# Identification of the maize Mediator CDK8 module, and *Dissociation* insertional mutagenesis of *ZmMed12a*

**Tania Núñez-Ríos[1†], Kevin R. Ahern[2], Ana Laura Alonso-Nieves[1], Daniel Lepe-Soltero[1], Carol Martínez-Camacho[1†], Marcelina García-Aguilar[1], Thomas P. Brutnell[3], C. Stewart Gillmor*[1] and Ruairidh James Hay Sawers*[1]**

[1] Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Unidad de Genómica Avanzada, Centro de Investigaciones y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, Guanajuato, México.
[2] Boyce Thompson Institute for Plant Research, Ithaca, New York, U.S.A.
[3] Donald Danforth Plant Science Center, St. Louis, Missouri, U.S.A.

[†]Current address: PLANAMERICA, Michoacán, México (TN-R); Universidad Intercultural de Chiapas, México (CM-C)

*Correspondence:

Stewart Gillmor & Ruairidh Sawers

Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Unidad de Genómica Avanzada, Centro de Investigaciones y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, Guanajuato, México.

stewart.gillmor@cinvestav.mx; rusawers@cinvestav.mx



## ABSTRACT

Mediator is a conserved transcriptional co-activator that links transcription factors bound at enhancer elements to RNA Polymerase II. Mediator-RNA Polymerase II interactions can be sterically hindered by the Cyclin Dependent Kinase 8 (CDK8) module, a submodule of Mediator that acts to repress transcription in response to discrete cellular and environmental cues. The CDK8 module is conserved in all eukaryotes and consists of 4 proteins: CDK8, CYCLIN C (CYCC), MED12, and MED13. In this study, we have characterized the CDK8 module of Mediator in maize. The maize genome contains single copy genes for *Cdk8*, *CycC*, and *Med13*, and two genes for *Med12*. Analysis of expression data for the CDK8 module demonstrated that all five genes are broadly expressed in maize tissues, with *ZmMed12a*, *ZmMed12b*, and *ZmMed13* exhibiting similar expression patterns. We performed a *Dissociation (Ds)* insertional mutagenesis, recovering two independent insertions in the *ZmMed12a* gene. One of these *Ds* insertions results in a truncation of the *ZmMed12a* transcript. Our molecular characterization of the maize CDK8 module, as well as transposon tagging of *ZmMed12a*, establish the basis for molecular and functional studies of these important transcriptional regulators in *Zea mays*.

**KEYWORDS** *Zea mays*, *Dissociation,* Mediator, CDK8 module, Med12, Med13, CycC

## INTRODUCTION

Transcriptional regulation plays an essential role in almost all aspects of development and physiology, including responses to the biotic and abiotic environment. One key regulator of transcription is Mediator, a multiprotein complex conserved from yeast to plants to animals, which was initially identified based on its requirement for transcription of virtually all protein-coding genes (Kelleher et al., 1990; Flanagan et al., 1991; Bourbon, 2008). The Core Mediator consists of Head, Middle and Tail domains, and typically functions as a transcriptional co-activator, linking transcription factors bound at upstream enhancer elements to RNA polymerase II (RNA pol II) (reviewed in Yin and Wang, 2014; Allen and Taatjes, 2015). The Head and Middle domains interact with RNA pol II, while the Tail domain is thought to interact with specific transcription factors (Tsai et al., 2014; Robinson et al., 2015; Plaschka et al., 2015; reviewed in Larivière et al., 2012). A fourth Mediator module shows transient association with Core Mediator and often acts to repress transcription. This Cyclin Dependent Kinase 8 (CDK8) module is composed of the proteins MED12, MED13, CYCLIN C (CYCC), and CDK8 (reviewed in Björklund and Gustafsson, 2005). In agreement with the variable association of the CDK8 module with Core Mediator, purification of Mediator from *Arabidopsis thaliana* yielded both conserved Core Mediator subunits, as well as subunits unique to Arabidopsis, but did not include components of the CDK8 module (Bäckström et al., 2007).

In yeast and animals, components of the CDK8 module can regulate transcription in several ways, with different subunits playing different roles. One mechanism for transcriptional repression involves steric inhibition, where the CDK8 module occupies the Core Mediator pocket that binds RNA pol II, thereby preventing interaction of Core Mediator and RNA pol II (Elmlund et al., 2006; Tsai et al., 2013). Transcriptional repression by this steric mechanism has the potential to be dynamic, as the occupancy of the RNA pol II binding pocket can be modulated during subsequent rounds of assembly of the Mediator-RNA pol II holoenzyme (reviewed in Allen and Taatjes, 2015). This steric mechanism involves all four units of the CDK8 module, with the MED13 subunit playing the most important role, interacting directly with the Middle domain of Core Mediator (Knuesel et al., 2009; Tsai et al., 2013). The MED13 subunit also serves an important function in regulation of CDK8 module stability: phosphorylation of a conserved phosphodegron site in MED13 can lead to recognition by a ubiquitin ligase complex, and subsequent degradation (Davis et al., 2013).

In Arabidopsis, components of the CDK8 module were initially identified by their requirement for development, and also affect the response to fungal pathogens and cellular stress. Mutations in *CDK8* were identified as enhancers of the phenotype of the floral homeotic mutant *hua1hua2,* and thus were named *hua enhancer 3 (hen3)*. *hen3* mutants affect floral organ identity, as well as leaf size and cell shape, and the HEN3 protein was demonstrated to have CDK8 kinase activity (Wang and Chen, 2004). CDK8 regulates

retrograde signaling from the mitochondria to the nucleus in response to $H_2O_2$ and cold stress (Ng et al., 2013). *CDK8*, as well as *MED12* and *MED13*, are also required for the response to both fungal and bacterial pathogens (Zhu et al., 2014).

Mutations in *MED12* and *MED13* were initially reported from a genetic screen for regulators of pattern formation in Arabidopsis embryogenesis, and were named *center city (cct)* and *grand central (gct),* to reflect the increased size of the shoot apical meristem (SAM) in these mutants. *cct* and *gct* mutants delay the timing of pattern formation during embryogenesis, rather than affecting pattern formation *per se*- the increased size of the SAM in *cct* and *gct* mutants can be attributed to its formation later in embryogenesis compared to the wild type (wt) (Gillmor et al., 2010). The delayed formation of the SAM may be related to auxin signaling, as both the *med13* allele *macchi-bou2 (mab2)*, and the *med12* allele *cryptic precocious* (*crp*) act as enhancers of a mutation in the auxin dependent kinase *PINOID* (Furutani et al., 2004; Ito et al., 2011; Imura et al., 2012). Importantly for mechanistic studies of CDK8 module function in Arabidopsis, Ito et al. (2011) demonstrated that the MED13 and CDK8 proteins are both able to interact with Cyclin C, as has previously been demonstrated in Drosophila (Loncle et al., 2007). Consistent with studies showing auxin-related phenotypes for mutants in *MED12* and *MED13*, a recent study showed that both of these genes, as well as *CDK8*, are involved in auxin transcriptional responses, and that the MED13 protein relays signals from the IAA14 protein to repress the auxin responsive transcription factors ARF7 and ARF19 (Ito et al., 2016).

In addition to affecting the timing of pattern formation in embryogenesis, *MED12* and *MED13* also regulate the timing of post-embryonic phase transitions in Arabidopsis. A dominant allele of *med12* (named *cryptic precocious* (*crp-1D*)) was isolated in a genetic screen for enhancers of the early flowering phenotype conditioned by overexpression of the florigen *FT* (Imura et al., 2012). Loss of function mutants in *crp/cct* and *gct* show late flowering due to overexpression of the floral repressor *FLOWERING LOCUS C (FLC)*, as well as decreased expression of the floral promoters *FLOWERING LOCUS (FT)*, *TWIN SISTER OF FT (TSF)*, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANTS 1 (SOC1)*, *APETALA 1 (AP1)* and *FRUITFULL (FUL)* (Imura et al., 2012; Gillmor et al., 2014). *cct* and *gct* mutants also misexpress seed specific genes during seedling development, and have an elongated vegetative phase due to overexpression of the microRNA miR156 (Gillmor et al., 2014), a master regulator of the vegetative phase in plants (Wu et al., 2009). Taken together, these results demonstrate that *MED12* and *MED13* act as master regulators of developmental timing in plants, regulating the timing of pattern formation in embryogenesis, the seed-to-seedling transition, vegetative phase change, and the transition to flowering (Gillmor et al., 2010; Ito et al., 2011; Imura et al., 2012; Gillmor et al., 2014).

Due to its importance in plant development and physiology, we have extended studies of the CDK8 module to the crop plant maize (*Zea mays*). Establishment of molecular and genetic resources for the study of the maize CDK8 module will allow evaluation of its role in the regulation of agricultural traits such as timing of flowering and seed development, as well as responses to biotic and abiotic stresses. One of the primary goals of this work was isolation of loss of function mutant alleles of maize CDK8 module-encoding genes. In maize, resources based on endogenous DNA transposons constitute the most accessible and widely-used technology for reverse genetics (McCarty and Meeley, 2009). The two major transposon

systems used for gene tagging in maize are *Activator*/*Dissociation* (*Ac*/*Ds*) and *Mutator* (*Mu*) (Candela and Hake, 2008). These systems consist of an autonomous or master element that encodes a transposase (TPase) and a second non-autonomous or receptor element. The receptor elements are frequently derived from a master element by mutations within the TPase gene. Lacking TPase, non-autonomous elements are stable, unless mobilized by TPase supplied *in trans* by an autonomous element (Kunze et al., 1997). *Ac* is a member of the hAT transposon superfamily (named after the founding members *hobo*, *Ac* and *Tam3*; Calvi et al., 1991) and moves via a cut-and-paste mechanism (Bai et al. 2007), with a preference for transposition to linked sites, making the system ideal for local mutagenesis (Greenblatt, 1984; Dooner and Belachew, 1989; Brutnell and Conrad, 2003). To exploit the *Ac*/*Ds* system for reverse genetics, *Ds* elements have been distributed throughout the genome to provide potential "launch pads" for mutagenesis of nearby genes (Vollbrecht et al. 2010).

In this study, we identify five genes encoding components of the CDK8 module in maize, present experimentally determined gene structures, and report expression of corresponding transcripts. We performed *Ds* mutagenesis of the gene *ZmMed12a*, identifying two novel insertional alleles, one of which results in a truncation of the ZmMed12a transcript. These insertional mutant alleles will enable determination of the biological roles of the CDK8 module in maize development and stress responses.

## MATERIALS AND METHODS

### Identification of maize CDK8 module genes

Maize CDK8 module genes were identified by BLAST searches using the predicted *Arabidopsis thaliana* protein sequences for HEN3/CDK8 (AT5G63610), CYCC1;1 (At5g48640), CCT/MED12 (At4g00450), and GCT/MED13 (At1g55325) available at TAIR (www.arabidopsis.org). Reciprocal BLAST searches were conducted between all maize and Arabidopsis sequences, to establish that the five maize genes *ZmCDK8*, *ZmCycC*, *ZmMed12a*, *ZmMed12b*, and *ZmMed13* were the only full length CDK8 homologs present in maize.

### Determination of coding sequences for *ZmCDK8*, *ZmCycC*, *ZmMed12a*, *ZmMed12b*, and *ZmMed13*

Multiple mRNA sequences with full-length coding sequences (as well as upstream and downstream untranslated regions) were identified from the NCBI database for both *ZmCDK8* and *ZmCycC*. For *CDK8*, cDNAs for two alternative splice products were identified: EU968864, NM_001157457 and BT018448 correspond to one splice variant, and BT039744 and XR_552425 correspond to the other splice variant. For *CycC*, three independent cDNAs (BT040922, BT033427, and XM008652706) were identified for the one splice variant (shown in Figure 1). Two independent cDNAs (AY105730 and EU972675) represented another *CycC* splice variant with an identical coding sequence but with slight differences in the 3'UTR. A third splice variant was represented by a single cDNA (BT036293); this mRNA has two upstream ORFs, and encodes a truncated CycC protein. For *ZmMed12a*, *ZmMed12b* and *ZmMed13*, partial sequences were obtained from the maize database (maizegdb.org), which were then confirmed and extended by RT-PCR using RNA extracted

from seedlings of the B73 inbred line. To confirm the *ZmMed12a,* Zm*Med12b,* and *ZmMed13* gene models, we amplified cDNA products covering the entire predicted coding regions. Given their large expected size, *ZmMed12a, ZmMed12b,* and *ZmMed13* cDNAs were amplified in multiple over-lapping fragments. Sequencing of cDNA products was generally consistent with gene models based on genomic sequence analysis, except in the case of *ZmMed13*, where a large intron not present in the maize genome sequence was discovered. Coding sequences were deposited in the NCBI database with the following accession numbers: *ZmMed12a* (KP455660), *ZmMed12b* (KP455661), and *ZmMed13* (KP455662).

In addition, numerous short genes that are predicted to encode highly truncated ZmMed12 proteins of 199 to 431 residues were identified (Núñez-Ríos, 2012). These short *ZmMed12* genes are predicted to encode the Med12 domain (pfam09497) and many have corresponding expressed sequence tags (EST) (B73 RefGen_v3), which do not cover the entire body of these short genes. Analysis of genomic sequences around these predicted coding sequences did not identify additional *Med12* exons (data not shown), suggesting that these are indeed truncated versions of *ZmMed12*, and not mis-annotated genes with nearby exons that would constitute the middle and C-terminal portions of Med12 proteins.

**Expression profiles of maize CDK8 module genes**

Expression data from 22 maize tissues were obtained from http://qteller.com/qteller3/ on August 2014, in the form of Fragments Per Kilobase of transcript per Million (FPKM). In order to look for correlations between pairs of genes across the tissues, the data was log2 transformed (first adding 1, to avoid the logarithm of 0) and normalized using the normalizeQuantiles function from the limma package (Bolstad et al., 2003).

The expression values were selected for the 5 CDK8 module genes: *CDK8* (GRMZM2G166771), *CycC* (GRMZM2G408242), *Med12a* (GRMZM2G114459), *Med12b.1* (GRMZM5G828278), *Med12b.2* (GRMZM5G844080), *Med13.1* (GRMZM2G053588), and *Med13.2* (GRMZM2G153792). Since *Med12b.1* and *Med12b.2* as well as *Med13.1* and *Med13.2* are spliced versions of the same gene, the geometric mean was calculated to obtain an averaged estimate of their expression. These data were employed to produce Figure 2A, using the heatmap.2 function from the gplots package (Warns et al., 2015). All pair-wise combinations of the 5 genes across all tissues were plotted using the generic plot function in R (R Core Team, 2015) (Figure S5). The Pearson correlations for all possible pairs of genes were calculated with the cor function, and these data were used as the empirical null to calculate p-values. Correlations for CDK8 module genes were calculated separately. The blob plot in Figure 2B was generated with the corrplot for R.

**Description of maize stocks**

All stocks were maintained in the common genetic background of a color-converted W22 inbred line (Dooner & Kermicle, 1971). A stable source of *Ac* transposase was provided by *Ac-immobilized* (*Ac-im*), an *Ac* derivative which has lost 10bp at the 5' end of the element, preventing excision (Conrad and Brutnell, 2005). Activity of *Ac* transposase was monitored using the mutable *Ds* reporter *r1-sc:m3* that carries a *Ds6*-like insertion in the *r1* locus that controls anthocyanin production in the aleurone and scutellum tissues (Alleman and

Kermicle, 1993): when *Ac* transposase is present, excision of *Ds* from *r1* restores gene function producing colored sectors (Brutnell & Dellaporta, 1994). The donor *Ds* (*dDs*) stock *dDs-B.S07.0835* was generated by isolation of novel transpositions from *r1-sc:m3* as previously described (Vollbrecht et al., 2010). Presence of *dDs-B.S07.0835* was assayed by PCR as previously described (Vollbrecht et al., 2010) using a combination of the *Ds* end primer JSR05 and a primer specific to the genomic site of B.S07.0835 (5'-GACGCACACACGTCAGTATAG-3'). To generate the test-cross population, plants verified as carrying the donor *dDs-B.S07.0835* with *Ac-im* in the genetic background were used as males to pollinate *r1-sc:m3*/*r1-sc:m3* female plants.

**Seedling screen for transposon insertions in *ZmMed12a***

Testcross progeny were germinated and screened for novel insertions of *Ds* in *ZmMed12a* using a PCR-based strategy. Tissue was collected between 7 and 10 days after planting from pools of 10-18 seedlings using a ≈3mm hole punch, and DNA was isolated following a CTAB-based extraction protocol (Weigel and Glazebrook, 2009). A total of 10 *ZmMed12a* gene-specific primers were designed, covering a region extending from 1.8kb upstream of the translational start to the stop codon. These were used in conjunction with the 5' and 3' *Ds*-end primers JSR01 and JGp3, respectively, to amplify DNA adjacent to novel *Ds* insertions in *ZmMed12a* (Table 1). Pools amplifying a product were de-convoluted by screening individuals separately; this second round of PCR used DNA extracted from a different seedling leaf than that sampled for the pool to reduce the chances of recovering somatic transposition events. The PCR products of the second PCR were cleaned (Sambrook and Russell, 2006) and the DNA concentration was adjusted for sequencing by the GENEWIZ Company (South Plainfield, New Jersey, USA). Seedlings carrying putative *med12a* insertional alleles were grown to maturity and propagated by both self-pollination and out-crossing to W22 and B73 inbred lines.

**RT-PCR analysis of *zmmed12a-1::Ds* and *zmmed12a-2::Ds* alleles**

DNA was extracted from 10 day old greenhouse grown seedlings of F2 populations segregating the *1::Ds* and *2::Ds* insertions. Seedlings were genotyped using primers to identify homozygous wild type and homozygous insertion alleles for *1::Ds* (primer pair A5.12F and A5.12R for wild type and A5.12F and JGp3 for *Ds* insertion) and *2::Ds* (primer pair C2.7F and C2.7R for wild type allele and C2.7R and JGp3 for *Ds* insertion). RNA was then extracted Trizol (Invitrogen) for wild type and homozygous insertion alleles. Reverse transcription was performed with SuperScript II (Invitrogen). PCR was performed with the following programs, using Kapa Taq Polymerase (Kapa Biosystems). TNC4-TNC5 primer pair: initial denaturation 95 °C 5'; 10 cycles of 95 °C 30'', 60 °C 30'' (-0.5 °C per cycle), 72 °C 45''; 27 cycles of 95 °C 30'', 55 °C 30'', 72 °C 45''; final extension 72 °C 5′. RS170-RS167 and ZmCDK primer pairs: initial denaturation 95 °C 5'; 30 cycles of 95 °C 30'', 60 °C 30'', 72 °C 1'; 72 °C 10′.

**Table 1: PCR primers used in this study**

| Name | Sequence | Purpose |
|---|---|---|
| B.S07.0835 | 5'-GACGCACACACGTCAGTATAG-3' | Donor Ds site |
| JGp3 | 5'-ACCCGACCGGATCGTATCGG-3' | Ds specific |
| JSR01 | 5'-GTTCGAAATCGATCGGGATA-3' | Ds specific |
| JSR05 | 5'-CGTCCCGCAAGTTAAATATGA-3' | Ds specific |
| 5'UTRF | 5'-TGCACTGCTGCTGTCCTATT-3' | *ZmMed12a* specific- Ds tagging |
| E03R | 5'-TGGTCCATAACTCAGACATACTTGT-3' | *ZmMed12a* specific- Ds tagging |
| E03F | 5'-CTCCCTAATACCCCTGTATTTCA-3' | *ZmMed12a* specific- Ds tagging |
| E07R | 5'-GCATTTGGTAGTAAACAAGAGATGG-3' | *ZmMed12a* specific- Ds tagging |
| E06F | 5'-CCTTGTTAGAATGCGGTCAA-3' | *ZmMed12a* specific- Ds tagging |
| E09.2R | 5'-TCAGGACGAACATACCTAAGCA-3' | *ZmMed12a* specific- Ds tagging |
| INT02F | 5'-ACCAAGTTTGTCAGGTCAACG-3' | *ZmMed12a* specific- Ds tagging |
| E10.2R | 5'-CTACCGAAAACCCATGTTGG-3' | *ZmMed12a* specific- Ds tagging |
| E10.2F | 5'-GCAGCTTTTGAGAGGTTTGAA-3' | *ZmMed12a* specific- Ds tagging |
| E12R | 5'-GCAACTTCCGTCAGCCTTAG-3' | *ZmMed12a* specific- Ds tagging |
| RS170 | CTGGCGAAAGCCTTTTTGAGAAGC | RT-PCR for Ds insertion |
| RS167 | CCCCCACAGGCCCTAACTAAAACA | RT-PCR for Ds insertion |
| TNC4 | CCATATGAGGAACTTCACTCCAG | RT-PCR for Ds insertion |
| TNC5 | ACCTGTACAGAAGTCTGTTAAGCAA | RT-PCR for Ds insertion |
| ZmCDKF | GGAAGGTATGCACAGGACAGAT | RT-PCR for Ds insertion |
| ZmCDKR | TTCAGCACAATCTTGGCAAAAC | RT-PCR for Ds insertion |

| | | |
|---|---|---|
| C2.7 F | ACCCAGGAATCCACTCACTTTT | Genotyping F2 for *2::Ds* |
| C2.7 R | TGCAATCAATAATAGCGTCCAG | Genotyping F2 for *2::Ds* |
| A5.12 F | AACGTGTAGACCTTGGGTTGAAT | Genotyping F2 for *1::Ds* |
| A5.12 R | AGGCGTATAGCGGCTAAGGA | Genotyping F2 for *1::Ds* |

## RESULTS

### The maize genome encodes all four components of the CDK8 module of Mediator

A previous effort to identify Mediator genes from many plant species identified a single maize homolog for all four CDK8 module genes (*CDK8*, *CYCC*, *MED12* and *MED13*) (Mathur et al., 2011). In order to conclusively define the number and identity of CDK8 module homologs in maize, we performed BLAST searches to identify all maize gene-models (B73 reference genome v3; www.maizesequence.org) whose putative protein products exhibit a high degree of similarity to the entire predicted *Arabidopsis* proteins of the CDK8 module of Mediator: CDK8 (encoded by *HEN3*) (Wang and Chen, 2004); CYCC1;1 or CYCC1;2 (Wang et al., 2004); MED12 (encoded by *CCT/CRP*) (Gillmor et al., 2010; Imura et al., 2012); and MED13 (encoded by *GCT/MAB2*) (Gillmor et al., 2010; Ito et al., 2011) (Table 2). Using the translated experimentally verified coding sequences for all maize CDK8 module genes (see below), all potential orthologous relationships were further validated by reciprocal searching of the *Arabidopsis* genome using maize sequences, and by inspection of the next-best-hit in both *Arabidopsis*-to-maize and maize-to-*Arabidopsis* searches (data not shown).

**Table 2: Components of Human, Arabidopsis, and Maize CDK8 modules**

| Human | Hs GenBank mRNA | Arabidopsis | *At* Model[1] | *At* GenBank mRNA | Maize | *Zm* Model[2] | *Zm* GenBank Locus | *Zm* GenBank mRNAs[4] |
|---|---|---|---|---|---|---|---|---|
| CDK8 | P49336 | *HEN3* | AT5G63610.1 | AAT36644 | *ZmCDK8* | GRMZM2G166771 | LOC100284562 | EU968864, NM_001157457, BT018448; BT039744, XR_552425 |
| CYCC | P24863 | *CYCC1;1* | AT5G48640.1 | BX833973 | *ZmCYCC* | GRMZM2G408242 | LOC100193909 | BT040922, BT033427, XM008652706; AY105730, EU972675; BT036293 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | *CYCC1;2* | AT5G48630.1 AT5G48630.2 | AY085977 BT024473 | N/A | N/A | N/A | N/A |
| MED12 | NP_005111 | *CCT/CRP* | AT4G00450.1 | AB690341 | *ZmMed12a* | GRMZM2G114459 | LOC103630556 | KP455660 |
| | | | | | *ZmMed12b* | GRMZM5G828278/ GRMZM5G844080[3] | LOC100384108 | KP455661 |
| MED13 | NM_005121 | *GCT/MAB2* | AT1G55325.2 | N/A | *ZmMed13* | GRMZM2G053588/ GRMZM2G153792[3] | LOC100279[9] / LOC103638[1][3] | KP455662 |

[1]TAIR gene models [www.arabidopsis.org] [2]Maize gene models B73 Reference Genome v3 [maizegdb.org] [3]Split gene annotation [4]Independent mRNAs containing full length coding sequences are listed for each splice product. Different spice products are separated by a semi-colon.

A single maize gene (GRMZM2G166771) was identified as a potential ortholog of *HEN3/CDK8*, and designated *ZmCDK8*. Two different full-length splice products were identified for this gene (EU968864 and BT039744), predicted to encode a full-length and a truncated maize CDK8 protein (Figure 1A; Figure S1). The full-length ZmCDK8 protein is 471 amino acids (AA), and shows 73% identity with the 470 AA Arabidopsis CDK8 protein, and 43% identity with the 464 AA human CDK8 protein (Figure S1). The smaller ZmCDK8 protein is 385 AA, primarily because of a truncation of the C terminal domain, and shows 75% identity with Arabidopsis CDK8, and 43% identity with human CDK8. This truncation occurs after the CDK8 kinase catalytic domain (cd07842), and is thus unlikely to interfere with the kinase function of the protein (Figure S1).

Although Arabidopsis CYCC is encoded by a tandem-duplicated gene pair (Wang et al., 2004), a single potential maize ortholog of *CYCC* (GRMZM2G408242) was identified, and designated *ZmCycC*. Figure 1B shows the splice product represented by the full-length cDNA clone BT040922 (Figure 1B). The 257AA BT040922 protein is 42% identical to human CycC and 67% identical to Arabidopsis CycC1;1 (Figure S2), and contains the Cyclin domain (cd00043) that is present in human and Arabidopsis CycC (Figure S2).

BLAST searches using the Arabidopsis CCT/MED12 protein identified two putative full-length maize genes (GRMZM2G114459 on chromosome 1, and the split gene GRMZM5G828278 / GRMZM5G844080 on chromosome 9), which were designated *ZmMed12a* and *ZmMed12b*. Partial cDNA sequences were publicly available for *ZmMED12a* and *ZmMed12b*; these sequences, as well as coding sequences predicted by the maize database, were used to experimentally determine mRNA sequences for both genes by RT-PCR. The exon-intron structure of both genes is very similar, with the only differences occurring in the length and position of exons 2, 3 and 4 (Figure 1C&D). These splicing differences lead to several small insertions or deletions in the N-terminal portions of the ZmMed12 proteins, with *ZmMed12a* encoding a protein of 2193AA, and *ZmMed12b* encoding a protein of 2202AA; the two ZmMed12 proteins are 91% identical (Figure S3). ZmMed12a is 19% identical to human Med12, and 46% identical to Arabidopsis MED12;

ZmMed12b is 20% identical to human Med12, and 46% identical to Arabidopsis MED12 (Figure S3). The region of highest identity is that comprising the Med12 domain (pfam09497), located at the N-terminus of the Med12 proteins (Figure S3).

A single maize gene was identified corresponding to *GCT/MED13* (split gene GRMZM2G053588 / GRMZM2G153792), and designated *ZmMed13*. Partial cDNA sequences were publicly available for *ZmMed13*; these sequences were used as the basis for RT-PCR experiments to identify full-length mRNA and coding sequences, which demonstrated that *ZmMed13* encodes a protein of 1892 AA, with 20% identity to human Med13, and 49% identity to Arabidopsis MED13 (Figure 1E & Figure S4).

**Maize CDK8 module genes are expressed throughout development**

In other organisms where the CDK8 module has been studied, the gene pairs *CDK8* and *CyclinC*; and *Med12* and *Med13*, have similar expression patterns and mutant phenotypes (Yoda et al., 2005; Loncle et al., 2007; Gillmor et al., 2010; Gillmor et al., 2014). In order to determine whether the *CDK8* / *CycC* and *Med12* / *Med13* genes have similar expression patterns in maize, we used publicly available RNA sequence data to quantify CDK8 module gene expression in different tissues and at different developmental stages (see Materials and Methods). As seen in the heatmap in Figure 2A, *CycC* was expressed at much higher levels in all tissues than the other CDK8 module genes, with *CDK8* and *Med12a* the next highest expressed genes, and *Med13* and *Med12b* with the lowest expression levels

In order to more precisely compare tissue-specific expression between the different CDK8 module genes, we made pairwise comparisons for all five genes (Figure 2B & Figure S5). Expression was most highly correlated for *Med13* and *Med12b* (Pearson's r = 0.93), where the expression ratio between the two genes was close to 1 (compare dotted red line for r, with solid black line representing a 1:1 expression ratio) (Figure 2B & Figure S5). *Med12a* and *Med12b* (r = 0.77); *Med12a* and *Med13* (r = 0.7); and *CDK8* and *Med12a* (r = 0.76) also had high Pearson's coefficients for pairwise comparisons (Figure 2B & Figure S5). By contrast, *CycC* showed almost no correlation with any of the other CDK8 module genes (Figure 2B & Figure S5). The fact that *CycC* shows little expression correlation with the other CDK8 module genes, and is expressed at higher levels than *CDK8*, and many times higher than *Med13*, *Med12a* and *Med12b*, suggests that *CycC* may play more varied roles in development and physiology than the other CDK8 module genes.

**Maize Med12 is encoded by the duplicated gene pair *ZmMed12a* and *ZmMed12b***

The high degree of similarity between *ZmMed12a* and *ZmMed12b* suggests that they are the result of a recent duplication event (Figure S6). *ZmMed12a* and *ZmMed12b* are located in homologous regions of the genome (1S and 9L, respectively), which derive from a polyploidy event that occurred 5-12 million years ago, sometime after the divergence of maize and sorghum lineages. Although gene loss has reduced the number of genes in present-day maize close to pre-duplication levels, in certain cases both syntenic paralogs have been retained (Schnable et al., 2011). Further inspection revealed a sorghum *Med12* gene (Sb01g050260; *SbMed12*) to be present in a region on Chromosome 1L syntenic to the two maize *ZmMed12* containing regions. Moving up- and downstream from *SbMed12*, micro-synteny was conserved, although, typically, for any given sorghum gene only one candidate

ortholog was identified in maize, in either the 1S or 9L region, presumably as the result of gene-loss within paralog pairs following whole genome duplication (Fig. 3).

**Reverse genetics strategies to target maize CDK8 components**

To initiate functional analysis of the maize CDK8 module, we identified publicly available seed stocks carrying *Ac/Ds* or *Mu* family transposons inserted into, or close to, maize CDK8 module encoding genes (Table 3). On the basis of this search, we selected *ZmMed12a* as our first target for reverse genetics: at ~56kb, the closest potential *Ds* donor was nearer to *ZmMed12a* than to any of the other genes. In addition, the availability of a well-characterized *med12* mutant in *Arabidopsis* provides possibility for comparative study (Gillmor et al., 2010; Imura et al., 2012; Gillmor et al., 2014). Finally, the retention of two *Med12* syntenic paralogs in maize suggests that the roles of *ZmMed12a* and *ZmMed12b* are functionally different, a question which can be addressed by characterization of maize *med12* mutant alleles.

**Table 3: Reverse genetics resources for maize CDK8 module genes**

| Maize | Maize Accession | Position (kb) | Closest *Ac/Ds*[1] | Uniform *Mu*[2,3] | *Mu* Illumina[2] |
|---|---|---|---|---|---|
| *ZmCDK8* | GRMZM2G166771 | Chr5: 45,538,294-45,544,117 | 8311.3 kb | 3' UTR | 5' UTR |
| *ZmCYCC* | GRMZM2G408242 | Chr7:137,095,918-137,100,956 | 420 kb | Upstream | 5' UTR |
| *ZmMED12a* | GRMZM2G114459 | Chr1: 2,088,572-2,102,312 | 56.2 kb | Upstream (3), 5' UTR (4), Int | |
| *ZmMED12b* | GRMZM5G828278 /GRMZM5G844080 | Chr9: 155,361,528-155,373,747 | 717.1 kb | 5' UTR (3), Ex | |
| *ZmMED13* | GRMZM2G053588 /GRMZM2G153792 | Chr9: 28,392,287-28,413,513 | 226.7 kb | | |

[1]acdstagging.org; [2]*Mutator* resources available at maizeGDB.org; [3]Upstream indicates within 1kb 5' of the translational start; numbers in parentheses indicate multiple insertions

**Identification of novel *Ds* insertions into *ZmMed12a***

To use the *Ac/Ds* transposon system to generate mutant alleles of *ZmMed12a*, we first obtained donor *Ds* (*dDs*) stocks carrying the *Ds* element *dDs-B.S07.0835*, located 56.2 kb from *ZmMed12a* (acdstagging.org). The position of the linked *Ds* element was confirmed by PCR assay (see Materials and Methods) (Conrad and Brutnell, 2005). Presence of *Ac-im* in testcross progenitor seed stocks was monitored by somatic excision of a second *Ds* from the *r1-sc:m3* marker locus, resulting in variegated spotting of the kernel aleurone and scutellar tissues (Figure 4A & B). Spotted kernels were planted and seedlings genotyped for the presence of *dDs* using a PCR assay (Materials and Methods). To generate novel germinal insertions into *ZmMed12a*, individuals carrying the *dD* and the *Ac-im* transposase source were used as males to pollinate T43 (*r-sc:m3*/*r-sc:m3*) females. A test cross population of 59

ears was obtained for the *ZmMed12a* screen (Figure 4A&B).

The test-cross population was screened for *Ds* insertions in *ZmMed12a* using combinations of gene specific and *Ds* specific PCR primers (see Materials and Methods). Pools of 10-18 seedlings were assayed for amplification of putative *Ds*-flanking junction products (see Figure 4C for example for the *Zmmed12a-2::Ds* insertion). Seedlings constituting the pools from which products were amplified were re-screened separately to identify positive individuals (Figure 4D). This second PCR was performed using DNA extracted from a leaf different from that used for the pool PCR to reduce the rate at which we recovered somatic transposition events. We screened a total of 3,049 seedlings and identified two novel insertions into *ZmMed12a*: *zmmed12a-1::Ds*, located 918bp upstream of the translational start, and *zmmed12a-2::Ds* located in exon 10 (Figure 1C). We performed additional PCR reactions to recover both flanks of the *zmmed12a-1::Ds* and *zmmed12a-2::Ds* insertions. Flanking DNA products were sequenced, confirming the location of the insertions and identifying characteristic 8bp target site duplications. The seedlings carrying the two novel *zmmed12a* insertional alleles were grown to maturity and propagated by both self-pollination and out-crossing. Progeny were germinated and genotyped, confirming the heritability of novel *Ds* insertions (Figure 4E).

**The *2::Ds* insertion results in a truncated *ZmMed12a* transcript**

In order to determine the effect of these novel *Ds* insertions on the *ZmMed12a* gene, we performed RT-PCR analysis of plants homozygous for the wild type and *Ds* alleles, using primer pairs that amplify fragments in exon 9 (downstream of the *1::Ds* insertion, and upstream of the *2::Ds* insertion), and exon 12 (downstream of both *Ds* insertions) (Figure 4F). Both primer sets amplified fragments from wild type and the *zmmed12a-1::Ds* allele, suggesting that the 1::Ds insertion has no significant effect on the ZmMed12a transcript, a result which is not surprising, since this Ds insertion is upstream of exon 1. In the case of the *zmmed12a-2::Ds* allele, the primer pair in exon 9 produced an amplification product, while the primer pair in exon 12 showed no aplification from homozygous *2::Ds* plants. This result demonstrates that the *2::Ds* insertion causes production of a truncated version of the *ZmMed12a* transcript, likely causing a loss of function of the *ZmMed12a* gene.

## DISCUSSION

In this study we have identified the five genes encoding the CDK8 module of Mediator in maize, determined their coding sequences, characterized their expression in maize tissues during development, and examined the synteny of maize and sorghum in the region of the *Med12* genes. Additionally, we have mutagenized the *ZmMed12a* gene using the *Ac/Ds* transposon system created by Vollbrecht et al. (2010).

In our analysis of CDK8 module genes, we identified two alternative transcripts for *CDK8* (Figure 1). One predicted CDK8 protein is significantly shorter than the other, lacking the C-terminal 86 AA. This truncation seems unlikely to affect enzyme activity *per se*, as the kinase domain is intact (Figure S1). However, the lack of this domain may alter regulation of the kinase activity. Alternatively, the truncation may modify the interaction of CDK8 with

CycC, or affect the formation of the four protein CDK8 complex. This complex sterically inhibits the interaction of Core Mediator with RNA pol II, by making direct contact with Core Mediator (Tsai et al., 2013). In the case of CycC, one only one isoform was represented by multiple independent cDNAs. Only single splice products were identified for *Med12a*, *Med12b* and *Med13* (Figure 1). One explanation for this is that there is indeed only one splice product for each gene in maize. It is also possible that the very large size of the mRNAs for these three genes (6-7 kb) makes cloning of multiple splice products difficult, due to technical difficulties in cloning such large cDNAs.

In our analysis of the relative expression of CDK8 module genes, we found *CDK8* and *CycC* to be more highly expressed in all tissues than *Med12a*, *Med12b* or *Med13*. In particular, *CycC* showed the highest expression in all tissues, consistently 3-4 times higher even than *CDK8* (Figure 2). This increased expression of *CycC* is consistent with roles of CycC beyond regulating transcription in tandem with CDK8 (the best known role for CycC) (Allen and Taatjes, 2015). In addition to regulation of transcription, CycC has been shown to promote the G0 to G1 cell cycle transition through phosphorylation of Retinoblastoma, allowing quiescent cells to enter the cell cycle. CycC achieves this through interaction with CDK3, a kinase that is not associated with transcriptional activation, but instead promotes cell cycle entry (Ren and Rollins, 2004). CycC has also been demonstrated to be a haploinsuficient tumor suppressor in mammals, whose loss of function in mice is lethal during embryogenesis (Li et al., 2014). The haploinsuficiency of CycC may require its mRNA or protein levels to be stably maintained, suggesting an explanation for its high levels in all the tissues that we examined (Figure 2 & Figure S6). *Med12a*, *Med12b*, and *Med13* show much lower expression levels, which also vary considerably between different tissues (Figure 2 & Figure S6). The similar expression profiles for *Med12* and *Med13* in maize are consistent with Arabidopsis, where similar expression profiles for these two genes were reported (Gillmor et al., 2010; Ito et al., 2011; Imura et al., 2012; Gillmor et al., 2014). The widely varying expression levels for *Med12* and *Med13* in different tissues are consistent with various roles for these genes in development, both in primordia (where they show the highest expression), as well as in differentiating and more mature tissue.

In Arabidopsis, *MED12* is a single copy gene, with mutant phenotypes in both development and pathogen responses (Gillmor et al., 2010; Imura et al., 2012; Gillmor et al., 2014; Zhu et al., 2014). In maize, however, two *Med12* genes were identified. Sometime after divergence with sorghum, the maize lineage underwent whole genome duplication (Schnable et al., 2011). While in the majority of cases resulting additional gene copies have been lost, for ~10% of the original gene set syntenic paralog pairs have been retained (Hughes et al., 2015). The genomic location of *ZmMed12a* and *ZmMed12b* is consistent with them representing such a paralog pair. In the region of synteny between maize and sorghum, other genes surrounding *Med12* have been reduced to a single copy, suggesting that the retention of both paralogs of *Med12* in maize may have functional significance. Our isolation of the *2::Ds* insertional allele of *ZmMed12a* will allow us to test the functional importance of this gene. The truncation of the *ZmMed12a* transcript in the *2::Ds* allele makes it very likely that this allele causes a loss of function: a T-DNA insertion in a similar location of the *CCT* (*MED12*) gene of Arabidopsis causes a strong loss of function phenotype, even when some aberrant transcript is produced (Gillmor et al., 2010; Gillmor et al., 2014)

One additional advantage of *Ds* as a mutagen is that novel transpositions occur into linked sites, meaning that the *Ds* insertions in *ZmMed12a* can be remobilized to create further allelic variation in *ZmMed12a*. In addition to mutant alleles that cause a complete loss of function, subsequent *Ds* mutagenesis of *ZmMed12a* may result in hypomorphic alleles that either reduce (but do not eliminate) the function of *ZmMed12a*, or that inactivate specific functional domains of Med12. Alleles that eliminate only certain parts of the Med12 protein could be especially useful in understanding the function of different domains of Med12, currently one of the most interesting, and least explored, aspects of Mediator biology.

## AUTHOR CONTRIBUTIONS

Study designed by TN-R, KA, TPB, SG and RS. Data acquired and/or analyzed by TN-R, KA, ALA-N, DL-S, CM-C, MG-A, SG, RS. Manuscript written by TN-R, SG and RS, and approved by all authors.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## FUNDING

This study was funded by CINVESTAV institutional funds to CSG and RJHS, CONACyT CB 2009 No. 133990 to MGA, CONACyT CB 2012 No. 151947 to RS, and NSF IOS-0922701 to TPB. CONACyT graduate fellowships supported TN-R, ALA-N (No. 339468), DL-S (No. 262808), and CM-C (No. 20729).

## ACKNOWLEDGEMENTS

Thanks to Cei Abreu-Goodger for advice on analysis of gene expression data, and to Jessica Carcaño-Macías for managing seed stocks.

## SUPPLEMENTAL MATERIAL

Supplemental Figures S1-S6.

## REFERENCES

Alleman, M., and Kermicle, J. L. (1993). Somatic variegation and germinal mutability reflect the position of transposable element Dissociation within the maize R gene. *Genetics* 135, 189–203.

Allen, B.L. and Taatjes, D.J. (2015). The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* 16, 155-166. doi:10.1038/nrm3951

Bäckström, S., Elfving, N., Nilsson, R., Wingsle, G., & Björklund, S. (2007). Purification of a plant mediator from *Arabidopsis thaliana* identifies PFT1 as the Med25 subunit. *Molecular Cell* 26, 717–729. http://doi.org/10.1016/j.molcel.2007.05.007

Bai, L., Singh, M., Pitt, L., Sweeney, M., and Brutnell, T. P. (2007). Generating novel allelic variation through Activator insertional mutagenesis in maize. *Genetics* 175, 981–992. doi:10.1534/genetics.106.066837.

Björklund, S., and Gustafsson, C.M. (2005). The yeast Mediator complex and its regulation. *Trends Biochem Sci* 30, 240-244. doi:10.1016/j.tibs.2005.03.008

Bolduc, N., Yilmaz, A., Mejia-Guerra, M. K., Morohashi, K., O'Connor, D., Grotewold, E., et al. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev* 26, 1685–1690. doi:10.1101/gad.193433.112.

Bolstad, B. M., Irizarry R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinformatics* 19, 185-193.

Bourbon, H.-M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Research* 36, 3993-4008. doi:10.1093/nar/gkn349

Brutnell, T. P., and Dellaporta, S. L. (1994). Somatic inactivation and reactivation of Ac associated with changes in cytosine methylation and transposase expression. *Genetics* 138, 213–225.

Brutnell, T.P. and Conrad L.J. (2003). Transposon tagging using Activator (Ac) in maize. *Methods Mol Biol* 236, 157–176. doi:10.1385/1-59259-413-1:157.

Calvi, B. R., Hong, T. J., Findley, S. D., and Gelbart, W. M. (1991). Evidence for a common evolutionary origin of inverted repeat transposons in Drosophila and plants: hobo, Activator, and Tam3. *Cell* 66, 465–471.

Candela, H., and Hake, S. (2008). The art and design of genetic screens: maize. *Nat Rev Genet* 9, 192–203. doi:10.1038/nrg2291.

Conrad, L. J., and Brutnell, T. P. (2005). Ac-immobilized, a stable source of Activator transposase that mediates sporophytic and gametophytic excision of Dissociation elements in maize. *Genetics* 171, 1999–2012. doi:10.1534/genetics.105.046623.

Davidson, R. M., Hansey, C. N., Gowda, M., Childs, K. L., Lin, H., Vaillancourt, B., et al. (2011). Utility of RNA Sequencing for Analysis of Maize Reproductive Transcriptomes. *The Plant Genome Journal* 4, 191–13. doi:10.3835/plantgenome2011.05.0015.

Davis, M. A., Larimore, E. A., Fissel, B. M., Swanger, J., Taatjes, D. J., & Clurman, B. E. (2013). The SCF-Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. *Genes & Development* 27, 151–156. doi:10.1101/gad.207720.112

523 Dooner, H. K., and Kermicle, J. L. (1971). Structure of the R tandem duplication in maize.
524 *Genetics* 67, 427–436.

525 Dooner, H.K. and Belachew, A. (1989). Transposition Pattern of the Maize Element Ac from
526 the Bz-M2(ac) Allele. 122, 447–457.

527 Elmlund, H., Baraznenok, V., Lindahl, M., Samuelsen, C.O., Koeck, P.J.B., Holmberg, S.,
528 Hebert, H., and Gustafsson, C.M. (2006). The cyclin-dependent kinase 8 module sterically
529 blocks Mediator interactions with RNA polymerase II. *Proc Natl Acad Sci USA* 103, 15788-
530 15793. doi:10.1073/pnas.0607483103

531 Flanagan, P. M., Kelleher, R. J., Sayre, M. H., Tschochner, H., & Kornberg, R. D. (1991). A
532 mediator required for activation of RNA polymerase II transcription in vitro. *Nature* 350,
533 436–438. http://doi.org/10.1038/350436a0

534 Furutani, M., Vernoux, T., Traas, J., Kato, T., Tasaka, M., & Aida, M. (2004). PIN-
535 FORMED1 and PINOID regulate boundary formation and cotyledon development in
536 Arabidopsis embryogenesis. *Development* 131, 5021–5030. doi:10.1242/dev.01388

537 Gillmor, C.S., Park, M.-Y., Smith, M.R., Pepitone, R., Kerstetter, R.A., and Poethig, R.S.
538 (2010). The MED12-MED13 module of Mediator regulates the timing of embryo patterning
539 in Arabidopsis. *Development* 137, 113-122. doi:10.1242/dev.043174

540 Gillmor, C.S., Silva-Ortega, C.O., Willmann, M.R., Buendía-Monreal, M., and Poethig, R.S.
541 (2014). The Arabidopsis Mediator CDK8 module genes CCT (MED12) and GCT (MED13)
542 are global regulators of developmental phase transitions. *Development* 141, 4580-4589. doi:
543 10.1242/dev.111229

544 Greenblatt, I. M. (1984). A chromosome replication pattern deduced from pericarp
545 phenotypes resulting from movements of the transposable element, modulator, in maize.
546 *Genetics* 108, 471–485.

547 Imura, Y., Kobayashi, Y., Yamamoto, S., Furutani, M., Tasaka, M., Abe, M., and Araki, T.
548 (2012). *CRYPTIC PRECOCIOUS/MED12* is a novel flowering regulator with multiple target
549 steps in Arabidopsis. *Plant Cell Physiol.* 53, 287-303. doi:10.1093/pcp/pcs002

550 Ito, J., Sono, T., Tasaka, M., and Furutani, M. (2011). *MACCHI-BOU 2* is required for early
551 embryo patterning and cotyledon organogenesis in Arabidopsis. *Plant Cell Physiol.* 52, 539-
552 552. doi:10.1093/pcp/pcr013

553 Ito, J., Fukaki, H., Onoda, M., Li, L., Li, C., Tasaka, M., Furutani, M., (2016). Auxin-
554 dependent compositional change in Mediator in ARF7- and ARF19-mediated transcription.
555 *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1600739113

556 Kelleher, R. J., Flanagan, P. M., & Kornberg, R. D. (1990). A novel mediator between
557 activator proteins and the RNA polymerase II transcription apparatus. *Cell*, 61, 1209–1215.

Knuesel, M. T., Meyer, K. D., Bernecky, C., & Taatjes, D. J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes & Development*, 23, 439–451. http://doi.org/10.1101/gad.1767009

Kunze, R., H. Saedler, and W.-E. Lonnig. (1997). Plant Transposable Elements. in *Advances in Botanical Research,* Academic Press ISBN 978-0-12-005927-0, vol 27, pp350-387.

Larivière, L., Seizl, M., and Cramer, P. (2012). A structural perspective on Mediator function. *Curr Opin Cell Biol*, 24, 305-313. doi:10.1016/j.ceb.2012.01.007

Li P., Ponnala L., Gandotra N., Wang L., Si Y., Tausta S. L., Kebrom T. H., Provart N., Patel R., Myers C. R., Reidel E. J., Turgeon R., Liu P., Sun Q., Nelson T. & Brutnell T. P. (2010). The developmental dynamics of the maize leaf transcriptome. *Nature Genetics* 42, 1060–1067. doi:10.1038/ng.703.

Li, N., Fassl, A., Chick, J., Inuzuka, H., Li, X., Mansour, M. R., et al. (2014). Cyclin C is a haploinsufficient tumour suppressor. *Nat Cell Biol* 16, 1080–1091. doi:10.1038/ncb3046.

Loncle, N., Boube, M., Joulia, L., Boschiero, C., Werner, M., Cribbs, D. L., & Bourbon, H.-M. (2007). Distinct roles for Mediator Cdk8 module subunits in Drosophila development. *EMBO Journal* 26, 1045–1054. doi:10.1038/sj.emboj.7601566

Mathur, S., Vyas, S., Kapoor, S., and Tyagi, A.K. (2011). The Mediator complex in plants: structure, phylogeny, and expression profiling of representative genes in a dicot (Arabidopsis) and a monocot (rice) during reproduction and abiotic stress. *Plant Physiol.* 157, 1609-1627. doi:10.1104/pp.111.18830

McCarty D. R., and Meeley R. B., (2009). Transposon resources for forward and reverse genetics in maize, in *Handbook of Maize: Genetics and Genomics*, ed. Bennetzen J. L., Hake S., editors. Springer, New York. pp. 561–584

Ng, S., Giraud, E., Duncan, O., Law, S. R., Wang, Y., Xu, L., et al. (2013). Cyclin-dependent kinase E1 (CDKE1) provides a cellular switch in plants between growth and stress responses. *Journal of Biological Chemistry* 288, 3449–3459. doi:10.1074/jbc.M112.416727

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., et al. (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature* 518, 376-380. doi:10.1038/nature14229

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ren, S., and Rollins, B. J. (2004). Cyclin C/cdk3 promotes Rb-dependent G0 exit. *Cell* 117, 239–251.

Robinson, P. J., Trnka, M. J., Pellarin, R., Greenberg, C. H., Bushnell, D. A., Davis, R., et al. (2015). Molecular architecture of the yeast Mediator complex. *Elife* 4. doi:10.7554/eLife.08719.

Sambrook, J., and Russell, D. W. (2006). Purification of PCR products in preparation for

595 cloning. *CSH Protoc* 2006, pdb.prot3825. doi:10.1101/pdb.prot3825.

596 Schnable, J. C., Springer, N. M., and Freeling, M. (2011). Differentiation of the maize
597 subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of*
598 *the National Academy of Sciences* 108, 4069–4074. doi:10.1073/pnas.1101368108.

599 Tsai, K.-L., Sato, S., Tomomori-Sato, C., Conaway, R. C., Conaway, J. W., & Asturias, F. J.
600 (2013). A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA
601 polymerase II interaction. *Nature Structural & Molecular Biology* 20, 611–619.
602 http://doi.org/10.1038/nsmb.2549

603 Tsai, K.-L., Tomomori-Sato, C., Sato, S., Conaway, R. C., Conaway, J. W., and Asturias, F. J.
604 (2014). Subunit architecture and functional modular rearrangements of the transcriptional
605 mediator complex. *Cell* 157, 1430–1444. doi:10.1016/j.cell.2014.05.015.

606 Vollbrecht, E., Duvick, J., Schares, J. P., Ahern, K. R., Deewatthanawong, P., Xu, L., et al.
607 (2010). Genome-wide distribution of transposed Dissociation elements in maize. *Plant Cell*
608 22, 1667–1685. doi:10.1105/tpc.109.073452.

609 Wang, W. and Chen, X. (2004). *HUA ENHANCER3* reveals a role for a cyclin-dependent
610 protein kinase in the specification of floral organ identity in Arabidopsis. *Development* 131,
611 3147-3156. doi:10.1242/dev.01187

612 Wang, X., Elling, A. A., Li, X., Li, N., Peng, Z., He, G., et al. (2009). Genome-wide and
613 organ-specific landscapes of epigenetic modifications and their relationships to mRNA and
614 small RNA transcriptomes in maize. *Plant Cell* 21, 1053–1069. doi:10.1105/tpc.109.065714.

615 Wang, G., Kong, H., Sun, Y., Zhang, X., Zhang, W., Altman, N., DePamphilis, C.W., and
616 Ma, H. (2004). Genome-wide analysis of the cyclin family in Arabidopsis and comparative
617 phylogenetic analysis of plant cyclin-like proteins. *Plant Physiol.* 135, 1084-1099. doi:
618 10.1104/pp.104.040436

619 Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T.,
620 Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., and Venables, B. (2013). gplots:
621 Various R programming tools for plotting data. R package version 2.12.1 available at
622 http://CRAN.R-project.org/package=gplots

623 Waters, A. J., Makarevitch, I., Eichten, S. R., Swanson-Wagner, R. A., Yeh, C.-T., Xu, W., et
624 al. (2011). Parent-of-origin effects on gene expression and DNA methylation in the maize
625 endosperm. *THE PLANT CELL ONLINE* 23, 4221–4233. doi:10.1105/tpc.111.092668.

626 Weigel, D., and Glazebrook, J. (2009). Quick miniprep for plant DNA isolation. *Cold Spring*
627 *Harb Protoc* 2009, pdb.prot5179–pdb.prot5179. doi:10.1101/pdb.prot5179.

628 Wu, G., Park, M. Y., Conway, S. R., Wang, J.-W., Weigel, D., & Poethig, R. S. (2009). The
629 sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell*
630 138, 750–759. doi:10.1016/j.cell.2009.06.031

631 Yin, J.-W. and Wang, G. (2014). The Mediator complex: a master coordinator of
632 transcription and cell lineage development. *Development* 141, 977-987.

633 doi:10.1242/dev.098392

634 Yoda, A., Kouike, H., Okano, H., Sawa, H., (2005). Components of the transcriptional
635 Mediator complex are required for asymmetric cell division in C. elegans. Development 132,
636 1885–1893. doi:10.1242/dev.01776

637 Zhu, Y., Schluttenhoffer, C. M., Wang, P., Fu, F., Thimmapuram, J., Zhu, J.-K., et al. (2014).
638 CYCLIN-DEPENDENT KINASE8 differentially regulates plant immunity to fungal
639 pathogens through kinase-dependent and -independent functions in Arabidopsis. *Plant Cell*
640 26, 4149–4170. doi:10.1105/tpc.114.128611

641

642 **FIGURE LEGENDS**

643 **Figure 1. The CDK8 module of maize consists of *CDK8*, *CyclinC*, *Med12a*, *Med12b*, and**
644 ***Med13*** (A) Exon-intron structure for two different splice products (EU968864 and
645 BT039744) of the *ZmCDK8* gene (GRMZM2G166771). EU968864 encodes a 471AA protein,
646 while BT039744 encodes a 385AA protein, truncated after the CDK8 kinase catalytic domain
647 (cd07842) (B) Exon-intron structure of mRNA sequence BT040922 for the *ZmCycC* gene
648 (GRMZM2G408242), encoding a predicted protein of 257AA. (C) *ZmMed12a*
649 (GRMZM2G114459) encodes a 2193 AA protein (mRNA sequence KP455660). The
650 location of the *Ds* insertions *Zmmed12a-1::Ds* (918bp upstream of ATG) and *Zmmed12a-*
651 *2::Ds* (exon 10, at bp 4,236 of coding sequence) are indicated. The orientation of each *Ds*
652 insertion is represented by the triangle above the gene (5'-3'), and below (3'-5'). (D)
653 *ZmMed12b* (split gene GRMZM5G828278 / GRMZM5G844080) encodes a 2202 AA
654 predicted protein (mRNA sequence KP455661). (E) *ZmMed13* (split gene
655 GRMZM2G053588 / GRMZM2G153792) encodes a 1892 AA protein (mRNA sequence
656 KP455662). Intron 11 of *ZmMed13* is of unknown size (dotted lines), as it spans a gap in the
657 maize genome sequence. Intron sizes for all genes were determined using corresonding maize
658 genomic sequence. Exons are represented by black boxes, untranslated regions by open boxes,
659 introns by solid black lines, and genomic sequence of ZmMed12a upstream of start codon as
660 solid grey line.

661 **Figure 2. CDK8 module genes are broadly expressed in development.** (A) Expression of
662 *Med13*, *Med12a*, *Med12b*, *CDK8* and *CycC* are shown as log2 (FPKM+1) (Fragments Per
663 Kilobase of exon per Million reads mapped). Data are from the following sources: Mature
664 tassel, Developing ear, Ovule, Seed 5 dap, Seed 10 dap, Embryo 25 dap, Endosperm 25 dap,
665 Silk, Developing tassel, Ear, Seedling leaves field, Seedling leaves gc (growth chamber) from
666 Davidson et al. (2011). Developing leaf and Mature leaf from Li et al. (2010). Seedling roots
667 and Seedling shoots from Wang et al. (2009). Embryo 14 dap and Endosperm 14 dap from
668 Waters et al. (2011). Shoot apex, Ear primordia, Tassel primordia and Leaf primordia from
669 Bolduc et al. (2012). (B) Correlation of expression patterns for pairwise combinations of
670 members of CDK8 module. Positive correlations are shown as blue circles, with larger circles
671 and darker blue signifying greater correlations between the two genes. Gene-by-gene
672 comparisons for all tissue samples are shown in Figure S5, from which r values to make this
673 plot were taken.

**Figure 3. Synteny between maize and sorghum genomic regions surrounding *Med12*.** The *Med12* gene is conserved across sorghum and maize syntenic regions. Upper and lower rows: annotated genes in syntenic regions on maize chromosome 1S (*Zm*1S) at ~2Mb (upper row), and maize chromosome 9L (*Zm*9L ) at ~155Mb (lower row). Middle row: annotated genes in the region of *SbMed12* at ~73Mb on sorghum chromosome 1L (*Sb*1L). Orthologous genes are connected by dashed lines. *Sb*1L and *Zm*9L run left to right, *Zm*1S runs right to left. Genes are shown as black boxes, and the chromosomes are represented by vertical lines. Regions shown to scale, with the right hand position corresponding to the chromosome location mentioned.

**Figure 4. Generation of *Ds* insertional alleles of *ZmMed12a*, and their effect on the *ZmMed12a* transcript.** (A) Crossing scheme for generating plants homozygous for the *r-sc:m3* reporter allele, and heterozygous for *Ac-immobilized* (*Ac-im*) and the *Ds* insertion linked to *ZmMed12a*. (B) The presence of *r-sc:m3* allows for selection of spotted F1 kernels, indicating the presence of *Ac-im*, required for remobilizing the *Ds* insertion linked to *ZmMed12a*. (C) Initial pools of 10-18 seedlings were screened by PCR for the presence of a *Ds* element in *ZmMed12a*, using gene specific primer E10.2 and *Ds* specific primer JGp3. A 1.8 kb fragment (yellow arrow) was amplified from pool C2. (D) Individual plants from pool C2 were tested for the presence of the same fragment, which was amplified from plant 7, denominated C2.7 (*Zmmed12a-2::Ds*) (E) This 1.8kb band segregated in nine progeny of the selfed plant C2.7, demonstrating that it is a heritable germinal insertion. (F) RT-PCR analysis of the effect of the *1::Ds* and *2::Ds* insertions on the *ZmMed12a* transcript. The *1::Ds* insertion has no detectable effect on the stability of the *ZmMed12a* transcript, while the *2::Ds* insertion creates a transcript that is truncated after the *Ds* insertion. The location of the primer pairs used to amplify the transcript are indicated in the gene diagram. *ZmCDK* (GRMZM2G149286) was used as a control gene. One representative experiment of three biological replicates is shown.

Figure 1.TIF



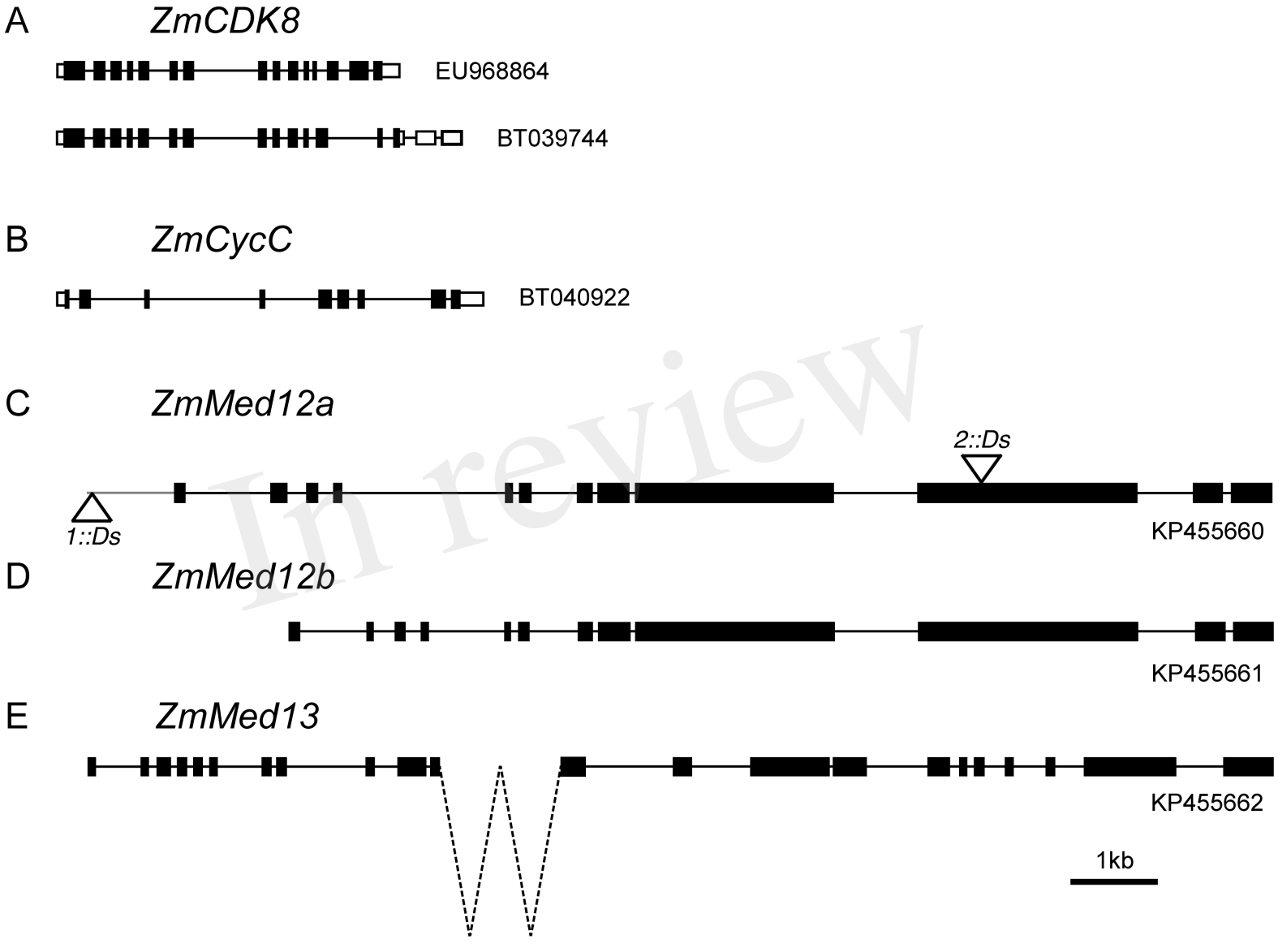


Figure 1

Figure 2.TIF



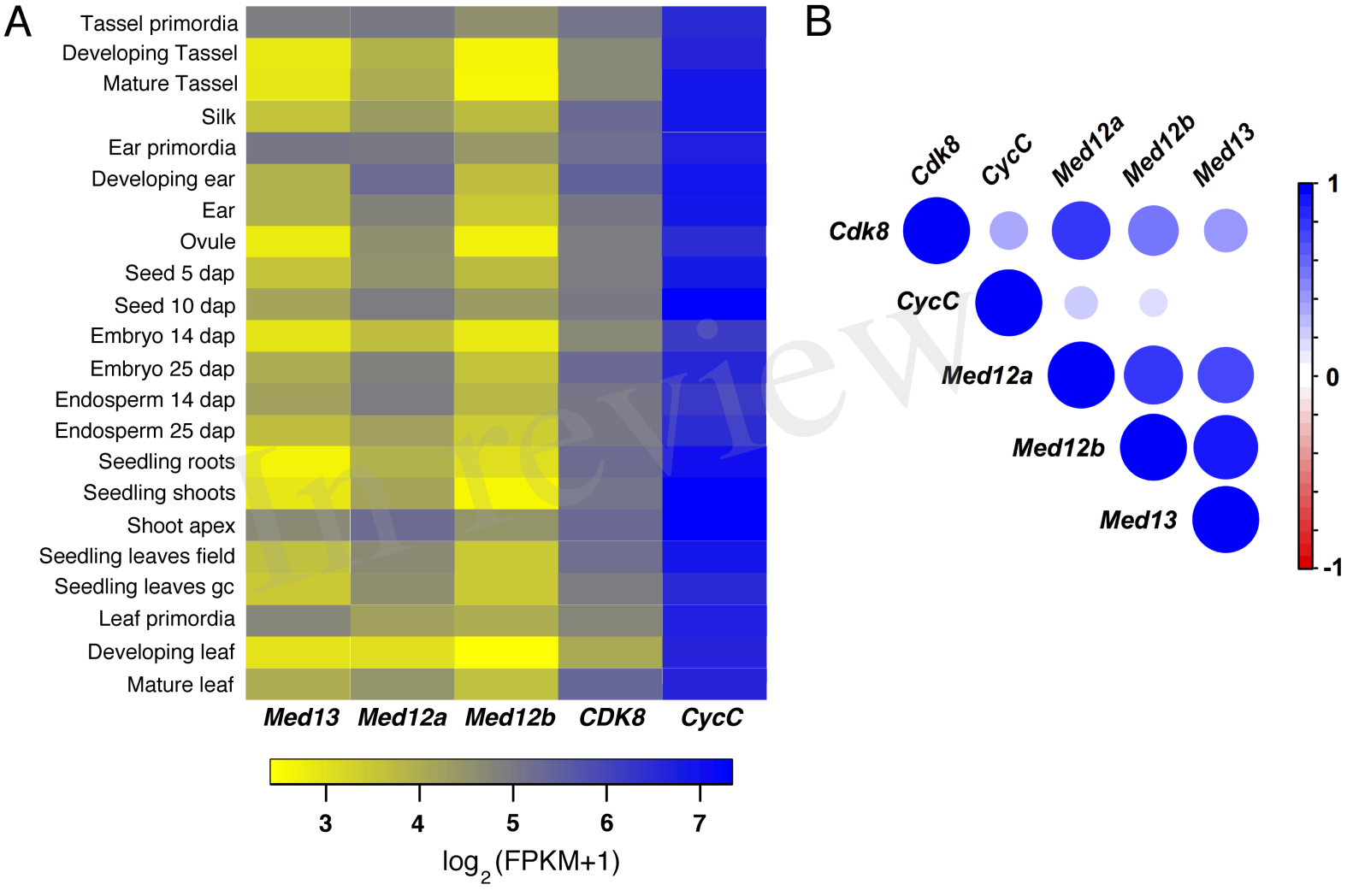


Figure 2

Figure 3.TIF



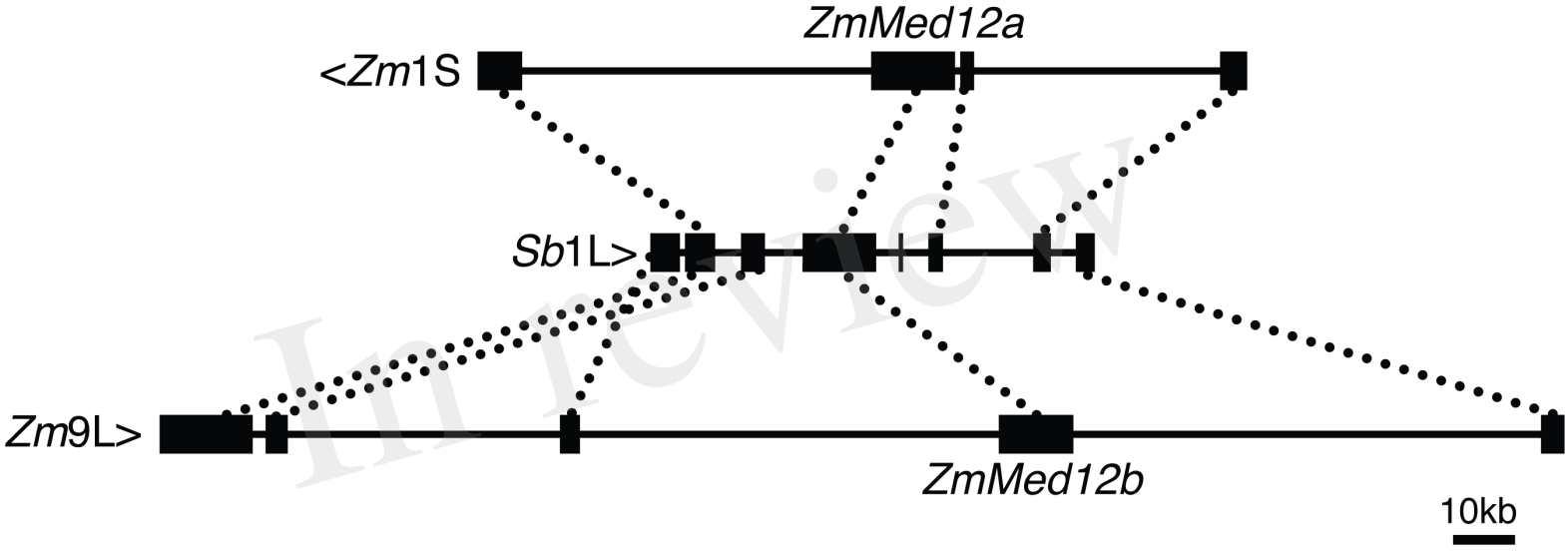




Figure 3

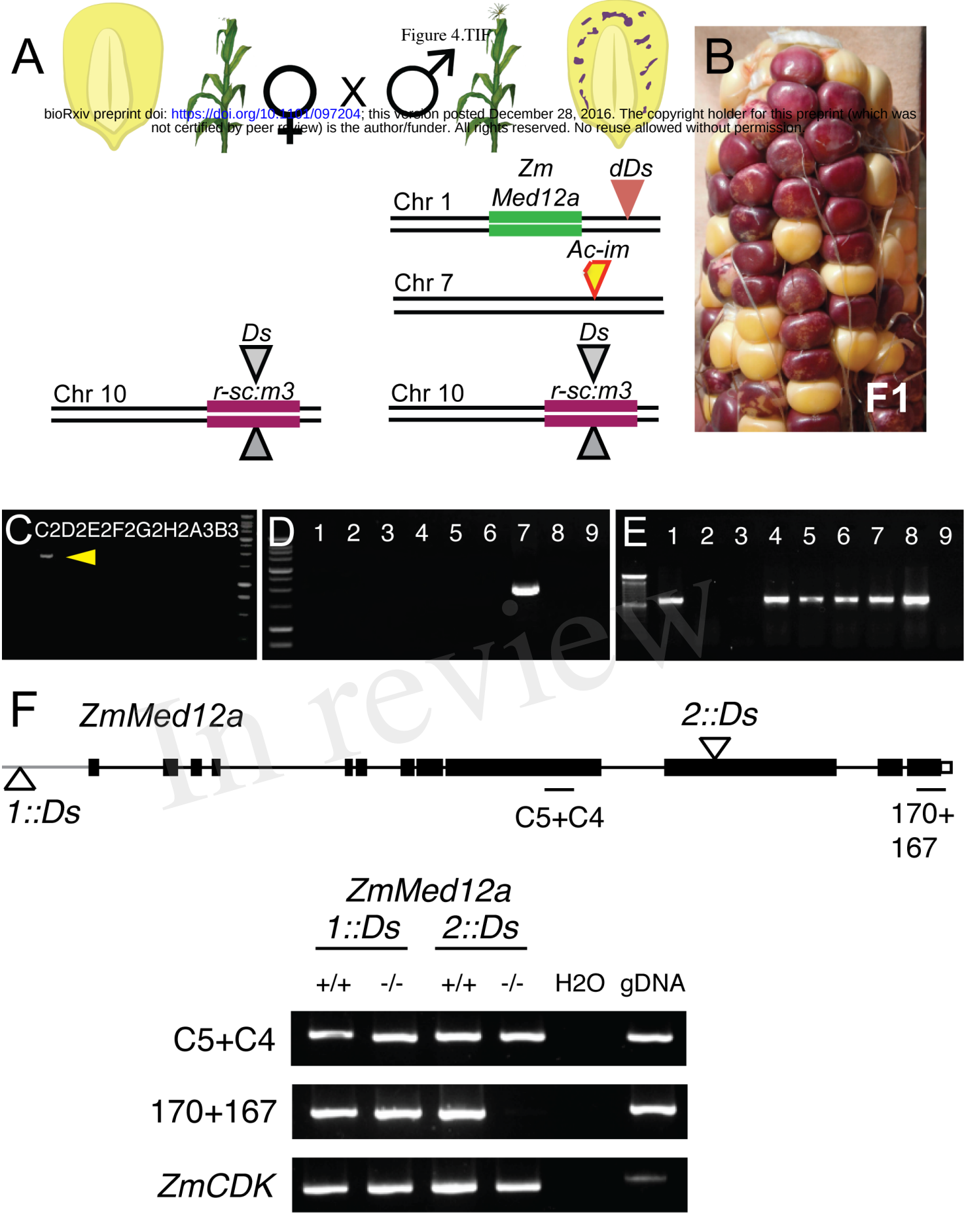




Figure 4