Supplementary material

Contrasted histories of organelle and nuclear genomes underlying physiological diversification in a grass species

Matheus E. Bianconi, Luke T. Dunning, Emma V. Curran, Oriane Hidalgo, Robyn F. Powell, Sahr Mian, Ilia J. Leitch, Marjorie R. Lundgren, Sophie Manzi, Maria S. Vorontsova, Guillaume Besnard, Colin P. Osborne, Jill K. Olofsson, Pascal-Antoine Christin

This Supplementary material contains six figures and two tables. Table S1 is in a separate .xlsx file.
**Fig. S1.** Habitat diversity of *Alloteropsis semialata*. a) inflorescence of a $C_4$ hexaploid population in Mozambique (MOZ1601); b) tussocks of a $C_3$ diploid population in South Africa (RSA2); c) grassland in northern Australia where $C_4$ diploids can be found (population AUS4 in Olofsson et al. 2019); d) $C_4$ hexaploids in the lowland dry miombo woodlands of northern Mozambique (MOZ1601); e) $C_4$ hexaploids in the highland wet miombo woodlands of northwestern Zambia (ZAM1507); f) miombo woodlands in central Tanzania, where diploid $C_3+C_4$ populations can be found (TAN2).
Fig. S2. Time-calibrated phylogeny of *Alloteropsis* based on mitochondrial (left) and plastid (right) genome sequences. Brown bars on nodes indicate 95% HPD. Circles highlight nodes with posterior probability $\geq 0.95$. The organelle clades are indicated with letters (A-G).
Fig. S3. Multigene coalescent species trees estimated from two datasets with less stringent filtering; a) alignments not trimmed (4,345 genes retained); b) sites with more than 70% of missing data were discarded (4,198 genes retained). Pie charts on nodes indicate the proportion of quartet trees that support the main (dark grey), first (pale blue) and second (light grey) alternative topologies. Local posterior probabilities are indicated near nodes. Branch lengths are in coalescent units.
Fig. S4. Multigene coalescent species tree based on 3,553 genes. The gene set is as in Fig. 2, but only diploid individuals were retained. Pie charts on nodes indicate the proportion of quartet trees that support the main (dark grey), first (pale blue) and second (light grey) alternative topologies. Local posterior probabilities are indicated near nodes. Branch lengths are in coalescent units.
Fig. S5. Principal component analysis using genome-wide nuclear data of *Allotropis semialata*. a) The first and second axes (left), and the second and third axes (right). Colours correspond to major nuclear clades as identified in Fig. 2; b) ranked eigenvectors.
**Fig. S6.** Genetic structure of *Allotropis semialata*. a) Individual-based admixture analysis assigning individuals to the two best numbers of clusters (*K*). Nuclear clades (I-IV) and organelle lineages (A-G) are indicated at the bottom. b) (left) Mean likelihood (± SD) over five runs for each number of clusters tested for the admixture analysis (*K*:1-10); (right) fit improvement for the admixture analysis, as calculated according to Evanno *et al.* (2005).
**Table S2.** Genome composition analysis.

| Accession  | Nuclear clade | I     | II    | III   | IV    | Total³ |
|------------|---------------|-------|-------|-------|-------|--------|
| RSA5-03    | I             | 4750  | 273   | 77    | 37    | 5137   |
| RSA8-06    | I             | 675   | 60    | 10    | 6     | 751    |
| RSA6-09    | I             | 2233  | 119   | 24    | 14    | 2390   |
| RSA2365-1  | I             | 1552  | 65    | 20    | 19    | 1656   |
| ZIM1504-61 | I             | 1150  | 225   | 55    | 49    | 1479   |
| TAN1-04B   | II            | 346   | 8304  | 769   | 226   | 9645   |
| RSA2365-1  | II            | 675   | 60    | 10    | 6     | 751    |
| RSA6-09    | II            | 2233  | 119   | 24    | 14    | 2390   |
| RSA2365-1  | III           | 2233  | 119   | 24    | 14    | 2390   |
| ZIM1504-61 | III           | 1150  | 225   | 55    | 49    | 1479   |
| TAN1-04B   | III           | 346   | 8304  | 769   | 226   | 9645   |
| RSA2365-1  | IV            | 1150  | 225   | 55    | 49    | 1479   |
| ZIM1504-61 | IV            | 346   | 8304  | 769   | 226   | 9645   |
| RSA2365-1  | IV            | 1150  | 225   | 55    | 49    | 1479   |
| ZIM1504-61 | IV            | 346   | 8304  | 769   | 226   | 9645   |

² Phylogenetic assignment of alleles to each of the four major nuclear clades of *Alloteropsis semialata* (number of reads, with corresponding percentage between parentheses); ³ total number of reads assigned; ⁴ polyploid (6x); ⁵ polyploid (12x).
Supplementary references

Dunning LT, Lundgren MR, Moreno-Villena JJ, Namaganda M, Edwards EJ, Nosil P, Osborne CP, Christin P-A. 2017. Introgression and repeated co-option facilitated the recurrent emergence of C₄ photosynthesis among close relatives. *Evolution* **71**: 1541–1555.

Dunning LT, Moreno-Villena JJ, Lundgren MR, Dionora J, Salazar P, Adams C, Nyirenda F, Olofsson JK, Mapaura A, Grundy IM, *et al.* 2019a. Key changes in gene expression identified for different stages of C₄ evolution in *Alloteropsis semialata*. *Journal of Experimental Botany* **70**: 3255–3268.

Dunning LT, Olofsson JK, Parisod C, Choudhury RR, Moreno-Villena JJ, Yang Y, Dionora J, Paul Quick W, Park M, Bennetzen JL, *et al.* 2019b. Lateral transfers of large DNA fragments spread functional genes among grasses. *Proceedings of the National Academy of Sciences, USA* **116**: 4416–4425.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* **14**: 2611–2620.

Lundgren MR, Besnard G, Ripley BS, Lehmann CER, Chatelet DS, Kynast RG, Namaganda M, Vorontsova MS, Hall RC, Elia J, *et al.* 2015. Photosynthetic innovation broadens the niche within a single species. *Ecology Letters* **18**: 1021–1029.

Lundgren MR, Christin P-A, Escobar EG, Ripley BS, Besnard G, Long CM, Hattersley PW, Ellis RP, Leegood RC, Osborne CP. 2016. Evolutionary implications of C₃–C₄ intermediates in the grass *Alloteropsis semialata*. *Plant, Cell & Environment* **39**: 1874–1885.

Lundgren MR, Dunning LT, Olofsson JK, Moreno-Villena JJ, Bouvier JW, Sage TL, Khoshravesh R, Sultmanis S, Stata M, Ripley BS, *et al.* 2019. C₄ anatomy can evolve via a single developmental change. *Ecology Letters* **22**: 302–312.

Olofsson JK, Bianconi M, Besnard G, Dunning LT, Lundgren MR, Holota H, Vorontsova MS, Hidalgo O, Leitch IJ, Nosil P, *et al.* 2016. Genome biogeography reveals the intraspecific spread of adaptive mutations for a complex trait. *Molecular Ecology* **25**: 6107–6123.

Olofsson JK, Dunning LT, Lundgren MR, Barton HJ, Thompson J, Cuff N, Ariyarathne M, Yakandawala D, Sotelo G, Zeng K, *et al.* 2019. Population-specific selection on standing variation generated by lateral gene transfers in a grass. *Current Biology* **29**: 3921–3927.