***Draft of 2018-2-16; main text currently ~1340 words without refs notes or captions***

**Reports** (up to ~2500 words including references, notes and captions or ~3 printed pages) present important new research results of broad significance. Reports should include an abstract, an introductory paragraph, up to four figures or tables, and about 30 references. Materials and Methods should usually be included in [supplementary materials](http://www.sciencemag.org/authors/instructions-preparing-initial-manuscript),

**Title:** should be no more than 96 characters (including spaces) **One Sentence Summary:** capturing the most important point should be submitted for Research Articles, Reports and Reviews. These should be a maximum of 125 characters and should complement rather than repeat the title **Abstract:** of Research Articles and Reports should explain to the general reader why the research was done, what was found and why the results are important. They should start with some brief BACKGROUND information: a sentence giving a broad introduction to the field comprehensible to the general reader, and then a sentence of more detailed background specific to your study. This should be followed by an explanation of the OBJECTIVES/METHODS and then the RESULTS. The final sentence should outline the main CONCLUSIONS of the study, in terms that will be comprehensible to all our readers. The Abstract is distinct from the main body of the text, and thus should not be the only source of background information critical to understanding the manuscript. Please do not include citations or abbreviations in the Abstract. The abstract should be 125 words or less.

**Main Text** is not divided into sub-headings for Reports. Subheadings are used only in Research Articles, and Reviews.

TITLE Photosynthetic proteins at continental scale

ONE SENTENCE SUMMARY (maybe put the imptce of the tech advance here, complementing the abstract?) – or else the major patterns – can only be one or other of these

ABSTRACT (125 words, 6-8 sentences; currently 139 words)

Wide-area vegetation functioning currently is interpreted and modelled via total N measurements, which have been achievable across many species and sites. Yet within total N lie many specific proteins with different specific functions.

Here we report single pass quantification of almost all individual leaf proteins across species spanning continental scale and much of the world’s bioclimate space.

Proteins responsible for C fixation per leaf area increased toward lower rainfall, as expected from least-cost theory for balancing nitrogen with water use. Proteins responsible for light capture increased toward lower irradiance, as expected for coordination of light capture with carboxylation capacity. Both protein groups increased toward lower temperature, as expected from the decline in activity per mol.

Proteomic methods adopted here for the first time have potential to advance many areas of ecology via broad-scale quantification of proteins with specific functions.

MAIN TEXT

Interpretation of what vegetation is doing across wide areas and many species has previously relied on measurements of leaf nitrogen content, both for comparative trait ecology (refs) and for models intended to capture contribution of vegetation to world carbon budgets (refs). This is because leaf nitrogen content is feasible to analyse across many hundreds of species and samples. Yet total leaf nitrogen includes many different proteins, as well as some non-protein N pools. Each protein or protein-complex carries out specific transactions, ranging from bulk transfers of carbon between atmosphere and vegetation down to defences against specific pathogens or herbivores.

To date, methods for precisely quantifying protein amounts in leaves have been too intensive to apply across large numbers of samples (see SuppMat for extended comment). Laborious benchtop assays have been needed to measure abundance of specific proteins of interest. Alternatively, amounts have been estimated indirectly using proxies: chlorophyll is used as a proxy for amount of light harvesting proteins, and abundance estimates of the carbon fixing enzyme Rubisco can be derived from leaf gas exchange measurements.

Here we apply new proteomics techniques that make possible one-pass quantification of almost all proteins in a leaf. We used quantitative mass spectrometry to analyse 320 leaves from 32 Australian eucalypt species distributed across eastern Australia from 5-27 °C of mean annual temperature and from 200-3200 mm of mean annual rainfall (Fig. 1). The resulting dataset describes protein abundances at all levels of functional organisation, from broad groupings down to individual protein subunits.

We detected up to 2581 individual proteins per leaf sample, among which the 500 most abundant proteins represented 90% (Fig 2c). This is a higher degree of dominance by the top few proteins than observed in [comparison] (Fig 2d), reflecting the specialist nature of leaves as photosynthetic organs. Leaf protein data were associated with physiological and trait measurements for the species, and with site properties. This made it possible to investigate many correlations both within and between these data categories, as illustrated by the correlation heat map Fig X. Absolute protein amounts per leaf area (below the diagonal) were mostly correlated with each other and with total protein and total N, but relative amounts showed few correlations. In other words, variation in total protein across species was typically the strongest influence on protein amounts.

The majority (64% on average, across all species; SD X%) of leaf protein was associated with photosynthesis: 36% was associated with the carbon fixing light independent reactions, 22% (SD X%) with light capture and 4% (SD X%) with photorespiration (Fig 2a). The most abundant individual protein complexes were Rubisco (30%, SD X%) of leaf protein, and photosystem II (X%, SD X%) (Fig 2b). Protein synthesis, folding and degradation contributed X% (SD X%) (Fig 2a). These quantities for the major categories are of the same order as obtained using classical biochemical assays on domesticated species such as *Spinacea oleracea*, *Phaseolus vulgaris*, and *Cucumis sativus* (best compilation by Evans & Seeman 1989).

Here we report mainly about photosynthesis proteins, but readers are welcome to explore other aspects using the interactive data explorer at proteography.org. ~~Photosynthesis is responsible for one of the most important and abundant sets of biochemical reactions within leaves as well as in the biosphere as a whole (Blankenship & Hartman 1998; Raven 2013, Evans & Seeman 1989)..~~

The two major processes in photosynthetic carbon assimilation are carboxylation of ribulose-1,6-bisphosphate (RuBP) by the enzyme Rubisco, and regeneration of RuBP via the light reactions of photosynthesis. At any given time a leaf can be limited either by the light reactions or by CO2 supply and the light-independent reactions ((Farquhar, von Caemmerer & Berry 1980; Farquhar, von Caemmerer S & Berry 2001), but allocation to different parts of the apparatus is thought to be optimized when carboxylation is co-limiting with regeneration of RuBP under average daytime conditions (Haxeltine & Prentice 1996; Chen et al. 2009; Maire et al. 2012, Niinemets & Tenhuenen 1997). This is known as the coordination hypothesis (Maire et al. 2012).

Both light capture proteins and light-independent reaction proteins per leaf area increased notably towards lower mean annual temperature (stat, Fig. 3b-I, stat Fig XX % blah). Their fractional contributions to total protein did not change along temperature gradients (Fig stat %). Increased protein content toward lower temperatures is thought to compensate for slower enzyme activity (Raven & Geider 1988). This effect has been observed for Rubisco in a number of studies (summarised by Hikosaka et al 2006).

Light capture protein per leaf area increased with lower irradiance (Fig. 3b-v, X% per Y irradiance; Fig 3blah stat). Differences in irradiance arose from three sources (detail in Methods). Some sites received more year-round sunlight than others. Individual shoots were in more or less shaded positions, estimated by fish-eye canopy photographs plus software. Along the length of each shoot, older leaves tended to be more shaded. Light-independent reaction protein per leaf area did not change systematically with irradiance, and consequently light capture contributed a larger fraction and light-independent reactions a smaller fraction of protein at lower irradiance (Fig, %, stat). A higher ratio of light-capture to carboxylation apparatus under lower light is what is expected under the coordination hypothesis (Niinemets 2007) (Farquhar et al. 1980).

Light-independent reaction proteins per leaf area increased towards drier sites (Fig x stat %). This was expected on the basis that species operating in drier atmospheres typically draw down leaf-internal CO2 more strongly, and this offsets what would otherwise be an increase in water expenditure via the stomata for a given rate of CO2 assimilation. This “least-cost” interpretation of how nitrogen and water are combined as inputs to the photosynthetic process (Wright et al. 2001a,b, Scalon & Wright 2017) has previously been supported via measurements of leaf-internal CO2 and of leaf total nitrogen (refs), but not by measurement of the specific proteins involved.

Increasing light independent reaction protein per area mainly took the form of increasing leaf mass per area (Fig. 3c-i), in other words there tended to be more mesophyll tissue under each unit of leaf area. Correspondingly, light independent reaction cycle proteins did not increase as a fraction of total protein (Fig 3b-iii). This contrasts with the outcome in relation to irradiance, where the ratio of light independent reaction to light capture adjusted, as expected from the coordination hypothesis.

The substantial scatter in the light independent reaction – LMA relationship indicated that LMA responded to other requirements in addition to carboxylation capacity. Light reaction proteins per area did not increase with LMA (Fig. 3c-ii), and were more weakly correlated than light independent reaction with total protein (R2). Correspondingly, light reaction as a fraction of total protein varied widely (0.09-0.23, 2.6-fold, versus 0.30-0.39 and 1.3-fold for light independent reaction fraction).

Increase of light independent reaction protein toward low temperatures was 46% over 22 C, increase toward dry environments was somewhat smaller, 37% over 4-fold range of precipitation. For protein amounts per leaf area, temperature and precipitation effects were simply additive (Fig Xa, b). Beginning from the high-temp high-precip (upper right) corner in Fig X, increases in protein per leaf area toward lower temp and lower precip came about because of strong increases in leaf mass per area in both those dimensions (Fig Xe, f). Concentrations per leaf mass actually increased along both those dimensions (Fig Xc, d), but this was outweighed by stronger increases in LMA. Across the diagonal from low temp high precip to high temp low precip (upper left to lower right in Fig X) protein amounts per area were unchanging, but the two contributors LMA and protein per mass were more complicated, with protein per mass lower at the extremes and intermediate in the centre, LMA higher at the extremes and intermediate in the centre.

The major working hypotheses invoked here – coordination theory for the shift in ratio of light reaction to carboxylation with irradiance; least-cost theory for the increased carboxylation per leaf area in drier environments; and overall increase in total protein toward lower temperatures – make predictions across the continent’s bioclimate space about protein amounts relative to each other or per leaf area, rather than per leaf mass. These working hypotheses are supported by the patterns observed.

Although this report has focused on major protein categories, the proteomic methods assembled here [for the first time?] also quantify more than 2000 individual proteins. For example in the correlation heatmap (Fig X), isoprene synthase and rubisco activase appear as well as broader categories. We believe this study will prefigure widespread use of one-pass protein quantification to study the ecology of specific proteins, both those that underpin well-understood processes and those where function is not yet clear.