BMC Genomics

RESEARCH

Open Access

Cross-species single-nucleus analysis reveals the potential role of whole-genome duplication in the evolution of maize flower development

Huawei Feng[1], Wenjuan Fan[1], Min Liu[2], Jiaqian Huang[1,3], Bosheng Li[1], Qing Sang[1*] and Baoxing Song[1,3*]

Abstract

Background The evolution and development of flowers are biologically essential and of broad interest. Maize and sorghum have similar morphologies and phylogeny while harboring different inflorescence architecture. The difference in flower architecture between these two species is likely due to spatiotemporal gene expression regulation, and they are a good model for researching the evolution of flower development.

Results In this study, we generated single nucleus and spatial RNA-seq data for maize ear, tassel, and sorghum inflorescence. By combining single nucleus and spatial transcriptome, we can track the spatial expression of single nucleus cluster marker genes and map single nucleus clusters to spatial positions. This ability provides great power to annotate the single nucleus clusters. Combining the cell cluster resolved transcriptome comparison with genome alignment, our analysis suggested that maize ear and tassel inflorescence diversity is associated with the maize-specific whole genome duplication. Taking sorghum as the outgroup, it is likely that the loss of gene expression profiling contributes to the inflorescence diversity between tassel and ear, resulting in the unisexual flower architecture of maize. The sequence of highly expressed genes in the tassel is more conserved than the highly expressed genes in the ear.

Conclusion This study provides a high-resolution atlas of gene activity during inflorescence development and helps to unravel the potential evolution associated with the differentiation of the ear and tassel in maize.

Keywords Maize, Sorghum, Inflorescence, Evolution, Single-nucleus RNA-seq, Whole-Genome Duplication

Background

Understanding the evolution of molecular mechanisms associated with organism development diversity has been of broad interest [1–3]. Closely related plant species can exhibit huge morphological trait differences. Flowers are responsible for producing the next generation of plants and of high morphological variation. The diversity and evolution of the internal organization of the flower are crucial in phylogenetic and taxonomic studies of angiosperms [4].

Maize (*Zea mays* L.) and sorghum (*Sorghum bicolor*) are two model and essential crop species belonging to the Andropogoneae tribe and share a common ancestor around 12 million years ago. The maize lineage has undergone a whole genome duplication (WGD) since diverging from sorghum [5] (termed as the recent WGD, herein), but subsequent chromosomal fusions resulted in these species having the same chromosome number ($n=10$). While the sub-genomes of maize are in extensive syntenic with the sorghum genome [6]. These two

*Correspondence:
Qing Sang
qing.sang@pku-iaas.edu.cn
Baoxing Song
baoxing.song@pku-iaas.edu.cn
[1] Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agriculture Sciences in Weifang, Weifang, Shandong 261325, China
[2] Baimaike Intelligent Manufacturing, Qingdao, Shandong 266500, China
[3] Key Laboratory of Maize Biology and Genetic Breeding in Arid Area of Northwest Region, Ministry of Agriculture, College of Agronomy, Northwest A&F University, Yangling, Shaanxi 712100, China

closely related species can look strikingly similar before flowering. However, they have distinct terminal inflorescence architectures. Maize is monoecious and diclinous, containing two distinctive inflorescences, tassel and ear, bearing male and female flowers, respectively. Unisexual flowers are highly advantageous for maize, enabling hybrid production through outcrossing and making commercial seed production easy [7].

The maize tassel originates from the shoot apical meristem (SAM). It consists of long, indeterminate branches at the base and a central spike with shorter branches housing spikelet pairs. In contrast, the ears are located laterally in the leaf axils and consist only of short branches [8]. The tassel and ear share a common structure, where an apical indeterminate inflorescence meristem (IM) produces a series of determinate spikelet-pair meristems (SPM), each of which generates two spikelet meristems (SM) that, in turn, initiate two floral meristems (FM) [9].

Interestingly, both tassel and ear initially develop as bisexual flowers, containing both male and female components [10, 11]. Shortly after initiation, the female parts (gynoecia) of the tassel and the male parts (stamens) of the ear gradually decay, resulting in the tassel florets maturing into unisexual male florets and the ear florets into unisexual female florets [11]. Like maize tassel, the sorghum features a determinate panicle at the end of the shoot meristem [12], typically with bisexual flowers.

We frequently observe that, under stress, maize could produce bisexual flowers which described as "tassel-seed (ts)". These flowers are characterized by branched inflorescences that have female branches enclosed in husk leaves, culminating in a spike of male flowers [13]. The survival of female floral parts on the tassel is likely environmentally triggered [14]. Several sex-determination mutants have been identified in maize, particularly noticeable in "ts" mutants where silks (pistils) and kernels (seeds) appear in the tassel [15]. For example, *ts1* encodes a lipoxygenase, also known as *ZmLOX8*, which plays a role in jasmonic acid (JA) biosynthesis [16] and *ts2* encodes a monocot-specific short-chain alcohol dehydrogenase [17]. The *ts5* phenotype arises from disrupted JA signaling during sexual differentiation, caused by the upregulation of *ZmCYP94B1* [18]. Other mutants featuring silks in the tassels include *ts4*, which encodes a miR172 microRNA, and *ts6*, which contains a mutation in the *ts4*-binding site [19]. A detailed investigation of the gene expression in flowers of maize and sorghum at single-cell resolution could shed light on the evolution of this vital reproductive organ.

In this study, we present a high-resolution cell atlas of transcriptome comparative analysis for maize tassel inflorescence and ear inflorescence utilizing single nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics approach with sorghum inflorescence as an outgroup. Our result suggested that the recent WGD is associated with the differentially expressed genes (DEGs) between maize ear and tassel. Taking sorghum as the outgroup, altered gene expression mainly contributed to the degeneration of gynoecia in the tassel and stamens in the ear.

Methods

Plant growth conditions and tissue collection stage

Maize B73 inbred and sorghum BTx623 plants were grown in the greenhouse with 16 h daytime, 26–28 °C, and 8 h night, 22–24 °C.

For single nucleus RNA-seq, spatial transcriptome sequencing, and in situ hybridization, the developing maize ears and tassels at the 5–10 mm stage, where major developmental and architectural decisions, including meristem initiation, maintenance and determinacy, organ specification, and differentiation of vascular and ground tissues, are being collected [20]. The sorghum inflorescence tissues were collected from BTx623 at stage 3 during floral transition [21]. All the samples were immediately frozen in liquid nitrogen and stored at -80 °C.

Nuclei isolation and 10 × single nucleus RNA-seq library construction

Nuclei isolation for maize ears, tassels and sorghum inflorescence tissues was performed as previously reported [22]. In brief, the tissues were chopped in ice-cold extraction buffer (0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 20 mM 4-(2-Hydroxyethyl)-1-piperaz ineethanesulfonic acid, 10 mM MgCl2, 0.5% Triton X-100, 1 mM DTT, 1 × protease inhibitor, and 0.4 U/μL RNase inhibitor). The nuclear solution was filtered through 40 μm cell strainer once and 20 μm cell strainer twice, and then centrifuged at 500 g for 5 min at 4 °C. The pellet was washed once with extraction buffer, and then resuspended in 1 × PBS solution. The nuclei were stained with 4,6-diamino-2-phenylindole (DAPI) and counted with a cytometer. The final concentration was diluted to 1000–1200 nuclei/μl. The snRNA-seq library was constructed with 10 × Genomic protocol (Chromium Single Cell 3' Reagent Kits v3.1User Guide) and sequenced by Illumina NovaSeq (paired-end, 2 × 150 bp).

Spatial transcriptome sequencing on BMKMANU S1000 spatial transcriptomics platform

The tissues were prepared, embedded, and cryosectioned according to the previously described [23] with modifications. In brief, the fresh tissues were soaked in 0.05% tween 20 for 15 min at room temperature and then were transferred into 75% ice-cold optimal cutting temperature (OCT) compound and vacuum for 10 min. Soak the tissues in OCT and put them on dry ice until fully frozen. The section thickness of 10 um is adopted in this study.

Tissue sectioning, Trypan staining, tissue permeabilization, fluorescent cDNA synthesis and imaging were performed on BMKMANU S1000 spatial transcriptomics platform according to the manufacturer's protocol. cDNAs of samples were sequenced on the Illumina NovaSeq 6000 platform.

Single nuclei and spatial RNA-seq analysis

We use Cell Ranger (version 7.1.0; 10×Genomics, US) to quantify the gene expression level by counting unique molecular identifiers (UMIs) for each gene from each sample separately. The maize genome assembly V5 and sorghum genome assembly V3 were used as the reference genomes. The reference genomes were indexed using the cellranger "mkref" command with default parameters. The reads were aligned to the reference sequence, and expression levels were determined for each cell and gene using the cellranger count command with modified STAR parameter "–alignIntronMin 20 –alignIntron Max 10000".

To filter out low-quality cells for each sample, the expression profiling was further processed using Seurat v4.3.0 for each sample. Rarely expressed genes were filtered using the parameter "min.cells = 3" and cells with < 200 genes detected were filtered using the parameter "min.features = 200". We calculated quantiles 0.01 and 0.99 for the number of features from each cell for each sample separately, and cells with features out of the range of quantiles 0.01 and 0.99 are further filtered. We used the DoubletFinder [24] package to remove potential doublets (two cells encapsulated in a single droplet) with default parameters.

To perform clustering, we conducted unsupervised clustering with a resolution of 0.5 using the Seurat R package following the standard manual. The clusters were then visualized using the Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) method. Finally, marker genes were automatically calculated by the Wilcoxon rank-sum test implemented in the "FindAllMarkers " function.

Each RNA capturing spot on the BMKMANU S1000 spatial transcriptome profiling chip is 2.5 μm in diameter and the center-to-center distance of two adjacent spots is 4.8 μm. The size of a BMKMANU S1000 chip is 6.5*6.5 mm. We merged the high-resolution stained images and sequencing data using the BSTMatrix system. We could use the merged dataset to conduct spatial RNA-seq analysis using different bin sizes, including bin1 (~ 4.8 μm), bin7 (~ 9.6 μm) and bin19 (~ 19.2 μm). There are too few UMIs at bin1 resolution. Bin7 (~ 9.6 μm) was chosen to represent each spot and construct the spatial transcriptome.

To map single-cell clusters to spatial positions, we normalized the counting matrix from snRNA-seq and spatial RNA-seq using the "SCTransform" function from the Seurat package separately. We performed dimensionality reduction for the spatial RNA-seq by PCA using the "RunPCA" function. We used the "FindTransferAnchors" function to identify a set of anchors between the snRNA-seq and spatial RNA-seq data. The transferring categorical information was obtained using the "TransferData" function.

To perform meta-snRNA clustering for maize ear and tassel, we remove batch effects using the multiset canonical correlation analysis (MultiCCA) approach [25] implemented in the Seurat package. Highly variable genes were identified across all the data sets by calculating the mean expression and dispersion (variance/mean) for each gene. With these genes, MultiCCA was performed using the "FindIntegrationAnchors" and "IntegrateData " functions to define a shared correlation space and generate an integrated gene expression matrix. The integrated gene expression matrix was used for standard clustering as introduced above. We used the "FindAllMarkers(only.pos = TRUE, min.pct = 0, logfc.threshold = 0.25, test.use = "MAST")" function to identify HEGs for each cluster. To perform meta-snRNA clustering for maize ear, tassel and sorghum, we used only genes with a single syntenic copy in maize obtained using the quota-alignment toolkit [26] (see the Homologous and gene family analysis part).

We mapped the maize ear, tassel and sorghum meta clustering label back to the three expression matrices normalized using the "usingScaleData(features = all.genes)" function. We used the "AverageExpression" function to get a normalized average expression value for each gene in each cluster. These normalized matrixes were used for dosage correlation analysis. To compare gene expression levels, we calculated the average dosage value of each gene in each cluster as its expression level.

In Situ Hybridization

In situ hybridization was performed as described in Bradley et al. [27]. The sequences of primers used to generate the probes are listed in table S7.

Homologous and gene family analysis

The syntenic orthologous between maize and sorghum was identified using the quota-alignment package [26] following the standard protocol, except that the value of parameter quota was set 2:1 to account for the evolutional event that the maize genome experienced one more round of whole duplication after diversity from maize.

The YABBY gene list was acquired by searching the keyword "yabby" from the MaizeGDB website (01 SEP 2023). Records without gene ID from the V5 assembly were ignored.

Using the quota-alignment result and following the prior studies [28, 29], we obtained a list of syntenic homeologues in the M1 and M2 genomes, and they were referred to as recent whole-genome duplicated homologs. To identify the other types of duplicated genes, we used the DupGen_finder-unique.pl script of the DupGen_finder package [31]. Duplicate pairs are derived from five. Gene pairs present in WGD classification while absent from the M1 and M2 gene pairs, were classified as ancestral WGD.

Functional annotation enrichment analysis

We conducted Gene ontology (GO) analyses using the biological process database of the ShinyGO v0.80 [32] online application with default parameters. The top 20 terms, ranked via fold enrichment, were used to draw the plots of Fig. 2.

Sequence conservation analysis

We conducted sequence alignment and calculated the alignment score for each sorghum-maize ortholog pair (using the CDS sequence) using the globalms function of the Bio.pairwise2 module(parameters: match = 1; mismatch = -1;open = -1; extend = -0.5) from the Biopython package [33]. The scores are then divided by the mean length of the CDS sequences of each sorghum-maize ortholog pair to get the CDS similarity.

To estimate protein substitution rates (Ka/Ks), we used the 'free-ratio' branch model (parameters: model = 1, Nssites = 0, CodonFreq = 2) from Bio.Phylo.PAML.codeml module from the Biopython package [33]. The 'free-ratio' branch model assumes an independent Ka/Ks ratio for each branch. The tree of ((maize, sorghum), millet), rice) was used.

Results

Construction cell atlas of maize ear and tassel inflorescences by integrating single nucleus and spatial transcriptomics

The architecture of the maize tassel and ear is determined when meristem initiation, maintenance, and determinacy decisions start to form specific organs [34]. For maize, at about V5 (5th leaf stage), a microscopically small tassel is initiated in the stem apex tip just under or at the soil surface. At V6, the tassel is above the soil surface. At V7-V8, the tassel size is about 2 to 5 mm. By V9, ear shoots, which are potential ears, can be seen upon dissection. An ear shoot forms at each above-ground node, except for the last six to eight nodes beneath the tassel, but only one or two of these ear shoots will mature into harvestable ears. The number of ovules (potential kernels) and the overall size of the ear are established by the V12 stage. To investigate the regulation of maize inflorescence development, we collected tassels and ears at a stage where major meristem initiations and architectural decisions, including IM, SPM, SM and FM for transcriptomic profiling [20]. For sorghum, we collected the inflorescence at stage 3 during the floral transition and ovary and stamen primordia started to be observed [21] for transcriptomic profiling.

The plant samples were processed following previously described methods to generate isolated single nuclei [22]. The snRNA-seq libraries were constructed with 10 × Genomic protocol and sequenced using Illumina NovaSeq. We used the Cell Ranger package [35] to quantify the gene expression level by counting unique molecular identifiers (UMI) for each gene from each sample. Data were filtered at both cell and gene levels using Seurat v4.3.0 [36] for each sample separately. To assign the information of cells to spatial locations, we utilized the BMKMANU S1000 techniques [22] to generate spatial transcriptome data for maize ear, tassel, and sorghum inflorescences at similar developmental stages as these used for snRNA-seq (Figs. 1A and S1).

Each spot on the S1000 spatial transcriptomics chip is 2.5 μm in diameter and the center-to-center distance between two adjacent spots is 4.8 μm. We merged the data of 7 spots as a single bin (bin7) for the following analysis with a resolution of 9.6 μm (details see Methods and Materials). Fewer UMI and genes were captured in the central region than in the peripheral region for maize ear and tassel (Figure S1), likely due to bigger vacuoles. Compared to single-cell RNA sequencing (scRNA-seq), snRNA-seq has the advantage of profiling expression for cells with a large range of sizes [37]. For the sorghum inflorescence, cells with putative bigger vacuole could also be observed in the central region, and we could observe spots with higher UMI density in stripe patterns (Figure S1), indicating an anatomical difference between maize and sorghum flower. We preprocessed the snRNA-seq data by filtering both low features and cells/spots. There were 6,824–8,000 cells left for each snRNA-seq sample and a median of 1,963–2,659 genes for each cell (Table S1). However, spatial transcriptomics has not yet been able to attain single-cell resolution, and each spot captured the expression profile of a small group of genes (Table S2). It can only identify the physical distribution of a cell cluster. SnRNA-seq captures abundant gene expression profiles and has a better ability to identify cell types within tissues [38]. Integration of snRNA-seq and spatial transcriptomics enables us to improve cell type

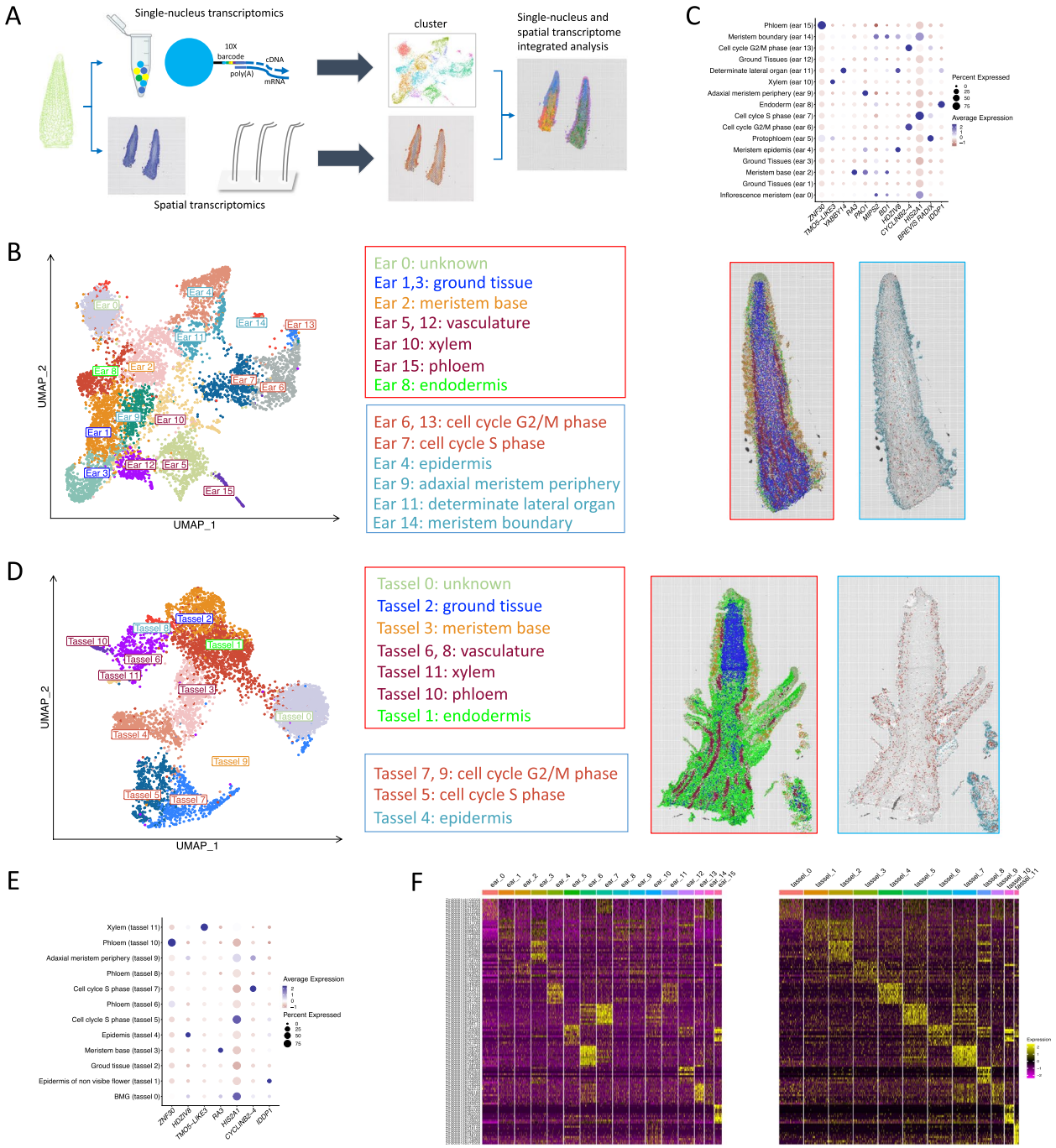

Fig. 1 Cell clustering and cell-type identification in maize ear and tassel based on spatial transcriptomics and marker genes. **A** A schematic diagram of spatial and single nuclear RNA sequencing, taking the maize ear for example. The size of the spatial transcriptome capture chip is 6.8 mm * 6.8 mm. The ear inflorescences were collected at a stage that could capture major meristem initiations and architectural decisions. **B** UMAP of ear single nucleus transcriptomic clusters and mapping single nucleus transcriptomic clusters to the spatial transcriptomics data using an anchor-based integration approach. The colors of the labels in the middle match that of their spatial clusters on the right side. **C** The expression profile of previously used marker genes in ear single nucleus transcriptomic clusters. **D** UMAP of tassel single nucleus transcriptomic clusters and mapping single nucleus transcriptomic clusters to the spatial transcriptomics data using an anchor-based integration approach. **E** The expression profile of previously used marker genes in tassel single nucleus transcriptomic clusters. **F** The expression spectrum of tassel single nucleus transcriptomic clusters specific genes in ear single nucleus transcriptomic clusters and tassel single nucleus transcriptomic clusters

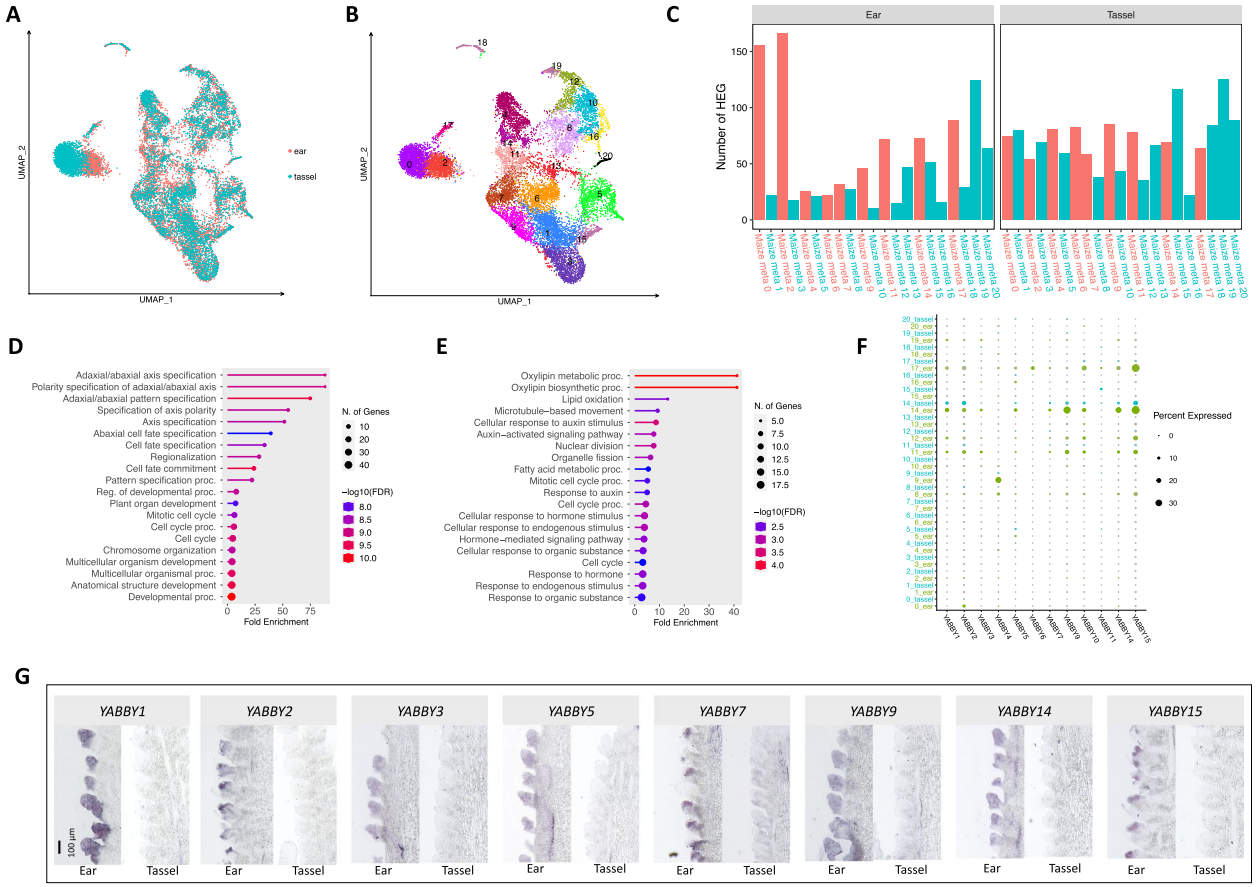

Fig. 2 Comparison of transcriptome between ear and tassel at single-cell resolution. **A** The snRNA meta clustering of combining maize ear and tassel. **B** The maize cluster 0 is highly specific for tassel and the maize cluster 2 is highly specific for ear. **C** The number of highly expressed genes (HEGs) comparing between ear and tassel in each snRNA cluster. Clusters in red can be mapped to an anatomically inflorescence meristem and spikelet region, and we focused on them for gene ontology (GO) enrichment analysis. **D** The enriched GO terms of HEGs in the ear. The specification of axis polarity and cell cycle terms are significantly enriched. **E** The enriched GO enrichment of HEGs in the tassel. The auxin signaling and cell cycle terms are significantly enriched. (F) *YABBY* genes generally have higher expression levels in the ear than in the tassel. **G** mRNA in situ hybridization of *YABBY* genes in ear and tassel

annotation and helps us better understand the function of each group of cells.

After the linear dimensional reduction for the snRNA-seq data, we used the DoubletFinder [24] package to filter putative doublets, RNA-seq read from two or more nuclei with the same barcode, and use the graph-based clustering approaches implemented in the Seurat package to conduct unsupervised clustering. The cluster results were visualized using a uniform manifold approximation and projection (UMAP) plot, which can provide insights into the biological similarity of adjacent clusters (Fig. 1B). To annotate snRNA clusters, we mapped cell clusters to the spatial regions using the anchor-based integration workflow [39] implemented in Seurat, which enables the probabilistic transfer of snRNA clusters to the spatial profiling. With the maize marker genes used by Xu et al. [40] (Fig. 1C), we were able to identify vasculature-phloem (marker gene *ZNF30,* cluster ear 15, Figure S2A), vasculature-xylem (marker gene *TMO5-LIKE3*, cluster ear 10, Figure S2B), determinate lateral organ (marker gene *YABBY14*, cluster ear 11, Figure S2C), meristem base (marker gene *RA3*, cluster ear 2, Figure S2D), adaxial meristem periphery (marker gene *PAO1* [40], cluster ear 9, Figure S2E), meristem boundary (marker genes *MIPS2* and *BD1*, cluster ear 14, Figure S2F), epidermis (marker gene *HDZIV8*, cluster ear 4, Figure S2G), cell cycle G2/M phase (marker gene *CYCLINB2-4*, cluster ear 6 and cluster ear 13, Figures S2H and S2I) and cell cycle S phase (marker gene *HISTONE2A*, cluster ear 7, Figure S2J). We found the *BREVIS RADIX* (Zm00001eb071850) was highly enriched in ear cluster 5, and we annotated ear cluster 5 as vasculature [41] (Figure S2K). *IDDP1* is expressed explicitly in cluster ear 8, and its Arabidopsis ortholog, *BLUEJAY*, has been reported as a marker

gene for endoderm [42]. The ear spatial data also suggested that *IDDP1* is expressed explicitly in the anatomical endodermal region (Figure S2L). Using the strategy of mapping snRNA clusters to spatial transcriptomics data, we annotated the clusters ear 1 and 3 as ground tissue (Figures S2M and S2N). We annotated cluster ear 12 as vasculature by mapping the snRNA cluster to spatial transcriptomics data (Figure S2O). The cluster ear 0 was mainly mapped to the inflorescence meristem and floret region of spatial transcriptomics data using the anchor-based integration workflow, and we termed cluster ear 0 as unknown (Figure S2P).

Using the same approach, we got 12 snRNA clusters (Figs. 1D and 1E) for tassel. The specifically expressed genes identified are largely shared between maize ear and tassel (Figs. 1F, S4 and S5). We identified vasculature-phloem (marker gene *ZNF30,* cluster tassel 10, Figure S3A), vasculature-xylem (marker gene *TMO5-LIKE3*, cluster tassel 11, Figure S3B), epidermis (marker gene *HDZIV8*, cluster tassel 4, Figure S3C), meristem base (marker gene *RA3*, cluster tassel 3, Figure S3D), cell cycle S phase (marker gene *HISTONE2A*, cluster tassel 5, Figure S3E), cell cycle G2/M phase (marker gene *CYCLINB2-4*, cluster tassel 7, Figure S3F) and endodermis (marker gene *IDDP1*, cluster tassel 1, Figure S3G). By mapping the snRNA clusters to the spatial transcriptomics data using an anchor-based integration workflow, we annotated tassel cluster 6 and tassel cluster 8 as vasculature (Figures S3H and S3I). The cluster tassel 2 was mainly mapped to the ground tissue of spatial transcriptomics data (Figure S3J) and shared cluster-specific genes with the ear ground tissue (ear cluster 12). The cluster tassel 9 was annotated as cell cycle G2/M phase (Figure S3K), with the observation that sharing cluster-specific genes with the ear cluster 13 and tassel cluster 7 (Figs. 1F and S4). The cluster tassel 0 was mainly mapped to the inflorescence meristem and floret region of spatial transcriptomics data (Figure S3L) and shared cluster-specific genes with ear cluster 0.

For the maize ear and tassel cell cluster annotations, spatial expression of marker genes and the probabilistic transfer of snRNA clusters to the spatial profiling largely confirm each other, demonstrating the power of integrating scRNA-seq and spatial data. The vasculature, vasculature-phloem and vasculature-xylem clusters are adjacent on the UMAP plot and in strip shapes on the spatial distribution. The meristem base, epidermis, meristem boundary, adaxial meristem periphery, and determinate lateral organ are mapped to the floret region. In summary, by combining the snRNA and spatial RNA-seq, here we classified the major cell types of maize ear and tassel inflorescence. The spatial data provided valuable hints to identify or confirm the annotation of snRNA clusters (Figure S5). These data provide reliable basics to compare the molecular mechanism underlying maize ear, maize tassel, and sorghum inflorescence developmental morphology.

Comparison of transcriptome between ear and tassel at cell cluster resolution

To investigate the potential gene expression profiling corresponding to the developmental diversity between the maize tassel and ear, we performed a meta-clustering by merging the snRNA data of these two tissues. Batch effects between the two replicates were removed using the batch effect correction algorithm implemented in the Seurat package [39]. Using an unsupervised approach as introduced above, we got 21 clusters (Fig. 2A, maize meta cluster 0 to maize meta cluster 20). The maize meta cluster 0 is enriched in tassel, while the maize meta cluster 2 is enriched in ear (Figs. 2A, B and S6). Each single cell cluster can be mapped to an anatomically specific spatial region (Figures S7 and S8). The data enable us to conduct single-cell cluster resolved differential gene expression analysis, providing a higher resolution than bulk RNA-seq.

Tassel and ear start similar developmental processes but form unique architectures. Tassel primordia also produce multiple branch meristems from the base, contributing to their distinct architecture. During the floral organ stage, the tassel and ear differentiate into male and female inflorescences, respectively, due to the degeneration of gynoecium primordia in the tassel florets and stamen primordia in the ear florets [43]. Various genes that regulate inflorescence structure and sex determination in maize have been identified, but they were identified mainly by forward genetics. To uncover more genes that regulate inflorescence architecture and sex determination, we performed differential gene expression analysis for each snRNA cluster separately and focused on the snRNA clusters mapped to the inflorescence meristem and spikelet regions (Figs. 2C and S7). There were 339 genes detected as highly expressed in the ear (ear HEGs, Table S3). Gene ontology (GO) enrichment analysis using ShinyGO v0.80 [32] indicated that the specification of axis polarity, etc., are significantly enriched (Fig. 2D, Table S4). 5 out of 6 genes related to this GO item belong to the *YABBY* family (*YABBY1*, *YABBY4*, *YABBY9*, *YABBY14*, and *YABBY15*). The genes in the *YABBY* transcription factor family, including *YABBY2, YABBY3, YABBY5, YABBY7* are also specifically highly expressed in maize ear (Fig. 2F, G). Genes of the*YABBY* transcription factor family have been reported to play an essential role in the development of the dimorphic, unisexual ear and tassel florets in maize (*DRL1-YABBY2*

and *DRL2-YABBY7*) [44]. CRABS CLAW (CRC), a gene closely related to *YABBY2* and *YABBY7* (Figure S9), is highly expressed in the carpel, and its mutant results in female-to-male sex transition in cucurbits [45]. The results of in situ hybridizations (Fig. 2G) showed that these *YABBY* mRNAs are specifically expressed in FM, and their signal was much stronger in the ear than in the tassel. These results suggest that the YABBYs might contribute to sex determination in maize.

There were 241 genes highly expressed in the tassel compared to the ear (tassel HEGs, Table S5). The GO term "oxylipin metabolic process" is enriched (Fig. 2E, Table S6) and all the tassel HEGs annotated as this GO term are *LOX* genes, including *LOX1_A* (Zm00001eb144930), *LOX1_B* (Zm00001eb144960), *LOX3*, *LOX5* and *LOX10* (Figure S10). *LOX* genes are essential components for the synthesis of JA. JA deficiency is responsible for the "ts" phenotype, where the famous *ts1* gene was *LOX8*. JA has been reported to be essential for the abortion of pistil primordia and the development of maize tassel [16]. JA is also required for the development of sorghum inflorescences [21, 46].

Gibberellins (GAs) have long been researched in sex determination in maize. The addition of exogenous GAs feminizes the tassel [47], whereas the reduction of GAs in the ear prevents stamen abortion, leading to flowers with male and female organs [48]. Our results showed that the transcriptional levels of GA biosynthesis gene, *GA20ox5* were ear-enriched, and the expression levels of GA deactivations such as *GA2ox2, GA2ox8* and *GIBBERELLIN-INSENSITIVE DWARF PROTEIN HOMOLOG 3 (GID3)* were highly expressed in the tassel.

Our results showed that the term of cell auxin signal is significantly enriched in tassel highly expressed genes (HEGs) (Fig. 2E). The auxin signaling modules play important roles in maize inflorescence architecture and tassel branching [49]. The tassel HEGs include Aux/IAA-transcription factors and also several auxin response factors (*ARFTF20*, *ARFTF23*). *IAA4* and *IAA32* are Aux/IAA-transcription factors and are significantly highly expressed in tassel. *BARREN INFLORESCENCE 1* and *BARREN INFLORESCENCE 4* (*BIF1* and *BIF4*) have been reported as a key component of the auxin hormone signaling pathway and their mutants developed tassels with fewer branches and spikelets than normal [49]. *BARREN INFLORESCENCE2 (BIF2)* plays key roles in axillary meristem and lateral organ initiation in maize inflorescence. The *bif2* mutation makes fewer branches, spikelets, florets and floral organs in the tassel [50, 51]. Interestingly, our results showed that the transcript levels of *BIF1* and *BIF2* were significantly higher in tassel. We further checked the expression pattern of DEGs annotated as related to the auxin signal and found they did not show strong single-cell cluster specification (Figure S11).

Collectively, these findings indicate that the cluster-resolved DEGs analysis could identify essential genes potentially relating to the degradation of gynoecia in the tassel, stamens abortion in the ear, and inflorescence branching.

Whole-genome duplication contributes to the evolution of maize inflorescence architecture

Gene duplication is a widely appreciated contribution to the evolution of phenotypic novelty within plants [52]. DEGs between ear and tassel are significantly enriched with genes with duplicated copies (Fig. 3A). To investigate the evolution of duplicated genes further, we classified homology gene pairs into five classifications, arising from whole-genome, tandem, transposon-mediated, proximal (separated by up to ten genes), and dispersed (separated by more than ten genes) duplicate pairs (details see Methods and Materials). The result suggested HEGs are significantly enriched in WGD genes (Fig. 3B). Tassel or ear HEGs are enriched in syntenic with sorghum. There are 21,810 out of 39,702 maize genes in syntenic with the sorghum genome. Out of 560 ear HEGs, 479 genes are syntenic with sorghum, which is significantly more than random expectation (p-value < 0.001, hypergeometric test). Similarly, 385 out of 459 tassel HEGs are syntenic with sorghum and significantly enriched (p-value < 0.001, hypergeometric test). We further classified maize duplicated genes into the recent WGD (the WGD happened with maize but not sorghum) and ancestral WGDs (detectable WGDs shared with sorghum). The syntenic genes, resulting from the recent WGD, made a prevalent contribution to expression domain expansion for floral development (Fig. 3C). The maize duplicated genes resulting from the recent WGD could be divided into two subgenomes, M1 and M2. The M1 genes are more conserved compared to the M2 genes and M1 genes have slightly higher expression levels [29], indicating functional conservation. We checked the distribution of DEGs between the tassel and ear on M1 and M2, and found the DEGs between the tassel and ear are dominated by M1 genes (Fig. 3D).

To further understand the gene regulation in maize inflorescence development, we extracted 9459 single-copy syntenic orthologues in maize and sorghum. We used single-copy syntenic orthologues to perform snRNA clustering for maize ear, maize tassel and sorghum with removed batch effects and got 18 meta-clusters (Figure S12, S13). We then conducted Pearson correlation analysis for the expression profile of syntenic homolog gene pairs across single nucleus clusters. For duplicated

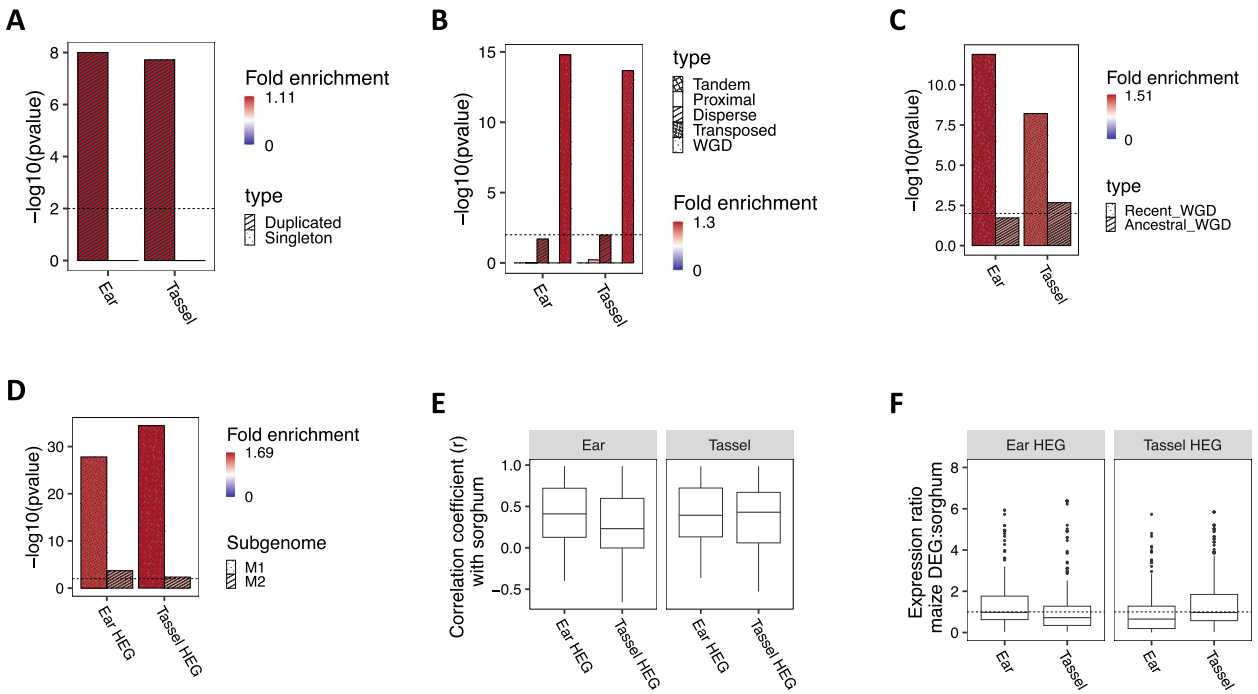

Fig. 3 The differentially expressed genes (DEGs) between the maize ear and the tassel are enriched with the recent whole-genome duplication (WGD). **A** The HEGs are enriched with duplicated genes. **B** The HEGs are enriched with duplicates from WGD. **C** Compared to the WGDs shared between maize and sorghum, the DEGs are enriched with the recent WGD specific to the maize genomes. **D** DEGs are significantly more enriched in the M1 subgenome of maize than in the M2 subgenome. A one-tailed hypergeometric test was used for the enrichment test without multiple testing adjustments. **E** The expression profile of ear HEGs in the ear snRNA clusters is significantly more similar to the sorghum than the expression profiles in the tassel snRNA clusters (mean r value of ear HEGs was 0.40 and tassel HEGs was 0.27, p-value < 0.001, t-test). The expression profile of tassel HEGs in the tassel snRNA clusters is more similar to the sorghum than the expression profiles in the ear snRNA clusters (mean r value of for tassel HEGs was 0.40 and for ear HEGs was 0.36, p-value = 0.09, t-test). **F** The ear HEGs in the ear have expression levels more similar to the sorghum than those in the tassel. The tassel HEGs in the tassel have expression levels that are more similar to those of the sorghum than in the ear

gene pairs resulting from different duplication events, gene pairs from the recent WGD have higher correlated expression profiles compared to other types of duplications (Figure S14), which fits the previous hypothesis that syntenic genomic block is an indication of regulation context and expression pattern conservation [53, 54]. The expression profiles of ear HEGs in the maize ear are in higher correlation with sorghum than in the tassel. While the expression profiles of tassel HEGs in maize tassel are in higher correlation with sorghum than that in the ear (Fig. 3E). We then investigated the gene expression dosage between these three tissues. Single-copy syntenic orthologous of the maize have a comparable dosage with sorghum (Figure S15). The dosage of ear HEGs in the ear has a ratio to the dosage in sorghum closer to 1, while the dosage of ear HEGs in the tassel is much lower than that in sorghum. Similarly, the dosage of tassel HEGs in tassel has a ratio to the dosage in sorghum closer to 1, while the dosage of tassel HEGs in the ear is much lower than that in sorghum (Fig. 3F). We assume that the tassel HEGs are functionally conserved in the tassel, while they partially lose function in the ear. And vice versa for ear HEGs. This result suggested the degradation of stamens in the ear is correlated with the decreased expression of a group of genes in the ear. The degradation of gynoecia in tassel is highly correlated with the decreased expression of a group of genes in tassel. Overall, the recent WGD is associated with the expression diversity between maize ear and tassel. Loss of gene expression profiling of one copy of WGD homology likely contributed to the differentiation between maize ear and tassel.

Tassel highly expressed genes are more conserved than ear highly expressed genes

The maize tassel is similar to the sorghum inflorescence by developing at the end of the shoot meristem and with branches. And the maize ear is likely an evolutionally new organ. To test this hypothesis, we compared the maize gene sequences with their sorghum orthologous. As expected, the similarity of CDS sequence of ear HEGs with their syntenic orthologous from the sorghum was significantly lower than that of tassel HEGs (Fig. 4 B). To

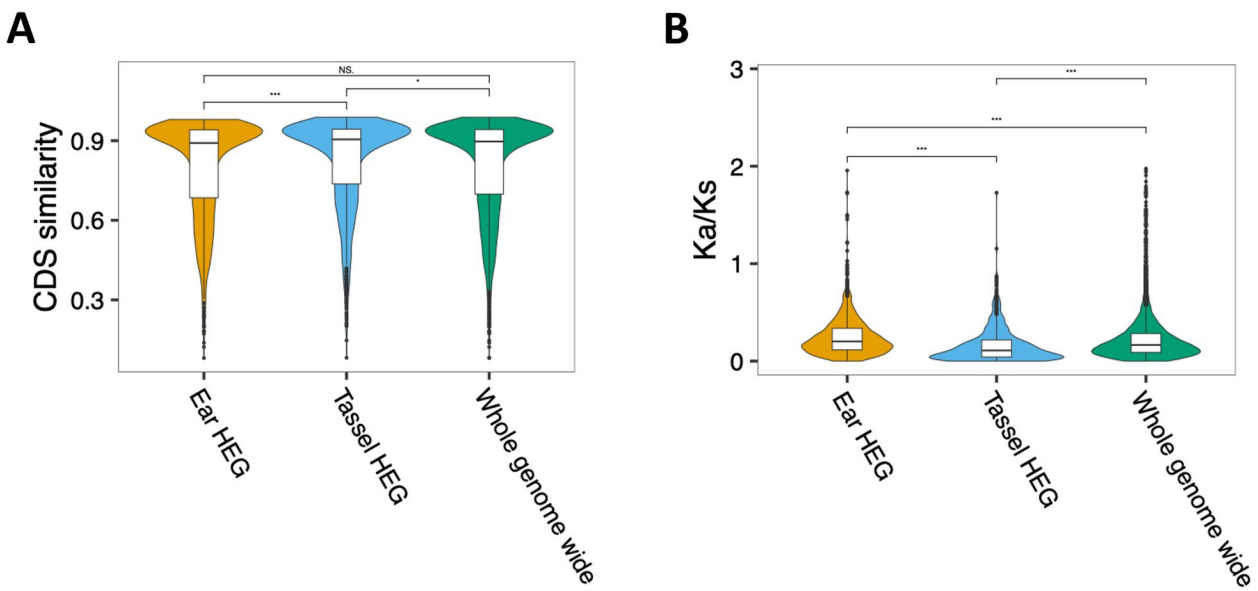

Fig. 4 Tassel highly expressed genes are more conserved than ear highly expressed genes. **A** The CDS similarities of Tassel HEG are significantly higher than ear HEGs. **B** Ear HEGs have significantly higher Ka/Ks ratios than tassel HEGs. *P*-values less than 0.0001 are shown as "****" via a Wilcoxon rank-sum test

further compare the evolution rate of ear HEGs and tassel HEGs, we calculated the lineage-specific Ka/Ks ratios for each maize gene using rice, sorghum and millet as outgroups. We found that ear HEG has significantly higher Ka/Ks ratios than whole genome wide background, while tassel HEG has significantly lower Ka/Ks values than whole genome wide background (Fig. 4A). Those results indicate that tassel HEGs are more conservative than ear HEGs (Fig. 4B).

Discussion

Overall, we provide a cellular atlas of the maize ear and tassel by combining single nuclei data with spatial transcriptomic data. Combining snRNA-seq with spatial transcriptomics allowed us to identify and annotate cell types for maize ear and tassel. An issue with plant single-cell transcriptome sequencing is the wide range of plant cell sizes. The size of animal cells is often between 10 μm and 30 μm, while plant cells can reach up to 100 μm [55–57]. Too large cells will not be captured and encapsulated into droplets [37]. Single nuclei reduce detection bias toward cell types that are easier to dissociate. Our maize ear single nuclei data supplements previously published single-cell data [40]. Different cell types play different biological roles in the development of plants and their ability to adapt to their surroundings [58]. The observation of "ts" suggests that, for maize, the evolution from the bisexual flower to the two monoecious flowers is more likely due to the evolution of gene expression regulation or gain of novel gene function while less likely due to gene absence. We identified cell types and their potential roles in inflorescence development. Our study and data provide important resources to learn the regulatory and evolution of maize inflorescence development.

We performed a meta-clustering to investigate the DEGs between maize tassel and ear at cell cluster resolution. By investigating DEGs between the ear and tassel, we can identify candidate genes related to maize ear and tassel differential development. The *YABBY* family transcription factors are highly expressed in the ear. Auxin and JA related genes are highly expressed in the tassel. Using the gene expression of sorghum as the outgroup, our analysis suggested that genes with lower expression levels in the ear are likely due to their diversified expression profile from the ancestry. A similar pattern was also observed for ear HEGs. This pattern is consistent with the developmental observations that both ear and tassel initiate as bisexual flowers, containing both male and female reproductive structures. The tassel's female components and the ear's male components gradually abort.

Conclusions

Our data and results provide novel insights into the evolution of unisexual flower development in maize. But please note the inflorescence of most grass plants on the top of the shoot. The molecular mechanisms that maize ear inflorescence could initialize laterally in the axils of leaves are still to be explored.

Abbreviations

WGD Whole genome duplication

SAM	Shoot apical meristem
IM	Inflorescence meristem
SPM	Spikelet-pair meristem
SM	Spikelet meristems
FM	Floral meristems
JA	Jasmonic acid
snRNA-seq	Single nucleus RNA sequencing
DEG	Differentially expressed gene

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11186-1.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Supplementary Material 7.

Acknowledgements
Not applicable.

Authors' contributions
Q. S. and B. S. designed the study and wrote the main manuscript. H. F., M. L. and J. H. performed the data analysis. W.F. and B.L. performed the molecular experiments.

Funding
This project is supported by the Shandong Provincial Natural Science Fund for Excellent Young Scientists Fund Program (Overseas) (2023HWYQ-109), Shandong Provincial Natural Science Foundation (ZR2023QC219) and Key R&D Program of Shandong Province(ZR202211070163).

Data availability
Sequencing data generated in this study have been deposited in the National Genomics Data Center with the primary accession code PRJCA023192 (https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA023192).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 21 September 2024 Accepted: 26 December 2024
Published online: 03 January 2025

References

1. Satterlee JW, Alonso D, Gramazio P, Jenike KM, He J, Arrones A, et al. Convergent evolution of plant prickles by repeated gene co-option over deep time. Science. 2024;385:1663.
2. Rudel D, Sommer RJ. The evolution of developmental mechanisms. Dev Biol. 2003;264:15–37.
3. Theissen G, Melzer R. Molecular mechanisms underlying origin and diversification of the angiosperm flower. Ann Bot. 2007;100:603–19.
4. Endress PK. Evolutionary diversification of the flowers in angiosperms. Am J Bot. 2011;98:370–96.
5. Swigonová Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, et al. Close split of sorghum and maize genome progenitors. Genome Res. 2004;14:1916–23.
6. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009;457:551–6.
7. Chakrabarty SK, Basu S, Schipprach W. Hybrid seed production technology. In: Seed Science and Technology. Singapore: Springer Nature Singapore; 2023. p. 173–212.
8. Chen Z, Gallavotti A. Improving architectural traits of maize inflorescences. Mol Breed. 2021;41:21.
9. Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, et al. Regulatory modules controlling maize inflorescence architecture. Genome Res. 2014;24:431–43.
10. Cheng P-C, Pareddy D. Morphology and development of the tassel and ear. In The Maize Handbook. Edited by: Freeling M, Walbot V. New York: Springer-Verlag; 1993;37–47.
11. Cheng PC, Greyson RI, Walden DB. Organ initiation and the development of unisexual flowers in the tassel and ear of Zea Mays. Am J Bot. 1983;70:450–62.
12. Smith CW, Frederiksen RA, editors. Sorghum: Origin, history, technology, and production. Nashville, TN: John Wiley & Sons; 2000.
13. Bortiri E, Hake S. Flowering and determinacy in maize. J Exp Bot. 2007;58:909–16.
14. Ortez OA, McMechan AJ, Hoegemeyer T, Ciampitti IA, Nielsen R, Thomison PR, et al. Abnormal ear development in corn: A review. Agron J. 2022;114:1168–83.
15. Nickerson NH, Dale EE. Tassel Modifications in Zea Mays. Ann Mo Bot Gard. 1955;42:195–211.
16. Acosta IF, Laparra H, Romero SP, Schmelz E, Hamberg M, Mottinger JP, et al. tasselseed1 is a lipoxygenase affecting jasmonic acid signaling in sex determination of maize. Science. 2009;323:262–5.
17. DeLong A, Calderon-Urrea A, Dellaporta SL. Sex determination gene TASSELSEED2 of maize encodes a short-chain alcohol dehydrogenase required for stage-specific floral organ abortion. Cell. 1993;74:757–68.
18. Lunde C, Kimberlin A, Leiboff S, Koo AJ, Hake S. Tasselseed5 overexpresses a wound-inducible enzyme, ZmCYP94B1, that affects jasmonate catabolism, sex determination, and plant architecture in maize. Commun Biol. 2019;2:114.
19. Chuck G, Meeley R, Irish E, Sakai H, Hake S. The maize tasselseed4 microRNA controls sex determination and meristem cell fate by targeting Tasselseed6/indeterminate spikelet1. Nat Genet. 2007;39:1517–21.
20. Laudencia-Chingcuanco D, Hake S. The indeterminate floral apex1 gene regulates meristem determinacy and identity in the maize inflorescence. Development. 2002;129:2629–38.
21. Jiao Y, Lee YK, Gladman N, Chopra R, Christensen SA, Regulski M, et al. MSD1 regulates pedicellate spikelet fertility in sorghum through the jasmonic acid pathway. Nat Commun. 2018;9:822.
22. Song X, Guo P, Xia K, Wang M, Liu Y, Chen L, et al. Spatial transcriptomics reveals light-induced chlorenchyma cells involved in promoting shoot regeneration in tomato callus. Proc Natl Acad Sci U S A. 2023;120:e2310163120.
23. Liu C, Leng J, Li Y, Ge T, Li J, Chen Y, et al. A spatiotemporal atlas of organogenesis in the development of orchid flowers. Nucleic Acids Res. 2022;50:9724–37.
24. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. 2019;8:329–337.e4.
25. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36:411–20.
26. Tang H, Lyons E, Pedersen B, Schnable JC, Paterson AH, Freeling M. Screening synteny blocks in pairwise genome comparisons through integer programming. BMC Bioinformatics. 2011;12:102.
27. Bradley D, Carpenter R, Sommer H, Hartley N, Coen E. Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plena locus of Antirrhinum. Cell. 1993;72:85–95.

28. Zhang Y, Ngu DW, Carvalho D, Liang Z, Qiu Y, Roston RL, et al. Differentially Regulated Orthologs in Sorghum and the Subgenomes of Maize. Plant Cell. 2017;29:1938–51.
29. Schnable JC, Springer NM, Freeling M. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci. 2011;108:4069–74.
30. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
31. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, et al. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol. 2019;20:38.
32. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020;36:2628–9.
33. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25:1422–3.
34. Kaplinsky NJ, Freeling M. Combinatorial control of meristem identity in maize inflorescences. Development. 2003;130:1149–58.
35. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049.
36. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–3587.e29.
37. Zhang Q, Kim SW, Gorham JM, DeLaughter DM, Ward T, Seidman CE, et al. Multiplexed Single-Nucleus RNA Sequencing Using Lipid-Oligo Barcodes. Curr Protoc. 2022;2:e579.
38. Wang Y, Luo Y, Guo X, Li Y, Yan J, Shao W, et al. A spatial transcriptome map of the developing maize ear. Nat Plants. 2024;10:815–27.
39. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177:1888–1902.e21.
40. Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, et al. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. Dev Cell. 2021;56:557–568.e6.
41. Marhava P, Bassukas AEL, Zourelidou M, Kolb M, Moret B, Fastner A, et al. A molecular rheostat adjusts auxin flux to promote root protophloem differentiation. Nature. 2018;558:297–300.
42. Dorrity MW, Alexandre CM, Hamm MO, Vigil A-L, Fields S, Queitsch C, et al. The regulatory landscape of Arabidopsis thaliana roots at single-cell resolution. Nat Commun. 2021;12:3334.
43. Vollbrecht E, Schmidt RJ. Development of the Inflorescences. In: Bennetzen JL, Hake SC, editors. Handbook of Maize: Its Biology. New York, NY: Springer New York; 2009. p. 13–40.
44. Strable J, Vollbrecht E. Maize YABBY genes drooping leaf1 and drooping leaf2 regulate floret development and floral meristem determinacy. Development. 2019;146:dev171181.
45. Zhang S, Tan F-Q, Chung C-H, Slavkovic F, Devani RS, Troadec C, et al. The control of carpel determinacy pathway leads to sex determination in cucurbits. Science. 2022;378:543–9.
46. Gladman N, Jiao Y, Lee YK, Zhang L, Chopra R, Regulski M, et al. Fertility of Pedicellate Spikelets in Sorghum Is Controlled by a Jasmonic Acid Regulatory Module. Int J Mol Sci. 2019;20:4951.
47. Nickerson NH. Sustained Treatment with Gibberellic Acid of Five Different Kinds of Maize. Ann Mo Bot Gard. 1959;46:19–37.
48. Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, et al. Cloning and characterization of the maize An1 gene. Plant Cell. 1995;7:75–84.
49. Galli M, Liu Q, Moss BL, Malcomber S, Li W, Gaines C, et al. Auxin signaling modules regulate maize inflorescence architecture. Proc Natl Acad Sci U S A. 2015;112:13372–7.
50. McSteen P, Malcomber S, Skirpan A, Lunde C, Wu X, Kellogg E, et al. barren inflorescence2 Encodes a co-ortholog of the PINOID serine/threonine kinase and is required for organogenesis during inflorescence and vegetative development in maize. Plant Physiol. 2007;144:1000–11.
51. McSteen P, Hake S. barren inflorescence2 regulates axillary meristem development in the maize inflorescence. Development. 2001;128:2881–91.
52. Flagel LE, Wendel JF. Gene duplication and evolutionary novelty in plants. New Phytol. 2009;183:557–64.
53. Song B, Buckler ES, Wang H, Wu Y, Rees E, Kellogg EA, et al. Conserved noncoding sequences provide insights into regulatory sequence and loss of gene expression in maize. Genome Res. 2021. https://doi.org/10.1101/gr.266528.120.
54. Lai X, Behera S, Liang Z, Lu Y, Deogun JS, Schnable JC. STAG-CNS: An Order-Aware Conserved Noncoding Sequences Discovery Tool for Arbitrary Numbers of Species. Mol Plant. 2017;10:990–9.
55. Niklas KJ. Size-dependent variations in plant growth rates and the "¾ power rule." Am J Bot. 1994;81:134.
56. Tárnok A. Visiting the plant kingdom. Cytometry A. 2009;75:973–4.
57. Liu S, Ginzberg MB, Patel N, Hild M, Leung B, Li Z, et al. Size uniformity of animal cells is actively maintained by a p38 MAPK-dependent regulation of G1-length. Elife. 2018;7:e26947.
58. Zeng H. What is a cell type and how to define it? Cell. 2022;185:2739–55.

Publisher's Note