Germline Fate through Activation of the Transcription Factor DUO1 Promoter

Benjamin Peters, Jonathan Casey, Jack Aidley, Stuart Zohrab, Michael Borg3, David Twell2*, and Lynette Brownfield1

Department of Biochemistry, University of Otago, Dunedin 9016, New Zealand (B.P., J.C., S.Z., L.B.); and Department of Genetics, University of Leicester, Leicester LE1 7RH, United Kingdom (J.A., M.B., D.T.)

ORCID IDs: 0000-0001-6168-4694 (B.P.); 0000-0003-4168-1503 (J.A.); 0000-0003-0483-1461 (D.T.); 0000-0002-3490-5444 (L.B.).

The development of the male germline within pollen relies upon the activation of numerous target genes by the transcription factor DUO POLLEN1 (DUO1). The expression of DUO1 is restricted to the male germline and is first detected shortly after the asymmetric division that segregates the germ cell lineage. Transcriptional regulation is critical in controlling DUO1 expression, since transcriptional and translational fusions show similar expression patterns. Here, we identify key promoter sequences required for the germline-specific regulation of DUO1 transcription. Combining promoter deletion analyses with phylogenetic footprinting in eudicots and in Arabidopsis accessions, we identify a cis-regulatory module, Regulatory region of DUO1 (RODI), which replicates the expression pattern of DUO1 in Arabidopsis (Arabidopsis thaliana). We show that ROD1 from the legume Medicago truncatula directs male germline-specific expression in Arabidopsis, demonstrating conservation of DUO1 regulation among eudicots. ROD1 contains several short conserved cis-regulatory elements, including three copies of the motif DNGTGGV, required for germline expression and tandem repeats of the motif YAACYGY, which enhance DUO1 transcription in a positive feedback loop. We conclude that a cis-regulatory module conserved in eudicots directs the spatial and temporal expression of the transcription factor DUO1 to specify male germline fate and sperm cell differentiation.

Pollen grains, the haploid male gametophytes, play a key role in flowering plant sexual reproduction through the generation and delivery of two sperm cells to the embryo sac to enable double fertilization. Pollen develops in the anther locules from haploid microspores formed after meiosis. The microspores complete a highly asymmetric division to produce two daughter cells, which immediately enter different developmental pathways and have distinct transcriptional profiles (for review, see Rutley and Twell, 2015). The larger vegetative cell has decondensed chromatin, exits the cell cycle and has a supportive role, whereas the smaller generative (or germ) cell is engulfed within the vegetative cell and divides to produce tricellular pollen containing two sperm cells. In some plants, such as Arabidopsis (Arabidopsis thaliana) germ cell division occurs before anthesis, while in others it occurs during pollen-tube growth. The asymmetric division of the microspore is crucial for the specification of distinct daughter cells as nearly symmetrical divisions result in two vegetative-like cells (Eady et al., 1995; Twell et al., 1998). This suggests that vegetative cell fate is the default developmental pathway, and that germ cell fate must be actively specified. For further reviews of pollen development refer to Borg et al. (2009), Borg and Twell (2011), Twell (2011), Berger and Twell (2011), and Brownfield and Twell (2016).

Following asymmetric microspore division, the R2R3 MYB transcription factor DUO POLLEN1 (DUO1) plays a key role in the development of the male germline in Arabidopsis (Durbarry et al., 2005; Rotman et al., 2005). DUO1 is responsible for the transcriptional control of a gene regulatory network, which includes the mitotic cyclins required for entry into pollen mitosis II and proteins required...
for sperm cell adhesion and fertilization, such as GEX2 and GCS1 (HAP2; Brownfield et al., 2009; Borg et al., 2011; Mori et al., 2006, 2014). DUO1 therefore coordinates an essential gene regulatory network required for male germline differentiation and plant fertility, and the control of DUO1 expression is key in understanding the genetic control of male gametophyte and germline development.

The expression of DUO1 is restricted to the male germline in Arabidopsis (Rotman et al., 2005; Brownfield et al., 2009). For example, a translational reporter fusion of DUO1 is first detectable in the newly formed germ cell soon after asymmetric microspore division with expression increasing during germline development, before declining in the sperm cells (Brownfield et al., 2009; Borg et al., 2014). The early activation of DUO1 is not only required to establish the germline soon after asymmetric division, but DUO1 expression must also be limited to male germ cells, as aberrant growth or cell death results from the ectopic expression of DUO1 in vegetative tissues or in pollen vegetative cells (Palatnik et al., 2007; Brownfield et al., 2009). Thus, tight spatial and temporal regulation is necessary to ensure that DUO1 is activated soon after asymmetric microspore division and is restricted to the germline. This strict control operates primarily at the level of transcription, since DUO1 promoter-reporter constructs and DUO1 protein fusion constructs both show male germline-specific expression (Rotman et al., 2005; Brownfield et al., 2009).

Transcriptional regulation is commonly achieved through the recognition and binding of transcription factors to short cis-regulatory elements (CREs) within promoter regions. Several CREs arranged into cis-regulatory modules important for the expression of pollen-specific genes in the vegetative cell have been defined (Twell et al., 1991; Eyal et al., 1995; Weterings et al., 1995; Bate and Twell, 1998; Rogers et al., 2001). These include MEF2-type CARG box motifs bound by pollen-specific MIKC* MADS box proteins (Verelst et al., 2007a, 2007b). Although germ cells possess unique transcriptional profiles (Borges et al., 2008; Russell et al., 2012), only a few CREs that are known to act in the germline have been characterized.

Haerizadeh et al. (2006) proposed that a CRE in the promoter of the Lily Generative Cell1 (LGC1) gene acts as a silencer to control germline specificity through binding of the repressor, GERMLINE RESTRICTIVE SILENCING FACTOR (GRSF), in nongermline cells. Another CRE important for the expression of genes in the male germline is the AACC motif bound by DUO1 in Arabidopsis (Borg et al., 2011). Among the targets of DUO1, the activity of the Arabidopsis male germline-specific HTR10 (MGH3) promoter (Okada et al., 2005; Brownfield et al., 2009) and those of the DUO1-ACTIVATED ZINC FINGER PROTEIN genes, DAZ1 and DAZ2, depend upon the presence of several DUO1 binding sites (Borg et al., 2011, 2014). Moreover, the OsGEX2 promoter, which also contains a conserved DUO1 binding site, confers sperm-cell-specific expression in rice pollen (Cook and Thilmony, 2012), similar to the activity of the AtGEX2 promoter in Arabidopsis (Engel et al., 2005).

More recently, a CRE that directs sperm-cell-specific expression in Arabidopsis was identified in the promoter of the Plumbago zeylanica isopentenyltransferase (PzIPT1) gene (Zhang et al., 2016). This study identified a regulatory region consisting of duplicated 6-bp Male Gamete Selective Activation (MGSA) motifs, located near the transcription start site. These MGSA motifs are sufficient to direct male germline-specific expression of the PzIPT1 promoter and were shown to be important for the activity of six other sperm-cell-expressed genes in Arabidopsis. In contrast to proposed repressive functions of GRSF in nongermline cells, both the MGSA and DUO1 binding sites led to transcriptional activation specifically in the male germline.

The early expression of DUO1 in the generative cell suggests that transcription factors bind at this stage to CREs in the DUO1 promoter to provide germline specificity. The DUO1 promoter contains a sequence motif similar to that in the silencer CRE of the lily LGCl promoter, to which the GRSF repressor can bind (Haerizadeh et al., 2006). However, mutation or removal of this site does not impact upon the specificity of DUO1 transcription, suggesting that release from GRSF-mediated repression does not determine the germline specificity of DUO1 (Brownfield et al., 2009). On the contrary, there is evidence that the ARID1 protein has a positive influence on DUO1 promoter activity in Arabidopsis. ARID1 is reported to bind the DUO1 promoter in a region approximately − 600 to − 300 bp upstream of the ATG, and in the absence of ARID1, the expression of DUO1 is reduced (Zheng et al., 2014). Since ARID1 is expressed in microspores and vegetative cells (Zheng et al., 2014) it is unlikely to be required for the germline-specific activation of DUO1. Moreover, reduced H3K9ac levels at the DUO1 promoter in the arid1-1 mutant suggest that ARID1 might act through derepression of DUO1 transcription, rather than through direct activation (Zheng et al., 2014). Given that sequences in the − 150 bp proximal DUO1 promoter are sufficient to direct male germline-specific transcription (Brownfield et al., 2009), positive transcriptional control mechanisms are likely to control the specificity of DUO1 expression.

Here, we combine quantitative analysis of a DUO1 promoter deletion series with phylogenetic footprinting to identify a conserved region that is sufficient for germline-specific expression throughout pollen development. This cis-regulatory module, named Regulatory region of DUO1 (ROD1), shows sequence conservation throughout the eudicots and is functionally conserved in the legume Medicago (Medicago truncatula). We show that ROD1 contains conserved short CREs that are essential for male germline expression as well as CREs involved in positive feedback regulation by DUO1. The conservation of ROD1 throughout the eudicots highlights the importance of the correct spatial and temporal expression of DUO1 for the generation of functional sperm cells in flowering plants.
RESULTS
Quantification of a Promoter Deletion Series Identifies Important Regions of the DUO1 Promoter

We previously defined the Arabidopsis DUO1 promoter to include the intergenic region between DUO1 (At3g60460) and the coding sequence of the adjacent gene (At3g60470; Brownfield et al., 2009). This promoter sequence begins -1253 bp upstream of the annotated ATG in TAIR10. The RACE on pollen RNA was used to delimit the DUO1 transcript and transcription start site to 52 bp upstream of this ATG (Supplemental Fig. S1). This revealed an in-frame ATG 9 bp upstream of the annotated ATG. We herein specify numbering of the DUO1 promoter relative to this first ATG (for a comparison of numbering see Supplemental Table S1).

To identify regions of the promoter regulating the germline expression of DUO1, we previously analyzed a 5' deletion series ranging from -1244 to -61 bp relative to the ATG and concluded that a -198-bp fragment was required for male germline-specific expression (Brownfield et al., 2009). Here, we extend this analysis by quantifying the fluorescence signal of the H2B-GFP reporter (H2B provides nuclear localization) for selected promoter deletion fragments with a further deletion (-143 bp) to improve resolution.

Analysis of the deletion series revealed a progressive decline in H2B-GFP fluorescence signal in sperm cells of mature pollen (Fig. 1). There was a significant reduction in signal from the full-length promoter to a -775 bp fragment, followed by a plateau between -775 bp and -455 bp (Fig. 1A: the slight drop between -570 bp and -455 bp is not statistically significant). Expression was significantly lower between -455 bp and -304 bp, and between -304 bp and -143 bp (Fig. 1A), reaching the lowest signal with the -143 bp promoter (Fig. 1B). No expression was detected for the shortest promoter fragment, indicating that sequences essential for DUO1 activation are found in the region between -143 and -61 bp upstream of the ATG.

Phylogenetic Footprinting Identifies a Conserved Region of the DUO1 Promoter

The strict spatial and temporal transcriptional activation of DUO1 in the male germline of flowering plants suggests that key regulatory sequences could be conserved. We therefore analyzed promoter regions in the 1000 bp upstream of the ATG for 30 DUO1 orthologs from the eudicots for overrepresented sequences using MEME (Bailey et al., 2009). The most highly significant sequence identified (E value 7.7e -146) is located within 300 bp upstream of the ATG in all except three cases.

Figure 1. Quantification of expression from a 5' deletion series of the Arabidopsis DUO1 promoter. A, The sperm cell fluorescence signal of plants harboring DUO1 promoter deletion fragments fused to H2B:GFP relative to the longest fragment. The data are represented as the median with error bars indicating the 95% confidence interval for each construct. Deletions with significant differences to the previous are indicated by bars above (***P < 0.001; Mann-Whitney U test). No statistical comparison could be made between the -143-bp and -61-bp constructs as no signal was detected for the -61 bp fragment. For each construct, fluorescence was measured for 189 to 214 pollen grains from 26 to 40 independent lines. B, Representative images of mature pollen with sperm-cell-specific expression viewed by epifluorescence microscopy. The length of the deletion fragment is indicated on each image. Scale bar = 10 μm.
Figure 2. The DUO1 promoter has regions with sequence conservation in eudicots and Arabidopsis accessions. A, Position-specific scoring matrix and E value of the overrepresented region in the DUO1 promoter from 30 diverse eudicot species identified through MEME. B, The position of the conserved region in the 30 eudicot DUO1 promoters. The relationship between the species is indicated on the left (adapted from Phytozome; Goodstein et al., 2012). Black lines represent DUO1 promoter from 1000 bp upstream of the ATG, and green boxes indicate the position of the conserved region, with the precise position for the Arabidopsis region indicated above. Gene identifiers are provided in Supplemental Table S2. C, The frequency of SNPs in the DUO1 promoter 1000 bp upstream of the ATG for 1135 Arabidopsis accessions. The conserved region in eudicots shown in A is indicated by a green box.
(Fig. 2, A and B; Supplemental Table S2). In Arabidopsis, this region is located between −153 and −83 bp upstream of the ATG. Moreover, the functional −143 bp deletion construct included the majority of this conserved region.

To further explore the conservation of DUO1 promoter sequences we utilized data from the Arabidopsis 1001 Genomes Project to examine conservation between Arabidopsis accessions (The 1001 Genomes Consortium, 2016). The regions 1000 bp upstream of DUO1 were analyzed for the frequency of single-nucleotide polymorphisms (SNPs) in 1135 accessions (Fig. 2C). The SNP frequency showed regional variation, with up to 60% of accessions possessing an alternative base in some regions while other regions have very low numbers of SNPs. One notable region low in SNPs is located between −167 to −69 bp, which overlaps the region conserved within the eudicots, and with the −143 bp deletion fragment. This conservation in eudicots and among Arabidopsis accessions is indicative of positive selective pressure and thus suggests it has functional importance.

Another region between −426 and −273 bp also showed a low SNP frequency, with only a few polymorphic sites in one or two accessions (Fig. 2C). While this region was not identified as being conserved in eudicots, the conservation within Arabidopsis accessions and the reduction in reporter fluorescence when this region was deleted (the −455 bp and −304 bp fragments in Fig. 1A), indicates that this region might enhance DUO1 expression. The region from −775 to −642 bp was also low in SNPs, but its removal did not significantly reduce expression (compare −755- to −570-bp fragments in Fig. 1A). Thus, this region does not appear to have an important role in regulating DUO1 expression and the conservation in Arabidopsis accessions might instead relate to the regulation of the adjacent At3g60470 gene that is in the opposite orientation.

A Small Region of the DUO1 Promoter Is Sufficient for Germline-Specific Expression within the Male Gametophyte

Taken together, the results of the DUO1 promoter deletion analysis and the identification of a highly conserved region in eudicot promoters suggest that the DUO1 promoter region from −153 to −61 bp is important for germline expression. We therefore tested if this region was sufficient for expression by fusing four copies of this region upstream of the minimal Cauliflower mosaic virus (CaMV) 35S promoter and an H2B-TdTomato reporter (Fig. 3A). The full-length CaMV 35S promoter and the minimal CaMV 35S promoter are not active in Arabidopsis pollen (Wilkinson et al., 1997; Zhang et al., 2016). Sperm-cell-specific expression was observed in mature pollen in 10 out of 10 transgenic lines, indicating that this region of DNA was sufficient to activate germline transcription (Fig. 3B). We also tested a shorter region containing just the region conserved in eudicots (−153 to −83 bp). Plants harboring this construct showed germline-specific expression in mature pollen in 10 transgenic lines examined (Fig. 3C), demonstrating that this conserved region alone is sufficient to direct germline-specific transcription.

We next asked if this region could replicate the expression pattern of the full-length DUO1 promoter during pollen development. We first monitored expression of the full-length (−1244 bp) DUO1 promoter driving expression of H2B-GFP (equivalent to the longest fragment in the deletion series) using confocal laser scanning microscopy (CLSM; Fig. 3D). Similar to our previous analysis of a translational fusion (Brownfield et al., 2009), GFP was not observed in polarized microspores, but a weak signal was detected, only in the germ cell nucleus, shortly after the asymmetric division. This signal increased during the bicellular stage and persisted in sperm cells present in tricellular mature pollen. We then analyzed the expression of the −153 to −61 bp fragment in developing pollen (Fig. 3E). This region duplicated the expression pattern of the full-length DUO1 promoter, with GFP signal being absent from polarized microspores, weak in the germ cell soon after asymmetric division, subsequently increasing in germ cells and persisting in sperm.

While the −143 bp promoter fragment was sufficient for germline expression, it showed reduced GFP signal compared with longer fragments in the deletion series (Fig. 1), suggesting the presence of other elements influencing DUO1 transcription. The two promoter regions that provided increases in expression in the deletion series (−455 to −304 bp and −304 to −143 bp) overlap with an area low in SNPs (−429 to −273 bp). We therefore tested the ability of these two regions to independently activate expression when placed upstream of min35S:H2B-TdTomato. No expression was detected in 10 lines for the region −304 to −144 bp (Fig. 3F). A very low level of GFP fluorescence was detectable in a few pollen grains from two lines for the −455 to −305 bp fragment (Fig. 3G) while no expression was detected in the other ten lines analyzed. To determine if extra copies would support an increase in expression or if potential CREs were disturbed at the margins of the fragment, we also made a tetramer of an extended −475 to −285 bp region fused to min35S:H2B-TdTomato. However, no expression was detected in 10 transgenic lines examined (Fig. 3H). While the deletion series and conservation within Arabidopsis accessions indicate that these regions have a role in DUO1 regulation, the lack of, or minimal activity of, these regions alone suggests that they may act primarily as enhancers.

Collectively, our analysis shows that the region −153 bp to −83 bp is the major region controlling the germline-specific expression of DUO1 in the Arabidopsis male gametophyte, and we have thus named it Regulatory region of DUO1 (ROD1).

ROD1 Is Functionally Conserved

ROD1 was partially identified based on its conservation among eudicots. To determine whether the
Figure 3. ROD1 is sufficient for germline-specific expression in the male gametophyte. A, Diagram showing the constructs used to test the activity of regions of the DUO1 promoter as four repeats (1, 2, 3, and 4) placed upstream of min35S:H2B-fluorescent reporter. Restriction enzyme and Gateway att sites used for cloning are indicated. B and C, Epifluorescence images showing sperm cell-specific expression in mature pollen containing pDUO1(-153 to -61)X4-min35S:H2B-TdTomato (B) and pDUO1(-153 to -83)X4-min35S:H2B-TdTomato (subsequently named ROD1) (C). D and E, Expression of the full-length DUO1 promoter (pDUO1(1244):H2B-GFP) (D) and pDUO1(-153 to -61)X4-min35S:H2B-GFP (E) in developing wild-type pollen, observed with CLSM. Left to right, polarized microspores (PM), early bicellular pollen (EBC), mid- to late-bicellular pollen (M/LBC), tricellular pollen (TC), and mature pollen (MP). Top, GFP; bottom, transmitted light. Detector settings were not consistent for all images to enable a range of signals to be shown, meaning a direct comparison of intensity cannot be made between images. White arrowheads, germ cell expression in early bicellular pollen; black arrows, the position of the microspore or the...
function of this region is conserved, we monitored the expression of the Medicago DUO1 (MtDUO1) promoter. Medicago is a legume and distantly related to Arabidopsis within in eudicots (Fig. 2B). We cloned an MtDUO1 fragment of −726 bp upstream of the ATG and fused this to an H2B-GFP reporter. Ten out of the 10 transgenic Arabidopsis lines we examined had sperm-cell-specific expression in mature pollen (Fig. 4A), demonstrating that the Medicago promoter is active in Arabidopsis. Furthermore, the MtDUO1 promoter showed a similar expression profile to that of the AtDUO1 promoter, with expression first detected specifically in the germ cell soon after asymmetric division before increasing in the germ-line (compare Figs. 3D and 4A).

We created a 5′ deletion series of the MtDUO1 promoter and quantified H2B-GFP expression in seven to 10 independent T1 transgenic lines for each construct (Fig. 4, B and C). Similar to the Arabidopsis deletion series, there was a progressive decline in reporter expression in sperm cells. Removal of a region from −726 to −517 bp led to a significant decrease in expression, suggesting this region contributes to transcriptional activation. The MtDUO1 promoter, truncated to either −516 bp or −247 bp, showed similar levels of expression, suggesting this region has little impact upon transcription. None of the 10 transgenic lines examined containing the −171-bp deletion showed GFP signal, which is consistent with the absence of MtROD1 from this construct and the requirement of this conserved region for expression in Arabidopsis.

We then asked if the MtROD1 was also sufficient for expression in Arabidopsis pollen by placing four copies of the MtROD1 region (−247 to −150 bp) upstream of min35S::H2B-GFP. Ten independent lines displayed sperm-cell-specific expression in mature pollen (Fig. 4D). Analysis of the activity of MtROD1 during pollen development revealed a similar pattern of expression to those of the Arabidopsis full-length promoter and to four copies of the AtROD1 fused upstream of min35S::H2B-GFP (Fig. 4D).

Together, these results suggest that the AtDUO1 and MtDUO1 promoters share a similar architecture with enhancer regions located upstream of ROD1, the latter being sufficient for germline-specific expression in pollen.

**ROD1 Contains Several CREs**

Since ROD1 is sufficient to direct germline expression, we sought to identify the functional CREs present within this key region. Visual inspection of ROD1 revealed several short-sequence motifs that display higher conservation than other regions (Fig. 5A). These motifs include an AG-rich region at the distal end, with the consensus GAGARAAA, although with some variation at 5′ nucleotide positions. This is followed by three adjacent copies of a GTGG core motif (only two copies in Eucalyptus grandis), which show some similarity in flanking nucleotides. While these motifs differ slightly at each location, the overall consensus is DNGTGGV. At the proximal end of ROD1, there are tandem copies of the consensus motif YAACGYG, with at least one repeat matching the consensus YAACCGY and the core DUO1 binding site AACCG (Borg et al., 2011). Although the spacing between the individual motifs is variable, they usually occur within 20 bp of the adjacent motif (Fig. 5A).

To test the functional significance of these common motifs within ROD1, we disrupted these using site-directed mutagenesis within an active −198 bp AtDUO1 promoter fragment (Fig. 5B). We first confirmed sperm-cell-specific expression of the unmodified pAtDUO1 (−198):H2B-TdTomato construct in 10 transgenic lines (Fig. 5C). Mutation of all sites completely abolished expression in 10 independent lines examined, indicating that at least one of these putative CREs is required for germline expression by ROD1 (Fig. 5C).

Mutation of the AG-rich motif had no impact on sperm-cell-specific expression in mature pollen of four transgenic lines (Fig. 5C), indicating this region is not essential for the activity of ROD1. However, mutation of the three DNGTGGV motifs abolished sperm cell expression in all ten transgenic lines examined (Fig. 5C). This loss of expression indicates that the DNGTGGV sites are essential for the activation of transcription via ROD1 and are likely to represent transcription factor binding sites. When the paired YAACGYG sites were mutated in the −198 bp AtDUO1 promoter fragment, sperm cell-specific expression was detected in mature pollen from all eight transgenic lines examined.

**The YAACGYG Motifs Are Involved in Autoregulation by DUO1**

The consensus sequence AACCG has been characterized as a DUO1 binding site (Borg et al., 2011). Since the YAACGYG motif is similar to the DUO1 binding site and is found in all eudicot DUO1 promoters and at least one of the two copies contains AACCG, we asked whether this site is involved in DUO1 autoactivation. We quantified the expression of the unmodified −198 bp DUO1 promoter fragment and the version with both YAACGYG sites mutated (Fig. 5D). This revealed a significant drop in expression upon ablation of the AACCG motifs, consistent with a role for this in enhancing expression through autoregulation by DUO1.

---

**Figure 3. (Continued.)**

Vegetative cell nucleus when discernable. F to H, Epifluorescence images showing no pollen expression (F and H) or weak sperm-cell expression (G; white arrows) with pDUO1(−304 to −144)::min35S::H2B-TdTomato (F); pDUO1(−455 to −305)::min35S::H2B-TdTomato (F); pDUO1(−455 to −305)::min35S::H2B-TdTomato (G); and pDUO1(−455 to −305)::min35S::H2B-TdTomato (H). Scale bar = 10 μm.
Figure 4. The Medicago DUO1 promoter and ROD1 direct male germline-specific expression in Arabidopsis. A, Expression of the full-length MtDUO1 promoter (pMtDUO1(-726):H2B-GFP) in developing wild-type pollen, observed with CLSM. Left to right, polarized microspores (PM), early bicellular pollen (EBC), mid- to late-bicellular pollen (M/LBC), tricellular pollen (TC), and mature pollen (MP). Top, GFP; bottom, transmitted light. Detector settings were not consistent for all images to enable a range of signals to be visualized. White arrowheads, germ cell expression in early bicellular pollen; black arrows, position of the microspore or the vegetative cell nucleus when discernable. B, The fluorescence level of MtDUO1 promoter deletion fragments fused to H2B:GFP in the sperm cells of mature pollen relative to the longest fragment. The data represents the median, with error bars indicating the 95% confidence interval for each construct. Deletions with significant differences from the previous are indicated by bars above (**P < 0.001; t test). No statistical comparison could be made between the −247 and −171 bp constructs.
In addition, we made a construct containing four copies of the region with the two YAACYGY motifs placed upstream of the min35S promoter linked to the H2B:TdTomato reporter. This construct was expressed in sperm cells of mature pollen for nine transgenic lines examined in a wild-type (Col-0) background. Expression from the construct was also monitored in pollen from nine duo1-4 heterozygous lines, which produce 50% wild-type (tricellular) pollen and 50% duo1 (bicellular) pollen. Fluorescence was detected in sperm cells from approximately half of the wild-type pollen grains, indicating a high frequency of expression in the 50% of wild-type pollen segregating with the transgene following meiosis. However, fluorescence was very rarely detected in single duo1 germ cells of bicellular pollen (Fig. 5, E and F). Thus, high activity of this construct is dependent on the presence of functional DUO1 protein, consistent with DUO1 binding to the tandem YAACYGY motifs in a positive feedback loop to increase expression.

Overall, this analysis indicates that DUO1 is a cis-regulatory module with distinct CRE architecture in eudicots with three DNGTGGV motifs required to initiate expression and two YAACYGY motifs bound by DUO1 to enhance expression. We also asked if a similar regulatory module is present in the promoters of DUO1 genes outside the eudicots. An altered module with reduced size was identified in DUO1 promoter regions from grasses, with two DNGTGGV motifs in close proximity and a single YAACYGY motif, although this was commonly in the reverse direction (Supplemental Fig. S2). In nongrass monocots and basal angiosperms an altered module was also present, with three DNGTGGV motifs and a single YAACYGY motif, but with a greater distance between motifs. In nonflowering plants, individual motifs are present, but these are not clearly organized into a ROD1-like module, indicating lack of conservation of this regulatory module in preangiosperms (Supplemental Fig. S3).

DISCUSSION

The transcription factor DUO1 is essential for the activation of the developmental program leading to differentiation of the male germline and sperm cells. Germ cells fail to divide in duo1 mutants and do not express a range of direct target genes, including those required for sperm cell function (Dubarry et al., 2005; Rotman et al., 2005; Brownfield et al., 2009; Borg et al., 2011). The expression of DUO1 is restricted to the male germline and is first activated shortly after asymmetric division of the microspore, before the germ cell detaches from the microspore wall (Brownfield et al., 2009). Further, ectopic expression of DUO1, either in sporophytic tissues or in the pollen vegetative cell, leads to growth defects and cell death, respectively (Palatnik et al., 2007; Brownfield et al., 2009). Here, we show that a small 71 bp region of the DUO1 promoter, ROD1, is both necessary (no expression when removed; Fig. 1) and sufficient (activates reporter expression; Fig. 3) for the germline-specific expression of DUO1. Moreover, this region replicates the expression pattern of the full-length DUO1 promoter during pollen development, being first activated shortly after the asymmetric division that is critical for the segregation and specification of the male germline.

**ROD1 Acts as a cis-Regulatory Module**

ROD1 contains a combination of short conserved-sequence motifs. At the distal end, an AG-rich motif is followed by three repeats of the consensus DNGTGGV motif, while two copies of the YAACYGY motif are present at the proximal end. The AG-rich motif is least conserved and this variation is tolerated, as this motif differs in the MtDUO1 promoter (Fig. 4A), but MiROD1 is still functional in Arabidopsis. Site-directed mutagenesis revealed that the DNGTGGV motifs are required for the activation of germline transcription by ROD1 and this does not depend upon the presence of the other motifs. Thus, the repeated DNGTGGV motif is an essential CRE for male germline-specific expression of DUO1 and its key role in germline specification. This motif has some sequence similarities to defined CREs, such as GTGGNG bound by the Medicago HD-PHD family member ALFIN1 (Bastola et al., 1998), and the ABA-response element YACGTGKC bound by bZIP transcription factors (Nakashima et al., 2009). Both of these elements appear to require repetition of the same or a similar motif. As these sequences display differences from the DNGTGGV motif in the DUO1 promoter and are bound by transcription factors from different families, it is difficult to predict which transcription factors may bind the DNGTGGV repeats in the DUO1 promoter. The requirement for multiple DNGTGGV motifs for DUO1 promoter activity in the germline immediately after asymmetric division allows us to speculate that DNA-binding regulatory factors may be segregated during asymmetric division, or their activity may be regulated in association with microspore polarity and asymmetric division. Once these are identified it will also be of interest to examine whether such factors are disturbed in mutants such as gem1 and gem2, which disrupt polarity and fail to establish germ cell fate (Park et al., 1998, 2004; Twell et al., 2002).
Figure 5. ROD1 contains multiple CREs. A, Manual alignment of ROD1 from DUO1 orthologs from various eudicots with the PSSM for short highly conserved sequences above and the consensus below. Nucleotides in three or more DUO1 promoters are represented in the consensus. Each motif type is shown in a different color. Species are the same as in Figure 1; however, for some polyploid species, the second DUO1 ortholog is shown. See Supplemental Table S2 for gene identifiers. B and C, Site-directed mutagenesis was used to disrupt the putative CREs in ROD1 in the 2198-bp AtDUO1 promoter fused to H2B-TdTomato. B, Schematic of constructs. Colors represent the motifs shown in A with the Arabidopsis core sequence indicated and the altered sequence shown as noncolored regions. C, Epifluorescence images of representative pollen grains for the mutated constructs. D, Fluorescence level of the construct with mutated YAACYGY motifs (2198 AACCC) relative to the unmodified 2198-bp DUO1 promoter fragment (2198). Fluorescence was measured from at least 200 pollen grains from 10 (2198) or five (2198 AACCC) transgenic lines, and data are presented as the mean ± se. ***indicates a significant difference (P < 0.01, Student’s t test). E, Epifluorescence images of representative pollen grains showing reporter fluorescence (left) and 4′,6-diamino-phenylindole stain (right).
**ROD1** is sufficient for the activation of germline transcription in the absence of either the AG-rich region or the YAACYGY motifs, showing that neither of these motifs are essential for **DUO1** promoter activity. However, mutation of the YAACYGY motif decreased expression, suggesting this region acts to enhance expression. Two copies of YAACYGY are present in **ROD1** in all eudicots, and at least one copy contains the **DUO1** binding site consensus AACCAG (Borg et al., 2011). Interestingly, the YAACYGY motif alone can only activate transcription in wild-type pollen that expresses **DUO1** protein. Taken together, these data indicate that the expression of **DUO1** is likely to be enhanced by a positive feedback loop in which **DUO1** binds to the tandem motifs in its own promoter to increase expression. Positive feedback has been observed for other R2R3-type MYB transcription factors, such as MYB10 involved in anthocyanin production in apple, where multiplication of the corresponding CRE leads to red-fleshed apples (Espley et al., 2007). Similarly, MYB23 provides a positive feedback loop in Arabidopsis root cell fate specification via MYB23 binding to its own promoter, together with its activation by the R2R3 MYB WEREWOLF (Kang et al., 2009).

Overall, our data lead to a model for the germline-specific activation of **DUO1** controlled by **ROD1**. Binding of a transcription factor to the DNGTGGV motifs activates expression at, or very soon after, asymmetric division of the microspore. This leads to the production of **DUO1** in the germ cell in early bicellular pollen, which in turn switches on transcription of its target genes (Brownfield et al., 2009; Borg et al., 2011). As the DNGTGGV motif is not overrepresented in the promoters of **DUO1** target genes (MEME analysis; data not shown), the transcription factor(s) that activate **DUO1** are not involved in enhancing expression of **DUO1** target genes. The binding of **DUO1** to its target promoters includes binding to its own promoter in a positive feedback loop. This increases expression of **DUO1**, ensuring sufficient protein is made to activate its multiple targets and commit the segregated generative cell to male germline fate. In addition, upstream enhancer elements in the **DUO1** promoter (−455 to −143) further increase **DUO1** expression (see Fig. 1A). These enhancer regions could function by promoting the formation of permissive chromatin, since ARID1, which is involved in chromatin remodeling, has been shown to bind chromatin within these regions (Zheng et al., 2014).

**Relationship with Other Known Male Germline-CREs**

Despite plant sperm cells having a distinct transcriptional profile (Borges et al., 2008; Russell et al., 2012), relatively few CREs important for male germline expression in Arabidopsis pollen have been reported. One is the **DUO1** binding site found in the promoters of genes directly regulated by **DUO1** (Borg et al., 2011) and shown here to likely be involved in **DUO1** autoactivation. Many **DUO1** target genes, such as **HTR10** (MGH3) and **GCS1** (HAP2), display germline-specific expression, and mutation of **DUO1** binding sites within the **HTR10** promoter reduces expression to low levels (Borg et al., 2011). Thus, the binding of **DUO1** to its corresponding CREs contributes to the germline specificity of a number of genes, as well as enhancing its own expression.

**Other germline CREs,** which direct sperm-cell-specific expression, were identified in the promoter of the **PzIPT1** gene (Zhang et al., 2016). Disruption of two 6-bp **Male Gamete Selective Activation** (MGSA) motifs located near the transcription start site inactivated **PzIPT1** promoter activity in sperm. Similar to **ROD1**, multiple copies of the **MGSA** motif placed upstream of the minimal CaMV 35S promoter were sufficient to direct sperm-cell-specific expression. **MGSA** motifs were found in other sperm-cell-expressed genes and shown to be required for the activity of six sperm-cell-expressed genes in Arabidopsis. However, unlike **ROD1** these **MGSA** motifs activate transcription only at late stages of pollen development (Zhang et al., 2016). Therefore, they are unlikely to play critical roles in establishing male germline fate but could promote the expression of genes required for sperm cell maturation through responses to the plant hormone cytokinin (Zhang et al., 2016).

A repressive model for controlling male germline specificity has also been proposed (Haerizadeh et al., 2006). In this model, the binding of the regulatory factor GRSF to a silencer element represses expression in non-male germline cells, such that transcriptional activation in germ cells results from derepression. However, mutation of the putative GRSF binding site in the **DUO1** promoter did not alter germ cell specificity (Brownfield et al., 2009), and the **PzIPT1** promoter does not contain the GRSF binding site (Zhang et al., 2016). Further, the activity profiles of the **DUO1** (Fig. 1) and **PzIPT1** (Zhang et al., 2016) promoter deletion series and the identification of regions that specially activate transcription in the germline (**ROD1** and **MGSA**) both indicate that transcriptional activation, rather than repression, controls male germline specificity in Arabidopsis.

**Conservation of ROD1 within Eudicots Reflects Its Functional Importance**

The architecture of **ROD1** suggests an ordered cis-regulatory module that is highly conserved in the **DUO1** promoter in Arabidopsis accessions and...
throughout eudicots, spanning 140 to 150 million years (Chaw et al., 2004). Moreover, the organization of conserved CREs in similar cis-regulatory modules has been reported in animal and plant promoters, respectively (Gupta and Liu, 2005; Picot et al., 2010). The pattern of an AG-rich region followed by three repeats of DNHTGVG and then two repeats of YAACYGG in a region spanning less than 80 bp is found in almost all eudicot DUO1 orthologs examined. A high level of sequence similarity in both coding and regulatory regions indicates functional conservation (Picot et al., 2010; Korkuc et al., 2014), which appears to be the case for ROD1, as MiROD1 functions in Arabidopsis. A reduced, altered region is also present in DUO1 promoters from grasses, which could potentially function as a cis-regulatory module for male germline expression of DUO1 in these species. In nonflowering plants, some of the individual motifs are present; however, they are not organized into a similar regulatory module, and it remains to be determined whether these short motifs are involved in regulating DUO1 expression.

The high level of conservation suggests that ROD1 in eudicots is under evolutionary constraint. This likely relates to the requirement for tight spatial and temporal control of DUO1 protein expression. As DUO1 is essential for the activation of a male germline developmental pathway, expression is required at a sufficient level in the germline to activate all target genes, and reduced expression may be selected against through negative effects on sperm cell fitness. Further, misexpression of DUO1 and the resulting ectopic activation of germline genes is known to be detrimental to other cell types, such as the pollen vegetative cell (Brownfield et al., 2009) and to sporophytic tissues (Palatnik et al., 2007). Thus, mutations resulting in “leaky” expression of DUO1, outside of the male germline are expected to reduce fitness. Interestingly, this high level of evolutionary conservation contrasts with the evolutionarily young transcriptome of angiosperm male gametophytes (including the germ-line), which represent an innovation incubator for the birth of new genes (Wu et al., 2014; Cui et al., 2015).

CONCLUSION

We have shown that expression of the male germline-specific transcription factor DUO1 is controlled by the conserved cis-regulatory module ROD1. We propose a model wherein DNHTGVG motifs bind an activator to stimulate germ cell expression of DUO1 soon after microspore division, and binding of DUO1 to tandem YAACYGG motifs maintains and enhances DUO1 transcription. The conservation of ROD1 among eudicots also highlights evolutionary constraints on the mechanisms controlling the level and specificity of DUO1 expression required to ensure gamete differentiation. Equally important are the mechanisms restricting DUO1 transcription to prevent inappropriate expression of male germline genes and its negative consequences. A key challenge that remains is to identify the regulatory proteins that bind ROD1 to activate DUO1 transcription.

MATERIALS AND METHODS

Plant Material and Transformation

Arabidopsis (Arabidopsis thaliana) plants were grown at 20°C to 21°C under a 16-h-light and 8-h-dark cycle, with approximately 70% relative humidity. Plants were transformed with Agrobacterium tumefaciens strain GV3101 using a standard (Clough and Bent, 1998) or modified (Martinez-Trujillo et al., 2004) floral dip method. Transformants were selected on soil with BASTA (200 mg/L glufosinate ammonium, DHAI PROCA) by spraying or subirrigation or on 0.5% Murashige and Skoog plates with kanamycin (50 mg/L). Vectors were transformed into either heterozygous duo1-4 plants (Borg et al., 2014) to generate wild-type and heterozygous duo1-4 T1 plants or into Col-0. Plants used for the Arabidopsis deletion series were as described in Brownfield et al. (2009).

Vector Construction

Gateway multisite cloning (Life Technologies) was used to generate vectors. Promoter regions were amplified from genomic DNA from Arabidopsis Col-0 or Medicago truncatula R108 plants by PCR with high-fidelity Taq polymerase (Phusion) with primers containing suitable attachment sites. For creation of tetramer (4x) promoter fragments, the four copies of DNA were amplified individually and flanked with different combinations of attachment sites and restriction sites (Fig. 3A). PCR fragments were purified (PCR kit), and the four pieces were digested in a single reaction then ligated using T4 ligase, and the ligation product was added to a PCR with full-length n/β-site primers. For site-directed mutagenesis, DNA fragments were synthesized by Integrated DNA Technologies with suitable attachment sites. Purified PCR products or synthesized gene fragments were cloned into pDONR2P4-I6Aatted with BP Clonase II (Life Technologies) and verified by sequencing, pENTR221: min35S-H2B, pENTR221:H2B, and pENTR2P3-TdTomato were as described in Borg et al. (2014), and pENTR2P3:TdGF was as described in Brownfield et al. (2015). pDONR221: min35S-H2B was made by fusion PCR combining the −4 to +18 region of the CaMV 35S promoter with H2B from pDONR221:H2B, with attachment sites added to the primers flanking the fusion product, cloned into pDONR221 with BP Clonase II and verified by sequencing. To create expression constructs, multipart LR reactions were performed using LR Clonase II Plus (Life Technologies) and the destination vector pB7m34GW (Karimi et al., 2005).

RACE

Mature pollen RNA was used to amplify the 5' and 3' ends of the DUO1 mRNA with a GeneRacer kit (Invitrogen) according to manufacturer’s instructions. The 5' cDNA end was amplified with DUO1 C3-R reverse primer (CTTGGAACATGCTCGATC) and nested with DUO1-RF (CGAA-CAATGCTCAAGAATCAGC). The amplification of 3' ends of the cDNA was performed with a DUO1 N1-F forward primer (ATGGCTAGGATTCTTTAGA) and the software used to measure the mean intensity. Background fluorescence was measured using NIS Elements BR version 4.00.03 or ImageJ with background subtracted by using the rolling ball algorithm in ImageJ (ball size 500.0 pixels). In-focus sperm cells were selected, commonly only one per pollen grain, and the software used to measure the mean intensity. Background fluorescence from a cytoplasmic region of each pollen grain was subtracted from each value. Normally distributed data were tested for significance with a Student’s t test on the mean of each independent line. Nonnormally distributed data were tested with a Mann-Whitney U test on the medians.

For analysis of developing pollen, spores were released from flower buds from different stages of development as in Brownfield et al. (2009). Confocal

Microscopy

For epifluorescence microscopy mature pollen was isolated in DAPI buffer as previously described (Park et al., 1998). Epifluorescence microscopy was performed on either a Nikon Eclipse 80i microscope fitted a DS-QiXM camera with illumination provided by a preciExcite LED light source or a Olympus IX71/IX51 inverted microscope using the Olympus DP controller with a 40X objective. The excitation filter range was 450 to 490 nm for GFP with an emission filter of 500 to 545 nm and for TdTomato 550 to 600 nm excitation and 615 to 665 nm emission. For analysis of fluorescence all images in an experiment were captured under standard exposure, ISO sensitivity, and pixel size. Fluorescence was measured using NIS Elements BR version 4.00.03 or ImageJ with background subtracted by using the rolling ball algorithm in ImageJ (ball size 500.0 pixels). In-focus sperm cells were selected, commonly only one per pollen grain, and the software used to measure the mean intensity. Background fluorescence from a cytoplasmic region of each pollen grain was subtracted from each value. Normally distributed data were tested for significance with a Student’s t test on the mean of each independent line. Nonnormally distributed data were tested with a Mann-Whitney U test on the medians.
laser scanning microscopy was performed on an Olympus Fluoview FV1000 microscope, with Olympus Fluoview software and Olympus BX61 camera with a 60 × oil objective. Dwell time was 20 μs and the pinhole 100 nm. For GFP excitation, a 473 nm laser was used, and detection between 485 and 494 nm with the sensitivity of the photon multiplier detector altered between 500 and 650 mV, depending on signal strength. For transmitted light, the detector was set at 140 mV. All images are an average of four scans (Kalman set to 4).

Bioinformatics

DUO1 orthologs were detected using a BLASTp search of the genomes available in Phytozone using default settings (Goodstein et al., 2012), with two exceptions: Picea abies and Marchantia polymorpha sequences were obtained from BlastP searches against NCBI databases. The hit with the lowest E value for each species was selected and each verified to contain the signature supermutermy Lys residue (K66) in the R2R3 MYB domain that distinguishes DUO1 from other R2R3 MYB family members (Rotman et al., 2005; Supplemental Fig. S4). Appropriate sequences were downloaded from Phytozone with details provided in Supplemental Table S2. To find overrepresented sequences, 1000 bp upstream of the ATG for each DUO1 ortholog was analyzed through MEME (Bailey et al., 2009) with settings adjusted to search for up to three motifs of 50 to 100 bp in length with one match per sequence. Sequence logos were created using Weblogo (Crooks et al., 2004).

For analysis of the frequency of SNPs in Arabidopsis accessions, VCF files for the 1000 bp upstream of the DUO1 ATG from 1135 accessions was downloaded from the 1001 Genomes Consortium (2016). In excel, the number of accessions containing an SNP at each position was summed and divided by the total number of accessions (1135).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers HM776521 (DUO1 transcript) and NP_191605.2 (DUO1 protein).

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Analysis of the DUO1 gene and transcript.

Supplemental Figure S2. ROD1 from the angiosperms.

Supplemental Figure S3. Location of candidate CRE in the DUO1 promoters of nonflowering plants.

Supplemental Figure S4. Alignment of part of the R2R3 domain of DUO1 orthologs.

Supplemental Table S1. Comparison of numbering of promoter fragments used the deletion series in this work and Brownfield et al. (2009).

Supplemental Table S2. Gene identifiers for DUO1 orthologs from the 30 eudicot species shown in Figures 2 and 5.

ACKNOWLEDGMENTS

We thank Murray Cadzow for help with SNP analysis of the DUO1 promoters of the Arabidopsis accessions.

Received July 29, 2016; accepted September 7, 2016; published September 13, 2016.

LITERATURE CITED

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME suite: tools for motif discovery and searching. Nucleic Acids Res 37: W202–W208

Bastola DR, Pethe VV, Winicov I (1998) Alfin1, a novel zinc-finger protein in alfalfa roots that binds to promoter elements in the salt-inducible MsPRP2 gene. Plant Mol Biol 38: 1123–1135

Bate N, Twell D (1998) Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol Biol 37: 859–869

Berger EF, Twell D (2011) Germline specification and function in plants. Annu Rev Plant Biol 62: 461–484

Borg M, Brownfield L, Khatab H, Sidorova A, Lingaya M, Twell D (2011) The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. Plant Cell 23: 534–549

Borg M, Brownfield L, Twell D (2009) Male gametophyte development: a molecular perspective. J Exp Bot 60: 1465–1478

Borg M, Rulley N, Kagale S, Hamamura Y, Gherghinoiu M, Kumar S, Sari U, Esparza-Franc MA, Sakamoto W, Rozwadowski K, et al. (2014) An EAR-dependent regulatory module promotes male germ cell division and sperm fertility in Arabidopsis. Plant Cell 26: 2098–2113

Borg M, Twell D (2011) Pollen: structure and development. In eLS. John Wiley and Sons, Chichester, UK, www.els.net, doi/10.1002/9780470159020.a0002039.pub2

Borges F, Comes G, Gardner R, Moreno N, McCormick S, Feijö JA, Becker JD (2008) Comparative transcriptomics of Arabidopsis sperm cells. Plant Physiol 148: 1168–1181

Brownfield L, Hafidh S, Borg M, Sidorova A, Mori T, Twell D (2009) A plant germline-specific integrator of sperm specification and cell cycle progression. PLoS Genet 5: e1000430

Brownfield L, Yi J, Jiang H, Minina EA, Twell D, Köhler C (2015) OReganles maintain spindle position in plant meiosis. Nat Commun 6: 6492

Brownfield L, Twell D (2016) Plant reproduction. In eLS. John Wiley and Sons, Chichester, UK, http://www.els.net, doi/10.1002/9780470159020.a0002046.pub2

Chaw S-M, Chang CC, Chen HL, Li WH (2004) Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. J Mol Evol 58: 424–441

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16: 735–743

Cook M, Thilmony R (2012) The OsGEX2 gene promoter confers sperm cell expression in transgenic rice. Plant Mol Biol Rep 30: 1138–1148

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14: 1188–1190

Cui X, Ly Y, Chen M, Nikoloski Z, Twell D, Zhang D (2015) Young genes out of the male: an insight from evolutionary age analysis of the pollen transcriptome. Mol Plant 8: 935–949

Durbarry A, Vizir I, Twell D (2005) Male germ line development in Arabidopsis. duo pollen mutants reveal gametophytic regulators of generative cell cycle progression. Plant Physiol 137: 297–307

Eady C, Lindsey K, Twell D (1995) The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. Plant Cell 7: 65–74

Engel ML, Holmes-Davis R, McCormick S (2005) Green sperm. Identification of male gamete promoters in Arabidopsis. Plant Physiol 138: 2124–2133

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kuttya-amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J 49: 414–427

Eyal Y, Curie C, McCormick S (1995) Pollen-specific genes reside in 30 bp of the proximal promoters of two pollen-expressed genes. Plant Cell 7: 373–384

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hcllsten U, Putnam N, et al. (2012) Phytozone: a comparative platform for green plant genomics. Nucleic Acids Res 40: D1178–D1186

Gupta M, Liu JS (2005) De novo cis-regulatory module elicitation for eukaryotic genomes. Proc Natl Acad Sci USA 102: 7079–7084

Haerizadeh F, Singh MB, Bhalla PL (2006) Transcriptional repression distinguishes somatic from germ cell lineages in a plant. Science 313: 496–499

Kang YH, Kirk V, Hulskamp M, Nam KH, Hagely K, Lee MM, Schiefelbein J (2009) The MYB23 gene provides a positive feedback loop for cell fate specification in the Arabidopsis root epidermis. Plant Cell 21: 1080–1094

Kariotis M, De Meyer B, Hillson P (2005) Modular cloning in plant cells. Trends Plant Sci 10: 103–105

Korkut P, Schippers JHM, Walther D (2014) Characterization and identification of cis-regulatory elements in Arabidopsis based on single-nucleotide polymorphism information. Plant Physiol 164: 181–200

292

Plant Physiol. Vol. 173, 2017
Martinez-Trujillo M, Limones-Briones V, Cabrera-Ponce JL, Herrera-Estrella L (2004) Improving transformation efficiency of Arabidopsis thaliana by modifying the floral dip method. Plant Mol Biol Rep 22: 63–70
Mori T, Igawa T, Tamiya G, Miyagishima SY, Berger F (2014) Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. Curr Biol 24: 170–175
Mori T, Kuroiwa H, Higashiyama T, Kuroiwa T (2006) GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. Nat Cell Biol 8: 64–71
Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2004) Improving transformation efficiency by modifying the attachment requires GEX2 for successful fertilization in Arabidopsis. Plant Physiol. Vol. 173, 2017 293
Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez-Russell SD, Gou X, Wong CE, Wang X, Yuan T, Wei X, Bhalla PL, Singh MB (2016) Male gametogenesis and germline specification in flowering plants. Sex Plant Reprod 24: 149–160
Park SK, Howden R, Twell D (1998) The Arabidopsis thaliana gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate. Development 125: 3789–3799
Picot E, Krusche P, Tiskin A, Carré I, Ott S (2010) Evolutionary analysis of regulatory sequences (EARS) in plants. Plant J 64: 165–176
Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D (2001) Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene g10. Plant Mol Biol 45: 577–585
Rotman N, Durbarry A, Wardle A, Yang WC, Chaboud A, Faure JE, Berger F, Twell D (2005) A novel class of MYB factors controls sperm-cell formation in plants. Curr Biol 15: 244–248
Russell SD, Gou X, Wong CE, Wang X, Yuan T, Wei X, Bhalla PL, Singh MB (2012) Genomic profiling of rice sperm cell transcripts reveals conserved and distinct elements in the flowering plant male germ lineage. New Phytol 195: 560–573
Rutley N, Twell D (2015) A decade of pollen transcriptomics. Plant Reprod 28: 73–89
The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. Cell 166: 481–491
Twell D, Park SK, Hawkins TJ, Schubert D, Schmidt R, Smertenko A, Hussey PJ (2002) MOR1/GEM1 has an essential role in the plant-specific cytokinetic phragmoplast. Nat Cell Biol 4: 711–714
Twell D (2011) Male gametogenesis and germline specification in flowering plants. Sex Plant Reprod 24: 149–160
Twell D, Park SK, Lalanne E (1998) Asymmetric division and cell-fate determination in developing pollen. Trends Plant Sci 3: 305–310
Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S (1991) Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Dev 5: 496–507
Verelst W, Saedler H, Münster T (2007a) MIKC* MADS-protein complexes bind motifs enriched in the proximal region of late pollen-specific Arabidopsis promoters. Plant Physiol 143: 447–460
Verelst W, Twell D, de Folter S, Immink R, Saedler H, Münster T (2007b) MADS-complexes regulate transcriptome dynamics during pollen maturation. Genome Biol 8: R249
Weterings K, Schrauwen J, Wullems G, Twell D (1995) Functional dissection of the promoter of the pollen-specific gene NTP303 reveals a novel pollen-specific, and conserved cis-regulatory element. Plant J 8: 55–63
Wilkinson JE, Twell D, Lindsey K (1997) Activities of CaMV 35S and nos promoters in pollen: Implications for field release of transgenic plants. J Exp Bot 48: 265–275
Wu D-D, Wang X, Li Y, Zeng L, Irwin DM, Zhang Y-P, The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. Cell 166: 481–491
Zhang J, Yuan T, Duan X, Wei X, Shi T, Li J, Russell SD, Gou X (2016) Cis-regulatory elements determine germline specificity and expression level of an isopentenyltransferase gene in sperm cells of Arabidopsis. Plant Physiol 170: 1524–1534
Zheng B, He H, Zheng Y, Wu W, McCormick S (2014) An ARID domain-containing protein within nuclear bodies is required for sperm cell formation in Arabidopsis thaliana. PLoS Genet 10: e1004421