Supporting Information

Genomic footprints of repeated evolution of CAM photosynthesis in a Neotropical species radiation

MARYLAURE DE LA HARPE[1,2#], MARGOT PARIS[1,2#*], JAQUELINE HESS[1#], MICHAEL HARALD JOHANNES BARFUSS[1], MARTHA LILIANA SERRANO-SERRANO[3], ARINDAM GHATAK[4,5], PALAK CHATURVEDI[4,5], WOLFRAM WECKWERTH[4,5], WALTER TILL[1], NICOLAS SALAMIN[3], CHING MAN WAI[6], RAY MING[7], CHRISTIAN LEXER[1,2]

[#] Marylaure de la Harpe, Margot Paris and Jaqueline Hess should be considered joint first author

*Corresponding author: margot.paris@unifr.ch

[1]University of Vienna, Faculty of Life Sciences, Department of Botany and Biodiversity Research, Division of Systematic and Evolutionary Botany, Rennweg 14, A-1030 Vienna, Austria, [2]University of Fribourg, Department of Biology, Unit of Ecology & Evolution, Chemin du Musée 10, CH-1700 Fribourg, Switzerland, [3]University of Lausanne, Faculty of Biology and Medicine, Department of Ecology and Evolution, Biophore, 1015 Lausanne, Switzerland, [4]University of Vienna, Faculty of Life Sciences, Department of Functional and Evolutionary Ecology, Division of Molecular Systems Biology (MOSYS), Althanstraße 14 (UZA I), 1090 Vienna, Austria, [5]University of Vienna, Vienna Metabolomics Center (VIME), Althanstraße 14 (UZA I), 1090 Vienna, Austria, [6]Michigan State University, College of Agriculture & Natural Resources, Department of Horticulture, 1066 Bogue Street, East Lansing, MI 48824, U.S.A, [7]University of Illinois at Urbana-Champaign, School of Integrative Biology, Department of Plant Biology, 265 Morrill Hall, MC-116, 505 South Goodwin Avenue, Urbana, IL 61801, U.S.A.

Supplemental Text

1) Carbon isotope phenotyping

Whole tissue carbon isotope ratios ($^{13}C/^{12}C$) can be used to characterize typical C_3 and CAM plants, especially in combination with distinct morphological and anatomical features as are present in bromeliads (**Fig. 1**). Carbon isotope ratios recovered for the studied species indicate a continuum of values ranging from typical C_3 to fairly strong CAM (**Fig. 1**), following commonly used thresholds. Many species in our sample set displayed typical C_3 carbon isotope ($\delta^{13}C$) phenotypes far beyond -20‰ and in fact reaching as far into the C_3 extreme as -30‰ (labelled green in **Fig. 1**). *Tillandsia australis (Taust),* our C_3 reference taxon also used for transcriptome-wide expression profiling (below), exhibited $\delta^{13}C$ values of

-26 to -29‰, clearly beyond the -20‰ threshold commonly used to classify C_3 plants. This species also exhibited all other phenotypic features expected for C_3 bromeliads, including tank-forming rosettes, no succulence, and absence of dense trichome cover. On the other end of the C_3/CAM continuum, *T. ionantha* exhibited a $\delta^{13}C$ value of only -13.9‰, indicating it represents a so-called ´strong´ CAM species (labelled in yellow in **Fig. 1**). The three CAM taxa sampled for expression profiling in our study (below) exhibited a broad range of CAM-like $\delta^{13}C$ values, from weak CAM in *T. floribunda* (*Tflor*, $\delta^{13}C$ = -18.6 to -20.3‰) to relatively strong CAM in *T. sphaerocephala* (*Tspha*, $\delta^{13}C$ = 15.2 to -16.2‰) and *T. fasciculata* (*Tfasc*, $\delta^{13}C$ = -14.5 to -18.3); species exhibiting pronounced, strong CAM typically exhibit $\delta^{13}C$ values that are less negative than -20‰.

Although measuring night-time acidity under drought stress is preferable for distinguishing true C_3 species from inducible, facultative CAM species, we opted for carbon isotope ratios based on clear restrictions presented by our study system. First, our sampling of highly divergent phenotypic forms made it challenging to derive comparable 'common garden' drought conditions across all species. Second, our work relied on precious living collections in botanical gardens and experimentation thus required careful consideration to avoid the loss of individual accessions. While we cannot exclude that some species classified as C_3-like here include facultative CAM plants, we use isotopic ratios as a proxy to partition species according to the extremes of the distribution of this continuous phenotypic trait for our evolutionary analyses. This is a conservative and pragmatic strategy, since phenotyping error would likely diminish the signal-to-noise ratio.

2) Phylogenetic analyses and ancestral state reconstruction (ASR)

A total of 177 Tillandsioideae species (203 accessions) were used for the ASR analyses. The table below describes in detail the taxonomic distribution of the selected species. We used the

number of accepted species names provided by Gouda & Butcher (http://bromeliad.nl/bromNames/). This list is used for the 'Encyclopaedia of Bromeliads' (http://bromeliad.nl/encyclopedia/) and is updated on a daily basis.

TABLE representing the number of terminals used, and species studied in comparison to known species. Most recent species numbers are from: Gouda, E.J. & Butcher, D. (cont. updated) *A List of Accepted Bromeliaceae Names* [http://bromeliad.nl/bromNames/]. University Botanic Gardens, Utrecht (accessed: 08-06-2020).

rank	**taxon**	**terminals**	**spp. studied**	**spp. known**	**sp. coverage**	**C_3/CAM**
family	Bromeliaceae	210	184	3628	5%	
subfamily	Tillandsioideae	203	177	1488	12%	
A. tribe	Catopsideae	4	4	18	22%	
1. genus	***Catopsis*** Griseb.	**4**	**4**	**18**	**22%**	**C_3 only**
B. tribe	Glomeropitcairnieae	2	2	2	100%	
2. genus	***Glomeropitcairnia*** (Mez) Mez	**2**	**2**	**2**	**100%**	**C_3 only**
C. tribe	Vrieseeae	51	41	405	10%	
a. subtribe	Vrieseinae	25	22	291	8%	
	Vriesea group	19	16	249	6%	
3. genus	***Vriesea*** Lindl.	**16**	**13**	**231**	**6%**	**C_3 only**
4. genus	***Stigmatodon*** Leme, G.K. Br. & Barfuss	**3**	**3**	**18**	**17%**	**C_3 mostly**
	Alcantarea group	6	6	42	14%	
5. genus	***Alcantarea*** (E. Morren ex Mez) Harms	**6**	**6**	**41**	**15%**	**C_3 only**
6. genus	***Waltillia*** Leme, Barfuss & Halbritt.	**0**	**0**	**1**	**0%**	**C_3 only**
b. subtribe	Cipuropsidinae	26	19	114	17%	
	Cipuropsis group	17	10	16	63%	
7. genus	***Cipuropsis*** Ule	**2**	**2**	**3**	**67%**	**C_3 only**
8. genus	***Goudaea*** W. Till & Barfuss	**6**	**2**	**2**	**100%**	**C_3 only**
9. genus	***Josemania*** W. Till & Barfuss	**2**	**2**	**5**	**40%**	**C_3 only**
10. genus	***Mezobromelia*** L.B. Sm.	**6**	**3**	**5**	**60%**	**C_3 only**
11. genus	***Zizkaea*** W. Till & Barfuss	**1**	**1**	**1**	**100%**	**C_3 only**
	Werauhia group	9	9	98	9%	
12. genus	***Werauhia*** J.R. Grant	**6**	**6**	**93**	**7%**	**C_3 only**
13. genus	***Jagrantia*** Barfuss & W. Till	**1**	**1**	**1**	**100%**	**C_3 only**
14. genus	***Lutheria*** Barfuss & W. Till	**2**	**2**	**4**	**50%**	**C_3 only**
D. tribe	Tillandsieae	146	130	1063	12%	
	Tillandsia group	131	115	843	14%	
15. genus	***Tillandsia*** L.	**107**	**93**	**746**	**12%**	**C_3/CAM**
16. genus	***Barfussia*** Manzan. & W. Till	**4**	**3**	**3**	**100%**	**C_3 only**
17. genus	***Lemeltonia*** Barfuss & W. Till	**4**	**3**	**7**	**43%**	**C_3 mostly**
18. genus	***Pseudalcantarea*** (Mez) Pinzón & Barfuss	**3**	**3**	**3**	**100%**	**C_3 only**
19. genus	***Racinaea*** M.A. Spencer & L.B. Sm.	**10**	**10**	**79**	**13%**	**C_3 only**
20. genus	***Wallisia*** (Regel) E. Morren	**3**	**3**	**5**	**60%**	**C_3 mostly**
	Guzmania group	15	15	219	7%	
21. genus	***Guzmania*** Ruiz & Pav.	**13**	**13**	**215**	**6%**	**C_3 only**
22. genus	***Gregbrownia*** W. Till & Barfuss	**2**	**2**	**4**	**50%**	**C_3 only**

Although it seems we studied only a small proportion of the large genera (e.g. *Vriesea*, *Werauhia*, *Guzmania*, *Racinaea*, *Tillandsia*), the samples provided in the analysis were carefully selected and cover almost the whole taxonomic and morphological range within these genera.

We used two complementary methods for species tree estimation: ASTRAL, a coalescent-based summary method (**Fig. 1**) and RAxML which infers maximum likelihood (ML) based phylogenetic trees (**SI Fig. 2**). The ML tree was inferred using the program RAxML v8.228 with a GTRGAMMA model and 100 bootstrap replicates to determine branch support. For detailed settings used for estimation of the ASTRAL tree please refer to Materials and Methods in the main text.

Both trees are identical, except for the position of *Tillandsia disticha* which in the ASTRAL tree is placed sister to subgenus *Tillandsia*. In the RAxML tree the split involving this species is inferred as basal to all main *Tillandsia* clades.

To establish the macro-evolutionary framework for this study, we used the R Package diversitree (v.0.9-11) to reconstruct the ancestral states of photosynthetic syndrome on the largest currently available phylogenetic tree for tillandsioid bromeliads (Barfuss et al. 2016). This maximum likelihood (ML) tree comprised 210 taxa, representing roughly 30% taxon sampling. We used the Multiple State Speciation and Extinction (MuSSE) algorithm implemented in diversitree (FitzJohn 2012), coding photosynthetic syndrome as either 1 for C_3 photosynthesis, 2 for intermediate or 'Winter Holtum Zone' (WHZ, δ13C from -23.0‰ to -19.0‰), and 3 for CAM photosynthesis. The three states were inferred from published data (Crayn et al. 2015) and newly collected carbon isotope ratio values for a total of 27 species (Fig. 1). Character states for all included species are listed in SI Table 6. The sampling.f parameter in diversitree was used to represent the estimated fraction of species included in the phylogeny for each photosynthetic state: for C_3, 112/524=0.213740458, for WHZ 30/107=0.280373832, for CAM 68/227=0.299559471. We used maximum likelihood to compare the following models: null (all birth and death rates equal between states), full (all rates of speciation and extinction depend on character state), lambda (diversification rate λ varies between states), mu (extinction rate μ varies between states), lambda.mu (λ & μ vary,

but character transitions are ordered), and fit.unordered (λ & μ are constant, with full flexible transition process). Lambda (variable λ between states) was inferred as the best model based on likelihoods and AIC values (logLik= 829.6978, AIC= -1649.396), and ancestral states were thus constructed within this model.

	lnLik	logLik()	AIC()
fit.null	837.84	837.8432 (df=3)	-1669.69
fit.full	874.3	874.2964 (df=12)	-1724.59
fit.lambda	**829.7**	**829.6978 (df=5)**	**-1649.4**
fit.mu	847.18	847.1797 (df=5)	-1684.36
fit.lambda.mu	853.29	853.2904 (df=7)	-1692.58
fit.unordered	857.52	857.5183 (df=8)	-1699.04

3) CNV analyses: Detailed implementation of the CNVkit analysis

Relative copy numbers were estimated in a two-step approach: i) estimation of base copy number (CN) state in *Alcantarea trepida* through comparison with *Ananas comosus.* ii) estimation of CN in *Tillandsia* samples relative to *A. trepida,* scaling of Tillandsia CN with *A. trepida* base CN. For both analyses we applied a stringent coverage-based filtering to exclude unmappable regions since we cannot distinguish them from lost genes. Coverage cutoffs were set as follows: filtered_exon_set_1 (Aco-Atre analysis) – retain only exons with mean coverage of at least five, filtered_exon_set2 (Atre-Tillandsia analysis) – retain only exons with mean coverage of at least five in five or more species. This resulted in a total of 19,298 discoverable genes in filtered_exon_set2 (roughly 2/3rds of the genome).

In order to derive meaningful log2 thresholds for CNV calling, we used deeply conserved single-copy orthologs in the pineapple to estimate variation in coverage. To this end, we used BUSCO v3[7] to obtain the set of single-copy orthologs using the *A. comosus* predicted proteome and the "embryophyta_odb9" database. To derive per-sample single copy

thresholds we calculated the weighted average log2 ratios across exon bins for each of the BUSCO genes. The resulting distributions are shown below.

Species	**1st Qu.**	**Median**	**Mean**	**3rd. Qu.**	**2.50%**	**97.50%**
Tillandsia leiboldiana	-0.1389	-0.02263	-0.01518	0.09641	-0.4906562	0.5907426
Tillandsia australis	-0.08921	-0.003557	-0.004812	0.07833	-0.4050659	0.4650078
Tillandsia propagulifera	-0.08921	-0.003557	-0.004812	0.07833	-0.3229344	0.3847263
Tillandsia floribunda	-0.09287	-0.003627	-0.00184	0.07432	-0.3525346	0.3999353
Tillandsia latifolia ssp. *latifolia*	-0.1014	-0.01013	-0.006266	0.08504	-0.4294942	0.6181706
Tillandsia trauneri	-0.1078	-0.008357	0.005459	0.0832	-0.3475898	0.6060711
Tillandsia hitchcockiana	-0.1047	-0.01453	-0.006099	0.07408	-0.3488612	0.5124046
Tillandsia sphaerocephala	-0.09096	-0.009717	0.001125	0.07442	-0.3531745	0.4825815
Tillandsia adpressiflora	-0.08112	-0.007189	-0.01488	0.06597	-0.3108748	0.4262556
Tillandsia somnians	-0.08239	-0.005033	-0.005541	0.07851	-0.3314174	0.550001
Tillandsia stenoura	-0.09341	-0.004734	0.02142	0.09441	-0.3748035	0.6693527
Tillandsia complanata	-0.08054	-0.0008737	0.02569	0.08393	-0.2753471	0.568465
Tillandsia fasciculata	-0.08956	-0.01163	0.007404	0.07773	-0.336556	0.5267081
Tillandsia juncea	-0.09474	-0.007876	-0.004926	0.06462	-0.3213492	0.4067321
Tillandsia stricta	-0.1053	-0.005714	-0.0399	0.0946	-0.4515171	0.4965152
Vriesea itatiaiae	-0.09843	-0.01375	-0.01111	0.06589	-0.3339326	0.3725004

Based on these results, we settled to set cutoffs of log2(allele_count-0.5/alleles) for copy number decrease and log2(allele_count+1/alleles) for copy number increase which corresponds to log2(1.5/2) and log2(3/2) for a single-copy locus with two alleles. This encompasses the empirically observed range of variation and is dynamically adjusted to accommodate increasing variation we expect to be associated with CN > 1 in the reference sequence *A. trepida*.

CAFÉ error model estimation and model selection

In order to account for inaccuracies in the CN estimates and differences in accuracy between species, e.g. due to variation in coverage, we estimated an error model to be applied to each species via a built-in estimator supplied with CAFÉ. This resulted in the following best-fit error models:

Sample Name	Error rate
Tillandsia fasciculata	0.03515625
Tillandsia trauneri	0.0309375
Tillandsia propagulifera	0.00140625
Tillandsia juncea	0.03515625
Tillandsia latifolia ssp. *latifolia*	0.01125
Tillandsia australis	0.00703125
Tillandsia hitchcockiana	4.34E-19
Tillandsia floribunda	0.00703125
Tillandsia leiboldiana	0.06328125
Tillandsia complanata	0.01125
Tillandsia somnians	0.00703125
Tillandsia adpressiflora	0.00140625
Tillandsia stenoura	0.00984375

We first ran CAFÉ using a single global rate model. Based on the observation of an apparent increase in the rates of duplication and loss in the subgenus *Tillandsia*, we also tested a two-rate model allowing for separate rates of evolution in this subgenus. Each model was run

three times to check convergence of ML estimates and significance was determined using the Akaike Information Criterion (AIC) as below:

Global model	**λ**	**μ**	**Score**
Run1	0.00101061203283	0.00027763568401	61203.9
Run2	0.00101060963310	0.00027763236492	61203.9
Run3	0.00101060978654	0.00027763425550	61203.9
Two-rate model	**λ**	**μ**	**Score**
Run1	0.00079544338562 0.00284107269136	0.00023883871686 0.00086474231103	60340.4
Run2	0.00079544582304 0.00284121037836	0.00023883335117 0.00086475802060	60340.4
Run3	0.00079543765336 0.00284102919941	0.00023884156640 0.00086464833310	60340.4

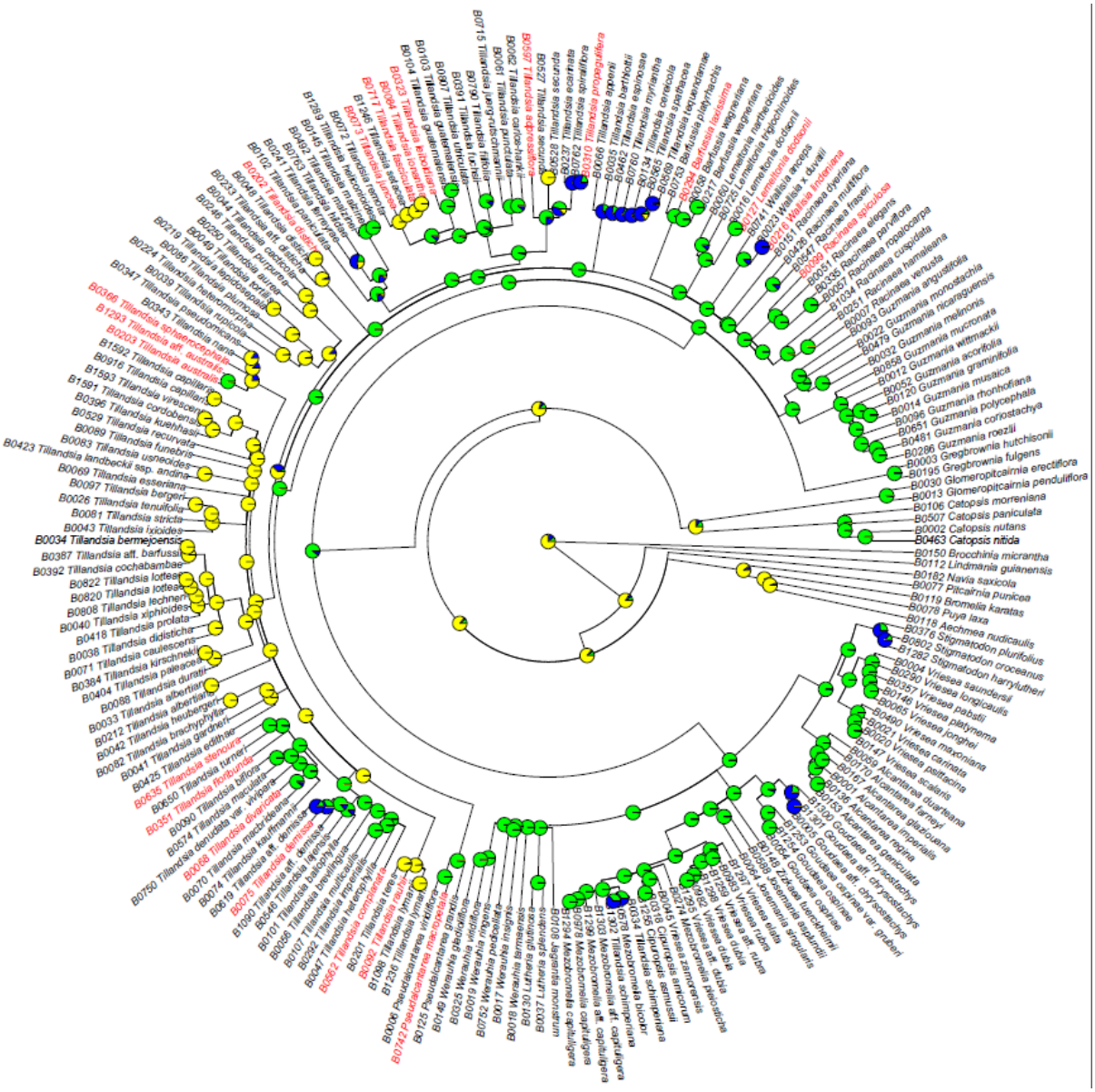
Model testing using AIC with the best run of the one- and two rate models:

	lnL	**#params**	**AIC**	**ΔAIC**
One rate	-61203.9	2	122411	1727
Two rate	-60340.4	4	120684	

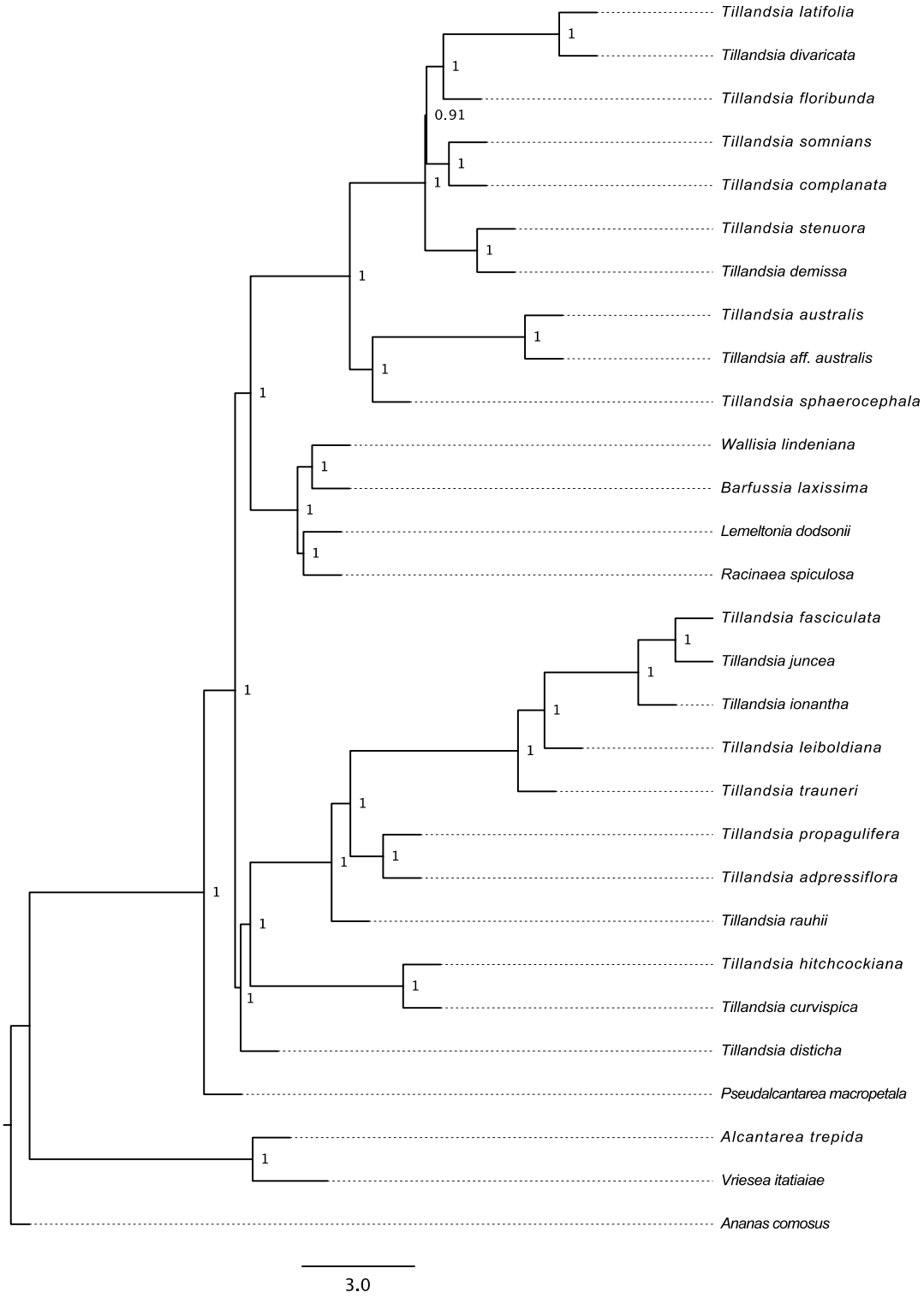
AIC = 2k - 2ln(L) where k = number of parameters.

Supplemental Figures

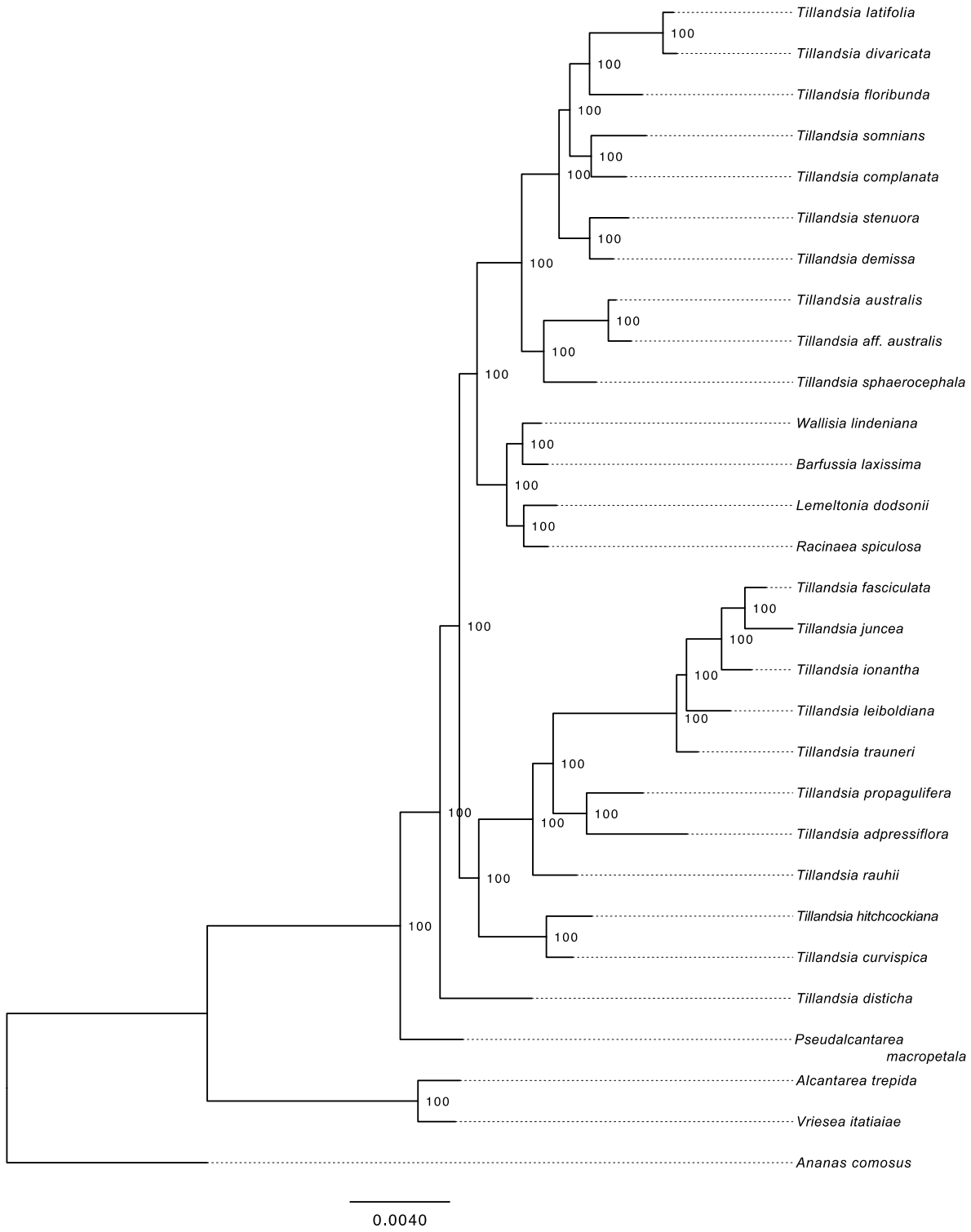
SI_Figure_1. Ancestral state reconstruction (ASR) of photosynthetic phenotype on the 210-taxa phylogenetic tree by Barfuss et al. (2016). Green, C_3; yellow, CAM; blue, Winter Holtum Zone. Taxa in red were subjected to whole genome sequencing (WGS) in the present study. Ancestral states inferred from this tree were carried over to the WGS tree presented in the main paper as described in the main text.



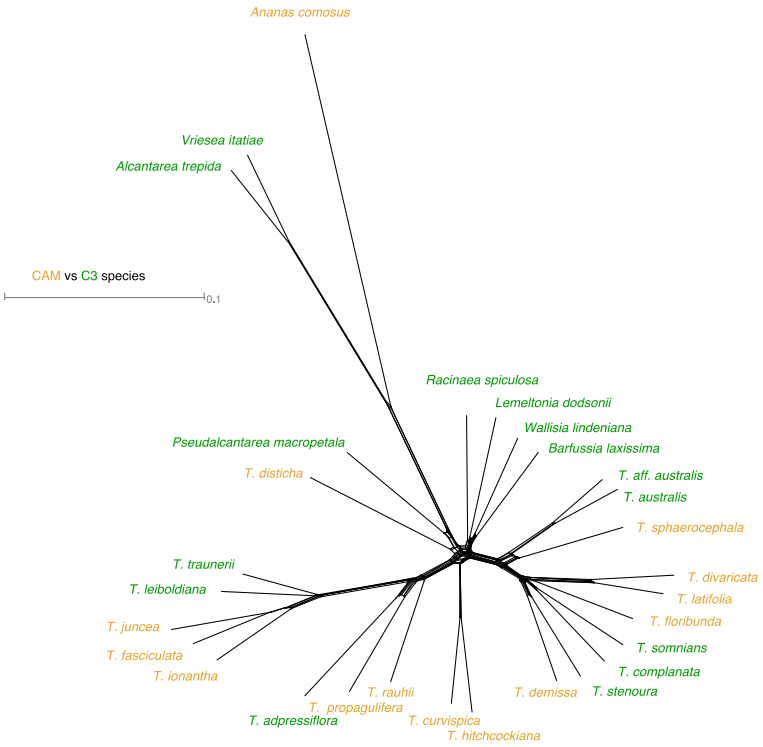
SI_Figure_2. Multispecies coalescent (ASTRAL) phylogenetic tree of 28 whole-genome sequenced tillandsioid bromeliad taxa (species of Tillandsia and related genera), including *Alcantarea trepida* and *Vriesea itatiaiae* as outgroups for phylogenetic analysis, and *Ananas comosus*, the species used to anchor most analyses presented in this study. Posterior probabilities are indicated for all branches.



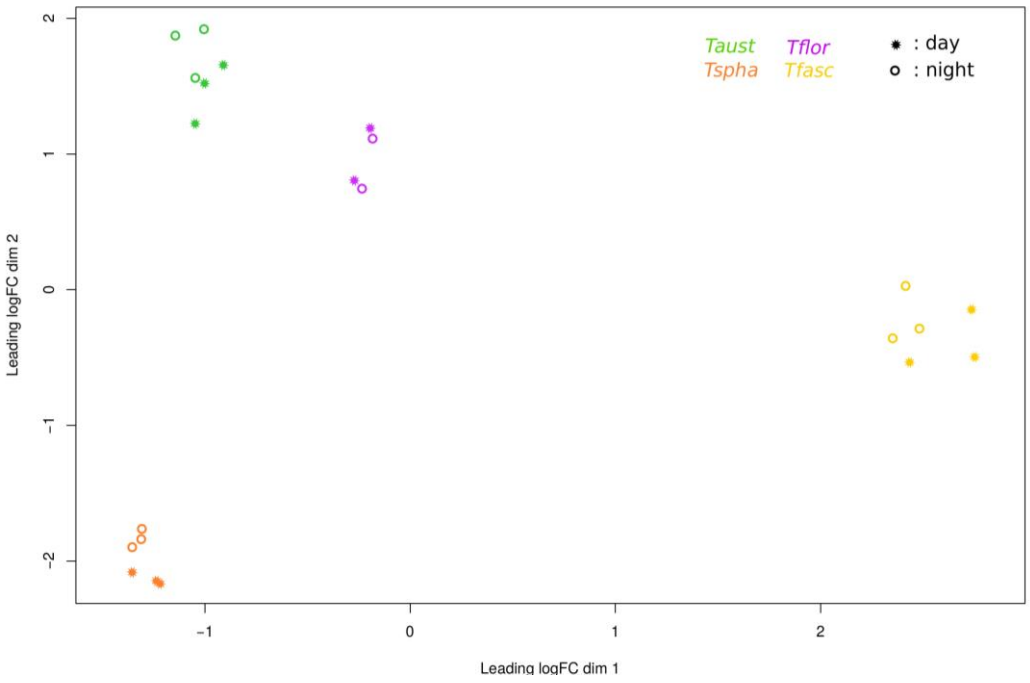
SI_Figure_3. Maximum likelihood (RAxML) phylogenetic tree of 28 whole-genome sequenced tillandsioid bromeliad taxa (species of Tillandsia and related genera), including *Alcantarea trepida* and *Vriesea itatiaiae* as outgroups for phylogenetic analysis, and *Ananas comosus*, the species used to anchor most analyses presented in this study. Bootstrap values of 100 were observed for all branches.



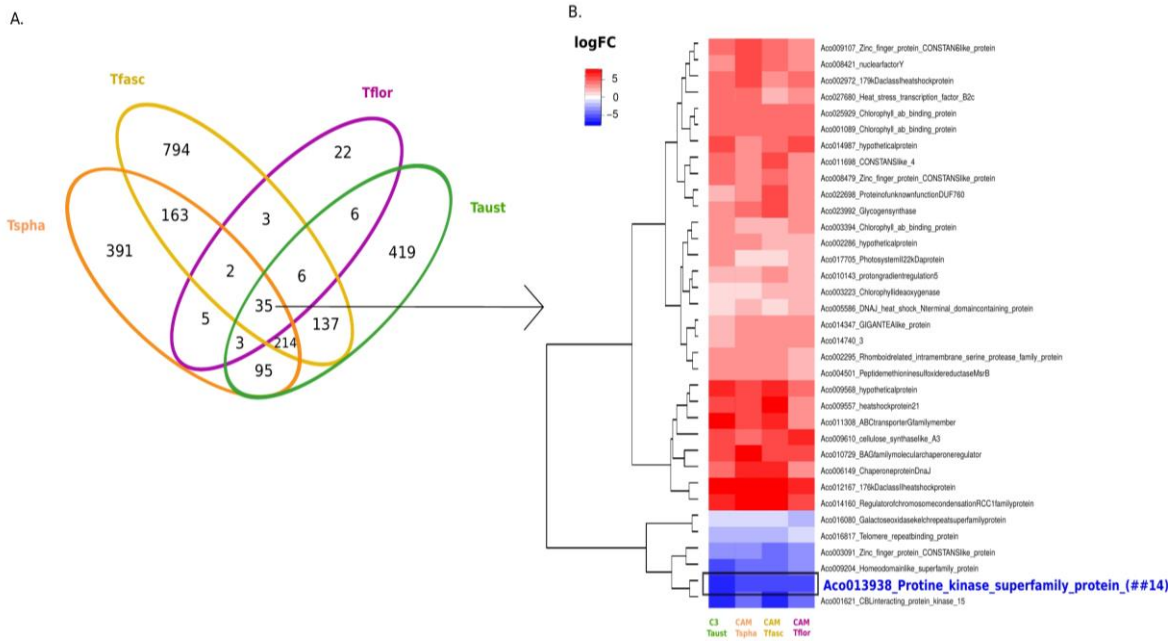
SI_Figure_4. SplitsTree network based on whole genome sequencing (WGS) of all sampled taxa, with C_3 and CAM taxa labelled in green and yellow respectively.



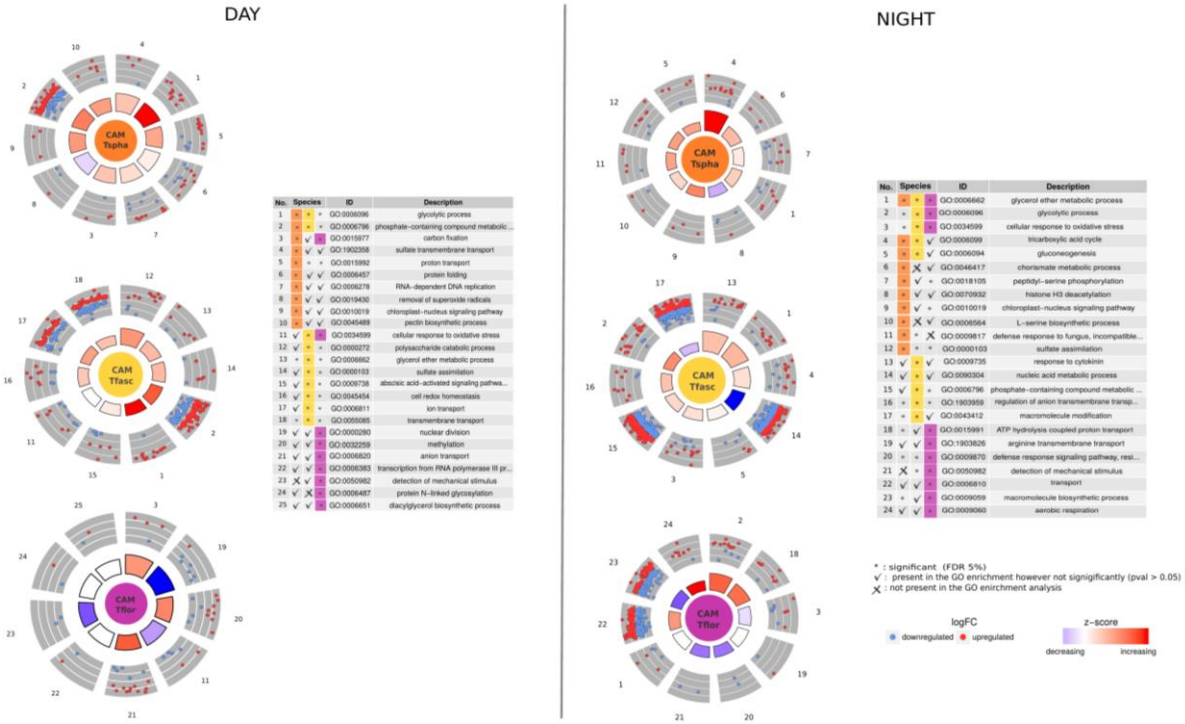
SI_Figure_5. Multi dimensional scaling (MDS) plot of RNA-seq count data. Top axes from Multi Dimensional Scaling (MDS) analysis of differential gene expression (DE) data.



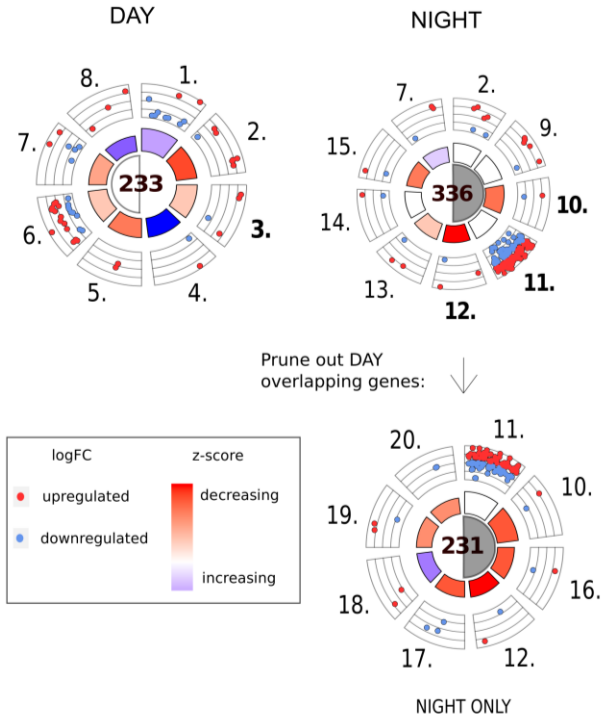
SI_Figure_6: Day/night RNAseq comparisons. Results from transcriptome-wide analysis of differential gene expression (DE) between two sampling time points, night (1AM) and day (11AM), referred to as "intraspecific day/night" test throughout the text. Shown are results at FDR<0.05 for logFC values >1 or <1 (i.e. pruning away logFC values close to zero). **A**, Venn chart depicting similarities and differences in temporal DE patterns between the studied species. **B**, Clustering heatmap for transcripts shared by all taxa (35 genes), corresponding to the central overlap field in the Venn chart. Red and blue colors in the heatmap indicate up- and down-regulation during the day, respectively. The key CAM enzyme PEPC kinase (PPCK; down-regulated during the day) is highlighted in blue font. Species designations are Taust, *T. australis* - C_3; Tspha, *T. sphaerocephala* – CAM; Tfasc, *T. fasciculata* – CAM; Tflor, *T. floribunda* - CAM. Color labels for species are consistent with figures relating to RNA-seq in the main text.



SI_Figure_7. Gene ontology (GO) enrichment results for interspecific CAM-C_3 comparisons. Results of gene ontology (GO) enrichment analysis for significant genes (FDR<0.05) from transcriptome-wide "interspecific day/night" DE tests. Shown are the 10 most significantly enriched GO terms for each species, excluding GO terms represented by less than two genes. The two subfigures entitled „DAY" and „NIGHT" are composed by the same graphical elements. First, rosette plots for genes within each of the enriched GO terms are presented, with red and blue dots indicating up- and down-regulation of single genes, respectively. Wedges in the inner portions of the rosettes designate z-scores based on logFC values for each gene in each group, thus indicating general trends of up- or down- expression in each group (red, increasing: blue, decreasing). Then, a table indicates the identities and descriptions of the top GO terms, including their occurrence in each species. *: significant enrichment at the 5% level; check-symbol: GO term found in a particular species but no significant enrichment; X: GO term not found in a particular species. Colored cells in the tables correspond to those GO terms depicted in the rosette plots.



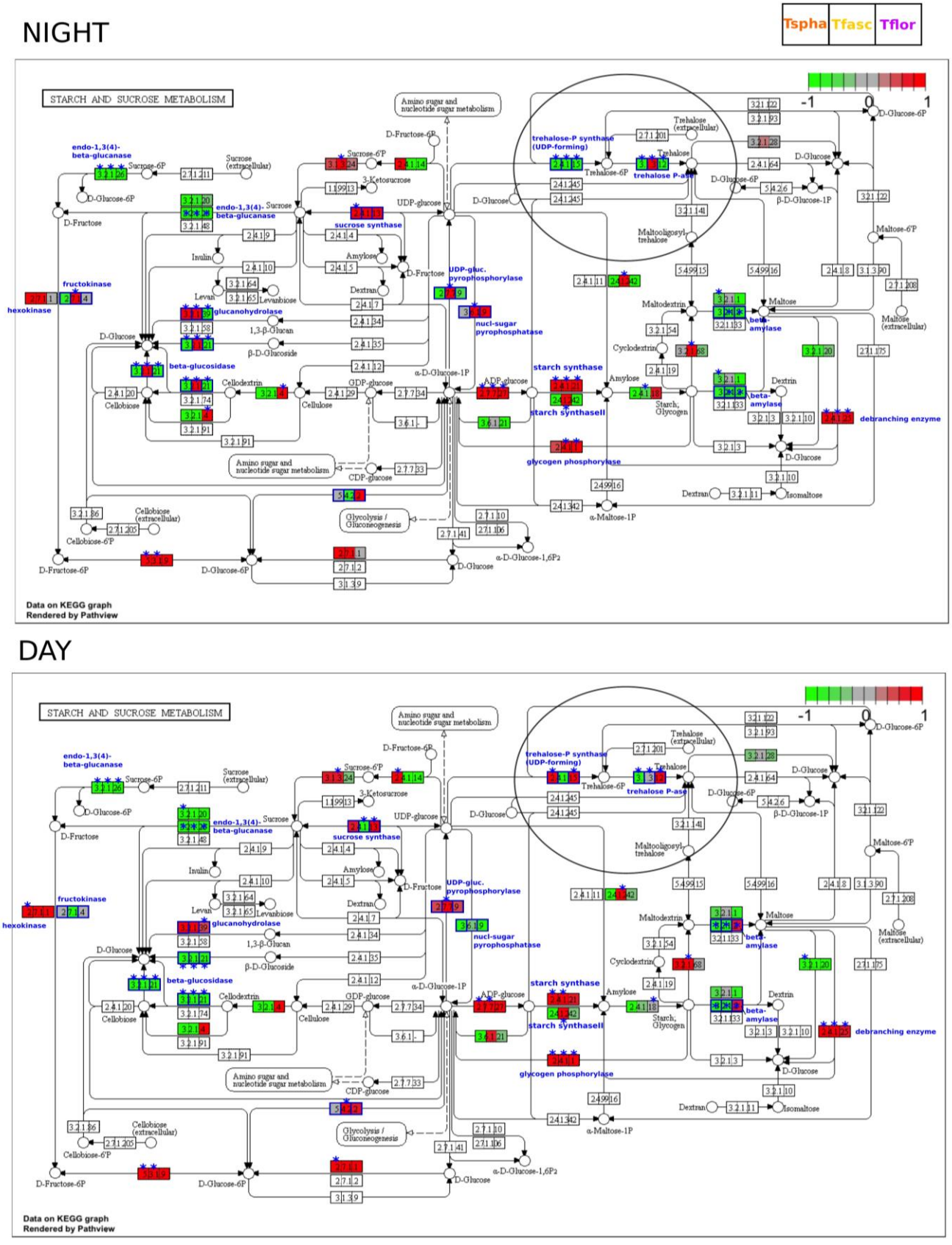
SI_Figure_8. GO enrichment analysis on the subsets of overlapping significant genes for the tree CAM species (see Figure 4, A). Enriched GO terms are presented for each of the three tested conditions: during the day (GO terms n° 1 to n° 8), during the night (n°2, n°7, and n° 9 to n° 15), and a subset of night-specific genes after pruning out the day-overlapping genes. For each subcategory (day, night, and night only) we depict the 8 most significantly enriched GO terms, with emphasis on the photosynthesis-related day-specific (bold, n°3) and night-specific (bold; n°10 to n°12) terms. Rosette plots highlight the relative contribution of up- and downregulated genes to each term and the overall trend (middle circle).



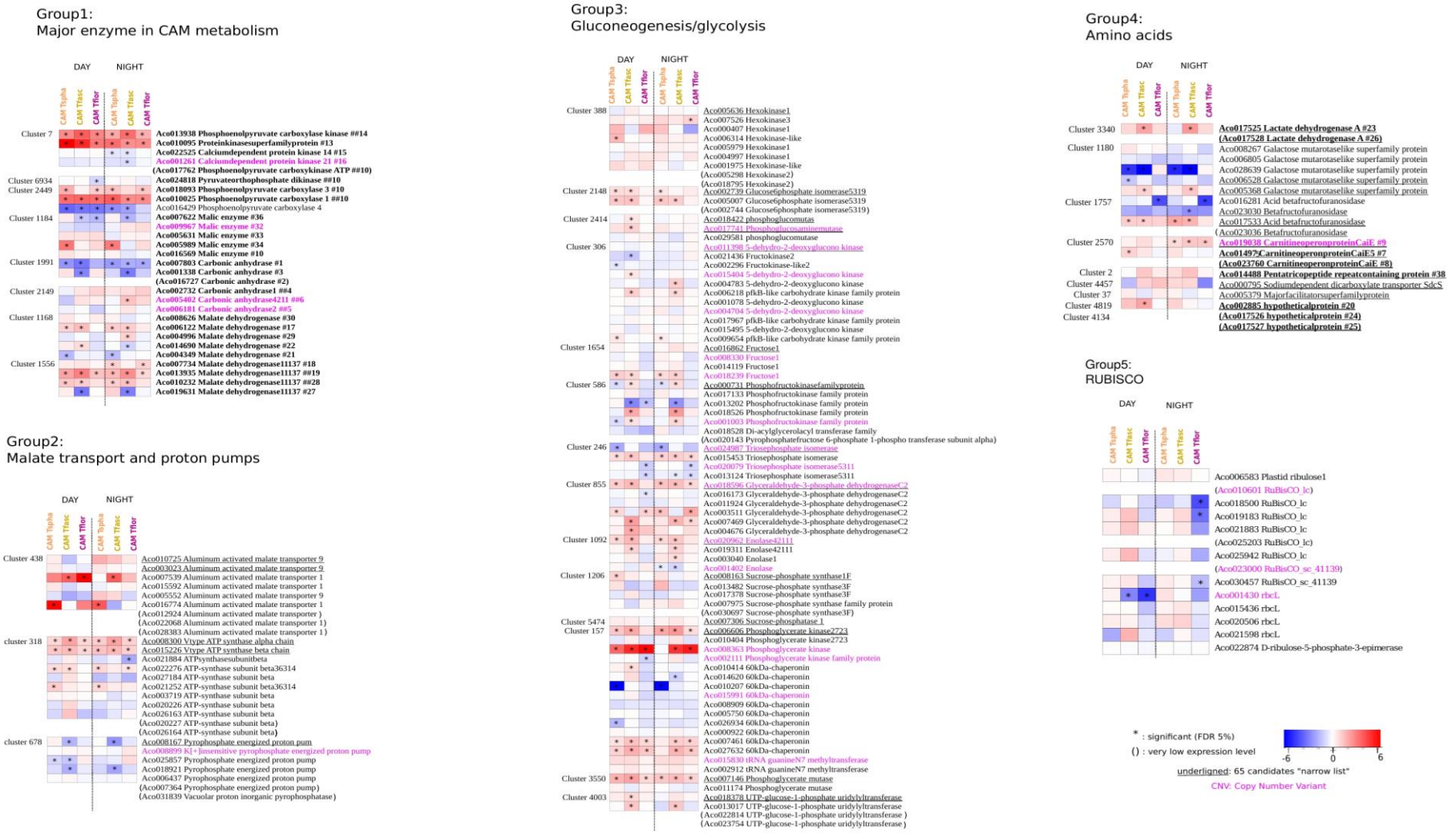
ID	Biological process
1. GO:0006457	protein folding
2. GO:0018105	peptidyl−serine phosphorylation
3. GO:0015977	**carbon fixation**
4. GO:0042744	hydrogen peroxide catabolic process
5. GO:0015992	proton transport
6. GO:0006099	tricarboxylic acid cycle
7. GO:0005975	carbohydrate metabolic process
8. GO:0015793	glycerol transport
9. GO:0034599	cellular response to oxidative stress
10. GO:0007035	**vacuolar acidification**
11. GO:0015770	**sucrose transport**
12. GO:0015991	**ATP hydrolysis coupled proton transport**
13. GO:0006662	glycerol ether metabolic process
14. GO:0009738	abscisic acid−activated signaling path.
15. GO:0006096	glycolytic process
16. GO:0045489	pectin biosynthetic process
17. GO:0055088	lipid homeostasis
18. GO:0046827	positive regulation of protein export
19. GO:0044571	[2Fe−2S] cluster assembly
20. GO:0019700	organic phosphonate catabolic process

248 **SI_Figure_9.** KEGG map for starch/sucrose metabolism (ko00500).

249



SI_Figure_10. Extended target gene list for interspecific C_3/CAM differential gene expression analyses with pineapple expression data. Heatmaps depicting inter-specific C_3/CAM DE patterns for homologue clusters of genes with potential involvement in CAM photosynthesis and carbohydrates metabolism based on evidence from pineapple, *Ananas comosus* (Ming et al. 2015; Wai et al. 2017). The original 65 CAM-related candidate genes used to identify the homologue clusters are underlined.



References

Aldous, S. H. *et al.* Evolution of the Phosphoenolpyruvate Carboxylase Protein Kinase Family in C_3 and C_4 Flaveria spp. *Plant Physiol.* (2014). doi:10.1104/pp.114.240283

Bräutigam, A., Schlüter, U., Eisenhut, M. & Gowik, U. On the Evolutionary Origin of CAM Photosynthesis. *Plant Physiol.* 174, 473–477 (2017).

Crayn, D. M., Winter, K. & Smith, J. A. C. Multiple origins of crassulacean acid metabolism and the epiphytic habit in the Neotropical family Bromeliaceae. *Proc. Natl. Acad. Sci.* **101,** 3703–3708 (2004).

Gouda, E. J. & Butcher, D. (cont. updated) *A List of Accepted Bromeliaceae Names*[http://bromeliad.nl/bromNames/]. University Botanic Gardens, Utrecht (*accessed: 08-06-2020*).

Males, J. Concerted anatomical change associated with crassulacean acid metabolism in the Bromeliaceae. *Functional Plant Biol.* **45**, 681–695 (2018).

Ming, R. *et al.* The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.* **47,** 1435–1442 (2015).

Mirarab, S. & Warnow, T. ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btv234

O'Brien, T. P., Feder, N., & McCully, M. E. Polychromatic staining of plant cell walls by toluidine blue O. *Protoplasma*, *59*(2), 368-373 (1964).

Pierce, S., Winter, K. & Griffiths, H. Carbon isotope ratio and the extent of daily CAM use by Bromeliaceae. *New Phytol.* **156,** 75–83 (2002).

Kraus, J. E., de Sousa, H. C., Rezende, M. H., Castro, N. M., Vecchi, C., & Luque, R. Astra blue and basic fuchsin double staining of plant materials. *Biotech. Histochem. 73*(5), 235-243 (1998).

Silvera, K. *et al.* Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology* **37,** 995–1010 (2010).

Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* (2014). doi:10.1093/bioinformatics/btu033

Wai, C. M. *et al.* Temporal and spatial transcriptomic and microRNA dynamics of CAM photosynthesis in pineapple. *Plant J.* **92**, 19–30 (2017).

Waterhouse, R. M. *et al.* BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2018).

Winter, K. & Holtum, J. A. M. How closely do the d13C values of Crassulacean acid metabolism plants reflect the proportion of CO2 fixed during day and night? *Plant Physiol.* **129,** 1843–1851 (2002).