. Here, we apply an evolutionary quantitative genetic approach to understand the constraints imposed by the genetic architecture of trait variation in teosinte, the wild ancestor of maize, and the consequences of domestication on genetic architecture. Using modern teosinte and maize landrace populations as proxies for the ancestor and domesticate, respectively, we estimated heritabilities, additive and dominance genetic variances, genetic-by-environment variances, genetic correlations, and genetic covariances for 18 domestication-related traits using realized genomic relationships estimated from genome-wide markers. We found a reduction in heritabilities across most traits, and the reduction is stronger in reproductive traits (size and numbers of grains and ears) than vegetative traits. We observed larger depletion in additive genetic variance than dominance genetic variance. Selection intensities during domestication were weak for all traits, with reproductive traits showing the highest values. For 17 of 18 traits, neutral divergence is rejected, suggesting they were targets of selection during domestication. Yield (total grain weight) per plant is the sole trait that selection does not appear to have improved in maize relative to teosinte. From a multivariate evolution perspective, we identified a strong, nonneutral divergence between teosinte and maize landrace genetic variance–covariance matrices (*G*-matrices). While the structure of *G*-matrix in teosinte posed considerable genetic constraint on early domestication, the maize landrace *G*-matrix indicates that the degree of constraint is more unfavorable for further evolution along the same trajectory.

maize | teosinte | domestication | evolution | selection

Crop domestication was central to Charles Darwin's formulation of his theory of evolution by natural selection (1), and it continues to inspire modern evolutionists to investigate a broad range of questions in evolutionary biology. The power of domestication as a model for evolution draws from several of its features compared with cases of natural evolution. All crops were domesticated within the last 12,000 y, providing definitive and restricted timeline for events (2). Many crops have well-documented archaeological records (3, 4). Unlike evolution in nature where the ancestral species is often unknown, the wild ancestors for crops are mostly known and available for comparative analyses. Given the recent origin of crops, they typically remain cross-compatible with their nearest wild relatives, allowing genetic analysis using crosses of domesticated by wild species. As a result, there is a substantial body of research on evolution under domestication for many crops including maize, rice, wheat, sorghum, barley, cotton, common beans, and sunflower (5–7).

The origin and evolutionary history of maize has been intensely studied (8). Phylogenetic analysis and archaeological data

revealed that maize originated from a single domestication event in southern Mexico about 9,000 y ago (9, 10). The direct ancestor of maize is a lowland wild grass known as teosinte (*Zea mays* ssp. *parviglumis*). Population genetics comparison of maize and teosinte revealed evidence for recent selection in multiple genomic regions, a moderate bottleneck causing loss in genetic diversity during domestication, and postdomestication gene flow from teosinte into maize that enhanced maize adaptation to diverse environments (11–14). Despite the loss of genetic diversity by selection and bottleneck, modern maize remains rich in genetic variation, facilitating its pre-Columbian adaptation to diverse habitats from Canada to Chile (15, 16).

Teosinte and maize differ profoundly in plant and inflorescence architecture (Fig. 1) (17, 18). Teosinte plants typically have many long branches, each bearing multiple small ears along their length and tassels at their tip. In contrast, maize plants typically have one or two short branches, each with a single ear at its tip. A teosinte plant can have several hundred ears, each with only 10 grains, whereas maize typically has only two ears, each with several hundred grains. These differences in growth form and ear size are related to the different reproductive strategies.

## Significance

Crop domestication is a well-established system for understanding evolution. We interrogated the genetic architecture of maize domestication from a quantitative genetics perspective. We analyzed domestication-related traits in a maize landrace and a population of its ancestor, teosinte. We observed strong divergence in the underlying genetic architecture including change in the genetic correlations among traits. Despite striking divergence, selection intensities were low for all traits, indicating that selection under domestication can be weaker than natural selection. Analyses suggest total grain weight per plant was not improved and that genetic correlations placed considerable constraint on selection. We hope our results will motivate crop evolutionists to perform similar work in other crops.

Author contributions: E.S.B., J.B.H., and J.F.D. designed research; C.J.Y., L.F.S., P.J.B., B.A.O., W.X., A.M.Y., M.R.T., W.W., L.L.D., M.A.N., J.d.J.S.-G., M.C.R., J.C.G., Q.S., E.S.B., J.B.H., and J.F.D. performed research; C.J.Y., L.F.S., P.J.B., and J.B.H. analyzed data; C.J.Y., P.J.B., M.C.R., J.B.H., and J.F.D. wrote the paper; J.d.J.S.-G. provided seeds for maize landrace; and M.C.R., J.C.G., and Q.S. organized the project.

Reviewers: L.H.R., University of British Columbia, Vancouver; and B.W., University of Arizona.

The authors declare no conflict of interest.

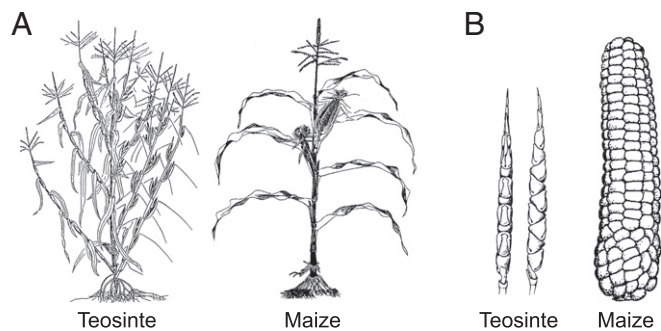
This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: The data reported in this paper have been deposited in the figshare database (<https://doi.org/10.6084/m9.figshare.7655588>).

<sup>1</sup>To whom correspondence should be addressed. Email: [jdoebley@wisc.edu](mailto:jdoebley@wisc.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820997116/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820997116/-DCSupplemental).

Published online March 6, 2019.



**Fig. 1.** Morphology of teosinte and maize. Differences in plant morphology between teosinte and maize are highlighted in *A*, while differences in ear morphology are shown in *B*. Teosinte plant has many branches with multiple ears on each branch and tassel at the tip of the branch; maize plant has few branches with a single ear on each branch and ear at the tip of the branch. Teosinte ear has few grains enclosed in fruit cases, while maize ear has many grains with exposed fruit cases. Adapted from ref. 19.

Teosinte has a plastic growth form, allowing the plant to branch prolifically in good environments to produce hundreds of ears or remain small and weakly branched with just a few ears in poor environments. Maize has a more fixed growth form, producing one or two large ears over a wide range of environments, which represents an adaptation for easy harvest of the grain by its human cultivators.

The genetic basis of the morphological differences between maize and teosinte has been intensely investigated. Quantitative trait locus mapping defined the genetic architecture underlying trait differences as complex and involving many loci, a few of which have been mapped to the underlying genes (19–28). Still, many questions regarding the genetics of maize domestication have never been addressed. Does teosinte possess heritable variation for domestication traits? To what extent was that genetic variation depleted by domestication? How did the genetic architecture of domestication traits within teosinte constrain or direct domestication? Given the genetic architecture in teosinte, how strong would selection need to be to complete domestication within the known time frame?

Here, we seek to add to the understanding of maize domestication using a quantitative genetic approach based on the genomic similarities among relatives to infer the genetic architecture of domestication traits within teosinte and maize landrace. We selected parent plants from a modern teosinte and a modern maize landrace population sampled from nearby locations. We mated the teosinte parents with one another to produce a cohort of 4,455 offspring with variable degrees of kinship, and similarly, we mated the maize landrace parents with one another to produce a parallel cohort of 4,398 maize offspring. We scored the offspring for multiple domestication traits and estimated the genetic relationships among the offspring with molecular marker data. We fitted a statistical model to the data that allowed us to estimate a variety of genetic parameters that define genetic architecture. This approach provides a first glimpse into a previously uninvestigated aspect of the story of maize domestication.

In this article, we report that the narrow-sense heritabilities ( $h^2$ ) for domestication traits are generally depleted in maize landrace compared with teosinte. In contrast, the proportion of genetic variance attributable to dominance and the proportion of phenotypic variance attributable to genetic-by-environment interaction increased in maize landrace compared with teosinte. Our estimates of the strength of selection during domestication indicate that reproductive (ear and grain) traits were under stronger selection than vegetative traits. The matrices of genetic variance–covariance among traits (*G*-matrices) within maize and teosinte are strongly differentiated, especially with respect to reproductive traits and this divergence was driven by selection. The teosinte *G*-matrix indicates that there was considerable

constraint on domestication early on, and the maize *G*-matrix indicates that constraints increased over time. Overall, genetic architecture diverged strongly over the domestication process, especially in regard to reproductive traits.

## Results

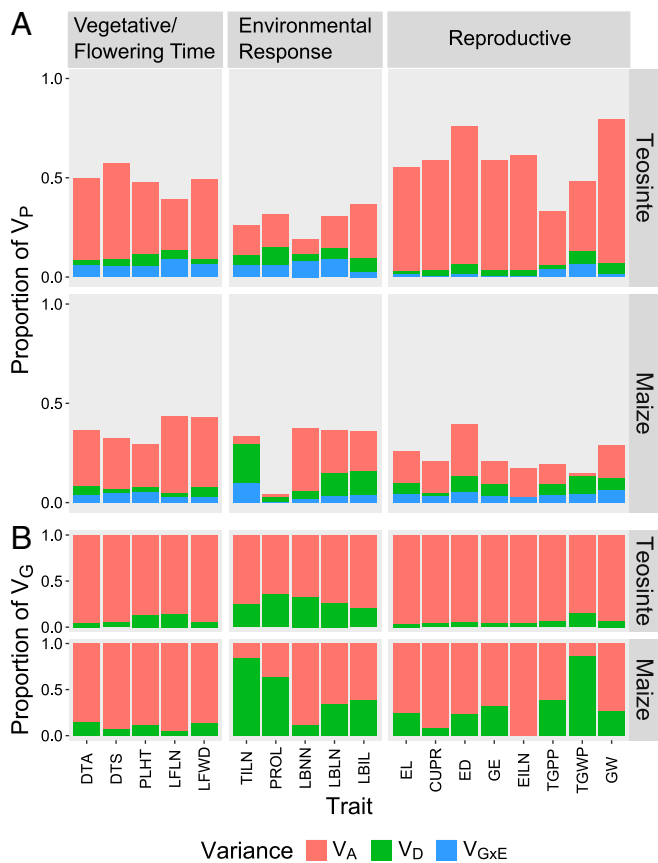
To infer the quantitative genetics of the maize ancestor, teosinte, and how this architecture was altered through domestication, we assayed the correlations among relatives for a set of 18 domestication traits (Table 1) in a large sample of teosinte plants of known pedigree derived from 49 founder teosinte parents chosen to serve as a proxy for the ancestral teosinte population from which maize was domesticated. For comparison, we also assayed the correlations among relatives for the same domestication traits in a large sample of maize landrace plants of known pedigree derived from 40 founder parents chosen to serve as a proxy for maize postdomestication but before the era of modern scientific breeding. For simplicity, we refer to each of the 18 traits by their acronyms as defined in Table 1. Both our teosinte and maize landrace parents were collected near the town of Palmar Chico in the State of Mexico in the region identified as the likely cradle of maize domestication (9, 10). Trait evaluations of the two subspecies were conducted in adjacent field blocks of a daylength-neutral environment over 2 y. Genome-wide DNA markers were used to estimate the additive and dominance-realized genomic relationships among all pairs of individuals in each population.

**Univariate Genetic Analyses.** Our teosinte population generally possesses higher levels of heritable variation than our maize landrace population for domestication traits, suggesting that domestication partially depleted genetic variance for traits that were the targets of human selection (Fig. 2*A* and *SI Appendix, Tables S1 and S2*). Using a variance component approach, we partitioned the phenotypic variance ( $V_P$ ) into additive genetic variance ( $V_A$ ), dominance genetic variance ( $V_D$ ), genetic-by-environment variance ( $V_{G \times E}$ ), and environmental variance ( $V_E$ ). To compare the  $V_A$  between teosinte and the maize landrace, we used the narrow-sense heritabilities ( $h^2 = V_A/V_P$ ), which is a  $V_P$ -standardized measure of the  $V_A$ . We observed higher  $h^2$  in teosinte ( $h^2 = 0.39 \pm 0.19$ , ranging from 0.07 to 0.73) than in maize landrace ( $h^2 = 0.19 \pm 0.11$ , ranging from 0.01 to 0.38) for 15 of 18 traits (Fig. 2*A*). Among the three predefined trait groups,

**Table 1. Trait abbreviations**

Trait	Acronym	Units	Trait group
Days to anthesis	DTA	days	Veg/FT
Days to silking	DS	days	Veg/FT
Plant height	PLHT	cm	Veg/FT
Leaf length	LFLN	cm	Veg/FT
Leaf width	LFWD	cm	Veg/FT
Tiller number	TILN	count	EnvRes
Prolificacy	PROL	count	EnvRes
Lateral branch node number	LBNN	count	EnvRes
Lateral branch length	LBLN	mm	EnvRes
Lateral branch internode length	LBIL	mm	EnvRes
Ear length	EL	mm	Rep
Cupules per row	CUPR	count	Rep
Ear diameter	ED	mm	Rep
Grains per ear	GE	count	Rep
Ear internode length	EILN	mm	Rep
Total grain per plant	TGPP	count	Rep
Total grain weight per plant	TGWP	g	Rep
Grain weight	GW	mg	Rep

List of 18 teosinte–maize landrace comparable traits and the corresponding acronyms, units, and trait groups. The trait groups are abbreviated as Veg/FT for Vegetative/Flowering Time, EnvRes for Environmental Response, and Rep for Reproductive.



**Fig. 2.** Variances for 18 teosinte and maize landrace comparable traits. Proportions of phenotypic variance ( $V_p$ ) attributed to additive genetic variance ( $V_A$ ), dominance genetic variance ( $V_d$ ), and genetic-by-environment variance ( $V_{G \times E}$ ) are shown in A. Proportions of genetic variance ( $V_g$ ) attributed to additive genetic variance ( $V_A$ ) and dominance genetic variance ( $V_d$ ) are shown in B.

Reproductive traits showed the strongest depletion in  $h^2$  from teosinte to maize landrace while  $h^2$  in the Vegetative/Flowering Time group showed little difference between teosinte and maize landrace (Fig. 2A). There is not a consistent pattern for  $h^2$  within the Environmental Response trait group, since  $h^2$  for PROL and TILN were depleted in the maize landrace, while  $h^2$  for LBNN was increased in the maize landrace and  $h^2$  for LBLN and LBIL remained similar (Fig. 2A). Despite the overall depletion of  $h^2$  in maize, our maize landrace still possesses considerable amount of  $h^2$  for most traits (Fig. 2A).

While additive genetic variation is reduced in maize landrace relative to teosinte, the proportion of the genetic variance attributable to dominance effects is generally increased, suggesting that dominance genetic variance depletes more slowly than additive genetic variance during domestication (Fig. 2B and *SI Appendix, Tables S1 and S2*). We observed lower  $V_d/V_g$  in teosinte ( $V_d/V_g = 0.14 \pm 0.11$ , ranging from 0.04 to 0.36) than maize landrace ( $V_d/V_g = 0.29 \pm 0.26$ , ranging from 0.00 to 0.86). Among the trait groups, Reproductive showed the most increase in  $V_d/V_g$  from teosinte to maize landrace, while Vegetative/Flowering Time showed the least difference in  $V_d/V_g$  between teosinte and maize landrace. Within Environmental Response, large increases in  $V_d/V_g$  can be seen with TILN and PROL. Of all 18 traits, only four traits (PLHT, LFLN, LBNN, and EILN) showed a slight decrease in  $V_d/V_g$  from teosinte to maize landrace.

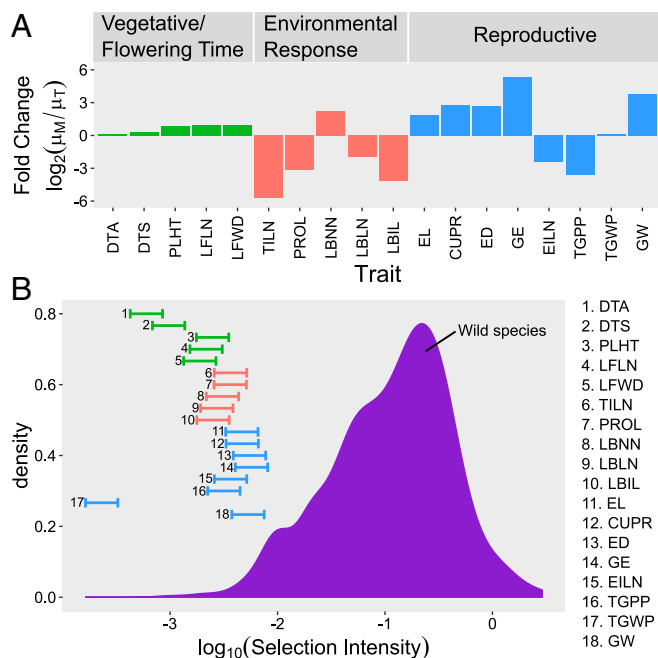
Maize landrace displayed a smaller proportion of the phenotypic variance attributable to genetic-by-environment interaction

( $V_{G \times E}/V_p$ ) in Vegetative/Flowering Time and Environmental Response traits but larger  $V_{G \times E}/V_p$  in Reproductive traits than teosinte, suggesting that genetic-by-environment interaction is also altered during domestication. Overall,  $V_{G \times E}/V_p$ , which is a  $V_p$ -standardized measure of the  $V_{G \times E}$ , is similar between teosinte ( $V_{G \times E}/V_p = 0.05 \pm 0.03$ , ranging from 0.01 to 0.09) and the maize landrace ( $V_{G \times E}/V_p = 0.04 \pm 0.02$ , ranging from 0.00 to 0.10). However, there are differences in  $V_{G \times E}/V_p$  between teosinte and the maize landrace among the trait groups. Teosinte has about 69% more  $V_{G \times E}/V_p$  for Vegetative/Flowering Time group and 65% more  $V_{G \times E}/V_p$  for Environmental Response group compared with the maize landrace. In contrast, teosinte exhibits about 49% less  $V_{G \times E}/V_p$  for Reproductive group compared with the maize landrace. Given that  $V_{G \times E}/V_p$  provides the necessary phenotypic plasticity for survival in diverse wild settings compared with more uniform farmland,  $V_{G \times E}/V_p$  is likely to be reduced during domestication. This expectation contradicts the greater  $V_{G \times E}/V_p$  for Reproductive traits in maize, an observation we explore further in *Discussion*.

Our estimates of the selection intensity ( $i$ ) during domestication are small, although highest among Reproductive traits except for TGWP. We computed the selection intensity for each trait and observed weak  $i$  across all traits (Fig. 3 and *SI Appendix, Table S3*). The magnitude of  $i$  ranges from 0.0002 to 0.0040, which equates to a selection differential of one phenotypic SD in about 250–6,000 generations. The difference in  $i$  among the traits suggests that some traits underwent stronger selection than others. Generally, the traits with high magnitude of  $i$  fall within the Reproductive group ( $|i| = 0.0022$ – $0.0040$ ) except for TGWP. Moderate magnitude of  $i$  are observed for the Environmental Response group ( $|i| = 0.0018$ – $0.0026$ ). The lowest magnitude of  $i$  is found for DTA, DTS, and TGWP ( $|i| = 0.0002$ – $0.0007$ ). These three traits were probably either indirectly selected or weakly selected at most. Overall,  $i$  correlates well with changes in  $h^2$ , suggesting that PROL, TILN, and all Reproductive traits (except TGWP) were key targets of the selection process.

**Multivariate Genetic Analyses.** Genetic correlations ( $r_g$ ) among traits indicate that the 18 traits fit the three predefined trait groups of Vegetative/Flowering Time, Environmental Response, and Reproductive such that genetic control of traits within the groups is due in part to shared sets of underlying genes in both teosinte and maize landrace. We constructed distance matrices as  $(1 - |r_g|)$  for teosinte and maize landrace and visualized the genetic relationships among traits using principal-coordinate analysis (PCoA) and neighbor-joining (NJ) trees (*SI Appendix, Fig. S1*). This analysis verified three groups of genetically related traits: Vegetative/Flowering Time, Environmental Response, and Reproductive. Traits within each group have shorter genetic distance (or higher absolute genetic correlation) than traits across groups. Each trait group is named after a common biological theme shared among the group members. For example, traits within the Vegetative/Flowering Time group are generally positively correlated, which translates to late flowering plants being taller and having bigger leaves. The Environmental Response group is composed of traits that are highly affected by environmental factors. For example, teosinte plants have many ears along long lateral branches under favorable conditions, but few ears on short branches in poor conditions.

Most genetic correlations among traits are conserved between teosinte and maize landrace with the exception for Reproductive traits, indicating that the underlying genetic network for Reproductive traits changed through domestication. Through an element-wise comparison of the genetic correlation matrices using the Mantel test (29), we showed that the genetic correlations ( $r_g$ ) for teosinte are overall correlated with those for maize landrace ( $r = 0.51$ ;  $P < 1 \times 10^{-4}$ ) (Fig. 4). The genetic correlations are even better preserved within the submatrices of each trait group. Strongest preservation of genetic correlations is observed within the Vegetative/Flowering Time group ( $r = 0.90$ ;  $P < 0.05$ ), followed by the Reproductive group ( $r = 0.79$ ;  $P < 0.01$ )



**Fig. 3.** Changes in trait means and selection intensities. In **A**, changes in trait mean ( $\mu$ ) are measured as  $\log_2(\mu_M/\mu_T)$  or fold change of maize landrace outcross mean ( $\mu_M$ ) over teosinte outcross mean ( $\mu_T$ ). In **B**, absolute selection intensities ( $|i|$ ) for 741 traits in wild species under natural selection are shown in a density plot [data from Kingsolver et al. (54)] and  $|i|$  for 18 traits in maize under domestication are shown in horizontal bars. Left end of the bars represents selection intensities estimated from 9,000 generations of selection, and right end represents selection intensities estimated from 4,500 generations of selection.

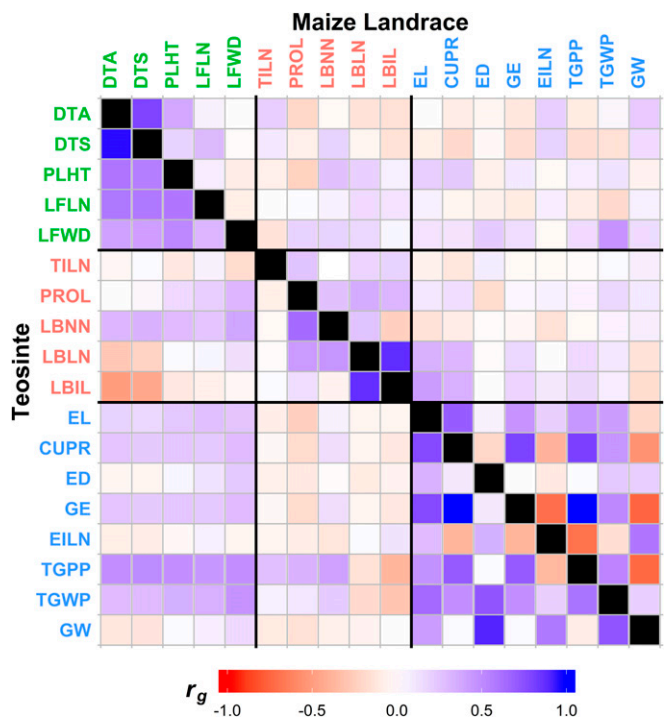
and the Environmental Response group ( $r = 0.77$ ;  $P < 0.05$ ). On the other hand, eigenstructure comparison of the teosinte and maize landrace genetic correlation matrices yielded slightly different results. The first two leading eigenvectors of the full genetic correlation matrices are  $89.6^\circ$  and  $88.4^\circ$  apart, respectively. Similar comparisons of the submatrices resulted in  $28.1^\circ$  and  $50.6^\circ$  for the Vegetative/Flowering Time submatrices,  $27.4^\circ$  and  $27.7^\circ$  for the Environmental Response submatrices, and  $87.8^\circ$  and  $86.7^\circ$  for the Reproductive submatrices. While some elements of genetic correlation matrices remained similar during domestication, the overall structure of the genetic correlation matrices did not, especially in the submatrix involving Reproductive traits.

An in-depth examination on the genetic correlation matrices revealed strong differences between teosinte and the maize landrace including change in sign of the correlations for some Reproductive traits, suggesting domestication modified the genetic networks underlying Reproductive trait variation. Out of the 153 genetic correlations, only 33 are significant and of the same sign in both teosinte and maize landrace. An additional 44 are nonsignificant in both maize and teosinte. There are 55 trait pairs significantly correlated in teosinte but not in maize landrace, and 17 are significant in maize landrace but not teosinte. Overall, teosinte shows more significant and stronger genetic correlations among traits than maize. Interestingly, there are four trait-pair correlations that are significant in both teosinte and maize landrace but with opposite signs, indicating a reversal during maize domestication. Two of the trait-pair correlations with opposite signs in teosinte and maize landrace are found within Reproductive traits, and the other two are between Reproductive and Vegetative/Flowering Time traits. Also, the genetic correlation between Reproductive and two other groups are stronger in teosinte but depleted in maize landrace.

The genetic variance-covariance ( $G$ ) matrices of teosinte and maize landrace are radically different, indicating that response to

selection in modern maize would be very different to that in teosinte. Although genetic correlations can influence evolution of multiple traits through indirect response to selection, it is the genetic variances and covariances that define the magnitude of the influence. We asked whether the  $G$ -matrices are conserved between teosinte and maize landrace. First, we tested for correlation between teosinte and the maize landrace  $G$ -matrices (*SI Appendix, Tables S4 and S5*) using the Mantel test (29) and found that the  $G$ -matrices are not significantly correlated ( $r = 0.03$ ;  $P = 0.21$ ). We tested for the structural similarity in teosinte and the maize landrace  $G$ -matrices using Flury hierarchy (*SI Appendix, Materials and Methods*) (30, 31) and found no similarity in structure, which suggests that teosinte and maize landrace  $G$ -matrices are completely unrelated and the structure of  $G$ -matrix diverged during domestication (*SI Appendix, Table S6*). We also tested for similarity in the predicted evolutionary responses from teosinte and maize landrace  $G$ -matrices using random skewers (*SI Appendix, Materials and Methods*) (32). We found that the overall predicted evolutionary responses are not significantly more correlated than random ( $r = 0.19$ ;  $P = 1.00$ ), again suggesting that teosinte and maize landrace  $G$ -matrices are quite different. Comparison of the teosinte and maize landrace  $G$ -matrices with a Bayesian approach (33) showed that the matrices are completely distinct ( $q = 0.00 \pm 0.00$ ).

The dissimilarity of  $G$ -matrices of teosinte and maize landrace is primarily due to changes in the submatrix for Reproductive traits, while the submatrix for Vegetative/Flowering Time traits is conserved. The random skewers analysis indicated there would be highly correlated predicted evolutionary responses for Vegetative/Flowering Time traits ( $r = 0.89$ ;  $P = 0.001$ ), but uncorrelated responses for Reproductive traits ( $r = 0.09$ ;  $P = 1.00$ ). Environmental Response traits are intermediate in this regard ( $r = 0.80$ ;  $P = 0.01$ ). Bayesian comparison analysis showed high similarity in the Vegetative/Flowering Time submatrices ( $q = 0.31 \pm 0.01$ )



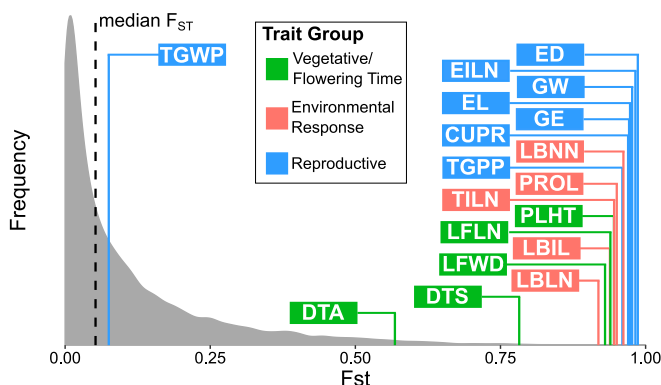
**Fig. 4.** Genetic correlations for 18 teosinte and maize landrace comparable traits. Genetic correlations for traits in teosinte are shown in the *Bottom Left* triangle of the matrix, and maize landrace is shown in the *Top Right* triangle of the matrix. Genetic correlations are colored according to the scale as shown in the *Bottom*. Values of the genetic correlations can be found in *SI Appendix, Tables S4 and S5*.

but no similarity in the Environmental Response ( $q = 0.00 \pm 0.00$ ) or Reproductive submatrices ( $q = 0.00 \pm 0.00$ ). These results again highlight that domestication has strongly modified the underlying genetic network for Reproductive traits such as GW and GE, but not that for flowering and plant size.

Change in the  $G$ -matrices of teosinte and maize landrace appears to be due to selection and not merely drift. We asked whether the difference in  $G$ -matrices can be solely attributed to neutral drift using the multivariate  $Q_{ST}$ - $F_{ST}$  test (34). The multivariate  $Q_{ST}$ - $F_{ST}$  test compares the proportionality of the between-population  $G$ -matrix ( $G_B$ ) to within-population  $G$ -matrix ( $G_W$ ). Under neutral trait evolution, the coefficient  $\rho_{ST}$  in  $G_B = \rho_{ST}G_W$  should be equal to  $2F_{ST}/(1 - F_{ST})$ , where  $F_{ST}$  is estimated from neutral loci (common SNP markers in this case); however, the coefficient calculated from the traits ( $\rho_{ST,G} = 314$ ; 95% CI, 190–908) is significantly higher than the expected coefficient calculated from neutral loci ( $\rho_{ST,N} = 0.372$ ; 95% CI, 0.363–0.381). Overall, we see that the teosinte and maize landrace  $G$ -matrices have very different structures and predicted evolutionary responses, and these differences cannot be explained by neutral drift alone.

Individual trait divergence between teosinte and maize landrace appears to have been driven primarily by selection with the exception of a single trait, total grain weight per plant (TGWP). We performed univariate  $Q_{ST}$ - $F_{ST}$  tests on the individual traits to ask whether neutral evolution can also be rejected at the individual trait level. In the univariate  $Q_{ST}$ - $F_{ST}$  test, the observed difference between  $Q_{ST}$  and  $F_{ST}$  is compared with a null distribution of  $Q_{ST}$ - $F_{ST}$  generated by parametric bootstrapping (SI Appendix, Materials and Methods) (35). The results of this analysis reject the neutral drift model for the differences in trait means between teosinte and maize landrace for all traits except TGWP, for which the population means were nearly equal (Fig. 5 and SI Appendix, Table S7).

Comparison between the trajectory of phenotypic change during domestication ( $Z$ ) and the teosinte genetic line of least resistance ( $g_{max,T}$ ) suggests that genetic correlations in teosinte posed considerable constraint during the early domestication process.  $Z$  defines the vector from the teosinte to the landrace centroids (trait means) through phenotypic space of 16 traits, while  $g_{max,T}$  is the first eigenvector of the teosinte genetic correlation matrix and explains the largest portion of the variance (27.2%) (SI Appendix, Fig. S2). While  $g_{max}$  is commonly calculated from  $G$ -matrix (36), we opted to calculate it from the scaled  $G$ -matrix such that each trait has a genetic variance of 1, which is essentially the genetic correlation matrix. Doing so allows us to calculate unbiased  $g_{max}$  since the first eigenvector can be skewed toward traits with highest magnitude. The  $Z$  vector is also scaled similarly by standardizing the trait mean differences by their genetic SDs. The degree of evolutionary constraint is measured by  $\theta_T$ , the angle between  $Z$



**Fig. 5.** Univariate  $Q_{ST}$ - $F_{ST}$  comparison for all 18 traits. The distribution of  $F_{ST}$  was estimated from 21,157 markers that are in common between teosinte and maize landrace. The  $Q_{ST}$  for each trait is shown as individual line along the horizontal axis and is colored according to trait group.

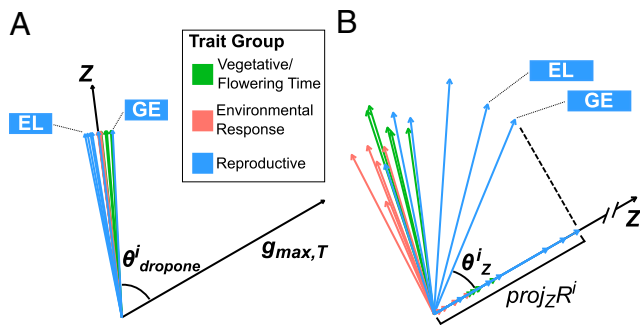
and  $g_{max,T}$ , where  $\theta_T$  ranges from  $0^\circ$  to  $90^\circ$ . Evolution is least constrained (small  $\theta_T$ ) when the multitrait selection response is closest to the direction of the maximum genetic variation, and the opposite is true (large  $\theta_T$ ) when the response is farthest to direction of the maximum genetic variation. Evolutionary constraint slows trait evolution since selection for improvement of one trait can be offset by decline in another due to a genetic correlation. Our estimate of  $\theta_T$  is  $67.3^\circ$ , which suggests that maize domestication proceeded despite strong constraint imposed by the  $G$ -matrix. Such constraint would slow progress toward the optimal phenotype and require a circuitous evolutionary path over the generations.

Because the genetic line of least resistance ( $g_{max,T}$ ) is informative only for the first eigenvector of the  $G$ -matrix, we also estimated the angles between observed trait responses to domestication ( $Z$ ) and each of the first five principal components of  $G$ , which accounted for 27.2, 18.8, 15.3, 9.7, and 7.3% of the variation (78.2% in total) (SI Appendix, Fig. S2). These angles were  $\theta_{T1} = 67.3^\circ, 79.4^\circ, 88.1^\circ, 54.1^\circ,$  and  $81.0^\circ$ , indicating moderate to strong constraint to the trajectory of domestication in the five most important directions of multivariate genetic variation.

In addition to the overall constraint seen between  $Z$  and  $g_{max,T}$ , there is variation in the degree to which different traits contribute to genetic constraint (Fig. 6A and SI Appendix, Table S8). By dropping one trait at a time and calculating the angle  $\theta_{dropone}^i$  between  $Z_i$  and  $g_{max,T,i}$ , we can estimate the genetic constraint from the remaining 15 traits after dropping trait  $i$ . If the genetic constraint decreases ( $\theta_{dropone}^i < \theta_T$ ) after dropping trait  $i$ , then it is said that trait  $i$  constrained evolution. However, if  $\theta_{dropone}^i > \theta_T$ , then it is said that trait  $i$  assisted evolution. Our results (Fig. 6A and SI Appendix, Table S8) showed that the response to selection was largely hindered by genetic correlations involving Vegetative/Flowering Time traits but assisted by the genetic correlations involving Reproductive traits.

The amount of constraint imposed by the  $G$ -matrix increased during the domestication process. To explore how evolutionary constraint changed during domestication, we also estimated  $\theta_M$ , the angle between the domestication trajectory ( $Z$ ) and the direction of maximum genetic variation in maize ( $g_{max,M}$ ), as a comparison with  $\theta_T$ .  $g_{max,M}$  explains 19.1% of the variance (SI Appendix, Fig. S2). Like teosinte, we also standardized the  $Z$  vector and  $G$ -matrix here for calculating  $\theta_M$ . We observed that  $\theta_M$  is  $74.3^\circ$ . The comparison of  $\theta_M$  and  $\theta_T$  indicates while there was substantial constraint early in domestication, it increased over time. It is not surprising to find  $\theta_M > \theta_T$  since the domestication process likely depleted variants that contributed beneficially to the structure of the  $G$ -matrix. Given that the  $G$ -matrix toward the end of domestication likely resembled the  $G$ -matrix of our maize landrace, trait evolution toward the optimum defined by the trait centroid for our maize landrace is likely to have slowed.

Selection on some individual traits would engender more rapid evolution in the direction of the domestication trajectory than others due to the structure of the  $G$ -matrix. Using the multivariate breeders' equation of  $G\beta = R$  (37), we can estimate the multivariate response ( $R$ ) based on teosinte  $G$ -matrix and hypothetical selection differentials ( $\beta$ ). We took a "what-if" approach to ask whether the individual traits differ in the extent that selection on any one trait would maximize gain along the evolutionary trajectory for all traits. We used 16 different  $\beta$  values, each  $\beta^i$  having one element of a value of 1 and 15 elements of a value of zero. The  $i$ th trait with a value of 1 in  $\beta^i$  would be directly selected while traits with a value of zero in  $\beta^i$  are only indirectly selected. If the traits are selected for lower value in maize, the corresponding elements in  $\beta^i$  would have a negative sign. Using the multivariate breeder's equation, we obtained  $R^i$  for each  $\beta^i$ . We then compared each  $R^i$  to the actual domestication trajectory ( $Z$ ) by measuring the angle ( $\theta_Z^i$ ) between  $R^i$  and  $Z$  and the scalar projection of  $R^i$  on  $Z$  ( $proj_Z R^i$ ) (Fig. 6B and SI Appendix, Table S9). While  $\theta_Z^i$  measures the deviation between



**Fig. 6.** Constraints and consequences of multivariate selection. Various possibilities due to multivariate selection in maize domestication are explored here. In *A*, individual trait contribution toward genetic constraint is identified by dropping *i*th trait from *Z* (actual domestication trajectory) and  $g_{max,T}$  (genetic lines of least resistance) and measuring the angle  $\theta_{dropone}^i$  between the two vectors. If  $\theta_{dropone}^i$  is smaller than  $\theta_T = 67.3^\circ$  (angle between *Z* and  $g_{max,T}$ ), then the *i*th trait is said to constrain evolution. If  $\theta_{dropone}^i$  is larger than  $\theta_T$ , then the *i*th trait is said to assist evolution. In *B*, multivariate response ( $R^i$ ) from hypothetical selection on a single *i*th trait is explored.  $R^i$  is compared with *Z* through the angle ( $\theta_z$ ) and scalar projection ( $|proj_Z R^i|$ ) of  $R^i$  on *Z* (*B*).  $\theta_z$  measures the deviation from *Z* by selecting on *i*th trait.  $|proj_Z R^i|$  measures the evolutionary gain along *Z* by selecting on *i*th trait.

$R^i$  and  $Z$ ,  $|proj_Z R^i|$  measures the gain along *Z* for each  $R^i$ . *Z* is also scaled similarly to the *G*-matrix, so that values of  $|proj_Z R^i|$  for different traits can be compared directly. Of all 16 traits, GE has the smallest  $\theta_z^i$  and largest  $|proj_Z R^i|$ , indicating that direct selection for more GE would give an overall response most closely aligned with the evolutionary trajectory as well as maximum gain along the trajectory. However, GE has the smallest  $\theta_{dropone}^i$ , which makes it unfavorable since selection for higher GE contributed to more genetic constraint (Fig. 6*A* and *SI Appendix, Table S8*). Of all traits, EL comes closest to GE in terms of  $\theta_z^i$  and  $|proj_Z R^i|$  with only one rank behind GE, and EL also has the largest  $\theta_{dropone}^i$ , which suggests that selection for higher EL contributed less than others to genetic constraint. Our results here suggest that if the ancient farmers were to domesticate teosinte by selecting for only a single trait, EL would be the ideal candidate as EL produced the maximum desired multivariate gains with the least genetic constraint.

## Discussion

**Change in Genetic Variances.** Our results demonstrate that maize landrace experienced a substantial loss in additive genetic variance, as measured by narrow-sense heritabilities ( $h^2$ ), during domestication compared with teosinte. The loss in  $h^2$  can be attributed to population bottleneck and selection during maize domestication (11, 38). Previous reports suggest that maize retained only ~80% of genomic nucleotide diversity and *cis* regulatory variation present in teosinte (13, 39). Given that selection likely brought beneficial alleles to fixation, it is not surprising to find an overall reduction in  $h^2$  in any domesticated relative of its wild progenitor.

Our estimates for the loss of  $h^2$  on individual traits are consistent with the morphological changes during maize domestication and how these morphological changes relate to the maize population's fitness as a crop. Modern maize typically possesses only a single stalk with few lateral branches that each carries a single large ear. This morphology was derived from that of teosinte, which is more branched and possesses many smaller ears. Given a strong reconfiguration of morphology during domestication, it is not surprising that the maize landrace is mostly depleted in  $h^2$  for traits such as PROL and TILN. Similarly, maize landrace has very low  $h^2$  for Reproductive traits, suggesting that many beneficial alleles for Reproductive traits were brought to

fixation during domestication. Being a wild plant, teosinte may maintain higher  $h^2$  than maize as the natural environment is apt to vary more across time and space than the cultivated field (40).

The observed high level of  $h^2$  in teosinte means that selection during domestication could have acted largely on standing variation in teosinte populations rather than depending on the occurrence of new mutation. Among the trait groups, Reproductive traits have the highest level of  $h^2$ , ranging from 0.27 to 0.73 (Fig. 2). Consequently, ancient farmers would have been able to capitalize on the abundance of standing variation to select a productive crop from a wild plant. Previous studies have also suggested that standing variation can lead to a faster evolution than new mutations (41, 42), thus highlighting the importance of standing variation in domestication.

Previously, Ladizinsky (43) and Lester (44) argued that a major force in domestication is the fixation of recessive loss-of-function alleles. There are several known genes that fit this model as exemplified by *nud* in barley (45), *Sh1* in sorghum (28), *Black hull4 (BH4)* in rice (46), and *ZmYAB2.1* in maize (47). However, there are also counterexamples in which the domestication alleles act additively and encode/express a functional protein, as seen for three genes in maize: *gal1* (24), *tb1* (23), and *gt1* (26). Intuitively, fixation of recessive loss-of-function alleles should deplete  $V_D/V_P$  because of reduction of  $V_D$  due to allele fixation. We observed no apparent depletion of  $V_D/V_P$  in maize landrace relative to teosinte (Fig. 2*A*); however,  $V_D/V_G$ , or the proportion of genetic variance due to dominance, increased for PROL, TILN, and Reproductive traits while remaining similar for others (Fig. 2*B*). The increase in  $V_D/V_G$  for those traits suggests that the additive genetic variance was depleted at a faster rate than the dominance genetic variance during domestication, congruent with the expected changes in  $V_D/V_G$  due to increase in frequency of recessive alleles (*SI Appendix, Fig. S3*) and as commonly observed in populations under selection (37). Overall, our data suggest that domestication was not biased toward fixation of recessive loss-of-function alleles.

Despite having a general depletion in additive genetic variance, the maize landrace still possesses some additive genetic variance that could potentially be useful for future crop improvement. Such variation may be preserved by multiple forces including mutation–selection balance (48), variable selection (49), and negative interactions between two alleles of different genes (2). As a result, maize landraces serve as a rich source of genetic diversity for breeding modern maize (12).

Additive genetic variation is important for long term adaptation, but it is genetic-by-environment (G×E) variation that is crucial for short-term adaptation to changing environments through phenotypic plasticity. Phenotypic plasticity is the ability of an individual genotype to express phenotypes differently as a response to environmental fluctuation (50, 51). Gage et al. (52) demonstrated that selection for increased yield in maize reduced G×E in favor of stability. This finding can be extended to selection during domestication. Unlike maize that is adapted to growth in relatively uniform agricultural fields, teosinte likely requires more G×E for adaptation to a broader range of environments in the wild. Our results showed higher G×E in teosinte than maize landrace, specifically in traits belonging to the Vegetative/Flowering Time and Environmental Response group.

However, there is an exception: teosinte possesses less G×E than maize landrace for traits that relate to ear and grain size. This observation could be explained as follows. Teosinte and maize may have evolved different strategies to cope with environmental fluctuations as related to reproductive traits. Teosinte adapts by varying the number of ears per plant while maintaining constant ear and grain size across environments. Teosinte may need to maintain an optimal grain size for seedling establishment. In contrast, maize produces a small and fixed number of large ears for easy harvest. Aborting an entire ear in poor environments when there are only two ears per plant places a heavy cost on reproduction. Thus, maize evolved to modulate grains per ear or grain weight instead of ear number.

**Strength of Selection During Domestication.** Crop domestication has been described as a process of slow evolution, where the selection force is similar or lower than selection force seen in natural selection (53), and maize domestication is no exception to that. Selection intensities for maize phenotypes during domestication are low ( $0.0001 < |i| < 0.0040$ ) and comparable to selection intensities in other domesticated crop species ( $|i| < 0.0040$ ) (53) as well as wild species under natural selection ( $0.001 < |i| < 2.911$ ) (54). In fact, selection intensities for all 18 traits fall on the lower end of the distribution compared with selection intensities for traits in wild species under natural selection (Fig. 3). This is true whether one calculates the average selection intensity over 9,000 or 4,500 generations of selection (Fig. 3 and *SI Appendix, Table S3*). Fugère and Hendry (55) recently showed that human intervention on trait evolution tends to result in weaker selection forces, which further strengthens the hypothesis that domestication is a process of slow evolution.

Our observation of weak selection intensities is consistent with the mild domestication bottleneck for maize and weak selection on known domestication genes. Evidence for weak selection on major domestication loci is suggested by the observation that some domestication alleles were not at fixation 5,000 y ago, 4,000 y after the domestication process started at 9000 BP (56). The archaeological record also shows slow and continuous change in ear traits over a 5,000-y period, suggesting a gradual rather than abrupt process by which maize became a productive food source (57, 58). Overall, a slow process by weak selection rather than abrupt changes by strong selection as previously suggested seems more likely (17, 59).

While selection intensities appear to have been weak overall during maize domestication, they vary among trait groups. We observed strongest selection intensities within the Reproductive group, followed by Environmental Response, and last, Vegetative/Flowering Time. Average selection intensity within the Reproductive group is more than twice that within the Vegetative/Flowering Time group. This result is consistent with the expectation that maize domestication focused on restructuring the architecture of the ear more than other aspects of the plant.

While selection intensity is highest in Reproductive traits, there is one exception—TGWP—which has the lowest selection intensity among all traits. Ancient farmers were successful in selecting for larger grains and ears (making harvesting seeds much easier), but not overall edible biomass as measured by TGWP, which remained the same between teosinte and maize landrace (Fig. 3 and *SI Appendix, Table S3*). TGWP may be constrained by the ability of the plant to convert solar energy into chemical energy. While domestication led to repartition of the solar energy captured from many small grains to fewer larger grains, TGWP itself was not modified. Considering that 9,000 y of selection may not have increased yield per plant, it is not too surprising to find that modern maize breeding has only been successful in increasing yield per area but not yield per plant (60).

**Conservation and Change in Genetic Correlations.** Moderate conservation between the teosinte and maize landrace genetic correlation matrices can be seen on a broad scale, with stronger conservation within Vegetative/Flowering Time and Environmental Response trait groups to little conservation within Reproductive trait group. Conservation of genetic correlations, in particular within the trait groups of Vegetative/Flowering Time and Environmental Response, suggests that genetic correlations can be resilient under selection. However, as seen within Reproductive trait group, it is possible to shift the genetic correlations under a sufficiently strong selection force. Change in allele frequency is expected whenever selection is present and may eventually lead to allele fixation. Naturally, genetic correlations diminish when alleles are fixed, as genetic correlations are measured from variation between a pair of traits. Consistent with Reproductive traits experiencing stronger selection than other domestication traits, we observe the largest change in genetic correlations within the Reproductive trait group.

While we observed both conservation and change in genetic correlations between teosinte and maize landrace, can we draw some inferences regarding the nature of the conservation and change? First, there is a general reduction in the magnitude of favorable correlations from teosinte to maize landrace (Fig. 4). This observation is consistent with the expectation that selection on correlated traits can drive favorable correlations to be less favorable (61–63). Second, the reduction in the magnitude of genetic correlations from teosinte to maize landrace is much stronger between trait groups than within them (Fig. 4). This difference suggests that domestication may have favored greater independence of trait groups, for example, disassociating correlations between reproductive traits and flowering time. Third, of many individual elements of the genetic correlation matrix, only four had a sign reversal between teosinte and maize landrace, and these correlations strictly involve Reproductive traits like ED, EL, CUPR, and GW (Fig. 4). When selection is strong, it is possible to reverse the sign of genetic correlations by increasing the contribution of unfavorable pleiotropy toward overall correlations through fixation of favorable pleiotropic alleles.

From a qualitative perspective, the genetic correlations in teosinte are largely in a favorable direction for domestication. Within the Reproductive group, we observed favorable genetic correlations among EL, CUPR, ED, GE, and GW that led to easier selection for larger ears. Favorable genetic correlations can also be seen among PLHT, LFLN, and LFWD, all of which led to easier selection for larger plants. In addition, PLHT, LFLN, and LFWD also displayed favorable genetic correlations with many Reproductive traits, which would have facilitated simultaneous selection for larger ears and plants, or in a broader sense, gigantism, a common feature of domestication (64).

**The G-Matrix and Genetic Constraint.** While there is some conservation in the genetic correlations, the differences between teosinte and maize landrace *G*-matrices are pronounced. We found no similarities in elements and structures of the teosinte and maize landrace *G*-matrices via Mantel test, Flury hierarchy, random skewers, and Bayesian estimation. Within the submatrices for each trait group, both random skewers and Bayesian estimation reaffirmed that the submatrices for Vegetative/Flowering Time and Environmental Response are conserved while the submatrix for Reproductive is not. These observations again suggest that Reproductive traits were the focus of selection during domestication. Overall, our findings corroborate the hypothesis that the *G*-matrix is not conserved over long-term evolution (65), which limit us in making further predictions on the evolution of maize from teosinte.

What do the structures of the teosinte and landrace *G*-matrices and the differences between them tell us about early and subsequent evolutionary potential and constraint during domestication? First, there appears to have been substantial evolutionary constraint imposed by the structure of the *G*-matrix at the initial phase of domestication as measured by the angle of 67.3° between the genetic lines of least resistance in teosinte ( $g_{max}$ ) and the axis along which teosinte evolved into maize. Unlike evolutionary stasis that brings evolution to a halt (66), evolutionary constraint still allows evolution to proceed, although at a slower rate (67). When an underlying genetic constraint (i.e., unfavorable genetic covariance) is present, evolution can proceed by changes in the frequencies of nonpleiotropic alleles (68) as has been previously shown by Marchini et al. (69). Furthermore, genetic constraint can be ameliorated over time by the decay of linkage disequilibrium between linked causative factors (68).

A comparison of the teosinte and maize landrace *G*-matrices also informs us that the degree of constraint increased over time. While the angle between the genetic lines of least resistance in teosinte ( $g_{max}$ ) and the axis along which teosinte evolved into maize is large (67.3°), indicating substantial constraint, the similar angle in maize landrace is 74.3°, indicating higher constraint. An increase in constraint is not unexpected in the case of maize domestication. Arnold (70) described the change in genetic constraint as a function of changes in selection and mutation/recombination,

which is essentially equivalent to mutation–selection balance. Increase in constraint from teosinte to maize suggests that domestication reduced advantageous genetic variances and covariances more rapidly than mutation could restore them.

Another way to think about the relationship between the  $G$ -matrix and the tempo of evolution comes from the multivariate breeder's equation of  $R = G\beta$ . Here, the evolution of maize ( $R$ ) during domestication is a compromise between genetic constraints ( $G$ ) and humans' needs ( $\beta$ ). Even though evolution along the genetic lines of least resistance ( $g_{max}$ ) provides the maximum evolutionary gain for multiple traits, if  $g_{max}$  does not align with human desires ( $\beta$ ), then the response ( $R$ ) will be attenuated. We have shown that slow trait evolution during the domestication of maize is likely due to the strong genetic constraint that increased over time from teosinte to maize landrace.

**Causes and Consequences.** Results from both univariate and multivariate  $Q_{ST}$ – $F_{ST}$  tests suggested that selection drove the change in trait values between teosinte and maize landrace.  $Q_{ST}$  measures genetic differentiation at trait level, while  $F_{ST}$  measures genetic differentiation at neutral loci. Under neutral evolution, traits are expected to evolve at the same pace as neutral loci, which is essentially  $Q_{ST} = F_{ST}$  (71). The multivariate  $Q_{ST}$ – $F_{ST}$  test indicates that the 18 traits as a whole underwent nonneutral evolution; however, further dissection of individual traits using the univariate  $Q_{ST}$ – $F_{ST}$  test reveals that one trait, TGWP, conforms to neutral expectations. This result implies that selection was not effective at increasing total yield. Maize domestication can be seen as having mostly improved harvestability by creating a plant with few ears with more and larger grains.

Our “what-if” analysis on the multivariate selection response revealed potential initial targets that would maximize evolutionary gain along the domestication trajectory for all traits. We compared the multivariate responses ( $R^i$ ) from selecting ( $\beta^i$ ) for a single trait to the actual evolutionary trajectory ( $Z$ ) by measuring the angle ( $\theta_z^i$ ) between  $R^i$  and  $Z$  and the scalar projection of  $R^i$  on  $Z$  ( $|proj_z R^i|$ ) (Fig. 6 and *SI Appendix, Table S9*).  $\theta_z^i$  measures how close is  $R^i$  to  $Z$ , while  $|proj_z R^i|$  measures the contribution of  $R^i$  toward evolving along  $Z$ . Of all traits tested, EL is likely the most ideal option with small  $\theta_z^i$ , large  $|proj_z R^i|$ , and large  $\hat{\sigma}_{dropone}^i$ . This result suggests that directional selection on EL alone would yield the most gain and be closest to the evolutionary trajectory with minimal constraint.

**Caveats.** Caution is essential when interpreting our results and generalizing from them based on many limitations of our data and analyses. (i) Our single teosinte and maize landrace populations are not apt to be adequate proxies for the ancestral teosinte and resulting maize domesticate. (ii) Our field trials were performed in Florida, which is a different environment from the Balsas river drainage in Mexico where maize was domesticated. (iii) All of our parameter estimates are based on the start (teosinte) and end (maize landrace) without knowing the exact changes in each generation, so our estimates only represent an average over many generations, and it is likely that selection, genetic and residual covariances, and response to selection were not consistent among all generations. (iv) Inferences drawn from the genetic correlation and  $G$ -matrices are sensitive to the traits included. While we strived to be comprehensive in analyzing all possible domestication traits, there are several obvious domestication traits like kernel row number, ear shattering, and glume architecture that could not be studied here due to the lack of variation in either teosinte or maize landrace. Despite these limitations, we hope our approach will at least promote further thinking and study about domestication and how it was constrained by the genetic architecture of the ancestor and how the genetic architecture of crop species evolved over time.

## Conclusions

Based on our overall results, we propose the following inferences about maize domestication. First, selection was clearly focused on ear architecture since most of the traits that define ear architecture suffered a loss of heritable variation and ear morphology changed dramatically. Second, the difference in plant architecture is likely a consequence or correlated response to selection on ear architecture to create a less branched plant with fewer, larger ears. Third, because of favorable genetic correlations among most selection targets, ancient farmers had multiple options to select for what a maize ear would become. For instance, EL is positively correlated to GE in teosinte, so selection could have been applied to either trait or even both traits concurrently to speed up the process. Fortunately, the relative differences among plants for these traits are easily observable by eye and do not require modern tools for selection. Last, evolution can proceed despite substantial genetic constraint since selection can be applied on traits that pose the least constraint.

## Materials and Methods

**Sample Population and Crossing Strategy.** We sampled a single population of teosinte (*Zea mays ssp. parviglumis*) and a single population of maize landrace Tuxpeño (*Zea mays ssp. mays*) from the Balsas region of Mexico as proxies for the founding teosinte population and first domesticated maize population, respectively. The teosinte population is chosen from individuals in the “Mound” population near Palmar Chico, Mexico (latitude, 18.6403°; longitude, –100.3570°; altitude, 1,008 m) that were previously sampled by van Heerwaarden et al. (72). The maize landrace population was chosen from individuals in a site less than 1 km away from the teosinte population (University of Guadalajara collection J5G-RMM-LCL-529; latitude, 18.6483°; longitude, –100.3542°; altitude, 983 m). We constructed our study populations by selfing and intermating 70 teosinte parent plants to obtain progeny with various allelic combinations from different parents. We also applied similar crossing scheme to 55 maize landrace parent plants. We obtained 49 selfed and 377 outcross families for teosinte (*SI Appendix, Table S10*), and 34 selfed and 89 outcross families for maize landrace (*SI Appendix, Table S11*).

**Field Design.** We conducted field evaluations for the teosinte and maize landrace progeny over two winter seasons (2013 and 2014) under short daylength (<12 h) in Homestead, Florida (latitude, 25.5044°; longitude, –80.5045°; altitude, 3 m). Within each season, we planted the seeds in a randomized design and grid of 100 plants by 30 rows along with borders surrounding the experimental section. Each plant was separated by 30 cm within rows and 76 cm between rows. As an exception due to poor germination of teosinte seeds, we sampled some of the border plants in the first season and increased the grid to 112 plants by 54 rows in the second season.

**Tissue Collection, DNA Isolation, and Genotyping.** DNAs from the parents were isolated using either DNeasy Plant Mini Kit (Qiagen) or modified CTAB protocol (73). DNAs from the progeny were isolated using DNeasy 96 Plant Kit (Qiagen). All DNA samples from the parents and progeny were genotyped using genotype-by-sequencing (GBS) (74). As per GBS protocol, all DNA samples were digested using ApeKI restriction enzyme and sequenced in 96-plex on Illumina HiSeq 2000, SE 1 × 100 bp (Illumina). Following that, genotypes were called from GBS raw sequencing reads using the TASSEL5-GBS Production Pipeline based on 955,690 SNPs in the ZeaGBSv2.7 Production TagsOnPhysicalMap (TOPM) file (75). The overall genotyping process from raw sequencing reads to final, clean, and imputed GBS dataset is highlighted in a flowchart (*SI Appendix, Fig. S4*). Final GBS dataset has been deposited in the figshare database (76).

**Phenotyping.** We collected phenotypic trait data for a total of 18 traits from 4,455 teosinte plants and 4,398 maize landrace plants. Descriptions and methods of measurement for each trait are summarized in Table 1 and *SI Appendix, Table S12*.

**Parentage Inference.** We inferred the parentage of each progeny for both teosinte and maize landrace by first estimating the realized additive genomic relationship matrix ( $A$ -matrix) (77, 78) in TASSEL5 (79) from the unfiltered raw SNP calls (including many missing calls) based on GBS data. We then identified the parents of each progeny from the  $A$ -matrix using a custom R script. A progeny is considered a self-fertilized offspring of a parent if there is only one unambiguous progeny–parent pair with a high additive re-



relationship coefficient; a progeny is considered an outcross of two parents if there are two unambiguous progeny–parent pairs with high and similar additive relationship coefficients.

**GBS Data Imputation.** Before imputation, we used the CrossMap (80) software to convert the GBS SNP positions from maize B73 reference AGPv2 coordinates to AGPv4 coordinates. Then, we applied the following filter on the raw GBS data: (i) no sites with minor allele frequency (maf) below 0.001, (ii) no sites with more than 20% missing data, (iii) no sites that are nonbiallelic, and (iv) no sites with insertion–deletion (indel) polymorphism. We imputed the GBS data for teosinte and maize landrace using the ParentPhasingPlugin and ImputeProgenyStatesPlugin as implemented in TASSEL5 (79). Briefly, the imputation is a two-part process with first identifying the parent phase using a heuristic approach, followed by inferring the parent state at each site using a hidden Markov model. Additional details on imputation and quality check are described in *SI Appendix, Materials and Methods* and Fig. S5.

**Univariate Analysis.** We fitted a common univariate linear mixed model for each trait using ASReml, version 4 (81), which implements restricted maximum-likelihood (REML) estimation of model parameters. Briefly, among the fixed effects in the model are year, inbreeding coefficient, shading (*SI Appendix, Fig. S6*), and field positions; among the random effects in the model are polygenic additive, dominance, and genetic-by-environment effects. The covariances of additive and dominance effects of each individual were modeled to be proportional to the realized additive and dominance relationship matrices, respectively. The covariances of genotype-by-environment effects were modeled to be proportional to the additive relationships of individuals tested in a common year and zero for pairs of individuals tested in different years. The full model is shown in *SI Appendix, Materials and Methods*. Based on the model's variance component estimates, we were able to estimate the proportions of phenotypic variance due to additive genetic variance (heritability), dominance genetic variance, and genetic-by-environment variance.

We estimated the selection intensity ( $i$ ) for each trait using results from the univariate analysis. We derived the following formula from the univariate breeder's equation (82) to estimate  $i$ :

$$i = \frac{Z_N - Z_0}{\sum_{n=0}^{N-1} \frac{V_{A,0} + n\Delta V_A}{\sqrt{V_{P,0} + n\Delta V_P}}}$$

The full derivation and notation for this formula can be found in *SI Appendix, Materials and Methods*. There are several assumptions in our estimate: (i) constant selection intensity; (ii) constant change in additive genetic variance ( $V_A$ ) and phenotypic variance ( $V_P$ ); and (iii) teosinte  $V_A$  and  $V_P$  as initial variances and maize landrace  $V_A$  and  $V_P$  as final variances. While these assumptions can be naive, they represent our best model for estimating  $i$ , unlike models proposed in the literature that assume constant environmental variance over generations (83–86).

**Multivariate Analysis.** We fitted a common bivariate linear mixed model for each pair of traits using ASReml, version 4 (81). All of the fixed and random effects were the same as its univariate counterpart, except that polygenic dominance and genetic-by-environment effects are removed as a compromise to computational speed. The full model is shown in *SI Appendix, Materials and Methods*. Based on the model's variance and covariance component estimates, we were able to estimate additive genetic correlation.

We confirmed the genetic relationship among the 18 traits based on the distance matrix of these traits. We calculated the distance matrices for teosinte and maize landrace as  $1 - |r_g|$ , where  $r_g$  is the genetic correlation. We performed PCoA and constructed NJ trees (*SI Appendix, Fig. S1*) to visualize the distance matrices.

We tested for similarity between teosinte and maize landrace genetic correlation matrices ( $r_g$ ) using Mantel's test (29). Mantel's test tests the null hypothesis that two matrices are not correlated. Similarly, we also applied Mantel's test to each trait group to see how each trait group similarity compares to overall similarity. The Mantel's test is performed using *mantel.test* function with 10,000 permutations implemented in the package "ape" (87) in R (88). We also checked for the similarity in genetic correlation matrices by comparing the angle between the first two leading eigenvectors of the matrices, as well as the three submatrices defined by the trait groups.

We also tested for similarity between teosinte and maize landrace  $G$ -matrices (*SI Appendix, Tables S4 and S5*) using Mantel's test (29), Flury hierarchy (30), random skewers (32), Bayesian estimation (33), and multivariate  $Q_{ST}$ – $F_{ST}$  test (34). Flury hierarchy tests for structural similarity between two matrices by comparing their eigenvectors and eigenvalues in a hierarchical way (30). Flury hierarchy is implemented in the common principal component (CPC) software (31). Random skewers utilizes the multivariate breeder's equation of  $R = G\beta$  by comparing  $R$  values calculated from two  $G$ -matrices and simulated  $\beta$ , where two  $G$ -matrices are considered similar if their  $\Delta z$  values are significantly correlated. This test is performed using *skewers* function with 1,000 simulations implemented in the "phytools" package (89) in R (88). The Bayesian estimation compares the probabilities of two matrices being derived from the samples of multivariate normal distribution that is specified by one of the matrices (33). Multivariate  $Q_{ST}$ – $F_{ST}$  test compares the genetic differentiation of two or more populations at multiple-trait level ( $Q_{ST}$ ) to genetic differentiation at neutral loci ( $F_{ST}$ ) where  $Q_{ST} = F_{ST}$  under the null hypothesis (34, 90). This test is performed using R scripts provided by Martin et al. (34). Similarly, the univariate  $Q_{ST}$ – $F_{ST}$  test is also performed for each trait using R scripts provided by Whitlock and Guillaume (35).

We quantified the genetic constraint from the  $G$ -matrix using the angle ( $\theta$ ) between genetic lines of least resistance ( $g_{max}$ ) (36) and actual domestication trajectory ( $Z$ ). To calculate  $\theta$ , we used the following formula:  $\theta = \cos^{-1}(\hat{Z} \cdot g_{max})$ , where  $\theta$  ranges from  $0^\circ$  to  $90^\circ$ .  $Z$  is a vector of differences in outcross trait means between teosinte and maize landrace.  $g_{max}$  is the first eigenvector of the  $G$ -matrix, which is taken from eigendecomposition of the  $G$ -matrix using *eigen* function in R (88). Since the eigenvectors are sensitive to the magnitude of each trait, we opted to standardize the  $G$ -matrix and  $Z$  such that each trait has a genetic SD or variance of 1. We computed  $\theta$  for both  $g_{max}$  values from teosinte and maize landrace  $G$ -matrices to identify the genetic constraints at the beginning and ending points of domestication. To compare individual trait contribution to genetic constraint, we calculated the angle  $\theta_{dropone}^i$  between  $Z$  and  $g_{max}$  for every  $i$ th trait that is dropped from  $Z$  and  $g_{max}$ .

Using the multivariate breeder's equation again, we took a "what-if" approach to identify potential selection targets during domestication. For this approach, we simulated 18 unique  $\beta$ , where each  $\beta^i$  had only a single element with a value of 1 and the remaining elements with a value of zero. We multiplied  $G$  by each  $\beta^i$  to obtain  $R^i$ , and we compared each  $R^i$  to  $Z$  by calculating the angle  $\theta_z$  between  $R^i$  and  $Z$  and the scalar projection of  $R$  on  $Z$  ( $|\text{proj}_Z R^i|$ ).  $\theta_z$  measures the deviation in direction from  $Z$ , while  $|\text{proj}_Z R^i|$  measures the amount of evolutionary gain along  $Z$ .

**Data and Code Availability.** Additional details on the materials and methods and all R scripts used in this project can be found in *SI Appendix, Materials and Methods*. Genotype files are available in Figshare, and phenotype files are available in *Datasets S1* and *S2*.

**ACKNOWLEDGMENTS.** We thank various members of the J.F.D. laboratory, J.B.H. laboratory, and Panzea, especially Adam Mittermaier, Eric Rentmeester, and Jason Brewer, for their assistance in this project. This work is supported by National Science Foundation Grant IOS 1238014.

- Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, London).
- Meyer RS, Purugganan MD (2013) Evolution of crop species: Genetics of domestication and diversification. *Nat Rev Genet* 14:840–852.
- Zeder MA, Bradley DG, Emshwiller E, Smith BD (2006) *Documenting Domestication: New Genetic and Archaeological Paradigms* (Univ of California Press, Berkeley, CA).
- Fuller DQ, et al. (2014) Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proc Natl Acad Sci USA* 111: 6147–6152.
- Gross BL, Olsen KM (2010) Genetic perspectives on crop domestication. *Trends Plant Sci* 15:529–537.
- Olsen KM, Wendel JF (2013) A bountiful harvest: Genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47–70.
- Larson G, et al. (2014) Current perspectives and the future of domestication studies. *Proc Natl Acad Sci USA* 111:6139–6146.
- Hake S, Ross-Ibarra J (2015) The natural history of model organisms: Genetic, evolutionary and plant breeding insights from the domestication of maize. *eLife* 4:e05861.
- Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080–6084.
- Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106:5019–5024.
- Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: A multilocus investigation of the domestication process in maize. *Mol Biol Evol* 21:1214–1225.
- Yamasaki M, Wright SI, McMullen MD (2007) Genomic screening for artificial selection during domestication and improvement in maize. *Ann Bot* 100:967–973.

13. Hufford MB, et al. (2012) Comparative population genomics of maize domestication and improvement. *Nat Genet* 44:808–811.
14. Hufford MB, et al. (2013) The genomic signature of crop-wild introgression in maize. *PLoS Genet* 9:e1003477.
15. Vigouroux Y, et al. (2002) Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proc Natl Acad Sci USA* 99:9650–9655.
16. Warburton ML, et al. (2008) Genetic diversity in CIMMYT nontemperate maize germplasm: Landraces, open pollinated varieties, and inbred lines. *Crop Sci* 48:617–624.
17. Iltis HH (1983) From teosinte to maize: The catastrophic sexual transmutation. *Science* 222:886–894.
18. Iltis HH (2000) Homeotic sexual translocations and the origin of maize (*Zea mays*, Poaceae): A new look at an old problem. *Econ Bot* 54:7–42.
19. Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte  $F_2$  population: Implications for the origin of maize. *Proc Natl Acad Sci USA* 87:9888–9892.
20. Doebley J (2004) The genetics of maize evolution. *Annu Rev Genet* 38:37–59.
21. Briggs WH, McMullen MD, Gaut BS, Doebley J (2007) Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* 177:1915–1928.
22. Shannon LM (2012) The genetic architecture of maize domestication and range expansion. PhD dissertation (University of Wisconsin–Madison, Madison, WI).
23. Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43:1160–1163.
24. Wang H, et al. (2005) The origin of the naked grains of maize. *Nature* 436:714–719.
25. Wang H, Studer AJ, Zhao Q, Meeley R, Doebley JF (2015) Evidence that the origin of naked kernels during maize domestication was caused by a single amino acid substitution in *tg1*. *Genetics* 200:965–974.
26. Wills DM, et al. (2013) From many, one: Genetic control of prolificacy during maize domestication. *PLoS Genet* 9:e1003604.
27. Dong Z, et al. (2017) Ideal crop plant architecture is mediated by *tassels replace upper ears1*, a BTB/POZ ankyrin repeat gene directly targeted by TEOSINTE BRANCHED1. *Proc Natl Acad Sci USA* 114:E8656–E8664.
28. Lin Z, et al. (2012) Parallel domestication of the *Shattering1* genes in cereals. *Nat Genet* 44:720–724.
29. Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220.
30. Flury B (1988) *Common Principal Components and Related Multivariate Models* (Wiley, New York).
31. Phillips PC, Arnold SJ (1999) Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury Hierarchy. *Evolution* 53:1506–1515.
32. Cheverud JM, Marroig G (2007) Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genet Mol Biol* 30:461–469.
33. Ovaskainen O, Cano JM, Merilä J (2008) A Bayesian framework for comparative quantitative genetics. *Proc Biol Sci* 275:669–678.
34. Martin G, Chapuis E, Goudet J (2008) Multivariate  $Q_{ST}$ – $F_{ST}$  comparisons: A neutrality test for the evolution of the g matrix in structured populations. *Genetics* 180:2135–2149.
35. Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: Comparing  $Q_{ST}$  to  $F_{ST}$ . *Genetics* 183:1055–1063.
36. Schluter D (1996) Adaptive radiation along genetic lines of least resistance. *Evolution* 50:1766–1774.
37. Walsh B, Lynch M (2018) *Evolution and Selection of Quantitative Traits* (Oxford Univ Press, Oxford).
38. Eyre-Walker A, Keightley PD (2009) Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol Biol Evol* 26:2097–2108.
39. Lemmon ZH, Bukowski R, Sun Q, Doebley JF (2014) The role of *cis* regulatory evolution in maize domestication. *PLoS Genet* 10:e1004745.
40. Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50:434–437.
41. Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends Ecol Evol* 23:38–44.
42. Matuszewski S, Hermisson J, Kopp M (2015) Catch me if you can: Adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics* 200:1255–1274.
43. Ladizinsky G (1985) Founder effect in crop-plant evolution. *Econ Bot* 39:191–199.
44. Lester RN (1989) Evolution under domestication involving disturbance of genic balance. *Euphytica* 44:125–132.
45. Taketa S, et al. (2008) Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc Natl Acad Sci USA* 105:4062–4067.
46. Zhu B-F, et al. (2011) Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant Physiol* 155:1301–1311.
47. Yang CJ, et al. (2016) A gene for genetic background in *Zea mays*: Fine-mapping enhancer of *teosinte branched1.2* to a YABBY class transcription factor. *Genetics* 204:1573–1585.
48. Crow JF, Kimura M (1970) *An Introduction to Population Genetics Theory* (Burgess Publishing Company, Minneapolis, MN).
49. Felsenstein J (1976) The theoretical population genetics of variable selection and migration. *Annu Rev Genet* 10:253–280.
50. Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. *J Exp Biol* 209:2362–2367.
51. Des Marais DL, Hernandez KM, Juenger TE (2013) Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. *Annu Rev Ecol Evol Syst* 44:5–29.
52. Gage JL, et al. (2017) The effect of artificial selection on phenotypic plasticity in maize. *Nat Commun* 8:1348.
53. Purugganan MD, Fuller DQ (2011) Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* 65:171–183.
54. Kingsolver JG, et al. (2001) The strength of phenotypic selection in natural populations. *Am Nat* 157:245–261.
55. Fugère V, Hendry AP (2018) Human influences on the strength of phenotypic selection. *Proc Natl Acad Sci USA* 115:10070–10075.
56. Vallebueno-Estrada M, et al. (2016) The earliest maize from San Marcos Tehuacán is a partial domesticate with genomic evidence of inbreeding. *Proc Natl Acad Sci USA* 113:14151–14156.
57. Benz BF, Cheng L, Leavitt SW, Eastoe C (2006) El Riego and early maize agricultural evolution. *Histories of Maize*, eds Staller J, Tykot R, Benz B (Academic, San Diego), pp 73–82.
58. Kennett DJ, et al. (2017) High-precision chronology for Central American maize diversification from El Gigante rockshelter, Honduras. *Proc Natl Acad Sci USA* 114:9026–9031.
59. Innan H, Kim Y (2004) Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci USA* 101:10667–10672.
60. Duvick DN (2005) The contribution of breeding to yield advances in maize (*Zea mays* L.). *Adv Agron* 86:83–145.
61. Itoh Y (1991) Changes in genetic correlations by index selection. *Genet Sel Evol* 23:301–308.
62. Strandén I, Mäntysaari EA, Mäki-Tanila A (1993) Change in genetic correlation due to selection using animal model evaluation. *J Anim Breed Genet* 110:412–422.
63. McMillan I, Fairfull RW, Friars GW (1995) The effect of simultaneous selection on the genetic correlation. *Theor Appl Genet* 91:776–779.
64. Harlan JR (1992) *Crops & Man* (American Society of Agronomy, Madison, WI).
65. Agrawal AF, Brodie ED, 3rd, Rieseberg LH (2001) Possible consequences of genes of major effect: Transient changes in the G-matrix. *Genetica* 112-113:33–43.
66. Burt DB (2001) Evolutionary stasis, constraint and other terminology describing evolutionary patterns. *Biol J Linn Soc Lond* 72:509–517.
67. Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science* 294:151–154.
68. Hoffmann AA (2014) Evolutionary limits and constraints. *The Princeton Guide to Evolution*, ed Losos JB (Princeton Univ Press, Princeton), pp 247–252.
69. Marchini M, et al. (2014) Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals. *BMC Evol Biol* 14:258.
70. Arnold SJ (1992) Constraints on phenotypic evolution. *Am Nat* 140:S85–S107.
71. Spitze K (1993) Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics* 135:367–374.
72. Van Heerwaarden J, et al. (2010) Fine scale genetic structure in the wild ancestor of maize (*Zea mays* ssp. *parviglumis*). *Mol Ecol* 19:1162–1173.
73. CIMMYT (2005) *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory* (CIMMYT, El Batán, Mexico), 3rd Ed.
74. Elshire RJ, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379.
75. Glaubitz JC, et al. (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346.
76. Yang CJ, et al. (2019) Data from “The genetic architecture of teosinte catalyzed and constrained maize domestication.” Figshare. Available at <https://doi.org/10.6084/m9.figshare.7655588>. Deposited February 24, 2019.
77. VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91:4414–4423.
78. Endelman JB, Jannink JL (2012) Shrinkage estimation of the realized relationship matrix. *G3 (Bethesda)* 2:1405–1413.
79. Bradbury PJ, et al. (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635.
80. Zhao H, et al. (2014) CrossMap: A versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* 30:1006–1007.
81. Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thomson R (2015) *ASReml User Guide, Release 4.1, Structural Specification* (VSN International, Hemel Hempstead, UK).
82. Lush JL (1937) *Animal Breeding Plans* (Iowa State College Press, Ames, IA).
83. Bulmer MG (1971) The effect of selection on genetic variability. *Am Nat* 105:201–211.
84. Verrier E, Colleau JJ, Foulley JL (1990) Predicting cumulated response to directional selection in finite panmictic populations. *Theor Appl Genet* 79:833–840.
85. Hospital F, Chevalet C (1993) Effects of population size and linkage on optimal selection intensity. *Theor Appl Genet* 86:775–780.
86. Roff DA (1997) *Evolutionary Quantitative Genetics* (Chapman and Hall, New York).
87. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
88. R Core Team (2018) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna), Version 3.5.1. Available at <https://www.R-project.org>. Accessed July 02, 2018.
89. Revell LJ (2012) An R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3:217–223.
90. Leinonen T, McCairns RJ, O’Hara RB, Merilä J (2013)  $Q_{ST}$ – $F_{ST}$  comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nat Rev Genet* 14:179–190.