Two teosintes made modern maize


INTRODUCTION AND RATIONALE: The drastic morphological differences between maize and its wild relatives gave rise to more than a century of debate about its origins. Today, the most widely accepted model is also the simplest—maize was domesticated once from the wild annual grass *Zea mays* ssp. *parviglumis* in the lowlands of southwest Mexico. More recently, however, genomic surveys of traditional maize varieties in both Mexico and South America have identified evidence for gene flow from a second wild relative, *Zea mays* ssp. *mexicana*, a weedy annual grass adapted to the central Mexican highlands. These results, combined with long-standing archaeological evidence of hybridization, challenge the sufficiency of a simple model of a single origin.

RESULTS: To elucidate the genetic contributions of *Zea mays* ssp. *mexicana* to maize, we analyzed >1000 wild and domesticated genomes, including 338 newly sequenced traditional varieties. We found ubiquitous evidence for admixture between maize and *Zea mays* ssp. *mexicana*, including in ancient samples from North and South America, diverse traditional varieties, and even modern inbred lines. These results are mirrored in a genotyping survey of >5000 traditional varieties representing maize diversity across the Americas. The only maize sample surveyed that lacks strong evidence for admixture with *Zea mays* ssp. *mexicana* is a single ancient South American sample N16, dating to ~5500 years before present.

We next fit graphs of population history to our data, revealing multiple admixture events in the history of modern maize. On the basis of these results, we propose a new model of maize origins, which posits that, some 4000 years after domestication, maize hybridized with *Zea mays* ssp. *mexicana* in the highlands of central Mexico. The resulting admixed maize then spread across the Americas, replacing or hybridizing with existing populations. The timing of this secondary dispersal is roughly coincident with archaeological data showing a transition to a staple maize diet in regions across Mesoamerica.

We then explored variation in ancestry along the maize genome. We found that 15 to 25% of the genome could be attributed to *Zea mays* ssp. *mexicana* ancestry. We identified regions in which *Zea mays* ssp. *mexicana* alleles had reached high frequency in maize, presumably as a result of positive selection. We investigated one of these adaptive introgressions in more detail, using CRISPR-Cas9 knockout mutants and overexpression lines to demonstrate the role of the circadian clock gene *ZmPRR37a* in determining flowering time under long-day conditions. Our results suggest that introgression at this locus may have facilitated the adaptation of maize to higher latitudes.

Finally, we explored the contributions of *Zea mays* ssp. *mexicana* alleles to phenotypic variation in maize. Admixture mapping identified at least 25 loci in modern inbred lines where highland teosinte ancestry associates with phenotypes of agronomic importance, from oil content to kernel size and disease resistance, as well as a large effect locus associated with cob diameter in traditional maize varieties. We then modeled the additive genetic variance of each phenotype, allowing us to estimate that *Zea mays* ssp. *mexicana* admixture explained a meaningful proportion of the additive genetic variation for many traits, including 25% of the variation for the number of kernels per row and nearly 50% of some disease phenotypes.

CONCLUSION: Our extensive population and quantitative genetic analysis of domesticated maize and its wild relatives uncovered a substantial role for two different wild taxa in making modern maize. We propose a new model for the origin of maize that can explain both genetic and archaeological data, and we show how variation in *Zea mays* ssp. *mexicana* is a key component of maize diversity, both at individual loci and for genetic variation underlying agronomic traits.

Our model raises a number of questions about how and why a secondary spread of maize may have occurred, but we speculate that the timing of admixture suggests a possible direct role for hybridization between maize and *Zea mays* ssp. *mexicana* in improving early domesticated forms of maize, helping to transform it into the staple crop we know today.

*These authors contributed equally to this work.

†Corresponding author: Email: rossibarra@ucdavis.edu (J.R.-I.); yjianbing@mail.hzau.edu.cn (J.Y.); ningy@mail.hzau.edu.cn (N.Y.)

Cite this article as N. Yang et al., Science 382, eadg8940 (2023). DOI: 10.1126/science.adg8940

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adg8940
The origins of maize were the topic of vigorous debate for nearly a century, but neither the current genetic model nor earlier archaeological models account for the totality of available data, and recent work has highlighted the potential contribution of a wild relative, *Zea mays* ssp. *mexicana*. Our population genetic analysis reveals that the origin of modern maize can be traced to an admixture between ancient maize and *Zea mays* ssp. *mexicana* in the highlands of Mexico some 4000 years after domestication began. We show that variation in admixture is a key component of maize diversity, both at individual loci and for additive genetic variation underlying agronomic traits. Our results clarify the origin of modern maize and raise new questions about the anthropogenic mechanisms underlying dispersal throughout the Americas.

The domestication of crops transformed human culture. For many crops, the wild plants that domesticates are most closely related to can be readily identified by morphological and genetic similarities. But the origins of maize (*Zea mays* ssp. *mays*) have long been fraught with controversy, even with its global agricultural importance, ubiquity, and extended scrutiny as a genetic model organism. Although there was general agreement that maize was most morphologically similar to North American grasses in the subtribe Tripsacinae (1, 2), none of these grasses bear reproductive structures similar to the maize ear, in which seeds are exposed along a compact, nonshattering rachis. The form is so radically distinct from its relatives that the maize ear has been called “teratological” (3) and a “monstrosity” (4).

Explanations for the ancestry of maize have long been contentious (3). A model popular for much of the 20th century, based on extensive evaluation of the morphology of archaeological samples, argued that modern maize was the result of hybridization between a now-extinct wild maize and another wild grass (6). This archaeological model, however, fails to explain cytological (7) or genetic (8, 9) data showing that maize is most closely related to the extant wild grass *Zea mays* ssp. *parviglumis* (hereafter *parviglumis*). Today, the most widely accepted model is also the simplest—maize was domesticated from a wild annual grass in the genus *Zea*, commonly known as teosinte. This idea, originating with Ascherson (10) and championed by George Beadle throughout the 20th century (4, 7), became firmly cemented in the literature after genetic analysis revealed clear similarities between maize and teosinte (8, 9, 11). Nonetheless, this simple genetic model is insufficient to explain disparities between genetic and geographic overlap between maize and *parviglumis* (12) or morphological support for admixture in archaeological samples (13–15).

Much of the early work on maize origins was complicated by the relatively poor characterization of the diversity of annual teosinte (16). In addition to the lowland *parviglumis*, the other widespread annual teosinte is *Z. mays* ssp. *mexicana* (hereafter *mexicana*), found throughout the highlands of Mexico. These taxa diverged 30,000 to 60,000 years ago (17, 18) and show clear morphological (19), ecogeographic (20, 21), and genetic (22, 23) differences, as well as local adaptation along elevation (24). In contrast to the overall genetic similarities between maize and *parviglumis*, some early genetic studies identified sharing of alleles between *mexicana* and highland maize (25), a result confirmed by extensive genome-wide data (26, 27). Maize and *mexicana* co-occur in the highlands of Mexico, but recent work has revealed *mexicana* ancestry far outside this range, including in ancient maize from New Mexico (28), modern samples in the Peruvian Andes (29), and individual alleles apparently selected broadly in modern maize (30, 31).

Admixture with *mexicana* is ubiquitous in modern maize

Archaeological data suggest that after its initial domestication in the lowlands of the Balsas River basin, maize was introduced to the highlands of central Mexico by ~6200 cal BP (calendar years before present) (32), where it first came into sympathy with *mexicana*. By this time, however, maize had already reached Panama (by ~7800 cal BP) (33) and even farther into South America (by ~6900 to 6700 cal BP) (34–36). Samples from South America that reflect dispersal events before maize colonization of the Mexican highlands should thus not exhibit evidence of admixture with *mexicana*. Indeed, tests of admixture find no evidence of *mexicana* ancestry in N16, a ~5500 cal BP maize cob from northern Peru (37).

To investigate evidence of *mexicana* admixture across a broad sampling of maize, we applied $f_2$ tests (38) using a sample of the diploid perennial teosinte *Zea diploperennis* (39) as the outgroup. The greatest diversity of maize is found in present-day Mexico, but whole-genome resequencing exists for only a handful of traditional Mexican maize (40). We therefore sequenced 267 accessions of open-pollinated traditional maize from across Mexico (Fig. 1A and data S1 and S2). Applying $f_2$ tests revealed significant admixture with *mexicana* in all maize except the ancient Peruvian sample N16 (Fig. 1B and data S3). Analysis of subsets of the data allowing use of a greater number of single-nucleotide polymorphisms (SNPs) reveals higher $f_2$ values for N16, but these are still substantially lower than for all other samples (data S4). We find evidence for *mexicana* admixture well outside of Mexico, including in modern samples from the US Southwest and the Andes, consistent with previous work (29), as well as a newly sequenced set of 73 traditional Chinese varieties that represent maize dispersal out of the Americas after European
colonization (data S1 and S2). We extended our search to ancient samples, again finding *mexicana* admixture in archaeological samples from the Tehuacán Valley in central Mexico dating to ~5300 cal BP (41) and both lowland and highland (>2000 m above sea level) samples from South America dating to ~1000 cal BP (42). Finally, we turned to modern breeding material, where, again, f<sub>4</sub> tests identify significant admixture in a diversity panel of >500 modern inbred lines (43). In sum, we find evidence of *mexicana* ancestry in all examined maize samples dating as early as ~5300 cal BP.

To investigate the importance of introgression to maize diversity more broadly, we ran STRUCTURE (44) to estimate *mexicana* ancestry in genotyping data from a much larger sample of 5373 traditional maize varieties and 310 wild samples of *mexicana* and *parviglumis* from across the Americas (45, 46). These maize samples also show ubiquitous evidence of *mexicana* admixture (Fig. 1C). More surprisingly, principal component analyses of these maize samples reveal that the major axis of genetic variation across maize in the Americas is nearly perfectly correlated with the proportion of genetic variation across maize in the Americas, resulting in admixture estimated for ~5000 field collections from the International Maize and Wheat Improvement Center (CIMMYT). Correlation (R<sup>2</sup> = 0.97) between the first principal component (PC1) of genetic diversity in 5373 CIMMYT traditional maize varieties and *mexicana* admixture (Fig. 1C). More surprisingly, principal component analyses of these maize samples reveal that the major axis of genetic variation across maize in the Americas is nearly perfectly correlated with the proportion of the genome showing *mexicana* ancestry [coefficient of determination (R<sup>2</sup>) = 0.97; Fig. 1D]. The subspecies *mexicana* and *parviglumis* diverged tens of thousands of years before domestication (18), and it thus seems plausible that differences in the proportion of these ancestries dominate more recently derived aspects of maize diversity.

**A novel model of maize origins**

We interpret the timing and universality of *mexicana* admixture as evidence supporting a novel model of maize origins (Fig. 2A). Consistent with previous work (42, 47), we propose that maize dispersed out of the Balsas River basin in Mexico after domestication from *parviglumis*, quickly reaching South America by at least ~6500 cal BP (48). Then, ~6000 cal BP, maize was adopted by peoples living in the highlands of central Mexico, resulting in admixture with the sympatric *mexicana* (12, 26, 27, 32). Ancient samples showing *mexicana* admixture suggest that maize spread rapidly from the highlands of Mexico, replacing or mixing with existing populations across the Americas, introducing *mexicana* alleles as it moved. As it entered into the lowlands of Mexico, maize must have once again come into contact with *parviglumis*. This model is consistent with the second wave of maize migration into South America posited by Kistler et al. (49) and supported by recent chloroplast data (50) but further explains the origin of that wave and the existence of *mexicana* alleles in South America.
far outside *mexicana*’s native range, by at least 1000 cal BP.

To formally evaluate our model, we fit admixture graphs to $f_2$ statistics from five traditional maize varieties from the Andes (40) and 118 of our newly sequenced traditional maize varieties from Mexico that were collected at low elevation (<1500 m above sea level) (51). Our fitted admixture graph (Fig. 2A) estimates that initial hybridization with *mexicana* was substantial, consistent with estimates from modern-day traditional maize varieties from the highlands of central Mexico (27). Subsequent admixture with *parviglumis* reduced the contribution of *mexicana* to maize ancestry out of the highlands, but our model estimates that *mexicana* ancestry still represents ~25% of the genome of extant traditional varieties in Mexico. A simplified version of the model estimates nearly identical ancestry proportions for modern maize inbred lines (fig. S1). Although these estimates are lower than those from reduced representation genotyping (Fig. 1C), genotyping SNPs overestimate genome-wide admixture proportions owing to their biased distribution across the genome (51) (fig. S2). A thorough evaluation of alternative graphs (52) found models with nominally better fits to the data (53), but none had significantly better out-of-sample predictive ability (data S5; lowest empirical bootstrap $P$ value of 0.38), and we were unable to identify a better-fitting graph with fewer admixture events. Although many of the alternative graphs include implausible phylogenetic histories, and some incorporate hybrid *parviglumis-mexicana* populations (54), all of the alternative graphs qualitatively support our proposed model in requiring postdomestication admixture with *mexicana*. Finally, we note that our model is consistent with an independent population genetic approach (55) that estimated the timing of *mexicana* admixture at 5716 years (±5614) (51), which is exceptionally close to the earliest archaeological evidence of maize in the highlands (32) and substantially later than the first evidence of domesticated maize (56).

Together, the confluence of archaeological and genetic data suggests that *mexicana* admixture was central to the widespread use and dispersal of maize in the Americas by ~4000 cal BP (Fig. 2B). The timing of admixture between maize and *mexicana* in the highlands of Mexico between 6000 and 4000 cal BP corresponds with observed increases in cob size and the number of seed rows in archaeological samples (57–59). Southward dispersal of maize varieties with *mexicana* admixture coincides with the appearance of improved maize varieties in Honduras (49) but contrasts with a coeval movement of peoples northward (60). Archaeological samples demonstrate the presence of maize as a staple grain in the neotropical lowlands of Central America subsequent to *mexicana* admixture, between 4700 and 4000 cal BP (59, 61, 62). Ultimately, all varieties of maize in Mesoamerica had *mexicana* admixture by ~3000 cal BP as it became a staple grain across the entire region (28, 58, 59, 63). Early Mesoamerican sedentary agricultural villages also began developing at this time, forming the basis for demographic expansion and the emergence of later state-level societies dependent on more-intensive forms of maize agriculture (64–67).

**Variation in admixture along the genome**

Having established a central role for both *parviglumis* and *mexicana* in the origins of modern maize diversity, we next explored variation in *mexicana* ancestry across the genome. Using unadmixed *parviglumis* and *mexicana* individuals (39) as references, we applied an ancestry hidden Markov model to identify regions of *mexicana* ancestry along individual maize genomes (51). In close agreement with our admixture graph, we estimate 15 to 25% average *mexicana* ancestry across 845 maize genomes (mean: 18%; data S6) (68). This variation in total ancestry among modern maize is much greater than that predicted from a single pulse of ancient admixture (51) and likely reflects a combination of selection as well as ongoing gene flow in parts of the range (27). *Mexicana* ancestry also varies considerably along the genome (Fig. 3, A and B). The vast majority of introgressed haplotypes are small—on the scale of 10 kb (fig. S3)—consistent with a relatively ancient origin. In addition to numerous small introgressed haplotypes, we identify signals consistent with an important role for inversion polymorphisms. These include the apparent presence of the large inversion *Inv4m*—a well-studied target of adaptive introgression in maize from highland environments (26, 27)—in two Chinese inbred lines and one traditional Mexican variety (Fig. 3A). We also see high levels of *mexicana* admixture in the region of *Inv5m*, a 50-Mb inversion common in *parviglumis* but rare in *mexicana* and entirely absent in maize (69) (fig. S4). Finally, we estimate drastically decreased levels of *mexicana* introgression for chromosomes 5, 8, and 9 (fig. S5), which we speculate may be attributable to the presence of a recently characterized genetic incompatibility on chromosome 5 (70) and multiple large *mexicana*-specific inversions on chromosomes 8 and 9 that could hinder introgression by repressing recombination (18, 22).

A detailed look at admixture along individual genomes enabled us to begin to investigate the functional significance of variation in *mexicana* admixture. First, consistent with the possibility that *mexicana* alleles may have served to complement recessive deleterious genetic variants that rose to appreciable frequency in early domesticated maize (40), we find significantly lower genetic load on introgressed *mexicana* haplotypes (51) (fig. S6). We then turned to individual loci, identifying regions of the genome in which high-confidence *mexicana* alleles (>90% posterior probability) were at high frequency (>80%) across all modern maize, consistent with recent positive selection (51) (Fig. 3, A and B, and fig. S7). We found these loci clustered into 11 regions, which overlap quantitative...
trait loci for agronomically relevant phenotypes (71) and include genes with well-studied functions in Arabidopsis such as disease resistance and floral morphology (data S7).

We focused on one region on chromosome 7, where we found a narrow peak of high-frequency mexicana alleles (>90% posterior probability) that exist in >80% lines of all modern maize along chromosome 4 (black points) and average mexicana ancestry (red). (Bottom) Mexican ancestry of three inbred lines in the region around chromosome inversion Inv4m.

Fig. 3. Variation and functional validation of mexicana admixture. (A) (Top) Number of high-confidence mexicana alleles (>90% posterior probability) that exist in >80% lines of all modern maize along chromosome 4 (black points) and average mexicana ancestry (red). (Bottom) Mexican ancestry of three inbred lines in the region around chromosome inversion Inv4m. (B) (Top) Number of high-confidence mexicana alleles (>90% posterior probability) that exist in >80% lines of all modern maize along chromosome 7 (black points) and average mexicana ancestry (red). (Bottom) Mexican ancestry in B73 across the ZmPRR37a gene model (black bar). The differences of days to anthesis for nontransgenic (NT) and overexpression (OE) lines of ZmPRR37a in (C) long-day (LD) conditions (2022, China, 124°49’E, 43°30’N) and (D) short-day (SD) conditions (2021, China, 108°43’E, 18°34’N). The data in (C) and (D) are means ± SE. The numbers in each column indicate the sample sizes. The level of significance was determined by a two-tailed Student’s t test. (E) Nontransgenic and two independent overexpression lines of ZmPRR37a grown in long-day conditions. Scale bar, 10 cm.

from a targeted mutagenesis library (74) and developed two transgenic overexpression lines (51). Consistent with its hypothesized role in response to day length, ZmPRR37a knockout mutants exhibited a significantly earlier flowering phenotype in long-day conditions (two-tailed Student’s t test; P values are indicated in fig. S9, A, B, and D) but showed no effect in short-day conditions (two-tailed Student’s t test; P values are indicated in fig. S9, A and C), and overexpression lines exhibited significantly later flowering in both long- and short-day conditions (two-tailed Student’s t test; P values are indicated in Fig. 3, C to E). Maize carrying the mexicana introgression at ZmPRR37a shows lower levels of expression than parviglumis (75), and our functional evaluation thus suggests that mexicana alleles at ZmPRR37a may have helped maize adapt to earlier flowering in long-day conditions as it expanded out of Mexico to higher latitudes.

Admixture with mexicana underlies phenotypic variation in maize

Admixture with teosinte has been associated with phenotypic variation for a number of traits in traditional maize (76), and mexicana gene flow has been instrumental in the phenotypic adaptation of maize to the highlands (26, 77–79). Our analysis of teosinte ancestry across named varieties replicates historical estimates based on morphology (fig. S10 and data S8), suggesting a broader role for mexicana ancestry in pattern- ing phenotypic variation in maize. Indeed, if mexicana admixture played a key role in the dispersal and use of maize, mexicana alleles should contribute to agronomically relevant phenotypic variation. We thus combined our
estimates of admixture with data from 33 phenotypes to perform multivariate admixture mapping across 452 maize inbreds (31). At a false discovery rate of 10%, we identified 92 associations, which we grouped into 22 peaks representing 25 candidate genes (Fig. 4, fig. S11, and data S9 and S10). These include a significant association with zeaxanthin—a carotenoid pigment that plays a role in light sensing and chloroplast movement (80)—approximately 1-kb downstream of the gene ZmZEP1, a key locus in the xanthophyll cycle that regulates zeaxanthin abundance in low-light conditions (fig. S12A). Haplotype visualization reveals clear sharing between maize and mexicana (fig. S12B), and the mexicana-like haplotype increases the expression of ZmZEP1 and reduces zeaxanthin content in maize kernels (fig. S12, C and D). We also see associations with well-known lipid metabolism genes such as dgat1 and fad2 (82). The mexicana allele at dgat1 is associated with a decrease in the proportion of linoleic acid but an increase in overall oil content, but variation in mexicana ancestry is not in linkage disequilibrium with the well-studied amino acid variant at this locus (83) (fig. S13). Although expression of dgat1 has been suggested to play a role in cold tolerance in maize and Arabidopsis (84, 85), a preliminary experiment in maize seedlings failed to identify differences in cold tolerance in lines of varying ancestry at dgat1 (fig. S14). Finally, in addition to identifying compelling candidate loci in modern inbreds, we applied a novel genotype-by-environment association mapping approach (86) in a large set of traditional maize varieties evaluated across 13 different common garden trials (87). We find a strong association on chromosome 1 (Fig. 4C), where mexicana ancestry increases cob size. The candidate gene closest to the associated SNP, Zm00001d029675, was recently identified as a target of selection during breeding efforts in both the United States and China (88).

While genome-wide association studies (GWASs) can identify individual loci with large effects, it is likely that mexicana admixture contributes important variation of smaller effect size to polygenic traits. To test this hypothesis, we used our inbred association panel and phenotype data to estimate the proportion of additive genetic variance contributed by mexicana (51) (data S11). We estimate that mexicana admixture explains a meaningful proportion of the additive genetic variation for many of these traits, including nearly 25% for the number of kernels per row, 15% for plant height, 10% for flowering time, and 15 to nearly 50% for multiple disease phenotypes (Fig. 4D).

**Discussion**

Conflicting archaeological, cytological, genetic, and geographic evidence led to two irreconcilable models for the origin of maize. In this study, with more than 1000 genomes of maize and teosinte, including 338 newly sequenced traditional varieties, we revisited the evidence for admixture between maize and its wild relative Zea mays ssp. mexicana. We propose a new model of maize origins, which posits that, after admixture with mexicana in the highlands of central Mexico, admixed maize spread across the Americas, either replacing or hybridizing with preexisting maize populations. While this model is consistent with both genetic and archaeological data,
it also raises a number of questions. Among these, most notable is perhaps the question of why and how this secondary spread occurred—was it due to some advantage of the admixed maize over earlier domesticated forms, or was the spread coincidental with demic or cultural exchange among human populations (67).

Changes in maize cob morphology and dietary isotope data from human populations in Central America indicate a transition between early cultivation and the use of maize as a staple grain between 4700 and 4000 cal BP (39). This timing suggests a possible direct role for hybridization between maize and *mexicana* in improving early domesticated forms of maize. To better understand why admixed maize may have been beneficial for early farmers, we sought to investigate associations between *mexicana* alleles and phenotypes in extant maize. We identified and functionally validated a locus important for photoperiodicity and flowering time and found candidate genes associated with important agronomic phenotypes, including nutritional content and the size of kernels and cobs. None of these loci individually, however, are likely sufficient to drive a large advantage of admixed maize. And although we show that, combined, alleles introgressed from *mexicana* explain a meaningful proportion of additive genetic variance for agronomic and disease resistance traits, it remains unclear whether this novel variation could drive rapid adoption of admixed maize. In addition to variation at these specific phenotypes, admixture may have played a role in the spread of maize by augmenting genetic diversity and ameliorating genetic load in early domesticated populations, perhaps even providing some generalized hybrid vigor. Indeed, we show that *mexicana* alleles carry less load than maize alleles (fig. S6), and maize-*mexicana* hybrids show extensive heterosis for both viability and fecundity. This process could be augmented by similar ecologies as well—the global ecological niche of cultivated maize more closely reflects that of *mexicana* than *parviglumis* (69), and, like maize, *mexicana* has successfully colonized novel habitats at higher latitudes (89). Modern ethnographic evidence is also consistent with these ideas, as farmers continue to introgress teosinte into their maize populations to make them “stronger” (16, 90, 91).

Introgression between relatives has long been recognized as a major source of plant adaptation (92), yet only with the advent of molecular markers have we begun to recognize the key role that gene flow from wild relatives has played in crop evolution (93). These results not only highlight the past importance of crop wild relatives but also point to their potential as a source of adaptive diversity for future breeding. Most importantly, the work presented here suggests that, for many crops, millennia of diligent efforts by early farmers have capitalized on this diversity and that an abundance of relevant functional diversity may already be segregating in traditional varieties or preserved ex situ in germplasm gene banks.

### Materials and methods summary

SNP data from 507 modern maize inbred lines, 90 *Z. mays* ssp. *Mexicana*, 75 *Z. mays* ssp. *parviglumis*, and two *Z. diploperennis* were obtained from version 1 of the ZEAMAP project (39). We also sequenced an additional 338 traditional maize varieties, including 267 from across Mexico and 71 from China (data S1 and S2) and collected DNA sequencing data of 30 published traditional varieties and 10 ancient maize samples (data S1). For these additional genomes, we called sites from the enlarged ZEAMAP of these lines. For ancient maize, we did complete quality control on the raw reads by cutting low-quality bases and removing the adaptors using fastq (97). Then we adopted an ancient DNA mapping method optimized for reducing reference sequence bias and improving the accuracy and sensitivity of ancient DNA sequence identification (98). We used mapDAMAGE2 (98) to estimate damage parameters from the bam files, and then we rescaled base scores >1.8. The functions of proteins with PEAK scores >1.8 and mapping quality of <20 supported by <2 reads and reads with mean Phredscore of <20 and mapping quality of <20 were estimated by STRUCTURE (55). The genome-wide patterns of introgression of all 845 maize were investigated by admixture graphs estimated and compared using ADMIXTOOLS2 (52). The timing of admixture between *mexicana* and maize was estimated by DATES (55). Admixture with *mexicana* in CIMMYT SeeD GBS samples was estimated by STRUCTURE (44). The genome-wide patterns of introgression of all 845 maize were investigated by ELAI (102). We defined high-frequentance *mexicana* alleles as those for which >80% of the 845 maize lines had ELAI scores >1.8. The functions of *ZmPRR37a* were investigated by transgenic overexpression or CRISPR-Cas9 gene editing. The constructed overexpression and gene-editing vectors were transformed into maize inbred line KN5855. Genome-wide association of *mexicana* ancestry was performed using JointGWAS (86) for 33 phenotypes in an inbred association panel (103) and a multisite set of phenotypic trials of traditional varieties (87). Variance partitioning of phenotypes of 507 maize lines was performed by LDAD (104) using the kinship calculated by OSCA (105) from ELAI scores. All details of the materials and methods, including those summarized above, are provided in the supplementary materials.

### REFERENCES AND NOTES

96. A. M. Poets, Z. Fang, M. T. Clegg, P. L. Morrell, Barley
95. E. Chen, X. Huang, Z. Tian, R. A. Wing, B. Han, The genomics
et al
94. J. Blanca
Introgressive Hybridization
91. C. G. Takahashi, L. L. Kalns, J. S. Bernal, Plant defense
89. V. Le Corre, M. Siol, Y. Vigouroux, M. I. Tenaillon, C. Délye,
pmid: 31035826
pmid: 35184177
pmid: 35851624
pmid: 32989136
pmid: 26293830
pmid: 25420534
pmid: 22574660
pmid: 10.1186/s12859-014-0356-4; pmid: 25420514
pmid: 31138268
pmid: 31138262
pmid: 10.1186/s13059-015-0712-3; pmid: 26293830
pmid: 10.1038/s41477-022-01190-2; https://doi.org/10.6084/m9.figshare.23635437.v1.
100. L. Kistler, bamMask.pl, Figshare, Software (2023); https://doi.org/10.6084/m9.figshare.21718076.v1.
98. M. Schubert
96. A. M. Poets, Z. Fang, M. T. Clegg, P. L. Morrell, Barley
95. E. Chen, X. Huang, Z. Tian, R. A. Wing, B. Han, The genomics
et al
94. J. Blanca
Introgressive Hybridization
91. C. G. Takahashi, L. L. Kalns, J. S. Bernal, Plant defense
89. V. Le Corre, M. Siol, Y. Vigouroux, M. I. Tenaillon, C. Délye,
pmid: 31035826
pmid: 35184177
pmid: 35851624
pmid: 32989136
pmid: 26293830
pmid: 10.1186/s13059-015-0712-3; pmid: 26293830
pmid: 10.1186/s13059-015-0712-3; pmid: 26293830
pmid: 10.1038/s41477-022-01190-2; https://doi.org/10.6084/m9.figshare.23635437.v1.
100. L. Kistler, bamMask.pl, Figshare, Software (2023); https://doi.org/10.6084/m9.figshare.21718076.v1.
98. M. Schubert
96. A. M. Poets, Z. Fang, M. T. Clegg, P. L. Morrell, Barley
95. E. Chen, X. Huang, Z. Tian, R. A. Wing, B. Han, The genomics
et al
94. J. Blanca
Introgressive Hybridization
91. C. G. Takahashi, L. L. Kalns, J. S. Bernal, Plant defense
89. V. Le Corre, M. Siol, Y. Vigouroux, M. I. Tenaillon, C. Délye,
pmid: 31035826
pmid: 35184177
pmid: 35851624
pmid: 32989136
pmid: 26293830
pmid: 10.1186/s13059-015-0712-3; pmid: 26293830
pmid: 10.1186/s13059-015-0712-3; pmid: 26293830
pmid: 10.1038/s41477-022-01190-2; https://doi.org/10.6084/m9.figshare.23635437.v1.
100. L. Kistler, bamMask.pl, Figshare, Software (2023); https://doi.org/10.6084/m9.figshare.21718076.v1.
Two teosintes made modern maize

Science 382 (6674), eadg8940. DOI: 10.1126/science.adg8940

Editor's summary
Domestication of plants and animals is often characterized by selection for specific traits interspersed with introduction of new desirable traits from wild relatives. Yang et al. examined genetic data from more than 1000 varieties of maize and related species to clarify the complex origins of this agricultural staple. They found evidence that after initial domestication, introgression from a relative of domesticated maize, Zea mays ssp. mexicana, occurred in the highlands of Mexico before propagating across Central America. Alleles from this wild relative affect photoperiodicity and flowering time, which suggests that traits from Zea mays ssp. mexicana may have been beneficial during domestication. These results demonstrate the importance of broad sampling in elucidating the history of domesticates.
—Corinne Simonti

View the article online
https://www.science.org/doi/10.1126/science.adg8940
Permissions
https://www.science.org/help/reprints-and-permissions