Supplemental Materials and Methods

Plant Growth and Sampling

Accessions of *Eragrostis nindensis* (PI 410063) and *Eragrostis tef* (PI 524434) were obtained from the USDA Germplasm Resources Information Network (www.ars-grin.gov). For the drought timecourse experiments, three seeds of *E. nindensis* were planted in 3.5" nursery pots filled with 125g of redi earth potting mix. Plants were grown for 60 days in a growth chamber under the following conditions: 12hr photoperiod, ~ 400 mol of light, 28°C/22°C day/night temperature. Pots were brought to a total weight of 200g by adding water at the start of the drought experiment. Water was then withheld for the remainder of the experiment for drought treated plants, but well-watered (WW) plants were maintained at a total pot weight of 200g daily. Leaf tissue was sampled both for relative water content (RWC) and electrolyte leakage assays every 4 hours beginning 48 hours after the start of the drought experiment. Three non-senescing (inner) leaves from each plant were randomly selected and excised at the mid-section. Samples were divided for relative water content and electrolyte leakage measurements. Inner leaves were collected for RNAseq, Bisulfite-seq and ChIP-seq experiments. Leaf samples for the D1 / WW, D2 and D3 timepoints were collected 56, 104 and 224 hours after cessation of watering respectively (ZT8 on day 3, day 4 and day 5). At each timepoint, inner leaf tissue was pooled from 3 plants per pot. Leaf tips were removed as they generally do not recover from desiccation. For the rehydration experiment, 102-day old plants were maintained as described above and slowly desiccated over 143 hours followed by application of water for rehydration. Leaf samples were collected at 0, 12, 24 and 48 hours post rehydration. At each rehydration timepoint, samples for RWC, electrolyte leakage, RNAseq, and Methyl-seq were collected. All samples for RNAseq were flash frozen in liquid nitrogen before storing at -80°C. Samples for RWC and electrolyte leakage were processed immediately after collection. Electrolyte leakage data was not collected for 48 hours post rehydration samples.

E. tef plants were grown in the same growth chamber as *E. nindensis* plants with the same photoperiod, light, and temperature conditions. Three *E. tef* seedlings per pot were grown for one month in 3.5" nursery pots using redi earth potting mix. Pots were brought to the same weight at the start of the experiment (260g) and water was withheld from plants designated for drought treatment while well-watered (WW) plants were maintained at 220g daily. The *E. tef* D1 and D2 samples were collected at 128 hours and 152 hours after equalizing the pot weights respectively.

Relative Water Content

Relative water content was measured according to a previously published protocol with minor modifications (1). Briefly, leaf strips were excised from the midpoint of 3 *E. nindensis* leaves or a single *E. tef* leaf and immediately placed in a sealed tube at ~12°C. The fresh weight of all the samples was recorded directly following sample collection. Samples were then floated in 5mL of deionized water at 4°C in the dark for 24 hours before measuring the turgid weight. Samples were then dried for 24-48 hours at 60°C (until the sample weight stabilized) to obtain dry

weights. Relative water content was calculated as [(fresh weight - dry weight) / (Turgid weight - dry weight)] * 100%.

Electrolyte leakage

Electrolyte leakage was measured according to the method outlined by A. Thalhammer (2). Briefly, fresh leaf samples were placed in 5ml of deionized water and equilibrated overnight at 4°C. Samples were brought to 25°C the following day and the conductivity ($conductivity_{fresh}$) was measured using a Mettler Toledo InLab 731-ISM conductivity probe. Samples were then boiled for 30 minutes to disrupt the cell membranes before cooling back to 25°C. The conductivity after post boiling ($conductivity_{boiled}$) was then measured. Electrolyte leakage percentage was calculated as the ($conductivity_{fresh}$ / $conductivity_{boiled}$) * 100%. Leaf tips and older leaves in *E. nindensis* do not always recover after desiccation, and differences in electrolyte leakage among such tissues in response to drying has been reported (3). Thus the discrepancy in electrolyte leakage between the two desiccated timepoints (D3 and R0) may be attributed to this phenomenon.

Nucleic acid extraction, library preparation, and sequencing

High molecular weight genomic DNA for PacBio and Illumina library prep was isolated from leaf tissue of young *E. nindensis* plants (~30 days old) using a modified nuclei prep (4). PacBio libraries were constructed using the manufacturer's protocol and were size selected for 25 kb fragments on the BluePippen system (Sage Science). Libraries were sequenced on a PacBio Sequel system. An Illumina DNAseq library was constructed for polishing the PacBio based assembly using 1ug of DNA the same high molecular weight DNA prep with the KAPA HyperPrep Kit (Kapa Biosystems). The Illumina DNAseq library was sequenced on an Illumina HiSeq4000 under paired end mode (150 bp) at the RTSF Genomics Core at Michigan State University.

RNA was extracted from the timepoints described above for *E. nindensis* and *E. tef* using the Omega Biotek E.Z.N.A. Plant RNA kit according to the manufacturer's protocol using ~200 mg of frozen tissue for each sample and quantified using the Qubit RNA HS and IQ assay kit (Invitrogen, USA). Each timepoint for RNA samples had three biological replicates. Stranded RNAseq libraries were constructed using 2ug of high-quality total RNA. The Illumina TruSeq stranded total RNA LT sample prep kit (RS-122-2401 and RS-122-2402) were used for library construction following the manufacturer's protocol. Multiplexed RNAseq libraries were quantified, pooled, and sequenced on an Illumina HiSeq4000 under paired-end 150nt mode at the RTSF Genomics Core at Michigan State University.

Chromatin Immuno Precipitation sequencing (ChIP-seq) library construction

Chromatin immunoprecipitation was performed according to a protocol modified from previously published protocols (5, 6). Briefly, nuclei were extracted from 2g of freshly ground

tissue. The nuclei were then digested with micrococcal nuclease (MNase, Sigma #N5386-500UN). Following digestion a portion of the chromatin was set aside as the input control sample. The remaining chromatin was incubated overnight with a commercial H3K4me3 antibody (Abcam #Ab8580) in a rProtein A agarose (Roche #11134515001) suspension. Following antibody incubation the chromatin was eluted and purified using a DNA Clean & Concentrator kit (Zymo Research D4003). DNA-seq libraries were then constructed using the same protocol described above.

Genome assembly

In total, we generated 64 Gb of PacBio data representing 63x coverage of the 1.0 Gb *E. nindensis* genome. PacBio reads were error corrected and assembled using Canu, followed by polishing with Pilon using high-coverage Illumina data. Canu parameters were optimized to accurately assemble all haplotypes, yielding an initial *E. nindensis* genome assembly with 16,706 contigs spanning 1.96 Gb, or roughly twice the haploid genome size, and a contig N50 of 220 kb (Supplemental Table 1). We utilized the Pseudohaploid algorithm (https://github.com/schatzlab/pseudohaploid) to filter out redundant haplotypes from the assembly , as previously described in (7). Briefly, Pseudohaploid filters out redundant haplotypes from the full assembly based on overlap to produce a 'pseudo' haploid reference. This filtering approach yielded a total haploid assembly of 986 Mb across 4,368 contigs with an N50 of 520kb. This assembly is referred to as *E. nindensis* V2.1.

The genome size of *E. nindensis* (PI 410063) was estimated using flow cytometry in two separate runs as previously described (8). The *E. nindensis* genome was assembled using Canu V1.8 (9) with polishing using Pilon V1.22 (10). Raw PacBio reads were used as input for Canu and the following parameters were modified to allow for more careful unitigging and haplotype assembly: minReadLength=5000, GenomeSize=1035Mb, corOutCoverage=200 "batOptions=-dg 3 -db 3 -dr 1 -ca 500 -cp 50". All other parameters were left as default. The output assembly graph was visualized using Bandage (11) to assess ambiguities in the graph related to repetitive elements, heterozygosity, and polyploidy. The resulting 1.96 Canu based assembly was roughly twice the estimated genome size (1.05 Gb) indicating that all four haplotypes were at least partially assembled for the allotetraploid genome. The draft Canu based contigs were polished reiteratively using Illumina paired end 150 bp data (~60x). Illumina reads were aligned to the draft contigs using bowtie2 (V2.3.0) (12) under default parameters and the resulting BAM file was used as input for Pilon. The following parameters were modified for Pilon and all others were left as default: --flank 7, --K 49, and --mindepth 10. Pilon was run recursively a total of 5 times using the updated reference for each iteration.

The *E. nindensis* genome assembly was further processed to create a pseudo-haploid representation of the genome where one of the haploypes was filtered out using the Pseudohaploid algorithm (http://github.com/schatzlab/pseudohaploid). To identify haplotype containing contigs, the genome was aligned against itself using the whole genome aligner nucmer from the MUMmer package (13). The following parameters were used for nucmer to

report all unique and repetitive alignments longer than 500 bp: nucmer -maxmatch -l 100 -c 500. This file was used as input for Pseudohaploid and the following parameters were changed in the create_pseudohaploid.sh script: MIN_IDENTITY: 95; MIN_LENGTH: 1000; MIN_CONTAIN: 90; MAX_CHAIN_GAP: 20000. Using these parameters filtered alignment chains with a minimum identity of 95%, minimum contig overlap between haplotypes of 90%, and maximum insertion size of 20kb were removed. This approach ensured that homeologous regions from the allopolyploid event were not filtered out and the strict overlap ensured that informative sequences were not purged from the assembly. The final, V2.1 assembly has a total size of 986 Mb across 4,368 contigs with an N50 of 520kb, which is similar to the expected haploid genome size.

Genome annotation

We annotated 116,452 genes in the *E. nindensis* genome using the . Of these, 79,755 were syntenic with *E. tef*, and 80,997 had at least one ortholog, syntenic or otherwise, in the *E. tef* genome. 84,603 genes have at least one pfam domain. Overall, 98,294 genes (84.4%) were orthologous to *E. tef* or contained pfam domains. We combined this set of genes with the 58,602 genes with detectable expression (defined as $\sum TPM > 1$ across all conditions sampled) to create a set of 107,683 “high confidence" gene models. Of these high confidence genes, 74.1% were syntenic with *E. tef,* and 80.0% of genes with detectable expression were syntenic, suggesting sufficient collinearity for genome wide comparisons. We used the Embryophyta Benchmarking Universal Single-Copy Orthologs (BUSCO) to evaluate the completeness of our annotation. We found copies of most of the 1440 Embryophyta BUSCOs (92.1% complete, 95.6% complete or fragmented). The majority were duplicated (65.6%; 946), which is consistent with the polyploid nature of *E. nindensis*.

The *E. nindensis* genome was annotated with MAKER-P v2.31.8 (14) using transcript evidence from RNAseq data and protein homology. A *de-novo* transcriptome was assembled with RNAseq reads from well-watered leaf tissue using Trinity v2.6.6 (15). This assembly was used as expressed sequence tag evidence (EST) in MAKER. A second transcriptome library was assembled from RNAseq data of well-watered and desiccated leaf tissue using StringTie v1.3.3 (16). Default parameters were used for StingTie and the ?merge option was turned on. This evidence was provided to MAKER in the “maker_gff" slot. In addition to expression evidence, protein annotations for *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor, Zea mays*, *Seteria italica* and *Eragrostis tef* were used as protein homology evidence. Transposable elements and repetitive sequences were annotated using a custom repeat library (described below). We ran three rounds of ab-intio gene prediction using the SNAP gene prediction program (17) with the output of the prior MAKER run used as training data. BUSCO v3.0.1 (18) was used to assess the annotation quality with the set of 1440 conserved single copy orthologs from the odb9 database (https://busco.ezlab.org/v2/).

Identification of repetitive elements

Long terminal repeat retrotransposons (LTR-RTs) were identified using LTR harvest (genome tools V1.5.8) (19) and LTR_finder V1.07 (20) and this list of candidate LTR-RTs were filtered and refined using LTR retriever V1.8.0 (21). Parameters or LTR harvest were modified as follows based on guidelines from LTR retriever: -similar 90 –vic 10 –seed 20 –minlenltr 100 –maxlenltr 7000 –mintsd 4 –maxtsd 6. The following parameters for LTR finder were modified: -D 15000 –d 1000 –L 7000 –l 100 –p 20 –C –M 0.9. The resulting candidate LTR-RTs from both these programs were used as input for LTR retriever. LTR retriever was run with default parameters. Elements were defined as intact if they were flanked by terminal repeats. The filtered, non-redundant library from LTR retriever was used as input for whole-genome annotation of retrotransposons using RepeatMasker (http://www.repeatmasker.org/) (22).

ChIP-seq data analysis

Raw sequencing reads were trimmed with Trimmomatic v0.38 and aligned to the *E. nindensis* reference genome using bwa mem v 0.7.17 with default parameters (23, 24). Peaks of enriched ChIP signal were called relative to the corresponding input control, which was digested by MNase but not incubated with the antibody using PePr (25). PePr accounts for the variance between replicates when calling peaks and only returns peaks that are significant after accounting for this variation. PePr was also used to identify differentially bound regions between well-watered (WW) and D3 samples. pyBedtools was used to identify the closest gene to each of these peaks including genes that overlapped the peak regions (26). The log2fold enrichment of read coverage was calculated across the genome using 10bp bins for each ChIP sample compared with the corresponding input using the bamCompare tool from deepTools v. 3.2.1 (27). The average log2Fold change was calculated across all three replicates of WW and D3 samples separately using WiggleTools (28). This average log2Fold change was plotted for the 2kb upstream and downstream regions of the transcriptional start site of the genes closest to the differential peaks using the computeMatrix and plotHeatmap tools from deepTools (27).

Bisulfite-seq data analysis

Bisulfite sequencing reads were trimmed using Trimommatic v0.38 (29). Reads were then aligned reads to bisulfite corrected *E. nindensis* reference and methylation states were called using Bismark v0.21.0 with default settings and a minimum depth of 3 reads (30). The average methylation percentage for each cytosine was calculated with a custom python script (available on gitHub). The resulting bedGraph files were then converted to bigWig format using USC genome browser's bedGraphToBigWig script (31). The methylation percentage across gene regions was calculated with deepTools computeMatrix run in scaleRegions mode with gene bodies scaled to 1000bp and a bin size of 10bp (27).

Comparative genomics

The python version of MCScan was used (https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version)) to identify pairwise syntenic orthologs between each of the analyzed grass species *(32)*. The *O. thomaeum* genome was used as a common anchor to the other grass genomes as it is the phylogenetically closest diploid species to *E. nindensis* and *E. tef* and it has a high quality chromosome scale genome assembly. A minimum cutoff of five genes was used to identify syntenic gene blocks. The syntenic gene lists from all pairwise comparisons were combined and filtered into two tables, with one including syntenic orthogroups (syntegroups) with at least one gene in the three Chloridoid grasses and the second containing syntegroups with genes present in all six grass species analyzed.

While the surveyed grass genomes were largely collinear, synteny based approaches were not able to identify conserved genes that had translocated or were found in regions with extensive genome rearrangements. Orthofinder was used to identify orthologous genes that were missed by synteny based approaches. Orthofinder (v2.2.6) (33) was run using default parameters with the diamond algorithm to identify orthologs in 22 species. Only orthogroups with at least one ortholog present in all species were included for the analyses.

Expression analysis

Raw fastq files were trimmed to remove sequencing adapters using Trimmomatic v0.38 (29). Gene expression was quantified using Salmon v0.13.1 run in quasi mapping mode (34). The transcript level estimates of expression were converted to gene level transcript per million counts using the R package tximport (35). DEseq2 was used to perform differential expression analysis using the model yij ~ μ + timepoint + eij (36). Each drought and rehydration timepoint was compared to well-watered to identify differentially expressed genes. The built-in wald test in the DEseq2 package was used to test whether the log2fold change of a given gene was equal to 0 (36). Genes with a Wald test, fdr corrected, p-value < 0.05 were considered differentially expressed.

Identification of seed specific genes

Previously published expression data for late maturity or dry seeds, and well-watered leaf tissue from four desiccation sensitive grass species (*E. tef, O. sativa S. bicolor, Z. mays*) were used to identify a set of conserved seed-related genes in grasses. All RNAseq data was downloaded from the Short Read Archive from NCBI. Seed data was reanalyzed from the following sources: *S. bicolor*, (BioProject: PRJDB3281 (37), *E. tef* (38), *Z. mays* (GEO: GSE27004 (39). The raw RNAseq data was quality filtered and quantified using the same pipeline described above. Differential expression analysis was performed using DEseq2 separately for each species in order to identify genes upregulated in seeds compared with well-watered leaves. Using this approach, 640 syntelog groups with conserved upregulation in seeds were identified among all

four grasses. This list of 640 syntelog groups clustered into 386 orthogroups and was used as our list of 'seed related' genes. An empirical approach was used to test if these orthogroups were overrepresented among upregulated genes during desiccation in *E. nindensis*. The empirical null distribution was simulated by randomly selecting (without replacement) 386 orthogroups from the set of 11,905 orthogroups not related to seed processes. A Z-score was calculated based on the observed overlap between upregulated genes and seed orthogroups and the null distribution. This was compared to a normal distribution to determine the probability of identifying at least the observed number of genes as overlapping between the sets. Leaf drought datasets were downloaded from the NCBI SRA and analyzed as described above. The following drought datasets were analyzed: *O. sativa* (BioProject: PRJNA420056 (40), *S. bicolor* (BioProject: PRJNA319738) (41), *Z. mays* (BioProject: PRJNA378714).

K-means clustering

We fit 2nd order polynomial curve to expression across the time series for each gene in *E. nindensis* and *O. thomaeum* separately using the poly1d function in numpy. We then clustered the coefficients for each gene in both species together using the k-means++ algorithm implemented in scikit-learn. We used a k value of 9 as determined by examining a plot of sum of squared distances between clusters for k values between 2 and 25. We chose the k value where the rate of decrease in sum of squared distance became linear.

GO Analysis

We annotated gene ontology terms in the *E. nindensis* and *E. tef* genomes using interproscan v72.0 and BLAST search against Arabidopsis proteins. We conducted enrichment analysis of gene ontology terms among differentially expressed genes using TopGO v1.0. To conduct comparisions of GO enrichment between *E. nindensis* and *E. tef*, we classified terms into the following groups: "seed related" (offspring terms of seed development GO:0048316), "stress related" (offspring terms of response to stress GO:0006950), "sugar related" (offspring terms of carbohydrate metabolic process GO:0005975), "lipid related" (offspring terms of lipid metabolic process GO:0006629), "light response" (offspring terms of response to light stimulus GO:0009416), and "other" (all other terms). We then compared the number of enriched term in each group between species.

References:

1. R. E. Smart, Rapid estimates of relative water content. *Plant Physiol.* **53**, 258–260 (1974).

2. A. Thalhammer, D. K. Hincha, E. Zuther, "Measuring Freezing Tolerance: Electrolyte Leakage and Chlorophyll Fluorescence Assays" in (Humana Press, New York, NY, 2014), pp. 15–24.

3. C. Vander Willigen, N. W. Pammenter, S. Mundree, J. Farrant, Some physiological comparisons between the resurrection grass, Eragrostis nindensis, and the related desiccation-sensitive species, E. curvula. *Plant Growth Regul.* **35**, 121–129 (2001).

4. H.-B. Zhang, X. Zhao, X. Ding, A. H. Paterson, R. A. Wing, Preparation of megabase-size DNA from plant nuclei. *Plant J.* **7**, 175–184 (1995).

5. K. Nagaki, *et al.*, Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of Arabidopsis thaliana centromeres. *Genetics* **163**, 1221–1225 (2003).

6. W. Zhang, T. Zhang, Y. Wu, J. Jiang, Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. *Plant Cell* **24**, 2719–2731 (2012).

7. L.-Y. Chen, *et al.*, The bracteatus pineapple genome and domestication of clonally propagated crops. *Nat. Genet.* (2019) https:/doi.org/10.1038/s41588-019-0506-8.

8. K. Arumuganathan, E. D. Earle, Estimation of nuclear DNA content of plants by flow cytometry. *Plant Mol. Biol. Rep.* **9**, 229–241 (1991).

9. S. Koren, *et al.*, Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).

10. B. J. Walker, *et al.*, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963 (2014).

11. R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352 (2015).

12. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

13. A. L. Delcher, S. L. Salzberg, A. M. Phillippy, Using MUMmer to identify similar regions in large sequence sets. *Curr. Protoc. Bioinformatics* **Chapter 10**, Unit 10.3 (2003).

14. M. S. Campbell, *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).

15. B. J. Haas, *et al.*, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

16. M. Pertea, *et al.*, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).

17. I. Korf, Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).

18. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

19. D. Ellinghaus, S. Kurtz, U. Willhoeft, LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).

20. Z. Xu, H. Wang, LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–8 (2007).

21. S. Ou, N. Jiang, LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. *Plant Physiol.* **176**, 1410–1422 (2018).

22. M. Tarailo-Graovac, N. Chen, Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **25**, 4–10 (2009).

23. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).

24. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

25. Y. Zhang, Y.-H. Lin, T. D. Johnson, L. S. Rozek, M. A. Sartor, PePr: a peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. *Bioinformatics* **30**, 2568–2575 (2014).

26. R. K. Dale, B. S. Pedersen, A. R. Quinlan, Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics* **27**, 3423–3424 (2011).

27. F. Ramírez, F. Dündar, S. Diehl, B. A. Grüning, T. Manke, deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–91 (2014).

28. D. R. Zerbino, N. Johnson, T. Juettemann, S. P. Wilder, P. Flicek, WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* **30**, 1008–1009 (2014).

29. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

30. F. Krueger, S. R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

31. W. J. Kent, A. S. Zweig, G. Barber, A. S. Hinrichs, D. Karolchik, BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).

32. Y. Wang, *et al.*, MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).

33. D. M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons

dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).

34. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

35. C. Soneson, M. I. Love, M. D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).

36. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

37. Y. Makita, *et al.*, MOROKOSHI: transcriptome database in Sorghum bicolor. *Plant Cell Physiol.* **56**, e6 (2015).

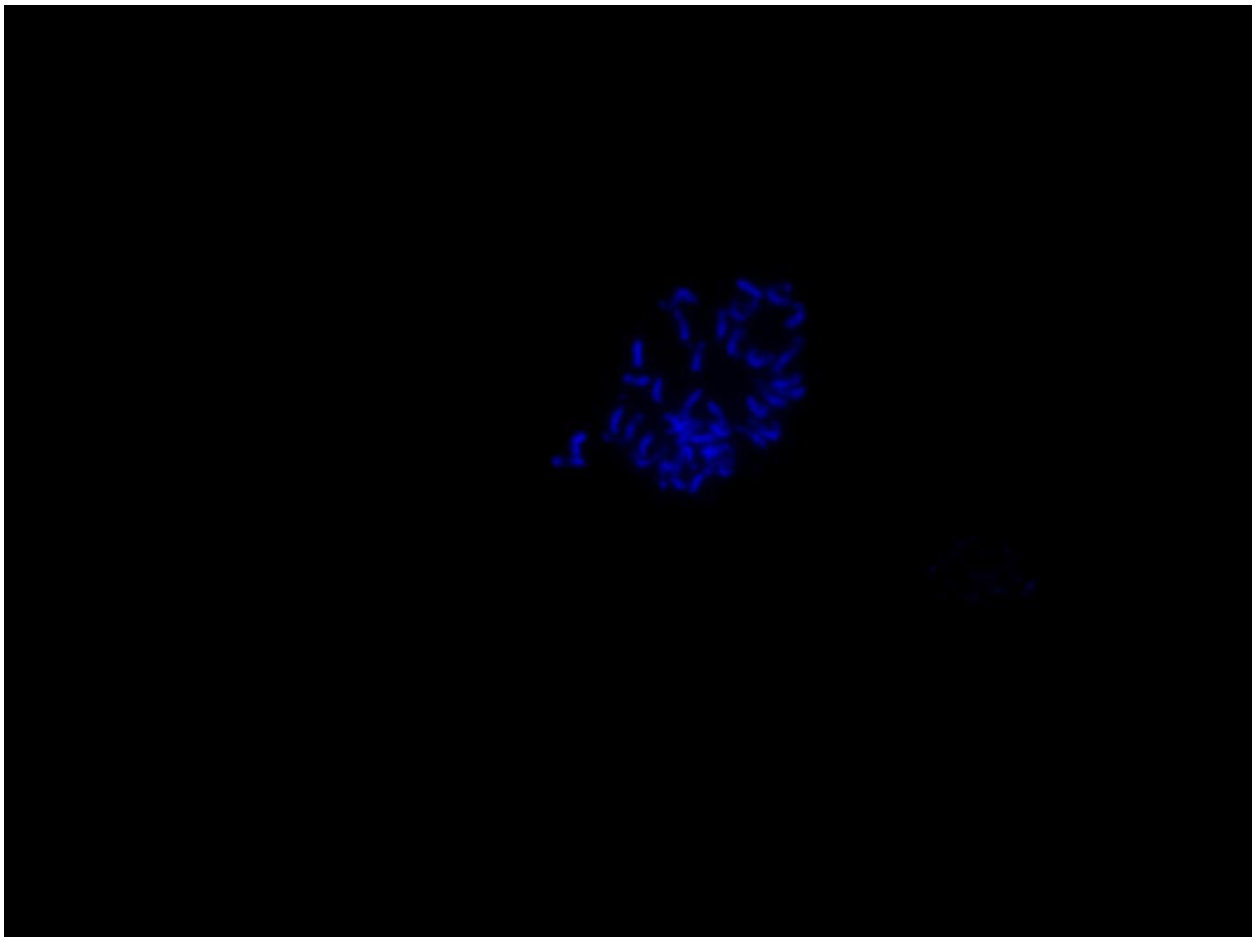
38. R. VanBuren, *et al.*, Exceptional subgenome stability and functional divergence in allotetraploid teff, the primary cereal crop in Ethiopia. *bioRxiv*, 580720 (2019).

39. R. S. Sekhon, *et al.*, Genome-wide atlas of transcription during maize development. *Plant J.* **66**, 553–563 (2011).

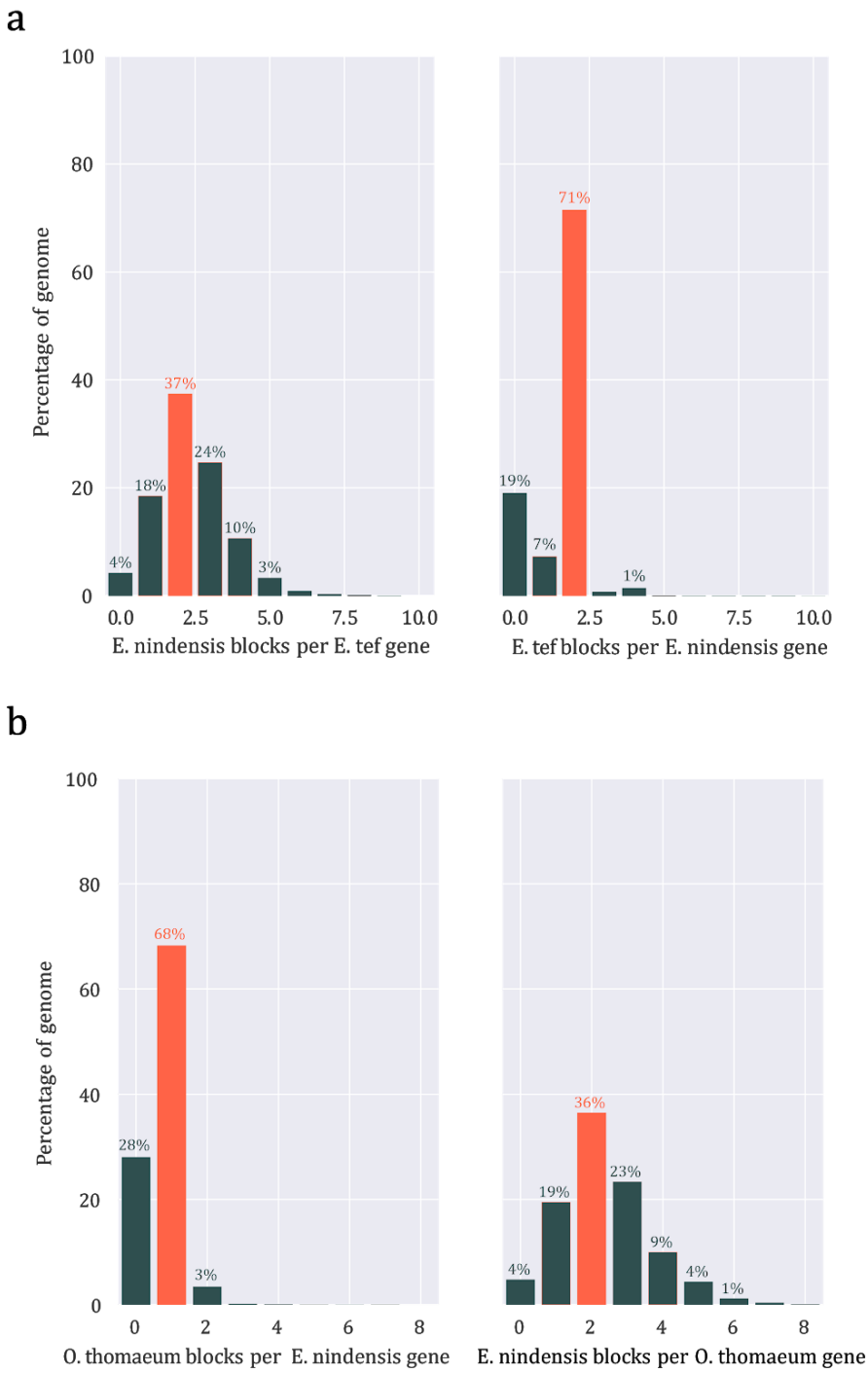
40. J. Fu, *et al.*, OsJAZ1 Attenuates Drought Resistance by Regulating JA and ABA Signaling in Rice. *Front. Plant Sci.* **8**, 2108 (2017).

41. A. Fracasso, L. M. Trindade, S. Amaducci, Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. *BMC Plant Biol.* **16**, 115 (2016).

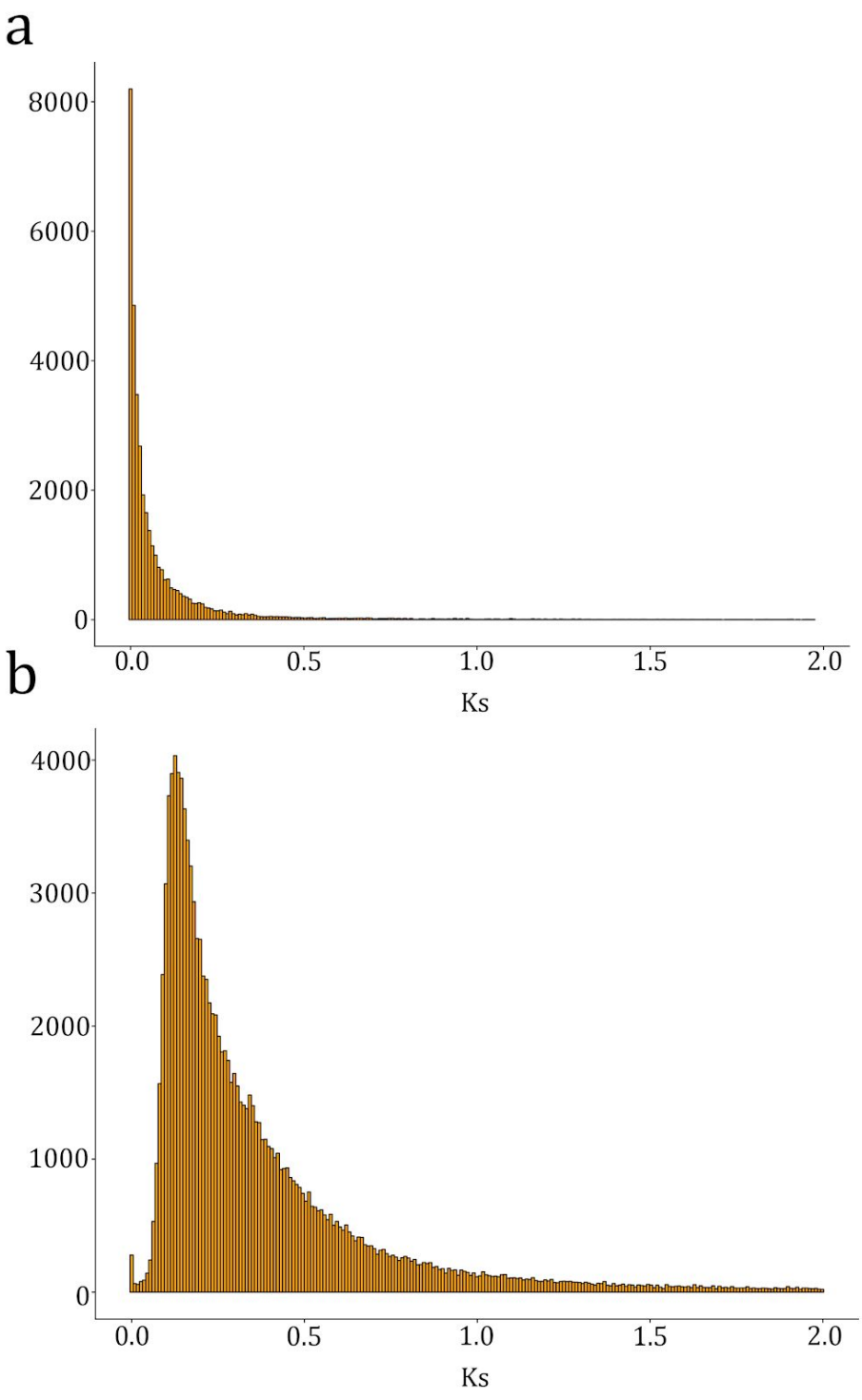
Supplemental Figures and Tables



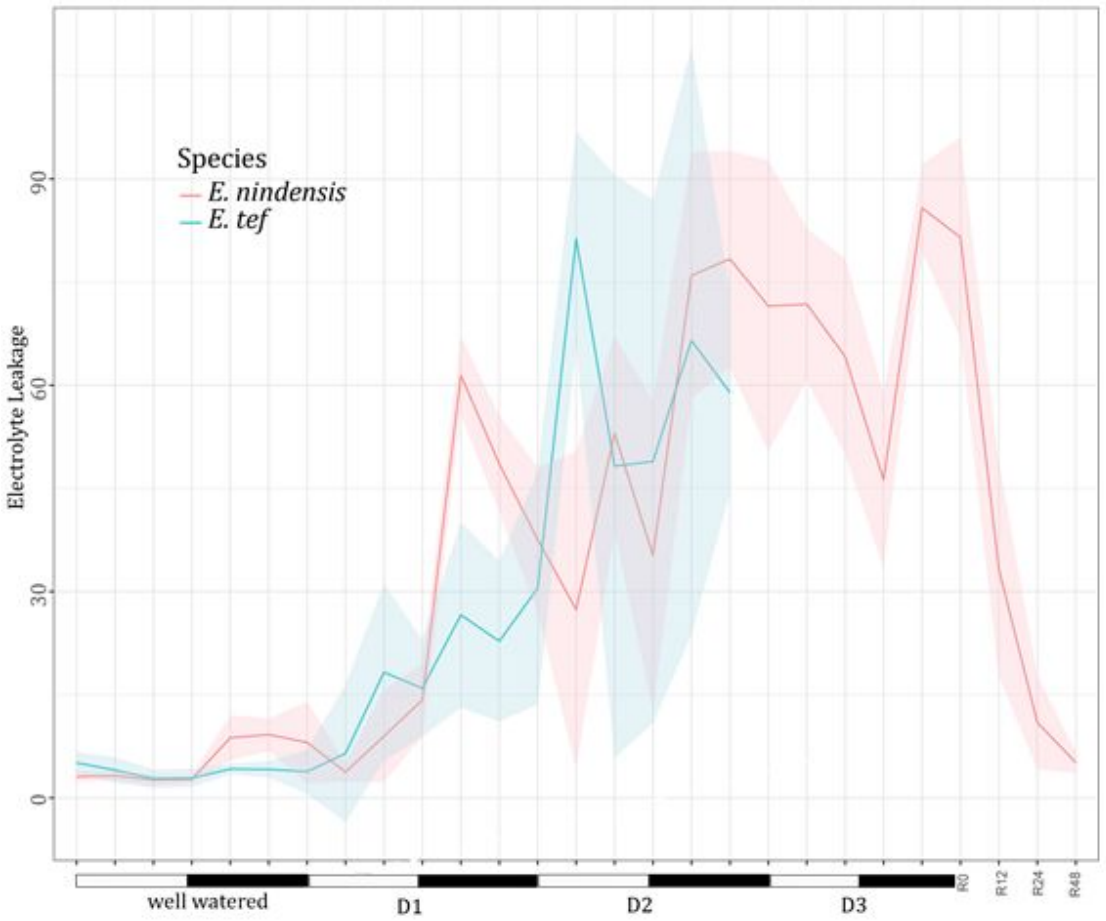
Supplemental Figure 1. Karyotype analysis of *E. nindensis*. Root tip cells were stained with DAPI revealing a karyotype of 2n=4x=40.



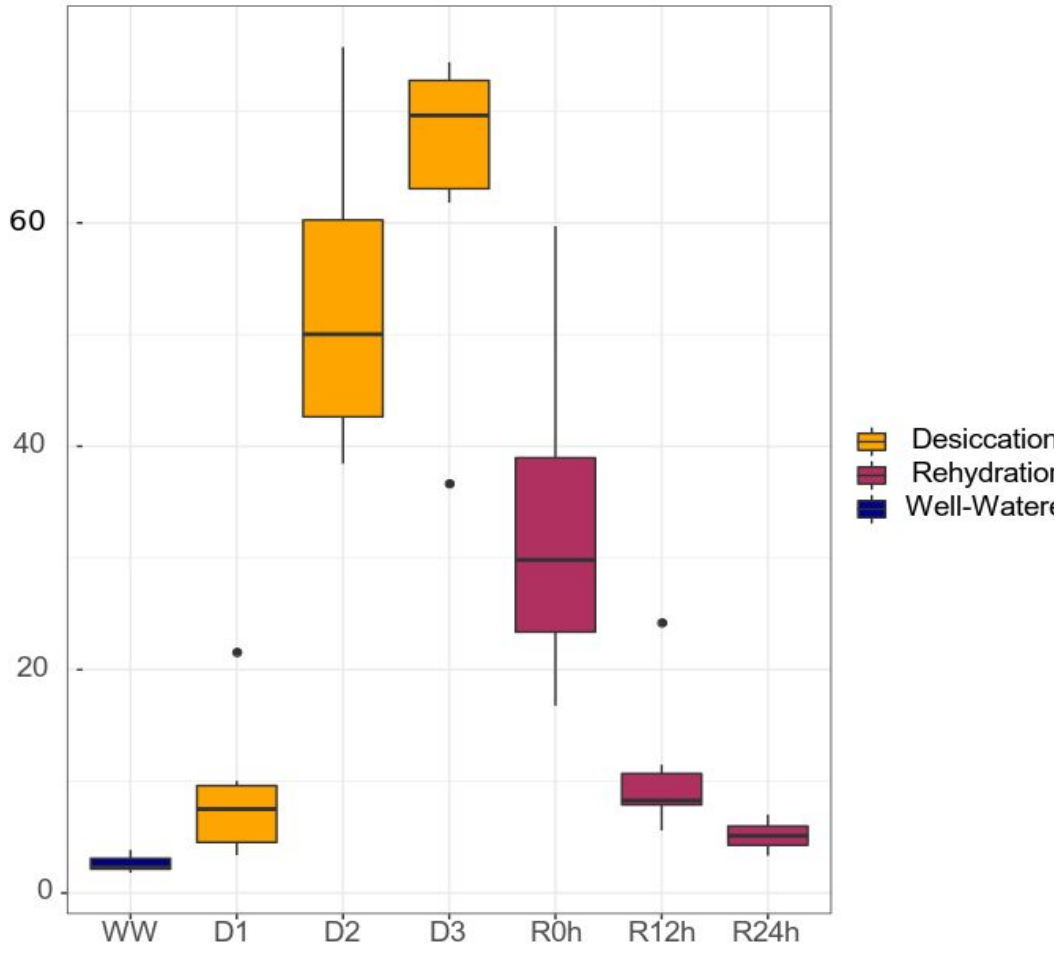
Supplemental Figure 2. Syntenic depth of *E. nindensis* compared to (a) *E. tef* and (b) *O. thomaeum.*



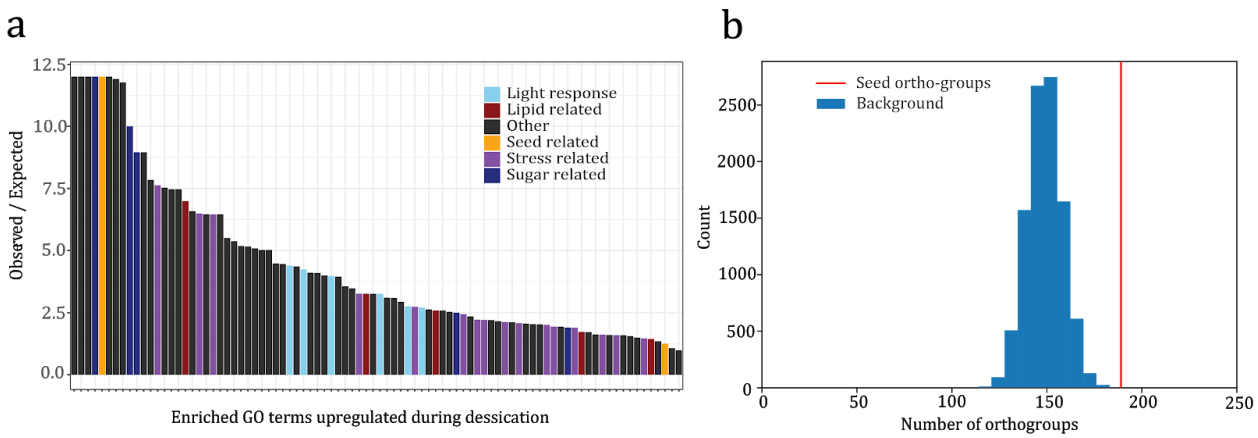
Supplemental Figure 3 Ks distribution of duplicated genes. (a) Ks of syntenic homeologs within the *E. nindensis* genome. (b) Ks of syntneic orthologs between the *E. tef* and *E. nindensis* genomes.



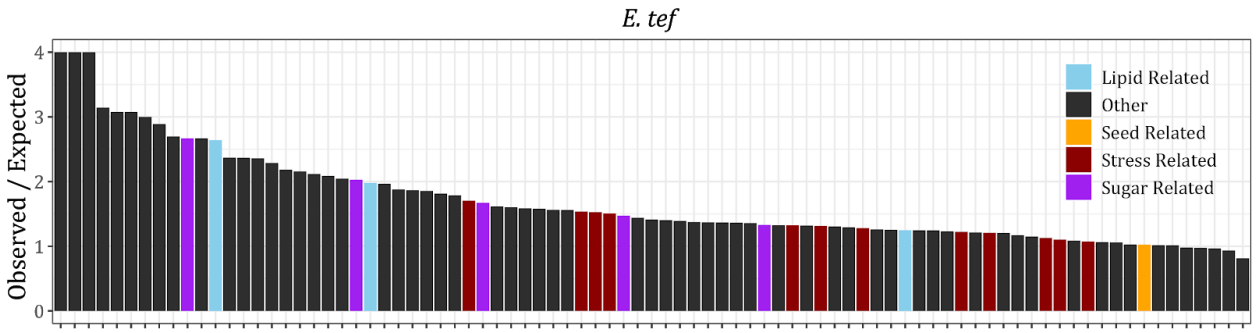
Supplemental Figure 4. Electrolyte leakage during the desiccation and rehydration timecourse of *E. nindensis* and *E. tef*. The electrolyte leakage in each sample is plotted as a percentage of the maximum leakage. The 95% confidence intervals are plotted for each timepoint.



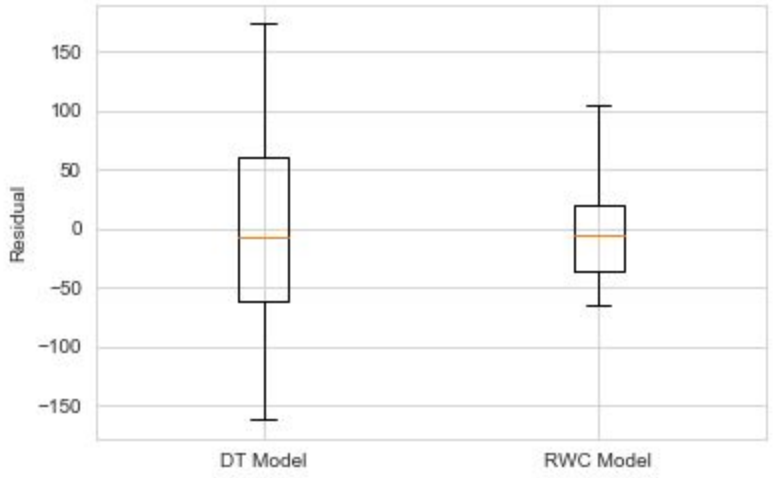
Supplemental Figure 5. Electrolyte leakage as a percentage of maximum leakage during the desiccation and rehydration timecourse of *E. nindensis*



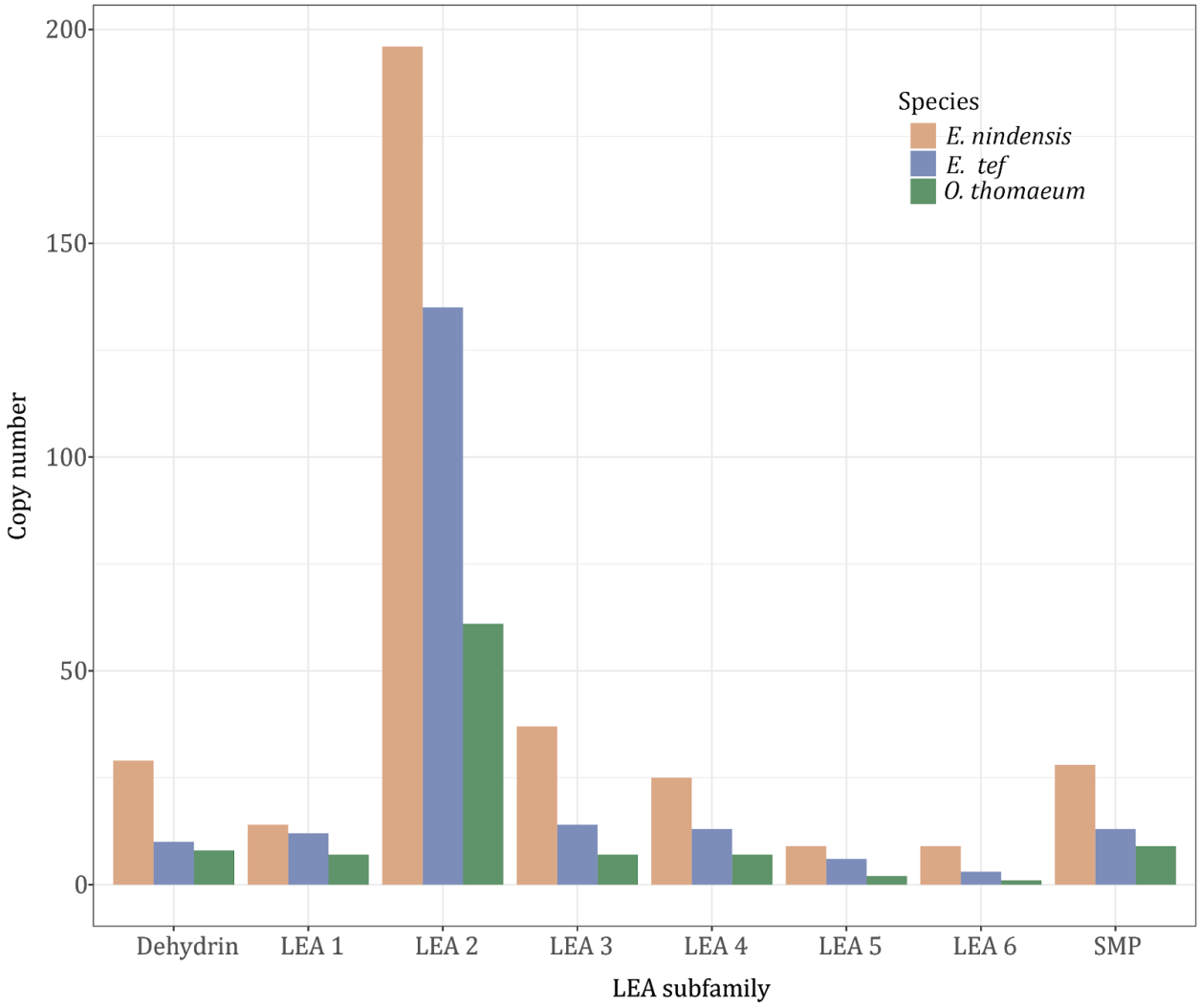
Supplemental Figure 6. Upregulation of seed pathway genes during desiccation. (a) Enriched GO terms associated with drought/desiccation. GO terms with seed or stress related functions are highlighted. (b) Enrichment of seed related genes in *E. nindensis* (red) compared to background (blue).



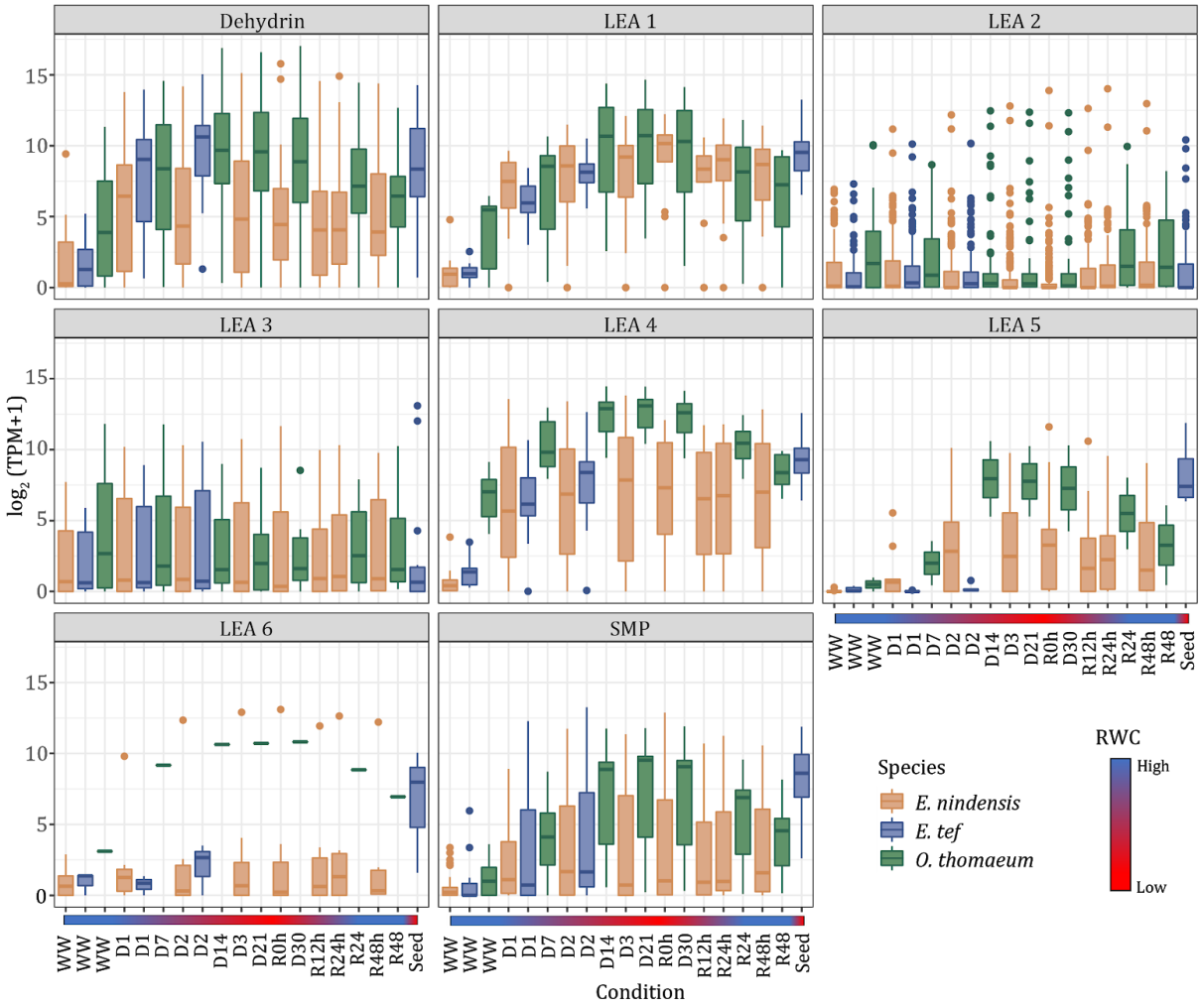
Supplemental Figure 7. Upregulated GO terms under drought in *E. tef*. Enriched GO terms involved in stress and seed development pathways are highlighted.



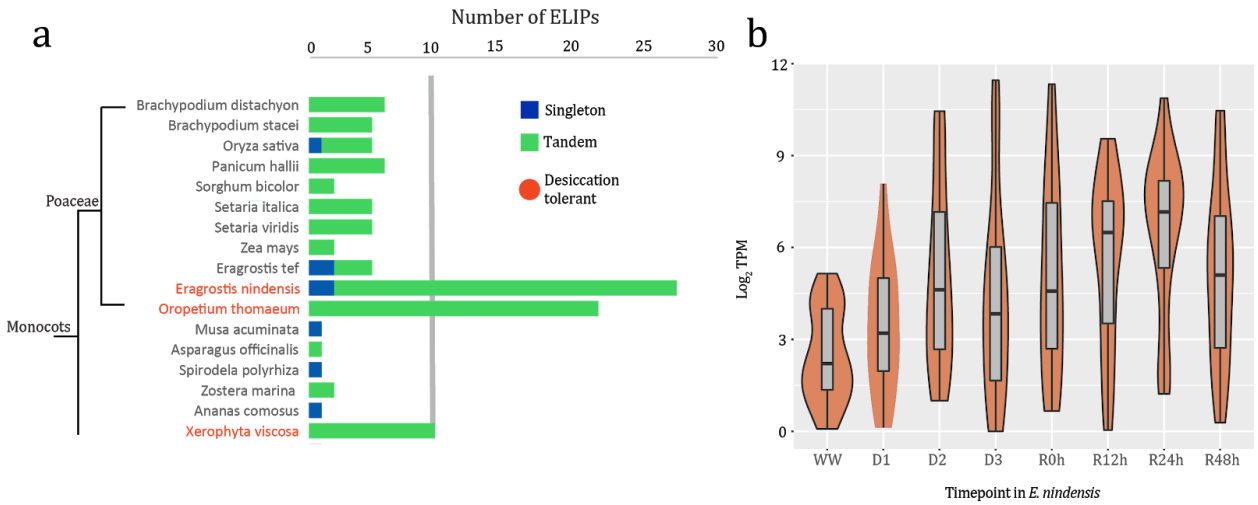
Supplemental Figure 8. Comparison of model residuals for RWC and Desiccation Tolerance (DT) models. The model residual variance for two models, one using the sample RWC and the other using the species status as desiccation tolerant or sensitive, to predict the number of seed related genes expressed. The RWC model has less residual variance and is a better fit.



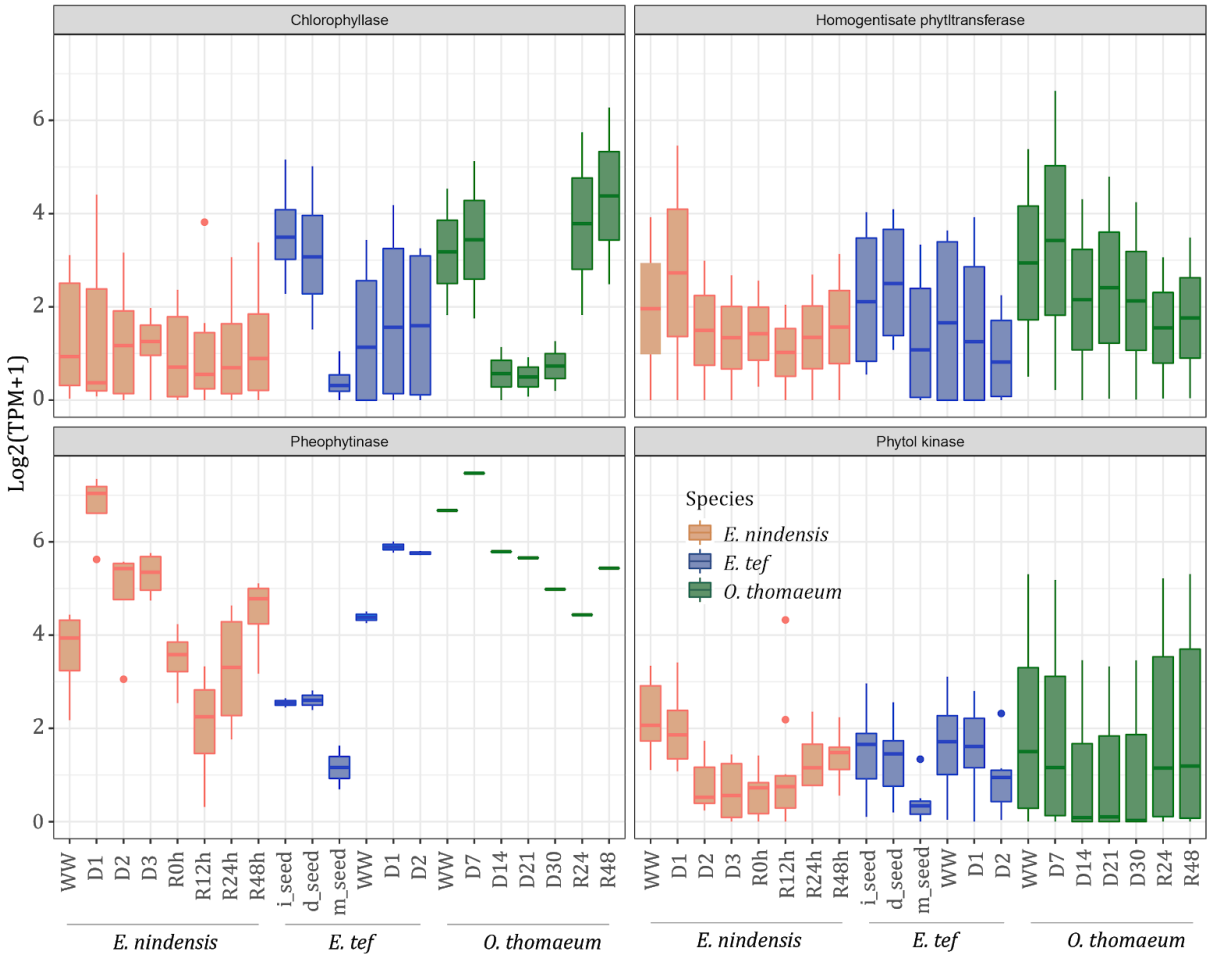
Supplemental Figure 9. LEA family composition in Chloridoideae grasses. The number of LEAs from each of the subfamilies is plotted for *E. nindensis*, *E. tef*, and *O. thomaeum*.



Supplemental Figure 10. Desiccation specific induction of LEA subfamilies. The expression of the eight subfamilies of LEAs is plotted for drought and rehydration datasets from *E. nindensis*, *E. tef*, and *O. thomaeum*. Expression of each LEA gene is plotted individually with boxplots showing the distribution within a subfamily. Drying was slower in *O. thomaeum* and the D7, D14, and D21 timepoints refer to days of drought where plants were desiccated by D14.



Supplemental Figure 11. Induction of photoprotective pathways during desiccation (a) Copy number of ELIPs in various desiccation tolerant and sensitive grasses. Tandemly duplicated ELIPs are plotted in green, and single copy or interspersed ELIPs are plotted in blue. Desiccation-tolerant species are highlighted in red. (B) Expression of ELIPs throughout desiccation and rehydration in *E. nindensis*.



Supplemental Figure 12. Unique desiccation associated expression of chlorophyll degradation enzymes in *E. nindensis.* The expression of several chlorophyll degradation enzymes is plotted for drought and rehydration datasets from *E. nindensis*, *E. tef*, and *O. thomaeum*. Expression of each gene is plotted individually with boxplots showing the distribution within an enzyme.

Supplemental Table 1. Poaceae genera containing desiccation tolerant species

Genus	Subfamily	Tribe	Example Species	Citation
Eragrostis	Chloridoideae	Eragrostideae	*E. nindensis*	(85)
Oropetium	Chloridoideae	Cyondonteae	*O. thomaeum*	(86)
Tripogon	Chloridoideae	Cyondonteae	*T. loliiformis*	(87)
Microchloa	Chloridoideae	Cyondonteae	*M. caffra*	(85)
Eragrostiella	Chloridoideae	Cyondonteae	*E. brachyphylla*	(86)
Micrachne (Brachyachne)	Chloridoideae	Cyondonteae	*M. patentiflora*	(85)
Sporobolus	Chloridoideae	Zoysieae	*S. stapfianus*	(85)
Micraira	Micrairoideae	Micraireae	*M. subulifolia*	(88)
Poa	Pooideae	Poeae	*P. bulbosa*	(89)

Supplemental Table 2. Assembly statistics for the *E. nindensis* genome.

Number of contigs	4,368
Contig N50	520 kb
Total length	986,209,651 bp
LTR composition	288 Mb (29%)
Number of gene models	116,452

Supplemental Table 3. Transcription factors with desiccation specific upregulation in *E. nindensis*

Orthogroup	AT BLAST hit	AT gene name	PFAM ID	PFAM Description
OG0002208	AT5G11260	HY5	PF00170	bZIP transcription factor
OG0001678	AT2G16770	ATbZIP23	PF00170	bZIP transcription factor
OG0000105	AT3G12250	TGA6	PF14144	Seed dormancy control
OG0000004	AT5G18270	ANAC087,	PF02365	No apical meristem (NAM) protein
OG0001862	AT1G17880	ATBTF3	PF01849	NAC domain
OG0000000	AT2G23340 AT2G33710 AT2G47520AT1G68550	DEAR3, ERF B-4, ERF B-2	PF00847	AP2 domain
OG0000041	AT1G51190 AT4G37750	PLT2, ANT	PF00847	AP2 domain
OG0002549	AT2G44730	Alcohol dehydrogenase transcription factor	PF13837	Myb/SANT-like DNA-binding domain
OG0000003	AT2G34140	CDF4	PF02701	Dof domain, zinc finger
OG0000024	AT5G25830	GATA12	PF00320	GATA zinc finger
OG0004606	AT3G20740	FIE1	PF00400	WD domain, G-beta repeat
OG0002396	AT1G32360	Zinc finger CCH type	PF00642	Zinc finger C-x8-C-x5-C-x3-H type (and similar)
OG0000390	AT1G18790	ATRKD1	PF02042	RWP-RK domain

Dataset S1. Relative water content and electrolyte leakage of *E. nindensis* and *E. tef* during dehydration timecourse. (see external file)

Dataset S2. Enriched GO terms upregulated or downregulated in E*. nindensis* under the drought/desiccation timecourse. (see external file)

Dataset S3. Enriched GO terms upregulated or downregulated in *E. tef* under the drought timecourse. (see external file)

Dataset S4. Enriched GO terms upregulated or downregulated in *E. nindensis* during the rehydration timecourse. (see external file)