Genome Biology

RESEARCH Open Access

Global genetic dissection of maize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator

Renyu Zhang (张人予) [1†], Xuan Zhang (张璇) [1†], Shenshen Wu (吴伸伸) [2†], Lichun Cai (蔡立春) [1†], Xiaowei Li (李晓伟) [1], Xiangyu Zhao (赵翔宇) [1], Zhenyuan Chen (陈真远) [1], Wenkang Chen (陈文康) [1], Jianghua Guo (郭江华) [1], Weiya Li (李威亚) [1], Ce Guo (郭策) [1], Dongzhe Yan (闫东哲) [1], Yangyang Li (李阳阳) [1], Yun Luo (罗芸) [2], Xingyu Ge (葛行宇) [1], Wenqiang Li (李文强) [2,6], Junhong Zhuang (庄军红) [1,3], Fang Yang (杨芳) [4], David Jackson[5], Jianbing Yan (严建兵) [2,6], Jiansheng Li (李建生) [1,3*], Ning Yang (杨宁) [2,6*] and Xiaohong Yang (杨小红) [1,3*]

[†]Renyu Zhang, Xuan Zhang, Shenshen Wu and Lichun Cai contributed equally to this work.

*Correspondence: lijiansheng@cau.edu.cn; ningy@mail.hzau.edu.cn; yxiaohong@cau.edu.cn

[1] Frontiers Science Center for Molecular Design Breeding, State Key Laboratory of Plant Environmental Resilience, and National Maize Improvement Center of China, China Agricultural University, Beijing 100193, China
[2] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China
Full list of author information is available at the end of the article

Abstract

Background: Understanding the genetic basis of trait divergence between maize and its wild progenitor teosinte is critical for elucidating crop domestication processes and accelerating maize improvement. This divergence is governed by multiple loci, including genes with pleiotropic effects and complex genetic interactions that collectively contribute to morphological variation.

Results: Using a maize–teosinte (Mo17–*mexicana*) introgression population, we identified 93 additive quantitative trait loci (QTLs) and nine epistatic interaction pairs affecting 20 agronomic traits, revealing a polygenic architecture underlying the domestication and improvement of these traits. The constructed QTL–trait network, together with identified genome-wide selection signals, revealed extensive genetic interconnections, with correlated traits sharing common loci, indicating a shared genetic basis for morphological divergence. Selection feature analysis further demonstrated that domestication, improvement, and *mexicana* introgression targeted key genomic regions associated with agronomic traits. Furthermore, we cloned an ear length QTL, *EL3-2*, which encodes a ULTRAPETALA (ULT) transcriptional regulator. *EL3-2* may function as a central pleiotropic regulator of domestication by modulating maize–teosinte divergence and regulating complex gene-expression networks across multiple developmental stages of maize.

Conclusions: This study reveals a complex genetic basis for maize–teosinte morphological divergence and highlights the role of pleiotropic regulators in crop domestication. Our findings provide new insights that bridge evolutionary genomics and breeding for complex traits in maize.

Keywords: Maize, *Mexicana*, Genetic architecture, Domestication and improvement, *Ear Length3-2*

Background

Maize (*Zea mays* L. ssp. *mays*) is one of the most productive and widely grown crops worldwide, serving as a vital source of food, animal feed, and industrial raw material. After its initial domestication from the wild teosinte *Zea mays* ssp. *parviglumis* (hereafter, *parviglumis*) in the lowlands of the Balsas River basin approximately 9,000 years before present (BP), maize experienced introgression from another wild teosinte *Zea mays* ssp. *mexicana* (hereafter, *mexicana*) in the highlands of central Mexico around 6,000 years BP [1–3]. This two-teosinte origin was pivotal for the expansion and adaptation of maize across diverse global environments [3–5]. During domestication and improvement, maize underwent dramatic changes in plant architecture, reproductive structure, kernel morphology, and adaptability [6–8]. Therefore, systematically understanding the genetic and molecular mechanisms underlying domestication and improvement traits will better enable directed manipulations of targeted traits to accelerate maize improvement.

Previous quantitative trait locus (QTL) mapping studies using various maize–teosinte populations identified hundreds of loci associated with domestication and improvement traits, providing critical insights into the genetic basis of morphological and adaptive shifts [9–16]. Multiple genetic approaches have further identified dozens of major domestication and improvement/adaptation genes. Notably, small alterations in single genes, such as *teosinte branched 1* (*tb1*) controlling branching, *grassy tillers1* (*gt1*) affecting prolificacy, and *teosinte glume architecture1* (*tga1*) controlling the formation of the stony fruit case, were found to drive key morphological transformations [17–20]. In addition to these large-effect loci, some morphological transitions between teosinte and maize involve the gradual accumulation of multiple genes with modest effects. A typical example is kernel row number, for which several loci/genes, including *KRN4*/*ub3*, *krn1*/*ids1*, *KRN2*, and *KRN5b*, have been identified, with each contributing an effect of fewer than two rows [21–24]. Similar to the domestication process underlying kernel row number enhancement, ear length and kernel number per row are quantitatively regulated by multiple loci. Genes such as *KERNEL NUMBER PER ROW6* (*KNR6*), *ZmACO2*, *YIGE1*, *EAR APICAL DEGENERATION1* (*EAD1*), and *ZmTPK2* have been identified through map-based cloning or association analyses, further highlighting the polygenic architecture underlying maize yield enhancement [25–29]. Recently, *tasselsheath4* (*tsh4*) was also found to be a domestication target, acting upstream of multiple domestication loci. *tsh4* functions to establish developmental boundaries and define meristem fates, and it plays a critical role in enhancing yield and improving crop architecture [30]. Following domestication, a subset of genes experienced intensive selection during maize improvement, facilitating adaptation to diverse agricultural practices. For instance, selection on *ZmCCT10*, *ZmCCT9*, *ZEA CENTRORADIALIS8* (*ZCN8*), and *ZmMADS69* contributed to maize adaptation to temperate regions with long-day conditions [31–34].

In addition to the cloning of QTLs, genome-wide surveys suggest that ~3.3%–10.9% of the genome, containing ~2,000–3,000 genes, underwent selection during maize domestication and improvement/adaptation [23, 35–38]. Studies leveraging larger populations and higher-density SNPs captured a broader range of genetic diversity between maize and teosinte [23, 38]. However, the functional roles of many genomic regions with strong selection signatures, particularly those exhibiting complex

inheritance patterns or pleiotropic effects, remain largely unknown. The integration of genomic approaches with traditional QTL cloning can help address the knowledge gaps and advance our understanding of the molecular basis of maize evolution.

The highland *mexicana* diverged from the lowland *parviglumis* approximately 30,000 years BP [38]. They both contributed to the origin of maize and shared similar teosinte-like traits, including plant, inflorescence, and kernel morphology. However, *mexicana* has abundant macrohairs and high pigmentation, and is well adapted to highland environments characterized by cool temperatures, high ultraviolet radiation, and low precipitation [39–43]. These morphological and physiological differences between *mexicana* and *parviglumis* reflect their distinct ecological adaptations and evolutionary histories. Thus, systematic investigation of the loci governing divergent traits between maize and *mexicana* can further highlight the potential for accelerating future breeding by reintroducing genetic diversity.

The construction of maize × *mexicana* introgression population and the availability of reference genomes for both maize and *mexicana* [5] can provide new opportunities to identify the genes underlying the divergent traits between maize and *mexicana*, assess the effects of *mexicana* alleles, and broaden the genetic diversity of cultivated maize. In this study, we dissected the genetic architecture of 20 divergent traits between wild and cultivated maize using a Mo17–*mexicana* introgression population (hereafter, TM population). We further constructed a QTL–trait network by integrating QTL–trait associations and QTL–QTL co-localizations. The QTLs were subsequently characterized using genome-wide signatures of selection during domestication and improvement of maize. Our studies enabled the identification and cloning of a key QTL associated with ear traits that is central to domestication and yield. These findings will enhance our understanding of maize domestication and improvement at the molecular level, deepen insights into the genetic control of ear development, and provide valuable genetic diversity for modern breeding.

Results

Genetic architecture of 20 divergent agronomic traits between maize and teosinte

The TM population was developed from a cross between the elite maize inbred line Mo17 and the teosinte accession X26-4 (*mexicana*, hereafter referred to as MEX) [5, 44] (Additional file 1: Fig. S1). The population was evaluated for a set of 20 agronomic traits, including three flowering time (days to heading, DH; days to anthesis, DA; days to silking, DS), seven plant architecture (plant height, PH; ear height, EH; leaf length, LL; leaf width, LW; leaf angle, LA; branch number, BN; tassel length, TL), six ear architecture (ear length, EL; kernel number per row, KNR; ear diameter, ED; kernel row number, KRN; cob diameter, CD; cob weight, CW), and four kernel weight and shape traits (hundred kernel weight, HKW; kernel length, KL; kernel width, KW; kernel thickness, KT) across up to 12 environments (Additional file 1: Fig. S2; Additional file 2: Table S1). A wide range of phenotypic variation was found across all traits, with most traits exhibiting continuous and approximately normal distributions (Fig. 1A–H; Additional file 2: Table S2). Analysis of variance (ANOVA) revealed significant effects of both genotype and environment across all traits. The average broad-sense heritability was 0.87, ranging

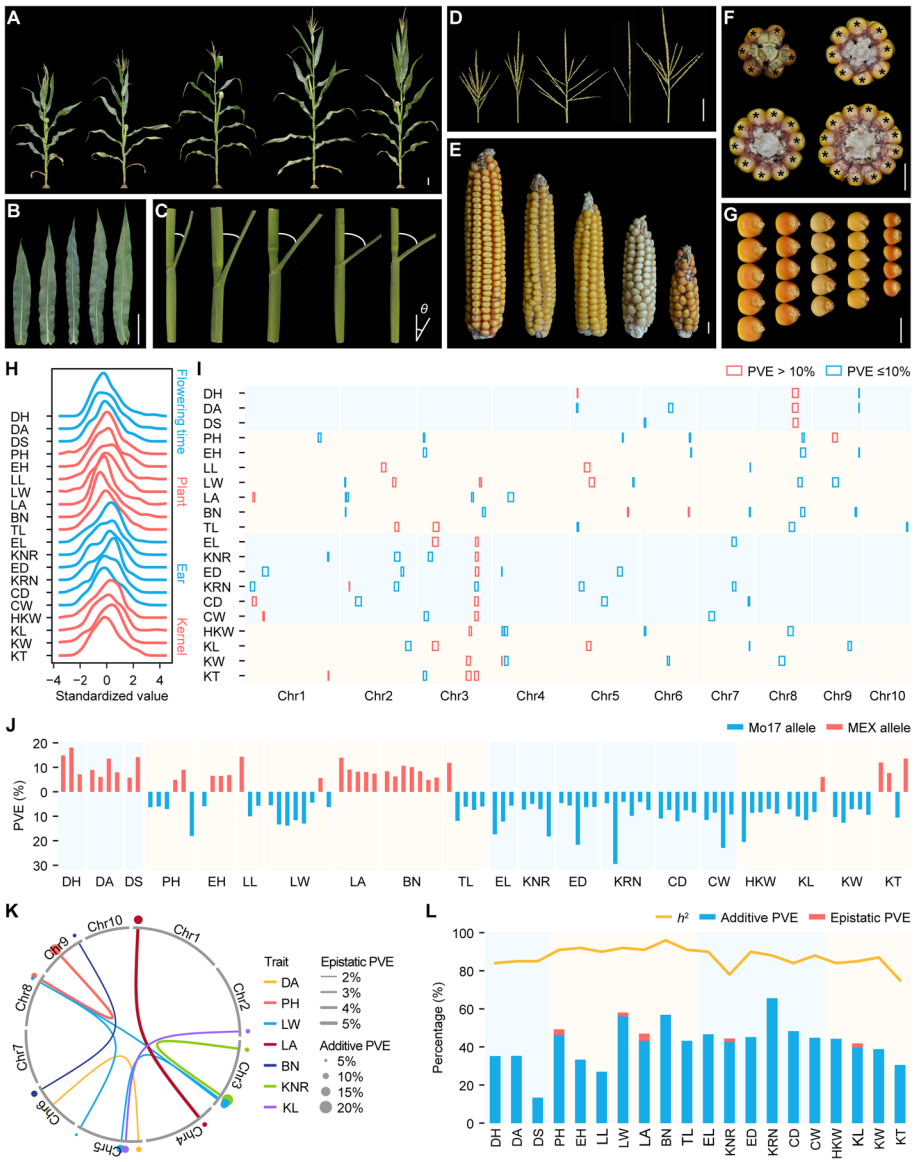

Fig. 1 Phenotypic variations and overview of QTLs identified for 20 agronomic traits in the TM population. **A–G** Plant (**A**), leaf (**B**), leaf angle (**C**), tassel (**D**), ear length (**E**), kernel row number (**F**), and kernel size (**G**) variations of the TM population. Scale bars, 10 cm in (**A**, **B**, and **D**), 1 cm in (**E**, **F**, and **G**). $\theta = 30°$ in (**C**). * indicates kernel rows in (**F**). **H** Phenotypic distributions of 20 agronomic traits. The *x*-axis indicates the standardized phenotypic value of each trait. **I** Distribution of additive QTLs on chromosomes. The confidence intervals of identified QTLs across the genome are outlined with red (PVE > 10%) and blue (PVE ≤ 10%) hollow rectangles. **J** Effect size (represented by PVE) and origin of the increasing alleles of the identified QTLs. Blue and red bars indicate the increasing alleles from Mo17 and MEX, respectively. **K** Pairwise epistatic interactions among additive QTLs. Dot size indicates the PVE of each additive QTL. QTLs that have epistatic interactions are connected by colored lines whose thickness indicates the PVE of the epistatic interactions. **L** Broad-sense heritability (h^2) (yellow line) and total PVE by additive QTLs (blue bars) and epistatic interactions (red bars) for each trait. In (**H–L**), DH, days to heading; DA, days to anthesis; DS, days to silking; PH, plant height; EH, ear height; LL, leaf length; LW, leaf width; LA, leaf angle; BN, branch number; TL, tassel length; EL, ear length; KNR, kernel number per row; ED, ear diameter; KRN, kernel row number; CD, cob diameter; CW, cob weight; HKW, hundred kernel weight; KL, kernel length; KW, kernel width; KT, kernel thickness

from 0.75 (KT) to 0.96 (BN) (Additional file 2: Table S2), indicating that most traits were highly heritable and predominantly controlled by genetic factors.

Using a high-density linkage map with 1,282 recombinant bins derived from 12,390 high-quality SNPs [44], we identified 93 QTLs associated with the 20 agronomic traits (Fig. 1I; Additional file 2: Table S3). The QTL number per trait ranged from 2 (DS) to 8 (LW), and the phenotypic variation explained (PVE) by individual QTL ranged from 4.1% (*qKRN2-2*) to 29.4% (*qKRN2-1*). Except for EH, all traits in the TM population were controlled by at least one large-effect QTL (PVE > 10%) (Fig. 1J), supporting the hypothesis that they underwent strong selection, exhibiting genetic architectures characterized by large-effect loci [45]. In particular, the enrichment of large-effect ear architecture trait QTLs highlights their central role in the morphological divergence between maize and teosinte. The origins of alleles increasing trait values varied among the 93 QTLs, of which 33 were derived from MEX (Fig. 1J). Notably, all alleles increasing flowering time originated from MEX, suggesting strong local adaptation to environmental cues such as photoperiod and elevation. In contrast, the alleles increasing ear architecture traits were exclusively contributed by Mo17, reflecting artificial selection for harvestability and improved yield. Among the seven plant architecture traits, QTLs associated with five traits carried increasing alleles contributed by both Mo17 and MEX. For LA and BN, the reducing alleles originated from Mo17, resulting in more upright leaves and fewer tassel branches, both of which are advantageous for modern breeding. These findings suggest that maize experienced strong directional selection for ideal plant architecture, such as more upright leaves to enhance light interception and reduced branching to optimize energy allocation to grain production.

In addition to additive QTLs, we identified 9 significant epistatic QTL pairs for seven traits, each explaining 2.1% (*Epi8*) to 4.3% (*Epi1*) of the phenotypic variation (Fig. 1K; Additional file 2: Table S4). The epistatic interactions collectively explained 0.0% (BN) to 3.8% (LA) of the phenotypic variations for each trait, considerably smaller than those explained by all additive QTLs, ranging from 13.3% (DS) to 65.6% (KRN). These findings suggest that additive effects play a more prominent role than epistatic interactions in shaping the genetic architecture of traits between maize and teosinte. Nevertheless, for all traits, the accumulative PVE from both additive QTLs and epistatic interactions remained lower than the corresponding broad-sense heritability (Fig. 1L). This discrepancy suggests that additional additive QTLs or epistatic interactions with minor effects remain undetected in the TM population, likely due to the limited population size.

QTL–trait network

Agronomic traits frequently exhibit substantial correlations due to synergistic selection during maize domestication and subsequent improvement [46]. Indeed, most trait pairs in a given category showed medium or high correlation, and clustered together in the TM population, such as three flowering time traits (DH, DA, and DS; $r = 0.80–0.93$), two ear length-related traits (EL and KNR; $r = 0.80$), three ear width-related traits (ED, KRN, and CD; $r = 0.54–0.73$), two plant height-related traits (PH and EH; $r = 0.68$), and three of the four kernel traits (HKW, KL, and KW; $r = 0.60–0.71$) (Fig. 2A). These results indicate that traits within the same category are

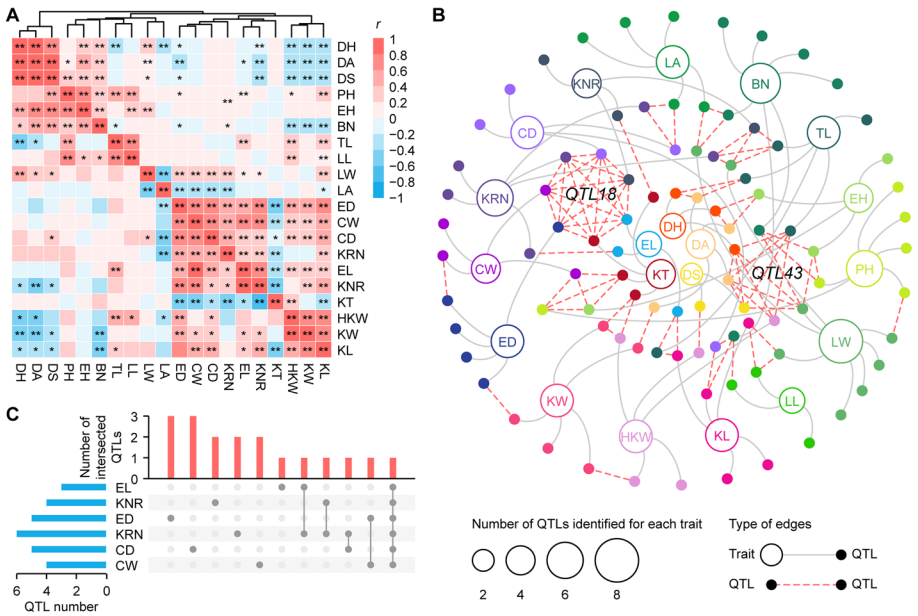

Fig. 2 Genetic network map of identified additive QTLs for 20 agronomic traits. **A** Phenotypic (upper triangle) and genetic (lower triangle) correlations among trait pairs. Hierarchical clustering based on phenotypic correlations is shown above the correlation matrix. The color scale indicates the magnitude of positive (red) and negative (blue) correlation coefficients. $^{**}P < 0.01$, $^{*}P < 0.05$. **B** The QTL–trait network is constructed based on the QTL results for 20 agronomic traits and the co-localization information of all 93 QTLs. Traits and QTLs are connected by the gray lines in the network when the QTLs are associated with the traits. Co-localized QTLs are linked by red dashed lines. **C** Upset plot showing pleiotropy of the QTLs for ear architecture traits. Horizontal bars indicate the total number of QTLs identified for each trait. Vertical bars corresponding to single or connected dark-gray dots represent the number of QTLs associated with a single trait or multiple traits, respectively

often controlled by shared genetic factors. In contrast, some plant architecture traits showed weak correlations (LL, LW, LA, and BN; $r = -0.42$–0.10), indicating an independent genetic basis (Fig. 2A). Consequently, 69.9% (65/93) of the identified QTLs influenced more than one trait in the TM population. After merging QTLs based on physical overlap of their 2-LOD confidence intervals, the 93 QTLs were consolidated into 48 distinct regions (Additional file 2: Table S3).

To capture the global genetic relationships among traits, we constructed a QTL–trait network by integrating the QTL–trait associations with QTL–QTL co-localizations (Fig. 2B). We found that each trait was connected to at least one QTL that was also linked to other traits, indicating a shared genetic basis among these traits. Traits with stronger correlations tended to share more common loci. For instance, 60.0% (3/5) of the loci detected for flowering time traits were shared by at least two flowering time traits, consistent with their high correlations (Fig. 2A and B; Additional file 2: Table S3). The largest-effect locus on chromosome 8 (*QTL43*) was shared for all flowering time traits (DH, DA, and DS), with PVE ranging from 13.5% (*qDA8-1*) to 18.0% (*qDH8-1*). Notably, this region harbored two well-characterized flowering time loci/genes, *ZCN8* [33] and *Vgt1*/*ZmRap2.7* [47]. *QTL43* also exhibited

pleiotropic effects on plant architecture and kernel traits. At this locus, the MEX allele increased all flowering time traits and four plant architecture traits (PH, EH, LW, and BN), while the Mo17 allele increased HKW and TL (Additional file 2: Table S3). However, the distinct peak positions of these QTLs suggest that different genetic factors at this locus might independently affect flowering time and plant architecture. Similarly, ear architecture traits also clustered tightly, showed strong correlation, and shared multiple common loci, reflecting strong connectivity within the QTL–trait network (Fig. 2A and B). A notable example is *QTL18*, which was associated with all six ear architecture traits (EL, KNR, ED, KRN, CD, and CW) as well as KT, with individual QTLs accounting for 9.8% (*qKRN3-1*)–22.9% (*qCW3-2*) of the phenotypic variation, highlighting its central role in ear architecture (Fig. 2B and C; Additional file 2: Table S3). At *QTL18*, the Mo17 allele generally increased all tested ear architecture traits, while the MEX allele increased KT (Additional file 2: Table S3). Notably, the peaks of the EL QTL (*qEL3-2*) and KNR QTL (*qKNR3-2*) differed from that of the KRN QTL (*qKRN3-1*), suggesting that the observed pleiotropy at *QTL18* may result from closely linked but different genetic variants. These findings offer valuable insights into the genetic basis of complex trait correlations and guide the use of QTLs in the improvement of multiple traits while balancing trade-offs.

Selection features of the detected QTLs for agronomic traits

To characterize the selection features of the detected QTLs, we analyzed the genomic regions affected by selection during maize evolution using resequencing data (~28 million SNPs) of 506 maize inbred lines, 253 maize landraces, and 151 teosinte (*parviglumis*

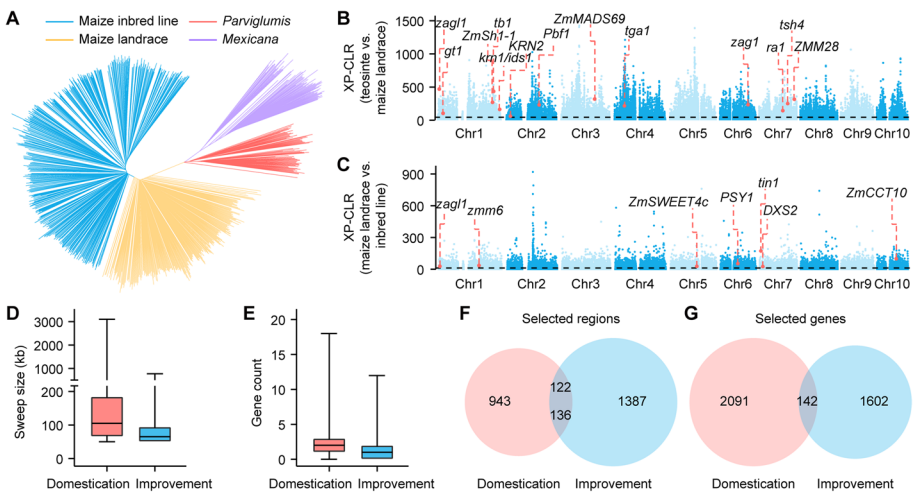

Fig. 3 Genome-wide analysis of selection during maize domestication and improvement. **A** The neighbor-joining phylogenetic tree of maize inbred lines, maize landraces, and teosinte (*parviglumis* and *mexicana*) accessions. **B** and **C** Genome-wide XP-CLR values between teosinte and maize landraces for selection during maize domestication (**B**), and between maize landraces and inbred lines for selection during maize improvement (**C**). The red dashed lines indicate the positions of known selected genes in maize. The black dashed lines indicate the selection thresholds that were set as the top 10% XP-CLR values. **D** and **E** Distributions of sweep size (**D**) and gene count (**E**) within sweeps in domestication and improvement scans. **F** and **G** Comparisons of selected regions (**F**) and genes (**G**) identified in domestication and improvement scans. In (**F**), 122 out of 1,065 domestication sweeps overlapped with improvement sweeps, while 136 out of 1,523 improvement sweeps overlapped with domestication sweeps

and *mexicana*) accessions [3, 38] (Fig. 3A; Additional file 2: Table S5). We estimated the cross-population composite likelihood ratios (XP-CLRs) [48], which were further cross-validated using permutation tests for nucleotide diversity (π) in the regions with the top 10% of the XP-CLR scores. By comparing teosinte and maize landraces, we identified 1,065 selected regions with domestication features, with a total of 175.0 Mb and a mean size of 164.3 kb, covering ~8.3% of the maize genome and 2,233 genes (Fig. 3B; Additional file 2: Tables S6 and 7). Among them, we identified 18 canonical domestication loci/genes, including *KRN2* [23], *krn1*/*ids1* [22], *ra1* [49], and *tsh4* [30], which regulate ear development (Fig. 3B; Additional file 2: Tables S6 and 7). Next, by comparing maize landraces and inbred lines, we identified 1,523 selected regions for improvement, covering ~5.9% (124.4 Mb) of the maize genome and 1,744 genes (Fig. 3C; Additional file 2: Tables S7 and 8). In this analysis, we identified 7 well-characterized improvement genes, such as *ZmCCT10* [31], *PSY1* [50], and *tin1* [51], which control flowering time, carotenoid content, and tiller number, respectively (Fig. 3C; Additional file 2: Tables S7 and 8). Consistent with previous studies [35], improvement-associated selective sweeps had a smaller average size and contained fewer genes than domestication-associated selective sweeps (Fig. 3D and E). We found that only 122 selected regions (11.5%) and 142 selected genes (6.4%) with domestication features also showed improvement features (Fig. 3F and G), which is lower than the proportion reported previously [35].

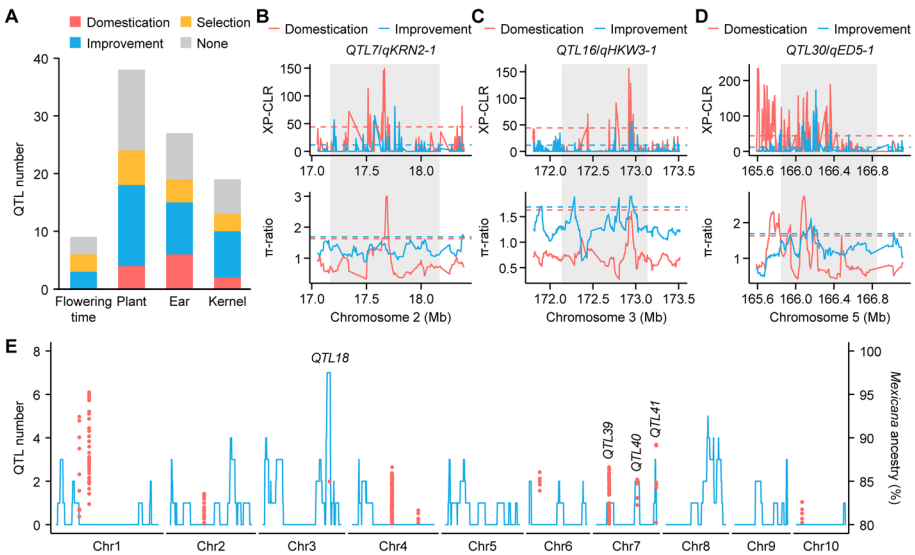

Fig. 4 Selection patterns of 93 QTLs identified for 20 agronomic traits. **A** Summary of QTLs with domestication or improvement features for the four trait categories. **B–D** XP-CLR (upper) and π-ratio (below) values across *qKRN2-1* with domestication feature (**B**), *qHKW3-1* with improvement feature (**C**), and *qED5-1* with both domestication and improvement features (**D**). For XP-CLR values (upper), the red and blue lines represent the comparison between teosinte and maize landraces (domestication) and between maize landraces and inbred lines (improvement), respectively. For π-ratio values (lower), the red and blue lines represent teosinte-to-maize landrace π-ratio (domestication) and maize landrace-to-inbred line π-ratio (improvement), respectively. The red and blue dashed lines represent the thresholds. **E** Summary of fixed *mexicana* alleles in modern maize that fell within QTLs identified in the TM population. Blue lines indicate QTL counts across the genome within 1-Mb windows. Red dots indicate genomic regions in which high-confidence *mexicana* alleles (>90% posterior probability) were present at high frequency (>80%) across a panel of 845 modern maize inbred lines [3]. The loci overlapping with fixed *mexicana* alleles are labeled

Next, we assessed selection features of the QTLs based on the co-localized selected regions. We found that 62 out of 93 QTLs underwent selection during maize evolution: 12 of the QTLs (e.g., *qKRN2-1*) were selected during domestication, 34 QTLs (e.g., *qHKW3-1*) were selected during improvement, and 16 QTLs (e.g., *qED5-1*) had both domestication and improvement features of selection (Fig. 4A–D; Additional file 2: Table S3). *QTL7*/*qKRN2-1*, the largest-effect QTL for KRN, is a typical domestication locus, and the underlying functional gene *KRN2* underwent selection only during maize domestication [23] (Fig. 4B). *QTL16/qHKW3-1*, the largest-effect QTL for HKW, had dramatically reduced nucleotide diversity only in the maize inbred lines compared with the landraces, and the XP-CLR values between teosinte and maize landraces and between maize landraces and inbred lines were high, indicating selection during maize improvement (Fig. 4C). *QTL30*/*qED5-1* for ED had high teosinte vs. maize landrace and maize landrace vs. inbred line XP-CLR values, and high teosinte-to-maize landrace and maize landrace-to-inbred line nucleotide diversity ratios, demonstrating both domestication and improvement features (Fig. 4D). These findings provide valuable insights into the evolutionary history of agronomic traits in maize and offer a framework for pinpointing candidate genes subject to selection.

To further elucidate the selection mechanisms underlying the tested traits, we compared nucleotide diversity ratios of QTLs with clear selection features across four trait categories. The teosinte-to-maize landrace nucleotide diversity ratio was the highest for the domestication QTLs controlling traits related to ear architecture, followed by kernel traits, plant architecture, and flowering time (Additional file 1: Fig. S3A). Conversely, the maize landrace-to-inbred line nucleotide diversity ratio was the highest for the improvement QTLs controlling flowering time, followed by plant architecture, kernel traits, and ear architecture (Additional file 1: Fig. S3B). These findings suggest that ear architecture experienced intensified selective pressures during the domestication phase, while flowering time underwent strong selection during the improvement phase.

In addition to natural and artificial selection, *mexicana* introgression is a major mechanism of maize domestication and improvement [5]. In the TM population, we identified 65 QTLs across 29 genomic regions that overlap with previously reported *mexicana* introgression signals (*mexicana* dosage > 0.9) in Mo17 [3] (Additional file 2: Table S3). Among these, 12 QTLs in four genomic regions (*QTL18*, *QTL39*, *QTL40*, and *QTL41*) coincide with high-frequency *mexicana* introgressed regions previously detected in modern maize [3] (Fig. 4E; Additional file 2: Table S3). These co-localized regions indicate that *mexicana* introgression contributed to trait variation associated with domestication and improvement/adaptation in modern maize. Additionally, among the 31 QTLs lacking detectable selection signals, 16 overlapped with selective sweeps previously identified between *parviglumis* and *mexicana* [38] (Additional file 2: Table S3), suggesting the presence of potentially beneficial alleles specific to *mexicana*.

EL3-2, a causal domestication gene underlying *QTL18*, has pleiotropic effects on multiple agronomic traits

Ear traits, as the determinant of maize grain yield, have undergone strong domestication and improvement. The aforementioned *QTL18* had pleiotropic effects on all six ear traits and KT, and was a selected locus with both domestication and improvement

features (Fig. 5A–C). To identify the gene(s) underlying this pleiotropic locus, we performed positional cloning for *qEL3-2* controlling EL using a heterogeneous inbred family, TM7, derived from the BC_2F_5 generation of the TM population (Additional file 1: Fig. S4A). Analysis of the TM7 progeny indicated that the MEX allele at the *qEL3-2* locus conferred shorter ears compared to the maize allele (Additional file 1: Fig. S4B). Using 17 molecular markers and 14,661 individuals, *qEL3-2* was delimited to a 64.1-kb genomic interval flanked by markers G68 and 3G76 through a recombinant-derived progeny testing strategy [23, 52] (Fig. 5D and E; Additional file 1: Fig. S5). This region harbored only one annotated gene, *Zm00001d043270*, herein named *Ear Length3-2* (*EL3-2*) (Fig. 5F). Comparisons of near-isogenic lines (NILs) derived from Rec-2 carrying Mo17 and MEX alleles at the *qEL3-2* locus (NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}) indicated that the Mo17 allele increased EL by 0.95–1.54 cm relative to the MEX allele across two environments (Fig. 5G and H; Additional file 2: Table S9). In addition, compared with NIL-*EL3-2*^{MEX}, NIL-*EL3-2*^{Mo17} significantly increased other ear traits, except KRN, indicating that *EL3-2* promotes ear enlargement. Conversely, *EL3-2* had bidirectional effects on kernel traits and moderate impacts on plant morphology (Fig. 5I and J; Additional file 2: Table S9). These findings suggest that *EL3-2* is the causal gene underlying the pleiotropic *QTL18.*

To validate the function of *EL3-2*, we overexpressed the coding sequences of NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} alleles in a maize inbred line and confirmed enhanced expression by real-time quantitative polymerase chain reaction (qPCR) (Additional file 1: Fig. S6). Compared to wild-type plants, both overexpression lines consistently reduced EL by 5.50–10.35 cm across three environments, with no significant

(See figure on next page.)

Fig. 5 *EL3-2* affected multiple agronomic traits and underwent selection during maize domestication. **A** LOD profiles of *QTL18* for multiple traits in the TM population. **B** and **C** XP-CLR (**B**) and π-ratio (**C**) values across *QTL18* with both domestication and improvement features. For XP-CLR values (**B**), the red and blue lines represent the comparison between teosinte and maize landrace (domestication) and between maize landrace and inbred line (improvement), respectively. For π-ratio values (**C**), the red and blue lines represent teosinte-to-maize landrace π-ratio (domestication) and maize landrace-to-inbred line π-ratio (improvement), respectively. The red and blue dashed lines represent the thresholds. **D** and **E** Fine mapping of *qEL3-2*. *qEL3-2* was preliminarily mapped between markers M1 and M10 on chromosome 3 and was fine mapped to a 187.7-kb genomic region flanked by markers G68 and M4 (**D**), and further refined to a 64.1-kb interval between markers G68 and 3G76 (**E**). The number of recombinants between adjacent markers is shown. **F** *EL3-2* gene structure. Black bars, exons; gray bars, UTRs. **G–I** EL quantification (**G**), ear (**H**) and plant (**I**) morphologies of NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}. **J** Comparison of 20 agronomic traits between NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} in two environments with two replications in each environment. Env1, Bayan Nur; Env2, Tieling. **K–M** EL quantification (**K**), ear (**L**) and plant (**M**) morphologies of wild type (WT) and two independent transgenic lines (*Ubi*::*EL3-2*^{Mo17} and *Ubi*::*EL3-2*^{MEX}). **N** Comparison of 20 agronomic traits between WT and two independent transgenic lines (*Ubi*::*EL3-2*^{Mo17} and *Ubi*::*EL3-2*^{MEX}) in three environments. Env1, Sanya; Env2, Beijing; Env3, Urumqi. **O** Molecular evolution of *EL3-2*. The nucleotide diversity (π) and Tajima's D with a 100-bp sliding window and a 25-bp step were calculated for the promoter and full-length gene, and coalescent simulations were performed to determine the significance (**$P < 0.01$). In (**H**, **I**, **L**, and **M**), scale bars are 1 cm for the ears and 10 cm for the plants. In (**E**, **G**, **J**, **K**, and **N**), data are presented as EL differences between NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} (relative to NIL-*EL3-2*^{Mo17}) in (**E**), as means ± SEM in (**G** and **K**), as percentage of trait changes between NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} (relative to NIL-*EL3-2*^{Mo17}) in (**J**), and as percentage of trait changes between WT and two independent transgenic lines (*Ubi*::*EL3-2*^{Mo17} and *Ubi*::*EL3-2*^{MEX}) (relative to WT) in (**N**); colors in (**J** and **N**) represent the direction and magnitude of percentage changes for each trait between genotypes, with red and blue indicating positive and negative percentage changes, respectively, relative to the reference genotype; **$P < 0.01$, *$P < 0.05$, NS, not significant (two-tailed Student's *t*-test)

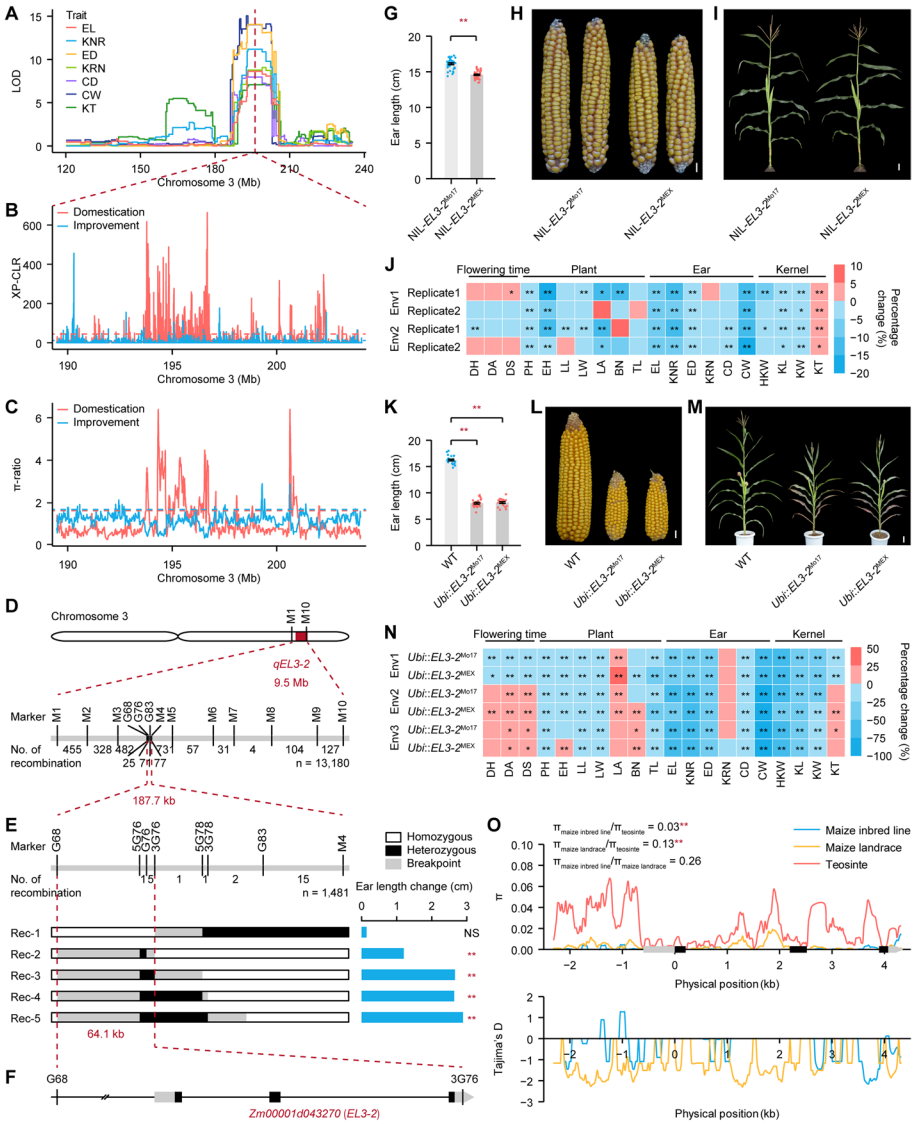

Fig. 5 (See legend on previous page.)

difference between *Ubi*::*EL3-2*^{Mo17} and *Ubi*::*EL3-2*^{MEX} transgenic plants (Fig. 5K and L; Additional file 2: Table S10), suggesting that *EL3-2* regulates ear development through variation in expression levels rather than through differences in coding sequences. Notably, overexpression of *EL3-2* also reduced plant and kernel sizes (Fig. 5M and N; Additional file 2: Table S10), indicating its negative pleiotropic effects on maize development.

To determine whether *EL3-2* underwent selection during maize evolution, we calculated nucleotide diversity and Tajima's D-statistic across its promoter and coding regions using 60 maize inbred lines, 19 maize landraces, and 29 teosinte accessions (Additional file 2: Table S5). We observed a dramatic reduction in nucleotide diversity in maize landraces and inbred lines relative to teosinte ($\pi_{\text{maize landrace/teosinte}} = 0.13$, $\pi_{\text{maize inbred line/teosinte}} = 0.03$), and the reduction in maize inbred lines was greater than that in maize landraces (Fig. 5O). Additionally, we observed a negative Tajima's D-statistic

in maize inbred lines and landraces across the promoter and coding regions of *EL3-2* (Fig. 5O). Coalescent simulation further supported the conclusion that the severe loss of genetic diversity in both the promoter and the coding regions could not be attributed solely to the maize domestication bottleneck (Fig. 5O). These findings suggest that the *EL3-2* locus underwent selection during maize domestication.

EL3-2 encodes a ULT transcription factor that regulates maize ear development

EL3-2 was predicted to encode a ULTRAPETALA (ULT)-related transcriptional regulator localized to the nucleus and cytoplasm (Additional file 1: Figs. S7–9). ULT homologs in *Arabidopsis* are trithorax group (trxG) factors that regulate chromatin structure and gene expression, and control various developmental processes and stress responses, particularly in determining shoot and floral meristem fate [53–58]. *EL3-2* encodes a protein of 235 amino acids that contains a SAND domain responsible for DNA binding and a B-box motif associated with protein–protein and protein–RNA interactions [53] (Additional file 1: Fig. S9). Sequence comparisons identified two amino acid changes between the Mo17 and MEX alleles (AA11: GTG-to-GCG resulting in V-to-A change, and AA114: TAC-to-AAC resulting in Y-to-N change) in the *EL3-2* coding region (Additional file 1: Fig. S9; Additional file 1: Fig. S10A). However, the similar phenotypes of *EL3-2*^{Mo17} and *EL3-2*^{MEX} overexpression lines (Fig. 5J–M) suggest that these amino acid differences do not alter protein function. Instead, *EL3-2* may regulate ear development through variation in expression levels rather than changes in protein-coding sequence. Indeed, a set of polymorphic sites, including complex transposable element variations, SNPs, and small InDels, was identified in the promoter region of *EL3-2*, along with two SNPs in the putative 5'UTR (Additional file 1: Fig. S10B and C). qPCR assays revealed that *EL3-2* was broadly expressed across tissues, with particularly elevated transcript levels in immature ears and tassels. Notably, the expression in inflorescence meristem (IM) and spikelet pair meristem (SPM) tissues was significantly lower in NIL-*EL3-2*^{Mo17} relative to NIL-*EL3-2*^{MEX} (Fig. 6A). These results suggest that *EL3-2* controls EL primarily through expression changes driven by promoter variation, while the variation in 5'UTR may contribute modestly. We next assessed *EL3-2* expression patterns in developing ears by RNA in situ hybridization. *EL3-2* was expressed strongly in initiating primordia as well as in spikelet meristems (SMs) and floral meristems (FMs) (Fig. 6B–F), suggesting a role in axillary meristem initiation and growth. We next examined whether *EL3-2* affected EL by altering IM size or SPM initiation, by measuring immature ears at multiple developmental stages (Fig. 6G). There was no significant difference in IM size between NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} at the V7 stage, whereas NIL-*EL3-2*^{Mo17} IM diameter was slightly wider compared to that of NIL-*EL3-2*^{MEX} at V8 (Fig. 6H and I). As ear development proceeded, the IM gradually disappeared and the number of SPMs stabilized, with NIL-*EL3-2*^{Mo17} exhibiting a significantly greater number of SPMs/SMs per row than NIL-*EL3-2*^{MEX} (Fig. 6J). These results suggest *EL3-2* expression did not substantially alter IM size, but rather affected SPM initiation, thereby influencing the number of florets per row and ultimately determining EL.

To investigate potential downstream targets of *EL3-2*, we conducted transcriptome profiling of immature ears from NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} at two key developmental stages (V7 and V14), and identified thousands of differentially expressed genes

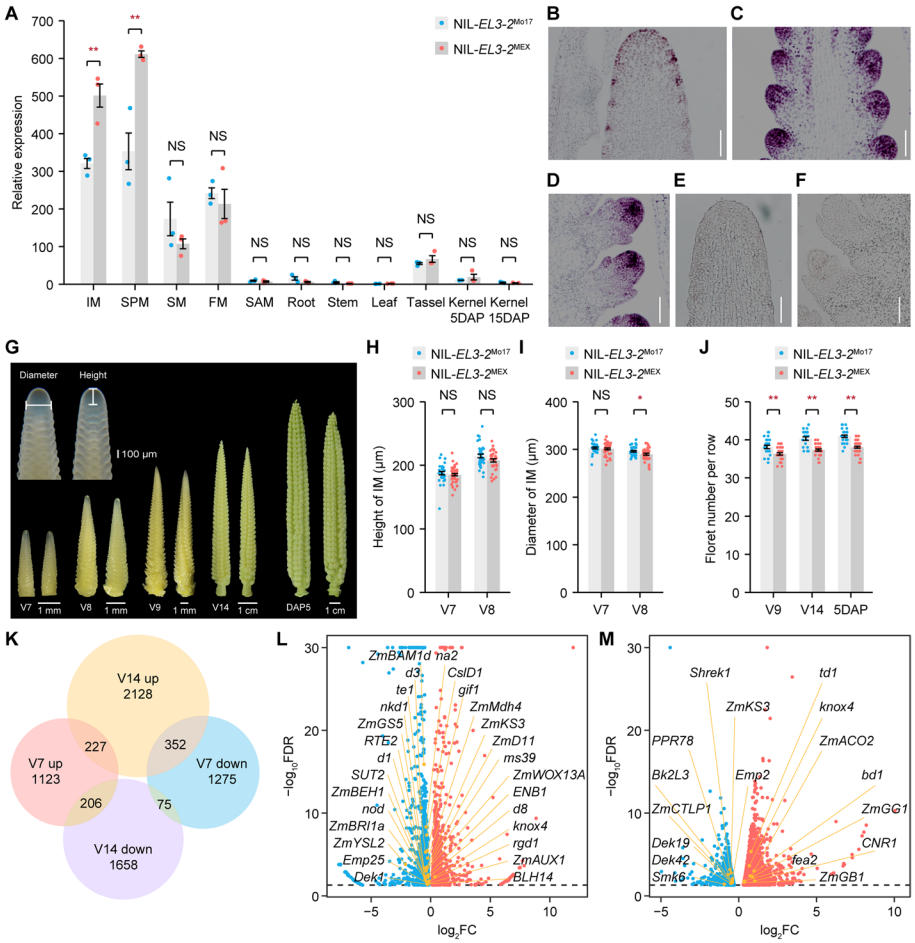

Fig. 6 Molecular characterization of *EL3-2*. **A** Expression pattern of *EL3-2* in various tissues of NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}. IM, inflorescence meristem; SPM, spikelet pair meristem; SM, spikelet meristem; FM, floral meristem; SAM, shoot apical meristem; DAP, days after pollination. **B–F** Representative images for in situ hybridization of *EL3-2* mRNA in the ear meristem. Hybridization with an antisense probe revealed strong signals in the initiating primordia (**B**), as well as the entire SM (**C**) and FM (**D**). No signal was detected in the negative controls hybridized with a sense probe (**E** and **F**). Scale bars, 100 μm. **G** Micrographs of immature ears of NIL-*EL3-2*^{Mo17} (left) and NIL-*EL3-2*^{MEX} (right) at different developmental stages. V7, V8, V9, and V14 represent the number of fully expanded leaves. 5DAP represents five days after pollination. The inset graph on the upper left represents the diameter and the height of IM at the V7 stage. **H–J** Comparison of IM height (**H**), IM diameter (**I**), and floret number per row (**J**) between NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}. **K** Venn diagram showing the number of up- and down-regulated genes of NIL-*EL3-2*^{MEX} relative to NIL-*EL3-2*^{Mo17} in immature ears at V7 and V14 stages. **L** and **M** Volcano plots of DEGs at V7 (**L**) and V14 (**M**) stages. The *x*-axis indicates log2-transformed fold changes in gene expression of NIL-*EL3-2*^{MEX} relative to NIL-*EL3-2*^{Mo17}, with positive and negative values indicating up- and down-regulated genes, respectively. The black dashed lines indicate the FDR thresholds. Some representative genes are labeled. In (**A**, **H**, **I**, and **J**), data are shown as means ± SEM; $^{**}P < 0.01$, $^{*}P < 0.05$, NS, not significant (two-tailed Student's *t*-test)

(DEGs) at both stages (Fig. 6K–M; Additional file 2: Table S11). Gene ontology (GO) enrichment analysis revealed stage-specific functions. At the early V7 stage when the ears had IMs and initiated spikelet primordia, enrichment was mainly in shoot and reproductive development, cell division, as well as multiple metabolic and biosynthetic processes, particularly among genes down-regulated in NIL-*EL3-2*^{MEX}. At the later developmental stage (V14) when the IM was consumed and ears had established kernel rows, up-regulated genes were enriched for categories related to regulatory

processes, while down-regulated genes were primarily involved in metabolic pathways (Additional file 1: Fig. S11). These patterns suggest that *EL3-2* may influence cell proliferation and organ differentiation during early ear development and contribute to continued ear growth and maturation at later stages. Notably, 8.9% (291/3,258) of DEGs at V7 and 8.7% (406/4,646) at V14 were located within selected regions identified in this study (Additional file 2: Table S11), highlighting the regulatory role of *EL3-2* in maize evolution.

Consistent with this, several key agronomic trait genes were significantly differentially regulated (Additional file 1: Fig. S12A). At V7, flowering time genes tended to be up-regulated in NIL-*EL3-2*^{MEX}, whereas the majority of genes with positive effects on plant, ear, and kernel development (*BLH14*, *CslD1*, *d1*, *d3*, *d8*, *Dek1*, *Emp25*, *ENB1*, *gif1*, *ms39*, *na2*, *nkd1*, *nod*, *rgd1*, *RTE2*, *SUT2*, *te1*, *ZmAUX1*, *ZmBAM1d*, *ZmBEH1*, *ZmBRI1a*, *ZmD11*, *ZmGS5*, *ZmKS3*, *ZmMdh4*, *ZmWOX13A*, and *ZmYSL2*) were down-regulated, potentially accounting for reduction in overall plant stature (Fig. 6L). At V14, several negative regulators of ear development (*bd1*, *fea2*, *td1*, *ZmACO2*, *ZmGB1*, and *ZmGG1*) were up-regulated, while genes positively regulating kernel development (*Dek19*, *Dek42*, *Emp2*, *PPR78*, *Shrek1*, *Smk6*, and *ZmCTLP1*) were down-regulated in NIL-*EL3-2*^{MEX} (Fig. 6M). In addition, multiple transcription factors and hormone-related genes were significantly differentially expressed at both stages (Additional file 1: Fig. S12B and C). Interestingly, we found that *EL3-2* affected the expression of several classical maize domestication or adaptation loci/genes, such as *tb1*, *tga1*, *tru1*, *krn1*/*ids1*, *KRN2*, *YIGE1*, *ZmACO2*, *d8*, *ZMM3*, and *ZmMADS69* (Additional file 2: Table S11). These transcriptomic findings suggest that *EL3-2* may function as a central regulator of maize development during domestication by modulating complex gene-expression networks across multiple developmental stages.

Discussion

Modern cultivated maize was domesticated from its wild ancestor teosinte and has undergone dramatic morphological changes through long-term natural and artificial selection [6]. Understanding the genetic architecture of these divergent traits provides crucial insights into the molecular mechanisms underlying maize domestication and improvement. In this study, we identified 93 additive QTLs and 9 pairs of epistatic interactions controlling 20 agronomic traits using a maize–*mexicana* introgression population. Consistent with previous studies in maize–teosinte populations [9–16], our analysis reveals a polygenic architecture characterized by both large-effect and small-effect loci for most traits in the TM population. This pattern contrasts with that of key domestication traits such as branch formation and glume development, which are typically controlled by large-effect loci/genes [17–20]. Out of the 8 detected large-effect QTLs (PVE > 15%), 5 were associated with ear architecture traits. This aligns with the expectation that maize domestication and improvement predominantly targeted ear restructuring, with higher priority than other morphological features [59].

The QTL–trait network and genomic selection signals illuminate the correlations and co-evolution of agronomic traits in maize, where phenotypic correlations largely mirror underlying genetic relationships, and co-evolving traits tend to share common QTLs/genes, thereby forming interconnected modules within the network. Remarkably, 41.7%

(20/48) of the loci exhibited extensive pleiotropy, and 68.8% (33/48) displayed clear selection signatures (Additional file 2: Table S3), revealing the tightly linked nature of morphological evolution during maize evolution. A high proportion of yield-related QTLs (ear and kernel architecture traits) were under selection and exhibited stronger signals during domestication, reflecting the substantial divergence in ear and kernel morphology from teosinte to modern maize, as well as intense selection for yield enhancement. The tight clustering of ear architecture traits in the correlation matrix and the dense interconnections among these traits in the QTL–trait network underscore the simultaneous selection for optimizing multiple yield components. This coordinated evolution is exemplified by *QTL18*, which affected all six ear traits and exhibited selection features during both domestication and improvement, illustrating how a pleiotropic locus can orchestrate complex morphological changes. However, the distinct peak positions and different selection features among the associated traits at *QTL18* raised the possibility that the observed pleiotropy might arise from multiple tightly linked loci rather than a single causal gene with pleiotropic effects. High-resolution fine mapping using EL as the target trait identified *EL3-2* as the causal gene underlying *QTL18*. Near-isogenic line analyses and transgenic validation confirmed *EL3-2* as a major determinant of EL and correlated ear enlargement traits, with no detectable effect on KRN. These results suggest that *QTL18* represents a complex locus harboring multiple tightly linked genes, including a pleiotropic gene affecting multiple traits, while the KRN-associated signal appears to be genetically independent. The dissection of *QTL18* demonstrates that high-resolution fine mapping, coupled with targeted gene engineering, provides an effective strategy to distinguish single-gene pleiotropy from tight linkage within complex QTL regions. Such pleiotropic effects likely stem from the evolutionary integration of developmental regulatory networks, enabling synchronized responses to selective pressures [60, 61]. Flowering time and plant architecture traits experienced relatively weaker selection, as a smaller proportion of their QTLs were under selection. Pleiotropic loci such as *QTL43* further underscore the genetic interconnectedness of these traits. Distinct peak positions and varying selection features of individual QTLs at *QTL43* similarly suggest tightly linked loci with distinct functions underlying correlated traits. Overall, the QTL–trait network provides a valuable framework for identifying candidate loci with pleiotropic effects and assessing their potential utility in breeding programs. It also offers practical implications for the simultaneous improvement of multiple agronomic traits. Integration of QTL–trait network information during the utilization of teosinte genetic resources can facilitate the effective introduction of favorable alleles while minimizing undesirable genetic variation.

Recent population genetics and genomics studies have illuminated the dual contributions of *parviglumis* and *mexicana* to modern maize [3]. Moreover, teosinte continues to serve as a valuable reservoir of beneficial alleles for modern maize improvement, even after domestication [62, 63]. Taking reduced flowering period and plant morphology, as well as increased ear and kernel size as favorable traits [64], the beneficial alleles of 22 QTLs were derived from MEX. Among the 16 QTLs without selection signals detected in the current study but overlapped with previously identified *parviglumis–mexicana* selective sweeps [38], five (*qPH3-1*, *qKT3-3*, *qLW3-1*, *qTL5-1*, and *qKL9-1*) carried favorable *mexicana* alleles. Considering the QTLs described here and previously

identified QTLs for the 20 agronomic traits [9–16], four of these five QTLs (*qPH3-1*, *qKT3-3*, *qTL5-1*, and *qKL9-1*) were unique to the maize–*mexicana* populations (Additional file 2: Table S3), suggesting potential *mexicana*-specific contributions. These results demonstrate that *mexicana* represents a crucial genetic resource for enhancing yield potential and broadening the genetic diversity available to modern maize breeding programs, beyond facilitating maize adaptation to new environments [5].

Genome-wide selection scans revealed uncharacterized genes within the QTL intervals that exhibited stronger selection signals than some classical maize domestication or adaptation genes, suggesting their potential but previously unrecognized contributions to maize evolution. Among these, the isolation and functional characterization of *EL3-2*, encoding a ULT transcription factor, provide new insights into the molecular basis of pleiotropy during maize evolution. The pleiotropic effects of *EL3-2* on multiple agronomic traits and strong selection signatures during domestication highlight its crucial role in ear morphological transformations. Interestingly, one of the rice homologs, *OsULT2* (*Os05g0502700*), also shows evidence of selection in its upstream regulatory regions, as detected by XP-CLR and nucleotide diversity ratio when comparing wild rice (*Oryza rufipogon*) with cultivar rice (*Oryza sativa* subsp. *japonica*) [23]. This finding indicates that *EL3-2* and its rice ortholog, *OsULT2*, may have experienced convergent selection during maize and rice domestication. Whether this regulatory divergence similarly contributes to the regulation of inflorescence development in rice warrants further investigation.

ULT transcription factors (*ULT1* and *ULT2*) in *Arabidopsis* regulate meristem activity and organ size via epigenetic mechanisms [53–58]. Consistent with these findings, we found that *EL3-2* modulates ear development by influencing spikelet pair meristem initiation. The reduced expression of *EL3-2* in maize relative to teosinte, driven by promoter variation, mirrors patterns observed for most domestication genes, in which transcriptional regulation plays a crucial role in phenotypic divergence [65]. The relationship between endogenous expression differences (Mo17 vs. MEX alleles) and the magnitude of the observed phenotypic effects in the NILs and overexpression lines suggests a general correspondence between increased *EL3-2* expression and enhanced suppression of ear elongation, providing further support for regulatory variation as the primary driver of phenotypic effects. Furthermore, transcriptomic analysis revealed that *EL3-2* appeared to act as a pleiotropic domestication regulator that modulates maize–teosinte divergence. This regulatory mechanism parallels that of *tsh4*, a recently characterized developmental regulator found to control meristem fate and influence multiple domestication loci [30]. Core regulatory nodes such as *EL3-2* and *tsh4* can drive complex morphological evolution by integrating diverse developmental processes. Notably, these conclusions are derived from transcriptional associations and therefore do not establish direct regulatory interactions. Future studies employing approaches such as chromatin immunoprecipitation sequencing (ChIP-seq) or DNA affinity purification sequencing (DAP-seq) will be essential to identify direct target genes and to elucidate the molecular mechanisms underlying the regulatory role of *EL3-2*.

Moreover, *EL3-2* provides a new target for maize molecular breeding. Its influence on multiple ear traits suggests strong potential for the coordinated improvement of yield components. Nevertheless, given its pleiotropic effects on other agronomic

traits, precise modulation will be essential to balance yield gains with overall plant performance. Recent advances in precision genome engineering provide unprecedented opportunities for the targeted utilization of such genes. These technologies allow breeders to fine-tune complex traits by optimizing the trade-off between beneficial improvements and potentially detrimental effects. For instance, CRISPR/Cas-mediated editing of regulatory elements enables subtle adjustment of gene expression [66, 67], and promoter engineering allows for tissue-specific or environmentally responsive gene activation [68].

Conclusions

Our findings offer a comprehensive framework for understanding the genetic basis of maize–teosinte morphological divergence. The combination of unique *mexicana* alleles and their pleiotropic network presents new opportunities for enhancing yield potential and genetic diversity in modern maize. Notably, *EL3-2*, as a regulatory hub gene and domestication locus with broad pleiotropy, emerges as a valuable target for future molecular breeding applications.

Methods

Plant materials and genotyping

A maize–teosinte introgression population (BC_2F_5; TM population) comprising 191 lines was developed by crossing an elite maize inbred line Mo17 with a teosinte accession X26-4 (accession number: PI 566686; *Zea mays* ssp. *mexicana*) (Additional file 1: Fig. S1), and was used to dissect the genetic architecture of agronomic traits [5, 44]. Together with the parental lines, the TM population was previously genotyped using the Illumina MaizeSNP50 array, and a high-density bin map with a total length of 1,748 cM was constructed using 1,282 recombinant bins derived from 12,390 high-quality SNPs [44]. All SNP positions based on the B73 reference genome version 2 were converted to their corresponding coordinates in version 4 using CrossMap [69].

A total of 511 maize inbred lines, 256 maize landraces, 70 *parviglumis*, and 84 *mexicana* (Additional file 2: Table S5) accessions were used for selection analyses. The SNPs from 506 maize inbred lines, 253 maize landraces, and 151 teosinte accessions were assessed from previous studies [3, 38]. After merging all SNPs, we filtered out the SNPs with minor allele frequency (MAF) < 0.01 and missing rate > 0.5, resulting in approximately 28 million high-quality SNPs for phylogenetic tree construction and genome-wide selection analysis. In addition, a set of 60 maize inbred lines, 19 maize landraces, 17 *parviglumis*, and 12 *mexicana* accessions was used for the molecular evolutionary analysis of *EL3-2*.

Field trials and phenotyping

The TM population was planted in a randomized complete block design with one replicate across 13 environments: Sanya, Hainan (18.38°N, 109.18°E) in 2011 and 2012; Nanning, Guangxi (22.61°N, 108.24°E) in 2011; Honghe Hani and Yi autonomous prefecture, Yunnan (23.71°N, 102.50°E) in 2011 and 2012; Chongqing (29.45°N, 106.37°E) in 2011 and 2012; Chengdu, Sichuan (30.44°N, 103.85°E) in 2011; Wuhan, Hubei

(30.47°N, 114.36°E) in 2011 and 2012; Hebi, Henan (35.67°N, 114.29°E) in 2011 and 2012; Beijing (40.14°N, 116.19°E) in 2012 (Additional file 1: Fig. S2). Each line was grown in a single-row plot (3.0-m rows, 0.25 m between plants, 0.7 m between rows). A total of 20 traits were measured, including three traits related to flowering time, seven traits related to plant architecture, six traits related to ear architecture, and four kernel traits (Additional file 2: Table S1).

Analysis of phenotypic data

All statistical analyses of phenotypic data were carried out using R (version 4.2.2; www.R-project.org). To estimate the genotype and environmental effects, ANOVA was performed for each trait using the *aov* function in R following the model: $y=\mu+\alpha_g+\beta_e+\varepsilon$, where α_g is the genotype effect of the g^{th} line, β_e is the effect of the e^{th} environment, and ε is the random error. The broad-sense heritability for each trait was calculated using the estimated variances as: $h^2=\sigma_g{}^2/(\sigma_g{}^2+\sigma\varepsilon^2/e)$, where $\sigma_g{}^2$ is the genetic variance, $\sigma\varepsilon^2$ is the residual variance, and e is the environment number. To reduce the environmental noise, the best linear unbiased predictor (BLUP) values for each line were calculated with the *lmer* function from the *lme4* package [70] in R using a linear mixed model considering both genotype and environment as random effects. The model is: $y_{ij}=\mu+e_i+g_j+\varepsilon_{ij}$, where y_{ij} represents the phenotypic value of line j in environment i, μ is the grand mean, e_i is the environmental effect, g_j is the genetic effect, and ε_{ij} denotes the random error. The BLUP values for each line were used to calculate summary statistics and Pearson correlation coefficients, and to perform QTL mapping. The phenotypic correlation analysis was performed using the *rcorr* function from the *Hmisc* package [71] in R. The genetic correlation analysis was performed using GCTA [72]. The correlation matrix with hierarchical clustering was visualized using the *pheatmap* function from the *pheatmap* package [73] in R.

QTL mapping

Using a high-density bin map previously constructed [44], we carried out QTL mapping for 20 agronomic traits using composite interval mapping (CIM) [74] implemented in Windows QTL Cartographer 2.5 [75]. The putative QTLs throughout the genome were scanned with 1-cM intervals. Model 6 of CIM was employed to control background effects through forward–backward stepwise regression, incorporating up to five control markers while excluding those within a 10-cM window around the test position. The logarithm of odds (LOD) scores for putative QTLs were determined via 1,000 permutations ($\alpha=0.05$), ranging from 3.2 (TL and CW) to 4.0 (DH, LA, and BN). For simplicity, an average cutoff of 3.4 was adopted as the global LOD threshold. The confidence interval of detected QTL was estimated with the 2-LOD supporting interval method [76]. For the QTLs with physical intervals exceeding 20 Mb, the peak-to-flank regions were restricted to 10 Mb to avoid merging distant QTLs into a single locus due to excessively large intervals. In the QTL co-localization analysis, QTLs detected for different traits were merged into a single locus if their physical positions overlapped.

Epistatic interactions

To test the epistatic effects of identified additive QTLs, we extracted the peak bin marker for each QTL and assigned all the heterozygous genotypes as missing values to ensure that only homozygous allelic interactions were tested. Then, a two-way ANOVA was performed to estimate the epistatic effects using the *aov* function in R. A threshold of $P < 0.05$ was used to determine putative epistatic interactions between pairwise QTLs for each trait.

PVE by all additive QTLs and epistatic interactions

The total PVE by additive QTLs and epistatic interactions was calculated for each trait using the *lm* function in R. Genotypic data for all significant additive QTLs and epistatic interactions were extracted and used as predictor variables. The variance of each predictor variable was calculated using the *aov* function in R. The total PVE for additive QTLs was calculated as the ratio of the sum of their variances to the total variance, while the total PVE for epistatic interactions was determined as the ratio of the sum of variances for all epistatic interactions to the total variance [77].

Construction of a QTL–trait network

The QTL–trait network was constructed based on the QTL mapping results using Cytoscape (version 3.10.3) [78]. In the network, traits and their corresponding QTLs were taken as nodes, while the links between traits and QTLs, as well as between overlapping QTLs based on physical positions, were taken as edges. The size of each node indicates the number of QTLs for each trait.

Phylogenetic tree construction

Around 28 million filtered SNPs from 506 maize inbred lines, 253 maize landraces, and 151 teosinte accessions were used to calculate genetic distance employing PLINK (version 1.90) [79]. The generated distance matrix was used to construct a phylogenetic tree using the neighbor-joining method implemented in MEGA (version 12.0.11) [80]. Finally, iTOL (version 6) [81] was used to visualize the tree.

Identification of regions and genes that have undergone selection

We identified the selected regions using ~28 million SNPs via an XP-CLR method [48], which is based on modeling the likelihood of multi-locus allele frequency differentiation between two populations, followed by cross validation on the basis of permutation tests for the ratio of nucleotide diversity between populations in the corresponding regions. The nucleotide diversity of cultivated and wild accessions was calculated using VCFtools (version 0.1.13) [82]. Genomic evidence of selection during domestication and improvement was evaluated through two contrasts: teosinte vs. maize landrace for domestication and maize landrace vs. inbred line for improvement. For XP-CLR analysis, the sliding window size was 50 kb, the step size was 5 kb, and the number of SNPs per window was fixed at 200. The top 10% of XP-CLR values were cross-validated by the ratio of nucleotide polymorphisms. The regions whose nucleotide polymorphism

ratio was higher than the top 10% value from 1,000 randomly selected regions with the same size window were selected. Finally, overlapping intervals were merged and defined as the final selective sweeps. The selected regions by comparing teosinte and maize landrace were defined as domestication features, and the selected regions by comparing maize landrace and inbred line were defined as improvement features. To identify genes that have undergone selection, we extracted the physical regions of all annotated genes in maize from the annotation files of B73 version 4 from the Gramene database (https://www.gramene.org). The genes that are located within the selected regions were regarded as having undergone selection. For the known selected genes (Additional file 2: Table S7), we also defined them as selected genes, although the selection signals were located in the intergenic regions upstream or downstream of these genes.

Selection features of the detected QTLs

We defined the selection features of detected QTLs by performing an overlap analysis between the selective sweeps and the QTL intervals. Given that the large QTL intervals hindered precise definition, we used the peak bin as the representative interval for each QTL. When detected QTLs overlapped with the domestication and improvement sweeps, we defined them as 'domestication' or 'improvement' QTLs, respectively. When the QTLs spanned both types of sweeps, we defined them as 'selection' QTLs.

Fine mapping of *qEL3-2*

A recombinant-derived progeny testing strategy [23, 52] was employed to fine map *qEL3-2*. We selected a heterogeneous inbred family, TM7, that was heterozygous at the *qEL3-2* locus (Additional file 1: Fig. S4A). Progenies from TM7 were used to validate the QTL effect, and heterozygous individuals were further self-pollinated to generate large F_2 populations for fine mapping. Across multiple generations, F_2 populations derived from TM7, comprising 13,180 individuals, were screened for recombinants with different breakpoints at the *qEL3-2* locus using 13 molecular markers (M1, M2, M3, G68, G76, G83, M4, M5, M6, M7, M8, M9, and M10; Additional file 2: Table S12). F_3 families derived from F_2 recombinants were planted in field trials and genotyped with these 13 molecular markers to identify homozygous recombinants (HRs) and homozygous nonrecombinants (HNRs). For each F_3 family, the significance of EL differences between HR and HNR plants was determined using a Student's *t*-test. If HR and HNR plants differed significantly in terms of EL, the parental F_2 recombinant was assumed to be heterozygous for the target QTL; otherwise, the parental F_2 recombinant was assumed to be homozygous. Integrating QTL information across all recombinants delimited *qEL3-2* to a 187.7-kb region flanked by markers G68 and M4 (Fig. 5D; Additional file 1: Fig. S5). To further refine the interval, a F_2 population containing 1,481 individuals was screened using six molecular markers (5G76, G76, 3G76, 5G78, 3G78, and G83; Additional file 2: Table S12) within the 187.7-kb genomic region. Using the same testing strategy, *qEL3-2* was further delimited to a 64.1-kb genomic region flanked by markers G68 and 3G76 (Fig. 5E), with G68 located in the intergenic region between genes *Zm00001d043269* and *Zm00001d043270*, and 3G76 located in the 3'UTR of *Zm00001d043270*. The fine-mapping populations were grown in the experimental fields of China Agricultural University in Beijing (40.14°N, 116.19°E) during the summer and

in Sanya, Hainan (18.38°N, 109.18°E) during the winter seasons in 2013–2016. NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}, used for the phenotypic assessment of 20 important agronomic traits (Additional file 2: Table S1), were grown in two environments with two replicates: Bayan Nur, Inner Mongolia (40.75°N, 107.43°E) and Tieling, Liaoning province (42.15°N, 123.62°E) in 2019.

Overexpression of *EL3-2* in maize

The full-length coding sequences of *EL3-2* were amplified from NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} cDNA, and cloned into the pBCXUN vector that had been cleaved with *Xcm*I under the constitutive *ubiquitin* promoter to produce constructs *Ubi*::*EL3-2*^{Mo17} and *Ubi*::*EL3-2*^{MEX}. The constructs were introduced into *Agrobacterium tumefaciens* strain EHA105 via heat shock, and then transformed into the immature embryos of maize inbred line LH244 via an *Agrobacterium*-mediated transformation system. The transgenic lines were generated at Center for Crop Functional Genomics and Molecular Breeding of China Agricultural University. The genotypes were confirmed by PCR, and *EL3-2* expression was quantified by qPCR. The primer pairs used for vector construction (G76oe1m and G76oe1t), genotyping (P1/P2), and qPCR analysis (G76c12 and ACT1c) are listed in Additional file 2: Table S12. The overexpression lines and corresponding wild-type plants were evaluated for 20 agronomic traits (Additional file 2: Table S1) across three environments: Sanya, Hainan (18.38°N, 109.18°E) in 2018, Beijing (40.14°N, 116.19°E) in 2019, and Urumqi, Xinjiang (43.95°N, 87.49°E) in 2019.

Nucleotide diversity and molecular evolution of *EL3-2*

To determine whether the *EL3-2* locus has undergone molecular evolution, the ~2.3-kb promoter, 5'UTR, and full-length coding sequence of *EL3-2* were amplified and sequenced in a set of 60 maize inbred lines, 19 maize landraces, and 29 teosinte accessions (Additional file 2: Table S5) using five primer pairs (Additional file 2: Table S12). PCR products from maize inbred lines were directly sequenced, whereas those from maize landraces and teosinte accessions were cloned into pEASY-T5 Zero vector (Beijing TransGen Biotech Co., Ltd.), and one clone per PCR product was randomly selected for sequencing. The sequences were aligned by MUSCLE [83] and curated in BioEdit [84], and gaps exceeding 100 bp were excluded. Nucleotide diversity and Tajima's D-statistic were calculated using DnaSP [85]. Finally, coalescent simulations were conducted for the sequenced regions of *EL3-2* using the simulator *ms* [86] to test whether the observed reduction of nucleotide diversity in maize compared to teosinte could be attributed solely to a domestication bottleneck.

Subcellular localization

The *EL3-2* coding sequence was amplified from Mo17 cDNA using gene-specific primer pair G76sl (Additional file 2: Table S12). The obtained PCR product was then cloned into the SUP1300 vector between the *Xba*I and *Spe*I restriction sites to generate the construct *Super*::*EL3-2-GFP*. Maize protoplasts were isolated from young leaves of B73 following the method described previously [87]. The *Super*::*EL3-2-GFP* construct was transformed into maize protoplasts and incubated in the dark at 22 °C for 16–20 h. GFP fluorescence was observed and imaged using a confocal microscope (Zeiss LSM710).

Phylogenetic analysis of EL3-2 and its homologs

Amino acid sequences of EL3-2 homologs were obtained for *Arabidopsis thaliana* (*At*), *Glycine max* (*Gm*), *Solanum lycopersicum* (*Sl*), *Solanum tuberosum* (*St*), *Brachypodium distachyon* (*Bd*), *Hordeum vulgare* (*Hv*), *Oryza sativa Japonica* (*Os*), *Setaria italica* (*Si*), *Sorghum bicolor* (*Sb*), and *Triticum aestivum* (*Ta*) from the Gramene database (http://www.gramene.org). These amino acid sequences were aligned using Clustal W [88], and a phylogenetic tree was constructed using the neighbor-joining method implemented in MEGA [80] with 1,000 bootstraps.

Sequence and structure analysis of the *EL3-2* locus

The Mo17 genomic sequence surrounding *EL3-2* was retrieved from the Zm-Mo17-REFERENCE-CAU-2.0 genome assembly [89] and the MEX sequence from the assembly of teosinte [5]. The sequences of the promoter, UTRs, and coding region of *EL3-2* were validated using PCR-based Sanger sequencing. Sequence alignment was performed using LAST (version 932) [90], and the repetitive elements were identified and annotated using CENSOR [91].

RNA isolation and expression analysis

Total RNA was isolated from various maize tissues using TRIzol Reagent (Invitrogen). First-strand cDNA was synthesized by reverse transcription using the PrimeScript™ II 1st Strand cDNA Synthesis Kit (TaKaRa). qPCR was performed using the SYBR® *Premix Ex Taq*™ II (TaKaRa) on the 7500 Real-Time PCR System (Applied Biosystems). The maize *ACTIN* gene (*Zm00001d010159*) was used as the internal reference. Relative gene-expression levels were quantified using the comparative C_T ($2^{-\Delta\Delta C}{}_T$) method [92]. Each tissue contained three biological replicates, with each replicate collected from at least five individual maize plants. For shoot apical meristem and immature ears sectioned by meristem types (IM, SPM, SM, and FM), each biological replicate was pooled from at least 30 plants.

In situ hybridization

The immature ears (2–4 mm) from NIL-*EL3-2*^{Mo17} were used for RNA in situ hybridization as described previously [93]. A fragment of the coding and 3'UTR sequence of *EL3-2* was amplified from cDNA using primer pair G76pb4 (Additional file 2: Table S12). Digoxigenin-labeled riboprobes in the sense and antisense orientations were transcribed using SP6/T7 Transcription Kit (Roche) according to the manufacturer's instructions. Slides were exposed for 12–15 h before mounting and imaging. Hybridized tissues were imaged with a Nikon ECLIPSE DIC microscope.

Transcriptome analysis

The immature ears from NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX} were harvested at the stage with 7 (V7) and 14 (V14) extended leaves with two and three biological replicates, respectively. The total RNA from the upper 1/5 portion of the ears was

used for transcriptome sequencing. RNA-seq libraries were constructed according to the standard Illumina RNA-seq protocol, followed by paired-end sequencing on the Illumina HiSeq-4000 platform. Sequence reads were filtered using Trimmomatic (version 0.32) [94]. HISAT2 [95] was used to align reads to the B73 reference genome version 5 (Zm-B73-REFERENCE-NAM-5.0). HTSeq software [96] was used to determine the number of uniquely mapped reads covering each gene model. The counts were normalized according to the library size, and genes with fewer than one count per million were filtered out. DEGs were then identified using edgeR [97], using a false discovery rate (FDR) threshold of 0.05. GO analysis was conducted using agriGO (version 2.0) [98], and the redundant GO terms were simplified using the *calculateSimMatrix* and *reduceSimMatrix* functions from the *rrvgo* package [99] in R.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-026-04139-2.

Additional file 1: Figs. S1–S12. Supplementary figures cited in the main text.

Additional file 2: Tables S1–S12. Supplementary tables cited in the main text.

Acknowledgements
We thank the Center for Crop Functional Genomics and Molecular Breeding at China Agricultural University for maize transformation and seed propagation.

Peer review information
Qingxin Song and Zhenrun Jerry Zhang were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
X.Y., J.L., and N.Y. conceived the research and designed the experiments. R.Z., X.Z., and L.C. performed most of the experiments. R.Z., X.Z., S.W., L.C., and N.Y. performed the data analysis. X.L., XY.Z., Z.C., W.C., J.G., WY.L., C.G., D.Y., YY.L., Y.L., X.G., and WQ.L. performed field trials and phenotyping. J.Z. generated transgenic seeds. F.Y. and D.J. guided the experiments of RNA-seq sampling and in situ hybridization. R.Z., X.Z., and X.Y. prepared the manuscript. F.Y., D.J., J.Y., N.Y., and J.L. edited the manuscript. All the authors read and approved the final manuscript.

Funding
This research was supported by the National Key Research and Development Program (2022YFD1201800) to X.Y., the National Natural Science Foundation of China (32225036, 32201792, 32301818, W2511027) to X.Y., R.Z., X.Z. and N.Y, Chinese Universities Scientific Fund (2025TC147, 2025TC029) to X.Y., the Pinduoduo-China Agricultural University Research Fund (PC2023A01003) to X.Y., and the US National Science Foundation (IOS-2129189) to D.J.

Data availability
The sequence data reported in this paper have been deposited in the National Genomics Data Center [100], Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation. The *EL3-2* cDNA sequences of NIL-*EL3-2*^{Mo17} and NIL-*EL3-2*^{MEX}, as well as the *EL3-2* genomic sequences from 60 maize inbred lines, 19 maize landraces, and 29 teosinte accessions, are available in GenBase [101] under accession numbers C_AA112353 [102], C_AA112354 [103], and C_AA112355–C_AA112462 [104], respectively. The raw RNA-seq data have been deposited in the Genome Sequence Archive [105] under accession number CRA028234 [106]. All microscopy images are available on Figshare [107] under an MIT license. The code used for phenotypic data analysis is hosted on GitHub [108] and archived on Zenodo [109] under an MIT license.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1] Frontiers Science Center for Molecular Design Breeding, State Key Laboratory of Plant Environmental Resilience, and National Maize Improvement Center of China, China Agricultural University, Beijing 100193, China. [2]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China.[3]Center for Crop Functional Genomics and Molecular Breeding, China Agricultural University, Beijing 100193, China. [4]State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Stress Biology, School of Agriculture and Biotechnology, Sun Yat-Sen University, Shenzhen 518107, China. [5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [6]Hubei Hongshan Laboratory, Wuhan, Hubei 430070, China

Received: 14 January 2026 Accepted: 29 May 2026
Published online: 10 June 2026

References

1. Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J. A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci U S A. 2002;99:6080–4.
2. Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R. Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. Proc Natl Acad Sci U S A. 2009;106:5019–24.
3. Yang N, Wang Y, Liu X, Jin M, Vallebueno-Estrada M, Calfee E, et al. Two teosintes made modern maize. Science. 2023;382:eadg8940.
4. Wang L, Beissinger TM, Lorant A, Ross-Ibarra C, Ross-Ibarra J, Hufford MB. The interplay of demography and selection during maize domestication and expansion. Genome Biol. 2017;18:215.
5. Yang N, Xu X, Wang R, Peng W, Cai L, Song J, et al. Contributions of *Zea mays* subspecies *mexicana* haplotypes to modern maize. Nat Commun. 2017;8:1874.
6. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell. 2006;127:1309–21.
7. Flint-Garcia SA. Kernel evolution: from teosinte to maize. In: Maize kernel development. CAB International; 2017. p. 1–15.
8. Zhang M, Kong D, Wang H. Genomic landscape of maize domestication and breeding improvement. Seed Biol. 2023;2:9.
9. Shannon LM. The genetic architecture of maize domestication and range expansion. PhD thesis.University of Wisconsin–Madison; 2012.
10. Xu G, Wang X, Huang C, Xu D, Li D, Tian J, et al. Complex genetic architecture underlies maize tassel domestication. New Phytol. 2017;214:852–64.
11. Chen Q, Yang CJ, York AM, Xue W, Daskalska LL, DeValk CA, et al. TeoNAM: a nested association mapping population for domestication and agronomic trait analysis in maize. Genetics. 2019;213:1065–78.
12. Fu Y, Xu G, Chen H, Wang X, Chen Q, Huang C, et al. QTL mapping for leaf morphology traits in a large maize-teosinte population. Mol Breed. 2019;39:103.
13. Liu L, Huang J, He L, Liu N, Du Y, Hou R, et al. Dissecting the genetic architecture of important traits that enhance wild germplasm resource usage in modern maize breeding. Mol Breed. 2019;39:157.
14. Wang X, Liao C, Wang X, Yang R, Zhai L, Huang J. Construction of maize–teosinte introgression line population and identification of major quantitative trait loci. Euphytica. 2021;217:179.
15. Wang Q, Liao Z, Zhu C, Gou X, Liu Y, Xie W, et al. Teosinte confers specific alleles and yield potential to maize improvement. Theor Appl Genet. 2022;135:3545–62.
16. Cui Y, Xiao X, Wang M, Zhu M, Yuyama N, Zheng J, et al. The construction of a maize-teosinte introgression population and quantitative trait loci analysis of their 21 agronomic traits. Plant Sci. 2024;348:112226.
17. Studer A, Zhao Q, Ross-Ibarra J, Doebley J. Identification of a functional transposon insertion in the maize domestication gene *tb1*. Nat Genet. 2011;43:1160–3.
18. Whipple CJ, Kebrom TH, Weber AL, Yang F, Hall DH, Meeley RB, et al. *grassy tillers1* promotes apical dominance in maize and responds to shade signals in the grasses. Proc Natl Acad Sci U S A. 2011;108:E506-12.
19. Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, et al. The origin of the naked grains of maize. Nature. 2005;436:714–9.
20. Wills DM, Whipple CJ, Takuno S, Kursel LE, Shannon LM, Ross-Ibarra J, et al. From many, one: genetic control of prolificacy during maize domestication. PLoS Genet. 2013;9:e1003604.
21. Liu L, Du Y, Shen X, Li M, Sun W, Huang J, et al. *KRN4* controls quantitative variation in maize kernel row number. PLoS Genet. 2015;11:e1005670.
22. Wang J, Lin Z, Zhang X, Liu H, Zhou L, Zhong S, et al. *krn1*, a major quantitative trait locus for kernel row number in maize. New Phytol. 2019;223:1634–46.
23. Chen W, Chen L, Zhang X, Yang N, Guo J, Wang M, et al. Convergent selection of a WD40 protein that enhances grain yield in maize and rice. Science. 2022;375:abg7985.
24. Shen X, Liu L, Tran T, Ning Q, Li M, Huang L, et al. *KRN5b* regulates maize kernel row number through mediating phosphoinositol signalling. Plant Biotechnol J. 2024;22:3427–41.
25. Jia H, Li M, Li W, Liu L, Jian Y, Yang Z, et al. A serine/threonine protein kinase encoding gene *KERNEL NUMBER PER ROW6* regulates maize grain yield. Nat Commun. 2020;11:988.
26. Ning Q, Jian Y, Du Y, Li Y, Shen X, Jia H, et al. An ethylene biosynthesis enzyme controls quantitative variation in maize ear length and kernel yield. Nat Commun. 2021;12:5832.
27. Luo Y, Zhang M, Liu Y, Liu J, Li W, Chen G, et al. Genetic variation in *YIGE1* contributes to ear length and grain yield in maize. New Phytol. 2022;234:513–26.

28. Pei Y, Deng Y, Zhang H, Zhang Z, Liu J, Chen Z, et al. EAR APICAL DEGENERATION1 regulates maize ear development by maintaining malate supply for apical inflorescence. Plant Cell. 2022;34:2222–41.
29. Luo Y, Liu Y, Liu J, Liu X, Li H, Wu Y, et al. Optimizing thiamine pyrophosphate metabolism enhances crop yield and quality. Nat Commun. 2025;17:165.
30. Dong Z, Hu G, Chen Q, Shemyakina EA, Chau G, Whipple CJ, et al. A regulatory network controlling developmental boundaries and meristem fates contributed to maize domestication. Nat Genet. 2024;56:2528–37.
31. Yang Q, Li Z, Li W, Ku L, Chao W, Ye J, et al. CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proc Natl Acad Sci. 2013;110:16969–74.
32. Huang C, Sun H, Xu D, Chen Q, Liang Y, Wang X, et al. *ZmCCT9* enhances maize adaptation to higher latitudes. Proc Natl Acad Sci. 2018;115:E334–41.
33. Guo L, Wang X, Zhao M, Huang C, Li C, Li D, et al. Stepwise *cis*-regulatory changes in *ZCN8* contribute to maize flowering-time adaptation. Curr Biol. 2018;28:3005-3015.e4.
34. Liang Y, Liu Q, Wang X, Huang C, Xu G, Hey S, et al. *ZmMADS69* functions as a flowering activator through the *ZmRap2. 7-ZCN8* regulatory module and contributes to maize flowering time adaptation. New Phytol. 2019;221:2335–47.
35. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia J-M, Cartwright RA, et al. Comparative population genomics of maize domestication and improvement. Nat Genet. 2012;44:808–11.
36. Liu H, Wang X, Warburton ML, Wen W, Jin M, Deng M, et al. Genomic, transcriptomic, and phenomic variation reveals the complex adaptation of modern maize breeding. Mol Plant. 2015;8:871–84.
37. Xu G, Zhang X, Chen W, Zhang R, Li Z, Wen W, et al. Population genomics of *Zea* species identifies selection signatures during maize domestication and adaptation. BMC Plant Biol. 2022;22:72.
38. Chen L, Luo J, Jin M, Yang N, Liu X, Peng Y, et al. Genome sequencing reveals evidence of adaptive variation in the genus *Zea*. Nat Genet. 2022;54:1736–45.
39. Iltis HH, Doebley JF. Taxonomy of *Zea* (*Gramineae*). II. Subspecific categories in the *Zea mays* complex and a generic synopsis. Am J Bot. 1980;67:994–1004.
40. Doebley JF. The maize and teosinte male inflorescence: a numerical taxonomic study. Ann Mo Bot Gard. 1983;70:32.
41. Iltis HH. Homeotic sexual translocations and the origin of maize (*Zea mays*, Poaceae). Econ Bot. 2000;54:7–42.
42. Lauter N, Gustus C, Westerbergh A, Doebley J. The inheritance and evolution of leaf pigmentation and pubescence in teosinte. Genetics. 2004;167:1949–59.
43. Hufford MB, Bilinski P, Pyhäjärvi T, Ross-Ibarra J. Teosinte as a model system for population and ecological genomics. Trends Genet. 2012;28:606–15.
44. Pan Q, Li L, Yang X, Tong H, Xu S, Li Z, et al. Genome-wide recombination dynamics are associated with phenotypic variation in maize. New Phytol. 2016;210:1083–94.
45. Brown PJ, Upadyayula N, Mahone GS, Tian F, Bradbury PJ, Myles S, et al. Distinct genetic architectures for male and female inflorescence traits of maize. PLoS Genet. 2011;7:e1002383.
46. Xue S, Bradbury PJ, Casstevens T, Holland JB. Genetic architecture of domestication-related traits in maize. Genetics. 2016;204:99–113.
47. Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, et al. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc Natl Acad Sci U S A. 2007;104:11376–81.
48. Chen H, Patterson N, Reich D. Population differentiation as a test for selective sweeps. Genome Res. 2010;20:393–402.
49. Sigmon B, Vollbrecht E. Evidence of selection at the *ramosa1* locus during maize domestication. Mol Ecol. 2010;19:1296–311.
50. Palaisa KA, Morgante M, Williams M, Rafalski A. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell. 2003;15:1795–806.
51. Zhang X, Lin Z, Wang J, Liu H, Zhou L, Zhong S, et al. The *tin1* gene retains the function of promoting tillering in maize. Nat Commun. 2019;10:5608.
52. Tian J, Wang C, Xia J, Wu L, Xu G, Wu W, et al. Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. Science. 2019;365:658–64.
53. Carles CC, Choffnes-Inada D, Reville K, Lertpiriyapong K, Fletcher JC. *ULTRAPETALA1* encodes a SAND domain putative transcriptional regulator that controls shoot and floral meristem activity in *Arabidopsis*. Development. 2005;132:897–911.
54. Carles CC, Fletcher JC. The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants. Genes Dev. 2009;23:2723–8.
55. Monfared MM, Carles CC, Rossignol P, Pires HR, Fletcher JC. The *ULT1* and *ULT2* trxG genes play overlapping roles in *Arabidopsis* development and gene regulation. Mol Plant. 2013;6:1564–79.
56. Pu L, Liu M-S, Kim SY, Chen L-FO, Fletcher JC, Sung ZR. EMBRYONIC FLOWER1 and ULTRAPETALA1 act antagonistically on *Arabidopsis* development and stress response. Plant Physiol. 2013;162:812–30.
57. Pu L, Sung ZR. PcG and trxG in plants–friends or foes. Trends Genet. 2015;31:252–62.
58. Fletcher JC. State of the art: trxG factor regulation of post-embryonic plant development. Front Plant Sci. 2017;8:1925.
59. Yang CJ, Samayoa LF, Bradbury PJ, Olukolu BA, Xue W, York AM, et al. The genetic architecture of teosinte catalyzed and constrained maize domestication. Proc Natl Acad Sci U S A. 2019;116:5643–52.
60. Studer AJ, Wang H, Doebley JF. Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. Genetics. 2017;207:755–65.
61. Bertolini E, Rice BR, Braud M, Yang J, Hake S, Strable J, et al. Regulatory variation controlling architectural pleiotropy in maize. Nat Commun. 2025;16:2140.
62. Sahoo S, Adhikari S, Joshi A, Singh NK. Use of wild progenitor teosinte in maize (*Zea mays* subsp. *mays*) improvement: present status and future prospects. Trop Plant Biol. 2021;14:156–79.

63. Tittes S, Lorant A, McGinty SP, Holland JB, Sánchez-González J de J, Seetharam A, et al. The population genetics of convergent adaptation in maize and teosinte is not locally restricted. Elife. 2025;12:RP92405.
64. Yao W, Ku L, Wang B, Ren Z, Su H, Li C, et al. Breeding ideotype maize with enhanced yield through genomics-guided pyramiding of favorable alleles. Nat Genet. 2026;58:916–27.
65. Meyer RS, Purugganan MD. Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet. 2013;14:840–52.
66. Chen K, Wang Y, Zhang R, Zhang H, Gao C. CRISPR/Cas genome editing and precision plant breeding in agriculture. Annu Rev Plant Biol. 2019;70:667–97.
67. Lyzenga WJ, Pozniak CJ, Kagale S. Advanced domestication: harnessing the precision of gene editing in crop breeding. Plant Biotechnol J. 2021;19:660–70.
68. Kummari D, Palakolanu SR, Kishor PBK, Bhatnagar-Mathur P, Singam P, Vadez V, et al. An update and perspectives on the use of promoters in plant genetic engineering. J Biosci. 2020;45:119.
69. Zhao H, Sun Z, Wang J, Huang H, Kocher J-P, Wang L. CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics. 2014;30:1006–7.
70. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67:1–48.
71. Harrell Jr F. Hmisc: Harrell Miscellaneous. R package version 5.2-2. 2025. https://github.com/harrelfe/hmisc.
72. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76–82.
73. Kolde R. pheatmap: Pretty Heatmaps. R package version 1.0.13. 2025. https://github.com/raivokolde/pheatmap.
74. Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994;136:1457–68.
75. Wang S, Basten C, Zeng Z. Windows QTL Cartographer 2.5. Raleigh, NC, USA: Department of Statistics, North Carolina State University; 2012.
76. Liu J, Huang J, Guo H, Lan L, Wang H, Xu Y, et al. The conserved and unique genetic architecture of kernel size and weight in maize and rice. Plant Physiol. 2017;175:774–85.
77. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet. 2013;45:43–50.
78. Shannon P. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
79. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
80. Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. MEGA12: molecular evolutionary genetic analysis version 12 for adaptive and green computing. Mol Biol Evol. 2024;41:msae263.
81. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78-82.
82. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27:2156–8.
83. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
84. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41:95–8.
85. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 2003;19:2496–7.
86. Hudson RR. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics. 2002;18:337–8.
87. Burdo B, Gray J, Goetting-Minesky MP, Wittler B, Hunt M, Li T, et al. The Maize TFome – development of a transcription factor open reading frame collection for functional genomics. Plant J. 2014;80:356–66.
88. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8.
89. Chen J, Wang Z, Tan K, Huang W, Shi J, Li T, et al. A complete telomere-to-telomere assembly of the maize genome. Nat Genet. 2023;55:1221–31.
90. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. Genome Res. 2011;21:487–93.
91. Kohany O, Gentles AJ, Hankus L, Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics. 2006;7:474.
92. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008;3:1101–8.
93. Jackson D. In situ hybridization in plants. In: Molecular plant pathology: a practical approach. Oxford University Press; 1991. p. 163–74.
94. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
95. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37:907–15.
96. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9.
97. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.
98. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 2017;45:W122-9.
99. Sayols S. rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. microPublication Biol. 2023;10:17912.
100. CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2025. Nucleic Acids Res. 2025;53:D30-44.

101. Bu C, Zheng X, Zhao X, Xu T, Bai X, Jia Y, et al. GenBase: a nucleotide sequence database. Genomics Proteomics Bioinformatics. 2024;22:qzae047.
102. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. Global genetic dissection ofmaize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. GenBase. 2026. https://ngdc.cncb.ac.cn/genbase/search/gb/C_AA112353.
103. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. Global genetic dissection ofmaize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. GenBase. 2026. https://ngdc.cncb.ac.cn/genbase/search/gb/C_AA112354.
104. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. Global genetic dissection of maize–teosinte divergence reveals EL3-2 as a pleiotropic domestication regulator. GenBase. 2026. https://ngdc.cncb.ac.cn/genbase/search/gb/C_AA112355. (accession numbers: C_AA112355–C_AA112462).
105. Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, et al. The genome sequence archive family: toward explosive data growth and diverse data types. Genomics Proteomics Bioinformatics. 2021;19:578–83.
106. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. Global genetic dissection of maize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. Genome Sequence Archive. 2026. https://ngdc.cncb.ac.cn/gsa/browse/CRA028234.
107. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. Microscopy imaging data for: Global genetic dissection of maize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. 2026. Figshare. https://doi.org/10.6084/m9.figshare.32147080.
108. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, et al. TM population phenotype analysis code for: Global genetic dissection of maize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. GitHub. 2026. https://github.com/rz435/TM_population_phenotype_analysis.
109. Zhang R, Zhang X, Wu S, Cai L, Li X, Zhao X, Chen Z, Chen W, Guo J, Li WY, Guo C, Yan D, Li Y, Luo Y, Ge X, Li WQ, Zhuang J, Yang F, Jackson D, Yan J, Li J, Yang N, Yang X. TM population phenotype analysis code for: Global genetic dissection of maize–teosinte divergence reveals *EL3-2* as a pleiotropic domestication regulator. Zenodo. 2026. https://doi.org/10.5281/zenodo.18989738.

Publisher's Note