SHORT COMMUNICATION

# A molecular phylogeny of the genus *Alloteropsis* (Panicoideae, Poaceae) suggests an evolutionary reversion from $C_4$ to $C_3$ photosynthesis

## Douglas G. Ibrahim[1], Terry Burke[1], Brad S. Ripley[2] and Colin P. Osborne[1],*

[1]*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK* and [2]*Botany Department, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*

• *Background and Aims* The grass *Alloteropsis semialata* is the only plant species with both $C_3$ and $C_4$ subspecies. It therefore offers excellent potential as a model system for investigating the genetics, physiology and ecological significance of the $C_4$ photosynthetic pathway. Here, a molecular phylogeny of the genus *Alloteropsis* is constructed to: (*a*) confirm the close relationship between the $C_3$ and $C_4$ subspecies of *A. semialata*; and (*b*) infer evolutionary relationships between species within the *Alloteropsis* genus.

• *Methods* The chloroplast gene *ndhF* was sequenced from 12 individuals, representing both subspecies of *A. semialata* and all four of the other species in the genus. *ndhF* sequences were added to those previously sequenced from the Panicoideae, and used to construct a phylogenetic tree.

• *Key Results* The phylogeny confirms that the two subspecies of *A. semialata* are among the most recently diverging lineages of $C_3$ and $C_4$ taxa currently recognized within the Panicoideae. Furthermore, the position of the $C_3$ subspecies of *A. semialata* within the *Alloteropsis* genus is consistent with the hypothesis that its physiology represents a reversion from $C_4$ photosynthesis. The data point to a similar evolutionary event in the *Panicum stenodes–P. caricoides–P. mertensii* clade. The *Alloteropsis* genus is monophyletic and occurs in a clade with remarkable diversity of photosynthetic biochemistry and leaf anatomy.

• *Conclusions* These results confirm the utility of *A. semialata* as a model system for investigating $C_3$ and $C_4$ physiology, and provide molecular data that are consistent with reversions from $C_4$ to $C_3$ photosynthesis in two separate clades. It is suggested that further phylogenetic and functional investigations of the *Alloteropsis* genus and closely related taxa are likely to shed new light on the mechanisms and intermediate stages underlying photosynthetic pathway evolution.

**Key words:** *Alloteropsis semialata*, Panicoideae; Poaceae, *ndhF*, $C_4$ photosynthesis, $C_3$ photosynthesis, photosynthetic pathway evolution, molecular phylogeny.

## INTRODUCTION

Despite the multitude of changes to leaf anatomy and biochemistry required for the transition from $C_3$ to $C_4$ photosynthesis, this evolutionary event is estimated to have occurred independently in >48 plant lineages (Sage, 2004). Such a remarkable example of convergent evolution suggests either the action of very strong selective pressures, or that genetic mechanisms underlying the $C_3$ to $C_4$ transition are less complex than currently thought. Investigation of these selective pressures and genetic mechanisms requires comparison of closely related $C_3$ and $C_4$ plants in order to minimize confounding effects of different genetic backgrounds. $C_3$ and $C_4$ eudicot species have been used for this purpose (e.g. Brown *et al.*, 2005), but studies are also required in the grasses, half of whose species use the $C_4$ pathway (Sage *et al.*, 1999). Grasses have great ecological and commercial importance, covering a fifth of the vegetated land surface (Matthews, 1983) and providing more than half of digestible energy in human diets, as well as pasture for livestock (Evans, 1993). However, closely related $C_3$ and $C_4$ grasses are rare because

the ancient origins of $C_4$ photosynthesis within the Poaceae mean that substantial divergence has occurred between $C_3$ and $C_4$ lineages (Giussani *et al.*, 2001; Grass Phylogeny Working Group, 2001; Christin *et al.*, 2008). *Alloteropsis semialata* is unique in having both $C_3$ and $C_4$ subspecies, and therefore could provide a novel model system for comparative studies of $C_3$ and $C_4$ grasses.

Gibbs-Russell (1983) designated the $C_3$ and $C_4$ variants of *A. semialata* as subspecies rather than separate species because herbarium specimens included apparent intermediates between the $C_3$ and $C_4$ forms and therefore suggested there was gene flow between them. However, recent molecular phylogenies have shown some grass genera to be polyphyletic (e.g. *Panicum*; Duvall *et al.*, 2001; Aliscioni *et al.*, 2003), indicating that species assignment based on classical taxonomy can be misleading. The taxonomic relationship between the $C_3$ and $C_4$ subspecies of *A. semialata* therefore requires corroboration from molecular studies. In total, the *Alloteropsis* genus consists of five species and belongs to the grass subfamily Panicoideae (Clayton and Renvoize, 1986). This subfamily has considerable commercial importance, since its species include key food crops, such as maize (*Zea mays*) and

* For correspondence. E-mail c.p.osborne@sheffield.ac.uk

sugarcane (*Saccharum* sp.), as well as biofuel crops such as switchgrass (*Panicum virgatum*) and *Miscanthus* (Clayton and Renvoize, 1986). Furthermore, Panicoideae is an intriguing group for studying photosynthetic pathway evolution since it has $C_3$ and all three sub-types of $C_4$ photosynthesis represented among its species (Hattersley and Watson, 1992). A recent phylogeny indicates that the genus *Alloteropsis* occurs within a clade consisting of both $C_3$ and $C_4$ species (Christin *et al.*, 2008). Therefore, establishing relationships among this group and the five *Alloteropsis* species could illuminate the evolutionary routes leading to particular photosynthetic pathways.

Here, the aim was to: (*a*) confirm the close relationship between the $C_3$ and $C_4$ subspecies of *A. semialata*; and (*b*) infer evolutionary relationships between species within the *Alloteropsis* genus. A molecular phylogeny was constructed using the chloroplast locus *ndhF*, chosen because it has previously been sequenced for >150 Panicoideae species, and shows sufficiently high rates of molecular evolution to allow resolution between grass taxa at the species level (Clark *et al.*, 1995; Giussani *et al.*, 2001; Aliscioni *et al.*, 2003).

## MATERIALS AND METHODS

### DNA preparation

The specimens used for sequencing are detailed in the Supplementary Information (Table S1, available online). For DNA extraction from living specimens or from silica-dried leaf material, leaves were ground in liquid nitrogen, and genomic DNA was purified using the Plant DNAzol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. Whole genomic DNA from herbarium specimens was obtained from the DNA-Bank at the Royal Botanic Gardens, Kew, UK.

The *ndhF* gene was amplified via PCR using a *Taq*-mediated protocol (BioTaq, Bioline, London). Primer sequences were specified by Olmstead and Sweere (1994) and Aliscioni *et al.* (2003), or optimized for the *Alloteropsis* genus specifically for this project (Supplementary Information Table S2, available online). For DNA extracted from live specimens, *ndhF* was amplified in two overlapping fragments, 1F/1318R and 972F/2110R or 1143F/2110R. DNA from the herbarium specimens was degraded, so *ndhF* in these cases was amplified in smaller, overlapping fragments, 1F/536R, 216F/757R, 536F/972R, 757F/1318R, 1143F/1660R and 1606F/2110R. PCR was performed in 30 or 50 μL reactions in a DNA engine tetrad thermocycler (MJ Research, Cambridge, MA). PCR products were separated from the reaction mixture by band cutting after electrophoresis on a 1·5 % agarose gel and purification using Wizard SV gel and a PCR clean-up system (Promega, Madison, WI). PCR products were then quantified by comparison with DNA ladders of known size and concentration (GeneRuler 1 kb or 100 bp DNA ladder, Fermentas, Vilnius, Lithuania).

PCR products were sequenced in both directions using fluorescent dye terminators (ABI PRISM BigDye Terminator v1.1 Cycle Sequencing Kit, Applied Biosystems, Foster City, CA) in 10 μL reactions following conditions recommended by the manufacturer. For the short (<600 bp) PCR products derived from herbarium specimens, sequencing primers used

were the same as the PCR primers. For longer fragments, internal sequencing primers were also included. Extension products were precipitated using the manufacturer's ethanol/EDTA/sodium acetate-mediated protocol, suspended in 10 μL of Hi-Di formamide (Applied Biosystems) and electrophoresed on an ABI 3730 capillary sequencer (Applied Biosystems). Forward and reverse sequences were manually edited and checked using Bioedit v.7.0.5.2. (Hall, 1999), and overlapping fragments were assembled into the full *ndhF* sequence.

### Phylogenetic analysis

For preliminary analysis, the 12 sequences obtained from *Alloteropsis* specimens were aligned manually with 137 of the 156 *ndhF* sequences used by Aliscioni *et al.* (2003). Taxa used by Aliscioni *et al.* (2003) were excluded if information on their photosynthetic pathway was unavailable, or if the available *ndhF* sequences were incomplete (*Panicum piausense*, *P. pedersenii* and *Tatianyx arnacites*). A full list of all taxa used, together with the accession numbers of *ndhF* sequences and information on the photosynthetic pathway, is provided (Supplementary Information Table S3, available online). Gaps were treated as missing data in all analyses. Alternative evolutionary models for the resulting 149 taxa matrix were compared using ModelTest 3·7 (Posada and Crandall, 1998), which showed that the transversional nucleotide substitution model with gamma distribution of variation among sites and fixed proportion of invariant sites (TVM I + G) was optimal. Distances were calculated under this model with the neighbour-joining (NJ) method (Saitou and Nei, 1987), as implemented in PAUP v.4.0 b10 (Swofford, 2002), and NJ bootstrapping was then conducted using 1000 replicates. *Thysanolaena maxima*, *Danthoniopsis dinteri*, *Zeugites pittieri* and *Chasmanthium laxum* were chosen as outgroup taxa, following Aliscioni *et al.* (2003).

The preliminary NJ tree confirmed the position of *Alloteropsis* within the well-supported $x = 9$ clade of the Panicoideae and also that multiple accessions of an individual *Alloteropsis* species or subspecies were monophyletic. Therefore, Bayesian inference (BI) and maximum parsimony (MP) trees were constructed using only taxa from the $x = 9$ clade. To avoid taxonomic bias in further analysis of the $x = 9$ clade, a single specimen of each *Alloteropsis* species or subspecies was chosen. For *A. cimicina* and *A. papillosa*, one sequence of each was chosen at random: *A. cimicina* specimen A and *A. papillosa* specimen B (Supplementary Information Table S1). For *A. semialata* subspecies, sequences were chosen to represent the accessions used for other experiments (Ripley *et al.*, 2007, 2008; Ibrahim *et al.*, 2008; Osborne *et al.*, 2008), and therefore *A. semialata* subsp. *semialata* specimen A and *A. semialata* subsp. *eckloniana* specimen B (Supplementary Information Table S1) were used, although selection of alternative sequences did not change tree topology (data not shown). Based on the NJ analysis, four species were chosen as outgroup taxa for analysis of the $x = 9$ clade: *Zea mays*, *Sorghum bicolour*, *Panicum pilosum* and *Paspalum vaginatum*.

For the BI tree, evolutionary models were selected using MrModelTest 2·2 (Nylander, 2004), which showed that the best model was the general time-reversible nucleotide

substitution model with gamma distribution of variation among sites and fixed proportion of invariant sites (GTR I + G). Outgroup taxa were chosen as above. This model was implemented in MrBayes v.3.1 (Huelsenbeck and Ronquist, 2001) using 1 million generations, four chains, two runs, and trees sampled every 100 generations. The temperature parameter for chain heating was set to 0·05 to improve the acceptance rate for swapping between chains, and other parameters used default settings. The first 2500 trees from each run were discarded and the remainder combined to form a consensus tree.

The MP tree used the same data set as the BI tree. A parsimony ratchet (Nixon, 1999) as implemented in PAUPrat (Sikes and Lewis, 2001) with 15 replicates of 200 iterations, and 15 % of characters re-weighted at each iteration, was used to find the optimal tree in PAUP* v.4.0 b10. A total of 2905 of the 3000 resulting trees were shortest and of equal length, and were combined to form a strict consensus tree. Bootstrapping was then performed using 1000 replicates with 1000 random addition sequence replicates per bootstrap replicate and tree bisection–reconnection (TBR) branch swapping. The number of rearrangements per addition sequence replicate in bootstrapping was limited to $10^6$ to save computing time.

## RESULTS

### Alloteropsis *sequences*

In total, *ndhF* sequences ranging in length from 2070 to 2084 bp were obtained from 12 *Alloteropsis* specimens (Supplementary Information Table S1). The sequence for *A. paniculata* excluded 21 bp between nucleotide positions 520 and 541 where sequencing failed. Sequences were identical in the three accessions of *A. cimicina*, and did not differ at more than two nucleotide positions between accessions within *A. papillosa*, *A. semialata* subsp. *semialata* or *A. semialata* subsp. *eckloniana.* In addition, the *ndhF* sequence for *A. semialata* subsp. *semialata* specimen A was completely identical to that reported for *A. semialata* by Christin *et al.* (2008).

### *Preliminary NJ tree*

For preliminary analysis, an NJ tree was constructed using *ndhF* sequences from all *Alloteropsis* specimens and 137 other Panicoideae species (Fig. 1). The data matrix contained 149 taxa and 2073 characters. The topology of the NJ tree is congruent with the *ndhF* tree resolved by Aliscioni *et al.* (2003), and the combined *ndhF* and *rbcL* tree resolved by Christin *et al.* (2008; Fig. 1). The three main clades, Andropogoneae, $x = 10$ Paniceae and $x = 9$ Paniceae, are therefore well supported (Fig. 1). *Alloteropsis* has a basal chromosome number of nine (Liebenberg and Fossey, 2001) and, as expected, is resolved within the $x = 9$ Paniceae clade (Fig. 1). Additionally, replicate specimens for each *Alloteropsis* taxon emerged as sister taxa on the NJ tree (Fig. 1). Further analysis was therefore conducted using one specimen of each *Alloteropsis* taxon together with other species resolved in the $x = 9$ clade.

Multiple transitions between $C_3$ and $C_4$ photosynthesis are apparent across the NJ tree and the small, well-supported clade of $C_3$ and $C_4$ *Panicum* species (*P. mertensii*, *P. caricoides* and *P. stenodes*), the *P. mertensii* group, is of particular interest (Fig. 1). This clade was also identified in the MP *ndhF* tree of Aliscioni *et al.* (2003); however, these authors did not describe the photosynthetic pathway of these species, which represent the most recently diverging examples of $C_3$ and $C_4$ *Panicum* species within the phylogeny (Fig. 1). Furthermore, placement of the $C_4$ species *P. stenodes* basal to the clade suggests that *P. mertensii* may represent a reversion from $C_4$ to $C_3$ photosynthesis (Fig. 1).

### *The* x = 9 *Paniceae clade*

The BI and MP trees for the $x = 9$ Paniceae clade used 71 taxa and 2073 characters. The data set contained 1512 invariant characters and 306 parsimony-informative characters. The MP ratchet recovered 2905 trees all with a length of 803 steps, consistency index (CI) of 0·501 and retention index (RI) of 0·749 when uninformative characters were excluded. The topologies of BI and MP trees were identical, and the consensus BI tree shown in Fig. 2 therefore includes both posterior probabilities for clades and bootstrap values from the MP analysis.

Tree topology in the $x = 9$ Paniceae is similar to the preliminary NJ tree, and shows that *Alloteropsis* forms a monophyletic group (Fig. 2). This clade forms a polytomy with two sister clades that were together designated the 'forest shade clade' by Giussani *et al.* (2001) (Fig. 2). The *Alloteropsis*– forest shade group itself is well supported, although relationships within the group and among major clades in the $x = 9$ Paniceae are poorly resolved (Fig. 2). Within the genus *Alloteropsis*, species are split among two sister groups, with *A. semialata* more closely related to *A. angusta* than the other three *Alloteropsis* species (Fig. 2). In addition, *A. semialata* is a monophyletic species with the $C_3$ and $C_4$ subspecies as sister taxa (Fig. 2). Furthermore, the branch length between the two subspecies of *A. semialata* is short, indicating that they have undergone recent divergence.

### *Photosynthetic pathway evolution in* Alloteropsis

Poor resolution between clades in the $x = 9$ Paniceae means that the photosynthetic pathway in the common ancestor of the *Alloteropsis*–forest shade group is undetermined (Fig. 2). However, regardless of whether the common ancestor is $C_3$ or $C_4$, the most parsimonious explanation of photosynthetic pathway evolution in the *Alloteropsis*–forest shade group is that the $C_3$ subspecies of *A. semialata* represents a reversion from $C_4$ photosynthesis (Fig. 3). If the ancestor to the group was $C_3$, and reversions from $C_4$ photosynthesis have not occurred, then there must have been a minimum of four origins of $C_4$ photosynthesis within the clade (Fig. 3A). In contrast, if reversions have taken place, then a single origin and one subsequent reversion from $C_4$ photosynthesis could have occurred (Fig. 3B). If the common ancestor to the group was $C_4$, then reversions of photosynthetic type must have happened, and at least two incidences of this would have occurred within the group (Fig. 3C). The hypothesis of reversions from $C_4$ to $C_3$ photosynthesis therefore requires two fewer transitions in the photosynthetic pathway to explain the *ndhF* data than the alternative hypothesis of multiple $C_4$ origins.
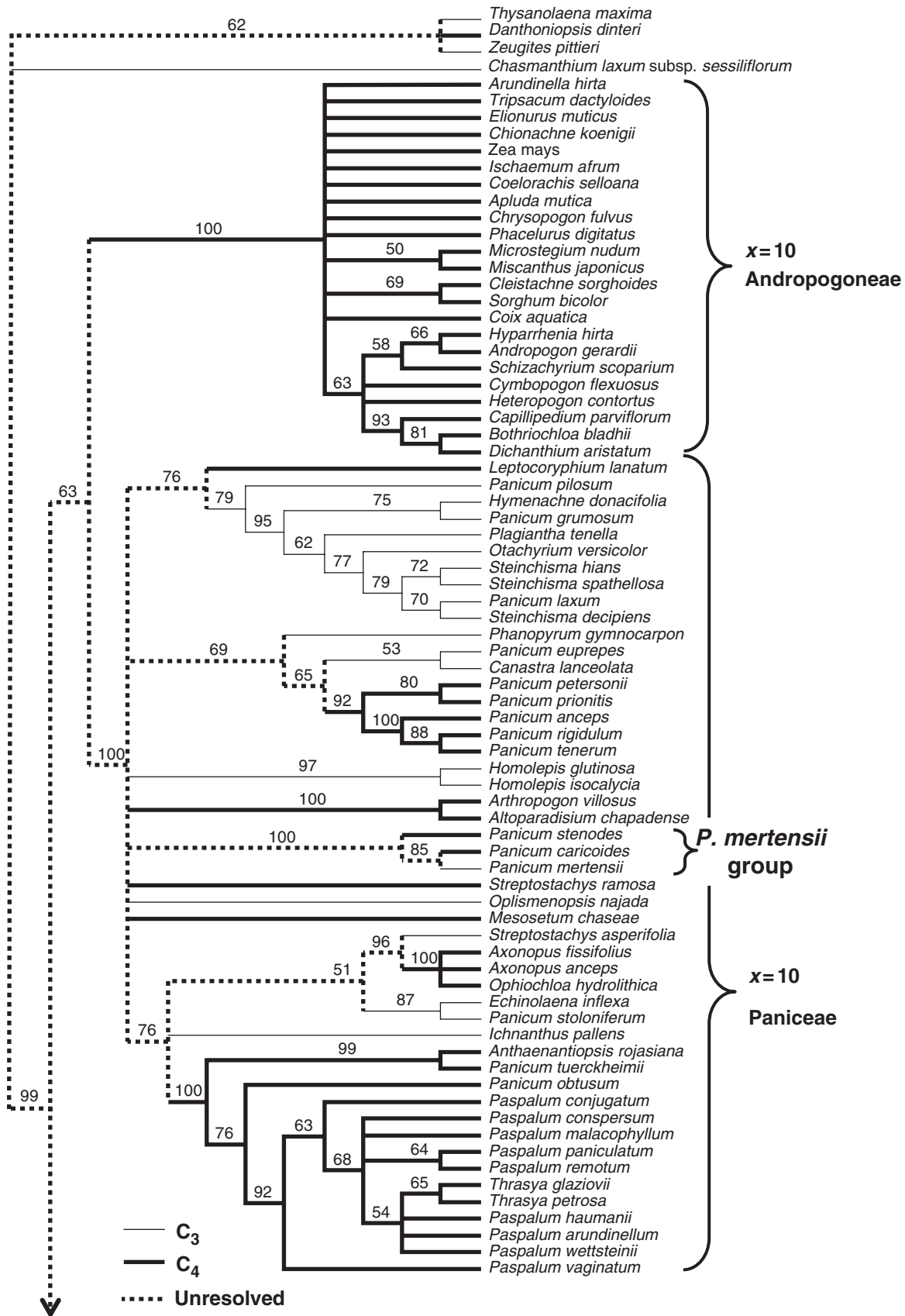
FIG. 1. Neighbour−joining cladogram for *ndhF* sequences in the Panicoideae. Percentage bootstrap values (1000 replicates) are shown above branches when they are >50 %. Major groupings of species are indicated. $C_3$ taxa are shown with narrow lines, $C_4$ taxa with bold lines, and the unresolved photosynthetic type of ancestral lineages with dotted lines. The figure continues on the next page.
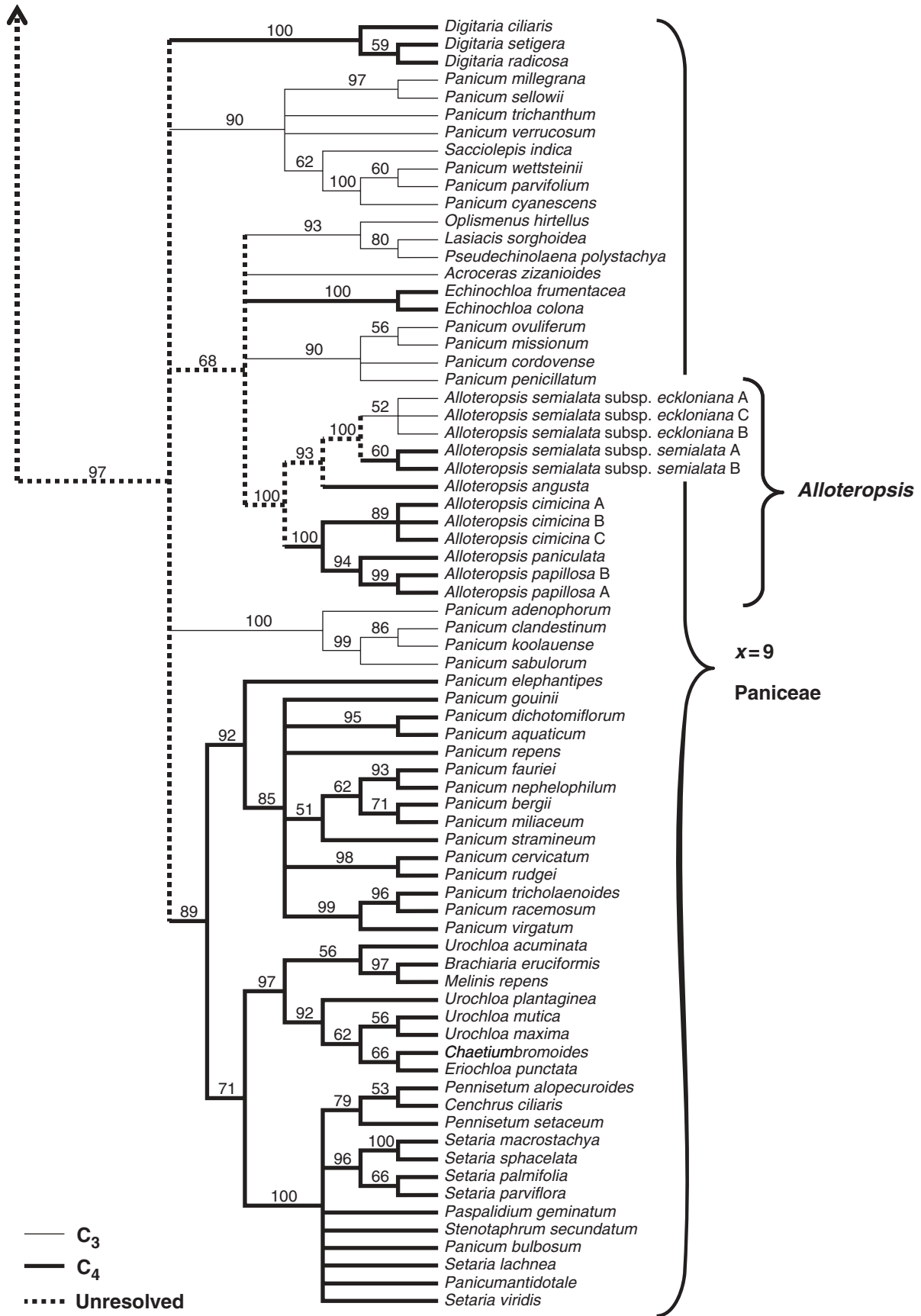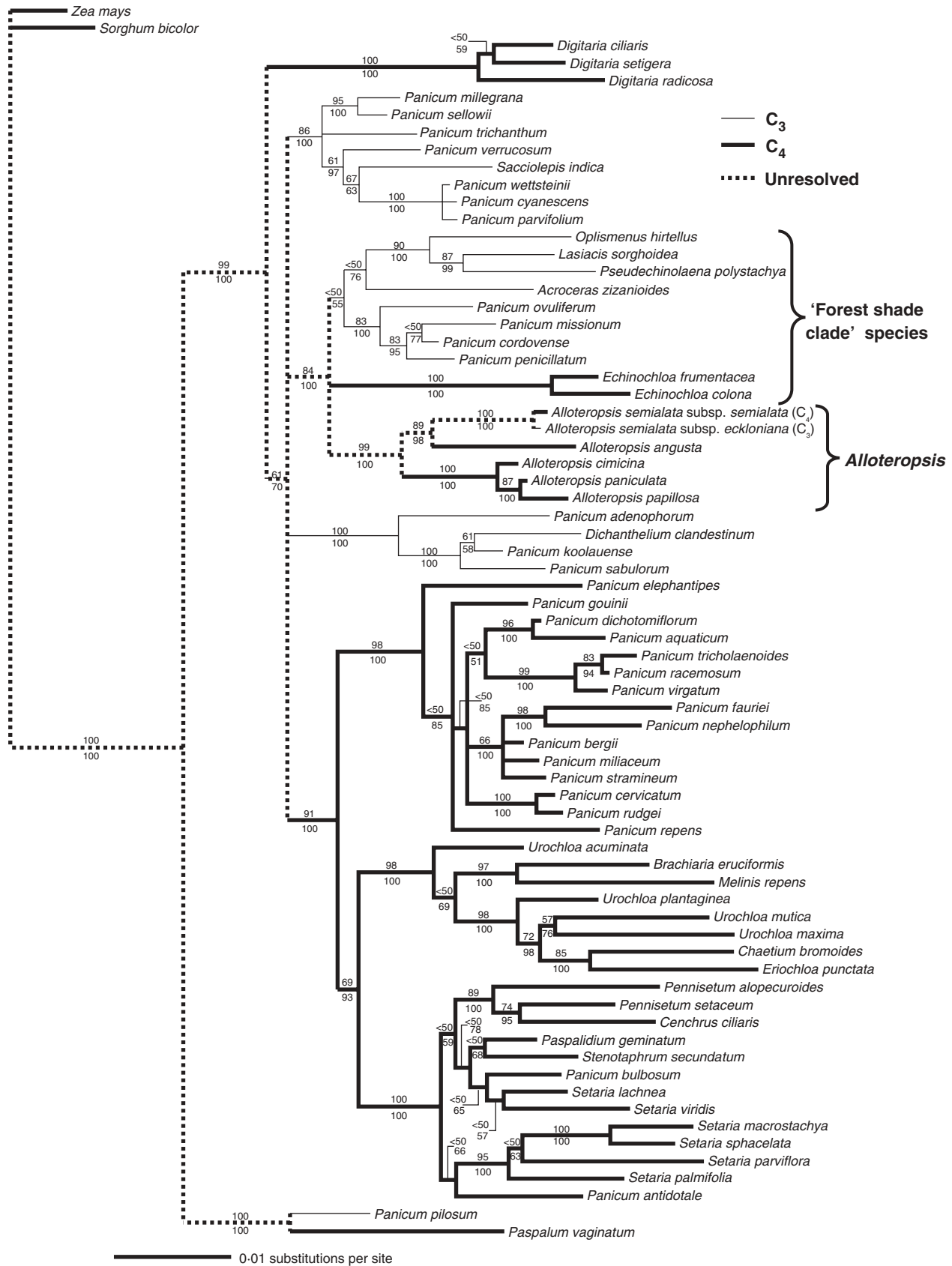
Fig. 1. Continued.

FIG. 2. Consensus Bayesian inference tree for the $x = 9$ clade of the Paniceae tribe. Values above and below branches indicate percentage bootstrap values from MP analysis and Bayesian posterior probabilities, respectively. C$_3$ taxa are shown with narrow lines, C$_4$ taxa with bold lines, and the unresolved photosynthetic type of ancestral lineages with dotted lines.

## (A) Ancestor C$_3$, reversions **not** allowed



## (B) Ancestor C$_3$, reversions allowed


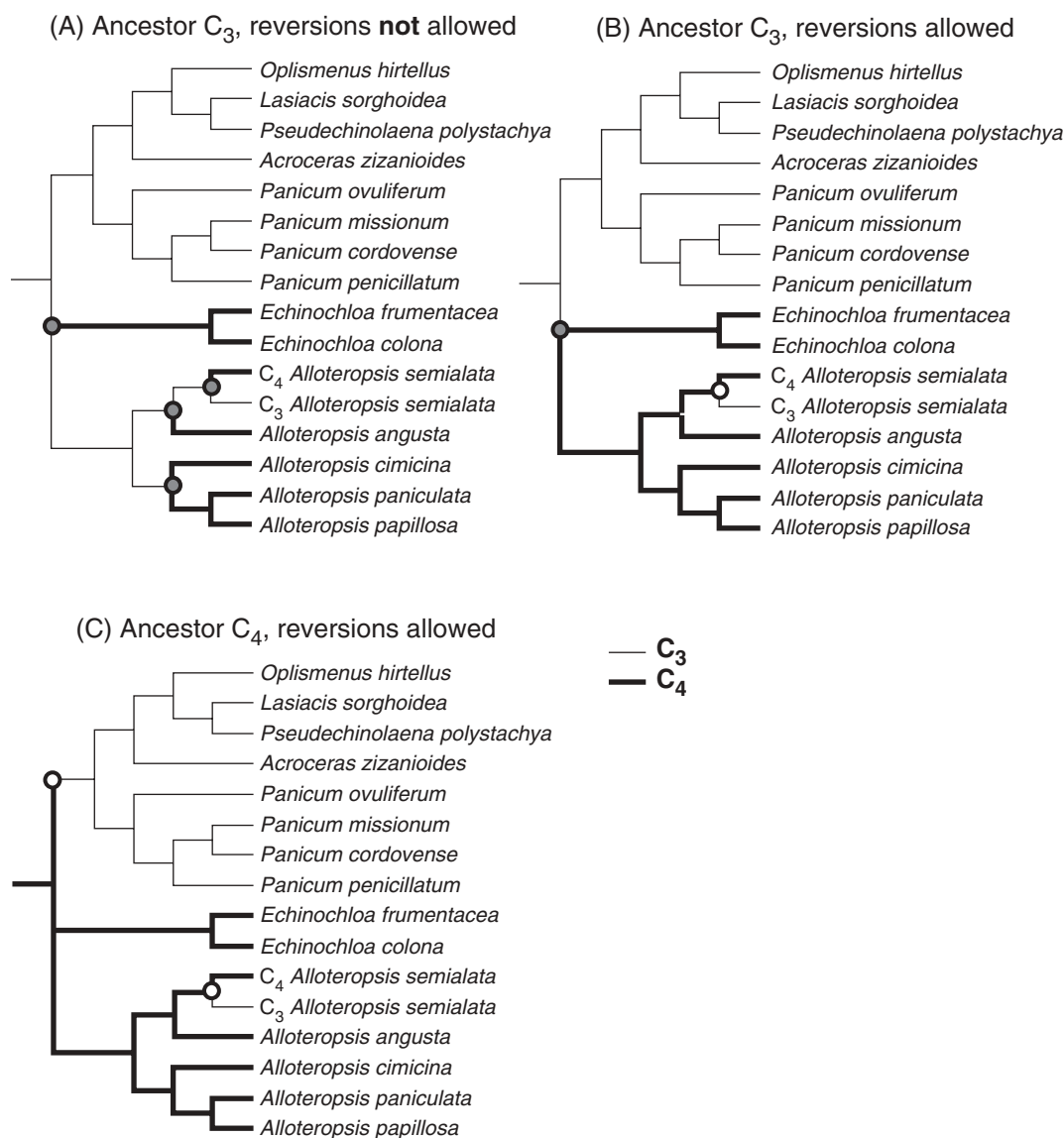
## (C) Ancestor C$_4$, reversions allowed



FIG. 3. Possible routes of photosynthetic pathway evolution in the *Alloteropsis*–forest shade group. The relationship between taxa is taken from Fig. 2. C$_3$ taxa are shown with narrow lines, and C$_4$ taxa with bold lines. Circles at nodes represent transitions in the photosynthetic pathway, where grey is a transition to C$_4$ and white is a transition to C$_3$.

The combined *ndhF* and *rbcL* phylogeny of Christin *et al.* (2008) indicates that *Echinochloa* occurs as sister to *Panicum ovuliferum*, and therefore represents a unique origin of C$_4$ photosynthesis. However, this relationship is only weakly supported (Christin *et al.*, 2008) and, regardless of the position of *Echinochloa*, the most parsimonious explanation for photosynthetic pathway evolution within the *Alloteropsis* genus requires a C$_4$ to C$_3$ transition to account for C$_3$ physiology in *A. semialata* subsp. *eckloniana* (Fig. 3).

### DISCUSSION

*Reversion from C$_4$ to C$_3$ photosynthesis*

The *ndhF* phylogeny confirms that the two subspecies of *A. semialata* are among the most recently diverging lineages

of C$_3$ and C$_4$ taxa currently recognized within the Panicoideae, and therefore that they can serve as a model system in which to compare gene expression, physiology and ecology of C$_3$ and C$_4$ grasses (Ripley *et al.*, 2007, 2008; Ibrahim *et al.*, 2008; Osborne *et al.*, 2008). Our data are also consistent with the hypothesis that the C$_3$ subspecies of *A. semialata* represents an evolutionary reversion from C$_4$ photosynthesis, and highlight the possibility of an additional reversal in the *P. stenodes*–*P. caricoides*–*P. mertensii* clade. These interpretations of the molecular data are consistent with observations of atypical leaf anatomy in the C$_3$ subspecies of *A. semialata*, where the inner bundle sheath cells contain a large number of chloroplasts, and mesophyll cells show a radiate arrangement around the vascular bundles (Ueno and Sentoku, 2006), traits that are usually associated with the C$_4$ pathway. Radiate mesophyll cells also occur in

*P. mertensii* (Renvoize, 1987). However, the relatively common occurrence of this anatomical arrangement among $C_3$ Paniceae species (Renvoize, 1987) suggests that caution should be exercised in interpreting it as evidence for photosynthetic pathway evolution.

Species of the *P. mertensii* clade do not share clear distinctive morphological characters; however, their leaf structure is intriguing because the $C_3$ species *P. mertensii* has lanceolate leaf laminae while its two $C_4$ counterparts have more linear laminae with a tendency to roll inwards (Aliscioni *et al.*, 2003). A similar difference in leaf morphology exists between the $C_3$ and $C_4$ subspecies of *A. semialata* (Gibbs-Russell, 1983), highlighting potential links between photosynthetic pathway evolution and changes to leaf anatomy and morphology.

Although the possibility of evolutionary reversions from $C_4$ to $C_3$ photosynthesis has been suggested for some time (Brown, 1977; Ellis, 1984; Hattersley and Watson, 1992; Pyankov *et al.*, 2001), the phenomenon has been controversial when interpreting molecular phylogenies (e.g. Duvall *et al.*, 2001; Giussani *et al.*, 2001). The data suggest that reversal of the photosynthetic pathway may occur, highlighting extraordinary lability for such a complex character. This interpretation assumes that the origin and loss of $C_4$ photosynthesis are equally likely. However, alternative scenarios are also possible: strong selection pressure towards gaining the pathway would reduce the probability of reversals; conversely, a complex character could be more likely to be lost than gained (e.g. Dollo parsimony). Such reversal of complex characters to ancestral states is receiving renewed interest in the wider field of evolutionary biology (Porter and Crandall, 2003; Whiting *et al.*, 2003), and reversions from the $C_4$ to $C_3$ photosynthetic pathway within the Panicoideae could serve as a further important example of this phenomenon.

### Evolution of the $C_4$ photosynthetic pathway

The most striking feature of the *Alloteropsis*–forest shade group is its remarkable degree of diversity in photosynthetic metabolism. As well as the $C_3$ and $C_4$ taxa indicated in Fig. 2, multiple subtypes of $C_4$ photosynthesis are present. *Echinochloa frumentacea* and *E. colona* are both of the NADP-malic enzyme (NADP-ME) subtype (Gutierrez *et al.*, 1974), while *A. semialata* subsp. *semialata* is of the phosphoenolpyruvate carboxylase (PCK) type (Prendergast *et al.*, 1987; Ueno and Sentoku, 2006). Intriguingly, even more diversity is suggested by intermediate biochemical features in *E. frumentacea*, which contains a significant amount of PCK enzyme despite being an NADP-ME-type $C_4$ species (Voznesenskaya *et al.*, 2006). Furthermore, *A. semialata* subsp. *semialata* exhibits higher NADP-ME activity than a 'classic' PCK species, *Urochloa maxima* (= *Panicum maximum*, Ueno and Sentoku, 2006). Photosynthetic biochemistry for other $C_4$ *Alloteropsis* species has not been analysed but, based on leaf anatomy, *A. papillosa*, *A. paniculata* and *A. cimicina* are predicted to be NAD-ME type (Hattersley and Watson, 1992), while *A. angusta* is expected to be NADP-ME type (Ellis, 1977). However, the inference of biochemistry in *A. angusta* should be treated with caution since it shares similar leaf anatomy to *A. semialata*, which was also

initially classified as NADP-ME on the basis of anatomical observations alone (Ellis, 1977).

The split of *Alloteropsis* into two sister groups in the *ndhF* phylogeny is consistent with their leaf anatomy, since the Kranz sheaths in both *A. semialata* and *A. angusta* are derived from the mestome sheath (XyMS– type anatomy; Hattersley and Watson, 1976), while in *A. papillosa*, *A. paniculata* and *A. cimicina* they are derived from the parenchyma sheath (XyMS+ type anatomy; Hattersley and Watson, 1976; Brown, 1977; Ellis, 1977, 1981). These differences in leaf anatomy correspond to the segregation of *Alloteropsis* into two species complexes, an 'Alloteropsis' group consisting of *A. semialata* and *A. angusta*, and a 'Coridochloa' group of *A. cimicina*, *A. paniculata* and *A. papillosa* (Hattersley and Watson, 1992), and the *ndhF* phylogeny therefore supports such a grouping.

XyMS– leaf anatomy is typically associated with a single bundle sheath and the NADP-ME subtype of $C_4$ biochemistry (Hattersley and Watson, 1992). However, *A. semialata* subsp. *semialata* is atypical because it has a complete layer of parenchyma cells surrounding the photosynthetically active mestome sheath (Ellis, 1974), and shows PCK-type biochemistry (Ueno and Sentoku, 2006). Leaf anatomy in *A. angusta* is similar to that in *A. semialata* subsp. *semialata*, but the parenchyma sheath cells are much reduced (Ellis, 1977). Placement of *A. semialata* and *A angusta* as sister species therefore suggests that their leaf anatomies represent intermediate stages during transition from typical XyMS+ to typical XyMS– anatomy through a pathway involving transfer of photosynthetic activity from parenchyma sheath cells to the mestome sheath, and subsequent loss of the parenchyma. In this scheme, *A. angusta* is further advanced towards conventional XyMS– physiology than the $C_4$ subspecies of *A. semialata*. However, this hypothetical evolutionary scenario depends critically on the type of photosynthetic biochemistry operating in *A. angusta*.

### Future directions

The present data suggest that coupling further phylogenetic and functional investigations of the *Alloteropsis*–forest shade group could provide dramatic and revealing insights into the genetic basis of photosynthetic pathway evolution. However, the current data should be treated with caution because phylogenies based on chloroplast genes reveal only maternal evolutionary history, and can differ from the underlying species phylogeny due to hybridization and introgression (Small *et al.*, 2004). These processes seem particularly likely in *Alloteropsis*, since previous surveys of chromosome number have shown that the $C_3$ subspecies of *A. semialata* is always diploid ($2n = 18$), while the $C_4$ subspecies is always hexaploid or higher ($2n = 54, 72, 108$), raising the possibility of a historical hybridization event accompanied by genome duplication (Liebenberg and Fossey, 2001). The possibility that $C_4$ photosynthesis in *A. semialata* subsp. *semialata* arose via hybridization with a $C_4$ sister species therefore cannot be excluded. However, even if this hybridization event occurred, the most parsimonious explanation for the $C_3$ photosynthetic pathway in *A. semialata* subsp. *eckloniana* remains unchanged; that it evolved via reversal from a $C_4$ progenitor. This scenario for the evolutionary loss of a complex trait, followed by its

re-acquisition, is unlikely, but not without precedent in evolutionary biology (Whiting *et al.*, 2003).

Clarification is also required of phylogenetic relationships among anatomical variants of *A. semialata* subsp. *semialata* (Renvoize, 1987), and the $C_3 - C_4$ intermediates that intergrade morphologically with the *A. semialata* subspecies (Hattersley and Watson, 1992). These occupy an approximately intermediate geographical distribution (Hattersley and Watson, 1992), and may represent intermediates in an actively evolving transition between photosynthetic types, or hybridization between the subspecies (Hattersley and Watson, 1992).

Further molecular markers from different cellular compartments will therefore be critical to improve robustness of the phylogeny, and unravel any hybridization events. In addition, the species of Paniceae sequenced thus far are biased towards New World taxa, and the addition of more African species would also be likely to improve resolution within the *Alloteropsis*–forest shade group. Finally, the atypical leaf anatomy and PCK-type $C_4$ biochemistry in *A. semialata* mean that assigning $C_4$ subtypes to the rest of the genus based on leaf anatomy is ill advised. It is suggested that the biochemical characterization of other *Alloteropsis* species, particularly *A. angusta*, is likely to provide novel insights into the $C_4$ evolutionary pathway.

## SUPPLEMENTARY DATA

Supplementary data is available at *Annals of Botany* Online and consists of the following: Table S1, details of *Alloteropsis* specimens used for sequencing; Table S2, primer sequences for amplification and sequencing of *ndhF*; and Table S3, details of taxa used in constructing the *ndhF* phylogeny.

## ACKNOWLEDGEMENTS

## LITERATURE CITED

**Aliscioni SS, Giussani LM, Zuloaga FO, Kellogg EA. 2003.** A molecular phylogeny of *Panicum* (Poaceae: Paniceae): tests of monophyly and phylogenetic placement within the Panicoideae. *American Journal of Botany* **90**: 796–821.

**Brown NJ, Parsley K, Hibberd JM. 2005.** The future of $C_4$ research – maize, *Flaveria* or *Cleome*? *Trends in Plant Science* **10**: 215–221.

**Brown WV. 1977.** The Kranz syndrome and its subtypes in grass systematics. *Memoirs of the Torrey Botanical Club* **23**: 1–97.

**Brown WV, Smith BN. 1975.** The genus *Dichanthelium* (Gramineae). *Bulletin of the Torrey Botanical Club* **102**: 10–13.

**Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen V, Salamin N. 2008.** Oligocene $CO_2$ decline promoted $C_4$ photosynthesis in grasses. *Current Biology* **18**: 37–43.

**Clark LG, Zhang WP, Wendel JF. 1995.** A phylogeny of the grass family (Poaceae) based on *ndhF* sequence data. *Systematic Botany* **20**: 436–460.

**Clayton WD, Renvoize SA. 1986.** *Genera Graminum*. Grasses of the world. *Kew Bulletin, Additional Series* **13**: 1–389.

**Duvall MR, Noll JD, Minn AH. 2001.** Phylogenetics of Paniceae (Poaceae). *American Journal of Botany* **88**: 1988–1992.

**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.

**Ellis RP. 1977.** Distribution of the Kranz syndrome in the southern African Eragrostoideae and Panicoideae according to bundle sheath anatomy and cytology. *Agroplantae* **9**: 73–110.

**Ellis RP. 1981.** *Relevance of comparative leaf anatomy in taxonomic and functional research on the South African Poaceae*. DSc Thesis, University of Pretoria, South Africa.

**Ellis RP. 1984.** *Eragrostis walterii*, a first record of non-Kranz leaf anatomy in the subfamily Chloridoideae (Poaceae). *South African Journal of Botany* **3**: 380–386.

**Evans LT. 1993.** *Crop evolution, adaptation and yield*. Cambridge: Cambridge University Press.

**Gibbs-Russell GE. 1983.** The taxonomic position of $C_3$ and $C_4$ *Alloteropsis semialata* (Poaceae) in southern Africa. *South African Journal of Botany* **2**: 250–250.

**Giussani LM, Cota-Sanchez JH, Zuloaga FO, Kellogg EA. 2001.** A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of $C_4$ photosynthesis. *American Journal of Botany* **88**: 1993–2012.

**Grass Phylogeny Working Group. 2001.** Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* **88**: 373–457.

**Gutierrez M, Gracen VE, Edwards GE. 1974.** Biochemical and cytological relationships in $C_4$ plants. *Planta* **119**: 279–300.

**Hall TA. 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.

**Hattersley PW, Watson L. 1976.** $C_4$ grasses: an anatomical criterion for distinguishing between NADP-malic enzyme species and PCK or NAD-malic enzyme species. *Australian Journal of Botany* **24**: 297–308.

**Hattersley PW, Watson L. 1992.** Diversification of photosynthesis. In: Chapman GP, ed. *Grass evolution and domestication*. Cambridge: Cambridge University Press, 38–116.

**Huelsenbeck JP, Ronquist F. 2001.** MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

**Ibrahim DG, Gilbert ME, Ripley BS, Osborne CP. 2008.** Seasonal differences in photosynthesis between the $C_3$ and $C_4$ subspecies of *Alloteropsis semialata* are offset by frost and drought. *Plant, Cell and Environment* **31**: 1038–1050.

**Liebenberg EJL, Fossey A. 2001.** Comparative cytogenetic investigation of the two subspecies of the grass *Alloteropsis semialata* (Poaceae). *Botanical Journal of the Linnean Society* **137**: 243–248.

**Matthews E. 1983.** Global vegetation and land use: new high-resolution data bases for climate studies. *Journal of Climate and Applied Meteorology* **22**: 474–487.

**Nixon KC. 1999.** The Parsimony Ratchet, a new method for rapid parsimony analysis. *Cladistics* **15**: 407–414.

**Nylander JAA. 2004.** *MrModeltest v2.2*. Uppsala University, Uppsala, Sweden: Evolutionary Biology Centre.

**Olmstead RG, Sweere JA. 1994.** Combining data in phylogenetic systematics – an empirical approach using 3 molecular data sets in the Solanaceae. *Systematic Biology* **43**: 467–481.

**Osborne CP, Wythe EJ, Ibrahim DG, Gilbert ME, Ripley BS. 2008.** Low temperature effects on leaf physiology and survivorship in the $C_3$ and $C_4$ subspecies of *Alloteropsis semialata*. *Journal of Experimental Botany* **59**: 1743–1754.

**Porter ML, Crandall KA. 2003.** Lost along the way: the significance of evolution in reverse. *Trends in Ecology and Evolution* **18**: 541–547.

**Posada D, Crandall KA. 1998.** MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

**Prendergast HDV, Hattersley PW, Stone NE. 1987.** New structural biochemical associations in leaf blades of C$_4$ grasses (Poaceae). *Australian Journal of Plant Physiology* **14**: 403–420.

**Pyankov VI, Artyusheva EG, Edwards GE, Black CC, Soltis DE. 2001.** Phylogenetic analysis of the tribe Salsoleae (Chenopodiaceae) based on ribosomal ITS sequences: implications for the evolution of photosynthesis types. *American Journal of Botany* **88**: 1189–1198.

**Renvoize SA. 1987.** A survey of leaf-blade anatomy in grasses XI. Paniceae. *Kew Bulletin* **42**: 739–768.

**Ripley BS, Gilbert ME, Ibrahim DG, Osborne CP. 2007.** Drought constraints on C$_4$ photosynthesis: stomatal and metabolic limitations in C$_3$ and C$_4$ subspecies of *Alloteropsis semialata*. *Journal of Experimental Botany* **58**: 1351–1363.

**Ripley BS, Abraham TI, Osborne CP. 2008.** Consequences of C$_4$ photosynthesis for the partitioning of growth: a test using C$_3$ and C$_4$ subspecies of *Alloteropsis semialata* under nitrogen limitation. *Journal of Experimental Botany* **59**: 1705–1714.

**Sage RF. 2004.** The evolution of C$_4$ photosynthesis. *New Phytologist* **161**: 341–370.

**Sage RF, Li M, Monson RK. 1999.** The taxonomic distribution of C$_4$ photosynthesis. In: Sage RF, Monson RK, eds. *C$_4$ plant biology*. San Diego: Academic Press, 551–584.

**Saitou N, Nei M. 1987.** The neighbor joining method – a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.

**Sikes DS, Lewis PO. 2001.** PAUPRat: PAUP* implementation of the parsimony ratchet. Beta software, version 1. Distributed by the authors. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA http://users.iab.uaf.edu/~derek_sikes/software2.htm (accessed 19 May 2008).

**Small RL, Cronn RC, Wendel JF. 2004.** Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* **17**: 145–170.

**Spangler R, Zaitchik B, Russo E, Kellogg E. 1999.** Andropogoneae evolution and generic limits in *Sorghum* (Poaceae) using *ndhF* sequences. *Systematic Botany* **24**: 267–281.

**Swofford DL. 2002.** *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.* Sunderland, MA: Sinauer Associates.

**Ueno O, Sentoku N. 2006.** Comparison of leaf structure and photosynthetic characteristics of C$_3$ and C$_4$ *Alloteropsis semialata* subspecies. *Plant, Cell and Environment* **29**: 257–268.

**Voznesenskaya EV, Franceschi VR, Chuong SDX, Edwards GE. 2006.** Functional characterization of phosphoenolpyruvate carboxykinase type C$_4$ leaf anatomy: immuno-, cytochemical and ultrastructural analyses. *Annals of Botany* **98**: 77–91.

**Whiting MF, Bradler S, Maxwell T. 2003.** Loss and recovery of wings in stick insects. *Nature* **421**: 264–267.