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]† ⓘ

[1]Department of Botany and Biodiversity Research, Division of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Vienna, Austria

[2]Department of Biology, Unit of Ecology & Evolution, University of Fribourg, Fribourg, Switzerland

[3]Department of Computational Biology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland

[4]Department of Functional and Evolutionary Ecology, Division of Molecular Systems Biology (MOSYS), Faculty of Life Sciences, University of Vienna, Vienna, Austria

[5]Vienna Metabolomics Center (VIME), University of Vienna, Vienna, Austria

[6]Department of Horticulture, College of Agriculture and Natural Resources, Michigan State University, East Lansing, Michigan

[7]Department of Plant Biology, School of Integrative Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois

Correspondence
Margot Paris, Department of Biology, Unit of Ecology & Evolution, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland.
Email: margot.paris@unifr.ch

Abstract

The adaptive radiation of Bromeliaceae (pineapple family) is one of the most diverse among Neotropical flowering plants. Diversification in this group was facilitated by shifts in several adaptive traits or "key innovations" including the transition from C_3 to CAM photosynthesis associated with xeric (heat/drought) adaptation. We used phylogenomic approaches, complemented by differential gene expression (RNA-seq) and targeted metabolite profiling, to address the mechanisms of C_3/CAM evolution in the extremely species-rich bromeliad genus, *Tillandsia*, and related taxa. Evolutionary analyses of whole-genome sequencing and RNA-seq data suggest that evolution of CAM is associated with coincident changes to different pathways mediating xeric adaptation in this group. At the molecular level, C_3/CAM shifts were accompanied by gene expansion of XAP5 CIRCADIAN TIMEKEEPER homologs, a regulator involved in sugar- and light-dependent regulation of growth and development. Our analyses also support the re-programming of abscisic acid-related gene expression via differential expression of ABF2/ABF3 transcription factor homologs, and adaptive sequence evolution of an ENO2/LOS2 enolase homolog, effectively tying carbohydrate flux to abscisic acid-mediated abiotic stress response. By pinpointing different regulators of overlapping molecular responses, our results suggest plausible mechanistic explanations for the repeated evolution of correlated adaptive traits seen in a textbook example of an adaptive radiation.

KEYWORDS

adaptive radiation, Bromeliaceae, circadian period length, copy number variation, drought, genome, transcriptome

1 | INTRODUCTION

Species radiations have traditionally been studied primarily from a macro-evolutionary perspective within a phylogenetic framework (Barraclough & Vogler, 2000; Cavender-Bares, Kozak, Fine, &

Kembel, 2009; Kozak & Wiens, 2010). The recent "-omics" revolution and novel analytical tools increasingly allow evolutionary biologists to address the dynamics of genomic variation across time scales (Dasmahapatra et al., 2012; Heled & Drummond, 2009; Nevado, Atchison, Hughes, & Filatov, 2016; Novikova et al., 2016). This potentially allows evolutionary geneticists to identify the major sources of genetic variation that fuel evolutionary responses to stresses and/or new environments (Pease, Haak, Hahn, & Moyle, 2016). Bromeliaceae (>3,000 species) represent a "textbook adaptive radiation" in flowering plants (Benzing, 2000). Numerous adaptive traits or "key innovations" vary among species of this large Neotropical family, including the epiphytic growth habit (life on trees), presence or absence of water-impounding leaves (tank-forming rosettes), absorptive trichomes, leaf succulence and Crassulacean acid metabolism (CAM) photosynthesis (Benzing, 2000; Crayn et al., 2004; Givnish et al., 2014). Chromosome counts in the family point to predominant diploidy with a remarkably conserved counts of $2n = 50$ or $2n = 48$ chromosomes and a comparatively compact range of DNA content from 0.85 to 2.23 pg/2C, suggesting a largely homoploid radiation (Gitaí, Paule, Zizka, Schulte, & Benko-Iseppon, 2014).

The highly species-rich genus, *Tillandsia* L., (ca. 746 species; subfamily Tillandsioidae) exhibits great variation in life habits (epiphytic, terrestrial and rock-growing), photosynthetic pathways (C_3 and CAM) and pollination syndromes (birds and insects) (Barfuss et al., 2016; Barfuss, Samuel, Till, & Stuessy, 2005; Benzing, 2000; Crayn et al., 2004; Givnish et al., 2014). The presence or absence of absorptive trichomes and leaf succulence, and several other adaptive phenotypic traits appear to have evolved in a contingent manner, giving rise to adaptive syndromes of correlated characters (Givnish et al., 2014; Males, 2016). Most obviously, these trait associations manifest themselves in the so-called "green" *Tillandsia* phenotypes and species adapted to cool, moist habitats, typically exhibiting C_3 photosynthesis, and widespread formation of tank rosettes. This contrasts with the so-called "grey" *Tillandsia* phenotypes and agrees with a strong tendency to express CAM photosynthesis, dense absorptive trichome cover and pronounced succulence, as well as strong association with warm habitats having high solar incidence in regions with low rainfall (Benzing, 2000; Givnish et al., 2014; Males & Griffiths, 2018). Although shifts in these adaptive traits have long been hypothesized to be drivers of adaptive radiation in this group (Barfuss et al., 2016; Benzing, 2000; Givnish et al., 2014), little to nothing is known about their genetic basis, with the notable exception of CAM photosynthesis.

The CAM pathway, known from at least 35 different plant families, represents an adaptation for increased water-use efficiency by shifting CO_2 assimilation to the night time, thus allowing stomata to be closed during the day and thereby reducing water loss (Borland et al., 2014; Silvera et al., 2010). CAM entails complex patterns of gene expression, post-translational regulation and metabolic fluxes that are only starting to be understood at a whole systems level (Wai et al., 2017). CAM represents an example for convergent evolution of functionally important phenotypes, but little is known about the underlying genetic and molecular mechanisms (Heyduk, Moreno-Villena, Gilman, Christin, & Edwards, 2019).

Here, we combine a range of complementary approaches to shed light on the genetic basis and evolution of an enigmatic adaptive trait syndrome, involving CAM, in the highly speciose Neotropical adaptive radiation of tillandsioid bromeliads. To this end, we reconstructed the ancestral states of photosynthetic syndromes based on the largest and most informative phylogenetic dataset assembled for this group to date (Barfuss et al., 2016) and then sequenced the genomes of 28 selected accessions of Tillandsioideae, including 25 species of *Tillandsia* sensu *lato* capturing most of the available variation in the CAM-adaptive syndrome. We first used the whole-genome phylogenetic data to ask whether repeated C_3/CAM trait shifts in this radiation are more likely to have arisen independently or by wide-spread gene flow. Next, we examined genome-wide signatures of branch-specific selection and gene duplication/loss accompanying CAM-related trait shifts. We then focused on time-dependent metabolic and transcriptomic changes in the representative species. We show that repeated xeric (heat/drought) adaptation including CAM involves not only structural genes in key pathways, but also remarkably pleiotropic regulators. These discoveries allow us to discuss the potential molecular mechanisms behind this enigmatic case of the repeated evolution of functionally important phenotypes.

2 | MATERIALS AND METHODS

2.1 | Ancestral state reconstruction

To establish a 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 accessions; including seven outgroup species from other bromeliad subfamilies. The 203 Tillandsioideae accessions represent a total of 177 species covering 21 out of the 22 genera of the subfamily (see supplemental information Data S1 for more detail about taxon sampling for this analysis). The taxonomic coverage of these genera varied from 6 to 100%. Most of the collected species (93 species) belong to the genus, *Tillandsia*—the only genus presenting frequent C_3 to CAM transitions. Despite the high number of *Tillandsia* species included in the analyses, we cover only 12.5% of this species-rich genus. Limited taxon sampling can affect statistical power and levels of inference (Folk et al., 2018; Litsios & Salamin, 2012). However, ancestral range reconstructions of Galliformes suggest that trees including important and representative taxa can recover comparable ancestral state reconstructions (ASRs) to those obtained using trees having complete taxon sampling (Wang, Kimball, Braun, Liang, & Zhang, 2012). Although it seems like 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 the Multiple State Speciation and Extinction (MuSSE) algorithm implemented in diversitree (FitzJohn 2012) for ASR, coding

photosynthetic syndrome as either 1 for C_3 photosynthesis, 2 for intermediate or "Winter Holtum Zone" ($\delta^{13}C$ from −23.0 to −19.0‰, Winter & Holtum, 2007) and 3 for CAM photosynthesis. The three states were inferred from published data (Crayn, Winter, Schulte, & Smith, 2015) and newly collected carbon isotope ratio values for 27 species (Figure 1; Table S1). We weighted the three character states based on their estimated fraction of sampled species and compared six different hypothesis-relevant models using ML as described in supplemental information (Data S1). The lambda model (diversification rate, λ, varies between states) received the strongest support (logLik = 829.6978, AIC = −1,649.396), thus ASR was carried out

within this model, and results from the published phylogenetic tree (Barfuss et al., 2016) were used to inform evolutionary patterns of trait transitions in the species included in our whole-genome phylogenetic analyses (Figure 1).

2.2 | Species sampling for whole-genome and transcriptome sequencing

A total of 28 accessions, representing 20 species of *Tillandsia* and seven species of closely related genera, were selected for further

FIGURE 1 Coalescent-based phylogeny of 28 whole-genome sequenced accessions for 25 tillandsioid bromeliad species and *Ananas comosus* outgroup. Annotations display grouping into C_3 (green stars) and CAM (yellow stars) species according to carbon isotope ($\delta^{13}C$) phenotypes. Pie charts on internal branches indicate the ancestral state of photosynthetic syndrome, coded as C_3 (green) Winter–Holtum Zone (blue) or CAM (yellow), inferred from the largest available phylogenetic tree available for this group to date (Barfuss et al., 2016; Figure S1). Size-reduced pie charts indicate nodes affected by missing species or minor topological differences with the larger tree used for ancestral state reconstruction (Figure S1); the closest and most likely ancestral states were applied parsimoniously for these cases. Carbon isotope ratios and night-time malate and citrate concentrations from targeted metabolite analysis are indicated on the right (numerical data provided in Tables S1 and S2). Samples used for RNA-seq expression profiling are indicated by a circle, samples with an asterisk represent species for which a close relative was used in metabolite analysis. Differences in six additional morphological/anatomical traits, including succulence, are indicated, as these may represent rate-limiting steps during the transition to strong CAM (Edwards, 2019) [Colour figure can be viewed at wileyonlinelibrary.com]

3
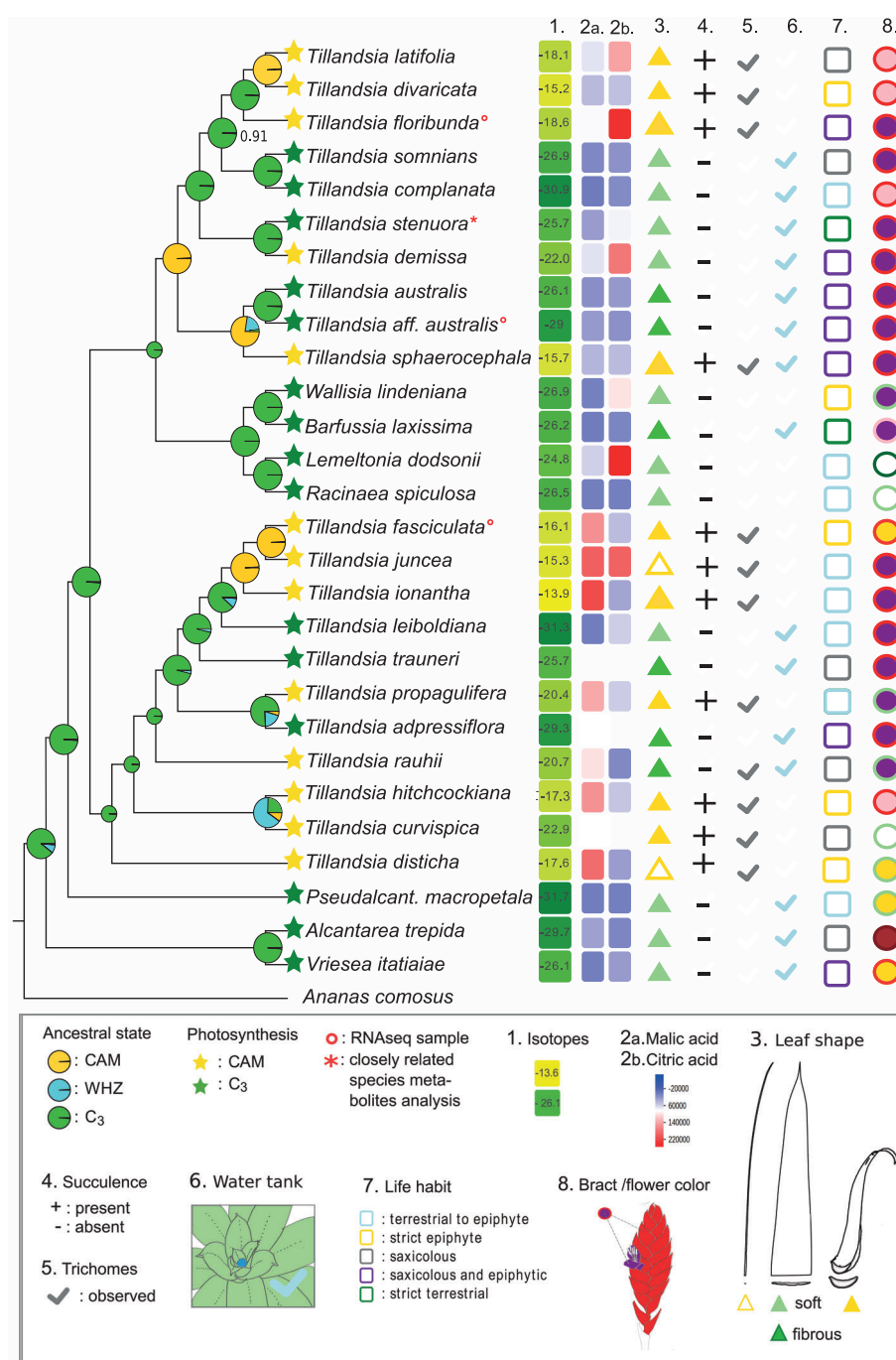
analyses at the whole-genome level. The *Tillandsia* species were chosen to achieve representative sampling of phenotypic variation and repeated character-trait shifts from the adaptive radiation of this group (Barfuss et al., 2016; Benzing, 2000). Using species from closely related genera, containing only C_3 species (see supplemental information in Data S1), we aimed to represent the range of photosynthetic syndromes along the dynamic zone of repeated transitions along the phylogenetic tree, from typical C_3 to typical CAM (Figure 1, Crayn et al., 2004, 2015). In addition, we included the outgroup species, *Ananas comosus*, using available Illumina paired-end sequences from cultivar N67-10 available on NCBI (accession DRX020985).

Specimens were sampled from the living collections of the Universities of Vienna (Austria) and Heidelberg (Germany) and in the Botanical Gardens of Rio de Janeiro (Brazil), Lyon (France), Porrentruy, and Geneva (both Switzerland). Information on morphological characters, habitats and natural distributions of the studied accessions is provided in Figure 1 and Table S2.

2.3 | DNA extraction and whole-genome sequencing

For each sample, 30 mg of leaf material was dried in silica gel before DNA extraction with a QIAGEN DNeasy® Plant Mini Kit, following supplier's instructions. After shearing of 1 μg of DNA with either a Covaris® or Bioruptor® instrument, sequencing libraries were prepared with the Illumina TruSeq® DNA PCR-Free Library Prep Kit. Libraries were sequenced paired-end 2×150bp or 2×125bp in seven different lanes of an Illumina HiSeq3000 sequencer. Mean depth coverage per sample was on average 17.7x (median 20.6x; Table S2).

2.4 | WGS data processing and variant calling

Reads were trimmed with condetri v2.2 (Smeds & Künstner, 2011) using 20 as a high-quality threshold parameter. The *Tillandsia adpressiflora* sample, with the highest number of reads (Table S2), was used to build a pseudo-reference genome following an iterative mapping strategy described in de La Harpe et al. (2019), using the annotated *A. comosus* reference genome (Ming et al., 2015) as a starting point. The 28 samples and *A. comosus* were mapped to the pseudo-reference using Bowtie2 v2.2.5 (Langmead & Salzberg, 2012) with the very-sensitive-local option. SNPs were called using GATK UnifiedGenotyper, using the EMIT_ALL_SITES and –glm SNP options, after realignment around indels and base recalibration with GATK v3.3 (McKenna et al., 2010). The EMIT_ALL_SITES output mode was used in order to obtain calls at both variant and invariant sites. A total of 78,620,138 positions were filtered with vcftools v0.1.13 (Danecek et al., 2011) before subsequent analyses, retaining only positions with quality >20, read depth >3 and a maximum of 50% of missing data. In addition, a matrix of 5,646,174 SNPs was obtained by selecting only polymorphic sites from the filtered position matrix, using a minimum

allelic count of 3x. Read number, coverage and mapping statistics were calculated using bedtools v2.24.0 (Quinlan & Hall, 2010) and vcftools v0.1.13 (Danecek et al., 2011).

2.5 | Phylogenomic analysis

In order to reconstruct evolutionary relationships among the target species, we used ASTRAL v5.6.1 (Mirarab & Warnow, 2015), a summary method based on the multispecies coalescent, to infer a species tree using 100-kb window trees. The individual window trees were estimated by a ML analysis in RAxML v8.228 (Stamatakis, 2014) under the GTRGAMMA substitution model with 100 rapid bootstrap replicates and using *A. comosus* as outgroup. All positions present in each 100-kb window, including variant and invariant sites, were used for these analyses. A total of 1753 100-kb window trees with average bootstrap values >70 were kept for the analyses (77.2% of the 2,272 estimated window trees), as weakly informative individual trees can potentially reduce the accuracy of species tree estimation with summary method based on coalescent analyses (Liu, Xi, Wu, Davis, & Edwards, 2015; Molloy & Warnow, 2018). We complemented the coalescent method by a ML tree reconstruction from the concatenated alignment of all SNPs, using RAxML v8.228 (Stamatakis, 2014), under the GTRGAMMA substitution model and support estimated with the -autoMRE option. We explicitly specified the number of invariable sites in our data set using the ascertainment bias correction option -asc-corr = felsenstein in RAxML. A phylogenetic network was constructed using the neighbour-net method implemented within SplitsTree v4.13.1 (Huson & Bryant, 2006) using an identity-by-descent distance matrix obtained with PLINK v1.07 (Purcell et al., 2007) and the 5,967,161 SNPs database.

2.6 | Carbon isotope phenotyping

We assessed the carbon isotope ratio ($^{13}C/^{12}C$) for all species used in the present study using 1 g of silica-dried material per sample. The measurements were carried out at the Institute of Earth Surface Dynamics (Faculty of Geosciences and Environment, University of Lausanne, Switzerland) following Spangenberg, Jacomet, and Schibler (2006). This approach makes use of flash combustion on an elemental analyser connected to a ThermoQuest/Finnigan Delta S isotope ratio mass spectrometer via a ConFlo III split interface. The carbon isotope ratios values (δ) correspond to the per mille (‰) deviation relative to the Vienna-Pee Dee belemnite standard (V-PDB).

2.7 | Targeted metabolite analysis

Dry silica samples were collected in the Bromeliaceae research greenhouses at Department of Botany and Biodiversity Research, University of Vienna, in July 2017 at distinct time points, namely at 1 a.m. and 11 a.m., following the same day/night sampling scheme also used

for gene expression profiling based on RNA-seq (above). We sampled green tissue from mature leaves, where possible from the exact same specimen also used for whole-genome sequencing (WGS) and RNA-seq. In four cases, for which the exactly same species were not available in the greenhouse, we sampled species known to be very closely related to the actual target taxa. All samples were processed at the Vienna Metabolomics Center (VIME, Department of Molecular Systems Biology, University of Vienna). All analysis steps, including plant metabolite extraction, sample derivatization, and GC–TOF-MS (gas chromatography coupled with time-of-flight mass spectrometry) were carried out as previously described (Weckwerth, Loureiro, Wenzel, & Fiehn, 2004). Data analysis was performed using ChromaTof (Leco) software. Briefly, representative chromatograms of different samples were used to generate a reference peak list, and all other data files were processed against this reference list. Deconvoluted mass spectra were matched against an in-house mass spectral library. Peak annotations and peak integrations were checked manually before exporting peak areas for relative quantification. Metabolite amounts are given in arbitrary units corresponding to the peak areas of the chromatograms. Data were analysed using partial least squares (PLS) discriminant analysis (DA) using the COVAIN toolbox for metabolomics data mining (Sun & Weckwerth, 2012).

2.8 | Positive selection on coding sequences

We tested for positive selection acting on the WGS data using the non-synonymous to synonymous substitution rate ratio (ω = dN/dS) tests using codeml in PAML version 4.8a (Yang, 2007). We filtered out gene alignments with less than 300 bp, more than 50% missing data and multiple stop codons. In addition, individual sequences with more than 90% missing data were removed from the codon alignment. We inferred gene phylogenies for each codon alignment using PhyML version 3.3.20170119 (Guindon & Gascuel, 2003) with HKY and GTR substitution models and GAMMA shape parameter. We computed a likelihood-ratio test (LRT) to select the best reconstruction for codeml analyses. We ran the cladeC model, containing five estimated parameters, denoted as follows: p0, p1, p2 (p2 = 1 – p0 – p1) for the proportion of sites in a site-class, and dN/dS parameters ω0, ω1 fixed to 1, ω2 background (C$_3$ plants) and ω2 foreground (CAM plants). This model tests whether selection has occurred in all branches leading to species with CAM metabolism (H1). The null model corresponded to the M2a-rel (Weadick & Chang, 2012) with four parameters, where ω2 background and foreground are fixed to be equal. Each model was run three times to overcome convergence issues, and the best likelihood run was used for model comparison and ω estimates. Model fit was compared using the corrected Akaike information criteria (AICc) with a significance threshold delta-AIC of 10 between M2a-rel and H1. For genes exhibiting stronger support for H1, we used the standard errors (SE) of ω estimates to determine the deviation from neutrality (ω = 1 is nearly neutral, therefore genes were discarded when ω ± SE included 1), and considered only genes with signatures of positive selection in any CAM lineage (ω2 in CAM > 1). Genes with dN/dS

estimates close to the optimization bound in the three replicates (i.e., parameter estimates close to 999) and SE larger than 10 were discarded.

2.9 | Inference of gene family evolution

Copy number variants were detected in 15 high-coverage individuals based on relative read-depth differences in exons using CNVkit (Talevich, Shain, Botton, & Bastian, 2016). Illumina reads were aligned to the *Tillandsia* pseudo-reference as above and filtered using SAMtools v.0.1.19 (Li et al., 2009) to remove duplicates (rmdup) and ambiguously mapping reads with mapping quality less than 10 (–q 10). To avoid conflating lack of mappability with gene loss, we only retained exons with a mean read coverage of at least five in five individuals for further analysis. We first called copy number (CN) status in *Al., trepida* by profiling against *A. comosus* as a reference. For CNV detection within the genus, *Tillandsia*, we then used *Al. trepida* as the reference sample since it better reflects the observed mapping biases against the pseudo-reference genome with the remaining *Tillandsia* samples (lack of coverage outside coding regions, systematic variation in coverage along the genome). CNVkit was run with an average antitarget size of 5,000 and an accessibility mask of 2 kb, generated based on the *A. comosus* hardmasked reference sequence, available from CoGe (https://genomevolution.org/coge/GenomeInfo.pl?gid= 25734). Gene gains and losses were called requiring consistent signal across at least four tiles (gainloss -m 4). Log2 ratios per gene were then filtered according to coverage-dependent, empirically determined thresholds for lower and upper bounds (Supplementary Text in Data S1). Log2 ratios, above and below the respective thresholds, were then translated into numbers of alleles by exponentiation and multiplication with the inferred allele number in *A. trepida*, divided by two and rounded to the next integer in order to calculate CNs per gene. We consider CNs to be an approximate and relative measure of CN variation with the genus, *Tillandsia*, rather than absolute estimates.

Homolog clusters in the pineapple were identified using MCL clustering of the *A. comosus*-predicted proteins with MCL-edge (Enright, Van Dongen, & Ouzounis, 2002), based on an all-vs-all BLASTp search with an e-value cutoff of $1e^{-5}$ and an inflation value of 3. Inferred CNs, for each gene in a cluster, were summed to obtain family-level gene pseudocounts across the high-coverage dataset for analysis of gene family evolution with CAFÉ v4.1 (Han, Thomas, Lugo-Martinez, & Hahn, 2013). We pruned species not included in the CNV analysis from the RAxML tree and used the Penalized Likelihood method with cross-validation implemented in the program, r8s, (Sanderson, 2003) to obtain an ultrametric tree for analysis with CAFÉ. All CAFÉ models were run with species-level error models inferred using the error model estimation script supplied with CAFÉ. Error estimates by species ranged between 0 in *A. trepida*, *T. sphaerocephala* and *T. hitchcockiana* to a maximum of 0.03 in *T. juncea* and *T. leiboldiana*. We estimated separate rates of gain (λ) and loss (μ) in either a global model with a single set of parameters across the

entire tree or a two-rate model with different sets of rate parameters for species in the subgenus *Tillandsia* and the rest of the tree. The two-rate model fit significantly better than the global model (delta-AIC 1727) and was the basis for inference of gene gains and losses along branches and ancestral copy numbers.

To test for associations between CN changes and evolution of CAM photosynthesis and correlated traits, we ran a permutation ANOVA (Nagy et al., 2017). Branch-wise turnover rates were calculated from ancestral CN estimates for each family and branches were labelled according to photosynthetic strategy. Due to uncertainties in ASR, we opted to only analyse terminal branches, mirroring the branch-wise tests for positive selection. Empirical p values were determined from 1,000 permutations of branch label swapping and corrected for multiple testing to control the false discovery rate (Benjamini & Hochberg, 1995).

2.10 | RNA sequencing

Four *Tillandsia* species were included in the RNAseq analyses; three different specimens were used to serve as biological replicates for the species, *T. australis* (C_3), *T. fasciculata* (CAM) and *T. sphaerocephala* (CAM), and only two replicates were available for the species, *T. floribunda* (CAM). All four species were kept under identical greenhouse conditions for at least 10 days prior to sampling (°C: min = 16.6, max = 39.0, M = 25.7 and %rF min = 30.8, max = 95.9, M = 67.5). For each sample and time point, up to 30 mg fresh leaf tissues were stabilized in RNAlater® immediately after sampling and kept at −20°C. Total RNA was carefully extracted under a sterilized fume hood with QIAGEN RNeasy® Mini Kit following the supplier's protocol. RNA libraries were prepared with the Illumina TruSeq® Stranded mRNA Library Prep Kit before sequencing pair-end 2×150bp on an Illumina HiSeq3000 sequencer.

Sequence quality was validated with FastQC v0.11.2 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/README.txt), and reads were trimmed with condetri v2.2 (Smeds & Künstner, 2011) using 20 as high-quality threshold parameter. Reads were mapped to a *Tillandsia adpressiflora* pseudo-reference genome (described below) with TopHat v2.1.0 (Kim et al., 2013, p. 2) using –*b2-very-sensitive* mapping parameters. Only uniquely mapped reads were kept for further analyses.

2.11 | Differential expression analysis of RNA-seq data

The number of reads mapping to reference genes was quantified with HTSeq-count v0.6.0 in default mode (Anders, Pyl, & Huber, 2015), producing the gene count data (27,024 genes in total) for the DE analyses. We filtered the database by removing genes with 1 and 0 counts, resulting in a database of 23,737 genes. We used edgeR v.3.12.1 (Robinson, McCarthy, & Smyth, 2010) for DE analysis. After filtering out lowly expressed genes (cpm < 2), we constructed a general linearized model (GLM) with species and time point as factors to implement the following tests: (a) "Intraspecific day/night" to detect genes with DE between day (11 a.m.) and night (1 a.m.) time points within each species, and (b) "Interspecific C_3/CAM" to detect DE genes between CAM and C_3 species at the day and night time points, respectively. Significance was tested using likelihood-ratio tests and p values were corrected for multiple testing using the FDR. To avoid potential complications arising from copy number variation, we removed genes with evidence for copy number variants (CNVs; above) before further analysis, unless otherwise noted. Gene Ontology (GO) terms were extracted for 20,165 genes (74.61% of the annotated genes) with Blast2GO (Conesa & Götz, 2008) in April 2017. GO enrichment analyses were performed with the R package, topGO v.2.22.0, (Alexa & Rahnenfuhrer, 2016) using Fisher's exact test (one-sided) and the weight01 algorithm. GOplot v1.0 (Walter, Sánchez-Cabo, & Ricote, 2015) was used to graphically represent the significantly enriched GO terms and visualize enrichment for the "biological processes" domain by calculating the z-score as z = (up − down)/√count).

3 | RESULTS

3.1 | Phenotypic variation captured

We reconstructed ancestral photosynthetic syndromes along the most informative and densely sampled phylogenetic tree available for tillandsioid bromeliads (Barfuss et al., 2016) accounting for possible differences in diversification and extinction rates, as well as modes of trait transitions (FitzJohn, 2012). The results indicated C_3 as ancestral and CAM as the derived state in the radiation of genus, Tillandsia sensu *lato*, with multiple repeated transitions between C_3 and CAM (Figures 1 and S1). The opposite pattern (CAM as ancestral) had previously been inferred based on much sparser phylogenetic sampling and without accounting for differences in rates of phenotypic diversification (Males & Griffiths, 2018).

The CAM species, sampled for this study, display several phenotypic traits thought to represent adaptations to xeric conditions (Benzing, 2000) that likely evolved in a correlated, contingent manner (Givnish et al., 2014). Notwithstanding taxon-specific idiosyncrasies due to leaf economic constraints (Males & Griffiths, 2018), there is a broad association between CAM and leaf succulence, reduced evaporation, dense trichomes and rosette morphology in CAM tillandsioids (Figure 1).

3.2 | Phylogenomic relationships among C_3 and CAM taxa

Coalescent-based reconstruction using ASTRAL resulted in a well-resolved tree, and we recovered major clades identified in previous molecular systematic work (Barfuss et al., 2016). The coalescent tree was broadly congruent with a ML phylogeny estimated using RAxML

(Figures S2 and S3; Text S1). Based on the traditional sorting of taxa into discrete C_3 and CAM phenotypes, as well as ASR (above), our phylogenetic tree suggests a minimum of five independent transitions from C_3 to "strong" (sensu Edwards, 2019; Winter, 2019) CAM photosynthesis among the species sampled (Figure 1). Placement of CAM and C_3 taxa in the SplitsTree network indicates no apparent signals of evolutionary reticulation at this phylogenetic scale (Figure S4). This is important information as it speaks against a role for interspecific gene flow as a likely cause of patterns of phenotypic trait differences observed within this study.

3.3 | Molecular phenotypes capturing photosynthetic syndromes

Carbon isotope ratios, recovered under greenhouse conditions, indicate a continuum of values ranging from typical C_3 to fairly strong CAM (Figure 1), following commonly used thresholds (Crayn et al., 2004; Silvera et al., 2010). 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 Figure 1). 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 Figure 1). Carbon isotope measurements were taken under standardized, well-watered conditions; thus, it is possible that our phenotyping potentially classified some facultative, presumably "weak" CAM species as C_3 (Silvera et al., 2010; Winter & Holtum, 2002). In effect, we used isotopic ratios as a proxy to partition species according to the extremes of the C_3/CAM distribution for evolutionary analyses. This is a pragmatic strategy, because commonly used tests for adaptive protein and gene family evolution benefit greatly from discrete categorization of phenotypes. It is also a conservative strategy, since phenotyping error (if present) would likely diminish the signal-to-noise ratio.

To support assignments of photosynthetic syndrome based on $\delta^{13}C$ values, we also surveyed metabolic phenotypes of all species included in this study using GC–TOF-MS for metabolomic profiling of green tissues sampled at 11 a.m. and 1 a.m., congruent with the gene expression profiling sampling. Partial least squares discriminant analysis (PLS-DA) of 32 putatively identified metabolic compounds (comprising mainly amino acids, carbohydrates and organic acids; Table S3) indicated broad metabolic differentiation between CAM and C_3 plants, especially along the second principal axis for both sampling time points, day and night (Figure 2). These patterns were driven primarily by organic acids, such as malic and fumaric acid, as expected for CAM plants (Benzing, 2000; Popp, Janett, Lüttge, & Medina, 2003), but also by soluble sugars. Closer inspection of organic acids revealed complex patterns of compound accumulation across taxa and sampling time points (Figure 1 and Table S3). Most conspicuously, species with strong CAM phenotypes, such as *T. ionantha* and *T. fasciculate*, showed strongly increased night-time malic acid abundances, compared with most C_3 species studied (Figure 1). In some of the weaker CAM species, such as *T. floribunda*, night-time malic acid

abundances were inconspicuous, but citric acid abundances were increased instead. *Tillandsia australis*, our C_3 reference species for expression profiling, did not exhibit increased accumulation of either CAM-related carbohydrate. A general pattern emerged with respect to carbohydrates (Figure 2c), with more abundant soluble sugars in C_3 compared to CAM plants.

3.4 | Branch-specific positive selection in coding sequences

To identify genes that underwent adaptive protein evolution during C_3/CAM transitions, we scanned the gene space of all genome-sequenced species for branch-specific signatures of positive selection (Zhang, Nielsen, & Yang, 2005). Stringently implemented tests for positive selection in the coding regions of 13,603 genes revealed 22 genes that have apparently undergone adaptive protein evolution along branches relevant to C_3/CAM shifts (Table S4). This includes two transcription factors and eight genes of relevance in the context of the xeric adaptive syndrome associated with CAM in tillandsioids. Eleven of the 22 genes under selection were identified as being differentially expressed in one or more of our CAM-related temporal and interspecific DE comparisons (Table S4; below). The 22 selected genes also include two genes involved in regulation of carbohydrate fluxes in plant tissues, an enolase (gene Aco020962) and a glucose6phosphate dehydrogenase (G6PD; gene Aco012435). The former encodes a homolog of AtENO2/LOS2, a bifunctional gene encoding both an enolase that catalyses the conversion of 2-phosphoglycerate to phosphoenolpyruvate, and MBP-1, a transcriptional repressor that plays a role in abscisic acid (ABA)-mediated response to abiotic stress (Kang et al., 2013).

3.5 | Gene family evolution

We used birth–death models implemented in CAFÉ (Han et al., 2013) to investigate the impact of gene duplications and losses on the repeated evolution of the CAM-correlated trait syndrome (Table S5). Copy number variant (CNV) calling in the subset of species sequenced with high coverage (Figure 3) identified gains in 2,808 and losses in 1,749 genes out of the 19,728 genes surveyed for CNVs. Clustering of the *A. comosus* proteome into gene families resulted in a total of 8,418 clusters, of which 7,216 had non-zero counts in at least one *Tillandsia* species and were retained for further analysis. Estimation of gene birth ($\lfloor$) and death ($\lceil$) rates revealed distinct evolutionary dynamics among different clades within *Tillandsia* (Figure 3). Rates of duplication and loss were almost threefold higher in the clade containing species of the subg. *Tillandsia* with a Central and North American distribution, including *T. fasciculata*, *T. juncea*, *T. leiboldiana* and *T. trauneri* (Granados Mendoza et al., 2017) when compared with the rest of the tree ($\lfloor_2 = 0.002841$, $\lceil_2 = 0.000865$ and $\lfloor_1 = 0.000795$, $\lceil_1 = 0.000239$, respectively).

Mining for gene families with an increase in duplication or loss rate in association with trait shifts (C_3 to CAM) along the tree resulted
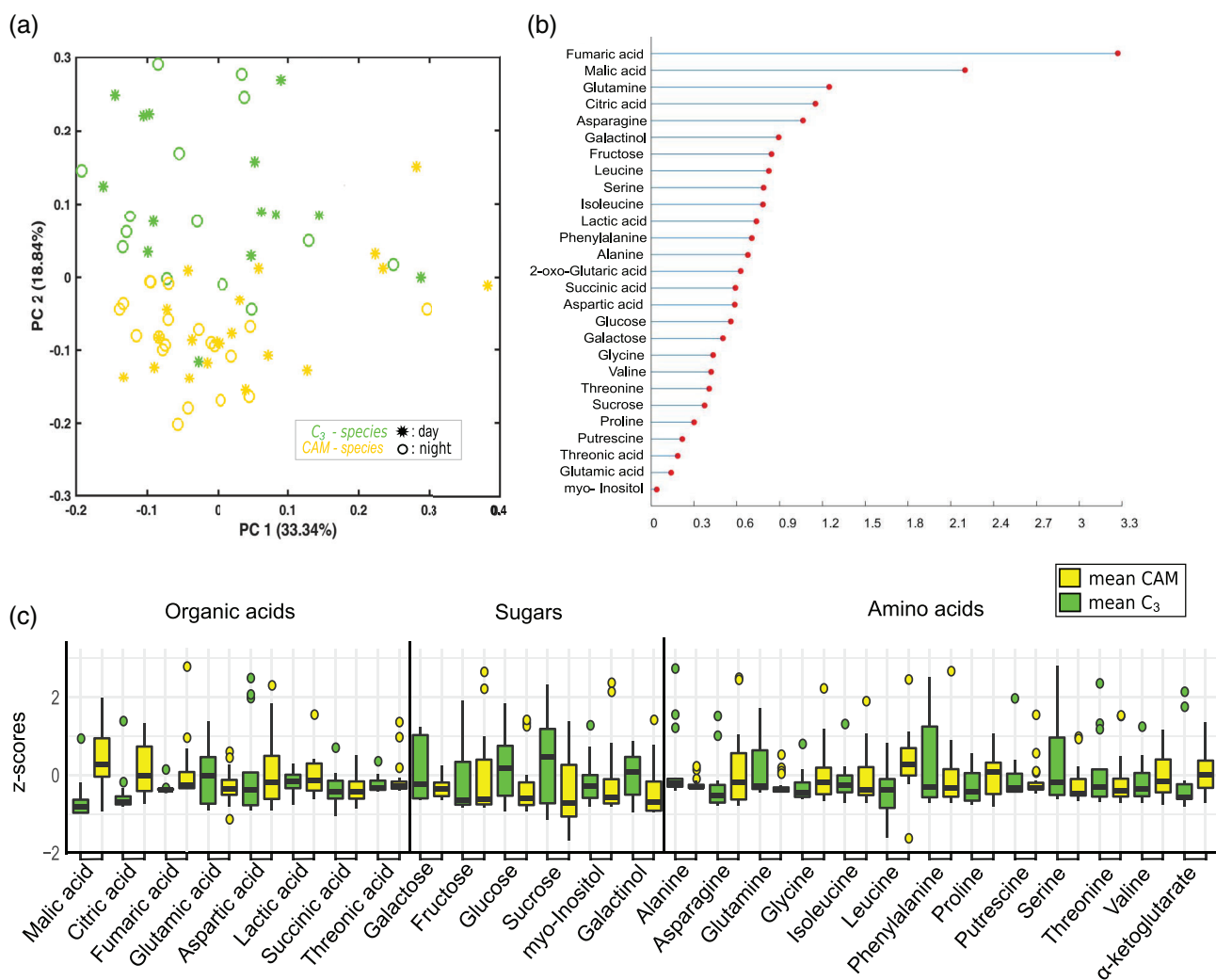
FIGURE 2 Partial least squares-discriminant analysis (PLS-DA) of targeted metabolite data. GC–TOF-MS analysis of green tissue sampled during the day (11 a.m.) and night (1 a.m.). (a), PLS-DA score plot, C_3/CAM day and night affiliation indicated in legend. (b), Loadings. (c), Boxplots of photosynthesis-relevant metabolites measured in C_3 and CAM plants at 1 a.m. The y-scale was truncated at a z-score of 3 for better visibility of interquartile ranges (boxes) and medians (solid lines) [Colour figure can be viewed at wileyonlinelibrary.com]

in five candidate gene families showing duplication rate increases and 12 families showing increased rates of loss along more than one CAM branch (Table 1). At least two of these are directly relevant to the adaptive trait syndrome associated with CAM plants: Cluster_3372 consists of trigger factor chaperones, a small family of proteins targeted to the chloroplast where they may contribute to energy metabolism by facilitating correct maturation of chloroplast-encoded proteins (Rohr et al., 2019). Deletion mutants of the *Arabidopsis thaliana* homolog, *AtTIG1*, show reduced linear electron flow under prolonged high-light intensities (Rohr et al., 2019), suggesting a possible link to adaptation to xeric, high-light environments. Cluster_7372, on the other hand, contains a single gene, *Aco019534*, which encodes a homolog of *XAP5 CIRCARDIAN TIMEKEEPER* (*XCT*) in *A. thaliana* (Figure 3). *AtXCT* is a regulatory protein involved in regulation of circadian period length, developmental processes, such as tissue greening and chloroplast development, and important for sugar- and light quality-dependent ethylene signalling related to growth (Ellison,

Vandenbussche, Van Der Straeten, & Harmer, 2011; Martin-Tryon & Harmer, 2008). Moreover, *AtXCT* is known to be a key player in regulation of all three major classes of small RNAs in *A. thaliana* (Fang, Shi, Lu, Chen, & Qi, 2015).

3.6 | Transcriptome-wide gene expression profiling across the C_3/CAM continuum

To identify genes potentially involved in C_3/CAM transitions and related trait shifts, we examined transcriptome-wide changes in gene expression by RNA sequencing (RNA-seq). One C_3 species (*Tillandsia australis*; *Taust* from here onwards) and three CAM species (*T. sphaerocephala*, *Tspha*; *T. fasciculata*, *Tfasc*; *T. floribunda*, *Tflor*) were selected for RNA-seq (Figure 1; Table S2). These species were chosen based on their range of carbon isotope phenotypes, the local availability of biological replicates, and because they represent C_3/CAM
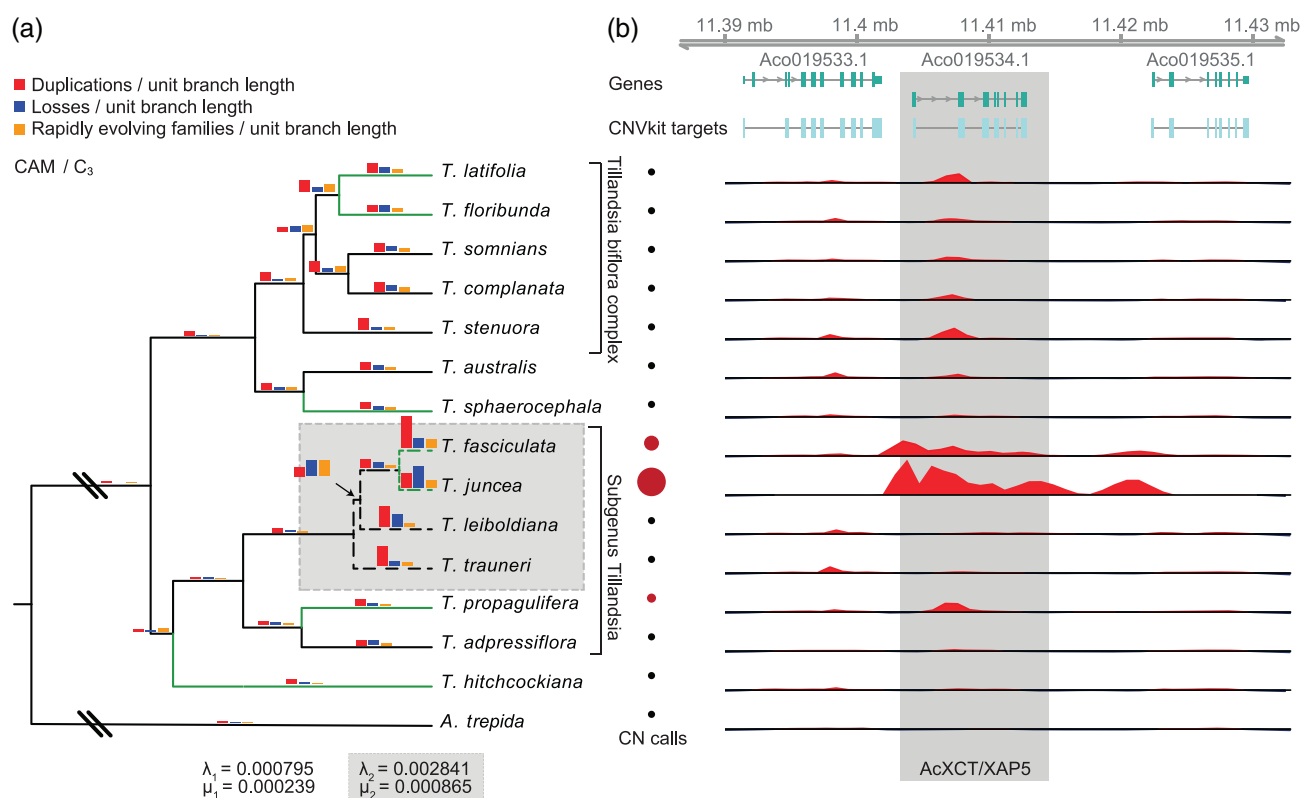
FIGURE 3 High-coverage samples included in analysis of copy number (CN) variants. (a), Bar plots indicate inferred numbers of gene duplications (red) and losses (blue), and numbers of families evolving at significantly different rates ($p < .01$) compared to the genomic background (orange), all scaled by the lengths of the respective branches. Genome-wide rates of duplication (λ) and loss (μ) are given for subgenus, *Tillandsia* ($\lambda2$, $\mu2$, grey shading) and the rest of the tree ($\lambda1$, $\mu1$). Colouring on branches indicates whether branches are CAM (green) or C_3 (black). (b), Estimated CNs and coverage distributions for *AtXCT* homolog, Aco019534, one of the six gene families whose rates of duplication and loss showed significant association with the CAM phenotype [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Gene families with putative association to C_3 vs CAM phenotype in *Tillandsia* spp.

Cluster	F	p value[*]	Direction	Annotation
cluster_3372	2.650	0	Gain	Trigger factor
cluster_1656	2.771	0	Gain	Apoptosis-inducing factor
cluster_289	2.650	0	Gain	RNA recognition motif-containing protein
cluster_3811	2.667	0	Gain	Protein of unknown function DUF1068
cluster_7372	1.253	0	Gain	XAP5 CIRCADIAN TIMEKEEPER
cluster_759	2.667	0	Loss	Ypt/Rab-GAP domain of gyp1p superfamily protein
cluster_3828	2.667	0	Loss	ER membrane protein complex subunit 2-A
cluster_773	1.642	0	Loss	FASCICLIN-like arabinogalactan 1
cluster_218	1.479	0	Loss	Oxidoreductase, aldo/keto reductase family protein
cluster_8075	4.155	0	Loss	Uncharacterized protein
cluster_3083	4.155	0	Loss	Acyl-CoA *N*-acyltransferase
cluster_933	2.667	0	Loss	Hypothetical protein
cluster_21	1.854	0	Loss	2-oxoglutarate 2OG and FeII-dependent oxygenase
cluster_8060	3.097	0	Loss	RNA recognition motif family protein
cluster_733	2.667	0	Loss	Growth-regulating factor 2
cluster_166	3.097	0.036	Loss	Lipoxygenase 3
cluster_4488	3.823	0.0495	Loss	Hypothetical protein

[*]*p* values adjusted for false discovery as described in text.

comparisons with different evolutionary distances between each CAM species and the C$_3$ reference taxon (Figure 1). *Taust*, our C$_3$ reference taxon, exhibited δ^{13}C values of −26 to −29‰, clearly beyond the −20‰ threshold commonly used to classify C$_3$ plants (Silvera et al., 2010). This species also exhibited all other phenotypic features typically expected for C$_3$ bromeliads, including tank-forming rosettes, no succulence and absence of dense trichome cover. The three CAM taxa, sampled for expression profiling, exhibited a broad range of CAM-like δ^{13}C values and morphological features (Table S2). Species were sampled at 1 a.m. and 11 a.m., corresponding to the highest and lowest peaks, respectively, of net CO$_2$ assimilation rates (phase I, carboxylation at night and phase III, decarboxylation at day) and apparent PEPC kinase (PPCK) activation state of the CAM species *T. usneoides* (Dodd, Borland, Haslam, Griffiths, & Maxwell, 2002). Between 18,439 and 20,378 genes were successfully recovered per individual sample, representing between 68.2 and 75.4% of the 27,024 annotated genes presented in the *A. comosus* reference genome assembly. Multi-dimensional scaling (MDS) of RNA-seq data revealed clear clustering of biological replicates within each species, and patterns of interspecific differentiation along the first (=horizontal) axis broadly reflected the known phylogenetic relationships among the studied taxa (Figure S5).

3.7 | Transcriptome-wide gene expression profiling: Intraspecific day/night comparisons

Comparisons of our day and night sampling time points within each species revealed that between 82 (*Tflor*) and 1354 (*Tfasc*) genes were differentially expressed at FDR 5% and log fold change (LFC) > 1 between day and night conditions across the four tested species (Figure S6). On average, 65.4% of them were over-expressed during the night, with few apparent differences shared among the three CAM species to the exclusion of our C$_3$ reference taxon (Figure S6a). Only two genes exhibited differential expression between day and night samples in all three CAM but not the C$_3$ species. One of these, *Aco003903*, encodes a regulatory protein homologous to *A. thaliana* *ABF2/ABF3*. In *A. thaliana*, this gene family is involved in abscisic acid (ABA)-mediated stress response to sugar-, salt- and osmotic-stress (Garcia, Lynch, Peeters, Snowden, & Finkelstein, 2008). The other gene (*Aco016050*) is of unknown function but encodes an ACT amino acid-binding domain (PF01842) and may, therefore, play a role in amino acid metabolism. Differential gene expression (DE) results and gene ontology (GO) enrichments yielded patterns of up- and down-regulated genes and expressed metabolic pathways consistent with expected differences between day and night (Figure S7; Text S1), including a range of photosynthesis-related GO terms.

3.8 | Transcriptome-wide gene expression profiling: Interspecific C$_3$/CAM comparisons

Comparisons between each of our CAM species and our C$_3$ reference taxon revealed: between 1302 (*Tflor*) and 2757 (*Tfasc*) genes that were DE between CAM and C$_3$ species during day conditions (11 a.m.), with 49.9% of genes, significantly up-regulated in CAM species. Under night conditions (1 a.m.), 51.7% of the genes were up-regulated in CAM species, affecting between 1460 (*Tflor*) and 3110 (*Tfasc*) genes (Figure 4a). Hundreds of DE transcripts were shared among all three CAM species in these comparisons, with more shared DE genes during the night (336) than during the day (233; Table S6). Large numbers of DE transcripts were unique to each CAM taxon, indicating abundant lineage-specific changes (Figure 4a). As expected, *Tfasc*, the species with the strongest CAM phenotype and the greatest phylogenetic distance to the C$_3$ reference taxon, exhibited by far the greatest number of unique DE transcripts (1,838 during the day and 1,881 during the night; Figure 4a).
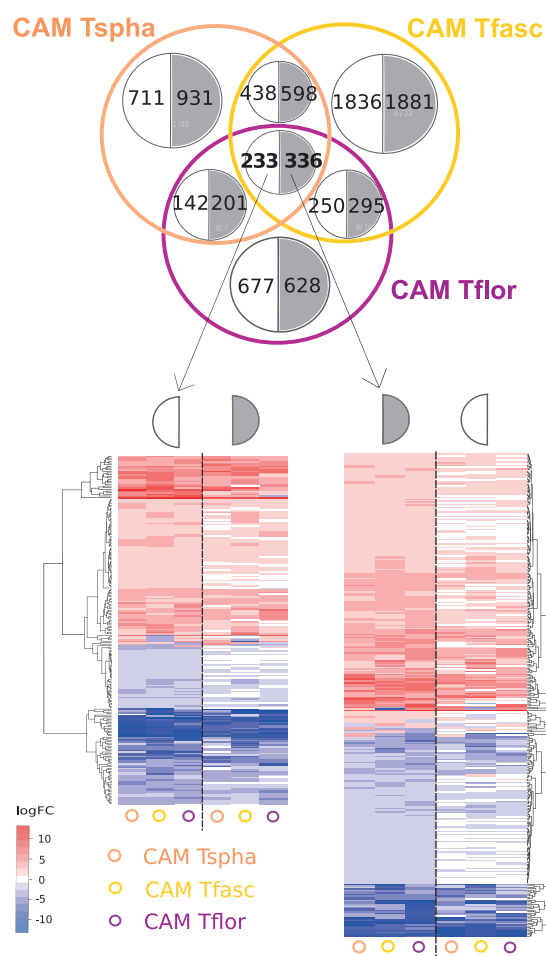
Genes with DE in interspecific C$_3$/CAM comparisons common to all three CAM species were enriched for a range of CAM-related metabolic processes in photosynthetic leaf tissues (Figures 4b and S8). For example, day time points were enriched for genes involved in carbon fixation (GO:0015977), carbohydrate metabolic process (GO:0005975) and the TCA cycle (GO:0006099). In turn, DE genes, at night, showed enrichment for vacuolar acidification (GO:0007035), sucrose transport (GO:0015770) and glycolysis (GO:0006096), among others. These results provide evidence for large-scale transcriptional reprogramming of CAM-related pathways shared among CAM *Tillandsia* species. A more detailed exploration of regulatory changes, shaping carbohydrate metabolism, is given in supplementary materials (Text S1; Figure S9).

Among the common C$_3$/CAM DE genes (up-regulated in all CAM compared to the C$_3$ species) there are also several genes involved in abscisic-acid signalling (GO:0009738), including *Aco005513*, a homolog of *AtPYL*, an ABA receptor involved in drought response and leaf senescence in *A. thaliana* (Zhao et al., 2016), and *Aco004854*, a further member of the ABF2/ABF3 gene family involved in ABA-mediated response to sugar, drought and osmotic stress (Garcia et al., 2008). DE genes also included five LEA genes (Table S6) likely involved in drought protection (Hong-Bo, Zong-Suo, & Ming-An, 2005) and *Aco004804*, a homolog of the *A. thaliana* gene, *YLS7*, identified as a QTL for drought resistance in this species (El-Soda, Kruijer, Malosetti, Koornneef, & Aarts, 2015).
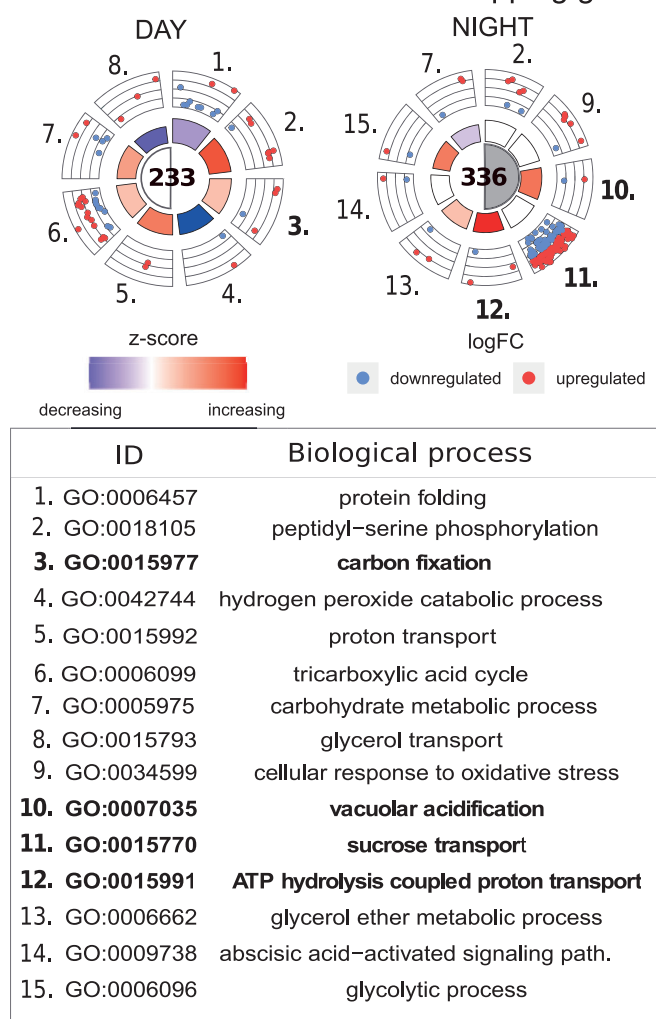
3.9 | Targeted DE analysis for CAM-related genes from pineapple

Results from targeted DE analysis for CAM-related genes identified in pineapple (Ming et al., 2015; Wai et al., 2017) mirrored those from transcriptome-wide gene expression profiling, again pointing to important roles for the upregulation of malate transporters and vacuolar proton pumps (enabling metabolite transport and leaf succulence) in CAM species and features of carbohydrate fluxes relevant to stress response. Day/night expression information, available for *A. comosus*, allowed us to examine similarities and differences among CAM-related diel expression patterns between *Tillandsia* spp. and the CAM plant pineapple at equivalent time points (Figure S10).

(a) Differential gene expression in CAM

(b) GO terms enrichments on overlapping genes

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

FIGURE 4 Transcriptome-wide analysis of differential gene expression (DE) between three CAM *Tillandsia* species and a C_3 reference taxon. (a) Venn chart depicting overlap in DE patterns for each of the three CAM taxa relative to the C_3 reference taxon during the day (11 a.m., numbers in light half-circles) and during the night (1 a.m., numbers in grey-shaded half-circles). Genes with FDR < 0.05 were considered significant. Heatmaps illustrate expression changes for the 233 and 336 transcripts with DE in all three CAM species at day and night, respectively, compared to the same set of transcripts at the respective alternative sampling time point (night and day). The similarities in expression patterns of most DE transcripts between day and night are clearly visible. (b) GO enrichment analysis on the subsets of overlapping significant genes during the day (GO terms n° 1 to n° 8), during the night (n°2, n°7, and n° 9 to n° 15). For each subcategory (day, night) we depict the eight 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) [Colour figure can be viewed at wileyonlinelibrary.com]

Perhaps, most obviously, our comparisons indicated diel cycling of the key post-translational regulator of CAM photosynthesis, phospho-enolpyruvate carboxylase kinase (PPCK, also commonly referred to as PEPC kinase, *A. comosus* gene model *Aco13938*; Figure S10). This important CAM-related gene was significantly upregulated at night in all three CAM *Tillandsia* species studied (*Tspha*, *Tflor*, *Tfasc*), following the pattern expected from the CAM plant, *A. comosus* (*Acomo*) Figure S10). Upregulation of PPCK at night was also significant for our C_3 *Tillandsia* reference taxon (*Taust*). This is suggestive of pre-adaptation of C_3 tillandsioids for the evolution of CAM and mirrors results obtained on C_3 to C_4 transition in *Flaveria* spp (Aldous et al., 2014), suggesting that PPCK night-time expression may be a conserved pattern among C_3 species. PPCK was significantly upregulated in interspecific C_3/CAM comparisons involving all three CAM *Tillandsia* taxa studied (*Tspha*, *Tflor*, *Tfasc*) relative to our C_3 reference (*Taust*) (Figure S10), consistent with the direct involvement of this important protein in CAM photosynthesis in *Tillandsia* spp. We also detected significant upregulation of a *PEPC* gene (*Aco010025*) in all three CAM *Tillandsia* spp., indicating that this is the homolog involved in nocturnal carbon fixation, consistent with its proposed function in pineapple.

4 | DISCUSSION

Our combined genomic, transcriptomic and metabolic data, including tests for adaptive protein evolution, gene family evolution and gene expression changes, point to extensive associations between CAM and numerous other plant traits linked to xeric adaptation, as previously suggested by evolutionary biologists (Benzing, 2000; Crayn et al., 2004; Givnish et al., 2014) and plant physiologists (Borland et al., 2014; Pierce, Winter, & Griffiths, 2002; Silvera et al., 2010; Vaasen, Begerow, & Hampp, 2006; Winter, Holtum, & Smith, 2015). Combining evolutionary genomic signatures with transcriptomic and targeted metabolite analyses indicated genetic and molecular components underpinning this integrated adaptive trait syndrome in this Neotropical adaptive radiation. Viewed together, our data paint a picture of complex interactions among evolutionary changes to gene expression levels, metabolite-based regulation and physiological transitions accompanied by both adaptive protein and gene family evolution.

Among our most salient results, we discovered the expansion of the *XAP5/XCT* gene in several CAM lineages (Figure 2, right panel). This gene encodes a highly pleiotropic regulator known to be involved in a variety of processes, including light-dependent gene expression and developmental processes in *A. thaliana* (Ellison et al., 2011; Fang et al., 2015; Martin-Tryon & Harmer, 2008). CAM genes are regulated in a circadian-clock dependent manner in the bromeliad, *A. comosus* (Ming et al., 2015) and duplication of *XCT* could have helped to mediate light-based reprogramming of transcription in CAM species.

Besides pleiotropic transcription factors, we also found indications for alterations to genes involved in regulation at the metabolic level. In particular, branch-specific tests for selection highlighted the enolase, *Aco020962*, a homolog of *AtENO2/LOS2*. This gene is of particular interest to correlated evolution involving CAM, drought response and carbon fluxes, since it provides a direct link between glycolysis/gluconeogenesis and the CAM pathway via PEP, catalysed by the enolase-coding isoform, and ABA-dependent response to abiotic stress mediated by the regulatory protein, MBP-1, transcribed from an alternative start codon at the same locus (Kang et al., 2013). Notably, this gene exhibited significant temporal and interspecific C_3/CAM expression changes in *T. fasciculata* (*Tfasc*), the strongest CAM species in our transcriptome study. CNV data indicate this gene was present in multiple copies in *Tillandsia* genomes, and it is thus possible that increased expression in the strong CAM plant *Tfasc* and a high estimate of non-synonymous substitutions were due to recent gene duplication, hypotheses that remain to be tested in the future.

Gene expression profiling using RNA-seq uncovered apparent regulatory changes shared among the three CAM species for hundreds of genes (Figure 4; Table S6), including significant upregulation of malate transporters and vacuolar proton pumps, commonly associated with the important adaptive plant trait succulence (Borland et al., 2014; Popp et al., 2003; Silvera et al., 2010) and the full expression of CAM (Heyduk et al., 2019), and extensive upregulation of transcripts involved in carbohydrate metabolism and fluxes (Figure S10) as expected for CAM plants (Ceusters, Borland, & De Proft, 2009).

Shared gene expression patterns, accompanying C_3/CAM shifts, suggest that the regulatory networks required for biochemical changes in CAM species are largely in place also in C_3 taxa, broadly consistent with the presence of a "C_3 + CAM" category of species with C_3 carbon isotope phenotypes and the ability to express small amounts of CAM (Edwards, 2019). In addition to differential expression of genes directly related to CAM photosynthesis, we also found changes in gene expression and/or gene family expansion for several transcription factors and a sensor involved in ABA signalling, providing possible candidates for regulators mediating this transition.

Several caveats are necessary when interpreting our results. Firstly, we used binary classifications of C_3 and CAM phenotypes, which greatly facilitated the study of adaptive protein and gene family evolution but prevented us from tackling the full range of intermediate and facultative phenotypes that exist between typical C_3 and strong CAM (Edwards, 2019; Silvera et al., 2010; Winter, 2019). Secondly, our RNA-seq experimental design, involving a single C_3 reference taxon (*T. australis*), did not allow us to fully disentangle truly convergent from lineage-specific changes in phenotypes. Thus, although our study contributes to a fuller understanding of the mechanistic basis of repeated adaptive trait shifts between C_3 and CAM, uncovering the genetic basis of convergent evolution (Heyduk et al., 2019; Yeaman, Gerstein, Hodgins, & Whitlock, 2018) remains a task for future research. In the same context, it appears that *T. australis* represents a reversal from CAM to C_3 (Figure 1). While this supports the relative ease of transition between these two states, it also suggests that more C_3 taxa, including species with truly ancestral C_3 state, should be included in future gene expression studies. Thirdly, our naturally limited taxon sampling for this combined WGS/transcriptomic/metabolic profiling effort may have led to an overestimation of the true number of changes accompanying C_3/ CAM trait shifts (Heyduk et al., 2019). Nevertheless, it is re-assuring that most species, categorized as C_3 or CAM, in our study, also exhibit the morphological/anatomical features commonly associated with these photosynthetic syndromes (e.g. absence or presence of succulence and dense trichome cover in C_3 and CAM species, respectively). It has recently been argued that anatomical features represent the actual rate-limiting step during evolution of strong CAM (Edwards, 2019), which suggests that our study indeed includes the full breadth of relevant phenotypes.

A persistent pattern recovered by both global and targeted gene expression profiling in our study was the general absence of a clear temporal signature of diel cycling of CAM genes that distinguishes CAM- and C_3-like species (Figure 4). Instead, our results point to sweeping regulatory changes to key CAM genes and many other pathways relevant for the CAM phenotype, often evident during both day and night. Consequently, our data suggest that rather than altering the timing of expression, relevant genes already follow CAM-like expression patterns in the C_3 species but their expression is amplified in CAM species. These observations provide mechanistic insights into the evolutionary mechanisms supporting the so-called C_3/CAM continuum (Silvera et al., 2010; Winter et al., 2015) and suggest that genes already following a CAM-like expression pattern are

preferentially co-opted into the CAM pathway, perhaps as suggested by Bräutigam, Schlüter, Eisenhut, and Gowik (2017); but see Edwards, 2019). This may explain the predominance of repeated co-option of the same homologs among our three CAM species surveyed and mirrors observations made in C_4 grasses, where recurrent co-option of specific C_4 homologs in independent origins of C_4 photosynthesis was driven by expression levels of the respective gene copies in photosynthetic leaves of the C_3 ancestor (Moreno-Villena, Dunning, Osborne, & Christin, 2018). In contrast to a recent high-profile study of the obligate CAM plant, *Kalanchoë* (Yang et al., 2017), our study also revealed a number of genes affected by both repeated (i.e., potentially convergent) protein sequence evolution *and* gene expression evolution, a pattern that appears to be rare in nature (Yang et al., 2017).

In summary, our combined whole-genome, transcriptome, and metabolic data point to a central role of highly pleiotropic regulators and tightly interconnected drought response-related pathways to the repeated origin of CAM photosynthesis in tillandsioid bromeliads, involving both "top down" and "bottom up" changes. We thus hypothesize that the striking correlated trait shifts seen in this textbook radiation are triggered by relatively few changes at highly pleiotropic regulators, rather than alternative mechanisms such as the genomic clustering of adaptive mutations. The repeated trait shifts studied here may be due to convergent evolution, or alternatively they may trace back to early evolutionary events at the onset of the radiation, effectively generating a genomic and transcriptomic toolkit that was later fine-tuned in individual lineages. We look forward to seeing these views corroborated or challenged by genomic, functional and ecophysiological data to emerge in years to come.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Christian Lexer, Marylaure De La Harpe, Margot Paris, Jaqueline Hess, and Nicolas Salamin conceived the study, Christian Lexer and Nicolas Salamin provided funding, Marylaure De La Harpe, Margot Paris, Michael Harald Johannes Barfuss, Arindam Ghatak, and Walter Till collected data, Marylaure De La Harpe, Margot Paris, Jaqueline Hess, Martha Liliana Serrano-Serrano, Palak Chaturvedi, and Wolfram Weckwerth analysed data, and Marylaure De La Harpe, Margot Paris, Jaqueline Hess and Christian Lexer wrote the paper with input and revisions from all co-authors.

DATA AVAILABILITY

The datasets generated during and/or analysed during the current study are available as extensive online supporting information, and all additional information, scripts and protocols are available from the corresponding author on request. All raw sequence reads will be deposited in NCBI-SRA (accession ID XXXX).

ORCID

Marylaure De La Harpe https://orcid.org/0000-0002-8353-714X
Margot Paris https://orcid.org/0000-0001-7328-3820
Jaqueline Hess https://orcid.org/0000-0003-3281-5434
Michael Harald Johannes Barfuss https://orcid.org/0000-0001-7172-9454
Martha Liliana Serrano-Serrano https://orcid.org/0000-0003-3313-725X
Arindam Ghatak https://orcid.org/0000-0003-4706-9841
Palak Chaturvedi https://orcid.org/0000-0002-5856-0348
Wolfram Weckwerth https://orcid.org/0000-0002-9719-6358
Nicolas Salamin https://orcid.org/0000-0002-3963-4954
Ching Man Wai https://orcid.org/0000-0002-2913-5721
Ray Ming https://orcid.org/0000-0002-9417-5789
Christian Lexer https://orcid.org/0000-0002-7221-7482

REFERENCES

Aldous, S. H., Weise, S. E., Sharkey, T. D., Waldera-Lupa, D. M., Stühler, K., Mallmann, J., ... Arsova, B. (2014). Evolution of the phosphoenolpyruvate carboxylase protein kinase family in C_3 and C_4 Flaveria spp. *Plant Physiology*, *165*(3), 1076–1091.

Alexa A and Rahnenfuhrer J. 2016. topGO: Enrichment analysis for gene ontology. Retrieved from ttp://bioconductor.org/packages/release/bioc/html/topGO.html

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169.

Barfuss, M. H., Samuel, R., Till, W., & Stuessy, T. F. (2005). Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) based on DNA sequence data from seven plastid regions. *American Journal of Botany*, *92*(2), 337–351.

Barfuss, M. H., Till, W., LEME, E. M., Pinzón, J. P., Manzanares, J. M., Halbritter, H., ... Brown, G. K. (2016). Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa*, *279*(1), 1–97.

Barraclough, T. G., & Vogler, A. P. (2000). Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, *155*(4), 419–434.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Benzing, D. H. (2000). *Bromeliaceae: Profile of an adaptive radiation*. Cambridge: Cambridge University Press.

Borland, A. M., Hartwell, J., Weston, D. J., Schlauch, K. A., Tschaplinski, T. J., Tuskan, G. A., ... Cushman, J. C. (2014). Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science*, *19*(5), 327–338.

Bräutigam, A., Schlüter, U., Eisenhut, M., & Gowik, U. (2017). On the evolutionary origin of CAM photosynthesis. *Plant Physiology*, *174*(2), 473–477.

Cavender-Bares, J., Kozak, K. H., Fine, P. V., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, *12*(7), 693–715.

Ceusters, J., Borland, A. M., & De Proft, M. P. (2009). Drought adaptation in plants with crassulacean acid metabolism involves the flexible use of different storage carbohydrate pools. *Plant Signaling & Behavior*, *4* (3), 212–214.

Conesa, A., & Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, *2008*, 1–12.

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

Crayn, D. M., Winter, K., Schulte, K., & Smith, J. A. C. (2015). Photosynthetic pathways in Bromeliaceae: Phylogenetic and ecological significance of CAM and C_3 based on carbon isotope ratios for 1893 species. *Botanical Journal of the Linnean Society*, *178*(2), 169–221.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

Dasmahapatra, K. K., Walters, J. R., Briscoe, A. D., Davey, J. W., Whibley, A., Nadeau, N. J., ... Salazar, C. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*(7405), 94.

de La Harpe, M., Hess, J., Loiseau, O., Salamin, N., Lexer, C., & Paris, M. (2019). A dedicated target capture approach reveals variable genetic markers across micro-and macro-evolutionary time scales in palms. *Molecular Ecology Resources*, *19*(1), 221–234.

Dodd, A. N., Borland, A. M., Haslam, R. P., Griffiths, H., & Maxwell, K. (2002). Crassulacean acid metabolism: Plastic, fantastic. *Journal of Experimental Botany*, *53*(369), 569–580.

Edwards, E. J. (2019). Evolutionary trajectories, accessibility and other metaphors: The case of C_4 and CAM photosynthesis. *New Phytologist*, *223*(4), 1742–1755.

Ellison, C. T., Vandenbussche, F., Van Der Straeten, D., & Harmer, S. L. (2011). XAP5 CIRCADIAN TIMEKEEPER regulates ethylene responses in aerial tissues of Arabidopsis. *Plant Physiology*, *155*(2), 988–999.

El-Soda, M., Kruijer, W., Malosetti, M., Koornneef, M., & Aarts, M. G. (2015). Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of Arabidopsis thaliana to drought. *Plant, Cell & Environment*, *38*(3), 585–599.

Enright, A. J., Van Dongen, S., & Ouzounis, C. A. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research*, *30*(7), 1575–1584.

Fang, X., Shi, Y., Lu, X., Chen, Z., & Qi, Y. (2015). CMA33/XCT regulates small RNA production through modulating the transcription of dicerlike genes in Arabidopsis. *Molecular Plant*, *8*(8), 1227–1236.

FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. Methods in ecology and evolution. In press. https://doi.org/10.1111/j.2041-210X.2012.00234.x.

Folk, R. A., Sun, M., Soltis, P. S., Smith, S. A., Soltis, D. E., & Guralnick, R. P. (2018). Challenges of comprehensive taxon sampling in comparative biology: Wrestling with rosids. *American Journal of Botany*, *105*(3), 433–445.

Garcia, M. E., Lynch, T., Peeters, J., Snowden, C., & Finkelstein, R. (2008). A small plant-specific protein family of ABI five binding proteins (AFPs) regulates stress response in germinating Arabidopsis seeds and seedlings. *Plant Molecular Biology*, *67*(6), 643–658.

Gitaí, J., Paule, J., Zizka, G., Schulte, K., & Benko-Iseppon, A. M. (2014). Chromosome numbers and DNA content in Bromeliaceae: Additional data and critical review. *Botanical Journal of the Linnean Society*, *176*(3), 349–368.

Givnish, T. J., Barfuss, M. H., Van Ee, B., Riina, R., Schulte, K., Horres, R., ... Winter, K. (2014). Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, *71*, 55–78.

Granados Mendoza, C., Granados-Aguilar, X., Donadío, S., Salazar, G. A., Flores-Cruz, M., Hágsater, E., ... Magallón, S. (2017). Geographic

structure in two highly diverse lineages of Tillandsia (Bromeliaceae). *Botany*, *95*(7), 641–651.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696–704.

Han, M. V., Thomas, G. W., Lugo-Martinez, J., & Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution*, *30*(8), 1987–1997.

Heled, J., & Drummond, A. J. (2009). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*(3), 570–580.

Heyduk, K., Moreno-Villena, J. J., Gilman, I. S., Christin, P. A., & Edwards, E. J. (2019). The genetics of convergent evolution: Insights from plant photosynthesis. *Nature Reviews Genetics*, *20*(8), 485–493.

Hong-Bo, S., Zong-Suo, L., & Ming-An, S. (2005). LEA proteins in higher plants: Structure, function, gene expression and regulation. *Colloids and Surfaces B: Biointerfaces*, *45*(3–4), 131–135.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267.

Kang, M., Abdelmageed, H., Lee, S., Reichert, A., Mysore, K. S., & Allen, R. D. (2013). At MBP-1, an alternative translation product of LOS 2, affects abscisic acid responses and is modulated by the E 3 ubiquitin ligase A tSAP 5. *The Plant Journal*, *76*(3), 481–493.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36.

Kozak, K. H., & Wiens, J. J. (2010). Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters*, *13*(11), 1378–1389.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, *9*(4), 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Litsios, G., & Salamin, N. (2012). Effects of phylogenetic signal on ancestral state reconstruction. *Systematic Biology*, *61*(3), 533–538.

Liu, L., Xi, Z., Wu, S., Davis, C., & Edwards, S. V. (2015). Estimating phylogenetic trees from genome-scale data. arXiv preprint arXiv: 1501.03578.

Males, J. (2016). Think tank: Water relations of Bromeliaceae in their evolutionary context. *Botanical Journal of the Linnean Society*, *181*(3), 415–440.

Males, J., & Griffiths, H. (2018). Economic and hydraulic divergences underpin ecological differentiation in the Bromeliaceae. *Plant, Cell & Environment*, *41*(1), 64–78.

Martin-Tryon, E. L., & Harmer, S. L. (2008). XAP5 CIRCADIAN TIMEKEEPER coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. *The Plant Cell*, *20*(5), 1244–1259.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.

Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., ... Zhang, J. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics*, *47*(12), 1435–1442.

Mirarab, S., & Warnow, T. (2015). ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, *31*(12), i44–i52.

Molloy, E. K., & Warnow, T. (2018). To include or not to include: The impact of gene filtering on species tree estimation methods. *Systematic Biology*, *67*(2), 285–303.

Moreno-Villena, J. J., Dunning, L. T., Osborne, C. P., & Christin, P. A. (2018). Highly expressed genes are preferentially co-opted for C_4 photosynthesis. *Molecular Biology and Evolution*, *35*(1), 94–106.

Nagy, L. G., Riley, R., Bergmann, P. J., Krizsán, K., Martin, F. M., Grigoriev, I. V., … Hibbett, D. S. (2017). Genetic bases of fungal white rot wood decay predicted by phylogenomic analysis of correlated gene-phenotype evolution. *Molecular Biology and Evolution*, *34*(1), 35–44.

Nevado, B., Atchison, G. W., Hughes, C. E., & Filatov, D. A. (2016). Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nature Communications*, *7*(1), 1–9.

Novikova, P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., … Holm, S. (2016). Sequencing of the genus Arabidopsis identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics*, *48*(9), 1077–1082.

Pease, J. B., Haak, D. C., Hahn, M. W., & Moyle, L. C. (2016). Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biology*, *14*(2), e1002379.

Pierce, S., Winter, K., & Griffiths, H. (2002). The role of CAM in high rainfall cloud forests: An in situ comparison of photosynthetic pathways in Bromeliaceae. *Plant, Cell & Environment*, *25*(9), 1181–1189.

Popp, M., Janett, H. P., Lüttge, U., & Medina, E. (2003). Metabolite gradients and carbohydrate translocation in rosette leaves of CAM and C_3 bromeliads. *New Phytologist*, *157*(3), 649–656.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140.

Rohr, M., Ries, F., Herkt, C., Gotsmann, V. L., Westrich, L. D., Gries, K., … Zimmer, D. (2019). The role of plastidic trigger factor serving protein biogenesis in green algae and land plants. *Plant Physiology*, *179*(3), 1093–1110.

Sanderson, M. J. (2003). r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, *19*(2), 301–302.

Silvera, K., Neubig, K. M., Whitten, W. M., Williams, N. H., Winter, K., & Cushman, J. C. (2010). Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology*, *37*(11), 995–1010.

Smeds, L., & Künstner, A. (2011). ConDeTri-a content dependent read trimmer for Illumina data. *PLoS One*, *6*(10), e26314.

Spangenberg, J. E., Jacomet, S., & Schibler, J. (2006). Chemical analyses of organic residues in archaeological pottery from Arbon Bleiche 3, Switzerland–evidence for dairying in the late Neolithic. *Journal of Archaeological Science*, *33*(1), 1–13.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Sun, X., & Weckwerth, W. (2012). COVAIN: A toolbox for uni-and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics*, *8*(1), 81–93.

Talevich, E., Shain, A. H., Botton, T., & Bastian, B. C. (2016). CNVkit: Genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Computational Biology*, *12*(4), e1004873.

Vaasen, A., Begerow, D., & Hampp, R. (2006). Phosphoenolpyruvate carboxylase genes in C_3, crassulacean acid metabolism (CAM) and

C_3/CAM intermediate species of the genus Clusia: Rapid reversible C_3/CAM switches are based on the C_3 housekeeping gene. *Plant, Cell & Environment*, *29*(12), 2113–2123.

Wai, C. M., VanBuren, R., Zhang, J., Huang, L., Miao, W., Edger, P. P., … Smith, J. A. C. (2017). Temporal and spatial transcriptomic and micro RNA dynamics of CAM photosynthesis in pineapple. *The Plant Journal*, *92*(1), 19–30.

Walter, W., Sánchez-Cabo, F., & Ricote, M. (2015). GOplot: An R package for visually combining expression data with functional analysis. *Bioinformatics*, *31*(17), 2912–2914.

Wang, N., Kimball, R. T., Braun, E. L., Liang, B., & Zhang, Z. W. (2012). Ancestral range reconstruction of Galliformes: The effects of topology and taxon sampling. *Journal of Biogeography*, *44*(1), 122–135.

Weadick, C. J., & Chang, B. S. (2012). An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Molecular Biology and Evolution*, *29*(5), 1297–1300.

Weckwerth, W., Loureiro, M. E., Wenzel, K., & Fiehn, O. (2004). Differential metabolic networks unravel the effects of silent plant phenotypes. *Proceedings of the National Academy of Sciences*, *101*(20), 7809–7814.

Winter, K. (2019). Ecophysiology of constitutive and facultative CAM photosynthesis. *Journal of Experimental Botany*, *70*(22), 6495–6508.

Winter, K., & Holtum, J. A. (2002). How closely do the δ13C values of crassulacean acid metabolism plants reflect the proportion of CO2 fixed during day and night? *Plant Physiology*, *129*(4), 1843–1851.

Winter, K., & Holtum, J. A. (2007). Environment or development? Lifetime net CO2 exchange and control of the expression of Crassulacean acid metabolism in *Mesembryanthemum crystallinum*. *Plant Physiology*, *143* (1), 98–107.

Winter, K., Holtum, J. A., & Smith, J. A. C. (2015). Crassulacean acid metabolism: A continuous or discrete trait? *New Phytologist*, *208*(1), 73–78.

Yang, X., Hu, R., Yin, H., Jenkins, J., Shu, S., Tang, H., … Heyduk, K. (2017). The Kalanchoë genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nature Communications*, *8*(1), 1–15.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591.

Yeaman, S., Gerstein, A. C., Hodgins, K. A., & Whitlock, M. C. (2018). Quantifying how constraints limit the diversity of viable routes to adaptation. *PLoS Genetics*, *14*(10), e1007717.

Zhang, J., Nielsen, R., & Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, *22*(12), 2472–2479.

Zhao, Y., Chan, Z., Gao, J., Xing, L., Cao, M., Yu, C., … Gong, Y. (2016). ABA receptor PYL9 promotes drought resistance and leaf senescence. *Proceedings of the National Academy of Sciences*, *113*(7), 1949–1954.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.