INTRODUCTION

The functioning of leaves is driven by proteins. 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. The capacity of leaves to perform these functions is typically determined by the amount of protein doing the work (ref – Michaelis & Menten?). And yet, interpretation of what vegetation is doing across wide areas and many species continues to rely on measurements of leaf nitrogen content, both in comparative trait ecology (refs) and in models intended to capture contribution of vegetation to world carbon budgets (refs). Leaf nitrogen content is feasible to analyse across many hundreds of species and samples, although it obviously represents a pool of many different proteins with many different functions.

To date, methods for precisely quantifying protein amounts in leaves been too intensive to apply across large numbers of samples, requiring laborious benchtop assays to measure abundance of specific proteins of interest. Alternatively, protein amounts have been estimated indirectly using proxy measurements: chlorophyll can be 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. As such, previous studies of how and why protein amounts vary in leaves have been limited in precision and in scope; see SuppMat for a more complete comment on this matter.

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 sampled from Tasmania, New South Wales, and Queensland, spanning large gradients of mean annual precipitation and temperature (200-3200 mm, 5-27 °C, respectively) (Fig. 1). The resulting dataset describes protein abundances at all levels of functional organisation, from broad groupings down to individual protein subunits. We provide a complete description of leaf protein allocation for all major protein functional categories, and describe the dominant biogeographical patterns in abundance of photosynthesis proteins.

RESULTS

*Protein composition of the average eucalypt leaf.*

Our mass spectrometric approach detected X individual proteins per sample, on average. These proteins accounted for 99.9% of sample mass, among which the top 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.

The majority (64%, SD X%) of leaf protein was associated with photosynthesis: 36% was associated with the carbon fixing Calvin Cycle, 22% (SD X%) with the light reactions 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 was the second most abundant top-level category at X% (SD X%) (Fig 2a). [need to decide here whether we’re saying that in the large, this mixture is in line with previous intensive results from few species, versus it’s really different, vide previous discussions with steve about uws results]

Although protein allocation is known to vary between species, our numbers for the major categories are not dissimilar to those obtained using classical biochemical assays on domesticated species such as *Spinacea oleracea*, *Phaseolus vulgaris*, and *Cucumis sativus* (Evans & Seeman 1989).

*Linking leaf protein abundances with environment and functional traits*

SENTENCE INTRODUCING HEATMAP, say we’ve got all the major protein functional categories plus some smaller categories of interest.

SAY SOMETHING ABOUT PER LEAF AREA VS PROPORTIONAL AMOUNTS

Per leaf area abundances of all major protein functional categories were correlated with each other, as well as with leaf nitrogen per area (N\_area), leaf mass per area (LMA), and maximum photosynthetic rate (Amax). Proportional abundances of protein functional categories (indicating investment in a defined function relative to investment in all other functions) were less clearly or consistently correlated.

*b.) first scatterplot panel*

**LIGHT**

We expected that allocation to photosystem complex proteins would be greatest where photosynthesis is light-limited (Niinemets 2007), and that investment in Calvin cycle enzymes would increase with light availability, since capacity for carboxylation of the small sugar molecule RuBP determines the rate of light-saturated photosynthesis (Farquhar et al. 1980).

Both PSarea and PSfrac showed a pronounced decline with increasing incident irradiance (Fig. 3b-v, X% per Y irradiance; Fig 3blah stat). No increase CCarea was observed with greater light availability, but CCfrac did increase marginally (Fig, %, stat); this fractional increase may be simple outcome of decreasing PSfrac, however.

Per leaf area photosystem protein abundance declined substantially with increasing MAT (Fig. 3b-i) and was also strongly correlated with total leaf protein abundance (Pearson’s r = 0.82).

**MAT**

Leaf protein content is known be higher in cool environments, to compensate for lower enzyme activity at lower temperatures (Raven & Geider 1988). This effect has been observed for Rubisco in a number of studies (summarised by Hikosaka et al 2006). MAYBE LESS OF A THING FOR PS DUE TO WEAKER TEMPERATURE DEPENDENCY OF LIGHT CAPTURE (OLD VERSION OF DOC?)

Both CCarea and PSarea declined notably with increasing MAT (stat, Fig. 3b-I, stat Fig XX % blah). Neither CCfrac nor PSfrac changed significantly over temperature gradients (Fig stat %).

**MAP**

We also expected that investment in Calvin cycle enzymes would increase towards drier sites. By effecting greater internal CO2 drawdown, rate of CO2 uptake can be maintained at lower stomatal conductance, reducing the water cost of photosynthesis for dryland plants (Wright et al. 2001a,b, Scalon & Wright 2017). No direct effect of precipitation on investment in photosystem proteins was expected, although cross-correlation between precipitation and vegetation canopy density could influence this relationship.

PSarea declined strongly with increasing MAP (Fig x stat %). No response of CCarea over the MAP gradient was observed (Fig. 3b-iii). Since MAP and incident irradiance were negatively correlated (i.e. denser canopies at wetter sites, Pearson’s r = -0.59) the lack of CCarea response to MAP could be explained by changing light conditions.

Proportional allocation of protein resources to Calvin cycle protein did not adjust over gradients of MAP (Fig. 3b-ii) but increased marginally (stat) with increasing incident radiation (Fig. 3b-vi). Proportional photosystem protein abundance increased with increasing MAP (Fig. 3b-iv) and decreased with increasing incident irradiation to a similar extent as the per leaf area measure (Fig . BLAH). This latter response may explain the observed decline in Calvin cycle proteins as incident irradiance increased. The range of interspecific variation in photosystem protein proportional abundance (0.09-0.23, 2.6-fold) was considerably higher than for Calvin cycle proteins (0.30-0.39, 1.3-fold). These observations provide robust evidence that eucalypt leaves specifically optimise protein allocation to light capture in response to environmental conditions (some stats and numbers), while adjustment of carboxylation capacity is largely achieved through bulk changes in per leaf area protein content.

It is worth noting at this point that CCarea was highly correlated with the total abundance of protein per area (Pearson’s r = 0.97), and environmental trends in Calvin cycle protein abundance were essentially identical to trends in leaf protein abundance (see sup info).

*c.) second scatterplot panel*

One obvious way Calvin Cycle protein per leaf area can change is via changes in depth of mesophyll and of leaf, and indeed adjustments in per leaf area Calvin cycle protein abundance occurred to some extent via changes in leaf mass per area (LMA) (Fig. 3c-i). The substantial scatter around the Calvin cycle – LMA relationship indicates that LMA responded to other requirements in addition to carboxylation capacity. Photosystem abundance did not increase per leaf area with increasing LMA (Fig. 3c-ii) and declined as a proportion of total leaf protein. Light harvesting capacity thus appears to be optimised for a given leaf area independently from leaf thickness.

Leaf nitrogen per area was a strong predictor of both Calvin cycle and photosystem protein abundance per leaf area, and no relative changes in these protein categories occurred with increasing nitrogen per area.

*d.) protein abundance/concentration/LMA multiple regressions*

We hypothesised that Calvin cycle protein abundance would be driven by temperature dependence of enzyme kinetics, and that maximisation of CO2 drawdown at low stomatal conductance in water-limited environments. Fig 3d-i shows that these demands were in fact complementary: leaves sampled at cold dry sites required the most protein, while leaves from warm wet sites experienced neither constraint, having both low protein content per area and low LMA.

The role of LMA versus protein concentration (Calvin cycle protein as a fraction of leaf dry mass) in determining per leaf area protein abundance depended interactively on MAP and MAT (Fig 3d-ii,iii). Low per leaf area Calvin cycle protein abundance at warm, wet sites was more closely associated with low LMA than low protein concentration, while high per leaf area Calvin cycle protein abundance at cold, dry sites was strongly associated with high Calvin cycle protein concentration.

Plants are able to build cheaper leaves at warm wet sites, where photosynthetic reaction kinetics are increased and plants are not water limited.