

Convergent evolution of desiccation tolerance in grasses

Rose A. Marks^{1,2,3*}, Llewelyn Van Der Pas⁴, Jenny Schuster^{1,2}, Ian S. Gilman^{1,2}, and Robert VanBuren^{1,2,*}

¹ Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA

² Plant Resilience Institute, Michigan State University, East Lansing, MI 48824, USA

³ Department of Plant Biology, University of Illinois, Urbana, IL, 61801, USA

⁴ Department of Molecular and Cell Biology, University of Cape Town, Rondebosch 7701, South Africa

*Corresponding authors: Rose A. Marks, marksros@msu.edu; Robert VanBuren vanbur31@msu.edu

Abstract

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.

38 Introduction

39 Anhydrobiosis, or life without water, is rare but widely distributed across life, spanning
40 microbial, animal, and plant lineages. Plants that can tolerate desiccation in their vegetative
41 tissues are known as resurrection plants due to their dramatic ability to revive from an extremely
42 dry state (water potential < -100 MPa or relative water content < 10%)¹. Desiccation tolerance
43 likely arose in plants during the Ordovician period and is thought to have played a critical role in
44 facilitating the transition from aquatic to terrestrial environments by early land plants². These
45 ancestral mechanisms of anhydrobiosis were retained in many non-seed plants (e.g., mosses,
46 liverworts, ferns, and fern allies) and there is a high frequency of vegetative desiccation
47 tolerance among extant bryophytes and pteridophytes³. In contrast, vegetative desiccation
48 tolerance was lost, or suppressed, in the common ancestor of seed plants, presumably in a
49 tradeoff for other systems of drought avoidance and escape, such as annual life histories, water
50 transport, and retention mechanisms including stomata, vasculature, and roots⁴. Desiccation
51 tolerance then re-evolved convergently in a subset of vascular plants, likely through the rewiring
52 of ancestral anhydrobiosis pathways maintained in seeds, spores, and pollen⁵⁻⁷. The retention
53 and re-evolution of desiccation tolerance seems to have been driven by a combination of
54 selective pressures in habitats with extreme water limitation, seasonal drought, and sporadic
55 water availability⁸. Consequently desiccation tolerance is more common in some lineages than
56 others, but diverse species of resurrection plants can often be found co-occurring in tightly
57 intertwined communities on rocky outcroppings in arid tropical and subtropical regions across
58 the world^{3,9}.

59 Despite more than 500 million years of evolution and divergence across extant
60 resurrection plants, multiple biochemical and physiological mechanisms of desiccation tolerance
61 are shared across distantly related species. For example, all surveyed resurrection plants
62 accumulate small non-reducing sugars and other osmoprotectants to vitrify the cytoplasm and
63 safeguard macromolecules during drying¹⁰. Dramatic shifts in carbohydrate and lipid
64 metabolism as well as the protection (or in some cases degradation) of photosynthetic apparatus
65 are also observed in all resurrection plants during drying¹¹⁻¹⁴. All surveyed desiccation tolerant
66 plants leverage robust anti-oxidant scavenging systems, mobilize numerous intrinsically
67 disordered and protective proteins, and have specialized cell wall properties that maximize
68 flexibility and mitigate the mechanical strain of shrinkage^{3,14,15}. These broad features of
69 anhydrobiosis are largely shared across organisms and tissues, but the specific metabolic
70 pathways, regulatory networks, and activated genes are notably complex and variable among
71 species^{3,10,16} and tissues¹⁷.

72 The recurrent evolution of desiccation tolerance offers an exciting opportunity to
73 understand how complex traits evolve independently across both broad and narrow
74 phylogenetic distances. The evolution of complex traits can occur via multiple pathways^{18,19},
75 and it is often assumed that when closely related taxa evolve the same trait independently, they
76 do so by leveraging the same genetic pathways (parallelism) due to internal constraints within
77 that lineage²⁰. In contrast, when distantly related taxa evolve the same trait independently they
78 are expected to leverage divergent pathways and genes (convergence), due to contrasting
79 genetic starting points^{19,21}. However, these patterns are not always observed in nature, and
80 contradictory examples exist, where distantly related taxa exhibit independent but identical
81 mutations and closely related taxa do not²¹. The recurrent evolution of desiccation tolerance at

82 multiple phylogenetic scales provides an ideal system to untangle the mechanisms of
83 convergent and parallel evolution. An important first step towards decoding the evolutionary
84 pathways to desiccation tolerance is characterizing the extent of shared genetic adaptations,
85 overlapping pathways, and lineage specific processes across resurrection plants .

86 Desiccation tolerance has received growing research attention in recent years and
87 several resurrection plants have emerged as models for understanding this remarkable trait ²².
88 Desiccation tolerance is found in at least ten angiosperm families, and is most common in
89 Poaceae, where it evolved independently at least six times across three subfamilies and is
90 found in dozens of grass species³. Thus, the grasses are an excellent system to test if the same
91 pathways, regulatory modules, and mechanisms were recruited during the recurrent evolution of
92 desiccation tolerance. Most genomic studies of resurrection plants have investigated only a
93 single species in isolation ^{17,23–27} or tolerant and sensitive taxon comparison ^{7,28,29}, but none
94 have identified core responses shared among independent lineages of resurrection plants.
95 Here, we quantify the extent of shared mechanisms of anhydrobiosis across resurrection
96 grasses and investigate the roles of parallel mutation and convergent pathway adaptation in the
97 evolution of desiccation tolerance. We present highly contiguous genome assemblies of three
98 resurrection grasses native to Sub-Saharan Africa coupled with comprehensive gene
99 expression datasets and supporting physiological data. We leveraged comparative genomic and
100 transcriptomic approaches to investigate the evolution of desiccation tolerance in these three
101 species. We also extend these analyses to other desiccation tolerant and sensitive grasses to
102 describe a core signature that defines desiccation tolerance.

103

104

105 **Results**

106

107 ***Comparative genomics of desiccation tolerant grasses***

108 We searched for signatures of convergent evolution across three grasses in two
109 Chloridoideae subtribes representing at least two independent origins of desiccation tolerance:
110 *Microchloa caffra* Nees. in subtribe Eleusininae and *Oropetium capense* Stapf. and *Tripogon*
111 *minimus* Steud. in the Tripogoninae subtribe. These three species have overlapping
112 distributions and tend to co-occur in shallow soils on rocky outcroppings, locally known as
113 ruwari, across Sub-Saharan Africa (Figure 1a). *Microchloa caffra*, commonly known as
114 pincushion grass, is distributed from Uganda to South Africa and is the largest of the three
115 species. *Oropetium capense* is smaller and grows as densely packed tufts on exposed rock
116 surfaces. *Tripogon minimus*. is a small but loosely tufted grass that occurs in shallow soils in
117 both western and southern Africa (Figure 1a). *Microchloa caffra* plants were collected from
118 Buffelskloof Private Nature Reserve in Mpumalanga and *O. capense* and *T. minimus* were
119 collected from Swebe Swebe Private Wildlife Reserve in Limpopo, South Africa.

120 We generated reference genome assemblies for each of the three grasses using PacBio
121 HiFi data. *O. capense* and *T. minimus* are diploid with haploid genome sizes of ~195 Mb based
122 on flow cytometry, and *M. caffra* is hexaploid with a 1.25 Gb haploid genome. Sequencing reads
123 were assembled using Hifiasm ³⁰, producing near complete reference assemblies for *O.*
124 *capense* and *T. minimus* and a highly contiguous draft assembly of *M. caffra* (Table 1). Six and
125 nine of the ten chromosomes were assembled telomere-to-telomere for *T. minimus* and *O.*

126 *capense* respectively, and the remaining chromosomes were split into two contigs. The *M.*
127 *caffra* genome assembly was more fragmented, with 118 contigs spanning 968 Mb and a contig
128 N50 of 16 Mb. The monoploid genome size of *M. caffra* is 322 Mb, which is roughly 30% larger
129 than *O. capense* and *T. minimus* (237 and 223 Mb, respectively), and this expansion was driven
130 largely by DNA transposons. All three species have a similar proportion of long terminal repeat
131 retrotransposons (22-27%), but 27% of the *M. caffra* genome is composed of DNA transposons
132 compared to 12% in *O. capense* and 16% in *T. minimus* (Table 1). Despite this expansion of
133 transposons in *M. caffra*, the three Chloridoid grasses have very compact genomes compared
134 to most grasses³¹. We used the MAKER-P pipeline to annotate these three genome
135 assemblies, with RNAseq data and protein homology as evidence. The *O. capense* and *T.*
136 *minimus* genome assemblies have 28,826 and 26,527 gene models respectively, which is
137 comparable to the well-annotated model resurrection plant *Oropetium thomaeum* (L.f.) Trin
138 (28,835)^{24,32}. The *M. caffra* genome assembly has 85,245 gene models, which matches the
139 expectations for a hexaploid genome (Table 1). We assessed annotation quality using the land
140 plant (Embryophyta) dataset of Benchmarking Universal Single-Copy Orthologs (BUSCO) and
141 found between 95.3-97.1% complete proteins across the three grasses, suggesting the genome
142 assemblies are largely complete and well-annotated (Table 1).

143 We leveraged comparative genomic approaches to identify evolutionary signatures
144 associated with desiccation tolerance and enable cross-species comparisons of gene
145 expression data. The three grass genomes are largely collinear with *O. thomaeum*, and have
146 considerable conserved gene content despite some notable structural rearrangements. Seven
147 pairs of *O. thomaeum* and *O. capense* chromosomes have near perfect synteny, with
148 chromosomes 8 and 9 showing a few large-scale inversions, and a telomeric translocation on
149 chromosome 2 (Supplemental Figure 1). *Tripogon minimus* has similar macrosynteny with *O.*
150 *thomaeum*, but has no rearrangements in chromosome 8. Synteny between *M. caffra* and *O.*
151 *thomaeum* is more fragmented because of phylogenetic divergence and each *O. thomaeum*
152 region has between 2-4 homeologous regions in *M. caffra* (Supplemental Figure 2). We
153 calculated the synonymous substitution rates (Ks) between homeologous gene pairs within *M.*
154 *caffra* to date the polyploid event(s). We observed a single Ks peak of 0.13 across all
155 homeologous gene pair combinations, suggesting the autohexaploidy event occurred ~4 million
156 years ago from rapidly successive polyploidy events (Supplemental Figure 2d). Using MCSan
157 with *O. thomaeum* as an anchor, we identified 18,428 syntenic orthologs (syntelogs) shared
158 among the three grasses, as well as previously published tolerant grasses *Eragrostis nindensis*
159 Ficalho & Hierr^{28 33}. These syntelogs were used to identify patterns of gene duplication
160 associated with desiccation tolerance across grasses and as anchor points to compare
161 expression of conserved genes across species.

162 To test for convergent evolution we characterized patterns of expansion and duplication
163 in gene families with important roles in desiccation tolerance. The genomes of all sequenced
164 resurrection plants have large tandem arrays of early light induced proteins (ELIPs)³⁴, and we
165 observed this same pattern across the desiccation tolerant grasses investigated here.
166 *Oropetium capense*, *T. minimus*, and *M. caffra* all have massive tandem arrays of 39, 31, and
167 58 ELIPs respectively, compared to an average of 4 in the genomes of desiccation sensitive
168 grasses³⁴. This expansion of ELIPs is similar to other chlorophyll retaining (homiochlorophyllus)
169 resurrection plants and is generally higher than chlorophyll degrading (poikiochlorophyllus)

170 species. ELIPs are universally highly expressed in the diploid resurrection grasses *O. capense*
171 and *T. minimus* during drying, desiccation, and early rehydration, but only a subset of the ELIPs
172 in the *M. caffra* tandem arrays have desiccation induced expression (Supplemental Figure 3).
173 We used CAFE³⁵ to test for changes in the dynamics of ELIP copy number evolution across
174 land plants. We found significant increases in the rate of ELIP expansion in all desiccation
175 tolerant lineages of plants (Figure 2c). Within the grass family, ELIP expansion occurred
176 independently in subtribes Eleusininae, Sporobolinae, Eragrostidinae, and Tripogonae, but
177 *Oropetium* and *Tripogon* share a single origin of desiccation tolerance (Figure 2c). Other gene
178 families with well-characterized roles in desiccation tolerance such as late embryogenesis
179 abundant (LEAs) and heat shock proteins (HSPs) show no expansion in resurrection plants
180 based on OrthoFinder and or CAFE (Supplemental Figures 4-7).

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

190
191

192 **Searching for overlapping signatures of desiccation tolerance**

193 We collected dehydration and rehydration timecourses of *O. capense*, *T. minimus*, and
194 *M. caffra* plants under similar conditions in a climate controlled growth chamber. Plants reached
195 desiccation after ~17-20 days of natural drying, with a relative water content (RWC) < 10% and
196 photosystem II efficiency (F_v/F_m) approaching 0.0 (Figure 1b-d). RWC and F_v/F_m recovered
197 within 12 hours of rehydration in *O. capense* and *T. minimus*, but F_v/F_m took longer to recover in
198 *M. caffra* (Figure 1b). We collected gene expression data (RNAseq) at six comparable
199 timepoints of drying and recovery for each of the three species. We quantified RNA abundance
200 and gene expression patterns across the dehydration-rehydration timecourse in each species
201 individually. RNAseq reads were pseudo-aligned to the respective genomes using Salmon (v
202 1.9.0)³⁶ and normalized counts were used for all downstream analyses. In general, gene
203 expression profiles were tightly associated with the hydration status of the plants. Correlation
204 matrices and principal component analysis (PCA) show tight clustering of samples by hydration
205 status, with hydrated, desiccated, and rehydrated samples forming distinct clusters for each
206 species (Supplemental Figure 8).

207 Using RWC as a covariate, we identified genes that were significantly up- and down-
208 regulated during dehydration and rehydration processes. Both dehydration and rehydration
209 induced substantial changes in gene expression in all three desiccation tolerant grasses, with
210 35-52% of genes showing differential abundance during dehydration and 23-47% during
211 rehydration (Figure 3a and Supplemental Figure 9). *Microchloa caffra* had significantly more
212 differentially expressed (DE) genes (Supplemental Figure 9) given its hexaploidy, but a lower
213 proportion of DE genes compared to the other two grasses (Figure 3a). Broadly, desiccation

214 and rehydration had inverse expression profiles, and most genes that increased in abundance
215 during dehydration, dissipated during rehydration and vice versa (Supplemental Figure 9).
216

217 To enable comparisons across species, we leveraged the 18,428 conserved syntelogs
218 and searched for overlapping patterns in the expression of these shared genes. There was
219 considerable overlap in gene expression across the three focal resurrection grasses, with ~18-
220 24% of all DE syntelogs showing similar expression across species (Figure 3b and
221 Supplemental Figure 10a). The proportions of DEGs shared across the three resurrection
222 grasses for both up- and down-regulated genes was considerably more than observed in
223 previous studies or expected due to chance. In order to differentiate between desiccation
224 tolerance mechanisms and more general drought responses, we identified the extent of shared
225 syntelog expression between these resurrection grasses and the desiccation sensitive species
226 *E. tef*, which was sampled along a similar dehydration timecourse in a previous study²⁸. There
227 was considerable overlap in syntelog expression between the resurrection grasses and *E. tef*
228 (Supplemental Figure 11), reflecting deeply conserved mechanisms of drought tolerance in
229 grasses. We also detected a large set of genes that were expressed exclusively in the
230 resurrection grasses, which likely play desiccation specific roles to survive anhydrobiosis.
231 Species-specific expression patterns are also evident, particularly for *E. tef*.

232 Dimensionality reduction and co-expression analyses also point towards parallel
233 mechanisms of desiccation tolerance in resurrection grasses. Samples clustered primarily by
234 hydration status and secondarily by species in PCA (Figure 3c,d and Supplemental Figure 12).
235 We defined co-expression modules for each species and screened for shared network level
236 responses within co-expressed genes. High confidence modules were defined for each species,
237 and we grouped these into three broad classes based on the expression pattern of each
238 module: 1) elevated expression in hydrated conditions, 2) elevated expression during
239 dehydration, and 3) elevated expression during rehydration (Figure 4d). We identified
240 substantial overlap in gene module conservation with elevated expression during dehydration,
241 but less overlap in modules with high expression during rehydration (Figure 4a). We then
242 identified enriched gene ontology (GO) terms for each co-expression module and performed
243 hierarchical clustering on the enrichment p-values of GO terms. Modules clustered by their
244 expression profile rather than species identity, suggesting that hydration status is more
245 predictive of gene expression than species identity (Figure 4c) and pointing towards a shared
246 signature of desiccation tolerance in resurrection grasses.

247
248

249 ***Functional characterization of the shared signatures of desiccation tolerance***

250 Our analyses of syntelog expression tested for ancestral conservation and parallelism,
251 but it is also possible that different lineages of resurrection plants may utilize similar metabolic
252 strategies for achieving desiccation tolerance but through divergent genes and pathways. To
253 investigate this possibility, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) to
254 assign each gene to a predicted enzymatic function and metabolic pathways and compared the
255 overlap in these functional predictions across species. We detected substantially higher overlap
256 in KEGG terms across species (~30-40%) compared to DE syntelogs (only 18-24%) (Figure 5b
257 and Supplemental Figure 10c). The increased similarity at a metabolic level suggests that while

258 these species do not always leverage parallel gene copies, they induce similar metabolic
259 mechanisms to survive anhydrobiosis, providing evidence of convergence across species.

260 We further investigated the functional roles of shared gene expression via GO
261 enrichment and KEGG analyses. We found that many hallmarks of desiccation tolerance were
262 shared across the three resurrection grasses, including the controlled downregulation of
263 photosynthesis and rapid induction of protective mechanisms. Enriched GO terms during
264 dehydration were related primarily to signaling and stress responses (e.g., stress perception
265 and ROS scavenging activities), developmental regulation (e.g., photoperiodism and
266 germination processes), cellular reorganization (e.g., lipid droplet formation, vesicle fusion,
267 endocytosis), and modifications to transcription and translation (e.g., RNA modifications,
268 splicing, and protein degradation). In contrast, enriched GO terms during rehydration are related
269 to photosynthesis and metabolism (e.g., fructose biosynthesis, cellulose biosynthesis, and light
270 harvesting), pigment metabolism (e.g., chlorophyll biosynthesis and anthocyanin metabolism),
271 protein modification (e.g., protein phosphorylation and proteolysis), and some residual stress
272 response (e.g., response to cold and non photochemical quenching) (Figure 5a). Hierarchical
273 clustering of enriched GO terms also highlighted the inverse relationship between dehydration
274 and rehydration process (Figure 5a and Supplemental Figure 10d).

275 To differentiate between desiccation tolerance mechanisms and more typical drought
276 tolerance responses, we compared the enriched GO terms for DE syntelogs uniquely induced in
277 the resurrection grasses vs. those shared with desiccation sensitive *E. tef* (Supplemental Figure
278 11). Many of the classic stress response terms were shared across all species, reflecting deeply
279 conserved responses to water deprivation. For example, all species showed metabolic arrest
280 during drying with a particular emphasis on photosynthetic shut down. All species exhibited an
281 increase in classic stress response terms such as response to heat, response to water
282 deprivation, response to hydrogen peroxide, and sucrose metabolic process. These processes
283 represent core mechanisms of water deficit tolerance that likely form the foundation of
284 desiccation tolerance. Building on this foundation, resurrection grasses appear to activate
285 additional processes that enable more extreme resilience. For example, the resurrection
286 grasses showed unique activation of nucleic acid processes including mRNA export, regulation
287 of chromosome condensation, and mRNA transcription by RNA polymerase II suggesting
288 greater overall regulation of transcription and translation. Several terms associated with the
289 circadian rhythm and hormonal signaling were also uniquely upregulated in the resurrection
290 grasses, indicating a central role of circadian clock processes in preparing for desiccation. The
291 resurrection grasses exhibited a unique downregulation of tissue and cellular developmental
292 processes, implying a tightly regulated cessation of metabolism at later stages of drying. Taken
293 together, this suggests that resurrection grasses build on a shared foundation of drought
294 tolerance to achieve desiccation tolerance via a highly organized shift in cellular processes.

295 KEGG annotations revealed characteristic desiccation tolerance mechanisms shared
296 across resurrection grasses. Metabolic pathways associated with photosynthetic energy
297 metabolism were significantly down regulated in all three grasses. Interestingly, we observed an
298 increase of malate to pyruvate catalysis with concomitant regeneration of NADPH, which could
299 be related to NADPH's REDOX potential for antioxidant enzymes such as glutathione
300 reductase. We also detected noticeable changes to carbohydrate and energy metabolism,
301 including a shift towards the production of raffinose and stachyose under dehydrating conditions
302 as seen in other resurrection plants (reviewed in ¹⁰). Central carbohydrate metabolism appeared

303 operational suggesting that at low water contents, other solvents, such as natural deep eutectic
304 solvents within the mitochondria may facilitate glycolysis, the TCA, and electron transport³⁷.
305 Amino acid metabolism favored degradative pathways with an increase in endoplasmic
306 reticulum-mediated ubiquitination and proteolysis, which could be serving a glucogenic role by
307 converting amino acids to pyruvate, or by generating an available amino acid pool for the rapid
308 assembly of thermo- and osmoprotective proteins. While amino acid metabolic pathways were
309 generally down-regulated, a few important pathways including glutathione metabolism were up-
310 regulated. Reduced glutathione (GSH) exerts numerous effects in the cell³⁸ from interaction
311 with hormones to acting as direct ROS quencher, and maintaining a steady supply of GSH is a
312 feature all three resurrection grasses share. Lipid metabolism showed a shift towards the
313 production of glycerolipids and glycerophospholipids which likely supports triacylglycerol
314 phosphatidylcholine production. The accumulation of phosphatidylcholine may further lead to
315 phosphatidic acid synthesis, which has been implicated in numerous plant processes from
316 signaling to storage³⁹⁻⁴¹. Pathways involved in the transcription and translation of genetic
317 information also showed an up-regulation of transcription factors, RNA polymerase, and
318 spliceosome activity, suggesting that active transcription and RNA processing are still occurring.
319 However, we observed substantial downregulation of ribosome activity, suggesting that RNA is
320 either differentially translated or delayed. Upon rehydration, up-regulated processes involved in
321 overall resumption of normal metabolic activity such as several photosystem I and II proteins,
322 light harvesting complexes, starch synthesis, and cell wall remodeling such as xyloglucan O-
323 acetyltransferase, expansin, and pectinesterase were observed.

324

325 ***Desiccation tolerance mechanisms are broadly conserved across grasses***

326 Desiccation tolerance evolved independently in at least four subtribes of Chloridoideae
327 (Eleusininae, Eragrostidinae, Sporobolinae, and Tripogoninae; Figure 2c), and we integrated
328 comparable desiccation and rehydration expression datasets from additional species to test for
329 patterns of convergence across grasses more broadly. Building on our detailed comparisons
330 across the three study species, we expanded our analysis to include publicly available RNAseq
331 samples from desiccation tolerant *O. thomaeum*⁴² and *E. nindensis*²⁸, leveraging syntelogs for
332 cross-species comparisons. Similar to the three species comparisons described above,
333 dimensionality reduction across the five species generally separated samples by hydration
334 status along PC1 and PC2 (Supplemental Figure 12). While PCA provided some degree of
335 separation; residual heterogeneity, experimental differences, noise, or species level differences
336 in the datasets might have obscured underlying conserved biology. To account for this, we
337 employed a topological data analysis (TDA) approach to discern the underlying structure of the
338 expression datasets. We utilized the Mapper algorithm, which condenses the dataset into a
339 scalable, navigable representation. The Mapper algorithm is particularly well-suited for genome
340 scale analyses, as the underlying datasets are often characterized by high dimensionality and
341 sparsity⁴³. For our gene expression data, we constructed Mapper graphs using a "stress lens"
342 with the well-watered condition as a reference point. This model represents the baseline for
343 gene expression and we quantified the residuals or deviation of each sample from the baseline,
344 which represent the degree of water stress or recovery.

345 The resultant Mapper graph illustrates a clear topological shape that delineates
346 desiccation processes across grasses (Figure 6). Each node on the graph represents a cluster

347 of similar RNAseq samples, and the node color depicts the identity of samples within that
348 cluster. Connections between nodes signify shared samples among the intersecting clusters.
349 These graphs reveal a compelling topological depiction of the gene expression variations
350 induced by water stress across different species. Similar topology was observed for both
351 targeted comparison of the three focal species (Figure 6a and b) and for the larger dataset
352 including *E. nindensis* and *O. thomaeum* (Figure 6c and d). In both instances, clear delineation
353 between samples of different hydration statuses are evident, while the species are intermixed.
354 We then added the desiccation sensitive sister species *E. tef* in a final analysis (Supplemental
355 Figure 13), which revealed a similar topology across all species with notable gaps in *E. tef*.
356 Broadly, this supports our finding that similar ancestral mechanisms are being recruited for
357 foundational drought tolerance mechanisms, which are enhanced in resurrection plants via the
358 independent recruitment of specific desiccation tolerance pathways.

359
360

361 ***Species specific mechanisms underlying desiccation tolerance***

362 Despite the considerable overlap in gene expression across all three focal species,
363 species-specific processes were also evident. In *M. caffra*, unique antioxidant responses were
364 induced including glutathione biosynthetic processes, glutamate decarboxylation, and L-
365 ascorbic acid biosynthesis. Other processes enriched uniquely in *M. caffra* included seed-
366 related terms such as seed oil body biogenesis and seed maturation. Several GO terms
367 associated with phytohormones were also uniquely induced in *M. caffra*, including overall
368 ethylene responses such as S-adenosylmethionine metabolic process, ethylene-activated
369 signaling pathway, and response to 1-aminocyclopropane-1-carboxylic acid, suggesting that
370 hormonal regulation might be exerting an effect on the partial breakdown of thylakoids and
371 photosynthetic machinery as seen in classical senescence⁴⁴. *Microchloa caffra* also exhibited
372 unique lipid, sphingolipid, riboflavin, and selenocompound metabolism, as well as
373 sesquiterpenoid and tripenoid biosynthesis. *Microchloa caffra* was the only species to have
374 multiple pathways involved in signal transduction up-regulated including phospholipase D and
375 calcium signaling. Uniquely down-regulated processes in *M. caffra* appear to center around
376 arresting growth and development, such as phototropism, gravitropism, leaf and root
377 morphogenesis, cell wall biogenesis, and regulation of auxin polar transport. There was also a
378 down-regulation of general amino acid-tRNA aminoacylation and nitrogen fixation and
379 assimilation.

380 *Oropetium capense* had fewer uniquely enriched processes compared to *M. caffra*, but
381 some notable patterns were detected. Uniquely upregulated processes in *O. capense* centered
382 around histone H3 and H4 acetylation, histone H3-K9 demethylation, and histone H2B
383 ubiquitination. The relative degree of acetylation of histones is directly related to the openness
384 of chromatin which impacts transcription in specific drought-responsive genes⁴⁵. Histone
385 demethylation⁴⁶ and H2B ubiquitination also regulate drought responsive genes⁴⁷.
386 Interestingly, terms associated with chloroplast mRNA processing, poly(A)+ mRNA export from
387 the nucleus, ribosome assembly, and regulation of translation were also upregulated in *O.*
388 *capense*, suggesting continued translation and active processing of mRNA from both the
389 chloroplast and nucleus, presumably through increased transcriptional regulation due to histone
390 modifications. *Oropetium capense* exhibited unique up-regulation of C5-branched dibasic

391 metabolism and down-regulation within galactose metabolism. Uniquely down-regulated
392 processes in *O. capense* were minimal, but included regulation of salicylic acid metabolic
393 process and auxin polar transport. Monoterpenoid biosynthesis was up-regulated for *O.*
394 *capense* and fatty acid degradation and steroid hormone biosynthesis was down-regulated.
395 Ferroptosis, an iron-dependent form of programmed cell death, was exclusively down-regulated
396 in *O. capense*.

397 *Tripogon minimus* also had fewer species-specific processes compared to *M. caffra*.
398 Uniquely up-regulated processes were centered around response to oxidative stress,
399 peroxisome organization, and removal of superoxide radicals. *Tripogon minimus* was the only
400 species to show upregulation of anthocyanin-containing compound biosynthetic process which
401 is a typical response seen in the homoiochlorophyllous resurrection plants⁴⁸. Similar to the
402 other two species, regulation of auxin-mediated signaling pathways were down-regulated as
403 was cellular response to salicylic acid stimulus. Other processes centered around mismatch
404 repair, chloroplast RNA processing, ribosome biogenesis, and plastid transcription.
405 Phosphonate and phosphinate, taurine and hypotaurine, and D-amino acid metabolism were
406 exclusively down-regulated in *T. minimus* whereas retinol metabolism was up-regulated.

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

415

416

417 Discussion

418 Our data suggest that the repeated evolution of desiccation tolerance within grasses
419 occurred via both parallel adaptations in the same ancestral genes and complementary
420 modifications to analogous pathways. We find evidence that core mechanisms of desiccation
421 tolerance are shared across resurrection grasses, and are supplemented with species-specific
422 adaptations. Many of these mechanisms overlap with typical drought responses and it is likely
423 that the evolution of anhydrobiosis builds on deeply conserved responses to water deficit shared
424 across all plants. Phenotypic and metabolic similarities in anhydrobiosis mechanisms have been
425 observed for decades, but the evolutionary pathways of convergence and parallelism have been
426 obscured by a lack of systems-level data and inconsistencies in experimental procedures⁴⁹.
427 Here, we leveraged large scale genomic and transcriptomic datasets in a replicated and
428 standardized framework to characterize signatures underlying the recurrent evolution of
429 desiccation tolerance within chloridoid grasses.

430 The adaptations required for desiccation tolerance appear to be sufficiently narrow, such
431 that not any organism can, or will, evolve desiccation tolerance⁴. The physiological changes
432 that occur during the final stages of desiccation are dramatic and specialized biochemistry and
433 molecular mechanisms are required to protect the cellular macromolecules for life without water.
434 Achieving anhydrobiosis requires tight coordination and orchestration of multiple physiological

435 processes, and there may be only a few trajectories to evolve this trait. However, desiccation
436 tolerance mechanisms overlap considerably with typical drought responses and many plants
437 possess the basic cellular machinery required to achieve desiccation tolerance²⁸. Desiccation
438 tolerance is likely an ancestral adaptation in plants that evolved during terrestrialization,
439 subsequently formed the basis of seed pathways, and was later rewired again in vegetative
440 tissues^{5,25,42,50}. While previous studies have found surprisingly little overlap in gene expression
441 across desiccation tolerant plants^{51,52}, our data suggest that the repeated evolution of specific
442 genetic, biochemical, and physiological traits required for anhydrobiosis, are highly convergent
443 and build on more broadly conserved water deficit responses.

444 Convergence is thought to be driven primarily by exposure to external selective
445 pressures that lead to the same emergent phenotype, while parallelism is thought to be
446 impacted more by internal constraints of the system¹⁸ through independent mutations in the
447 same ancestral gene^{19,21}. Because anhydrobiosis has evolved independently in both distantly
448 and closely related taxa, it is an ideal system in which to explore the roles of convergent and
449 parallel evolution. Numerous other independently evolved traits such as C4 and CAM
450 photosynthesis are highly complex, making their repeated evolution surprising⁵³ and difficult to
451 characterize. In the case of C4 photosynthesis, both mutations in the same genes and
452 recruitment of unique pathways occurred in distantly related lineages to enable the emergent C4
453 phenotype⁵⁴. Desiccation tolerance is similarly complex, involving the synchronized
454 orchestration of numerous pathways and genes, and it is likely that both external pressures (e.g.
455 selection in extremely xeric habitats) and internal constraints (lineage specific predispositions)
456 play a role in the recurrent evolution of desiccation tolerance. Here, we detected signatures of
457 both processes and identified far more overlap in gene expression across resurrection grasses
458 than expected by chance or detected in previous studies^{29,51}. The observed expansion of ELIP
459 tandem arrays coupled with activation of similar metabolic pathways driven by different gene
460 sets, suggests that both parallel and convergent processes contribute to the recurrent evolution
461 of desiccation tolerance in grasses.

462 Our systems-level analyses add to the growing literature on the mechanisms of
463 desiccation tolerance, and many of the patterns observed here corroborate previous findings
464^{17,25,29,34,51}. We show that desiccation induces a major and reversible shift in gene expression
465 where normal growth and development are halted and numerous protective mechanisms are
466 induced^{13,14,55-57}. Gene expression coalesced around a signature desiccation response during
467 drying with all three species initiating parallel processes⁵⁸. The resumption of species specific
468 processes related to growth and development was evident upon rehydration. The shared
469 pathways of anhydrobiosis observed in these grasses pull on the deeply conserved architecture
470 of drought tolerance coupled with convergent and parallel mutations that provide the necessary
471 protection to survive extreme desiccation. This reflects the relatively narrow set of regulatory
472 networks and pathways in plants that can enable the evolution of desiccation tolerance, but also
473 hints at multiple evolutionary paths to anhydrobiosis

474

475

476 **Methods**

477

478 *Field collections, plant growth, and maintenance*

479 Plants for the current study were collected from two research sites in South Africa:
480 Buffelskloof Nature Reserve in Mpumalanga (-25.30229 S, 030.50631 E) (*Microchloa caffra*)
481 and Swebe Swebe Private Wildlife Reserve in Limpopo (-23.7949 S, 028.0705 E) (*Oropetium*
482 *capense* and *Tripogon minimus*). Voucher specimens of each species were collected, pressed,
483 and deposited at the National Herbarium of South Africa in Pretoria (specimen numbers:
484 PRE1004810-0, PRE1004793-0, and PRE1004794-0). Seeds of each species were also
485 collected and transported to Michigan State University under United States Department of
486 Agriculture (USDA) permit #537-22-37-10071 and according to the specifications in a Material
487 Transfer Agreement established between Drs. Jill M. Farrant, Robert VanBuren, and Rose A.
488 Marks. Seeds were cold stratified at 4°C for two weeks and then germinated on our standard
489 propagation mix (50:50 sure-mix:redi-earth) and grown in a climate controlled growth chamber
490 with a 16 hour photoperiod and internal temperature of 28/18°C. Six weeks after germination,
491 individual seedlings were transplanted into separate pots and grown to maturity and a single
492 plant (genetic line) of each species was selected for downstream experimentation. Seeds of
493 each genetic line were collected, cold stratified, and ~70 seedlings from each species were
494 germinated. One seedling from each species was used for genome sequencing and was
495 transplanted into a larger pot. The remaining seedlings were used for the desiccation and
496 rehydration timecourses experiments and three seedlings were transplanted into 4" pots. The
497 three plants in each pot were pooled during sampling and treated as a single biological
498 replicate. These plants were grown for another two weeks prior to experimental treatments,
499 during which time they were maintained in constantly hydrated conditions in a growth chamber
500 set to the conditions described above.

501 502 *Dehydration treatment and sample collection*

503 After ~8 weeks of growth, plants were subjected to dehydration treatment. Prior to
504 treatment, any emerging reproductive tissues (e.g., panicles) were removed from plants. To
505 initiate dehydration treatment, plants were watered to full soil saturation and each pot was
506 weighed to ensure consistency across replicates. Water was then withheld until plants became
507 completely desiccated (between 2 and 3 weeks depending on the species). Plants were
508 sampled at targeted hydration states during the process of dehydration, including well watered,
509 partially dehydrated, fully desiccated, and rehydrated. We used visual cues to direct our
510 sampling and sampled plants at the first signs of visible leaf curling, partial pigmentation, deep
511 pigmentation, and full desiccation and validated the hydration status of tissues by measuring
512 relative water content (RWC). Plants were then rehydrated through a combination of watering
513 from the base and misting the aerial portions to simulate natural rainfall and sampled 24 and 48
514 hours post rehydration. We aimed to sample plants at biologically relevant water contents and
515 therefore directed our sampling by use of visual cues rather than a set number of hours. This
516 allowed us to compensate for different drying rates across species and plants due to subtle
517 variation in size, water use efficiency, and relative humidity in the growth chamber.

518 At each timepoint, we measured the photosynthetic efficiency (F_v/F_m) and RWC, and
519 harvested tissue for RNAseq. Briefly, F_v/F_m was measured on dark adapted leaves using a Opti-
520 Sciences OS30p+ chlorophyll fluorometer with the default test parameters. Relative water
521 content was measured using a set of 10-15 representative leaves from each pot / biological
522 replicate. Leaf mass was weighed immediately after collection (fresh weight), again after 48

523 hours submerged in dH₂O in darkness at 4°C (turgid weight), and finally after 48 hours in a 70°C
524 drying oven (dry weight). RWC was calculated as (fresh weight - dry weight)/(turgid weight - dry
525 weight). Tissue for RNAseq was collected by harvesting all the vegetative tissue from each pot
526 and flash freezing in liquid nitrogen. Tissue samples were stored in a -80°C freezer prior to
527 downstream processing.

528

529 *RNA extraction and sequencing*

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

540

541 *High molecular weight DNA extraction, and sequencing*

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

549

550 *Genome assembly*

551 We used flow cytometry to estimate genome sizes (2C DNA values) for the three
552 grasses. Healthy leaf tissue was harvested from each genotype. Nuclei were isolated and
553 stained according to standard protocols. The stained nuclei were then run on a BD Accuri™ C6
554 Plus Flow Cytometer at Plantploidy.com. *Hosta plantaginea* was used as an internal reference.
555 We built reference genomes for each species using high fidelity (HiFi) PacBio long read data. In
556 total, 70.1 Gb of HiFi reads were generated for *M. caffra*, 15.9 Gb for *O. capense*, and 20.2 Gb
557 for *T. minimus*, representing 56, 82, and 103 x genome coverage for each species respectively.
558 K-mer analysis revealed that *O. capense* and *T. minimus* have low within genome
559 heterozygosity and *M. caffra* is a highly heterozygous autopolyploid⁵⁹. PacBio reads were
560 assembled using hifiasm (v 0.18)^{30,60} with default settings for *O. capense* and *T. minimus* and
561 the number of haplotypes was set to 6 for *M. caffra* (flag: --n-hap 6). The resulting assemblies
562 were highly contiguous with six and nine of the ten chromosomes assembled telomere to
563 telomere for *T. minimus* and *O. capense* respectively, and 118 contigs across 968 Mb with an
564 N50 of 16 Mb for *M. caffra* (Table 1). Raw assemblies were filtered for non-plant contigs using a
565 representative microbial database with BLAST⁶¹. Full length chloroplast and mitochondrial

566 genomes were identified and retained, and any additional partial or rearranged organelle
567 genomes were removed.

568

569 *Genome annotation*

570 A library of repetitive elements was constructed for each of the three grass genomes
571 using the EDTA package (v2.0.0)⁶². EDTA comprehensively identifies DNA-based transposable
572 elements using HelitronScanner⁶³, and LTR retrotransposons using LTR_FINDER⁶⁴ and
573 *LTRharvest*⁶⁵. Protein coding genes were annotated using the MAKER-P pipeline (v2.31.10)⁶⁶
574 with the following sets of input data for training. Transcript evidence was generated using the
575 dehydration-rehydration timecourse RNAseq data from leaf tissue of each species described
576 below. Raw RNAseq reads were quality trimmed using fastp (v 0.23)⁶⁷ and aligned to the
577 unmasked genomes using the splice aware alignment program STAR (v2.6)⁶⁸. A set of non-
578 overlapping transcripts was identified from the aligned data using StringTie (v1.3.4)⁶⁹ with
579 default parameters. The resulting gff files were used as transcript evidence for MAKER. The
580 same protein evidence was used as training for each of the three grasses and this includes the
581 full annotations of *Oryza sativa*⁷⁰, *Arabidopsis thaliana*⁷¹, *Oropetium thomaeum*^{24,32}, and
582 *Eragrostis tef*³³. These datasets were used as input for MAKER and we utilized SNAP⁷² and
583 Augustus (version 3.0.2)⁷³ for *ab initio* gene prediction, performing two rounds of iterative
584 training to refine our models. To filter out repetitive element-derived proteins, we used BLAST
585 using a non-redundant transposase library against the raw gene models produced by MAKER.
586 We assessed the completeness of our assembly using the plant-specific embryophyte set of
587 Benchmarking Universal Single-Copy Orthologs (BUSCO v.2)⁷⁴. These high-confidence gene
588 models were used for all downstream analyses.

589

590 *Comparative genomics*

591 The three desiccation tolerant grass genomes were compared to each other and other
592 Chloridoid grasses using the MCScan toolkit (v1.1)⁷⁵ implemented in python
593 [[https://github.com/tanghaibao/jcvi/wiki/MCscan-\(Python-version\)](https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version))]. Syntenic orthologs were
594 identified across the three focal species, *E. nindensis*, *E. tef*, and *Oropetium thomeaum* using
595 the chromosome-scale *O. thomeaum* genome as an anchor. Syntenic blocks were identified
596 using gene models aligned using LAST with a minimum of five overlapping syntenic genes. The
597 macrosyntenic dot plots, histograms of depth, and microsynteny plots were generated using the
598 python version of MCScan. A set of 18,428 conserved syntenic orthologs across all six
599 desiccation tolerant grasses was created and used for downstream comparative genomic and
600 cross-species transcriptomic analyses. We identified orthologous genes across a subset of 33
601 land plant species to search for patterns of gene family expansion in desiccation tolerant
602 lineages as well as for downstream comparative genomic analyses. We included the following
603 species with desiccation tolerant species highlighted: *Ananas comosus*, *Arabidopsis thaliana*,
604 *Brachypodium distachyon*, *Eleusine coracana*, *Eragrostis curvula*, *Eragrostis nindensis* (DT),
605 *Eragrostis pilosa*, *Eragrostis tef*, *Hordeum vulgare*, *Lindernia brevidens* (DT), *Lindernia*
606 *subracemosa*, *Microchloa caffra* (DT), *Marchantia polymorpha* (DT), *Medicago truncatula*,
607 *Oropetium capense* (DT), *Oryza sativa*, *Oropetium thomaeum* (DT), *Physcomitrium patens*
608 (DT), *Sorghum bicolor*, *Setaria italica*, *Selaginella lepidophylla* (DT), *Solanum lycopersicum*,
609 *Selaginella moellendorffii*, *Sporobolus pyramidalis*, *Sporobolus stapfianus* (DT), *Setaria viridis*,

610 *Triticum aestivum*, *Tripogon minimus* (DT), *Vitis vinifera*, *Xerophyta viscosa* (DT), *Zostera*
611 *japonica*, *Zostera marina*, and *Zea mays*. Proteins were clustered into orthologous groups using
612 Orthofinder (v2.2.6)⁷⁶ with default parameters. For the orthogroup enrichment analysis, we
613 calculated a Z-score for each species within each orthogroup, compared it to a normal
614 distribution to obtain a p-value, and then adjust these p-values using the Benjamini and
615 Hochberg procedure to get q-values. We then searched for statistically enriched orthogroups
616 across all of the sequenced desiccation tolerant grasses. Using this approach, we identified
617 between 486 and 8,863 enriched orthogroups in the 33 species we included in our analysis, and
618 found none that are conserved across all desiccation tolerant grasses outside of ELIPs.

619

620 *ELIP gene family evolution*

621 To test the hypothesis that the ELIP gene family expansions are associated with the
622 evolution of desiccation tolerance, we used CAFÉ (v 5.1)³⁵, which analyzes changes in gene
623 family size in a phylogenetic framework. The input tree was created from the amino acid
624 sequences from 36 land plant species, with a focus on Chloridoid grasses (ELIPs count
625 phylogeny figure; Supplemental Table 1). Sequences were first clustered using Orthofinder (v
626 2.4.1)⁷⁶, filtered to remove any orthogroups that did not contain all taxa, and aligned using
627 MAFFT (v 7.305b)⁷⁷. No single copy orthologs were found containing all taxa for species tree
628 construction. Instead, we pruned gene trees and alignments to the largest subtree containing
629 unique taxa using PhyloPyPruner (v 1.2.4) (<https://gitlab.com/fethalen/phylopypruner>); where
630 paralogs were monophyletic within a species, we randomly pruned all but one sequence prior to
631 extracting the largest subtree. The resulting pruned gene trees and alignments were further
632 filtered to remove any trees no longer containing at least 19 taxa. This final set of 195
633 alignments were concatenated and used to construct a phylogeny using IQ-TREE (v 2.3.0)⁷⁸
634 and time calibrated fast least-squares dating⁷⁹.

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

642

643 *Transcriptomic analyses*

644 RNA sequencing reads were processed following a pipeline developed by the VanBuren
645 Lab (<https://github.com/pardojer23/RNAseqV2>). Briefly, sequence read quality was assessed
646 with fastQC (v 0.23) and reads were trimmed with trimmomatic (v 0.38)⁸⁰ to remove adapters
647 and low quality bases. Trimmed reads were sudo-aligned to reference genomes using Salmon
648 (v 1.9.0)³⁶, and the resulting quantification files were processed with tximport (v 3.18)⁸¹ and to
649 generate normalized expression matrices of transcripts per million (TPM). A Principal
650 Component Analysis (PCA) was used to visualize replicate and sample relationships within
651 each species using the respective TPM expression values. A cross-species PCA was performed
652 using the TPM matrix of conserved syntenic orthologs across all species. To effectively quantify
653 gene expression while acknowledging the complexities introduced by polyploidy, we summed

654 the expression levels of all homeologs in *E. nindensis*, *E. tef*, and *M. caffra* to obtain a single
655 gene expression value to enable interspecies comparisons. This approach is grounded in the
656 logic that a unified expression value not only simplifies the analysis but also encapsulates
657 potential functional diversifications among homeologs. This methodology has been applied and
658 validated in our previous research^{17,29,82,83}. We then computed the Z-score of the TPM for each
659 syntenic ortholog in each sample and ran PCA on the matrix of Z-scores of each syntenic
660 ortholog across all timepoints and species.

661

662 *Differentially expressed genes*

663 Differentially expressed genes (DEGs) were identified independently for each species
664 with DESeq2 R package (v 1.42.0)⁸⁴. Briefly, transcript abundance estimates from Salmon were
665 imported into DESeq2 using tximport to generate counts matrices. Count matrices were
666 normalized, and hierarchical clustering was conducted for basic quality control and visualization
667 of relationships across experimental timepoints and biological replicates. We tested multiple
668 models for differential expression in DESeq2, including models that identified DEGs by pairwise
669 comparisons of each timepoint against well-watered, and models that used the continuous
670 variables of RWC or F_v/F_m as covariates. DEGs identified by pairwise comparisons were
671 summarized into a nonredundant list of up and down regulated genes during dehydration and
672 rehydration. DEGs identified using the continuous variables are based on a significant linear
673 association (positive or negative) with RWC or F_v/F_m . When identifying DEGs, we included the
674 term “process” in our model to differentiate between dehydration and rehydration processes. To
675 select the best performing model, we quantified similarities and differences in the number and
676 identity of DEGs defined by each model. There was a high degree of overlap in genes identified
677 by all three models. Ultimately, we selected the model based on RWC because it performed
678 well and is easily comparable across experiments regardless of sampling time, consistency
679 across replicates, or differences in experimental design. These analyses produced species-
680 specific lists of DEGs during dehydration and rehydration with significant (FDR adjusted P-value
681 <0.05) associations with RWC. Log2foldchange values are calculated for one unit change in
682 RWC.

683 To gain insight into possible similarities and differences among the study species, we
684 looked at the overlap in DE syntenic orthologs. To do so, we used venn diagrams to identify the
685 shared syntelogs in up- and down-regulated genes during both dehydration and rehydration
686 across the three study species. We then compared the observed proportion of overlapping
687 DEGs in each category to the proportion of genes expected to overlap by chance (assuming
688 independent draws), and tested if these were significantly different using Fisher’s exact test.
689 This analysis was then extended to include DEGs identified in the desiccation sensitive sister
690 species *E. tef* to distinguish between typical drought vs. pure desiccation responses. We then
691 conducted targeted analyses to look at the functional roles of DE syntelogs that were uniquely
692 shared across the three resurrection species vs. those that were common with *E. tef*. We also
693 investigated the functional signatures of differentially abundant transcripts that were unique to
694 each species. To do so we pulled the lists of DE syntelogs that were only found in one of the
695 focal species.

696

697 *Functional annotation of DEGs*

698 We annotated DE syntelogs with KEGG and GO terms to describe generalized
699 metabolic and cellular processes responses shared across in the three study species. KEGG
700 annotations were generated for each species using BLASTKoala
701 (<https://www.kegg.jp/blastkoala/>) on the complete set of annotated peptide sequences. These
702 KEGG terms were then assigned to syntenic orthologs and DE KEGG terms that were shared
703 across all three species during dehydration and rehydration were identified and plotted using
704 venn diagrams. These shared DE KEGG terms were used to generate metabolic pathway maps
705 using the reconstruct function of the KEGGmapper tool
706 (<https://www.genome.jp/kegg/mapper/color.html>) for up- and down-regulated terms in
707 dehydration and rehydration. This returned a list of syntelogs per metabolic pathway and Brite
708 descriptions. The difference between the number of syntelogs assigned to each pathway for up
709 and down regulated genes was computed. The list was then sorted to determine pathways that
710 were primarily up or down regulated. Next, the assigned KO numbers were paired with
711 quantitative data on gene expression to identify which pathways were active at various
712 timepoints. To summarize patterns, genetic information processes (transcription, translation,
713 folding, sorting, and degradation, replication and repair, and processing in viruses) were
714 grouped together. Similarly Environmental information processing (membrane transport and
715 signal transduction) and cellular processes (transport and catabolism, cell growth and death,
716 cellular community- prokaryote and eukaryotes, and cell motility) were grouped together.
717 Pathways assigned to Organismal systems and Human disease were ignored. KEGG
718 annotation is not without its limits as single KEGG identifiers can be present in multiple
719 pathways.

720 GO terms were assigned through homology with the well annotated genome of sister
721 species *O. thomaeum*²⁶. This was done through a BLASTP (v 2.14.0)⁸⁵ search of all *O.*
722 *thomaeum* protein sequences against the protein sequences of each study species. Parameters
723 were set to return the single best match for each peptide and an e-value cutoff of 1e-10. We
724 assigned the GO terms from *O. thomaeum* to the homologous genes in our target species. We
725 then used TopGO R package (v 2.54.0) to identify significantly enriched GO terms (P-
726 value<0.05) within sets of DEGs for up- and down-regulated genes during dehydration and
727 rehydration in each target species and for the different sets of overlapping and unique syntelogs
728 identified via cross-species comparisons.

729 *Co-expression analyses*

731 To complement the above analyses, we generated co-expression networks using
732 Weighted Gene Co-expression Network Analysis (WGCNA) R package (v1.7)⁸⁶. While DE
733 analyses can be informative to identify and describe overarching patterns and large shifts in the
734 data, more nuanced patterns of gene expression can be obscured. To investigate the more
735 subtle temporal changes in gene expression, we used co-expression analyses to identify
736 modules of co-expressed genes for each species.

737 For each species, we created a signed co-expression network using WGCNA. Each
738 dataset was filtered to remove genes with no expression. To construct a weighted co-
739 expression network, we determined a soft thresholding power for each dataset. This power was
740 chosen to satisfy WGCNA's assumption that a weighted co-expression network is scale-free. An
741 adjacency matrix, representing the strength of connections between genes in the network, was

742 constructed for each network using the soft thresholding power. For module detection, this
743 matrix was then converted to a topological overlap matrix (TOM) and hierarchal clustering was
744 used on the TOM to group genes into modules based on similar expression patterns.
745 Additionally, we calculated connectivity of each gene within its network and its assigned module
746 using WGCNA's network analysis functions.

747 We identified shared and species-specific co-expressed genes using UpSet plots⁸⁷. For
748 all co-expressed genes, we identified the syntenic orthologs and computed the overlap across
749 species. We combined all modules for a species that showed increased expression during
750 dehydration, during rehydration, and under non-stressed conditions. We identified the sets of
751 shared syntelogs as well as those that were only found in a single species. We then ran GO
752 enrichment analysis of these sets of shared and unique co-expressed genes.

753 754 *Topological data analysis*

755 We employed a topological data analysis (TDA) approach following the pipeline
756 described at <https://github.com/PlantsAndPython/plant-evo-mapper> to discern the underlying
757 structure of the expression datasets. We utilized the Mapper algorithm, which condenses the
758 dataset into a scalable, navigable representation. The Mapper algorithm is particularly well-
759 suited for genome scale analyses, as the underlying datasets are often characterized by high
760 dimensionality and sparsity. For our gene expression data, we constructed Mapper graphs
761 using a "stress lens" formulated by applying a linear model using the well-watered condition as
762 a reference point. This model represents the baseline for leaf expression and we quantified the
763 residuals or deviation of each sample from the baseline, which represents the degree of water
764 stress or recovery. We generated three different mapper graphs, one was constructed using the
765 syntelog expression matrix from just the three focal resurrection grasses; *M. caffra*, *O. capense*,
766 and *T. minimus*. The second mapper graph was constructed using the syntelog expression
767 matrix that included two additional resurrection grasses (*E. nindensis* and *O. thomaeum*) and
768 the third graph included the desiccation sensitive species *E. tef*. For the mapper graph, we
769 specified different intervals and overlap for the 3 species comparisons and the 5 species
770 comparisons. For the three species, we specified 110 intervals with a 90% overlap and for the 5
771 species comparison we specified 120 intervals with 95% overlap.

772
773 **Acknowledgements:** This work was funded by NSF IOS-PRFB-1906094 to RAM, DBI-
774 2213983 to RAM and RV, and MCB-1817347 to RV. We thank landowners and stewards Pieter
775 and Jennie Pretorius, Ken Maude, Pieter and Nadine Vervoort, and Wayne and Messiah
776 Mudenda for assistance with field logistics and permission to collect plants. We thank the
777 University of Cape Town for access to facilities, Prof. Jill Farrant for her guidance, and Keren
778 Cooper for logistical support. We also thank the South African National Herbarium, Pretoria for
779 assistance identifying and vouchering of specimens and the United States Department of
780 Agriculture for providing import permits.

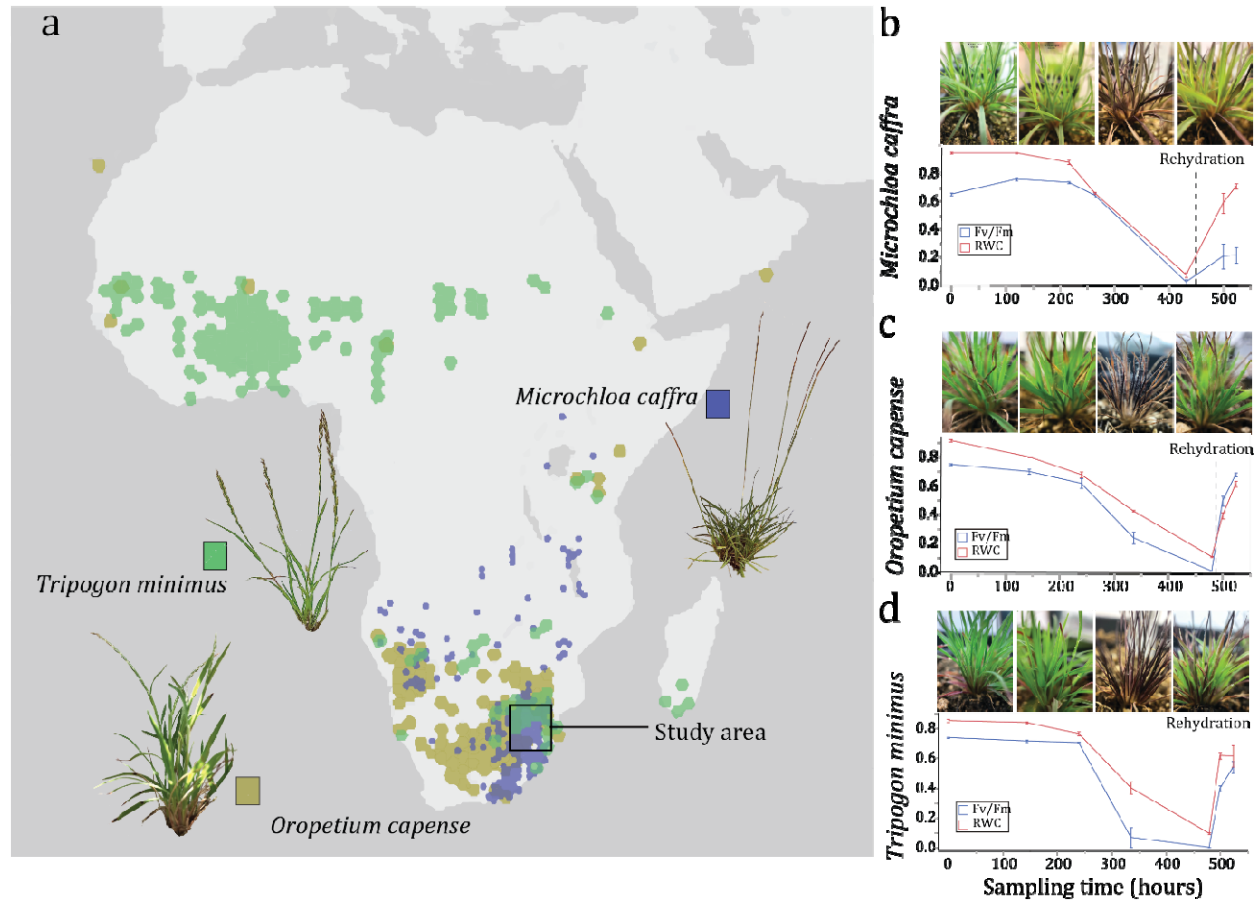
781
782 **Author contributions:** RAM and RV conceived of the study. RAM collected and curated data.
783 RAM, LVP, JS, ISG, and RV conducted data analyses and contributed to data interpretation and
784 conceptual framing of the manuscript. RAM and RV drew the figures. RAM, LVP, and RV wrote
785 the manuscript. All authors edited and reviewed the manuscript.

786

787 **Data availability:** Sequence data associated with this study are deposited at NCBI under
788 BioProject PRJNA1044305 and BioSamples SAMN38380430-92. Genome assemblies are
789 hosted on CoGe (<https://genomeevolution.org/>) under the following IDs: 65089 (*T. minimus*),
790 65046 (*O. capense*), and 64494 (*M. caffra*). Metadata and other data summaries associated
791 with this study are provided at Dryad <https://doi.org/10.5061/dryad.kh18932c4>.

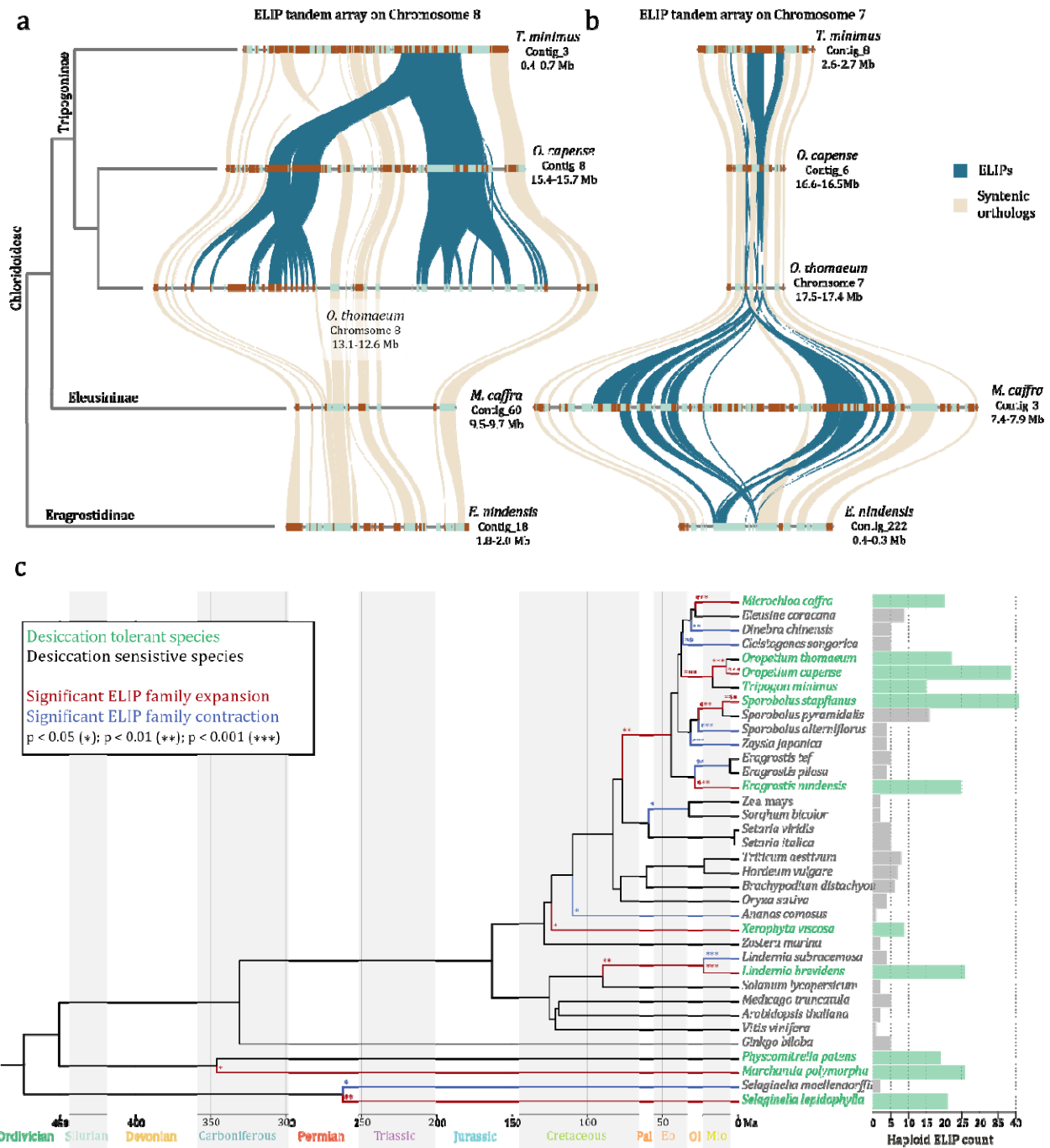
792

793
794



795
796 **Figure 1. Overview of species distribution and experimental design to test for convergent evolution in**
797 **grasses.** (a) Estimated distribution of the three desiccation tolerant grasses *Microchloa caffra*, *Oropetium capense*,
798 and *Tripogon minimus*. Distribution data were taken from GBIF.org (21 November 2023) GBIF Occurrence Download
799 <https://doi.org/10.15468/dl.5jf47y>. Collections for the current study were made in Mpumalanga and Limpopo
800 provinces of South Africa. Relative water content and F_v/F_m of plants during dehydration and rehydration timecourses
801 for (b) *M. caffra*, (c) *O. capense*, and (d) *T. minimus*. Three biological replicates were sampled at each timepoint for
802 each species. Error bars represent standard error of the mean.
803

804



805

806

Figure 2. Independent tandem gene duplication of ELIPs in different resurrection grass lineages.

807 Microsyntenic regions of the Chromosome 8 (a) and Chromosome 7 (b) ELIP tandem arrays is shown for resurrection

808 grasses in the Tropogoninae (*T. minimus*, *O. thomaeum*, and *O. capense*), Eleusininae (*M. caffra*), and

809 Eragrostidinae (*E. nindensis*) subtribes of Chloridoideae. Syntenic orthologs between the species are shown in beige

810 and the ELIPs are highlighted in blue. Only a single syntenic region for autopolyploids *M. caffra* (hexaploid) and *E.*

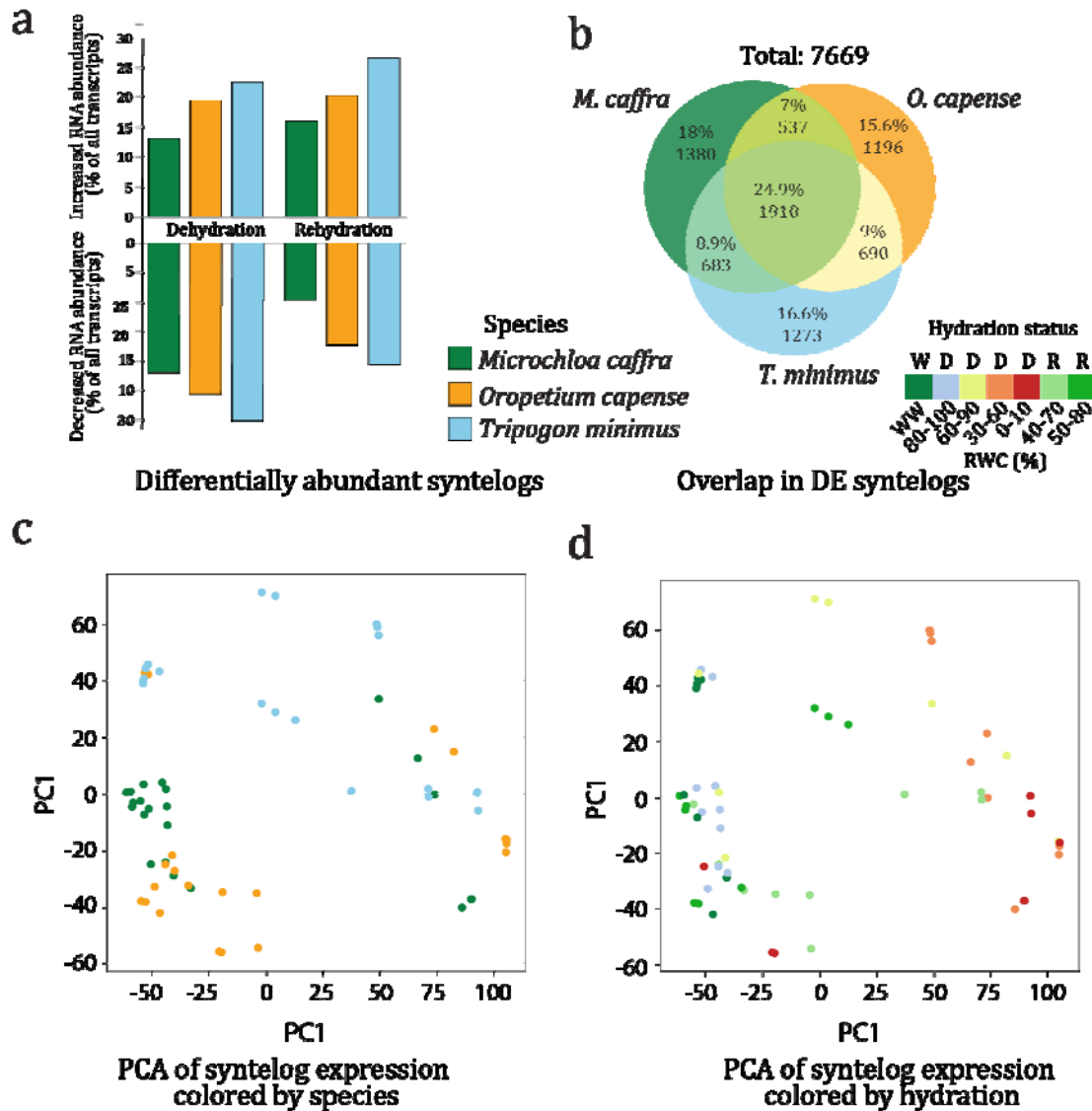
811 *nindensis* (tetraploid) is shown for simplicity, but each of the other haplotypes contain the same gene content in these

812 regions. (c) Evolutionary dynamics showing significant changes in the rates of gene family expansion (red) and

813 contraction (blue) of ELIPs inferred by CAFE. The haploid normalized number of ELIPs are plotted for desiccation

814 tolerant (green) and sensitive (gray) species.

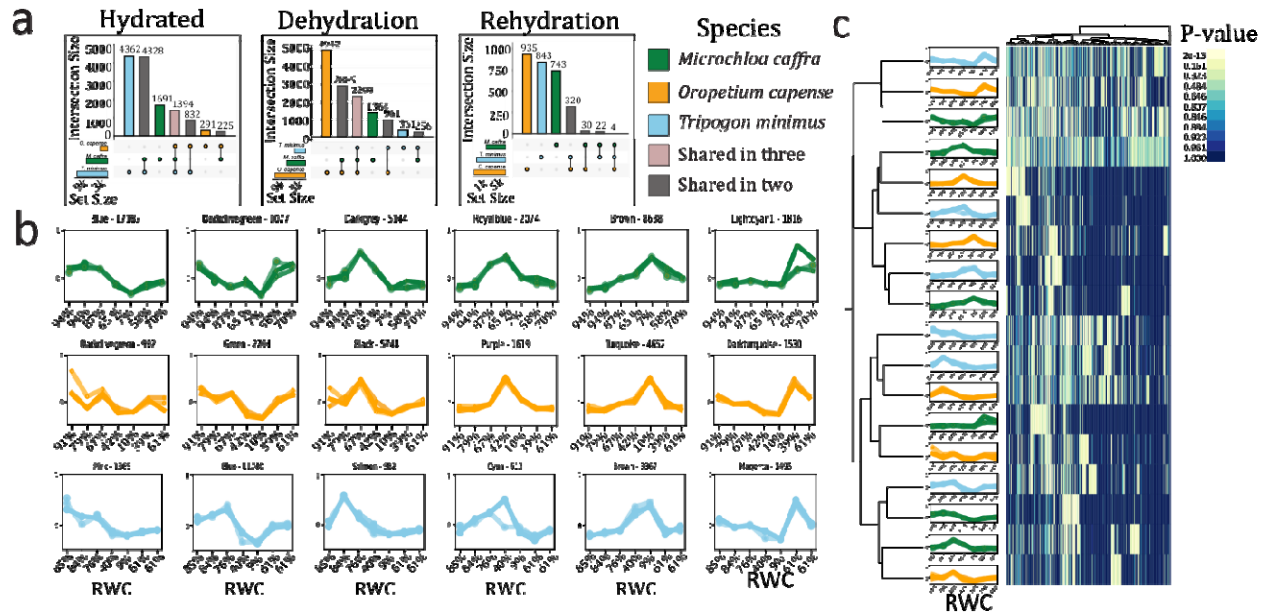
815
816



817
818
819
820
821
822
823
824

Figure 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. (c,d) Principal component analysis of z-score transformed expression values for conserved syntelogs across all three species. Samples are colored by hydration status in (c) and by species identity in (d).

825



826

827

828

829

830

831

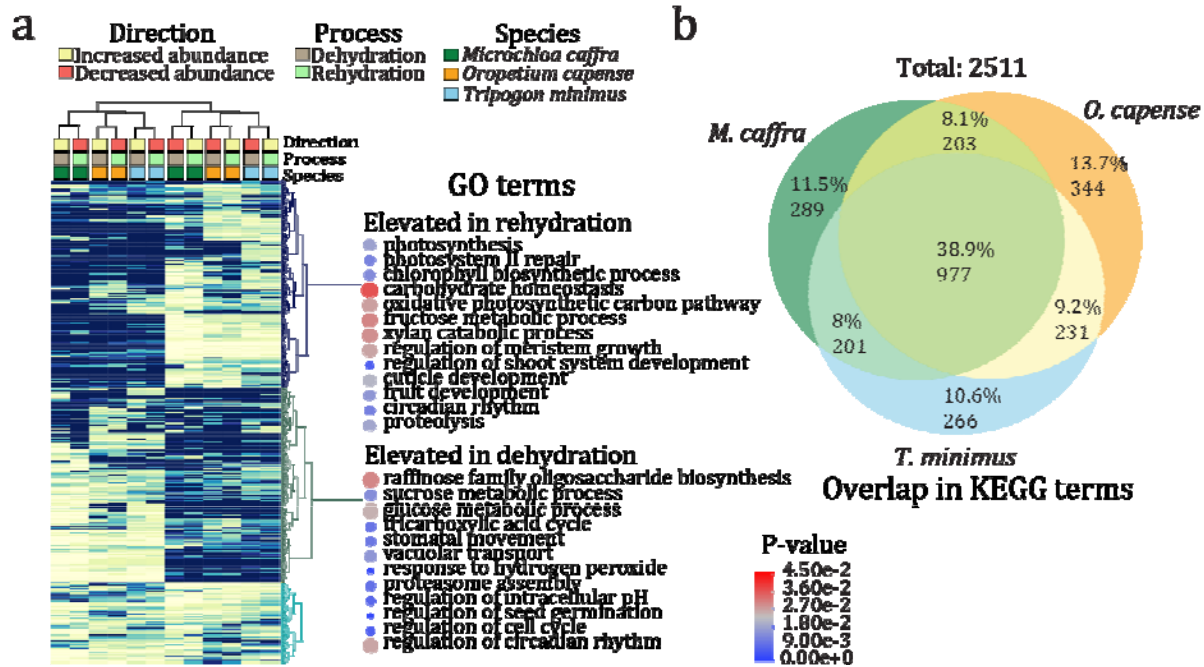
832

833

834

Figure 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. (b) Co-expression modules for each species. The X-axis shows the approximate relative water content (RWC) of samples at each timepoint. 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.

835



836

837

838

839

840

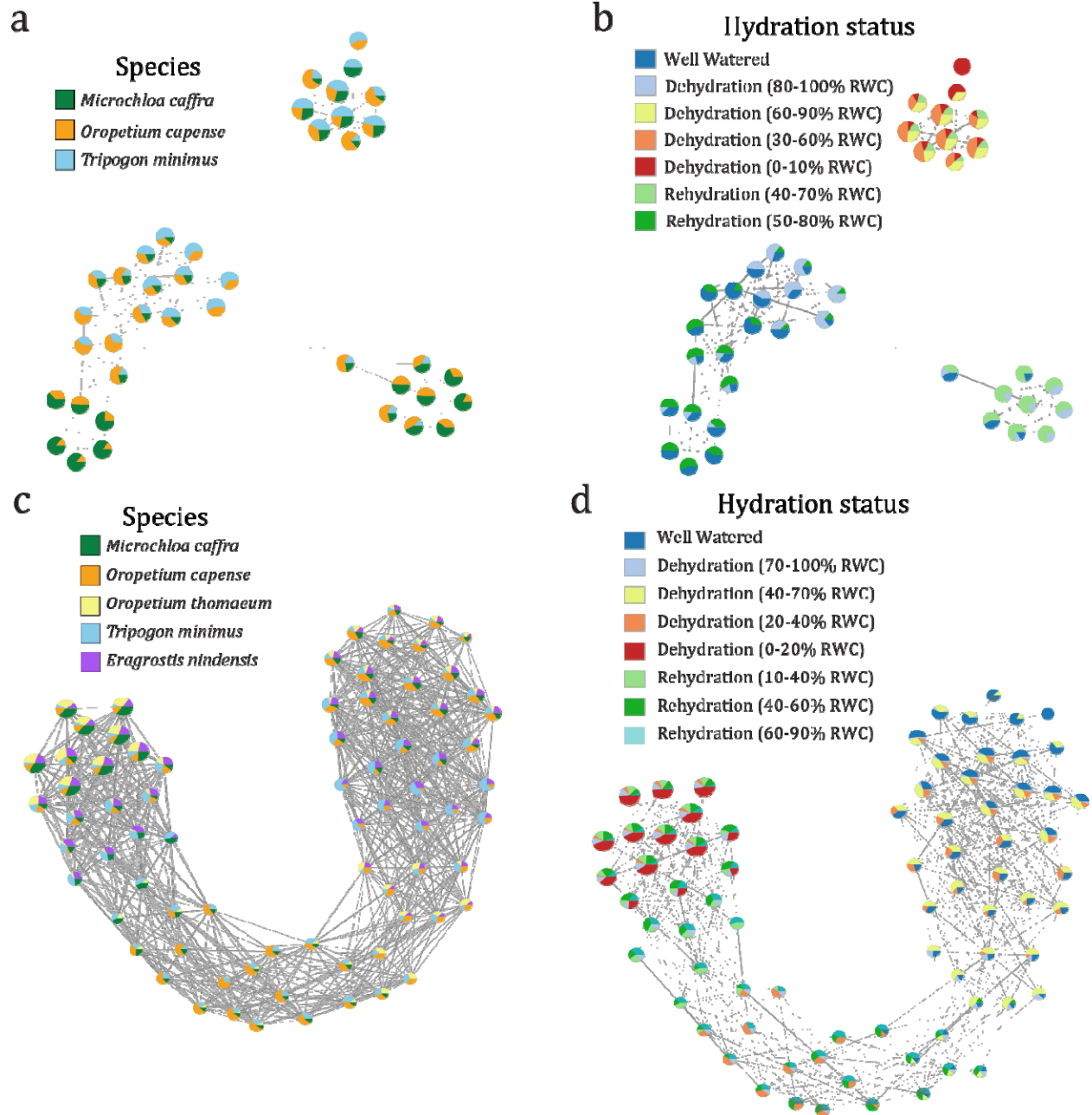
841

842

843

844

Figure 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 colored 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.



845
846 **Figure 6. Mapper graphs showing the topological shape of desiccation induced gene expression across**
847 **species.** Nodes within the graph represent clusters of RNAseq samples that are akin to one another, with the node
848 color indicating the identity of the samples contained within. Edges, or the connections between nodes, delineate
849 shared samples across intersecting clusters. Mapper graphs for the three species comparisons are shown in (a) and
850 (b), and Mapper graphs for the five species comparisons are shown in (c) and (d). Nodes within the graph are colored
851 by species (a and c) or hydration status (b and d).

852 **Table 1. Assembly stats of the three resurrection grasses**

Assembly stats	<i>O. capense</i>	<i>T. minimus</i>	<i>M. caffra</i>
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

853
854
855
856
857

858 **References:**

- 859 1. Bewley, J. D. Physiological Aspects of Desiccation Tolerance. *Annu. Rev. Plant Physiol.*
860 **30**, 195–238 (1979).
- 861 2. Oliver, M. J., Tuba, Z. & Mishler, B. D. The evolution of vegetative desiccation tolerance in
862 land plants. *Plant Ecol.* **151**, 85–100 (2000).
- 863 3. Marks, R. A., Farrant, J. M., Nicholas McLetchie, D. & VanBuren, R. Unexplored
864 dimensions of variability in vegetative desiccation tolerance. *Am. J. Bot.* (2021)
865 doi:10.1002/ajb2.1588.
- 866 4. Alpert, P. Constraints of tolerance: why are desiccation-tolerant organisms so small or
867 rare? *J. Exp. Biol.* **209**, 1575–1584 (2006).
- 868 5. VanBuren, R. Desiccation tolerance: Seedy origins of resurrection. *Nature Plants* **3**, 17046
869 (2017).
- 870 6. Costa, M. C. D. *et al.* Key genes involved in desiccation tolerance and dormancy across life
871 forms. *Plant Sci.* (2016) doi:10.1016/j.plantsci.2016.02.001.
- 872 7. VanBuren, R. *et al.* Desiccation Tolerance Evolved through Gene Duplication and Network
873 Rewiring in *Lindernia*. *Plant Cell* **30**, 2943–2958 (2018).
- 874 8. Alpert, P. The limits and frontiers of desiccation-tolerant life. *Integr. Comp. Biol.* **45**, 685–
875 695 (2005).
- 876 9. Porembski, S. & Barthlott, W. *Granitic and Gneissic Outcrops (inselbergs) as Centers of*
877 *Diversity for Desiccation-Tolerant Vascular Plants*. vol. 151 19–28
878 <https://link.springer.com/content/pdf/10.1023%2FA%3A1026565817218.pdf> (2000).
- 879 10. Dace, H. J. W. *et al.* A Horizontal View of Primary Metabolomes in Vegetative Desiccation
880 Tolerance. *bioRxiv* 2023.02.10.528018 (2023) doi:10.1101/2023.02.10.528018.
- 881 11. Crowe, J. H., Carpenter, J. F. & Crowe, L. M. THE ROLE OF VITRIFICATION IN
882 ANHYDROBIOSIS. *Annu. Rev. Physiol.* **60**, 73–103 (1998).
- 883 12. Hoekstra, F. A., Golovina, E. A. & Buitink, J. Mechanisms of plant desiccation tolerance.

- 884 *Trends Plant Sci.* **6**, 431–438 (2001).
- 885 13. Dinakar, C., Djilianov, D. & Bartels, D. Photosynthesis in desiccation tolerant plants: Energy
886 metabolism and antioxidative stress defense. *Plant Science* vol. 182 29–41 Preprint at
887 <https://www.ncbi.nlm.nih.gov/pubmed/22118613> (2012).
- 888 14. Oliver, M. J. *et al.* Desiccation Tolerance: Avoiding Cellular Damage During Drying and
889 Rehydration. *Annu. Rev. Plant Biol.* (2020) doi:10.1146/annurev-arplant-071219-105542.
- 890 15. Moore, J. P., Le, N. T., Brandt, W. F., Driouich, A. & Farrant, J. M. Towards a systems-
891 based understanding of plant desiccation tolerance. *Trends in Plant Science* vol. 14 110–
892 117 Preprint at <https://www.ncbi.nlm.nih.gov/pubmed/19179102> (2009).
- 893 16. Farrant, J. M. & Hilhorst, H. Crops for dry environments. *Curr. Opin. Biotechnol.* **74**, 84–91
894 (2021).
- 895 17. VanBuren, R. *et al.* Core cellular and tissue-specific mechanisms enable desiccation
896 tolerance in *Craterostigma*. *Plant J.* **114**, 231–245 (2023).
- 897 18. Pearce, T. Convergence and Parallelism in Evolution: A Neo-Gouldian Account. *Br. J.*
898 *Philos. Sci.* **63**, 429–448 (2012).
- 899 19. Stern, D. L. The genetic causes of convergent evolution. *Nat. Rev. Genet.* **14**, 751–764
900 (2013).
- 901 20. Wiley, E. O. Homoplasy. in *Encyclopedia of Genetics* (eds. Brenner, S. & Miller, J. H.) 969–
902 970 (Academic Press, New York, 2001).
- 903 21. Arendt, J. & Reznick, D. Convergence and parallelism reconsidered: what have we learned
904 about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32 (2008).
- 905 22. Tebele, S. M., Marks, R. A. & Farrant, J. M. Two Decades of Desiccation Biology: A
906 Systematic Review of the Best Studied Angiosperm Resurrection Plants. *Plants* **10**, (2021).
- 907 23. Xiao, L. *et al.* The resurrection genome of *Boea hygrometrica* : A blueprint for survival of
908 dehydration. *Proceedings of the National Academy of Sciences* **112**, 5833–5837 (2015).
- 909 24. VanBuren, R. *et al.* Single-molecule sequencing of the desiccation-tolerant grass

- 910 Oropetium thomaeum. *Nature* **527**, 508–511 (2015).
- 911 25. Costa, M.-C. D. *et al.* A footprint of desiccation tolerance in the genome of Xerophyta
912 viscosa. *Nature Plants* **3**, 17038 (2017).
- 913 26. VanBuren, R., Wai, C. M., Keilwagen, J. & Pardo, J. A chromosome-scale assembly of the
914 model desiccation tolerant grass *Oropetium thomaeum*. *Plant Direct* (2018).
- 915 27. VanBuren, R. *et al.* Extreme haplotype variation in the desiccation-tolerant clubmoss
916 *Selaginella lepidophylla*. *Nat. Commun.* **9**, 13 (2018).
- 917 28. Pardo, J. *et al.* Intertwined signatures of desiccation and drought tolerance in grasses.
918 *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10079–10088 (2020).
- 919 29. Montes, R. A. C. *et al.* A comparative genomics examination of desiccation tolerance and
920 sensitivity in two sister grass species. *Proceedings of the National Academy of Sciences*
921 **119**, e2118886119 (2022).
- 922 30. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo
923 assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
- 924 31. Marks, R. A., Hotaling, S., Frandsen, P. B. & VanBuren, R. Representation and
925 participation across 20 years of plant genome sequencing. *Nat Plants* **7**, 1571–1578
926 (2021).
- 927 32. VanBuren, R., Wai, C. M., Keilwagen, J. & Pardo, J. A chromosome-scale assembly of the
928 model desiccation tolerant grass *Oropetium thomaeum*. *Plant Direct* **2**, e00096 (2018).
- 929 33. VanBuren, R. *et al.* Exceptional subgenome stability and functional divergence in the
930 allotetraploid Ethiopian cereal teff. *Nat. Commun.* **11**, 884 (2020).
- 931 34. VanBuren, R., Pardo, J., Man Wai, C., Evans, S. & Bartels, D. Massive Tandem
932 Proliferation of ELIPs Supports Convergent Evolution of Desiccation Tolerance across Land
933 Plants. *Plant Physiol.* **179**, 1040–1049 (2019).
- 934 35. Han, M. V., Thomas, G. W. C., Lugo-Martinez, J. & Hahn, M. W. Estimating Gene Gain and
935 Loss Rates in the Presence of Error in Genome Assembly and Annotation Using CAFE 3.

- 936 *Mol. Biol. Evol.* **30**, 1987–1997 (2013).
- 937 36. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and
938 bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
- 939 37. du Toit, S. F., Bentley, J. & Farrant, J. M. Chapter Nine - NADES formation in vegetative
940 desiccation tolerance: Prospects and challenges. in *Advances in Botanical Research* (eds.
941 Verpoorte, R., Witkamp, G.-J. & Choi, Y. H.) vol. 97 225–252 (Academic Press, 2021).
- 942 38. Hasanuzzaman, M., Nahar, K., Anee, T. I. & Fujita, M. Glutathione in plants: biosynthesis
943 and physiological role in environmental stress tolerance. *Physiol. Mol. Biol. Plants* **23**, 249–
944 268 (2017).
- 945 39. Gasulla, F. *et al.* The role of lipid metabolism in the acquisition of desiccation tolerance in
946 *Craterostigma plantagineum*: a comparative approach. *Plant J.* **75**, 726–741 (2013).
- 947 40. Sun, M. *et al.* Phosphatidylcholine Enhances Homeostasis in Peach Seedling Cell
948 Membrane and Increases Its Salt Stress Tolerance by Phosphatidic Acid. *Int. J. Mol. Sci.*
949 **23**, (2022).
- 950 41. Zhang, X., Gao, Y., Zhuang, L. & Huang, B. Phosphatidic acid priming-enhanced heat
951 tolerance in tall fescue (*Festuca arundinacea*) involves lipidomic reprogramming of lipids for
952 membrane stability and stress signaling. *Plant Growth Regul.* **99**, 527–538 (2023).
- 953 42. VanBuren, R. *et al.* Seed desiccation mechanisms co-opted for vegetative desiccation in
954 the resurrection grass *Oropetium thomaeum*. *Plant Cell Environ.* **40**, 2292–2306 (2017).
- 955 43. Palande, S. *et al.* The topological shape of gene expression across the evolution of
956 flowering plants. *bioRxiv* 2022.09.07.506951 (2022) doi:10.1101/2022.09.07.506951.
- 957 44. Lim, P. O., Kim, H. J. & Nam, H. G. Leaf senescence. *Annu. Rev. Plant Biol.* **58**, 115–136
958 (2007).
- 959 45. Li, S. *et al.* Histone Acetylation Changes in Plant Response to Drought Stress. *Genes* **12**,
960 (2021).
- 961 46. Wang, Q. *et al.* JMJ27-mediated histone H3K9 demethylation positively regulates drought-

- 962 stress responses in Arabidopsis. *New Phytol.* **232**, 221–236 (2021).
- 963 47. Ma, S. *et al.* Reversible Histone H2B Monoubiquitination Fine-Tunes Abscisic Acid
964 Signaling and Drought Response in Rice. *Mol. Plant* **12**, 263–277 (2019).
- 965 48. Farrant, J. M. A comparison of mechanisms of desiccation tolerance among three
966 angiosperm resurrection plant species. *Plant Ecol.* (2000).
- 967 49. VanBuren, R. *et al.* Variability in drought gene expression datasets highlight the need for
968 community standardization. *bioRxiv* 2024.02.04.578814 (2024)
969 doi:10.1101/2024.02.04.578814.
- 970 50. Illing, N., Denby, K. J., Collett, H., Shen, A. & Farrant, J. M. The signature of seeds in
971 resurrection plants: A molecular and physiological comparison of desiccation tolerance in
972 seeds and vegetative tissues. *Integr. Comp. Biol.* **45**, 771–787 (2005).
- 973 51. Alejo-Jacuinde, G., González-Morales, S. I., Oropeza-Aburto, A., Simpson, J. & Herrera-
974 Estrella, L. Comparative transcriptome analysis suggests convergent evolution of
975 desiccation tolerance in Selaginella species. *BMC Plant Biol.* **20**, 468 (2020).
- 976 52. Ostría-Gallardo, E. *et al.* A comparative gene co-expression analysis using self-organizing
977 maps on two congener filmy ferns identifies specific desiccation tolerance mechanisms
978 associated to their microhabitat preference. *BMC Plant Biol.* **20**, 56 (2020).
- 979 53. Heyduk, K., Moreno-Villena, J. J., Gilman, I. S., Christin, P.-A. & Edwards, E. J. The
980 genetics of convergent evolution: insights from plant photosynthesis. *Nat. Rev. Genet.* **20**,
981 485–493 (2019).
- 982 54. Christin, P.-A., Salamin, N., Savolainen, V., Duvall, M. R. & Besnard, G. C4 Photosynthesis
983 evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* **17**, 1241–1247 (2007).
- 984 55. Vicré, M., Farrant, J. M. & Driouich, A. Insights into the cellular mechanisms of desiccation
985 tolerance among angiosperm resurrection plant species. *Plant, Cell and Environment* vol.
986 27 1329–1340 Preprint at (2004).
- 987 56. Gechev, T., Lyall, R., Petrov, V. & Bartels, D. Systems biology of resurrection plants. *Cell.*

- 988 *Mol. Life Sci.* (2021) doi:10.1007/s00018-021-03913-8.
- 989 57. Xu, X. *et al.* Molecular insights into plant desiccation tolerance: transcriptomics, proteomics
990 and targeted metabolite profiling in *Craterostigma plantagineum*. *Plant J.* (2021)
991 doi:10.1111/tpj.15294.
- 992 58. St Aubin, B., Wai, C. M., Kenchanmane Raju, S. K., Niederhuth, C. E. & VanBuren, R.
993 Regulatory dynamics distinguishing desiccation tolerance strategies within resurrection
994 grasses. *Plant Direct* **6**, e457 (2022).
- 995 59. Marks, R. A. *et al.* Polyploidy enhances desiccation tolerance in the grass *Microchloa*
996 *caffra*. *bioRxiv* 2023.06.20.545583 (2023) doi:10.1101/2023.06.20.545583.
- 997 60. Cheng, H. *et al.* Haplotype-resolved assembly of diploid genomes without parental data.
998 *Nat. Biotechnol.* **40**, 1332–1335 (2022).
- 999 61. Wheeler, D. L. *et al.* Database resources of the National Center for Biotechnology
1000 Information. *Nucleic Acids Res.* **36**, D13–21 (2008).
- 1001 62. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a
1002 streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).
- 1003 63. Xiong, W., He, L., Lai, J., Dooner, H. K. & Du, C. HelitronScanner uncovers a large
1004 overlooked cache of Helitron transposons in many plant genomes. *Proc. Natl. Acad. Sci. U.*
1005 *S. A.* **111**, 10263–10268 (2014).
- 1006 64. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR
1007 retrotransposons. *Nucleic Acids Res.* **35**, W265–8 (2007).
- 1008 65. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de
1009 novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).
- 1010 66. Campbell, M. S. *et al.* MAKER-P: a tool kit for the rapid creation, management, and quality
1011 control of plant genome annotations. *Plant Physiol.* **164**, 513–524 (2014).
- 1012 67. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
1013 *Bioinformatics* **34**, i884–i890 (2018).

- 1014 68. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- 1015 69. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-
1016 seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
- 1017 70. International Rice Genome Sequencing Project. The map-based sequence of the rice
1018 genome. *Nature* **436**, 793–800 (2005).
- 1019 71. Cheng, C.-Y. *et al.* Araport11: a complete reannotation of the Arabidopsis thaliana
1020 reference genome. *Plant J.* **89**, 789–804 (2017).
- 1021 72. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
- 1022 73. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron
1023 submodel. *Bioinformatics* **19 Suppl 2**, ii215–25 (2003).
- 1024 74. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M.
1025 BUSCO: assessing genome assembly and annotation completeness with single-copy
1026 orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
- 1027 75. Wang, Y. *et al.* MCScanX: a toolkit for detection and evolutionary analysis of gene synteny
1028 and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
- 1029 76. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome
1030 comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157
1031 (2015).
- 1032 77. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7:
1033 Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 1034 78. Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference
1035 in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
- 1036 79. To, T.-H., Jung, M., Lycett, S. & Gascuel, O. Fast Dating Using Least-Squares Criteria and
1037 Algorithms. *Syst. Biol.* **65**, 82–97 (2016).
- 1038 80. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
1039 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

- 1040 81. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-
1041 level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
- 1042 82. Palande, S. *et al.* Topological data analysis reveals a core gene expression backbone that
1043 defines form and function across flowering plants. *PLoS Biol.* **21**, e3002397 (2023).
- 1044 83. Pardo, J. *et al.* Cross-species predictive modeling reveals conserved drought responses
1045 between maize and sorghum. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2216894120 (2023).
- 1046 84. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
1047 RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- 1048 85. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment
1049 search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- 1050 86. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network
1051 analysis. *BMC Bioinformatics* **9**, 559 (2008).
- 1052 87. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of
1053 intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
- 1054