SIAMESE, a Plant-Specific Cell Cycle Regulator, Controls Endoreplication Onset in Arabidopsis thaliana


Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana 70803

Department of Köln, Botanical Institute III, 50931 Köln, Germany

Department of Plant Systems Biology, Flanders Interuniversitair Instituut for Biotechnology, Ghent University, B-9000 Ghent, Belgium

Department of Botany, Genetics Institute and Plant Molecular and Cell Biology Program, University of Florida, Gainesville, Florida 32611

Recessive mutations in the SIAMESE (SIM) gene of Arabidopsis thaliana result in multicellular trichomes harboring individual nuclei with a low ploidy level, a phenotype strikingly different from that of wild-type trichomes, which are single cells with a nuclear DNA content of ~16C to 32C. These observations suggested that SIM is required to suppress mitosis as part of the switch to endoreplication in trichomes. Here, we demonstrate that SIM encodes a nuclear-localized 14-kD protein containing a cyclin binding motif and a motif found in ICK/KRP (for Interactors of Cdc2 kinase/Kip-related protein) cell cycle inhibitor proteins. Accordingly, SIM was found to associate with D-type cyclins and CDKA;1. Homologs of SIM were detected in other dicots and in monocots but not in mammals or fungi. SIM proteins are expressed throughout the shoot apical meristem, in leaf primordia, and in the elongation zone of the root and are localized to the nucleus. Plants overexpressing SIM are slow-growing and have narrow leaves and enlarged epidermal cells with an increased DNA content resulting from additional endocycles. We hypothesize that SIM encodes a plant-specific CDK inhibitor with a key function in the mitosis-to-endoreplication transition.

INTRODUCTION

Cell differentiation is closely coordinated with cell cycle progression. In the simplest case, the cell cycle arrests concomitant with the onset of differentiation, but in many cell differentiation pathways, alternative versions of the cell cycle occur along with differentiation. One example is the altered division potential of transient amplifying cells, which are restricted in both their developmental potential and the number of times they can divide, relative to the undifferentiated and essentially immortal stem cells from which they derive (Watt and Hogan, 2000). Another example, common in both plants and animals, is the amplification of nuclear DNA by endocycles that continues during differentiation of many cell types, a process called either endoreplication or endoreduplication (Edgar and Orr-Weaver, 2001). The coordination of these modified cell cycles with differentiation remains poorly understood.

The regulation of cell cycle transitions in plants is similar to that of animals (reviewed in De Veylder et al., 2003; Dewitte and Murray, 2003; Inzé, 2005). Transitions between stages in the cell cycle are controlled by a class of Ser/Thr kinases known as cyclin-dependant kinases (CDKs). As suggested by their name, the kinase activity of CDKs depends on their association with a regulatory cyclin (CYC) protein. Cell cycle progression is regulated by periodic expression of cyclins and their ubiquitin-mediated proteolysis and by the phosphorylation of a variety of targets by CDK/cyclin complexes. The G1/S transition is regulated by phosphorylation of the retinoblastoma-related protein by a CDK/cyclin complex. The G2/M transition most likely requires both A-type and B-type CDKs, as well as CYCA and CYCB proteins, to form mitotic CYC/CDK complexes (De Veylder et al., 2003; Dewitte and Murray, 2003; Inzé, 2005).

Cell cycle progression is also regulated by inhibitors of CYC/CDK complexes. The only plant CDK inhibitors identified to date are a family of proteins distantly related to the Kip family of animal CDK inhibitors; these proteins are known as Kip-related proteins (KRP) (De Veylder et al., 2001) or Interactors of Cdc2 kinases (ICKs) (Wang et al., 1997). ICK/KRP proteins are generally thought to interact with CDKA and CYCDs (Wang et al., 1998; De Veylder et al., 2001), although two recent reports indicate that some family members may interact with CDKB as well (Nakai et al., 2006; Pettko-Szandtner et al., 2006). ICK/KRP proteins can inhibit CDK-associated histone H1 kinase activity in vitro or in vivo (Wang et al., 1997, 1998; De Veylder et al., 2001). Overproduction of these proteins in transgenic plants suppresses cell proliferation while increasing the length of the cell cycle and cell
size, resulting in smaller plants with serrated leaves (Wang et al., 2000; De Veylder et al., 2001). ICK/KRP overexpression differentially affects DNA content depending on the level of overexpression. Weak overexpression increases DNA content, while strong overexpression decreases DNA content (Verkest et al., 2005; Weinl et al., 2005). Together, these results suggest concentration-dependent roles for ICK/KRPs in blocking the G1/S cell cycle and blocking entry into mitosis but allowing S-phase progression.

During endoreplication cycles (endocycles), nuclear DNA is replicated without cytokinesis, resulting in cells with a DNA content greater than 2C. In angiosperms, endoreplication is particularly common and occurs in a wide variety of tissues and cell types, including agriculturally important tissues, such as maize (Zea mays) endosperm and cotton (Gossypium hirsutum) fibers (Kowles and Phillips, 1985; Van’t Hof, 1999). Often, there is a correlation between the final volume of a differentiated cell and its DNA content (Meralarogo et al., 1993; Hülskamp et al., 1994; Vlieghe et al., 2005). It is generally assumed that the function of endoreplication is gene amplification to supply the gene expression needs of large cells, but other explanations have been suggested (Nagl, 1976; Barlow, 1978). The primary functional features of the endocyte appear to be the absence of G2/M phase CDK activity, preventing mitosis, coupled with oscillations of G1/S CDK activity to allow relicensing of replication origins between each round of DNA replication (reviewed in Larkins et al., 2001). In maize endosperm, biochemical evidence has been obtained for two separable endoreplication-promoting factors: an activity inhibiting mitosis and an increase in S-phase-related protein kinase activity (Grafi and Larkins, 1995). Switching to endoreplication appears to involve downregulation of expression of CYCAs, CYCBs, and CDKB and activation of the anaphase promoting complex, which targets mitotic cyclins for degradation (reviewed in Dewitte and Murray, 2003; Inzé, 2005).

The shoot epidermal hairs (trichomes) of Arabidopsis thaliana are now well established as a model for the study of the plant cell cycle and cell differentiation. These trichomes are specialized branched single cells that extend out from the epidermis. During differentiation, trichome nuclei undergo endoreplication, resulting in a nuclear DNA content of 16C to 32C (Meralarogo et al., 1993; Hülskamp et al., 1994). Mutations exist that either increase or decrease the nuclear DNA content (Perazza et al., 1999), including KAKTUS, GLABRA3 (GL3), and TRIPTYCHON, which encode a HECT-class ubiquitin E3 ligase (El Reyf et al., 2003), a basic helix-loop-helix transcription factor (Payne et al., 2000), and an inhibitory Myb transcription factor, respectively (Hülskamp et al., 1994). Recessive mutations in the SIAMESE (SIM) gene of Arabidopsis have a unique cell cycle–related phenotype, the production of multicellular trichomes, the individual nuclei of which have reduced levels of endoreplication (Walker et al., 2000). These observations suggest that SIM is required to suppress mitosis as part of the switch to endoreplication in trichomes.

In this study, we report that SIM encodes a 14-kD protein that is part of a small Arabidopsis gene family comprised of four members. Homologs exist in other dicots and in monocots, though no obvious animal homologs have been identified. These proteins share one motif with the ICK/KRP cell cycle inhibitor proteins and have a potential cyclin binding motif. We identified protein–protein interactions between SIM and D-type cyclins as well as CDKA:1. SIM overexpression in transgenic plants results in small plants with serrated leaves containing enlarged cells with increased levels of nuclear DNA. Taken together, the SIM loss-of-function and gain-of-function phenotypes clearly point to an integral role for SIM in regulation of endoreplication.

RESULTS

Isolation of SIM

Wild-type trichomes are unicellular and occur singly on the leaf (Figure 1A). Three independent recessive sim mutant alleles have an essentially identical mutant phenotype of frequent multicellular trichomes and clusters of adjacent trichomes (Figures 1B and 1C). The sim-1 mutation was mapped to a region of ~407 kb on chromosome five. Two markers located 66.5 kb apart on the overlapping BAC clones T32M21 and T19N18 showed no recombination in 1088 F2 chromosomes. Further attempts to reduce the genetic interval were unsuccessful. The sim-2 allele originated in a T-DNA insertion population (Campisi et al., 1999). One of the two inserts present in the original sim-2 T-DNA line showed linkage between the sim mutant phenotype and to both kanamycin resistance and β-glucuronidase (GUS) expression from an enhancer trap contained in the T-DNA. The right border of this T-DNA was obtained by adaptor PCR and shown by DNA sequencing to be positioned 1189 bp upstream of the coding region of the gene At5g04470, a gene located on the T32M21 BAC clone. The left border was located 1461 bp downstream, within the At5g04470 coding region, resulting in a deletion of these upstream sequences and the first 272 bp of the At5g04470 coding region (Figure 2A). Sequencing of this gene for the original sim-1 allele revealed a point mutation changing the putative ATG start codon for At5g04470 to ATA, and a third allele, sim-3, contained a C → T mutation, resulting in a Pro-to-Ser amino acid change at position 36 (Figure 2A). In RT-PCR experiments, At5g04470 transcripts were detected in wild-type developing leaves but not in RNA isolated from the sim-2 deletion allele (see Supplemental Figure 1 online). Finally, the sim-1 mutant phenotype was rescued by the At5g04470 coding region under control of the trichome-specific GL2 promoter (Figure 1D) and by a genomic DNA fragment including 2870 bp upstream of the At5g04470 start codon, the entire coding region, and 250 bp downstream of the stop codon (Figure 1E). Taken together, these results established that At5g04470 is the SIM gene and that sim-1 and sim-2 are likely to be amorphic alleles.

Conceptual translation of the open reading frame reveals that SIM encodes a 14-kD protein of unknown function. The SIM gene family of Arabidopsis contains at least four members, and homologs exist in other plant species, including both monocots and dicots (Figure 2B). Although SIM has no overall similarity with proteins of known function, several conserved motifs give clues as to its function. Motif 4 of the SIM family is similar to motif 3 of the CDK-inhibitory ICK/KRP proteins (De Veylder et al., 2001; Figures 2B and 2C). Motif 3 of the SIM family (Figure 2B) is a putative cyclin binding motif known as the Cy or zRXL motif, where Z is basic or Cys, and X is usually a basic residue. This motif is implicated in binding of some CDK inhibitors, E2F and RB.
to CYCA, CYCE, and CYCD/CDK complexes (Adams et al., 1996; Wohlschlegel et al., 2001). Motif 1 and motif 2 show no obvious similarity to any domains with known functions, although the motif 1 residue P36, mutated to Ser in sim-3, is conserved in all of the homologs. SIM (amino acids 26 to 40), SIAMESE-RELATED2 (SMR2) (amino acids 32 to 46), SMR3 (amino acids 33 to 44), Zm SMR2 (amino acids 98 to 109), and Glycine max SMR (amino acids 21 to 33) also have PEST domains (PESTfind scores > +6.96) enriched in Pro, Glu, Ser, and Thr residues that often serve as proteolytic signals (Rogers et al., 1986). In addition, SMR1 contains a consensus CDK phosphorylation site (S/TPXK/R) at residues 16-19.

SIM and SMR Yellow Fluorescent Protein Fusions are Nuclear Localized

To determine the subcellular localization of the SIM protein, a 35S:EYFP:SIM (EYFP for enhanced yellow fluorescent protein) gene construct was introduced via biolistic bombardment directly into leaf epidermal cells of Arabidopsis. EYFP alone is cytoplasmically localized (Figure 3A), while an EYFP fusion to the transcription factor TGA5 is nuclear localized (Figure 3B). EYFP:SIM expression was detected in nuclei (Figure 3C). The subcellular localizations of the SMRs from Arabidopsis were also determined using biolistic bombardment of EYFP fusions into epidermal cells. In all instances, the proteins were localized to nuclei (Figures 3D to 3F).

SIM Interacts with D-Type Cyclins and CDKA;1 in Vivo and Regulates CYCB1;1 Expression

The presence of ICK/KRP-like domains and the Cy motif within the SIM protein suggested that it might associate with cyclins. To test this hypothesis, an ECFP:SIM (ECFP for enhanced cyan fluorescent protein) fusion protein was transiently expressed in the leaf epidermal cells of Arabidopsis, along with 35S:EYFP fusion proteins of several different core cell cycle proteins. Subsequently, protein–protein interaction in the cotransformed leaves was analyzed by the acceptor bleaching fluorescence resonance energy transfer (FRET) method. As a positive-control FRET protein pair, the Arabidopsis transcription factor TGA5 (At5g06960), whose self-interaction in plants was previously detected by FRET analysis (Cheng et al., 2003), was used. As a negative control, the noninteracting LexA-NLS (a bacteria protein fused to SV40 T-antigen nuclear localization signal) and TGA5 proteins were used (Kato et al., 2002). As an additional negative control, ECFP:SIM was tested for interaction with EYFP:LexA-NLS (Table 1). We observed that SIM interacted with the D-type cyclins CYCD2;1, CYCD3;2, and CYCD4;1 and with CDKA;1 (Table 1). By contrast, no significant association was observed between SIM and any of the A- or B-type cyclins tested nor with the B-type CDK, CDKB1;1 (Table 1). One homolog, SMR2, was shown to

Figure 1. sim Loss-of-Function Phenotype.

(A) Scanning electron micrograph of a wild-type trichome.
(B) Scanning electron micrograph of sim mutant trichome. Arrows indicate cell junctions.
(C) Light micrograph of sim loss-of-function phenotype.
(D) Complementation of sim loss-of-function phenotype by pGL2:SIM.

(E) Complementation of sim loss-of-function phenotype by genomic fragment.
(F) GUS staining pattern of sim-2 enhancer trap.
(G) pCYCB1;1:GUS expression in sim mutant trichomes. Bars in (A) and (B) = 200 μm.
Figure 2. SIM Encodes a Small Protein of Unknown Function Defining a Small Gene Family in Arabidopsis and Other Plants.

(A) The SIM locus (At5g04470). Sequence changes in mutant alleles are indicated; the sim-2 gene contains an insertion of the pD991 enhancer trap T-DNA that deletes 1461 bp, including 272 bp of coding sequence and 1189 bp of upstream sequence. simL and simR indicate the primers used for RT-PCR (see Supplemental Figure 1 online).

(B) Alignment of conceptual translation of SIM reading frame and related plant proteins. The regions numbered 1 to 5 denote conserved domains referred to in the text. Sl, Solanum lycopersicum; St, Solanum tuberosum; Zm, Zea mays; Os, Oryza sativa; Pt, Populus tremula; Gm, Glycine max.

(C) Similarity between SIM and D-type cyclin binding domain of ICK/KRPs.
interact with CYCD2;1, demonstrating that another SIM family member also associates with a D-type cyclin.

B-type cyclins are required for mitosis and are not normally expressed in wild-type trichomes (Schnittger et al., 2002a). To determine whether B-type cyclins are expressed in \textit{sim} mutant trichomes, a \textit{CYCB1;1}:GUS fusion gene including the \textit{CYCB1;1} promoter and the N-terminal portion of the coding region that encodes the cyclin destruction box was introduced into \textit{sim} plants by crossing. This fusion is thought to mimic the expression pattern of \textit{CYCB1;1} and has been used in other studies to identify G2/M cells (An-CoLo An-Carmona et al., 1999). We detected GUS expression in a fraction of developing \textit{sim} trichomes (Figure 1G), presumably those in G2/M, suggesting that \textit{CYCB1;1} is ectopically expressed in these cells. No GUS expression was seen in >1000 trichomes of wild-type plants containing this construct. SIM Family Expression in Plants

The existence of ESTs and our RT-PCR data indicate that SIM is expressed. Random amplification of cDNA ends (RACE) and sequencing of RT-PCR–derived cDNA was used to confirm the annotated transcript from public databases. 5’ RACE indicates that the 5’ terminus of the SIM transcript lies at chromosomal position 1267369 on chromosome five, 92 bp upstream of the start codon. 3’ RACE indicates that the 3’ terminus of this mRNA lies 217 bp downstream of the stop codon at chromosomal position 1266668. Sequencing of the complete PCR-amplified cDNA confirmed that, as annotated, the gene contains no introns.

Tissue-specific expression of the SIM family was analyzed using quantitative RT-PCR. As can be seen in Figure 4A, all family members are expressed to some degree in all tissues examined:

### Table 1. SIM Protein–Protein Interactions Determined by FRET

<table>
<thead>
<tr>
<th>Donor (CFP Fusion)</th>
<th>Acceptor (YFP Fusion)</th>
<th>FRET Efficiency (%)(^a)</th>
<th>(n)^b</th>
<th>(P)^c</th>
<th>Interaction?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGA5</td>
<td>TGA5</td>
<td>17.75</td>
<td>10</td>
<td>&lt;0.0001</td>
<td>Y</td>
</tr>
<tr>
<td>TGA5</td>
<td>LexA-NLS</td>
<td>1.10</td>
<td>10</td>
<td>1.000</td>
<td>N</td>
</tr>
<tr>
<td>SIM</td>
<td>LexA-NLS</td>
<td>0.79</td>
<td>10</td>
<td>0.9300</td>
<td>N</td>
</tr>
<tr>
<td>SIM</td>
<td>CDKB1;1</td>
<td>2.16</td>
<td>10</td>
<td>0.6400</td>
<td>N</td>
</tr>
<tr>
<td>SIM</td>
<td>CYCA2;2</td>
<td>2.90</td>
<td>10</td>
<td>0.4300</td>
<td>N</td>
</tr>
<tr>
<td>SIM</td>
<td>CYCB2;1</td>
<td>2.01</td>
<td>15</td>
<td>0.5200</td>
<td>N</td>
</tr>
<tr>
<td>SIM</td>
<td>CDKA;1</td>
<td>21.91</td>
<td>10</td>
<td>0.0003</td>
<td>Y</td>
</tr>
<tr>
<td>SIM</td>
<td>CYCD2;1</td>
<td>14.81</td>
<td>11</td>
<td>0.0003</td>
<td>Y</td>
</tr>
<tr>
<td>SIM</td>
<td>CYCD3;2</td>
<td>13.22</td>
<td>10</td>
<td>0.0190</td>
<td>Y</td>
</tr>
<tr>
<td>SIM</td>
<td>CYCD4;1</td>
<td>17.70</td>
<td>10</td>
<td>0.0120</td>
<td>Y</td>
</tr>
<tr>
<td>SMR2</td>
<td>CYCD2;1</td>
<td>27.43</td>
<td>10</td>
<td>0.0005</td>
<td>Y</td>
</tr>
</tbody>
</table>

\(^a\) The acceptor photobleaching method was used to determine the FRET efficiency (see Methods).

\(^b\) Number of nuclei analyzed.

\(^c\) Student’s t tests were performed for each data set. \(P\) indicates the statistically significant difference between each data set and the negative control values.
roots, stems, flowers, siliques, and rosettes. Although sim mutants have no obvious root phenotype (Walker et al., 2000), SIM expression is particularly strong in root tissue (Figure 4A). While the overall expression pattern of the four genes is quite similar at this whole-tissue level, the relatively high expression level of SMR1 in all inflorescence tissue (stem, unopened flowers, and siliques, Figure 4A) is notable. In situ hybridization to shoot apices shows that SIM is expressed throughout the shoot apical meristem and in leaf primordia (Figure 5), including developing trichomes (Figure 5A, closed arrows), and in procambial strands and developing vasculature (Figure 5A, open arrows).

As mentioned above, sim-2 mutants carry an insertion at the SIM locus that includes a GUS enhancer trap. The sim-2 mutants exhibit GUS expression in both developing and mature trichomes as well as in stipules (Figure 1F). Strong expression is also seen in the vasculature of both the aerial organs and the root (see Supplemental Figure 2 online). The relatively trichome-specific expression of the enhancer trap in developing leaves suggested that a trichome-specific enhancer was located either upstream or downstream of the SIM coding region. To test more directly whether some aspects of SIM expression were under the control of the trichome developmental pathway, we took advantage of Arabidopsis strains expressing various levels of function of the key trichome development transcription factor GL3 (Payne et al., 2000; Zhang et al., 2003). Plants doubly mutant for gl3 and its functional duplicate egl3 lack GL3 function and produce no trichomes, while the GL3OE line used here overexpresses the GL3 transcript and produces a greater number of larger, extra-branched trichomes. Thus, genes that are specifically expressed during trichome development should therefore show little or no expression in gl3 egl3 mutants and increased expression in the GL3OE line. SIM transcript levels increase with increasing GL3 function, suggesting that SIM expression is at least partially under direct or indirect control of GL3 (Figure 4B).

Plants Overexpressing SIM Have Greatly Enlarged Cells

To investigate the biological role of SIM in plant development, transgenic plants expressing SIM ectopically from the cauliflower mosaic virus 35S promoter were produced. Six transgenic lines

Figure 4. Expression of the SIM Gene Family in Various Arabidopsis Tissues.

(A) Absolute quantification of SIM family transcript levels by quantitative RT-PCR.

(B) Increase of SIM transcript levels in response to increasing levels of GL3 function. Expression of SIM transcripts in leaves of a gl3 egl3 line that lacks GL3 function, a Columbia (Col) wild-type line with normal GL3 function, and a line overexpressing GL3 was compared by quantitative RT-PCR. The values shown represent averages of three separate biological replicates ± SD.

Figure 5. Expression Pattern of SIM.

In situ RNA hybridizations to longitudinal sections of the shoot apex probed with DIG-labeled single-stranded antisense SIM (A) and sense probe of Ceratopteris richardii gene of unknown function (B; accession number CV735270). Closed arrow indicates a developing trichome. Open arrows indicate procambial strands and developing vasculature. Bars = 200 μm.
containing the 35S:SIM construct were generated. Plants from five of the six lines showed a similar phenotype. SIM-overexpressing plants are dramatically reduced in size compared with the wild type (Figure 6A). Quantitative RT-PCR revealed that these lines express ~50- to 90-fold more SIM transcripts than wild-type plants. Although the overall size of the plant is greatly reduced, 35S:SIM plants (Figures 6B, 6D, and 6F) contain abnormally large epidermal cells in comparison with the wild type (Figures 6C and 6E). As observed by scanning electron microscopy, cell patterning is highly irregular. Cells of SIM-overexpressing plants are

Figure 6. Phenotypic Analysis of SIM-Overexpressing Plants.

(A) Four-week-old wild-type (left) and 35S:SIM (right) plants with inflorescences removed.
(B) Second leaf of 35S:SIM plant shown in (A).
(C) Adaxial epidermal pavement cells of wild-type first leaf.
(D) Adaxial epidermal pavement cells of 35S:SIM first leaf.
(E) Cross section through wild-type first leaf.
(F) Cross section through 35S:SIM first leaf.
(G) DAPI-stained epidermal pavement cell nuclei of wild-type first leaf. Arrows indicate nuclei.
(H) DAPI-stained epidermal pavement cell nuclei of 35S:SIM first leaf. Arrows indicate nuclei.

All analyses were done using 4-week-old wild-type and 35S:SIM plants. Bars = 1 cm in (A), 1 mm in (B), 200 μm in (C) and (D), 50 μm in (E) and (F), and 22 μm in (G) and (H).
Plants Overexpressing SIM Undergo Increased Endoreplication

The enlarged epidermal cells in SIM-overexpressing plants contain enlarged nuclei relative to wild-type epidermal cells (Figures 6G and 6H), as expected from the previously described correlation between cell size and nuclear DNA content in Arabidopsis leaves (Melaragno et al., 1993). Analysis by flow cytometry confirms that the epidermal pavement cells of 35S:SIM plants undergo extra rounds of endoreplication (Figure 7A), with increased levels of 8C, 16C, and 32C cells clearly detected at 15 and 21 d after sowing (DAS). However, the extremely large epidermal cells observed in Figures 6B, 6D, and 6F represent only a small fraction of the cells in a leaf, even in regions containing a patch of enlarged epidermal cells, and these rare large nuclei might be missed in the flow cytometry experiments.

To estimate the DNA contents of the largest epidermal cells, in situ measurements were made of 4′,6-diamidino-2-phenylindole (DAPI)-stained nuclei of the largest class of adaxial epidermal cells on 35S:SIM plants and compared with equivalent measurements for wild-type adaxial epidermal cells (Figure 7B). In these experiments, the data are presented in terms of relative fluorescence units (RFUs) of DAPI fluorescence detected in the nuclei, but the data have been normalized to the mean value of 4.4C for the DNA content of wild-type Col epidermal nuclei (Melaragno et al., 1993); thus, the RFU values should represent the approximate DNA contents of the nuclei. On this basis, the nuclei of the enlarged epidermal pavement cells of 35S:SIM plants have DNA contents on average of 93.6 ± 45.9C (Figure 7B). However, the 35S:SIM nuclei clearly fall into two major clusters of ~42 and 85 RFU. Because these clusters are almost exactly twofold different in apparent DNA content, and given that in situ DNA contents are complicated by irregular nuclear shapes, background fluorescence, and other factors, a reasonable interpretation is that these two peaks represent nuclei with 32C and 64C DNA contents, respectively. In this case, the largest class of cells would consist primarily of 64C cells, with a substantial number of 32C cells and a few cells of higher C value. This interpretation is broadly consistent with the results of the flow cytometry study, although no 64C cells were detected by flow cytometry. The very largest cells in the adaxial epidermis of these leaves included only a few dozen cells per leaf, and it is possible that the nuclei of these cells were missed in the flow cytometry. However, it should be noted that the plants for flow cytometry and in situ observations were grown under different conditions; plants for flow cytometry were grown on plates for 21 d, while plants for the in situ measurements were grown on soil for 28 d. These differences in growth conditions may explain the differences in the highest ploidy detected.

**DISCUSSION**

We have isolated and characterized a cell cycle regulator that plays an integral role in controlling the onset of endoreplication in Arabidopsis. Loss of SIM function results in multicellular trichomes with decreased levels of endoreplication (Figures 1B and 1C; Walker et al., 2000). We have shown that SIM is encoded by At5g04470, a gene that previously had no known function. Three additional uncharacterized genes from Arabidopsis, as well as 10 others from a variety of plant species, have been identified as members of a family of proteins that share five distinct domains with SIM (Figure 2B). Sequence analyses reveal that one of these domains is a putative cyclin binding motif, the Cy motif found in some CDK inhibitors, Rb, and E2F (Adams et al.,
Another motif is shared with the CDK-inhibitory ICK/KRP proteins, which are known to bind to D-type cyclins (Wang et al., 1998; De Veylder et al., 2001). We demonstrate that EYFP fusions of SIM and its Arabidopsis homologs are localized to the nucleus (Figures 3C to 3F) and are expressed throughout the plant, including meristems, leaf primordia, and trichomes (Figures 4A and 5A).

Overexpression of SIM results in severely dwarfed plants, with varying degrees of enlarged cells having highly endoreplicated nuclei (Figures 6 and 7). Together with the loss-of-function phenotype, this gain-of-function phenotype supports a role for SIM in inhibiting mitosis, thereby promoting endoreplication. Our in vivo FRET experiments (Table 1) suggest that SIM interacts with one or more CYCD/CDKA;1 complexes but does not interact with B-type mitotic cyclins or with mitosis-specific CDKB-containing complexes. This result is consistent with the observation that sim mutants express a CYCB1;1:GUS reporter gene (Figure 1G) and also express CYCB1;2 transcripts (Schnittger et al., 2002a). These results indicate that SIM acts upstream of the G2 induction of CYCB expression.

The observation that SIM, a negative regulator of mitosis in endoreplicating trichomes, may interact with CYCD/CDK complexes was initially surprising. CYCD/CDK complexes are typically considered to function at the G1/S transition, promoting entry into S-phase (De Veylder et al., 2003; Dewitte and Murray, 2003; Inzé, 2005; Menges et al., 2006). However, there are several lines of evidence that are consistent with the hypothesis that SIM functions primarily or exclusively via its interactions with CYCD/CDK complexes. First, overexpression of a D-type cyclin in trichomes causes production of multicellular trichomes, which phenocopies the sim phenotype (Schnittger et al., 2002b). In sim mutants, D-cyclin overexpression produces an even stronger phenotype, consistent with SIM acting as a negative regulator of D-type cyclins. Second, the sim mutant phenotype is rescued when ICK1/KRP1, a CDK inhibitor known to interact with D-type cyclins, is ectopically expressed in trichomes (Weinl et al., 2005). Finally, like the ICK/KRP proteins, SIM is a small, nuclear-localized protein, and SIM shares one short motif with the ICK/KRP family of proteins (Figure 2C). Considering all of this evidence together, it is likely that SIM functions as a CDK inhibitor by interacting with CYCD/CDK complexes.

Our work and the work of Schnittger et al. (2002b) indicate that in addition to their known role at the G1/S transition, CYCD/CDKA complexes can promote mitosis in developing trichomes. In contrast with this observation, Murray and colleagues have shown that overexpression of CYCD3;1 from the 35S promoter leads to a

**Figure 7. DNA Contents of SIM-Overexpressing Plants.**

(A) Ploidy level distribution of the first leaves of wild-type (Col) and 35S:SIM plants at 9, 15, and 21 DAS as measured by flow cytometry. The indicated values are means ± SD (n = 3 to 5).

(B) Relative DNA contents of DAPI-stained epidermal pavement cell nuclei of wild-type (Col) and 35S:SIM enlarged cells of plants at 28 DAS. The degree of fluorescence is expressed as RFUs normalized to the mean fluorescence of Col. These RFU values have been adjusted to be comparable to published C values (see Methods), but they should only be interpreted as relative comparisons, not as measurements of absolute DNA contents.
decrease in cells in G1 and an increase in cells in G2 as well as an extended G2 phase and a delay in CYCB activation (Dewitte et al., 2003; Menges et al., 2006). These observations are consistent with a specific role of CYCD3;1 at the G1/S transition and do not support a role for this cyclin in promoting mitosis. One possible explanation is that endoreplicating trichome cells lack a cell cycle checkpoint that normally prevents D-cyclins from promoting mitosis, and SIM is needed to prevent mitosis. Alternatively, the in vivo target of SIM may be a specific CYCD/CDK complex that can play a role in promoting mitosis. Arabidopsis has 10 CYCD genes, and only a few have been functionally examined. It is not known which CYCD genes are expressed in developing trichomes, although CYCD3;1 is not (Schnittger et al., 2002b). It should be noted that while the acceptor photobleaching FRET method is unlikely to produce false positives, it is possible that true interactions might not be detected either due to the specific geometry of the complexes involved or to a low signal-to-noise ratio. Thus, we cannot rule out the possibility that SIM might interact with additional complexes among those tested.

The observation that SIM transcript levels show some degree of dependence on expression of the trichome developmental regulator GL3 (Figure 4B) suggests that regulation of SIM expression levels may play a role in the transition to endoreplication during trichome development. However, SIM transcript expression by itself is not sufficient to block mitosis; SIM transcripts are detected throughout the meristem and young leaf primordia (Figure 5A), yet these cells continue to divide. Also, the effect of SIM overexpression is tissue specific, with much larger cells occurring in the adaxial epidermis of leaves than in the abaxial epidermis or mesophyll (Figure 6F). These observations suggest that SIM expression is under posttranscriptional control, like many other cell cycle components, or that SIM function requires other trichome-specific components. Alternatively, these tissues may lack some activator or cofactor required for SIM function.

SIM is expressed in a wide variety of tissues (Figures 4A and 5A), yet the sim mutant phenotype has been detected only in trichome and hypocotyl cells (Walker et al., 2000). The existence of three Arabidopsis SIM homologs, the SMR genes, suggests that the SIM may play additional roles beyond those detected in sim homozygotes that are concealed by functional overlap with the other family members. SIM is particularly strongly expressed in roots (Figure 4A) and in procambial cells and developing vasculature (Figure 5A) and may play a role in root or vascular development. Further insight into the potential role of SIM and the SMR genes in the root comes from the cell-type-specific microarray expression study of Birnbaum et al. (2003). In this study, transcripts of SIM, SMR1, and SMR3 were detected in all tissue layers of the root, and consistent with our results in Figure 4A, SIM expression is the strongest, and SMR3 is much more weakly expressed than the other two genes. In the Birnbaum et al. (2003) study, expression was also examined in three developmental stages along the root axis. Stage one contains the root tip and would be expected to have the highest proportion of dividing cells; this assumption is supported by the relative transcript levels of mitotic cyclins, which were highest in this stage. Stage two included cells that were dividing less frequently and are beginning to differentiate, and stage three contained the zone of rapid cell elongation. Expression of SIM, SMR1, and SMR3 is much lower in stage one than in the other stages, and both SIM and SMR1 exhibit their highest expression level at stage two. Expression of SMR3 steadily increases from stage one, the root tip, to stage three, the expanding cells. This pattern of expression of SIM and its paralogs is consistent with a role in reduced division rates concomitant with cell differentiation and the onset of endoreplication.

One clue to other potential functions of SIM and its homologs is that a tomato (Solanum lycopersicum) homolog of SIM has been implicated in a signaling pathway involved in inflorescence development. SELF-PRUNING INTERACTING PROTEIN4 (SIP4; shown as SIIP4 in Figure 2B) was isolated in a yeast two-hybrid screen using the tomato SELF-PRUNING (SP) protein as bait (Pnueli et al., 2001). SP is the functional homolog of TERMINAL FLOWER (TFL) in Arabidopsis and CENTRORADIALIS (CEN) in Antirrhinum majus (Pnueli et al., 1998). Members of the TFL/SP/CEN family control inflorescence determinacy and architecture (Shannon and Meeks-Wagner, 1991; Alvarez et al., 1992; Bradley et al., 1997; Pnueli et al., 1998). These genes encode members of a plant family of proteins related to the animal phosphatidyethanolamine binding proteins, signaling proteins that appear to act via the kinase Raf1 in animals. A Ser/Thr kinase that was isolated in this yeast two-hybrid screen was shown to phosphorylate SIP4 on a Ser residue that is conserved in SIM and several other homologs (Pnueli et al., 2001; final S residue in motif 5 in Figure 2B). One of the first events classically observed in the transition to flowering is an increase in mitosis in the meristem (Steeves and Sussex, 1989), and it is tempting to speculate that other members of the SIM gene family could be involved in regulating mitotic cycling during this developmental transition.

Our results indicate that SIM encodes a cell cycle regulator that plays a key role in the establishment of endoreplication during trichome development. The SIM protein appears to act by regulating D-type cyclin-containing CDK complexes. Plant genomes have significantly larger gene families for most cell cycle components, suggesting that the plant cell cycle may have additional complexity (Vandepoele et al., 2002). For example, there are 10 CYCD genes in the Arabidopsis genome, whereas in mammals there are only three. Study of SIM and its homologs may give additional insights into the diversity of plant cell cycle responses and their integration with development, in addition to giving insights into the establishment of endoreplication.

METHODS

Isolation of SIM

The sim-2 allele, generated by insertional mutagenesis with the T-DNA pD991, originally segregated two inserts. Linkage between a single T-DNA insert and the sim phenotype and kanamycin resistance was established. The T-DNA right border junction was recovered from genomic DNA by an adaptor PCR method (Siebert et al., 1995) using pD991-specific primers available on the Jack lab website (www.dartmouth.edu/~tjack/index.html).

To pinpoint the position of the pD991 insert in entirety, PCR reactions with various combinations of primers in the At5g04470 region were performed. Failure to amplify a region in sim-2 DNA was indication of the insert; wild-type DNA was used as a positive control for the primers. Finally, the exact position of the left border of pD991 was identified by sequencing a DNA fragment that was PCR amplified using a primer specific for the left border, oligo 156 (5'-CCCTATACGACGGATCG-3').
and primer specific for a region of sim-2 that is able to be amplified, T32m21-23337L (5′-ACATACCTTGCGATGTCCTCGGC-3′).

For molecular complementation analysis, the genomic coding sequence of At5g04470, 2870 bp of upstream sequence, and 250 bp downstream sequence was PCR amplified from the BAC clone T32m21 using the primers simwhole3500 L (5′-ACATACCTTGCGATGTCCTCGGC-3′) and simR (5′-GCTGATCTCATCTTTGTTGACGAT-3′) to produce a sense or antisense probe. A DIG-labeled, single-stranded RNA was harvested from the shoot tissue of 3-week-old Wassilewskija and sim-2 plants using the Omniscript RT kit (Qiagen) to create pSIM3500. The pSIM3500 insert was subcloned as a BarnHI-XhoI fragment into these sites of the binary vector pBIN19 (Bevan, 1984) to create pSIM3500Bin, which was used to complement the sim phenotype.

For RT-PCR analysis of At5g04470 expression in Wassilewskija and sim-2 plants, RNA was harvested from the shoot tissue of 3-week-old Wassilewskija and sim-2 plants using the Plant RNaseasy kit (Qiagen) and treated with DNase according to the manufacturer’s instructions. cDNA was synthesized from this RNA using the Omniscript RT kit (Qiagen). The presence of the SIM transcript was measured by PCR amplification using this cdNA as template and the primers simL (5′-AGATCTGCCATCTTTGATTTCCC-3′) and simR (5′-GCTGATCTCATCTTTGTTGACGAT-3′) to assess the presence of the SIM transcript and the primers HIS4 L (5′-TCGTGGAAAGGGAGGAAAAGGT-3′) and HIS4 R (5′-CTAGCGTGCTCGGTGT-AAGTGAC-3′) to assess the presence of a control gene, HISTONE H4.

Expression Analysis

The uidA (GUS) expression in sim-2 plants was visualized using methods described previously (Larkin et al., 1996). In situ RNA hybridization was essentially performed as described previously (Larkin et al., 1993). DIG-labeled, single-stranded RNA probes were synthesized from PCR-derived template containing an appropriately positioned T7 promoter to produce a sense or antisense probe. A DIG-labeled, single-stranded sense strand of a Ceratopteris richardii (C-fern) gene of unknown function (accession number CV735270) was used as a negative control. For quantitative RT-PCR, RNA was extracted from various organs of wild-type (Col) plants to assess the expression of SIM and its homologs in these tissues using the Plant RNaseasy kit (Qiagen) and treated with DNase using a DNase kit (Qiagen) following the manufacturer's instructions. cDNA was synthesized from the RNA harvested from each tissue using an Omniscript RT kit (Qiagen). The absolute transcript levels for each gene in each tissue type (three replicates of each of the three

Generation of Transgenic Lines and Growth Conditions

The full-length SIM, At1g08180, At3g10525, and At5g02420 coding regions were PCR amplified from the BAC clones T32m21, T32g18, F13m14, and T22f11, respectively, in two-stage PCR reaction and inserted into the GATEWAY vector pDONR221 (Invitrogen) by attB recombination following the manufacturer’s protocol. Error-free entry clones were confirmed by sequence analysis before recombination into the following relevant destination vectors: overexpression, pK2GW7 (www.vib.be); YFP fusions, pDuEX-An (N. Kato, unpublished data); and CFP fusions, pDuEX-Dn (N. Kato, unpublished data). The resulting plasmids were introduced into Agrobacterium tumefaciens by transformation and subsequently into plants (ecotype Col) via the floral dip method (Clough and Bent, 1998). Transgenic plants were selected on kanamycin-containing medium and later transferred to soil. Plants were grown as previously described (Larkin et al., 1999). For YFP and CFP gene fusions, vectors were directly introduced to Arabidopsis thaliana leaves via particle bombardment using a PDS-1000/He biolistic particle delivery system (Bio-Rad), incubated overnight in water at room temperature with constant shaking, and visualized on a Leica TCS SP2 spectral confocal microscope. Scanning electron microscopy was performed as previously described (Larkin et al., 1999).

Nuclear DNA Measurements

DNA contents were measured as previously described (Walker et al., 2000), with the exceptions that nuclei were observed using a ×20 objective lens, and DNA values were normalized to reported wild-type epidermal cell nuclei values (Melaano et al., 1993). Flow cytometric analysis was performed as previously described (Verkest et al., 2005).

FRET Analysis

FRET efficiencies of protein pairs in Arabidopsis leaf epidermal cells were measured by an acceptor bleaching method (Szczesna-Skorupa et al., 2003). In this method, a protein pair is fused to CFP and YFP. YFP is selectively photobleached with a high intensity of the excitation laser, and changes of CFP intensity before and after the YFP photobleaching are monitored. If the protein pair interacts, CFP intensity will increase after YFP photobleaching due to a loss of the FRET. If the protein pair does not interact, CFP intensity will not change. Hence, the FRET efficiency in this method is quantified as:

$$ FRET_{\text{eff}} = \frac{D_{\text{post}} - D_{\text{pre}}}{D_{\text{post}}} $$

where $FRET_{\text{eff}}$ is FRET efficiency, $D_{\text{post}}$ is the fluorescence intensity of the donor (CFP) after acceptor (YFP) photobleaching, and $D_{\text{pre}}$ is the fluorescence intensity of the donor before acceptor photobleaching. FRET efficiency is considered positive when $D_{\text{post}} > D_{\text{pre}}$.

Arabidopsis leaves that transiently express fusion proteins were observed with a Leica TCS SP2 spectral confocal microscope with a ×40, 1.25-numerical aperture oil immersion objective lens and a double 458/514-nm dichroic mirror. The argon laser line of 458 nm was used to excite ECFP (photomultiplier tube window: 465 to 515 nm) and the 514-nm line to excite EYFP (photomultiplier tube window: 525 to 570 nm). To increase photon fluxes, a pinhole size of the confocal microscope was increased to 600 mm from the default size of 81.39 mm (Airy 1). The image was zoomed 3.5- to 5-fold, and the nucleus region where both ECFP and EYFP were detected was selectively bleached with the 514-nm light at 100% laser intensity. Three to ten percent of the 514-nm laser intensity was used to monitor changes of EYFP fluorescence intensity during the bleach. The nucleus region was bleached 20 times in ~2 min, and the ECFP intensity in the bleached region was measured before and after the EYFP bleaching. The FRET Wizard program in the Leica confocal microscope software (LCS 2.61.1537) was used to set experimental conditions and calculate FRET efficiencies.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: SIM (At5g04470), CAB85553; SMR1 (At3g10525), BAC42937; SMR2 (At1g08180), AA18255; SMR3 (At5g02420), CAB58579; Solanum lycopersicum SMR1, A1780963; S. lycopersicum SMR2, AW931119; S. lycopersicum SMR3, AA43410; S. tuberosum SMR1, BM110486; Zea mays SMR1, AZM4-61016; Z. mays SMR2, AZM4-26293; Oryza sativa EL2, TD3676; O. sativa SMR1, AA20082; Populus tremula SMR1, BU815024; and
Glycine max SMR1, AW704877. All of these accession numbers are from the National Center for Biotechnology Information database, with the exception of Z. mays SMR1 and Z. mays SMR2, which are from The Institute for Genomic Research.

Supplemental Data
The following materials are available in the online version of this article.

Supplemental Figure 1. At5g04470 Expression in sim-2 Plants.

Supplemental Figure 2. GUS Expression in sim-2 Plants.

Supplemental Figure 3. Expression of the 35S:GFP:SIM Fusion Protein.

ACKNOWLEDGMENTS
We thank Ying Xiao, David Burk, and Margaret C. Henk of the Socolofsky Microscopy Center for assistance with microscopy and M. David Marks (University of Minnesota, St. Paul, MN) for the sim-3 allele. We thank Juan-Antonio Torres-Acosta for materials and helpful discussions and Sara Maes, Ets Van Der Schuren, Steve Pollock, and Alice Simmons for technical assistance. We also thank Kristen Prufer and James Moroney for critical comments on the manuscript. This work was supported by National Science Foundation Grant I0B 0115684 to D.G.O., and the Louisiana Science Foundation Grant IOB 0444560 to J.C.L., National Science Foundation Grant DBI 0115684 to D.G.O., and the Louisiana Governor’s Biotechnology Initiative. Its contents are solely the responsibility of authors and do not necessarily represent the official views of the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Received June 13, 2006; revised August 15, 2006; accepted October 26, 2006; published November 10, 2006.

REFERENCES


SIAMESE, a Plant-Specific Cell Cycle Regulator, Controls Endoreplication Onset in *Arabidopsis thaliana*


*Plant Cell* 2006;18;3145-3157; originally published online November 10, 2006;
DOI 10.1105/tpc.106.044834

This information is current as of January 8, 2018

<table>
<thead>
<tr>
<th>Supplemental Data</th>
<th>/content/suppl/2006/11/07/tpc.106.044834.DC1.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 54 articles, 31 of which can be accessed free at: /content/18/11/3145.full.html#ref-list-1</td>
</tr>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>