Accepted Manuscript

Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development

Hongjun Liu, Xuerong Yang, Xinhui liao, Tao Zuo, Cheng Qin, Shiliang Cao, Ling Dong, Huangkai Zhou, Yongzhong Zhang, Sisi Liu, Yaou Shen, Haijian Lin, Thomas Lübberstedt, Zhiming Zhang, Guangtang Pan

PII: S0888-7543(15)00053-1
DOI: doi: 10.1016/j.ygeno.2015.03.005
Reference: YGENO 8728

To appear in: Genomics

Received date: 21 February 2015
Accepted date: 26 March 2015

Please cite this article as: Hongjun Liu, Xuerong Yang, Xinhui liao, Tao Zuo, Cheng Qin, Shiliang Cao, Ling Dong, Huangkai Zhou, Yongzhong Zhang, Sisi Liu, Yaou Shen, Haijian Lin, Thomas Lübberstedt, Zhiming Zhang, Guangtang Pan, Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development, Genomics (2015), doi: 10.1016/j.ygeno.2015.03.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development

Hongjun Liu¹, Xuerong Yang², Xinhui liao³, Tao Zuo⁴, Cheng Qin¹,⁵, Shiliang Cao⁶, Ling Dong⁷, Huangkai Zhou³, Yongzhong Zhang¹, Sisi Liu¹, Yaou Shen¹, Haijian Lin¹, Thomas Lübberstedt⁸, Zhiming Zhang¹*, Guangtang Pan¹*

Affiliations:
¹Maize Research Institute of Sichuan Agricultural University, Chengdu, 611130, China.
²Institute of Animal Nutrition, Sichuan Agricultural University, Ya’an 625014, China
³BGI-Shenzhen, Shenzhen, 518083, China
⁴Interdepartmental genetics program, Iowa State University, Ames, 50011, USA
⁵Zunyi Academy of Agricultural Sciences, Zunyi, 563102, China.
⁶Maize Research Institute of Heilongjiang Academy of Agricultural Sciences, 150086, Harbin, China
⁷Department of Agronomy, Northeast Agricultural University, 150030, Harbin, China
⁸Department of Agronomy, Iowa State University, Ames, 50011, USA

*Corresponding author
E-mail: pangt@sicau.edu.cn or zzmmaize@gmail.com

¶These authors contributed equally to this work.
Abstract

The present study profiled and analyzed gene expression of the maize ear at four key developmental stages. Based on genome-wide profile analysis, we detected differential mRNA of maize genes. Some of the differentially expressed genes (DEGs) were predicted to be potential candidates of maize ear development. Several well-known genes were found with reported mutants analyses, such as, *compact plant2* (*ct2*), *zea AGAMOUS homolog1* (*zag1*), *bearded ear* (*bde*), and *silky1* (*si1*). MicroRNAs such as *microRNA156* were predicted to target genes involved in maize ear development. Antisense transcripts were widespread throughout all the four stages, and are suspected to play important roles in maize ear development. Thus, identification and characterization of important genes and regulators at all the four developmental stages will contribute to an improved understanding of the molecular mechanisms responsible for maize ear development.

**Keywords:** maize; DGE; ear development; microRNA156; transcriptional factor
1. Introduction

Maize (Zea mays) is one of the most important crops and widely used model plant. Inflorescence and flower development are critically important for high yields in maize. Maize ears require a low concentration of nitrogen, thus making it more efficient and aiding in a sustainable production of crop without adding more fertilizer to the soil after harvest [1]. Various mutants have been discovered, providing insights into the molecular processes involved in the ear development [2-7]. However, understanding of the maize ear developmental dynamics at the transcriptome level is limited. Till now, only few studies have been conducted on the large-scale gene expression analyses of the maize ear, including, (i) evaluation of sequence-based expression profiles during reproductive organ development [8], (ii) study on the effect of water-deficiency on immature maize ear development [9], and (iii) discovery of novel microRNAs during maize ear development [10].

The B73 sequence assembly [11] enables analysis of maize ear development at a genome-wide transcriptome level. Owing to the dramatic decrease in the cost of sequencing and development of rapid and robust experimental procedures, it is now feasible to conduct a cost-efficient high-throughput profile analysis. For instance, by using digital gene expression (DGE) [12-15] and RNA-Seq [16-18] analyses, new genes have been discovered [19]. Furthermore, these technologies are useful for estimating the overall gene expression at different developmental stages or in different tissues [12,20], and in response to abiotic stresses [21,22]. Considering the significance of ears in maize production, it is of great importance to understand the molecular mechanisms involved in the maize ear development.
The objective of this study was to conduct a genome-wide comparative analysis of gene expression profiles to obtain an improved understanding of the molecular mechanisms of maize ear development during four developmental stages; the growth point elongation (I), spikelet differentiation (II), floret primordium differentiation (III), and floret organ differentiation phase (IV) [23] using a DGE approach. Ears of maize from all the four developmental stages were used to study the dynamics of mRNA expression. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed using randomly selected DEGs, in order to validate their expressions across different developmental stages. The K-means clustering method was employed to further determine the co-expression of genes involved in the maize ear development.

2. Results

2.1. Sequence alignment and expression analysis

Library construction and sequence analysis were conducted [20]. Around 16.8 million high-quality raw reads were generated. After performing quality-control measures, 16.1 million clean tags were obtained for all the four stages (S1 and S2 Tables). Briefly, after removing low-quality and contaminating reads, clean tags were retained for further analysis. Subsequently, the 16.1 million clean tags were aligned against the maize genome (B73 RefGen_v2). The percentage of clean tags in the raw data for each developmental stage was 93.62%, 96.81%, 96.54%, and 96.56%, respectively. About 69.86% clean tags were mapped to the B73 reference genome with an average of 76.10% genes covered. Incompleteness of maize genome sequence data was probably one of the reasons for the occurrence of unmapped tags. Most tags were aligned to genic regions and the genic distribution of reads from mRNA reference sequences in all the
four developmental stages (I-IV) showed that, a majority of tags (87.58%, 87.23%, 90.30%, and 90.95%, respectively) were mapped to exon regions and the remaining were distributed within introns, intergenic regions and repeat regions (S1 Fig.).

Majority of transcripts were expressed in all the four stages (Fig. 1A and S3 Table). The numbers of sense (Fig. 1A) and antisense (Fig. 1B) transcripts overlapping at all 4 stages were 11,970 and 4416, with a cutoff for gene expression at each stage of one tag per million (at least 4 reads). The number of genes that showed both sense and antisense expressions were 7230, 7052, 6918, and 6571 (Fig. 1C) for each developmental stage, respectively, and 10,456 in all stages in both sense and antisense expression. Of all the sense genes detected, only 74 genes were expressed uniquely in stage I, and was even lower than that of the other 3 developmental stages, suggesting the involvement of more genes in the maize ear development.

2.2. Analysis of differentially expressed genes and validation by qRT-PCR

Based on a cutoff of at least 4 reads per gene, all reads that were mapped to genes were used for differential expression analysis combined with the DGE method for a genome-wide comparative analysis of data for all the 4 developmental stages. Comparative gene expression analyses were used for estimation of gene expression levels in all the four developmental stages (S4 Table). We calculated the number of tags corresponding to each gene in each library to estimate the gene expression levels and compare the difference in fold-change between the developmental stages [20]. Transcripts that showed differential expression levels are shown in S2 Fig. The up-regulated and down-regulated genes indicate the DEGs (Fig. 2, S2 Fig., and S5 Table). In total, the number of DEGs between two stages were as follows: 3325 between stages I and II (36% up- and 64%
down-regulated in stage II), 4735 between stages I and III (57% up- and 43% down-regulated in stages III), 6398 between stages I and IV (46% up- and 54% down-regulated in stage IV), 3765 between stages II and III (71% up- and 29% down-regulated in stage III), 5178 between stages II and IV (60% up- and 40% down-regulated in stage IV), and 1698 between stages III and IV (35% up- and 65% down-regulated in stage IV).

To better understand the dynamic changes of gene expression in maize ear development during all the four developmental stages, further analyses of the DEGs were performed, especially of those genes in which up- or down-regulation gradually follow ear development (II vs. I, III vs. II, and IV vs. III; S5 Table). Among the DEGs identified, 1201, 2690, and 594 genes were up-regulated in stages II, III, and IV, respectively, compared with their own preceding stage. In contrast, the numbers of down-regulated genes were 2124, 1075, and 1104 in stages II, I, and IV, respectively (Fig. 2 and S5 Table). During the adjacent developmental stages, nearly two third of DEGs were up-regulated in developmental stage III (Fig. 2) vs. stage II or stage I. This suggests that DEGs were more abundant in stages III, indicating an active ear development during stage III (floret primordium differentiation phase). Furthermore, the expression patterns of 9 DEGs are illustrated in S3 Fig. Interestingly, we found some well-known genes with reported mutants analyses during maize inflorescence development, such as, compact plant2 (ct2), zea AGAMOUS homolog1 (zag1), bearded ear (bde), and silky1 (si1) [24].

Furthermore, homolog of ATFP3, an important factor binding to transition metal ions, was the most significant and annotated up-regulated gene (GRMZM2G155281, 2.72 log2 fold change, q-value < 0.001, S5 Table) in stage II than compared with stage I. *Pyrophosphatase (PPase)*, adapted to phosphate starvation, was significantly down-
regulated (1.27 log2 fold change, \(q\)-value < 0.001) in stage II as in contrast with stage I, whereas it was up-regulated (3.18 log2 fold change, \(q\)-value < 0.001) in stage III vs. stage II.

To confirm the expression patterns determined by RNA-sequencing (RNA_seq) analysis, we used qRT-PCR analyses to analyze the expression of 32 randomly selected genes (Fig. 3). Although the log2-fold values of the 32 genes showed slight variations as compared to the corresponding values from the qRT-PCR analyses, the expression data from the RNA-Seq analysis were closely positively correlated (most Pearson correlation coefficients were higher than 0.8) with those obtained from qRT-PCR (Fig. 3 and S4 Fig.), indicating the deep sequencing and RT-PCR data were well consistent.

2.3. Pathway enrichment analysis of DEGs using KEGG and global analysis of the gene expression profiles

To better understand the functional roles of DEGs responsible for the maize ear development, pathway enrichment analysis was performed to investigate the biological pathways of DEGs in all the four developmental stages. It was revealed that 3 pathways were affected by the up-regulated DEGs and 6 by the down-regulated DEGs (Table 1). As compared to stage I, ribosome pathway was significantly over-represented in up-regulated DEGs from stage II, and also in stage III vs. stage II samples. However, we did not found any significantly (\(q\) value < 0.05) pathways down-regulated for Stage III vs. Stage II and up-regulated for Stage IV vs. Stage III. Interestingly, spliceosome pathway, protein processing in the endoplasmic reticulum, and ubiquitin mediated proteolysis pathways were significantly over-expressed in the down-regulated DEGs of stage II vs. stage I, indicating more transcriptional and post-transcriptional events in stage II.
Furthermore, RNA transport pathway was significantly affected by the up-regulated DEGs in stage III vs. stage II. Photosynthesis-antenna proteins and photosynthesis pathways were over-expressed in the down-regulated DEGs in stage IV (vs. stage III). These results suggest that the ear development is most active during stage III, and in stage II, the active post-transcriptional events indicates the preparation for photosynthesis in the upcoming stage. This result is also in accordance with the high numbers of DEGs in stages II and III (Fig. 2).

We identified ~6800 out of 25,800 genes that were differentially expressed among developmental stages, representing 26% of the ear transcriptome. Using MapMan [25], we identified 18 clusters (K1 to K18; Fig. 4 and S6 Table). Most clusters showed significant enrichment for particular GO terms and KEGG pathways (S5 and S6 Fig.). To further understand the functional roles of genes in different clusters in a better manner, we looked for some particular clusters related to ear development. Of which, clusters K1 and K7 displayed similar expression patterns, consistent with the results of the GO enrichment analyses, across all the four developmental stages. Most over-represented GO terms such as, chromatin organization, organelle organization, and cellular component organization (S5 Fig), etc. were shared between clusters K1 and K7. In addition, ribosomal pathway was also over-represented in both the clusters (K1 and K7). Interestingly, the carboxy-lyase (GRMZM2G159149) in cluster K1 was significantly up-regulated in stage III, and significantly down-regulated in stages II and IV (S5 Table), indicating an important role of carboxy-lyase in the floret primordium differentiation phase during maize ear development.

2.4. Resolving transcription factors among differentially expressed genes
The primary objective was to identify genes that encoded transcription factors (TFs) and resolve the dynamics of TF accumulation during ear development. We retrieved putative orthologs of maize genes based on information from the Ensembl Compara gene trees [26] in Gramene (http://maizesequence.org), PlantGDB (http://plantgdb.org), and the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). We then queried known plant TFs in the Plant Transcription Factor Database (v2.0, http://planttfdb.cbi.edu.cn/) and maize transcription factor database (grassius.org). 1,522 Maize TFs with sequence similarities to known TFs that were expressed during at least one developmental stage were identified (S7 Table). In total, 242 TFs were differentially expressed between stages II vs. I, 345 between III vs. I, 445 between IV vs. I, 239 between III vs. II, 358 between IV vs. II, and 132 between IV vs. III (S8 Table). 13 out of 469 TFs from 10 TF families (B3, bZIP, ERF, HD-ZIP, LBD, MIKC, MYB, NAC, SBP, and TCP) were differentially expressed across comparisons of adjacent stages during all the four developmental stages of maize ear development (Table 3).

We also investigated the differentially expressed TFs along a wide range of abundances in stage II and a mutant in RAMOSA3 (RA3) gene (regulating the determinacy of axillary meristems), which has been reported previously in a study, on the developmental stages of maize [12]. The results are listed in S9 Table. As shown in the table, a total of 19 differentially expressed putative TFs were identified in our dataset, and almost all exhibited diverse expression patterns from ra3 mutants, which further illustrated that our dataset was highly correlated with the previous study [12]. The identified TFs were also differentially expressed across a wide range of abundances in ra3 mutant. Among the identified TFs, there were many members with function in
development and meristem maintenance (NAC, TCP, YABBY), and others in ethylene and stress response (ERF). Therefore, these TFs possibly contribute to the genetic control of maize ear development.

SBP-box genes, which encode a class of zinc finger-containing TFs, are important regulators with various functions during maize development [27-30]. In this study, SBP-domain protein 5 (GRMZM2G160917, 1.72 log2 fold change for III vs. II, 1.85 log2 fold change for II vs. I, q-value < 0.001) and SBP-domain protein 6 (GRMZM2G307588, 1.37 log2 fold change, q-value < 0.001) were significantly up-regulated in stages III & II, and stage IV compared to its previously adjacent stages, respectively (Fig. 5, S8 Table). Interestingly, both genes were predicted to be targets of microRNA156, which plays an important role in the maize ear development [10]. The obtained results indicate the potential roles of microRNA156 and SBP in maize ear development. As a TF, SBP (GRMZM2G109354) displayed significantly differential expression patterns during different developmental stages (Fig. 5C), suggesting a possibility that SBP may have different regulatory roles in each stage of maize ear development.

2.5. Antisense transcripts detected for many genes

Antisense transcripts have been identified and predicted from maize [31]. Antisense transcription is a common phenomenon in maize (Fig. 1B and 1C), widespread at all the four stages (S3 Table). S7 Fig. shows the relationship between sense and antisense transcripts at each developmental stage, respectively. The correlation coefficient (r) between sense and antisense transcripts were moderately positive at each stage, and highest at stage I (r=0.50). Interestingly, a large number of antisense transcripts were uniquely expressed in stage I (1232 genes), whereas, only few genes (692) were
expressed in the antisense direction in the other stages (Fig. 1C). The distribution pattern of antisense transcripts across all the four developmental stages was similar to the overall transcript distribution (both antisense and sense, Fig. 1C).

Annotation with GO terms was performed for the antisense transcripts that were detected at all the four developmental stages (S3 Table). More unique antisense transcripts were detected in stage I. The obtained unique antisense transcripts in stage I were assigned to 41 GO terms for cellular component such as mitochondria (GO:0005739), 144 for molecular function such as oxidoreductase activity (GO:0016491), and 109 for biological process such as protein catabolic process (GO:0030163).

3. Discussion

In this study, ~6800 DEGs were identified across various comparisons of developmental stages during maize ear development. Previous studies have suggested that the expression and function of the floral homeotic gene AGAMOUS (AG) were responsible for the normal floral development and floral organ identity transformation [32,33]. In the present study, the homolog of the Arabidopsis gene AGAMOUS (GRMZM2G052890) (9.21 log2 fold change, q-value < 0.001) was the most significantly up-regulated and annotated gene in stage III compared to stage II, which suggests an important role of that gene in ear development during stage III.

Among the DEGs validated by qRT-PCR, the gene GRMZM2G007025 displayed differential expression between stages III and IV. During stage III, GRMZM2G007025 was down-regulated (2.377 log2 fold change, q-value < 0.001) in contrast to stage II. It was up-regulated (2.382 log2 fold change, q-value < 0.001) during stage IV compared to
stage III. Interestingly, the mRNA level of GRMZM2G007025 was higher in stage IV than that of stage III (Fig. 3), which was consistent with our DGE results. This indicates that GRMZM2G007025 may be a candidate gene for better understanding the mechanism of maize ear development.

Metal ions such as zinc, copper, and iron are essential for plant growth. Dykema et al. (1999) [34] characterized ATFP3 (Arabidopsis thaliana farnesylated protein 3) as an important factor that binds to transition metal ions. In the current study, ATFP3 homolog was the most significant and annotated up-regulated gene in stage II, which indicated that additional metal ions might be required for maize ear development in stage II. In addition, soluble inorganic PPase (GRMZM2G104918) plays an important role in the adaptation of Phaseolus vulgaris to phosphate starvation [35]. Different expression patterns were found for PPase during developmental stages II and III. PPase might thus be a candidate gene for prediction of the physiological signal of phosphorus during maize ear development.

In plants, microRNAs play important regulatory roles in many aspects of plant biology, including metabolism, growth, and stress response [10]. Liu et al. reported that zma-miRNA156 was differentially expressed during maize ear development. In accordance with the previous study, 19 DEGs were predicted to be the targets of microRNA156 and microRNA319 (stage II vs. I), microRNA156, 160, 164, 167, 390, and 394 (stage III vs. II), and microRNA156, 160, 319, and 529 (stage IV vs. III) during the 4 developmental stages (Table 2). Hultquist and Dorweiler (2008) reported that decreased levels of microRNA156 might provide competency for SBP-box gene up-regulation in Tassel [29]. They also indicated that SBP-5 & SBP-6 appeared to be more
abundant in maize ear. SBP-5 & SBP-6 had a greater abundance in developing maize ear (Fig. 5). Thus, our result is in accordance with the previous study [29].

To further investigate the co-expression pattern of DEGs involved in maize ear development, we identified several clusters that were uniquely significantly enriched according to GO terms, such as, cellular aromatic compound metabolic process and microtubule-based process, in clusters K5 and K13 (S5 Fig.), respectively. These two clusters were also significantly involved in KEGG pathways; K5 was involved in phenylalanine, tyrosine, and tryptophan biosynthesis of secondary metabolites and K13 in proteasome pathway (S6 Fig.). 16 Genes were significantly differentially expressed in cluster K13 across all the four developmental stages (S6 Table), of which, most of the genes showed a unique expression pattern (either up- or down-regulated) during stage III.

In plants, protein phosphorylation is an important regulatory mechanism. Histidine-containing phosphotransfer proteins (HPts, GRMZM2G016439) participate in hormone signal transduction in higher plants [36]. These genes were significantly down-regulated in stage III vs. II, but up-regulated during stage II vs. I and stage IV vs. III, which suggested that phosphorylation and hormone signal transduction were more active in stages II and IV, and HPts might be a potential marker of phosphorylation status during maize ear development. However, this suggestion needs to be further investigated. Moreover, maize proteinase inhibitor (GRMZM2G028656) in cluster K2 was significantly down-regulated across all of the three comparisons, especially during stage IV (8.79 log2 fold change in stage IV vs. III, 1.13 log2 fold change in stage III vs. II, and 1.47 log2 fold change in stage II vs. I). Plant proteinase inhibitors play an important role in the insect resistance of transgenic plants [37]. The maize proteinase inhibitor was
significantly annotated in the serine-type endopeptidase inhibitor activity (GO:0004867), peptidase activity (GO:0008233), and response to wounding (GO:0009611).

To better explain the molecular mechanisms of maize ear development, 13 genes encoding transcriptional factors were identified across all the four stages. The B3 DNA-binding domain, a plant-specific domain, is found in all the flowering plants [38]. In the present study, B3 DNA binding domain-containing protein (GRMZM2G065496) was significantly up-regulated in stage III (vs. II, 1.51 log2 fold change, \(q\)-value < 0.001), and down-regulated in stage II (vs. I, 1.08 log2 fold change, \(q\)-value < 0.001), and stage IV (vs. III, 1.005 log2 fold change, \(q\)-value < 0.001), which indicates that the B3 DNA binding domain-containing protein may have an important role in the development of floret primordium differentiation (stage III).

Antisense transcripts have been reported in various biological processes, including translation regulation and RNA interference [31]. The antisense transcripts detected in this study may play important roles in maize ear development.

4. Conclusions

In this study, we undertook a comprehensive transcriptome analysis and identification of DEGs during maize ear development using an Illumina sequencing platform. In total, we identified 3325 genes, that were differentially expressed during the spikelet differentiation phase, 3765 genes during the floret primordium differentiation phase, and 1698 genes differentially expressed during the floret organ differentiation phase, compared to their previously adjacent stages, respectively. Some of the DEGs, such as the maize homologs of AGAMOUS (GRMZM2G052890) and ATFP3 (GRMZM2G155281) were predicted to be potential candidates of maize ear
development. We also found some well-known genes with reported mutants analyses during maize inflorescence development, such as, compact plant2 (ct2), *zea AGAMOUS homolog1* (zag1), bearded ear (bde), and silky1 (si1). In accordance with our previous study [10], several DEGs were predicted to be targets of microRNAs. MicroRNA156 appears to be a key microRNA in maize ear development. Predicted targets of microRNA156, the SBP-box genes, such as *SBP-5* and *SBP-6* were identified to play significantly important roles in different stages of maize ear development. K-means clustering revealed 18 major expression patterns. From the analysis of TFs, we also identified 13 TFs from 10 TF families (B3, bZIP, ERF, HD-ZIP, LBD, MIKC, MYB, NAC, SBP, and TCP) that were differentially expressed along 3 adjacent comparisons (II vs. I, III vs. II, and IV vs. III) of four developmental stages of maize ear development. Antisense transcripts were widespread in all the four stages, and especially in stage I, there were a large number of antisense transcripts. The antisense transcripts detected in this study may play important roles in early stage of maize ear development. Understanding the maize ear development is critical for improvement of maize production. Thus, identification and characterization of important genes and regulators at all the four developmental stages will contribute to an improved understanding of the molecular mechanisms responsible for maize ear development.

5. Materials and Methods

5.1. Plant cultivation and sample collection

Seeds of the maize inbred line 18-599 (Maize Research Institute, Sichuan Agricultural University, Chengdu, China) were grown in a growth chamber at 24°C/18°C (day/night) with 12 h illumination per day. Ears were collected as described previously [10,23] at
four developmental stages: the growth point elongation, spikelet differentiation, floret primordium differentiation, and the floret organ differentiation phases. In brief, ears were manually collected. All the samples were harvested and immediately frozen in liquid nitrogen, and stored at -80°C until used for RNA isolation.

5.2. Digital expression library preparation and sequencing

Total RNA from the maize inflorescences at each developmental stage were isolated using TRIzol (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer’s instructions. For RNA library construction and deep sequencing, equal quantities of RNA were pooled for each developmental stage. Approximately 6 µg of total RNA representing each library were sequenced using Illumina HiSeq™ 2000 System as described [15]. The DGE libraries were constructed using Illumina Gene Expression Sample Prep Kit according to the manufacturer’s instructions. Briefly, mRNA were isolated and purified from total RNA using Oligo (dT) magnetic beads, and synthesized to be the first and second-strand cDNA. Then the bead-bound cDNA was subsequently digested with restriction enzyme NlaIII, and two adaptors were ligated to the 5’ and 3’ ends of the tags, respectively. After 15 cycles of linear PCR amplification, 95bp fragments were purified. Then the single-chain molecules were fixed onto the Illumina Sequencing Chip (flowcell) and sequenced with the method of sequencing by synthesis. Finally, millions of raw reads of each library were generated with sequencing length of 35 bp.

5.3. Sequencing data and Differentially expressed gene analysis

Raw data were filtered to remove the adaptors, low quality tags, and tags with one copy number. Then clean reads were used for further analysis. In brief, clean reads were aligned to the maize reference genome [11] (B73 RefGen_v2) using SOAP2 software
[39], allowing only 1bp mismatch. Only the unique mapping tags (tags only mapped to unambiguous unique location of B73 genome) were retained for calculating the gene expression, and the unambiguous tags were annotated. To analyze the gene expression, the number of clean reads for each library was counted and then normalized to tags (reads) per million. To detect DEGs, statistical analyses among libraries were performed following the formula as described [20], where false discovery rate (FDR) was used to determine the threshold of the P value in multiple test and statistical analysis using the q-value package [40]. Significance of differential gene expression was determined at a false discovery rate < 0.001 and an absolute value of log2-ratio $\geq 1$.

5.4. Global analysis of differentially expressed genes

To further annotate and analyze the DEGs, a user-driven tool MapMan [25] was used to assign DEGs to functional categories of metabolic pathways or other processes, and the genes were grouped by developmental dynamics, using the gap statistic algorithm[41] to select the number of clusters (K, S8 Fig.) and K-means clustering algorithm as described [42]. To further identify the significantly enriched metabolic pathways in DEGs, Gene Ontology and enrichment analyses with KEGG annotation were performed using the genes within each cluster, where the formula used in this study is as follows [20]:

$$P = 1 - \sum_{i=0}^{m-1} \binom{M}{i} \frac{\binom{N-M}{n-i}}{\binom{N}{n}}$$
In this formula, N indicates the number of DEGs, n indicates the number of genes within each cluster in N. M indicates the number of the DEGs with specific GO/KEGG annotations, and m indicates the number of genes within each cluster in M.

5.5. Quantitative real-time PCR analysis

To validate the DEGs, quantitative real-time PCR was performed for 32 randomly selected DEGs, which were the same as those used for the DGE genome-wide comparative analysis. Real-time PCR was performed using the SYBR Premix Ex Taq™ protocol (TaKaRa Biotechnology, Dalian, China) on an Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). For each sample, measurements were performed in triplicate, with three biological replicates, and the average cycle thresholds (Ct) were used to determine fold-change. *18S rRNA* (forward primer: 5’-ATGTTCCGTGGCAAGATGAG-3’, reverse primer: 5’-CATTGTTGGGAATCCACTC-3’) was employed as an endogenous control. Primers were designed using the Primer Premier 5.0 (http://www.premierbiosoft.com/index.html) and Oligo 6 programs (http://www.oligo.net) (S10 Table). Thermal cycle conditions were as follows: 2 min at 95°C followed by 40 cycles of 15 s at 95°C, 15 s at 56-57°C, and 15 s at 72 °C. Statistical analysis was conducted using the $2^{-\Delta\Delta Ct}$ method as described previously [43].

5.6. Data access

RNA sequencing data have been deposited at NCBI under the accession number GSE49805.

Acknowledgments
This work was supported by the grants from the Major State Basic Research Development Program of China (973 Program) (No. 2014CB138200, http://www.973.gov.cn/AreaAppl.aspx), National Natural Science Foundation of China (31201221, http://www.nsfc.gov.cn), the Research Project on Rice Functional Genes related to the Maize Large-scale Gene Discovery and Functional Genomics Resource, Technology, Information Platform Construction Program (863) of China (SS2012AA100107, 2012AA10A300, http://www.863.gov.cn), and the Major Project of China on New varieties of GMO Cultivation (2014ZX08003-003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Reference


Figure Legends

Figure 1. **Comparison of four development stages of maize ear.** Comparison of genes expressed in sense (A) and antisense (B) directions in the four development stages. Overlaps show the number of genes shared between stages. (C) A Venn diagram shows the genes expressed in sense and antisense direction in each developmental stage, and all stages.

Figure 2. **Genes differentially expressed during various comparisons between developmental stages.** Data were shown as percentages (y-axis) and numbers (x axis) of differentially expressed genes within each comparison, respectively, with the color indicated genes up- (red) and down-regulated (green), respectively.

Figure 3. **Results of gene expression validated by quantitative real-time PCR analysis.** Genes were randomly selected based on sequencing results. For normalization, 18S rRNA was used as the internal control. The relative gene expression for each gene were normalized using the $2^{-\Delta \Delta Ct}$ method as described in Materials and Methods.

Figure 4. **Dynamic progression of maize ear transcriptome.** Dynamic progression of ear transcriptome. 18 clusters were identified along the four developmental stages from 6,800 differentially expressed genes. The 18 clusters are presented.

Figure 5. **miR156 and its predicted targets SBP-box genes in maize ear development (Adapted from Liu et al. (2014)).** (A) The development of the maize ear regulated by miR156 and its predicted targets SBP-box genes, SBP-5 and SBP-6. Both of SBP-5 and SBP-6 were found to be targets of miR156, during all four stages of maize ear development (Liu et al. 2014). (B) The expression pattern of SBP-5 and SBP-6. During ear development. SBP-5 was significantly up-regulated in stage II (spikelet
differentiation phase) and III (floret primordium differentiation phase), while SBP-6 was up-regulated instead of decreased expression of SBP-5 in stage IV (floret organ differentiation). (C) The expression pattern of SBP (GRMZM2G109354). As a transcriptional factor, SBP (GRMZM2G109354) is also found to be significantly up- or down-regulated during ear development, suggesting a potential regulatory role of SBP in the maize ear development. * indicated the significant difference in adjacent two developmental stages (p < 0.01, n = 3).
### Tables

**Table 1. List of over-represented pathways for DEGs**

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Pathway ID</th>
<th>DEGs tested</th>
<th>p value</th>
<th>q value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage II vs. Stage I (up-regulated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribosome</td>
<td>ko03010</td>
<td>43 (6.61%)</td>
<td>4.54E-05</td>
<td>4.86E-03</td>
</tr>
<tr>
<td><strong>Stage II vs. Stage I (down-regulated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural killer cell mediated cytotoxicity</td>
<td>ko04650</td>
<td>14 (1.25%)</td>
<td>6.09E-05</td>
<td>7.06E-03</td>
</tr>
<tr>
<td>Spliceosome</td>
<td>ko03040</td>
<td>47 (4.2%)</td>
<td>1.28E-04</td>
<td>7.42E-03</td>
</tr>
<tr>
<td>Protein processing in endoplasmic reticulum</td>
<td>ko04141</td>
<td>51 (4.56%)</td>
<td>5.80E-04</td>
<td>2.24E-02</td>
</tr>
<tr>
<td>Ubiquitin mediated proteolysis</td>
<td>ko04120</td>
<td>36 (3.22%)</td>
<td>1.25E-03</td>
<td>3.64E-02</td>
</tr>
<tr>
<td><strong>Stage III vs. Stage II (up-regulated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA transport</td>
<td>ko03013</td>
<td>62 (4.1%)</td>
<td>4.03E-04</td>
<td>2.73E-02</td>
</tr>
<tr>
<td>Ribosome</td>
<td>ko03010</td>
<td>77 (5.09%)</td>
<td>4.66E-04</td>
<td>2.73E-02</td>
</tr>
<tr>
<td><strong>Stage IV vs. Stage III (down-regulated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photosynthesis - antenna proteins</td>
<td>ko00196</td>
<td>6 (1.06%)</td>
<td>1.75E-04</td>
<td>1.09E-02</td>
</tr>
<tr>
<td>Photosynthesis</td>
<td>ko00195</td>
<td>14 (2.48%)</td>
<td>1.94E-04</td>
<td>1.09E-02</td>
</tr>
</tbody>
</table>

^Comparisons are presented as the current stage minus the previously adjacent stage.
Table 2. Predicted differentially expressed genes targeted by microRNA found by Liu et al. (2014).

<table>
<thead>
<tr>
<th>GeneID</th>
<th>microRNA</th>
<th>log2FC</th>
<th>q value</th>
<th>Gene Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage II vs. stage I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRMZM2G067624 miR156</td>
<td>1.97</td>
<td>0.00E+00</td>
<td>unknown [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G163813 miR156</td>
<td>-2.95</td>
<td>4.12E-38</td>
<td>unknown [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G160917 miR156</td>
<td>1.85</td>
<td>2.02E-8</td>
<td>SBP-domain protein 5 [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G089361 miR319</td>
<td>1.47</td>
<td>8.62E-10</td>
<td>hypothetical protein SORBIDRAFT_01g006020 [Sorghum bicolor]</td>
<td></td>
</tr>
<tr>
<td>stage III vs. stage II</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRMZM2G126018 miR156</td>
<td>1.02</td>
<td>2.67E-04</td>
<td>unknown [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G163813 miR156</td>
<td>1.49</td>
<td>1.89E-06</td>
<td>unknown [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G081406 miR160</td>
<td>3.52</td>
<td>8.60E-05</td>
<td>hypothetical protein SORBIDRAFT_04g026610 [Sorghum bicolor]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G153233 miR160</td>
<td>1.71</td>
<td>6.09E-07</td>
<td>hypothetical protein LOC100304210 [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G393433 miR164</td>
<td>2.59</td>
<td>6.36E-04</td>
<td>hypothetical protein [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G028980 miR167</td>
<td>1.43</td>
<td>1.57E-06</td>
<td>hypothetical protein SORBIDRAFT_04g004430 [Sorghum bicolor]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G155490 miR390</td>
<td>3.44</td>
<td>2.22E-04</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G443903 miR396</td>
<td>3.21</td>
<td>1.40E-04</td>
<td>putative pol protein [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>stage VI vs. stage III</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRMZM2G307588 miR156</td>
<td>1.37</td>
<td>1.52E-09</td>
<td>SBP-domain protein 6 [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G460544 miR156</td>
<td>-2.54</td>
<td>3.79E-51</td>
<td>unknown [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G081406 miR160</td>
<td>1.13</td>
<td>3.18E-07</td>
<td>hypothetical protein SORBIDRAFT_04g026610 [Sorghum bicolor]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G089361 miR319</td>
<td>-2.07</td>
<td>7.09E-10</td>
<td>hypothetical protein SORBIDRAFT_01g006020 [Sorghum bicolor]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G109843 miR319</td>
<td>-1.23</td>
<td>8.85E-10</td>
<td>hypothetical protein [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G131280 miR529</td>
<td>-1.36</td>
<td>2.06E-05</td>
<td>hypothetical protein LOC100277728 [Zea mays]</td>
<td></td>
</tr>
<tr>
<td>GRMZM2G136158 miR529</td>
<td>1.58</td>
<td>8.41E-05</td>
<td>hypothetical protein SORBIDRAFT_03g010740 [Sorghum bicolor]</td>
<td></td>
</tr>
</tbody>
</table>

a microRNAs identified by Liu H et al. (2014) [10].

b "FC" represents "fold change".
Table 3. Differentially expressed TFs across three comparisons of adjacent developmental stages.

<table>
<thead>
<tr>
<th>Gene</th>
<th>PlantTFDB_ID</th>
<th>TF_Family</th>
<th>IIvs.I</th>
<th>IIIvs.II</th>
<th>IVvs.III</th>
<th>Gene Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G065496</td>
<td>Zma029754</td>
<td>B3</td>
<td>-1.08</td>
<td>1.51</td>
<td>-1.005</td>
<td>B3 DNA binding domain containing protein [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G052102</td>
<td>Zma030311</td>
<td>bZIP</td>
<td>-1.40</td>
<td>-1.66</td>
<td>-3.92</td>
<td>hypothetical protein SORBIDRAFT_04g008840 [Sorghum bicolor]</td>
</tr>
<tr>
<td>GRMZM2G479885</td>
<td>Zma028890</td>
<td>bZIP</td>
<td>-1.08</td>
<td>1.73</td>
<td>1.05</td>
<td>unknown [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G061487</td>
<td>Zma003671</td>
<td>ERF</td>
<td>-2.79</td>
<td>2.68</td>
<td>-1.45</td>
<td>DRE binding factor 1 [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G056600</td>
<td>Zma017283</td>
<td>HD-ZIP</td>
<td>1.56</td>
<td>-1.20</td>
<td>2.74</td>
<td>hypothetical protein LOC100272620 [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G044902</td>
<td>Zma026448</td>
<td>LBD</td>
<td>-1.43</td>
<td>1.23</td>
<td>-1.19</td>
<td>hypothetical protein SORBIDRAFT_01g031790 [Sorghum bicolor]</td>
</tr>
<tr>
<td>GRMZM2G129034</td>
<td>Zma050175</td>
<td>MIKC</td>
<td>-2.20</td>
<td>2.90</td>
<td>2.38</td>
<td>unknown [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G137510</td>
<td>Zma056196</td>
<td>MIKC</td>
<td>-1.21</td>
<td>-1.38</td>
<td>-4.24</td>
<td>unknown [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G050550</td>
<td>Zma002240</td>
<td>MYB</td>
<td>-1.95</td>
<td>3.36</td>
<td>-1.27</td>
<td>sucrose responsive element binding protein [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G127379</td>
<td>Zma007036</td>
<td>NAC</td>
<td>-1.34</td>
<td>1.86</td>
<td>-2.49</td>
<td>unknown [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G347043</td>
<td>Zma057817</td>
<td>NAC</td>
<td>-1.88</td>
<td>-1.26</td>
<td>-1.81</td>
<td>NAC1 transcription factor [Zea mays]</td>
</tr>
<tr>
<td>GRMZM2G109354</td>
<td>Zma006127</td>
<td>SBP</td>
<td>-1.24</td>
<td>1.32</td>
<td>-1.25</td>
<td>H0215A08.3 [Oryza sativa (indica cultivar-group)]</td>
</tr>
<tr>
<td>GRMZM2G113888</td>
<td>Zma001368</td>
<td>TCP</td>
<td>-2.05</td>
<td>3.27</td>
<td>1.22</td>
<td>hypothetical protein LOC100272799 [Zea mays]</td>
</tr>
</tbody>
</table>

a,b,c The values indicate the log2 fold change
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Highlights

- The present study profiled and analyzed gene expression of the maize ear at four key developmental stages.
- Several well-known genes were found with reported mutants analyses, such as ct2, zag1, bde, and sil.
- MicroRNAs such as microRNA156 were predicted to target genes involved in maize ear development.
- 13 TFs from 10 TF families (B3, bZIP, ERF, HD-ZIP, LBD, MIKC, MYB, NAC, SBP, and TCP) were identified.
- Antisense transcripts were widespread throughout all the four stages in maize ear development.