Genome-wide identification, splicing, and expression analysis of the myosin gene family in maize (Zea mays)

Guifeng Wang*, Mingyu Zhong*, Jiajia Wang, Jushan Zhang, Yuaping Tang, Gang Wang† and Rentao Song†
Shanghai Key Laboratory of Bio-Energy Crops, School of Life Sciences, Shanghai University, No. 333 Nanchen Road, Shanghai, PR China

* These authors contributed equally to this work.
† To whom correspondence should be addressed. E-mail: wg@shu.edu.cn and rentaosong@staff.shu.edu.cn

Received 9 August 2013; Revised 17 November 2013; Accepted 20 November 2013

Abstract

The actin-based myosin system is essential for the organization and dynamics of the endomembrane system and transport network in plant cells. Plants harbour two unique myosin groups, class VIII and class XI, and the latter is structurally and functionally analogous to the animal and fungal class V myosin. Little is known about myosins in grass, even though grass includes several agronomically important cereal crops. Here, we identified 14 myosin genes from the genome of maize (Zea mays). The relatively larger sizes of maize myosin genes are due to their much longer introns, which are abundant in transposable elements. Phylogenetic analysis indicated that maize myosin genes could be classified into class VIII and class XI, with three and 11 members, respectively. Apart from subgroup XI-F, the remaining subgroups were duplicated at least in one analysed lineage, and the duplication events occurred more extensively in Arabidopsis than in maize. Only two pairs of maize myosins were generated from segmental duplication. Expression analysis revealed that most maize myosin genes were expressed universally, whereas a few members (XI-1, -6, and -11) showed an anther-specific pattern, and many underwent extensive alternative splicing. We also found a short transcript at the O1 locus, which conceptually encoded a headless myosin that most likely functions at the transcriptional level rather than via a dominant-negative mechanism at the translational level. Together, these data provide significant insights into the evolutionary and functional characterization of maize myosin genes that could transfer to the identification and application of homologous myosins of other grasses.

Key words: Alternative splicing, evolution, expression pattern, headless myosin, maize, myosin.

Introduction

In eukaryotic cells, motor proteins use the energy released from ATP hydrolysis to transport various intracellular cargos, including membranous organelles, protein complexes, and mRNAs, along tracks of cytoskeletal polymers (Lee and Liu, 2004). Of these systems, the actin–myosin system is required for the organization and dynamics of the endomembrane system and transport network in plant cells.

Plants harbour two unique myosin groups: class VIII and class XI. The latter is structurally analogous to class V in metazoans and fungi and contains a dilute domain in the globular tail (Li and Nebenfuhr, 2008). Class V myosin is a processive motor that is responsible for organelle and vesicle transport, and partitioning during cell division, mitotic spindle positioning, mRNA localization, and the establishment of cell polarity (Hammer and Sellers, 2012). Following the first cloned Arabidopsis myosin gene ATM1 (Knight and Kendrick-Jones, 1993), many advances have been achieved in understanding plant myosin function over the past decade. In Arabidopsis, the myosin family contains 17 genes, including 13 class XI (XI-A, -B, -C, -D, -E, -F, -G, -H, -I, ...
small, rounded cells (Vidali et al., 2010). Together, these data indicate that class XI myosins are involved in cytoplasmic streaming, organelle motility, and remodelling, and in plant growth and development.

In Poaceae, only limited progress has been achieved in myosin research, although grasses includes several agronomically important cereal crops with available whole-genome sequences. Twelve, seven, and nine class XI myosin genes exist in rice, sorghum, and brachypodium, respectively, and all contain two class VIII myosin genes (Jiang and Ramachandran, 2004; Peremyslov et al., 2011). Rice myosin XI-B is required for normal pollen development by localizing its protein in a photoperiod-sensitive manner (Jiang et al., 2007). Three myosin genes have been obtained in maize, and one of the XI members is associated with mitochondria, plastids, and the molecular chaperone subunit TCP-1α (Liu et al., 2001; Wang and Pesacreta, 2004). Previously, we functionally characterized opaque1, which encodes a myosin XI protein and influences protein body assembly by affecting ER morphology and motility (Wang et al., 2012a). Here, we employed bioinformatics and publicly available data to identify and analyse maize myosin genes at the genome-wide scale. These findings will provide an important blueprint for future maize myosin functional characterization that will transfer to other grass species for agricultural trait improvement, such as for Opaque1.

Materials and methods

Plant material

Maize inbred-line W22 plants were cultivated in the field at the campus of Shanghai University. Tissues (root, stem, the third leaf, silk, sheath, husk, tassel, and ear) were obtained from at least three healthy plants at the V12 stage as described previously (Wang et al., 2010), and developing kernels were collected at 3, 9, 15, 21, 27 and 33 d after pollination. Tobacco (Nicotiana benthamiana) plants were grown in a greenhouse under a 16/8 h day/night regime at a temperature of 20–25 °C (Wang et al., 2011). Maize callus was induced using F1 seeds of the hybrid Hi II line (Hi II pA×Hi II pB) and grown and maintained in an improved N6 medium.

Isolation and analysis of maize myosin genes

First, the conserved amino acids of the myosin head ATPase domain (Pfam: PF00063) and IQ motif (Pfam: PF00612) were used to search the maize genome database MaizeGDB (http://www.maizegdb.org/). Secondly, all Arabidopsis and rice myosin protein sequences were used as query sequences to search against the maize genome database and National Center for Biotechnology Information (NCBI) using the BLASTP program. The retrieved sequences were then assembled to remove redundancy. The Pfam (http://pfam.sanger.ac.uk/search) and SMART (http://smart.embl-heidelberg.de/) databases were used to confirm each predicted maize myosin sequence. For missannotated or split myosins, reverse transcription (RT)-PCR was used to combine the separated cDNA fragments with the primers described in Supplementary Table S1 (at JXB online).

Gene model and splicing analysis of maize myosin genes

The information for annotated maize myosin genes, including accession number, chromosomal location, open reading frame (ORF) length and exon–intron structure, were retrieved directly from the
B73 maize sequencing database (http://www.maizesequence.org/index.html), and the exon–intron organization of our filled, complete myosins was identified in the maize sequence database using the BLASTn program and constructed using the DNAMAN software. RepeatMasker searching was used to identify repetitive sequences that were present in large introns (>1kb) (Tarailo-Graovac and Chen, 2008). Maize RNA-seq transcriptome data were downloaded from the NCBI Short Read Archive (accession numbers SRX105522, SRX105660, SRX058602, SRX058603, SRX058601, and SRP006965; http://www.ncbi.nlm.nih.gov/sra). RNA-seq reads were mapped to the maize genome assemblies using the TopHat 2.0.9 software (http://tophat.cbcb.umd.edu/; Trapnell et al., 2009).

Phylogenetic analysis

The amino acid sequences of maize myosins, along with those of Arabidopsis thaliana, Oryza sativa, Sorghum bicolor and Saccharomyces cerevisiae, were submitted to the ClustaW program at the BCM search launcher (Baylor College of Medicine, Houston, TX) using their default settings (pairwise alignment options: gap opening penalty 10, gap extension penalty 0.1; multiple alignment options: gap opening penalty 10, gap extension penalty 0.2, gap distance 5, no end gaps and protein weight matrix using Gonnet) for multiple protein alignment. Based on the aligned protein sequences, the phylogenetic tree was constructed using the MEGA5.0 program (http://www.megasoftware.net/) and the neighbour-joining method, and the bootstrap test was carried out with 1000 replicates.

Chromosomal distribution and myosin gene duplication in maize

Based on a previously constructed syntenic map of maize, rice, and sorghum (Schnable et al., 2009), maize myosins were mapped on chromosomes by identifying their physical chromosome position, which was provided in the maize sequence database. Gene duplication events of maize myosin genes were investigated according to block pairs, as described previously (Wei et al., 2007).

RNA extraction and RT-PCR

Immature kernel RNA was isolated using a previously described protocol (Wang et al., 2012b), and RNA was extracted from tissues using TRIzol RNA extraction reagent (Tiangen). The residual DNA was removed by RNase-free DNase I (Takara) treatment, and approximately 2µg of total RNA from each sample was reverse transcribed to cDNA using RevertAid H Minus Reverse Transcriptase (Thermo Scientific). Semi-quantitative PCR was performed using a protocol described previously (Wang et al., 2010). PCR primers were designed using the Primer Premier 5.0 program (Premier Biosoft (Supplementary Table S1). To determine the transcription initiation site of Opaquel, 5’ RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) was performed using a First Choice RLM-RACE kit (Ambion) according to the manufacturer’s protocol. Maize public expression datasets were obtained from the Plant Expression Database (PLEXdb; http://www.plexdb.org/) and a heat map was created using the pheatmap (Pretty Heatmaps) package (R version 3.0.2, pheatmap version, 0.7.7; R Core Team, Vienna, Austria).

Subcellular localization

First, the enhanced YFP (eYFP) coding sequence was amplified from pB7WGY2.0 and inserted into pBl121 using the EcoRI–SalI and Xhol sites. The O1-head-IQ and O1-head cDNA fragments were amplified from maize kernel and inserted into the reconstructed pBl121 vector using the EcoRI and SalI sites, respectively. A transient expression assay was performed using a previously described method (Wang et al., 2012a; Gao et al., 2013), and images were obtained using a combination of 514 nm laser excitation and 530–580 nm long-pass emission filters using a Zeiss confocal microscope LSM 710 (Carl Zeiss). All images were analysed using Image J software (National Institutes of Health).

Analyses of headless myosin gene expression

A DNA fragment containing the β-glucuronidase (GUS) coding region and nopaline synthase (NOS) terminator was obtained from pBI121 and inserted into pUC18 using the SmaI and EcoRI sites. The P1 and P2 fragments were amplified from genomic DNA of B73 using the primers in Supplementary Table S1 and cloned into the above pUC18 harbouring GUS and the NOS terminator using the Xbal and SmaI sites. The resulting plasmids (1µg) were bombarded into maize calli with a Biolistic PDS-1000/He Gun System (Bio-Rad) and GUS activity was investigated by incubating the samples in a 5-bromo-4-chloro-3-indolyl β-glucuronide-containing standard histochemical solution overnight at 37°C.

Results

The maize myosin gene family has 14 members

With the complete maize genome sequence available (Schnable et al., 2009), BLASTP and TBLASTN searches were performed in MaizeGDB, using the head domain and full-length amino acid sequences of the rice (O. sativa) and Arabidopsis myosins. We first retrieved 22 sequences encoding myosin homologues (Supplementary Table S2 at JXB online). Fourteen of the sequences were incomplete because they contained only head or tail domains. Of these sequences, six pairs contained separated myosin head and dilute domains at the approximate positions on the chromosomes, which indicated that these coupled sequences were most likely split from a complete myosin gene.

RT-PCR was used to fill the gaps between the coupled sequences, and we successfully obtained the missing cDNA sequences of the six incomplete myosin genes. Therefore, 14 complete myosin genes were identified in the maize genome (Table 1), including the previously reported Opaquel (Wang et al., 2012a), ZMM1, ZMM2, and ZMM3 (Liu et al., 2001). This number was comparable to that of its monocot relatives Sorghum (12) and rice (14), but was less than that in Arabidopsis (17). The genomic region of the maize myosin genes ranged from 8.218 to 85.126 kb and encoded proteins of 990–2641 aa. Noticeably, the average size of maize myosins was 32.918 kb (class VIII, 12.177 kb; class XI, 38.575 kb), which was much larger than those in Sorghum (16.887 kb), rice (16.595 kb), and Arabidopsis (8.981 kb).

Maize myosin genes possess typical domains but complicated intron–exon organization

The Pfam (Punta et al., 2012) and SMART (Letunic et al., 2012) databases were used to identify the putative domains that were present in the 14 complete maize myosins (Fig. 1, Supplementary Table S3 at JXB online). The myosins all contained a large, ATPase motor domain and several IQ motifs, which were used for ATP hydrolysis and binding calmodulin, respectively. Except for the 990 aa member (GRMZM2G460396), maize myosins could apparently be divided into two classes (VIII and XI) according to the
Table 1. Myosins identified from the completed maize genome sequence

<table>
<thead>
<tr>
<th>Myosin name/new nomenclature</th>
<th>Chromosome</th>
<th>Length (kb)</th>
<th>Exon</th>
<th>Gene accession no.</th>
<th>No. of aa</th>
<th>MW (kDa)</th>
<th>pI</th>
<th>Evidence</th>
<th>Arabidopsis best hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZmVIII-1/</td>
<td>1</td>
<td>15.850</td>
<td>23</td>
<td>GRMZM2G113202</td>
<td>1191</td>
<td>133.472</td>
<td>8.53</td>
<td>Y(M3)</td>
<td>VIII-1</td>
</tr>
<tr>
<td>ZmMyo8A1</td>
<td>5</td>
<td>11.034</td>
<td>23</td>
<td>GRMZM2G139583</td>
<td>1194</td>
<td>134.072</td>
<td>8.50</td>
<td>Y</td>
<td>VIII-1</td>
</tr>
<tr>
<td>ZmVIII-2/</td>
<td>7</td>
<td>9.647</td>
<td>24</td>
<td>GRMZM2G057390</td>
<td>1238</td>
<td>139.825</td>
<td>9.13</td>
<td>Y</td>
<td>VIII-2</td>
</tr>
<tr>
<td>ZmVIII-3/</td>
<td>1</td>
<td>40.839</td>
<td>38</td>
<td>GRMZM2G471108</td>
<td>1506</td>
<td>170.675</td>
<td>7.70</td>
<td>Y</td>
<td>XI-B</td>
</tr>
<tr>
<td>ZmVIII-4/</td>
<td>3</td>
<td>44.509</td>
<td>52</td>
<td>KF493895</td>
<td>2641</td>
<td>299.535</td>
<td>5.93</td>
<td>Y</td>
<td>XI-K</td>
</tr>
<tr>
<td>ZmVIII-5/</td>
<td>4</td>
<td>26.818</td>
<td>38</td>
<td>GRMZM2G449909</td>
<td>1520</td>
<td>173.204</td>
<td>8.85</td>
<td>Y(O1)</td>
<td>XI-I</td>
</tr>
<tr>
<td>ZmVIII-6/</td>
<td>5</td>
<td>8.218</td>
<td>34</td>
<td>GRMZM2G460396</td>
<td>990</td>
<td>113.532</td>
<td>8.01</td>
<td>Y</td>
<td>XI-F</td>
</tr>
<tr>
<td>ZmVIII-7/</td>
<td>3</td>
<td>19.056</td>
<td>39</td>
<td>AC155377.1_FG001</td>
<td>1529</td>
<td>173.465</td>
<td>8.53</td>
<td>Y(M1)</td>
<td>XI-K</td>
</tr>
<tr>
<td>ZmVIII-8/</td>
<td>1</td>
<td>52.790</td>
<td>39</td>
<td>KF493893</td>
<td>1506</td>
<td>171.027</td>
<td>8.48</td>
<td>Y</td>
<td>XI-B</td>
</tr>
<tr>
<td>ZmVIII-9/</td>
<td>5</td>
<td>13.163</td>
<td>39?</td>
<td>KF493892</td>
<td>1554</td>
<td>175.882</td>
<td>8.48</td>
<td>Y(M2)</td>
<td>XI-K</td>
</tr>
<tr>
<td>ZmVIII-10/</td>
<td>6</td>
<td>63.910</td>
<td>45</td>
<td>KF493894</td>
<td>1973</td>
<td>224.613</td>
<td>6.87</td>
<td>Y</td>
<td>XI-I</td>
</tr>
<tr>
<td>ZmVIII-11/</td>
<td>7</td>
<td>59.851</td>
<td>45?</td>
<td>KF493896</td>
<td>1880</td>
<td>213.827</td>
<td>7.07</td>
<td>Y</td>
<td>XI-K</td>
</tr>
<tr>
<td>ZmVIII-12/</td>
<td>8</td>
<td>10.042</td>
<td>39</td>
<td>GRMZM2G435294</td>
<td>1529</td>
<td>173.418</td>
<td>8.92</td>
<td>Y</td>
<td>XI-E</td>
</tr>
</tbody>
</table>

remaining domains. Compared with class VIII, the class XI myosins were much longer and had an N-terminal SH3-like domain and a tail dilute domain. Similar to that in Arabidopsis and rice, the maize myosin head domain was located more towards the N terminus in class XI than in class VIII. Moreover, the class VIII myosins contained a coiled-coil domain and four IQ motifs, whereas a variable number of both domains were present in class XI members.

The intron–exon structures of the maize myosin genes were determined by comparison of the cDNA with genomic sequences. The results revealed that maize myosin genes consisted of 23 (ZmVIII-1 and ZmVIII-2) to 52 (ZmXI-2) exons, and the sizes varied from 14 to 1199 bp (Fig. 2, Table 1 and Supplementary Table S4 at JXB online). The class VIII myosin genes contained 23–24 exons, whereas class XI genes harboured 38–52 exons. Noticeably, the length of most exons appeared to be conserved in the same order in the different maize myosin genes.

Compared with the conserved exon sizes in the same order, the introns were more divergent in length (57 bp to 22.499 kb) and caused the large sizes of the maize myosin genes. RepeatMasker searching was used to identify repetitive sequences that were present in large introns (>1 kb) (Tarailo-Graovac and Chen, 2009). As shown in Supplementary Table S5 at JXB online, retroelements (L1/CIN4, Ty1/Copia, SINEs, RTE/Bov-B, and Gypsy/DIRS) and DNA transposons (hobo-Activator, Tourist/Harbinge, and Tc1-IS630-Pogo) were abundant in maize myosin introns. This finding revealed that transposable elements play an essential role in maize myosin size increase and intron–exon organization.

Maize myosins are less duplicated than those of Arabidopsis

The available myosin gene family in Arabidopsis, rice, and sorghum allowed us to investigate the evolutionary relationship between dicot and monocot myosin proteins. A neighbour-joining tree was constructed using the full-length protein sequences of 17 Arabidopsis, 13 rice, 9 sorghum, and 14 maize myosins, with two yeast class I myosins as an outgroup (Fig. 3). The results indicated that these myosins were divided distinctly into two groups, class VIII and class XI, which supports the suggestion that two myosin ancestors existed in planta before lineage-specific expansions (Peremyslov et al., 2011). Eleven myosins from maize, 13 from Arabidopsis, 11 from rice and seven from sorghum in class XI were found, whereas the smaller VIII group contained three maize, four Arabidopsis, and two rice and sorghum members. This result was generally consistent with the conclusion that was made previously using myosin motor domains (Peremyslov et al., 2011). The class VIII myosins were divided into two distinct subgroups (VIII-A and VIII-B), and class XI was split into the I, G, F, K, and E subdivisions (Peremyslov et al., 2011; Wang et al., 2012a). Interestingly, Arabidopsis XI-J was grouped alone into a clade without any close homologues.

It was clearly shown that class VIII myosins had undergone a single-gene duplication but that a burst of duplication occurred in class XI (Fig. 3). Two paralogous branches were formed within subclasses VIII-A and VIII-B in the dicot Arabidopsis compared with only one in monocots, excepting for the two in maize VIII-B. The largest subgroup, XI-G, contained six members in Arabidopsis but only three in the monocots rice, sorghum, and maize. Moreover, two paralogues were found within XI-K and XI-E in Arabidopsis, compared with only one in monocots. Consistently, subgroup XI-F was the only subdivision that was not duplicated in any of the plant lineages (Peremyslov et al., 2011). Noticeably, three maize and two rice myosins comprised a separate clade that was previously designated XI-K (Jiang and Ramachandran, 2004).

Chromosomal localization and myosin gene duplication in grass

To investigate the chromosomal distribution of the myosin family in maize, the loci were determined by directly...
identifying their physical positions provided in MaizeGDB. As shown in Fig. 4, the 14 myosins were mapped on eight out of the 10 maize chromosomes, excluding chromosomes 2 and 10. The largest number of myosin members (four) was present on chromosome 5, followed by two genes each on chromosomes 1, 3, and 7. The remaining chromosomes (4, 6, 8, and 9) each had a unique myosin.

Maize originated from an ancient allotetraploid and has undergone several rounds of whole-genome duplication events during its gene evolution (Wei et al., 2007; Schnable et al., 2009). Among these maize myosin genes, four sister pairs (ZmVIII-1 and ZmVIII-2, ZmXI-1 and ZmXI-5, ZmXI-2 and ZmXI-10, and ZmXI-4 and ZmXI-7) appeared to be generated from segmental duplication events due to their position on the same duplicated gene blocks within maize (Fig. 4). To confirm this possibility, we further analysed myosin gene evolution among maize, rice, and sorghum (Table 2). Thirteen of 14 maize myosins had collinear genes in rice (the exception was ZmXI-9), while all had syntenic members in sorghum. Two of the sister pairs (ZmVIII-1 and ZmVIII-2, and ZmXI-2 and ZmXI-10) each had a unique, syntenic myosin in rice and sorghum, which indicated that they were generated from segmental duplication after the divergence between maize and its relatives (rice and sorghum). The remaining two pairs (ZmXI-1 and ZmXI-7, and ZmXI-4 and ZmXI-5) were each located on the same duplicated blocks, but coupled, syntenic myosins were found in rice and sorghum. It is possible that the ancestor of grasses already contained two myosin genes at this region and that one copy of each duplicate myosin pair was lost at the duplicated blocks after the whole-genome duplication in maize (Schnable et al., 2011). In addition, three rice myosin genes (OsXI-D, OsXI-K, and OsXI-L) had no syntenic member in maize or sorghum. OsXI-L was most likely duplicated from OsXI-I, whereas OsXI-D and OsXI-K appeared to be duplicated or redistributed from the original loci to the current positions, and the original gene was partially or fully lost (Salse et al., 2008).

Expression pattern of maize myosin genes

To investigate the spatial patterns of myosin gene expression in maize, we detected their transcripts using RT-PCR in eight representative tissues, i.e. sheath, stem, leaf, tassel, husk, silk, root, and ear (Fig. 5A). The results revealed that the majority of maize myosin genes (excluding ZmXI-2 and ZmXI-10) were expressed at varying levels in all tested tissues. The ZmXI-2 transcripts were particularly abundant in root but absent in silk and tassel. In comparison, its duplicated sister, ZmXI-10, was expressed at low levels in all tested tissues and was undetectable in tassel. Three myosin
We also investigated the expression levels of maize myosins in a broad range of tissues and organs of the inbred line B73 using a NimbleGen microarray (Sekhon et al., 2011). For the currently misannotated full-length myosin genes, we used probes corresponding to their head and tail domain for further analysis. According to their expression levels, maize myosin genes could be clustered into three categories (Fig. 5C).

The first category was composed of eight genes (ZmVIII-1, -2, and -3; ZmXI-3, -4, -7, -9-tail and -10-tail) with high and ubiquitous expression levels, suggestingcompeting roles in plant growth and development. All maize class VIII myosins belonged to this category; however, their biological functions remain to be elucidated. Five class XI myosins (ZmXI-2, -5, -8, -9-head and -10-head) formed the second category with intermediate and differential transcription levels. The third category included ZmXI-1, -6 and -11, for which the expression levels were relatively low. However, all three genes were expressed preferentially in anther, especially ZmXI-11 (Fig. 5C), implying that these myosins are required for anther and/or pollen growth and development.

Alternative splicing is a common character of the maize myosin genes.

Alternative splicing plays an essential role in regulating gene expression during growth and development, as well as in response to biotic and abiotic stresses (Stamm et al., 2005; Syed et al., 2012). RNA-seq analysis has revealed that more than 61% of intron-containing genes in Arabidopsis undergo alternative splicing under normal growth conditions (Marquez et al., 2012). Due to their large sizes and the complicated gene models of myosins, we could speculate that maize myosin genes underwent extensive alternative splicing events. We then detected the splicing transcripts of the full-length cDNA for each maize myosin gene in leaf and kernel in two randomly selected windows. Six out of the 14 myosin genes had splicing transcripts in the given cDNA fragment in the two samples (Fig. 6 and Supplementary Fig. 1 at JXB online). Of these genes, four (ZmXI-2, -5, -7 and -10) possessed at least three splicing transcripts, while the remaining had two transcripts. Interestingly, ZmXI-11 was splice in a tissue-dependent manner. A 375 bp cDNA fragment of ZmXI-11 was unique to tassel, while the others had an unspliced, 637 bp transcript (Fig. 6). It is unknown whether this alternative splicing affects the subsequent translation. In the case of the maize ZmXI-11 gene, the spliced transcript encoded a complete myosin (1529 aa), whereas the unspliced transcript generated a truncated myosin of 1454 aa in length that lacked a portion of the DIL domain. Together with its preferential expression pattern, this finding suggests that ZmXI-11 most likely plays an essential role in microsporogenesis.

To evaluate further the gene models for maize myosins, transcript evidence created using the high-throughput RNA-seq transcriptome data from developing kernels was compared with all publicly available myosin genomic and transcript data using the TopHat 2.0.9 software (Trapnell et al., 2009). According to this method, we revised the gene models for 13 maize class VIII and class XI myosins, except for ZmXI-11 with few reads in the

Fig. 3. Phylogenetic relationships between the entire myosins of maize, rice, sorghum, and Arabidopsis. The phylogenetic tree was constructed using MEGA 5.0 software, and yeast class I myosins were used as outgroups (see Materials and methods). At, Arabidopsis thaliana; Os, Oryza sativa; Sb, Sorghum bicolor; Sc, Saccharomyces cerevisiae; Zm, Zea mays.

XI genes, ZmXI-3, ZmXI-5, and ZmXI-8, exhibited high transcription levels in silk. Noticeably, ZmXI-11 was highly expressed in tassel, whereas the remaining myosins were less abundant in tassel when compared with the other tissues. Interestingly, the short ZmXI-11 transcript was unique to tassel, while the long, unspliced transcript was common in the remaining tissues (Fig. 6). Furthermore, the two duplicated myosin gene pairs shared similar expression patterns in all tissues. We also unravelled the expression profiles of the maize myosin genes during kernel development (Fig. 5B). Transcripts of all 14 maize myosins were detectable with dynamic patterns. Most were expressed preferentially at the early stage, and peaked approximately 3 and 9 d after pollination. ZmXI-3, ZmXI-6, and ZmXI-7 exhibited a distinct pattern, with two expression peaks at the early and late stages of kernel development. The duplicated gene pairs also showed similar expression profiles.
RNA-seq data (Supplementary Fig. 2 at JXB online). In the case of ZmXI-2, the RNA-seq transcript evidence supported the results that ZmXI-2 was a complete myosin gene (Fig. 7A) and our RT-PCR detected splicing transcripts did exist in the transcriptomic data (Fig. 7B). In addition, the RNA-seq data revealed a novel exon splicing event in annotated exon 19 (Fig. 7B). All these splicing variants resulted in truncated proteins that only harboured the myosin head domain (Fig. 7C).

**O1, a functional myosin, is responsible for ER streaming and morphology that ultimately affects endosperm texture**

Previously, we isolated the maize endosperm mutant opaque1 (o1) using map-based cloning and found that it encoded a class XI myosin protein (Wang et al., 2012a). O1 was identical to ZmXI-4, and resulted in a transcript of 5131 bp, encoding a 173 kDa protein of 1520 aa. To study the association of ZmXI-4/O1 with actin filaments, we transiently expressed the ZmXI-4 (Head-6IQ/Head)–eYFP reporters, which harboured the head domain with/without the IQ motif fused to eYFP at its C terminus (Fig. 8A), for direct, microscopic observation. The signal of the ZmXI-4 head–neck domain protein showed a clear pattern of actin filaments throughout the entire cell (Fig. 8B) that also decorated the nucleus (Fig. 8C). When only the ZmXI-4 head domain was expressed, the YFP signal was similarly associated with actin filaments (Fig. 8D), indicating that the head domain was essential for myosin binding to actin filaments and that the IQ motif had no effect on this process. Together with overexpression of the ZmXI-4/O1 tail
Table 2. Analysis of gene duplication in maize myosins

The light grey and light and medium brown shading represent maize, rice, and sorghum myosin genes, respectively (from top to bottom); the other coupled shading represents the myosin genes in the same duplication block in maize.

<table>
<thead>
<tr>
<th>Maize myosin</th>
<th>Physical block</th>
<th>Chr</th>
<th>Block pairs</th>
<th>Rice syntenic region</th>
<th>Physical position</th>
<th>Chr</th>
<th>Sorghum syntenic region</th>
<th>Physical position</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZmVIII-1</td>
<td>46</td>
<td>1</td>
<td>1 La</td>
<td>Os10g34710 (OsVIII-A/B)</td>
<td>18.53</td>
<td>10</td>
<td>Sb01g018770 (SbVIII-1)</td>
<td>19.62</td>
</tr>
<tr>
<td>ZmVIII-2</td>
<td>212</td>
<td>5</td>
<td></td>
<td>Os02g37560 (OsVIII-B)</td>
<td>22.48</td>
<td>7</td>
<td>Sb02g035390 (SbVIII-2)</td>
<td>70.79</td>
</tr>
<tr>
<td>ZmXI-1</td>
<td>67</td>
<td>1</td>
<td>3 L</td>
<td>Os03g64290 (OsXI-G)</td>
<td>36.33</td>
<td>3</td>
<td>Sb01g000330 (SbXI-1)</td>
<td>0.18</td>
</tr>
<tr>
<td>ZmXI-5</td>
<td>206</td>
<td>5</td>
<td></td>
<td>Os03g53660 (OsXI-F/E)</td>
<td>30.76</td>
<td></td>
<td>Sb01g008180 (SbXI-2)</td>
<td>7.05</td>
</tr>
<tr>
<td>ZmXI-2</td>
<td>143</td>
<td>3</td>
<td>1 L</td>
<td>Os01g51630, Os01g51632, Os01g15634 (OsXI-A)</td>
<td>29.65*</td>
<td>1</td>
<td>Sb03g032760, Sb03g032770 (SbXI-5)</td>
<td>61.20*</td>
</tr>
<tr>
<td>ZmXI-10</td>
<td>359</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZmXI-3</td>
<td>150</td>
<td>3</td>
<td>1 L</td>
<td>Os01g40200 (Headless)</td>
<td>22.68</td>
<td>1</td>
<td>Sb03g026110 (SbXI-4)</td>
<td>52.49</td>
</tr>
<tr>
<td>ZmXI-4</td>
<td>182</td>
<td>4</td>
<td>2 Lb</td>
<td>Os02g53740 (OsXI-C/F)</td>
<td>32.90</td>
<td>2</td>
<td>Sb04g034830 (SbXI-6)</td>
<td>64.68</td>
</tr>
<tr>
<td>ZmXI-7</td>
<td>254</td>
<td>5</td>
<td></td>
<td>Os02g57190 (OsXI-B)</td>
<td>35.02</td>
<td></td>
<td>Sb04g037210 (SbXI-7)</td>
<td>66.88</td>
</tr>
<tr>
<td>ZmXI-6</td>
<td>217</td>
<td>5</td>
<td>10 La</td>
<td>Os10g25565 Os10g25566 (OsXI-C/K)</td>
<td>13.25*</td>
<td>10</td>
<td>Sb01g023160,</td>
<td>30.58*</td>
</tr>
<tr>
<td>ZmXI-8</td>
<td>287</td>
<td>6</td>
<td>5 L</td>
<td>Os05g46030 (OsXI-H)</td>
<td>26.68</td>
<td>5</td>
<td>Sb02g024840,</td>
<td>56.01*</td>
</tr>
<tr>
<td>ZmXI-9</td>
<td>300</td>
<td>7</td>
<td>7 S</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Sb02g010040 (SbXI-3)</td>
<td>14.57</td>
</tr>
<tr>
<td>ZmXI-11</td>
<td>368</td>
<td>9</td>
<td>6 S</td>
<td>Os02g34090 (OsXI-D)</td>
<td>20.36</td>
<td>2</td>
<td>Sb02g019860 (SbXI-L)</td>
<td>9.90</td>
</tr>
</tbody>
</table>

a Maize FPC contigs (http://www.genome.arizona.edu/fpc/maize/WebFPC/) have been numbered sequentially along each chromosome.
b Regions on chromosomes that have contiguous contigs that are syntenic with rice as defined by SyMAP.
c The physical position on each chromosome.
d The former letter represents the rice myosin gene name that was proposed by this study; the latter indicates that proposed by Jiang and Ramachandran (2004).
e The misannotated myosin genes.
Fig. 5. Expression patterns of maize myosin genes. (A) RNA expression levels of maize myosin genes in various tissues. (B) Expression profiles of myosin genes during maize kernel development. (C) Development- and anatomy-specific expression profiles of maize myosin genes. Public microarray data sets were obtained from PLEXdb. The genes are located on the right and the tissues are indicated at the bottom of each column. The colour bar represents the expression values. Head and tail represent the splicing cDNAs encoding myosin head and tail domains, respectively. (This figure is available in colour at JXB online.)
domain significantly inhibiting ER streaming in tobacco cells and the phenotypes of the opaque1 mutant, O1 was characterized as a functional myosin XI protein that was responsible for protein body assembly and endosperm texture determination (Wang et al., 2012a).

We further explored whether functional myosin headless variants were present in the genome and their roles during plant growth and development. Apart from the longest 5131 bp mature transcript of ZmXI-4/O1, we also identified a 2380 bp transcript that consisted of the second half of the longest transcript at the same locus (Fig. 9A) using two independent, 5′ RACE experiments (Supplementary Fig. 3 at JXB online). This short transcript contained an ORF of 1935 bp that conceptually encoded a 73 kDa headless myosin. Domain prediction analysis revealed that the deduced protein contained coiled-coil and DIL domains, and completely lost the N-terminal SH3-like domain, head domain, and IQ motif (Fig. 9B). To verify whether this headless myosin was a novel gene model or an alternative splicing variant of O1, the approximately 2 kb genomic regions that were located upstream of the transcription start sites of O1 (P1) and the headless myosin(P2) were fused to the GUS ORF. After bombardment into maize callus, substantial GUS staining was observed for the O1-prom::GUS construct, while the headless myosin promoter showed no activity in the callus or developing kernel (Fig. 9C, D). Curiously, the predicted headless myosin protein was also undetectable in all samples using an immunoblot assay (Fig. 9E). This finding suggested that this short transcript was a headless derivative of O1 and most likely functioned at the transcriptional level rather than in regulation of O1 activity via dominate-negative suppression in vivo.

Discussion

Myosins are cytoskeletal molecular motors that use the energy released from ATP hydrolysis to move along actin filaments. In a genome-wide screen, we identified 14 complete myosin genes that harboured conserved domains in maize. A comparable number of myosin genes exists in maize (14), rice (14), and Arabidopsis (17), although the maize genome size (2300 Mb) (Schnable et al., 2009) is ~5.3 and ~18.4 times larger than those of rice (430 Mb) (Burr, 2002) and Arabidopsis (125 Mb) (Arabidopsis, 2000), respectively. This huge discrepancy might partially be compensated by the larger average size of the maize myosin genes, which are nearly two and four times longer than those in rice and Arabidopsis, respectively. Moreover, the exon sizes of the maize myosin genes are similar to those in rice and Arabidopsis, but maize myosin genes contain larger introns because of the insertion of transposable elements (Supplementary Table S5). Overall, transposable elements comprise approximately 85% of the B73 maize genome (Schnable et al., 2009) and play important roles in genome organization and evolution (Fedoroff, 2012), as well as in gene-family expansion (Janousek et al., 2013), which is most likely what is occurring in the maize myosin gene family.

Phylogenetic reconstruction analyses using the sequence information of several complete genomes of the green algae, mosses, and higher plants indicate that two types of myosin (VIII and XI) are present in planta (Avisar et al., 2008b) and that three class VIII and 11 class XI myosin members are present in maize. In most cases, subgroups are duplicates of at least one lineage (Fig. 4). For example, subgroup VIII-A and VIII-B were duplicated after the separation between dicot...
Arabidopsis and monocots, while VIII-A was duplicated after the divergence between maize and its relatives. Prominently, only XI-F has not been duplicated within any lineage. The Arabidopsis T-DNA mutant of XI-F showed no discriminable phenotypes in vegetative growth (Peremyslov et al., 2008), and some unexpected clades, such as Arabidopsis XI-J and the maize- and rice-specific XI-K, were found. Thus, the specific function of these myosin genes still requires further experimental elucidation.

Two rounds of whole-genome duplications have occurred during cereal genome evolution: the first occurred before the divergence of maize, rice, and sorghum, whereas the second took place in the specification of the maize lineage (Salse et al., 2009). It appears that relatively more family members exist in maize than in other cereals; however, this is not the case for the maize myosin gene family (Fig. 5, Table 2). Overall, most of the myosins were present before the specification of maize, rice, and sorghum, but they were subjected to gene loss and redistribution within a specific lineage. For example, the rice syntenic gene (OsXI-D) of ZmXI-9 was lost in the syntenic region 7S but was redistributed to a new locus in rice chromosome 2. Only two myosin pairs (ZmVIII-1 and ZmVIII-2, and ZmXI-2 and ZmXI-10) were segmentally duplicated after the divergence of maize and its
relatives. Additionally, the *Arabidopsis* myosins have been amplified in most of the subgroups. Therefore, the events that triggered these lineage-specific duplications of the myosin genes and the subfunctionalization or neofunctionalization of their paralogous myosins following the duplications are becoming known.
Due to the global overlapping expression patterns of the Arabidopsis myosin genes (Peremyslov et al., 2011), only two (MYA2 and XI-K) of all 13 class XI myosin single mutants exhibited detectable defects under normal growth conditions (Peremyslov et al., 2008), whereas their double, triple, and quadruple mutants displayed similar but more severe phenotypes when compared with the single mutants (Prokhnvesky et al., 2008; Peremyslov et al., 2010; Ueda et al., 2010; Ojangu et al., 2012). Transcripts of all myosin family members were detected in all tested tissues and organs of maize, although their abundance was considerably varied (Fig. 6). It has been shown that the global expression patterns of a fraction of orthologous genes are conserved in animals (Zheng-Bradley et al., 2010) and plants (Davidson et al., 2012). In the case of orthologous myosin genes in maize and Arabidopsis, most display similar expression patterns. For example, the Arabidopsis orthologues (XI-C and XI-E) of the maize anther-expressed ZmXI-II displayed extremely low levels in the entire plant but exclusively high levels in the stamen/anther (Peremyslov et al., 2011), suggesting their important roles in pollen growth. In addition, the segmental duplicated maize myosin pairs shared correlated expression patterns when compared with other paralogues.

In humans, alternative splicing of myosin Va occurs in a region lying between the coiled-coil region of the IQ neck and the globular tail region in a tissue-specific manner (Hodi et al., 2006; Wagner et al., 2006; Roland et al., 2009; Wu et al., 2002), and usually three exons in this region are subject to alternative splicing. For example, exon B is required for the dynein light chain 2 (DLC2)-myosin Va interaction (Hodi et al., 2006; Wagner et al., 2006), whereas exon D is essential for Rab10 binding to myosin V tails in vivo (Roland et al., 2009). In the present study, we found that many maize class XI myosin genes had several splicing transcripts within the randomly selected cDNA fragment in leaf and kernel (Fig. 7). Further sequence analysis showed that most of these fragments mapped to the region between the IQ motif and the tail domain, suggesting a conserved mechanism that regulates myosin function in human and plants.

Opaque1 is the first characterized maize class XI myosin that is required for ER motility and protein body formation (Wang et al., 2012a). In addition to the full-length O1 cDNA, we isolated a short O1 transcript that conceptually encoded a truncated, headless O1 at the O1 locus (Fig. 9). Arabidopsis and Brachypodium possess and express a headless variant of XI-K that emerged via partial myosin XI duplication (Peremyslov et al., 2011). In comparison with HDK, the short myosin transcript is not an independent gene model but rather a headless variant of O1 and appears to function at the transcription level. It is well known that overexpression of the myosin XI tail domain can inhibit its corresponding, endogenous, complete myosin function via the competitive binding of adaptors that mediate the interaction between myosin and its cargo. However, in the case of the headless derivative of O1, whether another mechanism other than dominant-negative regulation of the function of myosin exists remains to be addressed.

**Supplementary data**

Supplementary data are available at *JXB* online.

- **Supplementary Table S1.** A list of primers used in this study.
- **Supplementary Table S2.** Sequences encoding myosin homologues retrieved from maize genome by BLASTP and TBLASTN searches.
- **Supplementary Table S3.** Major domains presented in maize myosin proteins.
- **Supplementary Table S4.** Analysis of intron and exon sizes and intron phases in maize myosins.
- **Supplementary Table S5.** Analysis of repetitive sequences presented in the large introns (>1 kb) in maize myosins using the RepeatMasker program.
- **Supplementary Fig. S1.** Sequencing analysis of the alternative splicing events of some maize myosin genes.
- **Supplementary Fig. S2.** Transcript evidence and splicing event for maize myosin genes.
- **Supplementary Fig. S3.** Determination of the headless derivative of O1 using two independent 5′ RACE experiments.

**Acknowledgements**

This work was supported by the National Natural Sciences Foundation of China (31370035, 31000747, and 31171559), Ministry of Science and Technology of China (2012AA10A305 and 2014CB138204), and Industry-Study-Research Program of Shanghai High Education Commission, Shanghai Key Basic Research Program (13JC140500).

**References**


Li JF, Nebenfuhr A. 2008. The tail that wags the dog: the globular tail domain defines the function of myosin V/XI. Traffic 9, 290–298.


