# Daysleeper : from genomic parasite to indispensable gene

Knip, M.

# Chapter 3

## DAYSLEEPER: a nuclear and vesicular-localized protein that is expressed proliferating tissues

**Marijn Knip, Steven Hiemstra, Afke Sietsma, Marina Castelein, Sylvia de Pater and Paul JJ Hooykaas**[§]

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

[§]Corresponding author



## Abstract

### Background

*DAYSLEEPER* is a domesticated transposase that is essential for development in *Arabidopsis thaliana* [1]. It is derived from a hAT-superfamily transposon and contains many of the features found in the coding sequence of these elements [1, 2]. This work sheds light on the expression of this gene and localization of its product in protoplasts and *in planta*. Using deletion constructs, important domains in the protein were identified.

### Results

*DAYSLEEPER* is predominantly expressed in meristems, developing flowers and siliques. The protein is mainly localized in the nucleus, but can also be seen in discrete foci in the cytoplasm. Using several vesicular markers, we found that these foci belong to vesicular structures of the trans-golgi network, multivesicular bodies (MVB's) and late endosomes. The central region as well as both the N- and the C-terminus are essential to DAYSLEEPER function, since versions of DAYSLEEPER deleted for these regions are not able to complement the *daysleeper* phenotype. Like hAT-transposases, we show that DAYSLEEPER has a functionally conserved dimerization domain [3].

### Conclusions

DAYSLEEPER has retained the global structure of hAT transposases and it seems that most of these conserved features are essential to DAYSLEEPER's cellular function. Although structurally similar, DAYSLEEPER seems to have broadened its range of action beyond the nucleus in comparison to transposases.

## Introduction

In the past two decades, transposable elements (TE's) have made a comeback into the spotlights. After being discovered by Barbara McClintock in the 1940's, transposons were long viewed as integral constituents of the so-called "junk-DNA" [4]. These genomic regions were generally viewed non-coding, non-functional sequences. In the past 20 years, however, the view of transposons has changed dramatically. TE's are now thought to be the most important drivers of genome evolution, since they are thought to be responsible for a plethora of ways to influence genes, gene expression, and genome structure [5–7]. TE's have contributed substantially to the protein coding capacity of their host genomes through the incorporation of transposon genes sequences into functional host genes [8]. In plants, a good example of molecular domestication of a transposase gene is the FAR1/FHY3 gene-family. This transcription factor gene family is derived from the transposase of a MULE-type DNA transposon, but is now involved in the far-red light response [9]. DNA transposons code for transposases that can recognize and excise the entire element from the genome in a cut-paste fashion. It is assumed that genes in the FAR1/FHY3 family have evolved a fashion of binding DNA and controlling expression that has directly evolved from the process in which transposases bind to the termini of the transposable elements they excise [9]. Many genes in various genomes have been uncovered over the years that are the result of molecular domestication of transposase genes [10]. *DAYSLEEPER* was described in 2005 as the first essential transposase-derived gene in Arabidopsis [1]. DAYSLEEPER structurally resembles a hAT transposase. It harbors an arginine and lysine-rich nuclear localization signal (NLS), "KRRKKKK", flanking a BED-type zinc finger and 6 identifiable hAT blocks (A to F), but lacks the amino acids essential for mobility. hAT Blocks D, E and F make up a hAT dimerization domain [2]. These hAT blocks are defining characteristics of hAT transposases in all species, although not all transposases possess all six blocks [2]. *DAYSLEEPER* is most likely derived from the Ac cluster elements within the hAT family [1]. DAYSLEEPER was identified by its ability to bind the promoter of the DNA-damage response gene Ku70 *in vitro* and is thought to influence transcription of other genes [1].

Here, we investigated the expression pattern of *DAYSLEEPER*, assessed functional complementation of the *daysleeper* phenotype with different deletions of the *DAYSLEEPER* coding sequence and studied its cellular localization using fluorescent protein fusions.

# Results

### *DAYSLEEPER* expression

To analyze the expression pattern of the *DAYSLEEPER* gene, qRT-PCR was performed to measure *DAYSLEEPER* transcript levels. *DAYSLEEPER* expression was found in all tissues analyzed. Expression levels were set against the expression levels found in material from one week old whole seedlings, using β -6-*TUBULIN* as a control (Figure 1). Relative expression in seedlings was 2 times higher as compared to leaf tissue of 4 week old plants. Expression in stem tissue was low. Higher expression was found in flowers and developing siliques (Figure 1). To obtain a more detailed expression pattern, promoter-reporter constructs were created and studied *in planta*. Analysis of plant lines containing a 1 kb or 3.6kb stretch of DNA directly upstream of the DAYSLEEPER start codon, including the 5' UTR, fused to a *mGFP5:gusA* gene-construct (pSDM4327 and pSDM4328), showed that the *DAYSLEEPER* promoter was most active in the root apical meristem, secondary root meristems and the root central stele (Figure 2A-E). In the upper part of the seedling, expression was found in the shoot meristem and



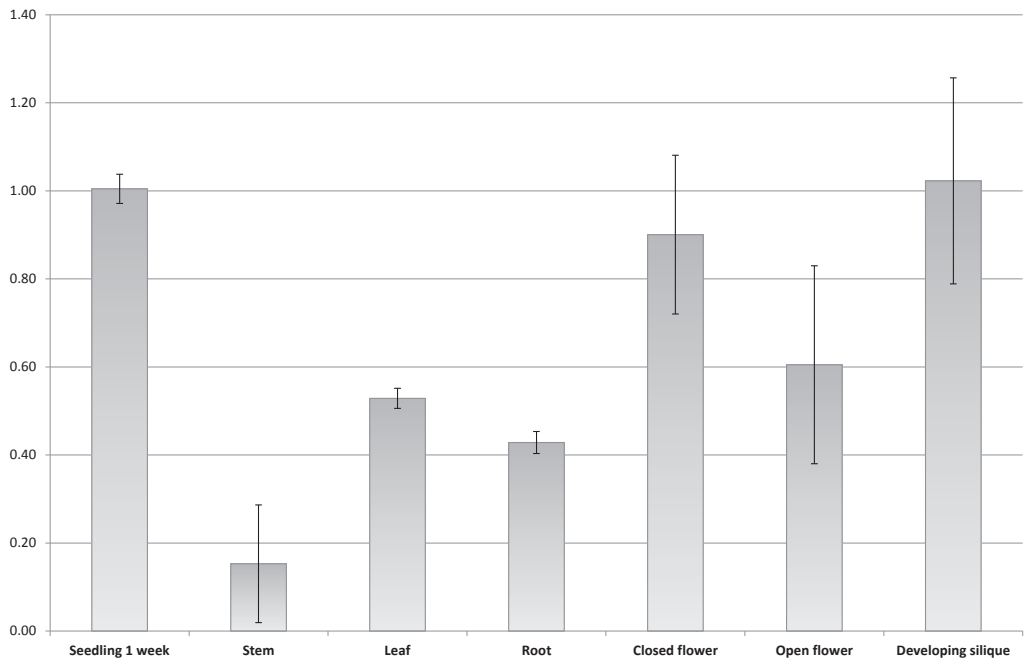


**Figure 1. qRT-PCR analysis of *DAYSLEEPER* expression in seedlings and various tissues of mature plants.** Expression of *DAYSLEEPER* in different organs of 4 week old mature plants is set against the expression of *DAYSLEEPER* in whole seedlings. Expression was normalized against β-*6-TUBULIN* expression.

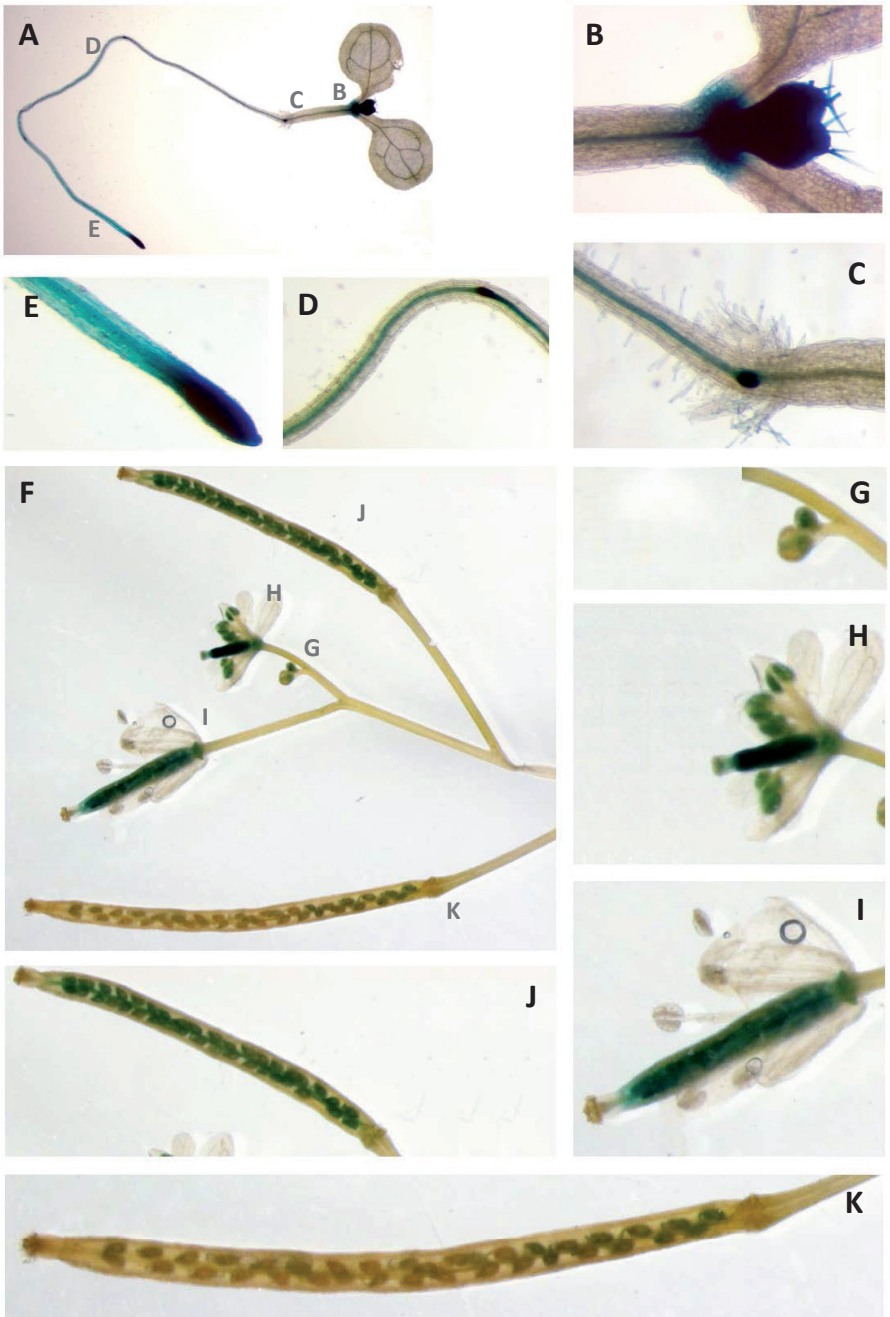


**Figure 2. GUS expression in seedlings, flowers buds, mature flowers and developing siliques**. **A** depicts a seedling, one week post germination, with enlarged areas shown in **B-E. B:** shoot apex and developing leaf**. C:** central stele, root crown. **D**: secondary stele with root meristem. **E**: root apex. **F** depicts an inflorescence displaying several flowers and developing fruits in different stages, with enlarged areas shown in **G-K**. **G:** closed flower buds. **H:** flower with developing fruit. **I:** developing fruit. **J:** young silique. **K:** mature silique.

the embryonic cotyledons (Figure 2 B). As the plant developed, expression was found mainly in proliferating tissues. Strong expression was found in the developing flower bud (Figure 2G). The developing pistil and the anthers displayed high expression levels as the flower developed (Figure 2H-I). In the anthers, expression diminished as the flower reached full maturation (Figure 2I). The expression in the pistil initially was rather uniform, but after fruit initiation was exclusively localized in developing seeds at later stages (Figure 2J-I). In mature siliques and seeds no expression activity was observed anymore (Figure 2 K). No difference in the pattern of *gusA* expression was detected between plants containing the 1kb or 3.6kb *DAYSLEEPER* upstream sequence fragments, although expression levels in plants containing the 1kb promoter fragment seemed higher overall (results not shown).



**DAYSLEEPER localization**

*Complementation of daysleeper with fluorescent protein fusion constructs*

*GFP:DAYSLEEPER:HA* and *DAYSLEEPER:YFP:HA* harboring plants were created in a *DAYSLEEPER (daysleeper)* heterozygous background. These fusion proteins were expressed under control of the native *DAYSLEEPER* promoter, including the 5' UTR. The *GFP:DAYSLEEPER:HA* construct was able to complement the *daysleeper* phenotype and fully restore the wild-type phenotype in the next generation, whereas the *daysleeper* phenotype could not be restored by the *DAYSLEEPER:YFP:HA* construct. Plant tissues were observed using confocal microscopy to study DAYSLEEPER localization *in planta.* GFP:DAYSLEEPER:HA was found in the nucleus of all cells of the plants (Figure 3A). Due to the depicted focal plane, Figure 3B and C do not show a nuclear signal for every cell, but nuclear signals were observed in all cells. Besides, in the elongation zone of the root and the root-tip, a cytosolic localization was also observed in vesicle-like structures (Figure 3B and C) in both primary and secondary roots. The DAYSLEEPER fluorescent signals in these vesicles partly overlapped with the staining pattern of the membrane-specific fluorescent dye FM4-64 (Figure 3B-C) [11]. This was most pronounced in the epidermis of the root.

*DAYSLEEPER localizes to multivesicular bodies, late endosomes and the trans-golgi network*

N-terminal TagRFP- or Cerulean- tagged DAYSLEEPER fluorescent fusions were expressed under control of the strong CaMV 35S promoter in protoplasts. We found

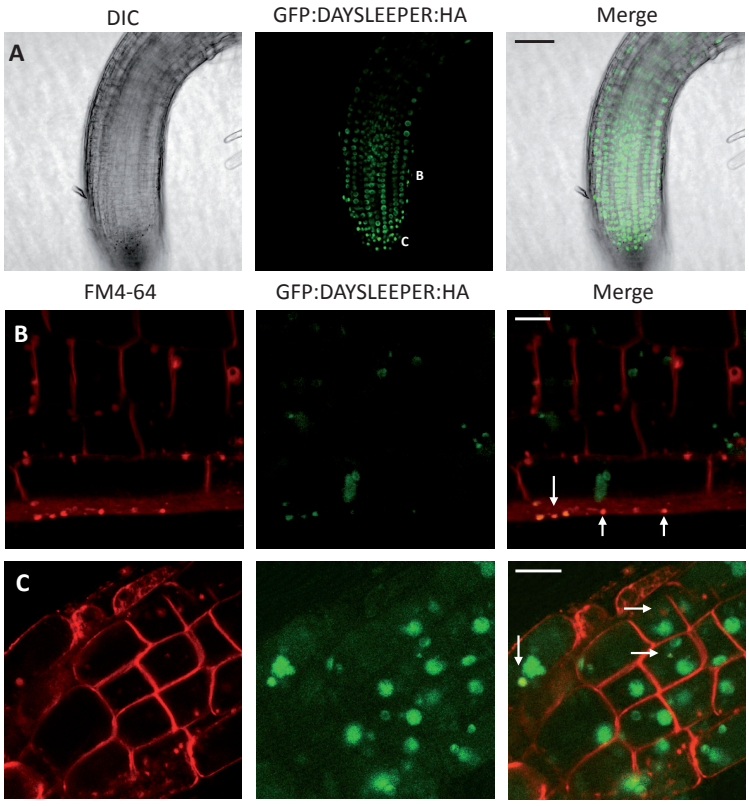


**Figure 3. Localization of fluorescent DAYSLEEPER fusions in Arabidopsis root cells. A.** An overview of the apical part of a seedling root, showing the predominant nuclear localization of GFP:DAYSLEEPER:HA. The white "B" and "C" indicate at which position of the root the pictures in panel **B** and **C** were taken (**B** and **C** are derived from different roots). **B.** depicts root epidermis and cortex cells and **C.** root cap cells, showing vesicles and membranes (FM4-64) or GFP fluorescence (DAYSLEEPER), or both (merge). Vesicular colocalizations are marked with an arrow. The scale bar in **A** represents 50µm. The scale bars in **B** and **C** represent 10µm.



that they localize in nuclei in all protoplasts and also in multi-vesicular structures and smaller vesicular structures outside the nucleus in 32% (SD=3%, n=3, 185 cells counted in total) of protoplasts (Figure 4 and 5E). To further investigate the observed vesicular localization of fluorescent DAYSLEEPER fusion proteins, we used fluorescent marker constructs for different vesicular organelles. We found DAYSLEEPER in multivesicular structures (Figure 4A), but also smaller vesicular structures (Figure 4C). Since we found that DAYSLEEPER interacts with a subunit of the ESCRTIII machinery, we suspected DAYSLEEPER localization to be in late endosomes and multi-vesicular bodies (MVBs) (Chapter 4, this thesis). In order to verify this localization, co-expression was performed in protoplasts with genes of which the localization of the corresponding proteins is well described, namely *SNX1*, *SYP61* and *RHA1*. SYP61 and SNX1 are both found in the trans-golgi network, but SNX1 localizes to sorting endosomes as well [12, 13]. RHA1 localizes to late endosomes and MVB's [14]. DAYSLEEPER-containing multivesicular structures are sometimes also positive for SNX1, which appears to be on the outer membrane of the multivesicular structure, while the DAYSLEEPER fusion protein seems to be internalized

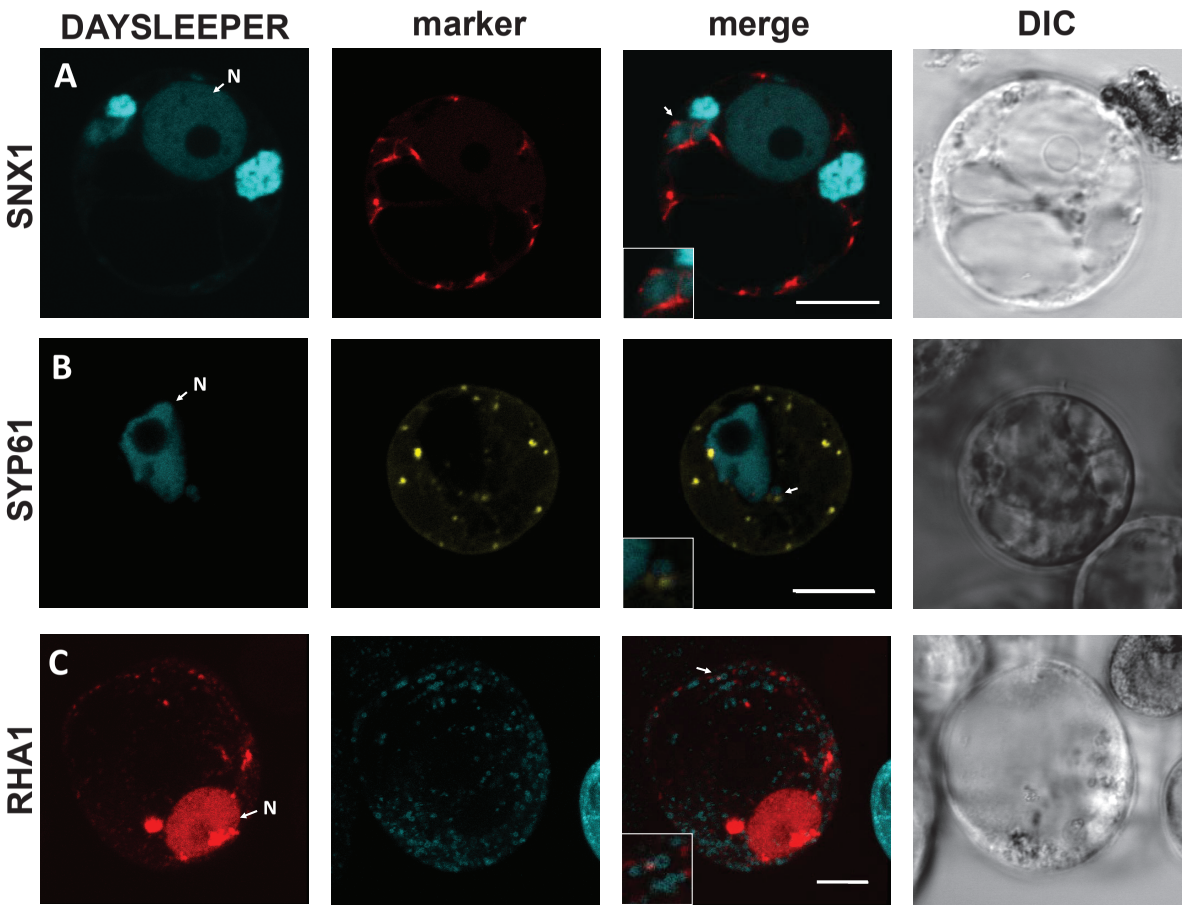


**Figure 4. Localization of fluorescent-tagged DAYSLEEPER in protoplasts.** Arabidopsis protoplasts transformed with **A.** *35S::Cerulean:DAYSLEEPER* and *35S::SNX1:mRFP1* constructs, **B**. *35S::Cerulean:DAYSLEEPER* and *35S::SYP61:YFP* constructs and **C.** 35S::*TagRFP:DAYSLEEPER* and *35S::RHA1:eCFP* constructs. From left to right, the panels depict fluorescent-tagged DAYSLEEPER (DAYSLEEPER), a fluorescent-tagged marker gene (marker), a composite image (merge) and a DIC image (DIC). The areas indicated by arrows are shown enlarged. The nucleus is designated with an "N". The scale bars represent 10µm.

(Figure 4A). For SYP61 and RHA1 colocalization with DAYSLEEPER-fusion constructs we observed similar patterns. We found vesicles that contained two fluorescent constructs, (Figure 4B and C) however the majority of vesicles were positive for either DAYSLEEPER-fusion protein or the marker gene. We propose that DAYSLEEPER-fusion proteins accumulate in larger vesicular structures, that are most likely MVBs (Figure 4).

**DAYSLEEPER structure**

*DAYSLEEPER can homodimerize through its hAT-dimerization domain*

Six conserved motifs of amino acid sequences, so-called hAT-blocks, have been found that are characteristic for hAT transposases in different species [15]. Ac-elements do not always possess all hAT-blocks; about 50% contain blocks A, B and/or C and about 75% contains blocks D, E and/or F [2]. *DAYSLEEPER* contains signatures of all blocks, although hAT block B is somewhat different from the reported consensus [2]. In DAYSLEEPER, this 13 amino acid (AA) sequence lacks a histidine at position 7, as well as the asparagine and leucine at position 10 and 11 [2]. The highly conserved tryptophan at position 26 is replaced by a leucine in DAYSLEEPER. Block F (28

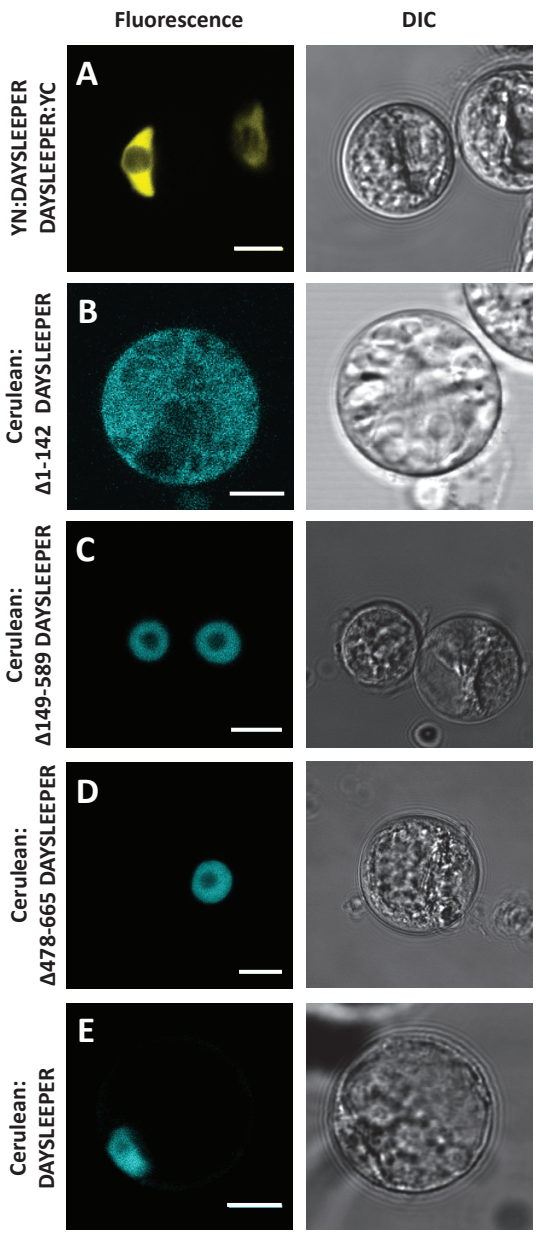


**Figure 5. DAYSLEEPER dimerization and localization of deletion constructs in protoplasts.** **A.** Bimolecular fluorescence complementation of YN:DAYSLEEPER and DAYSLEEPER:YC. **B**. Cerulean: Δ1-142 DAYSLEEPER. **C**. Cerulean: Δ149-589 DAYSLEEPER **D**. Cerulean: Δ478-665 DAYSLEEPER. **E**. Cerulean:DAYSLEEPER. Left panels: fluorescent-tagged DAYSLEEPER. Right panels: DIC images. The scale bars represent 10μm.

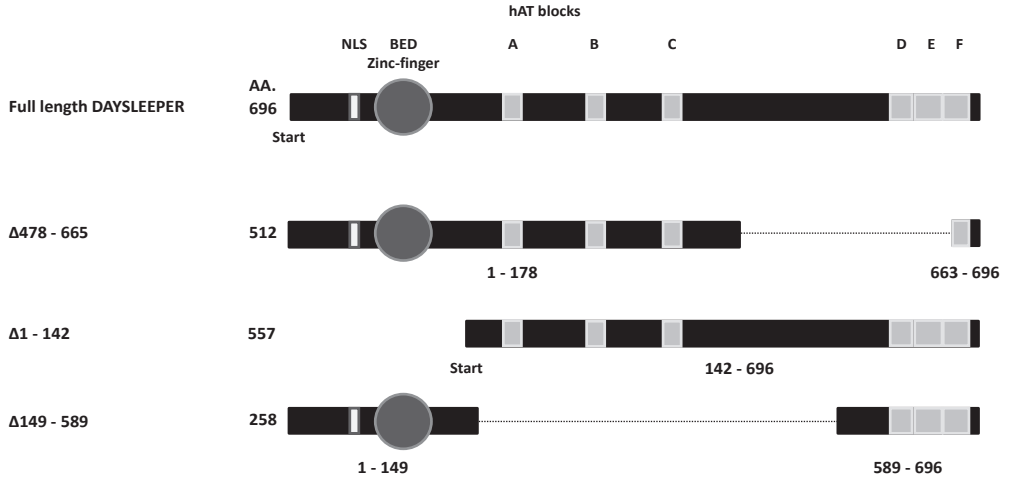


**Figure 6. DAYSLEEPER structure and deletions.** DAYSLEEPER is 696 amino acids long and has a NLS, a BED zinc finger and hAT blocks. The hAT blocks are located at the following positions: Block A: 231-243, B: 333-345, C: 397-406, D: 607-617, E: 621-645, F: 647-674. Three shortened coding sequences are displayed, which are deleted at either the N-terminus (Δ1-142), the central part (Δ149-589), or at the C-terminus (Δ478-665). Sizes of the coding sequences of different constructs are shown, as well as the position of deletions.

AA's) is clearly identifiable, but the first 6 AA's of this sequence are different from the consensus sequence. Of these six blocks, named A to F, hAT block D, E and F make up a dimerization domain in transposases. Since these blocks are largely conserved in DAYSLEEPER, we tested if DAYSLEEPER is able to dimerize via Bimolecular Fluorescence Complementation (BiFC) [16]. We created an N-terminal fusion of DAYSLEEPER fused to the N-terminal part of YFP (AA 1-155, YN) and a C-terminal fusion of DAYSLEEPER fused to the C-terminal half of YFP (AA 155-238, YC) [17]. Co-transformation of these constructs (*YN:DAYSLEEPER and DAYSLEEPER:YC*), both controlled by the strong CaMV 35S promoter, resulted in fluorescence reconstitution of YFP in *Arabidopsis* protoplasts (Figure 5A). A fluorescent signal was exclusively found in the nucleus. We also tested interaction of *YC:DAYSLEEPER* and *YN:DAYSLEEPER.* In this case localization was observed in the nucleus, but also in vesicular structures (results not shown). However this combination produced a signal only in a few non typical-cells, irrespective of the transformation efficiency of protoplasts. Constructs did not display fluorescence on their own, or in combination with a non-fused split-YFP moiety. This shows that DAYSLEEPER was able to dimerize. To test whether the hAT-blocks D, E and F were indeed responsible for dimerization, we co-transformed protoplasts with a *YN:DAYSLEEPER* construct and a Δ*478-665 DAYSLEEPER:YC* construct. The shortened version of DAYSLEEPER lacks a large

part of its C-terminus (Figure 6). We were not able to find a reconstituted YFP signal in these experiments (results not shown), indicating that the C-terminal hAT blocks D,E and F are indeed responsible for homodimerization.

*Full-length DAYSLEEPER coding sequence is required for complemention of the daysleeper phenotype*

In order to study the biological function of different regions of the *DAYSLEEPER* coding sequence *in vivo,* complementation was attempted of *daysleeper* mutant plants with 3 different deletion versions of the *DAYSLEEPER* coding sequence (Figure 6). All shortened coding sequences (Figure 6) were preceded by the native upstream sequence of *DAYSLEEPER*, including the 5' UTR. Deletion of either the C- or N-terminus (Δ1-142 and Δ478-665, respectively) or the central region (Δ149-589) resulted in non-functional DAYSLEEPER, since no complementation was observed. The mutant plantlets did not display any visible change in respect to plants not harboring partial *DAYSLEEPER* coding sequences. This indicated that the regions deleted in the three partial DAYSLEEPER versions are essential for its function.



*DAYSLEEPER's N-terminus is essential for nuclear localization*

Shortened versions of DAYSLEEPER (Figure 6) were fused to the fluorophore Cerulean (pSDM4367-4369) and were transformed into protoplasts (Figure 5BCD). Localization of DAYSLEEPER with central (Δ149-589) and C-terminal (Δ478-665) deletions were similar to the localization of fluorescent fusions of full-length DAYSLEEPER (Figure 5ACDE), whereas the N-terminal truncated construct (Δ1-142) never showed a nuclear signal and was uniformly localized in the cytosol (Figure 5B). The lack of a nuclear signal is in line with the location of the proposed NLS, discussed in the first paragraph, which is missing in the Δ1-142 deleted protein.

## Discussion

DAYSLEEPER does not seem to possess any known protein domains other than those also found in hAT transposases, but although DAYSLEEPER seems very similar to canonical hAT transposases, it must be noted that hAT blocks are rather loosely defined. DAYSLEEPER contains the signatures of these blocks, but little functional conclusions can be drawn from this, since the blocks are defined on the basis of homology and not functionality [2]. However, blocks D, E and F have been found to be functionally responsible for dimerization, which makes it likely that DAYSLEEPER would also be able to form dimers. We have shown that DAYSLEEPER indeed formed homodimers (Figure 6).



The localization of DAYSLEEPER was studied by fusion of the protein to fluorescent proteins. It appeared that the localization of DAYSLEEPER, partially depended on whether the fluorescent moiety was fused to its C- or N-terminus. When the fluorescent moiety was fused to the C-terminus of DAYSLEEPER, an exclusive nuclear localization was observed, in contrast to N-terminally tagged DAYSLEEPER fusion-proteins, which showed besides a vesicular localization. Although we did find cells expressing the construct, the N-terminal fluorescent-tagged Δ1-142 DAYSLEEPER fusion-protein appeared to be toxic to protoplasts, since the majority of cells died after transformation. The C-terminally fluorescent-tagged Δ1-142 DAYSLEEPER fusion-protein did not appear to display this toxicity, but also showed a cytoplasmic localization (data not shown). Dimerization of DAYSLEEPER was observed in nuclei of protoplasts. However, it must be noted that BiFC-pairs are stable once formed [24]. This property leaves the possibility that dimers are transported after they are formed, although dimerization was not observed outside the nucleus. In protoplasts, DAYSLEEPER:YFP fusions did not localize outside of the nucleus (data not shown), whereas N-terminally tagged constructs localized to both the nucleus (Figure 5E) and vesicular structures in ~30% of the cells (Figure 4). However, there was no difference detectable in localization between N- or C-terminal-tagged fusion proteins *in planta.* Although these constructs seemed to be localized similarly, C-terminal tagged DAYSLEEPER appeared non-functional, since it did not complement the *daysleeper* phenotype. We speculate that the difference in localization found in protoplasts are due to the fact that the studied proteins were expressed at a much higher level than is the case *in planta*, where *DAYSLEEPER*-constructs were expressed under control of the native *DAYSLEEPER*-promoter.

Extranuclear localization was unexpected in light of DAYSLEEPER's structure and suspected functionality. The protein was shown to bind to DNA suggesting that DAYSLEEPERS plays a role in the nucleus, as was found for most domesticated transposases [1, 10]. We speculate therefore that DAYSLEEPER is to some extent transported out of the nucleus by an interaction partner. We have found that DAYSLEEPER can bind Arabidopsis homologs of the ESCRTIII machinery subunit VPS2 (Chapter 4, this thesis). The ESCRTIII machinery is highly conserved and its main function is to snare off vesicles from membranes. We speculate that DAYSLEEPER's binding to VPS2-homologs in Arabidopsis might facilitate the translocation of DAYSLEEPER from the nucleus to vesicles. VPS2.2 has recently been shown to be partially localized to nuclei in Arabidopsis roots [18]. Further studies will have to reveal the functional implications of this interaction. Based on our co-localization experiments (Figure 4), we speculate that DAYSLEEPER is transported from the nucleus through the trans-golgi network and targeted to MVB's and late endosomes. We were not able to definitively show colocalization of MVB's with DAYSLEEPER using our marker-gene set, although in Figure 4A multi-vesicular structures can be seen, that are not stained by SNX1 and therefore might be MVB's. Future analysis of localization should discern the precise nature of DAYSLEEPER localization, by analyzing constructs *in planta*, instead of in a semi-artificial protoplast system.



## Conclusions

DAYSLEEPER is a predominantly nuclear protein that is expressed mainly in meristems, developing flowers and fruits. It is able to dimerize, most likely enabled by its hAT transposase-like dimerization domain. Although nuclear in most cells, vesicular localization was observed in root-tips and in protoplasts. We hypothesize that the N-terminal "KRRKKKK" nuclear localization motif of DAYSLEEPER is responsible for its nuclear localization, and that interaction with other factors allows it to be exported outside the nucleus. We propose DAYSLEEPER's vesicular localization is situated in the trans-golgi network, late endosomes and MVBs.

# Material and Methods

**Cloning**

PCR primers and vectors can be found in Table S1 and S2, respectively. PCR reactions were performed using the Phusion® polymerase (Finnzymes ®) with HF-buffer and recommended settings, unless stated otherwise. Cloned PCR products were sequenced. Restriction enzymes were obtained from Fermentas®, except for AlwNI, AvaI and HpaI which were obtained from New England Biolabs (NEB ®).



*cDNA isolation*

A λ-phage cDNA library constructed from auxin-treated Arabidopsis roots was used as a template for the amplification of the full length coding sequences of *DAYSLEEPER*, *RHA1* and *SNX1* [19].

*Binary vectors for promoter analysis using the gusA reporter gene*

*DAYSLEEPER* promoter constructs were made in pCAMBIA1304 (CAMBIA foundation) to obtain promoter-reporter gene fusions. First, to separate the *DAYSLEEPER* promoter from promoter elements present on the pCAMBIA1304 vector, λ phage HindIII DNA marker (New England Biolabs ®) was digested using KpnI and BamHI and the 5kb fragment was cloned into the respective sites in the multiple cloning sites (MCS) of pCAMBIA1304, resulting in pCAMBIA1304λ. Using primer combination MK3 and MK4, 6.1kb of the *DAYSLEEPER* upstream region was amplified from genomic DNA. This fragment was subsequently cloned into pJET1.2 (Fermentas®), giving rise to pJET1.2 6.1kb p*DAYSLEEPER*. Using NcoI and EcoRI, the 35S promoter of the pCAMBIA1304λ vector was replaced by 3.6kb of the *DAYSLEEPER* promoter, resulting in pSDM4328. The same was done using the enzymes NcoI and XbaI, giving rise to a pCAMBIA1304λ vector with 1kb of *DAYSLEEPER* upstream sequence cloned in its MCS, resulting in pSDM4327. In both plasmids (pSDM4327 and pSDM4328) the *mGFP5:gusA* coding sequence is preceded by *DAYSLEEPER* upstream sequence, which in turn is spaced from elements present in the pCAMBIA backbone by a 5kb stretch of λ DNA.

*Binary vectors for DAYSLEEPER complementation and YFP fusions*

*DAYSLEEPER* coding sequence was amplified using RedTAQ (Sigma Aldrich®) and primers PB1 and PB2 and cloned into pGEMtEASY (Promega®) to give rise to pGEMtEASY::At3g42170 (pSDM2099) and cloned into pAS2-1 (CLONTECH) to give rise to pSDM2304. In order to delete the central region of *DAYSLEEPER*, pSDM2304 was digested with Age1 and AlwNI (New England Bioscience; NEB®), blunted with T4 polymerase (Fermentas®) and ligated, giving rise to a shortened *DAYSLEEPER* coding sequence (Δ149-589) (pSDM4415).

To create a C-terminal deletion (Δ478-665), pSDM2099 was digested with the restriction enzyme AvaI and subsequently ligated, deleting the sequence between the 2 AvaI sites. The coding sequence was subsequently obtained with NcoI and SmaI and cloned into pAS2-1 (CLONTECH®), resulting in pSDM4416. pAS2-1 vectors were used in a yeast-two hybrid assay that is described in chapter 4 of this thesis. To join the Δ149-589 and Δ478-665 shortened coding sequences with the native *DAYSLEEPER* promoter, the pAS2-1 vectors containing the C-terminal and central truncated *DAYSLEEPER* coding sequence were cut using NcoI and HpaI. The 3.6kb fragment directly upstream of the ATG of the *DAYSLEEPER* coding sequence was inserted, after having been isolated using the same restriction enzymes. This 3.6 kb fragment was obtained from the *DAYSLEEPER* upstream fragment described in the paragraph "*Binary vectors for promoter analysis using the gusA reporter gene".*

To create an N-terminal deletion of the *DAYSLEEPER* coding sequence (Δ1-142), PCR primers MK39.1 and MK40 were used to amplify bases coding for amino acid 142 until the stop codon, adding an NcoI restriction site at the 5'end of the fragment and an EcoRI site flanking the stop codon at the 3'end. The resulting PCR fragment was cut with NcoI and EcoRI (NEB®) and used for cloning into the pJET1.2 6.1kb p*DAYSLEEPER* vector described in the paragraph "*Binary vectors for promoter analysis using the gusA reporter gene".* This plasmid was cut with NcoI and SmaI (NEB®) and ligated with the digested PCR fragment, to give rise to a vector with *DAYSLEEPER* upstream sequence directly fused to the N-terminal (Δ1-142) truncated coding sequence of *DAYSLEEPER*.

*Gateway cloning of binary vectors*

Using Gateway-compatible primers the promoter::coding sequence fusions described above were amplified. This was performed using a slightly modified standard PCR protocol using Phusion polymerase and HF-buffer (Finnzymes®). The annealing temperature was set at 65$^{O}$C for all reactions. The primers used to amplify the different fragments can be found in Table S1. PCR reactions were performed on ~0.5 ng plasmid template per reaction, except for the amplification of the native *DAYSLEEPER* upstream region and coding sequence, which were amplified directly from genomic DNA. All PCR fragments were subsequently cloned into the vector pDONR207 (Invitrogen®), using BP Clonase II (Invitrogen®). The resulting pENTR clones were recombined with a binary destination vector. The three *DAYSLEEPER* versions, Δ1-142, Δ149-589 and Δ478-665 with *DAYSLEEPER* upstream sequence were recombined into pEARLEYGATE301 [20], using LR Clonase II (Invitrogen®), giving rise to pSDM4323 to 4325. The full length genomic sequence of the *DAYSLEEPER* locus was recombined into pGREEN179YFP:HA [21] following the same method, giving rise to pSDM4322.



*Gateway cloning of protoplast vectors*

The pENTR clones described above were cloned into a pART7-derived plasmid containing the appropriate Gateway cassette in frame with a fluorophore coding region [22], using LR Clonase II (Invitrogen®), resulting in pSDM4337 and pSDM4341. For N-terminal fusions to fluorophores, HindIII fragments of the pSITEII 2C1 and 6C1 vectors [23] containing the expression cassettes were cloned into HindIII digested pSY vectors [17, 24]. This gave rise to the pSYSAT6 2xp35S TagRFP Gateway and pSYSAT6 2xp35S Cerulean Gateway vectors, respectively (pSDM4366 and pSDM4376).

*Cloning of the DAYSLEEPER coding sequence into the pSY vectors*

For the Bimolecular fluorescence-complementation assay in Arabidopsis protoplasts, the DAYSLEEPER coding sequence was isolated and restriction sites added (see: Table S1) using PCR, cloned into pJET1.2 (Fermentas ®) and sequenced. The DAYSLEEPER coding sequence was isolated from pJET1.2 using the appropriate restriction enzymes and subsequently cloned into the pSY728, 735, 736 and 738 vectors [17] (Table S1 and S2).

**qRT-PCR Analysis**

Seedlings were grown in vitro for 2 weeks, after being transferred to soil. Plants were grown with 12 hours of light at 21°C. Samples of *Arabidopsis thaliana* Col-0 were collected, flash-frozen in liquid nitrogen and stored at -80°C. The tissue was ground under liquid nitrogen in a TissueLyser II apparatus (Qiagen®). RNA was isolated with the RNeasy Mini Kit (Qiagen®) and 1µg was treated with DNase I (Ambion®), according to the recommended protocol, with the addition of 0.5 µL RNasin (Promega®) per reaction. From each sample, 0.5 µg was used for subsequent random-primed cDNA synthesis, using an iScript cDNA kit (BioRad®). qRT-PCR was performed on 1 µl 20x diluted cDNA, using a standard PCR reaction mix for Phusion DNA polymerase (Finnzymes), with the addition of 1.25 µL 500x diluted SYBR Green (BioRad®) in DMSO. To measure *DAYSLEEPER* transcript levels, primer combination MK1/MK2 was used. Transcript levels were normalized against expression of the housekeeping gene β-*6-TUBULIN* (At5g12250). Primers were adopted from Czechowski *et al*. [25] (Table S2). Experiments were performed in triplicates on a Chromo4 Real-Time PCR Detection system (Biorad®). Data were processed using the Opticon Monitor 3.1 software (Biorad®) and the GeNorm normalization procedure [26].



**Histochemical staining of *gusA* expressing plants**

Seedlings of 10 days old were grown on solid ½ MS medium [27] and stained with bromo-4-chloro-3indolyl-Beta-D-glucuronic acid (X-GLUC). Organs (eg. flower buds, leaves, siliques) of mature plants grown on soil were cut off and stained with X-GLUC staining. Seedlings and various tissues were fixed in 90% acetone for 1 hour at -20$^{O}$C, washed three times in 10mM EDTA, 100mM sodium phosphate (pH7.0), 2mM $K_3Fe(CN)_6$ and subsequently stained for 2 h in 10mM EDTA, 100mM sodium phosphate (pH7.0), 1mM $K_3Fe(CN)_6$, 1mM $K_4Fe(CN)_6$ containing X-GLUC (Duchefa®). Tissue was post-fixed in ethanol-acetate (3:1), cleared in 70% ethanol and stored in 100 mM sodium phosphate (pH7.0).

**Arabidopsis plant, protoplast transformation and microscopic analysis**

*Arabidopsis thaliana* ecotype *Columbia-0 (Col-0)* was used for floral dip transformation according to Clough and Bent, 1998 [28]. *Arabidopsis thaliana* (*Col-0)* suspension cells were used to isolate and transform protoplasts [29]. Protoplasts were observed after 16-18 hours of incubation at 25$^{O}$C in the dark on a Zeiss Observer (Zeiss ®) confocal

microscope.

**Observation of fluorescent constructs in Arabidopsis tissues and protoplasts**

Seedlings were taken directly from ½ MS solid plates and observed on a Zeiss Imager confocal microscope (Zeiss®) prepared on a glass slide with cover slip [27]. Older plants were dissected using a razor blade to allow observation of tissues using a glass slide and cover slip.



Fluorescent signals were visualized using a 63x oil objective on the Zeiss Imager and a 63x water objective with the Zeiss Observer confocal microscope. An Argon laser at 514nm for excitation and a 530/600 nm band pass emission filter was used for GFP and YFP signals. FM4-64 was also excited with the 514 nm Argon laser and the emission was collected using a 530/600 nm band-pass filter. Cerulean was excited using a 458nm laser and the emission was collected using a 475/525 nm band pass filter. TagRFP was visualized using a 543nm laser and a 560/615nm band pass filter. Images were processed using ImageJ (ImageJ, NIH) and Adobe Photoshop CS5 (Adobe ®).

## Author's contribution

MK performed most experiments and data processing in this study. SdP contributed technical advice. MC performed the analysis of the promoter::*gusA* experiment. AS, MK performed experiments for the analysis of DAYSLEEPER dimerization in protoplasts. SH performed most of the protoplast work, including cloning, transformations and microscopy. MK, SdP and PJJH contributed to the study design and writing of the manuscript. All authors have read and approved the final manuscript.



## Acknowledgements and Funding

We would like to thank Gerda Lamers for assistance with the microscopy. We thank Andrew Gleave (Auckland, New Zealand) for the pART7 vectors and Carlos Galvan-Malpudia and Remko Offringa for providing the pGREENII [21] and pART7 [30] gateway vectors. We thank Shaul Yalovsky (Tel Aviv University) for providing the protoplast transformation BiFC vectors. We thank Prof. Peter Pimpl (Universität Tübingen) for providing the *SYP61:YFP* construct. We are also thankful to Bert van der Zaal for providing the λ phage cDNA library and Felix Wittleben for his help with microscopy and cloning. This work is part of the Netherlands Organization for Scientific Research

(NWO) Earth and Life Sciences (ALW) research grant *817.02.003*.

# References


1. Bundock P, Hooykaas P: **An Arabidopsis hAT-like transposase is essential for plant development.** *Nature* 2005, **436**:282–4.

2. Rubin E, Lithwick G, Levy a a: **Structure and evolution of the hAT transposon superfamily.** *Genetics* 2001, **158**:949–57.

3. Yamashita D, Komori H, Higuchi Y, Yamaguchi T, Osumi T, Hirose F: **Human DNA replication-related element binding factor (hDREF) self-association via hATC domain is necessary for its nuclear accumulation and DNA binding.** *J. Biol. Chem.* 2007, **282**:7563–75.

4. Doolittle WF, Sapienza C: **Selfish genes, the phenotype paradigm and genome evolution.** *Nature* 1980, **284**:601–3.

5. Biemont C: **A Brief History of the Status of Transposable Elements: From Junk DNA to Major Players in Evolution**. *Genetics* 2010, **186**:1085–1093.

6. Kazazian HH: **Mobile elements: drivers of genome evolution.** *Science (80- )* 2004, **303**:1626–32.

7. Oliver KR, Greene WK: **Transposable elements: powerful facilitators of evolution.** *Bioessays* 2009, **31**:703–14.

8. Feschotte C, Pritham EJ: **DNA transposons and the evolution of eukaryotic genomes.** *Annu. Rev. Genet.* 2007, **41**:331–68.

9. Hudson ME, Lisch DR, Quail PH: **The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway.** *Plant J.* 2003, **34**:453–71.

10. Sinzelle L, Izsvák Z, Ivics Z: **Molecular domestication of transposable elements: from detrimental parasites to useful host genes.** *Cell. Mol. Life Sci.* 2009, **66**:1073–93.

11. Vida TA, Emr SD: **A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast.** *J. Cell Biol.* 1995, **128**:779–92.

12. Niemes S, Langhans M, Viotti C, Scheuring D, San Wan Yan M, Jiang L, Hillmer S, Robinson DG, Pimpl P: **Retromer recycles vacuolar sorting receptors from the trans-Golgi network**. *Plant J.* 2010, **61**:107–121.

13. Uemura T, Ueda T, Ohniwa RL, Nakano A, Takeyasu K, Sato MH: **Systematic analysis of SNARE molecules in Arabidopsis: dissection of the post-Golgi network in plant cells.** *Cell Struct. Funct.* 2004, **29**:49–65.

14. Lee G-J, Sohn EJ, Lee MH, Hwang I: **The Arabidopsis rab5 homologs rha1 and ara7 localize to the prevacuolar compartment.** *Plant Cell Physiol.* 2004, **45**:1211–20.

15. Essers L, Adolphs RH, Kunze R: **A highly conserved domain of the maize activator transposase is involved in dimerization.** *Plant Cell* 2000, **12**:211–24.

16. Hu C-D, Chinenov Y, Kerppola TK: **Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation.** *Mol. Cell* 2002, **9**:789–98.

17. Bracha-Drori K, Shichrur K, Katz A, Oliva M, Angelovici R, Yalovsky S, Ohad N: **Detection of protein-protein interactions in plants using bimolecular fluorescence complementation.** *Plant J.* 2004, **40**:419–27.

18. Ibl V, Csaszar E, Schlager N, Neubert S, Spitzer C, Hauser M-T: **Interactome of the plant-specific ESCRT-III component AtVPS2.2 in Arabidopsis thaliana.** *J. Proteome Res.* 2012, **11**:397–411.

19. Neuteboom LW, Ng JM, Kuyper M, Clijdesdale OR, Hooykaas PJ, van der Zaal BJ: **Isolation and characterization of cDNA clones corresponding with mRNAs that accumulate during auxin-induced lateral root formation.** *Plant Mol. Biol.* 1999, **39**:273–87.

20. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS: **Gateway-compatible vectors for**

**plant functional genomics and proteomics.** *Plant J.* 2006, **45**:616–29.

21. Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM: **pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation.** *Plant Mol. Biol.* 2000, **42**:819–32.

22. Dhonukshe P, Huang F, Galvan-Ampudia CS, Mähönen AP, Kleine-Vehn J, Xu J, Quint A, Prasad K, Friml J, Scheres B, Offringa R: **Plasma membrane-bound AGC3 kinases phosphorylate PIN auxin carriers at TPRXS(N/S) motifs to direct apical PIN recycling.** *Development* 2010, **137**:3245–55.

23. Tzfira T, Tian G-W, Lacroix B, Vyas S, Li J, Leitner-Dagan Y, Krichevsky A, Taylor T, Vainstein A, Citovsky V: **pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants.** *Plant Mol. Biol.* 2005, **57**:503–16.

24. Martin K, Kopperud K, Chakrabarty R, Banerjee R, Brooks R, Goodin MM: **Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta.** *Plant J.* 2009, **59**:150–62.

25. Czechowski T, Bari RP, Stitt M, Scheible W-R, Udvardi MK: **Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes.** *Plant J.* 2004, **38**:366–79.

26. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol.* 2002, **3**:RESEARCH0034.

27. Murashige T, Skoog F: **medium for rapid growth and bioassays with tobacco tissue cultures**. *Plant Physiol.* 1962, **15**:473–497.

28. Clough SJ, Bent AF: **Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana.** *Plant J.* 1998, **16**:735–43.

29. Schirawski J, Planchais S, Haenni AL: **An improved protocol for the preparation of protoplasts from an established Arabidopsis thaliana cell suspension culture and infection with RNA of turnip yellow mosaic tymovirus: a simple and reliable method.** *J. Virol. Methods* 2000, **86**:85–94.

30. Gleave AP: **Short communication A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome**. *Plant Molecular Biology* 1992, **7**:1203–1207.

# Supplemental Tables

| Primer | Description | Sequence |
|---|---|---|
| AS01 | DAYSLEEPER CDS cloning into pSY735/36, forward primer Sal1 | CAGTCGACTATGGAAGTGTACAATGACGATAC |
| AS02 | DAYSLEEPER CDS cloning into pSY735/36, reverse primer Spe1 | CTACTAGTCTATGCTTCAGATTTGATGGTAG |
| AS03 | DAYSLEEPER CDS cloning into pSY728/38, forward primer Nco1 | CGCGCCATGGATATGGAAGTGTACAATGACGATAC |
| AS04 | DAYSLEEPER CDS cloning into pSY728/38, reverse primer Not1 | GCGCGGCCGCTGTGCTTCAGATTTGATGGTAGCAC |
| B6Tf | Beta-6-Tubulin qRT-PCR forward | ACCACTCCTAGCTTTGGTGATCTG |
| B6Tr | Beta-6-Tubulin qRT-PCR reverse | AGGTTCACTGCGAGCTTCCTCA |
| MK01 | DAYSLEEPER qRT-PCR forward | ACAATGACAACAACCCACTG |
| MK02 | DAYSLEEPER qRT-PCR reverse | CACGAACGAGACAAAACCTG |
| MK03 | DAYSLEEPER promotor amplification forward | CCATGGTCTTTGCAACATAACATAAAAAGG |
| MK04 | DAYSLEEPER promotor amplification reverse | GTGATGGCATAGCATATTG |
| MK147 | SNX1 forward primer, gateway | GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGAGAGCACGGAGCAGCCGAGG |
| MK148 | SNX1 reverse primers, gateway | GGGGACCACTTTGTACAAGAAAGCTGGGTGGACAGAATAAGAAGCTTCAAGTTTG |
| MK151 | DAYSLEEPER reverse primer, gateway | CCACACCATCGACTTCTTCT |
| MK39.1 | DAYSLEEPER N-term. deletion forward | CCATGGCTGACACTCCGAGAAGG |
| MK40 | DAYSLEEPER N-term. deletion reverse | GGAATTCCTATGCTTCAGAT |
| MK60 | DAYSLEEPER forward primer, gateway | GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGAAGTGTACAATGACGATACTG |
| PB01 | DAYSLEEPER forward | ATGGAAGTGTACAATGACGATAC |
| PB02 | DAYSLEEPER reverse | CTATGCTTCAGATTTGATGGTAG |
| SH05 | RHA1 forward primer, gateway | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTAGCTCTGGAAACAAGAACA |
| SH06 | RHA1 reverse primers, gateway | GGGGACCACTTTGTACAAGAAAGCTGGGTCAGCACAACACGATGAACTCACTGCC |

**Table S1.** PCR Primers. Primer names, descriptions and sequences are shown.

| Collection number | Description | Purpose |
|---|---|---|
| pSDM2099 | pGEM-t Easy At3g42170 | |
| pSDM4337 | pART7 p35S Gateway FLAG mRFP1 SNX1 | Addition of a fluorescent tag on the C-terminus of a gene for protoplast transformation |
| pSDM4341 | pART7 p35S Gateway FLAG eCFP RHA1 | |
| pSDM4366 | pSYSAT6 2xp35S Cerulean Gateway DAYSLEEPER | Addition of a fluorescent tag on the N-terminus of a gene for protoplast transformation |
| pSDM4367 | pSYSAT6 2xp35S Cerulean Gateway DAYSLEEPER Δ1-142 | |
| pSDM4368 | pSYSAT6 2xp35S Cerulean Gateway DAYSLEEPER Δ478-665 | |
| pSDM4369 | pSYSAT6 2xp35S Cerulean Gateway DAYSLEEPER Δ149-589 | |
| pSDM4376 | pSYSAT6 2xp35S TagRFP Gateway DAYSLEEPER | |
| pSDM4384 | pSY 728 35S::EE:YN:DAYSLEEPER | Protoplast transformation vectors for BiFC assay |
| pSDM4385 | pSY 735 35S::DAYSLEEPER:YC:HA | |
| pSDM4386 | pSY 736 35S::DAYSLEEPER:YN:EE | |
| pSDM4387 | pSY 738 35S::HA:YC:DAYSLEEPER | |
| pSDM4322 | pGREEN179 Gateway YFP:HA pDAYSLEEPER::DAYSLEEPER | Plant transformation vectors for complementation assays |
| pSDM4323 | pEARLEYGATE301 pDAYSLEEPER::DAYSLEEPER N-term. Truncation :HA | |
| pSDM4324 | pEARLEYGATE301 pDAYSLEEPER::DAYSLEEPER Central. Truncation :HA | |
| pSDM4325 | pEARLEYGATE301 pDAYSLEEPER::DAYSLEEPER C-term. Truncation :HA | |
| pSDM4326 | pEARLEYGATE301 pDAYSLEEPER::GFP:DAYSLEEPER:HA | |
| pSDM4327 | pCAMBIA1304 Lambda 1kb pDAYSLEEPER::mGFP5:GUSa | Plant transformation vectors for promoter activity assay |
| pSDM4328 | pCAMBIA1304 Lambda 3.6kb pDAYSLEEPER::mGFP5:GUSa | |
| pSDM2304 | pAS2.1 At3g42170 (*DAYSLEEPER*) | |
| pSDM4415 | pAS2.1 DAYSLEEPER Δ149-589 | |
| pSDM4416 | pAS2.1 DAYSLEEPER Δ478-665 | |

**Table S2.** Plasmids used for localization of SLEEPER fusion proteins in protoplasts and complementation of the *daysleeper* phenotype in *Arabidopsis thaliana*. Collection number, brief description and purpose in this work are shown.