Supplemental information for:

## TeoNAM: a nested association mapping population for domestication and agronomic trait analysis in maize

Qiuyue Chen<sup>\*,†</sup>, Chin Jian Yang<sup>\*</sup>, Alessandra M. York<sup>\*</sup>, Wei Xue<sup>\*,1</sup>, Lora L. Daskalska<sup>\*</sup>, Craig A. DeValk<sup>\*</sup>, Kyle W. Krueger<sup>\*</sup>, Samuel B. Lawton<sup>\*</sup>, Bailey G. Spiegelberg<sup>\*</sup>, Jack M. Schnell<sup>\*</sup>, Michael A. Neumeyer<sup>\*</sup>, Joseph S. Perry<sup>\*</sup>, Aria C. Peterson<sup>\*</sup>, Brandon Kim<sup>\*</sup>, Laura Bergstrom<sup>\*</sup>, Liyan Yang<sup>\*,‡</sup>, Isaac C. Barber<sup>\*</sup>, Feng Tian<sup>†</sup>, John F. Doebley<sup>\*,2</sup>

\* Laboratory of Genetics, University of Wisconsin-Madison, WI 53706, USA

 <sup>†</sup> National Maize Improvement Center, Key Laboratory of Biology and Genetic Improvement of Maize (MOA), Beijing Key Laboratory of Crop Genetic Improvement, Joint International Research Laboratory of Crop Molecular Breeding, China Agricultural University, Beijing 100193, China

<sup>‡</sup>School of Life Science, Shanxi Normal University, Linfen, Shanxi 041004, China

<sup>1</sup> Present address: College of Agronomy, Shenyang Agricultural University, Shenyang, Liaoning 110866, China

<sup>2</sup> Corresponding author: Laboratory of Genetics, University of Wisconsin–Madison, 425 Henry Mall, WI 53706. E-mail: jdoebley@wisc.edu

This file includes: Figure S1-S51



**Figure S1.** Diagram of constructing TeoNAM. TeoNAM was constructed by crossing five wild teosinte inbred lines (TIL01, TIL03, TIL11, TIL14, and TIL25) to a cultivated maize inbred line W22, followed by one generation of backcross to the common recurrent maize parent and four generations of selfing. A total of 1257 BC<sub>1</sub>S<sub>4</sub> RILs were produced. ×, crossing;  $\otimes$ , selfing.



**Figure S2.** Statistics of the genotypes in TeoNAM. (A) The percentage of RIL with heterozygous genotypes at each SNP site across the genome. (B) The percentage of RIL with homozygous teosinte genotypes at each SNP site across the genome. The five subpopulations are shown in different colors. (C) Recombination rate (cM/Mb) was calculated for the TeoNAM using the composite genetic map, using a window of 10 Mb and 1 Mb steps. (D) The number of genes was calculated using a window of 10 Mb and 1 Mb steps. Centromeres are shown in red dots.



**Figure S3.** Statistics of the allele frequency in TeoNAM. Maize and teosinte alleles are shown in different colors.



**Figure S4.** The phenotypic distribution of 22 domestication and agronomic traits in density plot. The five subpopulations are shown in different colors. CULM, PLHT, LFLN, LFWD and PROL were not scored in all five subpopulations.



**Figure S5.** The phenotypic distribution of 22 domestication and agronomic traits in boxplot. The five subpopulations are shown in different colors. CULM, PLHT, LFLN, LFWD and PROL were not scored in all five subpopulations.



**Figure S6.** Separate QTL mapping for ASI in different subpopulations. The horizontal and vertical axis represent chromosomal genetic positions and LOD scores, respectively. Each QTL is indicated by chromosome @ genetic position. Multiple QTLs on the same chromosomes are shown in different colors. Same legends are also used for Figure S7 to Figure S27.



Figure S7. Separate QTL mapping for BARE in different subpopulations.



Figure S8. Separate QTL mapping for CULM in different subpopulations.



Figure S9. Separate QTL mapping for DTA in different subpopulations.



Figure S10. Separate QTL mapping for DTS in different subpopulations.



Figure S11. Separate QTL mapping for EB in different subpopulations.



Figure S12. Separate QTL mapping for ED in different subpopulations.



Figure S13. Separate QTL mapping for EL in different subpopulations.



Figure S14. Separate QTL mapping for GLCO in different subpopulations.



Figure S15. Separate QTL mapping for GLUM in different subpopulations.



Figure S16. Separate QTL mapping for KRN in different subpopulations.



Figure S17. Separate QTL mapping for KW in different subpopulations.



Figure S18. Separate QTL mapping for LFLN in different subpopulations.



Figure S19. Separate QTL mapping for LFWD in different subpopulations.



Figure S20. Separate QTL mapping for PLHT in different subpopulations.



Figure S21. Separate QTL mapping for PROL in different subpopulations.



Figure S22. Separate QTL mapping for REPE in different subpopulations.



Figure S23. Separate QTL mapping for SHN in different subpopulations.



Figure S24. Separate QTL mapping for STAM in different subpopulations.



Figure S25. Separate QTL mapping for TBN in different subpopulations.



Figure S26. Separate QTL mapping for TILN in different subpopulations.



Figure S27. Separate QTL mapping for YEPE in different subpopulations.



**Figure S28.** Distinct genetic architectures for different traits. The R<sup>2</sup> distribution for 13 additional traits are shown here in different panels. The horizontal axis indicates QTLs and the vertical axis indicates the phenotypic variation explained by each QTL (R<sup>2</sup>). Red number indicates variance explained by the QTL model for each trait.



**Figure S29.** Genomic distribution of phenotypic difference between teosinte allele and maize allele against 51,544 SNPs for DTA.



**Figure S30.** QTL characterization for agronomic trait DTS. (A) Genomic distribution of QTLs detected by JLM. The known candidate genes are shown above the corresponding QTLs in bold italic. (B) Heat map shows additive allele effects of teosinte relative to maize for QTLs detected by JLM. The allele effect of teosinte parent 8759 was estimated from the 866 maize-teosinte BC<sub>2</sub>S<sub>3</sub> RILs (Shannon 2012). Insignificant effects are shown as blank. Red and blue color indicates that the teosinte allele increases or decreases effect, respectively. (C) Manhattan plot shows QTLs detected by GWAS. The significance threshold at LOD=5 is indicated by black dotted line. The red stars indicate GWAS signals overlapping with QTLs by JLM. In (A) and (C), chromosomes in odd and even numbers are shown in blue and orange colors, respectively. Same legends are also used for Figure S31 to Figure S49.



Figure S31. QTL characterization for agronomic trait TBN.









Figure S32. QTL characterization for domestication trait TILN.



Figure S33. QTL characterization for domestication trait GLUM.



Figure S34. QTL characterization for domestication trait PROL.



Figure S35. QTL characterization for domestication trait SHN.







Figure S36. QTL characterization for domestication trait KRN.



Figure S37. QTL characterization for domestication trait REPE.



Figure S38. QTL characterization for agronomic trait ASI.







Figure S39. QTL characterization for agronomic trait BARE.



Figure S40. QTL characterization for domestication trait CULM.



Figure S41. QTL characterization for domestication trait EB.







Figure S42. QTL characterization for domestication trait ED.







Figure S43. QTL characterization for domestication trait EL.



Figure S44. QTL characterization for domestication trait GLCO.







Figure S45. QTL characterization for domestication trait KW.



Figure S46. QTL characterization for domestication trait LFLN.



Figure S47. QTL characterization for domestication trait LFWD.



Figure S48. QTL characterization for agronomic trait PLHT.







Figure S49. QTL characterization for agronomic trait YEPE.



**Figure S50.** Comparison of support interval size for QTL detected in TeoNAM (JLM results) and  $BC_2S_3$  (Shannon, 2012).



**Figure S51.** Comparison of  $R^2$  for QTL overlapped in TeoNAM (JLM results) and BC<sub>2</sub>S<sub>3</sub> (Shannon, 2012).