New Phytologist

Research

Syntrichia ruralis: emerging model moss genome reveals a conserved and previously unknown regulator of desiccation in flowering plants

Xiaodan Zhang[1]*, Jenna T. B. Ekwealor[2,3]*, Brent D. Mishler[4,5], Anderson T. Silva[6], Li'ang Yu[1], Andrea K. Jones[1], Andrew D. L. Nelson[1] and Melvin J. Oliver[7]

[1]The Boyce Thompson Institute, Cornell University, Ithaca, NY 14853, USA; [2]Department of Biology, Utah State University, Logan, UT 84322, USA; [3]Department of Biology, San Francisco State University, San Francisco, CA 94132, USA; [4]University and Jepson Herbaria, Berkeley, CA, 94720-2465, USA; [5]Department of Integrative Biology, University of California, Berkeley, CA, 94720-2465, USA; [6]Vytelle, Lenexa, KS 66219, USA; [7]Division of Plant Sciences and Technology and Interdisciplinary Plant Group, University of Missouri, Columbia, MO 65211, USA

Authors for correspondence:
Andrew D. L. Nelson
Email: an425@cornell.edu

Melvin J. Oliver
Email: olivermj@umsystem.edu

Received: *28 November 2023*
Accepted: *5 February 2024*

New Phytologist (2024) **243:** 981–996
doi: 10.1111/nph.19620

Key words: ABA-responsive abiotic stress, angiosperms, bryophytes, comparative transcriptomics, desiccation tolerance, *Syntrichia ruralis* genome, transcriptional regulation.

Summary

- Water scarcity, resulting from climate change, poses a significant threat to ecosystems. *Syntrichia ruralis*, a dryland desiccation-tolerant moss, provides valuable insights into survival of water-limited conditions.
- We sequenced the genome of *S. ruralis*, conducted transcriptomic analyses, and performed comparative genomic and transcriptomic analyses with existing genomes and transcriptomes, including with the close relative *S. caninervis*. We took a genetic approach to characterize the role of an *S. ruralis* transcription factor, identified in transcriptomic analyses, in *Arabidopsis thaliana*.
- The genome was assembled into 12 chromosomes encompassing 21 169 protein-coding genes. Comparative analysis revealed copy number and transcript abundance differences in known desiccation-associated gene families, and highlighted genome-level variation among species that may reflect adaptation to different habitats. A significant number of abscisic acid (ABA)-responsive genes were found to be negatively regulated by a MYB transcription factor (MYB55) that was upstream of the *S. ruralis* ortholog of ABA-insensitive 3 (ABI3). We determined that this conserved MYB transcription factor, uncharacterized in *Arabidopsis*, acts as a negative regulator of an ABA-dependent stress response in *Arabidopsis*.
- The new genomic resources from this emerging model moss offer novel insights into how plants regulate their responses to water deprivation.

Introduction

In the terrestrial realm, where aridity prevails and survival is an unyielding challenge, some plants can harness an extraordinary resilience known as desiccation tolerance (DT): the ability to endure extreme water loss without suffering irreparable damage or death. Specifically, the tolerance of desiccation is defined as the ability to survive equilibration of the water potential of tissues to that of the surrounding air (Bewley, 1979). Desiccation-tolerant plants are defined by their ability to survive tissue water potential of at least −100 MPa (equilibration to a relative humidity, RH, of 50%), then recover and resume normal growth once water becomes available again (Alpert & Oliver, 2002). A suite of adaptive traits that enables plants to survive in arid or water-scarce environments can be found across the land plant phylogeny (Stark, 2017; Oliver *et al.*, 2020). It has been postulated that vegetative DT (in vegetative, rather than reproductive, tissues; VDT) was a critical requirement for plants to colonize terrestrial habitats (Oliver *et al.*, 2005). Indeed, DT mechanisms are believed to have been adapted into new developmental contexts as the need arose, such as late embryogenesis in seed plants (Oliver *et al.*, 2005). Thus, understanding how DT is regulated can provide insights into land plant evolution.

The origin of early land plants coincided with a large degree of genetic novelty (Bowles *et al.*, 2020). A prime example includes the introduction of regulators (e.g. transcription factors) of critical phytohormone pathways such as abscisic acid (ABA), which itself has numerous roles governing drought and DT mechanisms. Recent studies demonstrated the presence of ABA and its physiological effects in the red alga *Cyanidioschyzon merolae*, suggesting an early evolutionary origin of these signaling pathways (Kobayashi *et al.*, 2016). The conservation of ABA signaling through evolution, especially in stress responses, is further evidenced by studies on various plant lineages. In seed plants, ABA has been extensively studied for its role in stress tolerance, particularly in response to drought and salinity (Cutler *et al.*, 2010).

*These authors contributed equally to this work.

The fundamental components of ABA signaling, such as PYR/PYL/RCAR receptors, PP2C phosphatases, and SnRK2 kinases, have also been identified across a wide range of plant species (Umezawa *et al.*, 2010). Since the emergence of land plants, these gene families have expanded, likely allowing for adaptation to increasingly diverse environments and developmental contexts (Donoghue *et al.*, 2021).

Bryophytes hold an important phylogenetic position as three of the four main living clades of land plants, along with the tracheophytes. Thus, the establishment of genetically tractable bryophyte models that exhibit VDT has been instrumental in understanding the significance of the genomic and molecular innovations that coincided with land plant emergence. Mosses, one of the bryophyte lineages, are excellent model systems for studying VDT (Cove *et al.*, 2009; Rensing *et al.*, 2020; Naramoto *et al.*, 2022; Li *et al.*, 2023; Yadav *et al.*, 2023), particularly as the molecular underpinnings of VDT may be shared across land plants (Phillips *et al.*, 2002; Oliver *et al.*, 2020). The principal bryophyte model that is widely used to study growth and development, *Physcomitrium patens* (Lang *et al.*, 2018; Rensing *et al.*, 2020), while drought-tolerant (Frank *et al.*, 2005), is generally sensitive to dehydration stress (Mishler & Oliver, 2009; Koster *et al.*, 2010). Although *P. patens* has the capability for DT if treated with ABA (Oldenhof *et al.*, 2006) or extremely slow-drying rates (Greenwood & Stark, 2014; Xiao *et al.*, 2018), it is not an ideal model for understanding tolerance to desiccation.

Two closely related desiccation-tolerant mosses, *Syntrichia caninervis* and *Syntrichia ruralis* (previously known as *Tortula caninervis and Tortula ruralis*; Schonbeck & Bewley, 1981), have emerged as important model systems for understanding mechanisms of abiotic stress tolerance and their evolution. Both species can survive internal water potentials well below −100 MPa and have served as a focus for research into the ecological, reproductive, physiological, biochemical, and molecular aspects of VDT in bryophytes (Stark *et al.*, 2005a,b; Proctor *et al.*, 2007; Oliver, 2009; Zhang *et al.*, 2011; Stark, 2017; Coe *et al.*, 2021). *Syntrichia* is a diverse genus of mosses that occur in dryland habitats across the world and is one of the most ecologically dominant groups of mosses across western and northern North America (Brinda *et al.*, 2021). Both *S. caninervis* and *S. ruralis* can dominate biological soil crust (biocrust) communities of North American drylands (Belnap & Lange, 2003). However, while *S. caninervis* can be considered a dryland biocrust specialist, *S. ruralis* grows in a wider range of habitats that are not restricted to the biocrust community and span a range of moisture gradients (Mishler, 2007). Thus, despite their relatively recent divergence (*c.* 5 Ma; Jauregui-Lazo *et al.*, 2023), the phenotypic variation within these two species argues that they are an excellent comparative resource for understanding how desiccation tolerance is experienced and regulated in land plants.

We report here the addition of a chromosomal-level genome assembly for *S. ruralis* derived from a clonally propagated male gametophyte and the insights this genome can provide to plant abiotic stress responses. Using comparative genomics and transcriptomics, we explored both the evolutionary dynamics of the two *Syntrichia* genomes and the ways in which the two mosses regulate their molecular responses to desiccation and rehydration. Specifically, we tested the hypothesis that the ability of *S. ruralis* to inhabit a broad range of stressful habitats will be reflected by genomic and transcriptomic differences relative to *S. caninervis*, which has a more restrictive habitat range. We used these comparisons to uncover critical elements that govern the plasticity of VDT in these two species. We then expanded our comparisons to flowering plants (i.e. *Arabidopsis thaliana*), where we used information from *S. ruralis* to identify a novel, but deeply conserved, ABA-associated transcription factor (TF). Our results highlight the importance of developing bryophyte models and the conservation of DT-associated pathways across land plants.

Materials and Methods

Syntrichia ruralis cultures

Syntrichia ruralis (Hedw.) F. Weber & D. Mohr shoots originated from a single specimen (Brinda 9108) collected from the Bow River Valley in Calgary Alberta (51.098875°N, 114.281461°W) and vouchered in the University of Nevada, Las Vegas herbarium (UNLV). In the laboratory, *S. ruralis* shoots, free of visible contamination with algae and protists, were vegetatively propagated on sterile fine sand collected from a dune near Moab, Utah (93.9% sand, 5.5% silt, and 0.6% clay with a pH of 8.4), and watered on alternating weeks with sterile distilled water and with a 30% inorganic nutrient solution (Hoagland & Arnon, 1938). This specimen is male (has antheridia reproductive structures) and was selected for its healthy appearance and sustained growth under culture conditions. Cultures were placed in a growth chamber set to a 12 h photoperiod (20°C : 8°C, light : dark), at *c.* 90 μmol m^{-2} s^{-1} photosynthetically active radiation (PAR). The single clonal line used for genome sequencing had been subcultured and grown to maturity through at least five asexual generations. Several gametophytes from the original clonal line were used to expand and establish subcultures for generating sufficient material for isolation of RNA for construction of transcriptome libraries for transcript abundance studies (RNA-seq). Expansion of cultures was achieved by isolating individual shoots following branching as well as by fragment regeneration. Cultures were examined every subculturing for lack of visible contamination by stereomicroscope.

Desiccation was achieved using a standardized slow dehydration protocol and drying curve (as described by Oliver (1991)) in which gametophytes in small wire baskets were placed over saturated ammonium nitrate (67% RH) in a closed glass desiccator placed in the same incubator as the moss cultures at 20°C. Under these conditions, the gametophytes reached a stable weight (equilibrium), measured gravimetrically, at 6 h and a water potential of −54 MPa within the light period (100 μmol m^{-2} s^{-1}) of the day : night cycle. Equilibrium drying was chosen to achieve desiccation, even though the moss can survive cellular water potentials of − 600 MPa (Alpert & Oliver, 2002), as it provides a highly reproducible drying rate and a precise level of dehydration, in this case −54 MPa, which is sufficient to induce quiescence. Rehydration was achieved by placing the desiccated gametophytes in a culture dish in the incubator at 20°C in the light and adding sufficient distilled water to ensure full hydration.

Genomic DNA isolation and Chicago library preparation and sequencing

Genomic DNA isolation, library preparation, sequencing, and assembly were conducted by Dovetail Genomics (Scotts Valley, CA) and as described for the *S. caninervis* genome (Silva *et al.*, 2021). Briefly, genomic DNA was isolated from 1 to 2 g of powdered flash frozen gametophytes using a standard CTAB-based procedure (Doyle & Doyle, 1987) and high molecular weight genomic DNA was precipitated, resuspended in Qiagen Buffer G2 with RNAse A and after incubation at 50°C for 30 min, the DNA was purified using Qiagen Genomic-tips (Qiagen). Chicago genomic DNA libraries were prepared as described by Putnam *et al.* (2016). The libraries were sequenced on an Illumina HiSeq platform to produce 389 million 150 bp, paired-end reads, which provided 153.2× physical coverage of the genome (1–100 kb pairs). A Dovetail HiC library was prepared as described by Lieberman-Aiden *et al.* (2009). For each library, chromatin was fixed in place in the nucleus by incubation of the gametophytes in 1% formaldehyde for 15 min under vacuum. The fixed chromatin was extracted from the treated gametophytes using the Dovetail TM Hi-C Kit (Dovetail Genomics, Scotts Valley, CA, USA). After digestion with DpnII, the 5″ overhangs filled in with biotinylated nucleotides and ligated. After ligation, cross-links were reversed, and the DNA purified from protein. The DNA was then sheared to *c.* 350 bp mean fragment size, and sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotin-containing fragments were isolated using streptavidin beads before PCR enrichment of each library. The libraries were sequenced on an Illumina HiSeq platform to produce 224 million 2 × 150 bp paired-end reads, which provided 23 246.36 × physical coverage of the genome (10–10 000 kb pairs).

Syntenic comparisons

To identify syntenic gene blocks between *S. ruralis*, *S. caninervis*, and *Ceratodon purpureus*, an all-against-all BLASTP analysis was conducted, using a stringent threshold (e-value $< 1 \times 10^{-5}$ and top five matches). Syntenic blocks were defined based on the presence of a minimum of five synteny gene pairs, employing the MCSCANX package with default settings (Wang *et al.*, 2012). Subsequently, adjacent blocks were merged, and large syntenic blocks were chosen for further analysis. These sizable syntenic blocks were utilized to infer the chromosome evolution and explore the syntenic relationships between *S. ruralis*, *S. caninervis*, and *P. patens*.

Phylogenetic analyses

Late embryogenesis abundant (LEA) and early light-induced protein (ELIP) genes were identified in *S. ruralis* and *S. caninervis* by HMMER (a software that uses probabilistic models called profile hidden Markov models; profile HMMs; Potter *et al.*, 2018) and NCBI BLAST. For the LEA genes, profile HMMs of eight LEA subfamilies (LEA_1, PF03760; LEA_2, PF03168; LEA_3, PF03242; LEA_4, PF02987; LEA_5, PF00477; LEA_6, PF10714; DHN, PF00257; and SMP, PF04927) were retrieved from the Pfam database (Finn *et al.*, 2014) and used to identify LEA domain-conserved proteins using `hmmsearch` in HMMER. Subsequently, sequences from *S. ruralis* and *S. caninervis* were aligned using CLUSTAL OMEGA (Sievers *et al.*, 2011). Then, the phylogenetic tree was constructed in a maximum likelihood framework with MEGA5 (Tamura *et al.*, 2011), and the resulting tree was visualized by Interactive Tree of Life online tool (iTOL; Letunic & Bork, 2021). For the ELIP family, profile HMMs of ELIPs were downloaded from the InterPro database (Hunter *et al.*, 2009) and used to screen for ELIP-conserved proteins with `hmmsearch`. The ELIPs identified were subject to the same BLAST confirmation and phylogenetic analyses as the LEA gene family.

RNA-sequencing

For RNA-sequencing, replicate gametophytic samples, in triplicate, were collected for each of six treatments: fully hydrated controls (Ctrl), slow-dried gametophytes (D; desiccated samples at −54 MPa), samples rehydrated for 1 h after slow-drying (R; fully rehydrated), heat stress at 30°C for 90 min, heat shock at 35°C, and cold treatment at 4°C for 90 min. The desiccated samples were chosen to ensure full capture of sequestered transcripts induced by desiccation (Wood & Oliver, 1999) and the 1 h rehydration time was chosen to fully capture the complete rehydration response (Oliver, 1991). Temperature stress growth conditions were similar to those in Silva *et al.* (2021). The gametophytes for this part of the study were obtained from the same clonal cultures used for the genome sequencing. RNA was extracted using the RNeasy (Qiagen) kit with the RLC buffer following the manufacturer's recommended protocol. RNA isolates were treated with DNase1 and cleaned using the DNA-free RNA Kit (Zymo Technologies, Irvine, CA, USA). RNA quality was assessed by the use of a fragment analyzer (Advanced Analytical Technologies, Ankeny, IA, USA) and concentration determined by use of a Nanodrop Spectrophotometer (ThermoFisher, Waltham, MA, USA). RNA libraries were created and individually barcoded from 2.7 μg of template total RNA utilizing the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA) as described in the manufacturer's recommended protocol. Libraries were pooled in groups of 12 and sequenced (12 samples per lane) on an Illumina HiSeq 2500 ultra high-throughput DNA sequencing platform (Illumina) at the DNA Core facility at the University of Missouri, Columbia, MO, USA.

Transcriptome analysis

The transcriptome analysis was conducted using the RNA-seq workflow `RMTA (Read Mapping Transcript Assembly`; Peri *et al.*, 2020), which first trimmed the adapter sequences using CUTADAPT with a minimum overlap of 10 bp and eliminated low-quality sequences from the raw data using TRIMMOMATIC with a sliding window of 4 : 15 and a minimum length of 36 bp. After preprocessing, the high-quality reads were aligned to the *S. ruralis* genome using HISAT2 (Pertea *et al.*,

2016), and the number of read counts mapping back to each gene was quantified using FeatureCounts (Liao *et al.*, 2014) with default parameters. Read normalization was performed with the DESeq2 package (Love *et al.*, 2014) in R (R Development Core Team, 2021), employing the Wald test with parameters for normalization and dispersion estimation set to DESeq2's defaults. For identifying differentially abundant transcripts, a threshold was set to an adjusted *P*-value < 0.05 and the absolute value of the $\log_2$ (fold change) of > 1, where treatments (Ctrl, D, and R) were compared pairwise. To visualize the overlap of genes with differential transcript abundance between treatments, Venn diagrams were generated by the R package ggVennDiagrams (Gao *et al.*, 2021). Gene ontology (GO) and enrichment analyses were performed using TopGO (Alexa & Rahnenführer, 2009). Syntelogs (syntenic homologs) were identified using SynMap2, and then 1 : 1 syntelogs were extracted from the final syntenic gene set output provided by SynMap2 in CoGe (the analysis can be recapitulated at the following link: https://genomevolution.org/r/1qb33; Haug-Baltzell *et al.*, 2017). UpSet plots were generated using the UpSetR package in R (Conway *et al.*, 2017). The pheatmap R package was used to process the log-transformed and normalized TPM data, with genes and samples as rows and columns, respectively (Kolde & Kolde, 2015).

Transcription factor analysis

Transcription factors (TFs) in *S. ruralis* were identified by querying the protein sequences against the PlantTFDB database (Tian *et al.*, 2020). The upstream region (2 kb) of the 3045 genes with differential abundance between Ctrl and D conditions was extracted using the BEDTools suite (Quinlan & Hall, 2010). These sequences served as the input data for performing an Analysis of Motif Enrichment (AME; McLeay & Bailey, 2010) with the default parameters, using the *Arabidopsis thaliana* DAP motif database (O'Malley *et al.*, 2016) to identify overrepresented motifs and corresponding *S. ruralis* TFs. Pairwise abundance correlation analysis was performed using Spearman's rank correlation coefficient in R, considering expression data across all treatment conditions, including Ctrl, D, and R. Only pairs with a Spearman's *rho* value of |0.75| and a *P*-value < 0.05 were selected for network construction. The regulatory network, illustrating the interactions between TFs and putative target genes, was generated by Cytoscape (Saito *et al.*, 2012). We utilized the built-in 'prefuse force directed layout' for network visualization.

Arabidopsis growth and genotyping conditions

For growth on soil (Cornell Mix), *Arabidopsis thaliana* wild-type (WT; Col-0) and *Atmyb55* mutant (Insertion ID: SAIL_555_H05, seed stock ID: CS823517; ABRC) T-DNA insertion lines were stratified at 4°C for 2 d and then grown in walk-in growth chambers at the Boyce Thompson Institute under standard Arabidopsis long-day growth conditions (16 h : 8 h, 22°C, light : dark). Plants were grown until flowering, whereupon genomic DNA was extracted from the youngest rosette leaves. DNA was extracted using the CTAB-based method as described by Doyle & Doyle (1987). To verify the T-DNA insertion, primers were designed using the T-DNA primer design website (http://signal.salk.edu/tdnaprimers.2.html) and used to genotype by PCR. Wild-type reaction = LP + RP, T-DNA insertion (mutant) reaction = RP + LB, where LP 5′-3′: TTTTCCTCTTATGTGGG AGGG; RP 5′-3′: AGCATCGGAATGAAAAATCG, and LB 5′-3′: ATTTTGCCGATTTCGGAAC. Seeds from individual confirmed homozygous mutant lines were collected and stored separately.

ABA treatment and MS plate growth conditions

Wild-type Columbia-0 (Col-0) and *Atmyb55* homozygous mutant seeds were stratified at 4°C for 48 h to synchronize germination and then sown on Murashige and Skoog (MS) agar plates supplemented with varying concentrations of ABA (0, 0.1, 0.3, and 0.5 μM). Seedlings on plates were then grown under 16 h : 8 h, 22°C, light : dark. Germination, defined as radicle emergence, was monitored and recorded at 12 h intervals over 6 d to evaluate the mutants' sensitivity to ABA during germination. To compare plant responses to ABA post-germination, WT and *Atmyb55* seedlings were grown on 0 μM ABA MS-agar plates for 1 wk (roughly four-leaf stage) under the same conditions. Then, the 1-wk-old seedlings were transferred to plates containing 0.3, 0.5, 1, and 2 μM ABA to assess the post-germination root growth response to ABA. Root elongation was assessed by capturing images weekly for 2 wk using an Epson V800 scanner, and the images were analyzed in ImageJ (Schneider *et al.*, 2012).

Results and Discussion

De novo assembly of *S. ruralis* genome

The draft genome of *S. ruralis* was assembled into 381.24 Mb, which consisted of 3211 scaffolds with an N50 value of 24.06 Mb. The *S. ruralis* reference genome represents the fourteenth publicly available moss genome, and so far, all have small genomes relative to vascular plants (Supporting Information Fig. S1). The 12 largest scaffolds, representing 80% of the assembly (305.19 out of 381.24 Mb), were considered near-chromosome level assemblies, with the consensus plant telomere repeat (TTTAGGG, Fajkus *et al.*, 2019; Fajkus *et al.*, 2021) present at eight of the 24 scaffold ends. The BUSCO analysis indicated 95% genome completeness of *S. ruralis* and recovered 404 of the 425 BUSCO groups in the Viridiplantae database (Table 1). The size of the 12 *S. ruralis* chromosomes ranged from 15.7 to 48.2 Mb (Table S1), with the largest being the putative sex chromosome (Chr. 12; Fig. 1). Of the additional 3199 unplaced scaffolds, 14 were > 1 Mbp. The average GC content for the 12 proposed chromosomes was 41.28%, whereas the average for the 14 largest unplaced scaffolds was 59.68%.

A total of 27 065 genes were predicted in the draft genome including 21 169 gene models associated with the 12 chromosomes and 5896 genes with the 3199 unplaced scaffolds (Table 1). The number of genes per chromosome ranged from 857 to 2888. Despite being the largest, Chr. 12 had the fewest

Table 1 Summary of *Syntrichia ruralis* genome assembly and annotation.

Genome assembly size without symbionts	305.19 Mb
Genome assembly size with symbionts	381.24 Mb
Number of chromosome level scaffolds	12
Number of total scaffolds	3211
Number of symbionts scaffolds	3199
Longest chromosomes	48.23 Mb
Scaffold N50 length	24.06 Mb
GC content	44.6%
Genome BUSCO assessment	95%
% ambiguous bases (Ns)	0.7%
Number of predicted gene models (likely bacterial derived)	21 169 (5896)
Repetitive regions in genome	45.1%
Mean transcript length (mRNA)	2398
Mean coding sequence length (CDS)	1231 bp
GC content of coding sequences	0.5362
Average number of exons per gene	5
Mean exon length	279 bp
Mean intron length	269 bp
InterProScan	16 393
Gene Ontology	14 252

predicted genes (Table S1; Fig. 1). Given their high GC content and our inability to assemble them with the rest of the chromosomes, we hypothesized that the unplaced chromosomes might be bacterial symbiont derived (Table S2). To test this, the 5896 genes from the unplaced scaffolds were used as queries in a BLASTP against the NCBI non-redundant nucleotide database. Significant hits were recovered for 5765 (97.8%) genes. Of these, 5272 genes (91.45%) were bacterial-associated, 343 genes (5.95%) were plant-related, and the rest (150, or 2.5%) were associated with fungi and single-cell eukaryotes (Table S3).

Evolution of the *S. ruralis* genome

Conservation of genome structure and collinearity between species can provide insight into major evolutionary events over history. First, CoGe's SynMap2 tool (Lyons & Freeling, 2008; Haug-Baltzell *et al.*, 2017) was used to assess synteny between close relatives *S. ruralis* and *S. caninervis*, which last shared a common ancestor *c.* 5 Ma (Jauregui-Lazo *et al.*, 2023; Fig. 2a). In addition to some smaller inversion events along *S. ruralis* Chrs. 1, 2, 9, and 10, this pairwise comparison also uncovered a chromosome rearrangement (either a chromosomal fusion or fission event) between *S. ruralis* Chr. 11 and *S. caninervis* chromosomes 2 and 8 (Fig. 2a). To infer the directionality in this chromosomal rearrangement, both mosses *C. purpureus* and *P. patens* were added to three-way genome-wide synteny comparisons with MCScan (Figs 2b, S2; *P. patens*). The inclusion of *Ceratodon purpureus* as an outgroup, which is reported to have last shared a common ancestor with *Syntrichia c.* 180 Ma (Jauregui-Lazo *et al.*, 2023) and thus is more closely related to *Syntrichia* than is *P. patens* (Zhong *et al.*, 2014), revealed that Chr. 11 in *S. ruralis* likely resulted from a chromosomal fusion

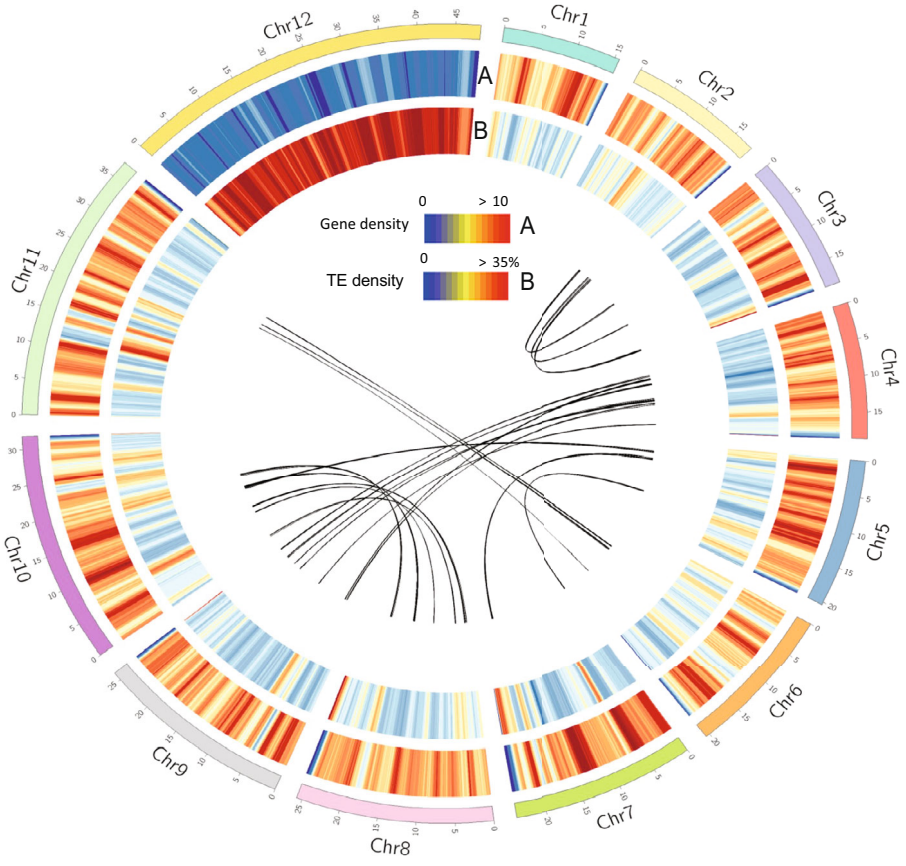


Fig. 1 Characteristics of the *Syntrichia ruralis* genome. From outer to inner: The 12 chromosomes; (A) gene density in 100-kilobase (kb) genomic regions; (B) transposable element (TE) density in 100-kb genomic regions; the lines in the center reflect inter-chromosomal genomic collinearity, as defined by MCScan.

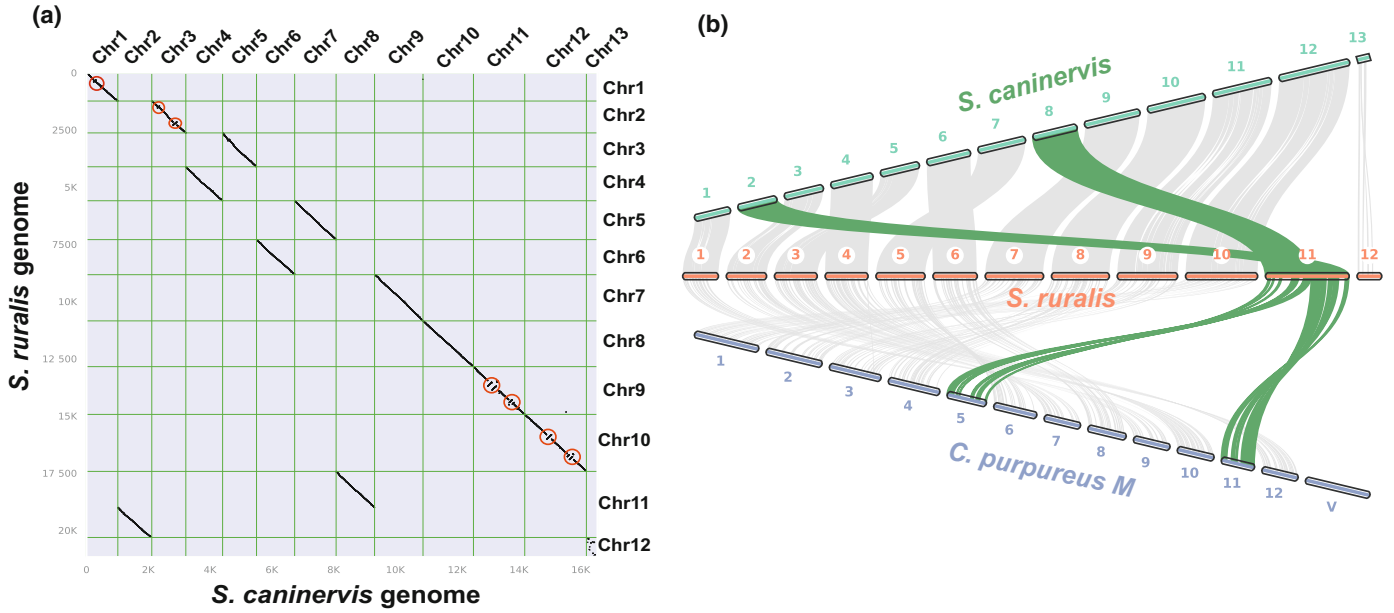


Fig. 2 Comparative genomics between *Syntrichia ruralis*, *Syntrichia caninervis*, and *Ceratodon purpureus*. (a) Dot plot graph of genome synteny revealed genome duplication and small-scale inversions between *S. ruralis* and *S. caninervis*. Several inversions (highlighted in red circles) were observed in *S. ruralis* chromosomes 1, 2, 9, and 10. (b) Large-scale genome conservation between *S. ruralis*, *S. caninervis* and *C. purpureus* generated by McScan. Each node (connection point) represents a genomic region, and the edges (lines) connecting the nodes represent syntenic blocks or collinear regions of syntenic genes. The length and thickness of the edges indicate the size and similarity of the syntenic blocks, respectively. Conserved syntenic blocks are depicted in shared colors. The green edges reflect a chromosomal rearrangement event (fusion event) between *S. caninervis* Chr. 2 and 8 and *S. ruralis* Chr. 11.

event of *S. caninervis* Chrs. 2 and 8 sometime after these two species last shared a common ancestor.

In haploid-dominant plants with separate sexes, sex is expressed in the gametophyte life stage and determined by a single sex-specific chromosome: U for females and V for males (Coelho *et al.*, 2018). In *S. ruralis*, Chr. 12, which is the largest in size (48.2 Mb; *c.* 1.3 times bigger than the largest autosome) but has the fewest number of genes (857 vs 2888 in Chr. 11), is proposed to be the putative sex chromosome, specifically the V (male) chromosome since the specimen used was phenotypically male. A previously identified gametolog (an allele of a gene that is shared between the non-recombining regions of the U and V sex chromosomes; Singh *et al.*, 2023) from *S. caninervis* (gene id: Sc_g00229; Ekwealor *et al.*, 2017; Silva *et al.*, 2021) was mapped to gene Sr_g00673 on Chr. 12 with 88% identity (e-value = 0), supporting this conclusion. Comparative analysis revealed that there were significantly fewer syntenic blocks between the *S. caninervis* U and *S. ruralis* V sex chromosomes than there were between homologous autosomes of these species, suggesting the presence of sex-specific genes on the sex chromosomes. It is also possible the observed lack of synteny may be due, at least in part, to algorithmic difficulty of McScan in aligning the high number of repetitive sequences on these two chromosomes. To address this possibility, a separate examination of the protein-coding sequences using an all-vs-all protein Blast (Blastp; e-value = 10 e−10, identity > 80%) revealed a small portion of homologous genes between the sex chromosomes of *S. ruralis* and *S. caninervis* (139 out of 857 genes in *S. ruralis*; Table S4), highlighting the stark differences between these two chromosomes. In *C. purpureus*, the U and V sex chromosomes share no obviously syntenic regions with each other even within the species (Carey *et al.*, 2021). These results suggest that the male V sex chromosome of *S. ruralis* and the female U chromosome of *S. caninervis* may have undergone distinct evolutionary processes that have resulted in differences in their gene content and structure, perhaps due to both differences in species and differences in sex. In the future, sequencing the V sex chromosome of *S. caninervis* and the U chromosome of *S. ruralis*, along with sex chromosomes of other *Syntrichia* species, will be useful for comparative purposes.

Transcriptomic response to desiccation and rehydration in *S. ruralis*

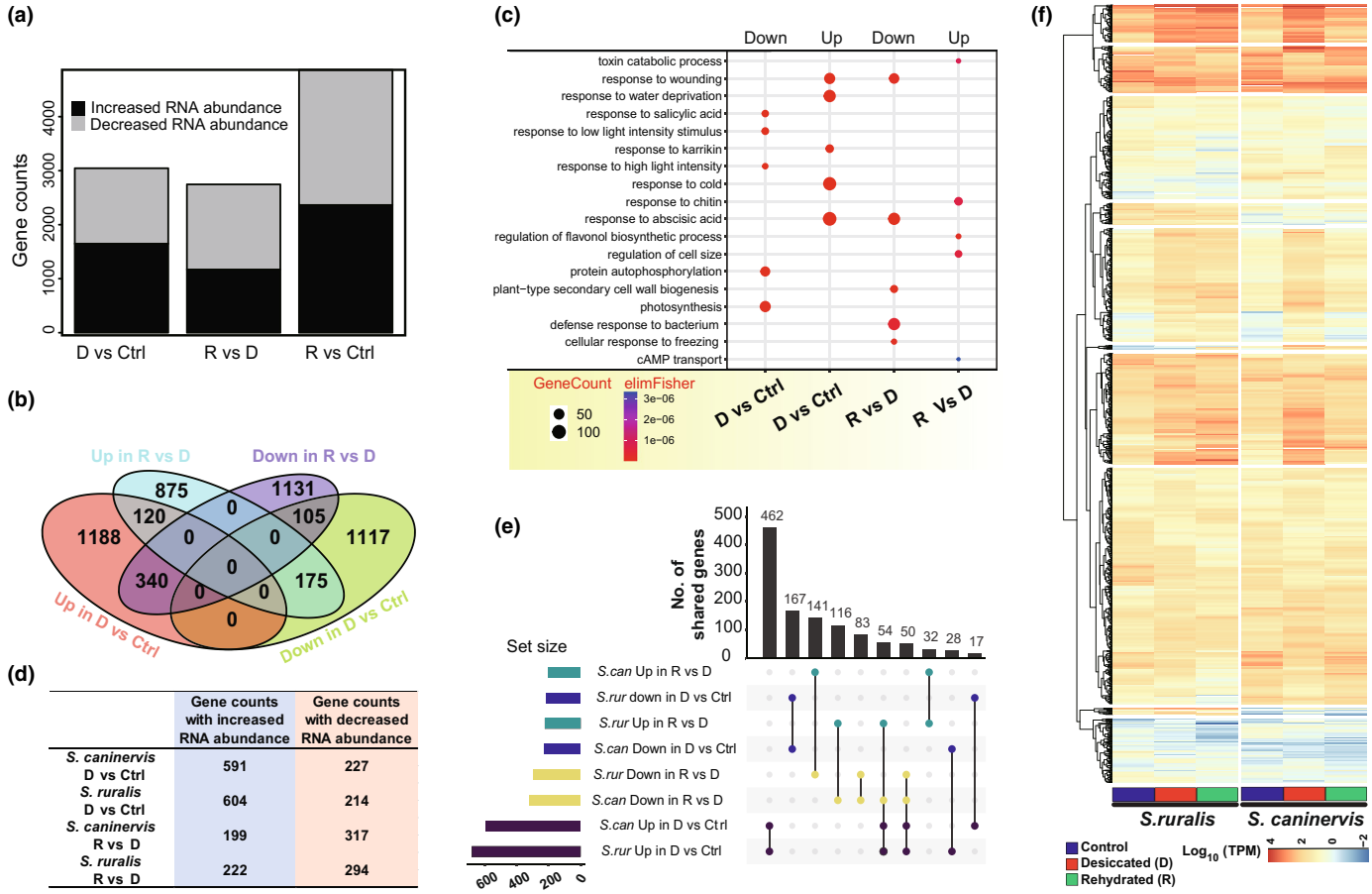
To investigate the potential molecular mechanism of desiccation tolerance of *S. ruralis*, a transcriptomic analysis was conducted under a variety of conditions (Table S5) including fully hydrated controls (Ctrl), desiccated (abbreviated as D), and rehydrated (abbreviated as R). An average of 16 million reads were sequenced per replicate ($n = 3$) per condition. Following QC, read mapping, and assigning reads to features (see Materials and Methods section; Table S5), a single replicate was removed due to poor clustering via both principal component analysis and pairwise Euclidean distance (Fig. S3). As observed increases in transcript abundance may reflect changes in RNA stability rather than gene expression, we use the more general term 'abundance' rather than 'expression' below.

A series of pairwise comparisons were then performed to examine the *S. ruralis* response to dehydration and rehydration. In Ctrl

vs D, D vs R, and Ctrl vs R comparisons, 3045, 2746, and 4867 genes showing differential transcript abundance were identified, respectively (adjusted *P*-value < 0.001, $|\log_2$ fold change$| > 2$). Specifically, when comparing D to Ctrl, transcripts for 1648 genes increased and transcripts for 1397 genes decreased in RNA abundance (Fig. 3a). When comparing R to D, transcripts for 1170 genes increased, whereas 1576 genes decreased in RNA abundance (Fig. 3a). While most of these genes were context-specific in their abundance change, a subset showed contrasting abundance patterns when comparing the transition from Ctrl to D and then to R (Fig. 3b). Transcripts representing 340 genes increased in abundance in D but decreased in abundance in R. By contrast, transcripts for 120 genes were increased in RNA abundance in both Ctrl vs D and D vs R comparisons, indicating that these genes may be necessary for both the dehydration and rehydration responses in *S. ruralis*.

Gene Ontology enrichment analysis was performed on genes showing differential RNA abundance in the Ctrl→D→R transition. Genes with increased RNA abundance in D (compared with control) were enriched in the GO terms: *response to abscisic acid*, *response to cold*, r*esponse to water deprivation*, *response to wounding*, and *response to karrikin* (Fig. 3c). The enrichment of these GO terms suggests an activation of specific stress response pathways in *S. ruralis* under dehydration conditions. For example, the response to karrikin, a group of compounds known for their role in smoke-induced seed germination and stress response in plants, is noteworthy (Nelson *et al.*, 2010). This either implies a unique adaptive mechanism in *S. ruralis*, potentially mirroring evolutionary adaptations to fire-prone environments, or that the karrikin-related pathways are ancient and adapted for different responses across land plants.

Interestingly, the *response to ABA* GO term was enriched in transcripts with significantly reduced abundance in rehydration, suggesting that these plants were no longer perceiving stress. The top five enriched GO terms for genes with reduced transcript abundance in D vs Ctrl were *photosynthesis*, *response to salicylic*



	Gene counts with increased RNA abundance	Gene counts with decreased RNA abundance
S. caninervis D vs Ctrl	591	227
S. ruralis D vs Ctrl	604	214
S. caninervis R vs D	199	317
S. ruralis R vs D	222	294

Fig. 3 Genes that display a change in transcript abundance in response to dehydration and rehydration of *Syntrichia ruralis and Syntrichia caninervis*. (a) The number of *S. ruralis* genes showing differential transcript abundance between two treatments according to a fold expression cutoff of ≥2 and a False Discovery Rate (FDR) ≤0.001. D, desiccated; R, rehydrated after the desiccated treatment. (b) Venn diagram illustrating the number of shared and unique genes showing differential transcript abundance across different treatments. Ctrl, control. (c) Gene Ontology (GO) enrichment analysis of genes showing differential RNA abundance in D vs Ctrl or R vs D. The top 5 most enriched GO terms are displayed. (d) Number of *S. ruralis* and *S. caninervis* homologous gene pairs with differential RNA abundance between D vs Ctrl or R vs D in both species. (e) Upset plot showing homologous genes with differential RNA abundance in D vs Ctrl or R vs D comparisons of the two species. Colored bars and dots denote similar comparisons/conditions between the two species (e.g. teal reflects transcripts that were more abundant in R vs D). transcripts that respond similarly between the two species under the same conditions. (f) Heatmap of Transcripts Per Million (TPMs) of homologous gene pairs under D vs Ctrl or R vs D conditions.

acid, *light stimuli* (two terms), and *protein autophosphorylation* (Fig. 3c). This reduction in photosynthesis-related transcripts during dehydration could reflect a conservation of energy and resources, a common strategy in plants under stress (Liu *et al.*, 2018). GO-enriched terms of genes showing decreased transcript abundance in rehydration were *response to ABA*, *plant-type secondary cell wall biogenesis*, *cellular response to freezing*, *response to wounding*, and *defense response to bacterium*, in which a majority of genes were associated with stress response (Fig. 3c). The presence of stress-related responses during rehydration likely indicates a dissipation of stress response-related transcripts that were elevated in rehydration.

The relatively recent divergence time and likely shared evolutionary origins of DT between *S. ruralis* and *S. caninervis* offers a unique opportunity to search for conserved, but also recent molecular adaptations to desiccation and rehydration that may coincide with different habitat preferences. To this end, publicly available RNA-sequencing data from *S. caninervis*, gathered from plants exposed to a similar set of desiccation and rehydration conditions (Silva *et al.*, 2021), were mapped to the *S. caninervis* genome. Following similar QC filtering as performed in *S. ruralis*, changes in RNA abundance were compared for syntelogs (i.e. both collinear and sequence conserved) between these two species. Syntelogs were identified for 13 658 of the 21 169 gene models in the *S. ruralis* nuclear genome, representing 64.5, and 82.5%, of annotated genes in *S. ruralis* and *S. caninervis*, respectively. There were 818 and 516 syntelogs showing differential RNA abundance in D vs Ctrl and R vs D of both *S. ruralis* and *S. caninervis*, respectively (Fig. 3d). In the D vs Ctrl comparison, a majority of syntelogs had similar patterns, with RNA abundance either increasing or decreasing similarly in both species (629/818; Fig. 3e). A closer examination of the genes falling within inversion events (Fig. 2a) revealed a number of heat shock and stress-responsive transcription factors with different responses to desiccation tolerance between *S. ruralis* and *S. caninervis* (Table S6). In the R vs D comparison, a majority of syntelogs (401/516) showed different RNA abundance patterns in the two species. Interestingly, 54 syntelogs exhibited increased RNA abundance in D compared with Ctrl in both species, but *S. caninervis* showed decreased RNA abundance and *S. ruralis* displayed increased RNA abundance in R vs D comparison (Figs 3e,f). The similarity between the two mosses in their response to desiccation is likely a reflection of the shared cellular requirements needed to protect cells from the rigors of severe water loss and survive drying (Oliver *et al.*, 2020). The differences between the two mosses in their response to rehydration likely reflects differences in the level of damage incurred upon drying by the two species (Oliver *et al.*, 1993) and perhaps is a measure of their overall DT.

Comparative genomic and transcriptomic analysis of the LEA and ELIP protein families

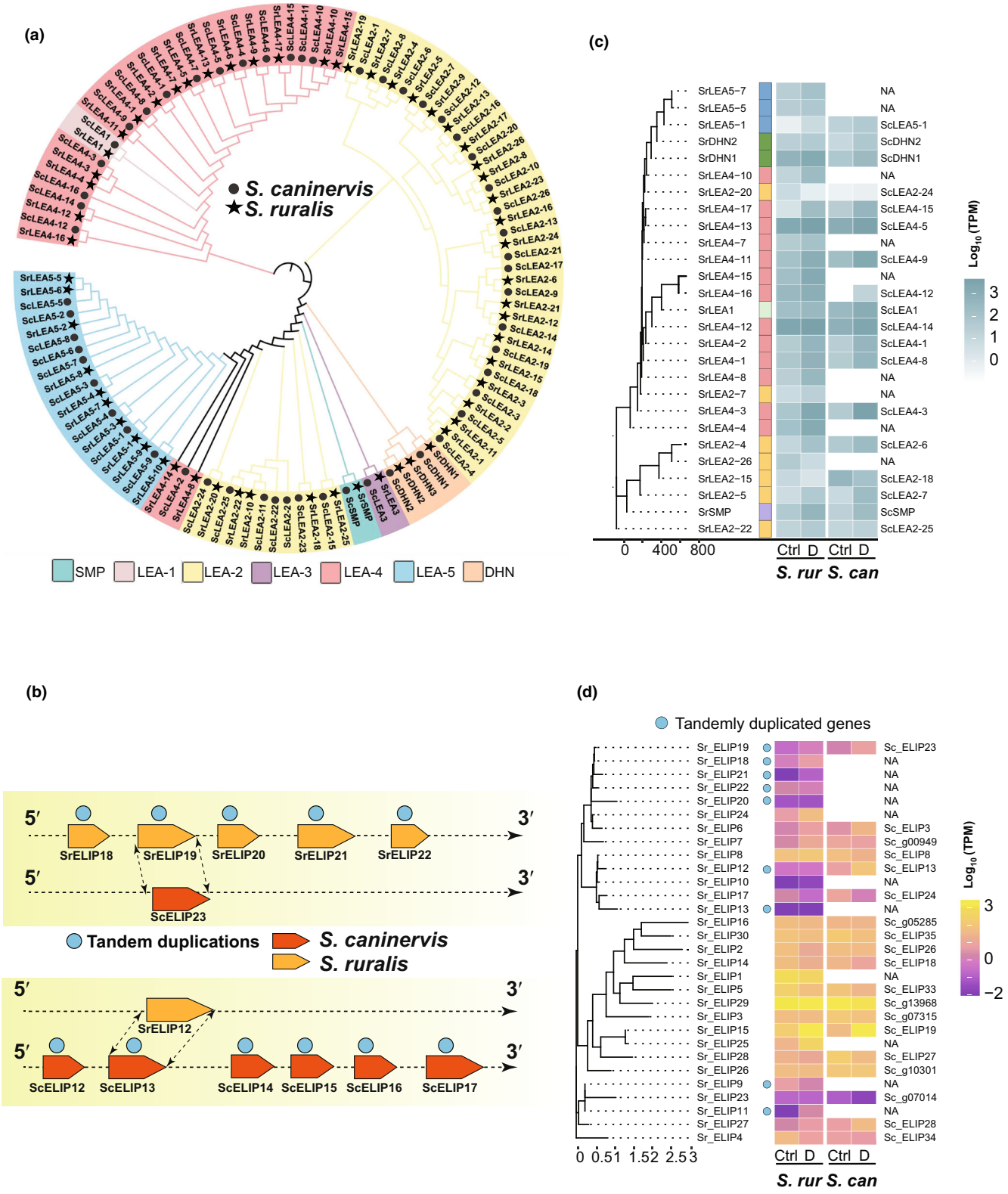
Certain gene families, such as late embryogenesis abundant (LEA) and early light-inducible proteins (ELIPs), have been associated with DT in a number of species (Tolleter *et al.*, 2007; VanBuren *et al.*, 2019). Late embryogenesis abundants, first identified during the late stages of seed development, are known for their protective functions in various cellular components under dehydration stress (Tolleter *et al.*, 2007; Hundertmark & Hincha, 2008). Similarly, ELIPs, a part of the chlorophyll *a*/*b*-binding (CAB) superfamily, are implicated in protecting photosynthetic apparatus from light-induced stress (Montané *et al.*, 1998; VanBuren *et al.*, 2019). Given their critical roles in stress tolerance, LEA and ELIP genes represent ideal candidates for understanding the molecular basis of DT in *S. ruralis*.

To better explore the evolution of these two important gene families in *S. ruralis*, candidate LEA and ELIP genes were first identified by protein sequence similarity using HMMER (Johnson *et al.*, 2010). A total of 59 LEA genes were identified in *S. ruralis*, clustering into seven of the eight characterized LEA subgroups based on sequence and structural motifs (Fig. S4; Singh & Graether, 2020). No LEA-6 was identified in *S. ruralis*, nor was one identified in *S. caninervis* (Fig. 4a; Silva *et al.*, 2021). A multiple sequence alignment of all *S. ruralis* and *S. caninervis* LEA protein sequences was used as input for inferring evolutionary relationships in this large gene family. The *S. ruralis* LEA-2 subgroup was the largest (25 genes), followed by LEA-4 ($n = 17$) and LEA-5 ($n = 10$; Fig. 4a). Multiple instances of species-specific gene duplication and then loss were evident in these larger subgroups, as inferred by an extant duplicate in one species adjacent to a node uniting a LEA from both species (asterisks, Fig. 4a). The ELIP gene family in *S. ruralis* is similarly large, comprising 30 members divided into four classes (I–IV; Fig. S5). As with the LEAs, both *S. ruralis* and *S. caninervis* have experienced lineage-specific duplication and loss events in the ELIP gene family. In addition, both species have experienced lineage-specific expansions that appear to have occurred through local tandem duplication events (Fig. 4b). For instance, nine of 30 *S. ruralis* ELIP genes are the product of tandem duplications, highlighting the volatility in these desiccation-associated gene families.

Transcriptomic analyses revealed how duplication (and loss) within these two gene families have impacted gene expression. Of the 59 identified *S. ruralis* LEA genes, transcripts of 27 of them were differentially abundant in desiccated conditions (Fig. 4c; Table S7). Of these 27, nine did not have an *S. caninervis* ortholog. Of the remainder, 12 had differentially abundant transcripts for an ortholog. A similar pattern was observed

Fig. 4 Comparison of late embryogenesis abundant (LEA) and early light-induced protein (ELIP) genes in *Syntrichia ruralis* and *Syntrichia caninervis*. (a) Phylogenetic relationship between LEA gene families in *S. ruralis* and *S. caninervis*. A maximum likelihood tree was generated based on a protein multiple sequence alignment. (b) The graph illustrated the occurrence of species-specific tandem duplication events. (c) Heatmap of the differentially abundant *S. ruralis* LEA genes and their *S. caninervis* orthologs, where present, under control (Ctrl) and desiccated (D) conditions. NA implies that a syntenic ortholog was not found in the *S. caninervis* genome. Normalized (transcript per million, TPM) values are shown. (d) Heatmap of the differentially abundant *S. ruralis* ELIP genes and their *S. caninervis* orthologs, where present, under control and slow dry conditions. Some homologous sequences were identified at syntenic locations in *S. caninervis* but were not predicted to be ELIPs based on absence of protein functional domains.

for the *S. ruralis* ELIP gene family, where 11/30 ELIP transcripts were differentially abundant under desiccation conditions (six up, five down; Fig. 4d; Table S8). Of these, three had *S. caninervis* orthologs whose transcripts were differentially abundant with similar changes in abundance in the two species. In sum, these two gene families are evolutionarily quite dynamic and look to be heavily (and differently) utilized during desiccation in these two species.



Putative transcription factors regulating the desiccation response in *S. ruralis*

To predict putative transcription factors in *S. ruralis*, the amino acid sequences of all nuclear-encoded genes were scanned using PlantTFDB (Tian *et al.*, 2020), resulting in a total of 636 TFs associated with 53 families (Table S9). These numbers are slightly higher than those previously identified in *S. caninervis* (542 in 50 TF families) and substantially lower than in *P. patens* (1136 in 53 families; Table S9; Silva *et al.*, 2021). When comparing TF family composition between *S. ruralis* and *S. caninervis*, the B3 and C2H2 families were expanded (although not significantly) in *S. ruralis*, whereas there were no families that had experienced a significant loss.

These *S. ruralis* TFs were then assessed for a possible role in the response to desiccation by examining their change in abundance between D and Ctrl conditions. A total of 126 (out of 636) TFs showed differential RNA abundance, with RNA levels increasing for 83 and decreasing for 53 (Fig. 5a). Ninety-four *S. ruralis* TFs (out of 126) have syntelogs in *S. caninervis*, of which 40 also showed differential transcript abundance in D conditions (Fig. 5a; Table S10). Interestingly, 32 differentially abundant *S. ruralis* TFs had no identifiable syntelog in *S. caninervis*, suggesting they may impart a species-specific response to desiccation (Fig. 5a).

To investigate the regulation of downstream dehydration-responsive genes by these 126 TFs, the 2 kb upstream region (promoter region) of the 3045 genes found to be differentially abundant between Ctrl and D conditions was extracted. Putative transcription factor binding sites were identified in these regions using Analysis of Motif Enrichment AME (AME; McLeay & Bailey, 2010) with TF-DNA binding data from the *A. thaliana* DAP motif database (O'Malley *et al.*, 2016). From the 3045 input genes, 121 overrepresented (enriched; *P*-value < 1e−60) motifs, corresponding to 120 *S. ruralis* TFs were selected for further analysis (Table S11). As expected based on the water-limited conditions, the enriched motifs were predicted to predominantly be associated with the APETALA2/ETHYLENE RESPONSE FACTOR (AP2/ERF) and AP2/ETHYLENE RESPONSE FACTOR BINDING PROTEIN (AP2/ERFBP) families of TFs, which are well-known regulators of plant abiotic stress responses. Of these 120 TFs, 14 corresponded to those that changed abundance in desiccated conditions (Fig. 5a,b), suggesting they may both be responding to and regulating the desiccation response in *S. ruralis*.

To assess their potential regulatory role, pairwise abundance correlation, Spearman's rank correlation coefficient (*r*), was calculated between these 14 *S. ruralis* TFs and their putative target genes across all six experimental conditions ($n = 17$ data points) and compared with a background set of genes. Seven of the 14 TFs had a median correlation above |0.5|, with seven of these TFs significantly more correlated with their target genes than background (Figs 5c, S6). Six of these *S. ruralis* TFs are homologous (based on amino acid sequence and domain similarities) to *A. thaliana* TFs that have reported functions in drought tolerance in vegetative tissues (ABA-Responsive Element Binding factor 2 [AREB2]/Sr_g16458), cell division during stress (GT2/Sr_g03037; HSFB4/Sr_g07011; MYBH/Sr_g09762; ERF6/Sr_g20827), and control gene expression during late embryogenesis–in particular the LEA gene family (ABA Insensitive 3 [ABI3]/Sr_g11739). Given their annotated functions in *A. thaliana*, SrABI3 and SrAREB2 are more likely to regulate ABA-responsive genes in *S. ruralis*. Several dehydration-responsive genes, including 5 LEAs, are among the 480 differentially abundant genes that contain ABI3 TF binding motifs and are correlated with SrABI3 abundance (Fig. 5d; Table S12). Of the 346 desiccated-responsive genes with an AREB2 TF binding site and expression correlation > |0.5|, 29 are annotated as ABA-responsive genes based on sequence or domain similarity (Table S12).

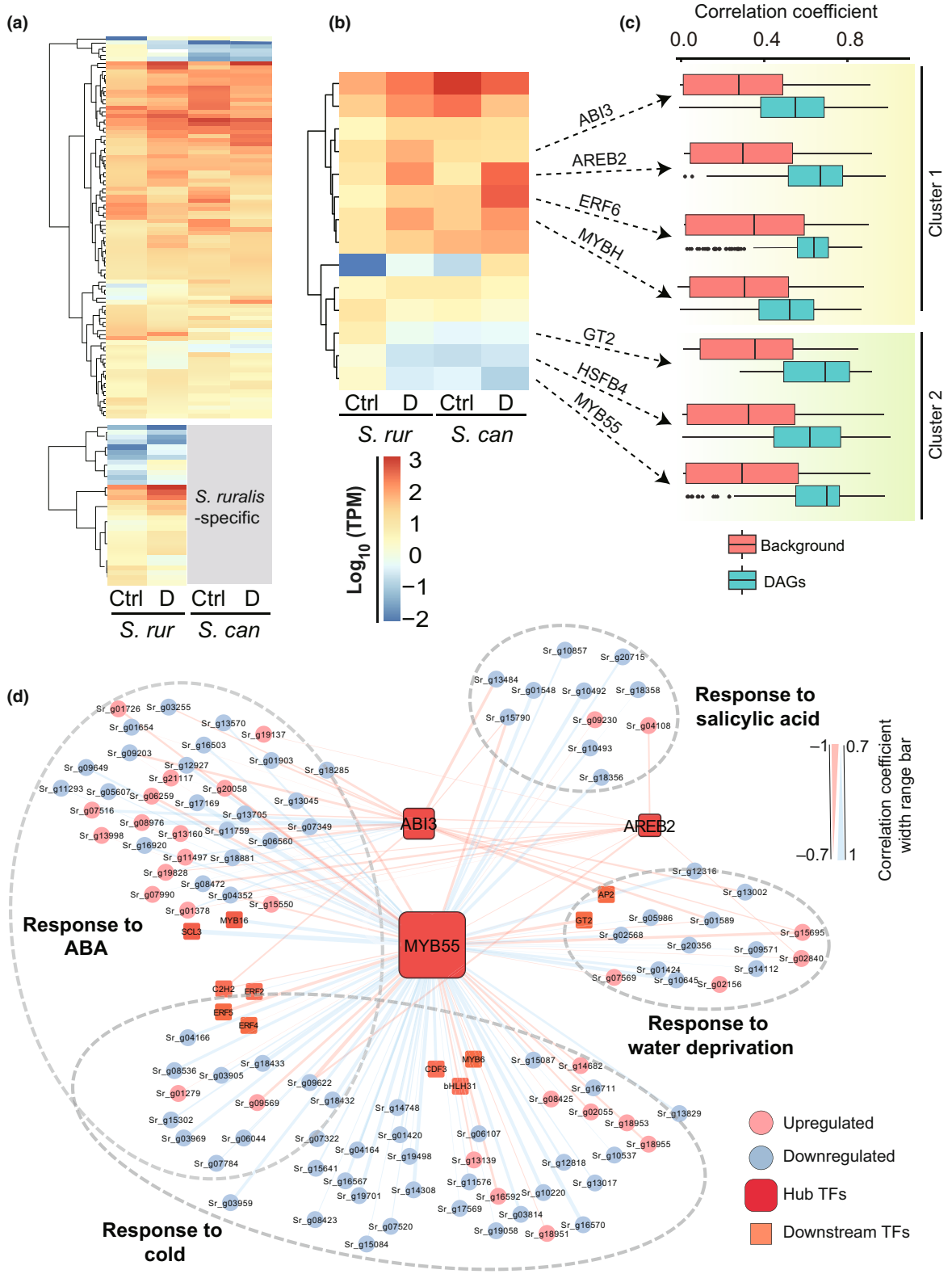
An uncharacterized regulator of dehydration in land plants

The remaining TF (Sr_g19809) is an R2R3 MYB transcription factor, a large family of regulators (both activators and repressors) well known for their diverse roles in development, metabolism, and stress responses in plants. In *S. ruralis*, *Sr_g19809* was downregulated during desiccation, and, due to correlation and TF binding site enrichment, appears to regulate SrGT2, SrHSFB4, and SrAREB2. Indeed, of the seven hub TFs, *Sr_g19809* appears to have the most interactions, with most of these being anticorrelated (Fig. 5d), suggesting it is a transcriptional repressor. In addition, *Sr_g19809* is also anticorrelated with the *S. ruralis* homolog of *A. thaliana* ABI3/FUS3 (AT3G24650 and AT3G26790, respectively), which are also known to play a role in DT during embryogenesis (Fig. S7; González-Morales *et al.*, 2016).

Given the repeated incorporation of specific components of the DT pathway across land plants (e.g. ABI3, LEAs, and ELIPs; Marks *et al.*, 2021), we investigated whether there were any MYB TFs similar to *Sr_g19809* in *A. thaliana*. Using publicly available transcriptomic data, including samples associated with seed development, ABA application, and drought treatment (Klepikova *et al.*, 2016; Zhao *et al.*, 2018), we searched for *A. thaliana* MYB

Fig. 5 Putative transcription factor (TF) regulators in response to dehydration in *Syntrichia ruralis*. (a) The expression heatmap of putative *S. ruralis* TFs and syntenic pairs in *Syntrichia caninervis* in desiccated and control conditions. (b) The expression heatmap of 14 hub TFs and *S. caninervis* syntenic pairs. (c) Expression correlation between differentially abundant TFs and putative target genes (correlation > 0.5) and randomly selected genes (background). Boxplots depict the distribution of expression correlation. Horizontal lines within boxplots indicate median correlation values while whiskers extend to 1.5 times the interquartile range. All seven comparisons are significant based on a Student's *t*-test (*P*-value < 0.05). (d) Regulatory network of hub TFs and putative downstream targets. The network was constructed by integrating differential gene abundance data with transcription factor binding motif enrichment analysis to identify potential regulatory relationships. Nodes represent TFs (squares) and target genes (circles), while edges indicate predicted interactions based on correlation analysis. The thickness of edges reflects the strength of correlation (|0.7–1|), with blue lines representing a positive correlation and pink lines a negative.

transcription factors anticorrelated in abundance with either AtFUS3 or AtABI3. This search uncovered a previously uncharacterized MYB TF, AtMYB55 (AT4G01680). *AtMYB55* has a similar R2R3 MYB binding domain as *Sr_g19809* (hereafter referred to as *SrMYB55*), is expressed during embryogenesis and germination, and shows strong negative correlation with *FUS3* under drought and ABA treatment (Fig S8A,B) but shows a strong negative correlation with *ABI3* during seed development (Fig. S8C). An investigation of *A. thaliana* Cistrome data (O'Malley *et al.*, 2016) identified a MYB55 binding site upstream of both ABI3 and FUS3,



suggesting AtMYB55 may be a direct negative regulator of both of these genes (Fig. S8D).

AtMYB55 is part of a clade of four similarly structured R2R3 MYB TFs in *A. thaliana* (Stracke *et al.*, 2001). To determine the specificity of the AtMYB55-FUS3/ABI3 regulatory interaction, we also examined the closest expressed homolog of *AtMYB55*, *AtMYB50* (AT1G57560), for its transcript abundance profile and potential ability to interact with ABI3 or FUS3. In contrast to *AtMYB55*, *AtMYB50* was not correlated with *ABI3* and *FUS3* (Fig. S8A–C), nor is a MYB50 binding site seen in the promoter elements of these two genes. In addition, *AtMYB50* is induced by the biotic-stress associated phytohormones JA and SA (Katiyar *et al.*, 2012) and thus likely works to regulate a distinct set of genes from *AtMYB55*.

Given that *SrMYB55* regulates a number of ABA-associated genes and *AtMYB55* is expressed during germination (Fig. S8D), we hypothesized that *Atmyb55* mutants might display a sensitivity to ABA at this developmental stage. We identified a T-DNA insertion line within the third exon of *AtMYB55* (Fig. 6a). Both WT and *Atmyb55* showed 100% germination after 24 h on plates without ABA (Fig. 6b). On the plates with 0.1, 0.3, and 0.5 μM ABA, WT seeds displayed delayed germination, but germinated at levels close to control plates after 72 h. By contrast, the *Atmyb55* background showed significantly delayed germination, and at the maximum ABA concentration (0.5 μM), never reached control levels of germination (Fig. 6c). To test whether this sensitivity was restricted to germination, WT and *Atmyb55* seedlings were germinated on 0 μM ABA plates and then after 1 wk (corresponding to roughly four true leaves for both backgrounds) transferred to plates containing varying concentrations (0.3–2.0 μM) of ABA. No variation in number or density of lateral roots, nor in length of primary root, was observed between WT and *Atmyb55* seedlings, suggesting the ABA sensitivity in this background is restricted to germination under normal growth conditions (Fig. S9). In sum, *MYB55* appears to be negatively regulating *ABI3/FUS3* in both *A. thaliana* and *S. ruralis*.

MYB55 transcript abundance appears to be positively correlated with water content in both species (Figs 5b, S8A–C), such that under normal water conditions, *MYB55* abundance is high and is reduced as water availability diminishes. Interestingly, this decrease in MYB55 abundance occurs in drying seeds as well as in vegetative tissues experiencing drought stress (Fig. S8A–C). A compelling model arises in which *MYB55* is tightly linked to perceived water stress in an ABA-dependent manner. As water stress increases, *MYB55* levels are decreased, allowing for increased levels of either *ABI3* or *FUS3*, depending on the developmental stage. While it remains to be seen if MYB55 directly binds the combined ABI3-FUS3 homolog in *S. ruralis*, the observed regulatory interactions in both species are suggestive of a conserved mechanism between *A. thaliana* and *S. ruralis*, and maybe across land plants in general. The similar regulatory interactions observed in *A. thaliana* and *S. ruralis*, therefore, might represent a fundamental survival strategy, rooted in early photosynthetic eukaryotes, and maintained throughout plant evolution.

Conclusions

In this study, we sequenced the genome of a desiccation-tolerant moss, *S. ruralis*, and then took a comparative transcriptomic approach to identify its regulatory response to water deprivation. Compared with *S. caninervis*, *S. ruralis* can inhabit a wide range of habitats and moisture gradients, and thus, we anticipated identifying both shared, and unique, molecular responses to desiccation. The transcriptomic analyses revealed that the major differences in the response to desiccation between the closely related *Syntrichias* occurred in the response to rehydration, a major aspect to overall DT and a major stress. We conclude that these differences likely indicate the different levels of DT exhibited by the highly tolerant *S. caninervis* and the more variable tolerance capabilities of *S. ruralis* (as described by Oliver *et al.*, 1993) as well as the overall plasticity of this trait in mosses. Unexpectedly, through our transcriptomic analyses, we identified a previously unknown regulatory component in the desiccation response pathway in *S. ruralis*: a MYB transcription factor that was anticorrelated with deeply conserved ABA-associated drought response and seed germination factors in angiosperms (e.g. ABI3 and FUS3). This TF appears itself to be functionally conserved

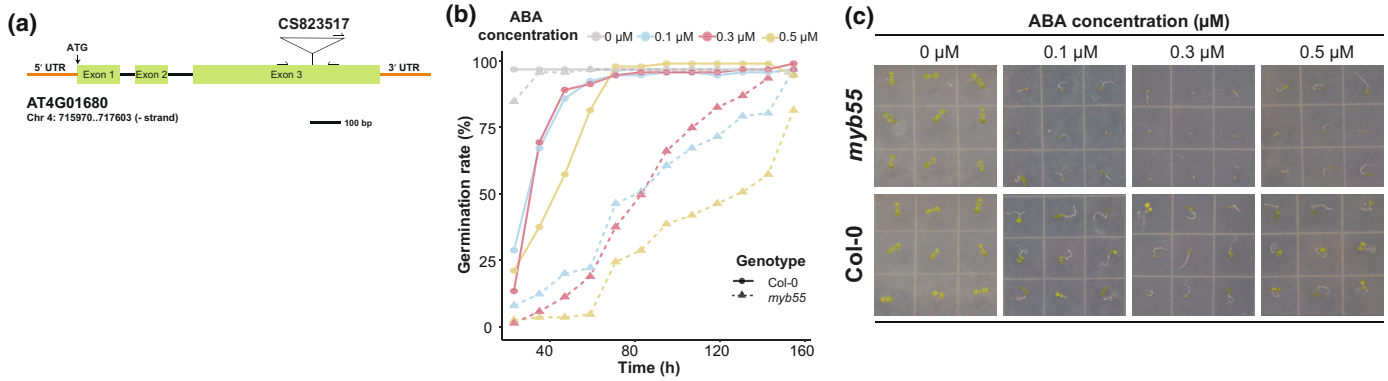


Fig. 6 *Atmyb55*, homolog of *Syntrichia ruralis* MYB55, is sensitive to abscisic acid (ABA) in *Arabidopsis thaliana*. (a) Schematic diagram of the T-DNA insertion of the *MYB55* gene. (b) Seed germination rate of *myb55* and Col-0 lines in response to different concentrations of ABA. Numbers of germinated seedlings were recorded from 0 to 156 h (h) after stratification on ½ MS agar plates containing 0, 0.1, 0.3, or 0.5 μM ABA. (c) Phenotypes of seedlings grown for 7 d after stratification on agar plates containing 0, 0.1, 0.3, or 0.5 μM ABA.

across land plants. In sum, we have developed a comparative genomic resource for studying desiccation tolerance and used this resource to uncover novel insights into desiccation tolerance across the land plant lineage.

Acknowledgements

We would like to acknowledge support from NSF Dimensions of Biodiversity Program grants 1638972 (to MJO) and 1638956 (to BDM), NSF IOS PGRP 2023310 (to ADLN) and NSF IOS PGRP 2102120 (to ADLN). JTBE was funded by the NSF CAREER Award 2144011. We would like to thank the Boyce Thompson Institute PGRP REU program for hosting AKJ (REU support for AKJ from NSF no. 2023310). We would like to thank Dr Jingjing Zhai (Cornell University) for bioinformatic analysis support. The authors would like to thank other members of the Nelson laboratory for their insightful feedback during the preparation of this manuscript. Authors would like to thank members of the Oliver lab, Kate Guill, and James Elde, for their technical expertise in maintaining moss cultures and the isolation of high molecular weight genomic DNA.

Competing interests

None declared.

Author contributions

XZ, JTBE, BDM, ADLN and MJO developed the project. ATS isolated nucleic acids for sequencing, assembled and annotated the genome. XZ, JTBE and LY performed bioinformatic analyses. XZ and AKJ performed the *Arabidopsis thaliana* analyses. XZ, JTBE, ADLN and MJO wrote the manuscript. XZ and JTBE contributed equally to this work.

ORCID

Jenna T. B. Ekwealor https://orcid.org/0000-0001-9014-8786
Andrea K. Jones https://orcid.org/0009-0007-1786-2626
Brent D. Mishler https://orcid.org/0000-0001-5727-4916
Andrew D. L. Nelson https://orcid.org/0000-0001-9896-1739
Melvin J. Oliver https://orcid.org/0000-0002-5991-5788
Anderson T. Silva https://orcid.org/0009-0008-2403-0438
Li'ang Yu https://orcid.org/0000-0002-9556-011X
Xiaodan Zhang https://orcid.org/0000-0001-8192-0666

Data availability

Syntrichia ruralis isolates are available by request from ADLN. The *S. ruralis* genome assembly and annotation are publicly available at CoGe (genomevolution.org, Genome ID 66749). DNA and RNA-sequencing data are available at NCBI (Bioproject PRJNA1042657). The raw DNA sequencing reads are available at the NCBI sequence read archive under Bioproject PRJNA1076122. Image files are stored at doi: 10.6084/m9.figshare.24581178.v1.

References

Alexa A, Rahnenführer J. 2009. Gene set enrichment analysis with topGO. *Bioconductor* **27**: 1–26.

Alpert P, Oliver MJ. 2002. Drying without dying. In: Black M, Pritchard HW, eds. *Desiccation and survival in plants: drying without dying*. Wallingford, UK: CABI Publishing, 3–43.

Belnap J, Lange OL. 2003. Structure and functioning of biological soil crusts: a synthesis. In: Belnap J, Lange OL, eds. *Biological soil crusts: structure, function, and management. Ecological studies, vol. 150*. Berlin, Germany: Springer-Verlag, 471–479.

Bewley JD. 1979. Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology* **30**: 195–238.

Bowles AMC, Bechtold U, Paps J. 2020. The origin of land plants is rooted in two bursts of genomic novelty. *Current Biology* **30**: 530–536.

Brinda JC, Jáuregui-Lazo JA, Oliver MJ, Mishler BD. 2021. Notes on the genus *Syntrichia* with a revised infrageneric classification and the recognition of a new genus *Syntrichiadelphus* (Bryophyta, Pottiaceae). *Phytologia* **103**: 90–103.

Carey SB, Jenkins J, Lovell JT, Maumus F, Sreedasyam A, Payton AC, Shu S, Tiley GP, Fernandez-Pozo N, Healey A *et al.* 2021. Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Science Advances* **7**: eabh2488.

Coe KK, Greenwood JL, Slate ML, Clark TA, Brinda JC, Fisher KM, Mishler BD, Bowker MA, Oliver MJ, Ebrahimi S *et al.* 2021. Strategies of desiccation tolerance vary across life phases in the moss *Syntrichia caninervis*. *American Journal of Botany* **108**: 249–262.

Coelho SM, Gueno J, Lipinska AP, Cock JM, Umen JG. 2018. UV chromosomes and haploid sexual systems. *Trends in Plant Science* **23**: 794–807.

Conway JR, Lex A, Gehlenborg N. 2017. upsetr: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**: 2938–2940.

Cove DJ, Perroud PF, Charron AJ, McDaniel SF, Khandelwal A, Quatrano RS. 2009. The moss *Physcomitrella patens*: a novel model system for plant development and genomic studies. *Cold Spring Harbor Protocols* **1**: pdb-emo115.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic acid: emergence of a core signaling network. *Annual Review of Plant Biology* **61**: 651–679.

Donoghue PJ, Harrison CJ, Jill C, Paps J, Schneider H. 2021. The evolutionary emergence of land plants. *Current Biology* **31**: R1281–R1298.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Ekwealor JT, Payton AC, Paasch AE, Fisher KM, McDaniel SF. 2017. Multiple factors influence population sex ratios in the Mojave Desert moss *Syntrichia caninervis*. *American Journal of Botany* **104**: 733–742.

Fajkus P, Kilar A, Nelson ADL, Holá M, Peška V, Goffová I, Fojtová M, Zachová D, Fulnečková J, Fajkus J. 2021. Evolution of plant telomerase RNAs: farther to the past, deeper to the roots. *Nucleic Acids Research* **49**: 7680–7694.

Fajkus P, Peška V, Závodník M, Fojtová M, Fulnečková J, Dobias Š, Kilar A, Dvořáčková M, Zachová D, Nečasová I *et al.* 2019. Telomerase RNAs in land plants. *Nucleic Acids Research* **47**: 9842–9856.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J *et al.* 2014. Pfam: the protein families database. *Nucleic Acids Research* **42**: D222–D230.

Frank W, Ratnadewi D, Reski R. 2005. *Physcomitrella patens* is highly tolerant against drought, salt and osmotic stress. *Planta* **220**: 384–394.

Gao CH, Yu G, Cai P. 2021. ggVennDiagram: an intuitive, easy-to-use, and highly customizable R package to generate Venn diagram. *Frontiers in Genetics* **12**: 706907.

González-Morales SI, Chávez-Montes RA, Hayano-Kanashiro C, Alejo-Jacuinde G, Rico-Cambron TY, de Folter S, Herrera-Estrella L. 2016. Regulatory network analysis reveals novel regulators of seed desiccation tolerance in

Arabidopsis thaliana. *Proceedings of the National Academy of Sciences, USA* **113**: E5232–E5241.

Greenwood JL, Stark LR. 2014. The rate of drying determines the extent of desiccation tolerance in *Physcomitrella patens*. *Functional Plant Biology* **41**: 460–467.

Haug-Baltzell A, Stephens SA, Davey S, Scheidegger CE, Lyons E. 2017. SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics* **33**: 2197–2198.

Hoagland DR, Arnon DI. 1938. Growing plants without soil by the water-culture method. *Circular California Agricultural Experimental Station* **367**: 1–24.

Hundertmark M, Hincha DK. 2008. LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* **9**: 1–22.

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L *et al.* 2009. InterPro: the integrative protein signature database. *Nucleic Acids Research* **37**: D211–D215.

Jauregui-Lazo J, Brinda JC, GoFlag Consortium, Mishler BD. 2023. The phylogeny of *Syntrichia*: an ecologically diverse clade of mosses with an origin in South America. *American Journal of Botany* **110**: e16103.

Johnson LS, Eddy SR, Portugaly E. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**: 431.

Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. 2012. Genome-wide classification and expression analysis of *MYB* transcription factor families in rice and *Arabidopsis*. *BMC Genomics* **13**: 544.

Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Pepin AA. 2016. A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *The Plant Journal* **88**: 1058–1070.

Kobayashi Y, Ando H, Hanaoka M, Tanaka K. 2016. Abscisic acid participates in the control of cell cycle initiation through heme homeostasis in the unicellular red alga *Cyanidioschyzon merolae*. *Plant and Cell Physiology* **57**: 953–960.

Kolde R, Kolde MR. 2015. Package "pheatmap". *R Package* **1**: 790.

Koster KL, Balsamo RA, Espinoza C, Oliver MJ. 2010. Desiccation sensitivity and tolerance in the moss *Physcomitrella patens*: assessing limits and damage. *Plant Growth Regulation* **62**: 293–302.

Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, Piednoel P, Gundlach H, Van Bel M, Meyberg R *et al.* 2018. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal* **93**: 515–533.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v.5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49**: 293–296.

Li X, Yang R, Liang Y, Gao B, Li S, Bai W, Oliver MJ, Zhang D. 2023. The ScAPD1-like gene from the desert moss *Syntrichia caninervis* enhances resistance to *Verticillium dahliae* via phenylpropanoid gene regulation. *The Plant Journal* **113**: 75–91.

Liao Y, Smyth GK, Shi W. 2014. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.* 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289–293.

Liu J, Moyankova D, Lin CT, Mladenov P, Sun RZ, Djilianov D, Deng X. 2018. Transcriptome reprogramming during severe dehydration contributes to physiological and metabolic changes in the resurrection plant *Haberlea rhodopensis*. *BMC Plant Biology* **18**: 351.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Lyons E, Freeling M. 2008. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal* **53**: 661–673.

Marks RA, Farrant JM, Nicholas McLetchie D, VanBuren R. 2021. Unexplored dimensions of variability in vegetative desiccation tolerance. *American Journal of Botany* **108**: 346–358.

McLeay RC, Bailey TL. 2010. Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* **11**: 165.

Mishler BD. 2007. Syntrichia. In: Flora of North America Editorial Committee, ed. *Flora of North America North of Mexico, vol. 27*. New York, NY, USA: Oxford University Press, 618–627.

Mishler BD, Oliver MJ. 2009. Putting *Physcomitrella* on the tree of life: the evolution and ecology of mosses. In: Knight C, Perroud PF, Cove D, eds. *The moss* Physcomitrella patens. *Annual plant reviews, vol. 36*. Wiley-Blackwell: Oxford, UK, 1–5.

Montané MH, Tardy F, Kloppstech K, Havaux M. 1998. Differential control of xanthophylls and light-induced stress proteins, as opposed to light-harvesting chlorophyll *a*/*b* proteins, during photosynthetic acclimation of barley leaves to light irradiance. *Plant Physiology* **118**: 227–235.

Naramoto S, Hata Y, Fujita T, Kyozuka J. 2022. The bryophytes *Physcomitrium patens* and *Marchantia polymorpha* as model systems for studying evolutionary cell and developmental biology in plants. *Plant Cell* **34**: 228–246.

Nelson DC, Flematti GR, Riseborough JA, Ghisalberti EL, Dixon KW, Smith SM. 2010. Karrikins enhance light responses during germination and seedling development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **107**: 7095–7100.

Oldenhof H, Wolkers WF, Bowman JL, Tablin F, Crowe JH. 2006. Freezing and desiccation tolerance in the moss *Physcomitrella patens*: an *in situ* fourier transform infrared spectroscopic study. *Biochimica et Biophysica Acta* **1760**: 1226–1234.

Oliver MJ. 1991. Influence of protoplasmic water loss on the control of protein synthesis in the desiccation-tolerant moss *Tortula ruralis*: ramifications for a repair-based mechanism of desiccation-tolerance. *Plant Physiology* **97**: 1501–1511.

Oliver MJ. 2009. Biochemical and molecular mechanisms of desiccation tolerance in bryophytes. In: Shaw J, Goffinet B, eds. *Bryophyte biology, 2nd edn.* Cambridge, UK: Cambridge University Press, 269–298.

Oliver MJ, Farrant JM, Hilhorst HWM, Mundree S, Williams B, Bewley JD. 2020. Desiccation tolerance: avoiding cellular damage during drying and rehydration. *Annual Review of Plant Biology* **71**: 435–460.

Oliver MJ, Mishler BD, Quisenberry JE. 1993. Comparative measures of desiccation-tolerance in the *Tortula ruralis* complex. I. Variation in damage control and repair. *American Journal of Botany* **80**: 127–136.

Oliver MJ, Velten J, Mishler BD. 2005. Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats? *Integrative and Comparative Biology* **45**: 788–799.

O'Malley RC, Huang SS, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. 2016. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**: 1280–1292.

Peri S, Roberts S, Kreko IR, McHan LB, Naron A, Ram A, Murphy RL, Lyons E, Gregory BD, Devisetty UK *et al.* 2020. Read mapping and transcript assembly: a scalable and high-throughput workflow for the processing and analysis of ribonucleic acid sequencing data. *Frontiers in Genetics* **10**: 1361.

Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* **11**: 1650–1667.

Phillips JR, Oliver MJ, Bartels D. 2002. Molecular genetics of desiccation and tolerant systems. In: Black M, Pritchard HW, eds. *Desiccation and plant survival*. New York, NY, USA: Oxford University Press, 319–342.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. *Nucleic Acids Research* **46**: W200–W204.

Proctor MC, Oliver MJ, Wood AJ, Alpert P, Stark LR, Cleavitt NL, Mishler BD. 2007. Desiccation-tolerance in bryophytes: a review. *The Bryologist* **110**: 595–621.

Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW *et al.* 2016. Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. *Genome Research* **26**: 342–350.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

R Development Core Team. 2021. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rensing SA, Goffinet B, Meyberg R, Wu SZ, Bezanilla M. 2020. The moss *Physcomitrium* (*Physcomitrella*) *patens*: a model organism for non-seed plants. *Plant Cell* **32**: 1361–1376.

Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T. 2012. A travel guide to Cytoscape plugins. *Nature Methods* **9**: 1069–1076.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.

Schonbeck MW, Bewley JD. 1981. Responses of the moss *Tortula ruralis* to desiccation treatments. II. Variations in desiccation tolerance. *Canadian Journal of Botany* **59**: 2707–2712.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J *et al.* 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**: 539.

Silva AT, Gao B, Fisher KM, Mishler BD, Ekwealor JTB, Stark LR, Li X, Zhang D, Bowker MA, Brinda JC *et al.* 2021. To dry perchance to live: insights from the genome of the desiccation-tolerant biocrust moss *Syntrichia caninervis*. *The Plant Journal* **105**: 1339–1356.

Singh KK, Graether SP. 2020. Conserved sequence motifs in the abiotic stress response protein late embryogenesis abundant 3. *PLoS ONE* **15**: e0237177.

Singh S, Davies KM, Chagné D, Bowman JL. 2023. The fate of sex chromosomes during the evolution of monoicy from dioicy in liverworts. *Current Biology* **33**: 3597–3609.

Stark LR. 2017. Ecology of desiccation tolerance in bryophytes: a conceptual framework and methodology. *The Bryologist* **120**: 130–165.

Stark LR, McLetchie DN, Mishler BD. 2005a. Sex expression, plant size, and spatial segregation of the sexes across a stress gradient in the desert moss *Syntrichia caninervis*. *The Bryologist* **108**: 183–193.

Stark LR, Nichols L II, McLetchie DN, Bonine ML. 2005b. Do the sexes of the desert moss *Syntrichia caninervis* differ in desiccation tolerance? A leaf regeneration assay. *International Journal of Plant Sciences* **166**: 21–29.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4**: 447–456.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. Mega5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.

Tian F, Yang D, Meng Y, Jin J, Gao G. 2020. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Research* **48**: D1104–D1113.

Tolleter D, Jaquinod M, Mangavel C, Passirani C, Saulnier P, Manon S, Teyssier E, Payet N, Avelange-Macherel MH, Macherel D. 2007. Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**: 1580–1589.

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. 2010. Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. *Plant and Cell Physiology* **51**: 1821–1839.

VanBuren R, Pardo J, Wai CM, Evans S, Bartels D. 2019. Massive tandem proliferation of ELIPs supports convergent evolution of desiccation tolerance across land plants. *Plant Physiology* **179**: 1040–1049.

Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H *et al.* 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40**: e49.

Wood AJ, Oliver MJ. 1999. Translational control in plant stress: formation of messenger ribonucleoprotein complexes (mRNPs) in *Tortula ruralis* in response to desiccation. *The Plant Journal* **18**: 359–370.

Xiao L, Yobi A, Koster KL, He Y, Oliver MJ. 2018. Desiccation tolerance in *Physcomitrella patens*: rate of dehydration and the involvement of endogenous ABA. *Plant, Cell & Environment* **41**: 275–284.

Yadav S, Basu S, Srivastava A, Biswas S, Mondal R, Jha VK, Singh SK, Mishra Y. 2023. Bryophytes as modern model plants: an overview of their development, contributions, and future prospects. *Journal of Plant Growth Regulation* **42**: 6933–6950.

Zhang J, Zhang YM, Downing A, Wu N, Zhang BC. 2011. Photosynthetic and cytological recovery on remoistening *Syntrichia caninervis* Mitt., a desiccation-tolerant moss from Northwestern China. *Photosynthetica* **49**: 13–20.

Zhao X, Li J, Lian B, Gu H, Li Y, Qi Y. 2018. Global identification of *Arabidopsis* lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nature Communications* **9**: 5056.

Zhong B, Fong R, Collins LJ, McLenachan PA, Penny D. 2014. Two new fern chloroplasts and decelerated evolution linked to the long generation time in tree ferns. *Genome Biology and Evolution* **6**: 1166–1173.

Supporting Information

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

Fig. S1 Comparison between genome size and annotated gene content for published moss genomes.

Fig. S2 Large-scale genome conservation between *Syntrichia ruralis*, *Syntrichia caninervis*, and *Physcomitrium patens* generated by McScan.

Fig. S3 Quality control of *Syntrichia ruralis* samples.

Fig. S4 Consensus patterns of *Syntrichia ruralis* late embryogenesis abundant subgroups.

Fig. S5 Phylogenetic relationship between early light-inducible proteins gene families in *Syntrichia ruralis* and *Syntrichia caninervis*.

Fig. S6 Density plot of correlation between *Syntrichia ruralis* putative transcription factors (TFs) and downstream targets.

Fig. S7 Heatmap of transcript abundance values for *Syntrichia ruralis FUS3*, *ABI3*, and *MYB55*.

Fig. S8 Heat map of transcript abundance data for *Arabidopsis thaliana MYB55*, *MYB50*, *LEC1*, *ABI3*, and *FUS3* during drought and abscisic acid treatment, as well as during seed development and germination.

Fig. S9 Phototypes of *myb55* and Col-0 seedings that germinated on plates with 0 μM abscisic acid (ABA) and transplanted and grown for a week on agar plates containing 0.3, 0.5, 1, and 2 μM ABA.

Table S1 Chromosome size and number of genes per chromosome in the *Syntrichia ruralis* genome.

Table S2 GC content and number of genes per scaffolds.

Table S3 Taxa classification of genes in unplaced scaffolds.

Table S4 Homologous genes in the sex chromosomes of *Syntrichia ruralis* and *Syntrichia caninervis*.

Table S5 Summary of *Syntrichia ruralis* transcriptomic data.

Table S6 Chromosomal inversion events contain transcription factors with different responses to desiccation tolerance between *Syntrichia ruralis* and *Syntrichia caninervis*.

Table S7 Characteristics of genes encoding late embryogenesis abundant proteins in *Syntrichia ruralis* and *Syntrichia caninervis*.

Table S8 Characteristics of genes encoding early light-inducible proteins in *Syntrichia ruralis*.

Table S9 Comparison of transcription factor family genes of *Syntrichia ruralis*, *Syntrichia caninervis* and *Physcomitrium patens*.

Table S10 *Syntrichia ruralis* transcription factors, and their syntelogs in *Syntrichia caninervis*, with differential RNA abundance in desiccation (D) relative to control.

Table S11 Summary of motif enrichment analysis.

Table S12 Expression correlation of 14 *Syntrichia ruralis* transcription factors and their putative target genes.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

See also the Commentary on this article by Marks, **243**: 828–829.