

## Letters

# A global database of C<sub>4</sub> photosynthesis in grasses

## Introduction

C<sub>3</sub>, C<sub>4</sub> or Crassulacean acid metabolism (CAM) photosynthetic pathways represent a fundamental axis of trait variation in plants, with importance at scales from genome to biome. Knowing the distribution of these pathways among wild species is a crucial first step in understanding the patterns and processes of photosynthetic evolution and its role in ecological processes at large scales (e.g. changes in the composition of biomes under global change). C<sub>4</sub> photosynthesis is most prevalent in the Poaceae (grasses), which account for about half of all C<sub>4</sub> species (Sage *et al.*, 1999a). Research on the evolution and ecology of these plants has undergone a renaissance during the last 7 yr, catalyzed by phylogenetic analyses showing multiple parallel C<sub>4</sub> origins (e.g. Christin *et al.*, 2007; Vicentini *et al.*, 2008; GPWG II, 2012), insights into the distribution of C<sub>4</sub> species and assembly of the C<sub>4</sub> grassland biome (Edwards & Still, 2008; Edwards & Smith, 2010; Edwards *et al.*, 2010), and efforts to introduce the C<sub>4</sub> pathway into rice (Hibberd *et al.*, 2008; von Caemmerer *et al.*, 2012). C<sub>4</sub> photosynthesis is an excellent model for investigating complex trait evolution, because of the broad knowledge base describing its biochemical basis, evolutionary history, and ecological interactions (Christin *et al.*, 2010).

## Why do we need a C<sub>4</sub> database?

Investigations of the evolution and ecological significance of C<sub>4</sub> photosynthesis are increasingly turning to large-scale comparisons of C<sub>3</sub> and C<sub>4</sub> species. These are straightforward for well-characterized common or model species. However, when comparisons are extended to include large numbers of nonmodel species, two important challenges arise. First, there are > 62 000 published scientific names for grasses corresponding to over 11 000 accepted species (Clayton *et al.*, 2002b onwards), making an average of five synonyms for each accepted name. This leads to problems when linking data based on alternative names for the same species concept, and to redundancy in published data surveys, when values for synonyms are presented as independent data. Second, although there have been extensive previous surveys of the photosynthetic pathway spanning the diversity of wild species (Hattersley & Watson, 1992; Sage *et al.*, 1999a), the rarity of most species means that this work is incomplete, and the synonymy problem makes it difficult to identify the gaps in these data.

Accounting for synonymy and spelling variants/mistakes has become one of the central challenges for the emerging fields of ecological and evolutionary informatics, in which data are synthesized across different sources on increasingly larger scales (Jones *et al.*, 2006; Sidlauskas *et al.*, 2009). In one infamous example, a 22.5 million record database of plant species occurrences and traits for the Americas contained more scientific names than there are thought to be plant species on Earth (Whitfield, 2011). However, this taxonomic impediment to data synthesis has been progressively broken down by a combination of new methodological developments for name matching (Patterson *et al.*, 2010; Boyle *et al.*, 2013; Chamberlain & Szocs, 2013; Kluyver & Osborne, 2013), and the compilation of nomenclatural databases by botanic gardens and natural history museums (e.g. The Plant List, 2010). Here, we showcase how such resources may be used to assemble and index databases of discrete traits for large numbers of species.

## Compilation and overview of the data

Our database of C<sub>3</sub> and C<sub>4</sub> photosynthetic types in grasses is based principally on published anatomical and stable carbon isotope evidence. We followed previous authors in assuming that all species within each genus shared the same photosynthetic pathway, unless the evidence suggested otherwise. However, we also measured δ<sup>13</sup>C for 99 species that had not previously been surveyed, including 96 species of *Panicum* s.l., *Acrostia gracilis*, *Lophopogon tridentatus* and *Theclachloa annua* (Supporting Information Table S1). We also obtained information on leaf anatomy and measured δ<sup>13</sup>C to check previous unverified reports of a C<sub>3</sub> species (*Stipagrostis paradisea*) in an otherwise C<sub>4</sub> genus (Sage *et al.*, 1999a), and a C<sub>4</sub> subspecies (*Chaetobromus dregeanus* ssp. *involutratus*) in an otherwise C<sub>3</sub> subfamily (Danthonioideae; Watson & Dallwitz, 1992 onwards). In both cases, our data contradicted previous reports, showing that the photosynthetic pathway of these taxa matches that of their close relatives; *S. paradisea* is C<sub>4</sub> and *C. dregeanus* ssp. *involutratus* is C<sub>3</sub> (Table 1).

The photosynthetic pathways of *Taeniorhachis repens*, *Veldkampia sagaiensis* and 39 rare species of *Panicum* s.l. remain unclassified, because we were unable to take samples of type specimens from herbarium collections. Most of these species are endemics of Madagascar (26 species), and the remaining 13 species are endemics of a small number of countries in Africa and Southeast Asia, and oceanic islands (Table S2). This means that the database is complete for most countries of the world.

Our approach has been to map the photosynthetic pathway data onto accepted species names in the Poaceae taxonomy of Clayton & Renvoize (1986) and Clayton *et al.* (2002b onwards), which is the most comprehensive treatment of accepted names and synonymy for grasses (see Methods S1 for full methodology). Coupling our

**Table 1** Photosynthetic type for genera previously reported to include a mixture of C<sub>3</sub>, C<sub>4</sub> and C<sub>3</sub>–C<sub>4</sub> intermediate species

| Taxon                                       | Type                               | References   |
|---|------------------------------------|--|
| 1. <i>Alloteropsis</i>                      |                                    |  |
| <i>A. angusta</i>                           | C <sub>4</sub>                     | Metcalf (1960), Ellis (1974), Hattersley <i>et al.</i> (1977), Watson & Dallwitz (1992 onwards)  |
| <i>A. cymicina</i>                          | C <sub>4</sub>                     |  |
| <i>A. paniculata</i>                        | C <sub>4</sub>                     |  |
| <i>A. papillosa</i>                         | C <sub>4</sub>                     |  |
| <i>A. semialata</i> ssp. <i>semialata</i>   | C <sub>4</sub>                     |  |
| <i>A. semialata</i> ssp. <i>eckloniana</i>  | C <sub>3</sub>                     |  |
| 2. <i>Aristida</i>                          |                                    |  |
| <i>A. longifolia</i>                        | C <sub>3</sub>                     | Cerro-Tlatilpa & Columbus (2009)<br>Watson & Dallwitz (1992 onwards)   |
| Other <i>Aristida</i> species               | C <sub>4</sub>                     |  |
| 3. <i>Chaetobromus involucreatus</i>        |                                    |  |
| ssp. <i>involucreatus</i> <sup>1</sup>      | C <sub>3</sub>                     | This study; H. P. Linder (pers. comm.)   |
| ssp. <i>sericeus</i>                        | C <sub>3</sub>                     | This study; H. P. Linder (pers. comm.)   |
| ssp. <i>dregeanus</i>                       | C <sub>3</sub>                     | This study; H. P. Linder (pers. comm.)   |
| 4. <i>Dregeochloa</i>                       |                                    |  |
| <i>D. calvinensis</i> <sup>2</sup>          | C <sub>3</sub> –C <sub>4</sub> (?) | Watson & Dallwitz (1992 onwards)   |
| <i>D. pumila</i>                            | C <sub>3</sub>                     |  |
| 5. <i>Eragrostis</i> <sup>3</sup>           |                                    |  |
| <i>Eragrostis walteri</i> <sup>4</sup>      | C <sub>3</sub>                     | Schulze <i>et al.</i> (1996); Ingram <i>et al.</i> (2011); Watson & Dallwitz (1992 onwards)  |
| Other <i>Eragrostis</i> species             | C <sub>4</sub>                     |  |
| 6. <i>Homolepis</i>                         |                                    |  |
| <i>Homolepis aturensis</i> <sup>5</sup>     | C <sub>3</sub> –C <sub>4</sub> (?) | Christin <i>et al.</i> (2013); Watson & Dallwitz (1992 onwards)  |
| Other <i>Homolepis</i> species              | C <sub>3</sub>                     |  |
| 7. <i>Neurachne</i>                         |                                    |  |
| <i>N. alopecuroides</i>                     | C <sub>3</sub>                     | Hattersley <i>et al.</i> (1982); Hattersley & Roksandic (1983); Christin <i>et al.</i> (2012)  |
| <i>N. annularis</i>                         | C <sub>3</sub>                     |  |
| <i>N. lanigera</i>                          | C <sub>3</sub>                     |  |
| <i>N. minor</i>                             | C <sub>3</sub> –C <sub>4</sub>     |  |
| <i>N. munroi</i>                            | C <sub>4</sub>                     |  |
| <i>N. queenslandica</i>                     | C <sub>3</sub>                     |  |
| <i>N. tenuifolia</i>                        | C <sub>3</sub>                     |  |
| <i>Paraneurachne muelleri</i> <sup>6</sup>  | C <sub>4</sub>                     |  |
| 8. <i>Panicum</i> s.l. <sup>3</sup>         |                                    |  |
| 252 species                                 | C <sub>3</sub>                     | Various sources, including this study (see Table S4 for full details)  |
| 169 species                                 | C <sub>4</sub>                     |  |
| <i>P. ruspolii</i>                          | C <sub>3</sub> –C <sub>4</sub> (?) |  |
| 9. <i>Steinchisma</i>                       |                                    |  |
| <i>S. cupreum</i>                           | C <sub>3</sub> –C <sub>4</sub>     | Brown (1977), Morgan & Brown (1979, 1980); Renvoize (1987); Watson & Dallwitz (1992 onwards)   |
| <i>S. decipiens</i>                         | C <sub>3</sub> –C <sub>4</sub>     |  |
| <i>S. exiguiflorum</i>                      | C <sub>3</sub> –C <sub>4</sub>     |  |
| <i>S. hians</i>                             | C <sub>3</sub> –C <sub>4</sub>     |  |
| <i>S. laxum</i>                             | C <sub>3</sub>                     |  |
| <i>S. spathellosum</i>                      | C <sub>3</sub> –C <sub>4</sub>     |  |
| <i>S. stenophyllum</i>                      | C <sub>3</sub> –C <sub>4</sub>     |  |
| 10. <i>Stipagrostis</i>                     |                                    |  |
| <i>S. paradisea</i> <sup>7</sup>            | C <sub>4</sub>                     | Renvoize (1986); this study<br>Watson & Dallwitz (1992 onwards)  |
| Other <i>Stipagrostis</i> species           | C <sub>4</sub>                     |  |
| 11. <i>Streptostachys</i> s.l. <sup>3</sup> |                                    |  |
| <i>S. acuminata</i>                         | C <sub>4</sub>                     | Morrone & Zuloaga (1991); Watson & Dallwitz (1992 onwards); Filgueiras <i>et al.</i> (1993);<br>De Olivera & Longhi-Wagner (2007); P-A. Christin (pers. comm.) |
| <i>S. asperifolia</i> <sup>5</sup>          | C <sub>3</sub> –C <sub>4</sub> (?) |  |
| <i>S. lanciflora</i>                        | C <sub>3</sub>                     |  |
| <i>S. macrantha</i>                         | C <sub>4</sub>                     |  |
| <i>S. ramosa</i>                            | C <sub>4</sub>                     |  |
| <i>S. rigidifolia</i>                       | C <sub>4</sub>                     |  |
| <i>S. robusta</i>                           | C <sub>3</sub>                     |  |

Taxonomy follows Clayton *et al.* (2002b onwards).

<sup>1</sup>Anatomical (H. P. Linder, pers. comm.) and  $\delta^{13}\text{C}$  evidence (Supporting Information Tables S1, S3) conflicts with a previous report that this subspecies is C<sub>4</sub> (Watson & Dallwitz, 1992 onwards).

<sup>2</sup>Anatomical evidence shows that in *D. calvinensis* most mesophyll cells are no more than one cell distant from bundle sheath cells, making it potentially a C<sub>3</sub>–C<sub>4</sub> intermediate.

<sup>3</sup>Genus known to be polyphyletic.

<sup>4</sup>Note that a recent phylogenetic treatment (Ingram *et al.*, 2011) places *E. walterii* outside the genus *Eragrostis*. However, its taxonomy has not yet been revised.

<sup>5</sup>Anatomical evidence showing a concentration of chloroplasts within large bundle sheath cells suggests that this species is potentially a C<sub>3</sub>–C<sub>4</sub> intermediate (Christin *et al.*, 2013; P-A. Christin, pers. comm.).

<sup>6</sup>Phylogenetic analysis places the genus *Paraneurachne* nested within the genus *Neurachne* (Christin *et al.*, 2012).

<sup>7</sup>Anatomical (Renvoize, 1986) and  $\delta^{13}\text{C}$  evidence (Tables S1, S3) conflicts with a previous report that this species is C<sub>3</sub> (Sage *et al.*, 1999a).

dataset with this synonymy allows users to return the photosynthetic type for all except 46 (corresponding to 41 accepted species) of the 62 678 published scientific names (accepted names and synonyms) for grasses (Clayton *et al.*, 2002b onwards). We have developed software tools to facilitate this task for users, which are detailed in the following.

The database covers 99.6% of the 11 087 grass species. It shows that 42% of these species use the  $C_4$  photosynthetic pathway and 57% the  $C_3$  pathway (Table S3; Notes S1). Six genera (*Alloterospis*, *Aristida*, *Eragrostis*, *Neurachne*, *Panicum* s.l., and *Streptostachys* s.l.) contain both  $C_3$  and  $C_4$  species (Tables 1, S4). Seven  $C_3$ – $C_4$  intermediate species (Table 1) are distributed between the genera *Neurachne* (one species) and *Steinchisma* (six species). Within the genus *Panicum* s.l., 169 species are  $C_4$ , 250 are  $C_3$ , and 41 remain unknown, with the photosynthetic type of *Panicum ruspolii* ambiguous on the basis of new  $\delta^{13}C$  measurements (Tables S1, S2, S4; Notes S1). The latter species may be a previously unrecognized  $C_3$ – $C_4$  intermediate, but further work is required to test this hypothesis. A number of further potential  $C_3$ – $C_4$  intermediates have been identified on the basis of anatomical observations (Tables 1, S3), and also need to be investigated physiologically. These are *Dregeochloa calviniensis* (most mesophyll cells are no more than one cell distant from bundle sheath cells; Watson & Dallwitz, 1992 onwards), *Homolepis aturensis* and *Streptostachys asperifolia* (concentration of chloroplasts in large bundle sheath cells; Christin *et al.*, 2013; P-A. Christin, pers. comm.). In total there are therefore 11 putative  $C_3$ – $C_4$  intermediates in the grasses.

## Caveats

A number of caveats are important when collating and using large trait databases of this kind. The assumption that all species within each genus share the same photosynthetic pathway is reasonable in most cases. However, significant and interesting exceptions, such as the  $C_3$  *Aristida* species in an otherwise  $C_4$  genus (Cerro-Tlatilpa & Columbus, 2009), raise the possibility of errors at the species level. Misclassification is most likely in lineages where multiple evolutionary transitions between photosynthetic pathways have occurred, especially in Paniceae and Paspaleae (Morrone *et al.*, 2012). The polyphyly of many grass genera accentuates this problem, most acutely illustrated by *Eragrostis walteri*, which was previously considered to be a  $C_3$  species within a wholly  $C_4$  genus (Table 1). Recent phylogenetic work has demonstrated that this species is actually a member of the  $C_3$  Arundinoideae lineage and misplaced within *Eragrostis* (Ingram *et al.*, 2011).

The polyphyly of grass genera means that Tables 1, S3 and S4 should be interpreted with caution. While they do catalogue the known distribution of  $C_4$  photosynthesis among taxa, they do not necessarily provide information about its evolutionary history. However, ongoing phylogenetic work is steadily resolving the polyphyly issue, which is most acute in the genus *Panicum*. We have used the conservative circumscription of *Panicum* s.l. adopted in GrassBase (Clayton *et al.*, 2002b onwards) and recently carried over to the World Checklist of Poaceae (Clayton *et al.*, 2012 onwards) and The Plant List (The Plant List, 2010), because these

online resources provide the most comprehensive, global list of accepted names and synonyms, and are regularly updated in the light of new publications. Using the software tools detailed in the following, it is straightforward to link the  $C_3$ / $C_4$  data listed for *Panicum* s.l. (see Table S4) to the new genus circumscriptions. The same applies to *Streptostachys* s.l. (Table 1).

## How to access the database

Easy routes for users to access information are crucial determinants of the usefulness and usage of data. Our database may be accessed via three routes. The first is static, but the second and third will report updates to the database as we make them.

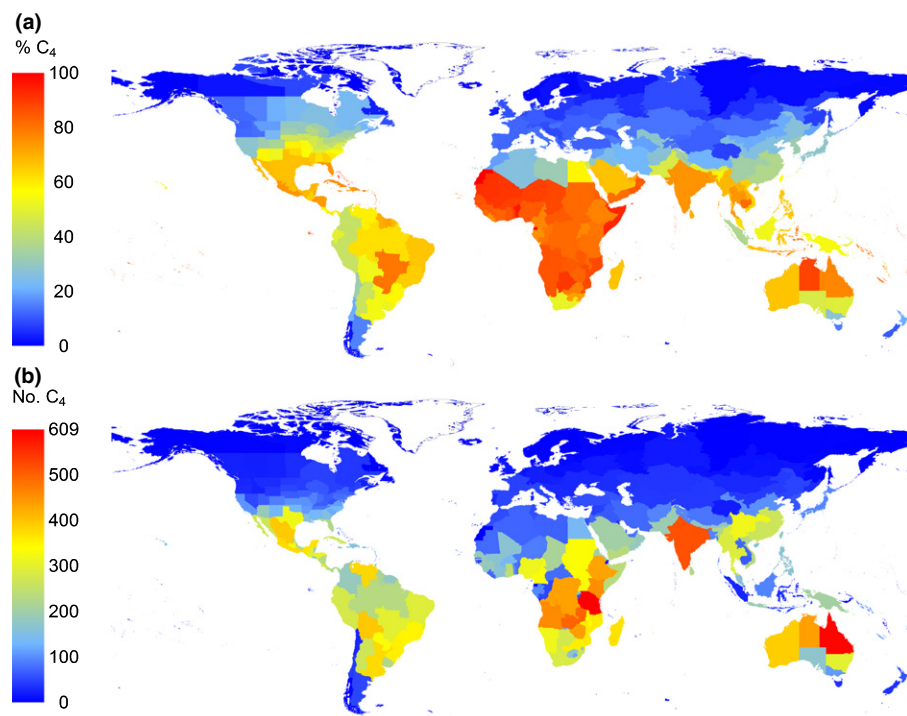
First, simple tables list photosynthetic pathway by accepted scientific name, and may be accessed in the Supporting Information (Tables S3, S4). These require the user to first prepare a list of accepted species names according to the taxonomy of Clayton *et al.* (2002b onwards) for the taxa of interest.

Second, the name-matching and data-linkage steps may be combined within the software package Taxonome (Kluyver & Osborne, 2013; <http://taxonome.bitbucket.org>; persistent URL <http://purl.org/NET/taxonome>). Taxonome links datasets using species names, handling both synonyms and spelling variants (including spelling mistakes). It deals rapidly with millions of names, and runs via either a simple Graphical User Interface (GUI) for basic functionality or python scripts for advanced users. A user first loads the Kew taxonomy and photosynthetic pathway database via a data file obtained from the Taxonome website. Custom lists comprising any published grass names may then be rapidly matched to this database, and outputted in CSV format.

Third, the photosynthetic pathway data are linked to the Kew taxonomy, together with morphological, phylogenetic, biogeographic and environmental data within the GrassPortal system (Osborne *et al.*, 2011; [www.grassportal.org](http://www.grassportal.org)). GrassPortal enables users to easily assemble large-scale, synthetic data products based on multiple original sources, and is accessed via an intuitive and simple GUI. Using this system, users are able to assemble a list of all grass species present in a particular geographic area, linked to photosynthetic pathway, growth form, and environmental niche data.

## Large-scale data synthesis

By carrying out technically challenging bioinformatic steps of data processing and linkage, services like GrassPortal open up new possibilities for a broad biological community to explore large-scale synthetic data products. For example, linkage of the photosynthetic pathway dataset with species occurrence data (Clayton *et al.*, 2002a) allows the distribution of  $C_4$  grass species to be mapped at the global scale (Fig. 1). This map improves the global coverage compared with previous data compilations, especially for Africa, South America and Southeast Asia (Sage *et al.*, 1999b). It particularly highlights the prevalence of  $C_4$  photosynthesis among African grasses (Fig. 1a), and the importance of central-east Africa, India and northern Australia as hotspots of  $C_4$  grass species richness (Fig. 1b). The new dataset also facilitates large-scale



**Fig. 1** Global map of  $C_4$  grass species distributions. (a) Percentage of grass species within each mapping unit that uses the  $C_4$  pathway; (b) the species richness of  $C_4$  grasses in each mapping unit. The map shows species distributions at the Taxonomic Databases Working Group (TDWG) level 3 'botanical country' scale, a biodiversity information standard corresponding largely to political countries, but with large countries subdivided into smaller mapping units (Brummitt *et al.*, 2001).

macroevolutionary analyses. For example, the Grass Phylogeny Working Group II (2012) used our data in phylogenetic analyses to discover multiple new  $C_4$  lineages, and to infer that evolutionary gains prevail over losses of this trait. Another recent study used our data in a macroevolutionary analysis to show an association between  $C_4$  photosynthesis and salt tolerance in grasses (Bromham & Bennett, 2014).

The integration of our  $C_4$  pathway data with information on geographical distributions, environmental niche, and phylogenetic relationships promises important novel insights into the ecological significance and evolution of this complex physiological and anatomical trait. More generally, it offers biologists an example of how functional trait data may be used in large-scale synthesis and analysis to advance our understanding of the ecological and evolutionary processes acting on organisms.

## Acknowledgements

We are grateful to Liliana Giussani and Pascal-Antoine Christin for their critical comments on the manuscript. This work builds on that of previous authors who have compiled comprehensive databases on grass leaf anatomy and photosynthetic type, including C. R. Metcalfe, Walter Brown, Roger Ellis, Paul Hattersley, Les Watson, and Rowan Sage, and owes a debt of gratitude to them. We thank Les Watson for his generosity in allowing us to use data from Grass Genera of the World (<http://delta-intkey.com/grass/>), and the following for financially supporting this work: C.P.O. was supported by a Royal Society University Research Fellowship and NERC grant number NE/I014322/1, A.S. by the European Union's Erasmus scheme, T.A.K. by a University of Sheffield Postgraduate Studentship,

and V.V. by the GrassPortal project supported by the e-content programme of the JISC. The development of GrassPortal was funded by the JISC, with additional support from the University of Sheffield, the Royal Botanic Gardens, Kew, and Knowledge-Now Limited. We thank Peter Linder of the University of Zurich for information on the leaf anatomy of Danthonioid grasses, Tony Verboom (University of Cape Town) for material of *Chaetobromus dregeanus* for isotopic analysis, Marjorie Lundgren for her help in acquiring material and Heather Walker for running the analyses. We also thank the following herbaria and their staff for generous help with plant material for the isotope survey: Skye Coffey (Western Australian Herbarium), Olof Ryding (Botanisk Museum, Koebenhavus Universitet), Mats Thulin (Uppsala University), Brendan Lepschi (Australian National Herbarium), Bryan Simon (Queensland Herbarium), Lyn Fish (SANBI) and the Missouri Botanical Garden.

**Colin P. Osborne<sup>1\*</sup>, Anna Salomaa<sup>1,2</sup>, Thomas A. Kluyver<sup>1</sup>, Vernon Visser<sup>1,3</sup>, Elizabeth A. Kellogg<sup>4</sup>, Osvaldo Morrone<sup>5</sup>, Maria S. Vorontsova<sup>6</sup>, W. Derek Clayton<sup>6</sup> and David A. Simpson<sup>6</sup>**

<sup>1</sup>Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK;

<sup>2</sup>Department of Biological and Environmental Science, University of Jyväskylä, PO Box 35, Jyväskylä 40500, Finland;

<sup>3</sup>Centre for Invasion Biology, Department of Botany and Zoology, University of Stellenbosch, Natural Sciences Building, Private Bag X1, Matieland 7602, South Africa;

<sup>4</sup>Donald Danforth Plant Science Center, 975 North Warson Road,



St Louis, MO 63132, USA;

<sup>5</sup>Instituto de Botánica Darwinion, Labardén 200, C.C. 22,  
B1642HYD, San Isidro, Buenos Aires, Argentina;

<sup>6</sup>Herbarium, Library, Art and Archives, Royal Botanic Gardens,  
Kew, Richmond, Surrey, TW9 3AE, UK

(\*Author for correspondence: tel +44 114 0146;  
email c.p.osborne@sheffield.ac.uk)

## References

- Boyle B, Hopkins N, Lu Z, Garay JAR, Mozzherin D, Rees T, Matasci N, Narro ML, Piel WH, McKay SJ *et al.* 2013. The taxonomic names resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* 14: 16.
- Bromham L, Bennett TH. 2014. Salt tolerance evolves more frequently in C<sub>4</sub> grass lineages. *Journal of Evolutionary Biology* 27: 653–659.
- Brown WV. 1977. The Kranz syndrome and its subtypes in grass systematics. *Memoirs of the Torrey Botanical Club* 23: 1–97.
- Brummitt RK, Pando F, Hollis S, Brummitt NA. 2001. *Plant taxonomic database standards no. 2. World geographical scheme for recording plant distributions*, 2<sup>nd</sup> edn. Pittsburgh, PA, USA: Published for the International Working Group on Taxonomic Databases For Plant Sciences (TDWG) by the Hunt Institute for Botanical Documentation, Carnegie Mellon University.
- von Caemmerer S, Quick WP, Furbank RT. 2012. The development of C<sub>4</sub> rice: current progress and future challenges. *Science* 336: 1671–1672.
- Cerros-Tlatilpa R, Columbus JT. 2009. C<sub>3</sub> photosynthesis in *Aristida longifolia*: implication for photosynthetic diversification in *Aristidoideae* (Poaceae). *American Journal of Botany* 96: 1379–1387.
- Chamberlain S, Szocs E. 2013. taxize – taxonomic search and retrieval in R. *F1000Research* 2: 191. [WWW document] URL <http://f1000research.com/articles/2-191/v2> [accessed 28 May 2014].
- Christin P-A, Freckleton RP, Osborne CP. 2010. Can phylogenetics identify C<sub>4</sub> origins and reversals? *Trends in Ecology and Evolution* 25: 403–409.
- Christin P-A, Salamin N, Savolainen V, Duvall MR, Besnard G. 2007. C<sub>4</sub> photosynthesis evolved in grasses via parallel adaptive genetic changes. *Current Biology* 17: 1241–1247.
- Christin P-A, Wallace MJ, Clayton H, Edwards EJ, Furbank RT, Hattersley PW, Sage RF, Macfarlane TD, Ludwig M. 2012. Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe Neurachninae. *Journal of Experimental Botany* 63: 6297–6308.
- Christin PA, Osborne CP, Chatelet DS, Columbus JT, Besnard G, Hodkinson TR, Garrison LM, Vorontsova MS, Edwards EJ. 2013. Anatomical enablers and the evolution of C<sub>4</sub> photosynthesis in grasses. *Proceedings of the National Academy of Sciences, USA* 110: 1381–1386.
- Clayton WD, Govaerts R, Harman KT, Williamson H, Vorontsova M. 2012 onwards. *World checklist of Poaceae*. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet. [WWW document] URL <http://apps.kew.org/wcsp/> [accessed 15 March 2012].
- Clayton WD, Renvoize SA. 1986. *Genera Graminum*. London, UK: HMSO.
- Clayton WD, Vorontsova MS, Harman KT, Williamson H. 2002a onwards. *GrassBase – the online world grass flora*. [WWW document] URL <http://www.kew.org/data/grasses-db.html> [accessed 25 September 2013].
- Clayton WD, Vorontsova MS, Harman KT, Williamson H. 2002b onwards. *GrassBase – the online world grass flora*. Synonymy. [WWW document] URL <http://www.kew.org/data/grasses-syn.html> [accessed 25 September 2013].
- De Olivera RP, Longhi-Wagner HM. 2007. New species of *Streptostachys* (Poaceae: Paniceae) from Brazil. *Kew Bulletin* 62: 493–497.
- Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, C<sub>4</sub> Grasses Consortium. 2010. The origins of C<sub>4</sub> grasslands: integrating evolutionary and ecosystem science. *Science* 328: 587–591.
- Edwards EJ, Smith SA. 2010. Phylogenetic analyses reveal the shady history of C<sub>4</sub> grasses. *Proceedings of the National Academy of Sciences, USA* 107: 2532–2537.
- Edwards EJ, Still CJ. 2008. Climate, phylogeny and the ecological distribution of C<sub>4</sub> grasses. *Ecology Letters* 11: 266–276.
- Ellis RP. 1974. The significance of the occurrence of both Kranz and non-Kranz leaf anatomy in the grass species *Alloteropsis semialata*. *South African Journal of Science* 70: 169–173.
- Filgueiras TS, Morrone O, Zuloaga FO. 1993. A new species of *Streptostachys* (Poaceae: Paniceae) from Brazil. *Novon* 3: 252–257.
- Grass Phylogeny Working Group II. 2012. New grass phylogeny resolves deep evolutionary relationships and discovers C<sub>4</sub> origins. *New Phytologist* 193: 304–312.
- Hattersley PW, Roksandic Z. 1983.  $\delta^{13}\text{C}$  values of C<sub>3</sub> and C<sub>4</sub> species of Australian *Neurachne* and its allies (Poaceae). *Australian Journal of Botany* 31: 317–321.
- Hattersley PW, Watson L. 1992. Diversification of photosynthesis. In: Chapman GP, ed. *Grass evolution and domestication*. New York, NY, USA: Cambridge University Press, 38–116.
- Hattersley PW, Watson L, Johnston CR. 1982. Remarkable leaf anatomical variations in *Neurachne* and its allies (Poaceae) in relation to C<sub>3</sub> and C<sub>4</sub> photosynthesis. *Botanical Journal of the Linnean Society* 84: 265–272.
- Hattersley PW, Watson L, Osmond CB. 1977. *In situ* immunofluorescent labelling of ribulose-1, 5-bisphosphate carboxylase in C<sub>3</sub> and C<sub>4</sub> plants. *Australian Journal of Plant Physiology* 4: 523–539.
- Hibberd JM, Sheehy JE, Langdale JA. 2008. Using C<sub>4</sub> photosynthesis to increase the yield of rice – rationale and feasibility. *Current Opinion in Plant Biology* 11: 228–231.
- Ingram AL, Christin P-A, Osborne CP. 2011. Molecular phylogenies disprove an hypothesized C<sub>4</sub> reversion in *Eragrostis walteri* (Poaceae: Chloridoideae). *Annals of Botany* 107: 321–325.
- Jones MB, Schildhauer MP, Reichman OJ, Bowers S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics* 37: 519–544.
- Kluyver TA, Osborne CP. 2013. Taxonome: a software package for linking biological species data. *Ecology and Evolution* 3: 1262–1265.
- Metcalfe CR. 1960. *Anatomy of the monocotyledons. I. Gramineae*. Oxford, UK: Clarendon Press.
- Morgan JA, Brown RH. 1979. Photosynthesis in grass species differing in carbon dioxide fixation pathways. II. A search for species with intermediate gas exchange and anatomical characteristics. *Plant Physiology* 64: 257–262.
- Morgan JA, Brown RH. 1980. Photosynthesis in grass species differing in carbon dioxide fixation pathways. III. Oxygen response and enzyme activities of species in the *laxa* group of *Panicum*. *Plant Physiology* 65: 156–159.
- Morrone O, Aagesen L, Scataglini MA, Salarato DL, Denham SS, Chemisquy MA, Sede SM, Giussani LM, Kellogg EA, Zuloaga FO. 2012. Phylogeny of the Paniceae (Poaceae: Panicoideae): integrating plastid DNA sequences and morphology into a new classification. *Cladistics* 28: 333–356.
- Morrone O, Zuloaga F. 1991. Revision del genero *Streptostachys* (Poaceae-Panicoideae), su posicion sistematica dentro de la tribu Paniceae. *Annals of the Missouri Botanical Gardens* 78: 359–376.
- Osborne CP, Visser V, Chapman S, Barker A, Freckleton RP, Salamin N, Simpson D, Uren V. 2011. *GrassPortal: an online ecological and evolutionary data facility*. [WWW document] URL [www.grassportal.org](http://www.grassportal.org) [accessed 12 June 2013].
- Patterson DJ, Cooper J, Kirk PM, Pyle RL, Remsen DP. 2010. Names are key to the new big biology. *Trends in Ecology and Evolution* 25: 686–691.
- Renvoize SA. 1986. A survey of leaf-blade anatomy in grasses: VIII. Arundinoideae. *Kew Bulletin* 41: 323–338.
- Renvoize SA. 1987. A survey of leaf-blade anatomy in grasses: XI. Paniceae. *Kew Bulletin* 42: 739–768.
- Sage RF, Li M, Monson RK. 1999a. The taxonomic distribution of C<sub>4</sub> photosynthesis. In: Sage RF, Monson RK, eds. *C<sub>4</sub> plant biology*. New York, NY, USA: Academic Press, 551–584.
- Sage RF, Wedin DA, Li M. 1999b. The biogeography of C<sub>4</sub> photosynthesis: patterns and controlling factors. In: Sage RF, Monson RK, eds. *C<sub>4</sub> plant biology*. New York, NY, USA: Academic Press, 313–373.
- Schulze ED, Ellis R, Schulze W, Trimborn P, Ziegler H. 1996. Diversity, metabolic types and  $\delta^{13}\text{C}$  carbon isotope ratios in the grass flora of Namibia in relation to growth form, precipitation and habitat conditions. *Oecologia* 106: 352–369.
- Sidlauskas B, Ganapathy G, Hazkani-Covo E, Jenkins KP, Lapp H, McCall LW, Price S, Scherle R, Spaeth PA, Kidd DM. 2009. Linking big: the continuing promise of evolutionary synthesis. *Evolution* 64: 871–880.

- The Plant List. 2010. *Version 1*. Published on the Internet. [WWW document] URL <http://www.theplantlist.org/> [accessed 15 March 2012].
- Vicentini A, Barber JC, Aliscioni SS, Giussani LM, Kellogg EA. 2008. The age of the grasses and clusters of origins of C<sub>4</sub> photosynthesis. *Global Change Biology* 14: 2963–2977.
- Watson L, Dallwitz MJ. 1992 onwards. *The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references*. [WWW document] URL <http://delta-intkey.com> [accessed 11 February 2012].
- Whitfield J. 2011. Species spellchecker fixes plant glitches. Online tool should weed out misspellings and duplications. *Nature* 474: 263.

## Supporting Information

Additional supporting information may be found in the online version of this article.

**Methods S1** Full methodology used to compile and index the dataset.

**Notes S1** Full list of literature sources used to compile the dataset presented in Tables S3 and S4.

**Table S1** Herbarium specimens of grass species analyzed for stable carbon isotope ratio

**Table S2** Uncharacterized species with countries of occurrence and synonyms (Clayton *et al.*, 2002a,b onwards)

**Table S3** Full genus-level dataset with numbers of species for each, and references for evidence

**Table S4** Full species-level dataset for the genus *Panicum* s.l., as circumscribed by Clayton *et al.* (2002b onwards), with the evidence and references used to ascribe photosynthetic type for each

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

**Key words:** C<sub>3</sub> photosynthesis, C<sub>4</sub> photosynthesis, comparative plant ecology, ecological informatics, evolutionary biology, grasses, Poaceae.



## About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <25 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **[www.newphytologist.com](http://www.newphytologist.com)** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office ([np-centraloffice@lancaster.ac.uk](mailto:np-centraloffice@lancaster.ac.uk)) or, if it is more convenient, our USA Office ([np-usaoffice@ornl.gov](mailto:np-usaoffice@ornl.gov))
- For submission instructions, subscription and all the latest information visit **[www.newphytologist.com](http://www.newphytologist.com)**