PNAS
www.pnas.org

Supplementary Information for

The genetic architecture of teosinte catalyzed and constrained maize domestication.

Chin Jian Yang, Luis Fernando Samayoa, Peter J. Bradbury, Bode A. Olukolu, Wei Xue, Alessandra M. York, Michael R. Tuholski, Weidong Wang, Lora L. Daskalska, Michael A. Neumeyer, Jose de Jesus Sanchez-Gonzalez, Maria Cinta Romay, Jeffrey C. Glaubitz, Qi Sun, Edward S. Buckler, James B. Holland, John F. Doebley.

Corresponding Author: John F. Doebley
Email: jdoebley@wisc.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1820997116

SUPPLEMENTARY TEXT

Materials and Methods

Sample Population and Crossing Strategy

We sampled a single teosinte (*Zea mays* ssp. *parviglumis*) population from the Balsas region of Mexico as a proxy for the founding population that gave rise to the first domesticated maize 10,000 years ago. This teosinte population is chosen from individuals in the "Mound" population near Palmar Chico, Mexico (lat. 18.6403˚, long. -100.3570˚, altitude 1008 m) that were previously sampled by van Heerwaarden *et al*. (2010). Initially, we grew 50 teosinte parent plants under short daylength (11 h of daylight) in our greenhouse. These 50 teosinte parents were each grown from a single seed collected from separate plants in Palmar Chico. Our initial strategy involved selfing and intermating these 50 parents to produce progeny with various allelic combinations from the different parents. Given the structure of teosinte flowers, controlled pollinations are difficult to make. Thus, we collected pollen from individual plants and applied it randomly to as many plants as we could, including itself. Parentage of the resulting progeny was determined with molecular markers (see below). Due to inadequate short-day induction, 20 parents produced only pollen but not silks. To correct this deficit, we repeated the process with 20 additional parents, of which 18 are half-sibs to the original 20 parent that produced only pollen. The flowering induction for the second set of parents was successful. Overall, we obtained self-fertilized ("selfed") families for 49 out of 70 parents with family size ranging from 3 to 95 progeny. We also obtained 288 outcross families out of 1415 possibilities with family size ranging from 1 to 75 progeny. Exact numbers of progeny in each teosinte family are summarized in *SI Appendix*, Table S10.

Similarly, we also sampled a single population of the maize landrace Tuxpeño (*Zea mays* ssp. *mays*) from a site less than one km from the teosinte population (University of Guadalajara collection JSG-RMM-LCL-529, lat. 18.6483˚, long. -100.3542˚, altitude 983 m) as a proxy for the first domesticated maize from the region where maize was domesticated. We grew 55 maize plants under short daylength (11 h of daylight) in a winter nursery. Unlike the teosinte crossing strategy, we divided the 55 parents into 11 groups of five and applied bulked pollen samples from within each group to female flowers of the same five. This process provided us with a mix of selfed and outcross plants similar to that obtained for teosinte. For the selfed families, we obtained 34 out of 55 possibilities with family size ranging from 1 to 125 progeny. For the outcross families, we obtained 55 out of 110 possibilities with family size ranging from 6 to 141 progeny. Of the 55 plants, 40 contributed to at least one progeny as male or female parent. Exact numbers of progeny in each maize landrace family are summarized in *SI Appendix*, Table S11.

Field Design

We conducted field evaluations for the teosinte progeny over two winter seasons (2013 and 2014) under short daylength in Homestead, FL (lat. 25.5044°, long. -80.5045˚, altitude 3 m, daylengths < 12 h for the entire growing season). Within each year, a randomized design was used, with individual plants as experimental units. We planted the seeds in a grid of 100 plants by 30 rows in the first season and 112 plants by 54 rows in the second season. Each plant was separated by 30 cm within rows and 76 cm between rows. Using the same seed source as the experimental plants, we also planted the borders surrounding the experimental plants. The borders included a row before the

first experimental row, a row after the last experimental row, rows between experimental rows and irrigation alleys, five plants before the start of each experimental row and five plants after the end of each experimental row. Each plant was individually staked and tagged to keep track of the plant identity for tissue and phenotypic data collection. Due to the lower than expected germination rate of teosinte seeds, we did not obtain sufficient plants in the main experimental plots during the first season and so we sampled some of the border plants as well.

We also conducted field evaluation for the maize landrace progeny alongside with the teosinte field evaluation in a similar design. We planted the seeds in a grid of 100 plants by 30 rows in both years with the same spacing as teosinte. The borders were also planted in similar fashion as teosinte. Unlike teosinte, we did not suffer from germination problems in the maize landrace and so we did not have to sample any of the border plants nor change the field layout in the second year.

Tissue Collection, DNA Isolation and Genotyping

Leaf tissue samples from the parents and progeny of teosinte and maize landrace were collected for DNA isolation using several different methods. For the teosinte parents, 100 – 350 mg leaf tissue samples were collected depending on the DNA isolation protocol used, which was either DNeasy® Plant Mini Kit (Qiagen Inc., Germantown, MD) or modified CTAB protocol (CIMMYT, 2005). Due to poor kit yield, DNAs were isolated and pooled from three to five leaf tissue samples for each of those affected parents. DNAs isolated from the modified CTAB protocol were sufficient so no pooling was required. For the maize landrace parents, 100 mg leaf tissue samples were collected and lyophilized prior to DNA isolation with the same kit but without pooling. For the progeny, small samples of leaf tissue (1 cm by 4 cm) were collected from each plant at approximately the 5-leaf stage. Roughly 1% of the progeny were randomly sampled twice as a control against tissue collection error. All of the tissue samples were lyophilized prior to DNA isolation using DNeasy® 96 Plant Kit (Qiagen Inc., Germantown, MD). All DNA samples from the parents and progeny were genotyped using Genotype-by-Sequencing (GBS) (Elshire *et al*. 2011). As per GBS protocol, all DNA samples were digested using *ApeKI* restriction enzyme and sequenced in 96-plex on Illumina HiSeq 2000, SE 1X100 bp (Illumina Inc., San Diego, CA). 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 (Glaubitz *et al*. 2014). The overall genotyping process from raw sequencing reads to final, clean and imputed GBS dataset is highlighted in a flowchart (*SI Appendix*, Fig. S4).

Phenotyping

We collected phenotype data for numerous traits in both teosinte and maize landrace populations. The trait abbreviations can be found in Table 1 and the trait details are summarized in *SI Appendix*, Table S12. We scored a total of 32 traits in teosinte and 43 traits in maize. Of these traits, 18 are shared by both teosinte and maize landrace and are the focus of our analyses (*SI Appendix*, Table S12). The remaining 14 traits for teosinte and 25 traits for maize landrace were not analyzed here since these traits were either invariable or not scored in one of the populations. In total, we collected phenotype data for 4,455 teosinte plants and 4398 maize landrace plants. Due to various reasons such as tractor damage to a plant or plant death, some of the plants do not have complete phenotype data.

Since it is impractical to obtain an accurate count of Total Ears per Plant (TEPP) for teosinte, we used a linear regression model to predict TEPP instead. TEPP itself is

not analyzed, but instead used for calculating TGPP and TGWP. We counted the actual TEPP for 200 plants over two years and regressed these actual counts on several other traits that we thought might be correlated to TEPP, including Plant Height (PLHT), Leaf Length (LFLN), Leaf Width (LFWD), Lateral Branch Node Number (LBNN), Lateral Branch Length (LBLN), Tiller Number (TILN), Culm Diameter (CULM), Branch Number (BRAN), Prolificacy (PROL), Days to Anthesis (DTA) and Days to Silk (DTS). From the initial model, we kept only traits that were significant ($p < 0.01$), resulting in the following prediction model that explained 67.8% of variance:

$$TEPP = -47.7 + 95.3 \cdot Year - 1.9 \cdot LBLN + 15.8 \cdot TILN + 27.2 \cdot BRAN + 22.1 \cdot PROL$$

Shapes of teosinte fruitcases are highly variable and yet the variations are often too subtle to be scored objectively by humans. In order to capture shape variation, we used the software SmartGrain (Tanabata *et al*. 2012) to analyze scanned images of teosinte fruitcases. First, we scanned the same 50 fruitcases (a few plants had fewer than 50 fruitcases) that were previously weighed for Grain Weight (GW). All the scans were done at 600 dpi on a pink background and exported as ".jpeg" format using Silverfast v8.0.1 (LaserSoft Imaging Inc., Sarasota, FL) and Epson Perfection V800 Photo Color Scanner (Epson America Inc., Long Beach, CA). The fruitcases were spread over a rectangular area of 215.9 mm by 94.2 mm to minimize contact among the fruitcases. We analyzed the scanned images using SmartGrain (Tanabata *et al*. 2012) where the scale was set to 0.0423 mm/pixel and individual fruitcases were identified based on color differences from the background. We also performed manual quality checks on each of the scanned images by excluding fruitcases that were in contact with one or more others and also fruitcases that did not lay flat. We scored a total of four traits using this method, including Fruitcase Length (FCLN), Fruitcase Width (equivalent to Ear Diameter, ED), Fruitcase Length-Width-Ratio (FCLW), and Fruitcase Triangularity (FCTR). Of these four traits, only ED is analyzed here.

Modelling Shading

A covariate for shading was modelled for each progeny (*x*) by considering the possible shading effects of neighboring plants that are 119cm or less away as described in *SI Appendix*, Fig. S6. We first calculated individual shading (s_i) from each neighboring plant using the following formula:

$$s_i = PLHT_i + \frac{1}{2} \cdot PLHT_i \cdot TILN_i$$

where s_i is the shading contributed by a neighboring plant *i*, $PLHT_i$ is the plant height of a neighboring plant *i*, and $TILN_i$ is the tiller number of a neighboring plant *i*. For example, s_{L1} is the shading on plant *x* by neighboring plant *L1* (*SI Appendix*, Fig. S6). Shading from border (non-experimental neighboring plants) was assumed to be the average shading of all progeny plants, $\overline{s_i}$. Shading from an empty plot or irrigation alley is 0. Specifically only in the first year teosinte data, shading from border was assumed to be half of $\overline{s_i}$ since the amount of border was sparser. Also, considering that neighboring plants that are farther away shade less than neighboring plants that are close by, we scaled s_i based on the distance of neighboring plant *i* from plant *x*. Summing all the scaled s_i values gave us the total shading (S_x) on plant *x*:

$$S_x = \sum_{i=1}^{20} \frac{d_i}{d} s_i$$

where d_i is the shortest distance of neighboring plant *i* from plant *x*, and *d* is the shortest possible distance of the closest neighboring plant to plant *x* (30cm).

Parentage Inference

Using raw GBS data of the parents and progeny, we inferred parentage of each progeny for both teosinte and maize landrace. Parentage inference was done in two parts, first by estimating the realized additive genomic relationship matrix (VanRaden 2008; Endelman and Jannink 2012) in TASSEL5 (Bradbury *et al*. 2007), followed by identifying the parents of each progeny using a custom R script. A progeny is considered a selfed of a parent if there is only one progeny-parent pair with the highest additive relationship value. A progeny is considered an outcross of two parents if there are two progeny-parent pairs with high and similar additive relationship values.

Once the parentage inference was complete, we compared the inferred parentage to the known maternal parentage of each progeny. Owing to our crossing strategy, we were able to trace each progeny to its maternal parent plant from which the seed was harvested. For the maize landrace, we also knew the five most likely paternal parents of each progeny based on our crossing design. Comparing the parentages from duplicate samples, we verified that there were no large-scale sample mix-ups. In a handful of cases where the inferred parentage did not meet the expected parentage, we investigated them on a step-by-step basis. First, we verified that the additive relationship values of the inferred progeny-parent pairs were reasonable and non-ambiguous. Second, we checked the planting and tissue collection notes for any known error. Third, we looked for adjacent seeds that would match up with the inferred parentage and suggest for seed or tissue mix-up. Once all the steps are taken, we corrected the progeny for which we could confidently identify the source of error. For progeny that we could not identify the source of error, we re-genotyped them using DNAs isolated from backup tissues and inferred the parentage again. If the re-genotyped progeny still had uncertain or unlikely parentage, then those progeny were removed from the dataset. If the re-genotyped progeny matched the expected parentage, then the GBS genotypes were corrected accordingly.

Uplifting from AGPv2 to AGPv4

The CrossMap (Zhao *et al*. 2013) software was used to convert the GBS SNP positions from maize B73 reference AGPv2 coordinates to AGPv4 coordinates. CrossMap requires an assembly chain file to do uplifting. However, at the time when this work was carried out, Ensembl Plants (http://plants.ensembl.org) only provided AGPv2->AGPv3 and AGPv3->AGPv4 chain files, and CrossMap's VCF functionality was not compatible with the AGPv3->AGPv4 chain file. A three-step approach was used for uplifting. First, the GBS AGPv2 VCF file was converted into the bed file format; second, CrossMap was used to uplift variant positions and allele strands to AGPv3 then to AGPv4 in bed file formats; third, a custom made PERL script was used to create a new GBS AGPv4 file in VCF format by combining data from the AGPv2 VCF file and AGPv2-to-v4 bed file.

GBS Data Imputation

We performed two rounds of imputation. Prior to the first imputation, we filtered the GBS data for teosinte and maize landrace separately. We removed the following: (1) sites with minor allele frequency (maf) below 0.001, (2) sites with missing rate above 20%, (3) sites that were non-biallelic, and (4) sites with insertion-deletion (indel) polymorphism. After the first imputation, we also removed sites that were incorrectly mapped in AGPv4, and progeny that had more than 100 crossovers total. Incorrectly mapped sites were only identified on Chromosome 6 between 58,557,140 and 61,656,987 bp and were discovered by a quality check (QC) described below. Similarly,

progeny with more than 100 crossovers were identified by a QC as described below. Our GBS data workflow is summarized in *SI Appendix*, Fig. S4.

When using GBS data for heterozygous individuals, imputing genotypes is an important but challenging first step due to the ambiguity of many of the genotypes. Any sequence-based variant calling method often provides low coverage genotype calls that are ambiguous because each sequence read only captures a single chromosome. For an individual, if a variant site is covered by only a single read, then it will not be known whether that individual is homozygous or heterozygous at that site. If a site is covered by two reads, both alleles will be observed only half the time. However, if parent phase is known, a hidden Markov model (HMM) can be used with GBS data to infer at almost every site in the genome which chromosomes a progeny has inherited from its parents. Based on that process, the correct genotype of the progeny can be inferred from the parent haplotypes.

Here, we briefly describe a two-part approach in imputing the progeny genotypes from raw GBS data, with the details to follow in the remaining paragraphs of this section. In the first part, we used a heuristic method to phase the parents with at least 10 progeny followed by a HMM to infer the parent states at all possible sites in the progeny. In the second part, we improved the parent phasing using the inferred parent states and re-imputed parent states in the progeny using the improved parent phasing. Since parent phasing is an important step in the imputation process, and yet no reliable method is available for our low coverage genotype calls, we created and applied two separate algorithms to phase the parents. One method used only selfed progeny and a second independent method used both selfed and outcross progeny. The results of the two methods were merged. For each individual in the population, the phased parent haplotypes were used with the Viterbi algorithm (Rabiner 1989) to determine the parent state at each site with sufficient data. Any unknown site flanked by two sites with the same parent state was assumed to have that state. At the end, the re-imputed parent states were used in conjunction with the phased parent haplotypes to impute the progeny genotypes.

Parents with sufficient number of selfed progeny (here we used a minimum of 10) can be phased relatively easily and reliably. The phasing process began by selecting all the selfed progeny of a single parent with all of the monomorphic sites within that family ignored. Using only polymorphic sites, each chromosome was divided into 50-site windows. For the first window, pairwise distances were calculated between every pair of genotypes as the number of nucleotide differences. Genotypes were clustered using those differences. The two largest clusters were taken to be the parental haplotypes. Because many heterozygous loci were randomly called as one or the other of the alleles, almost every heterozygous individual had a distinct genotype and the heterozygotes did not cluster together. The process was repeated for the second and subsequent windows and haplotypes from the same chromosome identified by keeping track of which individuals carried which haplotypes. Some regions had only a single haplotype because of segregation distortion or because the parental chromosomes were identical-by-descent (IBD). To handle these regions, the algorithm determined whether or not the window was IBD or whether it was likely that one haplotype had been eliminated through selection or sampling. All individuals in an IBD region are expected to form one large cluster. In a non-IBD region, about two thirds of the individuals are expected to be heterozygous and, as a result, not to cluster. As a result IBD could be distinguished from segregation distortion based on cluster size. In the case of IBD, the same haplotype was assigned to both parents. In the case of segregation distortion one parental haplotype was set to all missing values.

Additionally, we also separately phased each parent using both selfed and outcross progeny at sites with sufficiently high coverage (read depth > 6) to distinguish heterozygotes reliably. For most sites where there is sufficient depth for the parents and one progeny, the progeny can be phased, a process which provides one of the haplotypes from each of the parents. By clustering all of the haplotypes generated by all combinations with a single parent, the two chromosomes from that parent can be identified. This was repeated for windows of 40 polymorphic sites across a chromosome. Keeping track of which progeny contain each haplotype allows adjacent windows to be joined correctly.

Subsequently, we combined the haplotypes from the two independent phasing methods. Sites covered by only one method were kept. Sites covered by both methods were kept only when they agreed. In almost all cases, the agreement between the two methods was excellent. In a few cases where there was substantial disagreement, the parent had a large region in which the two chromosomes in a pair were IBD. As a result the haplotypes bordering that region were in weak LD and the haplotypes from the selfed-only progeny had been interchanged in the middle of the chromosome. When the haplotypes from the selfed and outcross progeny were used to correct the order, the two methods were brought into agreement.

The Viterbi algorithm describes a type of HMM that identifies the most likely genotype given the marker data. The "true" genotype is considered to be unknown or hidden. In this application, the algorithm was used to infer the parental chromosomes inherited by a single progeny at each position with both progeny and parent data. The possible (hidden) parent states are maternal chromosome 1, paternal chromosome 1; maternal chromosome 1, paternal chromosome 2; maternal chromosome 2, paternal chromosome 1; and maternal chromosome 2, paternal chromosome 2. The Viterbi algorithm requires a transition matrix and an emission matrix. The transition matrix was calculated for each pair of sites based on the probability of a recombination between those two sites. The emission matrix was calculated assuming that homozygous loci would be genotyped correctly with an error rate of 0.002 and the heterozygous loci would be genotyped as homozygous at a rate that depended on read depth. For example, at read depth 1, the probability that a heterozygous locus A/B would be genotyped as homozygous AA is 0.5 and as homozygous BB is 0.5. At read depth 2, the probability that a heterozygous locus A/B would be genotyped as heterozygous is 0.5, as homozygous AA is 0.25, and as homozygous BB is 0.25.

For a single progeny, once the parent states had been imputed at sites with both parent and progeny data, sites in the genome for which states had not been imputed and that were bordered by sites with the same state were assigned that state. In that way, parent states were assigned for all sites except those surrounding a recombination event. Once the states had been imputed, the progeny genotype was inferred by combining the phased haplotypes of the parents. As a result, the density of the imputed genotypes for a progeny depended on how many sites were phased in its parents.

The imputation algorithm is implemented as part of the TASSEL (Bradbury *et al*. 2007) code base as the ParentPhasingPlugin and the ImputeProgenyStatesPlugin. They can be run from the command line using the current version of TASSEL, which is available at http://www.maizegenetics.net/tassel. The source code is freely available at the TASSEL Wiki (https://bitbucket.org/tasseladmin/tassel-5-source/wiki/Home).

Quality Check on Imputation

Upon completing the first imputation, we performed a quality check (QC) using several metrics on our GBS data to ensure high quality data for our subsequent analyses (*SI Appendix*, Fig. S4-5). We first checked for crossover density along every

chromosome, where we expected to see low crossover density near centromeres and high crossover density away from centromeres. We identified a region on Chromosome 6: 58,557,140 - 61,656,987 bp with unusually high crossover density in both teosinte and maize landrace (*SI Appendix*, Fig. S5A). We removed SNPs in this region prior to the second imputation. We also looked at total crossover counts for each individual, where we identified six teosinte and 98 maize landrace individuals with 100 or more crossovers. These individuals were also removed prior to the second imputation. Since a small proportion of crossover errors are expected to have little impact on genome-wide association tests at individual markers, we used a liberal threshold for removing individuals with 100 or more crossovers, even though individuals with more than 40 imputed crossovers are likely to be imputed incorrectly, as maize gametes typically carry 20 or fewer recombinations (Anderson *et al*. 2004; Bauer *et al*. 2013; Sidhu *et al*. 2015). For a few self/outcross families, we plotted the minor allele frequencies (maf) along each chromosome (*SI Appendix*, Fig. S5C-E). Under Mendelian segregation, we expected to see maf of 0 and 0.5 in a selfed family, or 0, 0.25 and 0.5 in an outcross family. Regions showing deviations from those expected mafs would suggest either segregation distortion (SD) or identity-by-descent (IBD). While we identified regions with SD or IBD from the maf plots, none of these needed any correction. Lastly, we estimated the realized additive genomic relationship matrix from the imputed GBS data and plotted the matrix diagonals separately for self and outcross progeny (*SI Appendix*, Fig. S5F-G). The diagonals of the additive relationship matrix are estimates of 1 + *f* where *f* is the individual's inbreeding coefficient (Endelman and Jannink, 2012). Assuming no history of selfing among the parents, we expected the diagonals for selfed progeny (*f* = 0.5) to be centered around 1.5 and outcross progeny (*f* = 0) to center around 1.0. Initially the diagonal estimates fit the 1.5 (self) and 1.0 (outcross) expectations imperfectly, so we applied a conservative but not overly stringent post-imputation filtering criteria by removing progeny with more than 70% missing data and sites with more than 10% missing data. After this filtering, the expectations of diagonal values neear1.5 and 1.0 were met very well.

We also applied the QC to the output from the second imputation to ensure a high quality imputation. The second imputation resulted in 4,669 progeny and 349,964 sites for teosinte, and 4,792 progeny and 351,719 sites for maize landrace. After the second imputation, the previously identified erroneous spike in the crossover density plot was gone (*SI Appendix*, Fig. S5B). The individual crossover counts were satisfactory, where the crossover counts range from 10 to 95 (mean = 27.3 ± 8.8) for teosinte, 12 to 105 (mean = 28.42 ± 0.10) for maize landrace. Even though one individual did not pass the 100 crossovers threshold in the second imputation, we left the individual in the dataset since incorrectly imputed crossovers are unlikely to impact our downstream analyses. Similar to the first QC, we did not identify any issue with the maf distribution along chromosomes. We also applied the same post-imputation filtering of removing progeny with more than 70% missing data and sites with more than 10% missing data. After this step, we filtered out any monomorphic sites, leaving 4,455 progeny and 34,899 sites for teosinte, and 4,398 progeny and 40,255 sites for maize landrace. Lastly, we imputed most of the remaining missing data using LD-kNNi (Money *et al*. 2015) implemented in TASSEL5 (Bradbury *et al*. 2007) with default parameters. For teosinte, the final dataset for analysis had 0 to 2.38% missing data per site and 0 to 0.75% missing data per progeny. For maize landrace, these numbers are 0 to 3.62% missing data per site and 0 to 0.23% missing data per progeny.

Univariate Analysis

Realized additive and dominance genomic relationship matrices were estimated from the final GBS datasets using the observed allele frequencies method of VanRaden (2008) and Endelman and Jannink (2012) implemented in TASSEL5 (Bradbury *et al*. 2007). Realized dominance genomic relationship matrices were estimated using the method of Muñoz *et al*. (2014). The distribution of diagonal and off-diagonal elements from these matrices are shown in SI Appendix, Fig. S7. Reliable estimation of dominance variance components requires some close relatives with higher realized estimates of fraternity (off-diagonal elements of the dominance relationship matrix). Among all pairs of individuals in 0.8% (80,114 pairs) in the teosinte population and 1.6% (156,772 pairs) in the maize landrace population were full-sibs or S1 relatives from a common parent. These close relatives had realized dominance relationship coefficients centered around 0.5 (SI Appendix, Fig.S8), providing sufficient information for reliable estimation of dominance variance within each population.

A common univariate linear mixed model was fitted for each trait using ASReml version 4 (Gilmour *et al*. 2015), which implements restricted maximum likelihood estimation of model parameters.

$$Y_{ij} = \mu + E_i + (F_{ij} - \overline{F}_{..})\beta_F + (F_{ij} - \overline{F}_{i.})\beta_{Fi} + x_{Sij}\beta_S + x_{Bij}B(Y)_i + x_{Rij}\beta_{R1i} + x_{Rij}^2\beta_{R2i} + x_{Rij}^3\beta_{R3i} + x_{Rij}^4\beta_{R4i} + x_{Cij}\beta_{C1i} + x_{Cij}^2\beta_{C2i} + x_{Cij}^3\beta_{C3i} + x_{Rij}^4\beta_{C4i} + A_{ij} + D_{ij} + G \times E_{ij} + \varepsilon_{ij},$$

Where:
Y_{ij} is the observed phenotype on individual *j* in environment *i*.

The following fixed effects are included in the model:
E_i is the effect of environment (year) *i*,
F_{ij} is the marker-based inbreeding coefficient estimate for individual *j* in environment *i*,
$\overline{F}_{..}$ is the average inbreeding coefficient for all individuals across both years,
$\overline{F}_{i.}$ is the mean inbreeding coefficient for all individuals in environment *i*,
β_F is the average regression coefficient for phenotypes on the inbreeding coefficient (the estimate of inbreeding depression),
β_{Fi} is the interaction effect of inbreeding depression effect with years.
x_{Sij} is the deviation of the shading measurement on the *ij*th individual from the overall average shading measurement,
β_S is the average shading effect,
x_{Bij} is a dummy variable indicating if a plant is in an edge (border) row for teosinte plants or in a row adjacent to a tractor tire passing lane for maize landrace plants,
$B(Y)_i$ is the effect of border rows in the first year (since no plants were measured in border rows in the second year),
x_{Rij}^p and x_{Cij}^p are *p* = first to fourth order polynomials of the deviation in the row and column directions, respectively, of the *ij*th plant's position from the center of the field in year *i*,
β_{Rpi} and β_{Cpi} are the regression coefficients associated with the *p*th polynomials for row and column trend effects within year *i*, respectively.

The following random effects are included in the model:
A_{ij} is the polygenic additive effect of the *ij*th plant, with distribution $A_{ij} \sim MVN(0, \boldsymbol{A}\sigma_A^2)$, where $\boldsymbol{A}$ is the realized additive genomic relationship matrix and $\sigma_A^2$is the estimate of the additive genetic variance,

D_{ij} is the polygenic dominance effect of the *ij*th plant, with distribution $D_{ij} \sim MVN(0, \boldsymbol{D}\sigma_D^2)$, where $\boldsymbol{D}$ is the realized dominance genomic relationship matrix and σ_D^2 is the estimate of the dominance genetic variance,
$G \times E_{ij}$ is the interaction of polygenic effect of the *ij*th plant with environment *i*, with distribution $G \times E_{ij} \sim MVN(0, (\boldsymbol{A_1} \oplus \boldsymbol{A_2})\sigma_{AE}^2)$, where $\boldsymbol{A_i}$ is the realized additive relationship matrix for individuals grown in year *i* and $(\boldsymbol{A_1} \oplus \boldsymbol{A_2})$ is a block-diagonal structure that includes non-zero covariances for plants grown in the same year, but zero covariance for plants grown in different years.
ε_{ij} is the residual effect associated with the *ij*th plant, with heterogeneous variances across years: $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon i}^2)$.

Based on the model's variance component estimates, we estimated the following:
Narrow-sense heritability as $h^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2+\hat{\sigma}_D^2+\hat{\sigma}_{G\times E}^2+\hat{\sigma}_{\bar{\varepsilon}}^2}$, where $\hat{\sigma}_{\bar{\varepsilon}}^2$ is the average error variance across years.
Broad-sense heritability as $H^2 = \frac{\hat{\sigma}_A^2+\hat{\sigma}_D^2}{\hat{\sigma}_A^2+\hat{\sigma}_D^2+\hat{\sigma}_{G\times E}^2+\hat{\sigma}_{\bar{\varepsilon}}^2}$
Proportion of phenotypic variance due to dominance variance as ${}^{V_D}/_{V_P} = \frac{\hat{\sigma}_D^2}{\hat{\sigma}_A^2+\hat{\sigma}_D^2+\hat{\sigma}_{G\times E}^2+\hat{\sigma}_{\bar{\varepsilon}}^2}$
Proportion of phenotypic variance due to genetic-by-environment variance as ${}^{V_{G\times E}}/_{V_P} = \frac{\hat{\sigma}_{G\times E}^2}{\hat{\sigma}_A^2+\hat{\sigma}_D^2+\hat{\sigma}_{G\times E}^2+\hat{\sigma}_{\bar{\varepsilon}}^2}$

We inferred the evolutionary history for each individual trait by applying the univariate breeder's equation (Lush 1937) on our univariate analysis results. The univariate breeder's equation is given by the following (Equation 13.6b from Walsh and Lynch (2018)):

$$R = h^2 i\sqrt{V_P} = i\frac{V_A}{\sqrt{V_P}}$$

where R is the response or the difference in population mean before (z_{n-1}) and after (z_n) selection, h^2 is the narrow sense heritability, i is the selection intensity, V_P is the phenotypic variance and V_A is the additive genetic variance. We can generalize and re-express the univariate breeder's equation as:

$$z_n - z_{n-1} = i\frac{V_{A,n-1}}{\sqrt{V_{P,n-1}}}$$

For N generations of selection with a constant i, we have the following equations:

$$\begin{aligned} z_1 - z_0 &= i\frac{V_{A,0}}{\sqrt{V_{P,0}}} \\ z_2 - z_1 &= i\frac{V_{A,1}}{\sqrt{V_{P,1}}} \\ &\vdots \\ z_N - z_{N-1} &= i\frac{V_{A,N-1}}{\sqrt{V_{P,N-1}}} \end{aligned}$$

Summing up all the equations leads to:

$$z_N - z_0 = i\left(\frac{V_{A,0}}{\sqrt{V_{P,0}}} + \frac{V_{A,1}}{\sqrt{V_{P,1}}} + \cdots + \frac{V_{A,N-1}}{\sqrt{V_{P,N-1}}}\right) = i\sum_{n=0}^{N-1}\frac{V_{A,n}}{\sqrt{V_{P,n}}}$$

Or equivalently:

$$i = \frac{z_N - z_0}{\sum_{n=0}^{N-1}\frac{V_{A,n}}{\sqrt{V_{P,n}}}}$$

The phenotypic variance is often simplified as the sum of additive genetic and environmental variances as the other variances are assumed to be negligible. Furthermore, common literatures often assume that the environmental variance remains constant over multiple generations of selection and that the additive genetic variance always decreases under selection (Bulmer 1971; Verrier *et al*. 1990; Hospital and Chevalet 1993; Roff 1997). However, we found these assumptions to be unrealistic as our results clearly show that the environmental variances between teosinte and maize landrace are different and that the additive genetic variances do not necessarily decrease under selection (*SI Appendix*, Table S1-2). Environmental variance has been shown to increase with fixation of alleles that affect environmental sensitivity (Mackay and Lyman 2005). Such alleles may either be directly selected or indirectly selected via linkage disequilibrium during maize domestication. Additive genetic variance can increase with mutations and epistasis, especially in a population that has a long period of time to evolve. Thus, we opted against modeling the change in additive genetic variance using the methods described in the current literature. We instead modelled the change in genetic architecture of each quantitative trait due to selection as constant change in additive genetic and phenotypic variances per generation. Even though our models are likely naïve, our models are still probably closer to the reality than any model proposed in the current literature.

For our model, we assumed that the change in variance is constant every generation, where:

$$V_n = V_{n-1} + \Delta V$$
$$V_n = V_0 + n\Delta V$$

For N generations, the change in variance can be estimated as following:

$$\Delta V = \frac{V_N - V_0}{N}$$

Using our teosinte population as the starting point (generation 0) and maize landrace population as the ending point (generation N), we can easily estimate $V_{A,n}$ and $V_{P,n}$ for any $n < N$ generation and thus obtain the estimate of selection intensity (i) as following:

$$i = \frac{z_N - z_0}{\sum_{n=0}^{N-1} \frac{V_{A,0} + n\Delta V_A}{\sqrt{V_{P,0} + n\Delta V_P}}}$$

Where:

N is the number of generations of selection
z_N is the trait outcross mean for maize landrace
z_0 is the trait outcross mean for teosinte
$V_{A,0}$ is the additive genetic variance for teosinte
$V_{P,0}$ is the phenotypic variance for teosinte
$V_{A,N}$ is the additive genetic variance for maize landrace
$V_{P,N}$ is the phenotypic variance for maize landrace
ΔV_A can be estimated as $\Delta V_A = \left(V_{A,N} - V_{A,0}\right)/N$
ΔV_P can be estimated as $\Delta V_P = \left(V_{P,N} - V_{P,0}\right)/N$

Multivariate Analysis

Bivariate linear mixed model analyses were conducted using ASReml version 4 (Gilmour *et al*. 2015) for each pair of traits to estimate additive genetic correlations. Because computational demand increased dramatically for bivariate analysis compared to univariate analysis, we used a reduced multivariate analog of the univariate analysis model:

$$Y_{ijk} = \mu_k + E_{ik} + (F_{ij} - \overline{F}_{.})\beta_{Fk} + (F_{ij} - \overline{F}_{.})\beta_{Fik} + x_{Sij}\beta_{Sk} + x_{Bij}B(Y)_{ik} + x_{Rij}\beta_{R1ik} + x_{Rij}^2\beta_{R2ik} + x_{Rij}^3\beta_{R3ik} + x_{Rij}^4\beta_{R4ik} + x_{Cijk}\beta_{C1ik} + x_{Cijk}^2\beta_{C2ik} + x_{Cijk}^3\beta_{C3ik} + x_{Rij}^4\beta_{C4ik} + A_{ijk} + \varepsilon_{ijk},$$

Where $Y_{ijk}$is the measurement of trait *k* on individual *j* in environment *i*, and the fixed model terms are the same as in the univariate model, but they are nested within trait. The random terms in this model include the additive polygenic effect, A_{ijk}, which has distribution $A_{ijk} \sim MVN(0, \boldsymbol{A} \otimes \boldsymbol{T})$, where $\boldsymbol{T}$ is a 2×2 additive genetic variance-covariance matrix for traits 1 and 2:

$$\boldsymbol{T} = \begin{bmatrix} V_{A,1} & Cov_{A,1,2} \\ Cov_{A,1,2} & V_{A,2} \end{bmatrix}.$$

The random residual terms in this model are correlated across traits within a plant: $\varepsilon_{ijk} \sim MVN(0, \boldsymbol{I} \otimes \boldsymbol{E})$, where $\boldsymbol{I}$ is an identify matrix with dimension equal to the total number of plants measured and $\boldsymbol{E}$ is a 2×2 residual variance-covariance matrix:

$$\boldsymbol{E} = \begin{bmatrix} V_{Err,1} & Cov_{Err,1,2} \\ Cov_{Err,1,2} & V_{Err,2} \end{bmatrix}.$$

Restricted maximum likelihood estimates of the variance and covariance components were used to estimate the additive genetic correlation between traits *k* and *k'*:

$$r_{Akk'} = \frac{\hat{\sigma}_{Akk'}}{\sqrt{\hat{\sigma}_{Ak}^2 \hat{\sigma}_{Ak'}^2}}.$$

Significance of the genetic correlations can be determined by first applying Fisher's Z-transformation to the correlations and their standard errors, and then compared to the standard normal distribution for the corresponding two-tailed P-values of $2 \cdot P\left(\left|\frac{\tanh^{-1}(r_G)}{\tanh^{-1}\left(SE_{r_G}\right)}\right| > 1.96\right) < 0.05$. The genetic correlations and their standard errors can be found in *SI Appendix*, Table S4 and S5.

Using teosinte and maize landrace distance matrices, we confirmed the genetic relationship among the 18 traits. Elements of the distance matrix are calculated from the elements of the genetic correlation matrix as $1 - |r_g|$ where r_g is the genetic correlation between any two traits. To visualize the distance matrices, we used Principal Coordinate Analysis (PCoA) and Neighbor-Joining (NJ) tree. PCoA is performed using *cmdscale* function in R (R Core Team 2018) while NJ-tree is calculated using *nj* function implemented in the "ape" package (Paradis *et al*. 2004) in R (R Core Team 2018). Based on the results shown in Fig. 4 and *SI Appendix*, Fig. S1, we saw a good agreement between the PCoA and NJ plots and our pre-defined trait groups of Vegetative/Flowering Time, Environmental Response and Reproductive.

A clear result of the genetic correlation estimates is that GE and CUPR are identical in teosinte, whereas GE and TGPP are nearly identical in maize. Further analysis of the genetic correlation matrices and genetic covariance matrices ($\boldsymbol{G}$-matrices) is hindered by the singularity (or near-singularity) in the matrices caused by these identical or nearly identical traits. Therefore, for subsequent tests on the eigenstructure of genetic correlation matrices and $\boldsymbol{G}$-matrices, we dropped CUPR and TGPP from the matrices and analyzed the resulting sub-matrices of 16 traits.

We tested for conservation in genetic correlations by comparing teosinte and maize landrace genetic correlation matrices using Mantel test (Mantel, 1967). The Mantel test calculates correlation between the elements of two matrices and tests if the correlation is significantly different from zero. Therefore, significant Mantel test would suggest that two matrices are correlated and in our case, genetic correlations are preserved during domestication. Aside from testing the overall conservation in genetic correlations, we also applied Mantel test on genetic correlations within each trait group. These additional tests allow us to compare whether the overall genetic correlations or within trait group genetic correlations are better conserved. All Mantel tests are performed using *mantel.test* function with 10,000 permutations implemented in the package "ape" (Paradis *et al*. 2004) in R (R Core Team 2018).

Additionally, we also calculated the angle between the first two leading eigenvectors of the teosinte and maize landrace genetic correlation matrices as a supporting evidence for Mantel tests. The eigenvectors are identified using *eigen* function in R. The angle between eigenvectors measures the deviation between teosinte and maize landrace genetic correlations, i.e. the larger the angle, the less similar the genetic correlations are. Similar analysis was also repeated on the three submatrices based on trait groups.

We compared the structure of $\boldsymbol{G}$-matrices for teosinte and maize landrace using Mantel test (Mantel 1967), Flury hierarchy (Flury 1988), Random Skewers (Cheverud and Marroig 2007), Bayesian estimation (Ovaskainen et al. 2008), and multivariate $Q_{ST} - F_{ST}$ tests (Martin *et al*. 2008). For the first four tests, we used $\boldsymbol{G}$-matrices from teosinte ($\boldsymbol{G_T}$) and maize landrace ($\boldsymbol{G_M}$) that were previously calculated from our multivariate analysis. For the multivariate $Q_{ST} - F_{ST}$ test, we used $\boldsymbol{G}$-matrices from between-

population ($\boldsymbol{G_B}$) and within-population ($\boldsymbol{G_W}$). Formally, $\boldsymbol{G_B}$ and $\boldsymbol{G_W}$ ought to be calculated by combining the teosinte and maize landrace raw datasets and running the multivariate analysis again. We opted for a more time-efficient method by estimating the elements of $\boldsymbol{G_B}$ as $G_{B,ij} = \frac{1}{2}(\bar{T}_i - \bar{M}_i)(\bar{T}_j - \bar{M}_j)$ where $\bar{T}_{i/j}$ and $\bar{M}_{i/j}$ are teosinte and maize landrace outcross means for *i*th/*j*th trait estimated from the univariate analysis. We also estimated the elements of $\boldsymbol{G_W}$ as $G_{W,ij} = \frac{1}{2}(G_{T,ij} + G_{M,ij})$. Both $\boldsymbol{G_B}$ and $\boldsymbol{G_W}$ were bent to be positive definite (PD) using *nearPD* function implemented in the "Matrix" package (Bates and Maechler 2017) in R (R Core Team 2018). In addition, we also estimated the F_{ST} for teosinte and maize landrace using *varcomp.glob* function implemented in the "hierfstat" package (Goudet 2004) in R (R Core Team 2018). A total of 21,157 imputed GBS SNPs that are presumably neutral and shared between teosinte and maize landrace were used in estimating neutral F_{ST}.

Mantel test for the two $\boldsymbol{G}$-matrices was performed similarly to the previous Mantel test for genetic correlation matrices, with the exception that we did not perform Mantel test for each trait group within the $\boldsymbol{G}$-matrices.

Flury hierarchy tests for similarity between two matrices by comparing their eigenvectors and eigenvalues in a hierarchical way (Flury, 1988). The test is provided in the Common Principal Component (CPC) software (Phillips and Arnold, 1999). The CPC software allows us to test multiple hypotheses between different models like unrelated structure, partial common principal components, common principal components, matrix proportionality and matrix equality (Phillips and Arnold, 1999). Using the CPC software, we tested the $\boldsymbol{G}$-matrices from teosinte and maize landrace using jump-up approach. This approach tests between the model of unrelated structure and other higher models in the hierarchy. The null hypothesis can be rejected based on the first significant test ($P < 0.05$) starting from the bottom of the hierarchy and any subsequent significant test is ignored. Since the CPC software uses maximum likelihood method for matrix comparison, both teosinte and maize landrace $\boldsymbol{G}$-matrices are required to be strictly PD. To achieve that, we bent our $\boldsymbol{G}$-matrices to be PD using *nearPD* function implemented in the "Matrix" package (Bates and Maechler 2017) in R (R Core Team 2018).

Random Skewers tests for similarity between two matrices ($\boldsymbol{G}$) by comparing the predicted evolutionary responses ($\boldsymbol{R}$) calculated using the multivariate breeder's equation of $\boldsymbol{R} = \boldsymbol{G\beta}$ (Cheverud and Marroig 2007). Under the null hypothesis, correlation between teosinte and maize landrace $\boldsymbol{R}$ is no different from the correlation between $\boldsymbol{R}$ calculated from two random $\boldsymbol{G}$-matrices. As for the alternative hypothesis, correlation between teosinte and maize landrace $\boldsymbol{R}$ is higher than the null correlation. The $\boldsymbol{R}$ for teosinte, maize landrace and null distribution are generated by multiplying each $\boldsymbol{G}$ to *n*-randomly simulated **β**. The two random $\boldsymbol{G}$-matrices for the null distribution are made to have the same size as teosinte and maize landrace $\boldsymbol{G}$-matrices and the diagonal (variance) components are bound by the smallest and largest diagonal (variance) components of teosinte and maize landrace $\boldsymbol{G}$-matrices. Random Skewers is performed using *skewers* function implemented in the phytools package (Revell 2012) in R (R Core Team 2018). We applied 1000 simulations and obtained the correlation of those resulting $\boldsymbol{R}$, where the test significance ($P < 0.05$) suggests that the $\boldsymbol{R}$ under comparison are more correlated than $\boldsymbol{R}$ generated by random chance and thus the two $\boldsymbol{G}$-matrices are similar. In addition, we also sub-divided the $\boldsymbol{G}$-matrices into trait groups and tested each group with Random Skewers to identify group-specific difference in $\boldsymbol{R}$.

We also implemented a Bayesian method to compare the teosinte and maize genetic covariance matrices developed by Ovaskainen *et al*. (2008). This method takes random samples of vectors from a multivariate normal distribution described by the teosinte $\boldsymbol{G}$-matrix and computes the probability of the vector arising from the teosinte $\boldsymbol{G}$-

matrix (p_{TT}) and the probability of the vector arising from the maize $\boldsymbol{G}$-matrix (p_{TM}). The ratio $q = \frac{p_{TM}}{p_{TM} + p_{TT}}$ is a measure of the differentiation of the two matrices, ranging from 0 for completely distinct matrices to 0.5 for identical matrices. A posterior mean value of *q* was estimated by sampling 1000 random vectors from the teosinte $\boldsymbol{G}$-matrix and averaging over the resulting 1000 individual estimates of *q*. We used the formula for estimating *q* from Walsh and Lynch (2018), Appendix 3, with the following correction:

$$\hat{q}(f,g) = \frac{\widehat{1}}{n}\sum_{i=1}^{n} \frac{|\boldsymbol{G}_2|^{-\frac{1}{2}}\exp(-x_i^T \boldsymbol{G}_2^{-1} x_i/2)}{|\boldsymbol{G}_2|^{-\frac{1}{2}}\exp\left(-x_i^T \boldsymbol{G}_2^{-1} x_i/2\right) + |\boldsymbol{G}_1|^{-\frac{1}{2}}\exp(-x_i^T \boldsymbol{G}_1^{-1} x_i/2)}$$

The converse probabilities were also tested (probability of a vector sampled from the maize **G** matrix arising from either the maize or teosinte $\boldsymbol{G}$-matrix), with nearly identical results. The tests were also performed for sub-matrices of traits within each of the three trait groups.

Multivariate $Q_{ST} - F_{ST}$ test, similar to its original univariate counterpart (Spitze, 1993), assesses for neutral evolution by comparing genetic differentiation of two or more populations at multiple-trait level (Q_{ST}) to genetic differentiation at neutral loci (F_{ST}). Under the null hypothesis, any difference in the $\boldsymbol{G}$-matrices from different populations is solely attributable to neutral drift and thus $Q_{ST} = F_{ST}$ (Martin *et al*. 2008; Leinonen *et al*. 2013). Using a dual test developed by Martin *et al*. (2008), we can perform multivariate $Q_{ST} - F_{ST}$ test on our teosinte and maize landrace populations. The dual test is provided as R scripts in Martin *et al*. (2008) and the scripts were designed to use raw dataset from a breeding design as the input. Since we already had the necessary components calculated from our multivariate analyses, we instead modified the scripts to accommodate our dataset. For the first part of the test, the original null hypothesis of the $Q_{ST} - F_{ST}$ test can be re-expressed as $\boldsymbol{G_B} = 2F_{ST}/(1 - F_{ST})\boldsymbol{G_W} = \rho_{ST}\boldsymbol{G_W}$ (Martin *et al*. 2008). Neutrality is rejected if the observed coefficient $\rho_{ST,G}$ is significantly greater than the expected neutral coefficient $\rho_{ST,N}$, in which the significance is determined by non-overlapping 95% confidence intervals (CIs) for $\rho_{ST,G}$ and $\rho_{ST,N}$. CI for $\rho_{ST,G}$ is calculated using the R scripts (Martin *et al*. 2008) through maximum likelihood method. CI for $\rho_{ST,N}$ is calculated by bootstrapping the neutral F_{ST} values for 10,000 times and deriving the lower and upper bound of the interval from F_{ST} CI using $\rho_{ST,N} = 2F_{ST}/(1 - F_{ST})$ (Whitlock and Guillaume 2009). For the second part of the test, the estimates of mean square matrices between ($\boldsymbol{MS_B}$) and among ($\boldsymbol{MS_W}$) populations are compared using likelihood ratio test and expected to be proportional under the null hypothesis (Martin *et al*. 2008). To test for this, we again modified the R scripts to back-calculate the mean squares using $\boldsymbol{MS_B} = n_f\boldsymbol{G_B} + \boldsymbol{G_W}$ and $\boldsymbol{MS_W} = \boldsymbol{G_W}$, where n_f is the adjusted sample size to account for unbalanced sample sizes between populations, as provided in equation (9) in Martin *et al*. (2008).

Univariate $Q_{ST} - F_{ST}$ test assesses for neutral evolution of individual traits by comparing genetic differentiation of two or more populations at single-trait level (Q_{ST}) to genetic differentiation at neutral loci (F_{ST}). Under the null hypothesis, any difference in the trait additive genetic variance from different populations is solely attributable to neutral drift and thus $Q_{ST} = F_{ST}$ or $Q_{ST} - F_{ST} = 0$ (Leinonen *et al*. 2013). Using an improved parametric bootstrapping approach developed by Whitlock and Guillaume (2009), we can perform univariate $Q_{ST} - F_{ST}$ test on our teosinte and maize landrace populations. We modified the R scripts for the test from Whitlock and Guillaume (2009) to fit our test inputs as we already had the necessary components calculated from our

univariate analyses. We calculated the Q_{ST} for each trait as $\hat{Q}_{ST} = V_B(V_B + 2V_W)$ where V_B is the between-population additive genetic variance while V_W is the within-population additive genetic variance. For every i^{th} trait, we calculated V_B as $V_{B,i} = \frac{1}{2}(\bar{T}_i - \bar{M}_i)^2$ where $\bar{T}_i$ and $\bar{M}_i$ are teosinte and maize landrace outcross means respectively. We also calculated V_W as $V_{W,i} = \frac{1}{2}(V_{T,i} + V_{M,i})$ where $V_{T,i}$ and $V_{M,i}$ are teosinte and maize landrace additive genetic variances respectively. $\bar{F}_{ST}$ is derived by averaging previously calculated F_{ST} across all GBS loci. The null distribution of $Q_{ST} - F_{ST}$ is constructed by simulating each component in $V_{B0}/(V_{B0} + 2V_W) - F_{ST}$ 10,000 times (Whitlock and Guillaume 2009). For each trait null distribution, V_W is taken as is, while V_{B0} is estimated from V_W and F_{ST} by assuming neutrality. Under neutrality, $Q_{ST} = F_{ST}$ and thus $V_{B0}/(V_{B0} + 2V_W) = F_{ST}$. By rearranging the equation, we get $V_{B0} \cong \frac{2\bar{F}_{ST}V_W}{1-\bar{F}_{ST}}$. Within each simulation, V_{B0} and V_W can be sampled from $V_{B0}\frac{\chi^2(df_B)}{df_B}$ and $V_W\frac{\chi^2(df_W)}{df_W}$ respectively by assuming Lewontin-Krakauer distribution for V_{B0} and V_W (Whitlock 2008). F_{ST}, on the other hand, can be obtained by sampling with replacement from the previously calculated F_{ST} at multiple GBS loci. Finally, we computed the two-tailed P-value for each $\hat{Q}_{ST} - \bar{F}_{ST}$ from the null distribution.

For the analyses in the following section, we relied heavily on matrix algebra and hence, some useful terminologies are provided here:

Given two vectors of the same size, $\boldsymbol{a}$ and $\boldsymbol{b}$,

1. $\hat{\boldsymbol{a}}$ is the unit vector of $\boldsymbol{a}$ as given by $\hat{\boldsymbol{a}} = \frac{\boldsymbol{a}}{|\boldsymbol{a}|}$
2. $\boldsymbol{a} \cdot \boldsymbol{b}$ is the dot product of $\boldsymbol{a}$ and $\boldsymbol{b}$, which is the sum of product of the pairwise elements
3. $|\boldsymbol{a}|$ is the length of $\boldsymbol{a}$, which is the square root of the sum of squares of the elements

We compared the genetic lines of least resistance, $\boldsymbol{g_{max}}$ (Schluter 1996), to the actual domestication trajectory $\boldsymbol{Z}$. $\boldsymbol{g_{max}}$ is the eigenvector of $\boldsymbol{G}$ that accounts for the most variation in $\boldsymbol{G}$ while $\boldsymbol{Z}$ is a vector of difference in trait means between teosinte and maize landrace. Since the eigenvectors sensitive to the magnitude of each trait, we opted to standardized the $\boldsymbol{G}$-matrix and $\boldsymbol{Z}$ such that each trait has a genetic standard deviation or variance of 1. We first calculated $\boldsymbol{g_{max,T}}$ from the teosinte $\boldsymbol{G}$-matrix and compared the angle between $\boldsymbol{Z}$ and $\boldsymbol{g_{max,T}}$ using the following formula:

$$\boldsymbol{\theta = \cos^{-1} \hat{Z} \cdot \hat{g}_{max}}$$

The angle $\boldsymbol{\theta}$ ranges from 0˚ to 90˚ where larger angle means larger deviation in the direction between $\boldsymbol{Z}$ and $\boldsymbol{g_{max}}$, or equivalently, stronger evolutionary constraint. Additionally, we also repeated the same process using the $\boldsymbol{g_{max,M}}$ computed from the maize landrace $\boldsymbol{G}$-matrix. The angle $\boldsymbol{\theta}$ obtained from teosinte and maize landrace were compared.

We inferred the co-evolutionary history of multiple traits by using the multivariate version of the breeder's equation, as given by $\boldsymbol{R = G\beta}$. For n traits, $\boldsymbol{R}$ is a vector of responses $(R_1, R_2, \ldots, R_n)$, $\boldsymbol{G}$ is an $n \times n$ additive genetic variance-covariance matrix, and $\boldsymbol{\beta}$ is a vector of selection gradients $(\beta_1, \beta_2, \ldots, \beta_n)$. β_i elements are the partial linear regression coefficients of relative fitness on individual trait, where $\beta_i = 0$ represents no fitness advantage for any given trait value, $\beta_i < 0$ represents stronger fitness advantage for lower trait value and $\beta_i > 0$ represents stronger fitness advantage for higher trait value. Unlike its univariate counterpart, the change in $\boldsymbol{G}$ in every generation is likely

complicated and cannot be easily modelled. Instead, we applied a “what-if” approach on the multivariate breeder’s equation.

We first simulated 18 unique $\boldsymbol{\beta}$ where each $\boldsymbol{\beta^i}$ had only a single element with a value of one and the remaining elements with a value of zero. We then multiplied $\boldsymbol{G}$ by each $\boldsymbol{\beta^i}$ to obtain $\boldsymbol{R^i}$, which is the hypothetical overall response given selection on a single trait i. Under the assumption that maize domestication is largely driven by selection on a single trait, we compared each $\boldsymbol{R^i}$ to the actual domestication trajectory $\boldsymbol{Z}$. Again, the $\boldsymbol{G}$-matrix and $\boldsymbol{Z}$ are standardized such that each trait has a genetic standard deviation or variance of 1. The comparison between $\boldsymbol{R^i}$ and $\boldsymbol{Z}$ is measured by the angle between the two vectors ($\boldsymbol{\theta_Z}$) and also the scalar projection of $\boldsymbol{R^i}$ on $\boldsymbol{Z}$ ($\left|\boldsymbol{proj_Z R^i}\right|$). $\boldsymbol{\theta_Z}$ measures the deviation in the direction between $\boldsymbol{R^i}$ and $\boldsymbol{Z}$ and it ranges from 0˚ to 180˚ where larger angle means larger deviation. $\left|\boldsymbol{proj_Z R^i}\right|$ measures the amount of evolutionary gain contributed by $\boldsymbol{R^i}$ towards $\boldsymbol{Z}$ and so larger value means larger evolutionary gain.

The angles $\boldsymbol{\theta_Z}$ are calculated as following:

$$\boldsymbol{\theta_Z = \cos^{-1} \widehat{R}^i \cdot \widehat{Z}}$$

The scalar projection $\left|\boldsymbol{proj_Z R^i}\right|$ is calculated as following:

$$\left|\boldsymbol{proj_Z R^i}\right| = \boldsymbol{R^i \cdot \widehat{Z}}$$

R Scripts

```
################################################################################################
########
## R scripts used in this manuscript (roughly in the same order as presented in the
Results). ##
################################################################################################
########
library(Matrix)
library(reshape2)
library(ggplot2)
library(ggrepel)
library(ape)
library(phytools)
library(hierfstat)
library(mvtnorm)
library(gridExtra)
setwd("./")

### Read in genetic correlation (lower left triangle is teosinte; upper right triangle is
maize landrace)
corG <- read.delim("corG.txt", header=T, as.is=T)
corG <- corG[,-1]
rownames(corG) <- colnames(corG)
corG <- as.matrix(corG)

### Read in the standard errors for genetic correlation (lower left triangle is teosinte;
upper right triangle is maize landrace)
corG.se <- read.delim("corG_se.txt", header=T, as.is=T)
corG.se <- corG.se[,-1]
rownames(corG.se) <- colnames(corG.se)
corG.se <- as.matrix(corG.se)

### Isolate teosinte genetic correlations, mirror over to the lower/upper triangle and
fills in diagonals with 1.
rT <- corG
rT[upper.tri(rT)] <- NA
rT.temp <- rT
rT <- t(rT.temp)
rT[lower.tri(rT)] <- rT.temp[lower.tri(rT.temp)]
diag(rT) <- 1

### Isolate maize landrace genetic correlations, mirror over to the lower/upper triangle
and fills in diagonals with 1.
rM <- corG
rM[lower.tri(rM)] <- NA
rM.temp <- t(rM)
rM[lower.tri(rM)] <- rM.temp[lower.tri(rM.temp)]
diag(rM) <- 1

### Read in the variance components
varcomp <- read.delim("varcomp.txt", header=T, as.is=T)

### Construct the genetic variance-covariance (G) matrix for teosinte and maize landrace
stdT <- sqrt(varcomp[1:18,4])
stdM <- sqrt(varcomp[19:36,4])

GT <- rT
GM <- rM

for(i in 1:length(stdT)){
GT[i,] <- GT[i,]*stdT[i]
GT[,i] <- GT[,i]*stdT[i]
GM[i,] <- GM[i,]*stdM[i]
GM[,i] <- GM[,i]*stdM[i]
}

### Calculate the difference in trait means between teosinte and maize landrace
Z <- varcomp[19:36,9] - varcomp[1:18,9]
names(Z) <- rownames(GT)
Z.stdT <- Z/stdT
```

```
Z.stdM <- Z/stdM

### Subset the teosinte genetic correlations matrix into vegetative (veg), environmental
response (enr) and reproductive (rep).
vegT <- rT[1:5,1:5]
enrT <- rT[6:10,6:10]
repT <- rT[11:18,11:18]

### Subset the maize landrace genetic correlations matrix into vegetative (veg),
environmental response (enr) and reproductive (rep).
vegM <- rM[1:5,1:5]
enrM <- rM[6:10,6:10]
repM <- rM[11:18,11:18]

### Subset the teosinte genetic covariances matrix into vegetative (veg), environmental
response (enr) and reproductive (rep).
vegGT <- GT[1:5,1:5]
enrGT <- GT[6:10,6:10]
repGT <- GT[11:18,11:18]

### Subset the maize landrace genetic covariances matrix into vegetative (veg),
environmental response (enr) and reproductive (rep).
vegGM <- GM[1:5,1:5]
enrGM <- GM[6:10,6:10]
repGM <- GM[11:18,11:18]

### Create new datasets to exclude CUPR and TGPP for some of the multivariate analyses.
temp.re <- c("CUPR","TGPP")
varcomp.re <- varcomp[!(varcomp[,1]%in%temp.re),]; rownames(varcomp.re) <- NULL

### Genetic correlation matrices for 16 traits (CUPR and TGPP excluded).
rT.re <- rT[!(rownames(rT)%in%temp.re), !(colnames(rT)%in%temp.re)]
rM.re <- rM[!(rownames(rM)%in%temp.re), !(colnames(rM)%in%temp.re)]

### Genetic covariance matrices for 16 traits (CUPR and TGPP excluded).
GT.re <- GT[!(rownames(GT)%in%temp.re), !(colnames(GT)%in%temp.re)]
GM.re <- GM[!(rownames(GM)%in%temp.re), !(colnames(GM)%in%temp.re)]

### Difference in trait means between teosinte and maize landrace for 16 traits (CUPR and
TGPP excluded).
Z.re <- Z[!(names(Z)%in%temp.re)]
Z.stdT.re <- Z.stdT[!(names(Z.stdT)%in%temp.re)]
Z.stdM.re <- Z.stdM[!(names(Z.stdM)%in%temp.re)]

### Genetic correlation sub-matrices for 6 Reproductive traits (CUPR and TGPP excluded).
repT.re <- repT[!(rownames(repT)%in%temp.re), !(colnames(repT)%in%temp.re)]
repM.re <- repM[!(rownames(repM)%in%temp.re), !(colnames(repM)%in%temp.re)]

### Genetic covariance sub-matrices for 6 Reproductive traits (CUPR and TGPP excluded).
repGT.re <- repGT[!(rownames(repGT)%in%temp.re), !(colnames(repGT)%in%temp.re)]
repGM.re <- repGM[!(rownames(repGM)%in%temp.re), !(colnames(repGM)%in%temp.re)]

### Bent the genetic covariance matrices for 16 traits (CUPR and TGPP excluded) to be
Positive Definite.
GT.re.PD <- as.matrix(nearPD(GT.re, corr=F)[[1]])
GM.re.PD <- as.matrix(nearPD(GM.re, corr=F)[[1]])

### Bent the genetic covariance sub-matrices for 5 Vegetative traits to be Positive
Definite.
vegGT.PD <- as.matrix(nearPD(vegGT, corr=F)[[1]])
vegGM.PD <- as.matrix(nearPD(vegGM, corr=F)[[1]])

### Bent the genetic covariance sub-matrices for 5 Environmental Response traits to be
Positive Definite.
enrGT.PD <- as.matrix(nearPD(enrGT, corr=F)[[1]])
enrGM.PD <- as.matrix(nearPD(enrGM, corr=F)[[1]])

### Bent the genetic covariance sub-matrices for 6 Reproductive traits (CUPR and TGPP
excluded) to be Positive Definite.
repGT.PD <- as.matrix(nearPD(repGT.re, corr=F)[[1]])
repGM.PD <- as.matrix(nearPD(repGM.re, corr=F)[[1]])
```

```
################################################
### Custom functions used in this R script. ###
################################################

### The following function is used in PART 5.
### Function to compare the first two leading eigenvectors of two matrices by measuring
the angles between them.
### Also outputs the percent variance explained by the two leading eigenvectors for each
matrix.
theta.cal <- function(mat1, mat2){
mat1vec1 <- eigen(mat1)[[2]][,1]
mat1vec2 <- eigen(mat1)[[2]][,2]
mat2vec1 <- eigen(mat2)[[2]][,1]
mat2vec2 <- eigen(mat2)[[2]][,2]
out.theta.pve <- data.frame(label=c("theta","pve.mat1","pve.mat2"), eigenvector1=NA,
eigenvector2=NA)
out.theta.pve[1,2] <- if (acos(mat1vec1%*%mat2vec1)*180/pi < 90)
{acos(mat1vec1%*%mat2vec1)*180/pi} else {180 - acos(mat1vec1%*%mat2vec1)*180/pi}
out.theta.pve[1,3] <- if (acos(mat1vec2%*%mat2vec2)*180/pi < 90)
{acos(mat1vec2%*%mat2vec2)*180/pi} else {180 - acos(mat1vec2%*%mat2vec2)*180/pi}
out.theta.pve[2,2] <- 100*eigen(mat1)[[1]][1]/sum(eigen(mat1)[[1]][eigen(mat1)[[1]]>0])
out.theta.pve[2,3] <- 100*eigen(mat1)[[1]][2]/sum(eigen(mat1)[[1]][eigen(mat1)[[1]]>0])
out.theta.pve[3,2] <- 100*eigen(mat2)[[1]][1]/sum(eigen(mat2)[[1]][eigen(mat2)[[1]]>0])
out.theta.pve[3,3] <- 100*eigen(mat2)[[1]][2]/sum(eigen(mat2)[[1]][eigen(mat2)[[1]]>0])
return(print(out.theta.pve, row.names=F))
}

### The following function is used in PART 6.
### Function to compute 'q' from Ovaskainen's test for comparing two G-matrices.
### Originated from Ovaskainen et al. (2008)
### Modified by Walsh & Lynch (2018) Appendix 3, Example A3.1.
### Corrected by Jim Holland.
### 'q' is the probability that a vector of values sampled from one MVN distribution
could be closer to another MVN distribution.
### If two G matrices are identical, then q = 0.5.
### If two G matrices are completely unrelated, then q = 0.
### Nsamp refers to the number of samples drawn from a MVN distribution that is based on
G1.
q.Ova <- function(G1, G2, Nsamp){
X1 <- rmvnorm(n=Nsamp, sigma=G1, method="chol")
p1 <- apply(X1, MARGIN=1, FUN=dmvnorm, sigma=G1)
p2 <- apply(X1, MARGIN=1, FUN=dmvnorm, sigma=G2)
q.out <- p2/(p1 + p2)
q.summary <- c(mean(q.out), sd(q.out)/sqrt(Nsamp))
names(q.summary) <- c("q.mean", "q.se")
return(q.summary)
}

### The following function is used in PART 7.
### Function to compare the Blows' subspace of each teosinte and maize landrace
standardized G-matrices to the response Z.
blows.cal2 <- function(mat1, mat2, z1, z2, k, n){
  mat1val <- eigen(mat1)[[1]]
  mat2val <- eigen(mat2)[[1]]
  pve1 <- sum(mat1val[1:k])/sum(mat1val[mat1val>0])*100
  pve2 <- sum(mat2val[1:k])/sum(mat2val[mat2val>0])*100
  A1 <- eigen(mat1)[[2]][,1:k]
  A2 <- eigen(mat2)[[2]][,1:k]
  beta1 <- c(solve(mat1)%*%(z1/n))
  beta2 <- c(solve(mat2)%*%(z2/n))
  p1 <- c(A1%*%solve(t(A1)%*%A1)%*%t(A1)%*%beta1)
  p2 <- c(A2%*%solve(t(A2)%*%A2)%*%t(A2)%*%beta2)
  theta.beta1 <- acos(p1%*%beta1/(sqrt(sum(p1^2))*sqrt(sum(beta1^2))))*180/pi
  theta.beta2 <- acos(p2%*%beta2/(sqrt(sum(p2^2))*sqrt(sum(beta2^2))))*180/pi
  #project RESPONSE into the subspace
  p1.r <- c(A1%*%solve(t(A1)%*%A1)%*%t(A1)%*%z1)
  p2.r <- c(A2%*%solve(t(A2)%*%A2)%*%t(A2)%*%z2)
  #compute angle between response and its projection into G subspace
  theta.r1 <- acos(p1.r%*%z1/(sqrt(sum(p1.r^2))*sqrt(sum(z1^2))))*180/pi
```

```
  theta.r2 <- acos(p2.r%*%z1/(sqrt(sum(p2.r^2))*sqrt(sum(z2^2))))*180/pi
  #compute the angle between response vector and each eigenvector of G
  theta.r1.eig = acos(abs(t(z1)%*%A1/sqrt(sum(z1^2))))*180/pi
  theta.r2.eig = acos(abs(t(z2)%*%A2/sqrt(sum(z2^2))))*180/pi  #compute the angle between
beta and each eigenvector of G
  theta.b1.eig = acos(abs(t(beta1)%*%A1/sqrt(sum(beta1^2))))*180/pi
  theta.b2.eig = acos(abs(t(beta2)%*%A2/sqrt(sum(beta2^2))))*180/pi

  return(list(pve.mat1=pve1, pve.mat2=pve2, theta.beta1=c(theta.beta1),
theta.beta2=c(theta.beta2), theta.r1 = c(theta.r1),
              theta.r2 = c(theta.r2), theta.r1.eig = theta.r1.eig, theta.r2.eig =
theta.r2.eig, theta.b1.eig = theta.b1.eig, theta.b2.eig = theta.b2.eig) )
}

##############################################################################
### PART 1. Plot the proportions of phenotypic variance: additive/dominance/gxe. ###
### PART 2. Plot the proportions of genetic variance: additive/dominance.        ###
##############################################################################

### Prepare the data for plotting the proportions of phenotypic variances due to
additive/dominance/genetic-by-environment.
pheno.var <- varcomp[,c(1,2,3,10,11,12)]
names(pheno.var)[4:6] <- c("Va","Vd","Vge")
pheno.var <- melt(pheno.var, id.vars=c("Trait","Pop","Group"))

pheno.var$Trait <- as.factor(pheno.var$Trait)
pheno.var$Trait <-
factor(pheno.var$Trait,c("DTA","DTS","PLHT","LFLN","LFWD","TILN","PROL","LBNN","LBLN","LB
IL","EL","CUPR","ED","GE","EILN","TGPP","TGWP","GW"))

pheno.var$Pop <- as.factor(pheno.var$Pop)
pheno.var$Pop <- factor(pheno.var$Pop, c("Teosinte", "Maize Landrace"))

pheno.var$Group <- as.factor(pheno.var$Group)
pheno.var$Group <- factor(pheno.var$Group, c("Vegetative/Flowering Time","Environmental
Response","Reproductive"))

colnames(pheno.var)[4] <- "Variance"

### Prepare the data for plotting the proportions of genetic variances due to
additive/dominance.
gen.var <- varcomp[,c(1:5)]
gen.var[,4] <- gen.var[,4]/(varcomp[,4] + varcomp[,5])
gen.var[,5] <- gen.var[,5]/(varcomp[,4] + varcomp[,5])
names(gen.var)[4:5] <- c("Va","Vd")

gen.var <- melt(gen.var, id.vars=c("Trait","Pop","Group"))

gen.var$Trait <- as.factor(gen.var$Trait)
gen.var$Trait <-
factor(gen.var$Trait,c("DTA","DTS","PLHT","LFLN","LFWD","TILN","PROL","LBNN","LBLN","LBIL
","EL","CUPR","ED","GE","EILN","TGPP","TGWP","GW"))

gen.var$Pop <- as.factor(gen.var$Pop)
gen.var$Pop <- factor(gen.var$Pop, c("Teosinte", "Maize Landrace"))

gen.var$Group <- as.factor(gen.var$Group)
gen.var$Group <- factor(gen.var$Group, c("Vegetative/Flowering Time","Environmental
Response","Reproductive"))

colnames(gen.var)[4] <- "Variance"

#### Construct the plot.
pheno.var.plot <- ggplot(data=pheno.var, aes(x=Trait, y=value, fill=Variance)) +
geom_bar(stat="identity") +
scale_y_continuous(limits=c(0,1), breaks=c(0,0.25,0.50,0.75,1.00)) +
facet_grid(Pop~Group, scales="free_x", space="free_x", labeller=labeller(Pop=label_value,
Group=label_value)) +
theme(panel.grid=element_blank()) +
theme(strip.text=element_text(size=8), axis.text.x=element_text(angle=0, size=6)) +
theme(axis.title.x=element_text(size=10), axis.title.y=element_text(size=10)) +
```

```
ylab("Proportion of Phenotypic Variance")

### Construct the plot.
gen.var.plot <- ggplot(data=gen.var, aes(x=Trait, y=value, fill=Variance)) +
geom_bar(stat="identity") +
scale_y_continuous(limits=c(0,1), breaks=c(0,0.25,0.50,0.75,1.00)) +
facet_grid(Pop~Group, scales="free_x", space="free_x", labeller=labeller(Pop=label_value,
Group=label_value)) +
theme(panel.grid=element_blank()) +
theme(strip.text=element_text(size=8), axis.text.x=element_text(angle=0, size=6)) +
theme(axis.title.x=element_text(size=10), axis.title.y=element_text(size=10)) +
ylab("Proportion of Genetic Variance")

tiff("Figures/Fig02.tif", height=7.64, width=7, units="in", compression="lzw", res=1200)
grid.arrange(pheno.var.plot, gen.var.plot, layout_matrix=matrix(c(1,1,2),3,1))
dev.off()

#################################################################
### PART 3. Calculate the selection intensity for each trait. ###
#################################################################

### Set the number of generations of selection (assumption of 4500 or 9000 generations of
selection).
N1 <- 4500
N2 <- 9000

### Obtain the average change in additive genetic and phenotypic variance over N-
generations.
deltaVa1 <- (varcomp[19:36,4]-varcomp[1:18,4])/N1
deltaVp1 <- (varcomp[19:36,8]-varcomp[1:18,8])/N1
deltaVa2 <- (varcomp[19:36,4]-varcomp[1:18,4])/N2
deltaVp2 <- (varcomp[19:36,8]-varcomp[1:18,8])/N2

### Create dataframes to store our results.
tempdf <- data.frame(N=0:N1, h2_deltaVaVp_method=vector(length=N1+1),
Va_deltaVaVp_method=vector(length=N1+1), Vp_deltaVaVp_method=vector(length=N1+1))
h2vp1 <- replicate(18, tempdf, simplify=F)
tempdf <- data.frame(N=0:N2, h2_deltaVaVp_method=vector(length=N2+1),
Va_deltaVaVp_method=vector(length=N2+1), Vp_deltaVaVp_method=vector(length=N2+1))
h2vp2 <- replicate(18, tempdf, simplify=F)

### Note: Selection intensity, i=Z/(h2*sqrt(Vp))
### Obtain the starting values for h2*sqrt(Vp), Va and Vp.
for(i in 1:18){
h2vp1[[i]][1,2] <- varcomp[i,10]*sqrt(varcomp[i,8])
h2vp1[[i]][1,3] <- varcomp[i,4]
h2vp1[[i]][1,4] <- varcomp[i,8]

h2vp2[[i]][1,2] <- varcomp[i,10]*sqrt(varcomp[i,8])
h2vp2[[i]][1,3] <- varcomp[i,4]
h2vp2[[i]][1,4] <- varcomp[i,8]
}

### Calculate h2*sqrt(Vp), Va and Vp at every generation based on the average change in
variances.
for(i in 1:18){for(j in 2:(N1+1)){
h2vp1[[i]][j,2] <- (varcomp[i,4] + h2vp1[[i]][j,1]*deltaVa1[i])/sqrt(varcomp[i,8] +
h2vp1[[i]][j,1]*deltaVp1[i])
h2vp1[[i]][j,3] <- varcomp[i,4] + h2vp1[[i]][j,1]*deltaVa1[i]
h2vp1[[i]][j,4] <- varcomp[i,8] + h2vp1[[i]][j,1]*deltaVp1[i]
}}

for(i in 1:18){for(j in 2:(N2+1)){
h2vp2[[i]][j,2] <- (varcomp[i,4] + h2vp2[[i]][j,1]*deltaVa2[i])/sqrt(varcomp[i,8] +
h2vp2[[i]][j,1]*deltaVp2[i])
h2vp2[[i]][j,3] <- varcomp[i,4] + h2vp2[[i]][j,1]*deltaVa2[i]
h2vp2[[i]][j,4] <- varcomp[i,8] + h2vp2[[i]][j,1]*deltaVp2[i]
}}

### Create vectors to store the selection intensities.
i1 <- vector(length=18)
```

```
i2 <- vector(length=18)

### Calculate the selection intensities.
for(i in 1:18){
i1[i] <- Z[i]/sum(h2vp1[[i]][1:N1,2])
i2[i] <- Z[i]/sum(h2vp2[[i]][1:N2,2])
}

### Combine the selection intensities for two different Ns.
intensity <- data.frame(Trait=varcomp[1:18,1], intensity_N4500=i1, intensity_N9000=i2)

### Obtain selection intensities from wild species (Kingsolver et al. 2001, PubMed ID:
18707288, https://datadryad.org/resource/doi:10.5061/dryad.166/1)
wild <- read.delim("selection_intensity_wild.txt", header=T, as.is=T)

### Prepare the data for plotting our selection intensities against selection intensities
in wild species.
wild <- wild[!(wild[,1]==0),]
intensity$temp.y <- vector(length=18)
for(i in 1:18){intensity[i,4] <- 0.6/18*(19-i)+0.2}
intensity$group <- c(rep("Vegetative/Flowering Time", 5), rep("Environmental Response",
5), rep("Reproductive", 8))
intensity$group <- as.factor(intensity$group)
intensity$group <- factor(intensity$group, c("Environmental
Response","Vegetative/Flowering Time","Reproductive"))

### Construct the selection intensity plot.
tiff("Figures/selection_intensity.tiff", width=5.25, height=3, units="in", res=1200,
compression="lzw")
ggplot() +
geom_density(data=wild, aes(x=log10(Intensity)), colour="#7C26CB", fill="#7C26CB") +
xlab(expression(log[10]*(Selection ~ Intensity))) +
geom_segment(data=intensity, aes(x=log10(abs(intensity_N9000)),
xend=log10(abs(intensity_N4500)), y=temp.y, yend=temp.y, colour=group), size=3,
show.legend=F) +
geom_text(data=intensity, aes(x=log10(abs(intensity_N9000)), y=temp.y, label=Trait),
hjust=0,
nudge_x=c(0.06,0.06,0.03,0.04,0.02,0.05,0.03,0.03,0.04,0.06,0.10,0.02,0.08,0.08,0.05,0.03
,0.02,0.08), size=2) +
theme(panel.grid=element_blank(), axis.title=element_text(size=8),
axis.text=element_text(size=6))
dev.off()

### Prepare the data for plotting fold changes in trait means to accompany the selection
intensity plot.
fc.trait <- data.frame(trait=varcomp[1:18,1], group=varcomp[1:18,3],
fc=log2(varcomp[19:36,9]/varcomp[1:18,9]))

fc.trait$trait <- factor(fc.trait$trait,
c("DTA","DTS","PLHT","LFLN","LFWD","TILN","PROL","LBNN","LBLN","LBIL","EL","CUPR","ED","G
E","EILN","TGPP","TGWP","GW"))
fc.trait$group <- factor(fc.trait$group, c("Environmental Response","Vegetative/Flowering
Time","Reproductive"))

### Construct the plot of fold changes in trait means.
tiff("Figures/traitmean_fc.tiff", width=7, height=1.5, units="in", res=1200,
compression="lzw")
ggplot() +
geom_bar(data=fc.trait, stat="identity", aes(x=trait, y=fc, fill=group)) +
xlab("Trait") +
ylim(-6,6) +
scale_fill_discrete(name = "Trait Group") +
theme(panel.grid=element_blank(), axis.text=element_text(size=6),
axis.title=element_text(size=7)) +
theme(legend.text=element_text(size=7), legend.title=element_text(size=8)) +
ylab(expression(Fold ~ Change ~ log[2]*(mu[M]/mu[T])))
dev.off()


################################################################################
### PART 4. Show the relationship among trait groups (Vegetative/Flowering Time,   ###
```

```
### Environmental Response, Reproductive) via Principal Coordinate Analysis (PCoA) ###
### and Neighbor-Joining plots.                                                     ###
#########################################################################################
### Calculate the distance matrices for teosinte and maize landrace.
dT <- 1-abs(rT)
dT <- as.dist(dT, diag=F, upper=F)

dM <- 1-abs(rM)
dM <- as.dist(dM, diag=F, upper=F)

### Calculate the principal coordinate (PCoA), also known as multidimensional scaling
(MDS).
PCoA.T <- cmdscale(dT)
PCoA.M <- cmdscale(dM)

### Prepare the data for plotting PCoA/MDS.
PCoA.T <- data.frame(Classification=varcomp[1:18,3], x=PCoA.T[,1], y=PCoA.T[,2])
PCoA.T$Classification <- as.factor(PCoA.T$Classification)
PCoA.T$Classification <- factor(PCoA.T$Classification,c("Environmental Response",
"Vegetative/Flowering Time", "Reproductive"))

PCoA.M <- data.frame(Classification=varcomp[19:36,3], x=PCoA.M[,1], y=PCoA.M[,2])
PCoA.M$Classification <- as.factor(PCoA.M$Classification)
PCoA.M$Classification <- factor(PCoA.M$Classification,c("Environmental Response",
"Vegetative/Flowering Time", "Reproductive"))

### Make the PCoA plot for teosinte.
tiff("Figures/PCoA_teosinte.tiff", width=4.8, height=3, units="in", res=1200,
compression="lzw")
set.seed(89)
ggplot(data=PCoA.T) +
geom_point(aes(x, y), size = 1, color = "red") +
geom_label_repel(
    aes(x, y, fill = Classification, label = rownames(PCoA.T)),
    fontface = 'bold', color = 'white',
    box.padding = unit(0.35, "lines"),
    point.padding = unit(0.5, "lines"),
    min.segment.length = unit(0, "lines"),
    segment.color = 'grey50',
    size=2,
    ) +
xlab("Dimension 1") +
ylab("Dimension 2") +
theme_classic(base_size = 16) +
theme(legend.title = element_text(size=8, face="bold"), legend.text =
element_text(size=6)) +
theme(text = element_text(size=6))
dev.off()

### Make the PCoA plot for maize landrace.
tiff("Figures/PCoA_maize.tiff", width=4.8, height=3, units="in", res=1200,
compression="lzw")
set.seed(89)
ggplot(data=PCoA.M) +
geom_point(aes(x, y), size = 1, color = "red") +
geom_label_repel(
    aes(x, y, fill = Classification, label = rownames(PCoA.M)),
    fontface = 'bold', color = 'white',
    box.padding = unit(0.35, "lines"),
    point.padding = unit(0.5, "lines"),
    min.segment.length = unit(0, "lines"),
    segment.color = 'grey50',
    size=2,
    ) +
xlab("Dimension 1") +
ylab("Dimension 2") +
theme_classic(base_size = 16) +
theme(legend.title = element_text(size=8, face="bold"), legend.text =
element_text(size=6)) +
theme(text = element_text(size=6))
dev.off()
```

```
### Make a neighbor-joining (NJ) tree for teosinte.
njT <- list(nj(dT),varcomp[1:18,3])
njT[[2]] <- ifelse(njT[[2]]=="Environmental Response", "#F8766D",
ifelse(njT[[2]]=="Reproductive","#5C99FF",ifelse(njT[[2]]=="Vegetative/Flowering
Time","#00BA38",NA)))
tiff("SFig/NJ_teosinte.tiff", width=2000, height=2400, res=300, compression="lzw")
plot(njT[[1]], tip.color=njT[[2]], type="unrooted", lab4ut="axial")
dev.off()

### Make a neighbor-joining (NJ) tree for maize landrace.
njM <- list(nj(dM), varcomp[19:36,3])
njM[[2]] <- ifelse(njM[[2]]=="Environmental Response", "#F8766D",
ifelse(njM[[2]]=="Reproductive","#5C99FF",ifelse(njM[[2]]=="Vegetative/Flowering
Time","#00BA38",NA)))
tiff("SFig/NJ_maize.tiff", width=2000, height=2400, res=300, compression="lzw")
plot(njM[[1]], tip.color=njM[[2]], type="unrooted", lab4ut="axial")
dev.off()


##############################################################################################
######
### PART 5. Comparing the teosinte and maize landrace genetic correlations matrix and the
###
### three submatrices (Vegetative/Flowering Time, Environmental Response, and
Reproductive) ###
### via Mantel test and difference in their leading eigenvectors.
###
##############################################################################################
######

### Compare the full genetic correlations matrix (CUPR and TGPP excluded) from teosinte
and maize landrace.
### Mantel test
cor(rT.re[lower.tri(rT.re)], rM.re[lower.tri(rM.re)]); mantel.test(rT.re, rM.re,
nperm=10000)
#r=0.5091904
#Z=4.966196
#p=9.999e-05

### Leading eigenvectors
theta.cal(rT.re, rM.re)
#    label eigenvector1 eigenvector2
#    theta     89.61667     88.35276
# pve.mat1     27.24433     18.70337
# pve.mat2     19.10335     14.27192
#note that mat1vec1 and mat2vec1 are teosinte and maize landrace gmax respectively.

### Compare the vegetative genetic correlations matrix from teosinte and maize landrace.
### Mantel test
cor(vegT[lower.tri(vegT)], vegM[lower.tri(vegM)]); mantel.test(vegT, vegM, nperm=10000)
#r=0.899856
#Z=1.267712
#p=0.02289661

###Leading eigenvectors
theta.cal(vegT, vegM)
#    label eigenvector1 eigenvector2
#    theta     28.05962     50.62398
# pve.mat1     64.83790     15.67222
# pve.mat2     40.64529     21.12285

### Compare the environmental response genetic correlations matrix from teosinte and
maize landrace.
### Mantel test
cor(enrT[lower.tri(enrT)], enrM[lower.tri(enrM)]); mantel.test(enrT, enrM, nperm=10000)
#r=0.7707012
#Z=1.23834
#p=0.01469853

### Leading eigenvectors
```

```
theta.cal(enrT, enrM)
#    label eigenvector1 eigenvector2
#    theta     27.37998     27.74151
# pve.mat1     45.61217     26.91704
# pve.mat2     43.88581     24.59770


### Compare the reproductive genetic correlations matrix from teosinte and maize
landrace.
### Mantel test
cor(repT.re[lower.tri(repT.re)], repM.re[lower.tri(repM.re)]); mantel.test(repT.re,
repM.re, nperm=10000)
#r=0.7873143
#Z=1.882289
#p=0.00289971

### Leading eigenvectors
theta.cal(repT.re, repM.re)
#    label eigenvector1 eigenvector2
#    theta     87.81169     86.67189
# pve.mat1     54.14230     29.19786
# pve.mat2     41.97578     26.11113


### Construct a plot of the genetic correlation matrix.
corG.plot <- melt(corG, na.rm=T)
corG.se.plot <- melt(corG.se, na.rm=T)

corG.plot$Z <- atanh(abs(corG.plot[,3]))/atanh(corG.se.plot[,3])
corG.plot$label <- as.character(round(corG.plot[,3], 2))
corG.plot[nchar(corG.plot[,5])==1,5] <-
paste(corG.plot[nchar(corG.plot[,5])==1,5],".00",sep="")
corG.plot[nchar(corG.plot[,5])==3,5] <-
paste(corG.plot[nchar(corG.plot[,5])==3,5],"0",sep="")
corG.plot[grep("-",corG.plot[,5]),5][nchar(corG.plot[grep("-",corG.plot[,5]),5])==4] <-
paste(corG.plot[grep("-",corG.plot[,5]),5][nchar(corG.plot[grep("-
",corG.plot[,5]),5])==4], "0", sep="")
corG.plot[corG.plot[,4] < 1.96 ,5] <- NA #standard normal, two-tailed cutoff for P <
0.05.

temp.plot <- 0.5
for(i in 1:18){temp.plot <- c(temp.plot,i+0.5)}

tiff("Figures/corG_plot.tif", height=6.3, width=7, units="in", res=1200,
compression="lzw")
ggplot() +
geom_point(data=corG.plot, aes(x=Var2, y=Var1, colour=value), shape=15, size=10) +
geom_point(aes(x=1:18, y=18:1), shape=15, colour="black", size=10) +
scale_y_discrete(limits = rev(levels(corG.plot$Var1))) +
scale_x_discrete(position="top") +
scale_colour_gradient2(low="red", mid="white", high="blue", limits=c(-1,1), breaks=c(-
1,0,1), name=expression(bolditalic('r'[g]))) +
theme(panel.grid.major = element_blank()) +
theme(panel.background = element_rect(fill = "white", colour = "white")) +
theme(axis.text.x = element_text(angle=90, hjust=0)) +
theme(axis.ticks = element_blank()) +
geom_vline(xintercept=temp.plot, colour="grey") +
geom_hline(yintercept=temp.plot, colour="grey") +
geom_vline(xintercept=c(5.5,10.5), colour="black", size=1) +
geom_hline(yintercept=c(8.5,13.5), colour="black", size=1) +
xlab("Maize Landrace") +
ylab("Teosinte") +
theme(axis.title=element_text(size=18, face="bold")) +
theme(axis.text.x=element_text(colour=c(rep("#00BA38",5),rep("#F8766D",5),rep("#5C99FF",8
)), size=12, face="bold")) +
theme(axis.text.y=element_text(colour=c(rep("#5C99FF",8),rep("#F8766D",5),rep("#00BA38",5
)), size=12, face="bold")) +
guides(size=FALSE) +
geom_text(data=corG.plot, aes(x=Var2, y=Var1, label=label), size=2.5, nudge_y=0.25)
dev.off()
```

```
##############################################################################################
######
### PART 6. Comparing the teosinte and maize landrace genetic covariances matrix and the
###
### three submatrices (Vegetative/Flowering Time, Environmental Response, and
Reproductive) ###
### via Mantel test, Flury hierarchy, and random skewers.
###
##############################################################################################
######

################################ NOTES on using the CPC software
################################
# Many of the CPC versions have compatibilities issue (Windows, Mac, etc.)
#
# We used the Linux version
(https://pages.uoregon.edu/pphil/programs/cpc/linux/cpc.tar.gz).
#
# CPC requires the matrices to be PD; we only used the nearPD function when necessary,
i.e. non-PD matrices.         #
# After we export the ".dat" file, we manually edit the file into the following format
(description in parentheses): #
# 2 (number of matrices)
#
# 16 (number of traits, 5/5/8 for submatrices)
#
# 4455 (number of teosinte individuals)
#
# GT (teosinte G-matrix)
#
# 4398 (number of maize landrace individuals)
#
# GM (maize landrace G-matrix)
#
##############################################################################################
##############################

### Compare the FULL genetic covariances matrix from teosinte and maize landrace.
### Mantel test
cor(GT.re[lower.tri(GT.re)], GM.re[lower.tri(GM.re)]); mantel.test(GT.re, GM.re,
nperm=10000)
#r=0.02511114
#Z=44981.04
#p=0.2144786

### Flury hierarchy (Prepare file for CPC)
combined.G <- rbind(GT.re.PD, GM.re.PD)
write.table(combined.G, "cpc_covG.dat", row.names=F, col.names=F, quote=F, sep=" ")

### Random Skewers
skewers(GT.re, GM.re, nsim=1000, method="unifcorrmat")
#r=0.1853682
#p=1

### Ovaskainen's test
#Comparing full G-matrices between teosinte and maize landrace.
q.Ova(G1=GT.re.PD, G2=GM.re.PD, Nsamp=1000)
#q.mean   q.se
#     0      0


### Compare the VEGETATITVE genetic covariances matrix from teosinte and maize landrace.
### Mantel test
cor(vegGT[lower.tri(vegGT)], vegGM[lower.tri(vegGM)]); mantel.test(vegGT, vegGM,
nperm=10000)
#r=0.9011552
#Z=1190.493
#p=0.05169483

### Flury hierarchy (Prepare file for CPC)
```

```
combined.vegG <- rbind(vegGT.PD, vegGM.PD)
write.table(combined.vegG, "cpc_covG_veg.dat", row.names=F, col.names=F, quote=F, sep="
")

### Random Skewers
skewers(vegGT, vegGM, nsim=1000, method="unifcorrmat")
#r=0.8925812
#p=0.001

### Ovaskainen's test
#Comparing Vegetative submatrices between teosinte and maize landrace.
q.Ova(G1=vegGT.PD, G2=vegGM.PD, Nsamp=1000)
#     q.mean        q.se
#0.306186899 0.006744435


### Compare the ENVIRONMENTAL RESPONSE genetic covariances matrix from teosinte and maize
landrace.
### Mantel test
cor(enrGT[lower.tri(enrGT)], enrGM[lower.tri(enrGM)]); mantel.test(enrGT, enrGM,
nperm=10000)
#r=0.9624238
#Z=38991.71
#p=0.05029497

### Flury hierarchy (Prepare file for CPC)
combined.enrG <- rbind(enrGT.PD, enrGM.PD)
write.table(combined.enrG, "cpc_covG_enr.dat", row.names=F, col.names=F, quote=F, sep="
")

### Random Skewers
skewers(enrGT, enrGM, nsim=1000, method="unifcorrmat")
#r=0.8027262
#p=0.01

### Ovaskainen's test
#Comparing Environmental Response submatrices between teosinte and maize landrace.
q.Ova(G1=enrGT.PD, G2=enrGM.PD, Nsamp=1000)
#       q.mean          q.se
#2.246688e-100 2.246688e-100


### Compare the REPRODUCTIVE genetic covariances matrix from teosinte and maize landrace.
### Mantel test
cor(repGT.re[lower.tri(repGT.re)], repGM.re[lower.tri(repGM.re)]); mantel.test(repGT.re,
repGM.re, nperm=10000)
#r=0.1406578
#Z=4599.029
#p=0.3086691

### Flury hierarchy (Prepare file for CPC)
combined.repG <- rbind(repGT.PD, repGM.PD)
write.table(combined.repG, "cpc_covG_rep.dat", row.names=F, col.names=F, quote=F, sep="
")

### Random Skewers
skewers(repGT.re, repGM.re, nsim=1000, method="unifcorrmat")
#r=0.08677437
#p=1

### Ovaskainen's test
#Comparing Reproductive submatrices between teosinte and maize landrace.
q.Ova(G1=repGT.PD, G2=repGM.PD, Nsamp=1000)
#     q.mean        q.se
#1.09343e-12 1.08513e-12


##############################################################################################
#####
### PART 7. Comparing Blows' subspace (1-5 eigenvectors) for teosinte and maize landrace
G ###
```

```
################################################################################################
#####

### Compare Blows subspace of leading two eigenvectors to beta.
blows.cal2(rT.re, rM.re, Z.stdT.re, Z.stdM.re, 2, 9000)
# $`pve.mat1`
# [1] 45.9477
#
# $pve.mat2
# [1] 33.37526
#
# $theta.beta1
# [1] 89.97862
#
# $theta.beta2
# [1] 89.65813
#
# $theta.r1
# [1] 64.67802
#
# $theta.r2
# [1] 99.53384
#
# $theta.r1.eig
# [,1]     [,2]
# [1,] 67.27738 79.41768
#
# $theta.r2.eig
# [,1]    [,2]
# [1,] 74.26973 59.9559
#
# $theta.b1.eig
# [,1]     [,2]
# [1,] 89.98242 89.98783
#
# $theta.b2.eig
# [,1]     [,2]
# [1,] 89.87179 89.68308

### Compare Blows subspace of leading three eigenvectors to beta.
blows.cal2(rT.re, rM.re, Z.stdT.re, Z.stdM.re, 3, 9000)
# $`pve.mat1`
# [1] 61.23287
#
# $pve.mat2
# [1] 45.60396
#
# $theta.beta1
# [1] 89.97845
#
# $theta.beta2
# [1] 89.54612
#
# $theta.r1
# [1] 64.59429
#
# $theta.r2
# [1] 78.68411
#
# $theta.r1.eig
# [,1]     [,2]     [,3]
# [1,] 67.27738 79.41768 88.07245
#
# $theta.r2.eig
# [,1]    [,2]    [,3]
# [1,] 74.26973 59.9559 66.1641
#
# $theta.b1.eig
# [,1]     [,2]     [,3]
# [1,] 89.98242 89.98783 89.99727
#
```

```
# $theta.b2.eig
# [,1]     [,2]     [,3]
# [1,] 89.87179 89.68308 89.70146

### Compare Blows subspace of leading four eigenvectors to beta.
blows.cal2(rT.re, rM.re, Z.stdT.re, Z.stdM.re, 4, 9000)
# $`pve.mat1`
# [1] 70.96228
#
# $pve.mat2
# [1] 57.06434
#
# $theta.beta1
# [1] 89.92216
#
# $theta.beta2
# [1] 89.52534
#
# $theta.r1
# [1] 43.35824
#
# $theta.r2
# [1] 74.09296
#
# $theta.r1.eig
# [,1]     [,2]     [,3]     [,4]
# [1,] 67.27738 79.41768 88.07245 54.05518
#
# $theta.r2.eig
# [,1]    [,2]    [,3]     [,4]
# [1,] 74.26973 59.9559 66.1641 79.85031
#
# $theta.b1.eig
# [,1]     [,2]     [,3]     [,4]
# [1,] 89.98242 89.98783 89.99727 89.92521
#
# $theta.b2.eig
# [,1]     [,2]     [,3]     [,4]
# [1,] 89.87179 89.68308 89.70146 89.86109

### Compare Blows subspace of leading five eigenvectors to beta.
blows.cal2(rT.re, rM.re, Z.stdT.re, Z.stdM.re, 5, 9000)
#$`pve.mat1`
#[1] 78.22545
#
# $pve.mat2
# [1] 66.42401
#
# $theta.beta1
# [1] 89.91775
#
# $theta.beta2
# [1] 89.46881
#
# $theta.r1
# [1] 41.96595
#
# $theta.r2
# [1] 75.14963
#
# $theta.r1.eig
# [,1]     [,2]     [,3]     [,4]     [,5]
# [1,] 67.27738 79.41768 88.07245 54.05518 81.04738
#
# $theta.r2.eig
# [,1]    [,2]    [,3]     [,4]     [,5]
# [1,] 74.26973 59.9559 66.1641 79.85031 75.69734
#
# $theta.b1.eig
# [,1]     [,2]     [,3]     [,4]     [,5]
# [1,] 89.98242 89.98783 89.99727 89.92521 89.97344
```

```
#
# $theta.b2.eig
# [,1]     [,2]     [,3]     [,4]     [,5]
# [1,] 89.87179 89.68308 89.70146 89.86109 89.76155

##############################################################################################
########
# PART 8. Multivariate Qst-Fst Test using R scripts by Martin et al. (2008) PubMed ID:
18245845 #
##############################################################################################
########

### Step 5.1: Preparing additional files.
# Within-population G-matrix
Gw <- matrix(NA,18,18)
for(i in 1:18){for(j in 1:18){
Gw[i,j] <- mean(c(GT[i,j],GM[i,j]))
}}

# Between-population G-matrix
Gb <- matrix(NA,18,18)
for(i in 1:18){for(j in 1:18){
Gb[i,j] <- 0.5*Z[i]*Z[j]
}}

# Within-population additive genetic variance
Vw <- diag(Gw)

# Between-population additive genetic variance
Vb <- diag(Gb)

# Bent within- and between-population G-matrices to be positive definite
Gw <- as.matrix(nearPD(Gw, corr=F)[[1]])
Gb <- as.matrix(nearPD(Gb, corr=F)[[1]])

# "genoTM.txt" is a merged genotype file of both teosinte and maize landrace.
# Only common markers between teosinte and maize landrace can be used here.
# The file is coded such as row is individual, column is GBS marker.
# This file can be obtained by exporting the genotype file as "Table" format in TASSEL.
geno <- read.delim("genoTM.txt", header=T, as.is=T)

# Read in the marker names for the "genoTM.txt" file.
marker.names <- read.delim("markers.txt", header=F, as.is=T)

# Create a column for the population identifier in geno; top 4398 are maize landraces;
bottom 4455 are teosinte.
geno <- data.frame(Pop=c(rep(1,4398),rep(2,4455)), geno)

# Rename the columns
colnames(geno) <- c("Pop", "Individual", marker.names[,1])

# Recode genotypes into format that hierfstat recognizes
geno[geno=="A"] <- 11
geno[geno=="C"] <- 22
geno[geno=="G"] <- 33
geno[geno=="T"] <- 44
geno[geno=="R"] <- 13
geno[geno=="Y"] <- 24
geno[geno=="K"] <- 34
geno[geno=="M"] <- 12
geno[geno=="W"] <- 14
geno[geno=="S"] <- 23
geno[geno=="N"] <- "NA"


### Step 5.2: Calculate Fst
# WARNING: THIS STEP TAKES A VERY LONG TIME
# For a dataset of 8853 individuals and 21,157 markers, the following script took ~22hr
fst.dat <- varcomp.glob(as.factor(geno[,1]), geno[,-c(1:2)])
```

```
# It is advisable to save everything at this point to avoid redoing lengthy Fst
calculation.
save.image("temp.RData")


### Step 5.3: Bootstrapping Fst
### Fst bootstrapping function from Whitlock and Guillaume (2009)

fst.sample <- function(obs, nloci){
loc.smpl <- sample(1:nloci,size=nloci,replace=TRUE)
dat <- obs[loc.smpl,]                                  #select the sampled loci from the
input table
return( sum(dat[,1])/sum(dat[,1]+dat[,2]+dat[,3]) )    # Fst = a/(a+b+c); from
Weir&Cockerham 1984
}

# Set the bootstrap parameters
nboot <- 10000
nloci <- length(fst.dat[,1])
fst.est  = vector(length = nboot)

# Bootstrapping by sampling nloci from the neutral markers, with replacement
for(i in 1:nboot){
fst.repl = fst.sample(fst.dat, nloci)
fst.est[i] = fst.repl
}

# Mean of Fst
fst.mean <- mean(fst.est, na.rm=TRUE)
# 0.1567261

# Standard deviation of Fst
fst.sd <- sd(fst.est, na.rm=TRUE)
# 0.001606596

# 95% CI of Fst
fst.CI <- quantile(fst.est,c(0.025,0.975), na.rm=TRUE)
#      2.5%     97.5%
# 0.1536479 0.1598973

# mean of 2*Fst/(1-Fst)
2*fst.mean/(1-fst.mean)
# 0.3717086

# 95% CI of 2*Fst/(1-Fst)
c(2*fst.CI[1]/(1-fst.CI[1]),2*fst.CI[2]/(1-fst.CI[2]))
#      2.5%     97.5%
# 0.3630828 0.3806612


### Step 5.4: Calculating Qst

# Load in the Qst scripts by Martin et al (2008)
# Require the package corpcor.
# Link to the scripts as of 22 October 2018; http://mbb.univ-
montp2.fr/MBB/uploads/codes_Qst_Fst.rar
source("package neutrality test.r")

# Number of traits
ntrt <- length(Gw[,1])

# Number of individuals in teosinte and maize landrace
nind <- c(4455,4398)

# Number of populations
npop <- 2

# df for unbalanced design (Equation 9 of Martin et al (2008))
nf <- mean(nind)-1/npop*((mean(nind^2)-mean(nind)^2)/mean(nind))

# Back-calculate within-population mean square from Gw
```

```
MSw <- Gw

# Back-calculate between-population mean square from Gb
MSb <- nf*Gb + Gw

# df for within-population
dfw <- nind[1] + nind[2] - npop - (ntrt-1)^2

# df for between-population
dfb <- npop - 1

# Combine the inputs for calculating Qst
DF <- c(dfw,dfb)
G <- list(Gw,Gb)
MS <- list(MSw,MSb)

# Compute rho from test between G matrices, under neutrality: rho=2*Fst/(1-Fst)
testG <- k.prop(DF,G)
cat("rho_P (population) =", testG$rho[[2]], "\n95% CI for rho_P :", testG$CI[[2]], "\n")
#rho_P (population) = 314.6347
#95% CI for rho_P : 190.3003 907.6686

# Compute p-values from test between MS matrices
testMS <- k.prop(DF,MS)
cat("pBartlett=", testMS$pt1, "\npChi2=", testMS$pX,"\n")
#pBartlett= 0.4680047
#pChi2= 0.4680044


##############################################################################################
########
### PART 9. Univariate Qst-Fst Test using R scripts provided by Whitlock and Guillaume
(2009) PubMed ID: 19687138 ###
##############################################################################################
########
# Use the Fst calculated from PART 8.
Fst.dat <- fst.dat$loc

# Calculate the observed Fst
Fst.obs <- sum(Fst.dat[,1])/sum(Fst.dat[,1] + Fst.dat[,2] + Fst.dat[,3])
#0.1567383

# Calculate the observed Qst
Qst.obs <- Vb/(Vb + 2*Vw)

# [1] 0.5681129 0.7822637 0.9455266 0.9393080 0.9301162 0.9457644 0.9503338
# [8] 0.9617537 0.9187057 0.9385364 0.9731977 0.9698326 0.9865557 0.9715125
#[15] 0.9823915 0.9599634 0.0751920 0.9768512


#####################################################################
### Several functions for bootstrapping Qst, Fst, and Qst-Fst. ###
#####################################################################

### Qst bootstrapping function from Whitlock and Guillaume (2009)
qst.sample <- function(VarW, VarB, dfW, dfB){
VarW.sim <- VarW*rchisq(1, dfW)/dfW
VarB.sim <- VarB*rchisq(1, dfB)/dfB
return(VarB.sim/(VarB.sim+2*VarW.sim))
}

### Fst bootstrapping function from Whitlock and Guillaume (2009)
fst.sample <- function(obs, nloci){
loc.smpl <- sample(1:nloci,size=nloci,replace=TRUE)
dat <- obs[loc.smpl,]                                  #select the sampled loci from the
input table
return( sum(dat[,1])/sum(dat[,1]+dat[,2]+dat[,3]) )    # Fst = a/(a+b+c); from
Weir&Cockerham 1984
}

### Qst-Fst bootstrapping function
```

```
qstfst.bootstrap <- function(fst.dat, Vw, Vb, dfw, dfb, qst.obs, fst.obs){
# Set the bootstrap parameters
nboot <- 10000
nloci <- length(fst.dat[,1])
qst.est <- vector(length = nboot)
fst.est <- vector(length = nboot)
boot.est <- vector(length = nboot)

# Bootstrapping Qst and Fst
for(i in 1:nboot){
qst.repl <- qst.sample(Vw, Vb, dfw, dfb)
fst.repl <- fst.sample(fst.dat, nloci)
qst.est[i] <- qst.repl
fst.est[i] <- fst.repl
boot.est[i] = qst.repl - fst.repl
}

QstFst.obs <- qst.obs - fst.obs

# Summarizing results
bootstrap.results <- list(
    QstObs       = qst.obs,
    FstObs       = fst.obs,
    QstFstObs    = QstFst.obs,
    Pvalue       = 2*min(sum(boot.est < QstFst.obs),sum(boot.est > QstFst.obs))/nboot,
    boot         = mean(boot.est, na.rm=TRUE),
    boot.stdev   = sd(boot.est, na.rm=TRUE),
    boot.CI      = quantile(boot.est,c(0.025,0.975), na.rm=TRUE),
    Fst          = mean(fst.est, na.rm=TRUE),
    Fst.stdev    = sd(fst.est, na.rm=TRUE),
    Fst.CI       = quantile(fst.est,c(0.025,0.975), na.rm=TRUE),
    Qst          = mean(qst.est, na.rm=TRUE),
    Qst.stdev    = sd(qst.est, na.rm=TRUE),
    Qst.CI       = quantile(qst.est,c(0.025,0.975), na.rm=TRUE) )
return(bootstrap.results)
}

all.results <- list()

# Use the above function to bootstrap Qst-Fst.
for(i in 1:ncol(Gw)){
all.results[[i]] <- qstfst.bootstrap(Fst.dat, Vw[i], (2*Fst.obs*Vw[i])/(1-Fst.obs), 4455
+ 4398 - 2, 2 - 1, Qst.obs[i], Fst.obs)
}

# Qst-Fst bootstrapping results for all 18 traits.
for(i in 1:18){print(all.results[[i]][[4]])}
#[1] 0.016
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0
#[1] 0.986
#[1] 0

fst.plot <- all.results[[1]][[14]]
for(i in 2:18){fst.plot <- c(fst.plot, all.results[[i]][[14]])}


fst.plot <- fst.dat[,1]/(fst.dat[,1]+fst.dat[,2]+fst.dat[,3])
```

```
fst.plot[fst.plot<0] <- 0

names(Qst.obs) <- names(Z)
qst.plot <- data.frame(qst=Qst.obs, group=varcomp[1:18,3], xpos=NA, ypos=NA)
qst.plot$group <- as.factor(qst.plot$group)
qst.plot$group <- factor(qst.plot$group, levels=c("Environmental
Response","Vegetative/Flowering Time","Reproductive"))
qst.plot <- qst.plot[order(qst.plot$qst),]
qst.plot$xpos <- c(0.2,0.52,0.73,rep(0.89,15))
qst.plot$ypos <- c(9,seq(1,9,0.5))

tiff("Figures/fstqst.tiff", width=7, height=4, units="in", res=1200, compression="lzw")
ggplot() +
geom_density(aes(fst.plot), colour="#A9A9A9", fill="#A9A9A9") +
geom_segment(data=qst.plot, aes(x=qst,y=0,xend=qst,yend=ypos,colour=group),
show.legend=F, size=0.5) +
geom_label(data=qst.plot,
        aes(x=xpos, y=ypos, label=rownames(qst.plot), fill=group),
        nudge_y=c(0,0.5,0.5,rep(0,15)),
        show.legend=T,
        hjust=1,
        fontface = 'bold',
        color = 'white',
      size = 2.5) +
geom_segment(aes(x=median(fst.plot), y=0, xend=median(fst.plot), yend=10), linetype=2) +
theme(axis.text.y=element_blank(), axis.ticks.y=element_blank()) +
theme(panel.background=element_blank(), panel.grid=element_blank()) +
theme(legend.text=element_text(size=6), legend.title=element_text(size=8)) +
theme(axis.text.x=element_text(size=5), axis.title=element_text(size=7)) +
xlab("Fst") +
ylab("Frequency") +
scale_x_continuous(limits=c(0,1)) +
scale_y_continuous(expand=c(0,0)) +
expand_limits(y=0)
dev.off()

##########################################################
### PART 10. Gmax: Genetic Lines of Least Resistance. ###
##########################################################

### Gmax is calculated as the first eigenvector of a G-matrix (Schluter 1996, Pubmed ID:
28565589)
### Gmax is sensitive to magnitudes of traits, so we use the genetic correlation matrices
instead of G-matrices.
### Genetic correlation matrix is essentially G-matrices where each trait is standardized
by its genetic standard deviation.

### Teosinte Gmax and its percent variance explained (PVE).
GmaxT <- eigen(rT.re)[[2]][,1]

eigenvalueT <- eigen(rT.re)[[1]]
eigenvalueT[1]/sum(eigenvalueT[eigenvalueT>0])*100
#27.24433

### Maize landrace Gmax and its percent variance explained (PVE).
GmaxM <- eigen(rM.re)[[2]][,1]

eigenvalueM <- eigen(rM.re)[[1]]
eigenvalueM[1]/sum(eigenvalueM[eigenvalueM>0])*100
#19.10335

### Genetic constrain, as measured by angle between evolution trajectory (Z) and Gmax
### Theta for teosinte
thetaT <- acos(GmaxT%*%Z.stdT.re/sqrt(sum(Z.stdT.re^2)))*180/pi
cat(ifelse(thetaT>90, 180-thetaT, thetaT), "\n")
#67.27738

### Theta for maize landrace
thetaM <- acos(GmaxM%*%Z.stdM.re/sqrt(sum(Z.stdM.re^2)))*180/pi
cat(ifelse(thetaM>90, 180-thetaM, thetaM), "\n")
#74.26973
```

```
################################################################
### PART 11. Extension to Gmax: individual trait constraints. ###
################################################################

### Create a data frame to store the results from dropone analysis
dropone <- data.frame(Trait=varcomp.re[1:16,1], thetaD=NA)

### Calculate the theta (genetic constraint) from dropping one trait at a time.
for(i in 1:length(dropone[,1])){
temp.rT <- rT.re[-i,-i]
temp.GmaxT <- eigen(temp.rT)[[2]][,1]
temp.Z.stdT <- Z.stdT.re[-i]/sqrt(sum(Z.stdT.re[-i]^2))
dropone[i,2] <- acos(temp.GmaxT%*%temp.Z.stdT)*180/pi
dropone[i,2] <- ifelse(dropone[i,2]>90, 180-dropone[i,2], dropone[i,2])
}

### Compare the theta from dropone vs full 16-traits (teosinte).
dropone$diff <- dropone[,2]-67.27738

### Display the results from dropone analysis
print(dropone[order(dropone$diff),], row.names=F)

# The higher the decrease in theta, the more constraint that specific trait contributes.
# The higher the increase in theta, the more the specific trait aligns with desired
evolution.

# Trait   thetaD        diff
#    GE 62.82114 -4.45623585
#   DTA 63.22377 -4.05360998
#   DTS 63.35238 -3.92500055
#  PLHT 64.62573 -2.65165476
#  LFLN 65.03726 -2.24011950
#  LFWD 66.34697 -0.93041221
#  LBIL 66.52490 -0.75248410
#  PROL 67.00083 -0.27654897
#  LBLN 67.19671 -0.08066613
#  EILN 67.20138 -0.07599899
#  TILN 67.31564  0.03826208
#  LBNN 67.39876  0.12137953
#    GW 69.17319  1.89580571
#    ED 69.71071  2.43332908
#  TGWP 70.60707  3.32969345
#    EL 71.41978  4.14240418


################################################################################################
#######
### PART 12. What-if analysis: investigate how far can we get from single trait selection
on ###
### G-matrix, and the associated constrains. Note that comparison of this to gmax is
skipped.###
################################################################################################
#######

### Normalize the evolutionary trajectory (Z-vector) into length 1.
unit.Z.stdT <- Z.stdT.re/sqrt(sum(Z.stdT.re^2))

### Create a data frame to summarize the results from what-if analysis.
whatif <- data.frame(Trait=varcomp.re[1:16,1], Direction=Z.re/abs(Z.re),
ScalarProj1=vector(length=16), Theta1=vector(length=16),
ScalarProjGmax=vector(length=16), ThetaGmax=vector(length=16))

### Calculate the scalar projection and theta (constraint) between the response and
evolutionary trajectory (Z) when we select on ith trait (Column 3,4).
### Calculate the scalar projection and theta (constraint) between the response and Gmax
(teosinte) when we select on ith trait (Column 5,6).
for(i in 1:16){
GBeta <- whatif[i,2]*rT.re[,i]
unit.GBeta <- GBeta/sqrt(sum(GBeta^2))
```

```
whatif[i,3] <- GBeta%*%unit.Z.stdT
whatif[i,4] <- acos(unit.GBeta%*%unit.Z.stdT)*180/pi
whatif[i,5] <- GBeta%*%GmaxT
temp <- acos(unit.GBeta%*%GmaxT)*180/pi
whatif[i,6] <- ifelse(temp > 90, 180-temp, temp)
}

# Trait Direction ScalarProj1   Theta1 ScalarProjGmax ThetaGmax
#   DTA         1  0.30837688 80.39072     -1.5836123  30.99303
#   DTS         1  0.27975367 81.13283     -1.5434593  31.73983
#  PLHT         1  0.34050696 77.86373     -1.4322251  27.83552
#  LFLN         1  0.34843638 77.23302     -1.3835701  28.65863
#  LFWD         1  0.45225311 73.07725     -1.3344259  30.80987
#  TILN        -1  0.09386849 84.84097     -0.1823446  79.94024
#  PROL        -1  0.17160613 82.89246      0.1885700  82.18567
#  LBNN         1  0.26004793 79.86122     -0.8104191  56.72945
#  LBLN        -1  0.05404702 87.98209     -0.4160557  74.27254
#  LBIL        -1  0.17077299 83.63877     -0.8623796  55.97834
#    EL         1  0.97179312 53.76805     -1.2531874  40.34123
#    ED         1  0.49884978 72.23520     -0.8480237  58.75661
#    GE         1  1.07471697 45.55443     -1.1191808  43.18049
#  EILN        -1  0.20906572 80.99600      0.1391024  84.02296
#  TGWP         1  0.84177340 63.30709     -1.6329615  29.37568
#    GW         1  0.42008161 75.87721     -0.7302340  64.90313


### Plot the results from dropone and what-if analyses.
constraint <- merge(whatif, dropone, by="Trait", all=T, sort=F)

constraint$Trait <- as.factor(constraint$Trait)
constraint$Trait <- factor(constraint$Trait,
c("DTA","DTS","PLHT","LFLN","LFWD","TILN","PROL","LBNN","LBLN","LBIL","EL","ED","GE","EIL
N","TGWP","GW"))

constraint$Classification <- varcomp.re[1:16,3]
constraint$Classification <- as.factor(constraint$Classification)
constraint$Classification <- factor(constraint$Classification, levels=c("Environmental
Response", "Vegetative/Flowering Time", "Reproductive"))

### First, we plot the vector R (from whatif) against Z (evolutionary trajectory).
plotZ <- constraint[,c(1,3,4,9)]
names(plotZ)[2] <- "x"
plotZ$y <- plotZ[,2]*tan(plotZ[,3]*pi/180)
plotZ <- plotZ[,c(1,4,2,5)]
plotZ <- rbind(plotZ,plotZ)
plotZ[17:32,3] <- 0
plotZ[17:32,4] <- 0

plotZ.proj <- plotZ
plotZ.proj$y <- 0

temp.ends <- rep("last",16)
temp.ends2 <- rep("last",16)

tiff("Figures/constraintPlot_v2_thetaZ.tiff", units="in", height=3, width=3, res=1200,
compression="lzw")
ggplot() +
geom_line(aes(x=c(0,1.5),y=c(0,0)), size=0.75, arrow=arrow(length=unit(0.15,"cm"),
ends="last", type="open")) +
geom_line(data=plotZ, aes(x=x,y=y,group=Trait,colour=Classification), size=0.75, arrow =
arrow(length=unit(0.15,"cm"), ends=temp.ends, type="open")) +
geom_line(data=plotZ.proj, aes(x=x, y=y, group=Trait, colour=Classification), size=0.75,
arrow=arrow(length=unit(0.15,"cm"), ends=temp.ends2, type="open")) +
xlim(c(0,1.5)) +
ylim(c(0,2.0)) +
theme(axis.line=element_blank(), axis.text=element_blank(), axis.ticks=element_blank(),
axis.title=element_blank(), panel.background = element_blank()) +
guides(colour=F)
dev.off()

### Next, we plot the gmax calculated from dropone.
```

```
plotD <- constraint[,c(1,7,9)]
plotD$x <- cos(plotD[,2]*pi/180)
plotD$y <- sin(plotD[,2]*pi/180)
plotD <- plotD[,c(1,3,4,5)]
plotD <- rbind(plotD, plotD)
plotD[17:32,3] <- 0
plotD[17:32,4] <- 0

tiff("Figures/constraintPlot_v2_thetaDropone.tiff", units="in", height=3, width=3,
res=1200, compression="lzw")
ggplot() +
geom_line(aes(x=c(0,1.25),y=c(0,0)), size=0.5, arrow=arrow(length=unit(0.1,"cm"),
ends="last", type="open")) +
geom_line(aes(x=c(0,1.25*cos(pi*67.27738/180)),y=c(0,1.25*sin(pi*67.27738/180))),
size=0.5, arrow=arrow(length=unit(0.1,"cm"), ends="last", type="open")) +
geom_line(data=plotD, aes(x=x,y=y,group=Trait,colour=Classification), size=0.5, arrow =
arrow(length=unit(0.1,"cm"), ends="last", type="open")) +
xlim(c(0,1.25)) +
ylim(c(0,1.25)) +
theme(axis.line=element_blank(), axis.text=element_blank(), axis.ticks=element_blank(),
axis.title=element_blank(), panel.background = element_blank()) +
guides(colour=F)
dev.off()
################################################################################
```

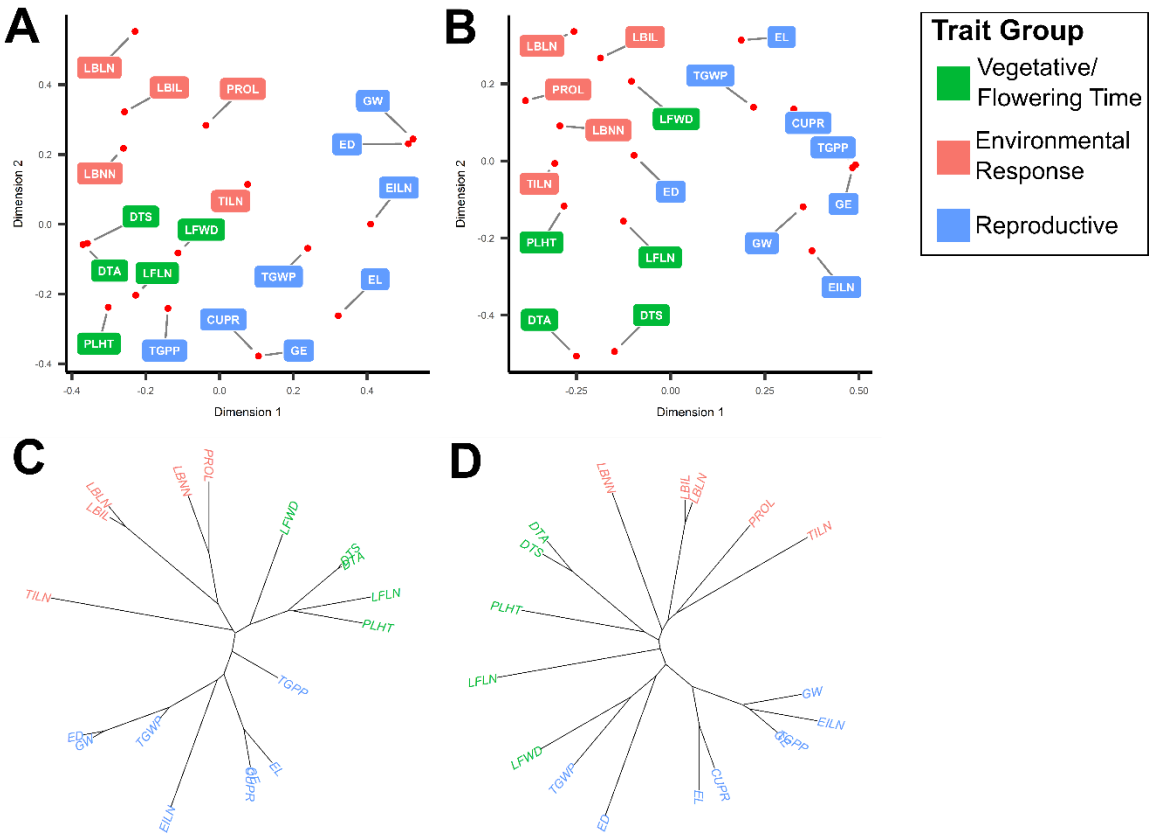

Fig. S1. Genetic relationship of 18 teosinte-maize landrace comparable traits. Plots of principal coordinate analysis (PCoA) [A,B] and neighbor-joining (NJ) tree [C,D] of 18 teosinte-maize landrace comparable traits are constructed from absolute distance calculated from 1 – $|r_g|$. Our pre-defined trait groups fit well with the genetic relationship of these 18 traits in teosinte [A,C] and maize landrace [B,D].

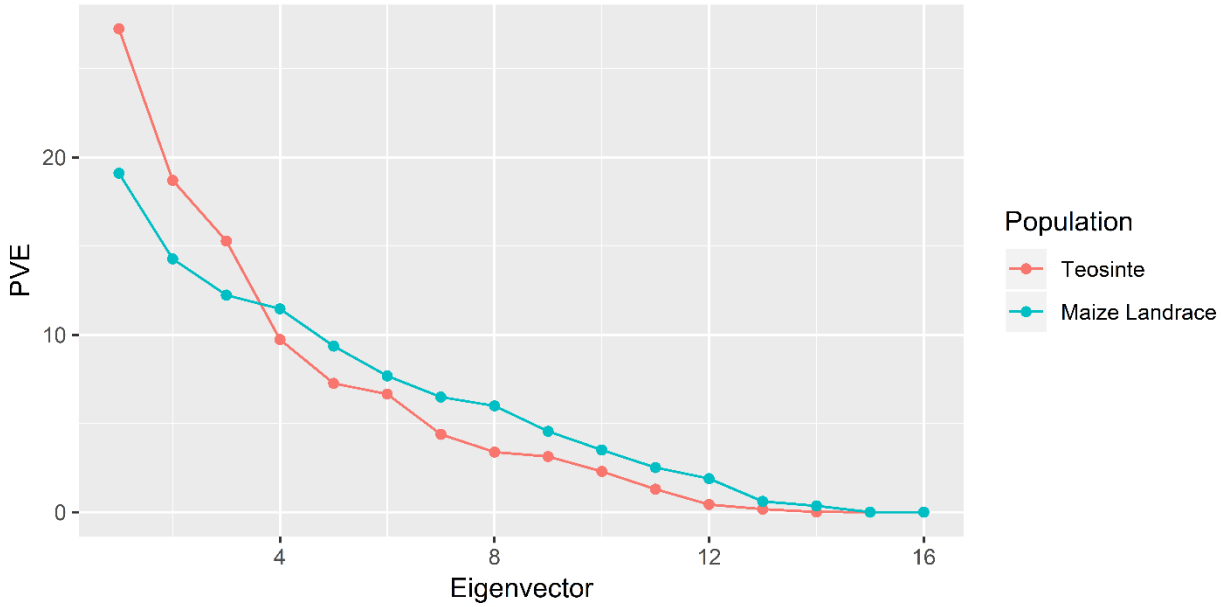

Fig. S2. Scree plot for eigenvectors of teosinte and maize landrace genetic correlation matrices. Percent variance explained (PVE) of all 16 eigenvectors are shown here. PVE for $\boldsymbol{g}_{max,T}$ is 27˚ and PVE for $\boldsymbol{g}_{max,M}$ is 19˚.

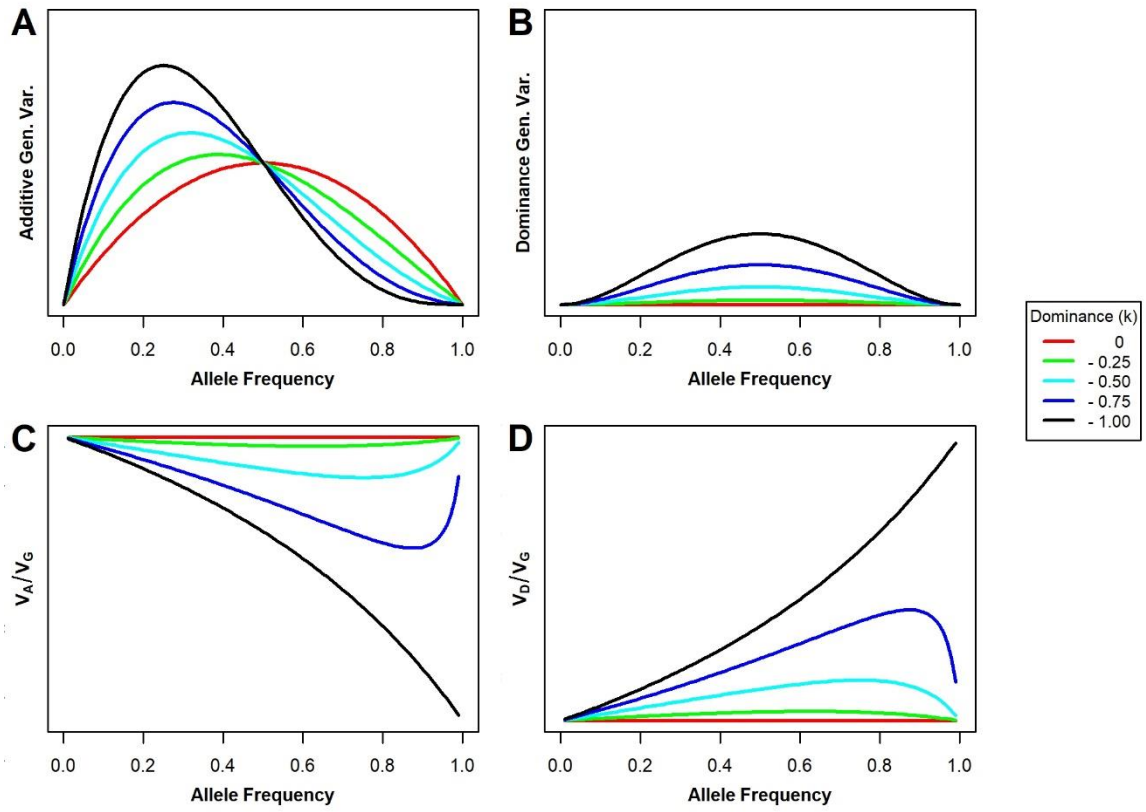

Fig. S3. Change in additive and dominance genetic variance as functions of allele frequency. When dominance is present, additive genetic variance is calculated as $V_A = 2p(1-p)a^2[1+k(2p-1)]^2$ while dominance genetic variance is calculated as $V_D = [2akp(1-p)]^2$. Using these two formulas, we plotted V_A, V_D, V_A/V_G and V_D/V_G for five different scenarios: no dominance (k=0), weakly recessive (k=-0.25), moderately recessive (k=-0.5), strongly recessive (k=-0.75) and complete recessive (k=-1.00).

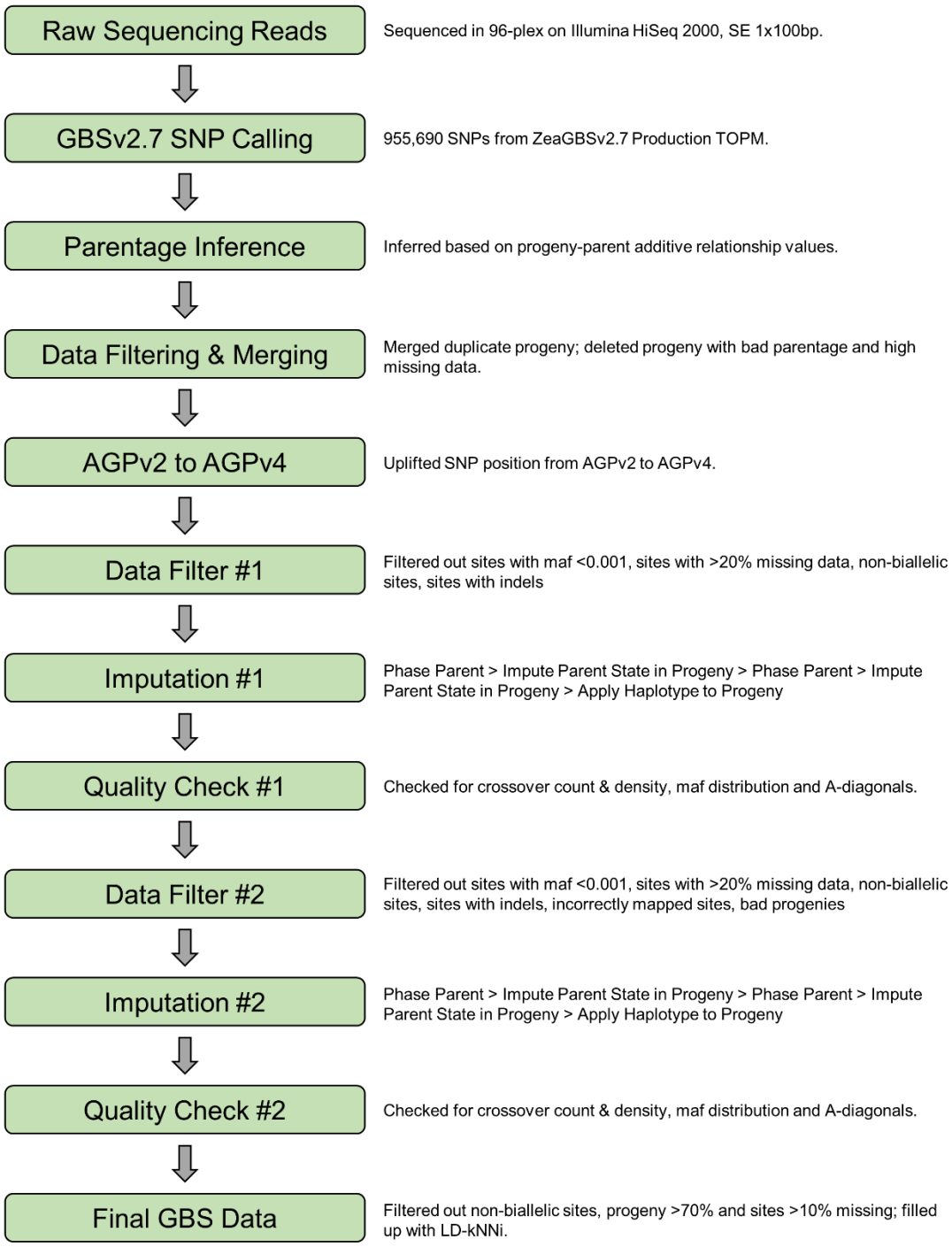

Fig. S4. Workflow from raw sequencing reads to final GBS data. Flowchart highlighting important steps along the path from raw sequencing reads to final GBS data. Steps include SNP calling, parentage inference, data filtering, coordinate conversion, imputation and quality check.

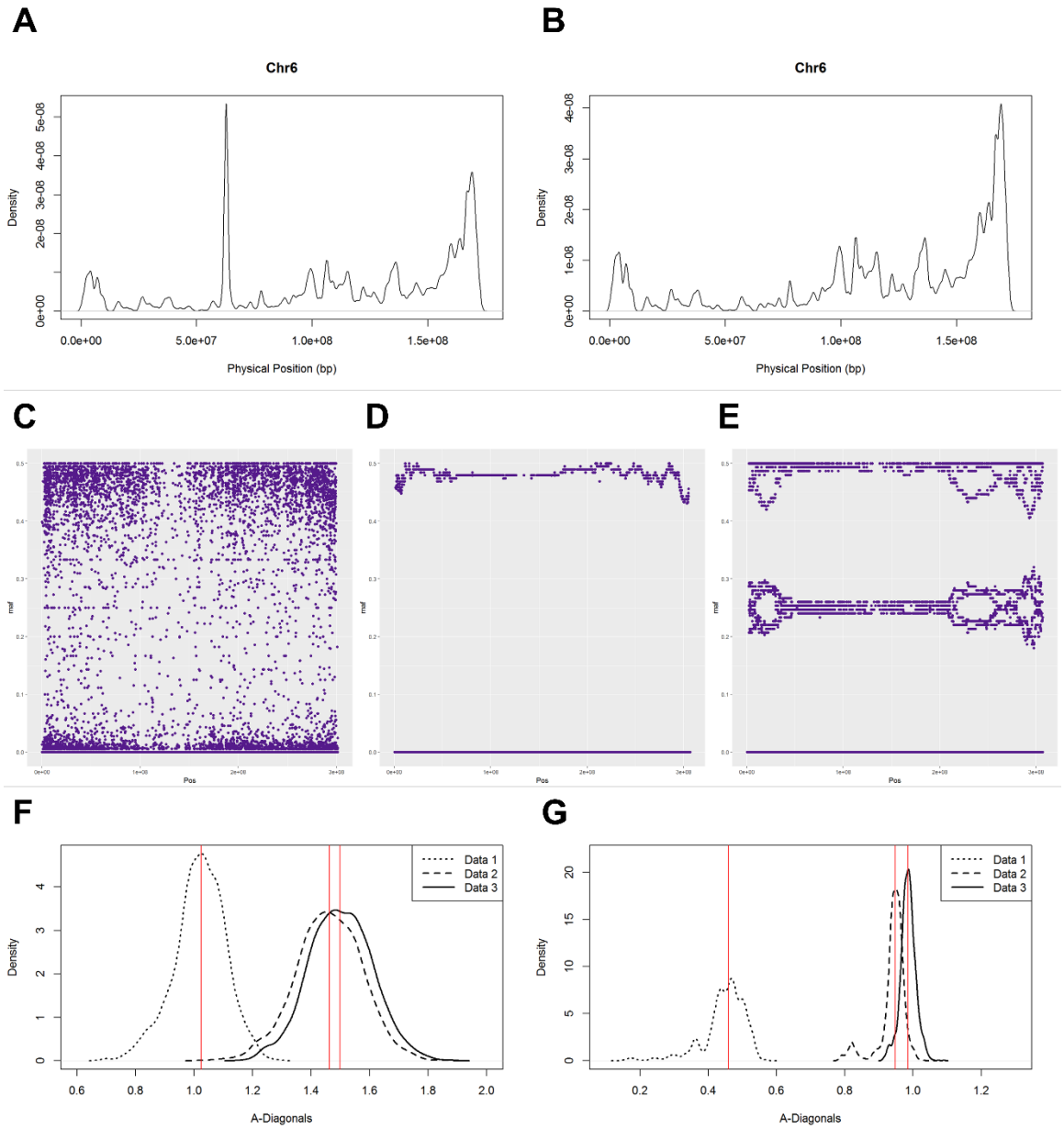

Fig. S5. Quality check (QC) plots. Various quality control plots are shown using the teosinte GBS data as example (maize landrace is similar and thus not shown). Crossover densities along chromosome 6 from the first imputation [A] and the second imputation [B] are presented. The crossover spike in the middle of chromosome 6 disappeared after we removed the bad SNPs prior to second imputation. Minor allele frequency (maf) distribution along chromosome 1 within the PC_O51_ID2 selfed family is shown as raw [C] and imputed [D]. Similar maf distribution is shown for the imputed PC_N13_ID1 x PC_N07_ID1 outcross family [E]. Under our expectation, maf for selfed family should be either 0 or ~0.5 and maf for outcross family should be 0, ~0.25 or ~0.5. Notice that the noise between 0 and 0.5 [C] is absent in the imputed GBS data [D]. Density plots of the diagonals of additive genomic relationship matrix are displayed for selfed [F] and outcross [G] progenies, where Data 1 is imputed and unfiltered, Data 2 is imputed and filtered (progenies <70% missing, sites <10% missing), and Data 3 is imputed, filtered (same), and imputed again with LD-kNNi. Medians are shown in red vertical lines, and Data 3 has the closest median to the expected value of 1.5 (selfed) and 1.0 (outcross).

L1 (119cm)	**C1** (91cm)	**R1** (119cm)
L2 (98cm)	**C2** (61cm)	**R2** (98cm)
L3 (82cm)	**C3** (30cm)	**R3** (82cm)
L4 (76cm)	***x***	**R4** (76cm)
L5 (82cm)	**C5** (30cm)	**R5** (82cm)
L6 (98cm)	**C6** (61cm)	**R6** (98cm)
L7 (119cm)	**C7** (91cm)	**R7** (119cm)

Fig. S6. Position and distance of neighboring plants under consideration for shading. Neighboring plants of plant *x* within 119cm or less away are considered in modelling for shading on plant *x*. This criterion results in 20 neighboring plants that are arbitrarily named after their positions: left (L1 – L7), center (C1 – C7), and right (R1 – R7). Shown in each box is the neighboring plant *i* and distance *d* in bracket.

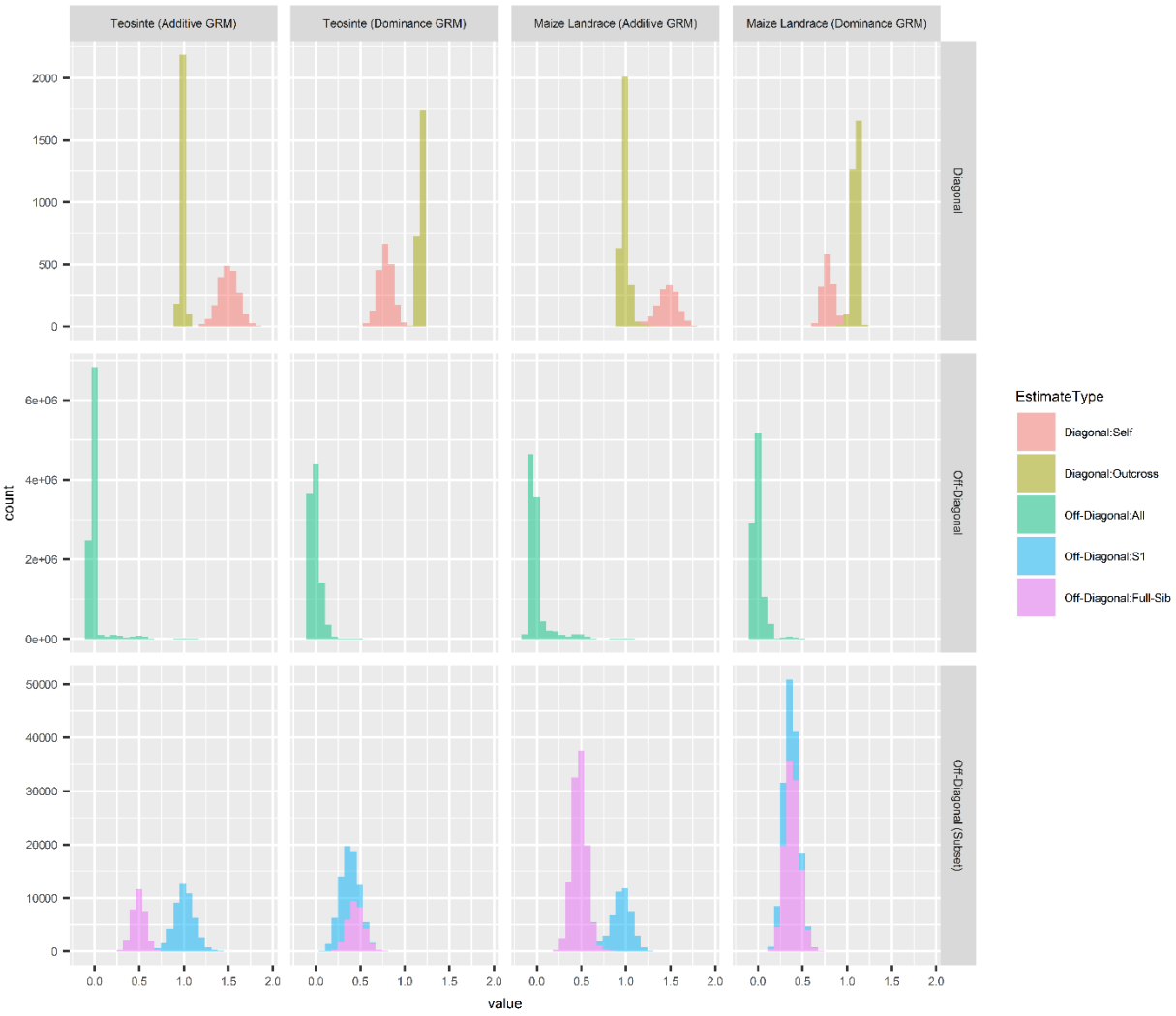

Fig. S7. Histograms of elements of the genomic relationship matrices (GRM). The columns are arranged from left to right in the following order: additive GRM elements for teosinte, dominance GRM elements for teosinte, additive GRM elements for maize landrace and dominance GRM elements for maize landrace. The rows are arranged from top to bottom in the following order: diagonal elements of the GRM, off-diagonal elements of the GRM, and off-diagonal elements of the GRM for full-sibs and S1 relatives from a common parent.

Table S1. Variance component estimates for teosinte. The following estimates (and their ratios) and their standard errors are displayed: additive genetics variance (V_A); dominance genetic variance (V_D); genetic-by-environment variance (V_{GxE}); residual variance (V_E); phenotypic variance (V_P); broad-sense heritability (H^2); narrow-sense heritability (h^2); outcross trait mean (μ).

Trait	V_A (se)	V_D (se)	V_{GxE} (se)	V_E (se)	V_P (se)	H^2 (se)	h^2 (se)	V_D/V_P (se)	V_{GxE}/V_P (se)	V_A/V_G (se)	V_D/V_G (se)	V_D/V_A (se)	μ (se)	units
DTA	20.517 (1.886)	1.186 (0.736)	3.159 (0.792)	25.401 (0.869)	50.262 (1.705)	0.432 (0.028)	0.408 (0.027)	0.024 (0.015)	0.063 (0.016)	0.945 (0.033)	0.055 (0.033)	0.058 (0.037)	68.482 (0.296)	day
DTS	20.226 (1.608)	1.328 (0.558)	2.425 (0.567)	17.866 (0.620)	41.846 (1.467)	0.515 (0.025)	0.483 (0.025)	0.032 (0.013)	0.058 (0.014)	0.938 (0.025)	0.062 (0.025)	0.066 (0.029)	66.869 (0.275)	day
PLHT	207.476 (19.647)	34.460 (9.339)	32.224 (8.319)	298.88 (10.599)	573.04 (18.521)	0.422 (0.027)	0.362 (0.026)	0.060 (0.016)	0.056 (0.015)	0.858 (0.036)	0.142 (0.036)	0.166 (0.049)	132.875 (1.292)	cm
LFLN	22.167 (2.620)	3.692 (1.373)	8.142 (1.691)	52.78 (1.833)	86.782 (2.573)	0.298 (0.027)	0.255 (0.026)	0.043 (0.016)	0.094 (0.019)	0.857 (0.049)	0.143 (0.049)	0.167 (0.067)	53.189 (0.470)	cm
LFWD	0.403 (0.036)	0.024 (0.014)	0.066 (0.016)	0.513 (0.018)	1.007 (0.033)	0.424 (0.027)	0.400 (0.026)	0.024 (0.014)	0.066 (0.016)	0.944 (0.031)	0.056 (0.031)	0.059 (0.035)	5.524 (0.043)	cm
TILN	1.403 (0.245)	0.490 (0.195)	0.581 (0.151)	0.000 (0.000)	9.508 (0.257)	0.199 (0.027)	0.148 (0.024)	0.052 (0.020)	0.061 (0.016)	0.741 (0.090)	0.259 (0.090)	0.349 (0.163)	7.145 (0.167)	count
PROL	1.956 (0.343)	1.112 (0.274)	0.702 (0.200)	8.096 (0.263)	11.867 (0.334)	0.259 (0.029)	0.165 (0.027)	0.094 (0.023)	0.059 (0.017)	0.638 (0.075)	0.363 (0.075)	0.569 (0.185)	9.763 (0.218)	count
LBNN	0.038 (0.012)	0.019 (0.009)	0.041 (0.011)	0.412 (0.012)	0.510 (0.013)	0.111 (0.025)	0.075 (0.023)	0.036 (0.018)	0.081 (0.020)	0.672 (0.141)	0.328 (0.141)	0.487 (0.312)	2.659 (0.036)	count
LBLN	4410.72 (797.599)	1645.46 (530.794)	2524.3 (609.560)	19347.0 (619.960)	27927.0 (782.940)	0.217 (0.029)	0.158 (0.027)	0.059 (0.019)	0.090 (0.022)	0.728 (0.077)	0.272 (0.077)	0.373 (0.144)	442.782 (9.289)	mm
LBIL	805.297 (94.630)	221.401 (58.727)	75.404 (31.164)	1883.20 (60.788)	2985.30 (90.500)	0.344 (0.028)	0.270 (0.026)	0.074 (0.019)	0.025 (0.010)	0.784 (0.051)	0.216 (0.051)	0.275 (0.083)	166.394 (3.185)	mm
EL	20.976 (1.560)	0.766 (0.491)	0.550 (0.257)	17.829 (0.589)	40.120 (1.426)	0.542 (0.023)	0.523 (0.023)	0.019 (0.012)	0.014 (0.006)	0.965 (0.022)	0.035 (0.022)	0.037 (0.024)	46.418 (0.243)	mm
CUPR	0.208 (0.015)	0.011 (0.005)	0.003 (0.002)	0.155 (0.005)	0.377 (0.014)	0.581 (0.022)	0.551 (0.022)	0.030 (0.013)	0.007 (0.005)	0.948 (0.022)	0.052 (0.022)	0.055 (0.024)	4.498 (0.026)	count
ED	0.066 (0.004)	0.005 (0.001)	0.002 (0.001)	0.023 (0.001)	0.095 (0.004)	0.740 (0.016)	0.692 (0.017)	0.048 (0.012)	0.017 (0.006)	0.935 (0.016)	0.065 (0.016)	0.069 (0.018)	3.971 (0.014)	mm
GE	0.830 (0.059)	0.045 (0.020)	0.011 (0.007)	0.622 (0.022)	1.509 (0.054)	0.581 (0.022)	0.551 (0.022)	0.030 (0.013)	0.007 (0.005)	0.948 (0.022)	0.052 (0.022)	0.055 (0.024)	8.996 (0.052)	count
EILN	0.143 (0.010)	0.007 (0.003)	0.002 (0.001)	0.096 (0.003)	0.249 (0.009)	0.605 (0.021)	0.576 (0.021)	0.029 (0.011)	0.008 (0.005)	0.953 (0.018)	0.047 (0.018)	0.050 (0.020)	5.180 (0.020)	mm
TGPP	328513 (39202.0)	25683.1 (20880.6)	53565.0 (16447.0)	816710 (25999.0)	1224500 (36948.0)	0.289 (0.028)	0.268 (0.027)	0.021 (0.017)	0.044 (0.013)	0.928 (0.057)	0.073 (0.057)	0.078 (0.066)	4336.062 (48.581)	count
TGWP	373.665 (39.458)	71.687 (22.543)	69.617 (18.067)	555.440 (20.389)	1070.400 (35.648)	0.416 (0.029)	0.349 (0.029)	0.067 (0.021)	0.065 (0.017)	0.839 (0.047)	0.161 (0.047)	0.192 (0.067)	113.948 (1.818)	mg
GW	17.446 (0.974)	1.270 (0.296)	0.440 (0.140)	4.907 (0.228)	24.061 (0.909)	0.778 (0.015)	0.725 (0.016)	0.053 (0.012)	0.018 (0.006)	0.932 (0.016)	0.068 (0.016)	0.073 (0.018)	26.300 (0.200)	mg

Table S2. Variance component estimates for maize landrace. The following estimates (and their ratios) and their standard errors are displayed: additive genetics variance (V_A); dominance genetic variance (V_D); genetic-by-environment variance (V_{GxE}); residual variance (V_E); phenotypic variance (V_P); broad-sense heritability (H^2); narrow-sense heritability (h^2); outcross trait mean (μ).

Trait	V_A (se)	V_D (se)	V_{GxE} (se)	V_E (se)	V_P (se)	H^2 (se)	h^2 (se)	V_D/V_P (se)	V_{GxE}/V_P (se)	V_A/V_G (se)	V_D/V_G (se)	V_D/V_A (se)	μ (se)	units
DTA	5.752 (0.684)	0.969 (0.368)	0.787 (0.279)	12.893 (0.404)	20.401 (0.649)	0.329 (0.028)	0.282 (0.027)	0.048 (0.018)	0.039 (0.014)	0.856 (0.052)	0.144 (0.052)	0.168 (0.070)	76.795 (0.192)	day
DTS	8.211 (1.089)	0.590 (0.522)	1.653 (0.566)	21.968 (0.682)	32.421 (1.024)	0.271 (0.029)	0.253 (0.028)	0.018 (0.016)	0.051 (0.017)	0.933 (0.057)	0.067 (0.057)	0.072 (0.066)	81.163 (0.229)	day
PLHT	172.123 (24.415)	21.645 (12.298)	43.257 (14.09)	575.27 (17.823)	812.290 (24.467)	0.239 (0.028)	0.212 (0.026)	0.027 (0.015)	0.053 (0.017)	0.888 (0.060)	0.112 (0.060)	0.126 (0.076)	247.67 (1.177)	cm
LFLN	49.342 (5.470)	2.885 (2.576)	3.569 (1.645)	73.042 (2.697)	128.840 (4.888)	0.405 (0.032)	0.383 (0.031)	0.022 (0.020)	0.028 (0.013)	0.945 (0.048)	0.055 (0.048)	0.059 (0.054)	100.236 (0.465)	cm
LFWD	0.616 (0.070)	0.094 (0.037)	0.050 (0.022)	0.997 (0.036)	1.757 (0.064)	0.404 (0.030)	0.350 (0.031)	0.054 (0.021)	0.028 (0.012)	0.867 (0.050)	0.133 (0.050)	0.153 (0.067)	10.730 (0.060)	cm
TILN	0.005 (0.004)	0.029 (0.004)	0.014 (0.003)	0.096 (0.003)	0.144 (0.004)	0.234 (0.029)	0.036 (0.025)	0.198 (0.029)	0.099 (0.022)	0.154 (0.099)	0.846 (0.099)	5.504 (4.202)	0.137 (0.022)	count
PROL	0.003 (0.002)	0.005 (0.004)	0.001 (0.001)	0.186 (0.005)	0.193 (0.005)	0.037 (0.016)	0.014 (0.011)	0.024 (0.020)	0.003 (0.007)	0.366 (0.352)	0.634 (0.352)	1.731 (2.627)	1.104 (0.018)	count
LBNN	1.886 (0.216)	0.236 (0.122)	0.104 (0.055)	3.731 (0.122)	5.956 (0.199)	0.356 (0.028)	0.317 (0.029)	0.040 (0.020)	0.017 (0.009)	0.889 (0.055)	0.111 (0.055)	0.125 (0.070)	12.496 (0.102)	count
LBLN	337.973 (2.168)	179.939 (41.944)	55.266 (22.285)	1005.20 (34.204)	1578.40 (49.597)	0.328 (0.029)	0.214 (0.030)	0.114 (0.026)	0.035 (0.014)	0.653 (0.073)	0.347 (0.073)	0.532 (0.171)	115.170 (2.061)	mm
LBIL	2.168 (0.366)	1.362 (0.303)	0.408 (0.157)	6.965 (0.240)	10.902 (0.341)	0.324 (0.029)	0.199 (0.030)	0.125 (0.027)	0.037 (0.014)	0.614 (0.075)	0.386 (0.075)	0.628 (0.200)	9.36 (0.176)	mm
EL	186.632 (37.552)	62.439 (28.642)	55.814 (22.781)	887.370 (29.223)	1192.30 (35.778)	0.209 (0.030)	0.157 (0.029)	0.052 (0.024)	0.047 (0.019)	0.749 (0.104)	0.251 (0.104)	0.335 (0.185)	169.204 (1.642)	mm
CUPR	11.459 (2.287)	1.009 (1.577)	2.638 (1.199)	58.205 (1.865)	73.312 (2.188)	0.170 (0.028)	0.156 (0.029)	0.014 (0.022)	0.036 (0.016)	0.919 (0.122)	0.081 (0.122)	0.088 (0.145)	31.886 (0.378)	count
ED	3.140 (0.466)	0.968 (0.309)	0.633 (0.215)	7.280 (0.271)	12.021 (0.419)	0.342 (0.032)	0.261 (0.033)	0.081 (0.025)	0.053 (0.018)	0.764 (0.069)	0.236 (0.069)	0.308 (0.119)	25.663 (0.172)	mm
GE	1791.92 (477.845)	865.488 (426.349)	563.669 (257.383)	12294 (415.18)	15515.0 (457.420)	0.171 (0.029)	0.116 (0.029)	0.056 (0.027)	0.036 (0.017)	0.674 (0.142)	0.326 (0.142)	0.483 (0.313)	358.677 (6.192)	count
EILN	0.016 (0.003)	0.000 (0.000)	0.003 (0.002)	0.093 (0.003)	0.112 (0.003)	0.144 (0.025)	0.144 (0.025)	0.000 (0.000)	0.028 (0.014)	1.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.963 (0.015)	mm
TGPP	1551.14 (472.909)	961.605 (449.348)	609.999 (269.911)	12909.0 (431.550)	16032.0 (461.980)	0.157 (0.029)	0.097 (0.029)	0.060 (0.028)	0.038 (0.017)	0.617 (0.155)	0.383 (0.155)	0.620 (0.408)	357.619 (6.364)	count
TGWP	25.343 (42.238)	158.484 (57.841)	87.020 (36.259)	1571.80 (51.044)	1842.70 (48.501)	0.100 (0.027)	0.014 (0.023)	0.086 (0.031)	0.047 (0.020)	0.138 (0.223)	0.862 (0.223)	6.254 (11.736)	122.003 (2.272)	mg
GW	1294.62 (269.151)	478.496 (207.141)	531.64 (170.945)	5722.00 (198.112)	8026.80 (249.464)	0.221 (0.032)	0.161 (0.031)	0.060 (0.026)	0.066 (0.021)	0.730 (0.105)	0.270 (0.105)	0.370 (0.197)	359.066 (4.338)	mg

Table S3. Trait mean and selection intensity. The mean and standard error for each trait in teosinte and maize landrace are shown here. Selection intensities are provided as estimates based on 9,000 generations of selection during domestication.

Trait	Units	Teosinte		Maize Landrace		Selection Intensity
		Mean	Std Err	Mean	Std Err	
DTA	days	68.4821	0.2963	76.7954	0.1920	0.0004
DTS	days	66.8688	0.2750	81.1632	0.2285	0.0007
PLHT	cm	132.8754	1.2919	247.6702	1.1773	0.0018
LFLN	cm	53.1889	0.4696	100.2361	0.4645	0.0015
LFWD	cm	5.5235	0.0425	10.7302	0.0600	0.0013
TILN	count	7.1449	0.1670	0.1373	0.0223	-0.0026
PROL	count	9.7625	0.2180	1.1040	0.0183	-0.0026
LBNN	count	2.6587	0.0363	12.4956	0.1024	0.0022
LBLN	mm	442.7818	9.2892	115.1695	2.0609	-0.0019
LBIL	mm	166.3942	3.1845	9.3601	0.1758	-0.0018
EL	mm	46.4175	0.2428	169.2042	1.6420	0.0033
CUPR	count	4.4977	0.0258	31.8860	0.3781	0.0033
ED	mm	3.9706	0.0140	25.6634	0.1722	0.0039
GE	count	8.9955	0.0516	358.6766	6.1921	0.0040
EILN	mm	5.1804	0.0204	0.9628	0.0146	-0.0026
TGPP	count	4336.0619	48.5813	357.6188	6.3639	-0.0022
TGWP	g	113.9475	1.8180	122.0025	2.2717	0.0002
GW	mg	26.3000	0.2000	359.0660	4.3380	0.0038

Table S4. Genetic correlations, variances and covariances for teosinte. Genetic correlations ($r_{G,ij}$) for all trait pairs and their standard errors are shown in the lower left triangle of the matrix; additive genetic variances ($V_{A,i}$) for each trait are shown in the diagonal of the matrix; genetic covariances ($Cov_{G,ij}$) for all trait pairs are shown in the upper right triangle of the matrix. Standard errors are not available for the $Cov_{G,ij}$ since they are calculated as $Cov_{G,ij} = r_{G,ij} \cdot \sqrt{V_{A,i} \cdot V_{A,j}}$ for each trait pair instead of taking the estimates directly from the bivariate REML outputs.

	DTA	DTS	PLHT	LFLN	LFWD	TILN	PROL	LBNN	LBLN	LBIL	EL	CUPR	ED	GE	EILN	TGPP	TGWP	GW
DTA	20.517 (1.886)	19.705	38.781	12.482	1.15	-0.214	0.016	0.269	-86.04	-63.91	3.921	0.504	-0.047	1.008	-0.135	1261.2	24.735	-2.321
DTS	0.967 (0.005)	20.226 (1.608)	35.875	12.241	1.166	0.086	0.223	0.289	-67.08	-57.51	3.394	0.452	-0.059	0.903	-0.150	1275.7	24.985	-2.532
PLHT	0.594 (0.046)	0.554 (0.046)	207.48 (19.65)	39.463	4.677	-1.948	3.125	0.799	15.115	-47.78	15.457	1.566	0.134	3.132	-0.131	3955.4	95.225	0.993
LFLN	0.585 (0.048)	0.578 (0.046)	0.582 (0.045)	22.167 (2.620)	0.931	0.301	1.333	0.220	11.569	-10.46	5.835	0.496	0.140	0.992	0.111	1183.1	31.072	1.316
LFWD	0.400 (0.052)	0.409 (0.050)	0.512 (0.042)	0.312 (0.054)	0.403 (0.036)	-0.137	0.276	0.047	5.593	-0.663	0.723	0.082	0.037	0.164	-0.012	173.40	5.665	0.412
TILN	-0.040 (0.071)	0.016 (0.069)	-0.114 (0.074)	0.054 (0.076)	-0.183 (0.073)	1.403 (0.245)	-0.132	-0.007	-1.101	0.565	-0.444	-0.018	-0.022	-0.036	-0.024	181.94	1.323	-0.527
PROL	0.002 (0.072)	0.036 (0.069)	0.155 (0.070)	0.202 (0.072)	0.310 (0.062)	-0.08 (0.088)	1.956 (0.343)	0.172	40.093	5.732	-1.575	-0.118	-0.039	-0.237	-0.044	262.79	2.874	-0.933
LBNN	0.304 (0.079)	0.329 (0.076)	0.284 (0.080)	0.239 (0.083)	0.381 (0.072)	-0.032 (0.100)	0.630 (0.060)	0.038 (0.012)	5.842	-0.321	0.048	0.013	0.000	0.025	-0.01	44.675	0.815	-0.058
LBLN	-0.286 (0.067)	-0.225 (0.066)	0.016 (0.073)	0.037 (0.076)	0.133 (0.069)	-0.014 (0.085)	0.432 (0.067)	0.451 (0.074)	4410.7 (797.6)	1644.4	-15.85	-1.425	-1.69	-2.851	0.302	-5851	-239.8	-16.17
LBIL	-0.497 (0.053)	-0.451 (0.053)	-0.117 (0.067)	-0.078 (0.070)	-0.037 (0.065)	0.017 (0.08)	0.144 (0.078)	-0.058 (0.093)	0.872 (0.024)	805.30 (94.63)	-6.719	-1.479	-0.427	-2.958	1.161	-6080	-162.4	0.616
EL	0.189 (0.053)	0.165 (0.051)	0.234 (0.052)	0.271 (0.055)	0.249 (0.051)	-0.082 (0.068)	-0.246 (0.065)	0.053 (0.08)	-0.052 (0.068)	-0.052 (0.062)	20.976 (1.560)	1.608	0.392	3.217	0.488	1228.0	56.882	8.065
CUPR	0.244 (0.052)	0.220 (0.049)	0.239 (0.051)	0.231 (0.055)	0.284 (0.049)	-0.034 (0.067)	-0.186 (0.065)	0.142 (0.077)	-0.047 (0.067)	-0.114 (0.061)	0.771 (0.020)	0.208 (0.015)	0.011	0.415	-0.067	183.99	4.224	0.047
ED	-0.04 (0.049)	-0.051 (0.046)	0.036 (0.049)	0.116 (0.052)	0.227 (0.046)	-0.073 (0.062)	-0.108 (0.061)	-0.010 (0.074)	-0.099 (0.062)	-0.058 (0.056)	0.333 (0.041)	0.093 (0.044)	0.066 (0.004)	0.022	0.033	1.252	3.568	0.972
GE	0.244 (0.052)	0.220 (0.049)	0.239 (0.051)	0.231 (0.055)	0.284 (0.049)	-0.034 (0.067)	-0.186 (0.065)	0.142 (0.077)	-0.047 (0.067)	-0.114 (0.061)	0.771 (0.020)	1.000 (0.000)	0.093 (0.044)	0.830 (0.059)	-0.134	367.98	8.449	0.093
EILN	-0.079 (0.053)	-0.088 (0.05)	-0.024 (0.054)	0.062 (0.056)	-0.050 (0.052)	-0.054 (0.066)	-0.083 (0.065)	-0.132 (0.077)	0.012 (0.066)	0.108 (0.060)	0.282 (0.045)	-0.390 (0.042)	0.342 (0.040)	-0.390 (0.042)	0.143 (0.010)	-79.60	1.450	0.918
TGPP	0.486 (0.060)	0.495 (0.057)	0.479 (0.052)	0.438 (0.059)	0.477 (0.05)	0.268 (0.071)	0.328 (0.068)	0.400 (0.078)	-0.154 (0.081)	-0.374 (0.068)	0.468 (0.051)	0.704 (0.036)	0.008 (0.056)	0.704 (0.036)	-0.367 (0.055)	328513 (39202)	6676.5	-225.5
TGWP	0.282 (0.058)	0.287 (0.055)	0.342 (0.053)	0.341 (0.057)	0.462 (0.047)	0.058 (0.071)	0.106 (0.071)	0.216 (0.082)	-0.187 (0.074)	-0.296 (0.065)	0.642 (0.037)	0.480 (0.044)	0.718 (0.029)	0.480 (0.044)	0.198 (0.054)	0.603 (0.038)	373.67 (39.46)	58.269
GW	-0.123 (0.048)	-0.135 (0.045)	0.016 (0.048)	0.067 (0.051)	0.156 (0.046)	-0.107 (0.061)	-0.160 (0.060)	-0.071 (0.073)	-0.058 (0.061)	0.005 (0.055)	0.422 (0.038)	0.024 (0.043)	0.906 (0.008)	0.024 (0.043)	0.581 (0.031)	-0.094 (0.055)	0.722 (0.028)	17.445 (0.974)

Table S5. Genetic correlations, variances and covariances for maize landrace. Genetic correlations ($r_{G,ij}$) for all trait pairs and their standard errors are shown in the lower left triangle of the matrix; additive genetic variances ($V_{A,i}$) for each trait are shown in the diagonal of the matrix; genetic covariances ($Cov_{G,ij}$) for all trait pairs are shown in the upper right triangle of the matrix. Standard errors are not available for the $Cov_{G,ij}$ since they are calculated as $Cov_{G,ij} = r_{G,ij} \cdot \sqrt{V_{A,i} \cdot V_{A,j}}$ for each trait pair instead of taking the estimates directly from the bivariate REML outputs.

	DTA	DTS	PLHT	LFLN	LFWD	TILN	PROL	LBNN	LBLN	LBIL	EL	CUPR	ED	GE	EILN	TGPP	TGWP	GW
DTA	5.752 (0.684)	5.445	11.954	0.938	-0.008	0.036	-0.025	-0.102	-6.455	-0.550	-0.144	-0.780	-0.181	-12.24	0.062	-9.531	0.404	19.443
DTS	0.792 (0.032)	8.210 (1.089)	7.098	6.044	-0.047	0.020	-0.011	0.755	-2.534	-0.644	-3.135	-1.845	-0.121	-20.77	0.069	-19.31	-2.293	15.217
PLHT	0.380 (0.077)	0.189 (0.084)	172.12 (24.42)	7.197	-0.914	-0.074	-0.160	4.717	47.635	1.244	37.083	9.748	-1.532	49.594	-0.045	39.322	7.661	-27.80
LFLN	0.056 (0.073)	0.300 (0.074)	0.078 (0.08)	49.342 (5.470)	-0.444	-0.002	0.003	0.478	19.306	1.220	5.604	-1.298	-0.804	-30.18	0.060	-26.34	-6.874	16.352
LFWD	-0.004 (0.071)	-0.021 (0.076)	-0.089 (0.081)	-0.080 (0.072)	0.616 (0.070)	-0.008	0.008	0.209	2.380	0.044	1.283	0.302	0.301	4.790	-0.003	4.574	1.817	4.240
TILN	0.209 (0.076)	0.097 (0.083)	-0.078 (0.087)	-0.004 (0.079)	-0.140 (0.077)	0.005 (0.004)	0.001	0.000	0.237	0.020	-0.063	-0.030	0.010	-0.091	0.000	-0.039	0.010	0.192
PROL	-0.204 (0.156)	-0.077 (0.17)	-0.238 (0.176)	0.009 (0.163)	0.191 (0.156)	0.256 (0.157)	0.003 (0.002)	0.018	0.339	0.024	0.072	0.023	-0.016	0.067	0.000	0.092	0.044	0.164
LBNN	-0.031 (0.073)	0.192 (0.076)	0.262 (0.075)	0.050 (0.072)	0.194 (0.069)	-0.002 (0.079)	0.261 (0.149)	1.886 (0.216)	6.509	-0.484	-2.739	-0.412	-0.041	-2.093	-0.028	-1.417	0.332	3.533
LBLN	-0.146 (0.073)	-0.048 (0.078)	0.198 (0.080)	0.150 (0.074)	0.165 (0.072)	0.179 (0.079)	0.359 (0.152)	0.258 (0.07)	337.97 (2.168)	23.681	85.316	18.937	-0.492	129.18	0.014	112.66	9.088	-93.86
LBIL	-0.156 (0.074)	-0.153 (0.078)	0.064 (0.085)	0.118 (0.076)	0.038 (0.076)	0.187 (0.08)	0.314 (0.165)	-0.239 (0.072)	0.875 (0.019)	2.168 (0.366)	8.539	1.702	0.012	11.002	0.017	9.482	0.555	-9.478
EL	-0.004 (0.085)	-0.080 (0.087)	0.207 (0.089)	0.058 (0.086)	0.120 (0.082)	-0.064 (0.095)	0.102 (0.177)	-0.146 (0.084)	0.340 (0.076)	0.424 (0.074)	186.63 (37.55)	32.122	1.346	265.09	0.350	239.91	28.513	-98.55
CUPR	-0.096 (0.088)	-0.190 (0.087)	0.220 (0.093)	-0.055 (0.091)	0.114 (0.087)	-0.125 (0.100)	0.135 (0.186)	-0.089 (0.089)	0.304 (0.084)	0.342 (0.084)	0.695 (0.052)	11.459 (2.287)	-1.246	114.08	-0.170	106.78	7.718	-67.25
ED	-0.042 (0.076)	-0.024 (0.08)	-0.066 (0.084)	-0.065 (0.077)	0.216 (0.072)	0.080 (0.083)	-0.181 (0.158)	-0.017 (0.077)	-0.015 (0.080)	0.004 (0.081)	0.056 (0.087)	-0.208 (0.094)	3.140 (0.466)	0.420	-0.025	0.607	2.098	13.364
GE	-0.121 (0.088)	-0.171 (0.09)	0.089 (0.100)	-0.102 (0.093)	0.144 (0.088)	-0.030 (0.103)	0.031 (0.195)	-0.036 (0.093)	0.166 (0.090)	0.176 (0.093)	0.458 (0.079)	0.796 (0.039)	0.006 (0.097)	1791.9 (477.8)	-3.813	1667.2	108.19	-1142
EILN	0.203 (0.091)	0.189 (0.096)	-0.027 (0.105)	0.067 (0.095)	-0.028 (0.094)	-0.030 (0.104)	0.049 (0.190)	-0.158 (0.094)	0.006 (0.098)	0.092 (0.100)	0.202 (0.115)	-0.395 (0.091)	-0.111 (0.097)	-0.710 (0.061)	0.016 (0.003)	-3.500	-0.103	2.655
TGPP	-0.101 (0.091)	-0.171 (0.091)	0.076 (0.102)	-0.095 (0.096)	0.148 (0.090)	-0.014 (0.105)	0.046 (0.198)	-0.026 (0.095)	0.156 (0.093)	0.164 (0.095)	0.446 (0.082)	0.801 (0.040)	0.009 (0.099)	1.000 (0.000)	-0.701 (0.065)	1551.1 (472.9)	103.42	-1042
TGWP	0.034 (0.114)	-0.159 (0.114)	0.116 (0.119)	-0.194 (0.119)	0.460 (0.098)	0.026 (0.128)	0.170 (0.208)	0.048 (0.117)	0.098 (0.114)	0.075 (0.119)	0.415 (0.099)	0.453 (0.099)	0.235 (0.107)	0.508 (0.096)	-0.162 (0.138)	0.522 (0.095)	25.343 (42.24)	35.828
GW	0.225 (0.083)	0.148 (0.088)	-0.059 (0.093)	0.065 (0.084)	0.150 (0.083)	0.074 (0.092)	0.089 (0.177)	0.072 (0.084)	-0.142 (0.087)	-0.179 (0.087)	-0.200 (0.095)	-0.552 (0.080)	0.210 (0.086)	-0.750 (0.059)	0.582 (0.078)	-0.735 (0.063)	0.198 (0.130)	1294.6 (269.2)

Table S6. Results from Jump-Up approach in Flury hierarchy. Lower model (unrelated structure) is compared against higher models of partial common principal component (PCPC), common principal component (CPC), proportionality and equality. Since the bottom-most comparison is significant, the test suggests that teosinte and maize landrace $\boldsymbol{G}$-matrices are unrelated.

Higher	**Lower**	χ^2	**df**	**P-values**
Equality	Unrelated	373168.025	136	0
Proportionality	Unrelated	373068.548	135	0
CPC	Unrelated	201724.419	120	0
PCPC(14)	Unrelated	201723.968	119	0
PCPC(13)	Unrelated	201390.193	117	0
PCPC(12)	Unrelated	201223.457	114	0
PCPC(11)	Unrelated	200907.826	110	0
PCPC(10)	Unrelated	200312.352	105	0
PCPC(9)	Unrelated	190861.067	99	0
PCPC(8)	Unrelated	190459.350	92	0
PCPC(7)	Unrelated	175875.368	84	0
PCPC(6)	Unrelated	158198.735	75	0
PCPC(5)	Unrelated	139626.102	65	0
PCPC(4)	Unrelated	113251.837	54	0
PCPC(3)	Unrelated	109270.987	42	0
PCPC(2)	Unrelated	84175.474	29	0
PCPC(1)	Unrelated	38611.482	15	0

Table S7. Q_{ST} estimates for all 18 traits and P-values from univariate Q_{ST}-F_{ST} test. For each trait, Q_{ST} is estimated as $Q_{ST} = V_B/(V_B+2V_W)$ where V_B is the between-population additive genetic variance and V_W is the within-population additive genetic variance. The difference between Q_{ST} and F_{ST} is then calculated as $Q_{ST} - F_{ST}$ where F_{ST} is 0.1567. The observed $Q_{ST} - F_{ST}$ is then compared to the null distribution of $Q_{ST} - F_{ST}$ under neutrality to obtain the two-tailed P-values.

Trait	**Q_{ST}**	**P-values**
DTA	0.5681	0.0160
DTS	0.7823	<0.0001
PLHT	0.9455	<0.0001
LFLN	0.9393	<0.0001
LFWD	0.9301	<0.0001
TILN	0.9458	<0.0001
PROL	0.9503	<0.0001
LBNN	0.9618	<0.0001
LBLN	0.9187	<0.0001
LBIL	0.9385	<0.0001
EL	0.9732	<0.0001
CUPR	0.9698	<0.0001
ED	0.9866	<0.0001
GE	0.9715	<0.0001
EILN	0.9824	<0.0001
TGPP	0.9600	<0.0001
TGWP	0.0752	0.9860
GW	0.9769	<0.0001

Table S8. Individual trait contribution towards genetic constraint. Genetic constraint is measured as an angle (θ) between the actual domestication trajectory (**Z**) and genetic lines of least resistance ($\mathbf{g}_{max}$). Using $\mathbf{g}_{max}$ calculated from the standardized teosinte **G**-matrix, θ_T is 67.3°. Individual trait contribution towards genetic constraint is measured by dropping each trait and calculating the angle ($\theta^i_{dropone}$) using similar method. $\theta^i_{dropone} < \theta_T$ implies that trait *i* constrained evolution while $\theta^i_{dropone} > \theta_T$ implies that trait *i* assisted evolution.

Trait	$\theta^i_{dropone}$	$\theta^i_{dropone} - \theta_T$
GE	62.82	-4.46
DTA	63.22	-4.05
DTS	63.35	-3.93
PLHT	64.63	-2.65
LFLN	65.04	-2.24
LFWD	66.35	-0.93
LBIL	66.52	-0.75
PROL	67.00	-0.28
LBLN	67.20	-0.08
EILN	67.20	-0.08
TILN	67.32	0.04
LBNN	67.40	0.12
GW	69.17	1.90
ED	69.71	2.43
TGWP	70.61	3.33
EL	71.42	4.14

Table S9. Comparing the responses from single trait selection. Responses ($\mathbf{R}_i$) are calculated from hypothetical selection of a single trait and compared to the actual domestication trajectory (**Z**). By selecting for trait *i*, $\left|proj_Z R^i\right|$ measures evolutionary gain along **Z** while θ_Z measures the deviation between $\mathbf{R}_i$ and **Z**.

Trait	Selection Direction	$\left\|proj_Z R^i\right\|$	θ_Z (°)
DTA	Positive	0.31	80.39
DTS	Positive	0.28	81.13
PLHT	Positive	0.34	77.86
LFLN	Positive	0.35	77.23
LFWD	Positive	0.45	73.08
TILN	Negative	0.09	84.84
PROL	Negative	0.17	82.89
LBNN	Positive	0.26	79.86
LBLN	Negative	0.05	87.98
LBIL	Negative	0.17	83.64
EL	Positive	0.97	53.77
ED	Positive	0.50	72.24
GE	Positive	1.07	45.55
EILN	Negative	0.21	81.00
TGWP	Positive	0.84	63.31
GW	Positive	0.42	75.88

Table S10. Progeny-parent information for teosinte.

| Maternal Parent | Paternal Parent | PC_I05_ID1 | PC_I06_ID1 | PC_I08_ID1 | PC_I11_ID1 | PC_I50_ID1 | PC_I52_ID1 | PC_I53_ID1 | PC_I58_ID1 | PC_J01_ID1 | PC_J04_ID1 | PC_J07_ID1 | PC_J08_ID1 | PC_J10_ID1 | PC_J12_ID1 | PC_J13_ID1 | PC_J14_ID1 | PC_J48_ID1 | PC_J50_ID1 | PC_J51_ID1 | PC_K02_ID1 | PC_K54_ID1 | PC_K55_ID1 | PC_K60_ID1 | PC_L06_ID1 | PC_L08_ID1 | PC_L10_ID1 | PC_L12_ID1 | PC_L14_ID1 | PC_L48_ID1 | PC_L56_ID1 | PC_M05_ID1 | PC_M15_ID1 | PC_M58_ID1 | PC_M59_ID1 | PC_N04_ID1 | PC_N07_ID1 | PC_N09_ID1 | PC_N10_ID1 | PC_N11_ID1 | PC_N13_ID1 | PC_N14_ID1 | PC_N48_ID1 | PC_N56_ID1 | PC_N57_ID1 | PC_N58_ID1 | PC_N60_ID1 | PC_O08_ID1 | PC_O10_ID1 | PC_O51_ID1 | PC_O59_ID1 | PC_I11_ID2 | PC_I50_ID2 | PC_I55_ID2 | PC_I58_ID2 | PC_J07_ID2 | PC_J14_ID2 | PC_J48_ID2 | PC_K55_ID2 | PC_L06_ID2 | PC_L12_ID2 | PC_L48_ID2 | PC_N03_ID2 | PC_N07_ID2 | PC_N10_ID2 | PC_N14_ID2 | PC_N57_ID2 | PC_N58_ID2 | PC_O08_ID2 | PC_O51_ID2 | PC_O59_ID2 |
|---|---|---

Table S10. Progeny-parent information for maize landrace.

Maternal Parent	Paternal Parent	164_1	164_2	164_4	164_5	164_8	164_9	164_11	164_14	165_2	165_4	165_5	165_6	165_7	165_10	166_1	166_3	166_8	166_9	166_10	166_11	166_12	167_1	167_3	167_5	167_6	167_8	167_9	167_12	167_13	168_2	168_5	168_8	169_1	169_6	169_7	169_11	169_12	170_1	170_2	170_4	170_6	170_7	170_11	170_12	170_13	171_3	171_4	171_5	171_6	171_7	171_8	171_9	171_10	171_11	171_13
	164_1	90	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	164_2	0	18	0	0	0	0	0	0	0	0	0	13	0	0	0	26	0	0	0	0	134	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	164_4	0	0	1	0	0	0	0	0	0	41	0	0	0	0	0	0	0	0	0	90	0	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	164_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	164_8	0	0	0	0	97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0
	164_9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	164_11	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
	164_14	91	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_4	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	87	0	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_6	0	10	0	0	0	0	0	0	0	0	0	5	0	0	0	12	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_7	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	165_10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_3	0	11	0	0	0	0	0	0	0	0	0	9	0	0	0	16	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_9	82	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	55	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166_12	0	7	0	0	0	0	0	0	0	0	0	8	0	0	0	10	0	0	0	0	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_3	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	52	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_5	74	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_8	0	0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0
	167_9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	167_12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	15	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0
	167_13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5	6	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0	0	0
	168_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	168_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	91	0	0	0	0	0	0	0	0	0
	168_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	169_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	169_6	74	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	169_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	169_11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	15	0	15	0	0	0	0	0	32	0	0	0	0	0	0	45	0	0
	169_12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	170_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	108	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	0
	170_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	125	0	13	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0
	170_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	0	0	0	13	0	28	0	0	0	0	0	20	0	0	0	0	0	0	58	0	0
	170_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88	0	24	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0
	170_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	20	0	39	0	0	0	0	0	25	0	0	0	0	0	0	50	0	0
	170_11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	66	0	5	0	35	0	0	0	0	0	0	0	0	0	0	0	0	0
	170_12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	5	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	2	0	0	0	0	0	0	0
	170_13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	171_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	171_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	139	0	0	0	0	0	0
	171_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	12	0	32	0	0	0	0	0	84	0	0	0	0	0	0	57	0	0
	171_6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	70	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0
	171_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	171_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	35	78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	22	0	0	0	0	0	0
	171_9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	171_10	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91	0	0	0	0	0	0
	171_11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	73	0	0	0	0	0	0	0	0	0	0	0	0	34	0	0	0	0	0
	171_13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	0	2	0	0	0	0	0	12	0	0	0	0	0	0	43	0	0

Table S12. Descriptions for 18 Comparable Teosinte and Maize Landrace Traits.

Trait	Acronym	Units	Teosinte Description	Maize Landrace Description
Days to Anthesis	DTA	days	Number of days from planting till the first day of at least one anther is shedding pollen; inspected visually.	Number of days from planting till the first day of 50% of anthers are shedding pollen; inspected visually.
Days to Silking	DTS	days	Number of days from planting till the first day of silks appearing; inspected visually.	Number of days from planting till the first day of silks appearing; inspected visually
Plant Height	PLHT	cm	Distance from ground to the node at the base of the flag leaf; measured using a meter stick.	Distance from ground to the node at the base of the flag leaf; measured using a meter stick.
Leaf Length	LFLN	cm	Length of the leaf on one of the 4^{th} to 6^{th} nodes from the top of the plant; measured using a tape measure.	Length of the leaf on the node bearing uppermost ear; measured using a tape measure.
Leaf Width	LFWD	cm	Width of the leaf on one of the 4^{th} to 6^{th} nodes from the top of the plant; measured using a tape measure.	Width of the leaf on the node bearing uppermost ear; measured using a tape measure.
Tiller Number	TILN	count	Number of tillers; counted visually.	Number of tillers; counted visually.
Prolificacy	PROL	count	Number of ears on the 2^{nd}/3^{rd} lateral branch from the top of the plant; counted visually.	Number of ears on the uppermost lateral branch; counted visually
Lateral Branch Node Number	LBNN	count	Number of nodes in the 2^{nd}/3^{rd} lateral branch from the top of the plant; counted visually.	Number of nodes/husks in the uppermost lateral branch; counted visually.
Lateral Branch Length	LBLN	mm	Length of the 2^{nd}/3^{rd} lateral branch from the top of the plant measured from main stalk to node below the terminal inflorescence; measured using a tape measure.	Length of the uppermost lateral branch measured from main stalk to node below the ear; measured using a tape measure.
Lateral Branch Internode Length	LBIL	mm	Average length between two nodes on a lateral branch; derived from LBLN/LBNN.	Average length between two nodes on a lateral branch; derived from LBLN/LBNN.
Ear Length	EL	mm	Length of an ear on the 2^{nd}/3^{rd} lateral branch from the top of the plant, averaged over two ears; measured using a ruler.	Length of the primary ear on the uppermost lateral branch; measured using a ruler.
Cupules per Row	CUPR	count	Number of cupules/fruitcases in a row; derived from GE/2.	Number of cupules/kernels in one row on the primary ear; counted visually.
Ear Diameter	ED	mm	Diameter of an ear as approximated by the average width of 10-50 fruitcases along the axis perpendicular to the fruitcase length; measured using SmartGrain.	Diameter of the primary ear on the uppermost lateral branch at its widest point; measured using a caliper.
Grains Per Ear	GE	count	Number of fruitcases in an ear on the 2^{nd}/3^{rd} lateral branch from the top of the plant, averaged over two ears; counted visually.	Total number of seeds of the primary ear on the uppermost lateral branch; counted visually.
Ear Internode Length	EILN	mm	Average length between two nodes on an ear; derived from EL/(CUPRx2).	Average length between two nodes on an ear; derived from EL/(CUPRxKRN/2) where KRN is number of rows on an ear.
Total Grain per Plant	TGPP	count	Predicted maximum number of fruitcases that a plant could have produced if all potential fruitcases were pollinated and developed to maturity; derived from GE x TEPP.	Total number of seeds on a plant; counted visually.
Total Grain Weight per Plant	TGWP	mg	Predicted total seed weight that a plant could have produced if all potential fruitcases were pollinated and developed to maturity; derived from GW x TGPP.	Total seed weight of a plant; measured on a weighing scale.
Grain Weight	GW	mg	Average weight of 50 seeds; derived from FCWT/2.	Average weight of 100 seeds, in which 50 seeds are from ears on uppermost lateral branch and another 50 are from all other ears; measured on a weighing scale.

References

1. Anderson LK, et al. (2004) Integrating genetic linkage maps with pachytene chromosome structure in maize. *Genetics* 166(4):1923-1933.
2. Bates D, Maechler M (2017) Matrix: Sparse and dense matrix classes and methods. R package version 1.2-10 (https://CRAN.R-project.org/package=Matrix).
3. Bauer E, et al. (2013) Intraspecific variation of recombination rate in maize. *Genome Biol* 14(9):R103.
4. Bradbury PJ, et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19):2633-2635.
5. Bulmer MG (1971) The effect of selection on genetic variability. *Am Nat* 105(943):201-211.
6. Cheverud JM, Marroig G (2007) Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genet Mol Biol* 30(2):461-469.
7. CIMMYT (2005) *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory Third Edition* (CIMMYT, Mexico).
8. Elshire RJ, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6(5):e19379.
9. Endelman JB, Jannink JL (2012) Shrinkage estimation of the realized relationship matrix. *G3* 2(11):1405-1413.
10. Flury B (1988) *Common principal components and related multivariate models* (Wiley, New York, USA).
11. Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thomson R (2015) *ASReml user guide release 4.1 structural specification* (VSN International Ltd, Hemel Hempstead, UK).
12. Glaubitz JC, et al. (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9(2):e90346.
13. Goudet J (2004) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol Ecol* 5(1):184-186.
14. Hospital F, Chevalet C (1993) Effects of population size and linkage on optimal selection intensity. *Theor Appl Genet* 86(6):775-780.
15. Leinonen T, Scott McCairns RJ, O'Hara RB, Merilä J (2013) Q_{ST}-F_{ST} comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nat Rev Genet* 14(3):179-190.
16. Lush JL (1937) *Animal Breeding Plans* (Iowa State College Press, Ames, Iowa, USA).
17. Mackay TFC, Lyman RF (2005) *Drosophila* bristles and the nature of quantitative genetic variation. *Phil Trans R Soc B* 360(1459):1513-1527.
18. Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27(2):209-220.
19. Martin G, Chapuis E, Goudet J (2008) Multivariate Qst-Fst Comparisons: A neutrality test for the evolution of the G matrix in structured populations. *Genetics* 180(4):2135-2149.

20. Money D, et al. (2015) LinkImpute: Fast and accurate genotype imputation for nonmodel organisms. *G3* 5(11):2383-2390.
21. Muñoz PR, et al. (2014) Unraveling additive from nonadditive effects using genomic relationship matrices. *Genetics* 198(4):1759-1768.
22. Ovaskainen O, Cano JM, Merilä J (2008) A Bayesian framework for comparative quantitative genetics. *Proc Royal Soc B* 275(1635):669-678.
23. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289-290.
24. Phillips PC, Arnold SJ (1999) Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury Hierarchy. *Evolution* 53(5):1506-1515.
25. R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (https://www.R-project.org).
26. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257-286.
27. Revell LJ (2012) An R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3(2):217-223.
28. Roff DA (1997) *Evolutionary Quantitative Genetics* (Chapman and Hall, New York, USA).
29. Schluter D (1996) Adaptive radiation along genetic lines of least resistance. *Evolution* 50(5):1766-1774.
30. Sidhu GK, et al. (2015) Recombination patterns in maize reveal limits to crossover homeostasis. *Proc Natl Acad Sci USA* 112(52):15982-15987.
31. Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135(2):367-374.
32. Tanabata T, Shibaya T, Hori K, Ebana K, Yano M (2012) *SmartGrain*: High-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol* 160(4):1871-1880.
33. van Heerwaarden J, et al. (2010) Fine scale genetic structure in the wild ancestor of maize (*Zea mays* ssp. *parviglumis*). *Mol Ecol* 19(6):1162-1173.
34. VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91(11):4414-4423.
35. Verrier E, Colleau JJ, Foulley JL (1990) Predicting cumulated response to directional selection in finite panmictic populations. *Theor Appl Genet* 79(6):833-840.
36. Walsh B, Lynch M (2018) *Evolution and Selection of Quantitative Traits* (Oxford University Press, Oxford, UK).
37. Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: Comparing Q_{ST} to F_{ST}. *Genetics* 183(3)1055-1063.
38. Whitlock MC (2008) Evolutionary inference from Q_{ST}. *Mol Ecol* 17(8):1885-1896.
39. Zhao H, et al. (2014) CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* 30(7):1006-1007.