The Plant Cell, Vol. 30: 2943–2958, December 2018, www.plantcell.org © 2018 ASPB.

LARGE-SCALE BIOLOGY ARTICLE

Desiccation Tolerance Evolved through Gene Duplication and Network Rewiring in *Lindernia*[OPEN]

Robert VanBuren,[a,b,1] Ching Man Wai,[a,b] Jeremy Pardo,[b,c] Valentino Giarola,[d] Stefano Ambrosini,[d] Xiaomin Song, and Dorothea Bartels[d,1]

[a] Department of Horticulture, Michigan State University, East Lansing, Michigan 48824
[b] Plant Resilience Institute, Michigan State University, East Lansing, Michigan 48824
[c] Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824
[d] IMBIO, University of Bonn, D-53115 Bonn, Germany

ORCID IDs: 0000-0003-2133-2760 (R.V.); 0000-0002-2913-5721 (C.M.W.); 0000-0003-3419-095X (J.P.); 0000-0001-6965-4799 (V.G.); 0000-0003-4079-9411 (S.A.); 0000-0003-3770-2404 (X.S.); 0000-0003-0505-3083 (D.B.)

Although several resurrection plant genomes have been sequenced, the lack of suitable dehydration-sensitive outgroups has limited genomic insights into the origin of desiccation tolerance. Here, we utilized a comparative system of closely related desiccation-tolerant (*Lindernia brevidens*) and -sensitive (*Lindernia subracemosa*) species to identify gene- and pathway-level changes associated with the evolution of desiccation tolerance. The two high-quality *Lindernia* genomes we assembled are largely collinear, and over 90% of genes are conserved. *L. brevidens* and *L. subracemosa* have evidence of an ancient, shared whole-genome duplication event, and retained genes have neofunctionalized, with desiccation-specific expression in *L. brevidens*. Tandem gene duplicates also are enriched in desiccation-associated functions, including a dramatic expansion of early light-induced proteins from 4 to 26 copies in *L. brevidens*. A comparative differential gene coexpression analysis between *L. brevidens* and *L. subracemosa* supports extensive network rewiring across early dehydration, desiccation, and rehydration time courses. Many *LATE EMBRYOGENESIS ABUNDANT* genes show significantly higher expression in *L. brevidens* compared with their orthologs in *L. subracemosa*. Coexpression modules uniquely upregulated during desiccation in *L. brevidens* are enriched with seed-specific and abscisic acid-associated *cis*-regulatory elements. These modules contain a wide array of seed-associated genes that have no expression in the desiccation-sensitive *L. subracemosa*. Together, these findings suggest that desiccation tolerance evolved through a combination of gene duplications and network-level rewiring of existing seed desiccation pathways.

INTRODUCTION

Comparative systems are a powerful tool for dissecting the molecular basis of complex biological traits. The origins of desiccation tolerance in resurrection plants are largely unknown, but the underlying genetic signatures could be traced using pairs of closely related desiccation-sensitive and -tolerant species. Such an approach has been applied to *Eragrostis* (Vander Willigen et al., 2001), *Selaginella* (Yobi et al., 2013), and *Sporobolus* (Oliver et al., 2011) at the morphological and biochemical levels to identify signatures that distinguish drought and desiccation responses. Detailed pairwise comparisons have identified changes in cell wall composition (Plancot et al., 2014), metabolite and osmoprotectant accumulation (Oliver et al., 2011; Yobi et al., 2013), and physical properties unique to desiccation-tolerant species. Although genomes are available for several resurrection plants (VanBuren et al., 2015, 2018; Xiao et al., 2015; Costa et al., 2017), genomic resources in these comparative lineages are limited, and no genomes of closely related desiccation-sensitive species have been sequenced. High-quality reference genomes are available for the desiccation-sensitive *Selaginella moellendorffii* (Banks et al., 2011) and the desiccation-tolerant *Selaginella lepidophylla* (VanBuren et al., 2018), but their estimated divergence 248 million years ago prevents detailed genomic comparisons (Baniaga et al., 2016).

Resurrection plants endure extreme and prolonged drought events through vegetative desiccation, entering a preserved and protected quiescent state that functionally mirrors seed dormancy in angiosperms. Desiccation tolerance was a critical adaptation during early land plant evolution, and many early-diverging fern, moss, and lycophyte lineages have retained or convergently evolved these ancestral resilience mechanisms (Proctor, 1990; Oliver et al., 2000; Lüttge et al., 2011). Vegetative desiccation tolerance is comparatively less common in angiosperms, and recent genomic and metabolic studies suggest that it evolved through rewiring seed desiccation pathways (Costa et al., 2017; VanBuren et al., 2017). Resurrection plants have a conserved set of molecular signatures associated with desiccation tolerance

[1] Address correspondence to bobvanburen@gmail.com or unb137@uni-bonn.de.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Robert VanBuren (bobvanburen@gmail.com).
[OPEN]Articles can be viewed without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.18.00517

IN A NUTSHELL

Background: Plants have evolved numerous strategies over the last ~400 million years to overcome water limitations. The most extreme way to cope with drought is to simply dry out when water is limited and enter a quiescent, near desiccated state until water becomes available. Resurrection plants can survive years without water and resume photosynthesis and growth within a few days after rehydration. Roughly 135 flowering plants are desiccation tolerant, and several resurrection plant species have emerged as models to dissect the genetic basis of this extreme adaptation. Genomes from resurrection plants have been sequenced, but insights are limited by a lack of dehydration-sensitive outgroups for comparisons. Here, we compared two closely related species with contrasting desiccation tolerance to identify elements that control this trait.

Question: We wanted to know what genomic changes gave rise to desiccation tolerance and test the hypothesis that this trait evolved through repurposing of existing seed desiccation pathways. We also aimed to test the role of gene duplication and network rewiring in the evolution of this complex trait.

Findings: We leveraged a unique comparative system of closely related, desiccation-tolerant, and desiccation-sensitive species to identify genetic changes associated with desiccation tolerance. Though the two *Lindernia* genomes we sequenced are highly similar, we identified differences in gene duplication associated with desiccation tolerance. This includes expansion of a high-light protection protein from 1 copy to 22 in the desiccation-tolerant *L. brevidens*. Several seed-specific pathways are activated only in *L. brevidens*, indicative of network rewiring and co-option of existing genes. Gene expression is stable after moderate dehydration in *L. brevidens*, which reflects the successful deployment of protective mechanisms. Dynamic and chaotic expression patterns in the desiccation-sensitive *L. subracemosa* reflect last-ditch efforts to avoid imminent death. Our findings suggest that the evolution of desiccation tolerance is complex and likely quantitative.

Next steps: We identified a series of gene duplications and rewired pathways that likely control desiccation tolerance. The next step is to functionally validate these interactions using the existing transformation system in *Lindernia*. Components of desiccation-related pathways will be useful for engineering improved drought tolerance into crop plants.

(Illing et al., 2005; Zhang and Bartels, 2018), but the underlying genomic basis of this trait is largely unknown. The cooption of seed- and desiccation-associated pathways in resurrection plants can occur through a broad range of mechanisms at the gene, pathway, or network level. Changes in gene regulation at network hubs could drive pathway-level rewiring to upregulate a cascade of desiccation-related mechanisms. Changes in *cis*-regulation at isolated nodes in a network could shift the stoichiometry or abundance of endpoint metabolites and proteins related to desiccation. Gene and genome duplication also can drive adaptive evolution by providing additional copies for pathway- and gene-level subfunctionalization and neofunctionalization.

Desiccation tolerance is prominent in Linderniaceae (order Lamiales) within the clade spanning *Craterostigma* and *Lindernia* (Rahmanzadeh et al., 2005). *Craterostigma plantagineum* is a model resurrection plant (Bartels and Salamini, 2001) native to rocky outcrops of sub-Saharan Africa. All *Craterostigma* and some desert-adapted *Lindernia* species are desiccation tolerant, but most *Lindernia* species are desiccation sensitive. *Lindernia brevidens* is unusual in that it displays desiccation tolerance despite an endemic habitat in the montane rainforests of eastern Africa that never experience seasonal drying (Phillips et al., 2008). Desiccation tolerance is likely an ancestral trait in this group (Fischer et al., 2013) and was retained in *L. brevidens* before its radiation to the tropical rainforest. *Lindernia* is paraphyletic (Fischer et al., 2013), and species outside of the clade containing *Craterostigma* and *L. brevidens*, such as *Lindernia subracemosa*, are desiccation sensitive. This diversity in desiccation tolerance makes *Lindernia* an excellent comparative system in which to test the contribution of gene duplication, *cis*-elements, and pathway rewiring in the evolution of desiccation tolerance. Here, we assembled high-quality reference genomes for the desiccation-tolerant *L. brevidens* and desiccation-sensitive *L. subracemosa*. Detailed comparative genomics and differential coexpression network analysis allowed us to survey the genetic basis of desiccation tolerance in *Lindernia*."

RESULTS

Comparative Grade Reference Genomes for *Lindernia*

C. plantagineum is a well-studied model for the evolution of desiccation tolerance in eudicots, but its highly complex, octoploid genome has hindered genome-scale analyses. *L. brevidens* and *L. subracemosa* are diploid with relatively small genomes (270 and 250 Mb, respectively), providing an excellent alternative system (Figure 1).

We generated high-quality reference genomes for both *Lindernia* species using a PacBio-based, single-molecule real-time sequencing approach. In total, we generated 21.7 and 17.9 Gb of filtered PacBio data, collectively representing 80.3× and 71.6× coverage for *L. brevidens* and *L. subracemosa*, respectively (Supplemental Figure 1). Raw PacBio reads were error corrected and assembled using the long-read assembler Canu (Koren et al., 2017), which is optimized to avoid collapsing highly repetitive and tandemly duplicated regions. Contigs were polished using high-coverage Illumina data with Pilon (Walker et al., 2014) to remove residual errors. The *L. brevidens* assembly spanned 265 Mb across 267 contigs with a contig N50 (length where half or more of the assembly is contained) of 3.6 Mb. The *L. subracemosa* assembly was slightly smaller, at 246 Mb with 328 contigs and an

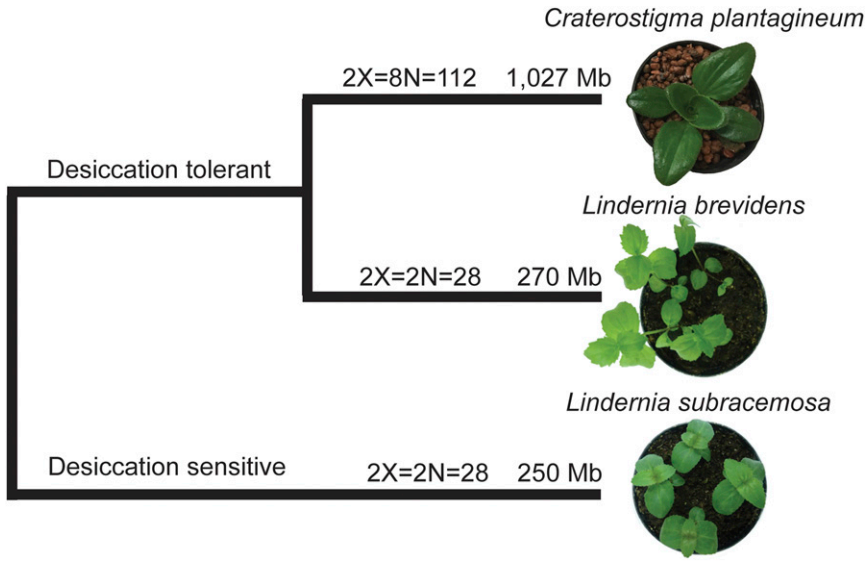

Figure 1. Comparative Desiccation Tolerance System within Linderniaceae.

Inferred phylogeny from Fischer et al. (2013) shows the two model desiccation-tolerant species (*C. plantagineum* and *L. brevidens*) and the desiccation-sensitive outgroup (*L. subracemosa*). Ploidy, karyotype, and genome size are shown on branches.

N50 of 1.9 Mb (Table 1). The total assembly sizes were consistent with the estimated genome sizes of 270 and 250 Mb based on flow cytometry. *Lindernia* species are primarily self-pollinated with low residual within-genome heterozygosity, which contributed to the high contiguity and relatively simple graph-based assembly structures (Supplemental Figures 2 and 3).

We used high-throughput chromatin conformation capture (Hi-C) to generate a chromosome-scale assembly of *L. brevidens*. The Hi-C-based Illumina reads were mapped to the draft assembly using bwa (Li, 2013) followed by filtering and proximity-based clustering using the Juicer pipeline (Durand et al., 2016) (Supplemental Table 1). This approach yielded 14 high-confidence clusters corresponding to the haploid chromosome number in *L. brevidens* ($2n = 2x = 28$; Figure 2). In total, 121 contigs were ordered and oriented into 14 scaffolds collectively representing 94.7% of the assembly (249 out of 263 Mb; Supplemental Table 2). This included anchoring 98.8% of the predicted gene models. The repetitive element density was inversely correlated with gene density, and most chromosomes contained large tracts of retrotransposons (RTs), which likely correlate with centromere position (Figure 3).

The genomes of *L. brevidens* and *L. subracemosa* were of similar size and the same karyotype, suggesting that they should have comparable repetitive elements and gene composition. Long terminal repeat (LTR)-RTs were the most abundant repetitive elements in both genomes, and they collectively spanned 34% (92.0 Mb) and 31% (77.4 Mb) of the *L. brevidens* and *L. subracemosa* genomes, respectively (Table 1). Despite the similar overall LTR composition, *L. subracemosa* had significantly more intact LTRs compared with *L. brevidens* (1972 versus 1025; Wilcoxon rank sum, $P < 0.05$). The distribution of LTR-RT insertion time was similar in both species, and most intact elements inserted within the last 1 million years (Figure 4). These findings suggest that LTR-RTs are similarly active in both genomes but may fractionate more quickly in *L. brevidens*.

The overall gene composition was similar in both *Lindernia* species, though *L. subracemosa* had more annotated gene

Table 1. *Lindernia* Genome Assembly Metrics

Metric	*L. subracemosa*	*L. brevidens* V1 (Draft)[a]	*L. brevidens* V2 (Chromosome Scale)
No. of contigs	328	327	144 (unplaced)[b]
No. of scaffolds	NA[c]	NA	14
Contig N50	1.9 Mb	3.7 Mb	18.7 Mb
Total length	246,514,821 bp	271,249,608 bp	263,221,403 bp
LTR composition	77.4 Mb (31%)[d]	92.0 Mb (34%)	92.0 Mb (34%)
N. of gene models	33,344	27,204	27,204

[a]*L. brevidens* V1 represents the contig-level PacBio-based assembly and V2 represents the chromosome-scale assembly of *L. brevidens* anchored using Hi-C data.
[b]Unplaced scaffolds were not anchored into the chromosome-scale assembly.
[c]NA, not applicable. The *L. subracemosa* and *L. brevidens* V1 assemblies are contig level and thus contain no scaffolds.
[d]Proportion (percentage) of the genome represented by LTRs.

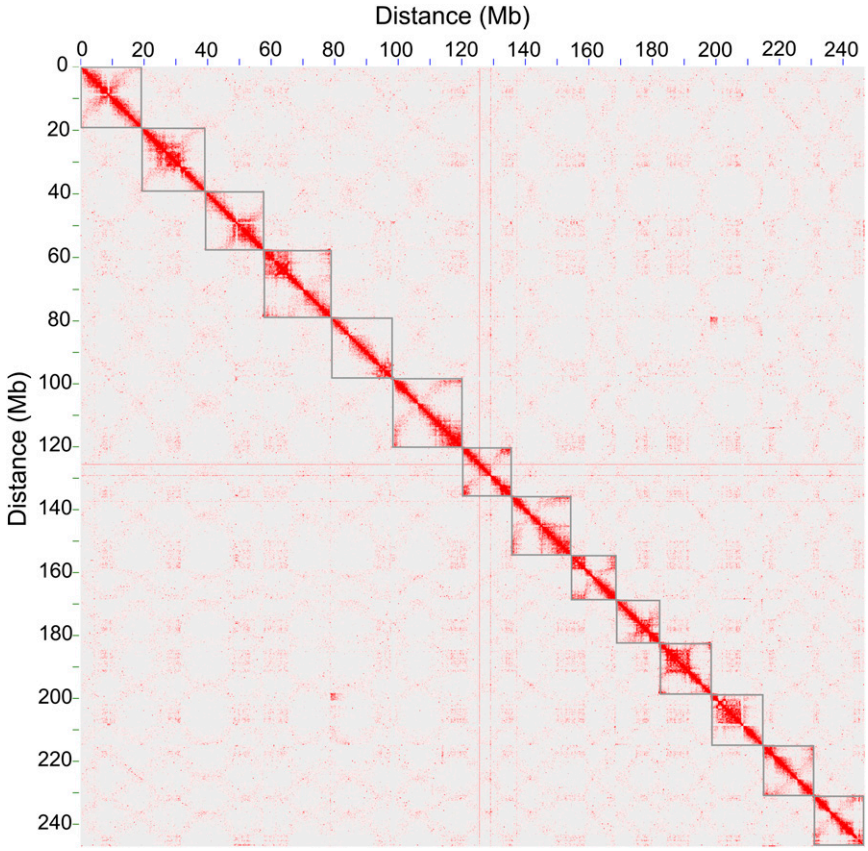

Figure 2. Hi-C Clustering for Pseudomolecule Construction in *L. brevidens*.

A postclustering heat map of Hi-C-based intrachromosomal interactions in *L. brevidens* is shown. Pseudomolecules corresponding to the 14 haploid chromosomes are delineated by gray boxes.

models. Ab initio gene prediction using the dehydration time-course RNAseq data and protein similarities to other angiosperms identified 27,204 and 33,344 gene models in *L. brevidens* and *L. subracemosa,* respectively (Table 1). We assessed annotation quality using the BUSCO pipeline and found 91% and 90% (1319 and 1298) of the 1440 genes in the Embryophyta data set present in the *L. brevidens* and *L. subracemosa* assemblies. This proportion is comparable with results from other recent PacBio-based genomes.

Comparative Genomics of *Lindernia*

The *L. brevidens* and *L. subracemosa* genomes were largely collinear based on whole-genome alignment, and 24,053 *L. brevidens* genes had syntenic orthologs in *L. subracemosa*. Roughly 70% of the genomes were conserved in 2:2 syntenic blocks, supporting a shared, ancient whole-genome duplication (WGD) event in both species (Figure 5; Supplemental Figures 4 and 5). Six of the seven ancestral homeologous chromosome pairs from the WGD were intact in *L. brevidens*, including modern chromosome pairs: 1 and 13, 2 and 14, 3 and 5, 6 and 9, 7 and 10, and 8 and 11 (Figure 5A). Two of the ancestral homeologous chromosomes were fused in modern chromosome 5, and chromosome 12 contained fragments from several ancestral chromosomes. Chromosomal rearrangements were difficult to identify in *L. subracemosa* given its contig-level assembly, but there were no obvious rearrangements based on macrosynteny with *L. brevidens* (Supplemental Figure 5). The ancestral subgenomes were heavily fractionated, and only 7742 gene pairs were retained in duplicate in *L. brevidens* and 8452 in *L. subracemosa* based on synteny. Gene-level fractionation was biased toward a dominant subgenome that contained significantly more genes (Figure 5B; Supplemental Figure 6).

We identified patterns of gene duplication and loss that may be related to the evolution of desiccation tolerance and other lineage-specific traits. Most gene pairs from the WGD were either retained in duplicate or fractionated to single copies in both species, including 11,874 single-copy genes (1:1) and 7568 duplicated genes (2:2) in both genomes (Table 2). We identified 3200 lineage-specific genes in *L. brevidens* (1:0 or 2:0) and 7067 lineage-specific genes in *L. subracemosa* (0:1 or 0:2) based on synteny. The higher number of lineage-specific genes in *L. subracemosa* was likely related to differences in total annotated gene number (27,204 versus 33,344). The lineage-specific genes in L. *brevidens* were enriched in Gene Ontology (GO) terms related to chlorophyll biosynthesis and metabolism, regulation of mitosis, and response

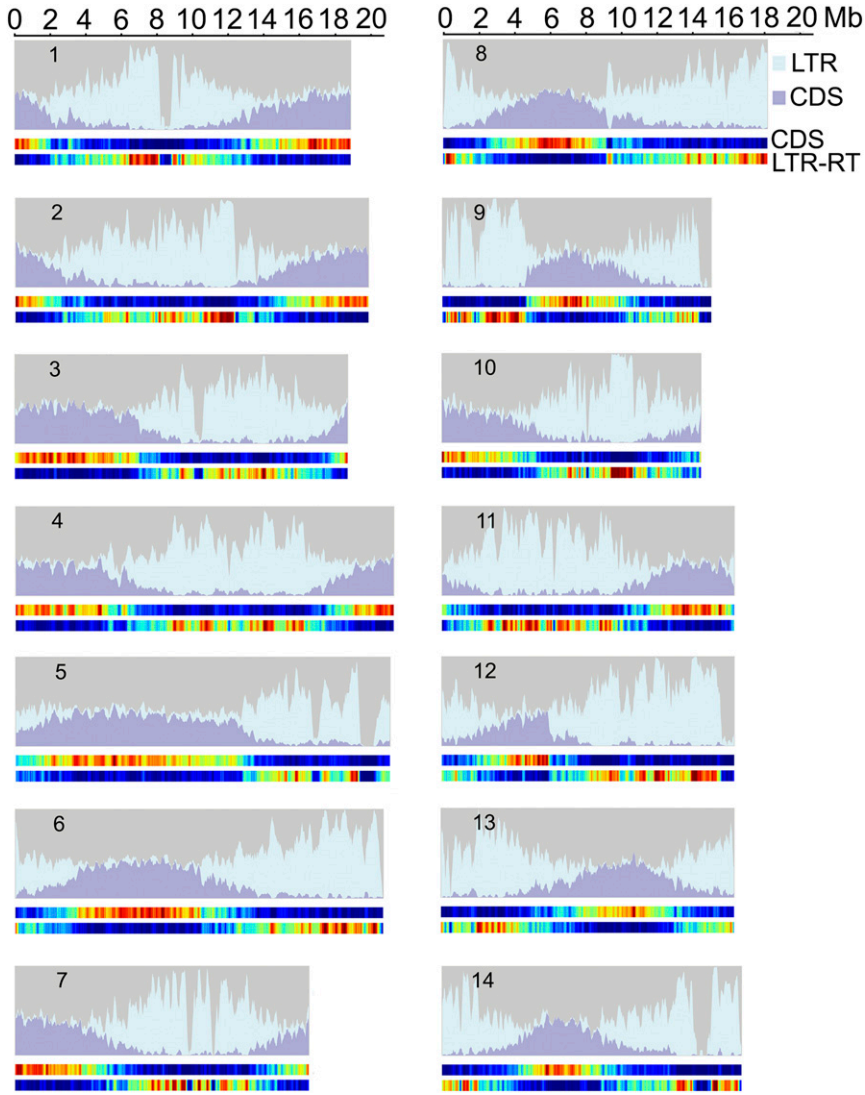

Figure 3. Landscape of the *L. brevidens* Genome.

LTR-RT and coding sequence (CDS) density are plotted in sliding windows of 50 kb with 25-kb step size for the 14 *L. brevidens* scaffolds (chromosomes). Red indicates high density and blue indicates low density of coding sequence and LTR-RTs in the heat map below each landscape.

to heat, which may suggest a role for these pathways in desiccation tolerance (Supplemental Table 3).

New genes can arise through tandem gene duplication, and tandem gene duplications are associated with adaptive evolution (Cannon et al., 2004), including of desiccation tolerance in other resurrection plants (VanBuren et al., 2015, 2018). *L. brevidens* and *L. subracemosa* had a similar overall number of tandem genes but major differences in array size. *L. brevidens* had 2673 tandem arrays containing 5345 genes with array sizes ranging from 2 to 24 members. *L. subracemosa* had 3404 tandem arrays across 6809 genes with array sizes ranging from 2 to 31 (Figure 6). Through cross-referencing with syntenic gene pairs, we found that most tandem arrays were conserved between *L. brevidens* and *L. subracemosa*. Only 153 tandem arrays were specific to *L. brevidens* and 247 arrays were specific to *L. subracemosa*. Although tandem gene arrays were generally conserved, array sizes were highly variable and few contained the same number of genes between species (Figure 6B). Together, these data suggest that most tandem gene duplication events are ancestral but that each species has undergone unique array expansion and contraction.

Global Expression Patterns and Desiccation-Related Network Rewiring

To construct a comparative framework of genes related to desiccation, we conducted parallel sampling of leaf tissue during desiccation and rehydration time courses in *L. brevidens* and *L. subracemosa*. Parallel sampling between species allowed us to distinguish between genes involved in typical dehydration

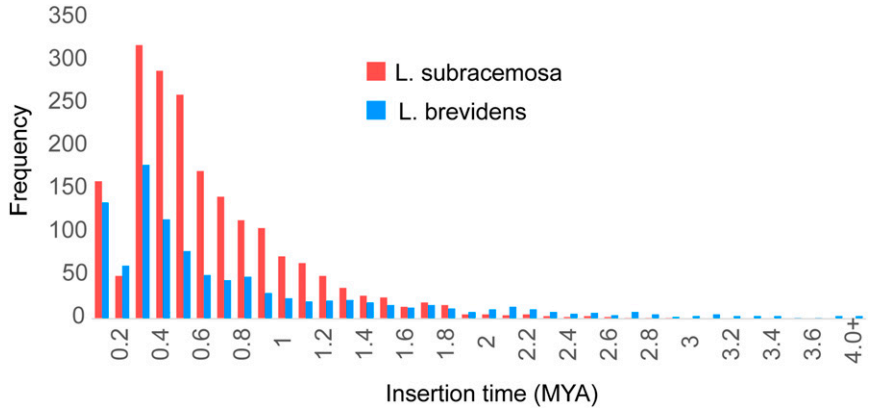

Figure 4. Insertion Time of Intact LTR-RTs.

The average insertion time calculated from the divergence of LTR pairs is plotted for the 1025 intact LTR-RTs in *L. brevidens* and the 1972 intact LTR-RTs in *L. subracemosa*. MYA, million years ago.

responses and those related specifically to desiccation tolerance. Sampling ranged from mild dehydration stress (relative water content [RWC] 53–56%; 3 d) through severe dehydration (RWC 23–27%; 7 d) and desiccation (RWC 6–9%; 10 and 14 d), followed by 24 and 48 h post rehydration (Figure 7). RWC was 53 to 56% at day 3 and fell below 10% after 10 d of drought in both species (Figure 7A). *L. subracemosa* plants were largely dead upon rehydration, and *L. brevidens* plants were mostly viable and physiologically active at 48 h post rehydration (RWC 44%) (Figure 7).

The greatest changes in gene expression occurred at two time points during the transition from mild to severe dehydration stress and from desiccated to rehydrated (Supplemental Table 4). The number of differentially expressed genes between well-watered and mild dehydration (F versus D3) were relatively similar in both species (5322 versus 4824 in *L. subracemosa* and *L. brevidens*). Many syntenic gene pairs had similar expression levels, with 581 upregulated and 133 downregulated in both species. Significantly more genes were differentially expressed between mild and severe dehydration stress (D3 versus D7), with 4329 and 9227 differentially expressed genes, respectively. A similar proportion of syntenic gene pairs was upregulated in both species at D7 (581), but significantly more gene pairs were similarly downregulated (1396) compared with mild dehydration stress. This pattern suggests that there is conservation of downregulated pathways in desiccation-sensitive and -tolerant species.

In *L. brevidens*, gene expression was relatively stable from severe dehydration to desiccation (D7, D10, and D14), whereas a high proportion of genes were differentially expressed in *L. subracemosa* during the transition to desiccation. This reflects the stability of desiccated *L. brevidens* and the imminent death of *L. subracemosa*. Few genes were similarly differentially expressed in both species under severe dehydration and desiccation (D7 versus D10: 50 and 62; D10 versus D14: 3 and 0, upregulated and downregulated, respectively; Supplemental Table 4). A substantial proportion of syntenic gene pairs (2065) were similarly upregulated in both species during early rehydration (24 h), supporting the conclusion that there is conserved activation of repair pathways. Expression changes in both species were minimal between 24 and 48 h post rehydration. Although *L. subracemosa* had some transcriptional response post rehydration, this was not sufficient to repair the extensive desiccation-induced damage. Together, the divergent expression patterns suggest that there is extensive upregulation of distinct pathways with desiccation-specific roles.

We conducted GO enrichment analysis of the gene pairs that were uniquely upregulated in *L. brevidens*, with no change or a decrease in expression in *L. subracemosa*. We reasoned that such genes are likely to be specific to the induction of desiccation tolerance. Most of the GO terms enriched among genes upregulated in mild dehydration stress (D3) were related to responses to abiotic stress and secondary metabolite biosynthesis (Supplemental Table 5), suggesting early activation of protective mechanisms. There were only a few GO terms enriched among genes upregulated in severe dehydration (D7) and desiccation (D10 and D14), including terms related to transport, vacuole organization, ion homeostasis, and RNA modification (Supplemental Table 5). Most GO terms of genes uniquely downregulated in *L. brevidens* under mild and severe dehydration stress were related to photosynthesis processes, suggesting that the photosynthetic apparatus is inactivated early under mild dehydration compared with *L. subracemosa* (Supplemental Table 6).

The large-scale expression changes unique to desiccation in *L. brevidens* may be driven by changes in *cis*-regulation. Genes with unique desiccation-related expression in *L. brevidens* were enriched with *cis*-regulatory elements associated with dehydration and abscisic acid (ABA)-mediated responses as well as seed development pathways (Supplemental Table 7). Enriched *cis*-elements associated with typical ABA-mediated dehydration responses included ABF1 and ABF2 (Yoshida et al., 2015) among others. Enriched seed maturation-associated *cis*-elements included bZIP53 (Alonso et al., 2009), ABA-responsive element binding protein3 (AREB3; Nakashima et al., 2009), and ABI5 (Lopez-Molina et al., 2001). *Cis*-elements at the interface of heat- and dehydration-mediated ABA responses, such as Heat Stress Factor A6b (HSFA6B; Huang et al., 2016) and HSF7, also were

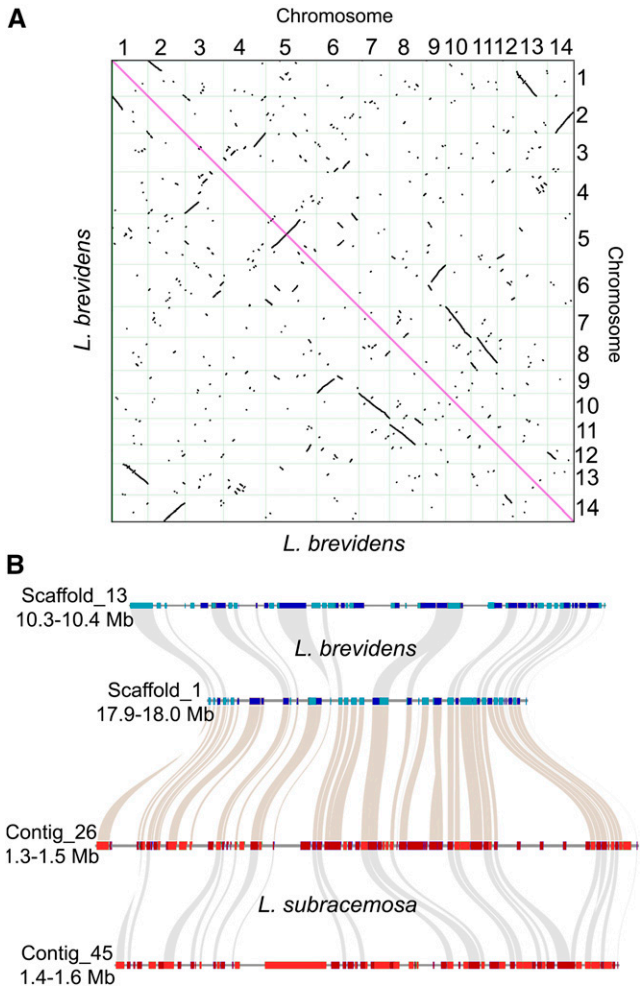

Figure 5. Comparative Genomics of *L. brevidens* and *L. subracemosa*.

(A) Syntenic dot plot of *L. brevidens* showing retained gene pairs from the recent WGD event. Each black dot represents a pair of retained genes. **(B)** Microsynteny between syntenic blocks of the *L. brevidens* (top) and *L. subracemosa* (bottom) genomes. Syntenic gene pairs between *L. brevidens* versus *L. subracemosa* are shown by brown connections, and retained WGD gene pairs within each genome are shown by gray connections. Genes are colored by orientation in *L. brevidens* (light blue are forward, dark blue are reverse) and *L. subracemosa* (light red are forward, dark red are reverse).

enriched in desiccation-related genes. Enriched *cis*-elements in genes downregulated during desiccation in *L. brevidens* had wide roles in plant growth and development, hormone responses, and photosynthesis (Supplemental Table 8).

To compare network-level gene expression in *L. brevidens* and *L. subracemosa*, we utilized a weighted correlation network analysis approach (Langfelder and Horvath, 2008) across the dehydration and rehydration RNAseq time courses. This comparative coexpression network approach allowed us to parse conserved dehydration-related patterns from desiccation-specific pathway rewiring. After filtering genes with low expression (see Methods), we constructed two coexpression networks, with 14,246 genes in 10 modules for *L. brevidens* and 14,075 genes in 9 modules for *L. subracemosa* (Figure 8; Supplemental Figures 7 and 8). Based on their temporal dynamics, coexpression modules could be broadly classified into three groups: (1) high expression in well-watered tissue but downregulation in dehydration/desiccation; (2) expression during early dehydration; and (3) sustained high expression throughout dehydration and desiccation (Figure 8). Modules 1, 2, 3, and 5 in the *L. brevidens* network and modules 2, 4, and 5 in *L. subracemosa* had high expression in well-watered and rehydrating conditions, with decreasing expression throughout dehydration/desiccation time points (Figure 8). Modules 7 and 8 in *L. brevidens* and modules 3, 7, and 8 in *L. subracemosa* were involved in early dehydration responses, with peak expression at 3 or 7 d of dehydration. Modules 4, 6, and 10 in *L. brevidens* and modules 1 and 3 in *L. subracemosa* had sustained dehydration and desiccation-induced expression.

We compared module overlap between the networks to identify patterns of conservation and species-specific divergence. Modules downregulated during severe dehydration and desiccation were

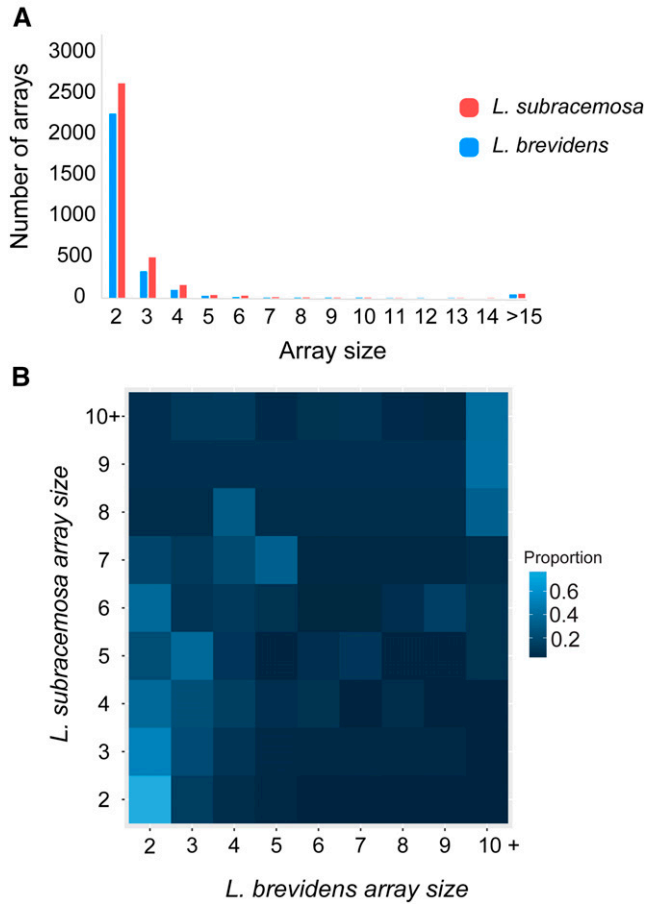

Figure 6. Comparison of Tandem Gene Arrays in *Lindernia*.

(A) Histogram of tandem array sizes.
(B) Heat map of tandem array size in syntenic orthologs between *L. brevidens* and *L. subracemosa*. Values are plotted as the proportion of tandem genes in each category against all the genes in that array size.

Table 2. Comparison of Biased Fractionation Following the Shared WGD in *Lindernia*

Gene Classification	Gene Ratio[a]	No. of Genes/Pairs
Single copy (both species)	1:1	11,874
Duplicate retained (both species)	2:2	7,568
Duplicate retained (*L. brevidens*)	2:1	1,276
Duplicate retained (*L subracemosa*)	1:2	3,286
L. brevidens specific (single copy)	1:0	3,026
L. subracemosa specific (single copy)	0:1	6,183
L. brevidens specific (duplicated)	2:0	174
L. subracemosa specific (duplicated)	0:2	884

[a]Gene ratios are shown as *L. brevidens*:*L. subracemosa*.

largely conserved between the two species, with 74% of gene pairs falling in the same group of modules (Supplemental Figure 9). Modules upregulated during dehydration and desiccation had comparatively little overlap between species, and only 43% of gene pairs were in overlapping modules (Supplemental Figure 9). This indicates that there was significant pathway rewiring during desiccation. Network-wide *cis*-regulatory element enrichment patterns mirrored the observations comparing pairwise differentially expressed genes. Desiccation-associated modules from the *L. brevidens* network were enriched in dehydration-associated ABA-responsive *cis*-elements and seed maturation-associated *cis*-elements, including bZIP53, AREB3, and ABI5 among others (Supplemental Table 9).

Unique Desiccation-Related Pathways in *L. brevidens*

The similarities between seed and vegetative desiccation suggest overlapping pathways, which is supported by expression data from several resurrection plant lineages (Costa et al., 2017; VanBuren et al., 2017). We identified a wide range of seed-specific genes and pathways that were expressed only under dehydration in *L. brevidens* compared with syntenic orthologs in *L. subracemosa* (Supplemental Table 10). Seed storage proteins serve as a reserve of nitrogen, carbon, and sulfur for germinating seeds, and they likely play a role in seed longevity (Nguyen et al., 2015). Orthologs to genes encoding 2S and 12S seed storage proteins were generally upregulated in *L. brevidens* under desiccation, and syntenic orthologs in *L. subracemosa* were not expressed or were expressed highly in well-watered conditions (Supplemental Table 10). Delay of germination1 (DOG1) is an essential component of seed dormancy regulation, and its expression affects hundreds of seed-related genes (Dekkers et al., 2016). *DOG1* was highly expressed in well-watered *L. brevidens* tissues but was downregulated during desiccation. The *L. subracemosa* *DOG1* transcript had a low basal-level expression in all time points.

Oil bodies are lipid organelles filled with triacylglycerols that function as high-density energy reserves during seed germination. Oil bodies accumulate in desiccated leaf tissue of *Oropetium* and likely play a role in desiccation tolerance. Oil bodies are enveloped with oleosin structural proteins that prevent membrane coalescence and protect membrane integrity during freeze-thaw cycles (Shimada et al., 2008). Oil body membranes also are studded with calcium binding caleosin proteins that are associated with oil body degradation (Poxleitner et al., 2006) and general stress response pathways (Shen et al., 2014). *L. brevidens* and *L. subracemosa* had similar numbers of genes for oleosin (eight versus seven) and caleosin (four versus three) proteins, although *L. brevidens* had more retained whole-genome and tandem duplicates (Figure 9). Most oleosin and caleosin genes had low or

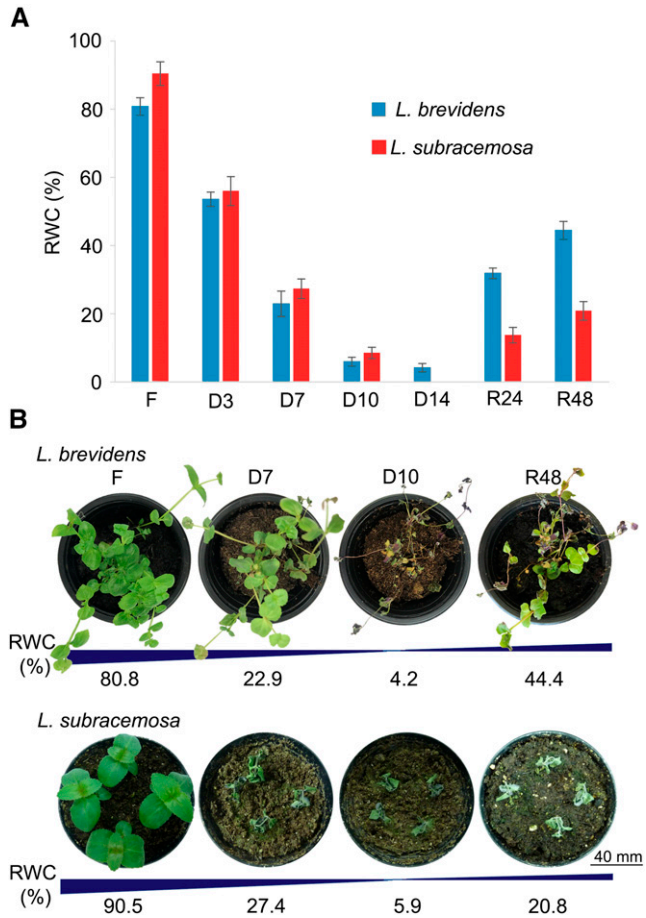

Figure 7. Overview of Desiccation and Rehydration Processes in *Lindernia*.

(A) RWC of fresh leaf tissues (F), 3, 7, 10, and 14 d of drought (D), and 24 and 48 h post rehydration (R) in *L. brevidens* and *L. subracemosa*. Error bars represent the SE with three replicates for each RWC measurement.
(B) Representative *L. brevidens* (top) and *L. subracemosa* (bottom) at various RWC.

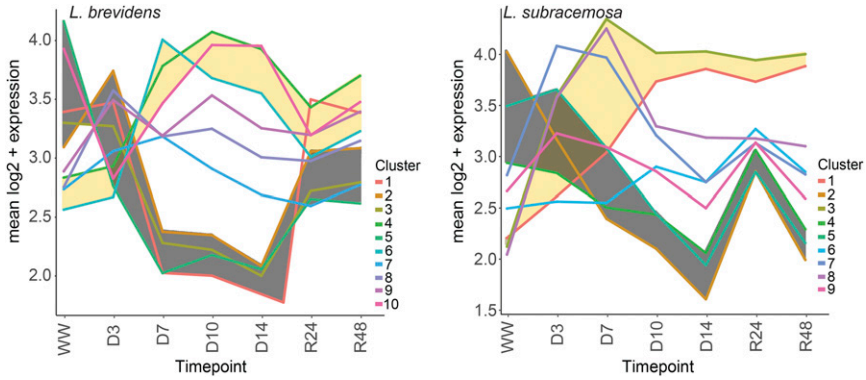

Figure 8. Comparative Coexpression Networks during Desiccation and Rehydration in *Lindernia*.

The mean expression of genes from modules in the coexpression network in *L. brevidens* (left) and *L. subracemosa* (right) are plotted for the seven time points. Conserved modules with upregulation during desiccation are highlighted in yellow in both networks, and modules with downregulation during desiccation are highlighted in gray.

undetectable expression in well-watered tissue, but several were induced during progressive dehydration and desiccation. Six oleosin genes in *L. brevidens* and three oleosin genes in *L. subracemosa* were upregulated in desiccating tissue, with most having a peak expression of less than 30 transcripts per million (TPM). *L. brevidens* had a pair of retained duplicated oleosin genes where one ortholog peaked at more than 500 TPMs in desiccating tissue, and the single syntenic ortholog in *L. subracemosa* had a relatively low expression. No caleosin genes were upregulated in *L. subracemosa* upon desiccation, but a pair of syntenic orthologs in *L. brevidens* were abundantly expressed (Figure 9).

Early light-induced proteins (ELIPs) are predicted to bind chlorophyll and function in photoprotection under high light and other abiotic stresses. The *L. subracemosa* genome had four genes encoding ELIP proteins, including a pair of syntelogs retained from the WGD event (Figure 10). Expression of two *ELIP* genes was hardly detectable during the surveyed time points, and two others were highly expressed during dehydration. The number of ELIPs in *L. subracemosa* was similar to that in other desiccation-sensitive angiosperms, and their dehydration-induced expression was consistent with the hypothesized protective mechanisms (Hayami et al., 2015). By contrast, the *L. brevidens* genome had undergone a dramatic expansion of ELIP genes, with 26 in total, including a large tandem array of 19 duplicates. This large tandem array was collinear to a pair of retained syntenic orthologs in *L. subracemosa* and a single retained gene copy in *L. brevidens*. Nearly all the *ELIP* genes in this array, and dispersed copies throughout the genome, were highly expressed during severe dehydration, desiccation, and rehydration, but they were hardly expressed in well-watered and mildly dehydrated tissue (Figure 10B). The tandem array was syntenic with the highly expressed ortholog in *L. subracemosa*, and the single-copy syntelog in *L. brevidens* and its syntenic ortholog in *L. subracemosa* were not expressed in dehydrated tissue. This suggests an ancestral subfunctionalization of this duplicated pair where only one gene copy was involved in dehydration-related responses. After the divergence of *L. brevidens* and *L. subracemosa*, the dehydration-specific syntelog likely underwent massive tandem proliferation in *L. brevidens*.

In Arabidopsis (*Arabidopsis thaliana*), STAY-GREEN (SGR) proteins are key regulators of chlorophyll degradation, and they are typically upregulated under abiotic stresses (Sakuraba et al., 2014b). Syntenic orthologs of SGR were highly expressed in both *Lindernia* species during dehydration/desiccation. STAY-GREEN

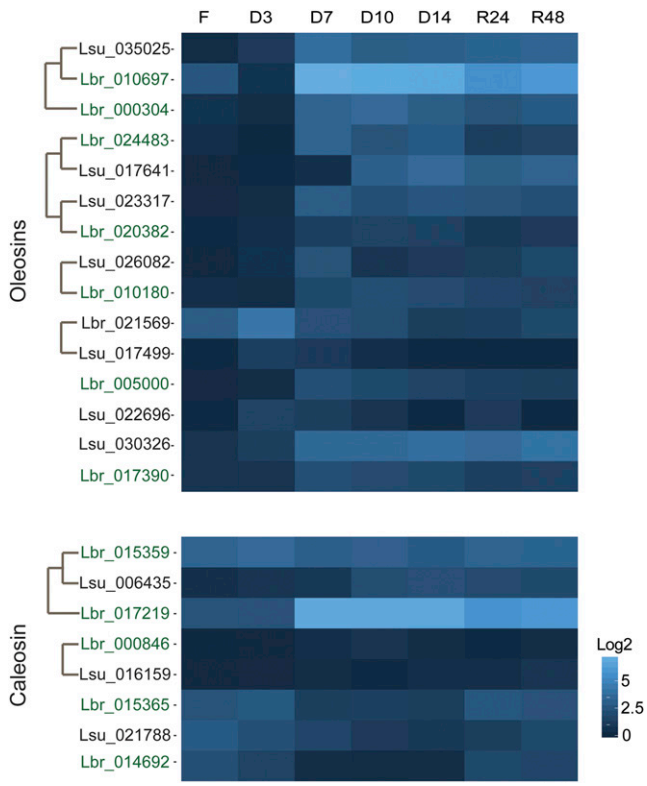

Figure 9. Subfunctionalization of Oleosin and Caleosin Genes in *L. brevidens*.

Heat maps of $\log_2$ transformed expression of genes encoding oleosins (top) and caleosins (bottom) in *L. brevidens* and *L. subracemosa*. Syntenic orthologs are connected by brown lines including 1:2 and 2:2 orthologs between species.

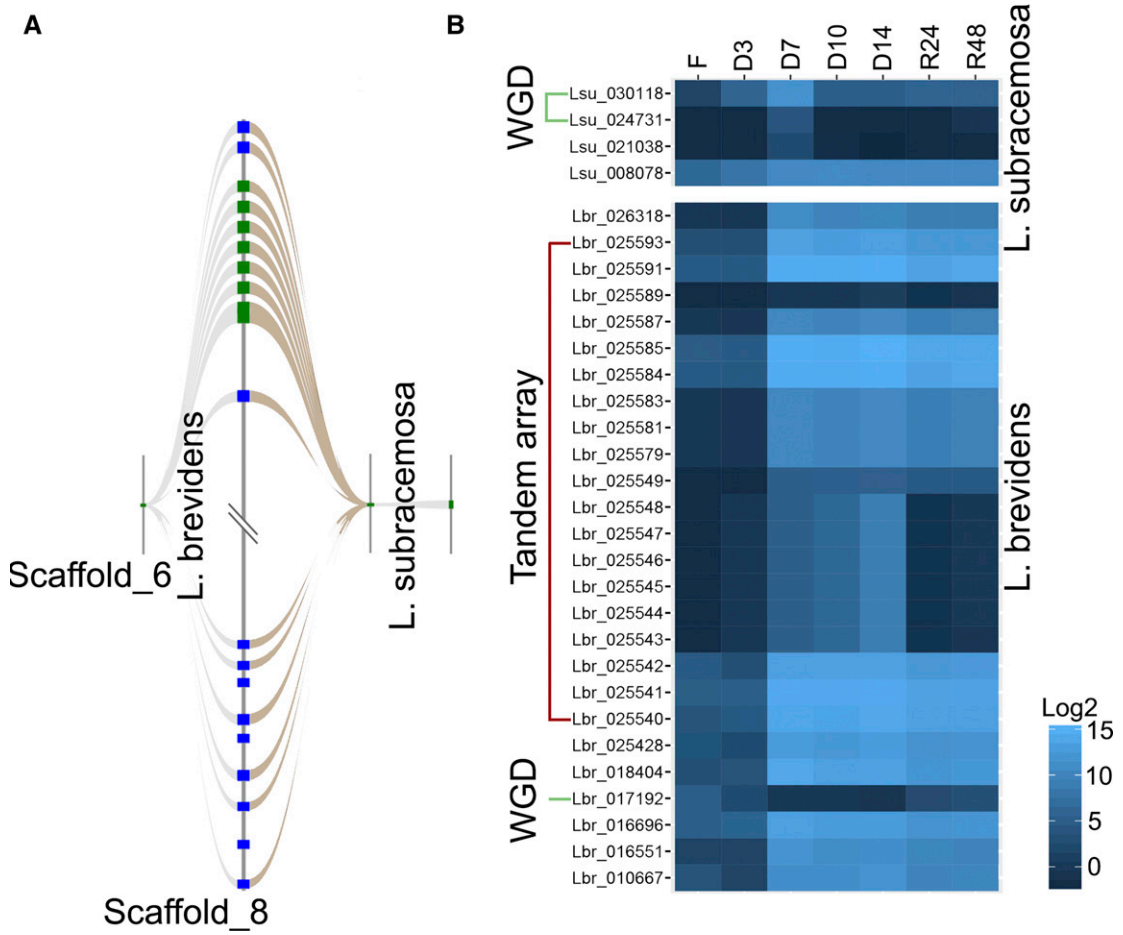

Figure 10. Tandem Proliferation of Desiccation-Associated *ELIP* Genes in *L. brevidens*.

(A) Microsynteny of a large tandem gene array in *L. brevidens* compared with the single-gene syntenic ortholog in *L. brevidens* and the two syntenic regions in *L. subracemosa*.

(B) Log_2 transformed expression patterns of *ELIP*s in *L. subracemosa* (top) and *L. brevidens* (bottom). The large tandem array and syntenic whole-genome duplicates are labeled.

LIKE (SGRL) proteins are negative regulators of chlorophyll degradation, and overexpression of SGRL2 leads to a stay-green phenotype (Sakuraba et al., 2014a). The syntenic ortholog of SGRL was highly induced during desiccation in *L. brevidens*, but the *L. subracemosa* ortholog had no detectable expression (Supplemental Table 10).

Carbohydrate metabolism is heavily shifted during desiccation, and sucrose, trehalose, and short-chain oligosaccharides function as osmoprotectants to stabilize cellular macromolecules. Sucrose is the most abundant carbohydrate in most resurrection plants, and the accumulation of sucrose distinguishes desiccation-sensitive and -tolerant *Eragrostis* species (Illing et al., 2005). *Craterostigma* and *L. brevidens* accumulate the unusual C8 sugar 2-octulose in photosynthetic tissues, which serves as a reservoir of sucrose accumulation during desiccation (Bianchi et al., 1991; Phillips et al., 2008). Transketolase7 and -10 catalyze the formation of octulose-8-phosphate in *Craterostigma* (Zhang et al., 2016), and the orthologous transketolase genes were highly expressed in *L. brevidens* leaf tissue (Supplemental Table 10). This included two pairs of syntenic 1:1 orthologs and a trio of retained 2:1 duplicates, with upregulation in well-watered and rehydrating tissue in *L. brevidens* and no or little expression in *L. subracemosa*.

LATE EMBRYOGENESIS ABUNDANT (LEA) proteins are predicted to have protective functions that are essential for desiccation tolerance (Hoekstra et al., 2001; Goyal et al., 2005; Hundertmark and Hincha, 2008). We identified 77 and 82 LEA protein-encoding genes in *L. brevidens* and *L. subracemosa*, respectively (Supplemental Table 11). Orthologs were assigned for 70 of these genes. About half of the identified LEAs were classified to the LEA_2 group, whereas the second largest LEA group was the LEA_4 group (14 and 12 genes in *L. brevidens* and *L. subracemosa*, respectively). Nine LEA genes in *L. brevidens* (one Dhn, five LEA_2, two LEA_4, and one LEA_5) were derived from gene duplication events, which suggests that these genes may have functions related to desiccation tolerance in *L. brevidens*. For example, *LEA5-2* occurs in one copy in *L. subracemosa* (*LsLEA5-2*) and has two orthologs in *L. brevidens* (i.e., *LbLEA5-2* and

LbLEA5-3). The *LsLEA5-2* gene showed negligible expression in *L. subracemosa* under control and dehydration conditions, whereas *LbLEA5-2* was among the highest expressed *LEA* genes upon dehydration in *L. brevidens* (Supplemental Data Set).

Almost one-third of *L. brevidens LEA* genes showed expression levels 30 times higher or more than *L. subracemosa* orthologs during late dehydration (10 and 14 d) (Supplemental Data Set). Most *L. subracemosa* orthologs had reduced or no expression in all surveyed time points, suggesting that there was a massive rewiring of expression networks (Supplemental Data Set). For example, *LEA1-3*, *LEA1-4*, *LEA2-14*, *LEA2-19*, *LEA4-1*, *LEA4-6*, *LEA4-7*, and *LEA5-2* showed very high expression in fully hydrated *L. brevidens* but no expression in fully hydrated *L. subracemosa*.

DISCUSSION

Genomic resources are abundant for resurrection plants, but the lack of suitable dehydration-sensitive outgroups has limited genomic insights into the origin and pathways controlling desiccation tolerance. Here, we leveraged a unique comparative system of closely related desiccation-tolerant and -sensitive species to identify gene- and pathway-level changes associated with the evolution of desiccation tolerance. This approach allowed us to distinguish dehydration pathways conserved in all plants from desiccation-specific processes observed only in resurrection plants.

L. brevidens and *L. subracemosa* have similar overall genome size and gene number, and most genes were likewise retained as singletons (1:1) or duplicates (2:2) after their shared WGD event. The genomes have no significant differences in architecture, rRNAs, repetitive element composition, or clustering of desiccation-related genes. These features were proposed previously to contribute to desiccation tolerance in other resurrection plant lineages (Xiao et al., 2015; Costa et al., 2017). Instead, our data indicate that desiccation tolerance in *L. brevidens* is driven by a complex cascade of *cis*-regulatory element-mediated pathway rewiring, tandem duplication, and preferential gene retention and neofunctionalization.

Gene expression patterns are dramatically divergent in dehydration and rehydration time-course data between *L. brevidens* and *L. subracemosa*. Only a few of the syntenic orthologs are similarly expressed in both species, and coexpression network modules are largely rewired. Early dehydration responses have surprisingly little overlap, suggesting that the gene expression program and signals leading to tolerance are already apparent upon mild dehydration. Gene expression is most dynamic between the day-3 and -7 time points, when the plants shift from moderate to severe dehydration stress. This likely reflects major shifts in leaf water potential, photosynthesis, oxidative stress, and cellular damage. Gene expression is stabilized after moderate drought and desiccation in *L. brevidens*, which reflects the successful deployment of protective mechanisms. By contrast, the dynamic and chaotic expression patterns in desiccating *L. subracemosa* may reflect last-ditch efforts to avoid imminent senescence.

Drought and seed development are linked by the common stress of water deficit. Vegetative and seed desiccation processes are strikingly similar, and overlapping pathways have been identified in resurrection plants (Costa et al., 2017; VanBuren et al., 2017). These include accumulation of osmoprotectants, expression of LEA proteins, and free radical scavenging systems as well as downregulation of photosynthesis and dismantling the photosynthetic apparatus. Drought responses and seed development are similarly regulated by ABA-related signaling, and both elicit comparable downstream responses (Nakashima and Yamaguchi-Shinozaki, 2013). Several important transcription factors involved in dehydration and seed-related processes are preferentially retained in *L. brevidens* compared with *L. subracemosa*, which may allow high-level pathway rewiring. Desiccation-related genes such as *ELIP*s or *LEA*s have increased in copy number in *L. brevidens* via tandem gene duplication. For genes encoding proteins with structural, enzymatic, or chaperone functions, tandem duplications may serve to increase their absolute abundance to surpass a threshold required for desiccation tolerance. The expansion of *ELIP*s has been observed in several resurrection plants, including *C. plantagineum* (Bartels et al., 1992), *S. lepidophylla* (VanBuren et al., 2018), and *Boea hygrometrica* (Xiao et al., 2015). ELIPs likely play an important role in protecting the photosynthetic apparatus and bind excess chlorophyll during prolonged desiccation (Alamillo and Bartels, 2001). The repeated duplication of *ELIP*s may be a hallmark of the convergent evolution of desiccation tolerance across land plants. Expression patterns also can be shifted by changes in *cis*-regulatory elements, as was observed previously in *LEA* genes from *C. plantagineum* and *L. brevidens* (van den Dries et al., 2011; Giarola et al., 2018). The enrichment of seed-related *cis*-regulatory elements in modules uniquely upregulated in *L. brevidens* is likely the result of novel *cis*-element acquisition in desiccation-related genes and activation of seed-related transcription factors.

Desiccation tolerance likely evolved from a complex, additive series of gene duplications and pathway rewiring rather than a simple master regulatory switch. Naturally drought-tolerant species could undergo favorable duplication of *ELIP*s or rewiring of LEA proteins to promote a quasi-desiccation-tolerant state. These responses could be refined further through the accumulation of additive features to surpass the threshold required for surviving anhydrobiosis. This step-wise hypothesis is supported by the continuum of desiccation tolerance, where the magnitude and duration of tolerance varies across species. *Craterostigma* can tolerate more rapid desiccation and recover more completely than *L. brevidens*, and older leaf tissue in *L. brevidens* is often desiccation sensitive.

The recovery rate in *L. brevidens* is related to environmental factors, including developmental stage, rate of drying, and dehydration priming. This comparatively weak desiccation tolerance may reflect relaxed selection in the drought-free rainforest habitat of *L. brevidens*. Desiccation tolerance is ancestral in the clade spanning *L. brevidens* and *Craterostigma* (Fischer et al., 2013), and some protective mechanisms were likely present in the shared ancestor with *L. subracemosa*. This also may explain the partial induction of seed and vegetative desiccation-associated pathways in *L. subracemosa* compared with the typical dehydration responses in other species. The trajectory from sensitive to

desiccation tolerant is a complex, multistep process, and future work in intermediate or weakly desiccation-tolerant species will help uncover the origins of this trait.

METHODS

Growth Conditions and Sampling

Lindernia brevidens and *Lindernia subracemosa* were grown as described previously (Phillips et al., 2008). Voucher specimens have been deposited: *L. brevidens* Kenya, Taita Hills, E. Fischer 8022 (KOBL;=Herbarium Koblenz) and *L. subracemosa* Rwanda, Uwinka, E. Fischer 1350 (BG-Bonn 19990-2, KOBL). Plants were propagated via cuttings and maintained under day/night temperatures of 22 and 18°C, respectively, under fluorescent lighting with an intensity of 80 μE m^{-2} s^{-1} and a 16/8-h photoperiod. *L. brevidens* and *L. subracemosa* were grown in the same chamber to minimize environmental variance. For the desiccation and rehydration time courses, plants were allowed to gradually dry for a period of 30 d, with sampling in triplicate with three independent plants at D3, D7, D14, D21, and D30. Plants were rehydrated and sampled at 24 and 48 h post rehydration. Samples were always taken at the same time of the day, 6 h after the onset of light, to minimize effects associated with circadian oscillation. Leaf tissue for RNAseq was flash frozen in liquid nitrogen and stored at −80°C. RWC measurements were calculated using the equation RWC = [(FW − DW)/(SW − DW)], where FW, DW, and SW indicate fresh weight of the leaf tissue, dry weight, and saturated weight. Dry weight was obtained after drying tissue at 80°C for 48 h, and saturated plant weights were obtained after submerging leaf tissue in water for 24 h. Three replicates of RWC measurements were collected for each time point.

Nucleic Acid Extraction, Library Construction, and Sequencing

High molecular weight genomic DNA for PacBio and Illumina sequencing was isolated from young leaf tissue of growth chamber-grown *L. brevidens* and *L. subracemosa*. DNA was isolated using a modified nucleus preparation (Zhang et al., 1995) followed by phenol chloroform purification to remove residual contaminants. PacBio libraries were constructed and size selected for 25-kb fragments on the BluePippen system (Sage Science) followed by purification using AMPure XP beads (Beckman Coulter). Libraries were sequenced on a Sequel platform with V4 software and V2 chemistry. In total, 2,054,566 filtered subreads spanning 21.7 Gb were sequenced for *L. brevidens* and 1,615,065 reads spanning 17.9 Gb were sequenced for *L. subracemosa*. This represents 80.3× and 71.6× coverage for *L. brevidens* and *L. subracemosa*, respectively. Illumina DNAseq libraries for error correction were constructed using the KAPA HyperPrep Kit (Kapa Biosystems) following the manufacturer's instructions. Libraries were sequenced on an Illumina HiSeq4000 device under paired-end 150-bp mode.

For RNAseq analysis, total RNA was extracted from 200 mg of ground *L. brevidens* and *L. subracemosa* leaf tissues using an Omega Bio-tek E.Z.N.A. Plant RNA Kit (Omega Bio-tek) according to the manufacturer's instructions. RNA quality was validated using gel electrophoresis and Qubit RNA IQ Assay (ThermoFisher). Two micrograms of total RNA was used to construct Illumina TruSeq-stranded mRNA libraries following the manufacturer's protocol (Illumina). Libraries were pooled and sequenced on the Illumina HiSeq4000 device under paired-end 150-bp mode. Three replicatfinnes were sequenced for each time point in each species.

Genome Assembly

Genome sizes for *L. brevidens* and *L. subracemosa* were estimated using flow cytometry as described previously (Arumuganathan and Earle, 1991). The flow cytometry-based estimates of 270 and 250 Mb for *L. brevidens* and *L. subracemosa*, respectively, were consistent with k-mer-based analysis using Illumina WGS data. *Lindernia* are mostly selfing, and unimodal k-mer distribution suggests low within-genome heterozygosity for both species. Raw PacBio reads were error corrected and assembled using Canu (V1.4; Koren et al., 2017). Based on previous experience (Edger et al., 2018; VanBuren et al., 2018), Canu produced the most contiguous and accurate assembly for homozygous, diploid species compared with other leading long-read assemblers. The following Canu parameters were modified and all others were left as default: minReadLength = 2500, GenomeSize = 270 Mb (or 250 Mb), minOverlapLength = 1000. Assembly graphs were visualized in Bandage (Wick et al., 2015). Draft contigs were polished with Pilon (V1.22; Walker et al., 2014) using 79× and 58× coverage of Illumina paired-end 150-bp data for *L. brevidens* and *L. subracemosa*, respectively. Illumina reads were quality trimmed using Trimmomatic (V0.33; Bolger et al., 2014) and aligned to the draft contigs using bowtie2 (V2.3.0; Langmead and Salzberg, 2012) with default parameters. Alignment rates in the first round of corrections were 96 and 97%, respectively, suggesting that both the *L. brevidens* and *L. subracemosa* assemblies were largely complete. Parameters for Pilon were modified as follows: –flank 7, –K 49, and –mindepth 20. Pilon was run recursively three times with minimal corrections in the third round, supporting accurate insertion/deletion correction.

Hi-C Library Construction and Analysis

The *L. brevidens* draft genome was anchored into a chromosome-scale assembly using a Hi-C proximity-based assembly approach. The Hi-C library was constructed using 0.2 g of young leaf tissue from well-watered *L. brevidens* plants with the Proximo Hi-C Plant kit (Phase Genomics) following the manufacturer's protocol. The final library was size selected for 300 to 600 bp and sequenced on the Illumina NexSeq 500 device under paired-end 75-bp mode. In total, 178 million reads were used as input for the Juicer and 3d-DNA Hi-C analysis and scaffolding pipelines (Supplemental Table 1; Durand et al., 2016; Dudchenko et al., 2017). Quality-filtered reads were aligned to the PacBio contigs using bwa (V0.7.16; Li, 2013) with strict parameters (-n 0) to prevent mismatches and nonspecific alignments, and the resulting sam files were used as input into the Juicer pipeline. Read pairs were merged and duplicates or near duplicates were removed prior to constructing the distance matrix. Contigs were ordered and oriented and assembly errors were identified using the 3d-DNA pipeline with default parameters (Dudchenko et al., 2017). The resulting hic contact matrix was visualized using Juicebox, and misassemblies and misjoins were manually corrected based on neighboring interactions. This approach identified 14 high-confidence clusters representing the haploid chromosome number in *L. brevidens*. The manually validated assembly was used to build pseudomolecules using the finalize-output.sh script from 3d-DNA, and chromosomes were renamed and ordered by size.

Genome Annotation

Prior to genome annotation, LTR-RTs were predicted using LTR harvest (genome tools V1.5.8; Ellinghaus et al., 2008) and LTR finder (V1.07; Xu and Wang, 2007), and the LTR library was refined using LTR retriever (V1.8.0; Ou and Jiang, 2018). RTs were classified as intact if they were flanked by full-length LTRs. The insertion time for each intact element was calculated using LTR retriever with the formula $T = K/2\mu$, where K is the divergence rate approximated by percentage identity and μ is the neutral mutation rate estimated as $\mu = 1 \times 10^{-8}$ mutations per bp per year.

The filtered, nonredundant LTR library from LTR retriever was used as input for whole-genome annotation of LTR-RTs using RepeatMasker (http://www.repeatmasker.org/; Chen, 2004). The *L. brevidens* and *L. subracemosa* genomes were annotated using the MAKER-P pipeline

(Campbell et al., 2014). Transcript-based evidence for gene predictions was produced using the desiccation/rehydration time-course RNAseq data. RNAseq reads were aligned to the *L. brevidens* and *L. subracemosa* genomes using the splice aware aligner STAR (V2.6; Dobin et al., 2013). Transcripts were identified using StringTie (V1.3.4; Pertea et al., 2015) with default parameters, and –merge flag was used to combine the output from individual libraries. The sets of nonredundant transcripts were used as EST evidence, and protein sequences from Arabidopsis (*Arabidopsis thaliana*; Lamesch et al., 2012) and UniprotKB plant databases (Boutet et al., 2007) were used as protein evidence. The custom LTR-RT library produced by LTR retriever and Repbase libraries were used for repeat masking. Ab initio gene prediction was done using SNAP (Korf, 2004) and Augustus (3.0.2; Stanke and Waack, 2003), with two rounds of iterative training for each species. The raw gene models were filtered to identify any residual repetitive elements using BLAST with a nonredundant transposase library. After filtering, a final set of 27,204 and 33,344 gene models was produced for *L. brevidens* and *L. subracemosa*, respectively. Annotation quality was assessed using BUSCO (V.2; Simão et al., 2015) with the plant-specific data set (embryophyta_odb9).

Comparative Genomics

Syntenic gene pairs within and between *L. brevidens* and *L. subracemosa* were identified using the MCScan toolkit (V1.1; Wang et al., 2012) implemented in python [https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version)]. Gene models were aligned using LAST, and hits were filtered to find syntenic blocks. Tandem gene duplicates were identified using all-versus-all BLAST with a minimum e-value of 1e-5 and maximum gene distance of 10 genes. Macrosynteny and microsynteny plots and syntenic block depths were plotted using the python version of MCScan. Genes were classified as lineage specific if they had no syntenic orthologs between the two species or hits from LASTAL with greater than 70% nucleotide identity. The WGD event within *Lindernia* was identified using a combination of synteny and synonymous substitution rate estimation between duplicated gene pairs. Comparison of *L. brevidens* and *L. subracemosa* in MCScan identified a 2:2 syntenic pattern with 7742 and 8452 duplicated gene pairs retained, respectively. Duplicated regions span 70% of the *L. brevidens* genome and 72% of the *L. subracemosa* genome. Most duplicated regions were retained in large blocks, allowing chromosome pairs to be identified. Substitution rate was estimated using KaKs_calculator with the NG model (Zhang et al., 2006), and peaks of 0.65 in *L. brevidens* and 0.69 for *L. subracemosa* were identified, which indicates that the event is shared between both species.

RNAseq Analysis

Paired-end Illumina RNAseq reads were trimmed by quality score and by adapter contamination using Trimmomatic (V0.33; Bolger et al., 2014) with default parameters. The expression level of each gene was quantified using the pseudoaligner Kallisto (Bray et al., 2016) against the final gene models for *L. brevidens* and *L. subracemosa*. Parameters were left as default with 100 bootstraps per sample. Expression was quantified in TPM, and a mean across the three replicates was used for single-gene analysis and to construct $\log_2$ transformed expression-based heat maps. Pairwise differentially expressed genes were identified using sleuth (Pimentel et al., 2017) implemented in R.

Coexpression Network Construction

The time-course RNAseq data were clustered into gene coexpression networks using the R package WGCNA (Langfelder and Horvath, 2008). Genes with less than an average TPM of 5 across all seven time points were filtered prior to network construction. A signed coexpression network was constructed for each species using a soft-thresholding power of 8 and a tree cut height of 0.15. All remaining parameters were left as default. In total, 14,246 genes were clustered into 10 modules for *L. brevidens* and 14,075 genes were clustered into 9 modules for *L. subracemosa*.

Cis-Element Identification

Cis-regulatory elements were identified using the Hypergeometric Optimization of Motif EnRichment program (V4.10; Heinz et al., 2010) using *cis*-elements from 529 plant transcription factors (O'Malley et al., 2016). *Cis*-elements were identified in the 1-kb region upstream of the transcriptional start site when known or directly upstream of the start codon of each gene model. Promoters of gene models with detectable expression (TPM > 1) were used as background. Enrichment tests were performed using syntenic gene pairs with differential expression specific to *L. brevidens* or gene models unique to modules upregulated or downregulated during desiccation in *L. brevidens*. $P < 0.00001$ was used as the cutoff for identifying enriched motifs in any comparison.

Identification of *LEA* Genes

LEA genes were retrieved from *L. brevidens* and *L. subracemosa* transcriptomes by BLAST and HMMER (http://hmmer.org/). Arabidopsis and *Craterostigma plantagineum* LEA protein sequences were used for BLAST searches. HMM profiles for the eight LEA families (DHN, PF00257; LEA_1, PF03760; LEA_2, PF03168; LEA_3, PF03242; LEA_4, PF02987; LEA_5, PF00477; LEA_6, PF10714; and SMP, PF04927) obtained from the Pfam database (http://pfam.xfam.org; Finn et al., 2016) were used with the program hmmscan to search for LEA domain-containing proteins. Proteins identified with hmmscan were queried against the nr databank to confirm their classification as LEAs. Ortholog pairs were additionally confirmed by pairwise sequence alignments of the predicted protein sequences using EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/).

Accession Numbers

The genome assemblies, raw PacBio data, Illumina DNAseq, and RNAseq data are available from the National Center for Biotechnology Information Short Read Archive. The RNAseq reads were deposited to the National Center for Biotechnology Information Short Read Archive under BioProject PRJNA488068. The genome assemblies for *L. brevidens* and *L. subracemosa* were deposited under BioProjects PRJNA489464 and PRJNA489465, respectively.

Supplemental Data

Supplemental Figure 1. Histogram of filtered PacBio subreads for *L. brevidens* and *L. subracemosa*.

Supplemental Figure 2. Graph-based assembly of the *L. brevidens* genome.

Supplemental Figure 3. Graph-based assembly of the *L. subracemosa* genome.

Supplemental Figure 4. Summary of genome-wide syntenic blocks in *lindernia*.

Supplemental Figure 5. Macrosyntenic dot plot between the *L. brevidens* and *L. subracemosa* genomes.

Supplemental Figure 6. Microsynteny showing a region with biased fractionation between *L. brevidens* and *L. subracemosa*.

Supplemental Figure 7. Weighted gene coexpression network in *L. brevidens*.

Supplemental Figure 8. Weighted gene coexpression network in *L. subracemosa*.

Supplemental Figure 9. Overlap between *L. brevidens* and *L. subracemosa* coexpression networks.

Supplemental Table 1. Statistics of read mapping, filtering, and interactions for the Hi-C data.

Supplemental Table 2. Summary of Hi-C-based scaffolding.

Supplemental Table 3. Summary of GO terms enriched in *L. brevidens*-specific genes.

Supplemental Table 4. Summary of differential expressed genes during dehydration and rehydration in the two *Lindernia* species.

Supplemental Table 5. Enriched GO terms in syntenic orthologs uniquely upregulated in *L. brevidens* with no change in expression in *L. subracemosa*.

Supplemental Table 6. Enriched GO terms in syntenic orthologs uniquely downregulated in *L. brevidens* with no change in expression in *L. subracemosa*.

Supplemental Table 7. Enriched *cis*-regulatory elements in genes uniquely upregulated under desiccation in *L. brevidens*.

Supplemental Table 8. Enriched *cis*-regulatory elements in genes uniquely downregulated under desiccation in *L. brevidens*.

Supplemental Table 9. Enriched *cis*-regulatory elements in desiccation-associated coexpression modules in *L. brevidens*.

Supplemental Table 10. Expression of desiccation-related genes.

Supplemental Table 11. Number of LEA genes in *L. brevidens* and *L. subracemosa*.

Supplemental Data Set. Expression of *LEA* genes.

ACKNOWLEDGMENTS

We thank Eberhard Fischer (University Koblenz Germany) for making *Lindernia* plants available originally. This work is supported by funding from the National Science Foundation (MCB-1817347 to R.V.).

AUTHOR CONTRIBUTIONS

R.V. and D.B. designed and conceived research. R.V., C.M.W., and J.P. annotated genome features. X.S. collected desiccation and rehydration data. C.M.W. constructed RNAseq, DNAseq, and Hi-C libraries. R.V., C.M.W., J.P., X.S., S.A., V.G., and D.B. analyzed data. R.V. wrote the article. All authors read and approved the final article.

Received July 6, 2018; revised October 2, 2018; accepted October 23, 2018; published October 25, 2018.

REFERENCES

Alamillo, J.M., and Bartels, D. (2001). Effects of desiccation on photosynthesis pigments and the ELIP-like dsp 22 protein complexes in the resurrection plant Craterostigma plantagineum. Plant Sci. **160:** 1161–1170. 11337073

Alonso, R., Oñate-Sánchez, L., Weltmeier, F., Ehlert, A., Diaz, I., Dietrich, K., Vicente-Carbajosa, J., and Dröge-Laser, W. (2009). A pivotal role of the basic leucine zipper transcription factor bZIP53 in the regulation of Arabidopsis seed maturation gene expression based on heterodimerization and protein complex formation. Plant Cell **21:** 1747–1761.

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

Baniaga, A.E., Arrigo, N., and Barker, M.S. (2016). The small nuclear genomes of Selaginella are associated with a low rate of genome size evolution. Genome Biol. Evol. **8:** 1516–1525.

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A., and Ashton, N.W. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science**332:** 960–963.

Bartels, D., and Salamini, F. (2001). Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*: A contribution to the study of drought tolerance at the molecular level. Plant Physiol. **127:** 1346–1353.

Bartels, D., Hanke, C., Schneider, K., Michel, D., and Salamini, F. (1992). A desiccation-related Elip-like gene from the resurrection plant Craterostigma plantagineum is regulated by light and ABA. EMBO J. **11:** 2771–2778.

Bianchi, G., Gamba, A., Murelli, C., Salamini, F., and Bartels, D. (1991). Novel carbohydrate metabolism in the resurrection plant Craterostigma plantagineum. Plant J. **1:** 355–359.

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

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., and Bairoch, A. (2007). Uniprotkb/swiss-prot. In Plant Bioinformatics. Methods in Molecular Biology, D Edwards, ed (New York City, NY: Humana Press), pp. 89–112.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. **34:** 525–527.

Campbell, M.S., Law, M., Holt, C., Stein, J.C., Moghe, G.D., Hufnagel, D.E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C.J., Ware, D., and Shiu, S.H., et al. (2014). MAKER-P: A tool kit for the rapid creation, management, and quality control of plant genome annotations. Plant Physiol. **164:** 513–524.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biol. **4:** 10.

Chen, N. (2004). Using Repeat Masker to identify repetitive elements in genomic sequences. Current Protocols in Bioinformatics **5:** 4.10.1–4.10.14.

Costa, M.D., Artur, M.A., Maia, J., Jonkheer, E., Derks, M.F., Nijveen, H., Williams, B., Mundree, S.G., Jiménez-Gómez, J.M., Hesselink, T., Schijlen, E.G., and Ligterink, W., et al. (2017). A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*. Nat. Plants **3:** 17038.

Dekkers, B.J., He, H., Hanson, J., Willems, L.A., Jamar, D.C., Cueff, G., Rajjou, L., Hilhorst, H.W., and Bentsink, L. (2016). The Arabidopsis DELAY OF GERMINATION 1 gene affects ABSCISIC ACID INSENSITIVE 5 (ABI5) expression and genetically interacts with ABI3 during Arabidopsis seed development. Plant J. **85:** 451–465.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics **29:** 15–21.

Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden,

A.P., and Aiden, E.L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science **356:** 92–95.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. **3:** 95–98.

Edger, P.P., VanBuren, R., Colle, M., Poorten, T.J., Wai, C.M., Niederhuth, C.E., Alger, E.I., Ou, S., Acharya, C.B., and Wang, J. (2018). Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. Gigascience **7:** 1–7.

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

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., and Tate, J., et al. (2016). The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res. **44:** D279–D285.

Fischer, E., Schäferhoff, B., and Müller, K. (2013). The phylogeny of Linderniaceae: The new genus Linderniella, and new combinations within Bonnaya, Craterostigma, Lindernia, Micranthemum, Torenia and Vandellia. Willdenowia **43:** 209–238.

Giarola, V., Jung, N.U., Singh, A., Satpathy, P., and Bartels, D. (2018). Analysis of pcC13-62 promoters predicts a link between cis-element variations and desiccation tolerance in Linderniaceae. J. Exp. Bot. **69:** 3773–3784.

Goyal, K., Walton, L.J., and Tunnacliffe, A. (2005). LEA proteins prevent protein aggregation due to water stress. Biochem. J. **388:** 151–157.

Hayami, N., Sakai, Y., Kimura, M., Saito, T., Tokizawa, M., Iuchi, S., Kurihara, Y., Matsui, M., Nomoto, M., Tada, Y., and Yamamoto, Y.Y. (2015). The responses of Arabidopsis early light-induced protein2 to ultraviolet B, high light, and cold stress are regulated by a transcriptional regulatory unit composed of two elements. Plant Physiol. **169:** 840–855.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell **38:** 576–589.

Hoekstra, F.A., Golovina, E.A., and Buitink, J. (2001). Mechanisms of plant desiccation tolerance. Trends Plant Sci. **6:** 431–438.

Huang, Y.-C., Niu, C.-Y., Yang, C.-R., and Jinn, T.-L. (2016). The heat stress factor HSFA6b connects ABA signaling and ABA-mediated heat responses. Plant Physiol. **172:** 1182–1199.

Hundertmark, M., and Hincha, D.K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. BMC Genomics **9:** 118.

Illing, N., Denby, K.J., Collett, H., Shen, A., and Farrant, J.M. (2005). The signature of seeds in resurrection plants: A molecular and physiological comparison of desiccation tolerance in seeds and vegetative tissues. Integr. Comp. Biol. **45:** 771–787.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: Scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. **27:** 722–736.

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

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., and Lee, C.H., et al. (2012). The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. Nucleic Acids Res. **40:** D1202–D1210.

Langfelder, P., and Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics **9:** 559.

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

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997.

Lopez-Molina, L., Mongrand, S., and Chua, N.-H. (2001). A post-germination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. Proc. Natl. Acad. Sci. USA **98:** 4782–4787.

Lüttge, U., Beck, E., and Bartels, D. (2011). Plant Desiccation Tolerance. (Berlin, Germany: Springer Science & Business Media).

Nakashima, K., and Yamaguchi-Shinozaki, K. (2013). ABA signaling in stress-response and seed development. Plant Cell Rep. **32:** 959–970.

Nakashima, K., Fujita, Y., Kanamori, N., Katagiri, T., Umezawa, T., Kidokoro, S., Maruyama, K., Yoshida, T., Ishiyama, K., Kobayashi, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2009). Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. Plant Cell Physiol. **50:** 1345–1363.

Nguyen, T.-P., Cueff, G., Hegedus, D.D., Rajjou, L., and Bentsink, L. (2015). A role for seed storage proteins in Arabidopsis seed longevity. J. Exp. Bot. **66:** 6399–6413.

O'Malley, R.C., Huang, S.-C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. Cell **165:** 1280–1292.

Oliver, M.J., Tuba, Z., and Mishler, B.D. (2000). The evolution of vegetative desiccation tolerance in land plants. Plant Ecol. **151:** 85–100.

Oliver, M.J., Guo, L., Alexander, D.C., Ryals, J.A., Wone, B.W., and Cushman, J.C. (2011). A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. Plant Cell **23:** 1231–1248.

Ou, S., and Jiang, N. (2018). LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. Plant Physiol. **176:** 1410–1422.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. **33:** 290–295.

Phillips, J.R., Fischer, E., Baron, M., van den Dries, N., Facchinelli, F., Kutzer, M., Rahmanzadeh, R., Remus, D., and Bartels, D. (2008). *Lindernia brevidens*: A novel desiccation-tolerant vascular plant, endemic to ancient tropical rainforests. Plant J. **54:** 938–948.

Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods **14:** 687–690.

Plancot, B., Vanier, G., Maire, F., Bardor, M., Lerouge, P., Farrant, J.M., Moore, J., Driouich, A., Vicré-Gibouin, M., Afonso, C., and Loutelier-Bourhis, C. (2014). Structural characterization of arabinoxylans from two African plant species *Eragrostis nindensis* and *Eragrostis tef* using various mass spectrometric methods. Rapid Commun. Mass Spectrom. **28:** 908–916.

Poxleitner, M., Rogers, S.W., Lacey Samuels, A., Browse, J., and Rogers, J.C. (2006). A role for caleosin in degradation of oil-body storage lipid during seed germination. Plant J. **47:** 917–933.

Proctor, M. (1990). The physiological basis of bryophyte production. Bot. J. Linn. Soc. **104:** 61–77.

Rahmanzadeh, R., Müller, K., Fischer, E., Bartels, D., and Borsch, T. (2005). The Linderniaceae and Gratiolaceae are further lineages distinct from the Scrophulariaceae (Lamiales). Plant Biol (Stuttg) **7:** 67–78.

Sakuraba, Y., Kim, D., Kim, Y.-S., Hörtensteiner, S., and Paek, N.-C. (2014a). Arabidopsis STAYGREEN-LIKE (SGRL) promotes abiotic stress-induced leaf yellowing during vegetative growth. FEBS Lett. **588:** 3830–3837.

Sakuraba, Y., Lee, S.-H., Kim, Y.-S., Park, O.K., Hörtensteiner, S., and Paek, N.-C. (2014b). Delayed degradation of chlorophylls and photosynthetic proteins in Arabidopsis autophagy mutants during stress-induced leaf yellowing. J. Exp. Bot. **65:** 3915–3925.

Shen, Y., Xie, J., Liu, R.D., Ni, X.F., Wang, X.H., Li, Z.X., and Zhang, M. (2014). Genomic analysis and expression investigation of caleosin gene family in Arabidopsis. Biochem. Biophys. Res. Commun. **448:** 365–371.

Shimada, T.L., Shimada, T., Takahashi, H., Fukao, Y., and Hara-Nishimura, I. (2008). A novel role for oleosins in freezing tolerance of oilseeds in Arabidopsis thaliana. Plant J. **55:** 798–809.

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

Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics **19** (suppl. 2): ii215–ii225.

VanBuren, R., Bryant, D., Edger, P.P., Tang, H., Burgess, D., Challabathula, D., Spittle, K., Hall, R., Gu, J., Lyons, E., Freeling, M., and Bartels, D., et al. (2015). Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. Nature **527:** 508–511.

VanBuren, R., Wai, C.M., Zhang, Q., Song, X., Edger, P.P., Bryant, D., Michael, T.P., Mockler, T.C., and Bartels, D. (2017). Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass *Oropetium thomaeum*. Plant Cell Environ. **40:** 2292–2306.

VanBuren, R., Wai, C.M., Ou, S., Pardo, J., Bryant, D., Jiang, N., Mockler, T.C., Edger, P., and Michael, T.P. (2018). Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. Nat. Commun. **9:** 13.

van den Dries, N., Facchinelli, F., Giarola, V., Phillips, J.R., and Bartels, D. (2011). Comparative analysis of LEA-like 11-24 gene expression and regulation in related plant species within the Linderniaceae that differ in desiccation tolerance. New Phytol. **190:** 75–88.

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

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE **9:** e112963.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., Kissinger, J.C., and Paterson, A.H. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. **40:** e49.

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

Xiao, L., Yang, G., Zhang, L., Yang, X., Zhao, S., Ji, Z., Zhou, Q., Hu, M., Wang, Y., Chen, M., Xu, Y., and Jin, H., et al. (2015). The resurrection genome of Boea hygrometrica: A blueprint for survival of dehydration. Proc. Natl. Acad. Sci. USA **112:** 5833–5837.

Xu, Z., and Wang, H. (2007). LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. **35:** W265–W268.

Yobi, A., Wone, B.W., Xu, W., Alexander, D.C., Guo, L., Ryals, J.A., Oliver, M.J., and Cushman, J.C. (2013). Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. Mol. Plant **6:** 369–385.

Yoshida, T., Fujita, Y., Maruyama, K., Mogami, J., Todaka, D., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2015). Four Arabidopsis AREB/ABF transcription factors function predominantly in gene expression downstream of SnRK2 kinases in abscisic acid signalling in response to osmotic stress. Plant Cell Environ. **38:** 35–49.

Zhang, Q., and Bartels, D. (2018). Molecular responses to dehydration and desiccation in desiccation-tolerant angiosperm plants. J. Exp. Bot. **69:** 3211–3222.

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

Zhang, Q., Linnemann, T.V., Schreiber, L., and Bartels, D. (2016). The role of transketolase and octulose in the resurrection plant *Craterostigma plantagineum*. J. Exp. Bot. **67:** 3551–3559.

Zhang, Z., Li, J., Zhao, X.-Q., Wang, J., Wong, G.K.-S., and Yu, J. (2006). KaKs_Calculator: Calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics **4:** 259–263.