RuBisCO adaptation is more limited by phylogenetic constraint than by catalytic trade-off

Jacques W. Bouvier¹,², David M. Emms¹, Timothy Rhodes³, Jochem R. Nielsen², Jai S. Bolton², Alice Eddershaw², Amelia Brasnett², Ana Unitt², Spencer M. Whitney², Steven Kelly¹

Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, United Kingdom

Doctoral Training Centre, University of Oxford, Keble Road, Oxford, OX1 3NP, United Kingdom

Research School of Biology, Australian National University, Canberra, Australian Capital Territory 2601, Australia

Name: Steven Kelly
Email: steven.kelly@plants.ox.ac.uk
Address: Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK

RuBisCO assimilates CO₂ to form the sugars that fuel life on earth. Correlations between RuBisCO kinetic traits across species have led to the proposition that RuBisCO adaptation is constrained by catalytic trade-offs. However, these correlations were founded on the assumption that kinetic measurements in different species are independent. Here, we show that this assumption is incorrect in angiosperms by evaluating the dependence of variation in RuBisCO kinetic traits on the phylogenetic tree that relates the enzymes. We show that there is significant phylogenetic signal in all carboxylase kinetic traits in angiosperms, and significant phylogenetic signal in the Michaelis constant for O₂ in species that conduct C₃ photosynthesis. When accounting for this non-independence, we show that the catalytic trade-off between carboxylase turnover and the Michaelis constant for CO₂ is weak (~30 % dependency) and that the correlations between all other RuBisCO kinetic traits are either not-significant or marginal (<9 % dependency). Finally, we demonstrate that phylogenetic constraints limit RuBisCO adaptation to a greater extent than catalytic trade-offs. Thus,
the biochemical landscape of RuBisCO adaptation in angiosperms is predominantly limited by phylogenetic constraint and a partial trade-off between carboxylase turnover and the Michaelis constant for CO₂.

**Introduction**

The vast majority of organic carbon on Earth entered the biosphere via the catalytic pocket of RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) (Beer et al., 2010). Whilst there are several metabolic contexts in which this RuBisCO-mediated reaction can occur, the most important of these in terms of global net primary production is photosynthesis (Andersson and Backlund, 2008). Here, RuBisCO plays a central role in the initial step of the Calvin-Benson-Bassham reductive pentose phosphate pathway, catalysing the fixation of CO₂ onto the acceptor molecule ribulose 1,5-bisphosphate (RuBP) to ultimately synthesise carbohydrates. In plants, this reaction is carried out by RuBisCO Form IB which is structurally composed of both large (RbcL) and small (RbcS) subunits in a hexadecameric configuration (Schneider, Lindqvist and Brändén, 1992; Tabita et al., 2008). Although this complex quaternary structure characteristic of all Form I (A-D) RuBisCO is the most common in nature (Andersson and Backlund, 2008), only the large subunit is essential for catalysis (Andrews, 1988; Lee and Tabita, 1990) as the active sites are located exclusively on this subunit (Whitney, Houtz and Alonso, 2011). In addition, disparate lineages of Form II, III and IV RuBisCO exist which catalyse CO₂ fixation and lack the small subunit entirely (Tabita et al., 2008). However, despite not being involved in the catalytic chemistry directly, the small subunit does have an indirect effect on the catalytic properties (Read and Tabita, 1992b, 1992a; Spreitzer, Peddi and Satagopan, 2005; Ishikawa et al., 2011; Joshi et al., 2015; Fukayama et al., 2019; Martin-Avila et al., 2020) and activity (Andrews, 1988; Lee and Tabita, 1990; Lee, Berka and Tabita, 1991) of RuBisCO, and thus is important to the function of Form I RuBisCO.

Perhaps unsurprisingly given its role as the entry point for carbon into the global food chain, RuBisCO is the most abundant protein on Earth (Ellis, 1979) with a global mass of ~0.7 gigatons (Bar-On and Milo, 2019). This extraordinary abundance is in part due to the inefficiency of RuBisCO as a catalyst. Specifically, RuBisCO has a low rate of CO₂ assimilation (Tcherkez, Farquhar and Andrews, 2006; Savir et al., 2010) and is poorly able to discriminate CO₂ and O₂ (Ogren and Bowes, 1971) causing it to catalyse both a carboxylation and an oxygenation reaction (Bowes, Ogren and Hageman, 1971;
The RuBisCO-mediated oxygenation of RuBP results in the production of 2-phosphoglycolate which must then be metabolised to recover the fixed carbon and avoid depletion of metabolite pools (Eckardt, 2005; Sharwood, 2017). This carbon scavenging process, known as photorespiration, consumes ATP and reducing power and also liberates ammonia which must then be re-assimilated (Peterhansel et al., 2010). Although the oxygenation reaction catalysed by RuBisCO is thought not to be deleterious in the anoxic environment prevalent when the enzyme first evolved (Nisbet et al., 2007; Erb and Zarzycki, 2018), under current atmospheric conditions it can comprise a quarter of all RuBisCO reactions in terrestrial plants (Ehleringer et al., 1991). Thus, despite oxygenation serving a number of beneficial functions (Busch, 2020), at its current rate it represents a substantial metabolic burden reducing the productivity of some plants by up to 50% (Ogren, 1984; Bauwe et al., 2012).

Given the high energetic cost incurred by the RuBisCO oxygenation reaction, a number of photoautotrophic organisms have evolved mechanisms to reduce the frequency of its occurrence. Collectively referred to as CO₂-concentrating mechanisms, these function to increase the concentration of CO₂ relative to O₂ in the vicinity of RuBisCO and thus increase the relative frequency of carboxylation reactions (Meyer and Griffiths, 2013). The simplest of these CO₂-concentrating mechanisms work by compartmentalising RuBisCO in a protein shell or sub-cellular micro-compartment within which inorganic carbon is concentrated using active transport (Raven, 1997; Giordano, Beardall and Raven, 2005). More complex CO₂-concentrating mechanisms, such as those found in plants that conduct C₄ photosynthesis or crassulacean acid metabolism, require anatomical and biochemical adaptations of cells and tissue types (Bräutigam et al., 2017; Edwards, 2019). These latter mechanisms utilise the oxygen insensitive enzyme phosphoenolpyruvate carboxylase to separate primary CO₂ assimilation from RuBisCO-mediated photosynthetic CO₂ reduction either spatially (C₄) (Sage, Sage and Kocacinar, 2012) or temporally (crassulacean acid metabolism) (Dodd et al., 2002). In doing so, both C₄ and crassulacean acid metabolism are able to largely circumvent the oxygenation reaction of RuBisCO resulting in highly successful adaptive phenotypes (Still et al., 2003; Lüttge, 2004).
The observation that natural selection has devised an array of CO₂-concentrating mechanisms rather than improve the CO₂ specificity of RuBisCO has led many to question whether altering the kinetics of the enzyme is possible (Ogren, 1984; Parry et al., 2007, 2013; Whitney, Houtz and Alonso, 2011; Sharwood, Ghannoum and Whitney, 2016; Araújo et al., 2017; Sharwood, 2017; Sharkey, 2020).

Here, the proposition that RuBisCO kinetics cannot be improved was supported by observations that the oxygenase and carboxylase activities of RuBisCO appear to be tightly linked (Badger and Lorimer, 1976; Chollet and Anderson, 1976). Subsequently, multiple studies have supported this suggestion by reporting strong antagonistic relationships between RuBisCO specificity (S_C/O), carboxylase turnover (k_{catC}) and the Michaelis constant (i.e., an inverse measure of substrate affinity for an enzyme) for CO₂ (K_C), as well as between K_C and the Michaelis constant for O₂ (K_O) (Tcherkez, Farquhar and Andrews, 2006; Savir et al., 2010). Collectively, these studies have led to the hypothesis that severe kinetic trait trade-offs hamstring the inherent efficiency by which the enzyme can catalyse CO₂ fixation, and that contemporary RuBisCO are near perfectly adapted within this heavily constrained catalytic landscape (Tcherkez, Farquhar and Andrews, 2006; Savir et al., 2010).

However, new evidence has started to overturn this paradigm of RuBisCO evolution. First, a recent analysis of the correlative nature of RuBisCO kinetics has demonstrated that associations between kinetic traits are weakened when a large number of species are considered (Flamholz et al., 2019). Further, long-standing engineering efforts to improve RuBisCO have begun to reveal signs of success, where the enzyme variants evolved by directed evolution show deviations from the proposed catalytic trade-offs between S_C/O, k_{catC} and K_C (Wilson et al., 2018; Zhou and Whitney, 2019). Thus together, these results indicate that RuBisCO kinetic traits are perhaps not as inextricably linked as originally thought, and suggest that there is scope for increasing the catalytic efficiency of the enzyme as has happened in nature for RuBisCO in some red algae (Andersson and Backlund, 2008).

Although the catalytic traits of RuBisCO appear to be correlated, there are flaws to inferring causality from this correlation. This is because previous analyses that have inferred correlations between kinetic traits have assumed that measurements of RuBisCO kinetic traits in different species are independent. However, this assumption has never formally been tested and is unlikely to be true because RuBisCO in all extant species are related to each other by descent from a single ancestral
gene. This means that the gene sequences encoding the enzyme in closely related species are more similar than the gene sequences from species drawn at random, a feature of RuBisCO which has been exploited in systematics and evolutionary analyses to serve as an accurate proxy for the phylogenetic relationship between species (Gielly and Taberlet, 1994; APG, 1998, 2016). By extension of this fact, it should be expected that closely related enzymes will generally also have similar kinetics, with the extent of this similarity being dependent on the underlying tree describing the relationship between species. This phenomenon, which is known as phylogenetic signal (also known as phylogenetic co-variation), can cause spurious correlations in measured trait values between species unless the structure of the phylogenetic tree is taken into consideration (Felsenstein, 1985; Grafen, 1989; Pagel and Harvey, 1989; Garland, 2001). Thus, as previous analyses of RuBisCO kinetics have not assessed whether phylogenetic signal exists in RuBisCO kinetic traits, nor accounted for any phylogenetic signal which may exist, it is possible that the observed catalytic trade-offs are, either wholly or in part, an artefact caused by phylogenetic signal. Here, we assess the phylogenetic signal and tree-like evolution of RuBisCO kinetic traits in a large dataset of angiosperms to re-evaluate the extent by which RuBisCO is constrained by phylogenetic effects and catalytic trade-offs. We demonstrate that there is a significant phylogenetic signal in RuBisCO specificity, carboxylase turnover and the Michaelis constant for CO$_2$ in angiosperms, and also a significant phylogenetic signal in the Michaelis constant for O$_2$ in species that conduct C$_3$ photosynthesis. This means that phylogenetic constraints cause the similarity of kinetic measurements between species to vary as a function of their evolutionary distance. When this phylogenetic dependency is correctly accounted for by using phylogenetic least squares regression, we reveal an attenuation in the interdependence between RuBisCO kinetic traits and show that previously inferred correlations have been over-estimated. Specifically, after phylogenetic correction we find evidence of a weak power-law correlation between carboxylase turnover and CO$_2$ affinity (~30 % dependency), and demonstrate that all other measured associations between specificity, carboxylase turnover and the Michaelis constant for both CO$_2$ and O$_2$ substrates are either non-significant or marginal (<9 % dependency). Moreover, we find that phylogenetic constraints, most likely resulting from low genetic variability and a slow rate of molecular evolution, constrain RuBisCO kinetic evolution to a greater extent than catalytic trade-offs. This new insight challenges the widely-
Results

RuBisCO kinetic data

A dataset comprising kinetic measurements for RuBisCO isolated from different photoautotrophs was obtained from (Flamholz et al., 2019). Measurements of specificity (S_{CO}) for CO₂ relative to O₂ (i.e., the overall carboxylation/oxygenation ratio of RuBisCO under defined concentrations of CO₂ and O₂ gases) in this dataset were normalised in order to overcome discrepancies between values determined using an oxygen electrode assay (Parry, Keys and Gutteridge, 1989) and high precision gas phase controlled 3H-RuBP-fixation assays (Kane et al., 1994) (see methods). Interrogation of this data was restricted to the angiosperms in order to minimize long-branch effects in subsequent analyses (Su and Townsend, 2015). It was also restricted to those species with measurements of S_{CO} (mol.mol⁻¹), maximum carboxylase turnover rate per active site (k_{catC}; s⁻¹), and the Michaelis constant (i.e., the substrate concentration in μM at half saturated catalysed rate) for both CO₂ (K_{C}) and O₂ (K_{O}). The Michaelis constant for CO₂ in 20.95 % O₂ air (K_{C,air}) was also inferred as a function of both K_{C} and K_{O} (see methods). Of the 137 species that satisfied these filtration criteria, 19 also had measurements of the Michaelis constant for RuBP (K_{RuBP}). From here on, these constants and rates are collectively termed kinetic traits, where S_{CO}, k_{catC}, K_{C} and K_{C,air} are referred to as carboxylase-related kinetic traits, and K_{O} as the oxygenase-related kinetic trait.

Significant changes in RuBisCO kinetics traits are with the evolution of C₄ photosynthesis

Consistent with previous analyses (Flamholz et al., 2019), all kinetic traits were log transformed to ensure they conformed to the distribution assumptions of the statistical analyses herein. Several dependencies in RuBisCO kinetic traits on photosynthetic type were found, with differences observed in carboxylase-related trait values between plant species that conducted C₃ and C₄ photosynthetic pathways (Figure 1). Specifically, the mean of the distribution of RuBisCO S_{CO} values in C₄ species (mean S_{CO} = 78.7 mol.mol⁻¹) was lower than that observed for RuBisCO in C₃ species (mean S_{CO} = 89.9 mol.mol⁻¹) (Figure 1; p < 0.001, t-test). Conversely, the mean of the distribution of RuBisCO k_{catC} values was higher in C₄ species (mean k_{catC} = 4.2 s⁻¹) than in C₃ species (mean k_{catC} = 8.9 s⁻¹).
= 3.2 s⁻¹) (Figure 1; p < 0.001, t-test). The means of the distributions of both $K_C$ and $K_C^{air}$ were also found to be higher in C₄ species (mean $K_C = 19.0$ µM, mean $K_C^{air} = 29.9$ µM) than in C₃ plants (mean $K_C = 15.4$ µM, mean $K_C^{air} = 23.6$ µM) (Figure 1; p < 0.05 and p < 0.05, t-test, respectively). In contrast, no significant difference was observed in $K_o$ between C₃ species (mean $K_o = 481.0$ µM) and C₄ species (mean $K_o = 466.7$ µM) (Figure 1; p > 0.05, t-test). However, variation in $K_o$ was found to be considerably greater in C₄ species (95 % CI = [379.1, 574.6]) than in C₃ species (95 % CI = [457.1, 506.0]) (p < 0.01; Levene test). Moreover, although the restricted number of $K_{RuBP}$ measurements did not allow statistical differences to be assessed between photosynthetic groups, the distribution of this trait appeared to show higher variability in C₄ species, similar to that observed for $K_o$. Taken together, this data supports a hypothesis that RuBisCO in C₄ plants have adapted to saturating CO₂ concentrations by increasing carboxylase turnover at the expense of lower specificity and CO₂ affinity, whilst RuBisCO affinity for O₂ (and possibly RuBP) appears to exhibit increased variability in C₄ species. These differences in the kinetic traits of RuBisCO between C₃ and C₄ plants were robust to phylogenetic correction (see below).

Owing to limited kinetic measurements for RuBisCO in C₃-C₄ intermediate and C₄-like species which respectively represent early and late transition states along the evolutionary continuum from C₃ to C₄ photosynthesis, it was not possible to assess changes in RuBisCO kinetics in these plants relative to the ancestral C₃ and derived C₄ photosynthetic types. Nevertheless, trait values of RuBisCO $S_{C/O}$ in both evolutionary intermediate C₃-C₄ and C₄-like states appear to closely resemble the distribution observed in C₄ species (Figure 1), thus indicating adaptation of this trait potentially occurred early during the evolution of C₄ photosynthesis.

**Significant phylogenetic signal exists in RuBisCO carboxylase-related kinetic traits**

To assess whether RuBisCO in different angiosperms display similar kinetics as a consequence of their phylogenetic relationship, the quantitative kinetic traits of the enzyme were analyzed in the context of the phylogenetic tree by which the species are related (Figure 2). Here, all kinetic traits were subject to interrogation for phylogenetic signal (Table 1) except for $K_{RuBP}$, which was omitted owing to the limited number of measurements available for this trait. For these analyses, several statistical tools varying in their approach to phylogenetic signal detection were implemented and the presence or absence of phylogenetic signal in each trait was judged by the majority result (i.e., the
result of ≥ 3 out of 5 methods tested). Out of the methods utilized, Pagel’s lambda (Pagel, 1999) and Blomberg’s K and K* (Blomberg, Garland and Ives, 2003) analyse the distribution of trait values in extant species using an explicit Brownian motion model of trait evolution in which the traits evolve stochastically on the underlying phylogenetic tree at a uniform rate and independently among branches. In contrast, Moran’s I (Gittleman and Kot, 1990) and Abouheif’s Cmean (Abouheif, 1999) do not invoke any specific aspect of evolutionary theory, but instead test for phylogenetic signal by assessing the correlation of trait values across evolutionary distance on the species tree using the concept of autocorrelation adopted from the field of spatial statistics (Cheverud, Dow and Leutenegger, 1985, 1986). For further discussion of the differences between these phylogenetic signal detection methods see (e.g., Münkemüller et al., 2012).

Irrespective of the methodological approach used for inference, a significant phylogenetic signal was observed in all of the carboxylase-related kinetic traits, $S_{\text{C/O}}$, $k_{\text{catC}}$, $K_{\text{C}}$, and $K_{\text{C}^\text{air}}$ (Table 1; Figure 2). However, the strength of this signal varied across the different methods (Table 1). In contrast, phylogenetic signal was not detected for the oxygenase-related kinetic trait $K_{\text{O}}$ (Table 1; Figure 2). These measurements of phylogenetic signal were demonstrated to not suffer from overfitting due to the use of the rbcL gene to infer the phylogenetic tree (Supplementary File 1, Figure S1, Figure S2, and Table S1). Overall, this means that carboxylase-related (but not oxygenase-related) kinetic traits in different species are similar as a consequence of their phylogenetic relationship. Therefore, conventional approaches to measure correlations that assume independence between observations of carboxylase-related kinetic trait values are invalid, and correlation coefficients computed using such approaches have likely been over-estimated.

As the presence of phylogenetic signal in RuBisCO carboxylase-related kinetic traits violates the assumption that species measurements are independent, kinetic variations observed between C$_3$ and C$_4$ individuals (Figure 1) were re-analysed accounting for the evolutionary relationships among angiosperms. Here, differences in the distributions of $S_{\text{C/O}}$, $k_{\text{catC}}$, $K_{\text{C}}$ and $K_{\text{C}^\text{air}}$ between C$_3$ and C$_4$ species (Figure 1) were found to be robust after corrections were applied based on the structure of the phylogenetic tree (Supplementary File 1, Table S2). Thus, there is genuine adaptation in RuBisCO kinetics that are associated with the evolution of C$_4$ photosynthesis, such that the
emergence of the C4 carbon concentrating mechanism is accompanied by a decreased specificity and CO2 affinity, and an increased carboxylase turnover.

**Significant phylogenetic signal exists in the RuBisCO K0 in C3 plants**

Based on the positions of C3-C4 intermediate, C4-like, and C4 species in the phylogenetic tree (Supplementary File 1, Figure S1), multiple independent evolutions toward C4 photosynthesis are present in the dataset. Furthermore, given that transition to C4 photosynthesis is found above to be associated with adaptive changes in RuBisCO kinetic traits including a reduction in $S_{CO}$, an increase in $k_{catC}$, $K_C$ and $K_C^c$, as well as increased variability in $K_O$ (Figure 1; Supplementary File 1, Table S2), it was hypothesised that a failure to account for kinetic differences associated with photosynthetic type may have confounded estimations of phylogenetic signal. For example, kinetic modifications associated with the evolution of C4 photosynthesis may cause larger differences in RuBisCO kinetics among closely related C3 and C4 species than expected based on evolutionary distance alone. Similarly, the independent evolution of C4 photosynthesis in distantly related plant lineages could also cause evolutionarily distant species to have similar kinetic trait values by convergence. To evaluate the extent by which these respective issues might have affected quantification of phylogenetic dependency, the phylogenetic signal of RuBisCO kinetic traits was reassessed using only the C3 species present in Supplementary Figure S1 (i.e., with C3-C4 intermediate, C4-like, and C4 species removed). In general, estimates of phylogenetic signal in the carboxylase-related kinetic traits in C3 species (Table 2) agreed with those observed when all species were considered on the phylogenetic tree (Table 1). Specifically, a phylogenetic signal of similar strength and significance was observed in $S_{CO}$, $k_{catC}$ and $K_C$ for each of the detection methods across both sets of analyses (Table 1 and Table 2). In addition, the discrepancies in signal strength between the statistical methods previously observed (Table 1) were recapitulated in the analysis using only C3 species (Table 2), thus indicating that these differences are not caused by a failure to control for photosynthetic type, but instead more likely represent distinctions in the assumptions and aspects of phylogenetic signal measured by each test (Hardy and Pavoine, 2012; Münkemüller et al., 2012). In summary therefore, there is a statistical phylogenetic co-variation in RuBisCO specificity, carboxylase turnover and the Michaelis constant for CO2 in angiosperms that is independent of photosynthetic type.
In contrast to the analysis of all species (Table 1) a significant signal was observed in $K_O$ when the analysis was restricted to $C_3$ species (Table 2). Thus, both the oxygenase-related and carboxylase-related traits of RuBisCO evolve in a tree-like manner in $C_3$ species. Furthermore, unlike the other carboxylase-related kinetic traits, the phylogenetic signal in $K_{C\text{air}}$ is found to increase in strength when the analysis is restricted to $C_3$ species. This result is a corollary of the fact that $K_{C\text{air}}$ is computed here from both $K_C$ and $K_O$. Taken together, these results reveal that all kinetic traits of RuBisCO have a significant phylogenetic dependency in $C_3$ species.

**Correlations between kinetic traits are weak in angiosperms and further relaxed after correcting for phylogenetic signal**

Given the finding that RuBisCO kinetic traits exhibit significant phylogenetic signal (Table 1; Figure 2), it is possible that previously reported associations between RuBisCO kinetic traits are an artefact of this signal. This is because prior analyses which did not account for the evolutionary relationships among (and non-independence of) species may suffer from over-estimated correlation coefficients as a consequence of the phylogenetic co-variation which is observed in the kinetic traits (Figure 3A).

In order to evaluate the severity by which phylogenetic co-variation may have influenced previous results, the correlations observed in the kinetic trait data using both phylogenetic and non-phylogenetic regression methods were compared (Figure 3B and 3C).

Using a standard non-phylogenetic approach the relationships between kinetic traits of RuBisCO were consistent in both standard linear and least squares regression models (Supplementary File 1, Figure S3). The direction of correlations observed (Figure 3B) match those previously reported (Flamholz et al., 2019). Specifically, significant positive power-law (log—log scale) relationships were found between $k_{\text{catC}}$ and both $K_C$ and $K_{C\text{air}}$ (Figure 3B and 3C). A significant positive correlation was also observed between the respective Michaelis constants for both $CO_2$ and $O_2$ substrates, $K_C$ and $K_O$ (Figure 3B and 3C). In addition, significant inverse power-law correlations were observed between $S_{C/O}$ and all other carboxylase-related kinetic traits investigated, including $k_{\text{catC}}$, $K_C$ and $K_{C\text{air}}$ (Figure 3B and 3C). In contrast, $K_O$ did not co-vary with either $S_{C/O}$ or $k_{\text{catC}}$ (Figure 3B and 3C), whilst $K_{\text{RuBEP}}$ did not appear to be tightly linked to any kinetic trait from the limited number of observations that are available (Supplementary File 1, Figure S3). Thus, all pairwise relationships between the
carboxylase-related kinetic traits $S_{\text{C/O}}$, $k_{\text{catC}}$ and either $K_C$ or $K_C^\text{air}$ were significant, whilst the oxygenase-related trait $K_O$ was only correlated with $K_C$.

Although kinetic trade-offs inferred using non-phylogenetic methods were concordant in direction with those previously described (Flamholz et al., 2019), they were substantially reduced in magnitude when the analysis was focussed solely on the angiosperms. For example, the strength of the correlation between $S_{\text{C/O}}$ and $K_C$ in angiosperm RuBisCO (9.9 % variance explained; Figure 3C) is attenuated by a factor of 4.4-fold when compared to that previously reported using a larger range of photoautotrophs (43.6 % variance explained (Flamholz et al., 2019)). Moreover, a 3.2-fold reduction was found in the dependency between $S_{\text{C/O}}$ and $k_{\text{catC}}$ in angiosperms (9.8 % variance explained; Figure 3C) in comparison to that reported based on the larger range of photoautotrophs (31.4 % variance explained (Flamholz et al., 2019)), whilst the antagonistic correlation observed between $K_C$ and $K_O$ (20.9 % variance explained; Figure 3C) was also weakened by a factor of 1.5-fold relative to previous reports (31.4 % variance explained (Flamholz et al., 2019)). In contrast, the dependency between $K_C$ and $k_{\text{catC}}$ was ~1.5-fold stronger when only angiosperms are assessed, increasing from 23.0 % (Flamholz et al., 2019) to 34.2 % in this study (Figure 3C). Thus overall, even in the absence of correctly accounting for the phylogenetic relationship between RuBisCO, the apparent catalytic trade-offs observed in angiosperms are weaker than previously thought. This supports the assertion that the biochemical landscape in angiosperms is relatively weakly constrained by catalytic trade-offs, and adds to a growing body of knowledge detailing differences in RuBisCO kinetics (Jordan and Ogren, 1981; Andersson and Backlund, 2008; Young et al., 2016; Heureux et al., 2017) and kinetic trade-offs (Iñiguez et al., 2020) across the RuBisCO tree of life.

However, significant phylogenetic signal is present in RuBisCO kinetic traits (Table 1). Thus, phylogenetic generalized least squares regression analyses (Felsenstein, 1985) were conducted in order to correctly estimate the magnitude of the catalytic trade-offs whilst accounting for phylogenetic co-variation. In comparison to non-corrected correlations, the phylogenetic regression caused a reduction in the majority of kinetic trade-offs (Figure 3C). The largest reduction was observed for the correlation between the Michaelis constants $K_C$ and $K_O$. Here, the correlation was reduced by 2.4-fold (variance explained = 8.7 %) relative to methods which do not correctly account for the non-
independence of these measurements (variance explained = 20.9\% ; Figure 3C). An analogous
correlation was also observed when analyses were limited to C\textsubscript{3} species (variance explained = 11.6
\% ; Figure 3C). Thus, there has only been a marginal trade-off between \(K\textsubscript{C}\) and \(K\textsubscript{O}\) during the evolution
of RuBisCO in angiosperms.

Phylogenetic correction also reduced the magnitude of the correlation between \(S\textsubscript{C/O}\) and each of the
other carboxylase-related traits (Figure 3C). Here the dependency between \(S\textsubscript{C/O}\) and \(k\textsubscript{catC}\) was
reduced from 9.8 to 8.1 \%, whilst that between \(S\textsubscript{C/O}\) and both \(K\textsubscript{C}\) and \(K\textsubscript{C}^\text{air}\) were reduced from 9.9 to
6.2 \%, and from 9.1 to 5.3 \%, respectively (Figure 3C). Furthermore, these correlations were not
significant when considering only C\textsubscript{3} species (Figure 3C). Thus, there has been either a marginal or
no trade-off between RuBisCO specificity and other carboxylase-related kinetic traits during the
evolution of RuBisCO in angiosperms.

In contrast to the weak trade-offs above, the trade-off between \(k\textsubscript{catC}\) and either \(K\textsubscript{C}\) or \(K\textsubscript{C}^\text{air}\) was robust
to phylogenetic correction. Here, the strength of the correlation between these kinetic traits
experienced a small proportional reduction when correctly accounting for phylogenetic covariation.
Specifically, the correlation between \(k\textsubscript{catC}\) and \(K\textsubscript{C}\) decreased from 34.2 to 30.9 \%, and the correlation
between \(k\textsubscript{catC}\) and \(K\textsubscript{C}^\text{air}\) decreased from 34.5 to 31.6 \% (Figure 3C). The phylogenetically-corrected
correlations between these kinetic traits were of a similar magnitude when only C\textsubscript{3} species were
considered (37.4 and 34.9 \% respectively (Figure 3C)). Consequently, these trade-offs are at least
3-fold stronger than any other trade-off between other RuBisCO kinetic traits. Thus, the trade-off
between CO\textsubscript{2} affinity and carboxylase turnover has been the largest catalytic constraint on the
adaptation of RuBisCO during the evolution of angiosperms.

The evolution of RuBisCO kinetics is more limited by phylogenetic constraints than by catalytic trade-offs
As each of the RuBisCO kinetic traits contained significant phylogenetic signal (Table 1 and 2), the
manner in which RuBisCO has adapted in angiosperms has been phylogenetically constrained. In
order to assess the relative strength these phylogenetic constraints, the variance in kinetic traits
partitioned between phylogenetic effects (i.e., the explanatory power of phylogenetic co-variation in
the goodness-of-fit model and a measure of phylogenetic constraint) and non-phylogenetic effects
(i.e., the remaining explanatory power of the regression model, accounted for by the sum of all other constraints such as random effects and all kinetic trait trade-offs including those not assessed here) was quantified. This revealed that phylogenetic constraints explained a significant proportion of variation in all carboxylase related kinetic traits in all species (Figure 4A). Moreover, consistent with analysis of phylogenetic signal in C_3 species above, phylogenetic constraints also explained a significant proportion of variation in the oxygenase-related kinetic trait in C_3 species (Figure 4B). Furthermore, phylogenetic constraints explained a larger proportion of variation in kinetic traits than any catalytic trade-off (Figure 4C and 4D). Consequently, phylogenetic constraints have been a larger limitation on RuBisCO adaptation than constraints attributable to kinetic traits (Cumulative variance for phylogenetic constraints = 24.8%, cumulative variance for catalytic trade-offs = 9%), with this effect being more pronounced if limited to the C_3 species in the dataset (Cumulative variance for phylogenetic constraints = 43.4%, cumulative variance for catalytic trade-offs = 8.2%). Thus, phylogenetic constraints have limited RuBisCO kinetic evolution to a greater extent than catalytic trade-offs.

**Discussion**

The evolutionary landscape of RuBisCO has long been proposed to be constrained by catalytic trade-offs. In support of this hypothesis, antagonistic correlations between RuBisCO kinetic traits inferred from studies comparing limited numbers of species are commonly cited (Tcherkez, Farquhar and Andrews, 2006; Savir *et al.*, 2010). Specifically, strong dependencies are thought to occur between RuBisCO specificity (S_{C/O}), carboxylase turnover (k_{catC}) and the Michaelis constants for CO_2 (K_C) and O_2 (K_O), respectively (Tcherkez, Farquhar and Andrews, 2006; Savir *et al.*, 2010). Combined, these trade-offs are hypothesized to limit the capacity of RuBisCO to assimilate CO_2 at high rates by curtailing the inherent carboxylase activity of the enzyme, whilst also causing it to catalyse a reaction with O_2 which is energetically expensive and results in a loss of fixed carbon (Bowes, Ogren and Hageman, 1971; Chollet, 1977). However, all trade-offs have been inferred based on the assumption that RuBisCO in different species are independent. Here, we show that this assumption was incorrect and that significant phylogenetic constraint, and thus significant phylogenetic signal (also known as phylogenetic covariation), is found in all carboxylase-related
kinetic traits in all species. In addition, significant phylogenetic constraint is also found in the oxygenase-related kinetic trait in C₃ species. When accounting for the phylogenetic signal caused by this constraint, we demonstrate that previously reported kinetic trait correlations have been over-estimated. Specifically, there is weak trade-off between carboxylase turnover and the Michaelis constant for CO₂ (~30 % dependency), and either non-significant or marginal trade-offs between all other kinetic traits (<9 % dependency). Finally, we demonstrate that phylogenetic constraints have played a larger role than catalytic trade-offs in limiting the evolution of RuBisCO kinetics in angiosperms.

The presence of phylogenetic constraint, and thus phylogenetic signal, in RuBisCO kinetics traits simply means that the kinetic traits of RuBisCO are more similar among close relatives, with this similarity changing as a function of the phylogenetic tree which relates the species. This could be considered to be unsurprising given the fact that all extant enzymes are related by the process of descent with modification from a single ancestral enzyme (Nisbet et al., 2007). However, such a tree-like pattern of evolution is not the rule for biological traits, and several mechanistic explanations have been put forward to account for the presence of such phylogenetic signal (Kamilar & Cooper, 2013; and references therein). In brief, it is often interpreted as suggestive of either neutral evolution under genetic drift (Felsenstein, 1985), or evolutionary stasis (Prinzing et al., 2001; Ackerly, 2004; Cavender-Bares et al., 2004; Moles et al., 2005; Swenson and Enquist, 2007) under which adaptive change is mitigated by processes including stabilizing selection, pleiotropy and a lack of molecular variability or phenotypic plasticity (Maynard Smith et al., 1985; Bradshaw, 1991; Edwards and Naeem, 1993; Wagner, 1995). Although it is possible that several of the above factors contribute to the phylogenetic constraint that gives rise to the phylogenetic signal, it is likely that low molecular variability is a key driver of this phenomenon in RuBisCO. This is because the catalytic subunit of RuBisCO is encoded by a single copy gene that is trapped in the haploid plastid genome and consequently is uniparentally inherited and does not recombine (Birky, 2001). In addition, the low rate of molecular evolution in RuBisCO is further entrenched by evolutionary constraints on the enzyme, including those imposed by the requirements for i) high levels of transcript and protein abundance (Kelly, 2018; Seward and Kelly, 2018), ii) maintaining complementarity to a wide array
of molecular chaperones which assist in protein folding and assembly (e.g. Raf1, Raf2, RbcX, BSD2, Cpn60/Cpn20) and metabolic regulation (e.g. RuBisCO activase) (Carmo-Silva et al., 2015; Aigner et al., 2017), and iii) the need to preserve overall protein stability within the molecular activity-stability trade-offs (Studer et al., 2014; Duraõ et al., 2015; Cummins, Kannappan and Gready, 2018).

Combined, these features create a perfect storm of limitations which constrain the molecular evolution of RuBisCO resulting in significant phylogenetic constraint that is independent of any catalytic requirements.

It should be noted that RuBisCO in angiosperms is a holoenzyme composed structurally of both large and small subunits (Schneider, Lindqvist and Brândén, 1992; Tabita et al., 2008). Thus, although the phylogenetic constraints identified in this study are related to the function of the complete holoenzyme, they apply predominantly to the rbcL gene encoding the large subunit. This is because the rbcS gene is nuclear encoded, multi-copy (in most plant species), undergoes recombination, and is less constrained by molecular chaperones. Moreover, as RbcS has an indirect influence on RuBisCO catalytic properties (Read and Tabita, 1992b, 1992a; Spreitzer, Peddi and Satagopan, 2005; Martin-Avila et al., 2020) and activity (Andrews, 1988; Lee and Tabita, 1990; Lee, Berka and Tabita, 1991), it is possible that this subunit may have helped the RuBisCO holoenzyme to evolve more rapidly and escape some phylogenetic constraint. However, the exact contribution of the RbcS to variation in holoenzyme kinetics, or indeed whether the influence of this protein has already been optimised, remain still to be fully understood.

When accounting for phylogenetic covariation in RuBisCO kinetic traits, the association between the Michaelis constants for CO₂ and O₂ substrates across all angiosperms was reduced to below 9 %. Thus, changes in CO₂ affinity have occurred with only weak antagonism of O₂ affinity. Moreover, the phylogenetically corrected dependency between specificity and both carboxylase turnover and the Michaelis constant for CO₂ were similarly relaxed to less than a 9 % dependency. However, no antagonism of any kinetic trait on specificity was found for RuBisCO in C₃ plants. Thus, natural selection has altered RuBisCO specificity, and thus the overall carboxylation/oxygenation ratio of RuBisCO, either largely or completely independent of changes in CO₂ affinity or carboxylase turnover, and vice versa. Therefore, catalytic trade-offs between CO₂ and O₂ affinity, and between...
specificity and other carboxylase-related kinetic traits, exert only very limited constraint on the
evolution of RuBisCO kinetics in angiosperms.

In contrast, whilst the biochemical landscape of RuBisCO is only weakly constrained by the majority
of catalytic trade-offs, a phylogenetically corrected ~30% dependency was observed between $k_{\text{catC}}$
and both $K_C$ and $K_{C^\text{air}}$ in angiosperms. Thus, there has been a more pronounced trade-off between
$\text{CO}_2$ affinity and carboxylase turnover during the evolution of RuBisCO. This result is compatible with
the mechanistic models of RuBisCO (Farquhar, 1979), and is supported by the recent discovery of
enzyme variants which exhibit the highest $k_{\text{catC}}$ ever recorded at the expense of poor $\text{CO}_2$ affinities
(i.e. $K_C$ values >250 µM) (Davidi et al., 2020). Nevertheless, the dependency between $\text{CO}_2$ affinity
and turnover, despite being the strongest pairwise correlation between kinetic traits observed, is still
only partial (~30% variance explained), and is substantially attenuated relative to the coefficients
conventionally cited from studies investigating limited numbers of species (Tcherkez, Farquhar and
Andrews, 2006; Savir et al., 2010). Thus, although selecting for a greater RuBisCO carboxylase
turnover is evolutionarily linked with a poorer affinity for $\text{CO}_2$ (higher $K_C$), and vice versa, significant
plasticity exists in this relationship among angiosperms such that variation in one kinetic trait only
explains 30% of variation in the other. This explains why there is variability in the carboxylation
efficiency among angiosperm RuBisCO (defined as $k_{\text{catC}}/K_C$), a core parameter which defines the
initial slope of the response of $\text{CO}_2$ fixation rate to changes in $\text{CO}_2$ concentration within the aerobic
environment of chloroplasts in $C_3$ species (Sharwood, 2017).

The phylogenetically resolved analysis of the evolution of RuBisCO kinetic traits also identified
changes in kinetics traits associated with the evolution of $C_4$ photosynthesis. Specifically, $S_{\text{CO}_2}$ was
lower in $C_4$ species than in $C_3$ species while $k_{\text{catC}}$, $K_C$ and $K_{C^\text{air}}$ were higher in $C_4$ species than in $C_3$
species. Moreover, variation in $K_O$ was found to be greater in $C_4$ species than in $C_3$ species. All of
these changes are consistent with the lack of catalytic trade-off, and are neutral or adaptive in a $C_4$
context. For example, any change in $K_O$ would effectively be neutral in the $\text{CO}_2$ saturated
environment of the bundle sheath chloroplast as it would have only a marginal effect on in vivo
carboxylation rate or carboxylation-to-oxygenation ratio, and thus would not cause a concomitant
change to organism fitness. In contrast, an increase in $k_{\text{catC}}$ in the same $\text{CO}_2$ saturated environment
would enable higher flux through RuBisCO and thus provide an energetic advantage to the cell. Accordingly, one would expect that elevated $K_C$ in $C_4$ species would occur by neutral drift (Kimura, 1991), and that an increase $K_C$ would occur given the weak antagonistic relationship between $k_{catC}$ and $K_C$. Thus, the adaptations to RuBisCO kinetics that occur concomitant with the evolution of $C_4$ photosynthesis are consistent with the change in $CO_2:O_2$ ratio and weak catalytic trade-off that exists between $k_{catC}$ and $K_C$. Moreover, the rate at which these changes occurred may have been facilitated by the increase rate of molecular evolution (Kelly, 2018) and diversification (Spriggs, Christinb and Edwards, 2014) that occurs concomitant with the evolution of $C_4$ photosynthesis.

Whilst every effort was taken to prevent systematic or methodological biases from influencing the results presented in this work, several factors outside of the control of the study may have led to the underestimation of phylogenetic dependency in traits. Specifically, experimental error in kinetic trait measurements may have constrained the detection of phylogenetic covariance, as has been shown in other studies (Rohlf, 2001). However, additional sources of measurement inaccuracy in the data analysed here may include i) biases introduced by a failure to account for the breadth of intraspecific variation among individuals, populations and ecotypes (Ives, Midford and Garland, 2007), ii) the fact that in vitro estimates of kinetics poorly correlate with the enzyme behaviour in vivo (e.g., Bar-Even et al., 2011), or finally iii) inconsistencies associated with measurements being compiled from numerous sources in the literature (Flamholz et al., 2019). However, toward partially addressing the latter inaccuracy, the $S_{CO}$ values used in our analyses have been normalised to avoid the known discrepancies between the rates measured using an oxygen electrode assay (Parry, Keys and Gutteridge, 1989) and those measured using high precision gas phase controlled $^3$H-RuBP-fixation assays (Kane et al., 1994) (see methods). Nevertheless, even though kinetic trade-offs measured here are reduced when compared to previous reports, it is possible that correlations have still been over-estimated due to the fact that the phylogenetic covariance inferred in this work represents a lower limit of the true signal due to these potential errors.

Given the importance of RuBisCO to life on Earth, the question as to why a “perfect” RuBisCO has not already materialized during evolution is legitimate. For example, although RuBisCO $K_C$ is thought to be near optimal in $C_3$ plants (e.g., Figure 2) in light of the ~8 µM chloroplastic concentration of
CO₂ and the inherent limitations of CO₂ as a substrate, including its inertness, hydrophobicity and low molecular mass (Andrews and Whitney, 2003; Bar-Even et al., 2011; Bathellier et al., 2018), the observed \( k_{\text{catC}} \) (\( \sim 3 \text{ s}^{-1} \)) is generally low compared to the turnover rates of other enzymes in primary metabolism (Bar-Even et al., 2011; Tcherkez, 2013; Davidi et al., 2018). In addition, all known RuBisCO variants catalyse a promiscuous and energetically costly reaction with O₂. However, the revised kinetic trait trade-offs described in this study suggest biochemical constraints on the optimisation of carboxylase and oxygenase kinetic traits are either partial or weak, and thus are not predominantly responsible for slow adaptation of RuBisCO in angiosperms. Instead, the findings presented here unearth a novel and important role of phylogenetic constraints, which are found to restrict the adaptive evolution of RuBisCO kinetics to a larger extent than catalytic constraints. These phylogenetic constraints are most likely the consequence of the inherent low level of molecular evolution of the catalytic large subunit of RuBisCO, as a consequence of both the genetics and heredity of chloroplasts and non-catalytic constraints on the enzyme, as discussed above. By inducing a considerable lag in adaptive evolution, this phylogenetic constraint slows the adaptation of RuBisCO and helps explain why the enzyme is better suited to former environmental conditions.

The phylogenetically resolved analysis of angiosperm RuBisCO kinetics presented here reveals that the biochemical landscape of RuBisCO is weakly constrained by catalytic trade-offs in angiosperms. Instead, phylogenetic constraints have provided a more substantial limitation to the evolution of RuBisCO kinetics. Accordingly, it should be feasible in the current synthetic biology revolution to circumvent these evolutionary barriers to RuBisCO adaptation. Indeed, promising steps toward this goal have been already demonstrated using directed evolution of the enzyme to generate variants with improved catalytic traits in non-photosynthetic archaea (Wilson, Alonso and Whitney, 2016) and plant-like Form I hexadecameric RuBisCO from photosynthetic bacteria (Zhou and Whitney, 2019) and cyanobacteria (Wilson et al., 2018). Thus, these findings presented here provide optimism for engineering RuBisCO in food, fibre and fuel crops to have improved catalytic efficiency.

**Materials and Methods**

**Kinetic data**

Kinetic measurements of RuBisCO (Form IB) from a large number of plant species were attained from (Flamholz et al., 2019). \( S_{\text{CO}} \) values measured using the O₂ electrode method which calculate
[CO$_2$] using a pKa of 6.11 (Parry, Keys and Gutteridge, 1989) were normalized relative to $S_{C/O}$ values quantified using high precision gas phase controlled $^3$H-RuBP-fixation assays (Kane et al., 1994).

Here, as the $S_{C/O}$ of wheat (*Triticum aestivum*) RuBisCO was measured using the method of (Kane et al., 1994) at 88 mol.mol$^{-1}$, and *T. aestivum* is also used as an enzyme standard in the O$_2$ electrode studies (Orr et al., 2016; Prins et al., 2016), multipliers of either 0.88 (Prins et al., 2016) or 0.85 (Orr et al., 2016) were applied to normalise all $S_{C/O}$ measurements to 88 mol.mol$^{-1}$ for *T. aestivum* RuBisCO and minimise methodological biases in the data.

All kinetic traits in the dataset were transformed on a log scale consistent with (Flamholz et al., 2019), and the distributional assumptions of each were verified for analyses herein. Only angiosperm species with experimental measurements of all four principal experimentally measured kinetic traits of interest ($S_{C/O}$, $k_{catC}$, $K_C$ and $K_O$) were taken forward for subsequent analysis. A small number of these species also had measurements available for $K_{RuBP}$. In addition, the measure of the Michaelis constant for CO$_2$ under 20.95 % ambient air, $K_{C_{air}}$, was inferred from $K_C$ and $K_O$ based on the formula $K_C + (K_C [O_2] / K_O)$, where 20.95 % [O$_2$] in water is 253µM. In cases where duplicate entries for a species were present in the kinetic dataset (including synonyms), the median value of their respective kinetic traits was used for inference. In this way, averages were also taken between *Triticum timonovum* and *Triticum timopheevii*, the former of which is a synthetic octoploid of the latter (Murashov and Morozova, 2008). The modified dataset containing corrected $S_{C/O}$ values and averages across duplicate entries for species is provided in Supplemental File 2.

**Phylogenetic tree inference**

As sequenced genomes or transcriptomes do not exist for many plants in the kinetic trait dataset, whole genome phylogenomic approaches could not be used to infer the species tree necessary in order to detect phylogenetic signal in measured kinetic traits of the RuBisCO holoenzyme. However, the *rbcL* gene that encodes the large subunit of RuBisCO has a long history of use in plants for phylogenetic inference of species relationships (Gielly and Taberlet, 1994; APG, 1998, 2016) and was available for all of the angiosperms that were used in the analysis. As such, *rbcL* was selected here for use in species tree inference. The complete coding sequences of *rbcL* for the 137 angiosperm species for which kinetic data was available can be found in Supplementary File 3. These sequences were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/) for all species except...
Flaveria brownii which was acquired from the 1KP database (Leebens-Mack et al., 2019). A multiple sequence alignment was performed using mafft L-INS-i (Katoh and Standley, 2013), and alignments were trimmed at the terminal ends to remove unaligned positions using AliView (Larsson, 2014). This trimmed nucleotide sequence alignment was used for subsequent phylogenetic analysis. A bootstrapped maximum likelihood phylogenetic tree was inferred by IQ-TREE (Nguyen et al., 2015) using the ultrafast bootstrapping method with 1000 replicates and the Shimodaira-Hasegawa approximate likelihood ratio branch test. The resultant maximum likelihood phylogenetic tree was rooted manually using Dendroscope (Huson and Scornavacca, 2012). A number of nodes (n = 18) in the complete tree (Supplementary File 1, Figure S1) exhibited terminal zero-length branches due to 100 % sequence identify with other closely related species. These species were condensed into single data points (as their rbcL are 100 % identical) and the mean of their kinetic traits was used. This reduced the dataset to 119 unique enzymes. The phylogenetic tree inferred from these taxa (Supplementary File 1, Figure S1) closely matched the topology of the phylogenetic tree expected from the angiosperm phylogeny with only a few alterations (APG, 2016). Moreover, the topology of the rbcL tree most accurately reflects the sequence similarity of the RuBisCO genes and thus it was deemed a suitable tree for investigation of phylogenetic signal and its effects on correlations between RuBisCO kinetic traits.

To confirm that the phylogenetic signal was not attributable to overfitting caused by the use of the rbcL gene to infer the phylogeny of RuBisCO, an analogous maximum likelihood phylogenetic tree was also inferred from a multiple sequence alignment in which codons containing non-synonymous nucleotide sequence changes were removed (Supplementary File 4). Due to the considerable loss of phylogenetic information accessible for tree building from this alignment, the species tree inferred using nucleotide sequences corresponding to ubiquitously conserved amino acid positions (Supplementary File 1, Figure S2) exhibited an additional number of species belonging to terminal zero-length branches (n = 13). As the sequences of these species are known to exhibit non-synonymous mutations which are not included in the tree, it is not appropriate to take means of their kinetic traits as above. As such, these data points were removed from analysis using only this tree. Use of this phylogenetic tree confirmed that the presence of phylogenetic signal in kinetic traits was
not due to overfitting, however as this tree was less accurate than the full-length alignment tree, it was not used for any subsequent analysis.

**Phylogenetic signal analysis**

The presence of phylogenetic signal in kinetic traits was assessed using five different phylogenetic signal detection methods (Gittleman and Kot, 1990; Abouheif, 1999; Pagel, 1999; Blomberg, Garland and Ives, 2003). Here, signal strength was estimated by assessing the distribution of trait values relative to the underlying species tree inferred from rbcL sequences in (Supplementary File 1, Figure S1) using methods which both depend on an explicit evolutionary model, such as Pagel’s lambda (Pagel, 1999) and Blomberg’s K and K* (Blomberg, Garland and Ives, 2003), as well as the spatial autocorrelation analyses of Moran’s I (Gittleman and Kot, 1990) and Abouheif’s Cmean (Abouheif, 1999). Implementation of these phylogenetic signal detection tools was performed using the phyloSignal function of the phylosignal package (Keck et al., 2016) in the R environment.

**Ancestral state estimation and mapping of kinetic traits to the phylogenetic tree**

Ancestral state estimation was conducted to visualise the evolution of RuBisCO kinetic traits on the phylogenetic tree which relates the species. For this purpose, the kinetic dataset was mapped and scaled onto the species tree by employing the function ggtree in the ggtree package (Yu et al., 2017). Here, terminal branches were coloured according to the measurement of the kinetic trait in the species which comprise the terminal branch, whereas internal branches were coloured based on values inferred in ancestral species using ancestral state estimation (Yu et al., 2017).

**Least squares and linear regression models**

All regression models between pairwise combinations of kinetic traits were fit in the R environment.

Phylogenetic generalized regression accounting for the phylogenetic non-independence between species was performed using the function pgls in the caper package (Comparative Analyses of Phylogenetics and Evolution in R) (Orme et al., 2014). In each case, the structure of the phylogenetic signal was corrected for using branch length transformations of the phylogenetic tree based on the mean maximum likelihood estimates of lambda calculated for the traits being interrogated with kapa and delta held constant. Phylogenetic corrections to differences in kinetic trait values between C₃ and C₄ plants based on the phylogenetic non-independence of species were also applied using the pgls function (Orme et al., 2014) with photosynthetic type incorporated as a factorial variable.
In order to partition the variance in RuBisCO kinetic traits explained by phylogenetic constraints as compared to non-phylogenetic constraints, the \textit{rr2} package (Ives and Li, 2018) was employed in R. To assess the contribution of phylogeny in explaining the co-variation in kinetic trait values, the explanatory power of the phylogenetic component was measured by comparing full and reduced phylogenetic regression models using the partial $R^2_{\text{pred}}$ inferential statistical, based on advice from (Ives, 2019) of this being the most intuitive and direct approach to understand how different model components explain variation. For this analysis, phylogenetic regression models were fit using the \textit{phylolm} function in the \textit{phylolm} package (Tung Ho and Ané, 2014) using Pagel’s lambda model for the error term.

\section*{Acknowledgements}
This work was funded by the Royal Society and the European Union’s Horizon 2020 research and innovation program under grant agreement number 637765. JWB, JRN, JSB, AE, AB and AU were funded by the BBSRC through BB/J014427/1. TR and SMW were supported by the Australian Government through the Australian Research Council Centre of Excellence for Translational Photosynthesis CE140100015

\section*{Author Contributions}
SK conceived the study. JWB conducted the analysis. JWB, TR, SMW, DME, and SK analyzed the data. SK and JWB wrote the manuscript. JRN, JSB, AE, AB and AU contributed to discussions and methodological development. All authors edited the manuscript.

\section*{References}


Savir, Y. et al. (2010) ‘Cross-species analysis traces adaptation of Rubisco toward optimality in a


Figure 1. The distributions of values for RuBisCO kinetic traits in angiosperms. Species are grouped by their photosynthetic types (rows). $S_{CO2}$: specificity. $k_{catC}$: carboxylase turnover. $K_C$: the Michaelis constant for the CO$_2$. $K_C^{air}$: the inferred Michaelis constant for CO$_2$ in 20.95% O$_2$ air. $K_O$: the Michaelis constant for O$_2$. $K_{RuBP}$: the Michaelis constant for ribulose 1,5-bisphosphate. Plants have been classified as those which perform C$_3$ photosynthesis (C$_3$), C$_4$ photosynthesis (C$_4$), C$_3$-C$_4$ intermediate (C$_3$-$C_4$), C$_4$-like (C$_4$-like). The x axis for all plots is on a log scale, where respective units are shown in column labels.
Figure 2. The evolution of RuBisCO kinetic traits in angiosperms. Phylogenetic tree of angiosperm species depicting the kinetic trait values in of the species used in this dataset and the maximum likelihood inferred ancestral kinetic traits for internal branches on the tree. Scale bars for colour schemes are presented next to each tree. Species names have been abbreviated for legibility and are provided in full in Supplementary File 1, Table S3.
**Figure 3.** The correlations between RuBisCO kinetic traits in angiosperms. A) Heatmap depicting the normalised variation in kinetic traits across the species set used in this study (± S.D. away from each respective kinetic trait mean). Species labels on the tree are colour coded by photosynthetic type (C₃: black, C₃-C₄ intermediates: red, C₄-like: blue, and C₄: green), and have been abbreviated for legibility (refer to Supplemental File 1 Table S3). B) Trends in relationships between all pairwise combinations of log transformed RuBisCO kinetic traits shown in A. C) Pairwise correlation coefficients (percent variance explained) and associated p-values between different RuBisCO kinetic traits assessed using non-phylogenetic least squares regression models or phylogenetic least
squares regression models. Phylogenetic least squares regression was fit to both the complete set of angiosperms in the dataset and the subset which perform C₃ photosynthesis. Significance values are represented as α levels, where; α = 0.001 if \( p < 0.001 \), α = 0.01 if \( 0.001 < p < 0.01 \), α = 0.05 if \( 0.01 < p < 0.05 \), and α = ns if \( p > 0.05 \).
Figure 4. The evolutionary constraints on RuBisCO kinetic adaptation. A) The variation (%) in RuBisCO kinetic traits across all species which can be explained by phylogenetic constraint or individual catalytic trade-off. B) as in A but for just the C₃ species. C) Boxplot of all variation explained by each kinetic trait in comparison to variation explained by phylogeny across all species. The phylogenetic constraints on the carboxylase-related traits (PhyCX) and phylogenetic constraints on the oxygenase-related trait (Phyox) are presented separately. D) As in C but for just the C₃ species.
### Table 1. Phylogenetic signal strength and associated significance level in RuBisCO kinetic traits in angiosperms using five different signal detection methods. Statistics are rounded to 3 decimal places, and significance values are represented as α levels, where; α = 0.001 if p < 0.001, α = 0.01 if 0.001 < p < 0.01, α = 0.05 if 0.01 < p < 0.05, and α = ns if p > 0.05.

<table>
<thead>
<tr>
<th>Kinetic Trait</th>
<th>C mean</th>
<th>I</th>
<th>K</th>
<th>K*</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stat</td>
<td>α</td>
<td>Stat</td>
<td>α</td>
<td>Stat</td>
</tr>
<tr>
<td>S_{C/O}</td>
<td>0.516</td>
<td>0.001</td>
<td>0.425</td>
<td>0.001</td>
<td>ns</td>
</tr>
<tr>
<td>k_{catC}</td>
<td>0.350</td>
<td>0.001</td>
<td>0.224</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>K_{C}</td>
<td>0.282</td>
<td>0.001</td>
<td>0.248</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>K_{Cair}</td>
<td>0.234</td>
<td>0.001</td>
<td>0.242</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>K_{O}</td>
<td>0.032</td>
<td>ns</td>
<td>-0.070</td>
<td>ns</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 2. Phylogenetic signal strength and associated significance level in RuBisCO kinetic traits in C_3 species using five different signal detection methods. Statistics are rounded to 3 decimal places, and significance values are represented as α levels, where; α = 0.001 if p < 0.001, α = 0.01 if 0.001 < p < 0.05, and α = ns if p > 0.05.

<table>
<thead>
<tr>
<th>Kinetic Trait</th>
<th>C mean</th>
<th>I</th>
<th>K</th>
<th>K*</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stat</td>
<td>α</td>
<td>Stat</td>
<td>α</td>
<td>Stat</td>
</tr>
<tr>
<td>S_{C/O}</td>
<td>0.533</td>
<td>0.001</td>
<td>0.453</td>
<td>0.001</td>
<td>0</td>
</tr>
<tr>
<td>k_{catC}</td>
<td>0.387</td>
<td>0.001</td>
<td>0.234</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>K_{C}</td>
<td>0.449</td>
<td>0.001</td>
<td>0.341</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>K_{Cair}</td>
<td>0.398</td>
<td>0.001</td>
<td>0.317</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>K_{O}</td>
<td>0.279</td>
<td>0.001</td>
<td>0.167</td>
<td>0.01</td>
<td>0</td>
</tr>
</tbody>
</table>