

# **Title: Ubiquitin-dependent chloroplast-associated protein degradation in plants**

**Authors:** Qihua Ling<sup>1†</sup>, William Broad<sup>1†‡</sup>, Raphael Trösch<sup>2‡</sup>, Mats Töpel<sup>2‡</sup>, Tijen Demiral Sert<sup>2‡</sup>, Panagiotis Lymperopoulos<sup>2‡</sup>, Amy Baldwin<sup>2‡</sup> & R. Paul Jarvis<sup>1,2\*</sup>

## **Affiliations:**

<sup>1</sup>Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK.

<sup>2</sup>Department of Biology, University of Leicester, Leicester LE1 7RH, UK.

†These authors contributed equally to this work.

‡Present addresses: UK Research and Innovation, 58 Victoria Embankment, London EC4Y 0DS, UK (W.B.); Department of Biology, University of Kaiserslautern, Paul-Ehrlich-Strasse 23, DE-67663 Kaiserslautern, Germany (R.T.); Department of Marine Sciences, University of Gothenburg, Box 460, SE-40530 Göteborg, Sweden (M.T.); Department of Biology, Faculty of Arts and Sciences, Harran University, 63300 Sanliurfa, Turkey (T.D.S.); Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Copenhagen, Denmark (P.L.); School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK (A.B.).

\*Corresponding author. E-mail: paul.jarvis@plants.ox.ac.uk

**Abstract:**

Chloroplasts contain thousands of nucleus-encoded proteins that are imported from the cytosol by translocases in the chloroplast envelope membranes. Proteolytic regulation of the translocases is critically important, but little is known about the underlying mechanisms. Applying forward genetics and proteomics in *Arabidopsis*, we identified two factors required for chloroplast outer envelope membrane (OEM) protein degradation: SP2, an Omp85-type beta-barrel channel of the OEM; and CDC48, a cytosolic AAA+ (ATPases associated with diverse cellular activities) chaperone. We show that both proteins act in the same pathway as the ubiquitin E3 ligase SP1, which regulates OEM translocase components. SP2 and CDC48 co-operate to bring about retrotranslocation of ubiquitinated substrates from the OEM (fulfilling conductance and motor functions, respectively), enabling their degradation by the 26S proteasome in the cytosol. Such chloroplast-associated protein degradation (CHLORAD) is vital for organellar functions and plant development.

**One sentence summary:**

A protein retrotranslocation system of chimeric prokaryotic/eukaryotic origin exists at the chloroplast envelope for proteasomal degradation.

## Introduction:

Chloroplasts are plant organelles responsible for the bulk of terrestrial photosynthetic primary production, and they evolved via endosymbiosis from a cyanobacterial organism more than 1 billion years ago (1). The modern chloroplast proteome comprises ~3000 proteins, most of which are nucleus-encoded and imported post-translationally by multiprotein translocases located in the organelle's outer and inner envelope membranes; these are termed TOC and TIC (translocon at the outer and inner envelope membrane of chloroplasts), respectively (2-5). The biogenesis and operation of chloroplasts requires not only the assembly of the constituent organellar proteins, but also their coordinated homeostasis. Turnover of internal chloroplast proteins is governed by several prokaryotic-type proteases inherited from the endosymbiont (6). In contrast, OEM proteins are degraded by the cytosolic ubiquitin-proteasome system (UPS) via poorly understood mechanisms (1).

The RING (really interesting new gene)-type ubiquitin E3 ligase SP1 is located in the chloroplast OEM where it mediates the ubiquitination of OEM components of the chloroplast protein import machinery (so-called TOC proteins), thereby promoting their degradation by the cytosolic 26S proteasome (7). The TOC components affected by SP1 include the receptors Toc159 and Toc33, and the channel protein Toc75. Such SP1-mediated regulation of the TOC apparatus changes the organellar proteome, which in turn influences the developmental fate and functions of the organelle (e.g., enabling plant adaptation to abiotic stress) (7, 8). While the role of SP1 in marking proteins for degradation is clear, other aspects of this chloroplast protein degradation system have remained obscure. Because TOC proteins are integral membrane components, additional factors are most likely required to overcome the physical and energetic barrier to their extraction from the membrane, prior to degradation in the cytosol, as is the case in other membrane-associated proteolytic systems (9-11).

## Results:

### Identification and phenotypic and molecular analysis of SP2

With the aim of achieving an improved understanding of the SP1-dependent proteolytic pathway, we revisited the forward-genetic screen (7) that originally identified SP1; this was a screen for extragenic suppressors of the *Arabidopsis* Toc33 mutation *plastid protein import1* (*ppi1*), which causes chlorosis owing to defective chloroplast protein import (12). In addition to *sp1*, mutants with lesions at a second, unlinked locus were identified in the screen; these were termed *suppressor of ppi1 locus2* (*sp2*). Double-mutant *sp2 ppi1* plants were larger and greener than the *ppi1* progenitor (Fig. 1, A to C), and exhibited substantial improvements in chloroplast development and chloroplast protein import capacity (Fig. 1, D to F); in all of these respects, the *sp2* mutants were phenotypically very similar to *sp1* mutants (7).

The *SP2* locus (At3g44160; Fig. 1G) was identified using a combination of genetic mapping and whole-genome sequencing of the three independent mutant alleles identified in the screen (fig. S1). The original alleles were phenotypically similar to each other, and to a T-DNA insertion mutant affecting the same gene (Fig. 1, H and I, and fig. S1C); unless specifically stated otherwise, the latter (*sp2-4*; a null allele) was used in all subsequent analyses. The encoded protein is a member of the Omp85 superfamily of beta-barrels, which are involved in protein biogenesis and transport and are widely distributed in the outer membranes of bacteria, mitochondria and chloroplasts (13, 14). The SP2 protein is of unknown function, but it is broadly conserved in the angiosperms (flowering plants) and is closely related to the chloroplast outer membrane protein OEP80 (a protein that has also been termed Toc75-V) (fig. S2) (15). The function of OEP80 is also uncertain (15), although it has been proposed to mediate outer membrane protein biogenesis (16, 17) by analogy with well-characterized homologues in bacteria (BamA, TamA) and mitochondria (Sam50/Tob55) (13, 14).

The SP2 protein is located in the chloroplast OEM (figs. S3 and S4) (15), and it was previously shown to have the capacity to form a membrane channel, like other members of the Omp85 superfamily (18). Unlike OEP80, SP2 lacks an N-terminal POTRA (polypeptide transport associated) domain (such domains typically mediate protein-protein interactions) suggesting that



the two proteins have functionally diverged (Fig. 1G and fig. S2B), a notion that is also supported by phylogenetic analyses of the proteins (fig. S2A) (17). Indeed, OEP80 and SP2 have diametrically opposing effects on TOC protein abundance, as is evident upon comparing published results (showing that OEP80 knockdown depletes TOC proteins (16)) with those discussed below.

### Analysis of the effects of SP2 on TOC proteins and plant development

To shed light on the role of SP2, further genetic analyses were conducted. In addition to *ppi1*, two other TOC mutations (hypomorphic alleles of the genes encoding Toc159 and Toc75) (16, 19) were phenotypically suppressed by *sp2* (fig. S5). In contrast, mutations that cause chlorosis owing to lesions affecting the TIC apparatus of the inner envelope membrane were not suppressed by *sp2* (fig. S6). Together, these data implied a close functional relationship between SP2 and the TOC apparatus; and this notion that was supported by the restored accumulation of Toc75 protein in *sp2 toc* double mutant plants (Fig. 2, A to D, and fig. S7). In all of these respects, the *sp2* mutants were phenotypically very similar to *sp1* mutants (7).

The close functional connection between SP2 and the TOC apparatus was further emphasized by the observation that overexpression of SP2 triggers the specific depletion of TOC proteins (Fig. 2, E and F), resembling closely the effect of SP1 overexpression (7). Moreover, SP2 (like SP1 previously (7)) was shown to interact physically with TOC components in co-immunoprecipitation experiments (fig. S4), suggesting that its effect on TOC accumulation is mediated through direct physical interaction with TOC proteins. The unlikely possibility that SP2 also influences peroxisomal protein import (20, 21) could be ruled out owing to the absence of an effect on the abundance of peroxisomal protein import machinery components (fig. S8).

Further similarities between SP2 and SP1 were observed when the expression profiles of the two genes were compared (fig. S9), and when *sp2* mutant and SP2 overexpressor plants were analysed physiologically in relation to leaf senescence (Fig. 2G) and salt stress tolerance (Fig. 2H and fig. S10). Activity of SP1 promotes both leaf senescence and abiotic stress tolerance (which it does by reconfiguring the chloroplast protein import machinery to produce the necessary organellar

proteome changes) (7, 8), and a very similar pattern of phenotypes was observed here for SP2 (Fig. 2, G and H).

### Identification of CDC48 as a mediator of OEM protein degradation

Identification of a channel-forming component (SP2) that putatively cooperates with SP1 in TOC protein degradation raised a parallel with endoplasmic reticulum (ER)-associated protein degradation (ERAD) (9, 10, 22), where polytopic membrane proteins form a retrotranslocon to enable substrate extraction from the membrane (23-25). In that system, the conserved eukaryotic, multifunctional AAA+ chaperone CDC48 (p97) (9, 10, 26) forms a cytosolic ATP-powered motor to drive such retrotranslocation. We sought novel regulators of the TOC apparatus by employing co-immunoprecipitation followed by mass spectrometry, and identified CDC48 as a minor TOC-associated component (fig. S11). The major *Arabidopsis* isoform, CDC48A (At3g09840) (27) (Fig. 3A), is ubiquitous in the cytosol, but its association with chloroplasts was evident upon analysing lysed, cytosol-free cells containing a CDC48 mutant with stabilized substrate binding (fig. S12).

As CDC48 is an essential component in *Arabidopsis*, we used plants that inducibly overexpress a dominant-negative (DN) form of the protein (Fig. 3A) (27) to assess its function. Expression of CDC48-DN caused chlorosis (Fig. 3B), indicating defective chloroplast biogenesis, as well as a build-up of reactive oxygen species (ROS), paralleling an effect of the *sp1* mutation (8) (fig. S13). In the case of *sp1*, such ROS overaccumulation was attributed to a failure to properly regulate the TOC apparatus and chloroplast protein import, leading to a deregulation of photosynthetic activity (8). Similarly, the effects of CDC48-DN observed here were linked to the specific over-accumulation of TOC proteins (Fig. 3, C and D), revealing phenotypic similarity to *sp2* mutants (Fig. 2, A to D) in addition to *sp1* (7). The SP1 protein, which is subject to UPS-dependent autoregulation (fig. S14) (7), also accumulated in response to CDC48-DN expression (Fig. 3E), indicating that SP1 is degraded via the same processes as TOC proteins.

Because CDC48 is an abundant cellular constituent distributed throughout the nucleocytosolic compartment, we wished to localize the aforementioned functional links to the chloroplast protein import machinery specifically to the OEM in intact cells. Using bimolecular fluorescence

complementation (BiFC; a method that employs reconstitution of YFP (yellow fluorescent protein) fluorescence to report on protein-protein interactions of interest), we demonstrated the physical interaction of CDC48 with Toc159 specifically at the chloroplast envelope (Fig. 3, F and G). In separate experiments, we further revealed that the CDC48-DN-triggered accumulation of OEM proteins (shown in Fig. 3, C and D) occurs specifically at the envelope *in vivo* (fig. S15). Together, these data supported a direct role for the CDC48 chaperone in OEM protein degradation, at the surface of the organelle.

### **Analysis of the functional and physical relationships between SP1, SP2 and CDC48**

Having identified two new components that apparently mediate OEM protein degradation, like SP1, we next addressed whether the three components co-act in a common pathway. We began by considering the relationship between SP1 and SP2. Absence of phenotypic additivity in *sp1 sp2* double mutants (in the *ppi1* background), in relation to plant greening and Toc75 protein accumulation (Fig. 4, A to D), supported the notion that SP1 and SP2 function together. Indeed, SP2 is essential for SP1 action, as the *sp2* mutation abrogated the effect of SP1 overexpression (fig. S16).

Moreover, whereas the overexpression of neither *SP1* nor *SP2* individually affected plant greening in the wild-type background under standard conditions (fig. S10), the simultaneous overexpression of both genes caused strong chlorosis linked to severe depletion of TOC proteins (Fig. 4, E to G), indicating functional interdependency between SP1 and SP2. Such interactions are often observed where the components interact physically, and may arise through mutual stabilization (28, 29). Indeed, SP1 and SP2 co-migrated (with each other and with the TOC apparatus) on native gels, suggesting their coexistence in high molecular weight complexes that either include or exclude the TOC apparatus (fig. S17A). In accordance with this interpretation, SP1 and SP2 interacted specifically in *in vitro* pull-down experiments (fig. S17B).

Next, the functional relationships between CDC48 and the two SP components were assessed. This was done by determining the dependence of *SP* overexpression effects upon CDC48 function. We observed that the ability of overexpression of either *SP1* or *SP2* to trigger TOC protein depletion was blocked when CDC48-DN was co-expressed (Fig. 5). These data indicated

that the functions of the SP components in OEM protein degradation require CDC48, and therefore that all three factors act in the same proteolytic pathway. In accordance with this conclusion, reciprocal co-immunoprecipitation assays demonstrated physical associations between CDC48, SP2, SP1 and TOC proteins (Fig. 6 and fig. S18).

### **Investigating the roles of SP2 and CDC48 in OEM protein retrotranslocation**

In ERAD, CDC48 drives the retrotranslocation of substrate proteins from the ER to the cytosol. To address the hypothesis that CDC48 is similarly involved in the extraction of substrates from the chloroplast OEM, we conducted retrotranslocation assays using two complementary methodologies (23, 30, 31).

First, we employed an *in vivo* assay based on assessment of the distribution of polyubiquitinated substrate between chloroplasts and the cytosol in intact, proteasome-inhibited cells. We observed that extraction of polyubiquitinated Toc33 to the cytosol was inhibited in cells expressing CDC48-DN, providing strong evidence that CDC48 indeed acts in OEM protein retrotranslocation (Fig. 7, A and B). Replication of this assay using *sp2* mutant and SP2 overexpressor cells revealed a similar requirement for SP2 in the extraction of polyubiquitinated Toc33 (Fig. 7, C and D). Moreover, when both of these assays were repeated using SP1 as a second, model substrate (selected owing to its high turnover rate (7)), essentially identical results were obtained (Fig. 8, A to D, and fig. S19).

To provide corroboration of the above-described *in vivo* assay data, we employed an *in vitro* retrotranslocation assay. This experiment tested the ability of CDC48-DN in a cytosol extract to interfere with the removal of polyubiquitinated SP1 from isolated chloroplasts (Fig. 8, E and F). The results closely paralleled those from the *in vivo* assays, as the presence of CDC48-DN was seen to inhibit OEM substrate extraction to the cytosol (in a manner that the control CDC48-WT protein did not). Thus, we concluded that both CDC48 and SP2 act in OEM protein retrotranslocation.

## Discussion:

This work identified SP2 and CDC48 as functional partners of the SP1 E3 ligase, and in so doing defined a multicomponent system for chloroplast protein degradation involving integral OEM and cytosolic factors. Both new components act in substrate protein retrotranslocation, and on the basis of their structural characteristics we conclude that they provide, respectively: a retrotranslocon for substrate conductance during extraction; and, a molecular motor needed to meet the energetic threshold of the process (fig. S20). We designate this system chloroplast-associated protein degradation (CHLORAD), in recognition of similarities with ERAD (including involvement of the UPS, CDC48 and retrotranslocation), but note that the central participation of a beta-barrel channel of prokaryotic ancestry, and a number of other key differences, set it apart as a mechanistically and evolutionarily unique process (9, 25).

A major function of the analogous ERAD system is the elimination of a broad range of misfolded ER proteins (9, 10, 22), which it performs alongside another role in the removal of undamaged proteins for regulatory reasons (32). The available evidence suggests that CHLORAD has a different focus, with its primary purpose being regulatory and centred on the TOC apparatus of the chloroplast protein import machinery. Moreover, while ERAD processes luminal proteins in addition to ER membrane proteins, there is presently no evidence to suggest that CHLORAD similarly acts on internal chloroplast proteins; this particular difference may be linked to the operation of a diversity of proteases of prokaryotic origin in the chloroplast interior (6). An early event in CHLORAD is substrate protein ubiquitination, mediated by SP1, although how substrates are identified and transferred to SP1 is currently unclear. The SP1 intermembrane space domain is involved in substrate binding (7), and the triggering factor appears to be related to developmental cues or environmental stress (7, 8); the stability of SP1 may also be relevant, as the protein is autoubiquitinated (7) and is itself degraded by CHLORAD (Figs. 3 and 8, and fig. S14). Whether CHLORAD also acts in the clearing of damaged proteins, in parallel with its regulatory functions, remains to be determined.

Whereas ERAD employs E3 ligases with multiple transmembrane spans that may simultaneously serve as channels (9), CHLORAD requires a separate channel-forming component. This is

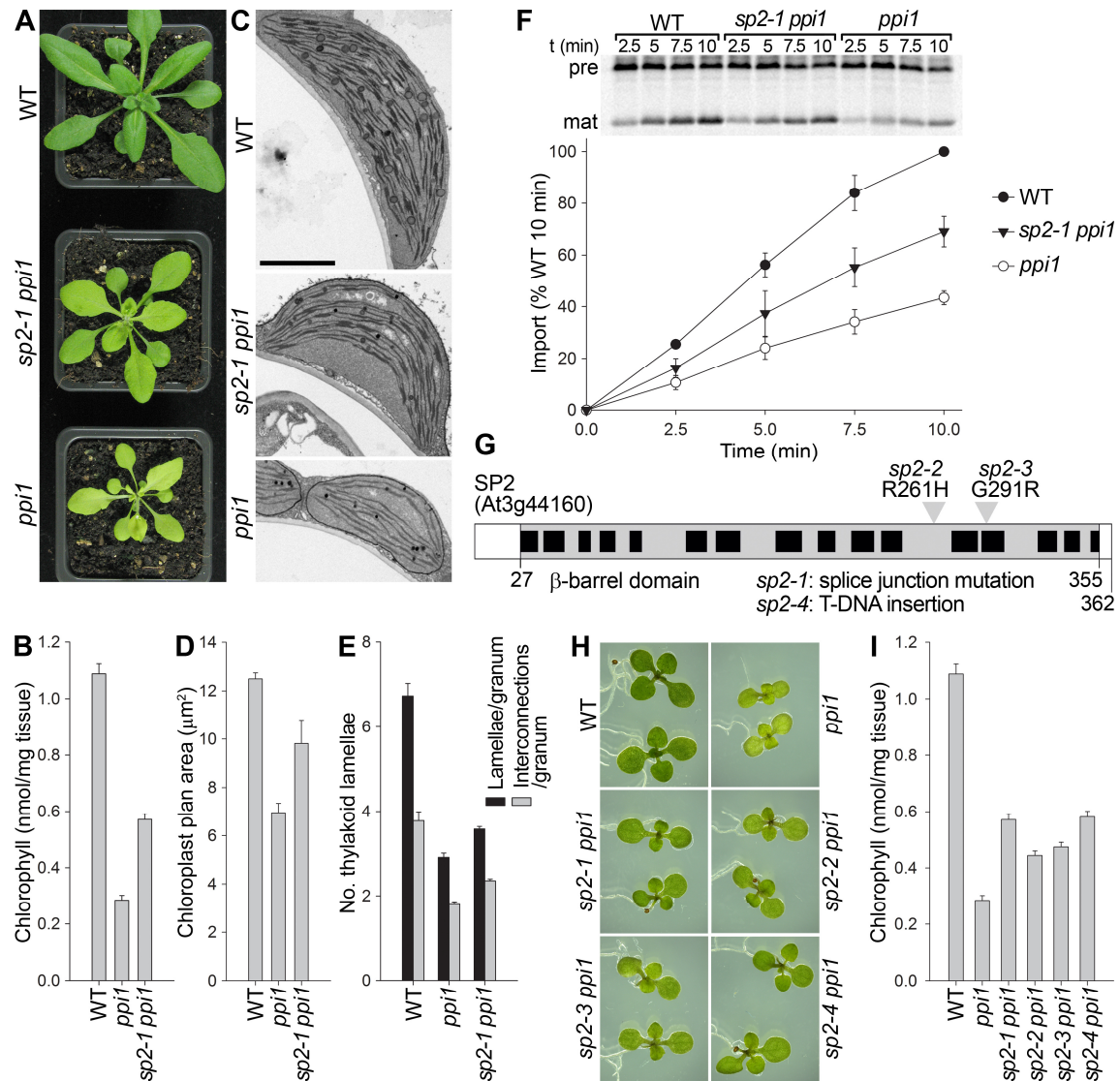
presumably because SP1 does not possess a sufficient number of transmembrane spans to enable channel formation in isolation. Our data indicate that CHLORAD is mediated by an SP1-SP2 complex (Fig. 6 and fig. S17), in which SP2 provides a channel, to enable a close link between ubiquitination and retrotranslocation. It is possible that the operation of the SP2 channel involves lateral opening to enable substrate entry from the membrane (24, 33). Interestingly, unlike any of the proposed ERAD channels (which are all based on alpha-helical transmembrane spans (25)), SP2 belongs to the Omp85 superfamily of beta-barrel proteins originating from prokaryotic cells (13, 14). This indicates that CHLORAD recruited a channel of prokaryotic origin, and that this component underwent neofunctionalization during the evolution that followed endosymbiosis. Consistent with this notion, SP2 is unusual in that it does not possess a partner-protein-interacting POTRA domain, which is a typical feature of Omp85-type proteins in other systems. The data imply that SP2 instead collaborates with SP1 to provide this recruitment function (7), and we hypothesize that the loss of the POTRA domain was a key step in the protein's acquisition of a new function. As has been proposed for ERAD (10), it is possible that more than one channel operates in CHLORAD (e.g., under different circumstances); this might account for the absence of a complete block in retrotranslocation in *sp2* mutants (Figs. 7 and 8).

Once CHORAD has been initiated, the conserved eukaryotic chaperone CDC48 is recruited to the chloroplast from the cytosol. This step may involve binding of CDC48 to the polyubiquitin chain on a substrate protein, with the assistance of unknown cofactors, or a putative chloroplast OEM tethering factor (9, 10, 22). Either way, the chaperone then drives substrate extraction to the cytosol for proteasomal degradation, which it does by overcoming the energetic barrier to the removal of proteins from the membrane (fig. S20). The functional cooperation of CDC48 with SP2 in this retrotranslocation process is remarkable, as it indicates that CHLORAD has a chimeric prokaryotic/eukaryotic origin, much like the chloroplast protein import machinery which also has an Omp85 superfamily protein at its core (i.e., Toc75) (2-5). Quite likely, other as yet unknown factors are involved in the CHLORAD process, and these may exist within the SP1-SP2 core complex or act peripherally in the broader proteolytic pathway.

Previously published data concerning SP1 function (7, 8), together with new results reported here concerning SP2 (Fig. 2, G and H), show clearly that the CHLORAD system is critically

important for plant growth and development, with roles linked to the reconfiguration of the organellar proteome and functions, and to organelle developmental fate. Not only do these observations provide a clear indication of the physiological importance of CHLORAD from a fundamental perspective, but they also suggest novel strategies to improve crop performance (e.g., under environmental stress). As humanity seeks to address the interrelated global challenges of food security and climate change, agriculture will need to carefully consider all such opportunities.





**Fig. 1. The *sp2* mutation suppresses the phenotype of the Toc33 knockout mutation, *ppi1*.**

(A) Visible phenotypes of 30-day-old *sp2-1 ppi1* suppressor mutant and control plants grown on soil.

(B) Chlorophyll contents of 10-day-old *sp2-1 ppi1* suppressor mutant and control seedlings grown in vitro.

(C to E) Transmission electron microscopy analysis of the ultrastructure of cotyledon chloroplasts in 10-day-old *sp2-1 ppi1* suppressor mutant and control plants grown in vitro. Typical organelles are shown (C). Scale bar, 2  $\mu\text{m}$ . These and other micrographs were used to estimate chloroplast cross-sectional area (D) and thylakoid development (E).

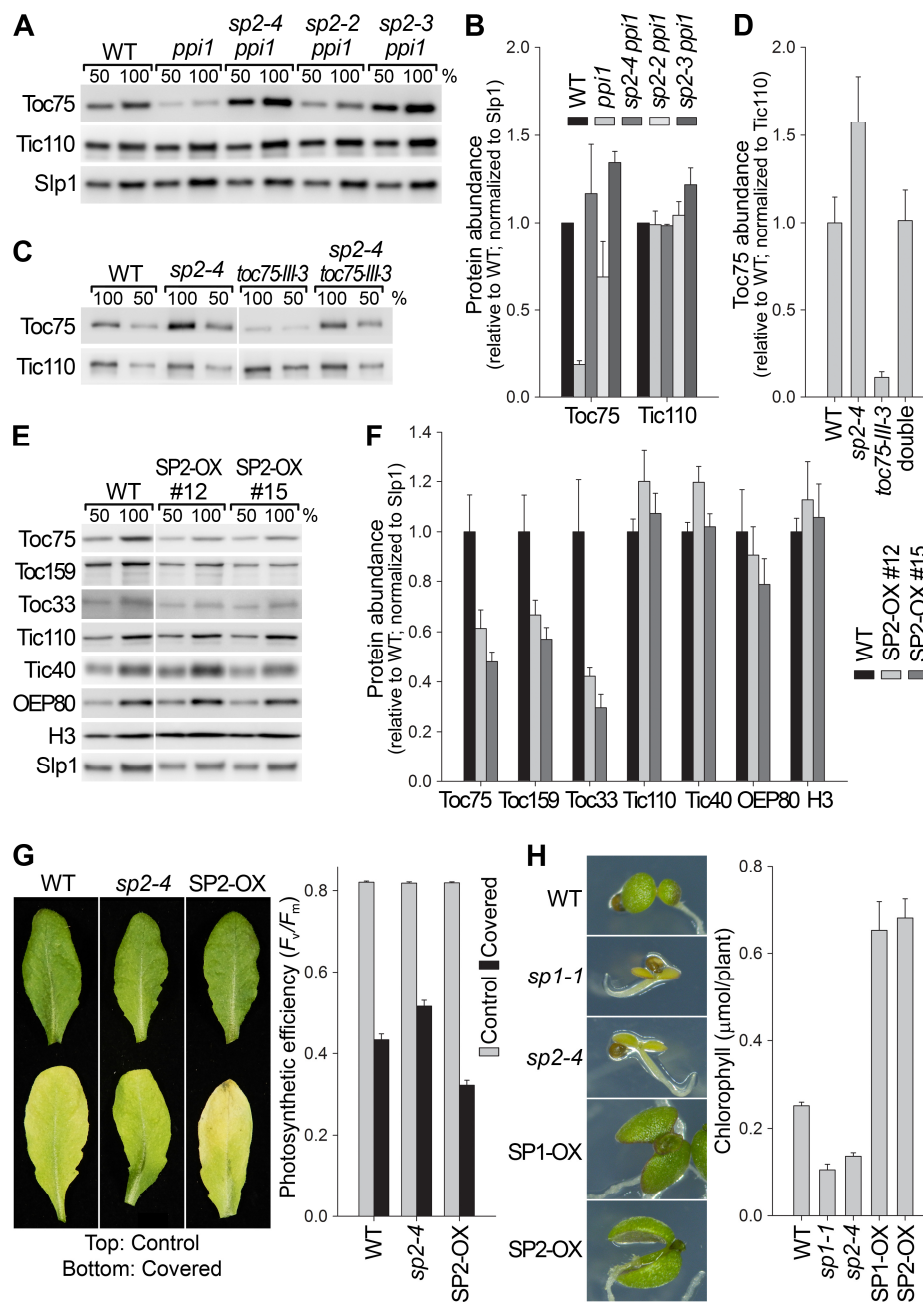
(F) Analysis of protein import into chloroplasts isolated from *sp2-1 ppi1* suppressor mutant and control plants, using <sup>35</sup>S-labelled Rubisco small subunit precursor protein as the import substrate, and corresponding quantification of the maturation (mat) of the radiolabelled precursor protein (pre). A



representative phosphor screen image is shown (top); times indicate minutes (min) after the start of each import reaction. Together with similar images from two additional experiments, this was used to conduct the quantitative analysis shown (bottom); the amount of imported, mature protein in each sample was expressed as a percentage of that present at the final time-point for wild type.

(G) Domain map of the SP2 protein. Grey box,  $\beta$ -barrel domain; black boxes, predicted transmembrane spans. The sites of amino acid substitutions in two *sp2* mutant alleles are indicated with grey triangles.

(H and I) Independent mutant alleles of *sp2* suppress *ppil* in similar fashion to *sp2-1*. The *sp2-2 ppil* and *sp2-3 ppil* mutants were identified in the EMS mutagenesis screen, whereas the *sp2-4* T-DNA insertion mutant was obtained from a stock centre and crossed to *ppil*. The plants were grown in vitro for 10 days before photography (H) and chlorophyll content analysis (I). The chlorophyll values for *sp2-2* and *sp2-3*, which cause missense mutations, were statistically significantly different from that for *sp2-1* (Student's t-test,  $p \leq 0.01$ ), suggesting that they are weak alleles. All values are means  $\pm$  SEM ( $n \geq 3$  experiments or samples).



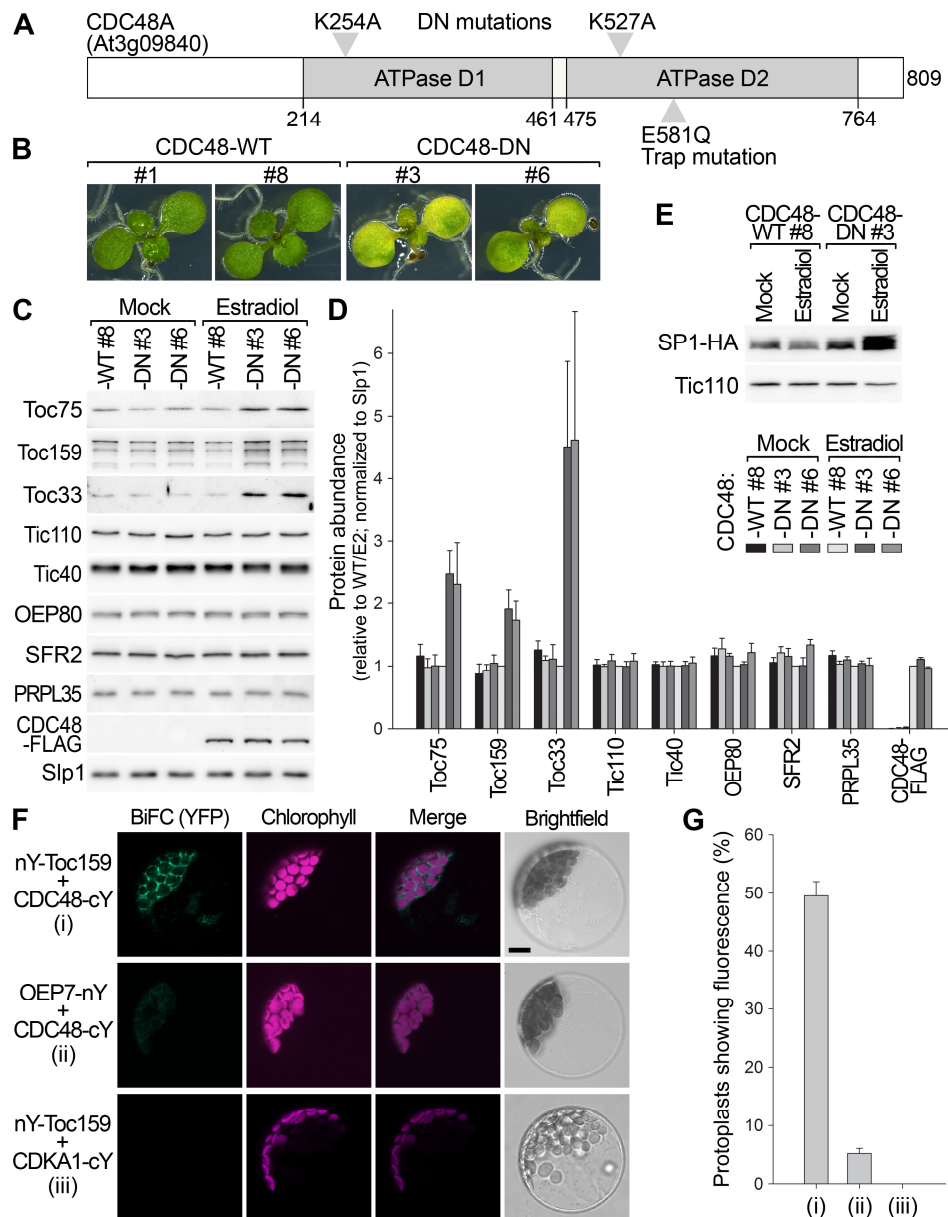
**Fig. 2. Functional analysis of SP2 reveals roles in chloroplast proteostasis and development.**

(A to F) Immunoblot analyses of total protein extracts (two loading amounts per sample) from the indicated genotypes, including *sp2 ppi1* suppressor mutants (A and B), *sp2* and *toc75-III-3* single and double mutants (C and D), and two different SP2 overexpressor (OX) lines (E and F). Slp1, mitochondrial stomatin-like protein1 (a non-chloroplastic membrane protein that served as loading control); H3, nuclear histone H3.

(G) Leaf senescence analysis of the indicated genotypes employing mature rosette leaves induced to senesce by covering with aluminium foil. Typical control (uncovered) and senescent (covered) leaves are

shown (left). Photochemical efficiency of photosystem II ( $F_v/F_m$ ) was measured to estimate the extent of senescence (right); the covered values for *sp2-4* and SP2-OX were statistically significantly different from that for wild type (Student's t-test,  $p \leq 0.0003$ ).

(H) Abiotic stress tolerance analysis of the indicated genotypes employing 14-day-old plants grown in vitro on NaCl medium. Typical plants (left) and chlorophyll contents (right) are shown. The chlorophyll values for the mutant and overexpressor plants were statistically significantly different from that for wild type (Student's t-test,  $p < 0.004$ ). All values are means  $\pm$  SEM ( $n \geq 3$  experiments or samples).



**Fig. 3. Functional analysis of CDC48 reveals roles in chloroplast proteostasis and development.**

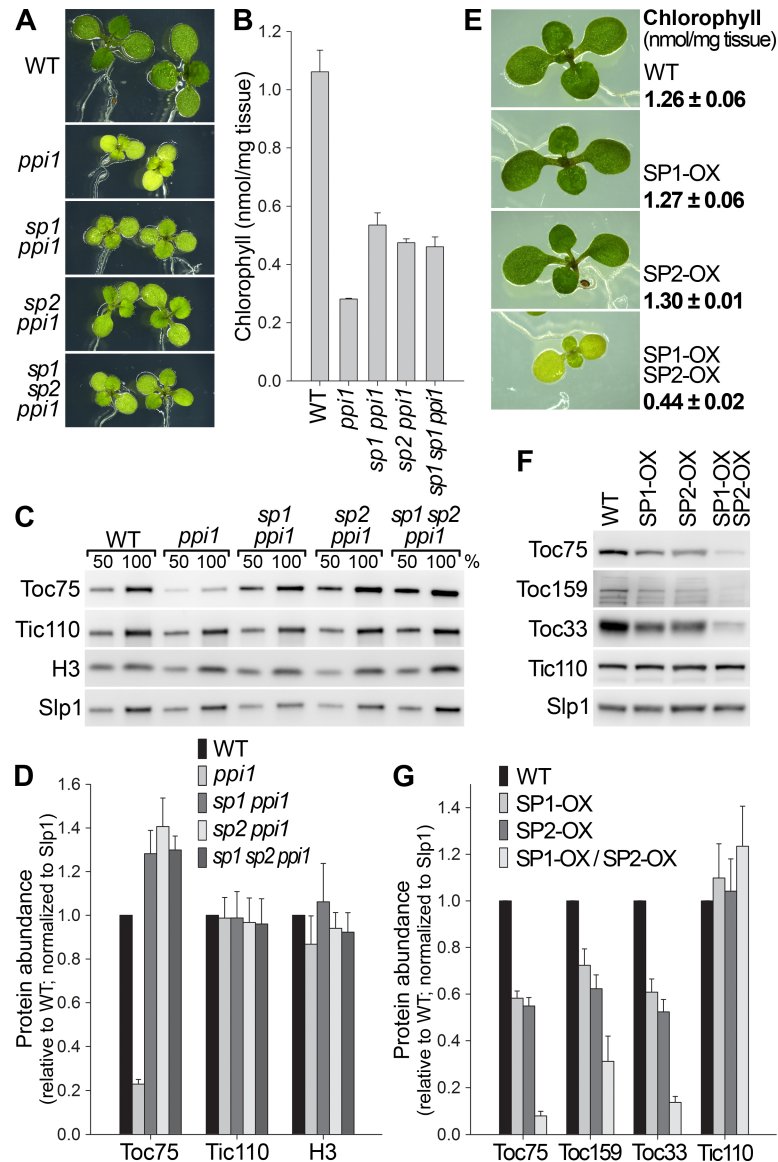
(A) Domain map of the CDC48A protein. Grey boxes, ATPase domains (D1, D2). The sites of the dominant-negative (DN) and Trap mutations are indicated with grey triangles.

(B) Visible phenotypes of 9-day-old plants expressing the CDC48-DN protein or an equivalent non-mutated control protein (CDC48-WT), both induced with estradiol for 2 days. Two independent transgenic lines are shown for each construct.

(C to E) Immunoblot analyses of total leaf (C and D) and protoplast (E) protein extracts from the CDC48-WT and CDC48-DN transgenic plants, following induction with estradiol (E2) or mock control treatment lacking the inducer. In C and D, to induce TOC component degradation (8), plants were simultaneously

treated with 200 mM mannitol for 2 days, before protein extraction, to trigger stress-dependent protein degradation (8); this was essential to observe the effects of CDC48-DN in this assay.

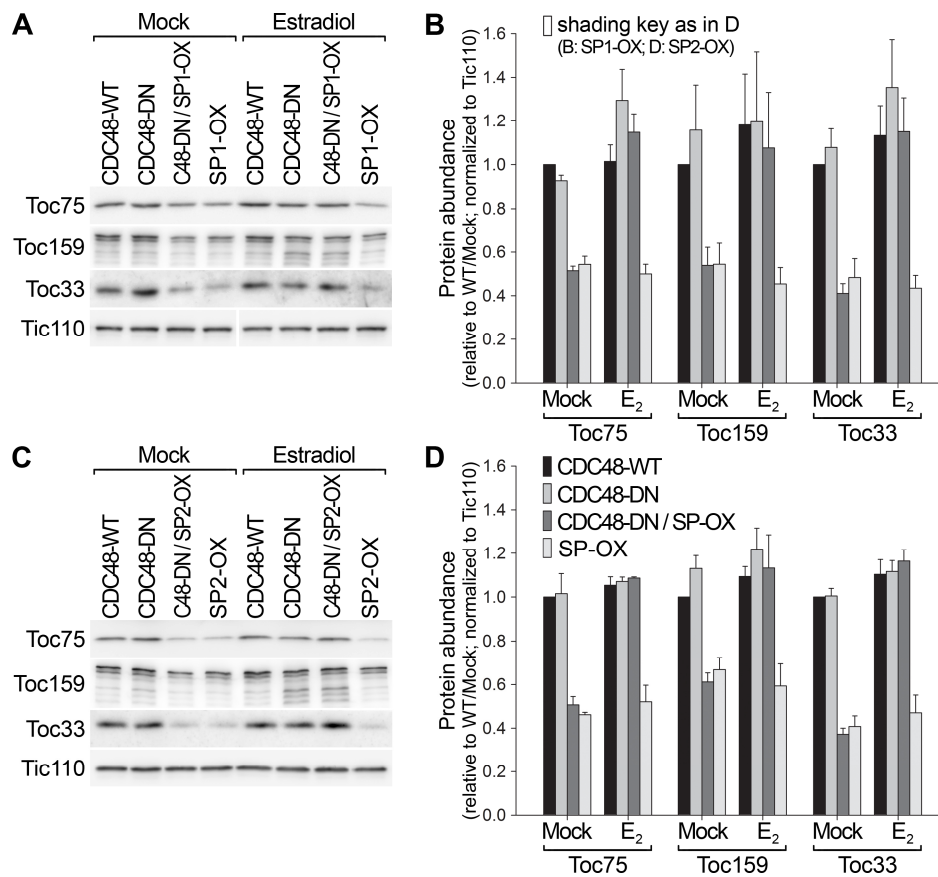
(F and G) BiFC analysis of the CDC48-TOC interaction. Reconstitution of YFP fluorescence was assessed following transient co-expression of the indicated pairs of fusion proteins, which carry complementary N- or C-terminal YFP fragments (nY and cY, respectively). OEP7 and CDKA1 acted as OEM and cytosolic controls, respectively. Representative images are shown (F). Scale bar, 10  $\mu$ m. In addition, the frequency of protoplasts showing a BiFC signal was quantified (G). All values are means  $\pm$  SEM (n = 3 experiments).



**Fig. 4. SP2 acts in the same pathway of TOC protein degradation as SP1.**

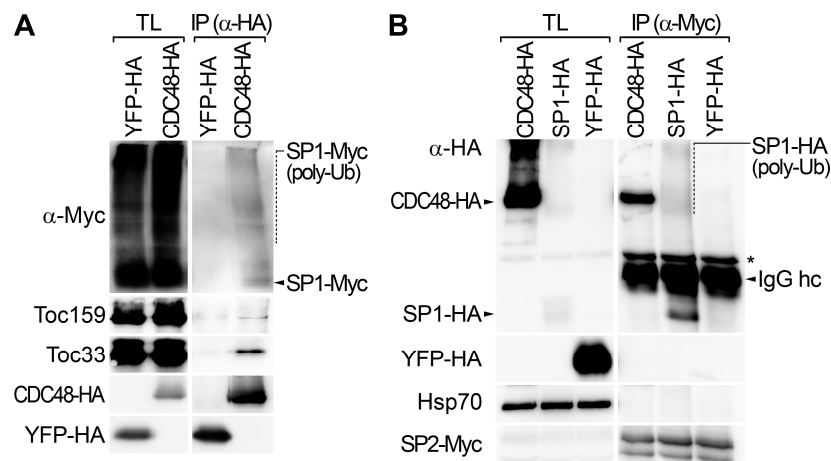
(A to D) Analysis of an *sp1 sp2 ppi1* triple mutant. Triple mutant plants were compared with both *sp ppi1* single mutants in relation to the extent of suppression of *ppi1*. No phenotypic additivity in the triple mutants was apparent upon analysing visible phenotypes (A), chlorophyll contents (B), or the abundance of Toc75 protein by immunoblotting (C and D); two loading amounts per sample were analysed.

(E to G) Analysis of plants simultaneously overexpressing (OX) both SP1 and SP2. Double OX plants were compared with both single SP-OX genotypes, revealing a synergistic interaction in relation to visible phenotype (E, left), chlorophyll content (E, right), and TOC protein depletion as analysed by immunoblotting (F and G).



**Fig. 5. CDC48 acts in the same pathway of TOC protein degradation as SP1 and SP2.**

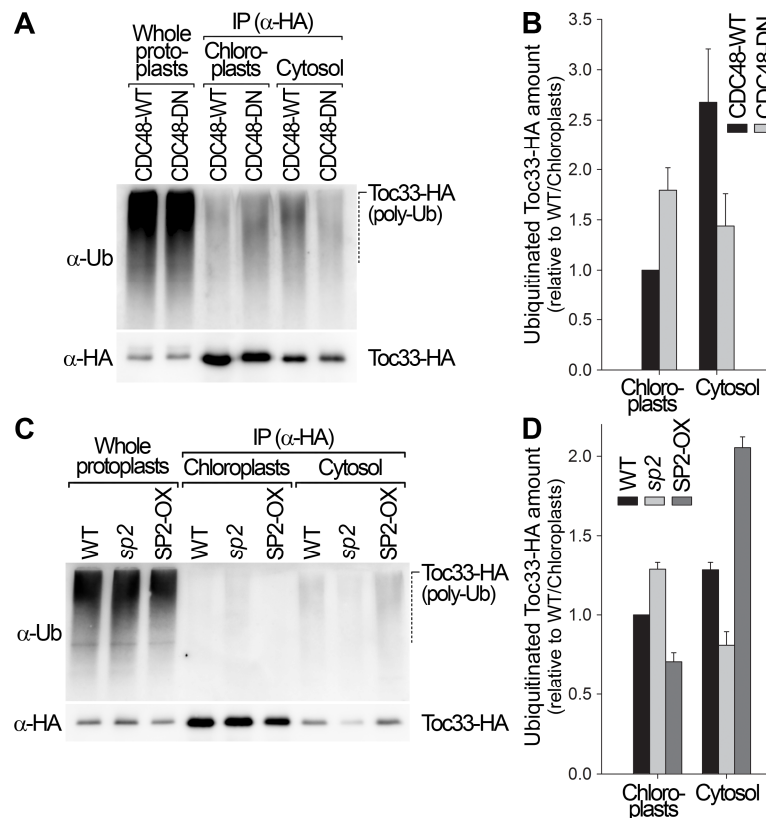
Analysis of the effects of CDC48-DN on the ability of SP1/2 overexpression to trigger TOC protein depletion. Transgenic plants carrying the following constructs were analysed by immunoblotting: CDC48-WT, CDC48-DN, SP1-OX, and the latter two in combination (A and B); and, CDC48-WT, CDC48-DN, SP2-OX, and the latter two in combination (C and D). In each case, the analysis was done following treatment with (or without; Mock) estradiol to induce expression of the CDC48 constructs. All values are means  $\pm$  SEM ( $n = 3-4$  experiments).



**Fig. 6. CDC48 and SP2 interact physically with each other and with OEM proteins.**

Co-immunoprecipitation (IP) of SP1, TOC components, and/or CDC48 with: haemagglutinin (HA)-tagged wild-type CDC48 (A) or Myc-tagged SP2 (B) from protoplast extracts. Cells were transfected with the following constructs: *CDC48-WT-HA* or *YFP-HA* plus *SP1-Myc* (A); or *SP2-Myc* plus *CDC48-WT-HA*, *SP1-HA*, or *YFP-HA* (B). In both cases, YFP-HA acted as a negative control. TL, total lysate; poly-Ub, poly-ubiquitinated form; hc, heavy chain. The asterisk indicates a non-specific cross-reacting band.

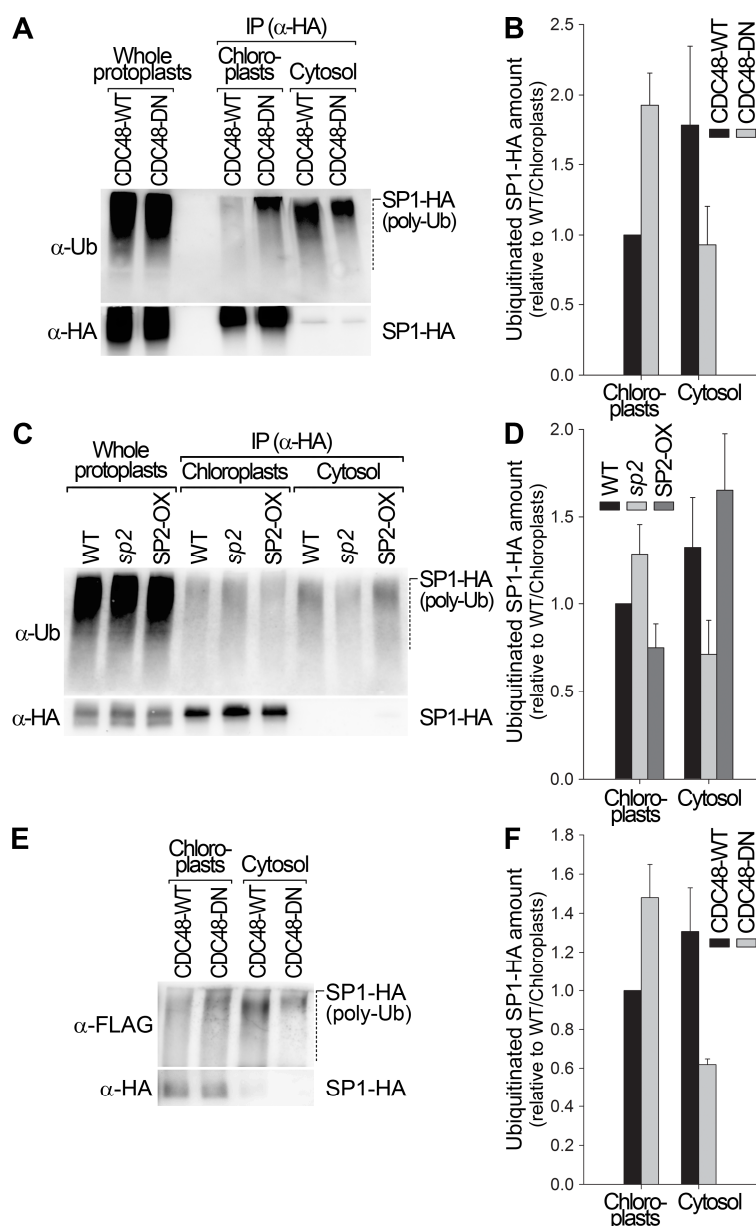




**Fig. 7. CDC48 and SP2 are required for the retrotranslocation of polyubiquitinated Toc33.**

(A and B) Analysis of the role of CDC48 in the retrotranslocation of Toc33-HA substrate using an in vivo retrotranslocation assay. Protoplasts isolated from CDC48-WT and CDC48-DN transgenic plants that were transiently expressing Toc33-HA were treated, following estradiol induction, with 5  $\mu$ M bortezomib proteasome inhibitor, and then separated into chloroplast and cytosol fractions. In this assay, retrotranslocation occurred in intact cells, and the retrotranslocated Toc33-HA was protected by Bortezomib inhibition, which initiated the experiment. Following fractionation, Toc33-HA was immunoprecipitated from both fractions, and detected by immunoblotting using  $\alpha$ -HA and  $\alpha$ -ubiquitin (Ub) (A). Retrotranslocation efficiency was assessed by quantifying the relative amounts of ubiquitinated Toc33 in the chloroplasts and cytosol (B).

(C and D) Analysis of the role of SP2 in the retrotranslocation of Toc33-HA substrate using an in vivo retrotranslocation assay performed as described in A and B using protoplasts from wild-type, *sp2* mutant, and SP2-OX plants (and without the need for estradiol induction). Typical immunoblotting results are shown (C), along with quantification (D). All values are means  $\pm$  SEM (n = 3 experiments).



**Fig. 8. CDC48 and SP2 are required for the retrotranslocation of polyubiquitinated SP1.**

(A to D) Analysis of the roles of CDC48 and SP2 in the retrotranslocation of SP1-HA substrate using an in vivo retrotranslocation assay. Protoplasts isolated from CDC48-WT and CDC48-DN transgenic plants (A and B), or from wild-type, *sp2* mutant, and SP2-OX plants (C and D), that were transiently expressing SP1-HA were treated (following estradiol induction in the case of A and B) with 5  $\mu$ M bortezomib proteasome inhibitor, and then separated into chloroplast and cytosol fractions. In this assay, retrotranslocation occurred in intact cells (as in Fig. 7), and the retrotranslocated SP1-HA was protected by Bortezomib inhibition, which initiated the experiment. Following fractionation, SP1-HA was immunoprecipitated from both fractions, and detected by immunoblotting using  $\alpha$ -HA and  $\alpha$ -ubiquitin

(Ub) (A and C). Retrotranslocation efficiency was assessed by quantifying the relative amounts of ubiquitinated SP1 in the chloroplasts and cytosol (B and D).

(E and F) Analysis of the role of CDC48 in the retrotranslocation of SP1-HA substrate using an in vitro retrotranslocation assay. This assay (in contrast with the in vivo assay in A to D) employed a cell-free reaction. Chloroplasts were prepared from estradiol-induced CDC48-DN protoplasts transiently expressing SP1-HA and FLAG-tagged ubiquitin. Cytosol was prepared from induced CDC48-WT or CDC48-DN protoplasts (that were not expressing SP1-HA or FLAG-tagged ubiquitin). Reaction mixtures composed of chloroplasts and either cytosol fraction (as indicated above the gel images) were incubated for 1 h, and then re-fractionated into chloroplast (left) and cytosol (right) samples; any SP1-HA detected in the cytosol at the end of the reaction must have arisen from the chloroplasts. SP1-HA was enriched from 90% of the protein extract of each fraction by IP, and detected by immunoblotting using anti-HA (to detect unmodified SP1-HA) and anti-FLAG (to detect ubiquitinated SP1-HA) (E), and the results were quantified (F). All values are means  $\pm$  SEM (n = 3-4 experiments).

## Materials and Methods:

### *Plant material and growth conditions*

All *Arabidopsis thaliana* plants were of the Columbia-0 (Col-0) ecotype, except the *ppi1* line used for the genetic mapping of *sp2* which was introgressed into Landsberg *erecta* (Ler) through seven outcrosses. The *sp1-1*, *sp1-3*, *ppi1*, *tic40-4*, *hsp93-V-1*, *ppi2-3* (*fts1*), *toc75-III-3* (*mar1*) mutants, and the 35S promoter-driven SP1 overexpressor (SP1-OX) transgenic line, have all been described previously (7, 12, 16, 19, 34, 35). The *sp2-4* (SALK\_137135) mutant was obtained from the Salk Institute Genomic Analysis Laboratory, and confirmed by PCR and RT-PCR analysis, as described previously (36); this mutant was phenotypically similar to the three chemically-induced *sp2* alleles identified in this study, and unlike a previously-described T-DNA mutant (37). When making double and triple mutants, the *sp1-3* and *sp2-4* alleles were employed (because their mutations are easier to detect by PCR); double mutants were selected and verified by phenotype analysis and by using PCR-based genotyping (table S1). Unless specifically stated otherwise, the *sp2-4* allele was employed in subsequent experiments. For consistency with previous work (7, 8), the *sp1-1* allele was employed in physiological experiments.

For in vitro growth, seeds were surface sterilized, sown on Murashige-Skoog (MS) agar medium in petri plates, cold-treated at 4°C, and thereafter kept in a growth chamber, as described previously (38). All plants were grown under a long-day cycle (16 h light, 8 h dark). For induction of CDC48-WT or CDC48-DN expression in the corresponding transgenic lines, 7-day-old plants were transferred onto MS agar medium supplemented with 4 µM estradiol (Sigma).

### *Physiological studies*

Chlorophyll measurements were performed by using a Konica-Minolta SPAD-502 meter (39), when analysing rosette-stage plants, or by photometric quantification following extraction in N,N'-dimethylformamide (DMF) as described previously (40), when analysing seedlings.

Dark treatments for the induction of senescence were conducted as previously described (7, 41). Developmentally-equivalent leaves of 28-day-old plants were wrapped in aluminium foil whilst still attached to the plant, and then left under standard growth conditions for 5 days. Photochemical efficiency of photosystem II ( $F_v/F_m$ ) was determined by measuring chlorophyll fluorescence using a CF Imager (Technologica, UK) as described previously (42). Five experiments were performed, and approximately five leaves (each one from a different plant) were analysed per genotype in each experiment.

Salt stress experiments were conducted as described previously with minor modifications (8). All seeds of the different genotypes used in this work were harvested at the same time. Seeds were germinated directly on MS agar medium (supplemented with 1% sucrose) containing 150-170 mM NaCl.

Stress tolerance was assessed by measuring chlorophyll accumulation after 14 days. Three experiments were performed, and ~25 seedlings per genotype were analysed in each experiment.

Detection of hydrogen peroxide was performed by staining with 3,3'-diaminobenzidine (DAB) (Sigma) as previously described (43). Seven-day-old plants grown on MS agar medium were transferred onto similar medium containing 4  $\mu$ M estradiol (for induction of CDC48-WT/-DN expression). The plants were left to grow for further 2 days before initiating DAB staining. Each experiment used approximately 5 seedlings per genotype. Three experiments were performed with the same result, and typical images are presented. The area of staining was quantified using ImageJ as described previously (8).

### ***Identification of the *sp2* mutants and genetic mapping***

The original *sp2* mutants (*sp2-1* [*sp2-310*], *sp2-2* [*sp2-416*] and *sp2-3* [*sp2-555*]) were identified by screening the M<sub>2</sub> progeny of 7,000 M<sub>1</sub> *ppi1* seeds that had been treated with 100 mM ethyl methanesulfonate for 3 h using a published procedure (7, 44). Allelism test crosses confirmed that all of these mutations are allelic. All three *sp2 ppi1* mutants were backcrossed to *ppi1* three times before phenotypic analysis. Initial mapping of *sp2* was conducted by analysing the greenest plants in F<sub>2</sub> populations from crosses between *sp2-1 ppi1* (Col-0) and *ppi1* introgressed into the *Ler* ecotype, using PCR markers that detect Col-0/*Ler* polymorphisms. In a mapping population of 190 such F<sub>2</sub> plants, six were heterozygous for the marker F21A17 at position 12285000 on the upper arm of chromosome 3, but homozygous for Col-0 downstream of that, suggesting that the suppressor mutation was in the downstream Col-0 region; F<sub>3</sub> seedlings from these six plants were grown and verified visually to be non-segregating, as expected for a homozygous *sp2-1 ppi1* double mutant. In a second mapping population of 192 plants, the *sp2* mutation was further mapped to the south of a more southerly marker, MJ16-2 at position 12597802 on the upper arm of chromosome 3. However, it was not possible to determine the position of the *sp2* locus precisely owing to the persistence of an “island” of Col-0 DNA in the *Ler*-introgressed *ppi1* line, near the *sp2* locus (around the chromosome 3 centromere). Thus, final identification of the gene was achieved by whole-genome sequencing.

### ***Whole-genome sequencing and assembly***

Approximately 100 mg of plant inflorescence tissue from each of the original *sp2* alleles (*sp2-1 ppi1*, *sp2-2 ppi1*, and *sp2-3 ppi1*), and from *ppi1-1*, was harvested and flash-frozen in liquid nitrogen. Total genomic DNA was then extracted using an E.Z.N.A. Plant DNA Kit (Omega Bio-tek) following the manufacturer's guidelines. The DNA samples were quantified by comparison with standards.

Library preparation and sequencing were conducted at the Earlham Institute (Norwich, UK). Approximately 1-5 µg genomic DNA per sample at a minimum concentration of 20 ng/µl was used in sequencing library preparation. Individual barcoded Illumina TruSeq DNA libraries were generated for each genotype. The four samples were then sequenced on one lane of Illumina HiSeq 2000, which generated between 32.2-42.3 million 100 bp paired-end reads for each sample. The first five bases of the 5' ends of the reads were removed using `fastx_trimmer v.0.0.14` ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)), and any bases with a Phred quality score below 15 were removed from the 3' end using `cutadapt (v.1.3)` (45). Illumina TruSeq adaptors were also removed using `cutadapt`, where a minimum overlap of 10 bases with the adaptors, a maximum error rate of 0.1, and a minimum final read length of 50 bases were set. Finally, `fastq_quality_filter v0.0.14` ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)) was used to remove sequences with a Phred score below 20 in more than 5% of the bases. Read pairs were identified using `pairSeq.py` (<https://github.com/topel-research-group/pairSeq>).

The reads from the *ppi1* single mutant were mapped to the TAIR10 *Arabidopsis thaliana* reference genome ([ftp://ftp.jgi-psf.org/pub/compugen/phytozome/v9.0/Athaliana/assembly/Athaliana\\_167.fa.gz](ftp://ftp.jgi-psf.org/pub/compugen/phytozome/v9.0/Athaliana/assembly/Athaliana_167.fa.gz)) of the Phytozome v.9.0 release (<http://www.phytozome.net/>) using `clc_mapper v. 4.0.13.86165` (<https://www.qiagenbioinformatics.com/>). A consensus sequence in FASTA format was then generated using `clc_find_variations v. 4.0.13.86165`. Also, the transcript sequences from the TAIR10 release were aligned to the reference genome in order to facilitate manual examination of identified mutations and visualization of whether a particular mutation occurs in an exon, intron, etc.

The three datasets from the individual *sp2* double mutants were independently aligned to the *ppi1* reference genome using `clc_mapper v. 4.0.13.86165`. The *ppi1* reads were also aligned to the same reference in order to identify any variable sites resulting from allelic variation in the *ppi1* line.

### ***In silico identification of mutations***

The mutagen ethyl methanesulfonate (EMS) used to generate the three *sp2* mutants reacts with guanine in the DNA molecule and is likely to (1.) cause point mutations that change guanine to adenosine (or cytosine to thymidine on the reverse strand). The respective mutations affecting the three *sp2* mutants were furthermore expected to (2.) occur in the same gene (or corresponding promotor region) in (3.) all three mutants but (4.) not necessarily in the exact same position. These four search criteria were implemented in the program “`find_sp2.py`” (<https://github.com/topel-research-group/sp2>) which takes as input a gff3 file with gene coordinates and the SNP variant output from `clc_mapper`, and outputs a list of names of mutated genes from each dataset. Genes found to be mutated in all the *sp2* datasets were then manually examined by visualizing the alignment data using the genome viewer IGV (v.2.3) (46).

This analysis showed that each *sp2* allele contains a G-to-A point mutation within the At3g44160 gene, just to the south of the chromosome 3 centromere. In *sp2-1*, a mutation was detected at the splice junction preceding the final exon; this was later shown to cause mis-splicing, frame-shifts, and premature termination, implying that *sp2-1* is a knockout allele. In *sp2-2* and *sp2-3*, the detected mutations were both predicted to cause an amino-acid substitution. Further details are provided in the fig. S1 legend. The transmembrane beta-strands and the three-dimensional structure of the SP2 protein were predicted using Phyre2, using the Intensive modelling mode, which produced a model based on five structures for bacterial TamA and BamA proteins (c4c00a, c5ekqA, c4k3bA, c4n75A, c4k3cA) to increase confidence (47).

### ***Phylogenetic analysis***

Sequences were obtained by BLAST searches of the Phytozome 12 database (48) (table S2). Sequences were aligned by multiple alignment using fast Fourier transform (MAFFT) (49), and manual alignment adjustments were made using Mesquite 1.12 (Tangient). Phylogeny was inferred using MrBayes 3.2 software (50). Two runs were performed in parallel, with each using 8 MCMC chains for 8 million generations and the temperature set to 0.2. The standard deviation of split frequencies (StdDev) was 0.001228 at the end of the analysis and therefore assumed to have converged. Trees were sampled every 1000 generations, reaching a total of 8000 trees. Burn-in was set to 25%, and so the first 2000 trees were discarded. The resulting phylogeny was a minimum 50% consensus of the remaining 6000 sampled trees. Parameters not mentioned were retained at the default setting.

### ***Gene identifiers***

The following gene sequences from *Arabidopsis thaliana* were employed experimentally in this study: SP2 (At3g44160); AtCDC48A (At3g09840); Toc33 (At1g02280); CDKA1 (At3g48750); Toc159 (At4g02510); OEP7 (At3g52420); OEP80 (At5g19620); SSU (At1g67090).

### ***Plasmid constructs***

All primers used are listed in table S1. The *SPI-HA*, *YFP-HA*, *FLAG-tagged ubiquitin*, *SPI-YFP*, *GST-SPIflex* and *YFP-Toc33* constructs have all been described previously (7, 21). The coding sequence (CDS) for the dominant-negative (DN) CDC48 mutant (AtCDC48A<sup>K254A,K527A</sup>) was amplified from a pre-existing plasmid (H6T7-DN-B, which employs an ethanol-inducible promoter) (27) using primers that add a C-terminal FLAG tag. All other *Arabidopsis* CDSs (including that for wild-type AtCDC48A, amplified with and without a FLAG tag) were PCR-amplified from Col-0 cDNA; and the CDS encoding the CDC48 “Trap” mutant (AtCDC48A<sup>E581Q</sup>) was generated by overlap-extension PCR (51). The Gateway



cloning system (Invitrogen) was used to make most of the constructs, and all entry clones were verified by DNA sequencing. To generate C-terminal 6×Myc tag fusion proteins, the *SP1* and *SP2* CDSs were cloned into the pE3c vector (52), and then subcloned into the p2GW7 35S-driven expression vector (53) for protoplast transfection (generating the *SP1-Myc* and *SP2-Myc<sup>a</sup>* constructs). The *SP2* CDS, with and without the Myc tag, was cloned into the pB2GW7 binary 35S-driven overexpression vector (53) for stable plant transformation (generating the *SP2-OX* and *SP2-Myc* constructs). The C-terminally FLAG-tagged wild-type and DN mutant CDC48 CDSs were cloned into the pMDC7 binary vector (54) for stable plant transformation, enabling estradiol-inducible transgene expression (providing more stable expression than the original ethanol-inducible system (27)) (generating the *CDC48-WT* and *CDC48-DN* constructs). Untagged wild-type and “Trap” mutant CDC48 CDSs were cloned into both: a modified p2GW7 plant expression vector providing a C-terminal haemagglutinin (HA) tag (generating the *CDC48-WT-HA* and *CDC48-Trap-HA* constructs); and the p2GWC7 (for wild type) or p2GWY7 (for “Trap”) plant expression vectors (53), which respectively provide a C-terminal CFP or YFP tag (generating the *CDC48-WT-CFP* and *CDC48-Trap-YFP* constructs). To generate C-terminally HA-tagged Toc33, the corresponding CDS was similarly cloned into the modified p2GW7 vector (generating the *Toc33-HA* construct). To generate N-terminally TAP-tagged Toc33, the corresponding CDS was cloned into the NTAPi binary vector (55) (generating the *TAP-Toc33* construct). To generate bimolecular fluorescence complementation (BiFC) constructs, selected gene sequences were cloned into pBlueScript II SK- using a single SmaI restriction site, sequenced, and then subcloned into: 5'-*KpnI* and 3'-*XmaI* sites of the pSAT4A-cEYFP-N1 vector (56) for *CDC48-cYFP*; 5'-*XhoI* and 3'-*EcoRI* sites of the pSAT4A-cEYFP-N1 vector (56) for *CDK1-cYFP*; 5'-*EcoRI* and 3'-*Sall* sites of the pSAT4-nEYFP-C1 vector (56) for *nYFP-Toc159*; and, 5'-*EcoRI* and 3'-*Sall* sites of the pSAT4A-nEYFP-N1 vector (56) for *OEP7-nYFP*.

### ***Transient assays and stable plant transformation***

Protoplast isolation and transient assays were carried out as described previously (7, 57). When required, MG132 (Sigma), epoxomicin (Merck), bortezomib (Selleckchem) (all three chemicals prepared as a 10 mM stock solution in DMSO), or E-64 (Melford) (prepared as a 10 mM stock solution in water) was added to the protoplast culture medium at 15 h following transfection, to a final concentration of 1-30 μM, 1-10 μM, 5 μM, or 1-10 μM, respectively; subsequently, the culture was incubated for a further 2-3 h before analysis. When using protoplasts isolated from the CDC48-WT and CDC48-DN transgenic lines, 10 μM estradiol (prepared as 10 mM stock solution in ethanol) was included in the culture medium throughout the incubation of protoplasts (either for 15 h in the case of transfected protoplasts, or for 2 d when transfection was not needed). For XFP fluorescence and immunoprecipitation assays, 0.1 ml (10<sup>5</sup>)



or 1 ml ( $10^6$ ) aliquots of protoplasts were transfected with 5  $\mu$ g or 100  $\mu$ g of DNA, respectively, and the fluorescence signals were analysed after 15-18 h.

Transgenic lines carrying the *SP2-OX*, *SP2-Myc*, *CDC48-WT* and *CDC48-DN* constructs were generated by *Agrobacterium*-mediated transformation (16, 36). Transformants were selected using MS medium containing either phosphinothricin (for the SP2 constructs) or hygromycin B (for the CDC48 constructs). At least 12 T<sub>2</sub> lines for each transformation were analysed, and at least two lines with a single T-DNA insertion (which showed a 3:1 segregation on selective MS medium in the T<sub>2</sub> generation) were chosen for further analysis.

### Microscopy

Transmission electron microscopy was performed as described previously (16). Measurements were recorded using at least 30 different plastids per genotype, and were representative of three individuals per genotype. Chloroplast cross-sectional area was estimated as described previously (16, 34), using the equation:  $\pi \times 0.25 \times \text{length} \times \text{width}$ . Numbers of thylakoid lamellae per granal stack, and of interconnections between granal stacks, were counted as previously described (7, 16) in at least 96 resolvable grana across three individuals per genotype.

All fluorescence microscopy and BiFC experiments were conducted at least twice with the same results, and typical images are presented. For the imaging of CFP, YFP and chlorophyll fluorescence signals, in most cases (except fig. S14) protoplasts were examined using a Zeiss LSM 510 META laser-scanning confocal microscope (Carl Zeiss Ltd.), as described previously (8). To visualize signals associated with chloroplasts without interference from cytosolic signals, protoplasts were ruptured by gently tapping the cover glass; this enabled the release of the cytosol and of intact chloroplasts. For fig. S14, fluorescence images were captured using a Nikon Eclipse TE-2000E inverted microscope as described previously (36).

For BiFC assays, plasmid DNA for two constructs (one cYFP fusion and one nYFP fusion) was co-transfected into wild-type protoplasts, and then YFP signals were analysed by confocal imaging. All images were captured using the same settings to enable comparisons. The frequency of protoplasts that successfully expressed a BiFC YFP fluorescence signal was determined by counting the number of positive cells, and the total number of cells, per microscope field (each field typically contained ~40 protoplasts), in approximately five microscope fields per transfection. Each combination of constructs was analysed three times when analysing the frequency.

### ***In vitro translation and in vitro pull-down analysis***

The SP2 and OEP80 CDSs were cloned into pBlueScript II SK- using a single SmaI restriction site and verified by DNA sequencing. The preSSU construct was described previously, as was the in vitro transcription/translation procedure (16, 58).

The GST-SP1flex and GST proteins were purified from bacteria as described previously, as was the procedure employed for in vitro pull-down analysis (7).

### ***Chloroplast isolation, protein import, and topology analysis***

Chloroplasts were isolated from 14-day-old in vitro grown plants (or, when stated, from protoplasts). Isolations, protein import, and protease treatments were performed as described previously (16, 38, 59-62). The presented chloroplast protein import data are representative of three independent experiments.

### ***Immunoblotting, immunoprecipitation and blue native (BN) PAGE***

Immunoblotting was performed as previously described (34, 63) with minor modifications. Total protein samples of 10-20 µg, prepared from seedlings, were typically analysed. Primary antibodies were as follows. To identify TOC proteins or components of the translocon at the inner envelope membrane of chloroplasts (TIC), we employed: anti-atToc75-III antibody (36); anti-atToc159 antibody (64); anti-atToc33 (G-domain) antibody (36); anti-atTic110 antibody (65, 66); and anti-atTic40 antibody (36). To identify non-TOC outer envelope membrane proteins, we employed: anti-OEP80 antibody (67); and anti-SFR2 antibody (68). To identify chloroplast stromal proteins, we employed: anti-cpHsc70 (AgriSera, AS08 348) (69); anti-Hsp93 (heat shock protein, 93 kD) antibody (16, 70); and anti-PRPL35 antibody (7). To identify proteins of other cellular compartments, we employed: anti-Slp1 (mitochondria) (71); anti-calreticulin (ER) (72, 73); anti-H3 histone (Abcam; nucleus) (36); and anti-PEX13 and anti-PEX14 (the latter, Agrisera; peroxisome) (20, 21). Other primary antibodies we employed were: anti-HA tag (Sigma); anti-c-Myc tag (Sigma); anti-GFP (detects both GFP and YFP; Sigma); and anti-FLAG tag (Sigma). As most of the proteins analysed in this study were membrane proteins, we employed Slp1 (a non-chloroplastic membrane protein) and Tic110 (an internal chloroplast membrane protein unaffected by CHLORAD (7, 8)) as loading controls.

Secondary antibodies were anti-rabbit IgG conjugated with horseradish peroxidase (Santa Cruz Biotechnology), or, in the case of anti-c-Myc and anti-FLAG, anti-mouse IgG conjugated with horseradish peroxidase (GE Healthcare). Chemiluminescence was detected using ECL Plus Western Blotting Detection Reagents (GE Healthcare) and an LAS-4000 imager (Fujifilm). Band intensities were quantified using Aida software (Raytest). Quantification data were based on results from at least three experiments all showing a similar trend. Typical images are shown in all figures.

For the immunoprecipitation of HA-tagged proteins, total protein (~500 mg) was extracted from protoplasts in IP buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) containing 0.5% plant protease inhibitor cocktail (PPIC, Sigma), and centrifuged at 20,000g for 10 min at 4°C. The clear lysate was then incubated with 50 µl EZview Red Anti-HA Affinity Gel (Sigma) for 2 h to overnight at 4°C with slow rotation. After six washes with 500 µl IP-washing buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% Triton X-100), bound proteins were eluted by boiling in 2× SDS-PAGE loading buffer (50 mM Tris-HCl, pH 6.8, 20% glycerol, 1% sodium dodecyl sulphate [SDS], and 0.1 M DTT) for 5 min, and analysed by SDS-PAGE and immunoblotting. A similar procedure was adopted for the immunoprecipitation of Myc-tagged proteins, except that 50 µl EZview Red Anti-c-Myc Affinity Gel (Sigma) was used instead of the anti-HA gel. When detecting ubiquitinated proteins, the IP buffer also contained 10 mM N-ethylmaleimide (NEM; Sigma).

Two-dimensional BN-PAGE was performed using a procedure described previously (74).

#### ***Tandem affinity purification (TAP) and mass spectrometry***

Chloroplasts were isolated from a complemented *ppi1* mutant line carrying the *TAP:Toc33* construct, and then used as starting material for TAP. The TAP procedure was performed as described previously (75), omitting the secondary affinity purification step which was not essential for our analysis. The Tobacco Etch Virus (TEV) nuclear-inclusion-a endopeptidase eluates were concentrated 1:10 by using Vivaspinn 500 ultrafiltration spin columns (Sartorius Stedim Biotech), boiled with 1 volume 2× SDS-PAGE loading buffer, and loaded on SDS-PAGE gels for analysis. Silver staining was used to visualize proteins and estimate their sizes and migration positions. For identification of CDC48, the 75-100 kD region of a Coomassie Brilliant Blue-stained SDS-PAGE gel slice was subjected to in-gel trypsin digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Scaffold (Proteome Software) and Mascot database searches were used to interpret the results.

#### ***In vivo retrotranslocation assays***

The method used was adapted and modified from similar approaches commonly applied in ERAD studies (23). Transfected protoplasts were used for these assays in order to facilitate the detection of substrate and enable efficient and uniform application of proteasome inhibitor; in addition, protoplastation applies a stress which triggers the TOC degradation process (8). First, SP1-HA and ubiquitin were transiently overexpressed in 10<sup>6</sup> protoplasts for each genotype to increase detection sensitivity for higher molecular weight (ubiquitinated) forms of SP1 substrate (76). The transformed protoplasts were incubated for 15 h, and then bortezomib was applied to a final concentration of 5 µM before an additional 3 h incubation. Subsequent fractionation steps to produce separate chloroplast and cytosol samples were all carried out on

ice or at 4°C, and used previously described procedures with modifications (59, 77). Protoplasts were pelleted by centrifugation at 100g for 2 min, and gently resuspended with protoplast-washing buffer (500 mM mannitol, 4 mM 4-morpholineethanesulfonic acid [MES]-KOH, pH 5.6). Then, the protoplasts were pelleted again and resuspended by gentle agitation in 500 µl HS buffer (50 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid [HEPES]-NaOH, pH 8.0, 0.3 M sorbitol) containing 0.5% PPIC and 5 µM bortezomib, and gently forced twice through 10 µm nylon mesh to release chloroplasts. The collected flow-through was centrifuged at 1,000g for 5 min to produce a chloroplast-containing pellet and a cytosol-containing supernatant (S1). The pellet was gently resuspended in 500 µl HS buffer, and the chloroplasts were purified by a two-step Percoll (Fisher Scientific) gradient (38). Intact chloroplasts were washed with 500 µl HS buffer, and then pelleted by centrifugation at 1,000g for 5 min. The S1 sample was centrifuged at 10,000g for 15 min. The resulting supernatant (S10) was recovered and ultracentrifuged at 100,000g for 1 h, producing a further supernatant (S100) that was concentrated to 50 µl by using Vivaspin 500 ultrafiltration spin columns; this was the cytosolic fraction. The pelleted chloroplasts were lysed in 100 µl denaturing buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 10 mM NEM, 1% SDS, 2% Sarcosyl, 5 mM dithiothreitol [DTT]) containing PPIC, while the cytosolic fraction was mixed with 50 µl 2×denaturing buffer containing PPIC. Finally, the SP1-HA protein was purified by immunoprecipitation using a previously described procedure to improve sensitivity of detection of the ubiquitinated protein (7). Experiments were repeated three times, and similar results were obtained.

### ***In vitro retrotranslocation assays***

This method was adapted and modified from previous reports (78, 79). Approximately  $4 \times 10^6$  protoplasts from CDC48-WT or CDC48-DN plants were incubated with 10 µM estradiol for 2 days to induce abundant expression of CDC48 protein in the cytosol. After incubation, the protoplasts were washed once with 2 ml protoplast-washing buffer, and then a cytosolic fraction was prepared, essentially as described in the section above except that the protoplasts were resuspended in import buffer (50 mM HEPES-NaOH, pH 8.0, 3 mM MgSO<sub>4</sub>, 0.3 M sorbitol, 5 mM MgATP, 20 mM gluconic acid [potassium salt], 10 mM NaHCO<sub>3</sub>, 0.2% bovine serum albumin) containing 5 µM bortezomib and complete EDTA-free protease inhibitor cocktail (PIC, Roche; following the manufacturer's instructions), instead of HS buffer. Cytosol was concentrated to approximately 10 mg/ml using Vivaspin 500 ultrafiltration spin columns, and was used immediately, or aliquoted and stored at -80°C. Chloroplasts were prepared from 10<sup>6</sup> CDC48-DN protoplasts transiently overexpressing SP1-HA and FLAG-tagged ubiquitin (and induced with 10 µM estradiol), using the procedure described in the section above. Chloroplast pellets were gently resuspended in reaction buffer (import buffer containing 30 mM MgATP, 0.1 µg/µl ubiquitin [Sigma], 5

$\mu$ M bortezomib, and PIC), and then immediately used in in vitro retrotranslocation reactions. Each reaction contained: reaction buffer (to a final reaction volume of 50  $\mu$ l), 10  $\mu$ l cytosol, and  $2 \times 10^7$  chloroplasts. Reactions were incubated at 25°C for 1 h with occasional agitation to resuspend the chloroplasts, and were then stopped by adding 1  $\mu$ l of 250 mM NEM to a final concentration of 5 mM. Subsequently, at 4°C, the chloroplasts were pelleted by centrifugation at 10,000g for 1 min, and the resulting supernatant was further centrifuged at 20,000g for 10 min to produce another supernatant corresponding to the cytosolic fraction. The SP1-HA protein was enriched from the resulting chloroplast and cytosol fractions by immunoprecipitation, as described in the section above. Ubiquitinated SP1 protein (either resident in the chloroplasts, or retrotranslocated into the cytosol) was detected by anti-FLAG immunoblotting; use of transiently-expressed FLAG-tagged ubiquitin in this way increased the detection sensitivity. Experiments were repeated three times, and similar results were obtained.

### ***Statistical analysis***

Statistical calculations (mean, standard error of the mean, t-test) were performed using Microsoft Excel software. Statistical significance of differences between two experimental groups was assessed by using a two-tailed Student's t-test. Differences between two datasets were considered significant at  $p < 0.05$ .

## References and Notes:

1. P. Jarvis, E. López-Juez, Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* **14**, 787-802 (2013).
2. L. G. L. Richardson, R. Singhal, D. J. Schnell, The integration of chloroplast protein targeting with plant developmental and stress responses. *BMC Biol.* **15**, 118 (2017).
3. E. Demarsy, A. M. Lakshmanan, F. Kessler, Border control: selectivity of chloroplast protein import and regulation at the TOC-complex. *Front. Plant Sci.* **5**, 483 (2014).
4. L. X. Shi, S. M. Theg, The chloroplast protein import system: from algae to trees. *Biochim. Biophys. Acta* **1833**, 314-331 (2013).
5. H. M. Li, C. C. Chiu, Protein transport into chloroplasts. *Annu. Rev. Plant Biol.* **61**, 157-180 (2010).
6. K. Nishimura, Y. Kato, W. Sakamoto, Essentials of proteolytic machineries in chloroplasts. *Mol. Plant* **10**, 4-19 (2017).
7. Q. Ling, W. Huang, A. Baldwin, P. Jarvis, Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. *Science* **338**, 655-659 (2012).
8. Q. Ling, P. Jarvis, Regulation of chloroplast protein import by the ubiquitin E3 ligase SP1 is important for stress tolerance in plants. *Curr. Biol.* **25**, 2527-2534 (2015).
9. X. Wu, T. A. Rapoport, Mechanistic insights into ER-associated protein degradation. *Curr. Opin. Cell Biol.* **53**, 22-28 (2018).
10. D. Zattas, M. Hochstrasser, Ubiquitin-dependent protein degradation at the yeast endoplasmic reticulum and nuclear envelope. *Crit. Rev. Biochem. Mol. Biol.* **50**, 1-17 (2015).
11. R. J. Braun, B. Westermann, With the help of MOM: mitochondrial contributions to cellular quality control. *Trends Cell Biol.* **27**, 441-452 (2017).
12. P. Jarvis *et al.*, An *Arabidopsis* mutant defective in the plastid general protein import apparatus. *Science* **282**, 100-103 (1998).
13. T. Ulrich, D. Rapoport, Biogenesis of beta-barrel proteins in evolutionary context. *Int. J. Med. Microbiol* **305**, 259-264 (2015).
14. I. E. Gentle, L. Burri, T. Lithgow, Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58**, 1216-1225 (2005).
15. P. M. Day, D. Potter, K. Inoue, Evolution and targeting of Omp85 homologs in the chloroplast outer envelope membrane. *Front. Plant Sci.* **5**, 535 (2014).
16. W. Huang, Q. Ling, J. Bédard, K. Lilley, P. Jarvis, In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. *Plant Physiol.* **157**, 147-159 (2011).
17. M. Töpel, Q. Ling, P. Jarvis, Neofunctionalization within the Omp85 protein superfamily during chloroplast evolution. *Plant Signal. Behav.* **7**, 161-164 (2012).
18. Y. C. Hsueh *et al.*, The chloroplast outer envelope protein P39 in *Arabidopsis thaliana* belongs to the Omp85 protein family. *Proteins* **85**, 1391-1401 (2017).
19. J. D. Woodson *et al.*, Ubiquitin facilitates a quality-control pathway that removes damaged chloroplasts. *Science* **350**, 450-454 (2015).
20. R. H. Pan, J. Satkovich, J. P. Hu, E3 ubiquitin ligase SP1 regulates peroxisome biogenesis in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **113**, E7307-E7316 (2016).
21. Q. Ling, N. Li, P. Jarvis, Chloroplast ubiquitin E3 ligase SP1: does it really function in peroxisomes? *Plant Physiol.* **175**, 586-588 (2017).
22. Y. Liu, J. Li, Endoplasmic reticulum-mediated protein quality control in *Arabidopsis*. *Front. Plant Sci.* **5**, 162 (2014).
23. S. Neal *et al.*, The Dfm1 Derlin is required for ERAD retrotranslocation of integral membrane proteins. *Mol. Cell* **69**, 306-320.e304 (2018).
24. S. Schoebel *et al.*, Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* **548**, 352-355 (2017).



25. K. Römisch, A case for Sec61 channel involvement in ERAD. *Trends Biochem. Sci.* **42**, 171-179 (2017).
26. Y. Ye, W. K. Tang, T. Zhang, D. Xia, A mighty "protein extractor" of the cell: structure and function of the p97/CDC48 ATPase. *Front. Mol. Biosci.* **4**, 39 (2017).
27. S. Park, D. M. Rancour, S. Y. Bednarek, In planta analysis of the cell cycle-dependent localization of AtCDC48A and its critical roles in cell division, expansion, and differentiation. *Plant Physiol.* **148**, 246-258 (2008).
28. E. Lecuyer *et al.*, Protein stability and transcription factor complex assembly determined by the SCL-LMO2 interaction. *J. Biol. Chem.* **282**, 33649-33658 (2007).
29. P. R. Johnson, R. Swanson, L. Rakhilina, M. Hochstrasser, Degradation signal masking by heterodimerization of MATalpha2 and MATa1 blocks their mutual destruction by the ubiquitin-proteasome pathway. *Cell* **94**, 217-227 (1998).
30. S. Xu, G. Peng, Y. Wang, S. Fang, M. Karbowski, The AAA-ATPase p97 is essential for outer mitochondrial membrane protein turnover. *Mol. Biol. Cell* **22**, 291-300 (2011).
31. K. Nakatsukasa, G. Huyer, S. Michaelis, J. L. Brodsky, Dissecting the ER-associated degradation of a misfolded polytopic membrane protein. *Cell* **132**, 101-112 (2008).
32. J. Stevenson, E. Y. Huang, J. A. Olzmann, Endoplasmic reticulum-associated degradation and lipid homeostasis. *Annu. Rev. Nutr.* **36**, 511-542 (2016).
33. M. G. Iadanza *et al.*, Lateral opening in the intact beta-barrel assembly machinery captured by cryo-EM. *Nat. Commun.* **7**, 12865 (2016).
34. S. Kovacheva *et al.*, *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J.* **41**, 412-428 (2005).
35. J. P. Stanga, K. Boonsirichai, J. C. Sedbrook, M. S. Otegui, P. H. Masson, A role for the TOC complex in Arabidopsis root gravitropism. *Plant Physiol.* **149**, 1896-1905 (2009).
36. A. R. Kasmati, M. Töpel, R. Patel, G. Murtaza, P. Jarvis, Molecular and genetic analyses of Tic20 homologues in *Arabidopsis thaliana* chloroplasts. *Plant J.* **66**, 877-889 (2011).
37. Y. C. Hsueh *et al.*, The outer membrane Omp85-like protein P39 influences metabolic homeostasis in mature *Arabidopsis thaliana*. *Plant Biol. (Stuttg.)* **20**, 825-833 (2018).
38. H. Aronsson, P. Jarvis, A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* **529**, 215-220 (2002).
39. Q. Ling, W. Huang, P. Jarvis, Use of a SPAD-502 meter to measure leaf chlorophyll concentration in *Arabidopsis thaliana*. *Photosynth. Res.* **107**, 209-214 (2011).
40. D. Constan, R. Patel, K. Keegstra, P. Jarvis, An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. *Plant J.* **38**, 93-106 (2004).
41. S. Schelbert *et al.*, Pheophytin pheophorbide hydrolase (pheophytinase) is involved in chlorophyll breakdown during leaf senescence in *Arabidopsis*. *Plant Cell* **21**, 767-785 (2009).
42. G. Gálvez-Valdivieso *et al.*, The high light response in *Arabidopsis* involves ABA signaling between vascular and bundle sheath cells. *Plant Cell* **21**, 2143-2162 (2009).
43. A. Daudi *et al.*, The apoplastic oxidative burst peroxidase in *Arabidopsis* is a major component of pattern-triggered immunity. *Plant Cell* **24**, 275-287 (2012).
44. O. H. M. Leyser, I. J. Furner, Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana*. *Development* **116**, 397-403 (1992).
45. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 3 (2011).
46. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).
47. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858 (2015).
48. D. M. Goodstein *et al.*, Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178-D1186 (2012).
49. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780 (2013).

50. F. Ronquist *et al.*, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539-542 (2012).
51. R. Higuchi, B. Krummel, R. K. Saiki, A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* **16**, 7351-7367 (1988).
52. M. J. Dubin, C. Bowler, G. Benvenuto, A modified Gateway cloning strategy for overexpressing tagged proteins in plants. *Plant Methods* **4**, 3 (2008).
53. M. Karimi, B. De Meyer, P. Hilson, Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103-105 (2005).
54. M. D. Curtis, U. Grossniklaus, A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462-469 (2003).
55. J. S. Rohila, M. Chen, R. Cerny, M. E. Fromm, Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* **38**, 172-181 (2004).
56. T. Tzfira *et al.*, pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. *Plant Mol. Biol.* **57**, 503-516 (2005).
57. F. H. Wu *et al.*, Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* **5**, 16 (2009).
58. A. Baldwin *et al.*, A molecular-genetic study of the Arabidopsis Toc75 gene family. *Plant Physiol.* **138**, 715-733 (2005).
59. L. M. Fitzpatrick, K. Keegstra, A method for isolating a high yield of *Arabidopsis* chloroplasts capable of efficient import of precursor proteins. *Plant J.* **27**, 59-65 (2001).
60. H. Aronsson, R. P. Jarvis, Rapid isolation of Arabidopsis chloroplasts and their use for in vitro protein import assays. *Methods Mol. Biol.* **774**, 281-305 (2011).
61. J. Froehlich, Studying Arabidopsis envelope protein localization and topology using thermolysin and trypsin proteases. *Methods Mol. Biol.* **774**, 351-367 (2011).
62. C. C. Chu, H. M. Li, Determining the location of an Arabidopsis chloroplast protein using in vitro import followed by fractionation and alkaline extraction. *Methods Mol. Biol.* **774**, 339-350 (2011).
63. S. Kovacheva, J. Bédard, A. Wardle, R. Patel, P. Jarvis, Further *in vivo* studies on the role of the molecular chaperone, Hsp93, in plastid protein import. *Plant J.* **50**, 364-379 (2007).
64. J. Bauer *et al.*, The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**, 203-207 (2000).
65. T. Inaba *et al.*, Arabidopsis Tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* **17**, 1482-1496 (2005).
66. H. Aronsson *et al.*, Nucleotide binding and dimerization at the chloroplast pre-protein import receptor, atToc33, are not essential *in vivo* but do increase import efficiency. *Plant J.* **63**, 297-311 (2010).
67. S. C. Hsu, M. Nafati, K. Inoue, OEP80, an essential protein paralogous to the chloroplast protein translocation channel Toc75, exists as a 70-kD protein in the *Arabidopsis thaliana* chloroplast outer envelope. *Plant Mol. Biol.* **78**, 147-158 (2012).
68. N. Fourrier *et al.*, A role for SENSITIVE TO FREEZING2 in protecting chloroplasts against freeze-induced damage in Arabidopsis. *Plant J.* **55**, 734-745 (2008).
69. P. H. Su, H. M. Li, Stromal Hsp70 is important for protein translocation into pea and *Arabidopsis* chloroplasts. *Plant Cell* **22**, 1516-1531 (2010).
70. E. Nielsen, M. Akita, J. Davila-Aponte, K. Keegstra, Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO J.* **16**, 935-946 (1997).
71. B. Gehl, C. P. Lee, P. Bota, M. R. Blatt, L. J. Sweetlove, An Arabidopsis stomatin-like protein affects mitochondrial respiratory supercomplex organization. *Plant Physiol.* **164**, 1389-1400 (2014).



72. J. Denecke *et al.*, The tobacco homolog of mammalian calreticulin is present in protein complexes in vivo. *Plant Cell* **7**, 391-406 (1995).
73. O. K. Teh, I. Moore, An ARF-GEF acting at the Golgi and in selective endocytosis in polarized plant cells. *Nature* **448**, 493-496 (2007).
74. S. Kikuchi, J. Bédard, M. Nakai, One- and two-dimensional blue native-PAGE and immunodetection of low-abundance chloroplast membrane protein complexes. *Methods Mol. Biol.* **775**, 3-17 (2011).
75. C. Andrès, B. Agne, F. Kessler, Preparation of multiprotein complexes from Arabidopsis chloroplasts using tandem affinity purification. *Methods Mol. Biol.* **775**, 31-49 (2011).
76. J. Bloom, M. Pagano, Experimental tests to definitively determine ubiquitylation of a substrate. *Methods Enzymol.* **399**, 249-266 (2005).
77. J. Ito *et al.*, Analysis of the Arabidopsis cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* **10**, 1571-1582 (2011).
78. R. M. Garza, B. K. Sato, R. Y. Hampton, In vitro analysis of Hrd1p-mediated retrotranslocation of its multispanning membrane substrate 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase. *J. Biol. Chem.* **284**, 14710-14722 (2009).
79. K. Nakatsukasa, J. L. Brodsky, In vitro reconstitution of the selection, ubiquitination, and membrane extraction of a polytopic ERAD substrate. *Methods Mol. Biol.* **619**, 365-376 (2010).
80. Y. D. Paila *et al.*, Multi-functional roles for the polypeptide transport associated domains of Toc75 in chloroplast protein import. *eLife* **5**, e12631 (2016).
81. Y. L. Chen, L. J. Chen, H. M. Li, Polypeptide transport domains of the Toc75 channel protein are located in the intermembrane space of chloroplasts. *Plant Physiol.* **172**, 235-243 (2016).
82. T. Hruz *et al.*, Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics* **2008**, 420747 (2008).
83. K. Nicolaisen *et al.*, The Omp85-type outer membrane protein p36 of *Arabidopsis thaliana* evolved by recent gene duplication. *J. Plant Res.* **128**, 317-325 (2015).
84. S. Dalal, M. F. Rosser, D. M. Cyr, P. I. Hanson, Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway. *Mol. Biol. Cell* **15**, 637-648 (2004).
85. S. Kikuchi, T. Hirohashi, M. Nakai, Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. *Plant Cell Physiol.* **47**, 363-371 (2006).
86. K. Y. Chen, H. M. Li, Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. *Plant J.* **49**, 149-158 (2007).

## Acknowledgements:

We thank M. Rashbrooke for assistance with initial analyses and rough mapping of *sp2*; Y. Zeng, J. Bédard and N. Li for technical assistance; N. Allcock and S. Hyman (University of Leicester EM Laboratory) for electron microscopy; A.R. Bottrill (University of Leicester Protein Nucleic Acid Chemistry Laboratory) for mass spectrometry; L. Dolan for comments on the manuscript; L.J. Sweetlove (Slp1), J. Denecke (calreticulin), and B. Bartel (PEX13) for antibodies; F. Wu for the Toc159 BiFC construct; S.Y. Bednarek for the H6T7-DN-B construct; G. Benvenuto (and Addgene) for the pE3c construct; J. Chory and J. Woodson for *ppi2-3 (fts1)* seeds; and SIGnAL and NASC for the *sp2-4* allele. **Funding:** Supported by grants from the BBSRC (BB/D016541/1, BB/K018442/1-2, BB/R009333/1) to R.P.J.; the Royal Society Rosenheim Research Fellowship to R.P.J.; a Department of Plant Sciences DPhil Studentship to R.P.J and W.B.; a Gatsby Sainsbury PhD Studentship to R.T.; and a Carl Tryggers Stiftelse för Vetenskaplig Forskning Fellowship (CTS 11:479) to M.T.. **Author contributions:** Q.L. and W.B. formulated the research plan, designed and performed experiments, and interpreted results. Q.L. carried out the functional analysis of CDC48, and contributed to the preparation of the manuscript. W.B. completed genetic, molecular, physiological and detailed functional studies of SP2, with assistance from Q.L.. Q.L. and W.B. contributed equally to the study. R.T. performed the genetic mapping of *sp2*, prepared samples for whole-genome sequencing, and conducted protein import assays with assistance from T.D.S.. M.T. analysed the whole-genome sequencing data and identified a candidate *sp2* locus. P.L. performed the TAP experiment. A.B. conducted the mutant screen and identified the original *sp2* alleles. R.P.J. conceived of the study, supervised the work, analysed the data, and prepared the manuscript. All authors discussed the results and commented on the manuscript. **Competing interests:** This work is the subject of pending UK patent applications GB 1803833.1 and GB 1803834.9 (inventor is R.P.J. in both cases), which cover the use of CHLORAD to manipulate plastid development in crops. **Data and materials availability:** Sequencing data used in this study, as well as the reference genome for *ppi1*, are available at the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>) under Bioproject PRJNA488548. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRoteomics IDentifications (PRIDE) repository (<https://www.ebi.ac.uk/pride/archive/>) with the identifier PXD010954. All other data are available in the manuscript or the supplementary materials.

**Supplementary Materials:**

Figures S1-S20

Tables S1-S2



## Supplementary Materials for

Ubiquitin-dependent chloroplast-associated protein degradation in plants

Qihua Ling, William Broad, Raphael Trösch, Mats Töpel, Tijen Demiral Sert, Panagiotis Lymperopoulos, Amy Baldwin & R. Paul Jarvis

Correspondence to: [paul.jarvis@plants.ox.ac.uk](mailto:paul.jarvis@plants.ox.ac.uk)

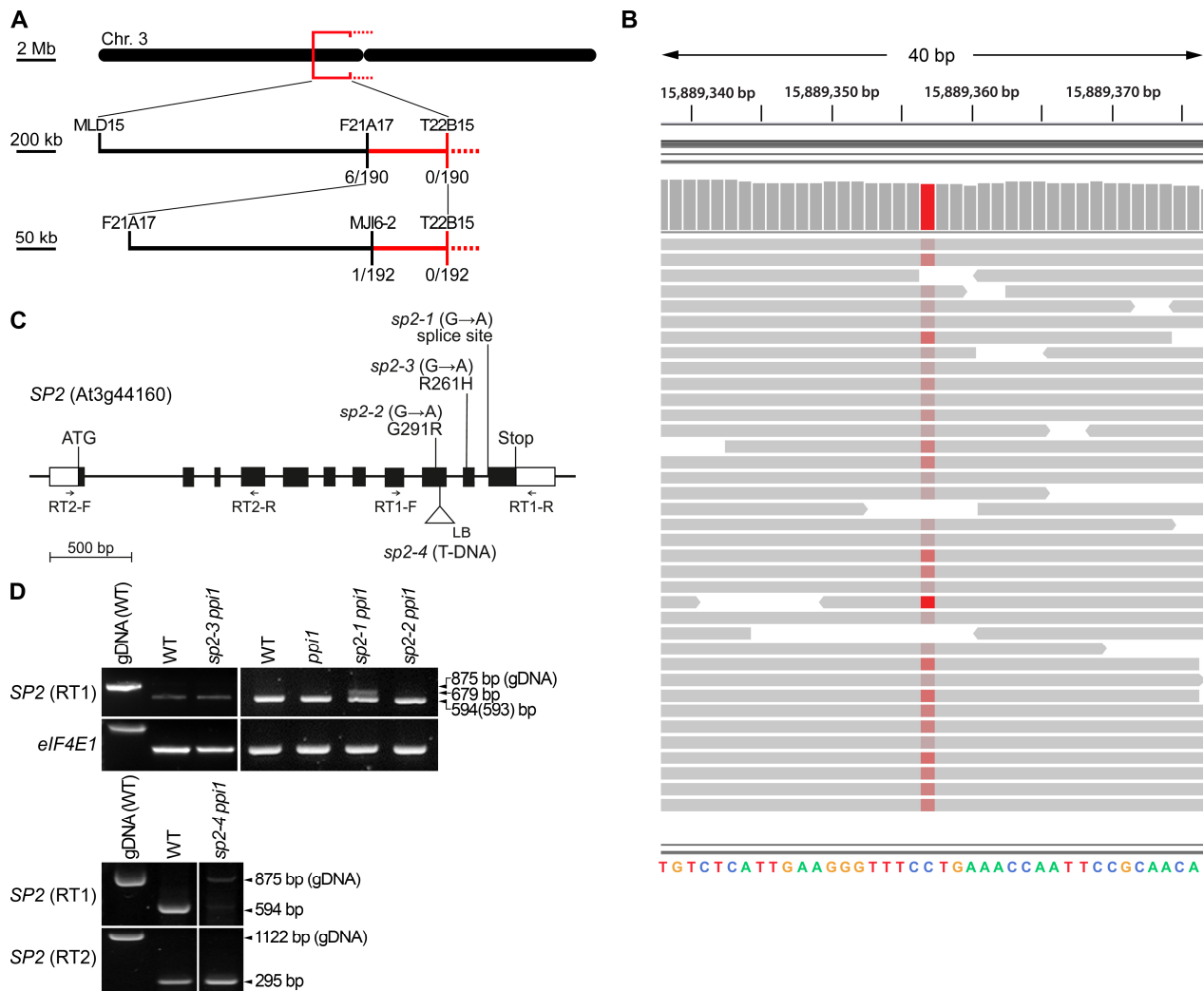
### **This PDF file includes:**

Figs. S1 to S20  
Table S1

### **Other Supplementary Materials for this manuscript include the following:**

Table S2

## Supplementary Figures



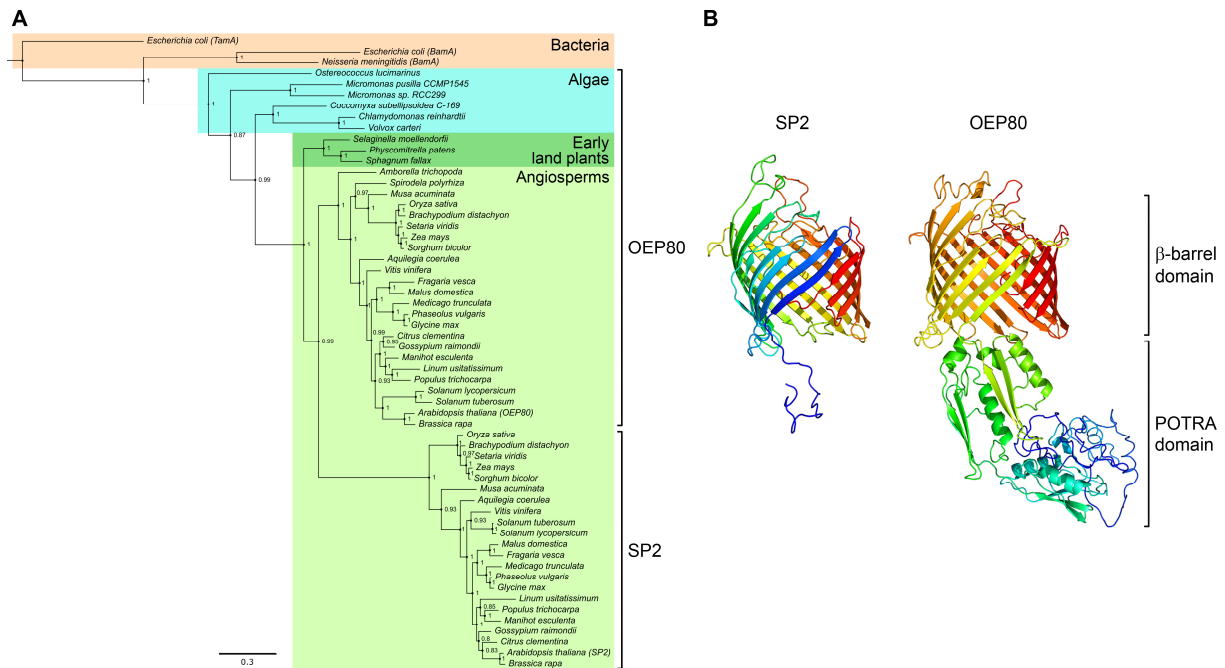
**Fig. S1. Identification of the *SP2* gene.**

(A) Genetic mapping of the *sp2* locus. Initial analysis of an F<sub>2</sub> population of 190 individuals from crosses between *sp2-1 ppi1* (Col-0 ecotype) and *ppi1* (introgressed into the *Ler* ecotype) placed the *sp2* locus to the south of marker F21A17 at position 12285000 on the upper arm of chromosome 3. A second mapping population of 192 plants placed the *sp2* mutation to the south of a more southerly marker, MJ16-2 at position 12597802 on the upper arm of chromosome 3. However, the genetic mapping failed to define a southern boundary concerning the location of *sp2*, owing to the existence of an “island” of Col-0 DNA in the *Ler*-introgressed *ppi1* line, around the chromosome 3 centromere. Numbers of recombinants at key markers are shown.

**(B)** Identification of the *sp2* mutations by whole-genome sequencing. The figure shows 40 bp of the alignment of *sp2-1 ppl1* reads to the *ppl1* reference genome. The vertical grey lines represent the read coverage at each base (this is 34 at the mutated site), and the horizontal grey lines indicate individual sequence reads. Red positions in the reads show the mutated site (C-to-T on the strand shown; the brightness of the red colour indicates the mapping quality) that is responsible for the *sp2* phenotype. The *SP2* (At3g44160) gene is coded on the reverse strand relative to the reference genome, and so the complement of the *SP2* gene, in the *ppl1* reference sequence, is shown at the bottom. Genomic position in the *ppl1* reference genome is indicated at the top of the figure. Similar results were obtained for the *sp2-2* and *sp2-3* alleles, although the mutation sites within the gene differed.

**(C)** Schematic representation of the *SP2* (At3g44160) gene, annotated with the positions of the *sp2* mutations. Information on the nature and consequence of each mutation is shown. The positions of PCR primers used in D are indicated with arrows. Black boxes and interconnecting black lines, exons and introns, respectively; white boxes, untranslated regions; LB, left border sequence of the SALK T-DNA insertion.

**(D)** Analysis of *SP2* mRNA expression in each of the *sp2* mutant alleles by RT-PCR. The primer pairs used in each case are indicated at left, and their positions are shown in C (table S1). Amplicon sizes are indicated at right (these were verified by DNA sequencing). The *sp2-1* allele carries a splice site mutation and exhibited slicing defects. Two different *sp2-1* amplicons were detected, and sequenced: the larger, minor amplicon retains a portion of intron 10, and the corresponding transcript is predicted to encode a truncated protein; the smaller, more abundant amplicon carries a single guanine nucleotide deletion (at position 924 of the CDS, in exon 10), causing a frameshift such that the corresponding transcript is also predicted to encode a truncated protein. This indicated that *sp2-1*, like *sp2-4*, is most likely a null allele. In contrast, *sp2-2* and *sp2-3* may be weak alleles, which is in line with the chlorophyll data shown in Fig. 11. gDNA, genomic DNA, RT1, RT-PCR primer pair 1; RT2, RT-PCR primer pair 2.

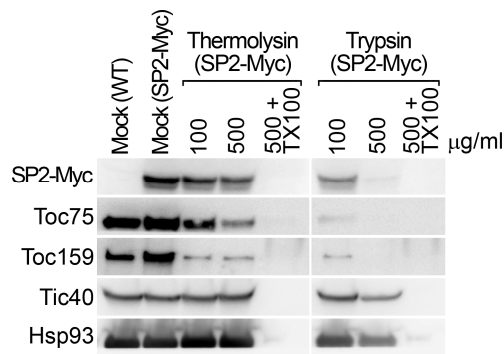


**Fig. S2. Phylogenetic analysis and predicted structure of the SP2 protein.**

(A) Bayesian inference phylogenetic analysis of SP2 and OEP80. Predicted amino acid sequences homologous to *Arabidopsis* SP2 and OEP80 from a variety of plants and algae were retrieved from the Phytozome 12 database (48) (table S2). Bacterial Omp85 sequences were also obtained, and were included in the analysis to act as an outgroup. Sequences were aligned using the MAFFT algorithm (49), and the phylogeny was inferred using MrBayes 3.2 software (50). The tree represents a 50% consensus of 6000 trees generated from two runs, each using 8 Markov chain Monte-Carlo (MCMC) chains for 8 million generations (the first 2000 trees were discarded as burn-in) where standard deviation of split frequencies was 0.001228 at the end of the analysis. Posterior probability values are shown. Scale bar, 0.3 changes per site. The long branch subtending the SP2 clade is consistent with previous results (17), and suggests that SP2 has undergone neofunctionalization.

(B) Structural models for the *Arabidopsis* SP2 and OEP80 proteins. Three-dimensional models for SP2 and OEP80 were derived by homology modelling using the crystal structures of bacterial TamA and BamA proteins, using the Phyre2 server (47). Both protein models are oriented with the N-terminal domain facing downwards. The beta-barrel and polypeptide transport associated (POTRA; OEP80 only) domains are indicated.



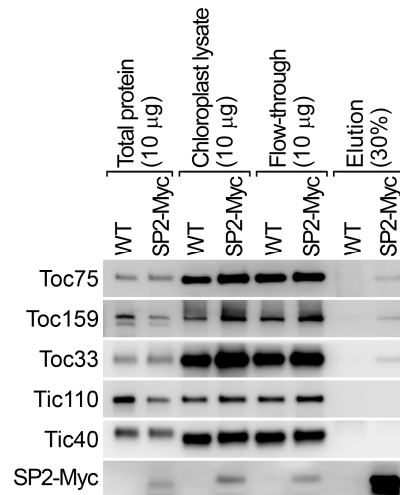


**Fig. S3. Analysis of SP2 chloroplast localization and topology.**

The SP2 CDS was fused with sequence encoding a C-terminal 6×Myc tag and cloned downstream of the strong, constitutive cauliflower mosaic virus (CaMV) 35S promoter in the pB2GW7 binary vector (53), and then the resultant *SP2-Myc* construct was used to stably transform *sp2-4 ppl1* and *sp2-4* plants.

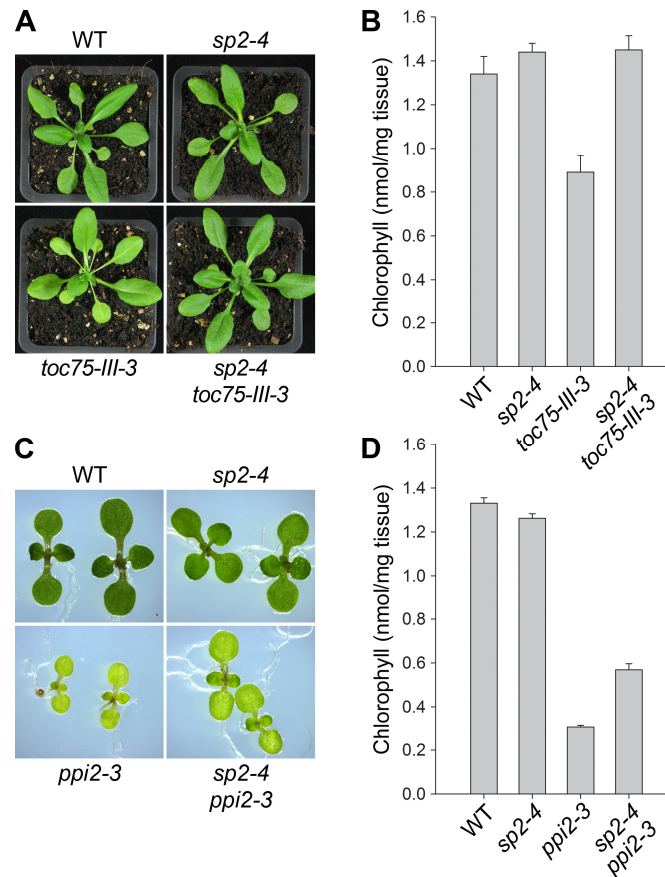
Approximately 12 T<sub>2</sub> transformants for each genotype were analysed, and from these, representative single-locus lines were selected for further investigation based on segregation of the resistance phenotype on petri plates containing phosphinothricin (those showing 3:1 resistant:sensitive ratios), *SP2* mRNA expression (those showing  $\geq 3$ -fold higher levels of expression than untransformed plants), and/or phenotype analysis (those showing a complemented phenotype). The *SP2-Myc* construct could fully complement the phenotype of *sp2-4 ppl1*, indicating that the Myc tag does not affect the function of SP2. Therefore, SP2-Myc *sp2-4* plants selected in this way were analysed here (and in fig. S4). Chloroplast localization and topology of SP2 were analysed by chloroplast isolation and subsequent protease treatment. Chloroplasts were prepared from 14-day-old wild-type (WT) and SP2-Myc *sp2-4* plants. The isolated SP2-Myc *sp2-4* chloroplasts were treated: with thermolysin or trypsin at different concentrations (100 or 500 µg/ml); with either protease plus 1% Triton X-100 detergent (TX100); or, with buffer lacking protease (Mock). Then, the samples (together with a wild-type chloroplast control; Mock) were analysed by SDS-PAGE and immunoblotting using antibodies against the Myc tag and a number of chloroplast proteins. Immunoblotting analysis showed that SP2-Myc is localized in chloroplasts. Specificity of the Myc signal was verified by comparing the WT and SP2-Myc Mock chloroplasts. Partial resistance to thermolysin indicated that the C-terminal tag of SP2-Myc is oriented towards the intermembrane space, as this protease does not penetrate the outer membrane (so that only protein domains exposed at the cytosolic surface are fully sensitive to thermolysin) (61); assuming that SP2 has an even number of transmembrane beta-strands, one may infer that its N-terminus is also located in the intermembrane space. In this assay, SP2 behaved similarly to Toc75, which is as expected as our Toc75 antibody reacts with the N-terminal POTRA domain that is localized in the intermembrane space (80, 81). Toc159 has a large cytosolic domain and was almost completely degraded by thermolysin, confirming efficacy of the treatments. Unlike thermolysin, trypsin can penetrate the outer membrane to the intermembrane space (61). Increasing

trypsin concentrations progressively depleted SP2-Myc, like Toc75, which indicated that SP2 is an outer membrane protein with similar topology as Toc75. In contrast, TIC-associated proteins located largely or wholly in the stroma (Tic40 and Hsp93, respectively) were not sensitive to trypsin, indicating that the protease was behaving as expected. When co-applied with Triton X-100 to solubilize the envelope membranes, both proteases completely degraded SP2-Myc, confirming that the signals observed were due to protection by the membranes and not a result of intrinsic protease-resistance.



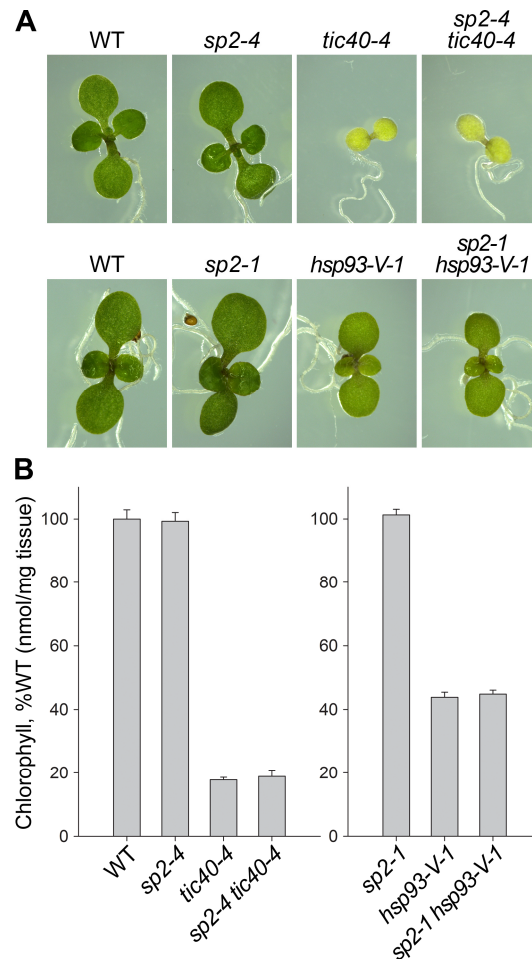
**Fig. S4. Enrichment of SP2 in isolated chloroplasts, and analysis of its interaction with TOC proteins by co-immunoprecipitation.**

Chloroplasts were isolated from 14-day-old wild-type and SP2-Myc *sp2-4* plants, and subsequently lysed with hypotonic buffer (25 mM HEPES, pH 8.0, 4 mM MgCl<sub>2</sub>); the chloroplast pellet was broken up with a sterile plastic pestle and then the sample was rotated at 4°C for 1 h to ensure efficient lysis. The membranes were recovered by ultracentrifugation at 110,000g for 1 h at 4°C, and were then subjected to immunoprecipitation using anti-c-Myc antibody. Input (i.e., before immunoprecipitation was initiated; Chloroplast lysate), flow-through, and immunoprecipitated (Elution) samples were then analysed by immunoblotting using antibodies against: the Myc tag, to verify the enrichment of SP2-Myc; the TOC components Toc159, Toc75 and Toc33, to detect putative SP2-partner interactions; and, Tic110 and Tic40, to assess whether the above interactions are specific. In addition, whole-plant protein extracts (Total protein) were prepared from wild-type and SP2-Myc *sp2-4* plants (equivalent to those from which the chloroplasts were isolated), and analysed in parallel. Enrichment of SP2-Myc in the chloroplast lysates (relative to the total protein samples), in a similar way to other chloroplast proteins, further confirmed that SP2 is localized in chloroplasts. The SP2-Myc *sp2-4* plants analysed here are described in the legend to fig. S3.



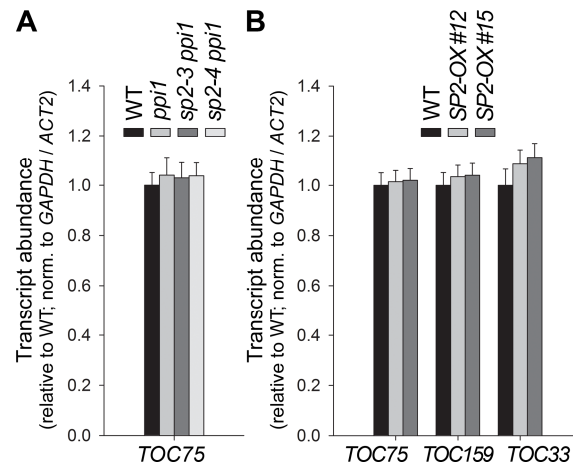
**Fig. S5. Genetic analysis of the effect of the *sp2* mutation on other TOC mutants.**

(A and C) Assessment of the effect of *sp2* on the chlorotic phenotypes of two additional TOC mutants. The *sp2-4* mutation was introduced into the *toc75-III-3* (*mar1*; affecting Toc75) (16) and *ppi2-3* (*fts1*; affecting Toc159) (19) mutant backgrounds by crossing. Homozygous double-mutant populations were identified and subsequently grown alongside wild-type and single-mutant controls under standard conditions on soil for 22 days (A), or on MS agar medium for 10 days (C), prior to photography; typical plants are shown. Both TOC mutations were suppressed by *sp2*, indicating that the effect of the latter is not specific to *ppi1*. (B and D) Chlorophyll contents in the genotypes shown in A and C were measured photometrically following DMF extraction. About ten 10-day-old seedlings per genotype were measured in each experiment, and the values shown are means  $\pm$  SEM derived from four experiments per genotype. The data indicate nmol total chlorophyll per mg tissue fresh weight.



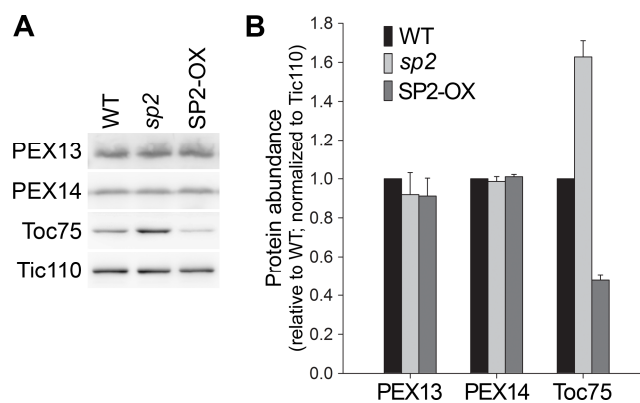
**Fig. S6. Genetic analysis of the specificity of suppression mediated by the *sp2* mutation.**

(A) Assessment of the effect of *sp2* on the chlorotic phenotypes of two mutants with protein import defects linked to the TIC apparatus of the inner envelope membrane. The *sp2-4* mutation was introduced into the *hsp93-V-1* background (34) by crossing. The *sp2-1* mutation was isolated from the *ppi1* background by crossing to wild type followed by PCR genotyping (table S1), and then introduced into the *tic40-4* background (34) by further crossing. Homozygous double-mutant populations were identified and subsequently grown alongside appropriate control genotypes under standard conditions on MS agar medium for 10 days prior to photography; typical plants are shown. Neither TIC-associated mutant was suppressed by *sp2*, indicating that the suppression effect of the latter is specific. (B) Chlorophyll contents in the genotypes shown in A were measured photometrically following DMF extraction. About ten 10-day-old seedlings per genotype were measured in each experiment, and the values shown are means  $\pm$  SEM derived from four experiments per genotype. For comparison purposes, the individual measured values (nmol per mg tissue fresh weight) were expressed as percentages of the corresponding wild-type value.



**Fig. S7. SP2 does not affect TOC component transcript levels.**

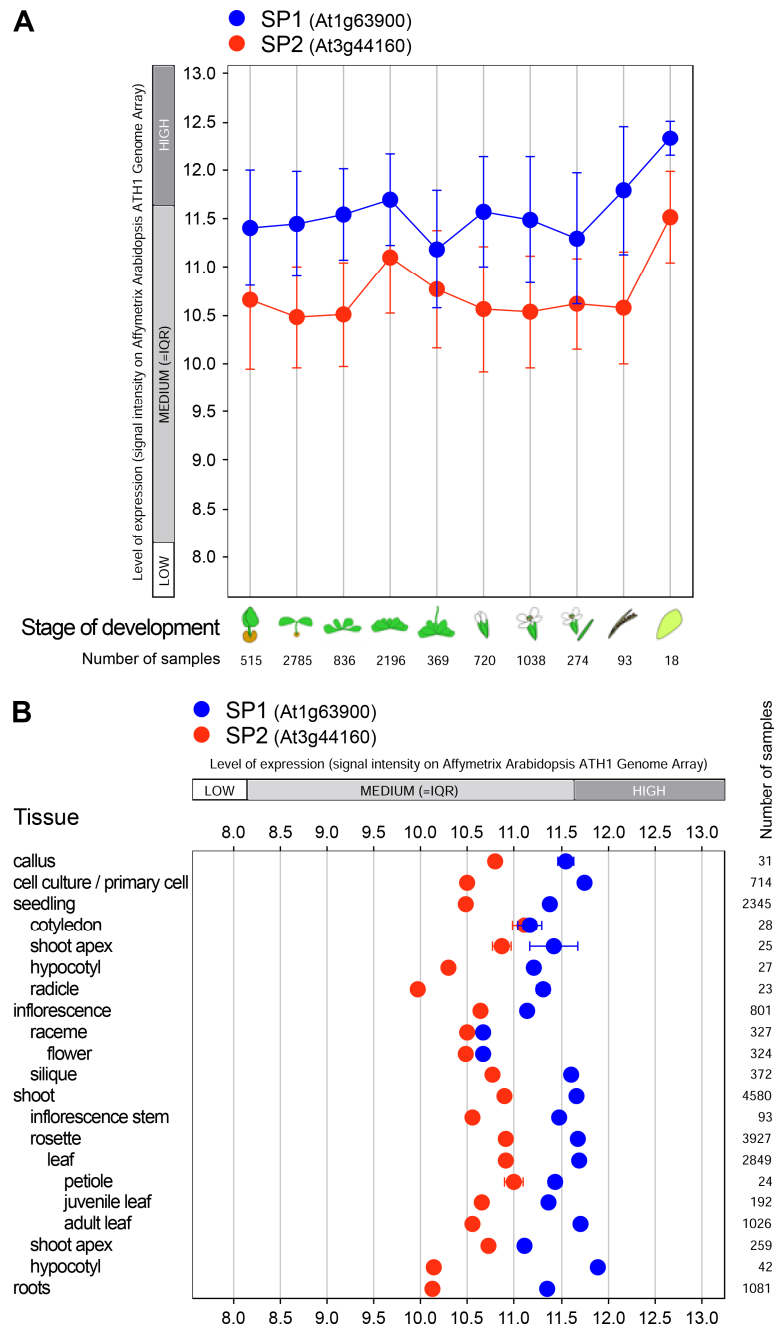
Total RNA samples isolated from 10-day-old plants of the indicated genotypes were analysed by quantitative real-time PCR using gene-specific primers for *atTOC75-III* (A), or *atTOC75-III*, *atTOC159* and *atTOC33* (B). Expression data for the *TOC* genes were normalized using equivalent data for two reference genes: *GAPDH* and *ACTIN2* (*ACT2*) (table S1). For each genotype tested, four biological replicates were analysed, with three technical replicates of each, and the values shown are means  $\pm$  SEM (n = 4 biological replicates).



**Fig. S8. SP2 does not affect the accumulation of peroxisomal protein import machinery components.**

Total protein extracts from 9-day-old seedlings were analysed by immunoblotting using previously-described antibodies (7, 20). As in an earlier study, Tic110 was employed as a sample normalization control (21). Specific immunoreactive bands derived from 3-4 biological replicates were quantified as previously described (7). Error bars indicate SEM (n=4).

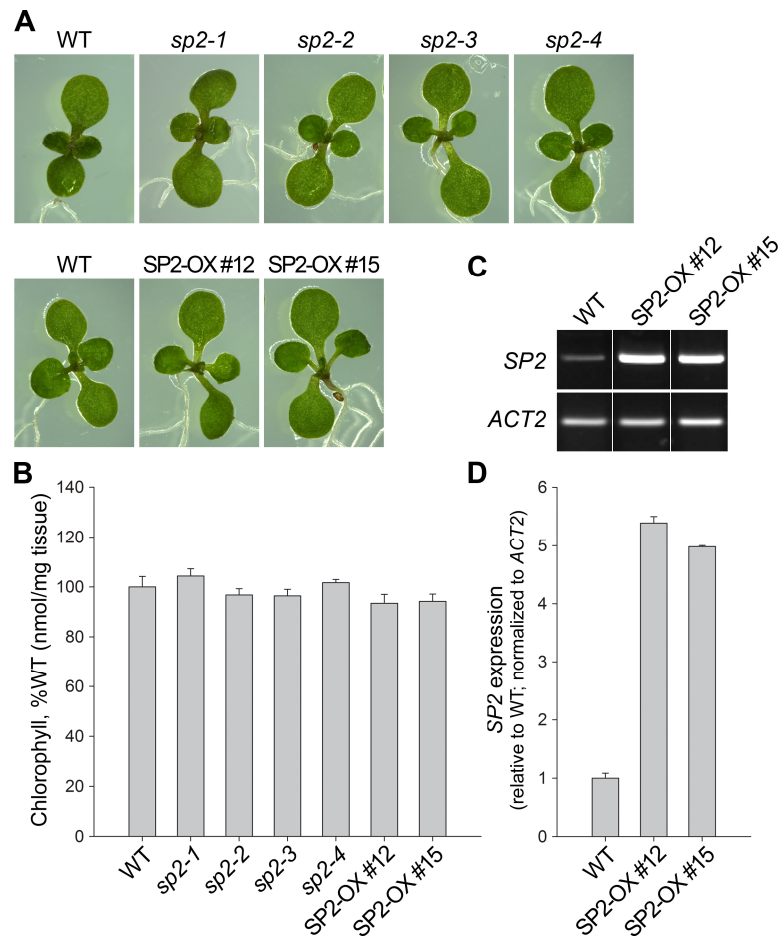




**Fig. S9. The expression profiles of the *SP1* and *SP2* genes are similar.**

The expression profiles shown are based on Affymetrix GeneChip data and were generated using the Development (A) and Anatomy (B) functions of Genevestigator (82). Data from ATH arrays are shown in scatter-plot diagrams. In A, the x-axis represents the following developmental stages, from left to right: germinating seed, seedling, young rosette, developed rosette, bolting, young flower, developed flower, flowers and siliques, mature siliques, and senescent leaves. For each data point, the number of samples is indicated. The values in the plots are means, and the error bars show standard errors. Medium expression

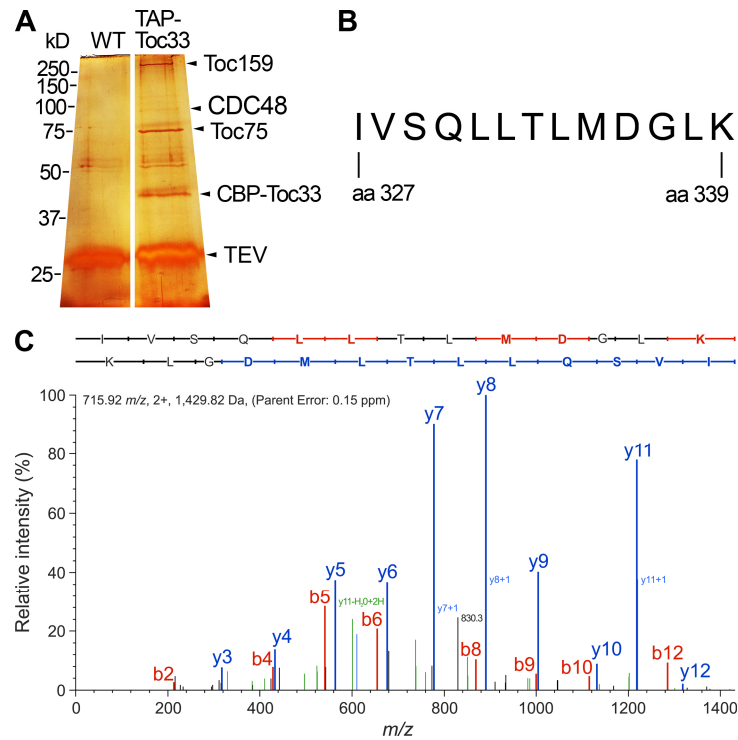
levels are defined as the interquartile range (IQR); values below the IQR are defined as “low expression” and values above the IQR as “high expression”. The presented data provide a more comprehensive and robust characterization of the expression of *SP2* (which a previous study investigated by analysing a promoter-reporter fusion (83)), and additionally provide information on the similarity of the expression pattern of *SP2* with that of *SP1*.



**Fig. S10. The *sp2* single-mutant and SP2 overexpressor plants appear phenotypically normal under standard growth conditions.**

(A and B) Phenotypic analysis of *sp2* single-mutant and SP2 overexpressor (OX) plants grown under normal (non-stressful) conditions in vitro. The three original *sp2* mutants were isolated from the *ppi1* background by crossing to wild type, followed by PCR genotyping (table S1). The SP2 CDS was cloned downstream of the 35S promoter in the pB2GW7 binary vector (53), and the resultant construct (SP2-OX) was used to stably transform wild-type plants. Using procedures similar to those described in the legend to fig. S3, representative single-locus lines were chosen for further analysis, based on the segregation of phosphinothricin resistance and the level of SP2 mRNA overexpression (see C and D); note that the selected transformants are also shown in Fig. 2 (#12 and #15), and in Figs. 4, 7 and 8 (#12 only). Images of typical plants (A) and chlorophyll content measurements determined photometrically following extraction in DMF (B) are shown. In each case 10-day-old seedlings grown under standard conditions on MS agar medium were analysed. For the chlorophyll assays, about 10 seedlings per genotype were measured in each experiment, and the values shown are means  $\pm$  SEM derived from four experiments per genotype. The data indicate nmol total chlorophyll per mg tissue fresh weight.

**(C and D)** Semi-quantitative RT-PCR analysis of *SP2* expression in wild-type and SP2-OX plants. Total RNA samples isolated from 10-day-old plants were analysed by RT-PCR using gene-specific primers for *SP2* and the reference gene *ACT2* (table S1). Amplifications employed a limited number of cycles to avoid saturation, and products were analysed by agarose gel electrophoresis and staining. A representative gel image is shown (C), along with quantification of four biological replicates (D). The amplicon bands were quantified using Aida software, and the data obtained for *SP2* were normalized relative to equivalent data for *ACT2*. Values shown are means  $\pm$  SEM (n = 4).

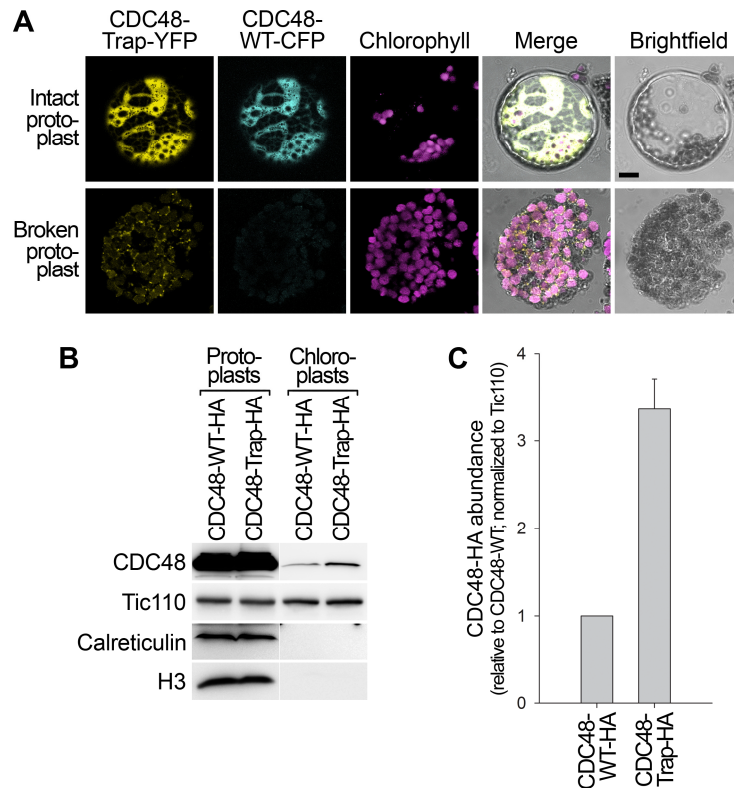


**Fig. S11. Identification of CDC48 in association with purified TOC complexes.**

(A) Silver-stained SDS-PAGE gel showing proteins co-immunoprecipitated with purified TAP-Toc33. The Toc33 CDS was fused with sequence encoding an N-terminal TAP tag upon cloning downstream of the 35S promoter in the NTAPi binary vector (55). The resultant *TAP-Toc33* construct was used to stably transform *ppil* plants. Using procedures similar to those described in the legend to fig. S3, representative, single-locus lines were selected for further analysis, based on segregation of phosphinothricin resistance, visible phenotype, and *TOC33* mRNA expression. The *TAP-Toc33* construct was shown to fully complement the chlorotic phenotype of *ppil*. Those *TAP-Toc33 ppil* lines which showed *TOC33* mRNA levels closest to those in wild type were selected for TAP analysis. Chloroplasts were isolated from 14-day-old wild-type and *TAP-Toc33 ppil* (*TAP-Toc33*) plants, and then subjected to lysis and membrane solubilization using 1% n-dodecyl  $\beta$ -D-maltoside (DDM) prior to complex purification using Protein A Sepharose beads (GE Healthcare). TAP-Toc33 was cleaved by TEV protease (Invitrogen), and the cleaved fraction (containing CBP-Toc33) was eluted. Concentrated elution samples were loaded for silver staining. Proteins specifically associated with Toc33 are indicated by the presence of unique bands in the TAP-Toc33 eluate but not in the wild-type eluate; these included known Toc33 partners, Toc159 and Toc75 (identities of which were confirmed by immunoblotting and mass spectrometry), and a putative CDC48 band.

(B and C) Identification of CDC48 by mass spectrometry. A single representative CDC48 peptide (B), and the associated fragmentation spectrum (C) are shown. The TAP-Toc33 elution sample shown in A was

loaded on a second gel for Coomassie Brilliant Blue staining, and the 75-100 kD region of the stained gel was analysed by in-gel trypsin digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS). The data were analysed using Scaffold Software and a Mascot database search, which in total identified 39 peptides for AtCDC48A (17 unique peptides), covering 48% of the whole AtCDC48A protein (the main CDC48 isoform in *Arabidopsis*; relatively few unique peptides for the other isoforms were identified). In the representative spectrum shown (C), the b and y ions are shown in red and blue, respectively, while the y11 di-charged ion is shown in green. The b ions appear to extend from the N-terminus, and the y ions appear to extend from the C-terminus. aa, amino acid.



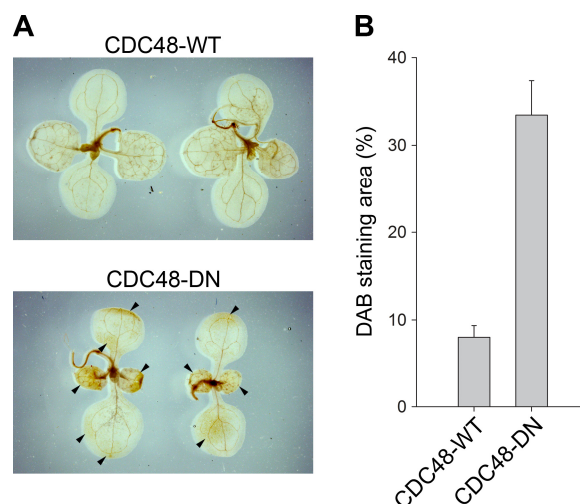
**Fig. S12. CDC48 associates with the chloroplast envelope in vivo.**

(A) Confocal microscopy analysis of the subcellular localization of CDC48. To increase the chance of detecting the fraction of CDC48 associated with chloroplasts (which may be transient and/or unstable), a “Trap” mutant (AtCDC48A<sup>E581Q</sup>, which stabilizes substrate binding (84)) was generated and employed. Mutant and wild-type CDSs were cloned upstream of sequences encoding YFP or CFP, respectively, in the p2GWY7 or p2GWC7 vectors (53), and then the constructs were used to co-transfect wild-type *Arabidopsis* protoplasts. After 15–18 h of incubation, the YFP/GFP fluorescent signals were analysed by confocal microscopy. Corresponding chlorophyll autofluorescence images were employed to orientate the YFP/GFP signals in relation to the chloroplasts. To eliminate interference from cytosolic CDC48 signals, intact protoplasts (upper panels) were broken by gently tapping the coverslip. Retention of YFP signals in association with chloroplasts upon rupturing the cell (lower panels) clearly indicated that the CDC48-Trap-YFP was associated with chloroplasts and localized around the periphery of the organelles in a pattern indicative of envelope localization. In contrast, the CDC48-WT-CFP signal displayed a cytosolic localization which disappeared in broken protoplasts. These results indicate that CDC48 associates transiently with the chloroplast envelope. Brightfield images enabled assessments of the integrity of the protoplasts. The representative images shown are all at the same magnification. Scale bar, 10  $\mu$ m.

(B and C) Immunoblotting analysis of the subcellular localization of CDC48 using isolated chloroplasts. The “Trap” mutant and wild-type CDC48 CDSs were cloned into a modified p2GW7 plant expression

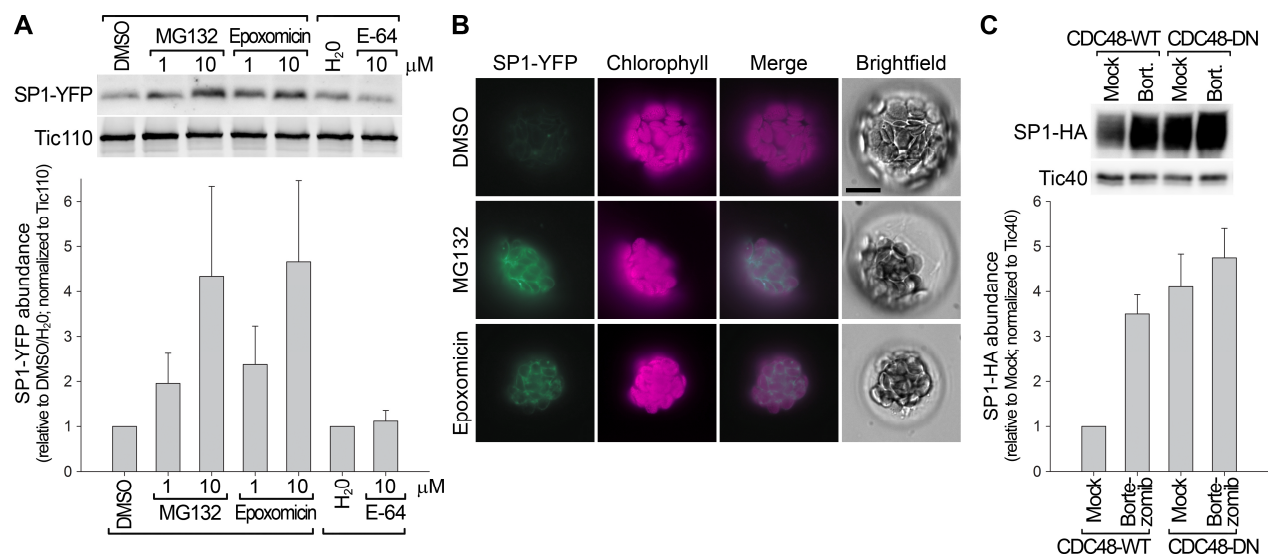


vector (53) that adds a C-terminal HA tag. The resultant constructs were used to transfect *Arabidopsis* protoplasts. Approximately 15 h later, chloroplasts were isolated as described previously (59), and then whole protoplast and isolated chloroplast protein samples were analysed by SDS-PAGE and immunoblotting (B), using antibodies against: the HA tag (to detect CDC48-HA proteins), Tic110 (a chloroplast marker), calreticulin (an ER marker), and H3 (a nuclear marker); the last two confirmed the purity of the isolated chloroplasts. Bands from the experiment shown, and those from two additional similar experiments, were quantified using Aida software, and the values obtained for the CDC48-HA protein were normalized using corresponding Tic110 data (C). The presented data are means  $\pm$  SEM (n = 3). Overall, the results showed that CDC48 is partially localized in chloroplasts, and that this association is more pronounced when CDC48 substrate binding is stabilized.



**Fig. S13. Expression of CDC48-DN triggers ROS overaccumulation.**

(A and B) Sequences encoding wild-type and dominant negative (DN) mutant (AtCDC48A<sup>K254A, K527A</sup>) CDC48, each with a C-terminal FLAG tag, were cloned into the pMDC7 binary vector (54) which enables estradiol-inducible transgene expression, and then the resultant constructs were used to stably transform wild-type plants. Using procedures similar to those described in the legend to fig. S3, representative, single-locus lines were selected for further analysis, based on segregation of hygromycin resistance and visible phenotype after induction. Immunoblotting using anti-FLAG antibody confirmed expression of the CDC48-FLAG proteins, and enabled selection of comparable CDC48-WT and -DN lines expressing similarly high CDC48-FLAG protein levels. Transformants grown under normal growth conditions for 7 days on MS agar medium were induced with estradiol for 2 days, and then stained with 3,3'-diaminobenzidine (DAB), to detect hydrogen peroxide accumulation. Representative images of stained plants are shown (A). Brown DAB staining (indicated by black arrowheads) was indicative of accumulation of the reactive oxygen species (ROS). The area of DAB staining was quantified using ImageJ software, and expressed as a percentage of the total surface area for each plant (B). Values shown are means  $\pm$  SEM ( $n = 6$ ). Similar to *spi* mutant plants grown under stress conditions (8), the CDC48-DN plants showed much higher ROS accumulation levels than control plants under normal growth conditions.



**Fig. S14. Proteasome- and CDC48-dependent regulation of SP1 abundance.**

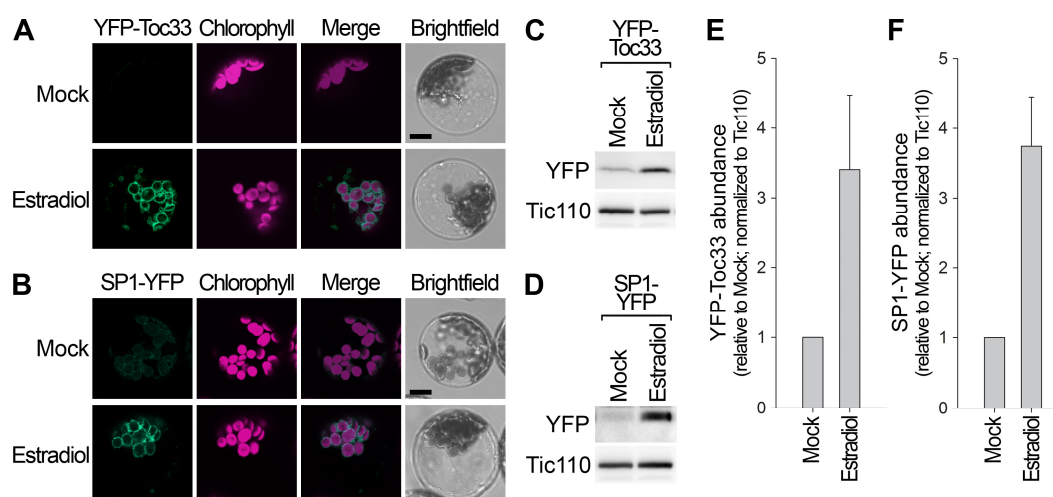
(A) Immunoblotting analysis of the effects of proteasome inhibitors on SP1 protein levels. *Arabidopsis* protoplasts that had been transfected with the *SP1-YFP* construct (7) were incubated in 1-10  $\mu$ M proteasome inhibitors (MG132, epoxomicin) or autophagy, cysteine protease inhibitor (E64) for 2 h. As MG132 and epoxomicin were dissolved in dimethylsulphoxide (DMSO), and E64 was dissolved in H<sub>2</sub>O, parallel incubations with each solvent alone were conducted as controls. Protein extracts of the cells were then analysed by immunoblotting using antibodies against: the YFP tag, to detect the SP1-YFP protein and assess its levels; and Tic110, as an internal loading control. Bands from the experiment shown, and those from two additional similar experiments, were quantified using Aida software, and the values obtained for SP1-YFP protein were normalized using corresponding Tic110 data. The presented data have been expressed relative to the corresponding control value for each treatment, and are means  $\pm$  SEM (n = 3). The SP1-YFP protein accumulated specifically after incubation with proteasome inhibitors, suggesting that SP1-YFP is degraded by the UPS and not by autophagy.

(B) Fluorescence microscopy analysis of the effects of proteasome inhibitors on SP1 protein levels. As in A, *SP1-YFP*-transfected protoplasts were incubated in 10  $\mu$ M MG132, epoxomicin or DMSO (as a control), and then the cells were analysed by fluorescence microscopy. Chlorophyll autofluorescence was employed to determine the localization of the YFP fluorescence signals relative to the chloroplasts, revealing clear chloroplast envelope localization of SP1-YFP in each case. The intensity of the envelope-localized YFP signals increased markedly upon treatment with proteasome inhibitors, indicating that SP1-YFP is degraded by the UPS at the envelope of chloroplasts. Brightfield images confirmed the intactness of the protoplasts. Scale bar, 10  $\mu$ m.

(C) Analysis of the combined effects of proteasome and CDC48 inhibition on SP1 protein levels.

Protoplasts from CDC48-WT and CDC48-DN plants were transfected with the *SP1-HA* construct (7) and

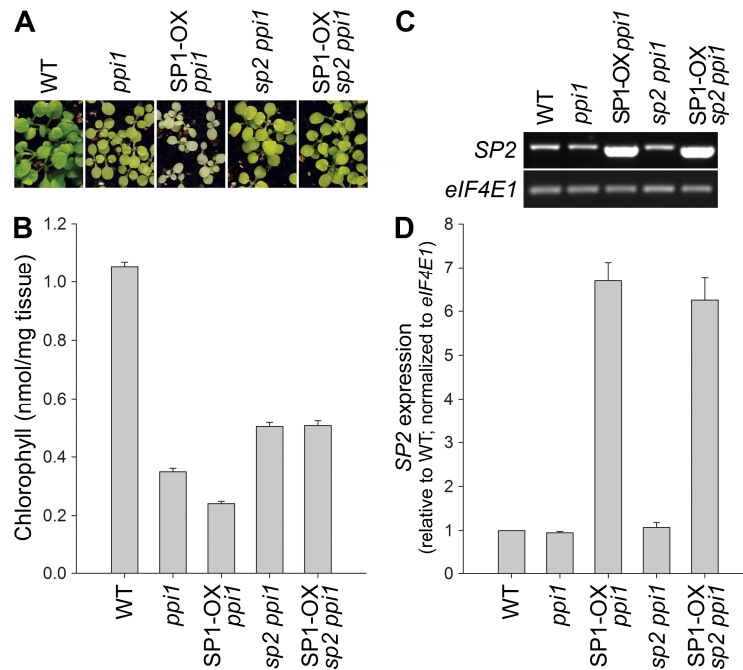
incubated for ~15 h with 10  $\mu$ M estradiol to induce CDC48 transgene expression. The protoplasts were further incubated in the absence (Mock) or presence of the proteasome inhibitor bortezomib for 2 h, before analysis by immunoblotting using antibodies against: the HA tag, to detect SP1-HA protein; and Tic40, as an internal loading control. Bands from the experiment shown, and those from four additional similar experiments, were quantified using Aida software, and the values obtained for SP1-HA protein were normalized using corresponding Tic40 data. The presented data are means  $\pm$  SEM (n = 5). The results suggest that the combined inhibition of both the proteasome and CDC48 does not have an additive effect on SP1 protein levels. The results suggested that CDC48 and the proteasome work in the same pathway to control SP1 protein stability.



**Fig. S15. Expression of CDC48-DN triggers the accumulation of OEM proteins at the chloroplast envelope in vivo.**

(A and B) Protoplasts from CDC48-DN plants were transfected with constructs (7, 21) encoding one of two different chloroplast outer membrane proteins, SP1-YFP (A) and YFP-Toc33 (B), either with (10  $\mu$ M estradiol) or without (Mock) the induction of CDC48-DN expression. The protoplasts were incubated for 15-18 h before analysis by confocal microscopy. The YFP signal intensity for each construct was greatly enhanced by the expression of CDC48-DN, confirming the role of CDC48 in the degradation of chloroplast proteins as shown in Figure 3. Chlorophyll autofluorescence was employed to determine the localization of the YFP fluorescence signals relative to the chloroplasts, revealing that the majority of the YFP signal in each case was localized at the chloroplast envelope. Brightfield images confirmed the intactness of the protoplasts. Scale bars, 10  $\mu$ m. These results support a role for CDC48 in degrading chloroplast-resident proteins, rather than un-imported chloroplast proteins that may exist in the cytosol.

(C to F) Protoplasts similar to those described above (in A and B) were analysed by immunoblotting using antibodies against: the YFP tag, to detect YFP-Toc33 and SP1-YFP protein; and Tic110, as an internal loading control (C and D). Bands from the experiments shown, and those from two additional similar experiments in each case, were quantified using Aida software, and the values obtained for the YFP fusion proteins were normalized using corresponding Tic110 data (E and F). The presented data have been expressed relative to the corresponding control value for each treatment, and are means  $\pm$  SEM ( $n = 3$ ).



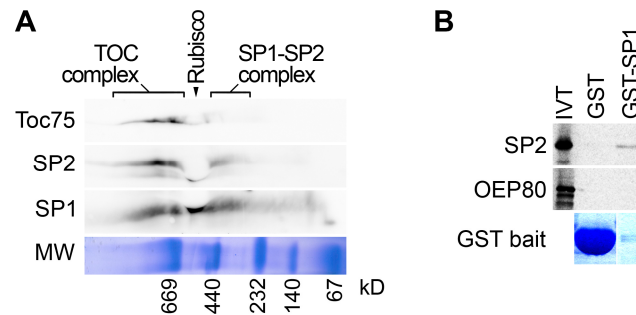
**Fig. S16. The function of SP1 depends on SP2.**

(A and B) The *sp2* mutation was introduced into the SP1-OX *ppi1* background (7) by crossing the latter with the *sp2-4 ppi1* double mutant. A resultant F<sub>2</sub> population was grown on phosphinothricin plates to select for the *SP1-OX* construct, and the phosphinothricin-resistant plants were PCR genotyped to identify those that were *sp2-4* homozygous (table S1). Individual F<sub>3</sub> families were grown on phosphinothricin plates to select those that were homozygous for the *SP1-OX* construct. The SP1-OX *sp2-4 ppi1* plants thus identified were then, together with corresponding control genotypes, grown on MS agar medium for 10 days before photography (A) and photometric chlorophyll content analysis following DMF extraction (B). About 10 seedlings per genotype were measured in each experiment, and the values shown are means  $\pm$  SEM derived from four experiments per genotype, and indicate nmol total chlorophyll per mg tissue fresh weight. As reported before (7), SP1-OX enhanced the chlorotic phenotype of *ppi1* by making the plants even paler. This effect of SP1-OX on *ppi1* was abolished by the *sp2-4* mutation, indicating that SP1 action is dependent on SP2.

(C and D) Semi-quantitative RT-PCR analysis of *SP1* expression in the plants described above (in A and B). Total RNA samples isolated from 10-day-old plants were analysed by RT-PCR using gene-specific primers for *SP1* and the reference gene *eIF4E1* (table S1). Amplifications employed a limited number of cycles to avoid saturation, and products were analysed by agarose gel electrophoresis and staining. A representative gel image is shown (C), along with quantification of three biological replicates (D). The amplicon bands were quantified using Aida software, and the data obtained for *SP1* were normalized

relative to equivalent data for *elF4E1*. Values shown are means  $\pm$  SEM (n = 3). The data confirmed that *SP1* was expressed at a similarly high level in the SP1-OX *ppi1* and SP1-OX *sp2-4 ppi1* plant lines.

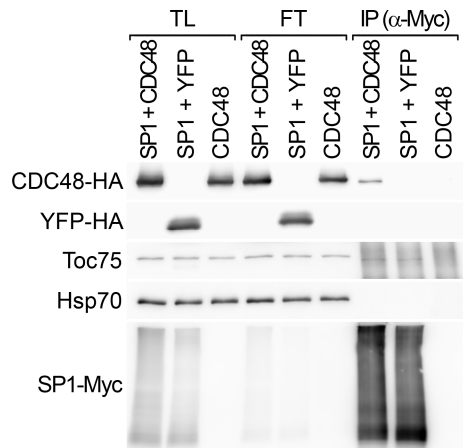




**Fig. S17. The SP1 and SP2 proteins associate to form a complex.**

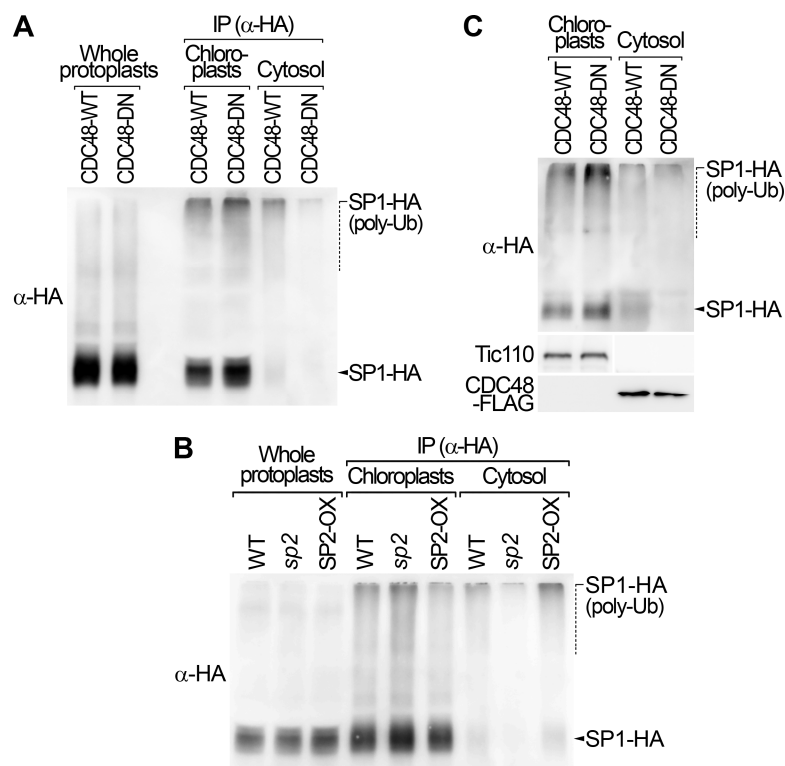
(A) Two-dimensional (2D)-blue native (BN)/SDS-PAGE analysis of SP1 and SP2. The SP2 CDS was fused with sequence encoding a C-terminal 6×Myc tag and cloned downstream of the 35S promoter in the p2GW7 vector. The resulting *SP2-Myc<sup>a</sup>* construct, together with the *SP1-HA* construct (7), were used to co-transform *Arabidopsis* protoplasts. Chloroplasts were isolated from the transformed cells as described previously (59), solubilized in BN-PAGE sample buffer containing 1% DDM, and subjected to 2D-BN/SDS-PAGE followed by immunoblotting using the indicated antibodies. Molecular weights (MW) of standards analysed in the first dimension are shown at the bottom in kD. The SP1-HA and SP2-Myc proteins co-migrated at two different positions: the first (>669 kD) overlapped with the TOC complex (85, 86) (and may indicate an intermediate step preceding TOC component degradation); the second (232-440 kD) may correspond to a resting or inactive core complex for CHLORAD (see fig. S20).

(B) In vitro pull-down analysis of the association between SP1 and SP2. In vitro translated (IVT), <sup>35</sup>S-radiolabelled SP2 or OEP80 (as a control OEM protein) were used as “prey” in pull-down assays with bacterially-expressed, purified GST-SP1 flex or an excess of GST (negative control) as the “baits”. Eluted GST proteins, along with any associated partner proteins, were resolved by SDS-PAGE and then analysed by phosphorimaging (to detect the radiolabelled “prey” proteins) or Coomassie Brilliant Blue staining (to detect the GST “baits”). The result showed that SP1 specifically associates with SP2 in vitro.



**Fig. S18. Analysis of the interaction between SP1 and CDC48 by co-immunoprecipitation.**

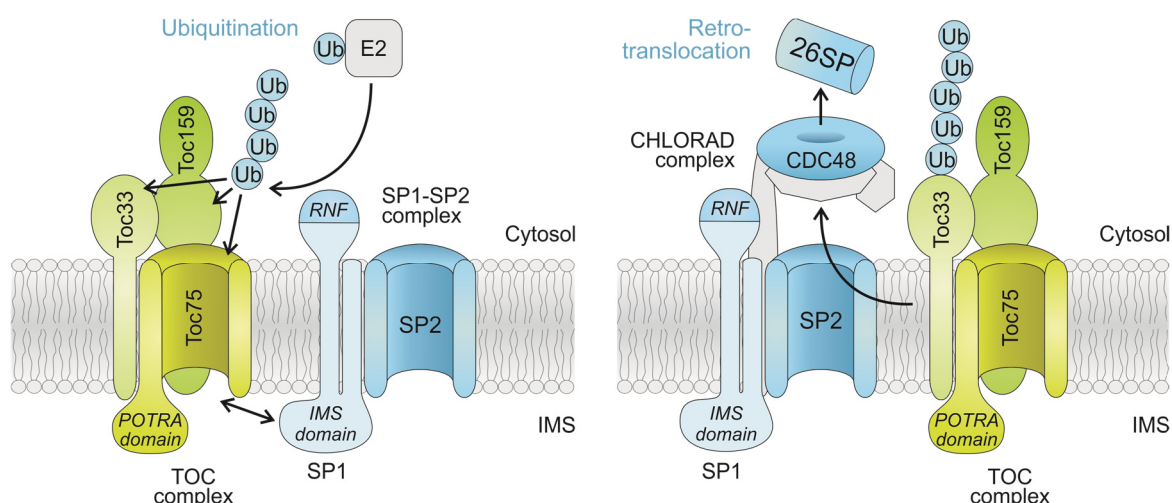
The SP1 CDS was fused with sequence encoding a C-terminal 6×Myc tag and cloned downstream of the 35S promoter in the p2GW7 vector. The resulting *SP1-Myc* construct was used, together with the *CDC48-HA* construct or a *YFP-HA* control construct (7), to co-transform *Arabidopsis* protoplasts. In parallel, the *CDC48-HA* construct was transformed alone, to confirm that CDC48-HA does not non-specifically associate with the binding matrix (agarose beads). The cells were analysed by immunoprecipitation using anti-c-Myc antibody. Total lysis (TL; before immunoprecipitation was initiated), flow-through (FT), and the immunoprecipitated (IP) samples were then analysed by immunoblotting using antibodies against: the Myc tag, to verify the enrichment of SP1-Myc; the HA tag, to assess for interaction of SP1-Myc with CDC48-HA or the YFP-HA negative control; Toc75, to detect an SP1-substrate interaction; and Hsp70, to act as a negative control and confirm that the detected interactions are specific.



**Fig. S19. CDC48 and SP2 are required for the retrotranslocation of polyubiquitinated SP1.**

(A and B) Analysis of the roles of CDC48 and SP2 in the retrotranslocation of SP1-HA substrate using an in vivo retrotranslocation assay. The data were generated as shown in Fig. 8 (A and C, respectively) except that SP1 was detected by immunoblotting using anti-HA only, even for the high molecular weight species. These results provided confirmation that the high molecular weight bands were indeed polyubiquitinated SP1.

(C) Analysis of the role of CDC48 in the retrotranslocation of SP1-HA substrate using an in vitro retrotranslocation assay. The data were generated as shown in Fig. 8E, except that SP1 was detected by immunoblotting using anti-HA only (after immunoprecipitation enrichment), even for high molecular weight species. The Tic110 and CDC48-FLAG control proteins were detected directly by immunoblotting (i.e., before the immunoprecipitation enrichment step, using the remaining 10% of protein extract). These results provided confirmation that the high molecular weight bands were indeed polyubiquitinated SP1.



**Fig. S20. Model for the roles of SP2 and CDC48 in the UPS-mediated regulation of chloroplast proteins.**

The model presents a system of chloroplast proteolysis via the UPS, which we designate Chloroplast-Associated Protein Degradation (CHLORAD). This system removes specific chloroplast OEM substrates (amongst which TOC components of the chloroplast protein import machinery are well-studied examples), for the regulation of plastid development and functions. Currently three core components of the CHLORAD system have been identified: SP1, a chloroplast OEM-anchored ubiquitin E3 ligase which is responsible for the ubiquitination of substrates; SP2, a chloroplast OEM-embedded  $\beta$ -barrel channel protein which is involved in the retrotranslocation of ubiquitinated substrates; and CDC48, a cytosolic AAA<sup>+</sup>-ATPase chaperone protein that associates peripherally with the chloroplast OEM and is also involved in retrotranslocation, providing a driving force for substrate extraction from the membrane. The SP1 protein possesses a cytosolic RING Finger domain (RNF; responsible for recruiting the E2 enzyme), a large intermembrane space domain (IMS; which binds to the substrate proteins), and two transmembrane domains. Substrate ubiquitination by SP1 is an early step in the process (left side); we propose that this is mediated by an SP1-SP2 complex to enable a close link between ubiquitination and retrotranslocation. Thereafter, cytosolic CDC48 is recruited to the chloroplast (right side), either by binding to the polyubiquitin chain with the assistance of cofactors or by employing a putative chloroplast OEM tethering factor (shown in grey). The chaperone then drives substrate extraction (retrotranslocation), overcoming the energetic barrier to the removal of proteins from the chloroplast membrane. When substrates are finally released into the cytosol, degradation by the 26S proteasome (26SP) ensues.

**Supplementary Table S1. Primers used during the course of the study.**

(A) Primers used to generate various constructs and mutations.

Primer name	Primer sequence (5' to 3')*	Used to generate ...
CDC48-CDS-F	AAAAAGCAGGCTCCATGTCTACCCCAGCTGAAT	... <i>CDC48</i> CDS for plant transformation, and C-terminal fusions
CDC48-CDS-R-FLAG	AGAAAGCTGGGTTCTATTTGTCGTCATCGTCTTTGTAGTCATTGTAGAGATCATCATCGT	... <i>CDC48</i> CDS fused with C-terminal FLAG for plant transformation
CDC48-nonstop-R	AGAAAGCTGGGTTATTGTAGAGATCATCATCGTCC	... <i>CDC48</i> CDS for C-terminal fusions
CDC48-E581Q-F	CCATGTGTTCTTTTCTTTGATCAGCTCGACTCC	... CDC48 Trap mutation
CDC48-E581Q-R	GGAGTCGAGCTGATCAAAGAAAAGAACACATGG	
SP1-BamHI-F	GGATCCATGATTCCTTGGGGTGGAG	... <i>SP1</i> CDS for cloning into pE3c vector
SP1-ns-NotI-R	GCGGCCGCGTGACGATATGTCTTAACC	
SP2-BclI-F	GGGTAACTCTGATCAATGGGAGCTCAGAAGAGTATCCA	... <i>SP2</i> CDS for cloning into pE3c vector
SP2-ns-NotI-R	AAACCGGTCTGCGGCCGCTGTTGATGAAGCAAGATTGGTG	
SP2-CDS-F	AAAAAGCAGGCTCCATGGGAGCTCAGAAGAGTATCCA	... <i>SP2</i> CDS for plant transformation
SP2-CDS-R	AGAAAGCTGGGTTTGTGATGAAGCAAGATTGGTG	
Toc33-CDS-F	AAAAAGCAGGCTCCATGGGGTCTCTCGTTCGTG	... <i>TOC33</i> CDS for cloning into NTAPi vector
Toc33-CDS-R	AGAAAGCTGGGTTTAAAGTGGCTTTCCACTTG	
CDC48-KpnI-F	AAGGTACCATGTCTACCCCAGCTGAATC	... <i>CDC48</i> CDS for cloning into pSAT4A-cEYFP-N1 vector
CDC48-ns-XmaI-R	AACCCGGGATTGTAGAGATCATCATCGTCC	
Toc159-EcoRI-F	AAGAATTCATGGACTCAAAGTCGGTT	... <i>TOC159</i> CDS for cloning into pSAT4-nEYFP-C1 vector
Toc159-SalI-R	AAGTCGACTTAGTACATGCTGTACTT	
OEP7-EcoRI-F	AAGAATTCATGGGAAAACTTCGGGA	... <i>OEP7</i> CDS for cloning into pSAT4A-nEYFP-N1 vector
OEP7-ns-SalI-F	TTGTCGACACAAACCCTCTTTGGATGT	
CDKA1-XhoI-F	CTCGAGATGGATCAGTACGAGAAAG	

CDKA1-ns-EcoRI-R	<b>GAATTC</b> AGGCATGCCTCCAAGATC	... <i>CDKA1</i> CDS for cloning into pSAT4A-cEYFP-N1
SP2-KpnI-F	<b>GGTACC</b> ATGTCTTTAATGTTTCCTGCTTTCAGGG	... <i>SP2</i> CDS for cloning into pBluescript SK <sup>-</sup> vector
SP2-R	CTATGTTGATGAAGCAAGATTGG	
OEP80-KpnI-F	<b>GGTACC</b> ATGCATTGTCACAACGATG	... <i>OEP80</i> CDS for cloning into pBluescript SK <sup>-</sup> vector
OEP80-R	TTAGTTCCGCAGACCAAC	
attB1	GGGGACAAGTTTGTACAAAAAGCAGGCT	... complete Gateway recombination sites for cloning into pDONR201
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGT	

**(B)** Primers used in RT-PCR experiments.

Primer name	Primer sequence (5' to 3')	Comments
SP2-RT1-F	GGTTGTGTCCAAGTGGCTTA	Used in Supplementary Fig. S1
SP2-RT1-R	TACAGCTTCTCCTTGGACTGT	
SP2-RT2-F	GGGTTCTCTTCATGGTTGATTCTTCTCT	
SP2-RT2-R	CCATTCAGGTCTAGGTCTCCTAAAA	
SP2-F2	GGTTGTGTCCAAGTGGCTTA	Used in Supplementary Fig. S10
SP2-R2	CTATGTTGATGAAGCAAGATTGGTG	
Toc159-QRT-F	AACTCTTGAAGTGGCTAATAAGT	
Toc159-QRT-R	ACAACCTCTGGCTCTACA	
Toc33-QRT-F	AATGGTGAAGCGTGGATC	
Toc33-QRT-R	TGCTCCTTGAATCATCTTAACG	
Toc75-III-QRT-F	TCGCATCTCCACTCAATC	
Toc75-III-QRT-R	GTCTCTGTATCTCGGTTAGG	
GAPDH-F	GTGGTGGATTTGGCTCACCT	
GAPDH-R	CTCATCAGCCGGGTTTGTCT	
SP1-F	GGTACAAGATAGTGC GTTGATG	
SP1-R	CTGCAGTCAGTGACGATATGTCTTAAC	

ACTIN2-F	TCAGATGCCCAGAAGTCTTGTTCC	
ACTIN2-R	CCGTACAGATCCTTCCTGATATCC	
eIF4E1-F	AAACAATGGCGGTAGAAGACACTC	
eIF4E1-R	AAGATTTGAGAGGTTTCAAGCGGTGTAAG	

(C) Primers used to genotype mutant plants by dCAPS (derived Cleaved Amplified Polymorphic Sequence) analysis.†

Primer name	Mutant	Primer sequence (5' to 3')*	Restriction enzyme
SP2-dC1-F	<i>sp2-1</i>	TGTTGCGGAATTGGTTTCA <b>T</b>	Enzyme: DdeI
SP2-dC1-R		GGTTGTGTCCAAGTGGCTTA	Digests: WT
SP2-dC2-F	<i>sp2-2</i>	CATTGGTGGGCTAGGCAGTG <b>ATC</b>	Enzyme: BclI
SP2-dC2-R		CATGCAACAGCTCCACGTACCAA	Digests: <i>sp2-2</i>
SP2-dC3-F	<i>sp2-3</i>	CCACAATCCAAGAAGATGGT <b>AC</b>	Enzyme: KpnI
SP2-dC3-R		GCCTTGTTGCCAATACAGAA	Digests: WT
Toc75-III-3-dCAPS-F	<i>toc75-III-3</i>	GCAAATCACAGGTGGAT <b>CTG</b>	Enzyme: DdeI
Toc75-III-3-dCAPS-R		TGCCCATGGAGGACTAGAAC	Digests: <i>toc75-III-3</i>

(D) Primers used to genotype T-DNA insertion mutant plants.

Primer name	Primer sequence (5' to 3')	Comments
Lba1	TGGTTCACGTAGTGGGCCATCG	Lba1 and SP2-R3 to identify the T-DNA; SP2-F3 and SP2-R3 to identify the SP2 gene region flanking T-DNA
SP2-R3	TACAGCTTCTCCTTGGACTGT	
SP2-F3	CGTGGAGCTGTTGCATGAT	

\*Nucleotides shown in red do not correspond to the target gene, but instead correspond to linker sequences or mutation sites.

†The PCR products amplified using dCAPS primers were digested with restriction enzyme, and thereafter resolved on 3% agarose gels. Genotypes were determined by comparing the sizes of the bands with controls corresponding to wild type and the homozygous mutant.