



Rubisco is evolving for improved catalytic efficiency and CO₂ assimilation in plants

Jacques W. Bouvier^a, David M. Emms^a, and Steven Kelly^{a,1}

Edited by Krishna Niyogi, University of California, Berkeley, CA; received November 30, 2023; accepted January 25, 2024

Rubisco is the primary entry point for carbon into the biosphere. However, rubisco is widely regarded as inefficient leading many to question whether the enzyme can adapt to become a better catalyst. Through a phylogenetic investigation of the molecular and kinetic evolution of Form I rubisco we uncover the evolutionary trajectory of rubisco kinetic evolution in angiosperms. We show that *rbcl* is among the 1% of slowest-evolving genes and enzymes on Earth, accumulating one nucleotide substitution every 0.9 My and one amino acid mutation every 7.2 My. Despite this, rubisco catalysis has been continually evolving toward improved CO₂/O₂ specificity, carboxylase turnover, and carboxylation efficiency. Consistent with this kinetic adaptation, increased rubisco evolution has led to a concomitant improvement in leaf-level CO₂ assimilation. Thus, rubisco has been slowly but continually evolving toward improved catalytic efficiency and CO₂ assimilation in plants.

rubisco | evolution | adaptation | kinetics | photosynthesis

Ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) converts atmospheric CO₂ into the sugars that fuel the majority of life on Earth. The enzyme evolved ~3 billion years ago when the atmosphere contained high levels of CO₂ ($\geq 10,000\%$ present atmospheric levels) and comparatively little O₂ ($\leq 0.1\%$ present atmospheric levels) (Fig. 1) (1–7). Since emergence, the enzyme has helped guide the atmosphere on a trajectory of increasing O₂ and declining CO₂ (1, 8) such that current concentrations of CO₂ (0.04%) and O₂ (20.95%) are inverted compared to when the enzyme first evolved (Fig. 1).

Although all extant rubisco are descended from a single ancestral rubisco-like protein (10–12), the enzyme is found in a variety of compositional forms across the tree of life (Fig. 1) (13, 14). The simplest manifestations are the Form II and Form III variants found in protists, archaea, and some bacteria which are composed of a dimer, or dimers, of the ~50 kDa rubisco large subunit (RbcL) (14–17). In contrast, Form I rubisco is a hexadecamer comprised of four RbcL dimers organized in an antiparallel core capped at either end by the ~15 kDa rubisco small subunit (RbcS) (14, 18). Of these three Forms, only Form I and II have been recruited for oxygenic photosynthesis (17), with Form I being responsible for the vast majority of global CO₂ assimilation (17, 19).

Within Form I rubisco the active site is located in RbcL (16, 20, 21). As a result, interspecific differences in Form I kinetics are primarily attributable to sequence variation in RbcL (22–33). Despite not playing a direct role in catalysis, RbcS influences the function of rubisco (34, 35) and its incorporation in the holoenzyme enables its higher kinetic efficiency (36). Specifically, RbcS enhances the stability and assembly of the holoenzyme complex (21, 37–42), improves the efficiency of CO₂ binding (43), and is thought to act as a reservoir for CO₂ accumulation (44). Accordingly, rubisco function is altered when RbcS is mutated (45–47), or when chimeric holoenzymes are created in vivo (48–52) and in vitro (53–58). Moreover, there is increasing recognition of the importance of both environment (59) and organ-specific (60, 61) differences in plant RbcS isoform expression on holoenzyme catalysis. However, even though RbcS influences holoenzyme function, sequence variation in RbcL remains the primary determinant of variation in kinetics (22–33).

Although there is kinetic variability between rubisco orthologs, the enzyme is considered to be an inefficient catalyst. For example, the maximum substrate-saturated turnover rate of Form I rubisco ($< 12 \text{ s}^{-1}$) (62) is slower than average (63). In addition, rubisco catalyzes a reaction with O₂ (64–66) that is competitive with CO₂ and results in the loss of fixed carbon via photorespiration (67–69). As a consequence, rubisco appears poorly suited to the current O₂-rich/CO₂-poor atmosphere (Fig. 1). Moreover, it appears that instead of improving enzyme function, multiple lineages have evolved alternative strategies to overcome rubisco's shortcomings. For example, higher rates of CO₂ assimilation are often achieved either by synthesizing large quantities of rubisco [$\sim 50\%$ of soluble protein in

Significance

Rubisco is the most abundant enzyme on Earth and is the source of almost all biological carbon. Here we uncover the trajectory of rubisco adaptive evolution in plants. We reveal that rubisco has experienced continuous directional selection toward higher carbon dioxide/oxygen specificity, carboxylase turnover, and carboxylation efficiency. Moreover, we find that this directional selection toward improved catalytic efficiency has resulted in a corresponding improvement in leaf-level CO₂ assimilation. Together, these findings have significant implications for our understanding of the past, present, and future of rubisco evolution.

Author affiliations: ^aDepartment of Biology, University of Oxford, Oxford OX1 3RB, United Kingdom

Preprint servers: This article is available as a preprint on bioRxiv, DOI: [10.1101/2022.07.06.498985](https://doi.org/10.1101/2022.07.06.498985).

Author contributions: J.W.B., D.M.E., and S.K. designed the research; J.W.B. and S.K. conceived the study; J.W.B. performed the research; J.W.B. and S.K. analyzed the data; and J.W.B. and S.K. wrote the paper.

Competing interest statement: S.K. is co-founder of Wild Bioscience Ltd.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

¹To whom correspondence may be addressed. Email: steven.kelly@biology.ox.ac.uk.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2321050121/-/DCSupplemental>.

Published March 5, 2024.

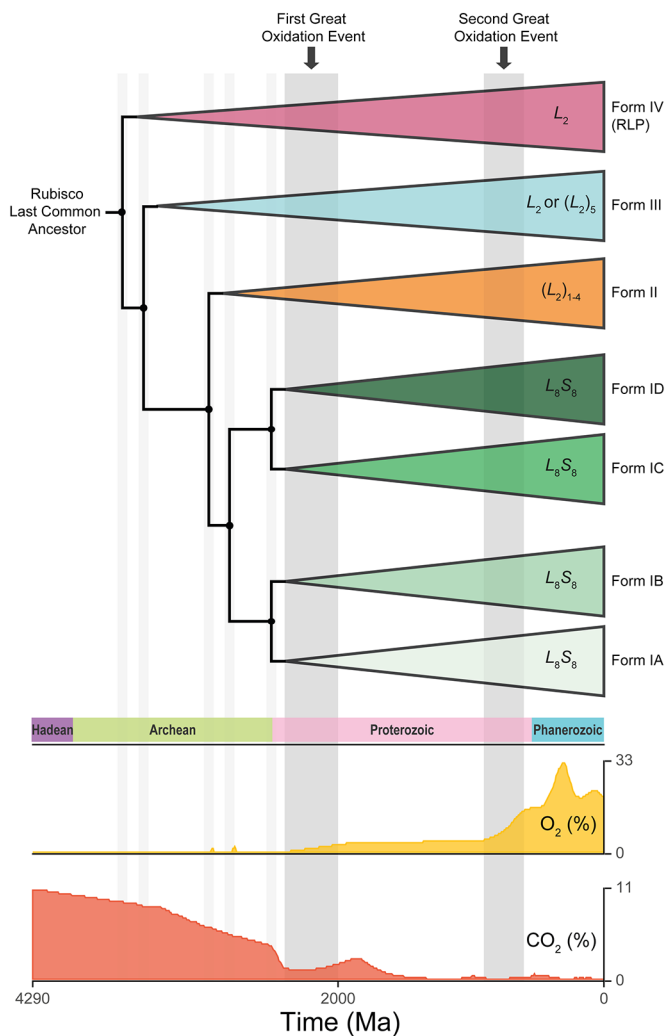


Fig. 1. The evolutionary history of rubisco in the context of atmospheric CO₂ (%) and O₂ (%) following divergence from the ancestral rubisco-like protein (RLP). Important branch points in the phylogeny at which rubisco diverged into different evolutionary lineages are indicated by gray vertical bars. To provide additional context, the time-period at which the First and Second Great Oxidation events occurred along this evolutionary trajectory are also labeled and referenced as gray vertical bars. Graphics of atmospheric CO₂ and O₂ levels were adapted from the *TimeTree* resource [<http://www.timetree.org>; (9)].

plants (70) and some microbes (71, 72)], or by operating CO₂-concentrating mechanisms (73–75). As a result, many have questioned whether the enzyme is already perfectly adapted and whether further kinetic improvements are possible (16, 65, 69, 76–80). Obtaining answers to these questions would shed light on the “rubisco paradox”—helping to explain why this enzyme of such paramount importance appears poorly adapted for its role.

The initial hypothesis that attempted to explain the above rubisco paradox proposed that rubisco is constrained by catalytic trade-offs that limit the enzyme’s adaptation. This theory was pioneered by two studies (81, 82) which found antagonistic correlations between rubisco kinetic traits and proposed that these trade-offs were caused by constraints on its catalytic mechanism. However, recent evidence has questioned this hypothesis as the sole mechanism to explain the rubisco paradox. Specifically, analysis of larger species sets have revealed that kinetic trait correlations are not strong (83–85). In addition, phylogenetic signal in rubisco kinetics causes kinetic trait correlations to be overestimated unless phylogenetic comparative methods are employed (22, 23). Thus, when larger datasets are analyzed with phylogenetic methods, the

strength of catalytic trade-offs are substantially reduced (22, 23). Instead, phylogenetic constraints have had a larger impact on limiting enzyme adaptation compared to catalytic trade-offs (22, 23). These recent findings motivate a reevaluation of the rubisco paradox, and an investigation of whether rubisco is evolving for improved catalysis and CO₂ assimilation in plants.

Here, we address these issues through a phylogenetic interrogation of the molecular and kinetic evolution of the Form I holoenzyme. We reveal that RbcL has evolved at a slower rate than >98% of all other gene/protein sequences across the tree of life. Through simultaneous evaluation of molecular and kinetic evolution of rubisco during the radiation of C₃ angiosperms, we reveal that the enzyme has been continually evolving toward improved CO₂/O₂ specificity, carboxylase turnover rate, and carboxylation efficiency. Furthermore, we demonstrate that enhanced rubisco evolution is associated with increased rates of CO₂ assimilation and higher photosynthetic nitrogen-use efficiencies. Thus, rubisco is not perfectly adapted, but is slowly evolving toward improved catalytic efficiency and CO₂ assimilation.

Results

RbcL Has Evolved Slower than RbcS and Has Experienced Stronger Purifying Selection. Sequences encoding Form I rubisco were obtained from the National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>). This dataset was filtered to retain sequences for a given species only if a full-length sequence for both *rbcL* and *rbcS* were present. Although *rbcL* exists as a single copy gene in all species, many species harbor multiple *rbcS* genes in their genomes. Thus, for each species, a single *rbcL* sequence and all available *rbcS* sequences were taken forward. In total, this resulted in a set of 488 *rbcL*/RbcL and 1,140 *rbcS*/RbcS sequences across 488 species (*SI Appendix, Supplemental File 1, Fig. S1, and Table S1*).

In order to compare the rate at which the two rubisco subunits have evolved, species were partitioned into distinct taxonomic groups comprising the red algae (*Rhodophyta*; *n* = 201), the SAR supergroup (*Stramenopiles, Alveolates, and Rhizaria*; *n* = 129), the bacteria (*Bacteria*; *n* = 78), the land plants (*Streptophyta*; *n* = 68) and the green algae (*Chlorophyta*; *n* = 12) (*SI Appendix, Fig. S1, Supplemental File 1, and Table S1*). Hereinafter, the total amount of molecular evolution of the nucleotide sequences (*rbcL* and *rbcS*) and the total amount of molecular evolution of the protein sequences (RbcL and RbcS) in a taxonomic group is referred to as “the extent of nucleotide evolution” and “the extent of protein evolution”, respectively. The term “the extent of molecular evolution” jointly refers to both.

Comparison of the two rubisco subunits revealed that the extent of molecular evolution in *rbcL*/RbcL is lower than that experienced by *rbcS*/RbcS (Fig. 2*A*). Specifically, the nucleotide and protein sequences of *rbcL*/RbcL have evolved at a rate equivalent to ~50% and ~25% of *rbcS*/RbcS on average across all taxa, respectively (Fig. 2*A* and *B*). This was not an artifact of the higher gene copy number of *rbcS*, as a 1,000 bootstrapped stratified sampling recovered the same result when only a single *rbcS*/RbcS sequence was randomly sampled per species (*Materials and Methods; Fig. 2A*). Therefore, *rbcL*/RbcL has explored less nucleotide and protein sequence space than *rbcS*/RbcS in the same sets of species over the same period of time (Fig. 2*A* and *B*). Furthermore, *rbcL* also experienced fewer amino acid changes per change in nucleotide sequence compared to *rbcS* (Fig. 2*A* and *C*), indicating a higher degree of purifying selection. Thus, *rbcL*/RbcL has evolved more slowly and has been subject to a higher degree of functional constraint on the encoded protein sequence than *rbcS*/RbcS.

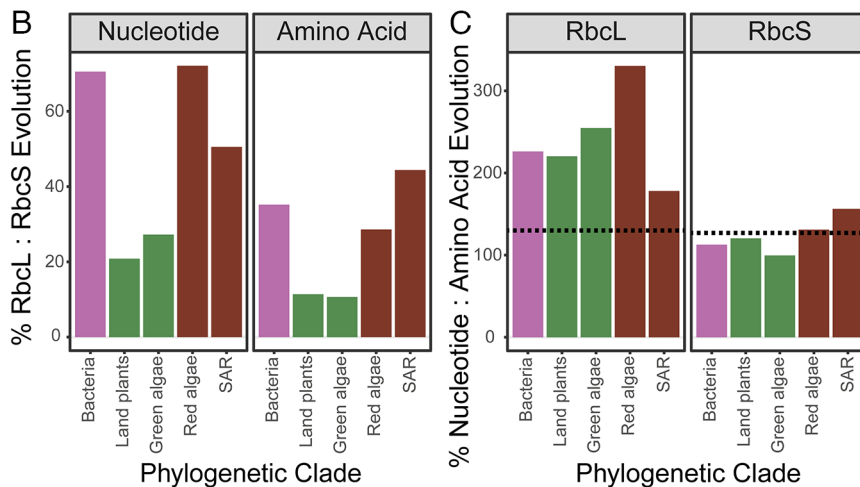
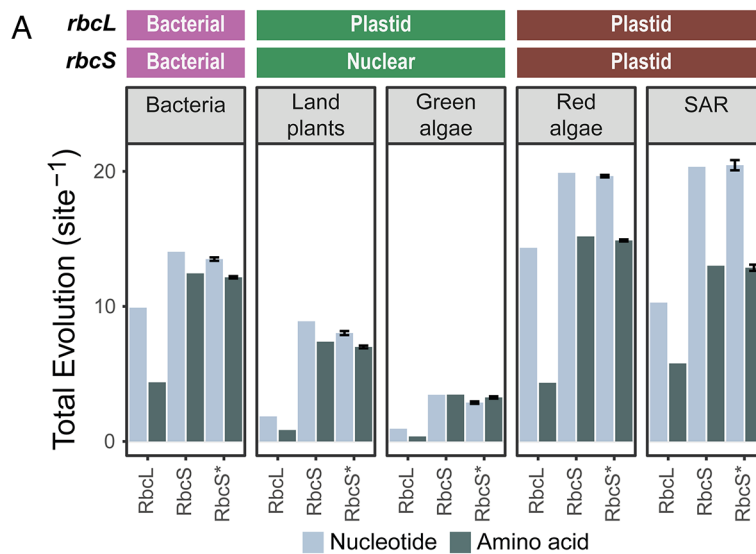


Fig. 2. The extent of molecular evolution in rubisco during the radiation of each taxonomic group. (A) Bar plot depicting the total amount of molecular evolution (substitutions per sequence site) in the nucleotide and protein sequences of Form I rubisco across taxonomic groups. RbcL: the extent of sequence evolution in the RbcL subunit. RbcS: the extent of sequence evolution in the RbcS subunit. RbcS*: the extent of sequence evolution in the RbcS subunit using 1,000 bootstrapped stratified sampling of *rbcS*/RbcS per species (*Materials and Methods*). The genome in which *rbcL* and *rbcS* genes reside within each group is indicated above the plot (bacterial, plastid, nuclear). Error bars represent ± 1 SD of the mean. (B) Bar plot depicting the percentage ratio (%) of nucleotide or amino acid evolution between each rubisco subunit (*rbcL* to *rbcS* and RbcL to RbcS, respectively) in each taxonomic group. The color of each bar is determined by the genome in which the *rbcL* and *rbcS* gene resides, following the color scale in A. (C) Bar plot depicting the percentage ratio (%) of nucleotide to amino acid evolution in each rubisco subunit (*rbcL* to RbcL and *rbcS* to RbcS, respectively) in each taxonomic group. The color of each bar is the same as described in B. Dashed lines indicate the expected ratio given a *rbcL* or *rbcS* sequence evolving in the absence of selection.

RbcL Is One of the Slowest-Evolving Genes in the Tree of Life. To evaluate the rate of molecular evolution in the context of all other genes in the species under consideration, the percentile rank of *rbcL*/RbcL and *rbcS*/RbcS was evaluated for all genes in all species (*Materials and Methods*). This revealed that 99.3% of all gene nucleotide sequences and 98.1% of all gene protein sequences evolved faster than *rbcL*/RbcL in the same sets of species over the same period of time (Fig. 3A and *SI Appendix, Supplemental File 1 and Table S2*). This held true even if *rbcL*/RbcL was only compared to the subset of genes that encode enzymes, with 99.2% of enzyme nucleotide sequences and 98.3% of enzyme protein sequences evolving faster than *rbcL*/RbcL (Fig. 3B and *SI Appendix, Supplemental File 1 and Table S3*). Furthermore, in land plants, *rbcL*/RbcL was also the slowest-evolving component of the Calvin–Benson–Bassham cycle (Fig. 3C and *SI Appendix, Supplemental File 1 and Tables S4 and S5*). This slow pace of evolution is not simply an artifact of being encoded in the plastid genome, as *rbcL*/RbcL was also one of the slowest-evolving genes/proteins in

bacteria which encode all of their genes in a single cytoplasmic genome. Thus, *rbcL*/RbcL is one of the slowest-evolving genes/enzymes in all species in which it is found, irrespective of the taxonomic group or genome in which it is encoded.

In contrast to *rbcL*/RbcL, considerable variability in the extent of molecular evolution in the rubisco small subunit was observed both within and between taxonomic groups (Fig. 3A and *SI Appendix, Supplemental File 1 and Table S2*). Analogous results in each taxonomic group were recovered when this analysis was restricted to the subset of genes that encode enzymes (Fig. 3B and *SI Appendix, Supplemental File 1 and Table S3*). Moreover, in land plants, *rbcS*/RbcS was the fastest-evolving component of the Calvin–Benson–Bassham cycle (Fig. 3C and *SI Appendix, Supplemental File 1 and Tables S4 and S5*). Thus, while the pace of molecular evolution in *rbcL*/RbcL is ubiquitously slow, the extent of molecular evolution of *rbcS*/RbcS is highly variable explaining the disparity in the rate of both subunits across the tree of life (Fig. 2B and *SI Appendix, Supplemental File 1 and Table S6*).

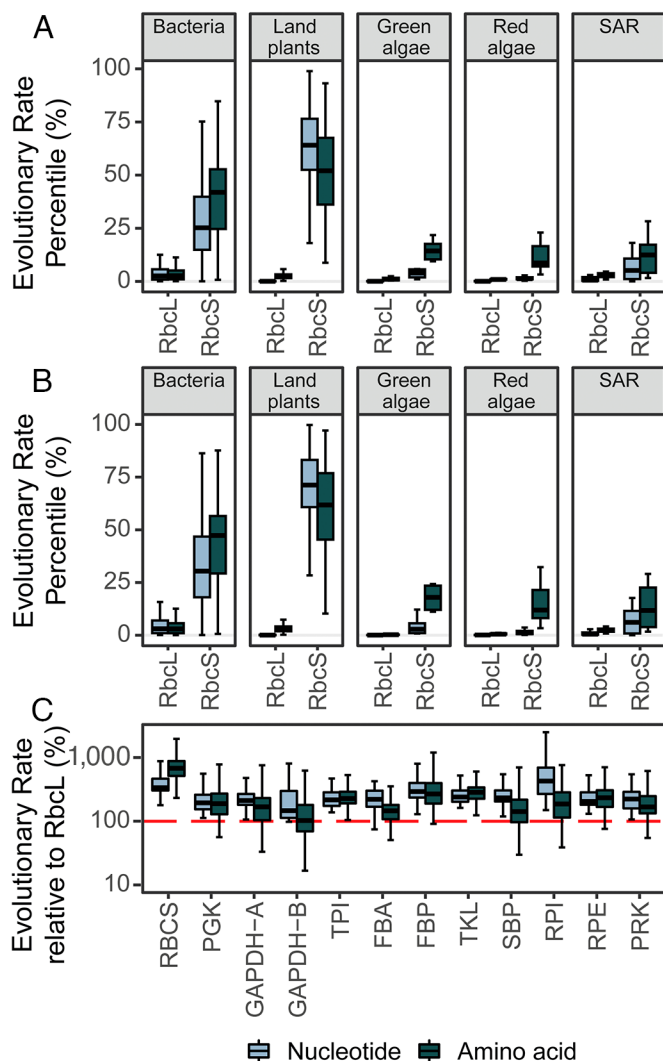


Fig. 3. The extent of molecular evolution in rubisco in the context of other genes. (A) Boxplot of the extent of molecular evolution (substitutions per sequence site) in the nucleotide and protein sequences of the *rbcL*/RbcL and *rbcS*/RbcS subunit expressed as a percentile (%) of that measured across all other genes and proteins, respectively. See also *SI Appendix, Supplemental File 1 and Table S2*. (B) As in A but calculating the percentile (%) extent of rubisco molecular evolution (substitutions per sequence site) relative to only the subset of genes and proteins in each species which encode enzymes. See also *SI Appendix, Supplemental File 1 and Table S3*. (C) Boxplot of the total amount of molecular evolution (substitutions per sequence site) in the nucleotide and protein sequences of each Calvin-Benson-Bassham cycle enzyme expressed as a percentage (%) of that measured in *rbcL*/RbcL (100%; red horizontal line) across land plants. Phosphoglycerate kinase: PGK. Glyceraldehyde-3-phosphate dehydrogenase A/B subunit: GAPDH-A/GAPDH-B. Triose phosphate isomerase: TPI. Fructose-bisphosphate aldolase: FBA. Fructose-1,6-bisphosphatase: FBP. Transketolase: TKL. Sedoheptulose-bisphosphatase: SBP. Ribose 5-phosphate isomerase: RPI. Ribulose-p-3-epimerase: RPE. Phosphoribulokinase: PRK. See also *SI Appendix, Supplemental File 1 and Tables S4 and S5*. The raw data for this figure can be found in *SI Appendix, Supplemental File 5*.

A similar variable rate was also observed for rubisco's ancillary chaperones (*SI Appendix, Supplemental File 1*). Thus, the rate of molecular evolution of *rbcL*/RbcL is ubiquitously low, and lower than *rbcS*/RbcS or any associated chaperone.

Rubisco Is Evolving for Improved Kinetic Efficiency in Plants.

Given that *rbcL* is among the slowest-evolving genes on Earth, the question arises as to whether its sequence evolution is adaptive and is improving the catalysis of the enzyme. We hypothesized

that if rubisco was undergoing directional selection for improved catalysis, then orthologs that have experienced the largest extent of molecular evolution would be the most efficient catalysts. To test this hypothesis, a dataset of kinetic measurements from C_3 angiosperms (22, 23, 83) was evaluated in the context of the molecular evolution of RbcL (Fig. 4A and B). This investigation focused on RbcL as it is the primary determinant of kinetics (22–33), and because sufficient sequence data for RbcS are unavailable. This analysis revealed that the more RbcL has evolved from the most recent common ancestral sequence, the better its CO_2/O_2 specificity ($S_{C/O}$; 10.1% variance explained, $P < 0.01$), CO_2 turnover rate (k_{catC} ; 4.6% variance explained, $P < 0.05$), and carboxylation efficiency (k_{catC}/K_C ; 3.8% variance explained, $P < 0.05$) (Fig. 4B). This result is not an artifact caused by potential systematic methodological biases associated with species sampling or potential uncertainties or errors in the underlying phylogenetic tree (*SI Appendix, Supplemental File 1*). Thus, rubisco has been adaptively evolving for improved $S_{C/O}$, k_{catC} , and k_{catC}/K_C during the radiation of the angiosperms.

Given that the origin of the angiosperms is estimated to have occurred 160 Mya (9) (Fig. 4A), it is possible to put the above kinetic change in the context of both molecular sequence changes and evolutionary time since the last common ancestor at the base of this clade (Table 1). As the large subunit acquired one nucleotide substitution every 0.9 My and one amino acid substitution every 7.2 My (*SI Appendix, Supplemental File 1 and Fig. S2*), each amino acid substitution resulted in an average increase in $S_{C/O}$ by $2.7 \times 10^{-1} \text{ mol mol}^{-1}$, in k_{catC} by $3.6 \times 10^{-2} \text{ s}^{-1}$, and in k_{catC}/K_C by $1.8 \times 10^{-3} \text{ s}^{-1} \mu\text{M}^{-1}$. This is equivalent to a relative improvement of 0.3% ($S_{C/O}$), 1.4% (k_{catC}), and 1.1% (k_{catC}/K_C) per amino acid substitution, and a relative improvement of 0.05% ($S_{C/O}$), 0.2% (k_{catC}), and 0.2% (k_{catC}/K_C) per million years. Thus, there has been improvement in rubisco kinetics during the radiation of the angiosperms at a rate that is dependent on the extent of its molecular sequence change.

Rubisco Is Evolving for Improved Leaf-Level CO_2 Assimilation.

Given that rubisco is evolving to become a better catalyst, we hypothesized that this adaptation would also drive adaptation in the rate of leaf-level CO_2 assimilation. To test this hypothesis, we analyzed a large dataset of photosynthetic measurements from C_3 angiosperms (86) in the context of the extent of their RbcL evolution (Fig. 5A–C). This revealed that the rate of leaf-level CO_2 assimilation was also dependent on the extent of molecular sequence change in rubisco, such that C_3 angiosperms with more evolved rubisco presented higher rates of CO_2 assimilation (A_{mass} ; 19.2% variance explained, $P < 0.001$) (Fig. 5B). This is not a consequence of increased nitrogen investment in the leaf, as the association between rubisco evolution and increased CO_2 assimilation is strengthened when measurements are controlled for leaf nitrogen content ($PNUE_{mass}$; 22.1% variance explained, $P < 0.001$) (Fig. 5B). Analogous results were obtained when measurements of CO_2 assimilation were evaluated on a leaf area basis (Fig. 5C). Together, these results are most parsimoniously explained by directional selection toward enhanced leaf-level CO_2 assimilation driven by the kinetic adaptation described above. Thus, the adaptive evolution of rubisco during the radiation of the angiosperms has resulted in the improvement of leaf-level CO_2 assimilation.

Discussion

Rubisco is the primary entry point for carbon into the biosphere and is responsible for fixing 250 billion tons of CO_2 annually (19). Despite this immense throughput, the enzyme is a surprisingly

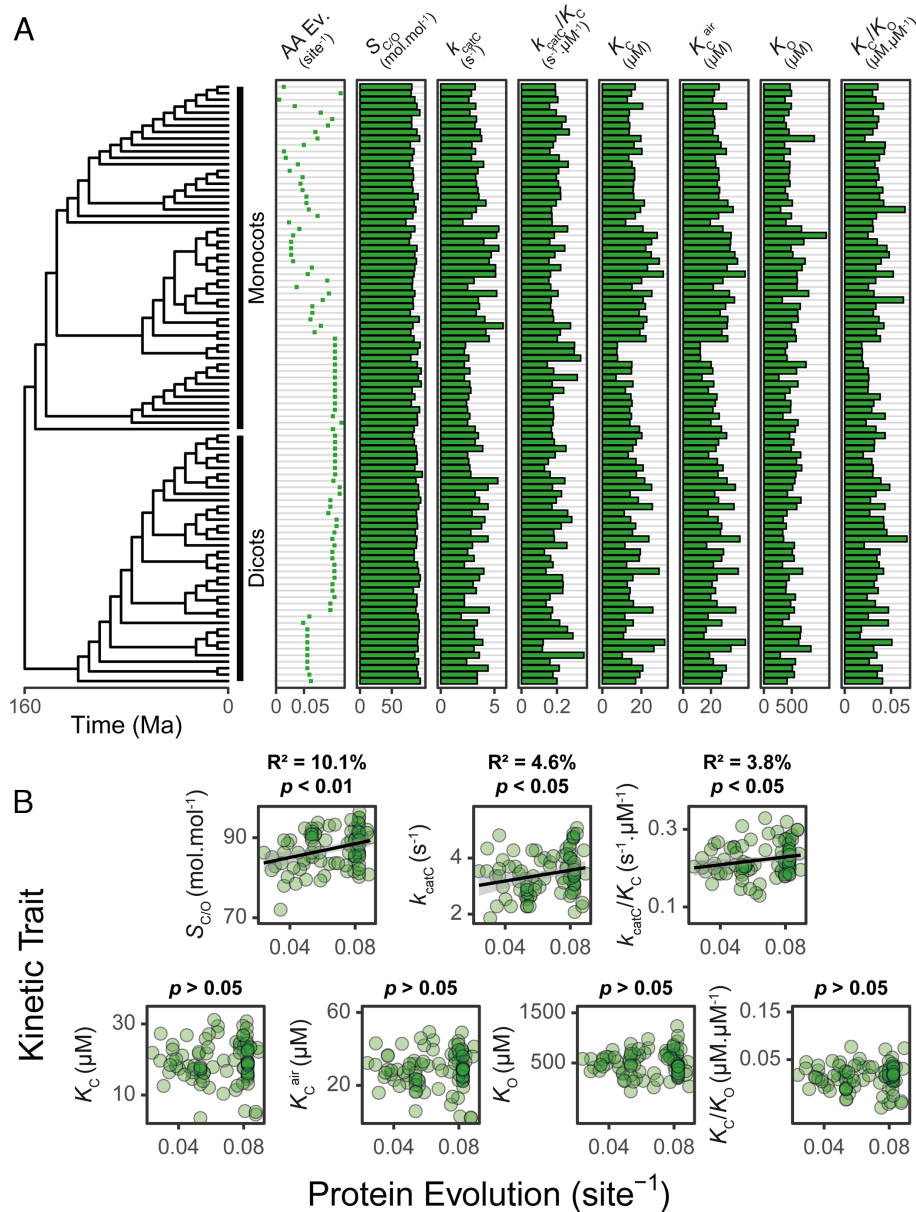


Fig. 4. The relationship between rubisco molecular and kinetic evolution in C_3 angiosperms. (A) The relationship between RbcL evolution and its corresponding kinetic trait values. AA Ev.: The extent of RbcL amino acid evolution that has occurred since the last common ancestor at the root of the angiosperm phylogeny. $S_{C/O}$: specificity. k_{catC} : carboxylase turnover per site. k_{catC}/K_C : carboxylation efficiency. K_C : the Michaelis constant for CO_2 . K_C^{air} : the inferred Michaelis constant for CO_2 in 20.95% O_2 . K_O : the Michaelis constant for O_2 . K_C/K_O : the ratio of the Michaelis constant for CO_2 compared to O_2 . (B) The relationship between the extent of RbcL protein evolution (substitutions per sequence site) and each rubisco kinetic trait in A as assessed using least squares regression models. The raw data can be found in *SI Appendix, Supplemental File 7*.

inefficient catalyst with a modest carboxylase turnover rate of $<12\text{ s}^{-1}$ (62) and a competing oxygenase activity that results in the loss of fixed carbon (65, 66, 87). This discord presents an evolutionary paradox that has attracted significant attention (16, 22, 23, 65, 69, 76–80). Here we demonstrate *rbcL* is one of the slowest-evolving genes on Earth. Despite this, we show that

rubisco has been evolving for improved CO_2/O_2 specificity ($S_{C/O}$), carboxylase turnover rates (k_{catC}), and carboxylation efficiencies (k_{catC}/K_C) in angiosperms. Moreover, we find that plants with more evolved rubisco exhibit higher leaf-level CO_2 assimilation and enhanced photosynthetic nitrogen-use efficiencies. Thus, rubisco has been continually evolving toward improved

Table 1. Rubisco kinetics in extinct and extant angiosperms

Rubisco	$S_{C/O}$ (mol mol ⁻¹)	k_{catC} (s ⁻¹)	k_{catC}/K_C (s ⁻¹ μM ⁻¹)	K_C (μM)	K_C^{air} (μM)	K_O (μM)	K_C/K_O (μM μM ⁻¹)
Last common angiosperm ancestor	81.1 ± 1.9	2.6 ± 0.3	0.16 ± 0.02	16.3 ± 2.1	24.8 ± 2.8	484.1 ± 56.4	0.034 ± 0.004
Extant angiosperms	87.1 ± 0.5	3.4 ± 0.1	0.20 ± 0.01	17.6 ± 0.5	26.4 ± 0.7	517.2 ± 14.7	0.035 ± 0.001

Kinetic trait values for the last common ancestor of the angiosperms were computed based on the estimated y intercept (mean ± 1 SE) of the linear regression analysis performed between the extent of RbcL protein evolution and each rubisco kinetic trait in Fig. 4B. Mean values of rubisco kinetic traits and associated variation (± 1 SE) in extant C_3 species are shown for comparison. The raw dataset used can be found in *SI Appendix, Supplemental File 7*.

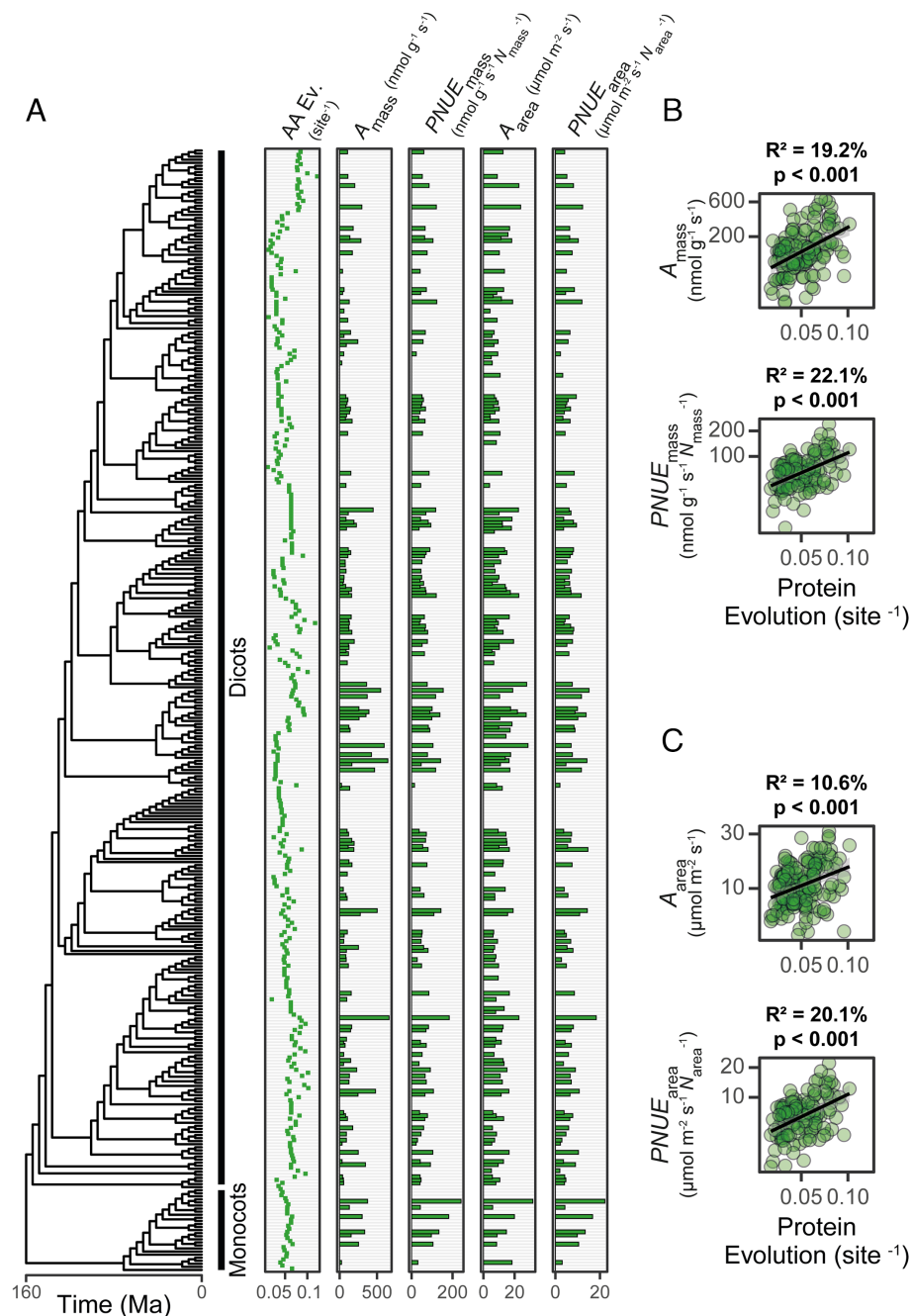


Fig. 5. The relationship between rubisco molecular evolution and CO₂ assimilation in C₃ angiosperms. (A) The relationship between the extent of RbL evolution and leaf-level CO₂ assimilation. AA Ev.: The extent of RbL amino acid evolution that has occurred since the most recent common ancestor at the root of the angiosperm phylogeny. A_{mass} : Photosynthetic rate per unit leaf mass. $PNUE_{\text{mass}}$: Photosynthetic nitrogen use efficiency rate, calculated as photosynthetic rate per unit leaf mass expressed per unit leaf mass nitrogen content (N_{mass} ; % N). A_{area} : Photosynthetic rate per unit leaf area. $PNUE_{\text{area}}$: Photosynthetic nitrogen use efficiency rate, calculated as photosynthetic rate per unit leaf area expressed per unit leaf area nitrogen content (N_{area} ; g m⁻² N). (B) The relationship between the extent of RbL protein evolution (substitutions per sequence site) and each photosynthetic trait in A evaluated on a mass-basis (A_{mass} , $PNUE_{\text{mass}}$) as assessed using least squares regression models. (C) As in B but for each photosynthetic trait evaluated on an area-basis (A_{area} , $PNUE_{\text{area}}$). The raw data can be found in [SI Appendix, Supplemental File 9](#).

catalytic efficiency and CO₂ assimilation during the radiation of the angiosperms.

A slow rate of molecular evolution in *rbcL* has long been assumed and has underpinned the use of this gene for systematics and phylogenetics (88–90). However, to our knowledge, there has been no contextualized measurement of the rate of *rbcL* evolution across the tree of life. The analysis presented here addresses this gap by revealing that *rbcL*/RbL has experienced a lower extent of molecular evolution than 99% of all gene nucleotide sequences and 98% of all gene protein sequences across all taxa in which it

is encoded. It is interesting to note that this result is not due to the presence of *rbcL* in the chloroplast genome, as *rbcL* is also one of the slowest-evolving sequences in bacteria which lack organellar genomes. Thus, RbL is universally one of the slowest-evolving sequences on Earth, irrespective of the taxon or genome in which it resides.

Although dissecting the factors which constrain the rate of *rbcL* evolution is beyond the scope of the current study, the slow pace of *rbcL* molecular evolution is most likely a consequence of several synergistic factors (91) including constraints imposed by expression

(92–96), selection to preserve protein function (97–101), and the requirements for protein–protein interactions in vivo (102–105). These factors would be particularly pertinent for rubisco given that it is the most abundant enzyme in organisms in which it is found (70, 72), it is subject to catalytic trade-offs (22, 23, 82, 83) and molecular activity–stability trade-offs (106–109), and given that it relies on multiple interacting partners and chaperones for folding, assembly and metabolic regulation (70, 110). Thus, a perfect storm of features exist which could limit the molecular evolution of *rbcl* and thereby cause it to be one of the slowest-evolving genes on Earth. Further work to elucidate the exact contribution of each of these biological determinants on rubisco’s rate of molecular evolution is warranted, building upon the work here and previous investigations (22, 23, 111).

Our integrated analysis of rubisco evolution revealed a continual improvement in $S_{C/O}$, k_{catC} , and k_{catC}/K_C during the radiation of C_3 angiosperms. Thus, although rubisco is slowly-evolving, sequence changes have enhanced the catalytic properties of the enzyme. In the context of the C_3 leaf, such directional selection toward improved $S_{C/O}$ is consistent with adaptation to maintain adequate carbon assimilation in response to declining atmospheric CO_2 and increasing atmospheric O_2 (Fig. 1). This evolutionary strategy has been proposed previously (82), and is suggested to apply broadly across photoautotrophs lacking a CO_2 concentrating mechanism (84, 112). In addition to adaptation for higher $S_{C/O}$, we also identify simultaneous improvement in k_{catC} and k_{catC}/K_C without antagonism in any other kinetic trait. These results are also consistent with the inferior $S_{C/O}$ and k_{catC}/K_C reported for extinct rubisco resurrected at the dawn of the Form IB (6) and Form I (36) lineages. It is noteworthy that on first appearances, all of these studies seem to contradict an analysis within the *Solanaceae* in which resurrected ancestral rubisco variants exhibited superior k_{catC} and k_{catC}/K_C values. However, in this instance the kinetic differences were proposed to be driven by sequence changes in RbcS (113), and therefore do not contradict the analysis of RbcL presented here or in other studies (6, 36). Thus, sequence change in RbcL during the radiation of angiosperms has driven the continual improvement of the enzyme in the presence of a declining atmospheric $CO_2:O_2$ concentration.

The “FvCB model” of photosynthesis (114), as well as a suite of other experimental (115–122) and computational (123, 124) studies all demonstrate that rubisco is a major rate-limiting factor for CO_2 assimilation under ambient steady-state conditions. The findings presented here link these mechanistic studies with evolutionary biology, and reveal that rubisco has experienced directional selection to improve kinetic efficiency and CO_2 assimilation. It is worth noting that the strength of the relationship observed between rubisco evolution and leaf-level CO_2 assimilation is greater in magnitude than that between evolution and enzyme kinetics. This phenomenon can be explained by the fact that rubisco is not only evolving for improved kinetic efficiency, but is also simultaneously undergoing optimization for multiple other properties such as stability, activatability, and chaperone affinity. This multi-objective optimization means that not all sequence changes are expected to alter kinetics. Instead, these sequence changes influence a much larger set of properties that together contribute to improvement in leaf-level CO_2 assimilation. Thus, our finding that rubisco evolution has a stronger correlation with leaf-level CO_2 assimilation than with any individual kinetic trait aligns with evolutionary expectations.

Ultimately, the global trends of rubisco evolution presented here change our understanding of the rubisco paradox. Rubisco is not locked in evolutionary stasis, but is instead slowly evolving toward improved CO_2 assimilation. These findings have significant

implications for our understanding of the past, present, and future potential of rubisco in natural and engineered contexts. Crucially, they also serve as a foundation for future work to further unravel the intricacies of rubisco evolution. For instance, an important focus of subsequent research would be to elucidate the adaptive trajectory of rubisco across different temporal, spatial, and phylogenetic scales. The present findings also motivate efforts to resolve the contribution of the rubisco small subunit and accessory chaperones to rubisco adaptation, given that these genes evolve faster than the large subunit and thus may have played an important role in the evolutionary dynamics of the enzyme over geological time scales.

Materials and Methods

Rubisco Sequence Data. All publicly available coding sequences of the *rbcl* and *rbcs* subunit genes in the NCBI database (<https://www.ncbi.nlm.nih.gov/>) as of July 2020 were downloaded (RbcL $n = 239,492$; RbcS $n = 2,061$). Manual inspection of nucleotide and translated protein sequences was performed to remove any duplicate, partial, chimeric, or erroneously annotated sequences. In addition, this dataset was further restricted to include only those species which possess a Form I rubisco and for which both a full-length *rbcl* and *rbcs* gene sequence could be obtained. Given that *rbcl* exists as a single copy gene in all species, only one *rbcl* sequence per species was retained for downstream analysis. In contrast, all possible full-length *rbcs* sequences were taken forward to account for the fact that *rbcs* is multicopy in some genomes.

Translated RbcL and RbcS protein sequences were aligned using the MAFFT-INS-i algorithm (125). The corresponding codon alignments of the nucleotide sequences were generated by threading the nucleotide sequences through the aligned protein sequences that they encode using PAL2NAL software (126). Multiple sequence alignments were trimmed to remove non-aligned codon or residue positions such that only ungapped columns remained. During this process, the putative transit peptide of *rbcs*/RbcS sequences in taxa in which this gene is encoded by the nuclear genome was computationally cleaved. Following these data processing steps, alignments were partitioned depending on species membership to either the bacteria (*Bacteria*; $n = 78$), land plants (*Streptophyta*; $n = 68$), green algae (*Chlorophyta*; $n = 12$), red algae (*Rhodophyta*; $n = 201$) or the SAR supergroup (*Stramenopiles*, *Alveolates*, and *Rhizaria*; $n = 129$) by use of the NCBI taxonomy browser (https://www.ncbi.nlm.nih.gov/Taxonomy/TaxIdentifier/tax_ididentifier.cgi). Any sequences belonging to species in either the *Haptophyta*, *Cryptophyta*, *Glaucocestophyta* or *Excavata* taxonomic groups were excluded from the dataset at this point due to insufficient data availability. In total, this resulted in a combined set of 488 *rbcl*/RbcL and 1,140 *rbcs*/RbcS gene and protein sequences across 488 species spanning 5 taxonomic groups (*SI Appendix, Fig. S1, Supplemental File 1, and Table S1*). The complete set of raw *rbcl*/RbcL and *rbcs*/RbcS sequences, as well as the complete set of aligned and trimmed *rbcl*/RbcL and *rbcs*/RbcS sequences can be found in *SI Appendix, Supplemental File 2*.

Rubisco Phylogenetic Tree inference. Maximum-likelihood *rbcl*/RbcL and *rbcs*/RbcS phylogenetic gene trees were inferred across all sequences within a taxonomic group by IQ-TREE (127) using the ultrafast bootstrapping method with 1,000 replicates and the Shimodaira–Hasegawa approximate-likelihood ratio branch test. The best fitting models of nucleotide (SYM+R8) and amino acid (LG+R5) sequence evolution were respectively determined as those which exhibit the lowest combined Bayesian information criterion rank score across the complete sets of both RbcL and RbcS sequences (127). Across all taxonomic groups, the models of nucleotide and amino acid sequence evolution were held constant between the gene trees for *rbcl* and *rbcs*, and RbcL and RbcS, respectively, such that branch lengths are comparable across both subunits. The complete set of these *rbcl*/RbcL and *rbcs*/RbcS phylogenetic gene trees used as the basis of the analysis herein can be found in *SI Appendix, Supplemental File 3*. To account for potential biases in our analysis caused by some species exhibiting multiple copies of *rbcs*, random stratified sampling of the non-gapped *rbcs*/RbcS sequence alignments was also conducted as described in *SI Appendix, Supplemental File 1*.

Quantification of the Total Extent of Nucleotide and Protein Molecular Evolution in *rbcl*/RbcL and *rbcs*/RbcS. The extent of molecular evolution in both rubisco subunits was assessed across all species in a given taxonomic group as the total length (sequence substitutions per aligned sequence site) of the

phylogenetic tree describing the evolutionary history of each respective gene. For this purpose, tree length was calculated as the combined sum of branch lengths leading from the root at the last common ancestor of the tree to the set of sequences at the terminal nodes. In this way, using the trees inferred across the complete cohort of *rbcl/Rbcl* and *rbcs/Rbcs* sequences in each taxonomic group, it was possible to capture all nucleotide and amino acid evolution which has arisen in each subunit since the most recent common ancestor of all sampled species in the group. The total tree length measure used here accounts for differences in alignment length between genes because the branch lengths in the tree are estimates of the number of substitutions per sequence site. This measure is also unaffected by species sampling or variation in model of sequence evolution as each of these factors was held constant in the analyses that were conducted. An identical analysis was also performed for each *rbcs/Rbcs* tree generated by stratified sampling, with mean and SD of estimates being calculated in this case across the 1,000 unique bootstrap replicate trees.

Orthogroup Classification and Phylogenetic Tree inference. Complete sets of representative gene models were acquired for a total of 32 of the bacteria species, 27 of the land plant species, 8 of the SAR species, 6 of the red algae species, and 4 of the green algae species analyzed in the present study, respectively (SI Appendix, Supplemental File 1 and Tables S1 and S7). A detailed description of the method used to download and prepare these gene models for downstream analysis is provided in SI Appendix, Supplemental File 1.

The complete set of translated proteomes for species in each respective taxonomic group were subject to orthogroup inference using OrthoFinder V2.5.2 (128, 129) software run with default settings and with the DIAMOND ultra-sensitive mode (130, 131). Protein sequences within each orthogroup were aligned using the MAFFT L-INS-I algorithm with 1,000 cycles of iterative refinement (125). The corresponding codon alignments of the nucleotide sequences were generated by threading the nucleotide sequences through the aligned protein sequences that they encode using PAL2NAL software (126). Alignments were trimmed to remove positions which contain gap characters. Sequences that were <50% of the median length of the cohort of all other sequences in the given orthogroup were excluded to avoid analysis of partial or truncated genes that could influence downstream analysis. All nucleotide and protein multiple sequence alignments which satisfied the above criterion and which possessed >50 ubiquitously aligned codon or amino acid positions were subject to bootstrapped maximum likelihood phylogenetic tree inference using IQ-TREE (127) following the exact method and evolutionary substitution models described above. In total, this resulted in a combined set of 16,631 orthogroup phylogenies comprising 5,126,017 ortholog pairwise comparisons across 351 species pairwise comparisons for the land plant clade, 6,953 orthogroup phylogenies comprising 153,288 ortholog pairwise comparisons across 28 species pairwise comparisons for the SAR clade, 5,422 orthogroup phylogenies comprising 642,057 ortholog pairwise comparisons across 496 species pairwise comparisons for the bacteria clade, 4,269 orthogroup phylogenies comprising 31,133 ortholog pairwise comparisons across 6 species pairwise comparisons for the green algae clade and 3,966 orthogroup phylogenies comprising 54,091 ortholog pairwise comparisons across 15 species pairwise comparisons for the red algae clade, from which to base the analyses herein. A further breakdown of these metrics for each species comparison can be found in SI Appendix, Supplemental File 4. The set of these respective genes within each species proteome that encode enzymes was determined following the method described in SI Appendix, Supplemental File 1.

Quantification of the Percentile Extent of Rubisco Molecular Evolution within Each Taxonomic Group. To evaluate the extent of molecular evolution in rubisco in the context of all other genes, only species in the rubisco sequence dataset possessing a publicly available whole-genome gene assembly were considered. Across each pairwise combination of species in a given taxonomic group which satisfied this criterion, the extent of *rbcl/Rbcl* and *rbcs/Rbcs* molecular evolution since the time point of species divergence was measured by computing the sum of branch lengths (sequence substitutions per aligned sequence site) separating these respective sequences in the rubisco phylogenetic trees previously inferred. Following this, the extent of molecular evolution separating all other pairs of orthologous (but not paralogous) gene and protein sequences for that given species pair was measured across all inferred orthogroup phylogenies, and the percentile rank rate of rubisco nucleotide or protein evolution was

computed relative to the cohort of these measurements. To assess the extent of rubisco molecular evolution in the context of all other enzymes, the exact same steps were followed but only the subset of genes and proteins predicted to encode enzymes were included. In both of the above analyses, a minimum threshold of 100 measurements for orthologous genes and protein sequences was ensured per species pair. In cases where multiple percentiles are calculated for a rubisco subunit in a given species pair (due to gene duplications in the *rbcs* of some species, or due to a single species gene assembly matching multiple sub-species in the rubisco sequence dataset) the mean percentile was taken. To provide further context on the relative pace of rubisco molecular evolution, the difference in evolutionary rate between rubisco and all other photosynthetic isoforms of the Calvin-Benson-Bassham cycle enzymes (SI Appendix, Supplemental File 1 and Table S8) was also determined across the land plants (the taxonomic group where this data is available), as described in SI Appendix, Supplemental File 1. In addition, quantification of the percentile extent of molecular evolution in rubisco chaperones (SI Appendix, Supplemental File 1 and Table S9) was performed following the same method and the same dataset for rubisco above, as described in SI Appendix, Supplemental File 1.

The full set of raw data generated from these analyses measuring the molecular evolution in rubisco and across all other genes and proteins, and all other enzyme-encoding genes and proteins in each taxa have been deposited in figshare (<https://doi.org/10.6084/m9.figshare.24994625>) (132). Individual processed datasets quantifying the relative extents of molecular evolution in rubisco and rubisco chaperones for each unique pairwise species comparison can be found in SI Appendix, Supplemental File 5. This processed data has also been included in a combined dataset format in SI Appendix, Supplemental File 6.

Integrated Analysis of Rubisco Molecular and Kinetic Evolution. To interrogate the relationship between the molecular and kinetic evolution of extant Form I rubisco, a dataset of rubisco kinetic traits was downloaded from Bouvier and colleagues (22, 23), as modified from that originally compiled by Flamholz and colleagues (83). For the purpose of this study, only species in this dataset with a complete set of experimentally determined measurements of rubisco specificity (S_{CO_2}) for CO_2 relative to O_2 (i.e., the overall carboxylation/oxygenation ratio of rubisco under defined concentrations of CO_2 and O_2 gases), maximum carboxylase turnover rate per active site (k_{catC}), and the respective Michaelis constant (i.e., the substrate concentration at half-saturated catalyzed rate) for both CO_2 (K_C) and O_2 (K_O) substrates were selected. For each of the 137 species which satisfied this criterion (all of which were angiosperm land plants), an estimate of the Michaelis constant for CO_2 in 20.95% O_2 air (K_{air}) was also available (22, 23). In addition, the ratio of the Michaelis constant for CO_2 relative to O_2 (K_C/K_O), as well as carboxylation efficiency defined as the ratio of the maximum carboxylase turnover to the Michaelis constant for CO_2 (k_{catC}/K_C), were inferred. Measurements of the Michaelis constant for RuBP (K_{RuBP}) were not considered owing to a limited sample size ($n = 19$). All *Limonium* species in the dataset were also ignored on the basis that trait values obtained across different studies have been deemed to be inconsistent (33, 133). In total, this left a dataset of rubisco kinetic trait measurements for 123 angiosperms. Of these, only the subset of 93 species which perform C_3 photosynthesis were considered for the purpose of the integrated molecular and kinetic evolution analysis herein. This is because of both a limited sample size of C_3 - C_4 species ($n = 6$), C_4 -like species ($n = 3$) and C_4 species ($n = 21$) in the kinetic dataset, and given that transition toward C_4 photosynthesis is associated with a change in rubisco kinetic evolution (22, 23) that would confound the directional selection analysis being conducted.

Coding sequences of the *rbcl* gene were obtained from Bouvier and colleagues (23) for each species in the kinetic dataset. In order to facilitate more accurate downstream phylogenetic tree inference across these sequences and to minimize the impact of long-branch effects (134), the complete set of publicly available *rbcl* coding sequences in land plants were also acquired in parallel from NCBI (<https://www.ncbi.nlm.nih.gov/>) using the query term "*rbcl*[Gene Name] AND *plants*[porgn:_txid3193]". These sequences thus obtained were subject to the exact same data processing steps to remove ambiguous, partial or chimeric sequences as performed previously for the *rbcl* sequences of species in the rubisco kinetic dataset (23). In total, this step resulted in an additional set of 29,218 full-length *rbcl* coding sequences to aid downstream phylogenetic inference. Protein sequences were inferred from each *rbcl* coding sequences via in silico translation. Next, the complete set of translated Rbcl sequences (including

the set of sequences from angiosperms in the rubisco kinetic dataset, as well as the set of all publicly available sequences for land plants) were respectively aligned using MAFFT L-INS-I (125), and a corresponding *rbcl* coding sequence alignment was generated using PAL2NAL software (126). The resulting multiple sequence alignments were trimmed to remove non-aligned residue positions and bootstrapped phylogenetic trees were inferred using IQ-TREE (127) following the exact method described above and using the best-fit models of sequence evolution previously inferred. To facilitate downstream analysis, the *rbcl* and RbCl gene trees were subsequently modified to keep only internal and terminal branches leading to the set of species in the rubisco kinetic dataset, with pruned trees manually rooted in Dendroscope (135).

To compute the relative extent of protein evolution which has occurred in each angiosperm in the kinetic dataset, the summed branch length (sequence substitutions per aligned sequence site) leading from the last common ancestor at the root of this clade to each respective terminal node in the RbCl phylogeny was measured. The kinetic trait values and extent of molecular evolution for all C_3 angiosperm rubisco can be found in *SI Appendix, Supplemental File 7*. The predicted kinetic trait values at the last common ancestor at the base of the angiosperm clade were inferred from the estimated y-intercept values from these regression models, as found in Table 1. The *rbcl*/RbCl phylogenetic gene trees used as the basis of this analysis, including the trees inferred across the full set of sequences, as well as the pruned versions of these trees containing only the subset of C_3 species in the kinetic dataset, can be found in *SI Appendix, Supplemental File 8*.

To confirm that the above analysis of rubisco molecular and kinetic evolution was not an artifact of potential methodological biases, this analysis was repeated using both a minimal subset of phylogenetically diverse species (*SI Appendix, Supplemental File 1 and Table S10*) so as to control for incomplete species sampling, as well as using alternate RbCl phylogenetic gene trees so as to control for any phylogenetic uncertainties. These supporting additional analyses confirmed that the results of the original analysis were valid and robust to potential methodological biases. A full description of the design, method, and results of these additional analyses is provided in detail in *SI Appendix, Supplemental File 1*.

Integrated Analysis of Rubisco Molecular Evolution and CO_2 Assimilation.

To investigate the relationship between rubisco molecular evolution and whole-plant photosynthetic performance, a comprehensive meta-dataset of photosynthetic measurements from species spanning the whole land plant phylogeny was provided by Gago and colleagues (86). This dataset contained measurements of light-saturated net photosynthetic rates expressed both per unit leaf mass (A_{mass}) and per unit leaf area (A_{area}), as well as measurements of total nitrogen content expressed both per unit leaf mass (N_{mass}) and per unit leaf area (N_{area}). In addition, for each unique species observation in this dataset with a corresponding measurement for both A_{mass} and N_{mass} or for both A_{area} and N_{area} , the mass-based and area-based photosynthetic nitrogen-use efficiencies were also derived using the calculations A_{mass}/N_{mass} ($PNUE_{mass}$) and A_{area}/N_{area} ($PNUE_{area}$), respectively. In cases where duplicate entries for a parameter were present across species, the mean value was taken so as to collapse the dataset to contain only a single row per species. Finally, although photosynthetic measurements were available from individuals belonging to all major land plant lineages (including the mosses, liverworts, fern allies, ferns, gymnosperms, and angiosperms), only

the subset of angiosperms for which a publicly available *rbcl* sequence could be obtained were taken forward. This is because various diffusional and biochemical factors other than rubisco are known to cause reduced photosynthetic capacities in non-angiosperm plants (86) that would bias the results of the current study. For the same reasons, only the subset of C_3 angiosperms in this dataset were taken forward to avoid picking up photosynthetic effects which result from CO_2 concentrating mechanisms that act upstream of rubisco. In total, this left a photosynthetic dataset of 366 C_3 angiosperms from which to base the analyses herein. This dataset included 272 unique species measurements for N_{mass} , 137 unique species measurements for A_{mass} and 118 unique species measurements for $PNUE_{mass}$, as well as 270 unique species measurements for N_{area} , 151 unique species measurements for A_{area} and 120 unique species measurements for $PNUE_{area}$, respectively.

To compute the relative extent of RbCl molecular evolution which has occurred in each angiosperm in the photosynthetic dataset, the exact same method was followed as described above. First, the full RbCl phylogenetic gene tree in *SI Appendix, Supplemental File 8* that was previously inferred from the complete set of publicly available RbCl sequences in NCBI was pruned so as to contain only terminal and internal branches corresponding to angiosperms in the photosynthetic dataset. Here, in situations where duplicate sequences in the alignment resulted in multiple terminal nodes for a given species, only a single node was retained based on the sequence which is first in the alphabetical order of the gene accession numbers. As above, this reduced RbCl tree was then manually rooted in Dendroscope (135), and the relative extent of RbCl protein evolution in each angiosperm was computed as the summed branch length (sequence substitutions per aligned sequence site) leading from the last common ancestor at the root of this clade to each respective terminal node. Finally, linear regression models were employed to assess the pairwise relationships between the variation in rubisco molecular evolution and each respective photosynthetic parameter. The resulting full integrated dataset containing photosynthetic measurements and comparable extents of RbCl molecular evolution for all 366 C_3 angiosperms can be found in *SI Appendix, Supplemental File 9*. The RbCl phylogenetic gene tree which has been pruned from that in *SI Appendix, Supplemental File 8* to contain the subset of C_3 angiosperms in the photosynthetic dataset used for the basis of this analysis can be found in *SI Appendix, Supplemental File 10*.

Data, Materials, and Software Availability. The raw data generated from the analyses measuring the molecular evolution in rubisco and across all other genes and proteins (including across all other enzyme-encoding genes and proteins) has been deposited in figshare and is available at: <https://doi.org/10.6084/m9.figshare.24994625> (132). All other data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. This work was funded by the Royal Society and the European Union's Horizon 2020 research and innovation program under grant agreement number 637765. J.W.B. was funded by the BBSRC through BB/J014427/1. This research was funded in whole, or in part, by the BBSRC number BB/J014427/1.

1. D. E. Canfield, The early history of atmospheric oxygen: Homage to Robert M. Garrels. *Annu. Rev. Earth Planet. Sci.* **33**, 1–36 (2005).
2. J. F. Kasting, Earth's early atmosphere. *Science* **259**, 920–926 (1993).
3. A. J. Kaufman *et al.*, Late archean biospheric oxygenation and atmospheric evolution. *Science* **317**, 1900–1903 (2007).
4. E. G. Nisbet *et al.*, The age of Rubisco: The evolution of oxygenic photosynthesis. *Geobiology* **5**, 311–335 (2007).
5. E. G. Nisbet, R. E. R. Nisbet, Methane, oxygen, photosynthesis, rubisco and the regulation of the air through time. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 2745–2754 (2008).
6. P. M. Shih *et al.*, Biochemical characterization of predicted Precambrian RuBisCO. *Nat. Commun.* **7**, 10382 (2016).
7. I. Slesak, H. Slesak, J. Kruk, RuBisCO early oxygenase activity: A kinetic and evolutionary perspective. *BioEssays* **39**, Article 1700071 (2017).
8. H. D. Holland, Volcanic gases, black smokers, and the great oxidation event. *Geochim. Cosmochim. Acta* **66**, 3811–3826 (2002).
9. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).
10. H. Ashida *et al.*, RuBisCO-like proteins as the enolase enzyme in the methionine salvage pathway: Functional and evolutionary relationships between RuBisCO-like proteins and photosynthetic RuBisCO. *J. Exp. Bot.* **59**, 1543–1554 (2008).
11. H. Ashida, A. Danchin, A. Yokota, Was photosynthetic RuBisCO recruited by acquisitive evolution from RuBisCO-like proteins involved in sulfur metabolism? *Res. Microbiol.* **156**, 611–618 (2005).
12. T. J. Erb, J. Zarzycki, A short history of RubisCO: The rise and fall (?) of Nature's predominant CO_2 fixing enzyme. *Curr. Opin. Biotechnol.* **49**, 100–107 (2018), 10.1016/j.copbio.2017.07.017.
13. D. M. Banda *et al.*, Novel bacterial clade reveals origin of form I Rubisco. *Nat. Plants* **6**, 1158–1166 (2020).
14. F. R. Tabita, S. Satagopan, T. E. Hanson, N. E. Kreeel, S. S. Scott, Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* **59**, 1515–1524 (2008).
15. I. Andersson, A. Backlund, Structure and function of Rubisco. *Plant Physiol. Biochem.* **46**, 275–291 (2008).
16. S. M. Whitney, R. L. Houtz, H. Alonso, Advancing our understanding and capacity to engineer nature's CO_2 -sequestering enzyme, Rubisco. *Plant Physiol.* **155**, 27–35 (2011).
17. F. R. Tabita, T. E. Hanson, S. Satagopan, B. H. Witte, N. E. Kreeel, Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 2629–2640 (2008).
18. G. Schneider, Y. Lindqvist, C. I. Brändén, RUBISCO: Structure and mechanism. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 119–143 (1992), 10.1146/annurev.bb.21.060192.001003.
19. C. B. Field, M. J. Behrenfeld, J. T. Randerson, P. Falkowski, Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**, 237–240 (1998).

20. B. Lee, F. R. Tabita, Purification of recombinant Ribulose-1,5-Bisphosphate carboxylase/oxygenase large subunits suitable for reconstitution and assembly of active L8S8 enzyme. *Biochemistry* **29**, 9352–9357 (1990).
21. T. J. Andrews, Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits. *J. Biol. Chem.* **263**, 12213–12219 (1988).
22. J. W. Bouvier, S. Kelly, Response to Tcherkez and Farquhar: Rubisco adaptation is more limited by phylogenetic constraint than by catalytic trade-off. *J. Plant Physiol.* **287**, 154021 (2023).
23. J. W. Bouvier *et al.*, Rubisco adaptation is more limited by phylogenetic constraint than by catalytic trade-off. *Mol. Biol. Evol.* **38**, 2880–2896 (2021).
24. A. Chatterjee, B. Anindya, Rubisco: Limitations and re-engineering for a better enzyme. *Int. Res. J. Plant Sci.* **2**, 022–024 (2011).
25. P. A. Christin *et al.*, Evolutionary switch and genetic convergence on rbcL following the evolution of C4 photosynthesis. *Mol. Biol. Evol.* **25**, 2361–2368 (2008).
26. M. V. Kapralov, D. S. Kubien, I. Andersson, D. A. Filatov, Changes in Rubisco kinetics during the evolution of C4 Photosynthesis in Flaveria (Asteraceae) are associated with positive selection on genes encoding the enzyme. *Mol. Biol. Evol.* **28**, 1491–1503 (2011).
27. M. A. J. Parry, P. J. Andralojc, R. A. C. Mitchell, P. J. Madgwick, A. J. Keys, Manipulation of Rubisco: The amount, activity, function and regulation. *J. Exp. Bot.* **54**, 1321–1333 (2003).
28. S. M. Whitney *et al.*, 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 (2011).
29. Y. Zhou, L. H. Gunn, R. Birch, I. Andersson, S. M. Whitney, Grafting Rhodospirillum rubrum with red algae Rubisco to accelerate catalysis and plant growth. *Nat. Plants* **9**, 978–986 (2023).
30. M. V. Kapralov, D. A. Filatov, Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC Evol. Biol.* **7**, 73 (2007).
31. M. T. Lin *et al.*, A procedure to introduce point mutations into the Rubisco large subunit gene in wild-type plants. *Plant J.* **106**, 876–887 (2011).
32. J. Galmés *et al.*, Expanding knowledge of the Rubisco kinetics variability in plant species: Environmental and evolutionary trends. *Plant Cell Environ.* **37**, 1989–2001 (2014).
33. J. Galmés *et al.*, Environmentally driven evolution of Rubisco and improved photosynthesis and growth within the C3 genus Limonium (Plumbaginaceae). *New Phytol.* **203**, 989–999 (2014).
34. R. J., Spreitzer Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **414**, 141–149 (2003).
35. D. J. Orr, M. A. J. Parry, Overcoming the limitations of Rubisco: Fantasy or realistic prospect? *J. Plant Physiol.* **254**, 153285 (2020).
36. L. Schulz *et al.*, Evolution of increased complexity and specificity at the dawn of form I Rubiscos. *Science* **378**, 155–160 (2022).
37. J. H. Fitch, S. Knight, I. Andersson, C. I. Branden, L. McIntosh, Residues in three conserved regions of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase are required for quaternary structure. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5768–5772 (1990).
38. T. Genkov, R. J. Spreitzer, Highly conserved small subunit residues influence Rubisco large subunit catalysis. *J. Biol. Chem.* **284**, 30105–30112 (2009).
39. M. Grabsztunowicz, Z. Górski, R. Luciński, G. Jackowski, A reversible decrease in ribulose 1,5-bisphosphate carboxylase/oxygenase carboxylation activity caused by the aggregation of the enzyme's large subunit is triggered in response to the exposure of moderate irradiance-grown plants to low irradiance. *Physiol. Plant.* **154**, 591–608 (2015).
40. J. Joshi, O. Mueller-Cajar, Y. C. Tsai, F. U. Hartl, M. Hayer-Hartl, Role of small subunit in mediating assembly of red-type Form I Rubisco. *J. Biol. Chem.* **290**, 10666–1074 (2015).
41. B. Lee, R. M. Berka, F. R. Tabita, Mutations in the small subunit of cyanobacterial ribulose-bisphosphate carboxylase/oxygenase that modulate interactions with large subunits. *J. Biol. Chem.* **266**, 7417–7422 (1991).
42. B. A. McFadden, C. L. Small, Cloning, expression and directed mutagenesis of the genes for ribulose bisphosphate carboxylase/oxygenase. *Photosynth. Res.* **18**, 245–260 (1988).
43. G. Schneider *et al.*, Comparison of the crystal structures of L2 and L8S8 Rubisco suggests a functional role for the small subunit. *EMBO J.* **9**, 2045–2050 (1990).
44. M. Van Lun, J. S. Hub, D. Van Der Spoel, I. Andersson, CO2 and O2 distribution in rubisco suggests the small subunit functions as a CO2 reservoir. *J. Am. Chem. Soc.* **136**, 3165–3171 (2014).
45. R. V. Kostov, C. L. Small, B. A. McFadden, Mutations in a sequence near the N-terminus of the small subunit alter the CO2/O2 specificity factor for ribulose bisphosphate carboxylase/oxygenase. *Photosynth. Res.* **54**, 127–134 (1997).
46. G. Voordouw, P. A. De Vries, W. A. M. Van Den Berg, E. P. J. De Clerck, Site-directed mutagenesis of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Anacystis nidulans*. *Eur. J. Biochem.* **163**, 591–598 (1987).
47. B. A. Read, F. R. Tabita, Amino acid substitutions in the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase that influence catalytic activity of the holoenzyme. *Biochemistry* **31**, 519–525 (1992).
48. T. P. Getzoff, G. Zhu, H. J. Bohnert, R. G. Jensen, Chimeric Arabidopsis thaliana ribulose-1,5-bisphosphate carboxylase/oxygenase containing a pea small subunit protein is compromised in carboxylation. *Plant Physiol.* **116**, 695–702 (1998).
49. C. Ishikawa, T. Hatanaka, S. Misoo, C. Miyake, H. Fukayama, Functional incorporation of sorghum small subunit increases the catalytic turnover rate of rubisco in transgenic rice. *Plant Physiol.* **156**, 1603–1611 (2011).
50. I. Kanevski, P. Maliga, D. F. Rhoades, S. Gutteridge, Plastome engineering of ribulose-1,5-bisphosphate carboxylase/oxygenase in tobacco to form a sunflower large subunit and tobacco small subunit hybrid. *Plant Physiol.* **119**, 133–141 (1999).
51. H. Matsumura *et al.*, Hybrid rubisco with complete replacement of rice rubisco small subunits by sorghum counterparts confers C4 plant-like high catalytic activity. *Mol. Plant* **13**, 1570–1581 (2020).
52. K. Sakoda *et al.*, Effects of introduction of sorghum RbcS with rice RbcS knockdown by RNAi on photosynthetic activity and dry weight in rice. *Plant Prod. Sci.* **24**, 346–353 (2021).
53. T. J. Andrews, G. H. Lorimer, Catalytic properties of a hybrid between cyanobacterial large subunits and higher plant small subunits of ribulosebisphosphate carboxylase-oxygenase. *J. Biol. Chem.* **260**, 4632–4636 (1985).
54. K. M. Horken, F. R. Tabita, The 'green' form I ribulose 1,5-bisphosphate carboxylase/oxygenase from the nonsulfur purple bacterium *Rhodospirillum rubrum*. *J. Bacteriol.* **181**, 3935–3941 (1999).
55. B. Lee, B. A. Read, F. R. Tabita, Catalytic properties of recombinant octameric, hexadecameric, and heterologous cyanobacterial/bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **291**, 263–269 (1991).
56. S. M. van der Vies, D. Bradley, A. A. Gatenby, Assembly of cyanobacterial and higher plant ribulose bisphosphate carboxylase subunits into functional homologous and heterologous enzyme molecules in *Escherichia coli*. *EMBO J.* **5**, 2439–2444 (1986).
57. Y. L. Wang, J. H. Zhou, Y. F. Wang, J. S. Bao, H. B. Chen, Properties of hybrid enzymes between *Synechococcus* large subunits and higher plant small subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Escherichia coli*. *Arch. Biochem. Biophys.* **396**, 35–42 (2001).
58. B. A. Read, F. R. Tabita, A hybrid ribulosebisphosphate carboxylase/oxygenase enzyme exhibiting a substantial increase in substrate specificity factor. *Biochemistry* **31**, 5553–5560 (1992).
59. A. P. Cavanagh, R. Slattery, D. S. Kubien, Temperature-induced changes in Arabidopsis Rubisco activity and isoform expression. *J. Exp. Bot.* **74**, 651–663 (2022), 10.1093/jxb/erac379.
60. E. Martín-Avila *et al.*, Modifying plant photosynthesis and growth via simultaneous chloroplast transformation of rubisco large and small subunits. *Plant Cell* **32**, 2898–2916 (2020).
61. K. Morita, T. Hatanaka, S. Misoo, H. Fukayama, Unusual small subunit that is not expressed in photosynthetic cells alters the catalytic properties of Rubisco in rice. *Plant Physiol.* **164**, 69–79 (2014).
62. M. R. Badger *et al.*, The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO2-concentrating mechanisms in algae. *Can. J. Bot.* **76**, 1052–1071 (1998).
63. A. Bar-Even *et al.*, The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
64. R. Chollet, The biochemistry of photorespiration. *Trends Biochem. Sci.* **2**, 155–159 (1977).
65. T. D. Sharkey, Emerging research in plant photosynthesis. *Emerg. Top. Life Sci.* **4**, 137–150 (2020).
66. G. Bowes, W. L. Ogren, R. H. Hageman, Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase. *Biochem. Biophys. Res. Commun.* **45**, 716–722 (1971).
67. F. A. Busch, Photorespiration in the context of Rubisco biochemistry, CO2 diffusion and metabolism. *Plant J.* **101**, 919–939 (2020).
68. N. A. Eckardt, Photorespiration revisited. *Plant Cell* **17**, 2139–2141 (2005), 10.1105/tpc.105.03873.
69. R. E. Sharwood, Engineering chloroplasts to improve Rubisco catalysis: Prospects for translating improvements into food and fiber crops. *New Phytol.* **213**, 494–510 (2017), 10.1111/nph.14351.
70. E. Carmo-Silva, J. C. Scales, P. J. Madgwick, M. A. J. Parry, Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant Cell Environ.* **38**, 1817–1832 (2015), 10.1111/pce.12425.
71. L. S. Sarles, F. R. Tabita, Derepression of the synthesis of D-ribulose 1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum*. *J. Bacteriol.* **153**, 458–464 (1983).
72. F. R. Tabita, Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynth. Res.* **60**, 1–28 (1999).
73. A. Flammholz, P. M. Shih, Cell biology of photosynthesis over geologic time. *Curr. Biol.* **30**, R490–R494 (2020).
74. M. Meyer, H. Griffiths, Origins and diversity of eukaryotic CO2-concentrating mechanisms: Lessons for the future. *J. Exp. Bot.* **64**, 769–786 (2013), 10.1093/jxb/ers390.
75. U. Schlüter *et al.*, Brassicaceae display variation in efficiency of photorespiratory carbon recapturing mechanisms. *J. Exp. Bot.* **74**, 6631–6649 (2023), 10.1093/jxb/erad250.
76. W. L. Ogren, Photorespiration: Pathways, regulation, and modification. *Annu. Rev. Plant Physiol.* **35**, 415–442 (1984).
77. D. J. Orr *et al.*, Engineering photosynthesis: Progress and perspectives. *F1000Res.* **6**, 10.12688/f1000research.12181.1 (2017).
78. M. A. J. Parry *et al.*, Rubisco activity and regulation as targets for crop improvement. *J. Exp. Bot.* **64**, 717–730 (2013).
79. M. A. J. Parry, P. J. Madgwick, J. F. C. Carvalho, P. J. Andralojc, Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. *J. Agric. Sci.* **145**, 31–43 (2007).
80. R. E. Sharwood, O. Ghannoum, S. M. Whitney, Prospects for improving CO2 fixation in C3-crops through understanding C4-Rubisco biogenesis and catalytic diversity. *Curr. Opin. Plant Biol.* **31**, 135–142 (2016), 10.1016/j.cpb.2016.04.002.
81. Y. Savir, E. Noor, R. Milo, T. Tlusty, Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3475–3480 (2010).
82. G. B. Tcherkez, G. D. Farquhar, T. J. Andrews, Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7246–7251 (2006).
83. A. I. Flammholz *et al.*, Revisiting trade-offs between rubisco kinetic parameters. *Biochemistry* **58**, 3365–3376 (2019), 10.1021/acs.biochem.9b00237.
84. C. Iñiguez *et al.*, Evolutionary trends in RuBisCO kinetics and their co-evolution with CO2 concentrating mechanisms. *Plant J.* **101**, 897–918 (2020).
85. P. L. Cummins, The coevolution of RuBisCO, Photorespiration, and carbon concentrating mechanisms in higher plants. *Front. Plant Sci.* **12**, 662425 (2021).
86. J. Gago *et al.*, Photosynthesis optimized across land plant phylogeny. *Trends Plant Sci.* **24**, 947–958 (2019).
87. W. L. Ogren, G. Bowes, Ribulose diphosphate carboxylase regulates soybean photorespiration. *Nat. New Biol.* **230**, 159–160 (1971).
88. APG, An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Garden* **85**, 531–553 (1998).
89. APG, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **181**, 1–20 (2016).
90. L. Gielly, P. Taberlet, The use of chloroplast DNA to resolve plant phylogenies: Noncoding versus rbcL sequences. *Mol. Biol. Evol.* **11**, 769–777 (1994).
91. C. Pál, B. Papp, M. J. Lercher, An integrated view of protein evolution. *Nat. Rev. Genet.* **7**, 337–348 (2006).
92. E. H. J. Robbins, S. Kelly, The evolutionary constraints on angiosperm chloroplast adaptation. *Genome Biol. Evol.* **15**, evad101 (2023).
93. D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14338–14343 (2005).
94. J. F. Gout, D. Kahn, L. Duret, The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. *PLoS Genet.* **6**, 20 (2010).
95. E. A. Seward, S. Kelly, Selection-driven cost-efficiency optimization of transcripts modulates gene evolutionary rate in bacteria. *Genome Biol.* **19**, 102 (2018).

96. D. P. Wall *et al.*, Functional genomic analysis of the rates of protein evolution. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5483–5488 (2005).
97. J. D. Bloom, D. A. Drummond, F. H. Arnold, C. O. Wilke, Structural determinants of the rate of protein evolution in yeast. *Mol. Biol. Evol.* **23**, 1751–1761 (2006).
98. J. Echave, Beyond stability constraints: A biophysical model of enzyme evolution with selection on stability and activity. *Mol. Biol. Evol.* **36**, 613–620 (2019).
99. A. E. Lobkovsky, Y. I. Wolf, E. V. Koonin, Universal distribution of protein evolution rates as a consequence of protein folding physics. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2983–2988 (2010).
100. M. Soskine, D. S. Tawfik, Mutational effects and the evolution of new protein functions. *Nat. Rev. Genet.* **11**, 572–582 (2010).
101. N. J. Tourasse, W. H. Li, Selective constraints, amino acid composition, and the rate of protein evolution. *Mol. Biol. Evol.* **17**, 656–664 (2000).
102. H. B. Fraser, D. P. Wall, A. E. Hirsh, A simple dependence between protein evolution rate and the number of protein-protein interactions. *BMC Evol. Biol.* **3**, 1–6 (2003).
103. H. B. Fraser, A. E. Hirsh, Evolutionary rate depends on number of protein-protein interactions independently of gene expression level. *BMC Evol. Biol.* **4**, 1–5 (2004).
104. R. Saeed, C. M. Deane, Protein protein interactions, evolutionary rate, abundance and age. *BMC Bioinformatics* **7**, 128 (2006).
105. Á. Tóth-Petróczy, D. S. Tawfik, Slow protein evolutionary rates are dictated by surface-Core association. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11151–11156 (2011).
106. P. L. Cummins, B. Kannappan, J. E. Gready, Directions for optimization of photosynthetic carbon fixation: Rubisco's efficiency may not be so constrained after all. *Front. Plant Sci.* **9**, <https://doi.org/10.3389/fpls.2018.00183> (2018).
107. P. Duraõ *et al.*, Opposing effects of folding and assembly chaperones on evolvability of Rubisco. *Nat. Chem. Biol.* **11**, 148–155 (2015).
108. R. A. Studer, P. A. Christin, M. A. Williams, C. A. Orengo, Stability-activity tradeoffs constrain the adaptive evolution of Rubisco. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2223–2228 (2014).
109. P. L. Cummins, B. Kannappan, J. E. Gready, Response: Commentary: Directions for optimization of photosynthetic carbon fixation: RuBisCO's efficiency may not be so constrained after all. *Front. Plant Sci.* **10**, 1426 (2019).
110. H. Aigner *et al.*, Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science* **358**, 1272–1278 (2017).
111. S. M. Whitney, R. Birch, C. Kelso, J. L. Beck, M. V. Kapralov, Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3564–3569 (2015).
112. R. E. M. Rickaby, M. R. Eason Hubbard, Upper ocean oxygenation, evolution of RuBisCO and the Phanerozoic succession of phytoplankton. *Free Radic. Biol. Med.* **140**, 295–304 (2019).
113. M. T. Lin, H. Salihovic, F. K. Clark, M. R. Hanson, Improving the efficiency of Rubisco by resurrecting its ancestors in the family Solanaceae. *Sci. Adv.* **8**, eabm6871 (2022).
114. G. D. Farquhar, S. von Caemmerer, J. A. Berry, A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* **149**, 78–90 (1980).
115. J. A. Bassham, G. H. Krause, Free energy changes and metabolic regulation in steady-state photosynthetic carbon reduction. *Biochim. Biophys. Acta* **189**, 207–221 (1969).
116. K. J. Dietz, U. Heber, Rate-limiting factors in leaf photosynthesis. I. Carbon fluxes in the calvin cycle. *Biochim. Biophys. Acta Bioenerg.* **767**, 432–443 (1984).
117. K. Fichtner *et al.*, Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with 'antisense' rbcS - V. Relationship between photosynthetic rate, storage strategy, biomass allocation and vegetative plant growth at three different nitrogen s. *Planta* **190**, 1–9 (1993).
118. R. T. Furbank, J. A. Chitty, S. Von Caemmerer, C. L. D. Jenkins, Antisense RNA inhibition of RbcS gene expression reduces rubisco level and photosynthesis in the C₄ plant *Flaveria bidentis*. *Plant Physiol.* **111**, 725–734 (1996).
119. G. S. Hudson, J. R. Evans, S. Von Caemmerer, Y. B. C. Arvidsson, T. J. Andrews, Reduction of ribulose-1,5-bisphosphate carboxylase/oxygenase content by antisense RNA reduces photosynthesis in transgenic tobacco plants. *Plant Physiol.* **98**, 294–302 (1992).
120. C. Maheshwari *et al.*, Targeted knockdown of ribulose-1,5-bisphosphate carboxylase-oxygenase in rice mesophyll cells. *J. Plant Physiol.* **260**, 153395 (2021).
121. W. P. Quick *et al.*, Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with 'antisense' rbcS - I. Impact on photosynthesis in ambient growth conditions. *Planta* **183**, 542–554 (1991).
122. M. Stitt *et al.*, Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with 'antisense' rbcS-II. Flux-control coefficients for photosynthesis in varying light, CO₂, and air humidity. *Planta* **183**, 555–566 (1991).
123. M. G. Poolman, D. A. Fell, S. Thomas, Modelling photosynthesis and its control. *J. Exp. Bot.* **51**, 319–328 (2000).
124. X. G. Zhu, E. De Sturler, S. P. Long, Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: A numerical simulation using an evolutionary algorithm. *Plant Physiol.* **145**, 513–526 (2007).
125. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013), [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).
126. M. Suyama, D. Torrents, P. Bork, PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609 (2006).
127. L. T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015), [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300).
128. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 1–14 (2019).
129. D. M. Emms, S. Kelly, OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 1–14 (2015).
130. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2014).
131. B. Buchfink, K. Reuter, H. G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
132. J. Bouvier, Gene/Enzyme molecular evolution rates across each major taxonomic group of photosynthetic organisms. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.24994625.v1>. Deposited 13 January 2024.
133. J. Galmés *et al.*, Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves. *Plant Cell Environ.* **28**, 571–579 (2005).
134. Z. Su, J. P. Townsend, Utility of characters evolving at diverse rates of evolution to resolve quartet trees with unequal branch lengths: Analytical predictions of long-branch effects. *BMC Evol. Biol.* **15**, 86 (2015), [10.1186/s12862-015-0364-7](https://doi.org/10.1186/s12862-015-0364-7).
135. D. H. Huson, C. Scornavacca, Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061–1067 (2012), [10.1093/sysbio/sys062](https://doi.org/10.1093/sysbio/sys062).