nature plants

Article

https://doi.org/10.1038/s41477-024-01729-5

Convergent evolution of desiccation tolerance in grasses

Received: 30 November 2023

Accepted: 21 May 2024

Published online: 21 June 2024

Check for updates

Rose A. Marks[1,2,3] ✉, Llewelyn Van Der Pas[4], Jenny Schuster[1,2], Ian S. Gilman[1,2] & Robert VanBuren[1,2,5,6] ✉

Desiccation tolerance has evolved repeatedly in plants as an adaptation to survive extreme environments. Plants use similar biophysical and cellular mechanisms to survive life without water, but convergence at the molecular, gene and regulatory levels remains to be tested. Here we explore the evolutionary mechanisms underlying the recurrent evolution of desiccation tolerance across grasses. We observed substantial convergence in gene duplication and expression patterns associated with desiccation. Syntenic genes of shared origin are activated across species, indicative of parallel evolution. In other cases, similar metabolic pathways are induced but using different gene sets, pointing towards phenotypic convergence. Species-specific mechanisms supplement these shared core mechanisms, underlining the complexity and diversity of evolutionary adaptations to drought. Our findings provide insight into the evolutionary processes driving desiccation tolerance and highlight the roles of parallel and convergent evolution in response to environmental challenges.

Anhydrobiosis, or life without water, is rare but widely distributed across life, spanning microbial, animal and plant lineages. Plants that can tolerate desiccation in their vegetative tissues are known as resurrection plants due to their dramatic ability to revive from an extremely dry state (water potential of <−100 MPa or relative water content (RWC) of <10%)[1]. Desiccation tolerance likely arose in plants during the Ordivocian period and is thought to have played a critical role in facilitating the transition from aquatic to terrestrial environments by early land plants[2]. These ancestral mechanisms of anhydrobiosis were retained in many non-seed plants (for example, mosses, liverworts, ferns and fern allies), and there is a high frequency of vegetative desiccation tolerance among extant bryophytes and pteridophytes[3]. By contrast, vegetative desiccation tolerance was lost, or suppressed, in the common ancestor of seed plants, presumably in a trade-off for other systems of drought avoidance and escape, such as annual life histories, water transport and retention mechanisms including stomata, vasculature and roots[4]. Desiccation tolerance then re-evolved convergently in a subset of vascular plants, likely through the rewiring of ancestral anhydrobiosis pathways maintained in seeds, spores and pollen[5–7]. The retention and re-evolution of desiccation tolerance seems to have been driven by a combination of selective pressures in habitats with extreme water limitation, seasonal drought and sporadic water availability[8]. Consequently, desiccation tolerance is more common in some lineages than others, but diverse species of resurrection plants can often be found co-occurring in tightly intertwined communities on rocky outcroppings in arid tropical and subtropical regions across the world[3,9].

Despite more than 500 Myr of evolution and divergence across extant resurrection plants, multiple biochemical and physiological mechanisms of desiccation tolerance are shared across distantly related species. For example, all surveyed resurrection plants accumulate small non-reducing sugars and other osmoprotectants to vitrify the cytoplasm and safeguard macromolecules during drying[10]. Dramatic shifts in carbohydrate and lipid metabolism as well as the protection (or in some cases degradation) of photosynthetic apparatuses are also observed in all resurrection plants during drying[11–14]. All surveyed desiccation-tolerant plants leverage robust anti-oxidant scavenging

[1]Department of Horticulture, Michigan State University, East Lansing, MI, USA. [2]Plant Resilience Institute, Michigan State University, East Lansing, MI, USA. [3]Department of Plant Biology, University of Illinois, Urbana, IL, USA. [4]Department of Molecular and Cell Biology, University of Cape Town, Cape Town, South Africa. [5]Department of Plant Biology, Michigan State University, East Lansing, MI, USA. [6]Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, MI, USA. ✉e-mail: marksros@msu.edu; vanbur31@msu.edu

https://doi.org/10.1038/s41477-024-01729-5

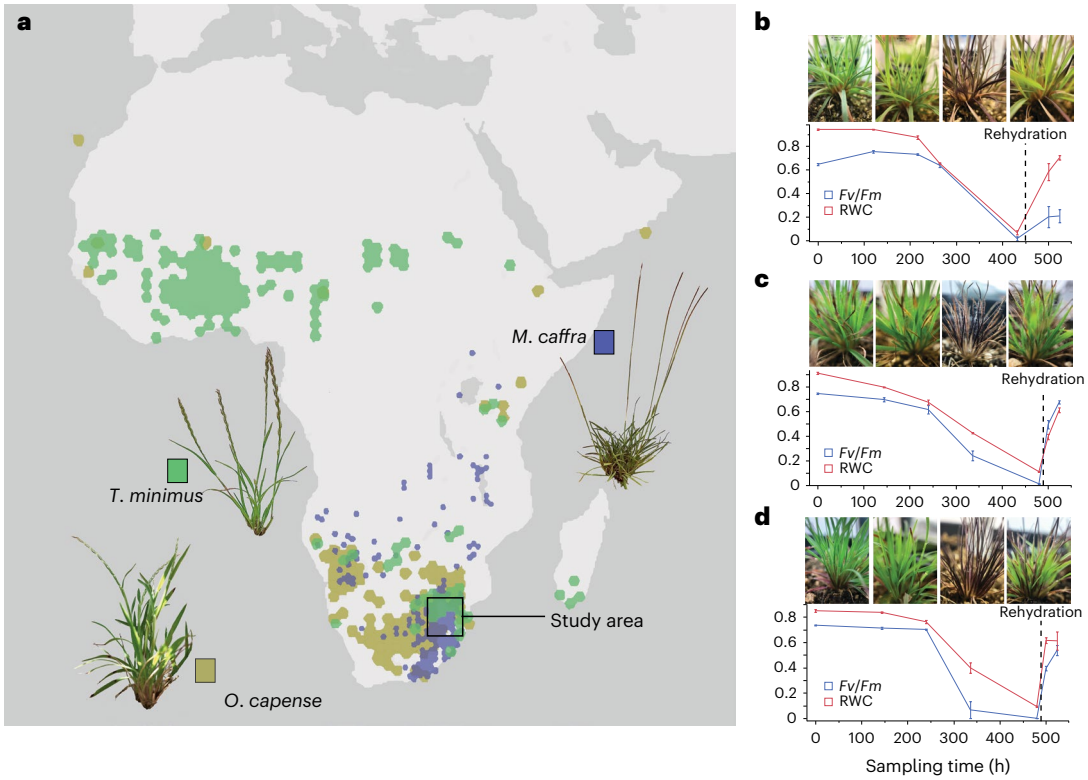


Fig. 1 | Overview of species distribution and experimental design to test for convergent evolution in grasses. a, Estimated distribution of the three desiccation-tolerant grasses: *M. caffra*, *O. capense* and *T. minimus*. Distribution data were taken from ref. 86. Collections for the current study were made in Mpumalanga and Limpopo provinces of South Africa. **b**–**d**, RWC and F_v/F_m of plants during dehydration and rehydration time courses for *M. caffra* (**b**), *O. capense* (**c**) and *T. minimus* (**d**). Three biological replicates were sampled at each time point for each species. Error bars represent standard error of the mean.

systems, mobilize numerous intrinsically disordered and protective proteins and have specialized cell wall properties that maximize flexibility and mitigate the mechanical strain of shrinkage[3,14,15]. These broad features of anhydrobiosis are largely shared across organisms and tissues, but the specific metabolic pathways, regulatory networks and activated genes are notably complex and variable among species[3,10,16] and tissues[17].

The recurrent evolution of desiccation tolerance offers an exciting opportunity to understand how complex traits evolve independently across both broad and narrow phylogenetic distances. The evolution of complex traits can occur via multiple pathways[18,19], and it is often assumed that when closely related taxa evolve the same trait independently, they do so by leveraging the same genetic pathways (parallelism) due to internal constraints within that lineage[20]. By contrast, when distantly related taxa evolve the same trait independently, they are expected to leverage divergent pathways and genes (convergence), due to contrasting genetic starting points[19,21]. However, these patterns are not always observed in nature, and contradictory examples exist, where distantly related taxa show independent but identical mutations and closely related taxa do not[21]. The recurrent evolution of desiccation tolerance at multiple phylogenetic scales provides an ideal system to untangle the mechanisms of convergent and parallel evolution. An important first step towards decoding the evolutionary pathways to desiccation tolerance is characterizing the extent of shared genetic adaptations, overlapping pathways and lineage-specific processes across resurrection plants.

Desiccation tolerance has received growing research attention in recent years, and several resurrection plants have emerged as models for understanding this remarkable trait[22]. Desiccation tolerance is found in at least ten angiosperm families and is most common in Poaceae, where it evolved independently at least six times across three subfamilies and is found in dozens of grass species[3]. Thus, the grasses are an excellent system to test whether the same pathways, regulatory modules and mechanisms were recruited during the recurrent evolution of desiccation tolerance. Most genomic studies of resurrection plants have investigated only a single species in isolation[17,23–27] or tolerant and sensitive taxon comparison[7,28,29], but none have identified core responses shared among independent lineages of resurrection plants. In this Article, we quantify the extent of shared mechanisms of anhydrobiosis across resurrection grasses and investigate the roles of parallel mutation and convergent pathway adaptation in the evolution of desiccation tolerance. We present highly contiguous genome assemblies of three resurrection grasses native to Sub-Saharan Africa coupled with comprehensive gene expression datasets and supporting physiological data. We leveraged comparative genomic and transcriptomic approaches to investigate the evolution of desiccation tolerance in these three species. We also extend these analyses to other desiccation-tolerant and desiccation-sensitive grasses to describe a core signature that defines desiccation tolerance.

Results

Comparative genomics of desiccation-tolerant grasses

We searched for signatures of convergent evolution across three grasses in two Chloridoideae subtribes representing at least two independent origins of desiccation tolerance: *Microchloa caffra* Nees in subtribe Eleusininae and *Oropetium capense* Stapf and *Tripogon minimus* Steud. in the Tripogoninae subtribe. These three species have overlapping distributions and tend to co-occur in shallow soils on rocky outcroppings, locally known as ruwari, across Sub-Saharan Africa (Fig. 1a). *M. caffra*, commonly known as pincushion grass, is distributed from Uganda to South Africa and is the largest of the three species. *O. capense* is smaller and grows as densely packed tufts on exposed rock surfaces. *T. minimus* is a small but loosely tufted grass that occurs in shallow soils in both western and southern Africa (Fig. 1a). *M. caffra* plants were

https://doi.org/10.1038/s41477-024-01729-5

Table 1 | Assembly stats of the three resurrection grasses

Assembly stats	*O. capense*	*T. minimus*	*M. caffra*
Ploidy	Diploid	Diploid	Hexaploid
Total assembly size (Mb)	237	223	968
Number of contigs	14	57	118
Contig N50	27,924,228	19,548,099	16,141,787
Contig L50	4	5	22
Number of genes	28,826	26,527	85,245
Complete BUSCO (%)	97.1	95.3	96.2
LTR elements (% of genome):	27.6	22.4	27.6
Ty1/Copia (%)	4.0	5.2	5.5
Gypsy/DIRS1 (%)	21.4	13.2	15.1
DNA transposons (%)	12.2	15.9	27.6
Total repeats (%)	41.7	39.4	56.1

Ty1/Copia and Gypsy/DIRS1 are groups of LTR retrotransposons that replicate via an RNA intermediate, with Ty1/Copia found in organisms like yeast and fruit flies, and Gypsy/DIRS1 identified in species including fruit flies and slime molds.

collected from Buffelskloof Private Nature Reserve in Mpumalanga, and *O. capense* and *T. minimus* were collected from Swebe Swebe Private Wildlife Reserve in Limpopo, South Africa.

We generated reference genome assemblies for each of the three grasses using PacBio high-fidelity (HiFi) data. *O. capense* and *T. minimus* are diploid with haploid genome sizes of ~195 Mb based on flow cytometry, and *M. caffra* is hexaploid with a 1.25 Gb haploid genome. Sequencing reads were assembled using hifiasm (v 0.18)[30], producing near-complete reference assemblies for *O. capense* and *T. minimus* and a highly contiguous draft assembly of *M. caffra* (Table 1). Six and nine of the ten chromosomes were assembled telomere to telomere for *T. minimus* and *O. capense*, respectively, and the remaining chromosomes were split into two contigs. The *M. caffra* genome assembly was more fragmented, with 118 contigs spanning 968 Mb and a contig N50 of 16 Mb. The monoploid genome size of *M. caffra* is 322 Mb, which is roughly 30% larger than *O. capense* and *T. minimus* (237 and 223 Mb, respectively), and this expansion was driven largely by DNA transposons. All three species have a similar proportion of long terminal repeat retrotransposons (22–27%), but 27% of the *M. caffra* genome is composed of DNA transposons compared to 12% in *O. capense* and 16% in *T. minimus* (Table 1). Despite this expansion of transposons in *M. caffra*, the three Chloridoid grasses have very compact genomes compared to most grasses[31]. We used the MAKER-P pipeline (v 2.31.10) to annotate these three genome assemblies, with RNA sequencing (RNA-seq) data and protein homology as evidence. The *O. capense* and *T. minimus* genome assemblies have 28,826 and 26,527 gene models, respectively, which is comparable to the well-annotated model resurrection plant *Oropetium thomaeum* (L.f.) Trin. (28,835)[24,26]. The *M. caffra* genome assembly has 85,245 gene models, which matches the expectations for a hexaploid genome (Table 1). We assessed annotation quality using the land plant (Embryophyta) dataset of Benchmarking Universal Single-Copy Orthologs (BUSCO) and found between 95.3% and 97.1% complete proteins across the three grasses, suggesting the genome assemblies are largely complete and well annotated (Table 1).

We leveraged comparative genomic approaches to identify evolutionary signatures associated with desiccation tolerance and enable cross-species comparisons of gene expression data. The three grass genomes are largely collinear with *O. thomaeum* and have considerable conserved gene content despite some notable structural rearrangements. Seven pairs of *O. thomaeum* and *O. capense* chromosomes have near-perfect synteny, with chromosomes 8 and 9 showing a few large-scale inversions and a telomeric translocation on chromosome 2 (Extended Data Fig. 1). *T. minimus* has similar macrosynteny with *O. thomaeum* but has no rearrangements in chromosome 8. Synteny between *M. caffra* and *O. thomaeum* is more fragmented because of phylogenetic divergence, and each *O. thomaeum* region has between two and four homeologous regions in *M. caffra* (Extended Data Fig. 2). We calculated the synonymous substitution rates (Ks) between homeologous gene pairs within *M. caffra* to date the polyploid event(s). We observed a single Ks peak of 0.13 across all homeologous gene pair combinations, suggesting the autohexaploidy event occurred ~4 Ma (million years ago) from rapidly successive polyploidy events (Extended Data Fig. 2d). Using MCScan (v 1.1) with *O. thomaeum* as an anchor, we identified 18,428 syntenic orthologs (syntelogs) shared among the three grasses, as well as previously published tolerant grasses *Eragrostis nindensis* Ficalho & Hiern[28,32]. These syntelogs were used to identify patterns of gene duplication associated with desiccation tolerance across grasses and as anchor points to compare expression of conserved genes across species.

To test for convergent evolution, we characterized patterns of expansion and duplication in gene families with important roles in desiccation tolerance. The genomes of all sequenced resurrection plants have large tandem arrays of early light-induced proteins (ELIPs)[33], and we observed this same pattern across the desiccation-tolerant grasses investigated here. *O. capense*, *T. minimus* and *M. caffra* all have massive tandem arrays of 39, 31 and 58 ELIPs, respectively, compared to an average of 4 in the genomes of desiccation-sensitive grasses[33]. This expansion of ELIPs is similar to other chlorophyll-retaining (homiochlorophylus) resurrection plants and is generally higher than chlorophyll-degrading (poikiochlorophyllus) species. ELIPs are universally highly expressed in the diploid resurrection grasses *O. capense* and *T. minimus* during drying, desiccation and early rehydration, but only a subset of the ELIPs in the *M. caffra* tandem arrays have desiccation-induced expression (Extended Data Fig. 3). We used CAFE (v 5.1)[34] to test for changes in the dynamics of ELIP copy number evolution across land plants. We found notable increases in the rate of ELIP expansion in all desiccation-tolerant lineages of plants (Fig. 2c). Within the grass family, ELIP expansion occurred independently in subtribes Eleusininae, Sporobolinae, Eragrostidinae and Tripogonae, but *Oropetium* and *Tripogon* share a single origin of desiccation tolerance (Fig. 2c). Other gene families with well-characterized roles in desiccation tolerance such as late embryogenesis abundant proteins and heat shock proteins show no expansion in resurrection plants based on OrthoFinder (v 2.4.1) and/or CAFE (Extended Data Figs. 4 and 5).

We identified the origin of duplicated ELIPs to test whether the same or different ancestral copies were duplicated in each lineage using a synteny-based approach. Tandem duplication of ELIPs within the Tripogoninae occurred on chromosome 8, and the Eleusininae and Eragrostidinae subtribe species have no syntenic ELIPs in this region, despite otherwise high collinearity (Fig. 2a). Most ELIPs in Eleusininae and Eragrostidinae species are found in large tandem arrays on chromosome 7, compared to 4–5 ELIPS within Tripogoninae (Fig. 2b). Together, phylogenetic and comparative genomics analyses suggest these grass lineages duplicated ELIPs independently, supporting the convergent evolution of desiccation tolerance within Chloridoideae.

Identifying overlapping signatures of desiccation tolerance

We collected dehydration and rehydration time courses of *O. capense*, *T. minimus* and *M. caffra* plants under similar conditions in a climate-controlled growth chamber. Plants reached desiccation after ~17–20 days of natural drying, with RWC < 10% and photosystem II efficiency, represented as F_v/F_m (e.g. the ratio of variable to maximum fluorescence approaching 0.0) (Fig. 1b–d). RWC and F_v/F_m recovered within 12 h of rehydration in *O. capense* and *T. minimus*, but F_v/F_m took longer to recover in *M. caffra* (Fig. 1b). We collected gene expression data (RNA-seq) at six comparable time points of drying and recovery

https://doi.org/10.1038/s41477-024-01729-5

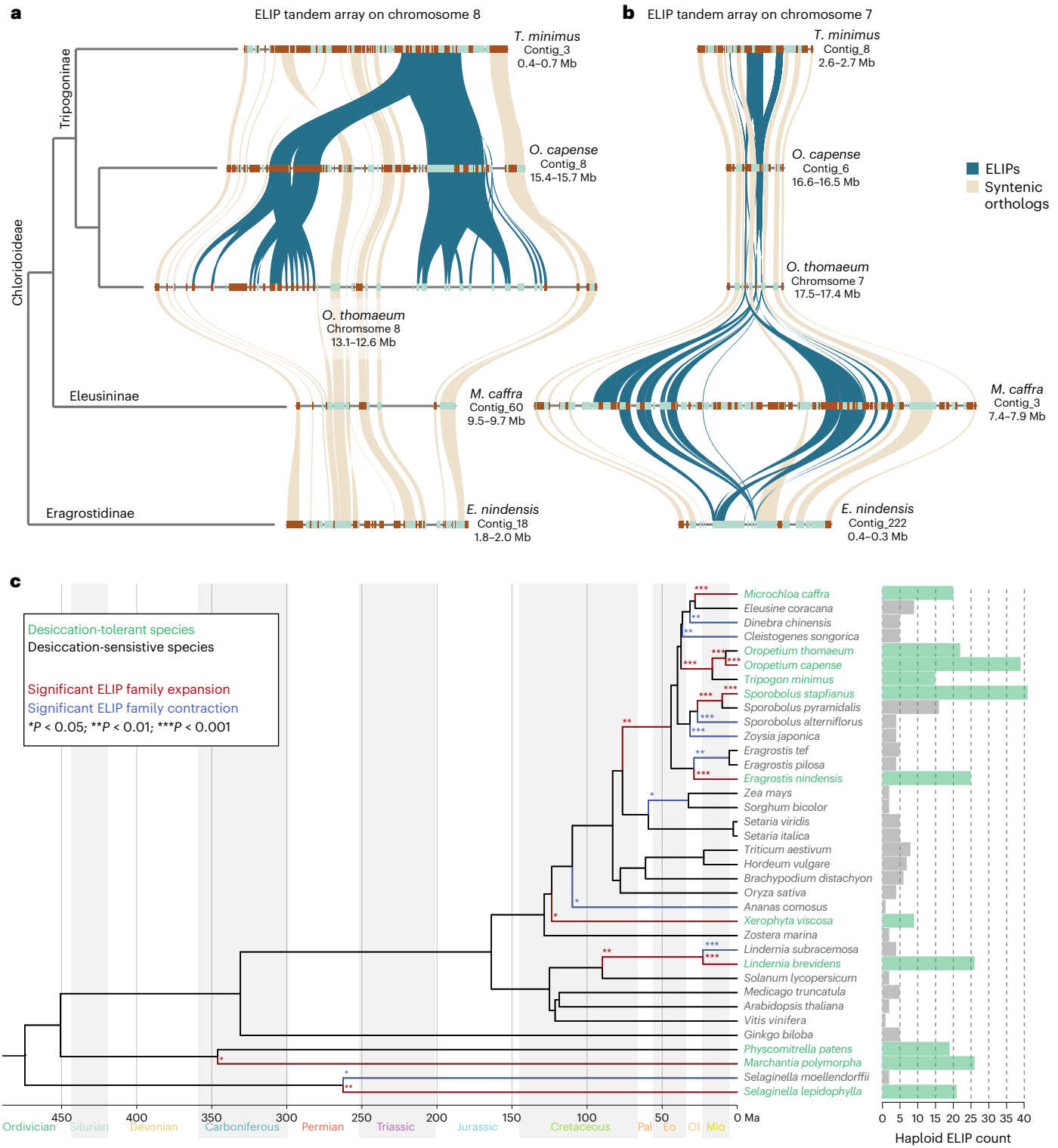


Fig. 2 | Independent tandem gene duplication of ELIPs in different resurrection grass lineages. a,**b**, Microsyntenic regions of the chromosome 8 (**a**) and chromosome 7 (**b**) ELIP tandem arrays are shown for resurrection grasses in the Tripogoninae (*T. minimus*, *O. thomaeum* and *O. capense*), Eleusininae (*M. caffra*) and Eragrostidinae (*E. nindensis*) subtribes of Chloridoideae. Syntenic orthologs between the species are shown in beige, and the ELIPs are highlighted in blue. Only a single syntenic region for autopolyploids *M. caffra* (hexaploid) and *E. nindensis* (tetraploid) is shown for simplicity, but each of the other haplotypes contain the same gene content in these regions. **c**, Evolutionary dynamics showing significant changes in the rates of gene family expansion (red) and contraction (blue) of ELIPs inferred by CAFE. The haploid normalized number of ELIPs are plotted for desiccation-tolerant (green) and desiccation-sensitive (grey) species. Pal; Paleocene, Eo; Eocene, Ol; Oligocene, Mio; Miocene.

for each of the three species. We quantified RNA abundance and gene expression patterns across the dehydration–rehydration time course in each species individually. RNA-seq reads were pseudo-aligned to the respective genomes using Salmon (v 1.9.0)[35], and normalized counts were used for all downstream analyses. In general, gene expression profiles were tightly associated with the hydration status of the plants.

https://doi.org/10.1038/s41477-024-01729-5

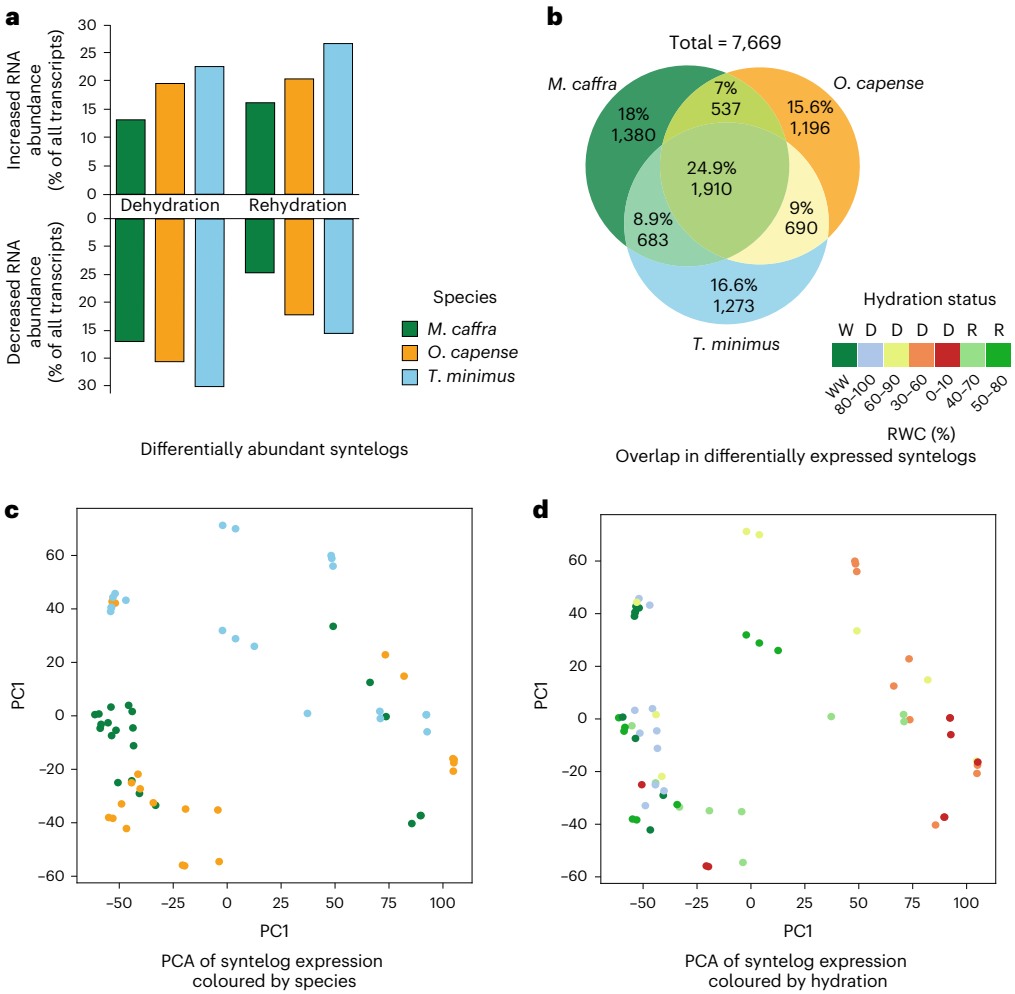


Fig. 3 | Overlapping expression dynamics of conserved genes across species. **a**, Barplot showing the percentage of DEGs in each species for up- and down-regulated genes in dehydration and rehydration conditions. **b**, Venn diagram showing the number of syntenic orthologs that increased in abundance during dehydration and overlap across species. The percentage and number of genes in each set are shown. Hydration stauts is abbreviated as follows: W or WW is well watered, D is drying, R is rehydration. **c**,**d**, PCA of *z*-score transformed expression values for conserved syntelogs across all three species. Samples are coloured by hydration status in **c** and by species identity in **d**.

Correlation matrices and principal component analysis (PCA) show tight clustering of samples by hydration status, with hydrated, desiccated and rehydrated samples forming distinct clusters for each species (Extended Data Fig. 6).

Using RWC as a covariate, we identified genes that were up- and down-regulated during dehydration and rehydration processes. Both dehydration and rehydration induced substantial changes in gene expression in all three desiccation-tolerant grasses, with 35–52% of genes showing differential abundance during dehydration and 23–47% during rehydration (Fig. 3a and Extended Data Fig. 7). *M. caffra* had more differentially expressed genes (DEGs) (Extended Data Fig. 7), given its hexaploidy, but a lower proportion of DEGs compared to the other two grasses (Fig. 3a). Broadly, desiccation and rehydration had inverse expression profiles, and most genes that increased in abundance during dehydration dissipated during rehydration and vice versa (Extended Data Fig. 7).

To enable comparisons across species, we leveraged the 18,428 conserved syntelogs and searched for overlapping patterns in the expression of these shared genes. There was considerable overlap in gene expression across the three focal resurrection grasses, with ~18–24% of all differentially expressed syntelogs showing similar expression across species (Fig. 3b and Extended Data Fig. 8). The proportions of DEGs shared across the three resurrection grasses for both up- and down-regulated genes was considerably more than observed in previous studies or expected due to chance. To differentiate between desiccation tolerance mechanisms and more general drought responses, we identified the extent of shared syntelog expression between these resurrection grasses and the desiccation-sensitive species *Eragrostis tef*, which was sampled along a similar dehydration time course in a previous study[28]. There was considerable overlap in syntelog expression between the resurrection grasses and *E. tef* (Extended Data Fig. 9), reflecting deeply conserved mechanisms of drought tolerance in grasses. We also detected a large set of genes that were expressed exclusively in the resurrection grasses, which likely play desiccation-specific roles to survive anhydrobiosis. Species-specific expression patterns are also evident, particularly for *E. tef*.

Dimensionality reduction and co-expression analyses also point towards parallel mechanisms of desiccation tolerance in resurrection grasses. Samples clustered primarily by hydration status and secondarily by species in PCA (Fig. 3c,d and Extended Data Fig. 10). We defined co-expression modules for each species and screened for shared network level responses within co-expressed genes. High-confidence modules were defined for each species, and we grouped these into three broad classes based on the expression pattern of each module: (1) elevated expression in hydrated conditions, (2) elevated expression during dehydration and (3) elevated expression during rehydration (Fig. 4d). We identified substantial overlap in gene module conservation with

https://doi.org/10.1038/s41477-024-01729-5

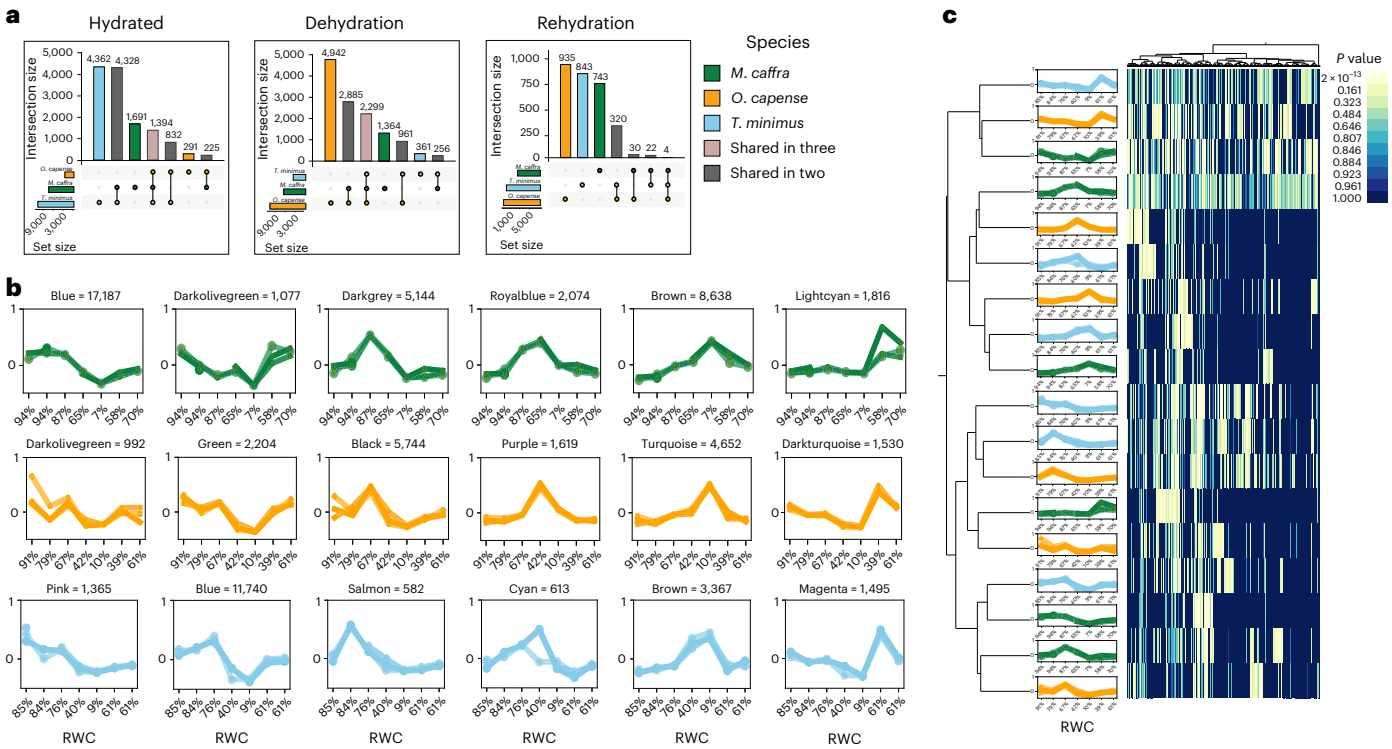


Fig. 4 | Comparative co-expression network dynamics across resurrection grasses. a, UpSet plots showing the number of shared and unique syntenic orthologs among co-expression modules characterized by elevated expression in hydrated, dehydrated or rehydrated conditions, with interaction size representing the number of syntenic orthologs in each categoriy. **b**, Co-expression modules for each species. The *x* axis shows the approximate RWC of samples at each time point. The module name and total number of genes in the module are listed above. **c**, Hierarchical clustering of enriched GO terms for each co-expression module. Secondary clustering performed on modules shows that modules are organized by expression profile rather than species.

elevated expression during dehydration but less overlap in modules with high expression during rehydration (Fig. 4a). We then identified enriched gene ontology (GO) terms for each co-expression module and performed hierarchical clustering on the enrichment *P* values of GO terms. Modules clustered by their expression profile rather than species identity, suggesting that hydration status is more predictive of gene expression than species identity (Fig. 4c) and pointing towards a shared signature of desiccation tolerance in resurrection grasses.

Shared signatures of desiccation tolerance

Our analyses of syntelog expression tested for ancestral conservation and parallelism, but it is also possible that different lineages of resurrection plants may utilize similar metabolic strategies for achieving desiccation tolerance but through divergent genes and pathways. To investigate this possibility, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) to assign each gene to a predicted enzymatic function and metabolic pathways and compared the overlap in these functional predictions across species. We detected substantially higher overlap in KEGG terms across species (~30–40%) compared to differentially expressed syntelogs (only 18–24%) (Fig. 5b and Extended Data Fig. 10). The increased similarity at a metabolic level suggests that while these species do not always leverage parallel gene copies, they induce similar metabolic mechanisms to survive anhydrobiosis, providing evidence of convergence across species.

We further investigated the functional roles of shared gene expression via GO enrichment and KEGG analyses. We found that many hallmarks of desiccation tolerance were shared across the three resurrection grasses, including the controlled down-regulation of photosynthesis and rapid induction of protective mechanisms. Enriched GO terms during dehydration were related primarily to signalling and stress responses (for example, stress perception and reactive oxygen species scavenging activities), developmental regulation (for example, photoperiodism and germination processes), cellular reorganization (for example, lipid droplet formation, vesicle fusion and endocytosis) and modifications to transcription and translation (for example, RNA modifications, splicing and protein degradation). By contrast, enriched GO terms during rehydration are related to photosynthesis and metabolism (for example, fructose biosynthesis, cellulose biosynthesis and light harvesting), pigment metabolism (for example, chlorophyll biosynthesis and anthocyanin metabolism), protein modification (for example, protein phosphorylation and proteolysis) and some residual stress response (for example, response to cold and non-photochemical quenching) (Fig. 5a). Hierarchical clustering of enriched GO terms also highlighted the inverse relationship between dehydration and rehydration process (Fig. 5a and Extended Data Fig. 10).

To differentiate between desiccation tolerance mechanisms and more typical drought tolerance responses, we compared the enriched GO terms for differentially expressed syntelogs uniquely induced in the resurrection grasses versus those shared with desiccation-sensitive *E. tef* (Extended Data Fig. 10). Many of the classic stress response terms were shared across all species, reflecting deeply conserved responses to water deprivation. For example, all species showed metabolic arrest during drying with a particular emphasis on photosynthetic shutdown. All species showed an increase in classic stress response terms such as response to heat, response to water deprivation, response to hydrogen peroxide and sucrose metabolic process. These processes represent core mechanisms of water deficit tolerance that likely form the foundation of desiccation tolerance. Building on this foundation, resurrection grasses appear to activate additional processes that enable more extreme resilience. For example, the resurrection grasses showed unique activation of nucleic acid processes including messenger RNA export, regulation of chromosome condensation and

https://doi.org/10.1038/s41477-024-01729-5

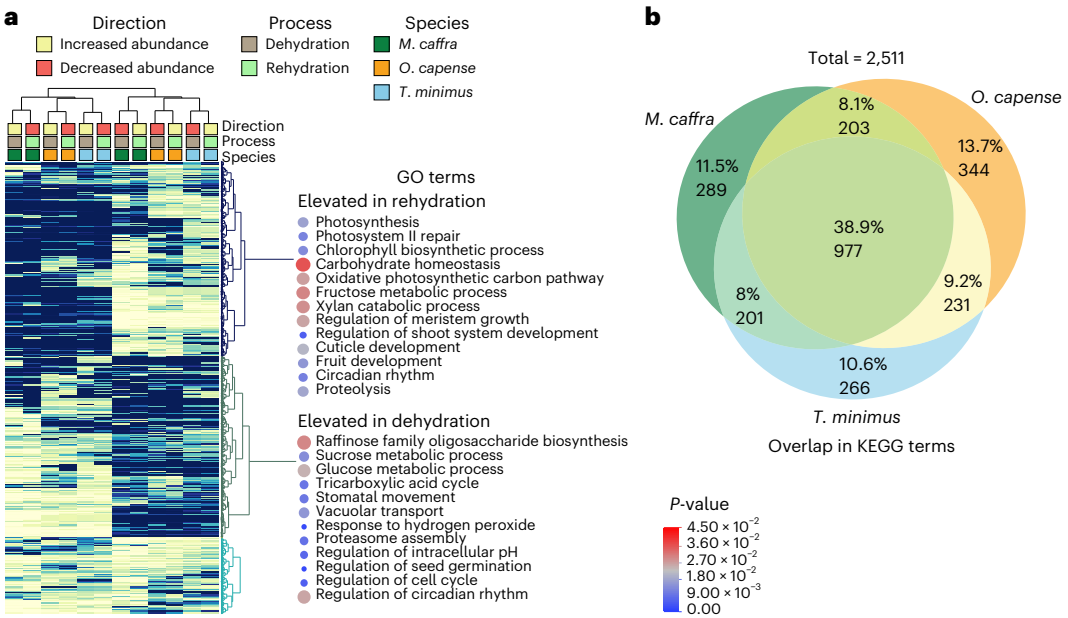


Fig. 5 | Overlapping gene functions during desiccation and rehydration in resurrection grasses. a, Hierarchical clustering of *P* values for enriched GO terms for each species and condition. Selected GO terms are highlighted for genes that increased in abundance during dehydration (decreased in abundance during rehydration) and genes that increased in abundance during rehydration (decreased in abundance during dehydration). Points are coloured by the average enrichment *P* value across all species and sized by the number of genes assigned to that GO term. **b**, Venn diagram showing the number of overlapping KEGG terms that increased in abundance during dehydration. The percentage and number of KEGG terms in each set are shown.

mRNA transcription by RNA polymerase II suggesting greater overall regulation of transcription and translation. Several terms associated with the circadian rhythm and hormonal signalling were also uniquely up-regulated in the resurrection grasses, indicating a central role of circadian clock processes in preparing for desiccation. The resurrection grasses showed a unique down-regulation of tissue and cellular developmental processes, implying a tightly regulated cessation of metabolism at later stages of drying. Taken together, this suggests that resurrection grasses build on a shared foundation of drought tolerance to achieve desiccation tolerance via a highly organized shift in cellular processes.

KEGG annotations revealed characteristic desiccation tolerance mechanisms shared across resurrection grasses. Metabolic pathways associated with photosynthetic energy metabolism were down-regulated in all three grasses. Interestingly, we observed an increase of malate to pyruvate catalysis with concomitant regeneration of reduced nicotinamide adenine dinucleotide phosphate, which could be related to reduced nicotinamide adenine dinucleotide phosphate's redox potential for antioxidant enzymes such as glutathione reductase. We also detected noticeable changes to carbohydrate and energy metabolism, including a shift towards the production of raffinose and stachyose under dehydrating conditions as seen in other resurrection plants (reviewed in ref. 10). Central carbohydrate metabolism appeared operational, suggesting that at low water contents, other solvents, such as natural deep eutectic solvents within the mitochondria, may facilitate glycolysis, the tricarboxylic acid cycle and electron transport[36]. Amino acid metabolism favoured degradative pathways with an increase in endoplasmic reticulum-mediated ubiquitination and proteolysis, which could be serving a glucogenic role by converting amino acids to pyruvate or by generating an available amino acid pool for the rapid assembly of thermo- and osmoprotective proteins. While amino acid metabolic pathways were generally down-regulated, a few important pathways including glutathione metabolism were up-regulated. Reduced glutathione exerts numerous effects in the cell[37] from interaction with hormones to acting as direct reactive oxygen species quencher, and maintaining a steady supply of reduced glutathione is a feature all three resurrection grasses share. Lipid metabolism showed a shift towards the production of glycerolipids and glycerophospholipids, which likely supports triacylglycerol phosphatidylcholine production. The accumulation of phosphatidylcholine may further lead to phosphatidic acid synthesis, which has been implicated in numerous plant processes from signalling to storage[38–40]. Pathways involved in the transcription and translation of genetic information also showed an up-regulation of transcription factors, RNA polymerase and spliceosome activity, suggesting that active transcription and RNA processing are still occurring. However, we observed substantial down-regulation of ribosome activity, suggesting that RNA is either differentially translated or delayed. Upon rehydration, up-regulated processes involved in overall resumption of normal metabolic activity such as several photosystem I and II proteins, light harvesting complexes, starch synthesis and cell wall remodelling such as xyloglucan O-acetyltransferase, expansin and pectinesterase were observed.

Desiccation tolerance mechanisms are conserved in grasses

Desiccation tolerance evolved independently in at least four subtribes of Chloridiodeae (Eleusininae, Eragrostidinae, Sporobolinae and Tripogoninae; Fig. 2c), and we integrated comparable desiccation and rehydration expression datasets from additional species to test for patterns of convergence across grasses more broadly. Building on our detailed comparisons across the three study species, we expanded our analysis to include publicly available RNA-seq samples from desiccation-tolerant *O. thomaeum*[41] and *E. nindensis*[28], leveraging syntelogs for cross-species comparisons. Similar to the three species comparisons described above, dimensionality reduction across the five species generally separated samples by hydration status along principal component 1 and principal component 2 (Extended Data Fig. 10). While PCA provided some degree of separation, the residual heterogeneity, experimental differences, noise or species level differences in the datasets might have obscured underlying conserved biology. To account for this, we used a topological data analysis approach to discern the underlying structure of the expression datasets. We utilized the Mapper algorithm, which condenses the dataset into a

https://doi.org/10.1038/s41477-024-01729-5

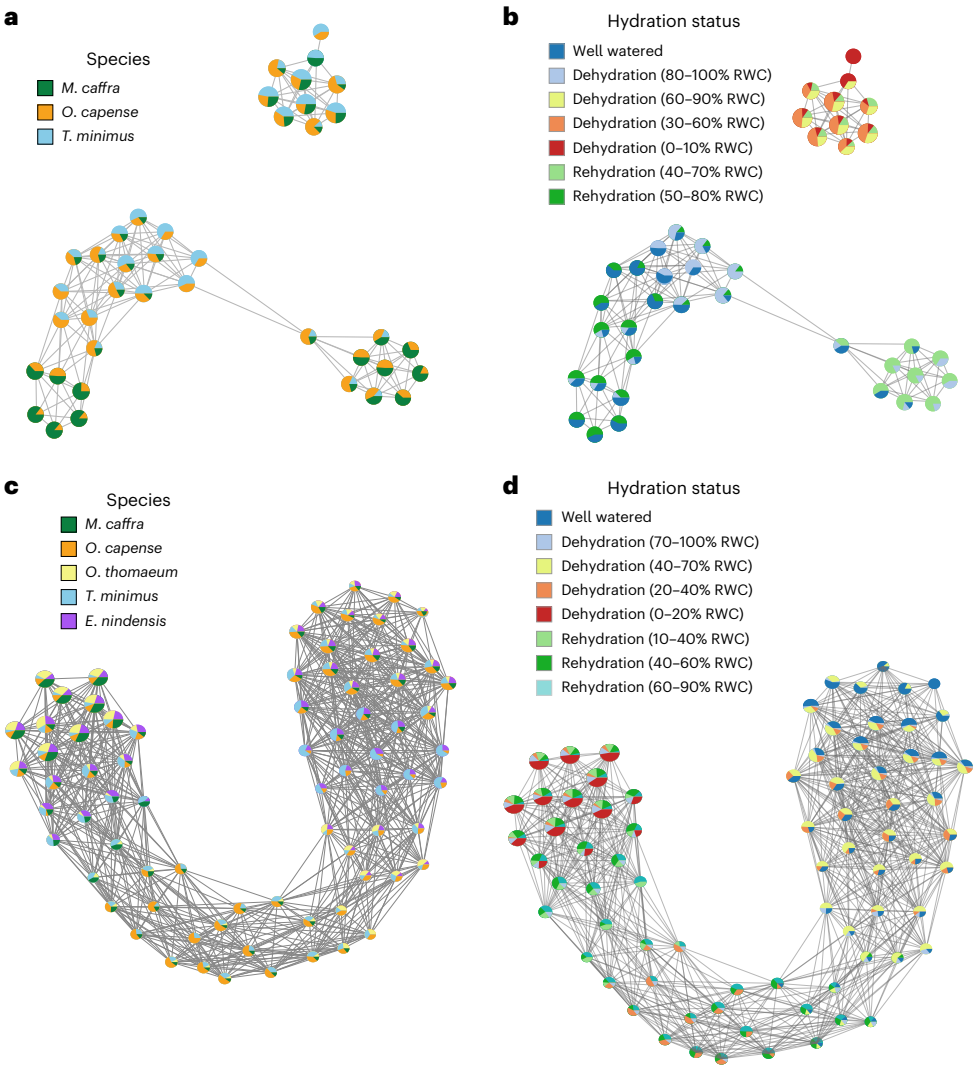


Fig. 6 | Mapper graphs showing the topological shape of desiccation-induced gene expression across species. Nodes within the graph represent clusters of RNA-seq samples that are akin to one another, with the node colour indicating the identity of the samples contained within. Edges, or the connections between nodes, delineate shared samples across intersecting clusters. **a**,**b**, Mapper graphs for the three-species comparisons. **c**,**d**, Mapper graphs for the five-species comparisons. Nodes within the graph are coloured by species (**a** and **c**) or hydration status (**b** and **d**).

scalable, navigable representation. The Mapper algorithm is particularly well suited for genome scale analyses, as the underlying datasets are often characterized by high dimensionality and sparsity[42]. For our gene expression data, we constructed Mapper graphs using a 'stress lens' with the well-watered condition as a reference point. This model represents the baseline for gene expression, and we quantified the residuals or deviation of each sample from the baseline, which represents the degree of water stress or recovery.

The resultant Mapper graph illustrates a clear topological shape that delineates desiccation processes across grasses (Fig. 6). Each node on the graph represents a cluster of similar RNA-seq samples, and the node colour depicts the identity of samples within that cluster. Connections between nodes signify shared samples among the intersecting clusters.

These graphs reveal a compelling topological depiction of the gene expression variations induced by water stress across different species. Similar topology was observed for targeted comparison of the three focal species (Fig. 6a,b) and for the larger dataset including *E. nindensis* and *O. thomaeum* (Fig. 6c,d). In both instances, clear delineation between samples of different hydration statuses are evident, while the species are intermixed. We then added the desiccation-sensitive sister species *E. tef* in a final analysis (Extended Data Fig. 10), which revealed a similar topology across all species with notable gaps in *E. tef*. Broadly, this supports our finding that similar ancestral mechanisms are being recruited for foundational drought tolerance mechanisms, which are enhanced in resurrection plants via the independent recruitment of specific desiccation tolerance pathways.

Species-specific mechanisms underlying desiccation tolerance

Despite the considerable overlap in gene expression across all three focal species, species-specific processes were also evident. In *M. caffra*, unique antioxidant responses were induced including glutathione biosynthetic processes, glutamate decarboxylation, and L-ascorbic acid biosynthesis. Other processes enriched uniquely in *M. caffra* included seed-related terms such as seed oil body biogenesis and seed maturation. Several GO terms associated with phytohormones were also uniquely induced in *M. caffra*, including overall ethylene responses such as S-adenosylmethionine metabolic process, ethylene-activated signalling pathway and response to 1-aminocyclopropane-1-carboxylic acid, suggesting that hormonal regulation might be exerting an effect on the partial breakdown of thylakoids and photosynthetic machinery as seen in classical senescence[43]. *M. caffra* also showed unique lipid, sphingolipid, riboflavin and selenocompound metabolism, as well

as sesquiterpenoid and terpenoid biosynthesis. *M. caffra* was the only species to have multiple pathways involved in signal transduction up-regulated including phospholipase D and calcium signalling. Uniquely down-regulated processes in *M. caffra* appear to centre around arresting growth and development, such as phototropism, gravitropism, leaf and root morphogenesis, cell wall biogenesis and regulation of auxin polar transport. There was also a down-regulation of general amino acid–transfer and nitrogen fixation and assimilation.

O. capense had fewer uniquely enriched processes compared to *M. caffra*, but some notable patterns were detected. Uniquely up-regulated processes in *O. capense* centred around histone H3 and H4 acetylation, histone H3–K9 demethylation and histone H2B ubiquitination. The relative degree of acetylation of histones is directly related to the openness of chromatin which impacts transcription in specific drought-responsive genes[44]. Histone demethylation[45] and H2B ubiquitination also regulate drought-responsive genes[46]. Interestingly, terms associated with chloroplast mRNA processing, poly(A)+ mRNA export from the nucleus, ribosome assembly and regulation of translation were also up-regulated in *O. capense*, suggesting continued translation and active processing of mRNA from both the chloroplast and nucleus, presumably through increased transcriptional regulation due to histone modifications. *O. capense* showed unique up-regulation of C5-branched dibasic metabolism and down-regulation within galactose metabolism. Uniquely down-regulated processes in *O. capense* were minimal but included regulation of salicylic acid metabolic process and auxin polar transport. Monoterpenoid biosynthesis was up-regulated for *O. capense*, and fatty acid degradation and steroid hormone biosynthesis was down-regulated. Ferroptosis, an iron-dependent form of programmed cell death, was exclusively down-regulated in *O. capense*.

T. minimus also had fewer species-specific processes compared to *M. caffra*. Uniquely up-regulated processes were centred around response to oxidative stress, peroxisome organization and removal of superoxide radicals. *T. minimus* was the only species to show up-regulation of anthocyanin-containing compound biosynthetic process which is a typical response seen in the homoiochlorophyllous resurrection plants[47]. Similar to the other two species, regulation of auxin-mediated signalling pathways were down-regulated as was cellular response to salicylic acid stimulus. Other processes centred around mismatch repair, chloroplast RNA processing, ribosome biogenesis and plastid transcription. Phosphonate and phosphinate, taurine and hypotaurine, and D-amino acid metabolism were exclusively down-regulated in *T. minimus*, whereas retinol metabolism was up-regulated.

Despite the unique pathways identified in each of the focal species, all three species appear to respond to drought and desiccation stress by leveraging similar mechanisms. The processes uniquely activated in each species are consistently centred around defence mechanisms, the induction of quiescence and reduction of normal growth and metabolism under desiccated conditions. While nuanced variation in metabolism and defence responses are evident, all species show well-known mechanisms of desiccation tolerance. Taken together, the three species appear to share a core set of conserved mechanisms which are then supplemented with convergent species-specific modules.

Discussion

Our data suggest that the repeated evolution of desiccation tolerance within grasses occurred via both parallel adaptations in the same ancestral genes and complementary modifications to analogous pathways. We find evidence that core mechanisms of desiccation tolerance are shared across resurrection grasses and are supplemented with species-specific adaptations. Many of these mechanisms overlap with typical drought responses, and it is likely that the evolution of anhydrobiosis builds on deeply conserved responses to water deficit shared across all plants. Phenotypic and metabolic similarities in anhydrobiosis mechanisms have been observed for decades, but the evolutionary pathways of convergence and parallelism have been obscured by a lack of systems-level data and inconsistencies in experimental procedures[48]. Here we leveraged large-scale genomic and transcriptomic datasets in a replicated and standardized framework to characterize signatures underlying the recurrent evolution of desiccation tolerance within chloridoid grasses.

The adaptations required for desiccation tolerance appear to be sufficiently narrow, such that not any organism can, or will, evolve desiccation tolerance[4]. The physiological changes that occur during the final stages of desiccation are dramatic, and specialized biochemistry and molecular mechanisms are required to protect the cellular macromolecules for life without water. Achieving anhydrobiosis requires tight coordination and orchestration of multiple physiological processes, and there may be only a few trajectories to evolve this trait. However, desiccation tolerance mechanisms overlap considerably with typical drought responses, and many plants possess the basic cellular machinery required to achieve desiccation tolerance[28]. Desiccation tolerance is likely an ancestral adaptation in plants that evolved during terrestrialization, subsequently formed the basis of seed pathways and was later rewired again in vegetative tissues[5,25,41,49]. While previous studies have found surprisingly little overlap in gene expression across desiccation-tolerant plants[50,51], our data suggest that the repeated evolution of specific genetic, biochemical and physiological traits required for anhydrobiosis are highly convergent and build on more broadly conserved water deficit responses.

Convergence is thought to be driven primarily by exposure to external selective pressures that lead to the same emergent phenotype, while parallelism is thought to be impacted more by internal constraints of the system[18] through independent mutations in the same ancestral gene[19,21]. Because anhydrobiosis has evolved independently in both distantly and closely related taxa, it is an ideal system in which to explore the roles of convergent and parallel evolution. Numerous other independently evolved traits such as C4 and Crassulacean acid metabolism photosynthesis are highly complex, making their repeated evolution surprising[52] and difficult to characterize. In the case of C4 photosynthesis, both mutations in the same genes and recruitment of unique pathways occurred in distantly related lineages to enable the emergent C4 phenotype[53]. Desiccation tolerance is similarly complex, involving the synchronized orchestration of numerous pathways and genes, and it is likely that both external pressures (for example, selection in extremely xeric habitats) and internal constraints (lineage-specific predispositions) play a role in the recurrent evolution of desiccation tolerance. Here we detected signatures of both processes and identified far more overlap in gene expression across resurrection grasses than expected by chance or detected in previous studies[29,50]. The observed expansion of ELIP tandem arrays coupled with activation of similar metabolic pathways driven by different gene sets suggests that both parallel and convergent processes contribute to the recurrent evolution of desiccation tolerance in grasses.

Our systems-level analyses add to the growing literature on the mechanisms of desiccation tolerance, and many of the patterns observed here corroborate previous findings[17,25,29,33,50]. We show that desiccation induces a major and reversible shift in gene expression where normal growth and development are halted and numerous protective mechanisms are induced[13,14,54–56]. Gene expression coalesced around a signature desiccation response during drying with all three species initiating parallel processes[57]. The resumption of species-specific processes related to growth and development was evident upon rehydration. The shared pathways of anhydrobiosis observed in these grasses pull on the deeply conserved architecture of drought tolerance coupled with convergent and parallel mutations that provide the necessary protection to survive extreme desiccation. This reflects the relatively narrow set of regulatory networks and pathways in plants that can enable the evolution of desiccation tolerance but also hints as multiple evolutionary paths to anhydrobiosis.

Methods

Field collections, plant growth and maintenance

Plants for the current study were collected from two research sites in South Africa: Buffelskloof Nature Reserve in Mpumalanga (−25.30229° S, 030.50631° E) (*M. caffra*) and Swebe Swebe Private Wildlife Reserve in Limpopo (−23.7949° S, 028.0705° E) (*O. capense* and *T. minimus*). Voucher specimens of each species were collected, pressed and deposited at the National Herbarium of South Africa in Pretoria (specimen numbers PRE1004810-0, PRE1004793-0 and PRE1004794-0). Seeds of each species were also collected and transported to Michigan State University under the US Department of Agriculture permit number 537-22-37-10071 and according to the specifications in a Material Transfer Agreement established between J. M. Farrant, R.V. and R.A.M. Seeds were cold stratified at 4 °C for 2 weeks and then germinated on our standard propagation mix (50:50 sure-mix to redi-earth) and grown in a climate-controlled growth chamber with a 16 h photoperiod and internal temperature of 28 °C/18 °C. One seedling from each species was used for genome sequencing, and the remaining seedlings were used for the desiccation and rehydration time courses experiments. Three plants in each pot were pooled during sampling and treated as a single biological replicate. These plants were grown for another 2 weeks before experimental treatments.

Dehydration treatment and sample collection

After ~8 weeks of growth, plants were subjected to dehydration treatment. Before treatment, any emerging reproductive tissues (for example, panicles) were removed from plants. To initiate dehydration treatment, plants were watered to full soil saturation, and each pot was weighed to ensure consistency across replicates. Water was then withheld until plants became completely desiccated (between 2 and 3 weeks depending on the species). Plants were sampled at targeted hydration states during the process of dehydration, including well watered, partially dehydrated, fully desiccated and rehydrated. We used visual cues to direct our sampling and sampled plants at the first signs of visible leaf curling, partial pigmentation, deep pigmentation and full desiccation and validated the hydration status of tissues by measuring RWC. Plants were then rehydrated through a combination of watering from the base and misting the aerial portions to simulate natural rainfall and sampled 24 h and 48 h post rehydration.

At each time point, we measured the photosynthetic efficiency (F_v/F_m) and RWC and collected tissue for RNA-seq. Briefly, F_v/F_m was measured on dark adapted leaves using an Opti-Sciences OS30p+ chlorophyll fluorometer with the default test parameters. RWC was measured using a set of 10–15 representative leaves from each pot/biological replicate. Leaf mass was weighed immediately after collection (fresh weight), again after 48 h submerged in distilled H_2O in darkness at 4 °C (turgid weight) and finally after 48 h in a 70 °C drying oven (dry weight). RWC was calculated as (fresh weight − dry weight)/(turgid weight − dry weight). Tissue for RNA-seq was collected by gathering all the vegetative tissue from each pot and flash freezing in liquid nitrogen. Tissue samples were stored in a −80 °C freezer before downstream processing.

RNA extraction and sequencing

Frozen leaf tissue was ground to powder by hand in a mortar and pestle with liquid nitrogen. RNA was extracted from each sample using Spectrum Plant Total RNA kit according to the manufacturer's instructions. Total RNA was then cleaned to remove impurities and contaminants using Zymo Clean & Concentrator kit. DNAse treatment was carried out during clean and concentration steps according to the manufacturer's instructions. Sample concentration was assessed on a qubit using the RNA broad range reagent set, purity was assessed with a nanodrop, and RNA integrity was visualized on an agarose gel. RNA-seq libraries were constructed by Novogene following a standard polyA+ enrichment strategy including fragmentation and complementary DNA synthesis. The resulting libraries were sequenced on an Illumina HiSeq 4000 under 150 bp paired end mode.

High-molecular-weight DNA extraction and sequencing

Tissue for whole genome sequencing was collected from a single mature plant of each species. Healthy green tissue was collected and flash frozen in liquid nitrogen. Tissue was ground by hand in a mortar and pestle for >20 min to liberate nuclei. Pure, high-molecular-weight genomic DNA was extracted by first isolating nuclei with the Circulomics Nuclei Isolation kit and then extracting DNA with the Circulomics Nanobind Plant Nuclei Big DNA kit. HiFi libraries were constructed from the genomic DNA and sequenced at the University of Georgia Sequencing Core on a PacBio Sequel II machine.

Genome assembly

We used flow cytometry to estimate genome sizes (diploid DNA content or 2C value) for the three grasses. Healthy leaf tissue was collected from each genotype. Nuclei were isolated and stained according to standard protocols. The stained nuclei were then run on a BD Accuri C6 Plus Flow Cytometer at Plantploidy.com. *Hosta plantaginea* was used as an internal reference.

We built reference genomes for each species using HiFi PacBio long read data. In total, 70.1 Gb of HiFi reads were generated for *M. caffra*, 15.9 Gb for *O. capense* and 20.2 Gb for *T. minimus*, representing 56, 82 and 103 x genome coverage for each species, respectively. K-mer analysis revealed that *O. capense* and *T. minimus* have low within genome heterozygosity and *M. caffra* is a highly heterozygous autopolyploid[58]. PacBio reads were assembled using hifiasm (v 0.18)[30,59] with default settings for *O. capense* and *T. minimus*, and the number of haplotypes was set to 6 for *M. caffra* (flag:–n-hap 6). The resulting assemblies were highly contiguous with six and nine of the ten chromosomes assembled telomere to telomere for *T. minimus* and *O. capense*, respectively, and 118 contigs across 968 Mb with an N50 of 16 Mb for *M. caffra* (Table 1). Raw assemblies were filtered for non-plant contigs using a representative microbial database with BLAST (v 2.10.0)[60]. Full-length chloroplast and mitochondrial genomes were identified and retained, and any additional partial or rearranged organelle genomes were removed.

Genome annotation

A library of repetitive elements was constructed for each of the three grass genomes using the EDTA package (v 2.0.0)[61]. EDTA comprehensively identifies DNA-based transposable elements using HelitronScanner[62] and long terminal repeat (LTR) retrotransposons using LTR_FINDER[63] and *LTRharvest*[64]. Protein coding genes were annotated using the MAKER-P pipeline (v 2.31.10)[65] with the following sets of input data for training. Transcript evidence was generated using the dehydration–rehydration time course RNA-seq data from leaf tissue of each species described below. Raw RNA-seq reads were quality trimmed using fastp (v 0.23)[66] and aligned to the unmasked genomes using the splice aware alignment program STAR (v 2.6)[67]. A set of non-overlapping transcripts was identified from the aligned data using StringTie (v1.3.4)[68] with default parameters. The resulting gff files were used as transcript evidence for MAKER. The same protein evidence was used as training for each of the three grasses, and this includes the full annotations of *Oryza sativa*[69], *Arabidopsis thaliana*[70], *O. thomaeum*[24,26] and *E. tef*[32]. These datasets were used as input for MAKER, and we utilized SNAP (2013 version)[71] and Augustus (v 3.0.2)[72] for ab initio gene prediction, performing two rounds of iterative training to refine our models. To filter out repetitive element-derived proteins, we used BLAST using a non-redundant transposase library against the raw gene models produced by MAKER. We assessed the completeness of our assembly using the plant-specific embryophyte set of BUSCO v.2 (ref. 73). These high-confidence gene models were used for all downstream analyses.

Comparative genomics

The three desiccation-tolerant grass genomes were compared to each other and other Chloridoid grasses using the MCScan toolkit (v 1.1)[74]

implemented in python [https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version)]. Syntenic orthologs were identified across the three focal species as well as *E. nindensis*, *E. tef* and *O. thomaeum* using the chromosome-scale *O. thomaeum* genome as an anchor. Syntenic blocks were identified using gene models aligned using LAST (v 914) with a minimum of five overlapping syntenic genes. The macrosyntenic dot plots, histograms of depth and microsynteny plots were generated using the python version of MCScan. A set of 18,428 conserved syntenic orthologs across all six desiccation-tolerant grasses was created and used for downstream comparative genomic and cross-species transcriptomic analyses. We identified orthologous genes across a subset of 33 land plant species to search for patterns of gene family expansion in desiccation-tolerant lineages as well as for downstream comparative genomic analyses. We included the following species with desiccation-tolerant (DT) species highlighted: *Ananas comosus*, *A. thaliana*, *Brachypodium distachyon*, *Eleusine coracana*, *Eragrostis curvula*, *E. nindensis* (DT), *Eragrostis pilosa*, *E. tef*, *Hordeum vulgare*, *Lindernia brevidens* (DT), *Lindernia subracemosa*, *M. caffra* (DT), *Marchantia polymorpha* (DT), *Medicago truncatula*, *O. capense* (DT), *O. sativa*, *O. thomaeum* (DT), *Physcomitrium patens* (DT), *Sorghum bicolor*, *Setaria italica*, *Selaginella lepidophylla* (DT), *Solanum lycopersicum*, *Selaginella moellendorffii*, *Sporobolus pyramidalis*, *Sporobolus stapfianus* (DT), *Setaria viridis*, *Triticum aestivum*, *T. minimus* (DT), *Vitis vinifera*, *Xerophyta viscosa* (DT), *Zostera japonica*, *Zostera marina* and *Zea mays*. Proteins were clustered into orthologous groups using Orthofinder (v 2.2.6)[75] with default parameters. For the orthogroup enrichment analysis, we calculated a *Z*-score for each species within each orthogroup, compared it to a normal distribution to obtain a *P* value and then adjusted these *P* values using the Benjamini and Hochberg procedure to get *q* values. We then searched for statistically enriched orthogroups across all of the sequenced desiccation-tolerant grasses. Using this approach, we identified between 486 and 8,863 enriched orthogroups in the 33 species we included in our analysis and found none that are conserved across all desiccation-tolerant grasses outside of ELIPs.

ELIP gene family evolution

To test the hypothesis that the ELIP gene family expansions are associated with the evolution of desiccation tolerance, we used CAFE (v 5.1)[34], which analyses changes in gene family size in a phylogenetic framework. The input tree was created from the amino acid sequences from 36 land plant species, with a focus on Chloridoid grasses. Sequences were first clustered using Orthofinder (v 2.4.1)[75], filtered to remove any orthogroups that did not contain all taxa and aligned using MAFFT (v 7.305b)[76]. No single-copy orthologs were found containing all taxa for species tree construction. Instead, we pruned gene trees and alignments to the largest subtree containing unique taxa using PhyloPyPruner (v 1.2.4) (https://gitlab.com/fethalen/phylopypruner); where paralogs were monophyletic within a species, we randomly pruned all but one sequence before extracting the largest subtree. The resulting pruned gene trees and alignments were further filtered to remove any trees no longer containing at least 19 taxa. This final set of 195 alignments were concatenated and used to construct a phylogeny using IQ-TREE (v 2.3.0)[77] and time-calibrated fast least-squares dating[78].

ELIP gene family counts per haploid genome for non-focal taxa were done using BLASTP with the *A. thaliana* (L.) Heynh. ELIP1 amino acid sequence as query for the remaining proteomes. We further investigated two other gene families with known roles in desiccation tolerance—heat shock proteins and late embryogenesis abundant proteins—along with 20 randomly selected orthogroups, to contextualize the tempo of ELIP evolution. These count data and the time-calibrated phylogeny were used as input for CAFE under a single lambda model.

Transcriptomic analyses

RNA-seq reads were processed following a pipeline developed by the VanBuren Lab (https://github.com/pardojer23/RNAseqV2). Briefly, sequence read quality was assessed with fastQC (v 0.23), and reads were trimmed with trimmomatic (v 0.38)[79] to remove adapters and low-quality bases. Trimmed reads were sudo-aligned to reference genomes using Salmon (v 1.9.0)[35], and the resulting quantification files were processed with tximport (v 3.18)[80] to generate normalized expression matrices of transcripts per million (TPM). PCA was used to visualize replicate and sample relationships within each species using the respective TPM expression values. A cross-species PCA was performed using the TPM matrix of conserved syntenic orthologs across all species. To effectively quantify gene expression while acknowledging the complexities introduced by polyploidy, we summed the expression levels of all homeologs in *E. nindensis*, *E. tef* and *M. caffra* to obtain a single gene expression value to enable interspecies comparisons. This approach is grounded in the logic that a unified expression value not only simplifies the analysis but also encapsulates potential functional diversifications among homeologs. This methodology has been applied and validated in our previous research[17,29,42,81].

Differentially expressed genes

DEGs were identified independently for each species with DEseq2 R package (v 1.42.0)[82]. Briefly, transcript abundance estimates from Salmon were imported into DEseq2 using tximport to generate counts matrices. We tested multiple models for differential expression in DEseq2, including models that identified DEGs by pairwise comparisons of each time point against well-watered, and models that used the continuous variables of RWC or F_v/F_m as covariates. DEGs identified by pairwise comparisons were summarized into a nonredundant list of up- and down-regulated genes during dehydration and rehydration. DEGs identified using the continuous variables are based on a significant linear association (positive or negative) with RWC or F_v/F_m. When identifying DEGs, we included the term 'process' in our model to differentiate between dehydration and rehydration processes. To select the best performing model, we quantified similarities and differences in the number and identity of DEGs defined by each model. There was a high degree of overlap in genes identified by all three models. Ultimately, we selected the model based on RWC because it performed well and is easily comparable across experiments regardless of sampling time, consistency across replicates or differences in experimental design. These analyses produced species-specific lists of DEGs during dehydration and rehydration with significant (FDR adjusted $P < 0.05$) associations with RWC. $\log_2$ fold change values are calculated for one unit change in RWC.

We then compared the observed proportion of overlapping DEGs in each category to the proportion of genes expected to overlap by chance (assuming independent draws) and tested whether these were significantly different using Fisher's exact test. This analysis was then extended to include DEGs identified in the desiccation-sensitive sister species *E. tef* to distinguish between typical drought versus pure desiccation responses. We then conducted targeted analyses to look at the functional roles of differentially expressed syntelogs that were uniquely shared across the three resurrection species versus those that were common with *E. tef*.

Functional annotation of DEGs

We annotated differentially expressed syntelogs with KEGG and GO terms to describe metabolic and cellular processes shared across the three study species. KEGG annotations were generated using BLASTKoala (https://www.kegg.jp/blastkoala/) for each species' annotated peptide sequences and assigned to syntenic orthologs. Shared differentially expressed KEGG terms across all three species during dehydration and rehydration were identified and plotted with Venn diagrams. These terms were used to generate metabolic pathway maps via KEGGmapper (https://www.genome.jp/kegg/mapper/color.html) for up- and down-regulated terms. A list of syntelogs per metabolic pathway and Brite descriptions was generated, and pathways were sorted by the difference in syntelog counts for up- and down-regulated genes.

KEGG orthology numbers were paired with gene expression data to identify active pathways at various time points. Genetic information processes and environmental information processing, along with cellular processes, were grouped, while pathways assigned to organismal systems and human disease were ignored. KEGG annotation limitations include the presence of single KEGG identifiers in multiple pathways.

GO terms were assigned through homology with the well-annotated genome of sister species *O. thomaeum* using BLASTP (v 2.14.0) with an *e* value cut-off of 1×10^{-10}, returning the single best match for each peptide. The GO terms from *O. thomaeum* were assigned to homologous genes in the target species. We used the TopGO R package (v 2.54.0) to identify significantly enriched GO terms ($P < 0.05$) among DEGs for up- and down-regulated genes during dehydration and rehydration in each target species and among different sets of overlapping and unique syntelogs from cross-species comparisons.

Co-expression analyses

We generated co-expression networks using Weighted Gene Co-expression Network Analysis (WGCNA) R package (v1.7)[83], filtering each dataset to exclude non-expressed genes. Each dataset was filtered to remove genes with no expression. A soft thresholding power was chosen to ensure a scale-free network, and an adjacency matrix was constructed. This was converted to a topological overlap matrix, and hierarchical clustering grouped genes into modules based on expression patterns. Gene connectivity within networks and modules was calculated. Shared and species-specific co-expressed genes were identified using UpSet plots[84], and syntenic orthologs were analysed for overlap across species. Modules with increased expression during dehydration, rehydration and non-stressed conditions were combined, and GO enrichment analysis was performed on shared and unique co-expressed gene sets.

Topological data analysis

We used a topological data analysis approach following the pipeline described at https://github.com/PlantsAndPython/plant-evo-mapper to discern the underlying structure of the expression datasets. We utilized the Mapper algorithm, which condenses the dataset into a scalable, navigable representation. For our gene expression data, we constructed Mapper graphs using a 'stress lens' formulated by applying a linear model using the well-watered condition as a reference point. This model represents the baseline for leaf expression, and we quantified the residuals or deviation of each sample from the baseline, which represents the degree of water stress or recovery. We generated three different mapper graphs from the syntelog expression matrix from the three focal resurrection grasses (*M. caffra*, *O. capense* and *T. minimus*), five resurrection grasses (*E. nindensis* and *O. thomaeum*) and the third graph which included the desiccation-sensitive species *E. tef*. For the mapper graph, we specified different intervals and overlap for the three-species comparisons and the five-species comparisons. For the three-species comparison, we specified 110 intervals with a 90% overlap, and for the five-species comparison, we specified 120 intervals with 95% overlap.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequence data associated with this study are deposited at NCBI under BioProject PRJNA1044305 and BioSamples SAMN38380430-92. Genome assemblies are hosted on CoGe (https://genomevolution.org/) under the following IDs: 65089 (*T. minimus*), 65046 (*O. capense*) and 64494 (*M. caffra*). Metadata and other data summaries associated with this study are available via Dryad at https://doi.org/10.5061/dryad.kh18932c4 (ref. 85).

Code availability

Code associated with this project is available via GitHub at https://github.com/bobvanburen/Marks_Convergent_DT_grasses.

References

1. Bewley, J. D. Physiological aspects of desiccation tolerance. *Annu. Rev. Plant Physiol.* **30**, 195–238 (1979).
2. Oliver, M. J., Tuba, Z. & Mishler, B. D. The evolution of vegetative desiccation tolerance in land plants. *Plant Ecol.* **151**, 85–100 (2000).
3. Marks, R. A., Farrant, J. M., Nicholas McLetchie, D. & VanBuren, R. Unexplored dimensions of variability in vegetative desiccation tolerance. *Am. J. Bot.* https://doi.org/10.1002/ajb2.1588 (2021).
4. Alpert, P. Constraints of tolerance: why are desiccation-tolerant organisms so small or rare? *J. Exp. Biol.* **209**, 1575–1584 (2006).
5. VanBuren, R. Desiccation tolerance: seedy origins of resurrection. *Nat. Plants* **3**, 17046 (2017).
6. Costa, M. C. D. et al. Key genes involved in desiccation tolerance and dormancy across life forms. *Plant Sci.* https://doi.org/10.1016/j.plantsci.2016.02.001 (2016).
7. VanBuren, R. et al. Desiccation tolerance evolved through gene duplication and network rewiring in *Lindernia*. *Plant Cell* **30**, 2943–2958 (2018).
8. Alpert, P. The limits and frontiers of desiccation-tolerant life. *Integr. Comp. Biol.* **45**, 685–695 (2005).
9. Porembski, S. & Barthlott, W. Granitic and gneissic outcrops (inselbergs) as centers of diversity for desiccation-tolerant vascular plants. *Plant Ecol.* **151**, 19–28 (2000).
10. Dace, H. J. W. et al. A horizontal view of primary metabolomes in vegetative desiccation tolerance. *Physiol. Plant.* **175**, e14109 (2023).
11. Crowe, J. H., Carpenter, J. F. & Crowe, L. M. The role of vitrification in anhydrobiosis. *Annu. Rev. Physiol.* **60**, 73–103 (1998).
12. Hoekstra, F. A., Golovina, E. A. & Buitink, J. Mechanisms of plant desiccation tolerance. *Trends Plant Sci.* **6**, 431–438 (2001).
13. Dinakar, C., Djilianov, D. & Bartels, D. Photosynthesis in desiccation tolerant plants: energy metabolism and antioxidative stress defense. *Plant Sci.* **182**, 29–41 (2012).
14. Oliver, M. J. et al. Desiccation tolerance: avoiding cellular damage during drying and rehydration. *Annu. Rev. Plant Biol.* https://doi.org/10.1146/annurev-arplant-071219-105542 (2020).
15. Moore, J. P., Le, N. T., Brandt, W. F., Driouich, A. & Farrant, J. M. Towards a systems-based understanding of plant desiccation tolerance. *Trends Plant Sci.* **14**, 110–117 (2009).
16. Farrant, J. M. & Hilhorst, H. Crops for dry environments. *Curr. Opin. Biotechnol.* **74**, 84–91 (2021).
17. VanBuren, R. et al. Core cellular and tissue-specific mechanisms enable desiccation tolerance in *Craterostigma*. *Plant J.* **114**, 231–245 (2023).
18. Pearce, T. Convergence and parallelism in evolution: a Neo-Gouldian account. *Br. J. Philos. Sci.* **63**, 429–448 (2012).
19. Stern, D. L. The genetic causes of convergent evolution. *Nat. Rev. Genet.* **14**, 751–764 (2013).
20. Wiley, E. O. Homoplasy. In *Encyclopedia of Genetics* (eds Brenner, S. & Miller, J. H.) 969–970 (Academic, 2001).
21. Arendt, J. & Reznick, D. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32 (2008).
22. Tebele, S. M., Marks, R. A. & Farrant, J. M. Two decades of desiccation biology: a systematic review of the best studied angiosperm resurrection plants. *Plants* **10**, 2784 (2021).
23. Xiao, L. et al. The resurrection genome of *Boea hygrometrica*: a blueprint for survival of dehydration. *Proc. Natl Acad. Sci. USA* **112**, 5833–5837 (2015).
24. VanBuren, R. et al. Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature* **527**, 508–511 (2015).

25. Costa, M.-C. D. et al. A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*. *Nat. Plants* **3**, 17038 (2017).
26. VanBuren, R., Wai, C. M., Keilwagen, J. & Pardo, J. A chromosome-scale assembly of the model desiccation tolerant grass *Oropetium thomaeum*. *Plant Direct* **2**, e00096 (2018).
27. VanBuren, R. et al. Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. *Nat. Commun.* **9**, 13 (2018).
28. Pardo, J. et al. Intertwined signatures of desiccation and drought tolerance in grasses. *Proc. Natl Acad. Sci. USA* **117**, 10079–10088 (2020).
29. Montes, R. A. C. et al. A comparative genomics examination of desiccation tolerance and sensitivity in two sister grass species. *Proc. Natl Acad. Sci. USA* **119**, e2118886119 (2022).
30. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
31. Marks, R. A., Hotaling, S., Frandsen, P. B. & VanBuren, R. Representation and participation across 20 years of plant genome sequencing. *Nat. Plants* **7**, 1571–1578 (2021).
32. VanBuren, R. et al. Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nat. Commun.* **11**, 884 (2020).
33. VanBuren, R., Pardo, J., Man Wai, C., Evans, S. & Bartels, D. Massive tandem proliferation of ELIPs supports convergent evolution of desiccation tolerance across land plants. *Plant Physiol.* **179**, 1040–1049 (2019).
34. Han, M. V., Thomas, G. W. C., Lugo-Martinez, J. & Hahn, M. W. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* **30**, 1987–1997 (2013).
35. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
36. du Toit, S. F., Bentley, J. & Farrant, J. M. Chapter Nine - NADES formation in vegetative desiccation tolerance: prospects and challenges. in *Advances in Botanical Research* (eds Verpoorte, R., Witkamp, G.-J. & Choi, Y. H.) vol. 97 225–252 (Academic, 2021).
37. Hasanuzzaman, M., Nahar, K., Anee, T. I. & Fujita, M. Glutathione in plants: biosynthesis and physiological role in environmental stress tolerance. *Physiol. Mol. Biol. Plants* **23**, 249–268 (2017).
38. Gasulla, F. et al. The role of lipid metabolism in the acquisition of desiccation tolerance in *Craterostigma plantagineum*: a comparative approach. *Plant J.* **75**, 726–741 (2013).
39. Sun, M. et al. Phosphatidylcholine enhances homeostasis in peach seedling cell membrane and increases its salt stress tolerance by phosphatidic acid. *Int. J. Mol. Sci.* **23**, 2585 (2022).
40. Zhang, X., Gao, Y., Zhuang, L. & Huang, B. Phosphatidic acid priming-enhanced heat tolerance in tall fescue (*Festuca arundinacea*) involves lipidomic reprogramming of lipids for membrane stability and stress signaling. *Plant Growth Regul.* **99**, 527–538 (2023).
41. VanBuren, R. et al. Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass *Oropetium thomaeum*. *Plant Cell Environ.* **40**, 2292–2306 (2017).
42. Palande, S. et al. Topological data analysis reveals a core gene expression backbone that defines form and function across flowering plants. *PLoS Biol.* **21**, e3002397 (2023).
43. Lim, P. O., Kim, H. J. & Nam, H. G. Leaf senescence. *Annu. Rev. Plant Biol.* **58**, 115–136 (2007).
44. Li, S. et al. Histone acetylation changes in plant response to drought stress. *Genes* 12 (2021).
45. Wang, Q. et al. JMJ27-mediated histone H3K9 demethylation positively regulates drought-stress responses in *Arabidopsis*. *New Phytol.* **232**, 221–236 (2021).
46. Ma, S. et al. Reversible histone H2B monoubiquitination fine-tunes abscisic acid signaling and drought response in rice. *Mol. Plant* **12**, 263–277 (2019).
47. Farrant, J. M. A comparison of mechanisms of desiccation tolerance among three angiosperm resurrection plant species. *Plant Ecol.* **151**, 29–39 (2000).
48. VanBuren, R. et al. Variability in drought gene expression datasets highlight the need for community standardization. Preprint at *bioRxiv* https://doi.org/10.1101/2024.02.04.578814 (2024).
49. Illing, N., Denby, K. J., Collett, H., Shen, A. & Farrant, J. M. The signature of seeds in resurrection plants: a molecular and physiological comparison of desiccation tolerance in seeds and vegetative tissues. *Integr. Comp. Biol.* **45**, 771–787 (2005).
50. Alejo-Jacuinde, G., González-Morales, S. I., Oropeza-Aburto, A., Simpson, J. & Herrera-Estrella, L. Comparative transcriptome analysis suggests convergent evolution of desiccation tolerance in *Selaginella* species. *BMC Plant Biol.* **20**, 468 (2020).
51. Ostria-Gallardo, E. et al. A comparative gene co-expression analysis using self-organizing maps on two congener filmy ferns identifies specific desiccation tolerance mechanisms associated to their microhabitat preference. *BMC Plant Biol.* **20**, 56 (2020).
52. Heyduk, K., Moreno-Villena, J. J., Gilman, I. S., Christin, P.-A. & Edwards, E. J. The genetics of convergent evolution: insights from plant photosynthesis. *Nat. Rev. Genet.* **20**, 485–493 (2019).
53. Christin, P.-A., Salamin, N., Savolainen, V., Duvall, M. R. & Besnard, G. C4 photosynthesis evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* **17**, 1241–1247 (2007).
54. Vicré, M., Farrant, J. M. & Driouich, A. Insights into the cellular mechanisms of desiccation tolerance among angiosperm resurrection plant species. *Plant Cell Environ.* **27**, 1329–1340 (2004).
55. Gechev, T., Lyall, R., Petrov, V. & Bartels, D. Systems biology of resurrection plants. *Cell. Mol. Life Sci.* https://doi.org/10.1007/s00018-021-03913-8 (2021).
56. Xu, X. et al. Molecular insights into plant desiccation tolerance: transcriptomics, proteomics and targeted metabolite profiling in *Craterostigma plantagineum*. *Plant J.* https://doi.org/10.1111/tpj.15294 (2021).
57. St Aubin, B., Wai, C. M., Kenchanmane Raju, S. K., Niederhuth, C. E. & VanBuren, R. Regulatory dynamics distinguishing desiccation tolerance strategies within resurrection grasses. *Plant Direct* **6**, e457 (2022).
58. Marks, R. A. et al. Higher order polyploids exhibit enhanced desiccation tolerance in the grass Microchloa caffra. *J. Exp. Bot.* **75**, 3612–3623 (2024).
59. Cheng, H. et al. Haplotype-resolved assembly of diploid genomes without parental data. *Nat. Biotechnol.* **40**, 1332–1335 (2022).
60. Wheeler, D. L. et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **36**, D13–D21 (2008).
61. Ou, S. et al. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).
62. Xiong, W., He, L., Lai, J., Dooner, H. K. & Du, C. HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. *Proc. Natl Acad. Sci. USA* **111**, 10263–10268 (2014).
63. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
64. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).

65. Campbell, M. S. et al. MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiol.* **164**, 513–524 (2014).
66. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
67. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
68. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
69. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005).
70. Cheng, C.-Y. et al. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* **89**, 789–804 (2017).
71. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
72. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003).
73. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
74. Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
75. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
76. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
77. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
78. To, T.-H., Jung, M., Lycett, S. & Gascuel, O. Fast dating using least-squares criteria and algorithms. *Syst. Biol.* **65**, 82–97 (2016).
79. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
80. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
81. Pardo, J. et al. Cross-species predictive modeling reveals conserved drought responses between maize and sorghum. *Proc. Natl Acad. Sci. USA* **120**, e2216894120 (2023).
82. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
83. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
84. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
85. Marks, R. Convergent evolution of desiccation tolerance in grasses. *Dryad* https://doi.org/10.5061/dryad.kh18932c4 (2023).
86. GBIF.org. GBIF Occurrence Download https://doi.org/10.15468/dl.5jf47y (2023).

Acknowledgements

This work was funded by the US National Science Foundation IOS-PRFB-1906094 to R.A.M., DBI-2213983 to R.A.M. and R.V., MCB-1817347 to R.V., and the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA 2022-67013-36118) to R.V. We thank landowners and stewards P. and J. Pretorius, K. Maude, P. and N. Vervoort, and W. and M. Mudenda for assistance with field logistics and permission to collect plants. We thank the University of Cape Town for access to facilities, J. Farrant for her guidance and K. Cooper for logistical support. We also thank the South African National Herbarium, Pretoria, for assistance identifying and vouchering of specimens and the US Department of Agriculture for providing import permits.

Author contributions

R.A.M. and R.V. conceived of the study. R.A.M. collected and curated data. R.A.M., L.V.D.P., J.S., I.S.G. and R.V. conducted data analyses and contributed to data interpretation and conceptual framing of the manuscript. R.A.M. and R.V. drew the figures. R.A.M., L.V.D.P. and R.V. wrote the manuscript. All authors edited and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

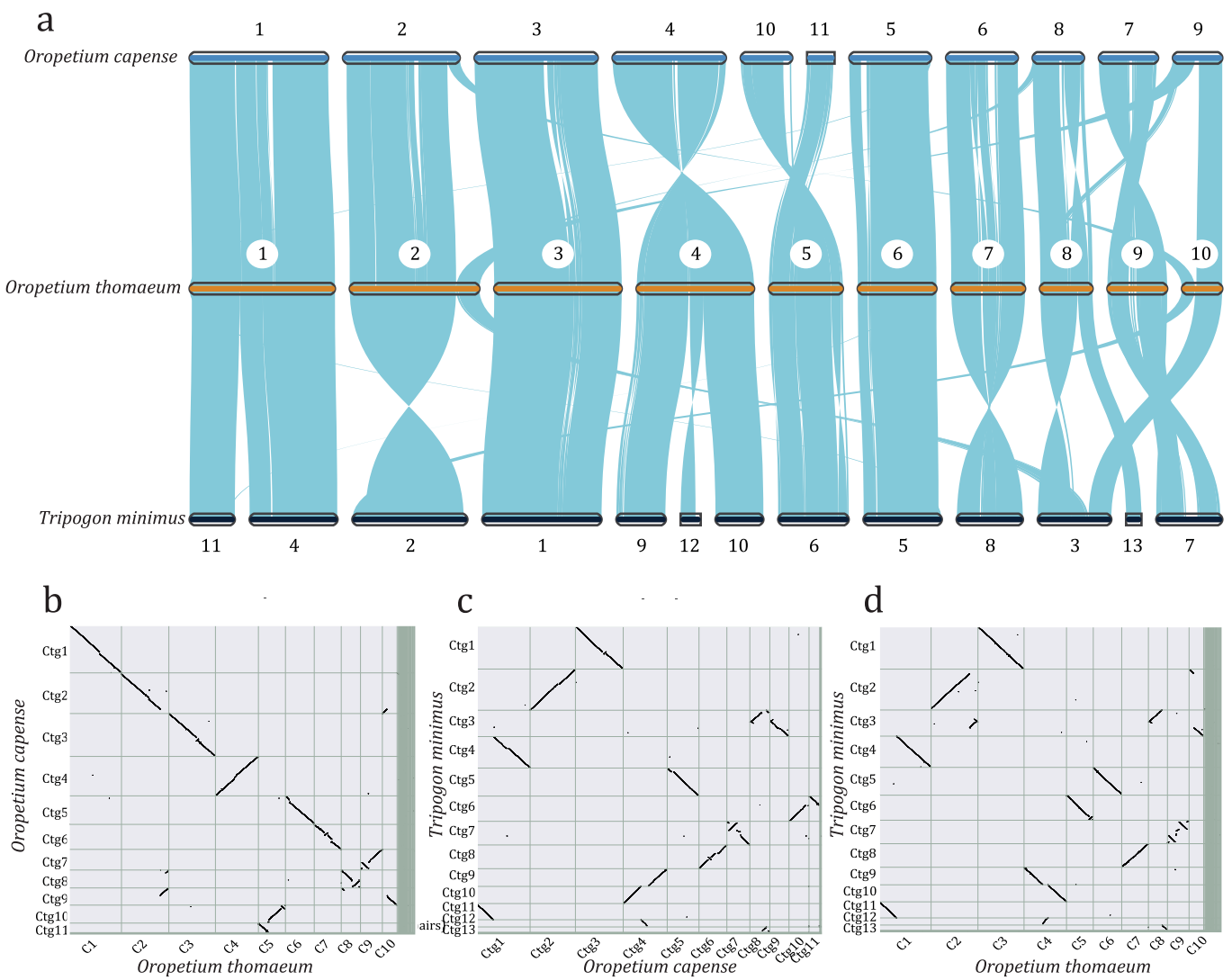
Extended data is available for this paper at https://doi.org/10.1038/s41477-024-01729-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41477-024-01729-5.

Correspondence and requests for materials should be addressed to Rose A. Marks or Robert VanBuren.

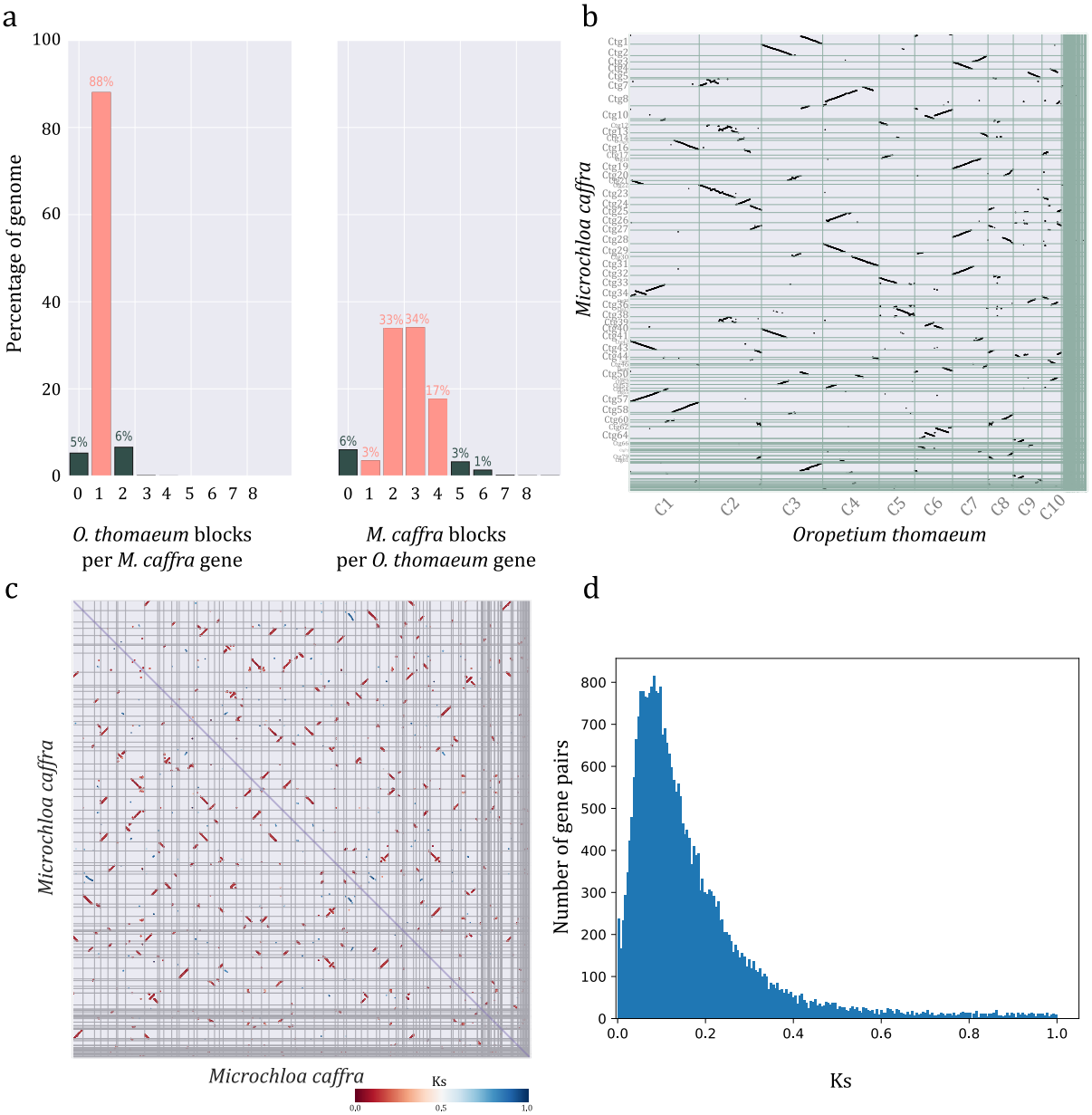
Peer review information *Nature Plants* thanks Jianquan Liu, Suhua Shi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41477-024-01729-5



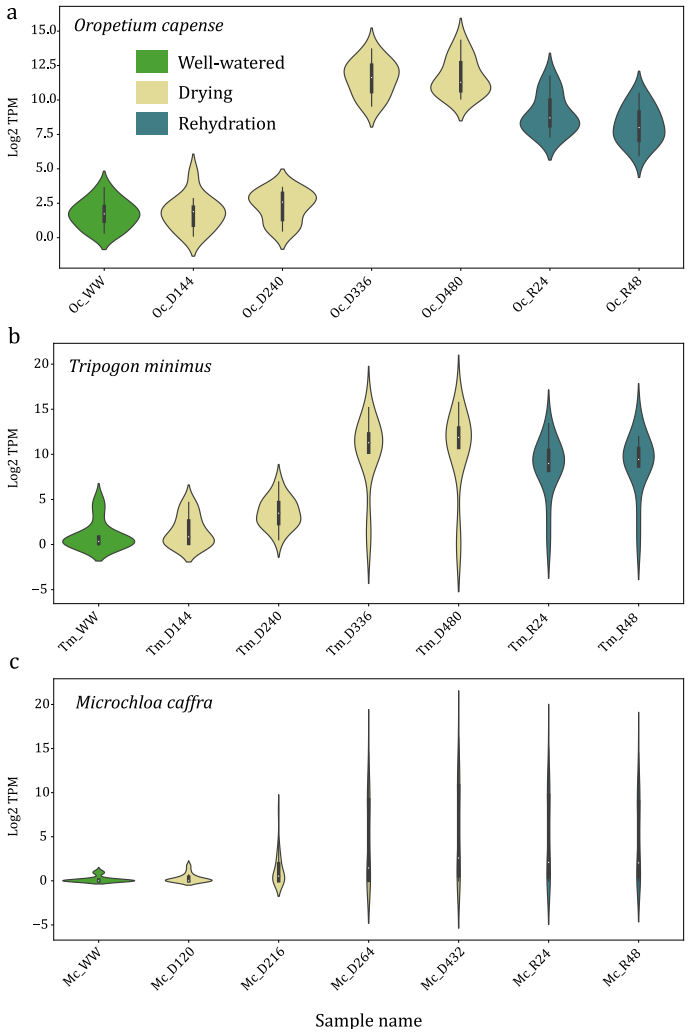
Extended Data Fig. 1. | Comparative genomics of the diploid desiccation tolerant grasses. (**a**) Large-scale genomic comparison of *O. capense* (top) *O. thomaeum* (middle) and *T. minimus* (bottom). (**b**) Macrosyntenic dot plot between the *O. capense* and *O. thomaeum* genomes. (**c**) Macrosyntenic dot plot between the *O. capense* and *T. minimus* genomes. (**d**) Macrosyntenic dot plot between the *T. minimus* and *O. thomaeum* genomes.

https://doi.org/10.1038/s41477-024-01729-5



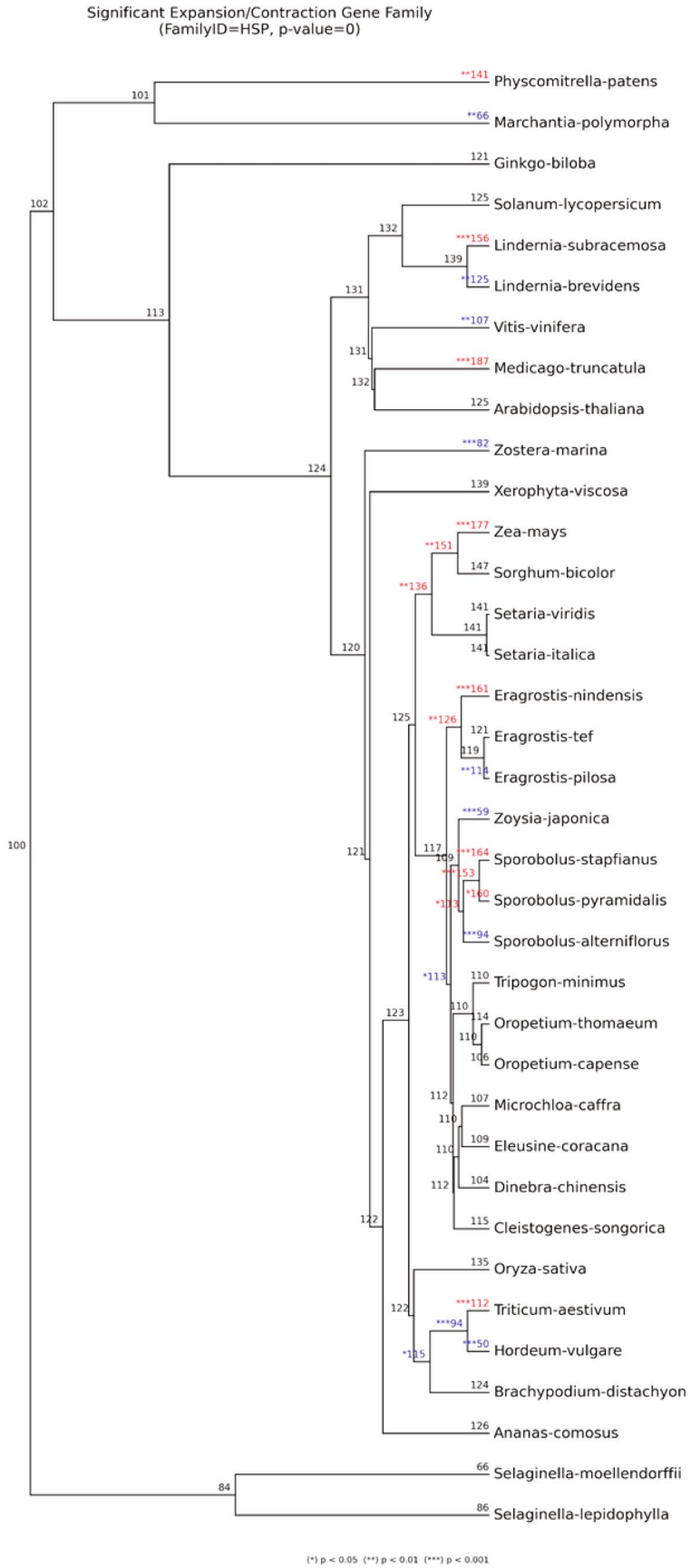
Extended Data Fig. 2. | Comparative genomics and polyploidy in *M. caffra*. (**a**) Syntenic depth of *O. thomaeum* blocks per *M. caffra* gene (left) and *M. caffra* blocks per *O. thomaeum* gene (right). (**b**) Syntenic dot plot between the *O. thomaeum* and *M. caffra* genomes where each dot represents a syntenic gene pair. (**c**) Macrosyntenic dot plot of *M. caffra* x *M. caffra* with syntenic gene pairs colored by Ks. (**d**) Histogram of Ks for homeologous gene pairs in *M. caffra*.

https://doi.org/10.1038/s41477-024-01729-5



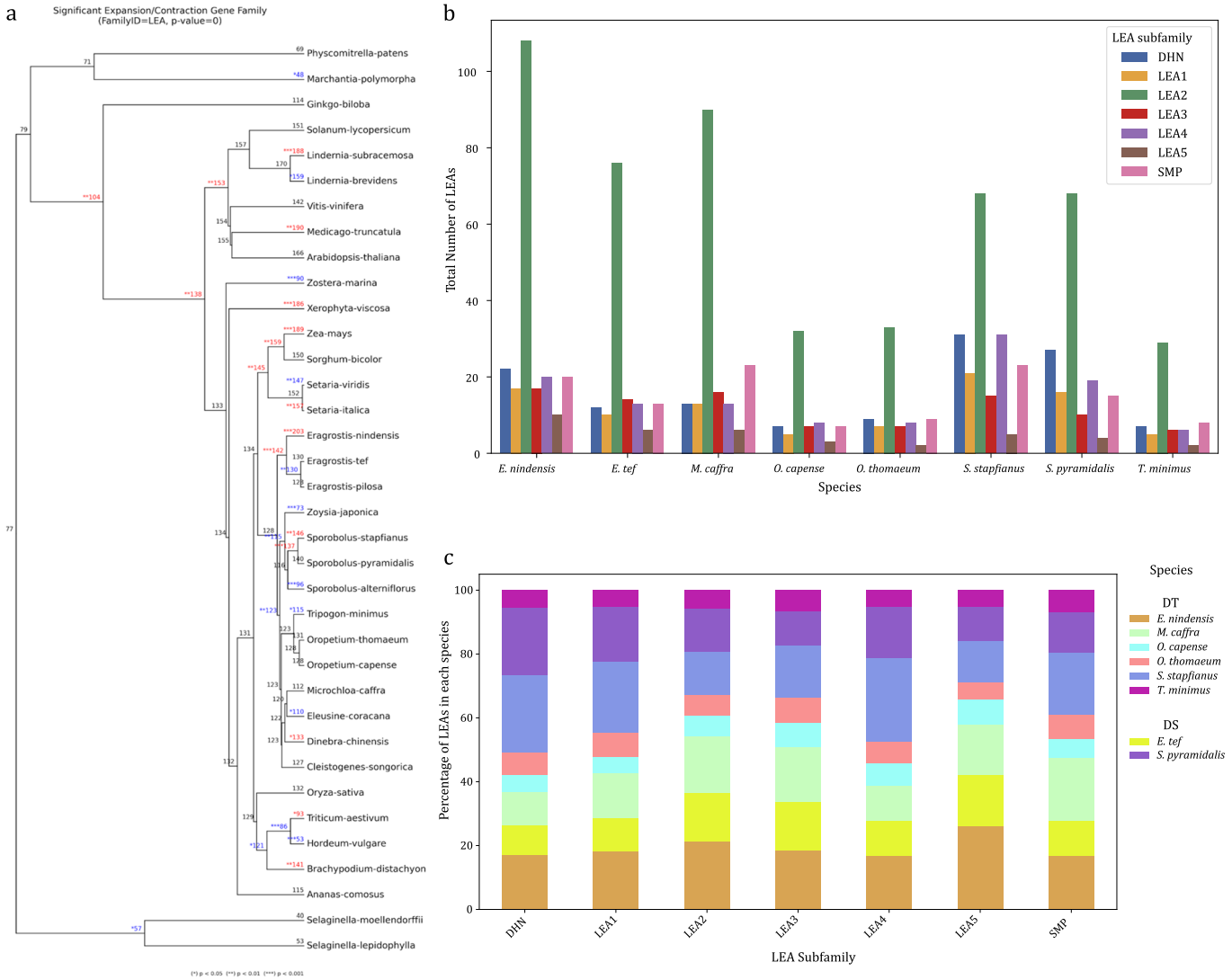
Extended Data Fig. 3. | Expression patterns of ELIPs across the three surveyed desiccation tolerant grasses. The Log2 transformed TPMs are plotted for all ELIPs in well-watered, dehydrated, and rehydrated samples for *O. capense* (**a**), *T. minimus* (**b**), and *M. caffra* (**c**). Three biological replicates were collected for each sample, and the expression of all ELIPs is shown in each violin box plot. The box plots represent minima, maxima, median (center line), first and third quartiles (bounds of the box), and whiskers (values within 1.5 times the IQR).

https://doi.org/10.1038/s41477-024-01729-5



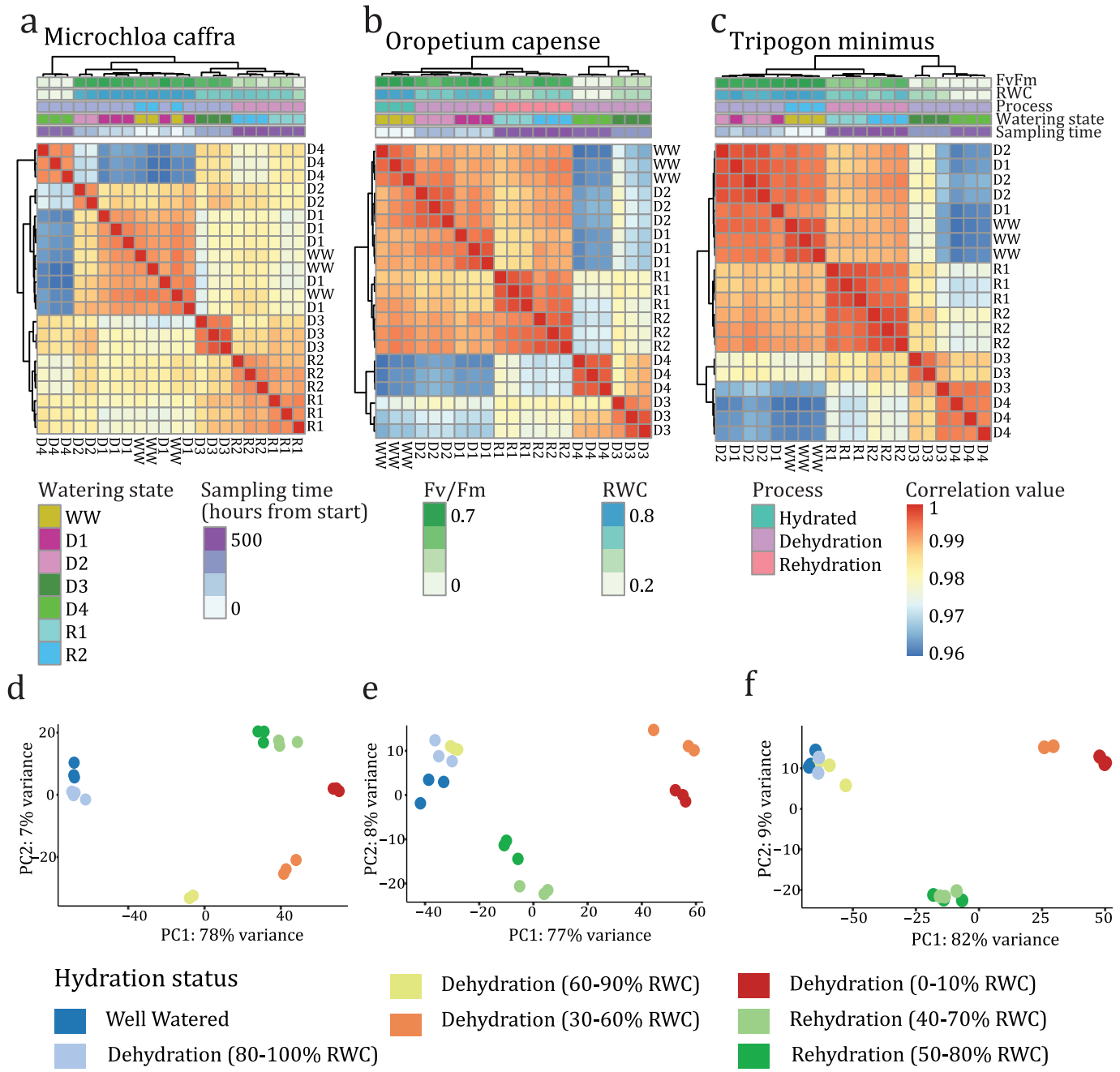
Extended Data Fig. 4. | HSP evolutionary dynamics showing significant changes in the rates of gene family expansion (red) and contraction (blue) inferred by CAFE. Numbers are node labels show the HSP copy number for each species (tips) or the modeled ancestral copy number (internal nodes). Café employs a likelihood ratio test (LRT) to compare models with and without gene gain and loss events.

https://doi.org/10.1038/s41477-024-01729-5



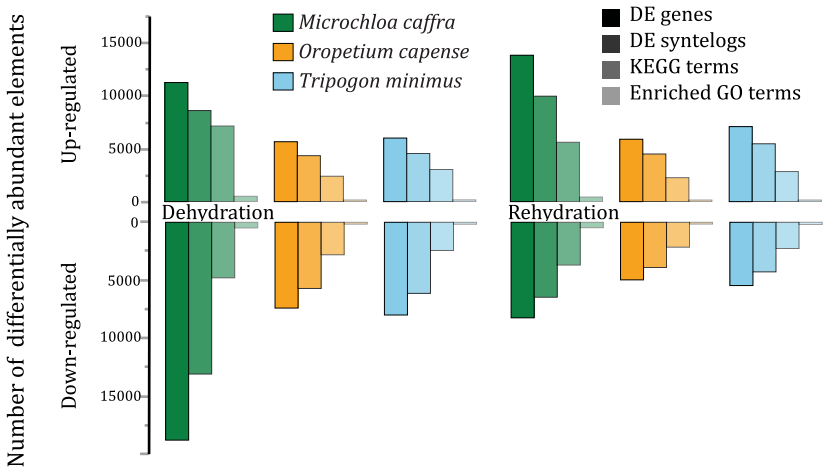
Extended Data Fig. 5. | LEA evolutionary dynamics. Significant changes in the rates of gene family expansion (red) and contraction (blue) inferred by CAFÉ are shown in (**a**). Numbers are node labels show the LEA copy number for each species (tips) or the modeled ancestral copy number (internal nodes). The total number of LEAs in each of the 7 subfamilies is plotted for six desiccation tolerant and two desiccation sensitive grasses (**b**) and the proportion of LEAs in each of the species (**c**).

https://doi.org/10.1038/s41477-024-01729-5



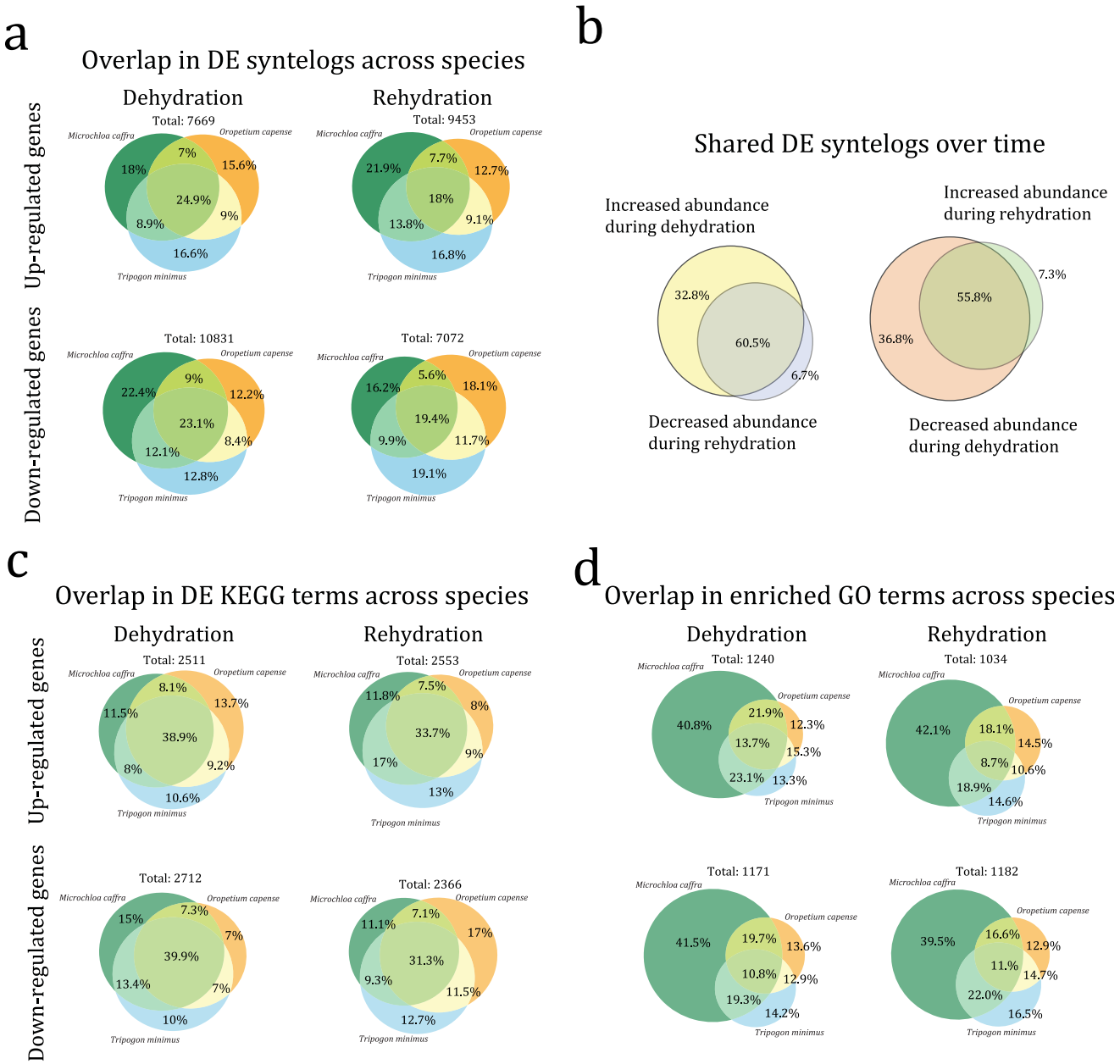
Extended Data Fig. 6. | Clustering and dimensionality reduction of desiccation and rehydration data for the three resurrection grasses. (**a-c**) Hierarchical clustering of gene expression across the dehydration timecourse for (**a**) *Microchloa caffra*, (**b**) *Oropetium capense*, and (**c**) *Tripogon minimus*. Samples cluster by sampling time, watering status, RWC, and Fv/Fm as expected. (**d-f**) Principal component analysis of gene expression data for the three resurrection grasses. The first two principal components are plotted using the raw TPMs for *Microchloa caffra* (**d**) *Oropetium capense* (**e**), and *Tripogon minimus* (**f**). Samples are colored by hydration status.

https://doi.org/10.1038/s41477-024-01729-5



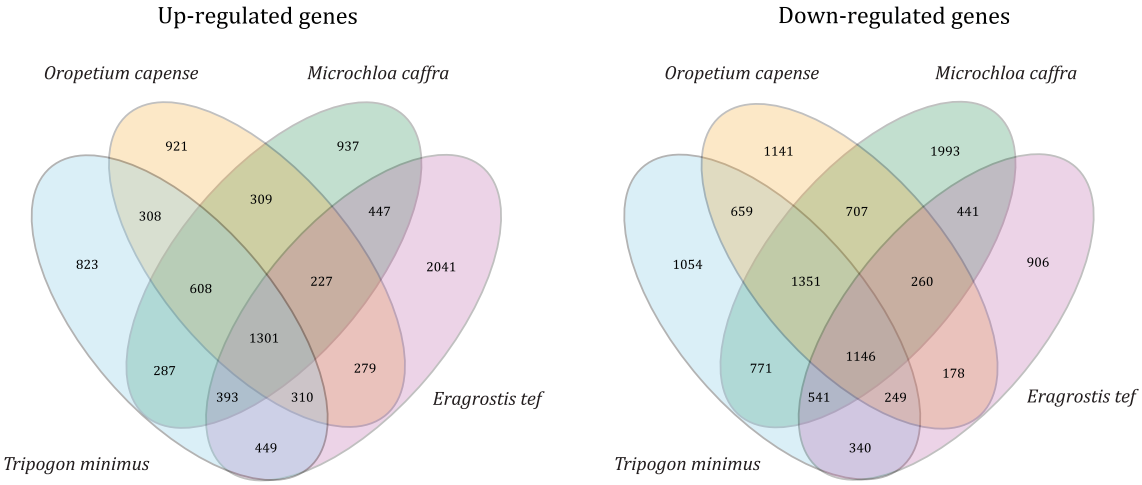
Extended Data Fig. 7. | Comparison of enriched genes or pathways during desiccation. Both up- and down-regulated genes, syntelogs, KEGG terms, and enriched GO terms in each species are shown under dehydration and rehydration conditions.

https://doi.org/10.1038/s41477-024-01729-5



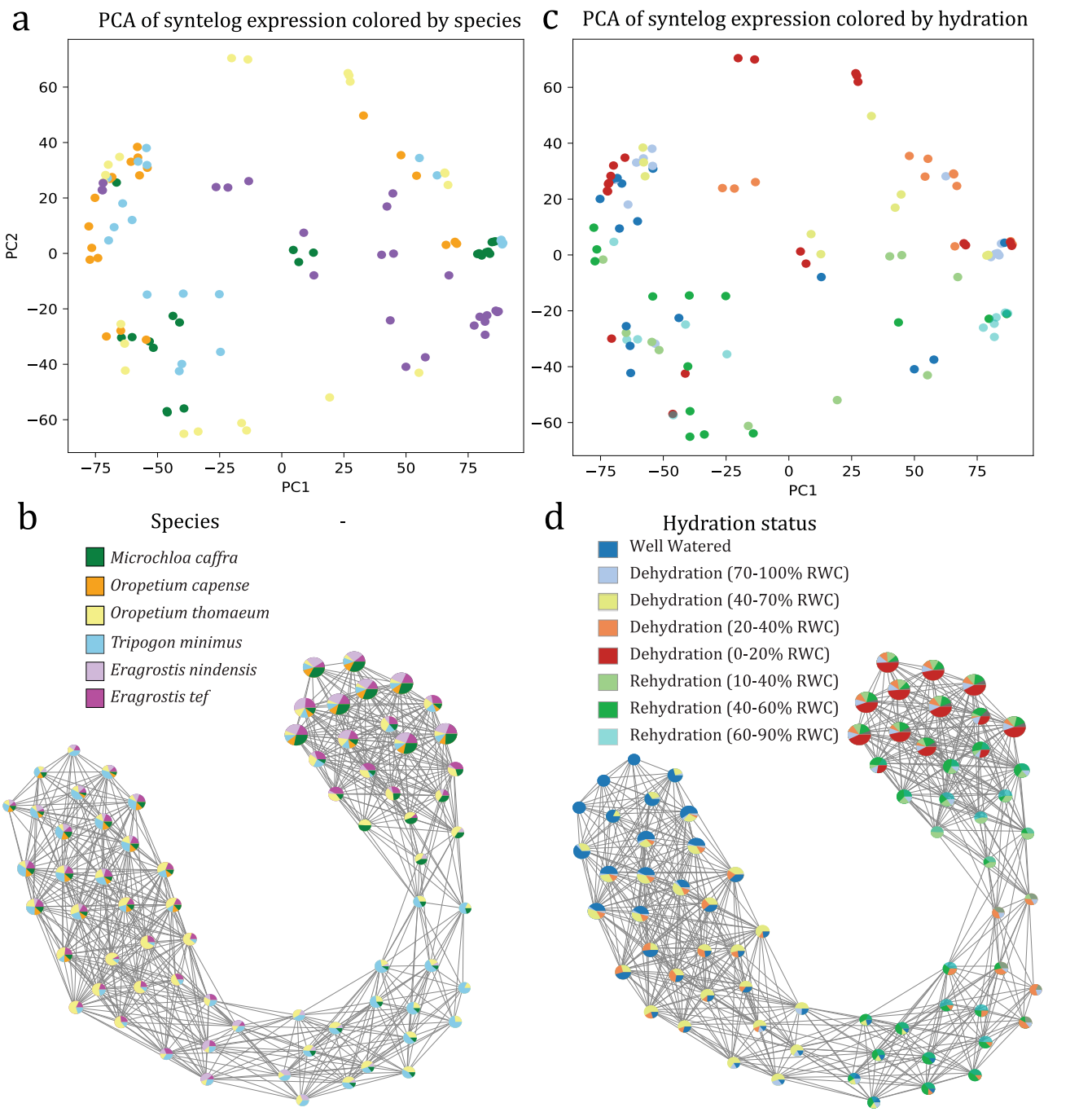
Extended Data Fig. 8. | Overlap in differentially expressed elements across species. Up- and down-regulated genes during dehydration and rehydration conditions for (**a**) syntelogs, (**b**) KEGG terms, and (**c**) enriched GO terms. (**d**) overlap in shared differentially expressed syntelogs across time.

https://doi.org/10.1038/s41477-024-01729-5



Extended data Fig. 9. | Overlap in differentially expressed syntelogs. Venn diagrams are shown for up and down regulated syntenic gene pairs in the three focal resurrection grasses and the desiccation sensitive grass *Eragrostis tef*.

https://doi.org/10.1038/s41477-024-01729-5



Extended data Fig. 10. | Expression patterns of syntelogs across the resurrection grasses. Principal component analysis on z-scores of syntelog expression for all 5 resurrection species with lots are colored by species identity (**a**) and hydration status (**c**). Topological data analysis of syntelog expression for 5 resurrection grasses and one desiccation sensitive species (*Eragrostis tef*). Nodes within the graph represent clusters of RNAseq samples that are akin to one another, with the node color indicating the identity of the samples contained within. Edges, or the connections between nodes, delineate shared samples across intersecting clusters. A TDA graph by species is shown in (b) and by hydration status in (**d**).

nature portfolio

Corresponding author(s): Robert VanBuren and Rose Marks

Last updated by author(s): May 12, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	no software was used for data collection
Data analysis	The following software with versions for our data analysis include: hifiasm (v 0.18), EDTA package (v2.0.0), MAKER-P pipeline (v2.31.10), fastp (v 0.23), STAR (v2.6), StringTie (v1.3.4), Augustus (version 3.0.2), BUSCO v.2, MCScan toolkit (v1.1), Orthofinder (v2.2.6), CAFÉ (v 5.1), Orthofinder (v 2.4.1), MAFFT (v 7.305b), PhyloPyPruner (v 1.2.4), IQ-TREE (v 2.3.0), fastQC (v 0.23), trimmomatic (v 0.38), Salmon (v 1.9.0), tximport (v 3.18), DEseq2 R package (v 1.42.0), BLASTP (v 2.14.0), TopGO R package (v 2.54.0), WGCNA R package (v1.7), mapper (0.2.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Sequence data associated with this study are deposited at NCBI under BioProject PRJNA1044305 and BioSamples SAMN38380430-92. Genome assemblies are

hosted on CoGe (https://genomevolution.org/) under the following IDs: 65089 (T. minimus), 65046 (O. capense), and 64494 (M. caffra). Metadata and other data summaries associated with this study are provided at Dryad https://doi.org/10.5061/dryad.kh18932c4.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For gene expression analysis, 9 plants were collected for each timepoint. Three of these plants were pooled together for each biological replicate for gene expression analyses, and values were averaged within each replicate for relative water content and fv/fm estimates. Three biological replicates is standard for RNAseq analysis and variance across replicates was surveyed using principal component analysis and various downstream statistical analyses.
Data exclusions	no data were excluded
Replication	3 replicates were collected for each physiology and gene expression timepoint, all replicates were successful.
Randomization	Plants were randomized within the growth chamber
Blinding	This is not relevant to our study as data was only collected for one genotype and timepoints for sampling were decided during experimental design.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	Plants for the current study were collected from two research sites in South Africa: Buffelskloof Nature Reserve in Mpumalanga (-25.30229 S, 030.50631 E) (Microchloa caffra) and Swebe Swebe Private Wildlife Reserve in Limpopo (-23.7949 S, 028.0705 E) (Oropetium capense and Tripogon minimus). Voucher specimens of each species were collected, pressed, and deposited at the National Herbarium of South Africa in Pretoria (specimen numbers: PRE1004810-0, PRE1004793-0, and PRE1004794-0). Seeds of each species were also collected and transported to Michigan State University under United States Department of Agriculture (USDA) permit #537-22-37-10071 and according to the specifications in a Material Transfer Agreement established between Drs. Jill M. Farrant, Robert VanBuren, and Rose A. Marks.
Novel plant genotypes	NA
Authentication	NA