

# Molecular and genetic analyses of Tic20 homologues in *Arabidopsis thaliana* chloroplasts

Ali Reza Kasmati<sup>1</sup>, Mats Töpel<sup>2</sup>, Ramesh Patel<sup>1</sup>, Ghulam Murtaza<sup>1,†</sup>, and  
Paul Jarvis<sup>1,\*</sup>

<sup>1</sup> Department of Biology, University of Leicester, University Road, Leicester LE1 7RH, United  
Kingdom, and

<sup>2</sup> Department of Plant and Environmental Sciences, Göteborg University, Box 461, SE-405 30  
Göteborg, Sweden.

\* For correspondence (fax +44-116-252-3330; tel +44-116-223-1296; e-mail rpj3@le.ac.uk).

† Present address: Department of Botany, City Campus, The University of AJK, Muzaffarabad 13100,  
Pakistan.

Running Title: *Arabidopsis tic20* mutants

Keywords: *Arabidopsis*, chloroplast protein import, protein targeting, Tic20, TOC/TIC machinery,  
translocon

Number of words: 6994

Number of data displays: 9 figures, 1 table, 6 supplementary figures, 2 supplementary tables.

**SUMMARY**

The Tic20 protein was identified in pea as a chloroplast protein import apparatus component. In Arabidopsis, there are four Tic20 homologues termed atTic20-I, atTic20-IV, atTic20-II, atTic20-V, all with predicted topological similarity to the pea protein (psTic20). Analysis of Tic20 sequences from many species indicated that they are phylogenetically unrelated to mitochondrial Tim17-22-23 proteins, and that they form two evolutionarily-conserved subgroups (characterized by psTic20/atTic20-I/IV [Group 1] and atTic20-II/V [Group 2]). Like psTic20, all four Arabidopsis proteins have a predicted transit peptide consistent with targeting to the inner envelope. Envelope localization of each one was confirmed by analysis of YFP fusions. RT-PCR and microarray data revealed that the four genes are expressed throughout development. To assess functional significance of the genes, T-DNA mutants were identified. Homozygous *tic20-I* plants had an albino phenotype that correlated with abnormal chloroplast development and reduced levels of chloroplast proteins. However, knockouts for the other three genes were indistinguishable from wild type. To test for redundancy, double and triple mutants were studied; apart from those involving *tic20-I*, none was distinguishable from wild type. The *tic20-I tic20-II* and *tic20-I tic20-V* doubles were albino, like the corresponding *tic20-I* parent. In contrast, *tic20-I tic20-IV* double homozygotes could not be identified, due to gametophytic and embryonic lethality. Redundancy between atTic20-I and atTic20-IV was confirmed by complementation analysis. Thus, atTic20-I and atTic20-IV are the major functional Tic20 isoforms in Arabidopsis, with partially overlapping roles. While the Group 2 proteins have been conserved over ~1.2 billion ( $1.2 \times 10^9$ ) years, they are not essential for normal development.

## INTRODUCTION

Most chloroplast proteins are translated on cytosolic ribosomes and imported into plastids (Inaba and Schnell, 2008; Jarvis, 2008; Kessler and Schnell, 2006; Smith, 2006; Soll and Schleiff, 2004). They are synthesized as precursors with N-terminal extensions called transit peptides, and are post-translationally imported after binding to the chloroplast outer envelope membrane. The cleavable transit peptide is essential for chloroplast targeting and translocation across the envelope membranes. Import is mediated by multiprotein complexes in the outer and inner envelope membranes, termed TOC and TIC (Translocon at the Outer/Inner envelope membrane of Chloroplasts), respectively.

Biochemical studies using pea chloroplasts identified several TOC and TIC components. The core TOC system consists of preprotein receptors (Toc159 and Toc34) and a transmembrane channel protein (Toc75). While several TIC components have been identified (Tic110, Tic62, Tic55, Tic40, Tic32, Tic22, Tic21 and Tic20), their roles are not well understood (Inaba and Schnell, 2008; Jarvis, 2008; Kessler and Schnell, 2006; Smith, 2006; Soll and Schleiff, 2004). Arguably the most crucial function of the TIC complex is channel formation, and yet there is uncertainty about the identity of the relevant component(s); it has been suggested that Tic20 (Kouranov *et al.*, 1998), Tic110 (Balsera *et al.*, 2009; Heins *et al.*, 2002), and Tic21 (Teng *et al.*, 2006) may each perform this role. Tic20 and Tic21 share topological similarity and complementary expression patterns, and so it was initially proposed that the former operates during early plant development with the latter taking over later on (Teng *et al.*, 2006). Nonetheless, both proteins have been detected together in a 1 MDa complex, and it is suggested that this corresponds to the core, channel-forming TIC complex (Kikuchi *et al.*, 2009). Interestingly, Tic110 was absent from this assembly, suggesting that it acts later by serving as a scaffold to coordinate stromal chaperones (*e.g.*, Hsp93) that bind to preproteins as they emerge from the translocon (Inaba *et al.*, 2003). Such chaperones may drive translocation and/or facilitate protein folding following transit peptide cleavage.

The pea Tic20 (psTic20) protein was identified by chemical crosslinking and shown to be close to preproteins engaged in import at an intermediate stage (Kouranov and Schnell, 1997; Ma *et al.*, 1996). It was initially proposed to have three transmembrane  $\alpha$ -helices, but it now seems clear that there

1  
2  
3 are four such domains (Chen *et al.*, 2002; Kouranov *et al.*, 1998). The predicted topology of Tic20  
4 makes it a good candidate for the protein-conducting channel. Initial phylogenetic studies suggested a  
5 distant relationship between Tic20 and bacterial branched-chain amino acid transporters (Reumann *et*  
6 *al.*, 1999). Because such a relationship was also reported to exist between these transporters and channel  
7 proteins of the mitochondrial protein import machinery (Tim22 and Tim23) (Rassow *et al.*, 1999), the  
8 notion of a channel-forming role for Tic20 was further supported. Nonetheless, there is a lack of *in vitro*  
9 evidence to substantiate this proposal (Soll and Schleiff, 2004).  
10  
11  
12  
13  
14  
15  
16  
17

18 Homologues of Tic20 have been identified in different species, including cyanobacteria, red  
19 and green algae, bryophytes, and the apicomplexan parasite, *Toxoplasma gondii* (Kalanon and  
20 McFadden, 2008; van Dooren *et al.*, 2008). *Toxoplasma gondii* Tic20 is an integral component of the  
21 innermost membrane of the complex plastid found in this organism (the apicoplast; the relevant  
22 membrane is equivalent to the inner envelope membrane of plant chloroplasts), and was predicted to be  
23 structurally similar to psTic20. Furthermore, analysis of a null mutant revealed that *T. gondii* Tic20 is  
24 essential for parasite viability, and that its loss severely impairs apicoplast protein import (van Dooren  
25 *et al.*, 2008).  
26  
27  
28  
29  
30  
31  
32  
33  
34

35 Four different Tic20-related sequences were reported to exist in the Arabidopsis genome, and  
36 named according to their chromosomal locations: atTic20-I, atTic20-IV, atTic20-II and atTic20-V  
37 (Bédard and Jarvis, 2005; Kalanon and McFadden, 2008). Protein sequence identities shared between  
38 these four and psTic20 were reported to be 62%, 35%, 25% and 25%, respectively, across the aligning  
39 regions (Bédard and Jarvis, 2005). Thus, while atTic20-I is likely the direct functional orthologue of  
40 psTic20, roles in protein import for the other Tic20 homologues cannot be excluded. Antisense  
41 technology was employed to elucidate the role of atTic20-I *in vivo* (Chen *et al.*, 2002). The antisense  
42 plants exhibited a pale-green phenotype, growth defects, underdeveloped chloroplasts, and reduced  
43 levels of chloroplast proteins. Protein import studies using chloroplasts isolated from atTic20-I  
44 antisense plants revealed a defect in preprotein translocation across the inner envelope (Chen *et al.*,  
45 2002). More recently, complete knockout of atTic20-I was reported to cause albinism (Teng *et al.*,  
46 2006), and this was correlated with an import defect with apparent specificity for photosynthesis-related  
47 preproteins (Kikuchi *et al.*, 2009).  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 In this study, we analysed all four Arabidopsis Tic20-related sequences with the aim of  
4 elucidating their phylogenetic and functional interrelationships.  
5  
6  
7  
8  
9

## 10 11 12 13 14 **RESULTS**

### 15 16 17 18 **Phylogenetic analysis of Tic20 homologues in Arabidopsis**

19  
20  
21  
22 According to our database searches, there are four Tic20 homologues in Arabidopsis (atTic20-I  
23 [At1g04940], atTic20-IV [At4g03320], atTic20-II [At2g47840] and atTic20-V [At5g55710]). Full-  
24 length, sequenced cDNA clones are available for each gene (accession numbers AK117165, AF361633,  
25 AY050346 and BT001996, respectively), confirming that they are all expressed. Protein sequences  
26 predicted using these cDNAs were analysed *in silico*. TargetP predicted that all four Tic20 homologues  
27 have a transit peptide with high confidence (Emanuelsson *et al.*, 2000), which is consistent with  
28 localization to the inner envelope membrane (Hofmann and Theg, 2005). Moreover, analysis using  
29 TMHMM indicated that the four proteins are topologically similar to psTic20 (Krogh *et al.*, 2001); *i.e.*,  
30 they each have four predicted transmembrane domains, and these are similarly located (Figure S1).  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

42 To gain insight into the evolutionary (and, by extension, functional) relationships between the  
43 four Arabidopsis proteins, the pea protein, and related proteins in other species, we conducted a  
44 phylogenetic analysis. Extensive database searches retrieved 59 sequences from 32 species, which were  
45 then analysed in a Bayesian framework. The resulting tree (Figure 1) shows that a gene duplication  
46 happened early in the evolution of eukaryote Tic20 proteins, resulting in two clades: Group 1 (including  
47 psTic20, atTic20-I and atTic20-IV) and Group 2 (including atTic20-II and atTic20-V), with  
48 representatives from red algae and viridiplantae in both groups (the position of the three *Phaeodactylum*  
49 and *Thalassiosira* proteins in Group 2 implies that there are two Tic20 types in the chromalveolates as  
50 well [see asterisks], although support for this is low and inconclusive). To assess whether the gene  
51 duplication happened in the eukaryotic lineage or in a pre-endosymbiotic organism, we conducted an  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 additional phylogenetic analysis of the proteins in Figure 1 together with an extensive set of sequences  
4 from cyanobacteria (data not shown). This showed that cyanobacterial proteins do not belong in either  
5 Group 1 or Group 2, and from this we conclude that the duplication happened after the endosymbiotic  
6 event leading to the formation of chloroplasts ~1.5 billion years ago (Ga), but before the split into red  
7 algae and viridiplantae (~1.2-1.5 Ga) (Butterfield, 2000; Reyes-Prieto *et al.*, 2007; Yoon *et al.*, 2004).  
8  
9

10  
11  
12 It is noteworthy that many plant and algal species are represented in Group 1 and Group 2  
13 (*e.g.*, red and green algae, bryophytes and angiosperms), as this suggests that the proteins in both clades  
14 play an important functional role, having been maintained over millions of years of evolution. Also  
15 worth noting is the long branch separating the land plant proteins (the clade including *Picea*,  
16 *Physcomitrella* and the angiosperms) from the rest of the clade, in Group 1 but not in Group 2 (see  
17 arrows in Figure 1). This indicates that an ancestral protein of Group 1 evolved much faster than the  
18 homologue in Group 2, and that this evolutionary burst happened prior to, or during, adaptation to land  
19 life in plants. Interestingly, the two major functional isoforms in *Arabidopsis* (atTic20-I and atTic20-IV;  
20 see later) as well as psTic20 are derived from this ancestral protein (Figure 1).  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

33  
34 A phylogenetic analysis of Tic20-related sequences was reported previously (Kalanon and  
35 McFadden, 2008), but this did not include such a large number of sequences and so was unable to reach  
36 the three main conclusions of our analysis, which can be summarized as follows: (1) Group 1 and  
37 Group 2 proteins separated ~1.2-1.5 Ga, after the evolution of chloroplasts; (2) red algae and  
38 viridiplantae, and possibly also chromalveolates, are represented in both groups; and (3) rapid evolution  
39 has occurred in Group 1.  
40  
41  
42  
43  
44  
45

46  
47 Despite earlier reports indicating a phylogenetic relationship between Tic20, bacterial  
48 branched-chain amino acid transporters, and the mitochondrial Tim17-22-23 proteins (Rassow *et al.*,  
49 1999; Reumann *et al.*, 1999), we could find no such link. Attempts to align the Tic20 dataset with these  
50 proteins (to include them in our phylogenetic analyses) failed, as homologous characters in the different  
51 datasets could not be identified. Nonetheless, we tested the proposed similarity as follows. First, we  
52 manually aligned a representative set of Tic20 and Tim17-22-23 proteins to the prokaryotic sequences,  
53 and adjusted the alignment according to that presented by Rassow *et al.* (1999). We then created four  
54 consensus sequences, one for each of the following groups: atTic20/psTic20, atTim17-22-23, scTim17-  
55  
56  
57  
58  
59  
60

1  
2  
3 22-23, and the prokaryote proteins (Figure S2). Consensus sequences are useful when comparing  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

22-23, and the prokaryote proteins (Figure S2). Consensus sequences are useful when comparing  
distantly-related proteins as they only show conserved positions less prone to accumulate mutations;  
*i.e.*, functional sites or positions required for maintenance of three-dimensional structure (Lesk and  
Fordham, 1996). Our analysis shows that no position shares the same residue in all four consensus  
sequences, and that the Tic20 sequence does not match with the prokaryotic sequence in any position.  
Therefore, we conclude that the similarity previously reported between Tic20, the prokaryotic branched-  
chain amino acid transporters, and Tim17-22-23 is more likely the result of convergent evolution than  
evidence of homology.

### Analysis of Tic20 homologue localization

To provide experimental support for the TargetP predictions, the subcellular localization of the  
Arabidopsis Tic20 homologues was assessed by analysing YFP fusion proteins. Full-length coding  
sequences for each of the four genes were inserted into a vector that adds a C-terminal YFP tag.  
Arabidopsis protoplasts were then transfected using these constructs, and analysed by fluorescence  
microscopy (Figure 2). The red fluorescent signal of chlorophyll shows the location of the chloroplasts  
in each case. For all four fusion proteins, the yellow-green fluorescence of YFP was observed in a ring-  
like pattern around the periphery of each chloroplast. This pattern was strongly reminiscent of the  
distributions seen for atTic110:YFP and atTic40:YFP fusion proteins in a previous study (Bédard *et al.*,  
2007), indicating that the Arabidopsis Tic20 homologues are all targeted to the chloroplast envelope.  
Considering that most outer envelope proteins do not possess a transit peptide (as the four Arabidopsis  
Tic20 homologues do) (Hofmann and Theg, 2005), and the aforementioned topological and  
phylogenetic similarities between the Arabidopsis proteins and psTic20, these results strongly support  
the hypothesis that Arabidopsis Tic20 homologues are localized in the chloroplast inner envelope  
membrane.

### Expression profiles of the Arabidopsis *TIC20* homologues

To gain insight into the functional relationships between the Arabidopsis Tic20 homologues, their developmental and tissue-specific gene expression patterns were studied by quantitative real time RT-PCR (Figure 3). The results indicated that *atTIC20-I* expression is highest in mature photosynthetic tissues (e.g., 14 dL and Rosettes; Figure 3), and relatively weak in non-photosynthetic tissues (5 dD and Roots; Figure 3). This suggests that *atTIC20-I* may be relatively more important for photosynthetic development, which is in-line with previous observations (Kikuchi *et al.*, 2009), including those made in relation to the dominant isoforms of the TOC receptors (Bauer *et al.*, 2000; Kubis *et al.*, 2003).

The expression pattern of *atTIC20-IV* was different, and somewhat complementary to that of the other Group 1 gene, with low levels in rosette leaves and relatively high levels in non-photosynthetic tissues (Figure 3). Thus, *atTIC20-IV* may be relatively more important for non-photosynthetic development. In general, this gene was expressed at lower levels than *atTIC20-I*, indicating that *atTic20-I* is the dominant isoform amongst the Group 1 proteins. Indeed, database searches using the BLAST program (Altschul *et al.*, 1990) detected 90 expressed sequence tags (ESTs) for *atTIC20-I* and just 25 for *atTIC20-IV*. Nonetheless, expression of *atTIC20-IV* did exceed that of *atTIC20-I* in roots (Figure 3).

The two Group 2 genes (*atTIC20-II* and *atTIC20-V*) were expressed at relatively high levels throughout development, and shared similar patterns of expression (Figure 3); EST numbers detected were 92 and 95, respectively. The expression profiles of *atTIC20-II* and *atTIC20-V* were both more similar to that of *atTIC20-I* than to that of *atTIC20-IV* (amongst the Group 1 genes); nonetheless, the photosynthetic *vs.* non-photosynthetic expression intensity differential was considerably less for *atTIC20-II* than for *atTIC20-V*, and so in this sense the profile of the former was relatively more similar to that of *atTIC20-IV*.

To corroborate the aforementioned conclusions, publicly-available microarray data were analysed using the Genevestigator tool (Zimmermann *et al.*, 2004). This revealed that three of the four Arabidopsis Tic20 genes are expressed throughout the plant's life-cycle (Figure S3); unfortunately, *atTIC20-I* could not be analysed due to the absence of a reliable probe-set from the Affymetrix

1  
2  
3 microarray (Teng *et al.*, 2006). The two Group 2 genes exhibited similar patterns of expression at  
4  
5 roughly equivalent levels. In agreement with the RT-PCR data, *atTIC20-IV* exhibited low levels of  
6  
7 expression in mature photosynthetic tissues, while markedly higher expression in mature siliques and  
8  
9 germinating seeds (neither of which was analysed by RT-PCR) suggested that *atTIC20-IV* may be  
10  
11 particularly important during seed development.  
12

13  
14 It should be borne in mind that relative transcript levels (as determined in these studies) do not  
15  
16 necessarily reflect protein abundances. Thus, in the future, when suitable isoform-specific antibodies  
17  
18 become available, it will be prudent to provide final confirmation of these expression patterns by  
19  
20 immunoblotting.  
21  
22  
23  
24  
25  
26

### 27 **Identification and analysis of Arabidopsis Tic20 T-DNA insertion mutants**

28  
29  
30

31 To elucidate the functional significance of the Arabidopsis Tic20 homologues *in vivo*, we identified two  
32  
33 independent T-DNA insertion mutants for each of the four genes. Firstly, all T-DNA insertion sites  
34  
35 were confirmed by genomic PCR, and by sequencing the T-DNA/gene junctions at one or both sides, as  
36  
37 indicated (Figure 4a). Segregation analysis was performed to ensure identification of only single-locus  
38  
39 insertion lines; a Mendelian ratio of three antibiotic-resistant plants to one antibiotic-sensitive plant  
40  
41 indicates the presence of a single T-DNA insertion (Table S2). Further segregation analysis identified  
42  
43 homozygous lines for analysis, and the zygosity of these was confirmed by genomic PCR (Figure S4).  
44  
45

46 The effect of each T-DNA insertion on Tic20 gene expression was assessed by RT-PCR  
47  
48 (Figure 4b). This confirmed that the relevant full-length mRNA was absent for all of the mutants except  
49  
50 *tic20-I-2*, *tic20-II-1* and *tic20-II-2*; these three displayed expression levels reduced to 29.0%, 8.4% and  
51  
52 58.0% of that in the wild type, respectively, and so are considered to be knockdown alleles (Figure 4c).  
53  
54 The other five mutants are all true knockout alleles.  
55  
56  
57  
58  
59  
60

### The *tic20-I* mutants exhibit albinism and severe defects in chloroplast biogenesis

Amongst the mutants identified, only *tic20-I-1* and *tic20-I-2* exhibited a phenotype that was obviously different from wild type. In populations segregating for these two mutations, significant numbers of individuals exhibited a striking albino-like appearance (Table S2). Co-segregation of the albino phenotypes with the T-DNA insertions was confirmed by PCR analysis; we genotyped 25 albinos and 38 green plants from *tic20-I-1* populations, and 38 albinos and 65 greens from *tic20-I-2* populations; all albinos were found to be homozygous mutant, whereas the greens were either heterozygous or wild type. The proportion of homozygous albino plants in such populations was frequently less than the 25% expected for normal inheritance (particularly for the knockout allele, *tic20-I-1*), suggesting that the homozygotes have reduced viability at an early developmental stage.

Plants carrying these two mutations are shown in Figure 5a; the *atToc159* knockout mutant, *plastid protein import 2 (ppi2)*, is shown for comparison purposes, as this mutant also exhibits an albino phenotype (Bauer *et al.*, 2000). Interestingly, we observed that *tic20-I-2* plants are slightly larger than *tic20-I-1* plants, and that they are somewhat greener in appearance (Figure 5a); this is consistent with the observation that *tic20-I-2* is a knockdown allele (Figure 4b). The *ppi2* mutant grew to a larger size than either of the *tic20-I* mutants, which may reflect the continued expression of related TOC receptor proteins in *ppi2* (Kubis *et al.*, 2004).

Impaired greening in the *tic20-I* mutants suggested defective chloroplast development, and so plastid ultrastructure was analysed by electron microscopy (Figure 6). In *tic20-I-1* plants, hardly any chloroplast development was observed: the plastids did not contain any thylakoid membranes, and a significant proportion were found to contain large inclusions or were surrounded by multilayered envelope membrane structures. By contrast, while chloroplasts in the *tic20-I-2* mutant were much smaller than those in wild type, they did exhibit significant thylakoid membrane development, especially after 14 days; this is consistent with the greener appearance of plants carrying this allele (Figure 5a). Taken together, these data support the notion that *atTic20-I* is essential for chloroplast biogenesis and development. Figure 6 also shows that the development of *ppi2* plastids is severely

1  
2  
3 disturbed, as was reported previously (Bauer *et al.*, 2000; Kubis *et al.*, 2004), although the defect is less  
4  
5 severe than that in *tic20-I-1* as inclusions and multilayered structures were not apparent.  
6

7  
8 To further characterize the chloroplast developmental defect in the *tic20-I* mutants, levels of  
9  
10 several chloroplast proteins were assessed by immunoblotting. These studies focused on components of  
11  
12 the protein import machinery (Figure 7a,c), and various components of the photosynthetic or  
13  
14 biochemical apparatus (Figure 7b,d). In each case, the visibly similar *ppi2* mutant was included as a  
15  
16 control. Surprisingly, levels of most of the translocon components investigated were not dramatically  
17  
18 affected by the *tic20* mutations. For the Toc proteins, this may reflect the fact that the TIC machinery is  
19  
20 not required for their biogenesis. On the other hand, the Tic110 and Tic40 proteins are predicted to  
21  
22 require a functional TIC apparatus for proper import and assembly, via the post-import pathway (Li and  
23  
24 Schnell, 2006; Tripp *et al.*, 2007). While these two Tic proteins are somewhat depleted in *tic20* plants,  
25  
26 the effect is not strong, and so it is possible that compensatory mechanisms (*e.g.*, reduced turnover) are  
27  
28 activated that help to maintain their levels and so compensate for the loss of Tic20. Dramatic effects  
29  
30 were observed in relation to components of the photosynthetic apparatus (LHCP, OE33 and FNR), with  
31  
32 amounts reduced to less than 10% of that seen in the wild type. In addition, the tetrapyrrole biosynthetic  
33  
34 enzyme, CPO, was also strongly depleted in the *tic20* mutants, providing a further indication that plastid  
35  
36 biogenesis is strongly disrupted.  
37  
38

#### 39 40 41 42 43 44 **Genetic interactions between *tic20-I* and the other *tic20* mutations**

45  
46  
47  
48 To study functional relationships between *atTic20-I* and the other three genes, *tic20-I* heterozygotes  
49  
50 (both alleles) were crossed to the other mutants. Individual green plants from resulting F<sub>2</sub> (or F<sub>3</sub>, in the  
51  
52 case of *tic20-I-1 tic20-V-1*) generations were genotyped, and plants that were heterozygous for *tic20-I*  
53  
54 and homozygous for each of the other *tic20* mutations were identified. For all such double mutants  
55  
56 involving *tic20-II* and *tic20-V*, albino plants were observed in the subsequent generation, following self-  
57  
58 pollination, at the expected frequency (Table 1); for each of the four relevant crosses, ten albino plants  
59  
60 were genotyped and shown to be double homozygotes (data not shown). Moreover, these double-

1  
2  
3 homozygous albino plants were not phenotypically different from the corresponding *tic20-I* single-  
4 mutant parent (Figure 5b). This suggests that atTic20-I shares little or no functional redundancy with  
5 the Group 2 proteins.  
6  
7

8  
9  
10 By contrast, in the F<sub>3</sub> progeny of the *tic20-I tic20-IV* double mutants, no albinos were  
11 observed (Table 1). Our failure to identify any albinos amongst >3000 F<sub>3</sub> or F<sub>4</sub> individuals suggested  
12 that the double homozygotes are not viable. To ensure that the double homozygotes were not amongst  
13 the green individuals, ~30-60 such plants from each cross were PCR-genotyped; none of them was  
14 found to be double homozygous (data not shown). To investigate the possibility that the double  
15 homozygous genotypes were embryo lethal, the siliques of individuals that were heterozygous for *tic20-*  
16 *I* and homozygous for *tic20-IV* were studied. In each case, we observed a significant number of very  
17 small, brown aborted seeds, and a much larger proportion of even smaller, white failed ovules (Figure  
18 8a). Detailed scoring revealed the frequencies of these defective structures to be ~5-10% and ~40-50%,  
19 respectively (Figure 8b). This implied a strong but incomplete defect in female gametophyte  
20 development, and furthermore indicated that any double homozygotes that do form arrest during  
21 embryo and seed development.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

35  
36 To test the hypothesis that the *tic20-I tic20-IV* double mutations affect female gametophytic  
37 transmission, we conducted reciprocal crossing experiments between the double mutants and wild type  
38 (Figure 8c). The results revealed essentially normal transmission through the male gametes  
39 (transmission efficiency was 93.6%), and a strong defect in transmission through the female gametes  
40 (34.4%) (Howden *et al.*, 1998). These results are consistent with phenotypic observations made upon  
41 analysing the double-mutant siliques (Figure 8a,b), and with the conclusions drawn there from.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

### 53 **Analysis of single, double and triple mutations affecting atTic20-IV, atTic20-II and atTic20-V**

54  
55  
56

57 None of the *tic20-IV*, *tic20-II* and *tic20-V* single mutants was detectably different from wild type. To  
58 investigate the possibility that this might be due to functional redundancy, we proceeded to identify a  
59 series of double- and triple-mutant combinations. For all tested combinations, doubly or triply  
60

1  
2  
3 homozygous mutant plants could be identified. However, none of these plants exhibited an obvious  
4  
5 mutant phenotype, and in each case chlorophyll content and photosynthetic performance was similar to  
6  
7 that in wild type (Figure S5). This suggests that the three genes do not play major roles in chloroplast  
8  
9 development under the conditions tested, and implies that they share little or no functional redundancy.  
10  
11 It is not surprising that no genetic interactions were detected between *tic20-IV* and the Group 2  
12  
13 mutations, as the *tic20-I* mutations also did not exhibit detectable interaction with either *tic20-II* or  
14  
15 *tic20-V* (Figure 5b; Table 1). It is possible that our failure to identify a complete knockout for *atTic20-II*  
16  
17 is responsible for the lack of any detectable interactions amongst the Group 2 mutants. However, this  
18  
19 seems unlikely as in the *tic20-II-1* allele at least, a very strong reduction in transcript abundance was  
20  
21 observed (Figure 4). A more likely explanation is that the Group 2 proteins only play a significant role  
22  
23 under non-standard conditions that were not assessed in our studies.  
24  
25  
26  
27  
28  
29  
30

### 31 **Analysis of functional interrelationships by complementation analysis**

32  
33  
34  
35 To further investigate the possibility of redundancy between the different Arabidopsis Tic20  
36  
37 homologues, we attempted to complement the *tic20-I-1* mutant with different Tic20 overexpressor  
38  
39 constructs. Heterozygous *+/tic20-I-1* plants were transformed with constructs comprising the *atTIC20-I*,  
40  
41 *atTIC20-IV*, *atTIC20-II* and *atTIC20-V* coding sequences under the control of the cauliflower mosaic  
42  
43 virus 35S promoter. For each construct, a total of ~8-10 transformants were identified, and from these  
44  
45 four independent lines were selected on the basis of overexpression of the relevant *TIC20* transgene, as  
46  
47 estimated by RT-PCR. The genotypes of the transformants were all verified by genomic PCR.  
48  
49

50  
51 As expected, the 35S:*TIC20-I* control construct was very effective at mediating  
52  
53 complementation of *tic20-I-1* (Figure S6). In accordance with the phylogenetic and double-mutant  
54  
55 analyses indicating a close relationship between *atTic20-I* and *atTic20-IV* (Figures 1 and 8), the  
56  
57 35S:*TIC20-IV* construct also mediated significant complementation (Figure 9). However, while the  
58  
59 35S:*TIC20-IV* transformants grew to a significantly larger size than the untransformed control plants,  
60  
this construct was not able to restore normal greening to *tic20-I*, unlike the control 35S:*TIC20-I*

1  
2  
3 construct. This indicated significant but incomplete functional redundancy between atTic20-I and  
4  
5 atTic20-IV.  
6

7  
8 By contrast, 35S:*TIC20-II* and 35S:*TIC20-V* constructs did not mediate any detectable  
9  
10 complementation of *tic20-I-1*, in any of the transformants identified (Figure S6). One caveat is that we  
11  
12 were unable to identify transformants (for either of these constructs) that exhibited very high levels of  
13  
14 overexpression, in contrast with the situation for the Group 1 constructs. The reason for this is  
15  
16 uncertain, although it may relate to transcript stability. Nonetheless, our genetic analyses (Figures 5b  
17  
18 and S5) do suggest that, even with higher levels of overexpression, a positive complementation result  
19  
20 would be unlikely in these experiments.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## DISCUSSION

Four Arabidopsis Tic20 genes were identified previously based on similarity with the original isolate from pea (Bédard and Jarvis, 2005; Kalanon and McFadden, 2008). Here we show that these four homologues are expressed, that they are likely to be topologically similar to the pea protein, and that they are targeted to the chloroplast envelope (Figures S1 and 2). Phylogenetic analysis of a large number of sequences revealed two evolutionarily conserved sub-classes, termed Group 1 (characterized by psTic20, atTic20-I and atTic20-IV) and Group 2 (characterized by atTic20-II and atTic20-V) (Figure 1). The former is more important in Arabidopsis, as the loss of these proteins results in albinism or developmental arrest during gametophyte or embryo development. The albino phenotype of the *tic20-I* T-DNA lines (Figure 5) is consistent with previous results (Teng *et al.*, 2006), and with the observation that antisense-mediated down-regulation of *atTIC20-I* inhibits greening (Chen *et al.*, 2002). Our genetic and complementation analyses demonstrated that the two Arabidopsis Group 1 proteins share functional redundancy (Figures 8 and 9), although the redundancy is incomplete as overexpression of atTic20-IV could only partially complement *tic20-I-1*.

All four Arabidopsis Tic20 genes are expressed throughout development (Figures 3 and S3). Nonetheless, *atTIC20-I* and *atTIC20-IV* do exhibit quite different patterns of expression, suggesting that the former is relatively more important for photosynthetic development, and that the latter acts during non-photosynthetic growth and seed development. High expression of *atTIC20-IV* in seeds is consistent with the observation that its inactivation, in conjunction with *tic20-I*, causes developmental arrest during female gametogenesis or embryogenesis. That the two genes exhibit distinct expression patterns, biased in favour of either photosynthetic or non-photosynthetic growth, parallels observations made in relation to the TOC receptors, Toc159 and Toc33 (Bauer *et al.*, 2000; Gutensohn *et al.*, 2000; Jarvis *et al.*, 1998; Kubis *et al.*, 2003; Kubis *et al.*, 2004). Analyses of different isoforms of these receptors indicated that the dominant one in each case (atToc159 and atToc33, respectively) is specialized for the import of highly-abundant, photosynthesis-related preproteins, and that the lesser isoforms (atToc132, atToc120 and atToc34) preferentially mediate import of low-abundance, housekeeping preproteins. In this regard, it is noteworthy that analysis of *tic20-I* plants using an *in vivo* import assay revealed a

1  
2  
3 strong effect on the import of a photosynthetic precursor, but little effect on a housekeeping protein  
4  
5 (Kikuchi *et al.*, 2009).  
6

7  
8 One hypothesis accounting for the sub-functionalization of TOC receptors holds that this  
9  
10 prevents damaging competition effects between preproteins; in the absence of such client-specific  
11  
12 receptor complexes, the bulk-flow of photosynthetic precursors might interfere with the import of other  
13  
14 preproteins (Jarvis, 2008; Kessler and Schnell, 2006). It has been suggested that the different import  
15  
16 pathways converge at the TIC machinery, based on observations that some TIC components (*e.g.*,  
17  
18 Tic110) have not undergone similar sub-functionalization in Arabidopsis (Kovacheva *et al.*, 2005).  
19  
20 However, recent evidence suggests that there may be at least two different TIC complexes, and that the  
21  
22 one containing Tic110 acts downstream of that containing Tic20 (Kikuchi *et al.*, 2009). If this is the  
23  
24 case, it is possible that the aforementioned client-specific import pathways do extend to the inner  
25  
26 membrane (at Tic20), and that convergence only happens later at Tic110. That is, there may be distinct  
27  
28 Tic20-containing channel complexes (*e.g.*, containing either atTic20-I or atTic20-IV), but just a single  
29  
30 Tic110-containing chaperone complex.  
31  
32

33  
34 Several TOC/TIC components are known to be essential during embryogenesis (Baldwin *et al.*  
35  
36 *et al.*, 2005; Hust and Gutensohn, 2006; Inaba *et al.*, 2005; Kovacheva *et al.*, 2005; Patel *et al.*, 2008), but  
37  
38 there have been few reports of lesions in chloroplast proteins leading to gametophyte arrest. Examples  
39  
40 are the *gpt1* mutations that affect the glucose 6-phosphate/phosphate translocator (GPT) of the inner  
41  
42 envelope membrane (Niewiadomski *et al.*, 2005). However, in contrast with the Group 1 Tic20 mutants,  
43  
44 *gpt1* mutants exhibited defective transmission through both male and female gametes; these effects  
45  
46 were attributed to a reduced supply of reducing equivalents via the oxidative pentose phosphate  
47  
48 pathway, which in turn affects fatty acid synthesis leading to defective membrane biogenesis. That the  
49  
50 gametophytic defects in *tic20-I tic20-IV* are less severe suggests that the Tic20 complex is not essential  
51  
52 for the biogenesis of inner membrane carrier proteins such as GPT, or that pre-meiotically synthesized  
53  
54 Tic20 persists long enough to mediate sufficient GPT biogenesis.  
55  
56

57  
58 Another example of plastid-linked defective gametogenesis occurs in the Arabidopsis Hsp93  
59  
60 double mutant (Kovacheva *et al.*, 2007). Transmission of *hsp93-V-1 hsp93-III-2* through female  
gametes was reduced to 46.8%, and consequently a significant number of failed ovules were observed

1  
2  
3 in heterozygous double-mutant plants. These results are similar to those reported here, although the  
4  
5 female transmission efficiency defect was stronger in the case of *tic20-I tic20-IV* (at 34.4%), and so a  
6  
7 greater proportion of defective reproductive structures (mostly failed ovules) was observed. This  
8  
9 severity of the Group 1 Tic20 double-mutant phenotype during reproductive growth (Figure 8) parallels  
10  
11 that seen in single-mutant *tic20-I* plants (which are even more sick than the *ppi2* albino; Figure 5), and  
12  
13 suggests that Tic20 plays a crucial role in the import mechanism (*e.g.*, translocation channel formation).  
14  
15

16 The roles of the Group 2 Tic20 proteins remain uncertain, as no mutant phenotypes were  
17  
18 detected in any of the mutant genotypes analysed (Figure S5). Nonetheless, because these genes are  
19  
20 expressed at high levels (Figure 3) and have been conserved over millions of years of evolution in many  
21  
22 species (Figure 1), it seems likely that they do perform some important role. One possibility is that they  
23  
24 are somehow analogous to Tim17 in mitochondria; the role of this protein is uncertain, although it has  
25  
26 been proposed to regulate channel formed by the related protein, Tim23 (Neupert and Herrmann, 2007).  
27  
28 In contrast with the Group 2 Tic20 proteins, however, Tim17 is essential for viability in yeast. It is  
29  
30 conceivable that the roles of *atTic20-II* and *atTic20-V* become critical only under non-standard  
31  
32 conditions that were not tested during the course of this study.  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## EXPERIMENTAL PROCEDURES

### Phylogenetics and consensus sequence analysis

In total, 59 sequences from the Tic20 protein family were obtained by BLAST queries of databases at NCBI (<http://www.ncbi.nlm.nih.gov/>), DOE Joint Genome Institute (<http://genome.jgi-psf.org/>), and the *Cyanidioschyzon merolae* genome project (<http://merolae.biol.s.u-tokyo.ac.jp/>) (Table S1). The sequences were aligned with mafft-linsi v.6.717b (Katoh *et al.*, 2005) and the N-terminal part of the matrix (starting at position 83 of the psTic20 sequence) was analysed with MrBayes v.3.2 (Huelsenbeck and Ronquist, 2001) under the mixed amino acid model. For further details, see Appendix S1.

Representative sequences from the Tic20 and Tim17-22-23 protein families were aligned to four prokaryotic sequences using mafft-linsi v.6.717b (Katoh *et al.*, 2005) and manually adjusted according to the alignment presented by Rassow *et al.* (1999). Consensus sequences were created in SeaView v.4.2 (Gouy *et al.*, 2010) using the default settings. For further details, see Appendix S1.

### Plant growth conditions

*Arabidopsis thaliana* plants were Columbia-0 ecotype, and were grown as described previously (Aronsson and Jarvis, 2002). To select for T-DNA insertions, antibiotics were added to the MS medium: kanamycin monosulfate, 50 µg/ml (*tic20-I-1*, *tic20-I-2*, *tic20-II-1* and *tic20-V-2*); DL-phosphinothricin, 10 µg/ml (*tic20-IV-1* and pB2GW7 transformants); hygromycin B, 15 µg/ml (*tic20-IV-2* and *tic20-V-1*); and sulfadiazine, 11.25 µg/ml (*tic20-II-2*).

### Protoplast transfection and subcellular localization analysis

Full-length Tic20 coding sequences, each one lacking the native stop codon, were PCR-amplified and inserted into the p2GWY7 vector (Karimi *et al.*, 2005); for primer sequences, see Appendix S1. Protoplasts were prepared from 14-day-old, wild-type plants, and transfected, as described previously (Bédard *et al.*, 2007). Fluorescence microscopy employed a Nikon Eclipse TE-2000E inverted microscope equipped with filters for analysing YFP (exciter HQ500/20x, emitter HQ535/30m) and chlorophyll autofluorescence (exciter D480/30x, emitter D660/50m) (Chroma Technologies).

### Quantitative real-time (QPCR) and standard RT-PCR

RNA was isolated using an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), and treated with DNase I. Reverse transcription was performed as described previously (Kovacheva *et al.*, 2005; Kovacheva *et al.*, 2007); see Appendix S1 for more detail.

For QPCR, three biological replicates were analysed, and each one was measured in triplicate using a Chromo4 Gradient Cycler (MJ Research) and SYBR Green Jump Start Taq Ready Mix (Sigma). Data were normalized using similarly-derived *ACTIN2* (At3g18780) data.

For semi-quantitative analysis, RT-PCR products were resolved by electrophoresis and stained with SYBR Safe (Invitrogen). To avoid saturation, only 25 amplification cycles were employed. Bands were quantified using ImageQuant (GE Healthcare), and data were normalized using equivalent *eIF4E1* values (Rodriguez *et al.*, 1998).

### Identification and analysis of the Tic20 mutants

The *tic20-I-1*, *tic20-I-2*, *tic20-II-1* and *tic20-V-2* mutants were from the Salk Institute Genomic Analysis Laboratory (lines SALK\_039676, SALK\_071877, SALK\_064931 and SALK\_013444, respectively)

1  
2  
3 (Alonso *et al.*, 2003; Teng *et al.*, 2006). The *tic20-IV-1* mutant was from Syngenta (line Garlic\_97\_F10)  
4  
5 (Sessions *et al.*, 2002). The *tic20-IV-2* and *tic20-V-1* mutants were from the Csaba Koncz laboratory  
6  
7 (pool 113, line 11324 and pool 53, line 5290, respectively) (Ríos *et al.*, 2002). The *tic20-II-2* mutant  
8  
9 was from GABI-Kat (line 879A01) (Rosso *et al.*, 2003). Mutant genotypes were assessed by PCR (see  
10  
11 Figure S4 and Appendix S1). The location of each T-DNA insertion was determined precisely (Figure  
12  
13 4a) by sequencing junction-spanning PCR products.  
14  
15

16 Transmission electron microscopy was performed as described previously (Aronsson *et al.*,  
17  
18 2007), with minor modifications (see Appendix S1). Immunoblotting employed previously described  
19  
20 procedures (Kovacheva *et al.*, 2005; Kovacheva *et al.*, 2007), using secondary antibodies conjugated to  
21  
22 either alkaline phosphatase or horseradish peroxidase. For further details, see Appendix S1.  
23  
24  
25  
26  
27  
28

### 29 **Complementation experiments**

30  
31  
32

33 Full-length Tic20 coding sequences, each one retaining the native stop codon, were PCR-amplified and  
34  
35 inserted into the pB2GW7 vector (Karimi *et al.*, 2005); for primer sequences, see Appendix S1.  
36  
37 Constructs were stably introduced into wild-type plants using the floral dip method (Clough and Bent,  
38  
39 1998). Transformants were identified by screening on MS medium containing phosphinothricin.  
40  
41 Homozygotes were identified in the T<sub>3</sub> or T<sub>4</sub> generations.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## ACKNOWLEDGEMENTS

We thank Natalie Allcock and Stefan Hyman for electron microscopy, Sybille Kubis for assistance with the YFP studies, and Anthony Wardle for technical support. We thank Jocelyn Bédard (atToc33, atTic40) and Feijie Wu (atToc75-III, atToc120) for preparing antigens for immunization, and Bernhard Grimm, Neil Hoffman, Felix Kessler, and Henrik Scheller for providing antibodies. This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) PhD studentship, by the Royal Society Rosenheim Research Fellowship, and by BBSRC grants 91/C18638, BBS/B/03629, BBS/C006348/1, BB/D016541/1 and BB/F020325/1 (to P.J.).

## TABLES

**Table 1.** Segregation of the albino phenotype in various *tic20-1* double-mutant populations.

**FIGURE LEGENDS**

**Figure 1.** Phylogenetic analysis of Tic20-related sequences from different species.

Amino acid sequences of Tic20 homologues were aligned and used to produce a phylogenetic tree; the N-terminal part of the matrix, equivalent to the first 82 residues of psTic20 (Kouranov *et al.*, 1998), was excluded before the analysis. Eukaryote sequences form two major clades, termed Group 1 and Group 2. Arrows indicate branches of differing lengths leading to land plant proteins in the two groups. Chromalveolate sequences are indicated with asterisks. Branch lengths in the phylogram are proportional to the number of expected changes (see scale). Posterior probability values appear above the branches.

**Figure 2.** Subcellular localization of the Arabidopsis Tic20 proteins as assessed by YFP fusion-protein analysis.

Wild-type Arabidopsis protoplasts were transfected with the indicated plasmids (atTic20-I:YFP, atTic20-IV:YFP, atTic20-II:YFP and atTic20-V:YFP) and then analysed for YFP fluorescence (green, left panels) and chlorophyll autofluorescence (red, centre-left panels), as well as under bright field illumination (right panels). An overlay of the YFP and chlorophyll images is presented (centre-right panels). Scale bars = 10  $\mu$ m.

**Figure 3.** Expression of the Tic20 genes in different tissues and at different developmental stages.

Quantitative RT-PCR analysis of total-RNA from whole seedlings grown *in vitro* for five days in the dark (5dD), or five and 14 days in the light (5dL and 14dL, respectively), as well as from three different tissues of mature plants grown on soil (roots, rosette leaves, and siliques). RNA samples were representative of ~10-30 seedlings (5dD, 5dL and 14dL), or 5-25 mature plants (roots, rosettes and siliques). Tic20 data were normalized relative to the control gene, *ACTIN2* (At3g18780), and then expressed relative to the *atTIC20-I* 5dL value. Data shown were derived from three independent

1  
2  
3 amplifications done on three biological replicates. Panels (a) and (b) contain the same data presented in  
4  
5 different ways.  
6  
7  
8  
9  
10

11 **Figure 4.** Molecular characterization of the Tic20 T-DNA insertion lines.

12  
13 (a) Diagrams showing the structure of each gene and the locations of T-DNA insertions. Protein-coding  
14 exons are represented by black boxes, untranslated regions by white boxes, and introns by thin lines  
15  
16 between the boxes. Locations of RT-PCR primers are shown by arrows beneath each gene model. T-  
17  
18 DNA insertion sites are indicated precisely, but insertion sizes are not to scale. ATG, translation  
19  
20 initiation codon; Stop, translation termination codon; p(A), polyadenylation site; LB, T-DNA left  
21  
22 border; RB, T-DNA right border.  
23  
24  
25  
26

27 (b) Expression of Tic20 genes in wild-type and mutant plants was analysed by RT-PCR. Locations of  
28  
29 the amplification primers are shown in (a). Similar analysis of *atTOC33* and of the translation initiation  
30  
31 factor gene, *eIF4E1* (At4g18040), was used to normalize loading. Amplicon sizes are indicated at right  
32  
33 (in kb). RNA samples were from whole, 10-day-old homozygous plants grown *in vitro*, and were  
34  
35 representative of ~20-30 seedlings.  
36  
37

38 (c) Semi-quantitative RT-PCR analysis of *tic20-I-2*, *tic20-II-1* and *tic20-II-2*. This was necessary  
39  
40 because data in (b) revealed that these mutants are not null; it was conducted essentially as described in  
41  
42 (b). Values shown are means ( $\pm$ SE) derived from four independent amplifications, and are expressed as  
43  
44 percentages of the wild-type value. Data were normalized relative to *eIF4E1*. PCR amplifications were  
45  
46 performed over 25 cycles.  
47  
48  
49  
50  
51  
52

53 **Figure 5.** Albino phenotypes of *tic20-I* single- and double-mutant plants.

54  
55 Populations segregating for the indicated *tic20-I* mutation, either in the wild-type background (a) or in  
56  
57 the indicated homozygous *tic20-II* or *tic20-V* mutant background (b), were plated on standard MS  
58  
59 medium. After eight days, homozygous albino mutants were rescued to medium containing 3% sucrose  
60  
(they are unable to survive photoautotrophically) and grown until they were 43 days old. Representative

1  
2  
3 plants were photographed. A similarly-grown *ppi2* mutant is shown for comparison (a). Plants were  
4  
5 illuminated with dim light ( $\sim 10 \mu\text{mol}/\text{m}^2/\text{sec}$ ), under a standard long-day cycle, to aid growth.  
6  
7  
8  
9

10  
11 **Figure 6.** Plastid ultrastructure in the *tic20-I* single mutants.

12  
13 Cotyledons of 10- and 14-day-old plants (10d and 14d, respectively) were analysed by transmission  
14  
15 electron microscopy. Seeds were plated on standard MS medium; after five days, homozygous albino  
16  
17 mutants were rescued to medium containing 3% sucrose. On average,  $\sim 30$  whole-chloroplast  
18  
19 micrographs from each of three independent plants per genotype per age (a minimum of 55 chloroplasts  
20  
21 per genotype per age) were analysed, and used to select the representative images shown. Size bars =  
22  
23  
24  
25 1.0  $\mu\text{m}$ .  
26  
27  
28  
29  
30

31  
32 **Figure 7.** Chloroplast protein levels in Tic20-deficient mutants.

33  
34 (a,b) Equal protein samples (20  $\mu\text{g}$ ) from 10-day-old homozygous, MS-grown plants of the indicated  
35  
36 genotypes were analysed by immunoblotting using the antibodies shown. LHCP, light-harvesting  
37  
38 chlorophyll-binding protein; OE33, oxygen-evolving complex, 33 kD subunit; FNR, ferredoxin-NADP  
39  
40 reductase; CPO, coproporphyrinogen oxidase; H3, histone H3.

41  
42 (c,d) Chloroplast protein bands in (a) and (b) (as well as those in other, similar experiments) were  
43  
44 quantified using Aida software (Raytest), and then the data were normalized using equivalent H3 data.  
45  
46 Values shown are means ( $\pm\text{SE}$ ) derived from 2-6 independent measurements (experiments with just two  
47  
48 repeats were exclusively those displaying the clearest results, shown in [b]).  
49  
50  
51  
52  
53  
54

55  
56 **Figure 8.** Genetic analysis of *tic20-I tic20-IV* double mutants.

57  
58 (a,b) Siliques of mature, soil-grown plants of the indicated genotypes were opened and inspected. In  
59  
60 *tic20-I tic20-IV* double-mutant plants (genotype:  $+/\text{tic20-I}; \text{tic20-IV}/\text{tic20-IV}$ ), a large proportion of  
failed ovules and a smaller number of aborted seeds was observed; a representative *tic20-I-1 tic20-IV-1*

1  
2  
3 silique is shown (a). White arrows indicate failed ovules; black arrows indicate aborted seeds.  
4  
5 Frequencies of failed ovules, aborted seeds, and normal seeds in double-mutant and control genotypes  
6  
7 were recorded (b); data shown are means ( $\pm$ SE) derived from analyses of, on average, 30 siliques (a  
8  
9 minimum of 400 seeds) per genotype.  
10

11 (c) Reciprocal crossing analysis. Transmission of the *tic20-I tic20-IV* double mutations, through the  
12  
13 female and male gametes, was assessed by crossing three of the four double-mutant genotypes shown in  
14  
15 (b) to wild type, multiple times, in both directions; simultaneous inheritance of both mutations in the F<sub>1</sub>  
16  
17 progeny was assessed by PCR genotyping or antibiotic resistance. Transmission efficiencies were  
18  
19 calculated for each double-mutant genotype, and then those three values were used to derive the means  
20  
21 shown ( $\pm$ SE).  
22  
23  
24  
25  
26  
27

28 **Figure 9.** Complementation of *tic20-I-1* by overexpression of *atTIC20-IV*.

29  
30 (a,b) Expression of the 35S:*TIC20-IV* transgene in three independent transformants (3-10, 8-3 and 11-  
31  
32 4). Total-RNA samples were extracted from pools (~20-30 individuals) of 10-day-old, T<sub>4</sub>-generation,  
33  
34 MS-grown seedlings that were homozygous for the relevant transgene, as well as from similar wild-type  
35  
36 plants. RT-PCR employed *atTIC20-IV* primers indicated in Figure 4a, and 25 amplification cycles.  
37  
38 Amplicon sizes are indicated to the right of the gel image (in kb). The chart shows mean fold-change  
39  
40 values ( $\pm$ SE), relative to wild type, derived from four independent amplifications; values have been  
41  
42 normalized using *eIF4E1* data.  
43  
44

45 (c) Rosette size in the transgenic lines. Diameters of the rosettes (at their widest points) of the plants  
46  
47 described in (d) below, as well as of other similar plants, were measured. Values shown are means  
48  
49 ( $\pm$ SE) derived from ~8-24 different plants per genotype.  
50

51 (d) Appearance of the transgenic lines. Populations segregating for *tic20-I-1* (and which either lacked  
52  
53 the 35S:*TIC20-IV* transgene [*tic20-I-1* control], or were homozygous for the indicated transgene [T<sub>4</sub>  
54  
55 generation]) were plated on standard MS medium; after five days, homozygous albino individuals were  
56  
57 rescued to medium containing 3% sucrose and grown until they were 38 days old. Representative plants  
58  
59 were photographed.  
60

1  
2  
3 **SUPPLEMENTARY MATERIAL**  
4  
5  
6

7 Additional supporting information may be found in the online version of this article:  
8  
9

10  
11 **Figure S1.** Annotated alignment of the Arabidopsis and pea Tic20 proteins.  
12

13 **Figure S2.** Assessment for possible homology between Tic20, bacterial amino acid transporter, and  
14 Tim17-22-23 sequences.  
15  
16

17 **Figure S3.** Expression of the Arabidopsis Tic20 genes as assessed using publicly-available microarray  
18 data.  
19  
20

21 **Figure S4.** Analysis of the Tic20 T-DNA mutants by genomic PCR.  
22

23 **Figure S5.** Phenotypic analysis of various *tic20* single, double and triple mutant plants.  
24  
25

26 **Figure S6.** Transgenic overexpression of *atTIC20-I*, *atTIC20-II* and *atTIC20-V* in *tic20-I-1* mutant  
27 plants.  
28  
29

30 **Table S1.** The Tic20-related amino acid sequences used for the phylogenetic analysis.  
31

32 **Table S2.** Segregation analysis of the T-DNA-associated antibiotic resistance marker in each one of the  
33 Tic20 mutants.  
34  
35

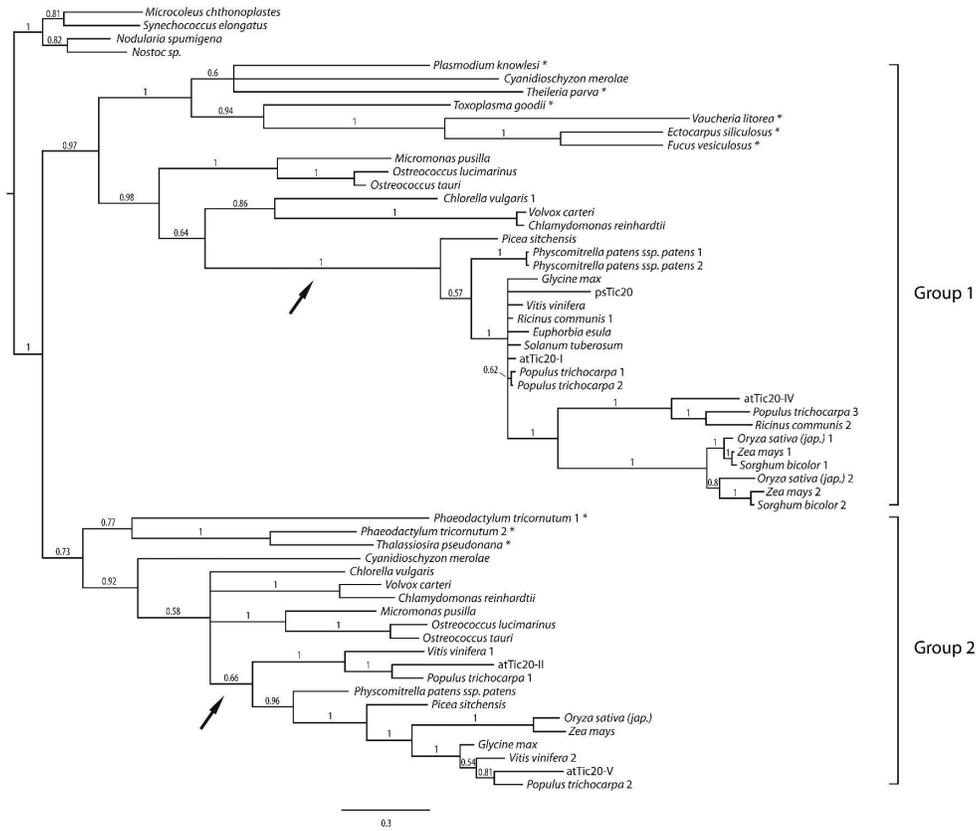
36 **Appendix S1.** Supplementary information for experimental procedures.  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## REFERENCES

- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
- Aronsson, H. and Jarvis, P. (2002) A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* **529**, 215-220.
- Aronsson, H., Boij, P., Patel, R., Wardle, A., Töpel, M. and Jarvis, P. (2007) Toc64/OEP64 is not essential for the efficient import of proteins into chloroplasts in *Arabidopsis thaliana*. *Plant J.* **52**, 53-68.
- Baldwin, A., Wardle, A., Patel, R., Dudley, P., Park, S.K., Twell, D., Inoue, K. and Jarvis, P. (2005) A molecular-genetic study of the *Arabidopsis* Toc75 gene family. *Plant Physiol.* **138**, 715-733.
- Balsera, M., Goetze, T.A., Kovács-Bogdán, E., Schürmann, P., Wagner, R., Buchanan, B.B., Soll, J. and Bölter, B. (2009) Characterization of Tic110, a channel-forming protein at the inner envelope membrane of chloroplasts, unveils a response to Ca<sup>2+</sup> and a stromal regulatory disulfide bridge. *J. Biol. Chem.* **284**, 2603-2616.
- Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D. and Kessler, F. (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**, 203-207.
- Bédard, J. and Jarvis, P. (2005) Recognition and envelope translocation of chloroplast preproteins. *J. Exp. Bot.* **56**, 2287-2320.
- Bédard, J., Kubis, S., Bimanadham, S. and Jarvis, P. (2007) Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). *J. Biol. Chem.* **282**, 21404-21414.
- Butterfield, N.J. (2000) *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* **26**, 386-404.
- Chen, X., Smith, M.D., Fitzpatrick, L. and Schnell, D.J. (2002) In vivo analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* **14**, 641-654.
- Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735-743.
- Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005-1016.
- Gouy, M., Guindon, S. and Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221-224.
- Gutensohn, M., Schulz, B., Nicolay, P. and Flügge, U.I. (2000) Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. *Plant J.* **23**, 771-783.
- Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Küchler, M., Hörmann, F., Sveshnikov, D. and Soll, J. (2002) The preprotein conducting channel at the inner envelope membrane of plastids. *EMBO J.* **21**, 2616-2625.
- Hofmann, N.R. and Theg, S.M. (2005) Chloroplast outer membrane protein targeting and insertion. *Trends Plant Sci.* **10**, 450-457.

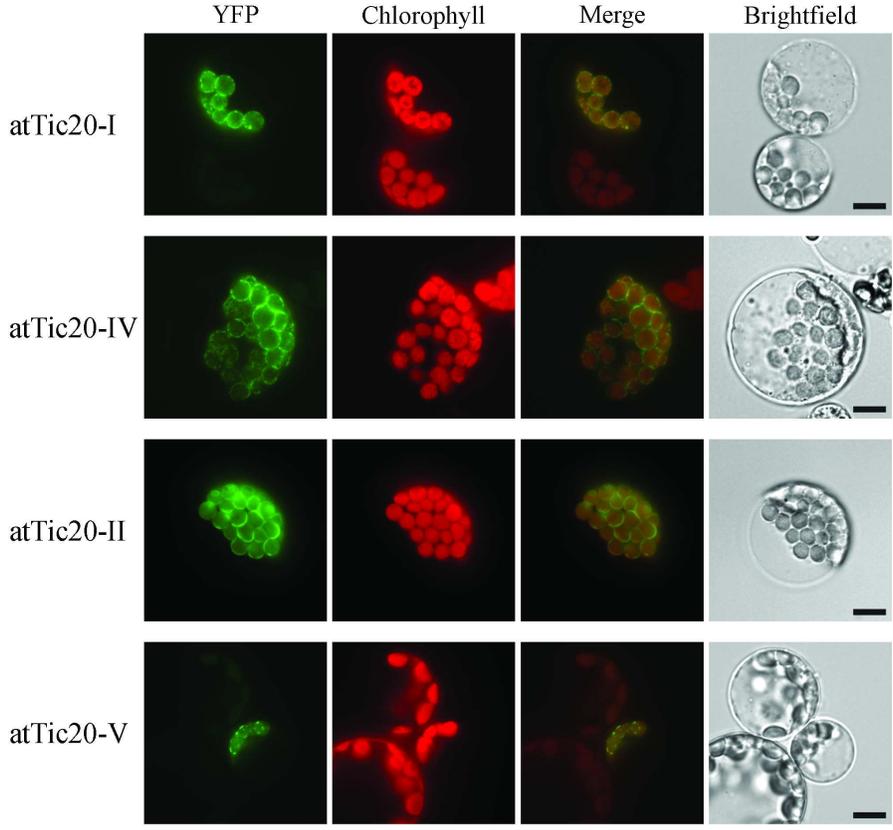
- 1  
2  
3 **Howden, R., Park, S.K., Moore, J.M., Orme, J., Grossniklaus, U. and Twell, D.** (1998) Selection of  
4 T-DNA-tagged male and female gametophytic mutants by segregation distortion in  
5 *Arabidopsis*. *Genetics* **149**, 621-631.
- 6 **Huelsenbeck, J.P. and Ronquist, F.** (2001) MRBAYES: Bayesian inference of phylogenetic trees.  
7 *Bioinformatics* **17**, 754-755.
- 8 **Hust, B. and Gutensohn, M.** (2006) Deletion of core components of the plastid protein import  
9 machinery causes differential arrest of embryo development in *Arabidopsis thaliana*. *Plant*  
10 *Biol. (Stuttg.)* **8**, 18-30.
- 11 **Inaba, T. and Schnell, D.J.** (2008) Protein trafficking to plastids: one theme, many variations. *Biochem*  
12 *J.* **413**, 15-28.
- 13 **Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F. and Schnell, D.J.** (2003) atTic110 functions as a  
14 scaffold for coordinating the stromal events of protein import into chloroplasts. *J. Biol. Chem.*  
15 **278**, 38617-38627.
- 16 **Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F. and Schnell, D.J.** (2005)  
17 *Arabidopsis* Tic110 is essential for the assembly and function of the protein import machinery  
18 of plastids. *Plant Cell* **17**, 1482-1496.
- 19 **Jarvis, P.** (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants (Tansley Review).  
20 *New Phytol.* **179**, 257-285.
- 21 **Jarvis, P., Chen, L.J., Li, H., Peto, C.A., Fankhauser, C. and Chory, J.** (1998) An *Arabidopsis*  
22 mutant defective in the plastid general protein import apparatus. *Science* **282**, 100-103.
- 23 **Kalanon, M. and McFadden, G.I.** (2008) The chloroplast protein translocation complexes of  
24 *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants,  
25 green algae and red algae. *Genetics* **179**, 95-112.
- 26 **Karimi, M., De Meyer, B. and Hilson, P.** (2005) Modular cloning in plant cells. *Trends Plant Sci.* **10**,  
27 103-105.
- 28 **Katoh, K., Kuma, K., Toh, H. and Miyata, T.** (2005) MAFFT version 5: improvement in accuracy of  
29 multiple sequence alignment. *Nucleic Acids Res.* **33**, 511-518.
- 30 **Kessler, F. and Schnell, D.J.** (2006) The function and diversity of plastid protein import pathways: a  
31 multilane GTPase highway into plastids. *Traffic* **7**, 248-257.
- 32 **Kikuchi, S., Oishi, M., Hirabayashi, Y., Lee, D.W., Hwang, I. and Nakai, M.** (2009) A 1-  
33 megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein  
34 import at the inner envelope membrane. *Plant Cell* **21**, 1781-1797.
- 35 **Kouranov, A. and Schnell, D.J.** (1997) Analysis of the interactions of preproteins with the import  
36 machinery over the course of protein import into chloroplasts. *J. Cell Biol.* **139**, 1677-1685.
- 37 **Kouranov, A., Chen, X., Fuks, B. and Schnell, D.J.** (1998) Tic20 and Tic22 are new components of  
38 the protein import apparatus at the chloroplast inner envelope membrane. *J. Cell Biol.* **143**, 991-  
39 1002.
- 40 **Kovacheva, S., Bédard, J., Wardle, A., Patel, R. and Jarvis, P.** (2007) Further *in vivo* studies on the  
41 role of the molecular chaperone, Hsp93, in plastid protein import. *Plant J.* **50**, 364-379.
- 42 **Kovacheva, S., Bédard, J., Patel, R., Dudley, P., Twell, D., Ríos, G., Koncz, C. and Jarvis, P.**  
43 (2005) *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein  
44 import. *Plant J.* **41**, 412-428.
- 45 **Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L.** (2001) Predicting transmembrane  
46 protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.*  
47 **305**, 567-580.
- 48 **Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D. and**  
49 **Jarvis, P.** (2003) The *Arabidopsis* *ppi1* mutant is specifically defective in the expression,  
50 chloroplast import, and accumulation of photosynthetic proteins. *Plant Cell* **15**, 1859-1871.
- 51 **Kubis, S., Patel, R., Combe, J., Bédard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Ríos,**  
52 **G., Koncz, C. and Jarvis, P.** (2004) Functional specialization amongst the *Arabidopsis* Toc159  
53 family of chloroplast protein import receptors. *Plant Cell* **16**, 2059-2077.
- 54 **Lesk, A.M. and Fordham, W.D.** (1996) Conservation and variability in the structures of serine  
55 proteinases of the chymotrypsin family. *J. Mol. Biol.* **258**, 501-537.
- 56 **Li, M. and Schnell, D.J.** (2006) Reconstitution of protein targeting to the inner envelope membrane of  
57 chloroplasts. *J. Cell Biol.* **175**, 249-259.
- 58  
59  
60

- 1  
2  
3 **Ma, Y., Kouranov, A., LaSala, S.E. and Schnell, D.J.** (1996) Two components of the chloroplast  
4 protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the  
5 recognition and translocation of precursor proteins at the outer envelope. *J. Cell Biol.* **134**, 315-  
6 327.
- 7  
8 **Neupert, W. and Herrmann, J.M.** (2007) Translocation of proteins into mitochondria. *Annu. Rev.*  
9 *Biochem.* **76**, 723-749.
- 10 **Niewiadomski, P., Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U.S., Rosso, M.G., Ache,**  
11 **P., Flügge, U.I. and Schneider, A.** (2005) The Arabidopsis plastidic glucose 6-  
12 phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac  
13 development. *Plant Cell* **17**, 760-775.
- 14 **Patel, R., Hsu, S., Bédard, J., Inoue, K. and Jarvis, P.** (2008) The Omp85-related chloroplast outer  
15 envelope protein OEP80 is essential for viability in Arabidopsis. *Plant Physiol.* **148**, 235-245.
- 16 **Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M. and Soll, J.** (1999) The preprotein translocase of  
17 the mitochondrial inner membrane: function and evolution. *J. Mol. Biol.* **286**, 105-120.
- 18 **Reumann, S., Davila-Aponte, J. and Keegstra, K.** (1999) The evolutionary origin of the protein-  
19 translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial  
20 homolog. *Proc. Natl. Acad. Sci. USA* **96**, 784-789.
- 21 **Reyes-Prieto, A., Weber, A.P. and Bhattacharya, D.** (2007) The origin and establishment of the  
22 plastid in algae and plants. *Annu. Rev. Genet.* **41**, 147-168.
- 23  
24 **Ríos, G., Lossow, A., Hertel, B., Breuer, F., Schaefer, S., Broich, M., Kleinow, T., Jasik, J.,**  
25 **Winter, J., Ferrando, A., Farras, R., Panicot, M., Henriques, R., Mariaux, J.B.,**  
26 **Oberschall, A., Molnar, G., Berendzen, K., Shukla, V., Lafos, M., Koncz, Z., Redei, G.P.,**  
27 **Schell, J. and Koncz, C.** (2002) Rapid identification of *Arabidopsis* insertion mutants by non-  
28 radioactive detection of T-DNA tagged genes. *Plant J.* **32**, 243-253.
- 29 **Rodriguez, C.M., Freire, M.A., Camilleri, C. and Robaglia, C.** (1998) The *Arabidopsis thaliana*  
30 cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast  
31 complementation and are differentially expressed during plant development. *Plant J.* **13**, 465-  
32 473.
- 33 **Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B.** (2003) An *Arabidopsis*  
34 *thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse  
35 genetics. *Plant Mol. Biol.* **53**, 247-259.
- 36 **Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P.,**  
37 **Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison,**  
38 **D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.** (2002) A  
39 high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* **14**, 2985-2994.
- 40 **Smith, M.D.** (2006) Protein import into chloroplasts: an ever-evolving story. *Can. J. Bot.* **84**, 531-542.
- 41 **Soll, J. and Schleiff, E.** (2004) Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* **5**, 198-208.
- 42 **Teng, Y.S., Su, Y.S., Chen, L.J., Lee, Y.J., Hwang, I. and Li, H.M.** (2006) Tic21 is an essential  
43 translocon component for protein translocation across the chloroplast inner envelope  
44 membrane. *Plant Cell* **18**, 2247-2257.
- 45  
46 **Tripp, J., Inoue, K., Keegstra, K. and Froehlich, J.E.** (2007) A novel serine/proline-rich domain in  
47 combination with a transmembrane domain is required for the insertion of AtTic40 into the  
48 inner envelope membrane of chloroplasts. *Plant J.* **52**, 824-838.
- 49 **van Dooren, G.G., Tomova, C., Agrawal, S., Humbel, B.M. and Striepen, B.** (2008) *Toxoplasma*  
50 *gondii* Tic20 is essential for apicoplast protein import. *Proc. Natl. Acad. Sci. USA* **105**, 13574-  
51 13579.
- 52  
53 **Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G. and Bhattacharya, D.** (2004) A molecular timeline  
54 for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809-818.
- 55 **Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W.** (2004)  
56 GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol.*  
57 **136**, 2621-2632.
- 58  
59  
60



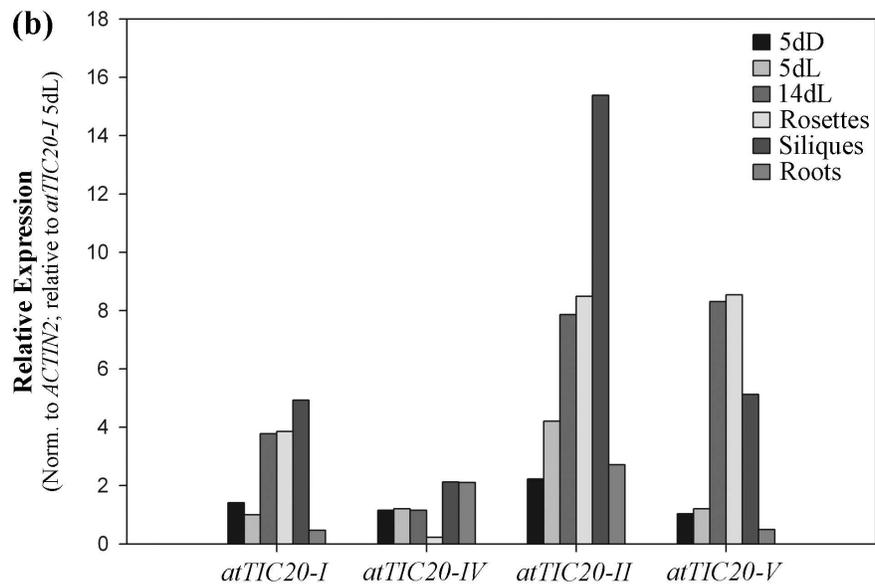
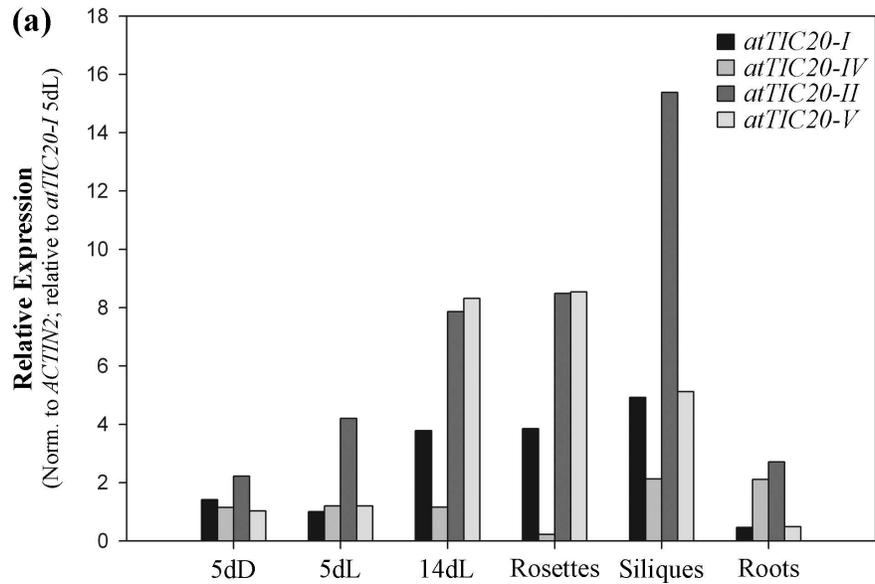
231x192mm (600 x 600 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

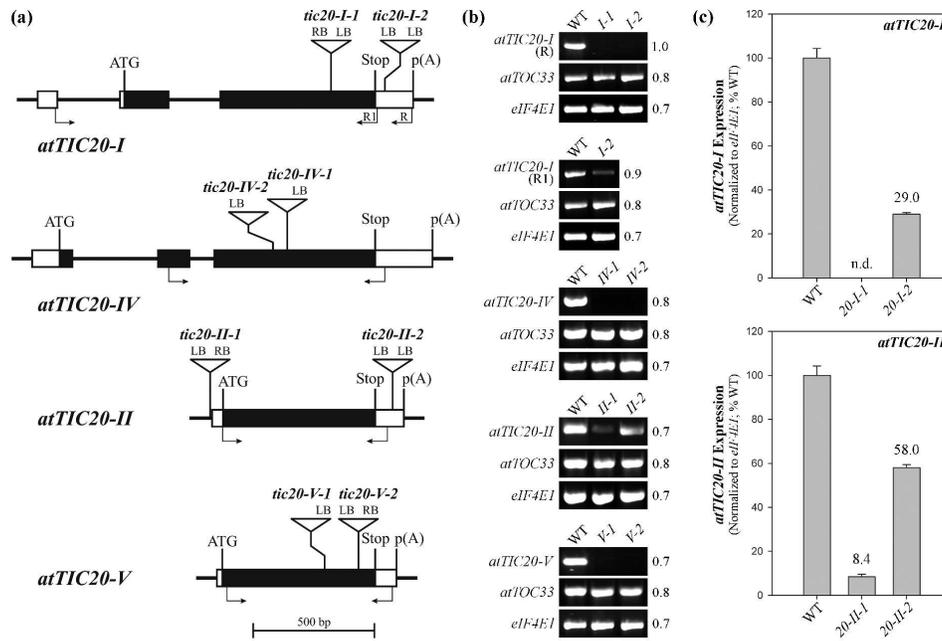


95x90mm (600 x 600 DPI)

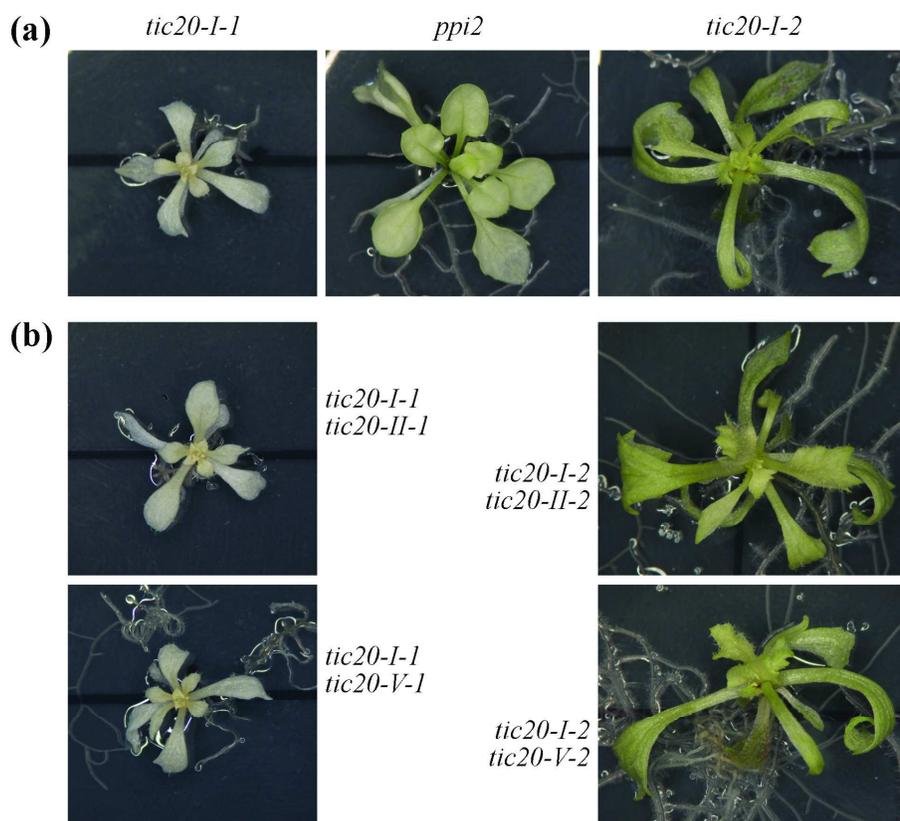




82x111mm (600 x 600 DPI)



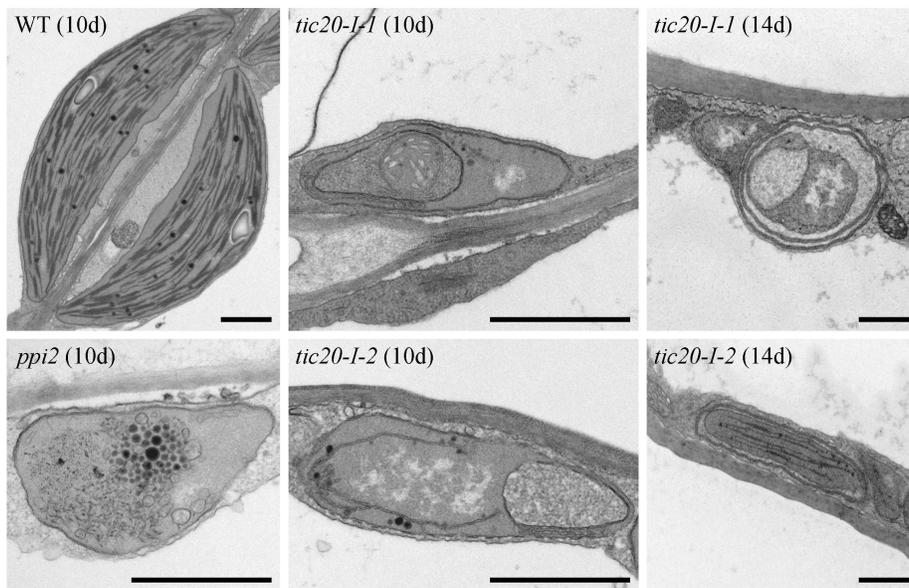
167x117mm (600 x 600 DPI)



89x86mm (600 x 600 DPI)

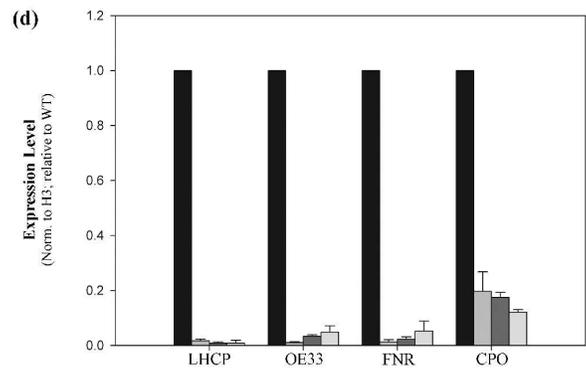
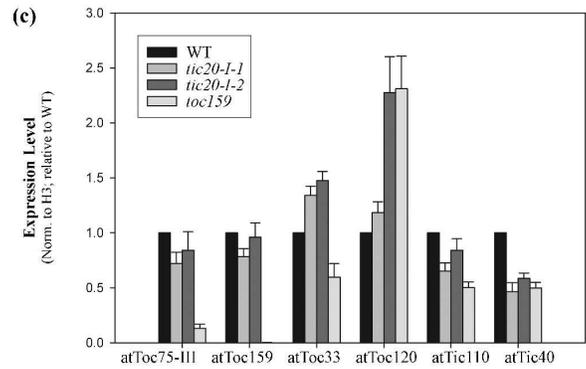
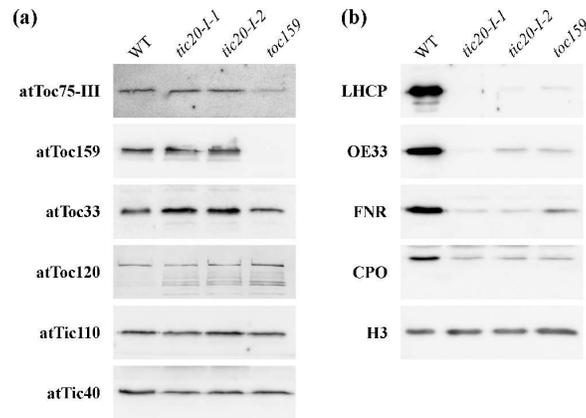


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



119x79mm (600 x 600 DPI)

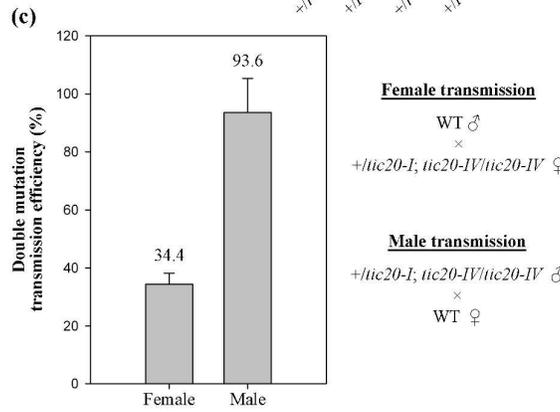
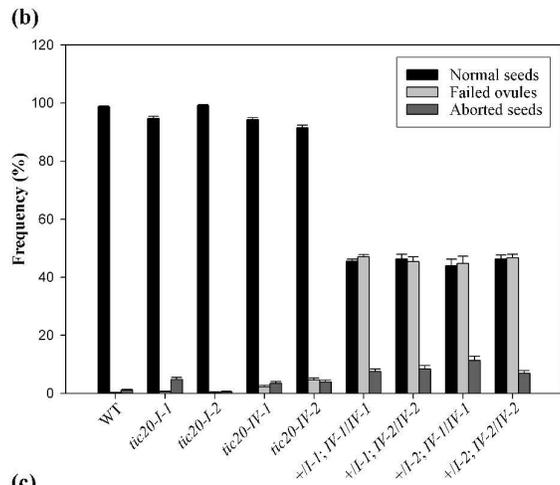
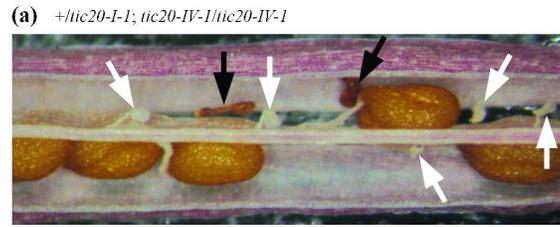
CONFIDENTIAL



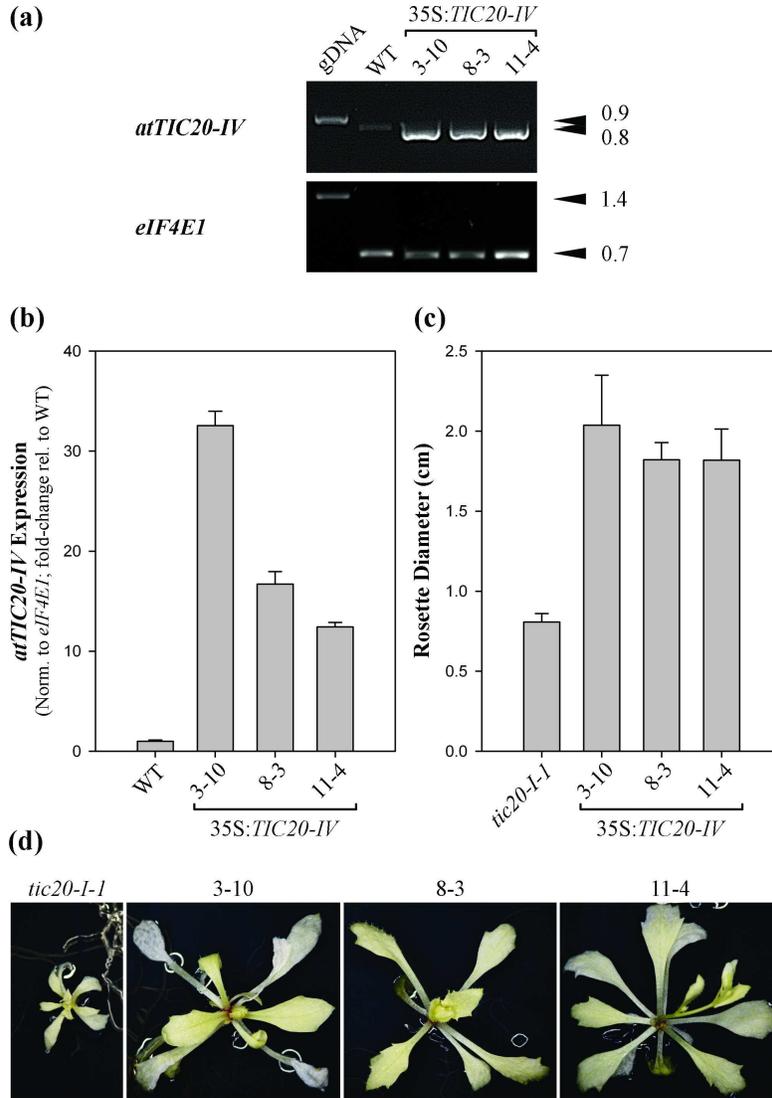
89x176mm (600 x 600 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



96x167mm (600 x 600 DPI)



94x131mm (600 x 600 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60