Babiychuk, E., Vandepoele, K., Wissing, J., Garcia-Diaz, M., De Rycke, R., Akbari, H., Joubès, J., Beeckman, T., Jänsch, L., Frentzen, M., Van Montagu, M.C.E., Kushnir, S.

Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family

(2011) Proceedings of the National Academy of Sciences of the United States of America, 108 (16), pp. 6674-6679.

Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family

Elena Babiychuk[a,b], Klaas Vandepoele[a,b], Josef Wissing[c], Miguel Garcia-Diaz[d], Riet De Rycke[a,b], Hana Akbari[e], Jérôme Joubès[f], Tom Beeckman[a,b], Lothar Jänsch[c], Margrit Frentzen[e], Marc Van Montagu[b], and Sergei Kushnir[a,b,1]

[a]Department of Plant Systems Biology, VIB, 9052 Gent, Belgium; [b]Department of Plant Biotechnology and Genetics, Ghent University, 9052 Gent, Belgium; [c]Abteilung Zellbiologie, Helmholtz-Zentrum für Infektionsforschung GmbH, 38124 Braunschweig, Germany; [d]Pharmacological Sciences, Stony Brook University, Stony Brook, NY 11794-8651, USA; [e]RWTH Aachen University, Institute of Biology I, Botany, 52056 Aachen, Germany; and [f]Université Victor Ségalen Bordeaux 2, Laboratoire de Biogenèse Membranaire, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5200, 33076 Bordeaux Cedex, France.

Running title: Plastids in cell proliferation

Character count: 38,028

Footnotes

Author contributions: E.B., K.V., L.J., M.F., and S.K., designed research; E.B., K.V., J.W., M.G.-D., R.D.R., H.A., J.J., and T.B., performed research; E.B., K.V., J.W., M.G.-D., L.J., M.F., and S.K., analyzed data; E.B., K.V., M.G.-D., M.F., M.C.E.V.M., and S.K., wrote the paper.

The authors declare no conflict of interest.

To whom correspondence should be addressed at : Department of Plant Systems Biology, VIB-Ghent University, Technologiepark 927, B-9052 Gent, Belgium. Tel. +32 9 3313800; Fax: +32 9 3313809; E-mail: sergei.kushnir@psb.vib-ugent.be.

This article contains supporting information online at www.pnas.org/cgi/content/full/……

ABSTRACT

Plastids are DNA-containing organelles unique to plant cells. In *Arabidopsis*, one third of the genes required for embryo development encode plastid-localized proteins. To help understand the role of plastids in embryogenesis and postembryonic development, we have characterized proteins of the mTERF family, which in animal models comprises DNA-binding regulators of mitochondrial transcription. 11 of the 35 *Arabidopsis* mTERF proteins are plastid-localized. Genetic complementation shows that at least one plastidic mTERF, BELAYA SMERT' (BSM) is required for embryogenesis. The main postembryonic phenotype of genetic mosaics with the *bsm* mutation are severe abnormalities in leaf development. Mutant *bsm* cells are albino, compromised in growth and suffer defects in global plastidic gene expression. The *bsm* phenotype could be phenocopied by inhibition of plastid translation with spectinomycin. Plastid translation is essential for cell viability in dicotyledonous species such as tobacco, but not in monocotyledonous maize. Here, genetic interactions between *BSM* and the *ACC2* gene encoding for plastid homomeric acetyl-CoA carboxylase suggest that there is a functional redundancy in malonyl-CoA biosynthesis that permits *bsm* cell survival in *Arabidopsis*. Overall our results indicate that biosynthesis of malonyl-CoA and plastid-derived systemic growth-promoting compounds are the processes that link plant development and plastid gene expression.

malonyl-CoA | mosaics | mTERF | organellar gene expression |splicing

\body

Plastids are DNA-containing organelles of endosymbiotic origin that are defining for plant cells (1). Plastid homeostasis is continually monitored and gene expression regulated by reading the activities of tetrapyrrole biosynthesis, monitoring plastid gene expression machinery, the abundance of reactive oxygen species, plastidic redox status and the ATP/ADP balance of organelles (2). Consequently, this so-called retrograde signaling also affects the progression of developmental programs (3). Among 339 non-redundant *Arabidopsis* genes required for proper embryo formation, 108 encode plastid-targeted proteins (4). To study the role of plastid gene expression in embryogenesis and postembryonic development we have characterized proteins of the mTERF family, a group of proteins named after the human mitochondrial transcription termination factor mTERF1. These proteins have a modular architecture based on repetitions of a 30-amino-acid mTERF motif (5). Studies of the crystal structure of mTERF1 have revealed that these mTERF motifs associate to create a helical structure apparently involved in nucleic acid binding (6). Vertebrates have four mTERF paralogs. In humans mTERF1 is a sequence-specific DNA-binding protein responsible for mitochondrial transcription termination at a site adjacent to the mitochondrial rRNA genes (6). mTERF2 can bind to mitochondrial DNA (7) and at least in mouse appears to influence transcription (8). mTERF3 acts as a specific repressor of mammalian mtDNA transcription initiation *in vivo* (9).

Here, we show that 11 of the 35 annotated *Arabidopsis* mTERFs are targeted to plastids. Genetic complementation indicates that early embryo arrest is a characteristic phenotype of mutation in mTERF/At4g02990, while i*n vitro* cell culture and analysis of genetic mosaics reveals a postembryonic phenotype of a growth-compromised albino,

which suggested the gene name *BELAYA SMERT'* or "white death" in Russian. Exogenously supplied phytohormones and plastidic homomeric acetyl-CoA carboxylase ACC2 promoted the survival of *bsm* mutant cells, suggesting that biosynthesis of growth factors and malonyl-CoA underpin the non-autonomous and autonomous cell functions of plastids in *Arabidopsis* development.

Results

***Arabidopsis* has at least eleven plastid-localized mTERF proteins.** Flowering plants have the highest number of mTERF genes amongst eukaryotes (Table S1). Green fluorescent protein (GFP) fusions of all members of the *Arabidopsis* mTERF family showed that 11 mTERF proteins were targeted to chloroplasts and 17 to mitochondria, while one fusion was distributed equally between the nuclei and the cytoplasm (Fig. S1). Six gene fusions showed no GFP expression. The detection of conserved protein motifs revealed the presence of two motifs specific for Arabidopsis mTERF proteins targeted to the mitochondria (Fig. S2; Table S2). Integration of genome organization information (10) reveals species-specific gene family expansion via tandem duplication both in rice and *Arabidopsis* (with 45% and 25% of mTERF homologs in tandem, respectively) (Fig. S3). In contrast, plastid-targeted mTERFs belong to a single subtype, probably representing the ancestral family composition.

To determine the involvement of plastid *mTERF* in embryo development, we have surveyed T-DNA insertion mutant alleles. Plants with mutations in four genes developed seeds with arrested embryos (Table S3). The *bsm* mutant has been first identified and first characterized in detail.

***bsm* cells are albino and dependent on exogenous phytohormones.** In a progeny of *BSM/bsm* plants, 25% of the seeds had embryos arrested at the late globular stage of development (Fig. 1*A*). Immature mutant seeds remained white, indicating a failure of the endosperm and embryo to differentiate chloroplasts. These developmental defects were complemented by both wild-type BSM protein as well as its GFP and tandem affinity purification (TAP) tag fusions.

Dry seeds with arrested embryos were cultured *in vitro*. While wild type seeds germinated within 2 to 3 days, the first germination events with *bsm* seeds were only observed after 3 weeks, and lead to poor-growing, malformed albino seedlings (Fig. 1*B*). Therefore, to induce cell proliferation and *de novo* organogenesis these *bsm* seedlings were transferred to medium supplemented with cytokinin and auxin, resulting in stable shoot cultures that maintained a slow growth over several years (Fig. 1*C*). DNA gel blot hybridization showed that the shoot cultures were homozygous for the mutant *bsm* allele and contained chloroplast DNA (Fig. 1*D*). We concluded therefore that phytohormones promote the growth of albino *bsm* cells.

BSM gene deficiency affects processing and steady-state levels of plastid transcripts. Plastids possess two types of DNA-dependent RNA polymerases, a nuclear-encoded bacteriophage-type RNA polymerase (NEP) and a bacterium-type multi-subunit RNA polymerase (PEP). The PEP catalytic core requires products of the plastid genes *rpoA*, *rpoB*, *rpoC1* and *rpoC2* (11). Plastid-encoded class I genes such as *rbcL* are transcribed predominantly by PEP. Class II genes are transcribed by both PEP and NEP, whereas a few class III genes, such as *rpoB* and *accD* are thought to be transcribed exclusively by NEP (12). In *Arabidopsis*, rRNA genes belong to class II, but PEP is responsible for most transcriptional activity of *rrn16S* and *rrn23S* genes (12).

Mature 16S rRNA and 23S rRNA were not detectable in *bsm* cells (Fig. 1*E*). The PEP-dependent expression of the protein coding genes *rbcL* and *atpA* was similarly affected, while *clpP* transcripts levels increased (Fig. 1*E*). A similar response characterized albino tobacco plants mutated in PEP RNA polymerase (13).

To understand the changes in *clpP* transcript size, we analyzed it's splicing by RT-PCR. In *Arabidopsis*, the coding region of *clpP* is disrupted by two introns that belong to group IIb and group IIa classes. In albino shoots, RT-PCR showed that the second *clpP* group IIa intron was not spliced (Fig. 1*F*). The last gene exon codes for the 85-amino-acid residues that include two residues of a *clpP* catalytic triad, indicating that the mutant polypeptide is catalytically inactive (14). The plastid ClpPR protease complex is reported to be essential in both tobacco and *Arabidopsis* (15).

Protein-coding regions of plastid genes for the AtpF subunit of the ATP synthase and ribosomal proteins Rpl2 and Rps12 are interrupted by group IIa introns (16). The production of the *rps12* open reading frame also requires a *trans*-splicing event of a IIb intron. Results show that *rps12 trans*-splicing was normal but that *atpF*, *rpl2* and *rps12* group IIa introns were not spliced in *bsm* cells (Fig. 1*G*). These non-spliced transcripts encode mutant proteins, of which $AtpF_{1\text{-}48}$ and $Rpl2_{1\text{-}130}$ are most probably non-functional. Arabidopsis plants with defective *atpF* splicing are albino (17).

RNA editing increases the complexity of the plastid transcriptome. Editing can influence both the protein activity and be affected by pigment deficiency (18, 19). Sequencing of cloned cDNA-derived PCR fragments suggested that the *clpP* and *accD* transcripts were fully edited in the mutant.

Splicing of clpP group IIa intron is BSM-dependent. Plastomes of most flowering plants encode a MatK maturase that is thought to be a *trans*-acting splicing factor for

group IIa introns (20). Splicing of the *clpP* second intron is thought to be MatK-independent (16, 21). Inhibiting plastid translation in wild-type cells with the antibiotic spectinomycin (22) allowed us to replicate the growth-compromised, albino phenotype of *bsm* (Fig. 2 *A*-*C*). Spectinomycin abolished splicing of *atpF*, *rpl2*, *rps12* group IIa introns, but not of the *clpP* second intron (Fig. *2D*). Thus, processing of the *clpP* transcript is BSM-dependent and probably does not require MatK, suggesting a more direct role for BSM in *clpP* second intron splicing.

BSM is an mTERF-like protein. BSM protein sequence was aligned to create a structural model of the BSM protein with I-TASSER (23), taking advantage of the recently solved crystal structures of human mTERF1 (6). BSM is slightly larger than mTERF1, and the model suggests that BSM consists of a central core that is structurally homologous to the mammalian mTERFs flanked by N- and C-terminal extensions (Fig. 3*A*). The mTERF fold appears to have evolved to mediate protein-nucleic acid interactions. In this respect, the model predicts that BSM should be capable of binding double-stranded DNA and suggests a similar mode of interaction to that observed in human mTERF1 (Fig. 3*B*). Nevertheless, conservation of the critical residues that enable human mTERF1 to bind and recognize DNA (6) were not observed in BSM, probably reflecting a completely different nucleic acid binding specificity. Moreover, the N- and C-terminal extensions in BSM could conceivably confer additional functionalities to the protein, although it is also possible that they simply allow BSM to interact with a DNA sequence longer than the 21 base pairs contacted by human mTERF1. To verify these conclusions, bacterially produced 6xHis-BSM was refolded on Ni^{+}-beads and incubated with radioactively labeled, double-stranded restriction fragments of cloned *Arabidopsis* chloroplast DNA. While 6xHis-BSM bound DNA, no preferential retention of cpDNA

fragments was observed (Fig. 3*C*).

Next we characterized BSM protein fusions produced *in planta*. BSM-GFP was targeted to chloroplasts (Fig. 3*D*). When purified chloroplasts were treated with protease, BSM-TAP proteolysis was facilitated by non-ionic detergents, indicating that the protein is localized within plastids (Fig. 3*E*).

mTERFs are components of mitochondrial nucleoid in animals (7, 24). Chloroplast nucleoids are known to partition both into Triton-insoluble 20.000 x *g* chloroplast preparations (25), and in soluble >3MDa complexes of chloroplast (26). In cell extracts prepared using non-ionic detergents such as Triton X-100 or Nonidet P-40 (Fig. 3*F*), half of the BSM-TAP was retained in a soluble complex that did not enter the Blue-Native (BN) polyacrylamide gels at a >2 MDa cut-off. The remainder of the BSM-TAP partitioned into Triton-insoluble 20.000 x *g* chloroplast preparations. This partitioning was disrupted by digitonin and on BN gels, digitonin-extracted 70 kDa BSM-TAP co-migrated with a discrete band of 200 kDa (Fig. 3*F*), representing either a heteromeric complex or a homomeric trimer. Digitonin sensitivity indicates that BSM could be a component of a peripheral membrane complex. Notably, chloroplast nucleoids are thought to be anchored to the membranes (27).

Previous studies have shown that mTERF pTAC15/At5g54180 co-purifies with transcriptionally active plastid chromosomes (28). BSM was identified as a component of the soluble nucleoid fraction (26). Plants with a mutation in mTERF At2g03050 have reduced rRNA levels and protein synthesis rates in plastids (29). Taken together, these data suggest that BSM is likely to be implicated in organellar gene expression, possibly in a similar way to the mTERF proteins in metazoans.

Genetic interactions between *BSM* and *ACC2*. Tobacco proteins encoded by the plastid genes *ycf1, ycf2, clpP, accD, matK* and several components of the translational machinery have been shown to be essential for cell viability (30,31). Our data suggest that a major deficiency of protein translation exists in *bsm* plastids. To understand why *bsm* cells are still viable, we applied genetic tests to investigate a possible functional redundancy in the *accD* gene function.

The plastid *accD* gene encodes the β-carboxyltransferase subunit of the heteromeric acetyl-CoA carboxylase (He-ACC), while the other three subunits are encoded for by the nuclear *CAC* genes. He-ACC produces malonyl-CoA that is used for *de novo* biosynthesis of fatty acids, which in plant cells occurs almost exclusively in plastids (32). However, since *bsm* cells are viable, mutant plastids most probably have sufficient malonyl-CoA for fatty acid synthesis. Indeed, analysis of fatty acid composition and content of albino *bsm/bsm* cells showed that there were no significant differences between mutant and wild type cells apart from lower levels of polyunsaturated C_{18} fatty acid in the mutant, and a correspondingly higher level of stearic acid (18:0) and oleic acid (18:1) (Fig. S4).

Malonyl-CoA in plastids can also be synthesized by the eukaryotic type, homomeric ACC (Ho-ACC) (32). Analysis of sequenced plant genomes indicates that three groups of plant species exist regarding capacity for malonyl CoA biosynthesis (Fig. 4*A*). Some species rely exclusively on a He-ACC (grapevine type), others only on a Ho-ACC (corn type), and some use both enzymes (canola type) (33). The *Arabidopsis* genome contains two genes for Ho-ACC that are positioned on chromosome 1 in a tandem duplication. The ACC1/GURKE/PASTICCINO3 is a cytosolic enzyme (34). As compared to ACC1, the ACC2 encoded by the gene *At1g36180* is 107-amino-acid residues longer at the N terminus. The current gene model (The Arabidopsis Information

Resource) indicates that *At1g36180* can produce 2 proteins (Fig. 4*B*), but that only the 2356-amino-acid-long ACC2.1 is expected to be fully active. The functional significance of the alternatively-spliced, C-terminally truncated ACC2.2 is unclear. We produced $ACC2_{1-204}$-GFP and $ACC2_{1-544}$-GFP in leaf protoplasts and both protein fusions localized to the chloroplasts (Fig. 4*C*), suggesting that ACC2.1 is indeed a plastid protein.

To assess whether *BSM* and *ACC2* interact genetically, we confirmed two T-DNA insertion mutations in the *ACC2* gene as annotated in the SALK Institute mutant collection (35). T-DNAs were located in the 21st exon and the 27th exon of the mutant *acc2-2* and *acc2-1* gene alleles, respectively (Fig. 4*B*; Fig. S5). Homozygous *acc2* plants had no obvious mutant phenotype when compared to the wild type, indicating that He-ACC is the major plastid acetyl-CoA carboxylase in *Arabidopsis*. Embryos from flowering *acc2-1/acc2-1;bsm/BSM* or *acc2-2/acc2-2;bsm/BSM* plants were compared with *ACC2/ACC2;bsm/BSM* plants that had been selected among the same F2 segregating populations. The *acc2-1/bsm* and *acc2-2/bsm* embryo arrested earlier than the *ACC2/bsm* embryos (Fig. 4*D*) and, in addition, seeds with arrested *acc2-1/bsm* embryos did not germinate *in vitro* (Fig. 4*E*), indicating synthetic lethality between the *bsm* and *acc2* mutations. This indicates therefore that cytosolic ACC1 cannot compensate for a deficiency in plastid malonyl-CoA biosynthesis in *bsm* plastids and that functional redundancy in malonyl-CoA biosynthesis in plastids of *Arabidopsis* promotes *bsm* cell survival.

***BSM* mosaics reveal a non-cell-autonomous role of plastids in development.** The *bsm* cell proliferation defects were partially alleviated by external application of cytokinin and auxin. Since phytohormones have a systemic action, we asked whether our results were indicative of a non-cell-autonomous function of plastids during development. To test this,

we characterized postembryonic mutant phenotypes in mosaic plants generated by heat-inducible CRE-*loxP* recombination (Fig. 5*A*).

When *BSM* gene loss was induced in 4-week-old plants grown on soil, leaves preformed at the time of heat-shock application remained green (Fig. 5 *B*). Young leaf primordia at the time of the heat shock, however, developed a proximo-distal gradient of pigment deficiency (Fig. 5*C*). The subsequent three to four leaf primordia developed serrated, small albino leaves and the youngest emerging leaves were rod-like albino structures (Fig. 5 *D*). Mosaics were photoautotrophic plants that often developed large masses of short albino shoots (Fig. 5*E*). Heat stress applied to 5- to 6-week-old plants resulted in the development of mosaics with albino inflorescence shoots (Fig. 5*F*). Albino flowers on primary shoots had sepals, petals, anthers, and carpels while in flowers of side shoots, only filamentous structures usually developed instead of the internal whorls of petals and anthers. No seeds were obtained following either self-pollination or cross-pollination with wild-type pollen, indicating major problems with gametophyte development.

To characterize the timing of BSM decay after induction of gene deletion, the fate of BSM-TAP was analyzed. BSM protein had a half-life of 3 days in green leaves of mosaics (Fig. 5*G*) but was undetectable in albino tissues. Therefore, the relative stability of BSM protein present at the time of heat shock application maintains wild type functionality in green tissues, suggesting that wild-type plastids in mosaic plants produce growth-promoting substances that had not been supplied to the *in vitro* grown *bsm* cells.

Characterization of foliar cell types showed that significantly less mesophyll cells were present in albino leaves (Fig. S6*A*), and while stomata complex cells and trichomes were recognizable, the epidermis pavement cells did not attain a wild-type interlocking jigsaw puzzle shape (Fig. S6*B*). In albino leaves, mutant *bsm* plastids differentiate as

double membrane bound, often doughnut-shaped, small organelles without a differentiated membrane system (Fig. S6*C*). Albino leaves were also found to accumulate several abundant polypeptides (Fig. S6*D*), identified as thioglucoside glucohydrolases TGG1 and TGG2, or myrosinases. In *Arabidopsis*, myrosinases are specific markers of the myrosin cells of vascular bundles (36) and guard cells of the stomata complex (37). TGG1 is also required for key abscisic acid responses of guard cells (37). The observed proteome phenotype was not however caused by ectopic gene activation (Fig. S6 *E* and *F*).

Discussion

Plastids are the metabolic factories of plant cells (38). Disruption of amino acid, vitamin or nucleotide biosynthesis in plastids generally leads to arrested embryo development (39). We show that loss of plastidic BSM protein leads to an arrest of embryogenesis and abnormal postembryonic development. *bsm* cells are albino and defective in photosynthesis, but can still be maintained *in vitro* on a minimal medium, indicating that they are not auxotrophic for basic metabolites.

Our results show that there is negligible protein translation in *bsm* plastids. Since seed abortion at the transition stage of embryogenesis in *Arabidopsis* is often caused by mutations in chloroplast ribosomal proteins (39) and aminoacyl-tRNA synthetases (40), defects in plastid gene expression could be the cause of arrested *bsm* embryo development. Although plastid translation is essential for cell viability in tobacco (30), this is not the case in monocotyledonous grasses (1,20). The dicot/monocot dichotomy is thought to depend on the plastome gene content (1), because grass genomes do not encode for Ycf1, Ycf2 and AccD proteins. Indeed, it has been proposed that *accD* is the

only truly essential gene of the tobacco plastome (31), implying that the loss of plastid translation results in auxotrophy for malonyl-CoA in dicots.

Genetic interactions between *BSM* and *ACC2* genes suggest that functional redundancy in plastid malonyl-CoA biosynthesis underlies *bsm* cell survival. Thus alternative modes of plastid malonyl-CoA biosynthesis (Fig. 4*A*) can explain the different sensitivities of dicot plant species to genetically- or pharmacologically-induced loss of plastid translation. Metazoans depend on Ho-ACC for their malonyl-CoA. Ho-ACC operates in the plant cell cytosol (34). The evolutionary forces that have lead to the retention of either He-ACC and/or Ho-ACC in plastids are not understood. However, biochemical enzyme properties (32) or protein neofunctionalization can have influenced the capacity for ecological adaptation. Interestingly, the seed oil-producing crops canola and soybean have retained genes for both the He-ACC and Ho-ACC plastid isoforms.

Materials and methods

Plant material and growth conditions. The ecotypes and genetic markers were C24, C24 *bsm/BSM* (this study), Col *acc2-1/acc2-1*, and Col *acc2-2/acc2-2* (SALK Institute mutant collection and this study). Plants were grown as described (41). Mutant cell/shoot *bsm/bsm* lines were maintained on a shoot-inducing medium (MS salts, 20 g/L sucrose, 1 mg/L 6-benzyladenine, and 0.1 mg/L α-naphthalene acetic acid). To induce mosaics, plants were heat-stressed at 37°C.

Transient assays and microscopy analysis. GFP protein fusions were expressed transiently (http://genetics.mgh.harvard.edu/sheenweb/). Tissue sections were analyzed by light and transmission electron microscopy (41); embryos by Nomarski optics

(http://www.seedgenes.org/Tutorial.html); protoplasts in isotonic solutions, epidermis peels or hand sections of leaves in water by laser scanning confocal microscopy (41). Surface morphology was examined with a table-top electron microscope TM-1000 (Hitachi).

Other methods. Sequences were analyzed as described (10). GATEWAY recombinational cloning was used routinely (Invitrogen). mTERF primer sequences are in a Table S4. The binary vector for mosaic analysis was pB6Act1 (42). Lipids were extracted from mutant and wild-type calli and their fatty acid compositions and contents were quantified via their methyl esters as described (43). The NativePAGE Novex Bis-Tris Gel system (Invitrogen) was used for BN-PAGE.

TAP-tags were visualized using the peroxidase-antiperoxidase soluble complex P1291 (Sigma-Aldrich) at a 1/1000 dilution. The GFP antigen was detected with 1/2000-diluted primary rabbit anti-GFP antibody AB3080 (Chemicon International) and 1/10000-diluted secondary ECL anti-rabbit IgG horseradish peroxidase-linked whole antibody from donkey NA934V (GE-Healthcare). Peroxidase activity was detected with the Western Lightning Chemiluminescence Reagent Plus (Perkin Elmer Life) and high-performance chemiluminescent film, Hyperfilm ECL (GE-Healthcare). ^{32}P-labeled DNA probes were prepared with nucleic acid labeling Readiprime II random prime labeling system (GE-Healthcare). Gene fragments for labeling were amplified by PCR with gene-specific primers (Table S4).

ACKNOWLEDGEMENTS. We thank Hans-Peter Braun and Mark Davey for critical reading of the manuscript; Mansour Karimi and Bernd Reiss for plasmids; Duncan Rix and Dow AgroSciences for chemicals; John Ohlrogge, Christian Schmitz-Linneweber,

Ian Small, Thomas Börner for discussions; and Martine De Cock and Karel Spruyt for help in preparing the manuscript. This work was supported by the VIB Institutional Valorization Funds; the European Union project REGIA (to S.K.); a grant from the Deutsche Forschungsgemeinschaft (to M.F.). M.G-D was supported by R00 ES015421. K.V. is a postdoctoral fellow of the Research Foundation-Flanders. This study is dedicated to the memory of Katrien van Leeuwen.

References

1. Stern DB, Hanson MR, Barkan A (2004) Genetics and genomics of chloroplast biogenesis: maize as a model system. *Trends Plant Sci* 9:293-301.
2. Woodson JD, Chory J (2008) Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet* 9:383-395.
3. López-Juez, E (2009) Steering the solar panel: plastids influence development. *New Phytol* 182:287-290.
4. Hsu S-C, Belmonte MF, Harada JJ, Inoue K (2010) Indispensable roles of plastids in *Arabidopsis thaliana* embryogenesis. *Curr Genomics* 11:338-349.
5. Roberti M, et al. (2009) The MTERF family proteins: mitochondrial transcription regulators and beyond. *Biochim Biophys Acta* 1787:303-311.
6. Yakubovskaya E, Mejia E, Byrnes J, Hambardjieva E, Garcia-Diaz M (2010) Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell* 141:982-993.
7. Pellegrini M, et al. (2009) MTERF2 is a nucleoid component in mammalian mitochondria. *Biochim Biophys Acta* 1787:296-302.
8. Wenz T, Luca C, Torraco A, Moraes CT (2009) mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription. *Cell Metabolism* 9:499-511.
9. Park CB, et al. (2007) MTERF3 is a negative regulator of mammalian mtDNA transcription. *Cell* 130:273-285.
10. Proost S, et al. (2010) PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell* 21:3718-3731.
11. Suzuki JY, et al. (2004) Affinity purification of the tobacco plastid RNA polymerase and *in vitro* reconstitution of the holoenzyme. *Plant J* 40:164-172.

12. Ishizaki Y, et al. (2005) A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J* 42:133-144.

13. Krause K, Maier RM, Kofer W, Krupinska K, Herrmann RG (2000) Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. *Mol Gen Genet* 263:1022-1030.

14. Peltier J-B, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ (2001) Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. *J Biol Chem* 276:16318-16327.

15. Zybailov BG, et al. (2009) Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism. *Mol Cell Proteomics* 8:1789-1810.

16. Tillich M, Krause K (2010) The ins and outs of editing and splicing of plastid RNAs: lessons from parasitic plants. *New Biotechnol* 27:256-266.

17. Asakura Y, Barkan A (2006) Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol* **142:**1656-1663.

18. Tseng C-C, et al. (2010) Editing of *accD* and *ndhF* chloroplast transcripts is partially affected in the *Arabidopsis vanilla cream1* mutant. *Plant Mol Biol* 73:309-323.

19. Robbins JC, Heller WP, Hanson MR (2009) A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the *Arabidopsis* chloroplast *accD* transcript. *RNA* **15**:1142-1153.

20. Vogel J, Börner T, Hess WR (1999) Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res* 27:3866-3874.

21. Zoschke R, et al. (2010) An organellar maturase associates with multiple group II introns. *Proc Natl Acad Sci USA* 107:3245-3250.

22. Zubko MK, Day A (1998) Stable albinism induced without mutagenesis: a model for ribosome-free plastid inheritance. *Plant J* 15:265-271.

23. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725-738.

24. Bogenhagen DF, Rousseau D, Burke S (2008) The layered structure of human mitochondrial DNA nucleoids. *J Biol Chem* 283:3665-3675.

25. Phinney BS, Thelen JJ (2005) Proteomic characterization of a Triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. *J Proteome Res* 4:497-506.

26. Olinares PDB, Ponnala L, van Wijk KJ (2010) Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Mol Cell Proteomics* 9:1594-1615.

27. Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. *EMBO J* 12:555-561.

28. Pfalz J, Liere K, Kandlbinder A, Dietz K-J, Oelmüller R (2006) pTAC2,-6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18:176-197.

29. Meskauskiene R, et al. (2009) A mutation in the Arabidopsis mTERF-related plastid protein SOLDAT10 activates retrograde signaling and suppresses $^1O_2$-induced cell death. *Plant J* 60:399-410.

30. Rogalski M, Ruf S, Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. *Nucleic Acids Res* 34:4537-4545.

31. Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco plastid *acc*D gene is essential and is required for leaf development. *Plant J* 44:237-244.

32. Ohlrogge J, Browse J (1995) Lipid biosynthesis. *Plant Cell* 7:957-970.

33. Schulte W, Töpfer R, Stracke R, Schell J, Martini N (1997) Multi-functional acetyl-CoA carboxylase from *Brassica napus* is encoded by a multi-gene family: Indication for plastidic localization of at least one isoform. *Proc Natl Acad Sci USA* 94:3465-3470.

34. Baud S, et al. (2004) *gurke* and *pasticcino3* mutants affected in embryo development are impaired in acetyl-CoA carboxylase. *EMBO Rep* 5:515-520.

35. Alonso JM, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301:653-657.

36. Ueda H, et al. (2006) AtVAM3is required for normal specification of idioblasts, myrosin cells. *Plant Cell Physiol* 47:164-175.

37. Zhao Z, Zhang W, Stanley BA, Assmann SM (2008) Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20:3210-3226.

38. Neuhaus HE, Emes MJ (2000) Nonphotosynthetic metabolism in plastids. *Annu Rev Plant Physiol Plant Mol Biol* 51:111-140.

39. Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D (2011) Identification of nuclear genes encoding chloroplast-localized proteins required for embryo

development in *Arabidopsis thaliana*. *Plant Physiol*:in press (doi: 10.1104/pp.110.168120).

40. Berg M, Rogers R, Muralla R, Meinke D (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. *Plant J* 44:866-878.

41. Kushnir S, et al. (2001) A mutation of the mitochondrial ABC transporter Sta1 leads to dwarfism and chlorosis in the Arabidopsis mutant *starik*. *Plant Cell* 13:89-100.

42. Babiychuk E, et al. (2008) Allelic mutant series reveal distinct functions for *Arabidopsis* cycloartenol synthase 1 in cell viability and chloroplast differentiation. *Proc Natl Acad Sci USA* 105:3163-3168.

43. Babiychuk E, et al. (2003) *Arabidopsis* phosphatidylglycerophosphate synthase 1 is essential for chloroplast differentiation, but is dispensable for mitochondrial function. *Plant J* 33:899-909.

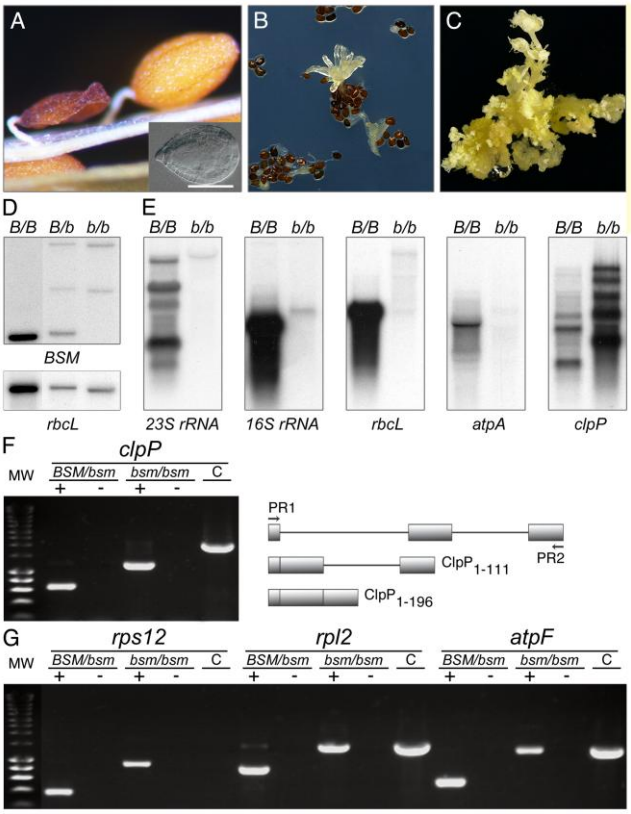


Fig. 1. Mutant *bsm* phenotype and plastid gene expression.
(A) Shrunken dry mutant seed, arrested embryo (inset) and a wild-type seed. *(B)* Asynchronous germination of the seeds with arrested embryos. *(C)* Mixed callus/shoot *bsm* culture on a medium supplemented with phytohormones. *(D)* Albino cells are *bsm/bsm* homozygous. *Hin*dIII-digested genomic DNA from albino cells (*b/b*) was compared to DNA of green homo- (*B/B*) and heterozygous (*B/b*) plants by hybridization analysis with radioactive probes detecting nuclear *BSM* and plastid *rbcL* genes. *(E)* Analysis of plastid transcripts. RNA was extracted from green plants (*B/B*) or *bsm* albino shoot cultures (*b/b*). Analyzed plastid genes were *rrn16S* and *rrn23S* for ribosomal RNAs (*23S rRNA*, *16S rRNA)*; the large subunit of Rubisco (*rbcL*); the α-subunit of the plastidial ATPase (*atpA*); ClpP subunit of ClpPR protease (*clpP*). *(F) bsm* cells do not splice the second *clpP* group IIa intron. DNA products were amplified by PCR with specific primers (PR1 and PR2) for *clpP/AtCg00670*. The templates were the first-strand cDNA (+); RNA, to control for DNA contamination (-); genomic DNA, to provide a size marker for a PCR fragment representing fully non-spliced *clpP* transcript (C). Drawing illustrates the exon-intron gene structure; organization of processed transcripts and the lengths of encoded polypeptides. *(G) atpF*, *rpl2* and *rps12* intron splicing. The labeling is as in *(F).* The first, group IIb intron of *rps12* is encoded by two different genes, which requires a *trans*-splicing event for its processing. This is a reason why PCR product is generated only with cDNA, but not with DNA as a template.

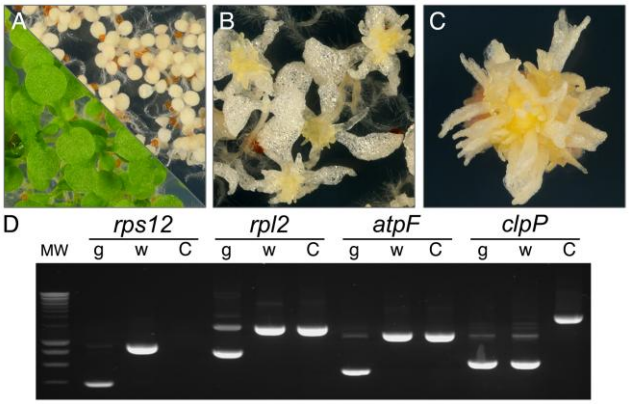


Fig. 2. Analysis of spectinomycin-treated plants of wild-type genome composition.

(A) Effect of spectinomycin on wild type plants. Seeds were placed on a medium without or with 500 mg/L of spectinomycin. Seven-day-old seedlings are shown. Treated seedlings are albino, because inhibition of plastid translation prevents biogenesis of chlorophyll-containing photosynthetic complexes. *(B*) Ninety-day-old plantlets developed in the presence of spectinomycin. While by this time green plants completed life cycle, albino seedlings were comparable to *bsm* seedlings of similar age (Fig. 1*B*). *(C)* Thirty-day-old shoot in vitro culture. Transfer of spectinomycin-induced albinos to a medium supplemented with phytohormones promoted the growth and established albino shoot cultures composed from undifferentiated calli and *de novo* developed shoots. *(D)* RT-PCR analysis of splicing. Splicing analysis of *atpF*, *rpl2*, *rps12*, and *clpP* introns is detailed in Fig. *1F* and *G*. RNA was extracted from albino shoots (w) or green shoots (g). Control PCR on genomic DNA (C).

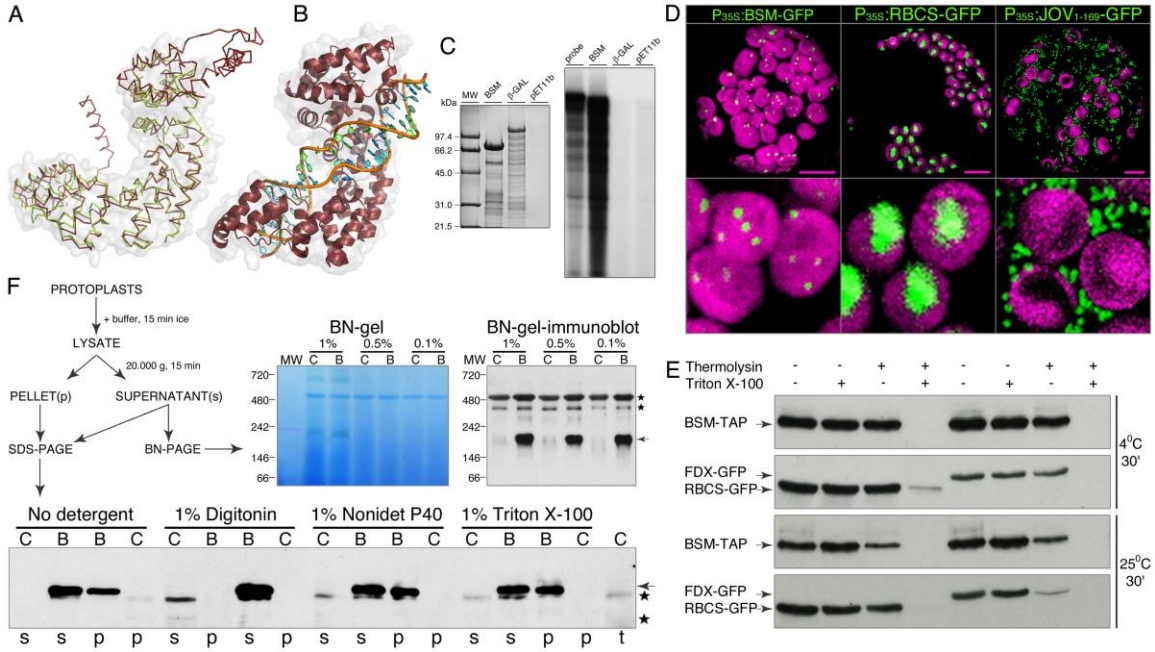


Fig. 3. Structural homology modeling and biochemical properties of BSM.
(A) Overlay between the BSM model and human mTERF1. The C-α trace of A model of BSM constructed with I-TASSER is shown (red) in an overlay with human mTERF1 (yellow, PDB 3MVA). The central core of the BSM protein is predicted to fold very similarly to mTERF1. The rmsd was of 2.28 Å for 280 C-α atoms. The molecular surface of human mTERF1 is shown transparent. *(B)* Predicted DNA binding mode for BSM. The central core of BSM is shown in ribbon representation (red) together with the DNA molecule from the human mTERF1 cocrystal (3MVA). The model indicates that the BSM predicted fold is consistent with double-strand DNA binding. The molecular surface of the BSM core is shown transparent. The two DNA strands are shown in green or blue with an orange coil. *(C)* DNA binding. Six-histidine tagged BSM (BSM) and bacterial β-galactosidase (β-GAL) were produced in *Escherichia coli* and purified on Ni^{+}-resin. The empty pET11b plasmid was used as control (pET11b). Beads with absorbed proteins were incubated with ^{32}P-labeled *Not*I-*Nco*I restriction fragments (probe) of cloned chloroplast DNA of *Arabidopsis*. Beads were washed, bound DNA was purified and resolved by agarose gel electrophoresis. *(D)* BSM-GFP localization in chloroplasts. The control protein fusions RBCS-GFP and JOV_{1-169}-GFP visualized chloroplast stroma and mitochondria. Shown are the entire protoplast and a close-up of chloroplasts from the same cell. Auto-fluorescence of chlorophyll visualize chloroplasts, which are false colored in magenta. Green is GFP fluorescence. Bars = 100 µm. *(E)* BSM-TAP is localized inside of chloroplasts. Percoll-purified chloroplasts of transgenic plants that co-produce BSM-TAP and small subunit of Rubisco (RBCS-GFP) or ferredoxin (FDX-GFP) were incubated in an isotonic buffer at 4°C or at 25°C for 30 min. Samples were differentially supplemented with a protease thermolysin and non-ionic detergent Triton X-100. GFP fusions of proteins known to be localized in chloroplast stroma were used as controls and biological repeats. *(F)* BSM-TAP is involved in intermolecular interactions. Live protoplasts from leaves of wild-type plants (C) or genetically complemented plants (B) that produced the BSM-TAP were used as starting material. On the 3-12% BN gel and its blot with separated samples from protoplasts lysed in the presence of 1, 0.5 or 0.1% of digitonin, the abundant protein complex migrating at approximately 500 kDa corresponds to a heteromeric ribulose-1,5-bisphosphate carboxylase-oxygenase (Rubisco) and provides a gel loading control. Asterisks mark non-specific signals also present in control, arrow points BSM-TAP.

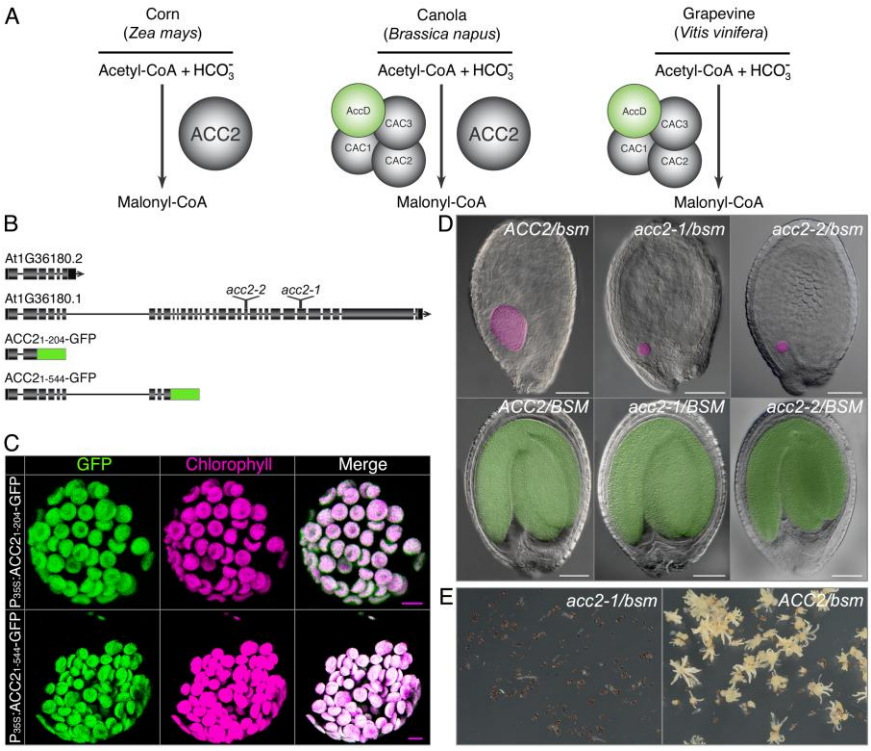


Fig. 4. Genetic interactions between *BSM* and *ACC2*.

(A) Modes of malonyl-CoA biosynthesis in plastids of higher plants. ACC2 – an eukaryote-type, homomeric acetyl-CoA carboxylase encoded by genes in the nuclear genome. Prokaryote-type, heteromeric acetyl-CoA carboxylase is composed of four subunits. CAC1, CAC2 and CAC3 are the nuclear genome-encoded polypeptides. AccD is a product of a plastid *accD* gene. *(B)* The *ACC2* gene model in the genome annotation. Two alternative transcripts are encoded by the *ACC2 Arabidopsis* gene locus. Positions of the T-DNAs in the analyzed mutant alleles *acc2-1* and *acc2-2* are indicated. *(C)* Localization of $ACC2_{1\text{-}204}$-GFP and $ACC2_{1\text{-}544}$-GFP in chloroplasts of *Arabidopsis*. Bar = 10 μm. *(D)* Embryos of the indicated genotypes in almost mature seeds that are turning brown. Images were false-colored in accordance with the *BSM* gene state (magenta and green for the *bsm* mutant and the *BSM* wild-type embryos, respectively). *(E)* Seeds with arrested *acc2-1/acc2-1;bsm/bsm* and *ACC2/ACC2;bsm/bsm* embryos after 4 weeks on a shoot-inducing medium. Arrested *acc2-1/acc2-1;bsm/bsm* embryos fail to develop shoots.

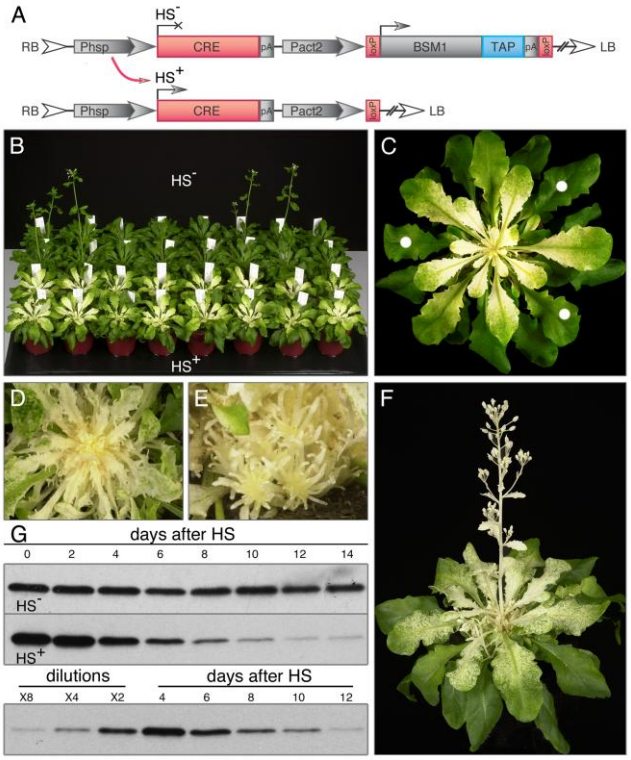
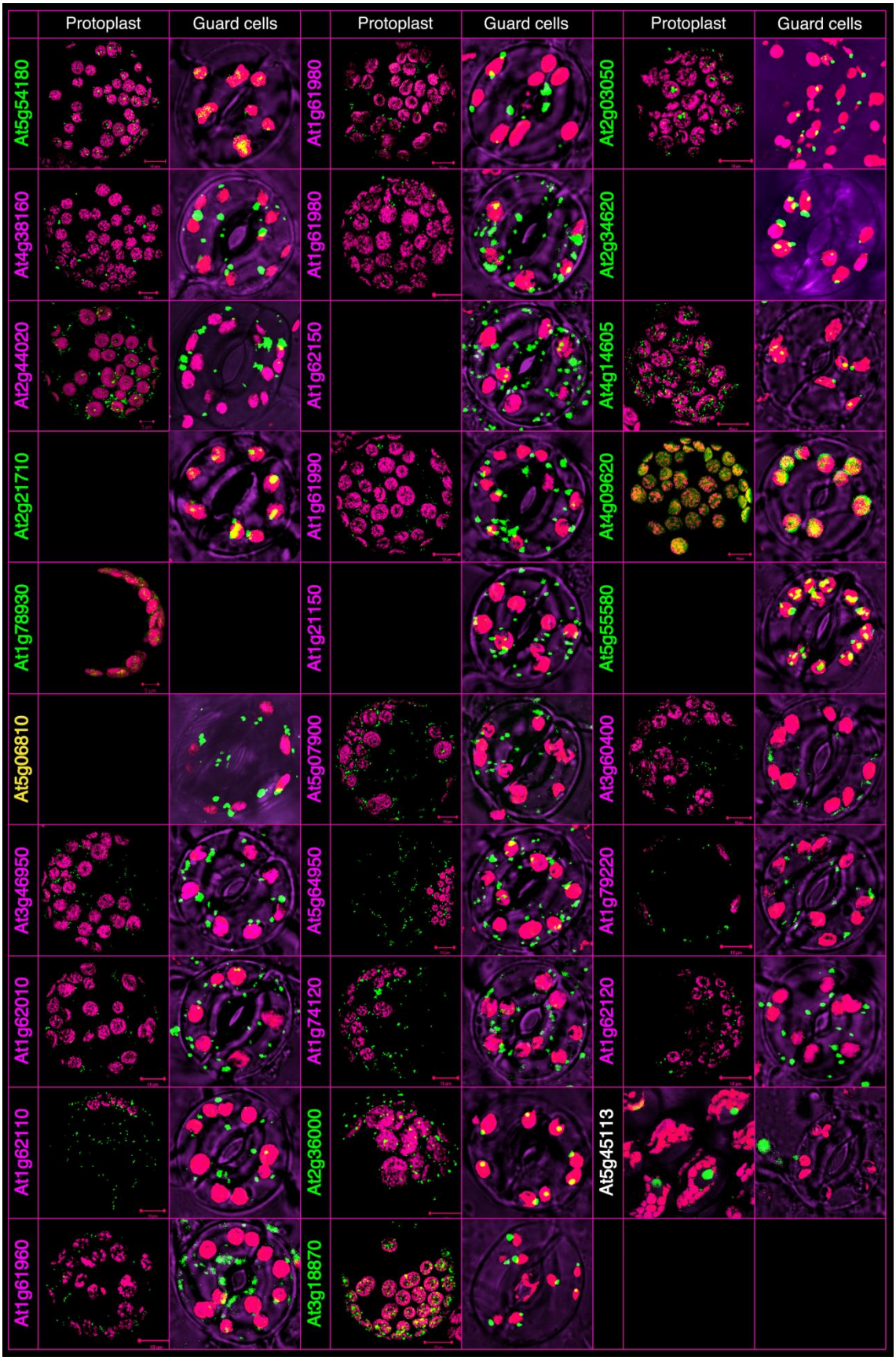


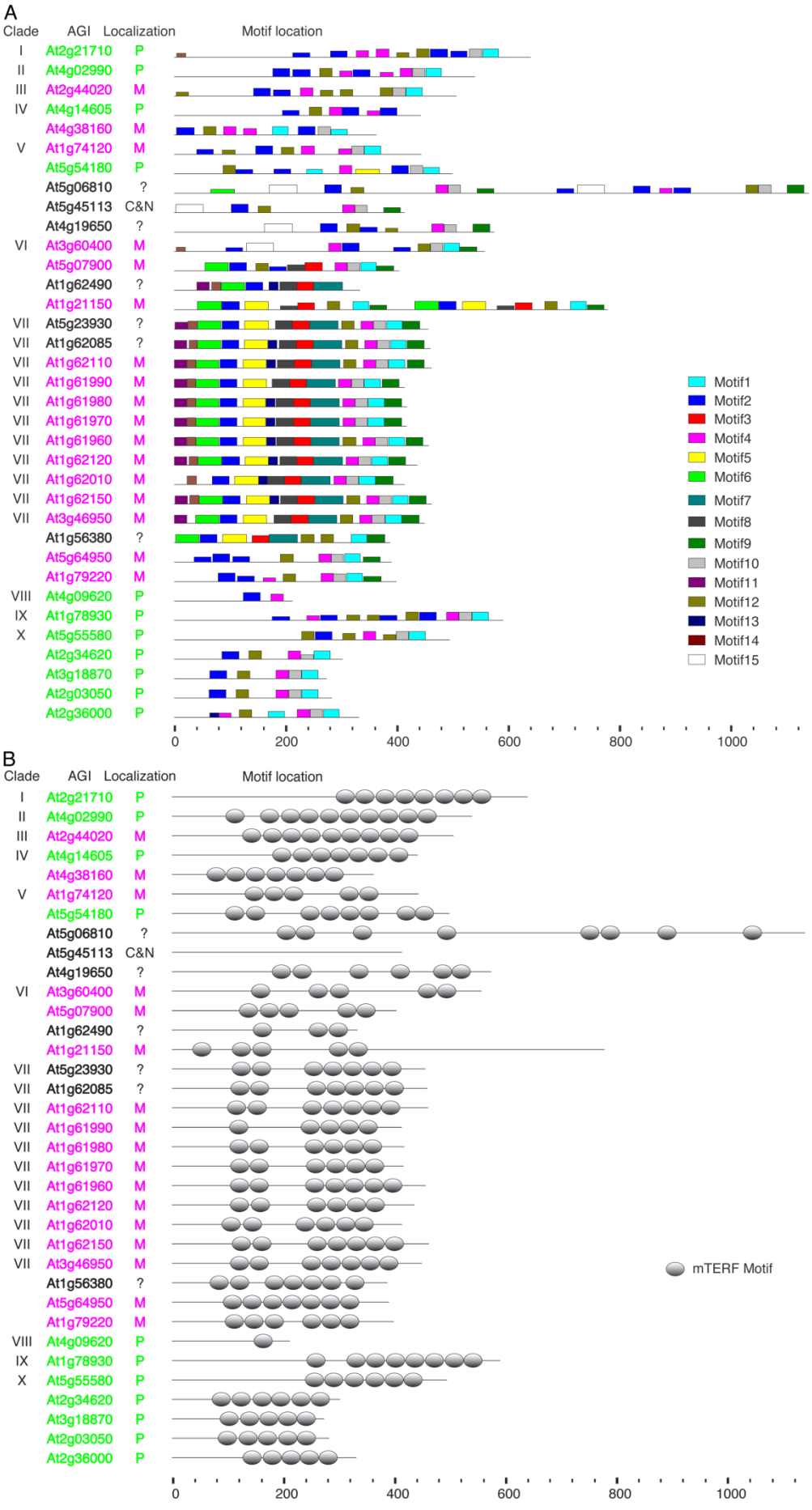
Fig. 5. Mosaic plants.
(A) Design of the gene complementation constructs. Right (RB) and left (LB) borders delineate the T-DNA integrated into the plant nuclear genome. ORFs are: the *BSM* ORF (BSM1); TAP-tag (TAP); CRE recombinase (CRE). The *cis*-elements are: (Pact2) - promoter sequences of the *Arabidopsis ACTIN2* (*At3g18780*); (pA) - the nopaline synthase polyadenylation signals; (loxP) - *loxP* recombination substrates. Heat stress-inducible expression was directed by the (Phsp) *Arabidopsis HSP18.2* promoter. Heat stress (37°C) induced CRE excise ORF from the T-DNA *in planta*. Eventual decay of the *BSM* transcript and protein leads to the loss of the gene function. The *BSM*, *BSM-TAP* and *BSM-GFP* ORF's were tested in this experimental design. For simplicity, only *BSM-TAP* ORF is drawn here. *(B)* Mosaic plants in the two front rows (HS+) and non-stressed siblings used as control (HS-) in the back rows. *(C)* Close-up of a plant from *(B)*. Green leaves indicated with dots were harvested for the time-course analysis in *(G)*. *(D,E)* The same plant as in *(C)* 2 weeks and 4 weeks later, respectively. Abnormal leaf shapes and massive shoot proliferation are illustrated. *(F)* Mosaic developed from a heat-stressed plant that transitioned to flowering. *(G)* Protein gel-blot analysis of the BSM-TAP from green leaves of mosaic (HS+) and control (HS-) plants. The type of leaf material harvested for the time course is illustrated in *(C)*. To estimate the half-life time as shown on the lower panel, protein extracts from plants 4 days after HS were serially diluted 2-fold with an extract from a wild-type plant.



Supplementary Figure S1. Localization of *Arabidopsis* mTERF proteins.

Gene-coding regions of the *Arabidopsis mTERF* genes were PCR-amplified with primers

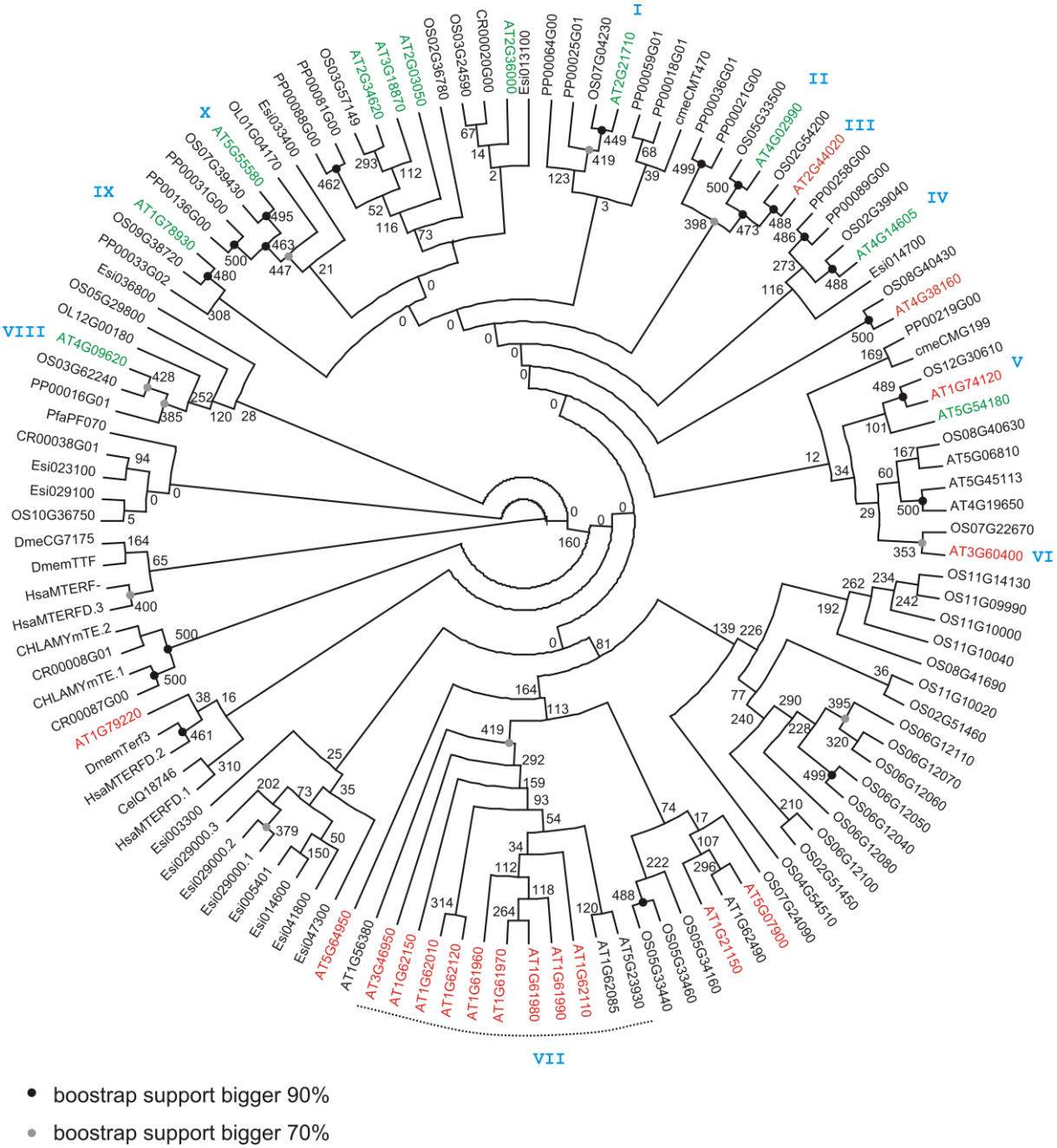
(Supplementary Table S4) designed according to TAIR genome annotation v.9 and used to prepare a collection of entry clones that were converted into expression clones by the GATEWAY recombinational cloning. Protein-GFP fusions were expressed transiently in protoplasts or stably in transgenic *Arabidopsis* plants. GFP fluorescence (green) and autofluorescence of chlorophyll in chloroplasts (magenta) are shown in protoplasts from transient expression assays (Protoplast) or in guard cells of transgenic plants (Guard cells). mTERFs are identified by their AGI codes in the annotated Arabidopsis genome and colored according to protein localization: in the chloroplasts (green); in the mitochondria (magenta); in the cytosol/nuclei (white). At5g06810 is in yellow and represents a test for a possible gene mis-annotation. We cannot detect GFP fluorescence in cells transformed with At5g06810-GFP prepared with primers based on the annotated gene sequences. Analysis of conserved protein motifs (Fig. S2) suggested that the annotated *At5g06810* locus actually consists of two tandemly duplicated genes that were fused in one by the genome annotation software. With this assumption in mind, a new reverse PCR primer was used to amplify the coding region of the first putative gene, and the resulting GFP fusion localized to the mitochondria. This result suggested that (i) the lack of GFP localization data for some genes analyzed in our study could be due to the mis-annotations of mTERF genes in the *Arabidopsis* genome; (ii) analysis of conserved protein motifs is indicative of protein localization (Fig. S2).


A
Clade AGI Localization Motif location
I At2g21710 P
II At4g02990 P
III At2g44020 M
IV At4g14605 P
At4g38160 M
V At1g74120 M
At5g54180 P
At5g06810 ?
At5g45113 C&N
At4g19650 ?
VI At3g60400 M
At5g07900 M
At1g62490 ?
At1g21150 M
VII At5g23930 ?
VII At1g62085 ?
VII At1g62110 M
VII At1g61990 M
VII At1g61980 M
VII At1g61970 M
VII At1g61960 M
VII At1g62120 M
VII At1g62010 M
VII At1g62150 M
VII At3g46950 M
At1g56380 ?
At5g64950 M
At1g79220 M
VIII At4g09620 P
IX At1g78930 P
X At5g55580 P
At2g34620 P
At3g18870 P
At2g03050 P
At2g36000 P
Motif1
Motif2
Motif3
Motif4
Motif5
Motif6
Motif7
Motif8
Motif9
Motif10
Motif11
Motif12
Motif13
Motif14
Motif15
0
200
400
600
800
1000
B
Clade AGI Localization Motif location
I At2g21710 P
II At4g02990 P
III At2g44020 M
IV At4g14605 P
At4g38160 M
V At1g74120 M
At5g54180 P
At5g06810 ?
At5g45113 C&N
At4g19650 ?
VI At3g60400 M
At5g07900 M
At1g62490 ?
At1g21150 M
VII At5g23930 ?
VII At1g62085 ?
VII At1g62110 M
VII At1g61990 M
VII At1g61980 M
VII At1g61970 M
VII At1g61960 M
VII At1g62120 M
VII At1g62010 M
VII At1g62150 M
VII At3g46950 M
At1g56380 ?
At5g64950 M
At1g79220 M
VIII At4g09620 P
IX At1g78930 P
X At5g55580 P
At2g34620 P
At3g18870 P
At2g03050 P
At2g36000 P
mTERF Motif
0
200
400
600
800
1000

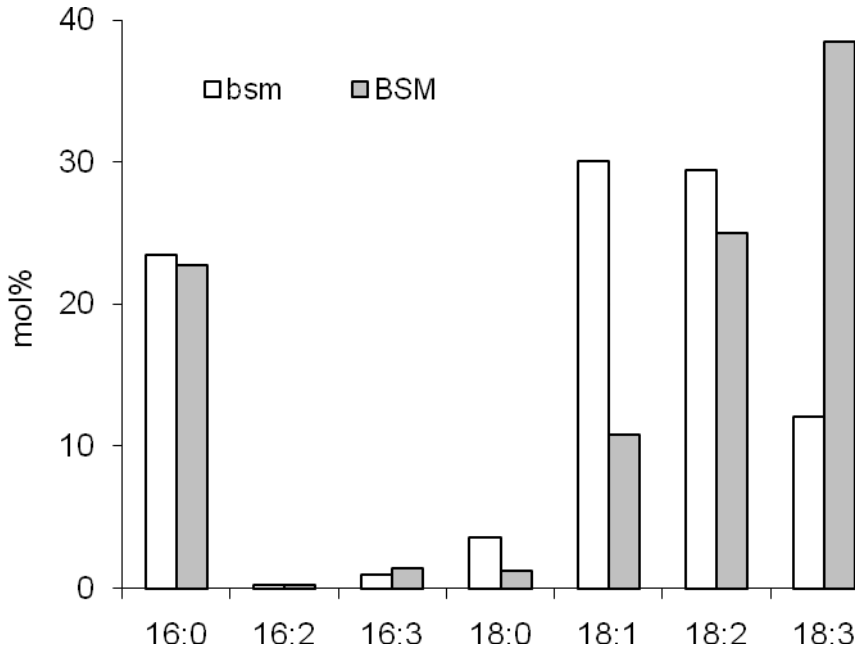
Supplementary Figure S2. Schematic representation of conserved protein motifs in Arabidopsis mTERF proteins. *(A)* Protein motifs identified with MEME software (1). *(B)* Structural protein motifs identified with a Simple Modular Architecture Research Tool SMART (http://smart.embl-heidelberg.de/) (2). *Arabidopsis* mTERFs are ordered by phylogenetic relationships (clade) using protein sequences that are indicated by their AGI codes. The localization of each corresponding protein-GFP fusions is indicated: Chloroplast (P); mitochondria (M); nucleus and cytoplasm (C&N). Question marks (?) correspond to cases when no GFP fluorescence was detectable in our assays. Domains 1,2,4,10 and 12 overlap with the PFAM mTERF domain while motifs 3 and 8 are strongly associated with mitochondrial targeting. The X axis indicates the length of polypeptides as numbers of amino acid residues. For sequence analysis, protein sequences were retrieved by sequence similarity searches (HMM search with PF02536 and queries with InterPro IPR003690) from PLAZA (3), EnsemblGenomes (4), and Orthomcldb (5), and aligned with T_Coffee (6). After the alignment had been edited and the non-conserved positions been removed, a phylogenetic tree was computed with PhyML (7) by applying the JTT substitution model, 500 bootstrap samples, estimated proportion of invariable sites, four substitution categories, estimated gamma distribution parameter, the BIONJ distance-based tree as starting tree and without tree optimization (default parameters for protein sequences). For protein motifs identified with MEME (1), motif regular expressions are reported (Table S2).

References

1. Bailey TL, et al. (2009) MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 37: W202-W208.
2. Schultz J, et al. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. Proc Natl Acad Sci USA 95:5857-5864.
3. Proost S, et al. (2010) PLAZA: a comparative genomics resource to study gene and genome evolution in plants. Plant Cell 21: 3718-3731.
4. Kersey PJ, et al. (2009). Ensembl Genomes: extending Ensembl across the taxonomic space. Nucleic Acids Res 38: D563-D569.
5. Chen F, et al. (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res 34: D363-D368.
6. Notredame C, et al. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205-217.
7. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696-704.



Supplementary Figure S3. Maximum likelihood phylogenetic tree of *mTERF* genes. Green and red genes indicate plastidial and mitochondrial targeting, respectively, while clade VII marks a group of Arabidopsis mTERF genes including a tandem gene cluster on chromosome 1. Species abbreviations: AT *Arabidopsis thaliana*, Hsa *Homo sapiens*, Dme *Drosophila melanogaster*, Cel *Caenorhabditis elegans*, Pfa *Plasmodium falciparum*, Esi *Ectocarpus siliculosus*, OS *Oryza sativa*, PP *Physcomitrella patens* and CHLAMY CR *Chlamydomonas reinhardtii*. Phylogenetic inference supports the existence of at least ten mTERF subtypes conserved between monocots and dicots (indicated in blue). Numbers on the nodes refer to bootstrap support values (500 samples).



Supplementary Figure S4. Analysis of the fatty acid composition in albino mutant and wild-type plants. Total lipids were extracted from calli, lipid extracts were transmethylated and fatty acid methylesters were analyzed by gas chromatography as described (43) (16:0, palmitic acid; 16:2, *cis*-7,10-hexadecadienoic acid; *cis*-7,10,13-hexadecatrienoic acid; 18:0, stearic acid, 18:1, oleic acid, 18:2, linoleic cid, 18:3, α-linolenic acid).

T-DNA1, ACC2-2/GSP1 flanking sequence

5'CCAAACCCTTCCGTGGAAGTTTCCCACGATTAGGGCTTCCAACGGCAATCTCC
GGTAAAGTTCATCAGAGATGTGCTGCAACTTTAAATGCTGCACGTATGATTCTTG
CTGGCTATGATCATAAAGTAGATGAG(exon#19)gtaaacgctatttggttttcctatgtgattcaactt
ttctattaaataatttgattctgagattgctcatatgttttgcagGTTCTTCAAGACTTGCTTAATTGCCTTG
ATAGCCCTGAACTCCCATTTCTTCAGTGGCAAGAGTGCTTTGCAGTTCTGGCGA
CACGACTACCTAAAGATCTCAGGAACATG(exon#20)gtaaacacttagatactattcgtaatcc
gttttcctttactgtgatcttgttttgagtttctagtcttttaatgctttttgtttataattctttatcagCTAGAATTGAAATAC
AAGGAATTTGAGATTATCTCCAAAACCTCCCTGACCCCAGAT*atatattgtggtgtaaaca*
aattgacgcttagacaacttaataacacattgcggacgttttatggatactggggtggttttcttttcaccagtgagacggg
*caaca*3'

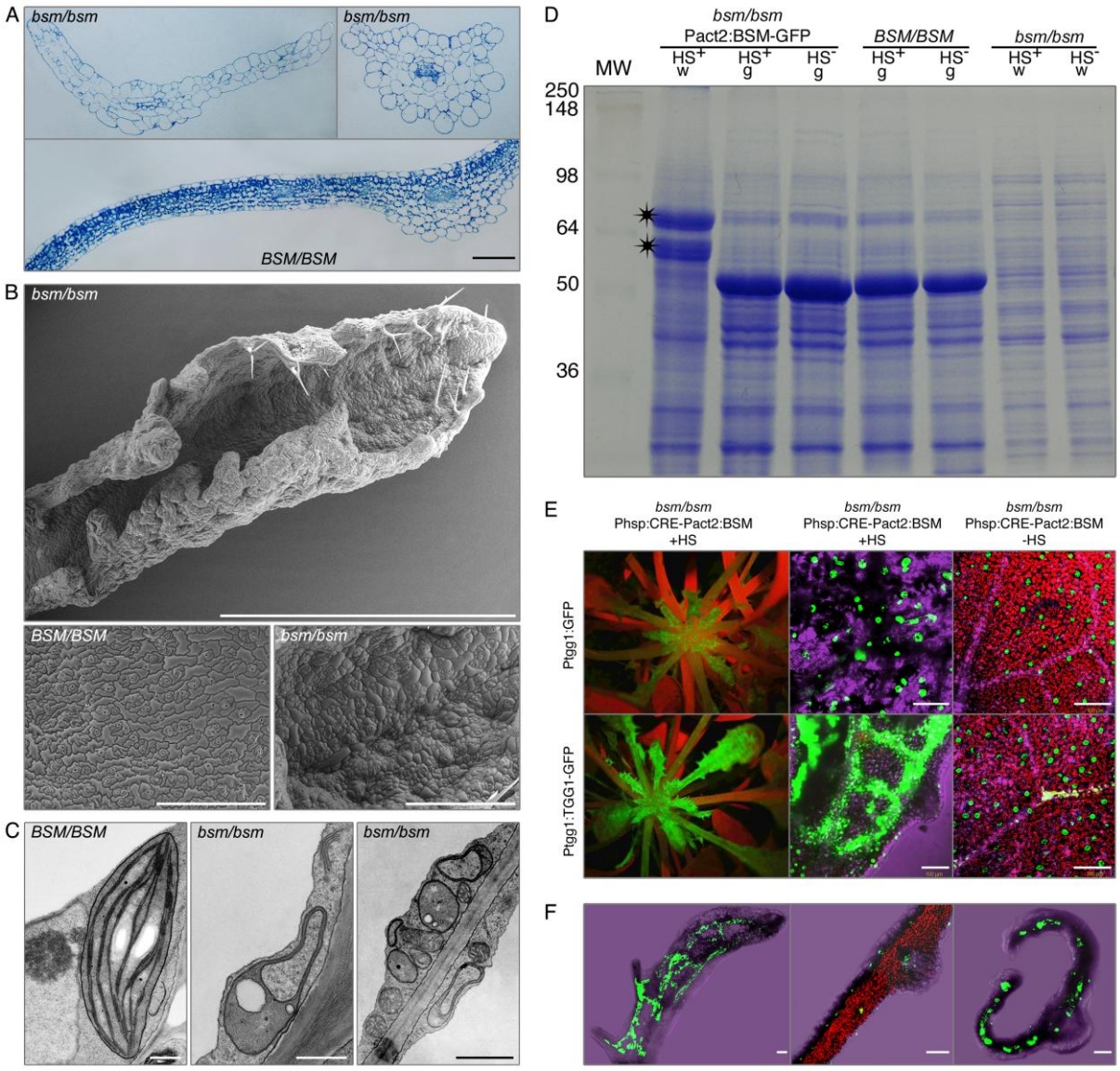
T-DNA2, ACC2-2/GSP1 flanking sequence

5'*tcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaaaaccaccccagtacatta*
*aaaacgtccgcaatgtgttattaagttgtctaagcgtcaatttgtttacaccacaatata***acacattgagacaacttga**A
GGCATCCTTGAG(exon#21)gtgattagctactgcttcttattctagacctactgttctgtattctactctaaaatac
tttcttttcggtccatttttgcaacagGCTCATCTATCTTCTTGTGATGAGAAAGAGAGAGGTTC
CCTTGAAAGGCTCATTGAACCGTTGATGAGCCTTGTGAAGTCTTATGAAGGTGG
TAGAGAAAGTCATGCCCGTCTTATTGTTCATTCCCTCTTTGAAGAATACCTATCA
GTTGAAGAATTATTCAATGATAACATGCTG(exon#22)gtaatatatggttcaattgttatcaaacg
ggtttgtgattaatagttatgtgtgaggagtcttaatatgcattttcttatcttgaaaa3'

T-DNA, ACC2-1 flanking sequence

5'GAAAACGGGGGACTATGGTTATAATCAAATCGCTGCAGTTTCTGCCGAGTATAATAAATGCATC
ATTGAGAGAAACAAACCATAGCCACTGTGAATATGCCCGAGCTCCTTTATCTGGTAATATGATGCA
CATTGCTGTTGTGGGCATCAACAATCAGATGAGTCTGCTTCAGGACAG(exon#26)gtacttgacagtcaac
acatcatgactgtttacaccagcttaggtcttgtagtgtctaacctaatctgttttttttttggtttggttgtctcagTGGGGACGA**cttagacaact**
aacacattgccctggcgtgacaaagagcaaacaaattaaggga*aaattgacgcttagacaacttaataacacattgcggacgttt*
*atatgtactggggtggttttc*3'

Supplementary Figure S5. T-DNA flanking sequences as found in *acc2-1* and *acc2-2* mutant alleles. T-DNA/*Arabidopsis*-DNA flanking sequences were amplified with *ACC2* gene-specific primers and a T-DNA primer. The *acc2-1* is known as SALK_148966C and *acc2-2* is SALK_110264. Two PCR fragment were amplified with DNA from *acc2-2* plants, suggesting insertion of two T-DNAs in inverted orientation. The sequences generated from amplified fragments are shown. The exon sequences of the *ACC2* gene (capitals) and intron sequences of the *ACC2* gene are presented as well as the sequences that form the T-DNA vector pROK2 (lowercase italics) and the novel sequences that are often generated during the process of T-DNA integration (bold).



Supplementary Figure S6. Cellular *bsm* phenotype.

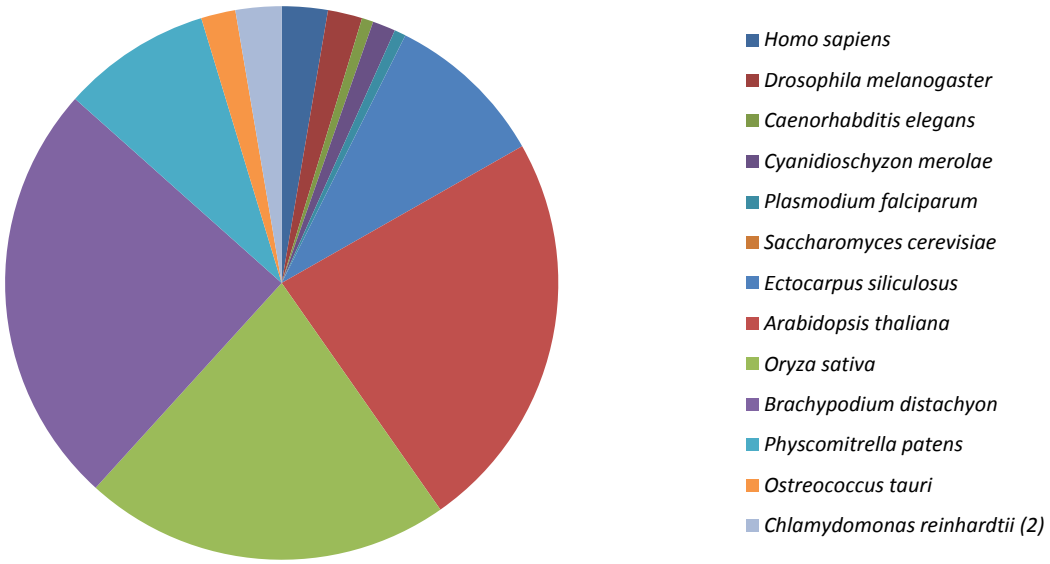
(A) Light microscopy of leaf sections. Albino leaves of mosaic plants (top row) and section of a green leaf from the same plant (bottom). Bar = 100 µm. *(B)* Scanning microscopy of leaf surfaces. Shown are an albino leaf of a mosaic plant; adaxial surface of wild type and albino leaves. *(C)* Ultrastructural analysis of plastid morphology. Chloroplast in a wild-type leaf cell and extreme mutant phenotypes of plastids in cells of pure albino, deformed leaves. *(D)* SDS-PAGE analysis of protein extracts from leaves of heat stress (HS^+)-induced mosaic plants (*bsm/bsm*; Pact2:BSM-GFP) grown on soil. Wild-type plants (*BSM/BSM*); *in vitro* calli/shoot cultures of homozygous mutant albinos (*bsm/bsm*) and not heat-stressed (HS^-) complemented plants were used as controls. Asterisks mark abundant polypeptides that accumulated in albino leaves of mosaic plants and that were identified by mass spectrometry as TGG1 and TGG2 myrosinases. The abundant polypeptide of 50 kDa in extracts of green leaves is the large

subunit of Rubisco. The pigmentation of leaves, green (g) or white (w) is indicated above the lanes of the gel. For the MS-MS identification of polypeptides**,** peptides were analyzed and the MALDI spectra were acquired with an Ultraflex TOFTOF instrument (Bruker Daltonics). LC-MS/MS analyses of desalted peptides were done on an Acquity ultraperformance LC system (Waters) connected to a Q-TOF micro mass spectrometer (Waters). Spectra were evaluated with the Mascot software (Matrix Science) and the UniProtKB/Swiss-Prot database (UniProt release 2010_07). Proteins were only accepted as identified when at least three unique peptides had an individual score above 22. *(E)* Patterns of GFP and chlorophyll distribution in induced mosaic (+HS) and control (-HS) plants. Myrosinase TGG1 (*At5g26000*) reporters (Ptgg1:GFP and Ptgg1:TGG1-GFP) highlight guard cells of stomata complexes in epidermis and myrosin cells of vascular bundles. Albino leaves do not show red chlorophyll fluorescence. Bar = 100 μm. *(F)* Hand sections of leaves from the mosaic plant that expresses the P_{tgg1}*:TGG1-GFP* reporter. Serrated albino mosaic leaf and cross-section of a similar leaf compared to a green leaf (middle). TGG1-GFP coincides with cells of the vascular bundles. Bar = 100 μm.

Supplemental Table S1. Numbers of *mTERF* genes in different organisms with sequenced genomes.

Species	Source	Query	#Genes [1]
Homo sapiens	Ensembl 58	IPR003690	4
Drosophila melanogaster	Ensembl 58	IPR003690	3
Caenorhabditis elegans	Ensembl 58	IPR003690	1
Cyanidioschyzon merolae	Orthomcldb	PF02536	2
Plasmodium falciparum	Ensembl Genomes	IPR003690	1
Saccharomyces cerevisiae	Ensembl Genomes	IPR003690	0
Ectocarpus siliculosus	BOGAS	IPR003690	14
Arabidopsis thaliana	PLAZA	IPR003690	35
Oryza sativa	PLAZA	IPR003690	32
Brachypodium distachyon	PLAZA	IPR003690	37
Physcomitrella patens	PLAZA	IPR003690	13
Ostreococcus tauri	PLAZA	IPR003690	3
Chlamydomonas reinhardtii [2]	PLAZA	IPR003690	4

[1] *P* value < 0.01 hmmsearch mTERF PFAM
[2] CR00087G00170 = CHLAMY_mTERF1; CR00008G01660 = CHLAMY_mTERF2.



Supplemental Table S2. MEME protein motifs in *Arabidopsis* mTERF genes

Motif	Protein sequence (regular expression)*
1	**PQ[VY][LF][GS][YF]SLEKR[IT][VK]PR**[CH]NV[IL]KALMSKGL
2	NPDSVL[SN]L**L**RS[HY]**GF[TS]**DSQIS[SR]I[IL][TKR]x**YP**[RQ][LV]
3	QPVCGKE[KNR]FE[EA]SLKKVVEMGFDPTTS[KT]FV
4	**GFS**R[DE][ED][VFI]Ax[MI][VI]K[RK]F**PQ**[CLI][LI]G[YL]**S**
5	LQFLQSRGASSSELTEIVSTVPKILG[KM][RK][GE][GH]K[ST][IL]S[RV]YYDF[VI]K
6	AAD[VL]S[LP]RD[GS]RKG[KN][NS]FTVSYLVDSLGL[PTA][KT]KLAESIS[KRM]KVSFE
7	AL[RC]V[VIL]Y[RG][LMF]S[DE]KT[IL]EE[KR][VF]N[VA][YC]K[RS][LF]G[FL][DAST]V[DEGN]DVW[AE][MIV]FKK[WC]P[SN][FS]L[NKT][YV]SE[KN]KI[LIT][QN]
8	[GS][KN][QL][EG]NKIRN[VI][SL]VLRELG[VM]P[QH][KR]LL[FL][PS]LLIS
9	ELP[PS][MI]SSVL[TV][CS]TD[EQ]VFL[KN]R[YF]VM[KN]HD[DE][KL]Q[LP]
10	[ED]x[LVI][KE][KP]KTE[FY]LVK[EK]M[GN][WR]P[LV][KE][AE][LV]
11	MY[SA]LILHG[RK][RK][LS]V[QE]LQK[WC]R[HN][LF][RS]
12	**G**[LIF]S[EK]E[ED][IV]x[RS][LMIV][LV]K[KR]C**P**[EQ][IC][LI][GT]S**S**
13	I[IT][EL][AL]DKSSK[YF]E[KT]LC[HQ]
14	[VL]Q[NK][AG]S[AP][FL]SNSFSS[AV][SA]
15	[ED][NA][FAY][HF][AV][LF][CSW][NGY][YF]G[IF][PG][RW][DEGN]K[ILM]G[KR][LIM][YF]KE[AE]R[EL][IV]F[VGR][YQ][ER][NPST]G[VEM][LI][AE][SM][KR][LI][EKLR][GKPS][YF][EK][DIN][LI][GV][FL][SR][KT]

* Bold and red residues correspond with highly conserved positions in the PFAM mTERF seed alignment

Supplementary Table S3. Mutant phenotypes of plastidial mTERF family genes

N	Lab ID[a]	AGI	Name	NASC ID[b]	Mutant phenotype[c]	TAP[d]	Reference
1	TERF1	At5g54180	pTAC15	N656192	-	Yes	1
2	TERF3	At4G02990	BSM	-	Embryo arrest	Yes	This study
3	TERF5	At2g21710	EMB2219	N16044	Embryo arrest	Yes	2
4	TERF6	At1g78930		N653232	Embryo arrest	Yes	This study
5	TERF25	At2g36000		-	-	No	This study
6	TERF26	At3g18870		N310733	-	No	This study
7	TERF27	At2g03050	SOLDAT10	N641368	Embryo arrest	Yes	3
8	TERF28	At2g34620		N597699	-	Yes	This study
9	TERF29	At4g1605		N819625	Gametophyte lethal	Yes	This study
10	TERF30	At4g09620		N654119		Yes	This study
11	TERF31	At5g55580		N656192	Could be yellowish	Yes	This study

[a] Temporary identification numbers used in our work.
[b] NASC, Nottingham Arabidopsis Stock Center collection number.
[c] Mutant phenotypes of the loss of function alleles as indicated herein were only confirmed by genetic complementation for the *BSM* gene.
[d] Expression of the TAP fusion in *bsm* complementation test.

References

1. Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K.-J., and Oelmüller, R. (2006). pTAC2,-6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. Plant Cell 18: 176-197.
2. Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T.C., McElver, J., Aux, G., Patton, D., and Meinke, D. (2004). Identification of genes required for embryo development in Arabidopsis. Plant Physiol. 135: 1206-1220.
3. Meskauskiene, R., Würsch, M., Laloi, C., Vidi, P.-A., Coll, N.S., Kessler, F., Baruah, A., Kim, C., and Apel, K. (2009). A mutation in the Arabidopsis mTERF-related plastid protein SOLDAT10 activates retrograde signaling and suppresses $^1O_2$-induced cell death. Plant J. 60: 399-410.

Supplementary Table S4. Sequences of primers used to generate *Arabidopsis* mTERF family orfeome; primers for probes; mutant alleles; gene expression cassetes and RACE PCR

Primer name	AGI of the analyzed gene and/or other	Name of the analyzed gene	Primer sequence from 5' to 3 '
TERF1DIR	At5g54180	PTAC15/TERF1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGTGATCTTATCTCTCGT
TERF1REV	At5g54180	PTAC15/TERF1	GGGGACCACTTTGTACAAGAAAGCTGGGTATAAACAAATCATTTCGATAT
TERF2DIR	At4g38160	PDE191/TERF2	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGAGGTGACAAATACGAG
TERF2REV	At4g38160	PDE191/TERF2	GGGGACCACTTTGTACAAGAAAGCTGGGTAGAAGCTTGCAGTTACCTCCG
TERF3DIR	At4g02990	BSM/TERF3	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGAAGATTAGGTTCTGTAATGGCTT
TERF3REV	At4g02990	BSM/TERF3	GGGGACCACTTTGTACAAGAAAGCTGGGTTATGCAAACTCCTCGTCGTCATCA
TERF4DIR	At2g44020	TERF4	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGAGCTATCTTCTTAGGAGGAA
TERF4REV	At2g44020	TERF4	GGGGACCACTTTGTACAAGAAAGCTGGGTACAAAGTAAGAGTGCGTCTGTA
TERF5DIR	At2g21710	EMB2219/TERF5	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGCTTCTCCACTGCAACGT
TERF5REV	At2g21710	EMB2219/TERF5	GGGGACCACTTTGTACAAGAAAGCTGGGTATTCTGTCAAATCCTCTTCT
TERF6DIR	At1g78930	TERF6	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGAGCTTTTACATACGTC
TERF6REV	At1g78930	TERF6	GGGGACCACTTTGTACAAGAAAGCTGGGTACGAGATTTCACTGTGAGTT
TERF7DIR	At4g19650	TERF7	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGAAGAAGGAGAAGGACGTT
TERF7REV	At4g19650	TERF7	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGTAGATGATTGGTTTATAC
TERF8bDIR	At5g06810	TERF8b	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTCTTTTATGCGAAACCCTAGATTCA
TERF8BREV	At5g06810	TERF8b	GGGGACCACTTTGTACAAGAAAGCTGGGTAACACACAAGTAGTTTCTTC
TERF9DIR	At3g46950	TERF9	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTTTTCTCTAATTCTCCAT
TERF9REV	At3g46950	TERF9	GGGGACCACTTTGTACAAGAAAGCTGGGTATGAAACACGACCTCTGTTG
TERF10DIR	At1g62010	TERF10	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGAATTCTCTTATACTCGGT
TERF10REV	At1g62010	TERF10	GGGGACCACTTTGTACAAGAAAGCTGGGTATTGTTCTAGGCGTGCCTTC
TERF11DIR	At1g62110	TERF11	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATTCTGTGATACTCCAT
TERF11REV	At1g62110	TERF11	GGGGACCACTTTGTACAAGAAAGCTGGGTATGATGATACACGACCTCTG
TERF13DIR	At1g61960	TERF13	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATGCTCTGATACACCAT
TERF13REV	At1g61960	TERF13	GGGGACCACTTTGTACAAGAAAGCTGGGTACTTCCCATTTTCTCCAGTG
TERF14DIR	At1g61980	TERF14	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATTCTCTGATTCGCCAT
TERF14REV	At1g61980	TERF14	GGGGACCACTTTGTACAAGAAAGCTGGGTATGAAGCACGATAGATAGCC
TERF15DIR	At1g61970	TERF15	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATGCTCTGATACTCCAT
TERF15REV	At1g61970	TERF15	GGGGACCACTTTGTACAAGAAAGCTGGGTATGAAGCACGATAGATAGCC
TERF16DIR	At1g62150	TERF16	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTGTTCTTCTCTAGTTCCT
TERF16REV	At1g62150	TERF16	GGGGACCACTTTGTACAAGAAAGCTGGGTATGAGACACGGTCTTGGTTG
TERF17DIR	At1g61990	TERF17	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATTCTCTGATTCTCCAT
TERF17REV	At1g61990	TERF17	GGGGACCACTTTGTACAAGAAAGCTGGGTATGAAACACTACCTTTGGTG
TERF18DIR	At5g23930	TERF18	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTTTTCATTAATACTCCAT
TERF18REV	At5g23930	TERF18	GGGGACCACTTTGTACAAGAAAGCTGGGTACAATGCTATCTTATTTATT
TERF19DIR	At1g62490	TERF19	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTCGGAATCTAAAACCCAGA
TERF19REV	At1g62490	TERF19	GGGGACCACTTTGTACAAGAAAGCTGGGTAATTAAGCTCAATATCCTCT
TERF20DIR	At1g56380	TERF20	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCTGCTGATGATGTGATC
TERF20REV	At1g56380	TERF20	GGGGACCACTTTGTACAAGAAAGCTGGGTACTGATCTTTGGTGAAGATAG
TERF21DIR	At1g21160	TERF21	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTCGAGCTTATTACACTGT
TERF21REV	At1g21160	TERF21	GGGGACCACTTTGTACAAGAAAGCTGGGTATCGATTACTTGTAAAGGGTT
TERF22DIR	At5g07900	TERF22	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCGGTCGCAATACCTCAAG
TERF22REV	At5g07900	TERF22	GGGGACCACTTTGTACAAGAAAGCTGGGTAGAGTTTGGTGCATCCTAGA
TERF23DIR	At5g64950	TERF23	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGCAATTTCTAGCTTCTCT
TERF23REV	At5g64950	TERF23	GGGGACCACTTTGTACAAGAAAGCTGGGTAGATCTTGATACGATCAAGT
TERF24DIR	At1g74120	TERF24	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCCTCAAAACTCAAAACCT
TERF24REV	At1g74120	TERF24	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGCAAGTGACTCTATAAAG
TERF25DIR	At2g36000	TERF25	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGCAACAAGAAGCTCTCTC
TERF25REV	At2g36000	TERF25	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGAATCTTTTTGTACAGTA
TERF26DIR	At3g18870	TERF26	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCTGTTATTGCTTCACTT
TERF26REV	At3g18870	TERF26	GGGGACCACTTTGTACAAGAAAGCTGGGTATGGCTTCCATTTTGAGTAGA
TERF27DIR	At2g03050	SOLDAT10/TERF27	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGATAGCAAGGTGTTCTCT

TERF27REV	At2g03050	SOLDAT10/TERF27	GGGGACCACTTTGTACAAGAAAGCTGGGTATCTTCTTTCAGCAGACCTA
TERF28DIR	At2g34620	TERF28	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTCTGCAGCATTGAGCTC
TERF28REV	At2g34620	TERF28	GGGGACCACTTTGTACAAGAAAGCTGGGTACACGTTTGCCACTGAAGAGGG
TERF29DIR	At4g14605	TERF29	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGCAAAGTCTTAGTCAACT
TERF29REV	At4g14605	TERF29	GGGGACCACTTTGTACAAGAAAGCTGGGTATATCTCGGCTTGATCCTCT
TERF30DIR	At4g09620	TERF30	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGAGATGGTGGGAAACTC
TERF30REV	At4g09620	TERF30	GGGGACCACTTTGTACAAGAAAGCTGGGTAAAACCGAACCCAACACCGAGT
TERF31DIR	At5g55580	TERF31	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCGGGTTTCTCACTGTAC
TERF31REV	At5g55580	TERF31	GGGGACCACTTTGTACAAGAAAGCTGGGTATCCTCTCTTGTCATACTTGT
TERF32DIR	At3g60400	TERF32	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTTCATGGTTCGATTGAA
TERF32REV	At3g60400	TERF32	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGATACAGTGTCTCTGGTTT
TERF33DIR	At1g79220	TERF33	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGCATTCGACCGGGAAGAG
TERF33REV	At1g79220	TERF33	GGGGACCACTTTGTACAAGAAAGCTGGGTAAAAGGGGAATCCTTTACGG
TERF34DIR	At1g62120	TERF34	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTATTCTCTGATTCTCCAT
TERF34REV	At1g62120	TERF34	GGGGACCACTTTGTACAAGAAAGCTGGGTATTGTTCTAGGCGTGTCTTCT
TERF36DIR	At5g45113	TERF35	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGTTTGAGAACTTTCATGTT
TERF36REV	At5g45113	TERF35	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGTAGATGCTTTCTTTATAT
ACC2GSP1	At1g36180	ACC2	TGCGTGGTCTGGTTTCATGGCAGCA
ACC2GSP2	At1g36180	ACC2	ACACTGGCCCAGACGATCACGAGA
ACC2ORFDIR1	At1g36180	ACC2	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGAGATGAGAGCTTTGGGTTCTTC
ACC2ORFREV1	At1g36180	ACC2	GGGGACCACTTTGTACAAGAAAGCTGGGTAATTAGTTCCTCCCGGAACCTCAACAAA
ACC2ORFREV3	At1g36180	ACC2	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTCATCAGGATCCTCGCTTGTCACAC
ACC2ORFREV4	At1g36180	ACC2	GGGGACCACTTTGTACAAGAAAGCTGGGTAGCAATAAATGGTTCTTGTTTTCGTGATTG
ACC2-2GSP1	At1g36180	ACC2	CCAAACCCTTCCGTGGAAGTTTCCCA
ACC2-2GSP2	At1g36180	ACC2	CAAGAGAAACCTGGTGTGCGAAGGA
SALK_lLBb1MOD	T-DNA	pROK2	GGCAAACCAGCGTGGACCGCTTGCTG
ESPORFDIR1	At1g54040	ESP1/ESR1/TASTY	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCTCCGACTTTGCAAGGCCAGT
ESPGECSEQ1	At1g54040	ESP1/ESR1/TASTY	ACGGGTGGGTACTCACCTAA
ESPTERDIR1	At1g54040	ESP1/ESR1/TASTY	TTGCAGCCCCTAGGTGCGTGTCAAGATTTGTGTTGTGTGTGTGG
ESPTERREV1	At1g54040	ESP1/ESR1/TASTY	TACTGGACCATTTCCATGGGTCCAA
ESPPROREV1	At1g54040	ESP1/ESR1/TASTY	CGCACCTAGGGGCTGCAATATAAGTATAAAAGTGTTTT
ESPORFREV1	At1g54040	ESP1/ESR1/TASTY	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGCTGAATTGACCGCATAGAAGTAGAGA
ESPPRODIR1	At1g54040	ESP1/ESR1/TASTY	TACTCCTAAGGCACTGGGCTCTACAAAA
TGG1ORFDIR1	At5g26000	TGG1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGAAGCTTCTTATGCTCGCCTTT
TGG1ORFREV1	At5g26000	TGG1	GGGGACCACTTTGTACAAGAAAGCTGGGTACTTGATGACTTTACTGAGGAAACAG
TGG1PRODIR1	At5g26000	TGG1	CATGCGACAATGCGACATGCGGA
TGG1TERDIR1	At5g26000	TGG1	CACTACTAATAAACCTAGGAATATCCAATCCACTATATGTCCAC
TGG1TERREV1	At5g26000	TGG1	AAAAAACACATGTCACATGTGGTCTCA
TGG1PROREV1	At5g26000	TGG1	GATATTCCTAGGTTTATTAGTAGTGTGTATGTTTTTG
5RACE_CRterf1_1	Chlamydomonas mTERF 154170	CR_TERF1	CTCCGACAGCAGCGCCCGCCA
5RACE_CRterf1_2	Chlamydomonas mTERF 154170	CR_TERF1	AGCAGCGCCCGCCACTTGAA
3RACE_CRterf1_1	Chlamydomonas mTERF 154170	CR_TERF1	CAAGTGGCGGGCGCTGCTGT
3RACE_CRterf1_2	Chlamydomonas mTERF 154170	CR_TERF1	GCGGGCGCTGCTGTCGGAG
3RACE_CRterf1_3	Chlamydomonas mTERF 154170	CR_TERF1	GGCCAGGTGTTCCTGTTCCTGAA
3RACE_CRterf1_4	Chlamydomonas mTERF 154170	CR_TERF1	GCTGGGTGTGCCCAACGACCAAAT
CR_5RACE_terf2_1	Chlamydomonas mTERF 155303	CR_TERF2	CCTCAGGCTTGAAGCCCAGCA
CR_5RACE_terf2_2	Chlamydomonas mTERF 155303	CR_TERF2	GGCTTGAAGCCCAGCAGGTGT
CR_3RACE_terf2_1	Chlamydomonas mTERF 155303	CR_TERF2	GTGCCCACACCTGCTGGGCTT
CR_3RACE_terf2_2	Chlamydomonas mTERF 155303	CR_TERF2	CACACCTGCTGGGCTTCAAGCCT
5P_RACE	5'/3'RACE primers		CGACTGGAGCACGAGGACACTGA
5P_RACE_NEXT	5'/3'RACE primers		GGACACTGACATGGACTGAAGGAGTA
3P_RACE	5'/3'RACE primers		GCTGTCAACGATACGCTACGTAACG
3P_RACE_NEXT	5'/3'RACE primers		CGCTACGTAACGGCATGACAGTG
CR_TERF1_REV1	Chlamydomonas mTERF 154170	CR_TERF1	GGGGACCACTTTGTACAAGAAAGCTGGGTAGATGGCGGGGTAGCGCTCCAGCA
CR_TERF1_REV2	Chlamydomonas mTERF 154170	CR_TERF1	GGGGACCACTTTGTACAAGAAAGCTGGGTACTCACCTCCTCCCCACCCGCCGCT
CR_TERF1_DIR1	Chlamydomonas mTERF 154170	CR_TERF1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAGTCACCTGTATGAAACAATGAG
CR_TERF2_REV1	Chlamydomonas mTERF 155303	CR_TERF2	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTACACCCCCAGGAAGTCCAGCCA

CR_TERF2_REV2	Chlamydomonas mTERF 155303	CR_TERF2	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGTACCCTCGTAAATGCCCAGCTT
CR_TERF2_DIR2	Chlamydomonas mTERF 155303	CR_TERF2	GGGGACAAGTTTGTACAAAAAAGCAGGCTGACACCATAAGAACACATATATTGT
CR_TERF2_DIR0	Chlamydomonas mTERF 155303	CR_TERF2	GGGGACAAGTTTGTACAAAAAAGCAGGCTAACAAGACGACACCATAAGAACACA
CLPP_GATE5P	AtCg00670	Arabidopsis clpP	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGCCTATTGGCGTTCCAAAAGT
CLPP_GATE3P	AtCg00670	Arabidopsis clpP	GGGGACCACTTTGTACAAGAAAGCTGGGTTTATTGAACCGCTACAAGATCAACAAT
CLPP_RT	AtCg00670	Arabidopsis clpP	TTATTGAACCGCTACAAGATCAA
23SRNA_5P	AtCg01180	RRN23S, 23S ribosomal RNA	CGAGGAAAGGCTTACGGTGGATA
23SRNA_3P	AtCg01180	RRN23S, 23S ribosomal RNA	TCAGTCGGTTCGGACCTCCACTT
5SRNA_5P	AtCg00970	RRN5S, plastidial 5S ribosomal RNA	ATTCTGGTGTCCTAGGCGTAGAGGAACAAC
5SRNA_3P	AtCg00970	RRN5S, plastidial 5S ribosomal RNA	TCCTGGCGTCGAGCTATTTTTCC
ACCD_GATE_5P	AtCg00500	Arabidopsis accD	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGAAAAATCGTGGTTCAATTTTATGTTT
ACCD_GATE_3P	AtCg00500	Arabidopsis accD	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATTTGTGTTCAAAGGAAAAAAAGCATGG
ACCD_RT	AtCg00500	Arabidopsis accD	TTTAATTTGTGTTCAAAGGAAAAAAAGCA
ACCD_SEQ1	AtCg00500	Arabidopsis accD	GAACTAATAGGAATCGTAGTAA
ACCD_SEQ2	AtCg00500	Arabidopsis accD	GGAATACCGTTTAGTTGACCT
rrn16S5	AtCg00920	RRN16S, 16S ribosomal RNA	CTGGCTCAGGATGAACGCTG
rrn16S3	AtCg00920	RRN16S, 16S ribosomal RNA	CTGGGATTTGACGGCGGACT
atpA5	AtCg00120	Arabidopsis atpA	ATGGTAACCATTAGAGCCGACGA
atpA3	AtCg00120	Arabidopsis atpA	ACTACCTGAGCCACGGAAGAAG
rbcL5	AtCg00490	Arabidopsis rbcL	ATGTCACCACAAACAGAGACTAAAG
rbcL3	AtCg00490	Arabidopsis rbcL	CATTCATAAACTGCTCTACCATAGT
RBCS5P	At5g38430	Ribulose bisphosphate carboxylase small chain protein, isoform 1b	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCTTCCTCTATGCTCTC
RBCS3P	At5g38430	Ribulose bisphosphate carboxylase small chain protein, isoform 1b	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGCATCAGTGAAGCTTGGGGG
FDX5P	At1g60950	Ferrodoxin precursor	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCAAATGGCTTCCACTGCTCTCTCAA
FDX3P	At1g60950	Ferrodoxin precursor	GGGGACCACTTTGTACAAGAAAGCTGGGTAAACAATGTCTTCTTCTTTGTGGGT