**1.) Quantifying leaf proteins at the continental scale.**

A total of 324 photosynthetically active Eucalypt leaf samples were collected from 32 species; four species were recorded at multiple location. For each species-location combination, three canopy leaves were collected from each of three individuals to make a total of nine samples.

a.) Sampling locations (triangles) were located along three latitudinal bands, spanning broad gradients of rainfall and temperature. The resulting coverage of climate space represents of much of the vegetated area of the Australian continent;

b.) Mean annual temperature (oC) and mean annual precipitation (mm, log scaled) of sampling sites (triangles) were distributed orthogonally with respect to one another (r = );

2.) **Protein composition of the average eucalypt leaf.**

a.) We used a hierarchical protein functional annotation system (MAPMAN/Mercator, ref) to assign proteins to functional groupings. Here we show the average abundances of proteins associated with all major functional groupings in eucalypt leaves (left) and within photosynthesis (right); angular fraction indicates the proportion of protein associated with a named functional category. % values represent averages of leaf 324 samples across 32 eucalypt species. The hierarchical annotation scheme is represented by the layers of the plot: the innermost layer corresponds to the broadest categories in the scheme, e.g. ‘photosynthesis’. Moving outwards, protein amounts are annotated to progressively more specific functions, e.g. ‘light harvesting complex II’. The majority of protein in leaves is associated with photosynthesis, of which the Rubisco large and small subunits comprise on average 30 % and photosystem II, 18 % on average. Protein synthesis, folding and degradation is the second largest top-level category at X % on average.

b.) The 500 most abundant proteins account for 90 % of the protein in leaves (500th protein shown by grey crosshairs). The steep initial slope of this curve contrasts with those associated with less specialised cells (e.g. mammalian cell, yeast).

**3.) Linking leaf protein abundances with environment and functional traits.**

a.) We are able to assess wide-area environmental patterns of leaf proteins with specific function, at any level of protein organisation. Here we show a heatmap of correlations between environmental variables, functional traits, gas exchange measurements and major protein functional categories. Pearson correlations between pairs of variables are represented by coloured tiles where p < 0.05. Protein abundances on a per leaf area basis (mg protein / m2 leaf area) are used to calculate correlations presented in the bottom/right diagonal and proportional protein abundances (i.e. fraction of total leaf protein abundance) are used to calculate correlations presented in the top/left diagonal.

Protein abundances on a per leaf area basis tend to be strongly cross-correlated. A strong positive relationship with leaf nitrogen content is apparent, as expected, together with a somewhat weaker correlation with leaf mass per area. Abundance of proteins associated with individual protein functional categories decline with mean annual temperature and precipitation – a trend which is underpinned by the negative relationship between total leaf protein and these environmental variables (Fig 1e).

Proportional protein abundance of a protein functional category indicates investment in a defined function relative to investment in all other functions, and can be viewed as an allocation trait. A number of trends in protein allocation were apparent across environmental gradients and in relation to functional traits. For example, allocation to light capturing protein (represented by the ‘photosystems’ category), was negatively related to measures of light availability (incident irradiance and canopy gap fraction). Proportional abundances also offer a clearer means to look at how abundances of proteins associated with different functions are related. For example, protein allocation to photorespiration strongly tracks allocation to Calvin cycle proteins, indicating that greater capacity for carboxylation requires a greater capacity to deal with the consequences of photorespiration.

b.) We selected several relationships for deeper analysis which are of current interest to the vegetation modelling community, but which to date have only been investigated via proxies. Trends in abundance of photosystem proteins [symbol, colour] and Calvin cycle enzymes [symbol, colour] are shown across gradients of: i,ii.) canopy-corrected mean annual irradiance (MJ/m2/year), R2 = , modelled change (%) = , p = ; iii,iv.) mean annual precipitation (MAP, mm/year) R2 = , modelled change (%) = , p = ; v,vi.) mean annual temperature (MAT, oC) R2 = , modelled change (%) = , p = .

Points represent the average protein abundance for individual species \* site combinations (n = 9, 3 leaves from each of 3 individuals). SEM error bars are presented for protein abundances and for canopy-corrected irradiances, since mean canopy openness values are derived from measurements of three individuals. Model fits (OLS regression) are shown where p < 0.05. A full table of univariate OLS regression statistics associated with Fig Xb and Xc is provided in the supplementary materials.

The top row of the lower panel (i, iii, v) shows models fit using protein abundances expressed on a proportional basis; the bottom row (ii, iv, vi,) shows models fit using protein abundances expressed on a per leaf area basis (mg protein / m2 leaf area).

c.) Influence of leaf traits on photosynthetic protein abundance: i.) Neither photosystems nor Calvin cycle enzymes change proportionally in response to total leaf protein, p = , R2 = ; ii.) on a per leaf area basis, abundances of both functional categories strongly track leaf total protein, although there is more variation associated with photosystem protein abundances than Calvin cycle protein abundances. iii.) Proportional abundance of photosystem proteins declines as leaf mass per area (LMA, g/m2) increases (R2 = , p = ), but no such trend is apparent for Calvin cycle proteins (p = , R2 = ); iv.) On a per leaf area basis, abundance of Calvin cycle proteins increases with LMA (R2 =, p = ), while photosystem protein abundance does not change (R2 = , p = ).

d.) Multiple regression models visualised using coloration to indicate the modelled magnitude of protein abundance in two-dimensional environmental space. Curved contours indicate significant interaction effects between predictors. A full table of multiple regression statistics is presented in the supplementary materials.

i.) Proportional abundance of photosystem proteins modelled by MAT and incident irradiance (adjusted multiple R2 (adj.R2) = , p(interaction) = , etc.). The strong univariate relationship between photosystem proportional abundance (PS(prop)) and incident irradiance is modulated by MAT. At warm sites, PS(prop) follows our expectations and is highest at warm high irradiance sites, and lowest at warm low irradiance sites. No relationship between irradiance and PS(prop) is apparent at cool sites, however.

ii.) Per leaf area abundance of photosystems proteins (PS(area)) modelled by MAT and MAP (adjusted multiple R2 (adj.R2) = , p(interaction) = , etc.). MAP and MAT interactively explain variation in PS(area), and a univariate relationship between PS(area) and MAP is only apparent at cool sites. This effect may result from the strong influence of temperature on total leaf protein concentration.

iii.) Proportional abundance of Calvin cycle proteins (Calv(prop))modelled by MAT and incident irradiance (adjusted multiple R2 (adj.R2) = , p(interaction) = , etc.). Calv(prop) is related to MAT and MAP by a simple additive relationship, similar in nature to and likely not independent of MAP/MAT influence on total leaf protein (Fig 1e) .

iv.) Per leaf area abundance of Calvin cycle proteins (Calv(area)) modelled by MAT and MAP (adjusted multiple R2 (adj.R2) = , p(interaction) = , etc.). Calv(area) is interactively related to incident irradiance and MAT: high temperature, high irradiance sites are associated with the highest Calv(area), while at low temperatures, higher irradiance predicts lower Calv(area). Care should be taken in interpretation here, however, as there were few low temperature sites with high incident irradiance to inform the model.