Plant families exhibit unique geographic trends in C₄ richness and cover

Samantha E.M. Munroe*,1,2, Francesca A. McInerney³, Greg R. Guerin¹,2, Jake W. Andrae³, Nina Welti⁴, Stefan Caddy-Retalic¹,5, Rachel Atkins³, & Ben Sparrow¹,2

¹ School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

² Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide, South Australia 5005, Australia

³ School of Physical Sciences and the Sprigg Geobiology Centre, The University of Adelaide, Adelaide, South Australia 5005, Australia

⁴ CSIRO Agriculture and Food, Urrbrae, South Australia 5064, Australia

⁵ School of Life and Environmental Sciences, University of Sydney, Sydney NSW 2006 Australia

*corresponding author: samantha.munroe@adelaide.edu.au
Abstract
Numerous studies have analysed the relationship between C4 plant cover and climate. However, few have examined how different C4 taxa vary in their response to climate, or how environmental factors alter C4:C3 abundance. Here we investigate (a) how proportional C4 plant cover and richness (relative to C3) responds to changes in climate and local environmental factors, and (b) if this response is consistent among families. Proportional cover and richness of C4 species were determined at 541 one-hectare plots across Australia for 14 families. C4 cover and richness of the most common and abundant families were regressed against climate and local parameters. C4 richness and cover in the monocot families Poaceae and Cyperaceae increased with latitude and were strongly positively correlated with January temperatures, however C4 Cyperaceae occupied a more restricted temperature range. Seasonal rainfall, soil pH, soil texture, and tree cover modified proportional C4 cover in both families. Eudicot families displayed considerable variation in C4 distribution patterns. Proportional C4 Euphorbiaceae richness and cover were negatively correlated with increased moisture availability (i.e. high rainfall and low aridity), indicating they were more common in dry environments. Proportional C4 Chenopodiaceae richness and cover were weakly correlated with climate and local environmental factors, including soil texture. However, the explanatory power of C4 Chenopodiaceae models were poor, suggesting none of the factors considered in this study strongly influenced Chenopodiaceae distribution. Proportional C4 richness and cover in Aizoaceae, Amaranthaceae, and Portulacaceae increased with latitude, suggesting C4 cover and richness in these families increased with temperature and summer rainfall, but sample size was insufficient for regression analysis. Results demonstrate the unique relationships between different C4 taxa and climate, and the significant modifying effects of environmental factors on C4 distribution. Our work also revealed C4 families will not exhibit similar responses to local perturbations or climate.
Introduction

C₃ and C₄ plants have distinct geographical distributions in large part due to differences in anatomy, physiology, and biochemistry [1]. Under conditions such as high heat or hypersalinity, C₃ plants experience high levels of photorespiration [oxygen fixation by rubisco; 2, 3, 4]. Consequently, C₃ plants may not thrive in hot or dry environments. C₄ photosynthesis is distinguished by a series of biochemical and anatomical adaptations that concentrate and isolate CO₂ with rubisco, helping to eliminate photorespiration and increase nutrient and water-use efficiency [5, 6]. As a result, C₄ plants can dominate warm environments and are essential to the development of tropical and subtropical savannas, grasslands, and shrublands [5, 7].

Broad patterns in C₃ and C₄ cover are predominantly estimated using temperature and precipitation models [8, 9]. However, the relative importance of different environmental factors is still debated. As a result, numerous approaches have been developed to estimate C₄ distribution. The leading method used to predict C₃:C₄ grass cover is the physiological temperature crossover model [10, 11]. This model incorporates the antagonistic effects of increasing temperature and rising CO₂ levels on C₃ and C₄ plants. Increases in CO₂ levels are expected to favour C₃ species, potentially leading to increased C₃ cover [12-14], while rising temperatures should favour C₄ species [15]. The crossover approach predicts that at modern CO₂ levels, the photosynthetic yields (i.e. light use efficiencies) of C₄ grasses surpasses those of C₃ grasses at a crossover temperature of approximately 22°C [10, 11, 16]. However, more recent research has challenged the underlying physiological assumptions of this approach, arguing that CO₂ levels have a reduced influence relative to other environmental factors [15, 17]. Alternative models to calculate C₄ cover include summer maximum temperatures [18] and seasonal rainfall patterns [19-21]. The method deemed to be the most
accurate often differs between regions and the types of data used to develop the model. Therefore, there remains uncertainty on how to best predict C₄ abundance, and how C₄ and C₃ species respond to environmental change.

Most C₄ cover–climate models are designed to estimate C₄ grass distribution. Few investigations have considered how additional C₄ taxa vary in their particular relationship to climate, a critical feature in anticipating responses to environmental change. Although most C₄ plants are indeed grasses (4500 species), C₄ lineages are also found among sedges (1500 species) and eudicots [1200 species; 5]. Work that is available has revealed distinct C₄ taxa exhibit disparate geographical trends [22, 23]. For example, Wang and Ma (24) used flora survey data to examine C₄ distribution in China and found that while C₄ grass occurrence was primarily associated with changes in precipitation, the distributions of C₄ chenopods and sedges were associated with changes in aridity and temperature, respectively. These studies demonstrate different models and approaches should be developed and applied across distinct lineages to provide the most accurate estimates of vegetation abundance.

Methods to predict C₄ species distribution are based on the assumption that species occurrence and cover are shaped by their physiological responses to climate. However, local environmental factors can significantly modify C₄ distribution, influence local competitive dynamics, and inhibit or maintain C₃ and C₄ coexistence [25-27]. For example, Griffith, Anderson (28) determined the proportional cover of C₄ grass in over 40,000 vegetation plots in North America. The best models to describe C₃:C₄ grass distribution included temperature, and precipitation, but also local environmental factors such as tree cover and fire. Local factors, such as soil nitrogen supply, have also been shown to reverse the responses of C₃ and
C₄ plants to atmospheric CO₂ concentrations [29]. Investigating the combined effects of climate and local factors on C₄ distribution at large-scales is essential to accurately predicting C₄ abundance now and into the future.

Large-scale studies of C₄ occurrence and cover that consider both climate and other environmental factors have been limited due to a lack of data. Since 2009, the Terrestrial Ecosystem Research Network (TERN) has collected plot-based environmental data across all major biomes and dryland habitats in Australia. Using a consistent point-intercept approach, TERN systematically surveys one-hectare plots to determine the relative cover of plant taxa. In addition, the TERN plant sample library can be used to identify the photosynthetic pathway of unassigned species via stable carbon isotope analysis [30, 31]. By combining these two novel resources, TERN’s monitoring program provides a rare opportunity to investigate the proportional cover and richness of different C₄ lineages at a large scale. This work is particularly valuable for Australia, where only a few studies have examined the distribution of C₄ grasses and sedges [32-35], and no studies have surveyed the distribution of C₄ plants in other taxa.

In this study, we used TERN vegetation surveys to compare proportional vegetation cover and richness of C₄ species in 14 plant families in Australia. We statistically assessed correlations between proportional C₄ richness and cover with climate and edaphic variables in the four most common and abundant families: Poaceae, Cyperaceae, Chenopodiaceae and Euphorbiaceae. Specifically, we asked how proportional C₄ plant richness and cover (relative to C₃) responds to changes in climate and local environmental factors, and if this response was consistent among families. The analysis presented here sheds new light on how C₄
occurrence and cover responds to particular climate variables, how local environmental factors modify C₄ abundance, and whether responses are consistent among broad taxonomic groups. This work will inform improved strategies to estimate C₄ abundance, and increase understanding of the impacts local conditions have on different C₄ taxa.

**Materials and Methods**

**Data collection**

Proportional C₄ vegetation richness and cover (relative to combined C₃ and C₄ cover and richness) for each family were determined at 541 one-hectare plots systemically surveyed by TERN between 2011 and 2017. TERN plots are not surveyed on an annual basis. Instead, plots have typically been surveyed once, with an intention to revisit at least once per decade. Surveys occur between February to November. The season of survey is chosen to be the most appropriate for a given system to capture species diversity and provide the best chance of accurate species identification. Most plots are located within the Australian rangelands, which are characterised by weathered features, old and typically infertile soils, highly variable rainfall, and variable plant communities. Plot locations are selected to avoid major anthropogenic influences (such as roads, cattle yards, fences, bores, etc.) or sources of anthropogenic disturbance (e.g. livestock grazing). Surveyed vegetation types include woodlands and savannas, tussock and hummock grasslands, and shrublands.

Plots (1 ha, 100 x 100 m) are permanently established sites located in a homogenous area of terrestrial vegetation. Transects (10 x 100 m long) were laid out within each plot in a 5 x 5 grid pattern. Parallel transects were spaced 20 meters apart. Species were recorded at each point (1 m) within the transect, resulting in 1010 survey points per plot. Point-intercept data
were used to calculate proportional C$_4$ vegetation cover and presence/absence data were used to calculate proportional C$_4$ species richness (see statistical analysis). Ground observers vouchered all species present within each plot and vouchers were identified at major herbaria across Australia. Full survey protocols are detailed in the TERN Ausplots Rangelands manual [36, 37].

**Statistical Analysis**

Data were analysed in the R statistical environment [38]. Point-intercept data were imported using the ‘ausplotsR’ package [39, 40]. Species cover (%) was calculated from point-intercepts at each TERN plot using the *species_table* function. The *species_table* function calculates species cover (%) as the total number of times a species was detected or ‘counted’ across all transects, divided by the total number of transect points in the plot (1010), and multiplied by 100. If more than one survey was available, we calculated the average species cover for each plot. Species were assigned a photosynthetic pathway using [41]. The cover values of species with the same pathway were summed to determine the total C$_4$ and C$_3$ plant cover in each family at each plot. Finally, proportional C$_4$ cover was calculated at each plot for each family as a proportion of C$_3$ and C$_4$ species cover by:

**Eq 2** \[ \text{Proportional C}_4 \text{ cover} = \frac{\text{C}_4 \text{ species cover}}{\text{C}_4 \text{ species cover} + \text{C}_3 \text{ species cover}} \]

Plot C$_4$ and C$_3$ richness was calculated as the number of unique C$_4$ and C$_3$ taxa at the species level that were detected across all transects in a plot. Proportional C$_4$ richness for each family was calculated as a proportion of the combined species richness of C$_3$ and C$_4$ species by:

**Eq 3** \[ \text{Proportional C}_4 \text{ species richness} = \frac{\text{C}_4 \text{ richness}}{\text{C}_4 \text{ richness} + \text{C}_3 \text{ richness}} \]
We decided to examine C₄ richness and cover as a proportion of total C₃ and C₄ species richness and cover to prevent our analysis from being influenced by factors that affect plant abundance. Much of Australia has a highly arid climate with limited plant cover, and these habitats can’t be easily directly compared to temperate or wet-tropical regions.

To investigate the unique cover patterns of C₄ species in different plant families, proportional C₄ richness and cover were regressed against climatic and local environmental parameters that are considered potential drivers of C₄ plant distribution [5, 28, 42]. Climate variables included mean annual temperature (MAT), mean annual precipitation (MAP), mean January minimum and maximum temperature [18, 33], January precipitation, mean annual aridity index (precipitation/potential evaporation), C₄ growing season water availability [see description below; 19, 32], and a variable generated to represent the Collatz, Berry (11) crossover temperature model (see below). Local edaphic variables included soil sand and clay content (%), pH, and available water capacity (the amount of water soil can store that is available to plants, AWC). Edaphic values were averaged over a soil depth of 0 to 15 cm. We also included tree cover (%), which was calculated as the total species cover (%) of trees equal to or taller than 5 m at each plot. Climate data were based on 1970–2018 records and were sourced from the Australian Gridded Climate Data set (Bureau of Meteorology, accessed through the TERN Data Discovery Portal). The aridity index used data spanning 1976–2005 and was sourced from Harwood, Donohue (43). Soil variables were accessed from the Soil and Landscape Grid of Australia [44].

C₄ growing season water availability (hereafter, seasonal water availability or SWA) was calculated according to Murphy and Bowman (32). This approach determines the proportion
of precipitation that occurs during C\textsubscript{3} and C\textsubscript{4} growing seasons as defined by temperature and was calculated by:

\begin{equation}
\text{Eq 4} \quad \text{SWA} = \frac{\text{Precipitation in } C_4 \text{ growing season}}{(\text{Precipitation in } C_4 \text{ growing season} + \text{precipitation in } C_3 \text{ growing season})}
\end{equation}

To apply the Collatz, Berry (11) crossover temperature approach as consistently as possible, we regressed the mean annual proportion of C\textsubscript{4} favoured months rather than the mean absolute number of C\textsubscript{4} favoured months. A given month was determined to favour C\textsubscript{3} growth when the mean daytime temperature was \(\leq 22\, ^\circ\text{C}\) and precipitation was \(\geq 25\, \text{mm}\). A given month was determined to favour C\textsubscript{4} growth when the mean daytime temperature was \(> 22\, ^\circ\text{C}\) and precipitation was \(\geq 25\, \text{mm}\). As previously mentioned, most of Australia has a highly arid climate and large areas of the country receive < 25 mm of precipitation per month. Therefore, comparing the absolute number of C\textsubscript{4} favoured months in each plot would have confounded comparisons between dry and wet habitats.

To investigate the partial effects of different climate and local environmental variables, we considered several possible parametric and non-parametric approaches. Because our data are proportional (i.e. range from 0 to 1 and included true values of 0 and 1) and were derived from discrete counts (i.e. the number of C\textsubscript{4} plants out of the total number of plants in the transect), a logistic regression with a weighted response variable was determined an appropriate method to explore the relationship between proportional C\textsubscript{4} plot values and climate and local variables [45]. Models were constructed using the \textit{glm} function. Strong covariance among variables or the inclusion of too many variables can lead to overfitting [46]. Therefore, models were limited to variables that had Pearson pairwise correlations < 0.7.
and a maximum of five predictors. Although we acknowledge interactions are probable in these complex systems, preliminary work showed that our sample size was not large enough to support interactions. Models were compared using a backwards, step-wise comparison process between each model using quasi-Akaike information criterion (QAIC) to account for over-dispersion. Spatial autocorrelation was tested using Moran’s I [47]. In cases where we identified significant (p<0.05) spatial autocorrelation, a spatial autocovariate term was included as a fixed covariate in each model. Autocorrelation was detected in Poaceae and Chenopodiaceae regressions, therefore spatial autocovariate terms were included as fixed covariates in Poaceae and Chenopodiaceae models. The spatial autocovariate terms were calculated as the distance-weighted mean of neighbouring proportional C4 values. The terms were calculated using the R package “spdep” [48]. A pseudo- $R^2$ (McFadden’s $R^2$) was used as a measure of explained variation [49]. To determine the relative importance of each predictor in the best fit models, each predictor was removed from the best fit model (i.e. a leave-one-out approach) to compare the change in QAIC to the full model. Models were visualised using the “visreg” package in R [50].

**Results**

**Proportional C4 cover and richness**

Using Munroe, McInerney (41), we determined the photosynthetic pathway of 2484 of the 2605 species identified within TERN plots. Most unassigned species were rarely encountered (i.e. only recorded at one point in a single plot) and thus had a limited effect on analysis. Of the 2484 assigned species, 347 (13.9%) were C4 species and 2101 were C3 (84.5%). The remaining species were C3-CAM (18), CAM (7), C3-C4 (7), and C4-CAM (4). C4 species were distributed amongst 14 families and 85 genera (SI Appendix 1). To enable consistent
comparisons with previous work, we evaluated Chenopodiaceae independent of Amaranthaceae [e.g. 24, 51]. C₃-CAM, CAM, C₃-C₄, and C₄-CAM species were excluded from statistical analysis.

C₄ plants were detected in 451 plots (85.7%). Proportional C₄ richness ranged from 0 to 88% (24.8% ±19.1), where 81% of plots had ≥ 5% C₄ richness, and 16% of plots had ≥ 50% C₄ richness. Proportional C₄ cover ranged from 0 to 98% (36.5% ± 30.7%, mean ± standard deviation). 72% of TERN plots had ≥ 5% proportional C₄ cover and 35% of plots had ≥ 50% proportional C₄ cover. We calculated proportional C₄ richness and cover at TERN plots for all 14 families in which C₄ species were identified (SI Appendix 2). Proportional C₄ richness and cover in Aizoaceae, Amaranthaceae, Cyperaceae, Poaceae, and Portulacaceae increased along a south to north trajectory. Although not tested for directly, these trends indicate C₄ cover in these families increased with increases in temperature, C₄ growing season, and summer rainfall (Figure 1; Supplemental Material 3). Proportional C₄ cover and richness in Boraginaceae and Zygophyllaceae was also typically higher in northern Australia. However, C₃ species in these families were also common in some northern plots. The remaining eudicot families exhibited no obvious geographical or climate-based trends in proportional C₄ distribution. C₃ Asteraceae species were found distributed across TERN plots, with only one C₄ species being detected in western Australia in two plots. Both C₃ and C₄ Cleomaceae were found in northern subtropical plots. C₄ Euphorbiaceae and Chenopodiaceae were found in a wide range of climates, including the relatively cool, temperate areas of southern Australia. There were sufficient data for logistic regression analysis of Poaceae, Cyperaceae, Chenopodiaceae, and Euphorbiaceae. For all other families, species were detected in too few plots (< 50) or had too little cover (< 1%) to conduct effective quantitative analysis.
Poaceae: Mean proportional C₄ Poaceae richness was 82.8% (± 34.0) and mean proportional C₄ Poaceae cover was 83.9% (± 34.9%). Most plots containing Poaceae were characterised by either 0 or 100% proportional C₄ Poaceae richness and cover (Figure 2,3, SI Appendix 3). Mean January maximum temperature, January precipitation, tree cover, and sand (%) were the best predictors of proportional C₄ Poaceae richness (Table 1). Proportional C₄ Poaceae richness increased with January maximum temperature (Figure 2). Higher January precipitation and low tree cover were correlated with comparatively small increases in proportional C₄ Poaceae richness. Sand (%) also had minor modifying effect on proportional C₄ Poaceae richness, where increased sand (%) were associated with decreased proportional richness. QAIC comparisons using our leave-one-out approach indicated January maximum temperature was the most important predictor of C₄ grass richness trends, while tree cover was of least importance in the model (SI Appendix 4).
Table 1. Binomial logistic regression model selection results of proportional C\textsubscript{4} richness for Poaceae, Cyperaceae, Chenopodiaceae, and Euphorbiaceae. Models were ranked using QAIC values, ΔQAIC is the difference between the model’s QAIC and the lowest QAIC of the candidate set. Only models with a ΔQAIC < 2 are shown. The best fit models are marked by an asterisk (*). R\textsuperscript{2} is the McFadden’s R\textsuperscript{2} value of each model. Spatial autocovariate terms were included as fixed covariates in the Poaceae and Chenopodiaceae models. Predictor variables are mean annual temperature (MAT), mean annual precipitation (MAP), mean January minimum temperature (Jan. Min), mean January maximum temperature (Jan. Max), January precipitation (Jan. Precip), mean annual aridity index (Aridity), season water availability (SWA), mean annual proportion of C\textsubscript{4} favoured months (Collatz) soil sand and clay content (% sand, clay), pH, and available water capacity (AWC).

<table>
<thead>
<tr>
<th>Family</th>
<th>Model</th>
<th>QAIC</th>
<th>ΔQAIC</th>
<th>R\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaceae</td>
<td>Jan. Max + Jan. Precip + sand + Tree Cover</td>
<td>448.1</td>
<td>0.00</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Jan. Max + Jan. Precip + sand + Tree Cover + AWC</td>
<td>448.4</td>
<td>0.26</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Jan. Max + Jan. Precip + sand + Tree Cover + pH</td>
<td>450.0</td>
<td>1.85</td>
<td>0.67</td>
</tr>
<tr>
<td>Cyperaceae</td>
<td>Collatz + pH + Tree Cover*</td>
<td>100.4</td>
<td>0.00</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Collatz + pH + Tree Cover + clay</td>
<td>101.1</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Collatz + pH + Tree Cover + sand</td>
<td>101.4</td>
<td>1.01</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Collatz + pH + Tree Cover + Jan. Precip</td>
<td>102.3</td>
<td>1.91</td>
<td>0.70</td>
</tr>
<tr>
<td>Chenopodiaceae</td>
<td>Jan. precip + sand*</td>
<td>484.5</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + AWC</td>
<td>486.3</td>
<td>1.78</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + MAT</td>
<td>486.3</td>
<td>1.81</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + pH</td>
<td>486.5</td>
<td>1.95</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + Jan. Min.</td>
<td>486.5</td>
<td>1.96</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + SWA</td>
<td>486.5</td>
<td>1.97</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + Jan. Max</td>
<td>486.5</td>
<td>1.98</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + Tree Cover</td>
<td>486.5</td>
<td>1.98</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Jan. precip + sand + Collatz</td>
<td>486.5</td>
<td>2.00</td>
<td>0.09</td>
</tr>
<tr>
<td>Euphorbiaceae</td>
<td>AWC + MAP + SWA</td>
<td>181.2</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + Jan. Min.</td>
<td>119.3</td>
<td>0.00</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + MAT</td>
<td>120.2</td>
<td>0.93</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + Collatz</td>
<td>120.4</td>
<td>1.13</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + Jan. Min. + pH</td>
<td>120.5</td>
<td>1.27</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + pH</td>
<td>120.6</td>
<td>1.39</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + MAT + pH</td>
<td>120.9</td>
<td>1.61</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + Jan. Max. + SWA</td>
<td>121.0</td>
<td>1.76</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + clay + Jan. Min. + pH</td>
<td>121.1</td>
<td>1.80</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + clay + SWA</td>
<td>121.2</td>
<td>1.95</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP</td>
<td>121.2</td>
<td>1.96</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>AWC + MAP + sand + SWA</td>
<td>121.3</td>
<td>1.99</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Mean January maximum temperature, January precipitation, AWC, and tree cover were the best predictors of proportional C\textsubscript{4} Poaceae cover (Table 2; SI Appendix 4). January maximum temperature was positively correlated with increases in proportional C\textsubscript{4} Poaceae cover (Figure 3). Low tree cover, increased January precipitation, and lower AWC was correlated with small increases in proportional C\textsubscript{4} Poaceae cover. However, only January maximum
temperature, January precipitation, and tree cover were consistently included in models with
a QAIC<2; the removal of AWC had little effect on the overall R² and QAIC values (SI
Appendix 4). QAIC comparisons using our leave-one-out approach also indicated January
maximum temperature was the most important predictor of C₄ grass cover trends (SI
Appendix 4). Both the proportional C₄ grass cover and richness models predicted indicated
that below 30°C January maximum temperature, C₃ grasses dominated survey plots (>80%
C₃), while at plots with January maximum temperatures above 33°C, C₄ grasses dominated
(>80% C₄). Temperatures between 30°C and 33°C supported mixed C₄/C₃ grass plots. There
was a significant positive correlation between proportional C₄ Poaceae richness and cover
(Spearman’s rank correlation, ρ=0.81, P<0.05, Supplemental Material 5). Visual inspection
of partial residuals indicated we achieved a better model fit for proportional C₄ Poaceae
richness than cover.
Table 2. Logistic regression model results of proportional C₄ cover for Poaceae, Cyperaceae, Chenopodiaceae, and Euphorbiaceae. Models were ranked using QAIC values, ΔQAIC is the difference between the model’s QAIC and the lowest QAIC of the candidate set. Only models with a ΔQAIC <2 are shown. The best fit models are marked by an asterisk (*). R² is the McFadden’s R² value of each model. Spatial autocovariate terms were included as fixed covariates in the Poaceae and Chenopodiaceae models. Predictor variables are mean annual temperature (MAT), mean annual precipitation (MAP), mean January minimum temperature (Jan. Min), mean January maximum temperature (Jan. Min), January precipitation (Jan. Precip), mean annual aridity index (Aridity), season water availability (SWA), mean annual proportion of C₄ favoured months (Collatz) soil sand and clay content (% sand, clay), pH, and available water capacity (AWC).

<table>
<thead>
<tr>
<th>Family</th>
<th>Model</th>
<th>QAIC</th>
<th>ΔQAIC</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaceae</td>
<td>AWC + Jan. Max + Jan. Precip + Tree cover*</td>
<td>483.8</td>
<td>0.0</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Jan. Max + Jan. Precip + Tree cover</td>
<td>484.7</td>
<td>0.9</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+ Jan. Max + Jan. Precip + Tree cover+sand</td>
<td>484.9</td>
<td>1.1</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+Jan. Max + Jan. Precip + Tree cover+MAP</td>
<td>485.1</td>
<td>1.3</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+Jan. Max + Jan. Precip + Tree cover + pH</td>
<td>485.2</td>
<td>1.3</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+MAT+ Jan. Max + Jan. Precip + Tree cover</td>
<td>485.6</td>
<td>1.7</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+Aridity + Jan. Max + Jan. Precip + Tree cover</td>
<td>485.7</td>
<td>1.8</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>AWC+clay+ Jan. Max + Jan. Precip + Tree cover</td>
<td>485.7</td>
<td>1.9</td>
<td>0.68</td>
</tr>
<tr>
<td>Cyperaceae</td>
<td>Jan. Min. + pH + Tree Cover+ Jan. Precip*</td>
<td>210.4</td>
<td>0.00</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Jan. Min. + pH + Tree Cover+ Jan. Precip + sand</td>
<td>211.9</td>
<td>1.5</td>
<td>0.85</td>
</tr>
<tr>
<td>Chenopodiaceae</td>
<td>Jan. Max + pH + sand + Tree cover *</td>
<td>322.0</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>pH + sand + Tree cover</td>
<td>322.6</td>
<td>0.58</td>
<td>0.14</td>
</tr>
<tr>
<td>Euphorbiaceae</td>
<td>AWC + Aridity + Jan. Max + sand+ SWA</td>
<td>181.2</td>
<td>0.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

**Cyperaceae:** Mean proportional C₄ Cyperaceae richness was 52.8% (± 48.6) and mean proportional C₄ Cyperaceae cover was 57.6% (± 48.2 %). Most plots containing Cyperaceae were defined by bimodal values of 0 or 100% proportional C₄ Cyperaceae cover and richness (Figure 4,5). The mean annual proportion of C₄ favoured months [11], tree cover, and pH were included in the best fit model to predict predictors of proportional C₄ Cyperaceae richness. Proportional C₄ Cyperaceae richness was positively correlated with an increased proportion of C₄ favoured months. Higher C₄ Cyperaceae proportional richness was positively correlated with low tree cover and higher pH values (less acidic; Figure 4). QAIC comparisons using a leave-one-out approach indicated the proportion of C₄ favoured months was the primary predictor of C₄ Cyperaceae richness trends (SI Appendix 4).
Mean January minimum temperature, January precipitation, tree cover and pH were included in the best fit model to predict proportional C\textsubscript{4} Cyperaceae cover. Proportional C\textsubscript{4} Cyperaceae cover was positively correlated with January minimum temperature and January precipitation (Figure 5). Low tree cover and higher pH were also positively correlated with increases in proportional C\textsubscript{4} Cyperaceae cover. QAIC comparisons using a leave-one-out approach indicated January minimum temperature was the primary predictor of C\textsubscript{4} Cyperaceae cover trends, followed by \% tree cover (SI Appendix 4). The proportional C\textsubscript{4} Cyperaceae cover model predicted that below 20°C January minimum temperature, C\textsubscript{3} Cyperaceae dominated survey plots (>80\% cover), while at plots with January minimum temperatures above 23°C, C\textsubscript{4} Cyperaceae dominated (>80\% cover). Temperatures between 20°C and 23°C supported mixed C\textsubscript{4}/C\textsubscript{3} Cyperaceae plots. There was a positive correlation between proportional C\textsubscript{4} Cyperaceae richness and cover (Spearman’s rank correlation, \(\rho=0.99\), \(P<0.05\); Supplemental Material 5).

\textit{Chenopodiaceae}: Mean proportional C\textsubscript{4} Chenopodiaceae richness was 15.8\% (± 22.1) and mean proportional C\textsubscript{4} Chenopodiaceae richness was cover was 19.0\% (± 29.5). Proportional C\textsubscript{4} Chenopodiaceae richness and cover was generally low (Figure 6,7). Plots that included Chenopodiaceae commonly contained a mix of C\textsubscript{3} and C\textsubscript{4} species. January precipitation and sand (\%) were included in the best fit model to predict proportional C\textsubscript{4} Chenopodiaceae richness (Table 1; Figure 6) but the model had low explanatory power and poor model fit. Lower sand content and January precipitation were weakly correlated with higher proportional C\textsubscript{4} Chenopodiaceae richness (SI Appendix 4). Mean January maximum temperature, tree cover, pH, and sand (\%) were included in the best fit model to predict proportional C\textsubscript{4} Chenopodiaceae cover (Table 2; Figure 7). However, examination of partial model residuals and \(R^2\) values indicates that even the best fit model performed poorly and
had little explanatory power. Increased proportional C\textsubscript{4} Chenopodiaceae cover was only weakly associated with decreasing tree cover, lower soil pH and sand content (SI Appendix 4). There was a positive correlation between proportional C\textsubscript{4} Chenopodiaceae richness and cover (Spearman’s rank correlation, \(\rho=0.92, P<0.05\); Supplemental Material 5).

\textit{Euphorbiaceae}: Mean proportional C\textsubscript{4} Euphorbiaceae richness was 69.1\% (± 44.8\%) and mean proportional C\textsubscript{4} Euphorbiaceae cover was 68.5\% (± 45.3\%). The total number of Euphorbiaceae species per plot was low (≤ 2), and they accounted for a small amount of the total vegetation cover at each plot (<1-20\%). MAP, SWA, and AWC were included in the best fit model to predict proportional C\textsubscript{4} Euphorbiaceae richness (Table 1). Proportional C\textsubscript{4} Euphorbiaceae richness was negatively correlated with MAP, SWA and AWC (Figure 8), however the model fit and explanatory power of the C\textsubscript{4} Euphorbiaceae richness logistic regression was relatively poor. SWA, sand (%), aridity, mean January maximum temperature, and AWC were included in the best fit model to predict proportional C\textsubscript{4} Euphorbiaceae cover. Increased January maximum temperature, higher aridity index values (i.e. wetter conditions), higher AWC, and high soil sand content (%) were negatively correlated with proportional C\textsubscript{4} Euphorbiaceae cover (Figure 9), while increased SWA was positively correlated with proportional C\textsubscript{4} Euphorbiaceae cover. Leave-one-out QAIC comparisons indicated aridity and SWA were the most important predictors of proportional C\textsubscript{4} Euphorbiaceae cover (SI Appendix 4). There was a positive correlation between proportional C\textsubscript{4} Euphorbiaceae richness and cover (Spearman’s rank correlation, \(\rho=0.99, P<0.05\); Supplemental Material 5).

**Discussion**

Our analysis of proportional C\textsubscript{4} richness and cover revealed different families exhibit divergent responses to both climate and local environmental conditions. Although
temperature was a key driver of proportional C$_4$ monocot distribution, local environmental factors also had significant modifying effects on C$_4$ richness and cover. Eudicot families displayed unique and sometimes contrasting C$_4$ distribution patterns. While proportional C$_4$ cover and richness in families such as Aizoaceae and Portulacaceae increased with latitude, suggesting there was a strong relationship with C$_4$ in these families and temperature, climate had a limited apparent influence on proportional C$_4$ Chenopodiaceae distribution. Cumulatively, our work demonstrates different C$_4$ lineages will display diverse responses to climate, as well as local environmental variation, and that climate patterns alone cannot explain trends in C$_4$ distribution.

Proportional C$_4$ Poaceae cover and richness increased with increases in January maximum temperatures and summer rainfall. Our results are congruent with other studies from Australia and around the world which have similarly concluded summer temperatures are strongly correlated with absolute and proportional C$_4$ grass cover and richness [18, 33, 35, 52-54]. Our results are also consistent with von Fischer, Tieszen (18), who also found summer rainfall had a significant relationship with North American C$_4$ biomass. Angelo and Daehler (54) also found precipitation had a weak but significant relationship with C$_4$ grass cover along tropical elevation gradients in Hawaii. Interestingly, the adaptation of the Collatz, Berry (11) crossover approach was not the best model to explain C$_4$ distribution in Poaceae. More importantly, it was difficult to apply the crossover approach consistently across Australia. Dry regions of Australia often receive less than the 25 mm of precipitation per month needed to assign plots a C$_3$ or C$_4$-dominate status. As a result, for large areas of Australia, a traditional crossover approach was not feasible. This demonstrates the inherent limitations of the crossover method in Australia. For these reasons, we argue this metric should be avoided when predicting C$_4$ grass distribution in Australia, and in other highly arid regions.
Regression analysis also indicated that proportional C₄ grass richness and cover was influenced by local environmental factors. Increased tree cover had a negative effect on proportional C₄ Poaceae cover and richness, which may reflect the increased difficulty for all C₄ plants to grow under shade [28, 42, 55-57]. Although regression analysis also indicated AWC was an important predictor of proportional C₄ Poaceae cover, the overall impact of AWC is questionable and should be interpreted with caution given the poor fit of the data and the fact that its removal from the best fit model had a limited impact on R² or QAIC values. It is possible that AWC, soil texture (i.e. sand content), and tree cover are acting as useful indicators of local water availability (where increased tree cover is typically associated with greater moisture availability and rainfall). Local water availability can have a substantial effect on C₃:C₄ competitive dynamics. Increased moisture availability, particularly in warm mesic habitats, may mitigate the benefits of increased water-use efficiency in C₄ species [5] and provide C₃ grasses with a competitive advantage, ultimately supporting species co-existence [26]. However, capturing the nuanced influence of these local factors requires further study at a finer scale.

Like Poaceae, C₄ Cyperaceae cover was correlated with increases in summer temperatures and C₄ growing season length. However, C₄ Cyperaceae had a more restricted temperature range than C₄ Poaceae, indicating C₄ Poaceae can occupy relatively cooler climates. These results are consistent with Stock, Chuba (58), who found the transitional temperatures of proportional C₄ Poaceae and Cyperaceae richness in South Africa were 23°C and 34°C, respectively. Our models provided limited evidence of a strong relationship between proportional C₄ Cyperaceae cover or richness and rainfall. Similar to Poaceae, proportional
C₄ Cyperaceae cover increase slightly as January precipitation increased, but the overall effect was minor, and precipitation had no apparent influence on proportional C₄ Cyperaceae richness. Previous work examining Cyperaceae occurrence in other countries has also found temperature is the strongest driver of proportional C₄ Cyperaceae richness, with moisture availability having little impact [22, 24, 58, 59]. The absence of moisture-C₄ Cyperaceae richness correlations has been attributed to the large number of Cyperaceae that prefer wet habitats [23, 24]. Our work supports the hypothesis that precipitation has a more limited impact on the relative occurrence of C₄ sedges as compared to temperature.

The significant influence of pH on proportional C₄ Cyperaceae richness and cover could suggest soil biochemistry plays a role in C₃:C₄ sedge dynamics. Alkaline soils are less soluble than acidic soils, which can limit nutrient availability [60, 61]. The relative impact of alkaline-stress on C₄ and C₃ plants has not been widely explored. However, C₄ plants are considered more resistant to stress and thus may be more tolerant of alkaline soils [5, 62, 63]. pH is often correlated with other important conditions including salinity and soil fertility, thus this trend may reflect several interacting factors not explicitly considered here [29, 63, 64]. Although the causal relationship between these factors remains unclear, our results suggest that local factors modify C₄ sedge abundance and richness and should be explored more deeply in future work.

Large-scale evaluations of C₄ eudicot richness and cover are uncommon. Our results show several C₄ eudicot lineages, chiefly Aizoaceae, Amaranthaceae, and Portulacaceae, follow expected geographical trends, where proportional C₄ richness and cover increase from South to North, potentially in response to increasing temperatures and summer rainfall. However, other eudicot families, such as Asteraceae and Cleomaceae, did not display any clear
latitudinal trends, suggesting C₄ cover in these families is not driven by climate variables, but more likely by local environmental factors.

Despite reports of correlation with aridity in other regions [22, 24, 51, 65], none of the climate or local environmental variables we considered were strongly correlated with proportional C₄ Chenopodiaceae richness or cover. A potentially critical factor we were unable to consider due to lack of data was soil salinity. The evolution of C₄ photosynthesis has long been considered a pre-adaption to arid and saline habitats [1, 66, 67], although additional work indicates adaptations to saline soils may have actually promoted C₄ evolution in Chenopodiaceae [68]. As a result, salinity may be a key controlling factor of the distribution patterns of all C₄ species in Australia [69]. Additional local factors such as soil nutrient content and salinity should be incorporated in future studies of C₄ eudicot distribution where possible. Moreover, while plot location selection procedures were designed to reduce anthropogenic influences on native richness and cover estimates, the C₃/C₄ patterns reported here are not necessarily in equilibrium. Historical long-term grazing has led to substantial changes in species composition in some areas. For example, chenopods have been lost from large areas of Australia due to overgrazing [70]. Therefore, the environmental trends in C₃/C₄ richness and cover observed in this study, or lack thereof in the case of chenopods, may be a partial reflection of human influences on the landscape.

Increased rainfall and moisture availability were negatively correlated with proportional C₄ Euphorbiaceae cover and richness, suggesting C₄ Euphorbiaceae were more prevalent in dryer, although not necessarily hotter, conditions. This may explain why we did not detect a clear latitudinal trend in proportional C₄ Euphorbiaceae cover and richness, as compared to
monocot and some other eudicot taxa. Rainfall seasonality also appeared important, where plots with proportionally high C₄ Euphorbiaceae cover and richness were only found in areas with greater than 50% SWA. Similar to our findings, Stowe and Teeri (65) found summer pan evaporation was most closely associated with North American C₄ Euphorbiaceae distribution. In Egypt, C₄ *Euphorbia* also occurred mainly in arid environments [71]. As ours is one of the few studies to examine C₄ Euphorbiaceae dynamics at a large scale, these results are useful in understanding what factors control their distribution, not only in Australia, but globally. However, because Euphorbiaceae were less common in TERN plots compared to the other primary taxa investigated here, these results should be interpreted carefully and warrant further investigation.

Plots containing Poaceae and Cyperaceae exhibited a rapid transition from C₃ to C₄ dominated plots, with few plots containing a mix of both pathways. The bimodal distributions observed here may reflect the relatively rapid geographic transition in temperature and summer to winter dominated rainfall in Australia, leaving a limited habitat range where both species can survive. This finding has significant implications, since the geographic location of this narrow threshold may be highly sensitive to climate change [35]. Griffith, Anderson (28) also noted this bimodal trend in C₃ and C₄ grass distribution in North America, and suggested these extremes were maintained by local disturbance, mainly fire, which can support dominant relationships of one group over the other. Fire-prone environments may favour increased C₄ grass cover in Australia [72, 73] however, we were unable to test for this possibility in this study due to lack of data. Anecdotal observations by ground observers at some temperate TERN plots that experience regular burning have noted greater C₄ cover compared to undisturbed areas, but a more precise fire history for these areas is needed to validate these observations.
The proportional cover and richness of C$_4$ species are rarely simultaneously evaluated, although C$_4$ grass presence and richness is commonly used to validate C$_4$ cover at large scales [7, 11, 19, 35]. Proportional C$_4$ richness and cover were positively correlated in each family, suggesting richness can be used to broadly support cover estimates. However, the strength of these correlations was largely driven by the bimodal distribution in some families. There was still considerable variance between proportional C$_4$ Poaceae and Chenopodiaceae richness and cover at mixed C$_3$/C$_4$ plots. Although our analysis was only designed to explore the influence of different factors, our work nonetheless indicated our ability to predict proportional C$_4$ monocot richness is more accurate compared to proportional C$_4$ monocot cover. These findings suggest the disproportionate relationship between C$_4$ cover and richness is due to differences in the ability of species to occupy versus dominate an area. For example, given both proportional C$_4$ Poaceae richness and cover were predominantly influenced by January maximum temperature, differences in model fit are likely determined by the modifying effects of water availability and other local environmental factors. Overall, our results demonstrate that C$_4$ cover and richness provide unique assessments of species distribution and responses to the environment and thus should be measured together whenever possible.

This study is the first to compare the influence of both climate and local ecology on C$_4$ grass, sedge, and eudicot cover and richness at a continental scale in Australia. Our results make clear that broadscale as well as localised environmental factors have divergent impacts on C$_4$ taxa. While monocot lineages generally followed expected temperature-driven trends, there was considerable variability among eudicots families. Quantifying these differences is critical
to predicting $C_4$ cover under different environmental scenarios, or estimating plant resilience to small-scale perturbations. Future work can leverage the TERN datasets presented here to investigate more fine-scale taxonomic patterns, such as the influence of $C_4$ subtype on species distribution and their relationship to climate [74, 75]. Moreover, as TERN continues to expand its plot network and range of local environmental data, it may soon be possible to investigate the influence of fire salinity on $C_4$ at both local and national scales.

Acknowledgements: We acknowledge the TERN Ecosystem Surveillance field team and the support of TERN by the Australian government through the National Collaborative Research Infrastructure Strategy. We thank Tom Saleeba and David Summers. Additional financial support for this project was from the AMP Foundation and the AMP Tomorrow Fund awarded to SM, and Australian Research Council Future Fellowship (FT110100100793) awarded to F.A.M., and Australian Government Research Training Program and University of Adelaide Faculty of Sciences Divisional scholarships awarded to J.W.A.
References


59. Kalapos T, Balogh-


Fig 1 Proportional C₄ cover and richness at TERN Monitoring plots for (a, b) Poaceae, (c, d) Cyperaceae, (e, f) Chenopodiaceae, and (g, h) Euphorbiaceae.
Fig 2 Binomial logistic regressions of proportional C$_4$ Poaceae richness. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot (A-D), independent variables not depicted on the x-axis are held constant at their median value.
**Fig 3** Binomial logistic regressions of proportional C₄ Poaceae cover. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot (A-D), independent variables not depicted on the x-axis are held constant at their median value.
Fig 4 Binomial logistic regressions of proportional C₄ Cyperaceae richness. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot (A-C), independent variables not depicted on the x-axis were held constant at their median value.
Fig 5 Binomial logistic regressions of proportional C₄ Cyperaceae cover. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot (A-D), independent variables not depicted on the x-axis were held constant at their median value.
**Fig 6.** Binomial logistic regressions of proportional C₄ Chenopodiaceae richness. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot, independent variables not depicted on the x-axis were held constant at their median value.
Fig 7 Binomial logistic regressions of proportional C₄ Chenopodiaceae cover. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot (A-D), independent variables not depicted on the x-axis were held constant at their median value.
Fig 8. Binomial logistic regressions of proportional C₄ Euphorbiaceae richness. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot, independent variables not depicted on the x-axis were held constant at their median value.
**Fig 9.** Binomial logistic regressions of proportional C4 Euphorbiaceae cover. Grey points are partial residuals, blue lines and shaded bands are predicted outcomes of the regression and 95% confidence intervals respectively, and rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). For each plot, independent variables not depicted on the x-axis were held constant at their median value.
SI Appendix Captions

**SI Appendix 1:** Total number of species recorded in TERN survey plots in each photosynthetic pathway belonging to families with C₄ species.

**SI Appendix 2:** Proportional C₄ richness and cover at each TERN plot for all 14 families in which C₄ species were identified.

**SI Appendix 3:** Maps of proportional C₄ cover and richness at TERN Monitoring plots for all 14 families in which C₄ species were identified.

**SI Appendix 4:** Binomial logistic regression model results of proportional C₄ cover and richness analysis, and results of leave-one-out comparisons of the best fit models to predict proportional C₄ cover and richness.

**SI Appendix 5:** Proportional C₄ richness versus cover