Trends in Plant Science

CellPress
REVIEWS

Opinion

The Hidden Face of Rubisco

Mathieu Pottier,[1] Dimitri Gilis,[2] and Marc Boutry[1,*]

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) fixes atmospheric CO_2 into organic compounds and is composed of eight copies each of a large subunit (RbcL) and a small subunit (RbcS). Recent reports have revealed unusual RbcS, which are expressed in particular tissues and confer higher catalytic rate, lesser affinity for CO_2, and a more acidic profile of the activity versus pH. The resulting Rubisco was proposed to be adapted to a high CO_2 environment and recycle CO_2 generated by the metabolism. These RbcS belong to a cluster named T (for trichome), phylogenetically distant from cluster M, which gathers well-characterized RbcS expressed in mesophyll or bundle-sheath tissues. Cluster T is largely represented in different plant phyla, including pteridophytes and bryophytes, indicating an ancient origin.

Highlights

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is a highly abundant chloroplast enzyme that fixes atmospheric CO_2 into organic compounds. It is composed of eight copies of a large subunit (RbcL) and eight copies of a small subunit (RbcS).

Genes that encode a particular RbcS (named RbcS-T) have been recently discovered in several species from different plant phyla. They are phylogenetically distant from well-characterized genes encoding RbcS (named RbcS-M), which are typically expressed in mesophyll and bundle sheaths.

Biochemical characterization of RbcS-T-containing Rubisco has revealed higher catalytic activity, higher K_m for CO_2, and an acidic shift of the activity versus pH profile.

A New Type of RbcS

CO_2 fixation into organic compounds is essential for life on Earth, and ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is a major player in this reaction. This enzyme, which, in eukaryotes, is localized to the chloroplast, catalyzes the conversion of CO_2 and ribulose-1,5-bisphosphate, a C5 compound, into two molecules of 3-phosphoglycerate, a C3 compound. Its abundance in plant tissues has made this enzyme a paradigm for structural, enzymatic, and molecular genetic characterization, as well as for protein evolution (for reviews, see [1–5]). Rubisco, which is by far the most abundant enzyme in photosynthetic tissues, has slow kinetics (k_{cat} usually lower than 5 s^{-1}) and is responsible for a side oxygenase reaction in which ribulose-1,5-bisphosphate is converted into phosphoglycolate, resulting in lower CO_2 fixation efficiency.

In plants, Rubisco is made of eight copies each of a large subunit (RbcL) encoded by the chloroplast genome and of a small subunit (RbcS) encoded by the nuclear genome. Two RbcL form a top to tail dimer, which is repeated four times (Figure 1). There are two active sites per dimer, located at the interface between two RbcL. The RbcS are located above and below the four RbcL dimers, and are thought to play a role in assembling, and maintaining together, the large subunits. However, RbcS are also indirectly involved in the catalytic reaction (see [6] for review).

While RbcL is usually encoded by a single chloroplast gene, RbcS is encoded by a nuclear family of several genes, which share high amino acid identity (usually between 90% and 100%) and are considered functionally equivalent in a given species. However, RbcS variants (amino acid identity < 65%) were recently identified, which clearly differ from the RbcS previously characterized in the same species [7–10]. This is the case in rice, where OsRbcS1 is phylogenetically distant from other rice RbcS and whose expression is restricted to nonphotosynthetic tissues such as leaf sheath, culm, anther, and root central cylinder. *OsRBCS1* overexpression in the leaf blade resulted in a Rubisco enzyme with a higher catalytic rate and higher K_m for CO_2 [8]. *OsRBCS1* homologs were identified in *Setaria italica* (foxtail millet), *Solanum lycopersicum* (tomato), *Lotus japonicus* (a wild legume), *Vitis vinifera* (grape), as well as in *Selaginella*

[1]Institut des Sciences de la Vie, University of Louvain, 1348 Louvain-la-Neuve, Belgium
[2]Bioinformatique génomique et structurale, Université Libre de Bruxelles, 1050 Brussels, Belgium

*Correspondence:
marc.boutry@uclouvain.be (M. Boutry).

 https://doi.org/10.1016/j.tplants.2018.02.006

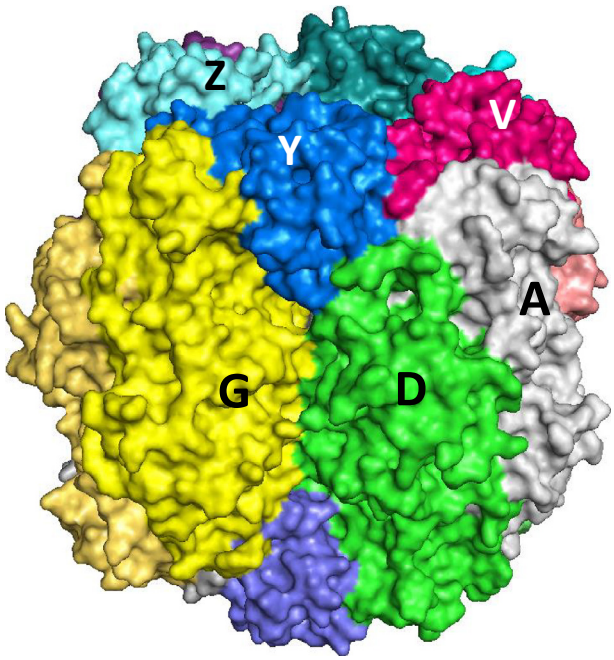


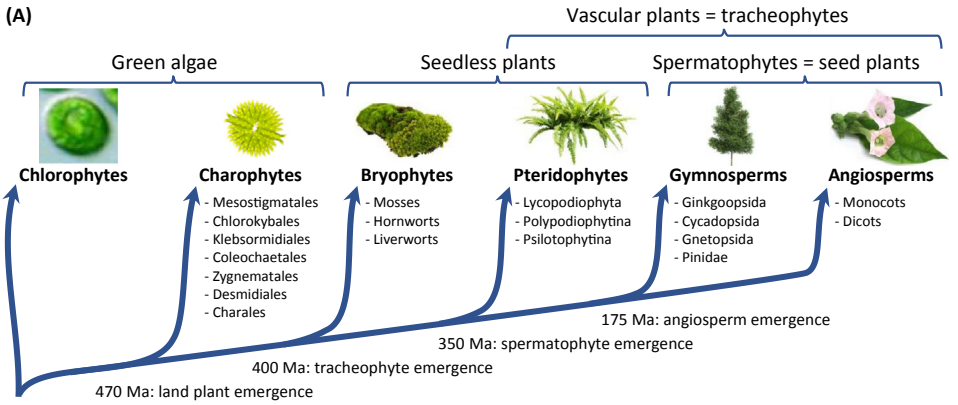
Figure 1. Model of Rbc-T from *Nicotiana tabacum*. NtRbc-T was modeled (Modeller 9.17, (https://salilab.org/modeller, [48])) using the *N. tabacum* 4rub structure [22] replacing the sequence of NtRbcS-M by that of NtRbc-T (GenBank accession id: DV157962). The different subunits are represented with different colors. Labeled subunits (Z, Y, V, G, D, and A) belong to the subcomplex shown in the Supplemental Information online, Figure S2A.

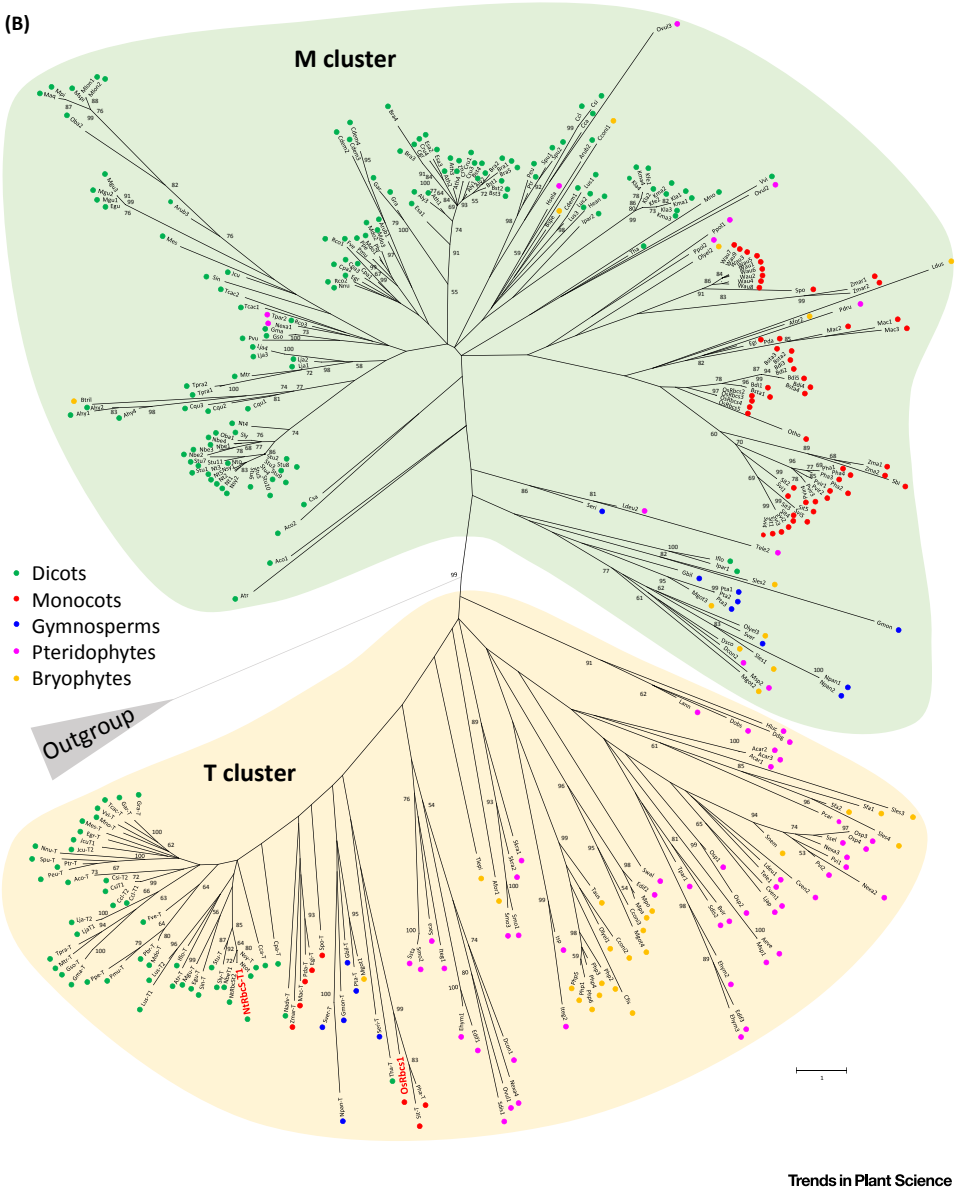
moellendorffii, a pteridophyte species. All were found to be mainly expressed in nonphotosynthetic tissues [9]. Another example was found in *Nicotiana tabacum* (cultivated tobacco), in which a particular RbcS named NtRbcS-T (T for trichome) is expressed exclusively in the secretory cells of glandular trichomes. This Rubisco was found to have higher K_m, higher V_{max}, and a more acidic profile of the pH-dependent activity than Rubisco from leaf mesophyll tissues [7]. Since RbcL is encoded by a single gene and is thus common for all the Rubisco enzymes within a given species, the kinetic differences found between trichomes and mesophyll tissues are uniquely due to the RbcS isoform.

Genomic data mining resulted in the identification of T-type RbcS in 34 flowering plant species [7,9]. It thus appears that, although identified in both monocot and dicot species, the T-type RbcS has remained so long unidentified because its expression is limited to particular tissues and is thus hidden by the large expression of the M type in mesophyll and bundle-sheath tissues. Here, we will address the following questions. (i) How widely spread is the T-type RbcS in land plants? Can it be found in phyla such as pteridophytes (e.g., ferns) or bryophytes (e.g., mosses; Figure 2A), which appeared before the spermatophytes (seed plants)? (ii) What are the structural features of RbcS-T that might explain the altered kinetic properties of Rubisco that contains this subunit? (iii) In relationship with the modified enzymatic properties of the T-type Rubisco, what are the physiological roles of this variant enzyme?

RbcS-T and RbcS-M Are Phylogenetically Distant

To date, T-type RbcS and therefore T-type Rubisco complexes have been little characterized at the expressional and functional levels. However, next-generation sequencing has provided a





(See figure legend on the bottom of the next page.)

wealth of genomic and transcriptomic data that have revealed the presence of T-type *RbcS* genes in many species. Thus, 38 genes from 34 flowering plant species including both monocot and dicot plants had been previously identified [7,9]. To determine the distribution of the T-type RbcS within the whole plant kingdom, a more exhaustive search was performed, taking advantage of the soaring genomic database. Plant *RbcS* sequences were retrieved from several genomic and cDNA databases [11–15] (see Supplemental Information online, Table S1). Phylogenetic analysis was first conducted on RbcS sequences from land plants including spermatophytes (also named seed plants), as well as pteridophytes and bryophytes, the latter two constituting the seedless plants (Figure 2A,B).

In spermatophytes, T-type *RbcS* were found in all subclasses of gymnosperms as well as of angiosperms (monocot and dicot plants). About 24% of the spermatophyte *RbcS* belong to the T type (Figure 2B). *RbcS-M* genes were consistently present in multicopies, while *RbcS-T* genes were mostly found as a single copy or as two copies in a few cases. Not all species contain T-type *RbcS*. Indeed, no *RbcS-T* sequences were identified among the 24 *RbcS* sequences retrieved from the seven sequenced Brassicaceae genomes. Moreover, only three of ten Poaceae species with their genome sequenced contain T-type *RbcS*, one being rice, a plant with C3-type photosynthesis and the others being *Panicum hallii* and *S. italica*, two plants with C4-type photosynthesis [9,16,17]. Overall, 55 of 87 spermatophyte species surveyed possess a T-type *RbcS*, while all of them have M-type *RbcS* (except for the nonsequenced genome species *Nuphar advena*, for which only a T-type *RbcS* was found). Although we need to be cautious (draft genomes are missing genes), these data might suggest that T-type *RbcS* genes were present at the onset of spermatophytes but disappeared several times independently during spermatophyte evolution.

Phylogenetic analysis also included *RbcS* sequences of more ancient tracheophyte species commonly referred to as pteridophytes as well as of nonvascular plants called bryophytes (Figure 2A,B). Although databases contain very few genome sequences of these seedless species, they contain many expressed sequence tag (EST) sequences, thus corresponding to expressed genes. T-type *RbcS* were identified in both phyla (see Supplemental Information online, Table S1). Interestingly, we found more T-type *RbcS* sequences (71) than M-type *RbcS* sequences (23), a result opposite to what was observed in spermatophytes. One hypothesis might be that *RbcS-T* expression in these species is not restricted to specific tissues, as it seems to be the case for spermatophyte species. Another hypothesis might be that pteridophyte and bryophyte species contain, on the whole, more T-type than M-type *RbcS* genes,

Figure 2. Distribution of the T-type RbcS into the Plantae Kingdom. (A) Overview of the phylogenetic relationships among the Plantae kingdom from green algae to land plants. Timescale for plant evolution according to Clarke *et al.* [49]. (B) Phylogenetic analysis of land plant RbcS sequences. All RbcS sequences from land plant species including spermatophytes, pteridophytes, and bryophytes are split into two clusters (M and T) defining M-type RbcS and T-type RbcS. The M cluster encompasses all well-characterized RbcS from spermatophytes, while the T cluster includes newly identified NtRbcS-T and OsRbcS1 (highlighted in red) from *Nicotiana tabacum* and *Oryza sativa*, respectively [7,8]. Although most spermatophyte RbcS are found in the M cluster, pteridophyte and bryophyte RbcS are mainly found in the T cluster. The abbreviations and the list of the corresponding RbcS sequence accessions can be found in the Supplemental Information online, Table S1. The evolutionary history was inferred using the maximum likelihood method based on the Le–Gascuel (LG) model [50]. Bootstrap values (400 iterations) greater than 50% are indicated [51]. The evolutionary distances were computed using the LG model. Initial trees for the heuristic search were obtained by applying the neighbor-joining method to a matrix of pairwise distances estimated using a Jones–Thornton–Taylor model. A discrete gamma distribution was used to model the evolutionary rate differences among sites (shape parameter = 10). The rate variation model allowed for some sites to be evolutionarily invariable. Proteobacteria RbcS were used as an outgroup. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site, except for the outgroup for which branches were collapsed. Similar evolutionary history was obtained using the neighbor-joining method [52] (data not shown). Analyses were conducted in MEGA7 [53].

while the opposite was clearly observed in the sequenced spermatophyte species. Besides, most of these seedless species (45 of 50) displayed T-type *RbcS* sequences. However, only a few of them (14 of 50) had both T-type and M-type RbcS. This tendency is mainly based on EST sequences and needs to be confirmed when whole genome sequences become available for many pteridophytes and bryophytes. Nevertheless, we can already conclude that *RbcS-T* are widespread in land plants, including ancient phyla. This raises the question as to whether a T-type RbcS might predate the terrestrial conquest.

To address this possibility, we analyzed *RbcS* sequences from chlorophytes, which contain most of the extant green algae (e.g., *Chlamydomonas*), and from charophytes (e.g., *Klebsormidium nitens*; [18]), other green algae, which gave rise to all land species (Figure 2A) [19]. *RbcS* sequences from chlorophytes were clearly confined into an independent cluster (see Supplemental Information online, Figure S1). On the contrary, RbcS sequences from charophyte species were closer to the M and T clusters, which is in agreement with the fact that land plants derived from charophytes [19]. However, it would be too speculative to more precisely associate charophyte *RbcS* with either the M or the T cluster (see Supplemental Information online, Figure S1).

Structural Analysis of the T-Type Rubisco from *N. tabacum*

While the 3D structure of several M-type Rubisco has been experimentally determined (reviewed in [20]), there is no structure available for the T-type, due to its recent discovery and its expression restricted to tissues not amenable for large-scale purification. Since RbcL is encoded by a single gene, and is thus common to both the T- and M-type Rubisco, RbcS is the only component that differs between both enzymes. Alignment of the 55 known T-type sequences of spermatophytes (seed plants) with M-type sequences from the same species pointed to 12 residues that are highly conserved (>75%) in each type but different between them (Figure 3). In addition, there is one additional residue (Position 46) in the M-type. We might speculate that some of these positions are responsible for the modified kinetics, namely, higher K_m and V_{max}, as well as the acidic shift of the activity versus pH profile, observed for the T-type compared with the M-type Rubisco in tobacco [7].

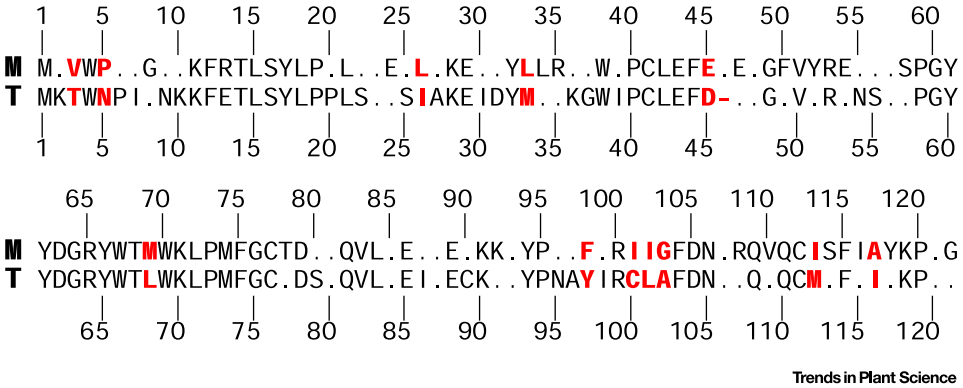


Figure 3. Alignment of the M- and T-Type RbcS Consensus Sequences. RbcS amino acid residues conserved in at least 75% of the 55 T-type sequences of spermatophytes listed in the Supplemental Information online, Table S1, are indicated. For the M-type, one RbcS sequence was retrieved from the same species as those in which a T-type RbcS had been identified. Conserved residues that differ between M and T are in red. A dot indicates that the position is not conserved. The red dash indicates a missing residue between Positions 45 and 46 of T.

Although there is no direct interaction between RbcS and the Rubisco active site, the small subunits contribute to the kinetic properties of the enzyme. For instance, point mutations in the small subunit at a remote distance from the catalytic site were shown to modify the Rubisco kinetic parameters of the alga *Chlamydomonas* and the cyanobacterium *Synechococcus* (reviewed in [2,3,6]). Replacing the *Chlamydomonas* RbcS by that of spinach, sunflower, or *Arabidopsis* resulted in a functional hybrid enzyme with modified kinetics [21]. Yet, despite these functional differences, and although RbcS sequences are clearly less conserved than RbcL sequences, their structures do show a common core structure [20]. It therefore makes sense to model the T-type Rubisco starting from the *N. tabacum* Rubisco 3D structure obtained for the M type (PDB code 4rub; [22], Figure 1), so as to determine the spatial position of the T-type-specific residues and, in particular, those that interact with the large subunits (see Supplemental Information online, Figure S2). Indeed, an RbcS subunit contacts three RbcL besides two other RbcS. Among the 12 type-specific positions, only three (for both M- and T-types) are predicted to contact an RbcL subunit: $_{M}$Pro5/$_{T}$Asg5, $_{M}$Glu45/$_{T}$Asp45, and $_{M}$Met69/$_{T}$Leu68. The most variable region between RbcS from various organisms lies in a loop extending from Residue 46 to 67 (M-type) or 66 (T-type; Figure S2). This concerns both the sequence and the length. For instance, the *Chlamydomonas* loop is 6, 7, or 18 residues longer than that of land plants (RbcS-M, RbcS-T) or the cyanobacterium *Synechococcus*, respectively. When the *Chlamydomonas* loop was replaced with that of spinach or *Synechococcus*, the kinetic parameters were modified [23]. However, the crystal structure of the chimeric enzymes did not show any alteration of the RbcS Cα backbone, with the exception of the loop region, which was shortened. Structural alterations of the large chain were limited to the side chains of some residues at the interface between the small and large subunits [23]. At the beginning of this loop, we can highlight the type-specific Residue 45 ($_{M}$Glu and $_{T}$Asp), which is predicted to interact with an RbcL (Figure S2). Interestingly, next to this residue lies the M-type additional residue (Position 46).

In conclusion, RbcS sequence alignment allowed for the identification of several residues specifically conserved in either the M- or T-type of RbcS. It can be expected that some of these, especially those localized at the interface between the small and large subunits, might be responsible for the different kinetic parameters observed between the M- and T-type Rubisco. The functional replacement of the *Chlamydomonas* RbcS by the tobacco RbcS-M or RbcS-T [7] represents an interesting tool to identify, through site-directed mutagenesis, those positions that play a major role in the distinct kinetics conferred by the two RbcS types. Another approach consists of performing molecular dynamics simulations, which might explain the influence of remote residues on the catalytic site. Determining the 3D structure of a T-type Rubisco would be the ultimate strategy to identify the particular structural features of this enzyme (see Outstanding Questions).

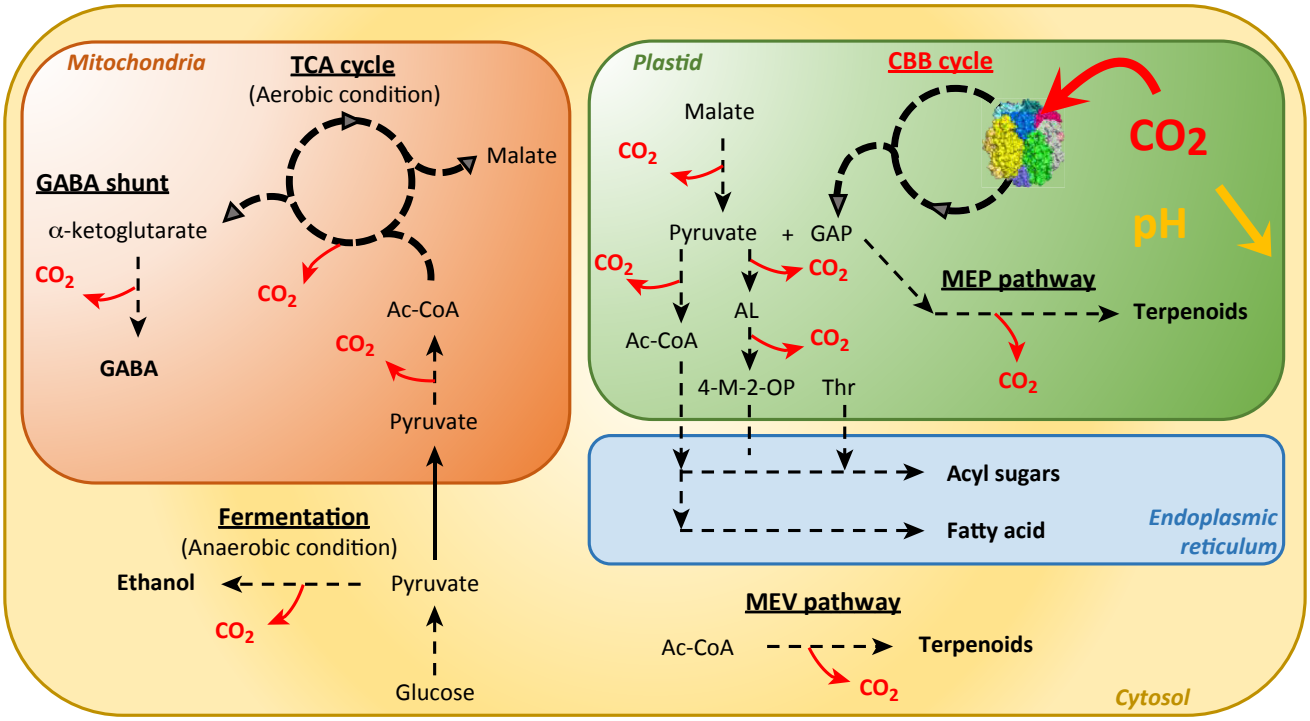
Physiological Roles of T-Type Rubisco

Despite the large presence of T-type RbcS in various plant phyla, little information is available concerning their biochemical properties and physiological roles. The main reason is that Rubisco has been mainly characterized in angiosperm species, where, at the whole leaf level, M-type *RbcS* are predominantly expressed. Indeed, although *RbcS-T* genes have been identified, notably through genomic data, their expression at the RNA or protein level does not seem to have drawn much attention, in contrast to M-type *RbcS*, which are largely expressed in chlorophyllian tissues.

The first biochemical characterization of a T-type RbcS was made by Morita *et al.* [8] who identified five *RbcS* genes in rice, one of which (*OsRbcS1*) actually fits in the T-type cluster and

is expressed in nonphotosynthetic tissues such as leaf sheath, culm, anther, and root central cylinder. Its overexpression in the leaf blade resulted in increased catalytic turnover and reduced affinity (increased K_m) for CO_2, thus reminiscent of high-catalytic-activity Rubisco typically found in algae or C4 plants, in which this enzyme operates in a high CO_2 microenvironment [8]. These authors further propose that this particular Rubisco recycles CO_2 released by metabolic pathways. More recently, the same team identified T-type *RbcS* genes (called *OsRbcS1*-like *RbcS*) in foxtail millet (expressed in seeds), tomato (stamen, pistil, and green fruit), *L. japonicus* (root, nodule, seed, and different floral organs), grape (mature leaf and green berry), and *S. moellendorffii* (rhizome and root) [9]. These authors proposed that T-type RbcS operate mainly in nonphotosynthetic organs where they are involved in metabolic pathways other than photosynthetic CO_2 fixation (Figure 4).

A more recent study identified in tobacco a T-type *RbcS*, *NtRbcS-T*, which is specifically expressed in glandular trichomes. As no M-type *RbcS* is expressed in these trichomes, the



Trends in Plant Science

Figure 4. Examples of Metabolic Pathways That Could Provide a CO_2-Rich Environment and Lead to the Acidification of the Plastid Stroma, Both Fitting the T-Type Rubisco Activity. T-type *RbcS* are mostly expressed in tissues in which CO_2 concentration is expected to be high [7,9]. Several metabolic pathways could lead to high CO_2 release, depending on the tissue. In rice, tomato, and *Lotus japonicus*, T-type *RbcS* is expressed in stamen, where high respiratory activity, especially in anther tapetum cells, is essential for pollen development and fertility [9,34]. In *L. japonicus*, T-type *RbcS* was also expressed in nodules where N_2 reduction requires intense respiration and high CO_2 fixation [9,54]. In tomato, T-type *RbcS* was also expressed in green fruits where high production of GABA associated with CO_2 release takes place though the GABA shunt [9,55]. In developing seeds of some oil-producing species such as soybean, a Rubisco plays a role in recycling the CO_2 released from the high respiratory activity related to the fatty acid synthesis [32,33]. In tobacco, T-type *RbcS* was specifically expressed in glandular trichomes where an important production of specialized metabolites, namely, diterpenes (MEP pathway) and acyl sugars, accounts for a large release of CO_2 within chloroplasts [7]. Similarly, other plant species with glandular trichomes were shown to express a T-type *RbcS* and to produce huge amounts of other terpenoids. This is the case of tomato, some trichomes of which mainly produce sesquiterpenes through the MEP pathway, known to release CO_2 in the cytosol [29]. 4-M-2-OP, 4-methyl-2-oxopentanoate; Ac-CoA, acetyl-CoA; AL, acetolactate; CBB, Calvin–Benson–Bassham; GABA, gamma-aminobutyric acid; GAP, glyceraldehyde-3-phosphate; MEP, methylerythritol 4-phosphate; MEV, mevalonate; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; TCA, tricarboxylic acid; Thr, threonine.

authors directly compared T-type Rubisco (from glandular trichomes) and M-type Rubisco (from leaf tissues cleared of trichomes). The former had a more acidic activity versus pH profile than the latter [7]. *NtRbcS-T* or *NtRbcS-M* (one of the *RbcS* expressed in the mesophyll) were individually expressed in a *Chlamydomonas* strain deleted of its two *RbcS* genes and which was previously used to functionally express *RbcS* from different plant species (e.g., [21,24,25]). T- and M-type Rubisco were purified and it turned out that NtRbcS-T conferred higher K_m for CO_2 and higher V_{max} [7], which is in line with the OsRbcS1 properties [8]. Glandular trichomes synthesize large amounts of specialized metabolites such as diterpenes and sucrose esters, which protect the plant against pathogens [26]. This metabolism is associated with high release of CO_2, especially in the plastids [7] (Figure 4). However, the thick cell wall and cuticle of glandular trichomes prevent gas exchange with the atmosphere and lead to high cellular CO_2 concentration [27], which possibly results in lowering the pH compared with the stroma pH (~8) of typical photosynthetic tissues when the photosynthetic chain is active. In this context, NtRbcS-T might adapt Rubisco to this particular environment, namely, high CO_2 concentration and lower pH, where the enzyme might be involved in recycling the CO_2 released during the large production of specialized metabolites.

In agreement with this, Balcke *et al.* [28] recently suggested that photosynthetic-type VI trichomes from tomato have a recycling mechanism allowing preserving the CO_2 released from terpenoid and lipid biosynthesis as well as from any other metabolic pathways producing CO_2 [28]. Interestingly, a T-type RbcS was identified in tomato [7]. Balcke *et al.* showed through isotopic labeling that the initial fixation of carbon required for the high production of specialized metabolites in these tomato trichomes takes place in leaf mesophyll cells. Therefore, reduced Calvin–Benson–Bassham cycle and associated Rubisco activities in trichomes aim mainly at saving CO_2, increasing productivity of these sink organs [28,29].

Most of the spermatophyte species with a T-type RbcS identified possess secretory organs such as glandular trichomes, colleters, nectaries, or secretory cavities [7]. Interestingly, a transcriptomic analysis of the secretory cavities of grapefruit showed that an *RbcS-T* gene was highly expressed in the cavity cells that synthesize and secrete essential oils [10]. Determining whether a T-type Rubisco operates in secretory organs of other species is required to generalize the hypothesis that a T-type Rubisco plays a major role in secretory organs with a very active CO_2-generating metabolism. However, T-type RbcS are not restricted to secretory organs since they were identified in various nonphotosynthetic tissues such as leaf sheath, culm, root, seed, nodule, nonmature fruit, and/or different floral organs [8,9]. A common feature might be the distance between these organs and the stomata involved in gas exchange with the atmosphere.

Can the presence of a T-type Rubisco be predicted in other cases? In developing seeds of some oil-producing species, Rubisco plays a role in recycling the CO_2 released from the high respiratory activity related to the fatty acid synthesis [30–33]. A T-type Rubisco might be involved in this activity in soybean in which a T-type *RbcS* has been identified [7]. However, no T-type *RbcS* has been identified in the oilseed rape genome, nor in other Brassicaceae genomes. It cannot be excluded that in species lacking T-type RbcS, an M-type RbcS might have evolved to be adapted to a high CO_2 environment.

RbcS-T expression was observed in anthers of tomato, *L. japonicus*, and rice [8,9]. Because pollen grains consume a large amount of carbon and energy during germination, the starch level constitutes a checkpoint during pollen maturity [34]. Similar to seed filling, efficient starch

loading from anther tissues, the tapetum in particular, into the pollen grain may require a CO_2 recycling mechanism to save carbon released from the metabolism.

Stomata are absent in the epidermal layer of woody tissues, limiting gaseous diffusion in metabolically active greenish bark sublayers [35]. Nevertheless, inorganic carbon refixation has also been described in these tissues, thanks to the high CO_2 concentration released from respiration [35–38]. This CO_2 recycling mechanism was shown to significantly contribute to the growth of branches in *Eucalyptus miniata* [37]. T-type *RbcS* genes have been identified in gymnosperm species as well as in angiosperm woody species (including *Eucalyptus*) and might thus play a role in this mechanism.

To broaden the characterization of T-type Rubisco, it would be interesting to biochemically characterize T-type Rubisco in more ancient phyla such as pteridophytes and bryophytes, and determine if the enzyme shares the same parameters as those displayed by tobacco and rice Rbc-T. From a physiological point of view, it would be interesting to determine in these phyla whether T-type Rubisco functions as a net carboxylase or as a CO_2 recycling enzyme. More T-type than M-type *RbcS* were identified in these phyla. We might speculate that *RbcS-T* could be the predominant, and possibly the exclusive, form in some of these species for which no *RbcS-M* genes were identified. Following this hypothesis, it would make sense that these species, which appeared before the large increase in the atmospheric O_2 concentration, had a Rubisco more adapted to an atmosphere with a lower O_2 (i.e., higher catalytic rate and lower capacity to discriminate between O_2 and CO_2). This is even more relevant for bryophytes, which grow close to the soil in a high CO_2 environment due to the soil respiration [39]. Consistent with this, the M type would have evolved to better discriminate CO_2 from O_2 when the latter had its concentration increased considerably in the atmosphere. Clearly, expression and biochemical analysis of T- and M-type Rubisco from ancient phyla is required to evaluate this hypothesis and better identify the function of these two clusters in these phyla.

Concluding Remarks

The T-type *RbcS* genes are widely spread, not only in seed plants, where they are mainly expressed in restricted tissues but also in pteridophytes and bryophytes, thus designating *RbcS-T* as an ancestral type in eukaryotes. The few reports on the biochemical characterization of RbcS-T containing Rubisco in land species revealed an enzyme with higher catalytic rate, lower CO_2 affinity, and acidic shift of the activity versus pH profile. This suggests that the RbcS-T might adapt Rubisco to high CO_2 concentration typically found in cell types where photosynthesis is low or absent but where intense metabolic pathways that generate CO_2 take place (Figure 4). However, more experimental work with other seed species is required to generalize this hypothesis. In addition, biochemical analysis of T- and M-type Rubisco from pteridophytes and bryophytes is required to determine their respective roles in these ancient phyla. Sequence comparison of T- and M-type RbcS has pointed out several T-type-specific residues, some of which might interact with RbcL and be responsible for the modified activity. However, the 3D structure of a T-type Rubisco should be experimentally determined to further explore the particular structural features of this enzyme (see Outstanding Questions). As Rubisco is often the limiting step in photosynthesis, this enzyme has been a target for metabolic engineering (Box 1). It would therefore be interesting to determine whether T-type RbcS could be used to improve the catalytic properties of Rubisco.

Outstanding Questions

What are the structural features that are responsible for the biochemical properties of the T-type Rubisco? Considering that in seed plants the T-type enzyme is restricted to particular tissues, its purification and crystallization might represent a major issue.

In seed plants, T-type Rubisco seems to be restricted to particular tissues. How diverse are the physiological roles of T-type Rubisco in these tissues?

What are the distribution and expression profiles of T-type *RbcS* in seedless species? Considering that in some pteridophyte and bryophyte species only T-type *RbcS* have been identified, can a T-type Rubisco be the major isoform in some of these species?

What are the biochemical properties of T-type Rubisco in pteridophytes and bryophytes? In these species, does it function as a net CO_2 fixing or as a CO_2 refixing enzyme?

Can T-type RbcS be used to engineer plants for improved photosynthesis?

Box 1. Engineering Rubisco?

Given the slow catalytic turnover of Rubisco and its wasteful side activity, which uses O_2 instead of CO_2, this enzyme is usually considered as a limiting step for photosynthesis and thus food production. The growth of the world population and the issues related to climate changes (photorespiration increases with temperature) have put pressure for improving the photosynthesis yield. Several trails have been proposed, such as increasing the CO_2 concentration close to Rubisco by mimicking the C4 metabolism, directly improving the Rubisco activity (increased V_{max} and/or better discrimination between CO_2 and O_2; for review, see [40–42]). The large subunit, which bears the active site, is an obvious target. For instance, replacement in tobacco (a C3 plant) of RbcL Met-309 by Ile (a residue found in some C4 plants) provided Rubisco with C4-like catalytic properties [43]. However, the plastid localization of the *RbcL* gene makes genetic engineering more complicated, as transformation of the chloroplast genome is still difficult in many species. Several reports have shown that RbcS also influences the catalytic activity. For instance, expression of a sorghum (C4 plant) RbcS increased the Rubisco catalytic turnover in transgenic rice, a C3 plant [44]. A wider range of T-type RbcS will have to be characterized to determine whether it might be used for engineering Rubisco. The nuclear localization of *RbcS* genes makes genetic transformation easier. Yet, one drawback is that RbcS is encoded by a gene family, and thus several genes have to be replaced to obtain a homogenous Rubisco population. Recent development of clustered regularly interspaced short palindromic repeats–CRISPR-associated protein 9 (CRISPR–Cas9)-based genome editing is expected to facilitate gene replacement at several loci. Finally, one should not forget that Rubisco assembly requires chaperones and that its functioning depends on proteins involved in its maintenance (for review, see [41,45]). This might represent an additional challenge when a heterologous Rubisco subunit is not adapted to the assembly and maintenance machinery of the host species [46,47].

Acknowledgments

This work was supported by the Belgian Fund for Scientific Research, the Interuniversity Poles of Attraction Program (Belgian State, Scientific, Technical, and Cultural Services), and an EU Marie Skłodowska-Curie fellowship (Project ID: 658932) to M.P.

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tplants.2018.02.006.

References

1. Spreitzer, R.J. and Salvucci, M.E. (2002) Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.* 53, 449–475
2. Andersson, I. and Backlund, A. (2008) Structure and function of Rubisco. *Plant Physiol. Biochem.* 46, 275–291
3. Tabita, F.R. *et al.* (2007) Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* 71, 576–599
4. Parry, M.A.J. *et al.* (2003) Manipulation of Rubisco: the amount, activity, function and regulation. *J. Exp. Bot.* 54, 1321–1333
5. Portis, A.R. and Parry, M.A.J. (2007) Discoveries in Rubisco (ribulose 1, 5-bisphosphate carboxylase/oxygenase): a historical perspective. *Photosynth. Res.* 97, 121–143
6. Spreitzer, R.J. (2003) Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* 414, 141–149
7. Laterre, R. *et al.* (2017) Photosynthetic trichomes contain a specific Rubisco with a modified pH-dependent activity. *Plant Physiol.* 173, 2110–2120
8. Morita, K. *et al.* (2014) Unusual small subunit that is not expressed in photosynthetic cells alters the catalytic properties of Rubisco in rice. *Plant Physiol.* 164, 69–79
9. Morita, K. *et al.* (2016) Identification and expression analysis of non-photosynthetic Rubisco small subunit, OsRbcS1-like genes in plants. *Plant Gene* 8, 26–31
10. Voo, S.S. *et al.* (2012) Assessing the biosynthetic capabilities of secretory glands in citrus peel. *Plant Physiol.* 159, 81–94
11. Goodstein, D.M. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, 1178–1186
12. Matasci, N. *et al.* (2014) Data access for the 1,000 Plants (1KP) project. *Gigascience* 3, 17
13. Xie, Z. *et al.* (2008) A systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. *Plant J.* 54, 349–361
14. Ahkami, A. *et al.* (2015) Multiple levels of regulation determine monoterpenoid essential oil compositional variation in the mint family. *Mol. Plant* 8, 188–191
15. NCBI Resource Coordinators (2017) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 45, D12–D17
16. Brutnell, T.P. *et al.* (2010) *Setaria viridis*: a model for C4 photosynthesis. *Plant Cell* 22, 2537–2544
17. Lowry, D.B. *et al.* (2015) The genetics of divergence and reproductive isolation between ecotypes of *Panicum hallii*. *New Phytol.* 205, 402–414
18. Hori, K. *et al.* (2014) *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5, 3978
19. Delwiche, C.F. and Cooper, E.D. (2015) The evolutionary origin of a terrestrial flora. *Curr. Biol.* 25, R899–R910
20. Andersson, I. and Taylor, T.C. (2003) Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* 414, 130–140
21. Genkov, T. *et al.* (2010) Functional hybrid rubisco enzymes with plant small subunits and algal large subunits: engineered rbcS cDNA for expression in *Chlamydomonas*. *J. Biol. Chem.* 285, 19833–19841
22. Schreuder, H.A. *et al.* (1993) Formation of the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase by a disorder-order transition from the unactivated to the activated form. *Proc. Natl. Acad. Sci. U. S. A.* 90, 9968–9972

23. Karkehabadi, S. *et al.* (2005) Chimeric small subunits influence catalysis without causing global conformational changes in the crystal structure of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Biochemistry* 44, 9851–9861
24. Wachter, R.M. *et al.* (2013) Activation of interspecies-hybrid Rubisco enzymes to assess different models for the Rubisco-Rubisco activase interaction. *Photosynth. Res.* 117, 557–566
25. Meyer, M.T. *et al.* (2012) Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19474–19479
26. Wagner, G.J. *et al.* (2004) New approaches for studying and exploiting an old protuberance, the plant trichome. *Ann. Bot.* 93, 3–11
27. Kandra, L. and Wagner, J. (1988) Studies of the site and mode of biosynthesis of tobacco trichome exudate components. *Arch. Biochem. Biophys.* 265, 425–432
28. Balcke, G. *et al.* (2017) Multiomics of tomato glandular trichomes reveals distinct features of central carbon metabolism supporting high productivity of specialized metabolites. *Plant Cell* 29, 960–983
29. Tissier, A. (2018) Plant secretory structures: more than just reaction bags. *Curr. Opin. Biotechnol.* 49, 73–79
30. Goffman, F.D. (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. *Plant Physiol.* 138, 2269–2279
31. Schwender, J. *et al.* (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432, 779–782
32. Ruuska, S.A. (2004) The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiol.* 136, 2700–2709
33. Allen, D.K. *et al.* (2009) The role of light in soybean seed filling metabolism. *Plant J.* 58, 220–234
34. Datta, R. *et al.* (2002) Starch biosynthesis during pollen maturation is associated with altered patterns of gene expression in maize. *Plant Physiol.* 130, 1645–1656
35. Cernusak, L.A. *et al.* (2001) Carbon isotope discrimination in photosynthetic bark. *Oecologia* 128, 24–35
36. Pfanz, H. *et al.* (2002) Ecology and ecophysiology of tree stems: corticular and wood photosynthesis. *Naturwissenschaften* 89, 147–162
37. Cernusak, L.A. and Hutley, L.B. (2011) Stable isotopes reveal the contribution of corticular photosynthesis to growth in branches of *Eucalyptus miniata*. *Plant Physiol.* 155, 515–523
38. Berveiller, D. *et al.* (2007) Tree stem phosphoenolpyruvate carboxylase (PEPC): lack of biochemical and localization evidence for a C4-like photosynthesis system. *New Phytol.* 176, 775–781
39. Hanson, D.T. *et al.* (2014) Photosynthesis in bryophytes and early land plants. In *Advances in Photosynthesis and Respiration* (Hanson, D.T. and Rice, S.K., eds), pp. 95–112, Springer Science + Business Media
40. Sharwood, R.E. *et al.* (2016) Prospects for improving CO_2 fixation in C3-crops through understanding C4-Rubisco biogenesis and catalytic diversity. *Curr. Opin. Plant Biol.* 31, 135–142
41. Sharwood, R.E. (2017) Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. *New Phytol.* 213, 494–510
42. Wilson, R.H. and Whitney, S.M. (2017) Improving CO_2 fixation by enhancing Rubisco performance. In *Directed Enzyme Evolution: Advances and Applications* (Alcalde, M., ed.), pp. 101–126, Springer International Publishing
43. Whitney, S.M. *et al.* (2011) Isoleucine 309 acts as a C4 catalytic switch that increases ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) carboxylation rate in Flaveria. *Proc. Natl. Acad. Sci. U. S. A.* 108, 14688–14693
44. Ishikawa, C. *et al.* (2011) Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. *Plant Physiol.* 156, 1603–1611
45. Bracher, A. *et al.* (2017) Biogenesis and metabolic maintenance of Rubisco. *Annu. Rev. Plant Biol.* 68, 29–60
46. Yeates, T.O. and Wheatley, N.M. (2017) Putting the RuBisCO pieces together. *Science* 358, 1253–1254
47. Aigner, H. *et al.* (2017) Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science* 358, 1272–1278
48. Webb, B. and Sali, A. (2014) Protein structure modeling with MODELLER. In *Protein Structure Prediction* (Kihara, D., ed.), pp. 1–15, Springer
49. Clarke, J.T. *et al.* (2011) Establishing a time-scale for plant evolution. *New Phytol.* 192, 266–301
50. Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320
51. Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution (N. Y.)* 39, 783–791
52. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425
53. Kumar, S. *et al.* (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33, 1870–1874
54. Lundquist, P.O. (2005) Carbon cost of nitrogenase activity in Frankia-*Alnus incana* root nodules. *Plant Soil* 273, 235–244
55. Carrari, F. *et al.* (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol.* 142, 1380–1396