Received: 23 October 2023 | Revised: 5 September 2024 | Accepted: 12 September 2024

DOI: 10.1111/mec.17536

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Evolution of gene networks underlying adaptation to drought stress in the wild tomato *Solanum chilense*

Kai Wei[1] | Saida Sharifova[2] | Xiaoyun Zhao[1] | Neelima Sinha[3] | Hokuto Nakayama[4] | Aurélien Tellier[1] | Gustavo A. Silva-Arias[1,5]

[1]Department of Life Science Systems, School of Life Sciences, Technical University of Munich, Freising, Germany

[2]Department of Life Sciences, Graduate School of Science, Arts and Technology, Khazar University, Baku, Azerbaijan

[3]Department of Plant Biology, University of California Davis, Davis, California, USA

[4]Department of Biological Sciences, The University of Tokyo, Tokyo, Japan

[5]Instituto de Ciencias Naturales, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia

Correspondence
Aurélien Tellier, Department of Life Science Systems, School of Life Sciences, Technical University of Munich, Liesel-Beckmann Strasse 2, 85354 Freising, Germany.
Email: aurelien.tellier@tum.de

Gustavo A. Silva-Arias, Instituto de Ciencias Naturales, Facultad de Ciencias, Universidad Nacional de Colombia, Sede Bogotá, Av. Carrera 30 # 45-03, 111321 Bogotá, Colombia.
Email: gasilvaa@unal.edu.co

Funding information
Fulbright Visiting Scholar Program sponsored by the U.S. Department of State; Deutsche Forschungsgemeinschaft, Grant/Award Number: 317616126 (TE809/7-1); Chinese Scholarship Council; Division of Integrative Organismal Systems, Grant/Award Number: IOS-1238243, IOS-1856749 and IOS-211980; Deutscher Akademischer Austauschdienst; JSPS KAKENHI, Grant/Award Number: JP19H05672, JP19K23742, JP20K06682 and JP20KK0340; Technische Universität München

Handling Editor: Alison Gonçalves Nazareno

Abstract

Drought stress is a key limitation for plant growth and colonization of arid habitats. We study the evolution of gene expression response to drought stress in a wild tomato, *Solanum chilense*, naturally occurring in dry habitats in South America. We conduct a transcriptome analysis under standard and drought experimental conditions to identify drought-responsive gene networks and estimate the age of the involved genes. We identify two main regulatory networks corresponding to two typical drought-responsive strategies: cell cycle and fundamental metabolic processes. The metabolic network exhibits a more recent evolutionary origin and a more variable transcriptome response than the cell cycle network (with ancestral origin and higher conservation of the transcriptional response). We also integrate population genomics analyses to reveal positive selection signals acting at the genes of both networks, revealing that genes exhibiting selective sweeps of older age also exhibit greater connectivity in the networks. These findings suggest that adaptive changes first occur at core genes of drought response networks, driving significant network re-wiring, which likely underpins species divergence and further spread into drier habitats. Combining transcriptomics and population genomics approaches, we decipher the timing of gene network evolution for drought stress response in arid habitats.

KEYWORDS
Atacama desert, drought adaptation, gene network evolution, Selective sweep, transcriptome analysis, wild tomato

Kai Wei and Saida Sharifova contributed equally to this work.

1 | INTRODUCTION

Drought stress is a major environmental constraint that negatively influences plant development and prevents plant growth, resulting in decreased yield in agriculture and limiting the colonization of arid or hyper-arid habitats (Ciais et al., 2005; Juenger, 2013). Plants respond to water insufficiency through multiple strategies underpinned by various physiological and developmental processes, such as storage of internal water to avoid tissue damage, increasing the metabolic activity of some tissues (Rodrigues et al., 2019) and tolerance to sustain growth processes (Basu et al., 2016). Drought response involves numerous quantitative and polygenic traits encoded in (complex) gene regulatory networks. To improve crops and predict the evolutionary responses of plant species under the current and predicted global water deficits, it is thus of interest to pinpoint and decipher the evolutionary history of the relevant gene networks underpinning the adaptation of wild plants to arid or hyper-arid habitats (Gehan et al., 2015).

Comparative transcriptomic analyses involving the inference of gene co-expression patterns demonstrate that many gene co-expression networks (GCNs) are conserved through the tree of life (Crow et al., 2022; Gerstein et al., 2014; Stuart et al., 2003; Zarrineh et al., 2014). Physiological, structural and regulatory innovations to cope with drought stress have arisen throughout the history of plants, many of them even predating the emergence of land plants (Bowles et al., 2021; De Vries & Archibald, 2018; Jill Harrison, 2017; Mustafin et al., 2019; Wang et al., 2020). Several conserved GCNs underlie fundamental biological processes such as protein metabolism, cell cycle, photosynthesis or wood formation (Ficklin & Feltus, 2011; Stuart et al., 2003; Zinkgraf et al., 2020).

A key question in functional and evolutionary genomics is thus to link GCN long-term evolution with the (relatively) short-scale processes of population adaptation and lineage divergence in order to assess the relative importance of contingency, exaptation and evolution of novel genes (duplication and neofunctionalization) in the colonization of novel habitats. Two main hypotheses are formulated. First, highly conserved sub-networks (so-called hubs or kernels) evolve under strong purifying selection to ensure the functionality of the GCNs (Josephs et al., 2017; Mähler et al., 2017; Masalia et al., 2017; Papakostas et al., 2014). Thus, genetic variation is only found in (less connected) genes at the periphery of the GCNs, these genes being the targets of positive selection (Erwin, 2020; Flowers et al., 2007; Kim et al., 2007; Luisi et al., 2015). However, this argument is likely to apply when the novel habitats may not differ much from the original one (in terms of climate, soil, etc.) so that only minor adjustments in the GCNs are enough to provide adaptation. This also aligns with the so-called developmental systems drift hypothesis (DSD; True & Haag, 2001). DSD predicts that GCN re-wiring would only occur at 'flexible' (sub-)modules as the accumulation of neutral variation would keep the network function intact until a new unique viable function (phenotype or developmental pathway) appears. Second, despite the general belief that genes with higher connectivity evolve at a slower rate, there is also evidence that changes in central genes (with higher connectivity) can be responsible for the short-term adaptive response (Jovelin & Phillips, 2009; Luisi et al., 2015) and promote re-wiring of the GCN (Koubkova-Yu et al., 2018). Thus, highly connected genes may be targets of positive selection during environmental change, even though these genes experience purifying selection in stable environments (Hämälä et al., 2020). In this latter case, we expect a correlation between the age of the positive selection events and the connectivity of a gene in a GCN network, but the absence of such a correlation under the first hypothesis.

To test these two hypotheses, we quantify the selective forces (positive vs. purifying selection) acting on different components of the networks (hub vs. peripheral genes) across wild tomato species adapted to contrasting drought conditions. Wild tomatoes, and *S. chilense* in particular, are good models of interest as their diversification correlates with exploring wide environmental gradients along the Pacific coast of South America (Haak et al., 2014; Nakazato et al., 2010). Wild tomato species and relatives such as *S. chilense*, *S. pennellii* and *S. sitiens* are well-established systems to study tolerance strategies to survive in extreme environments (Barrera-Ayala et al., 2023; Blanchard-Gros et al., 2021; Bolger et al., 2014; Kashyap, Prasanna, et al., 2020; Martínez et al., 2014; Molitor et al., 2024; Tapia et al., 2016). Since the divergence (with possible recurrent gene flow) from its sister species *S. peruvianum* (found at low altitudes in Peru) approximately 500 Kya (Städler et al., 2008), *S. chilense* has colonized and adapted to very-dry-to-arid environments of Southern Peru and Northern Chile (up to the margins of the Atacama desert) (Böndel et al., 2015). Populations of *S. chilense* are challenged by prolonged drought periods (including in the putative region of species origin, near to population LA1963), with a gradient of the most severe drought conditions increasing towards occurring in the southern part of the species range (Wei et al., 2023). In a previous study, we assayed for footprints of positive selection in 30 fully sequenced genomes of *S. chilense* to identify candidate genes underpinning adaptation across the species range. We found genes with putative functions related to root hair development and cell homeostasis likely involved in drought stress tolerance (Wei et al., 2023). Although previous empirical studies have focused on a few abiotic stress response genes (Böndel et al., 2015, 2018; Fischer et al., 2011, 2013; Mboup et al., 2012; Nosenko et al., 2016), we still lack a comprehensive understanding of the genomic and gene pathways (networks) underpinning adaptation to drought in *S. chilense*.

Therefore, we identify drought stress-responsive gene co-expression networks by combining multiple analyses of transcriptome data of *S. chilense* and focusing on two networks involved in cell cycle and metabolic processes. Furthermore, we infer the evolutionary processes at these two networks across three different evolutionary levels (tree of life/plants, species and populations) by calculating transcriptome indices, which indicate the evolutionary age and sequence divergence of the genes involved in drought response. We then analyse the timing of the emergence of adaptive variation in the identified drought-responsive genes of these networks and the correlation to gene connectivity.

2 | MATERIALS AND METHODS

2.1 | Plant material and drought stress experiment

Seeds of *Solanum chilense* accession no. LA1963 were acquired from the Tomato Genetics Resource Center (TGRC), University of California at Davis. Seeds were soaked in 50% household bleach (2.7% sodium hypochlorite) for 30 minutes and rinsed thoroughly with water according to instructions provided by TGRC. The rinsed seeds were sown into pots containing sterilized soil with perlite and sand (1:2) and grown under standard controlled conditions, showing optimal growth in previous studies (22C day/20C night, 16 h light/8 h dark photoperiod; Böndel, 2014; Gao et al., 2015; Kashyap, et al. 2020a; Blanchard-Gros et al., 2021; Kahlon et al., 2023). On the 24th day after sowing, 88 plants were randomly distributed into two groups and watered with a sufficient volume to reach the bottom of containers (30–40 mL). The first group of plants was maintained under normal watering conditions and watered with a sufficient volume of water (50–55 mL) on 4, 7 and 11 days after the start of the experiment. A moderate water stress regime was imposed on the second group of plants by stopping irrigation for 7 days, followed by re-watering with 25 mL of water. On day 12, newly expanded leaf (1–1.5 cm length) and shoot apices with immediately surrounding leaf primordia (shoot apices and P1-P5 leaf primordia) from each group were dissected carefully using razor blades and immediately grounded into fine powder in liquid nitrogen for RNA extraction. Four groups of biological replicates (viz. control_leaf, drought_leaf, control_shoot apex and drought_shoot apex) were used for all RNA-Seq experiments from each tissue type. Each leaf and shoot apex sample replicates included the pooled tissues from five and six plants respectively.

2.2 | RNA extraction and cDNA library construction

Libraries were constructed and named as follows: leaves under control (optimal watering) condition (CL-A to D), shoot apices under control condition (CSA-E to H), leaves under drought condition (DL-I to L) and shoot apices under drought condition (DSA-M to P). Tissues were lysed using zircon beads in a lysate-binding buffer containing sodium dodecyl sulphate. mRNA was isolated from 200 µL of lysate per sample with streptavidin-coated magnetic beads for indexed non-strand-specific RNA-Seq library preparation according to the method described by Kumar et al. (2012): 1 µL of 12.5 µM of 5-prime biotinylated polyT oligonucleotide and streptavidin-coated magnetic beads were used to capture mRNA and isolate captured mRNAs from the lysate respectively. An equal amount of mRNA from each experimental group was used to construct 16 libraries. The rapid version of RNA-sequencing method (Townsley et al., 2015) was used for library construction. Each sample was barcoded using standard Illumina adaptors 1–16 to allow up to 16 samples to be pooled in one sequencing lane on Illumina HiSeq4000. The libraries were eluted from the pellet with 10 µL of 10 mM Tris (pH 8.0) and pooled as described by Kumar et al. (2012). Quantification and quality assessment of resulting libraries were performed on Fragment Analyzer (FGL_DNF-474-2- HS NGS Fragment 1-6000 bp Mthds) and sequenced using the Illumina HiSeq 4000 platform to generate 100 bp single-end reads at the Vincent J. Coates Genomic Sequencing Facility at UC Berkeley.

2.3 | Transcriptome and genome data processing and mapping

For the transcriptome data, the adapters were removed from raw reads by two consecutive rounds using BBDuk v38.90 (Bushnell, 2014). Two sets of parameters were used in two rounds respectively: first round 'ktrim=r k=21 mink=11 hdist=2 tpe tbo minlength=21 trimpolya=4'; second round 'ktrim=r k=19 mink=9 hdist=1 tpe tbo minlength=21 trimpolya=4'. Then, low-quality reads were also removed with BBDuk using parameters 'k=31 hdist=1 qtrim=lr trimq=10 maq=12 minlength=21 maxns=5 ziplevel=5'. The clean reads of each sample were mapped to the *S. chilense* reference genome (Silva Arias et al., 2023) using BBMap in BBTools. The SAM files were then converted and sorted to BAM files using Samtools v1.11 (Wysoker et al., 2009). The number of reads mapped to each gene was counted via featureCounts v2.0.1 in each sample (Liao et al., 2014). The gene expression level was normalized using the transcripts per kilobase million (TPM) method to eliminate the differences between samples (Wagner et al., 2012).

The relationships among transcriptome samples were evaluated using the TPM values. The correlation coefficient between two samples was calculated to assess repeatability between samples using Pearson's test. Principal component analysis (PCA) was performed using the *prcomp()* function in R (R Core Team, 2023) based on TPM values.

Complementarily, to assess the generality of our results, we also analysed drought and control transcriptomic samples of *S. pennelliii* (PRJEB5809; Bolger et al., 2014) and *S. lycopersicum* (PRJNA812356; Yang et al., 2022) (Dataset S1) to compare to our transcriptomic analyses and assess the generality of our results.

2.4 | Identification of differentially expressed genes and gene co-expression analysis

Differential expression analysis of groups among the different conditions and tissues was performed using the DESeq2 R package (Love et al., 2014). The raw read counts were inputted to detect differentially expressed genes (DEGs). Genes meeting the joint criteria of p-value ≤.001, absolute value of log2FoldChange ≥1 and a false discovery rate (FDR) adjusted $p \leq .001$ were classified as DEGs.

To identify the gene co-expression networks, weighted gene correlation network analysis (WGCNA) was constructed using TPM values to identify specific modules of co-expressed genes associated

with drought stress (Langfelder & Horvath, 2008). We first checked for genes and samples with too many missing values using the *goodSamplesGenes()* function in the WGCNA R package. We then removed the offending genes (the last statement returns 'FALSE'). To construct an approximate scale-free network, a soft threshold power of 5 was used to calculate the adjacency matrix for a signed co-expression network. Topological overlap matrix (TOM) and dynamic-cut tree algorithm were used to extract network modules. We used a minimum module size of 30 genes for the initial network construction and merged similar modules exhibiting >75% similarity. To discover modules significantly related to drought, module eigengenes were used to calculate correlations across samples with different conditions. The visualization of networks was created using Cytoscape v3.8.2 (Shannon et al., 2003).

2.5 | Identification of transcript factor families and transcript factor-binding sites

The protein sequences were obtained from the reference genome and annotation 'gff' file with GffRead (Pertea & Pertea, 2020) and were used to identify transcription factors (TF) families using the online tool PlantTFDB v5.0 (Guo et al., 2007). Furthermore, the upstream 2000 bp sequences of the transcription start sites (TSS) were extracted as the gene promoter from the reference genome to detect transcription factor-binding sites (TFBS). The TFBS dataset of the related species *S. pennellii* was also downloaded from the Plant Transcriptional Regulatory Map (PlantRegMap, http://plantregmap.gao-lab.org/) as background of TFBS identification (Tian et al., 2020). Then, the TFBS of *S. chilense* was identified using the FIMO program in motif-based sequence analysis tools MEME Suit v5.3.2 (Bailey et al., 2015). The TFBS was extracted with $p < 1e\text{-}5$ and $q < .01$.

2.6 | Gene ontology (GO) analysis

We first constructed the dataset of assigned GO terms for the protein sequences of all used genes by PANTHER v16.0 (Mi et al., 2013). Then, the GO enrichment analysis of drought-responsive genes was performed using clusterProfiler v3.14.2 (Yu et al., 2012). Benjamini–Hochberg method was used to calibrate *P* value, and the significant GO terms were selected with *p*-value below .05.

2.7 | Construction of phylostratigraphic map

We performed phylostratigraphic analysis based on the following steps. First, the phylostrata (PS) were defined according to the full linkage of *S. chilense* from the NCBI taxonomy database. The similar PS were merged and finally, 18 PS were generated. Second, the protein sequences were blasted to a database of non-redundant (nr) proteins downloaded from NCBI (https://ftp.ncbi.nlm.nih.gov/blast/db/) with a minimum length of 30 amino acids and an E-value below 10^{-6} using blastp v2.9.0 (Camacho et al., 2009). Third, each gene was assigned to its PS by the following criterion: if no blast hit or only one hit of *S. chilense* with an E-value below 10^{-6} was identified, we assigned the gene to the youngest PS18. When multiple blast hits were identified, we computed the oldest common ancestor (LCA) for multiple hits using TaxonKit v0.8.0 (Shen & Ren, 2021) and then assigned the LCA to a specific PS.

2.8 | Construction of divergence map

Following four steps, we performed divergence stratigraphy analysis to construct a sequence divergence map of *S. chilense* using the function *divergence_stratigraphy()* of the R package '*orthologr*' (Drost et al., 2015). (1) The coding sequences for each gene of *S. chilense* and *S. pennellii* (NCBI assembly SPENNV200) were extracted from their reference and annotation files. (2) We identified orthologous gene pairs of both species by choosing the best blast hit for each gene using *blastp*. We only considered a gene pair as orthologous when the best hit has an E-value below 10^{-6}; otherwise, it is discarded. (3) Codon alignments of the orthologous gene pairs were performed using PAL2NAL (Suyama et al., 2006). Then, the Ka/Ks values of the codon alignments were calculated using Comeron's method (Comeron, 1995). (4) All genes were sorted according to Ka/Ks values into discrete deciles called divergence stratum (DS).

2.9 | Estimation of transcriptome age index and transcriptome divergence index

The transcriptome age index (TAI) was computed based on phylostratigraphy and expression profile, which assigned each gene to different phylogenetic ages by identification of the homologous sequences in other species (Domazet-Lošo et al., 2007). The evolutionary age of each gene was quantified by combining its PS and expression level to obtain the weighted evolutionary age. Finally, the weighted ages of all genes were averaged to yield TAI, which was defined as the mean evolutionary age of a transcriptome (Domazet-Lošo & Tautz, 2010). A lower value of TAI described an older mean evolutionary age, whereas a higher value of TAI denoted a younger mean evolutionary age implying that younger evolutionary genes were preferentially expressed in the corresponding sample or condition (Domazet-Lošo & Tautz, 2010; Piasecka et al., 2013). The transcriptome divergence index (TDI) represents the mean sequence divergence of a transcriptome quantified by divergence strata (DS) and gene expression profile (Quint et al., 2012). The genes were assigned to different DS and then weighted by their expression level to yield the TDI. A lower value of TDI described a more conserved transcriptome (in terms of sequence dissimilarity), whereas a higher value of TDI denoted a more variable transcriptome. Here, we calculated TAI and TDI profiles in different samples using the *PlotSignature()* function of the myTAI R package.

2.10 | Population genetics analysis and detection of positive selection on drought-responsive genes

Whole-genome sequence data from six populations of *S. chilense* (five individuals each) previously analysed in Wei et al. (2023; BioProject PRJEB47577) were used to calculate population genetics statistics for coding and promoter region sequences for all genes identified in the GCNs. Single nucleotide variants (SNPs) were based on the short-read alignment to the new reference genome for *S. chilense* (Silva Arias et al., 2023) using the same methods in Wei et al. (2023). Population genetics statistics, namely nucleotide diversity (π) and Tajima's D, were calculated with ANGSD v0.937 (Korneliussen et al., 2014) over gene and promoter regions. These statistics were first calculated per site in gene and promoter region, and then we used an R script (https://gitlab.lrz.de/population_genetics/s.chilense-drought-transcriptome) to obtain statistics for each gene and the corresponding promoter regions in bins of 100 bp. A PCA on SNP data from 30 whole genomes was also performed using GCTA v1.91.4 (Yang et al., 2011). The genetic structure inference was performed using ADMIXTURE v1.3.0 (Alexander et al., 2009).

The selective sweeps are identified based on the whole-genome SNPs of population C_LA1963. The sweep regions under positive selection were identified using the same pipeline as in our previous study (Wei et al., 2023). Based on neutral simulations, the same thresholds of 5.42 and 6.84 are used to identify selective sweeps for SweeD and OmegaPlus respectively. In addition, the sweep ages are also estimated using McSwan (Tournebize et al., 2019) with the same parameters as in Wei et al. (2023).

2.11 | Estimation of allele age

We implemented genealogical estimation of variant age (GEVA) to date genomic variants at the drought-responsive genes (Albers & McVean, 2020). We generated input for GEVA using a recombination rate of 3.24×10^{-9} per site per generation based on the overall recombination density in *S. lycopersicum* [1.41 cM/Mb] (Anderson & Stack, 2002; Nieri et al., 2017) and within the possible range of rates used in Wei et al. (2023). We used a population size (N_e) of 20,000 and a mutation rate of 5.1×10^{-9} per bp (Roselius et al., 2005; Wei et al., 2023), and then relied on the recombination clock to estimate the age of alleles (tMRCA).

3 | RESULTS

3.1 | Drought experiments and transcriptome analyses

Plants of *S. chilense* growing under well-watered or moderate-water-stress regimes (hereafter, control and drought) show clear morphological differences during tissue collection. Plant growth and ramification are boosted in well-watered groups, while plants under drought are smaller and slow growing. Hence, on day 12, newly expanded leaf and shoot apices are collected for the expression analysis of stress-responsive genes, and four biological replicates are used for all RNA-Seq experiments from each tissue type.

We analyse transcriptome data from 16 libraries (4 for control_leaf, 4 for drought_leaf, 4 for control_shoot apex and 4 for drought_shoot apex), each library including 5–6 biological replicates per condition and tissue, and then aligned to the reference genome of *S. chilense* (Dataset S1). A total of 27,832 genes are identified to be expressed in the 16 libraries (Dataset S2), of which 1536 genes are uniquely expressed in drought conditions and 1767 genes in control conditions. A PCA based on the gene expression profiles reveals consistent clustering primarily associated with the experimental conditions (control and drought) and secondarily with the developmental stages (leaf and shoot apex) (Figure 1a). PC1 accounts for 28.17% of the expression variability and separates the libraries from the two experimental conditions, indicating transcriptome remodelling between drought and control conditions. Libraries from different developmental stages are separated along the PC2 axis (accounting for 18.24% of the variance), supporting tissue age transcriptome specificity. Consistently, the transcriptome similarity and hierarchical clustering (Figure 1b) analyses also group the transcriptomes, mainly according to water-deficit intensity rather than tissue type.

3.2 | Identification of gene networks involved in drought stress

We identify three sets of differentially expressed genes (DEGs) from three drought/control comparison groups (full data set, only leaf and only shoot apex tissues) (Figure 2a; Dataset S3; log2FoldChange ≥1, FDR $p \leq .001$). A total of 4905 DEGs are obtained, of which 2484 DEGs (1235 up-regulated and 1249 down-regulated in drought transcriptome) are shared across the three comparison groups (Figure 2b). We deduce that these shared DEGs correspond to a core functional drought-responsive network.

To construct the gene co-expression networks (GCNs), we do not directly use the identified DEGs in WGCNA because it would violate the assumption of scale-free topology upon which this method relies. Therefore, a set of 16,181 of all expressed genes is used as input for WGCNA, yielding a clustering into eight co-expression modules named after different colours (genes in the grey module have not been clustered in any module). The module sizes range from 183 to 5364 genes (Figure 2c; Dataset S4). Among the identified co-expression modules, the blue module (3852 genes) shows a significantly positive correlation with the control condition and a negative correlation with the drought condition (Figure 2c; Kendall's test, $p = 2.20e\text{-}11$). In contrast, the turquoise module (5364 genes)

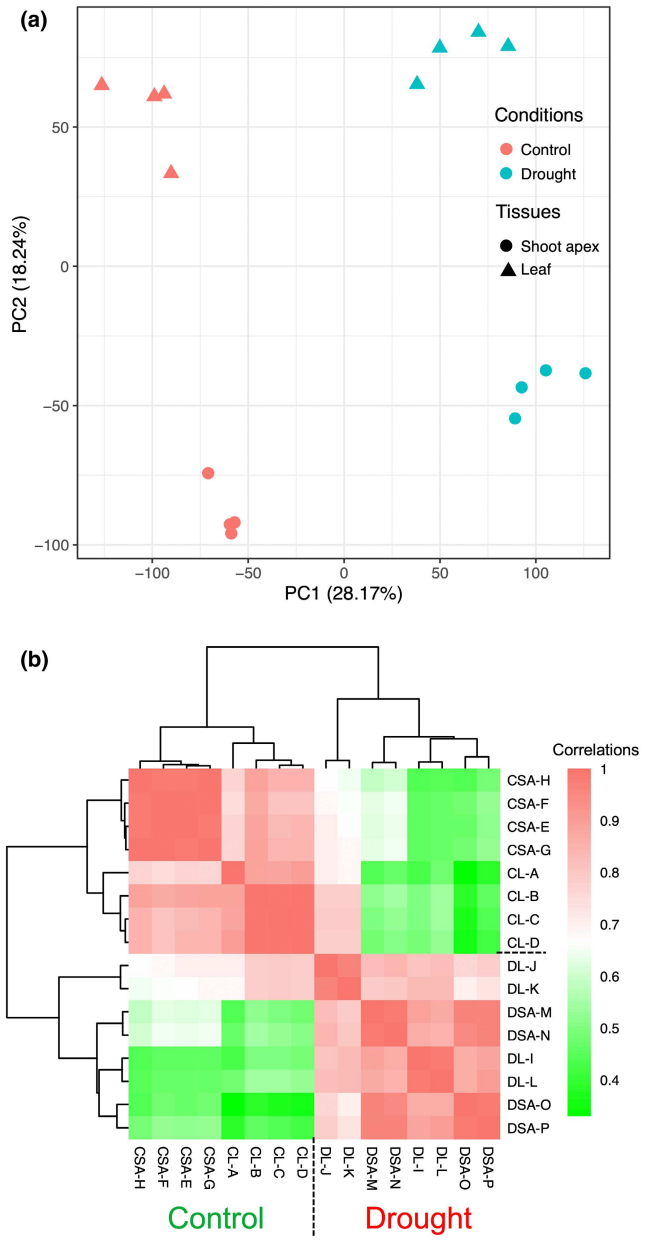

FIGURE 1 Exploratory analyses of RNA-Seq differential expression patterns in 16 libraries of *Solanum chilense*. (a) PCA reveals stronger clustering associated with the experimental conditions. (b) Heatmap plot of sample correlation (Pearson's test) and hierarchical clustering analysis based on Euclidean distance reveal exact drought specificity. The colour scale indicates correlation coefficients from high values in red to low values in green. RNAseq libraries abbreviations: CL-A to CL-D, leaves in the control condition; CSA-E to CSA-H, shoot apex in the control condition; DL-I to DL-L, leave in drought condition; DSA-M to DSA-P, shoot apex in drought condition.

is significantly positively correlated with the drought condition and negatively correlated with the control condition (Figure 2c; Kendall's test, $p = 2.34e\text{-}13$). In addition, the genes within blue and turquoise modules show higher connectivity than other modules (Figure S1; Kolmogorov–Smirnov test, $p = 2.41e\text{-}17$).

Next, we check the overlap between 2484 DEGs and co-expression modules to confirm that blue and turquoise modules are associated with drought stress in *S. chilense* (Table S1). DEGs share far more genes with the blue and turquoise modules than other co-expression modules (93% of DEGs are found in the blue and turquoise modules). This confirms that blue and turquoise modules encompass two sets of co-expressed drought stress-responsive genes. The overlapping DEGs and module genes are extracted to constitute two high-confidence subsets, the blue and turquoise modules, comprising 1223 and 1079 genes respectively.

3.3 | Regulatory functions in the identified GCNs are confirmed by predicted TF and TFBS interactions

To independently support regulatory relationships among genes identified in the two co-expression networks, we identify the transcription factors (TF) and transcription factor-binding sites (TFBS) for the two high-confidence subsets of genes. Therefore, we identify the TF and their TFBS (Table S2) and hereafter rename these sets as sub-blue (686 genes) and sub-turquoise (948 genes), respectively (Dataset S5). The genes in the sub-blue and sub-turquoise networks not only show specific co-expression patterns but also agree on predicted TF and TFBS interactions. The co-expression network reconstructed for the set of genes of the sub-turquoise network exhibits higher connectivity than the sub-blue network (Figure S2; Kolmogorov–Smirnov test, $p = .002$).

3.4 | Identified GCNs in *S. chilense* are also found in *S. pennellii* and cultivated tomato

To verify the robustness of these two drought-responsive networks across different tomato species (and thus the generality of our results), we employ the same pipeline to reconstruct GCNs from transcriptome data of *S. pennellii* (PRJEB5809; Bolger et al., 2014) and *S. lycopersicum* (PRJNA812356; Yang et al., 2022) under drought conditions (Dataset S1). Those transcriptomes exhibit differences attributed to watering conditions (Figure S3). We find that 74% of our DEGs (1837) in *S. chilense* overlap with the combined DEG set of *S. pennellii* and *S. lycopersicum* (Figure S4a,b). The GCNs from the combined transcriptome profiles of the two other tomato species consistently show the same two GCNs (Figure 2c, Figure S4c), so that 576 (84%) and 778 (82%) genes overlap between *S. chilense* and the other tomato species for the sub-blue and sub-turquoise networks, respectively (Figure S4d,e). Furthermore, nearly 60% of our two drought-responsive network genes (sub-blue and sub-turquoise) also overlap with DEGs of drought transcriptomes of *S. lycopersicum* (Nicolas et al., 2022). These overlap rates suggest that our two drought-responsive networks, albeit found using transcriptomics of one population of *S. chilense*, are present and perform similar functions in different tomato species. We thus focus, thereafter, on the sub-blue (686 genes) and sub-turquoise (948 genes) networks.

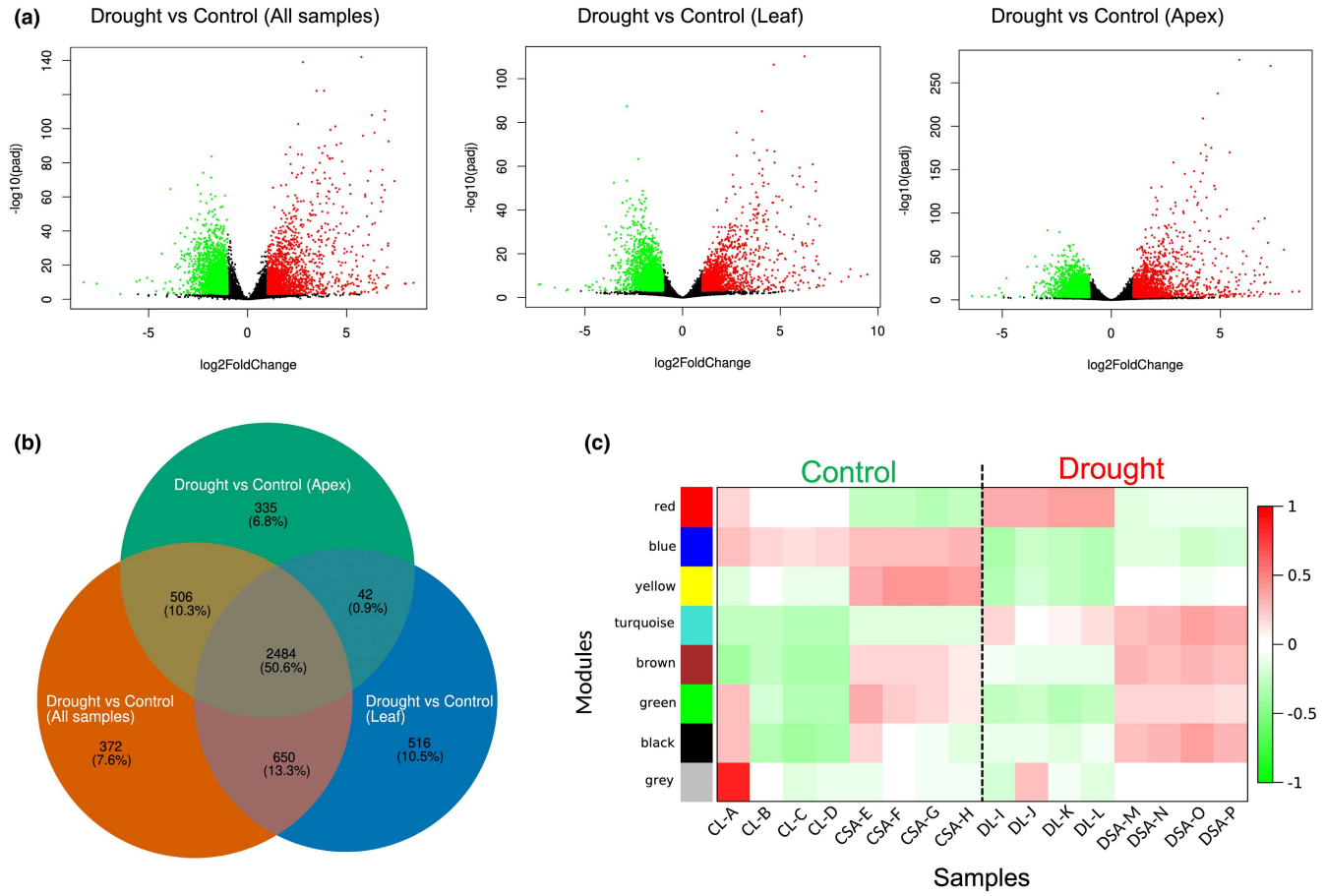

FIGURE 2 Identification of drought-response networks in *Solanum chilense*. (a) Differentially expressed genes (DEGs) identified from three comparison groups from left to right: 8 control versus 8 drought samples, 4 control leaves versus 4 drought leaves and 4 control shoot apices versus 4 drought shoot apices. Red indicates significantly upregulated genes, and green indicates significantly downregulated genes between control and drought samples using fold change higher than 2 ($p \leq .001$). (b) Venn diagram representing 2484 shared DEGs in three comparison groups. (c) Correlation between samples' expression patterns for the eight co-expression modules. A vertical dotted line separates control and drought samples. The colour scale indicates correlation coefficients from positive coefficient in red to negative coefficient in green. No correlation is indicated in white.

3.5 | Drought-responsive GCNs enrich for cell cycle and metabolic processes

We assess whether the two identified gene networks show functional differences. The gene ontology (GO) enrichment reveals that the sub-blue network is significantly enriched ($p < .05$) in cell cycle biological processes, including replication and modification of genetic information, ribosome production and assembly and cytoskeleton organization, among others (Figure 3a; Table S3). Conversely, the sub-turquoise network is enriched in response to physiological and metabolic processes to water shortage and heat, including some metabolic processes, signal pathways and changes of stomata and cuticle, among other processes (Figure 3a; Table S3). These functional differences suggest that genes in the two sub-networks are activated and expressed in different cellular compartments. The sub-blue network genes are linked to key components for cell division processes, such as nucleus (including nucleolus), chromosome, nuclear envelope and ribosome (Figure 3b; Table S4). The sub-turquoise network is enriched in components related to cellular processes, such as metabolism, membrane complexes and membrane structures (Figure 3b; Table S4). Indeed, modulation in the cell cycle and fundamental metabolism are two major response strategies to drought stress (Gupta et al., 2020; Nicolas et al., 2022; Yang et al., 2021). We focus, thereafter, on these two sub-networks, and therefore, the sub-blue network is referred to as the *cell cycle network* and the sub-turquoise as the *metabolic network*.

3.6 | The transcriptome ages of genes in the metabolic GCN are younger than those in the cell cycle GCN

We build phylostratigraphic profiles for all genes of the two GCNs, summarizing the gene emergence in 18 stages of plant evolution or phylostrata (PS): PS1 representing the emergence of oldest (ancestral) genes (at the time of the first cellular organisms) to PS18 for the most recent genes (present only in *S. chilense*). The PS18

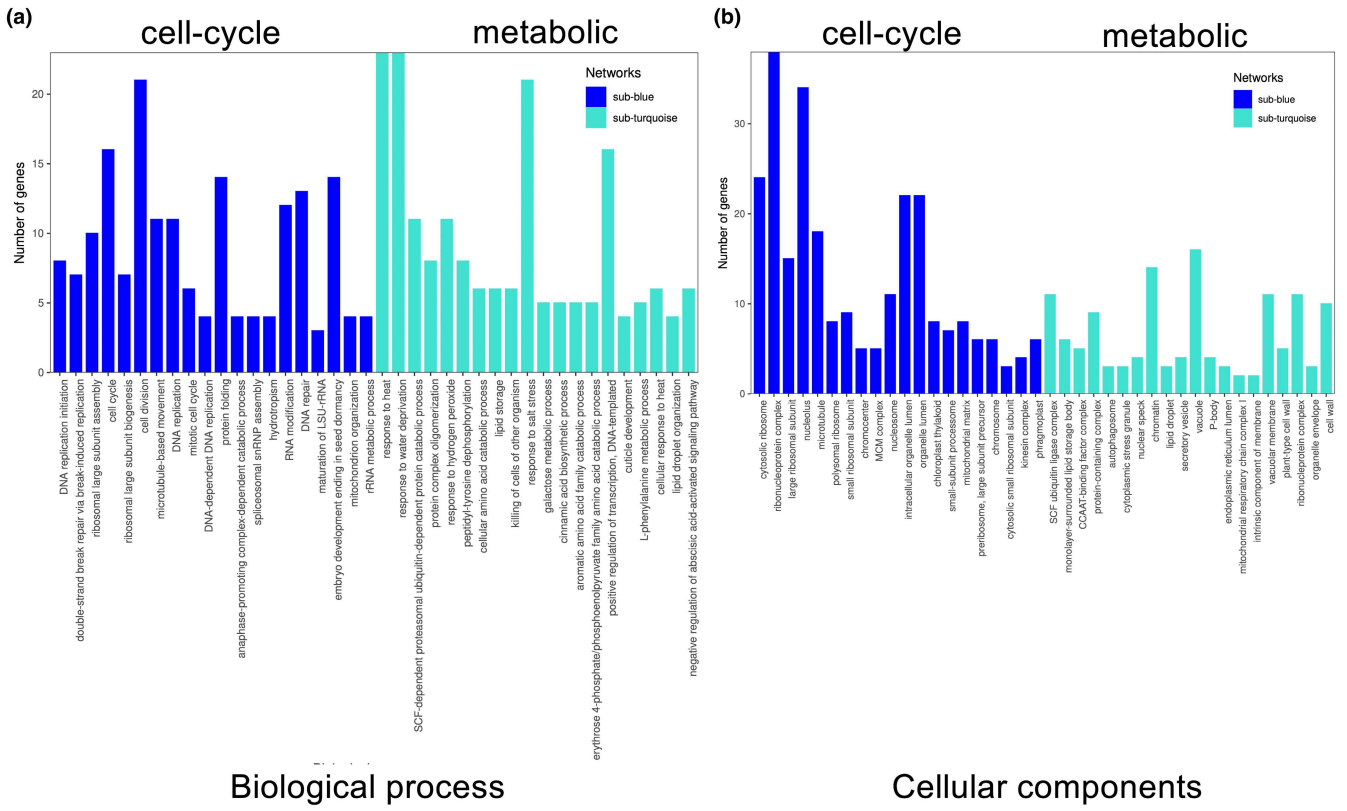

FIGURE 3 Gene ontology (GO) term enrichment in the cell cycle and metabolic drought response networks. (a) Top 20 terms of biological process. (b) Top 20 terms of cellular component.

shares no homologous genes with any other species in the non-redundant protein databases of NCBI (Figure 4a,b; Dataset S6). Most genes in the two GCNs (76.8% in metabolic and 65.5% in cell cycle network) are assigned to three main PS: cellular organisms (PS1), land plants (Embryophyta; PS5) and flowering plants (Magnoliopsida; PS8) (Figure 4a). This implies that the two drought-responsive GCNs have an ancient origin, and their components have remained relatively conserved. While most drought-responsive pathways likely emerged during the colonization of land by plants (PS5), others could derive from exaptation processes from GCNs involved in the core cell process (PS1) or the reproductive organ differentiation (PS8). Interestingly, the cell cycle network shows an older origin (43.7% of genes assigned to the PS1-3), while the metabolic network presents a larger proportion of genes (48.5%) originating in PS8 (Figure 4a,b). Under drought conditions, we also find that cell cycle network genes of almost all PS ages are down-regulated, while genes of the metabolic network are up-regulated (Figure S5).

Furthermore, we estimate the age of cell cycle and metabolic GCNs using the transcriptome age index (TAI). A higher TAI value implies that evolutionary younger genes are preferentially expressed at the corresponding condition/developmental stage. We observe higher TAI values in drought samples, supporting that the drought-responsive genes exhibit a younger transcriptome age than genes expressed under control conditions. Moreover, TAI values of the metabolic GCNs are significantly higher than those of the cell cycle (Figure 4c; Kolmogorov–Smirnov test, $p = 12.51e\text{-}7$), supporting the previous result that transcriptome ages of the genes in the cell cycle are older than that of the metabolic GCNs. We do not find a significant difference in TAI values between control and drought samples based on 1000 randomly selected from non-drought-responsive genes (Figure S6a; Kolmogorov–Smirnov test, $p = .34$), while in cell cycle and metabolic networks, the mean evolutionary ages of the transcriptomes are significantly different between drought and control conditions (Figure 4c; Kolmogorov–Smirnov test, $p = .03$).

The contributions of the different PS to the TAI profiles differ between the two networks (Figure 4d,e): early divergent genes (PS1 to PS7) show more constant transcriptome age in all conditions, and the genes found in PS1, PS5 and PS8 are likely important in both GCNs. The late-emerging genes (PS8 to PS18) increasingly contribute to the differential expression patterns between control and drought samples with age. This indicates that younger drought-responsive genes are differentially expressed under drought stress in both GCNs (Domazet-Lošo & Tautz, 2010; Piasecka et al., 2013). Remarkably, in contrast to the cell cycle network genes (PS9-18, Figure 4d,e), the youngest genes in PS18 present a higher contribution in the metabolic GCN, suggesting these genes are likely involved in *S. chilense* adaptation to drought conditions.

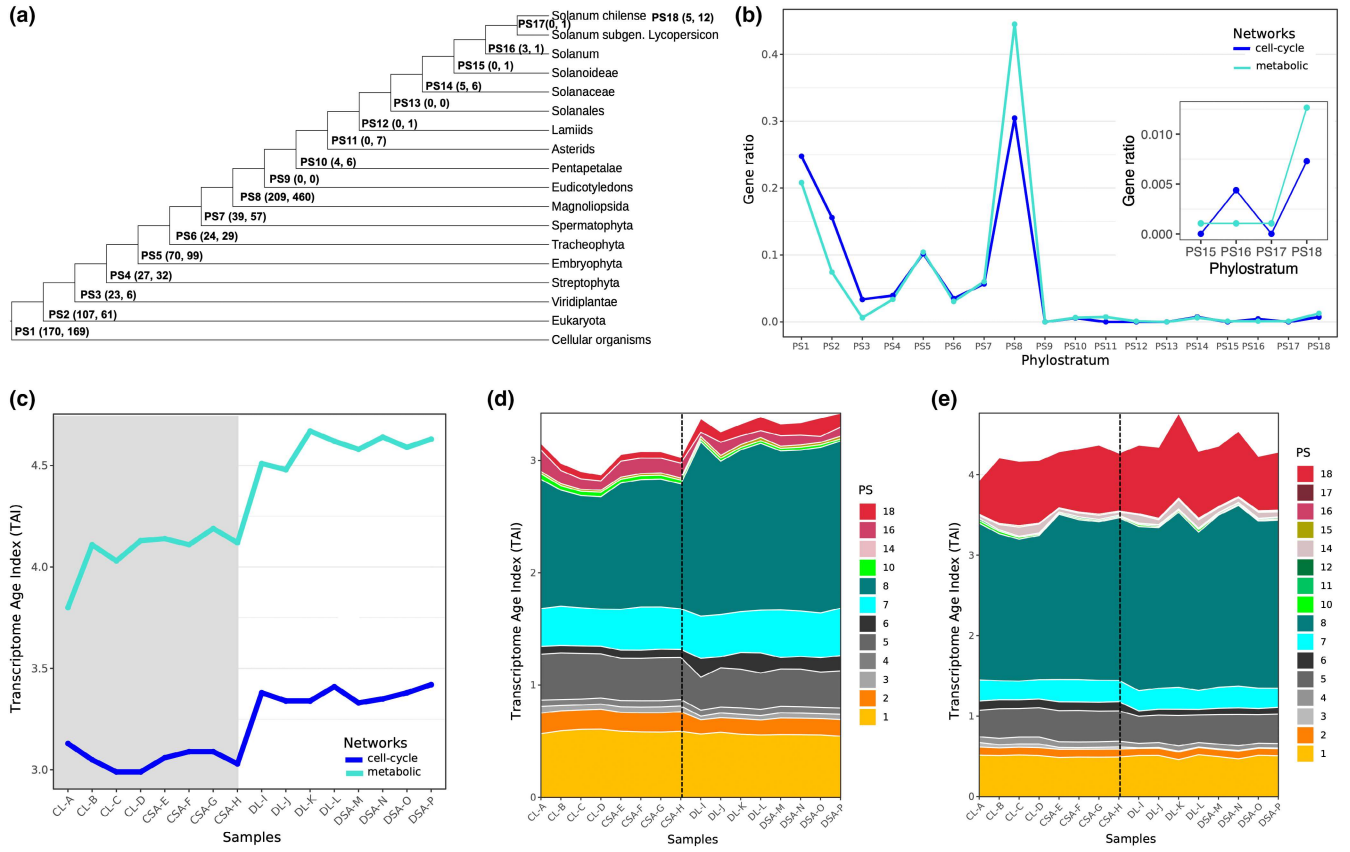

FIGURE 4 Transcriptome age index (TAI) profiles of cell cycle and metabolic networks. (a) Phylostratigraphic map of two networks and phylogeny used in the search for the evolutionary origin of *Solanum chilense* genes. Numbers in parentheses denote the number of genes assigned to each phylostratum (PS) in cell cycle and metabolic network respectively. (b) Gene ratio in each PS for two networks. (c) TAI profiles of two networks in control samples (grey background) and drought samples (white background). (d) TAI contributions split according to different PS in the cell cycle network. A vertical dotted line separates control and drought samples. (e) TAI contributions split according to different PS in the metabolic network.

3.7 | Genes of the metabolic GCN show more recent divergence and weaker purifying selection than the cell cycle GCN

A total of 10 divergence strata (DS) are constructed from the transcriptome differentiation index (TDI) based on the sequence divergence between *S. chilense* and *S. pennellii* (Figure 5a, S7; Dataset S6). The distributions of the Ka/Ks ratio per gene for both GCNs indicate the action of purifying selection, which confirms the conservation of most drought-responsive genes at the species level. Consistent with the phylostratigraphic patterns (PS, see above), the strength of purifying selection at the cell cycle GCN (Ka/Ks = 0.279 ± 0.333) is higher than at the metabolic GCN (Ka/Ks = 0.329 ± 0.331) (Kolmogorov–Smirnov test, $p = 2.34e\text{-}11$; Figure 5a; Table S5). In addition, higher TDI values are observed in the drought samples (Figure 5b), suggesting that the drought-responsive genes exhibit a more conserved transcriptome profile under control compared to drought conditions (Kolmogorov–Smirnov test, $p = .004$). No significant difference is found between control and drought samples based on 1000 random genes (Kolmogorov–Smirnov test, $p = .17$; Figure S6b). This result supports that different selection pressures act on two GCNs across conditions. In accordance with the TAI results, the metabolic GCN appears to exhibit a higher transcriptome divergence than the cell cycle GCN (Figure 5b; Kolmogorov–Smirnov test, $p = 2.25e\text{-}7$). Moreover, the low TDI in the cell cycle GCN and larger TDI differences between drought and control transcriptomes also suggest that cell cycle regulation is likely an ancestral strategy of stress response (pre-speciation of *S. chilense*). The transcriptome of the cell cycle GCN may have been evolving and changing in older times to reach a conserved structure in recent times, while the metabolic GCN may appear to have experienced changes and re-wiring in recent times.

The contributions of the low-divergence DS classes (low Ka/Ks in DS1 to DS5) in the cell cycle GCN (~ 50% of the genes) are larger than in the metabolic GCN (DS1 to DS5 about 30%), especially in DS1 (lowest Ka/Ks ratio; Figure 5c,d). This indicates that purifying selection is pervasive in genes of the cell cycle GCN. In contrast, the metabolic GCN genes show about 70% contributions in high DS (higher Ka/Ks ratio in DS6 to DS10), especially in DS10 (highest Ka/Ks ratio), indicating that genes in the metabolic GCN evolve under weaker purifying selection and/or recent evolutionary changes occurred.

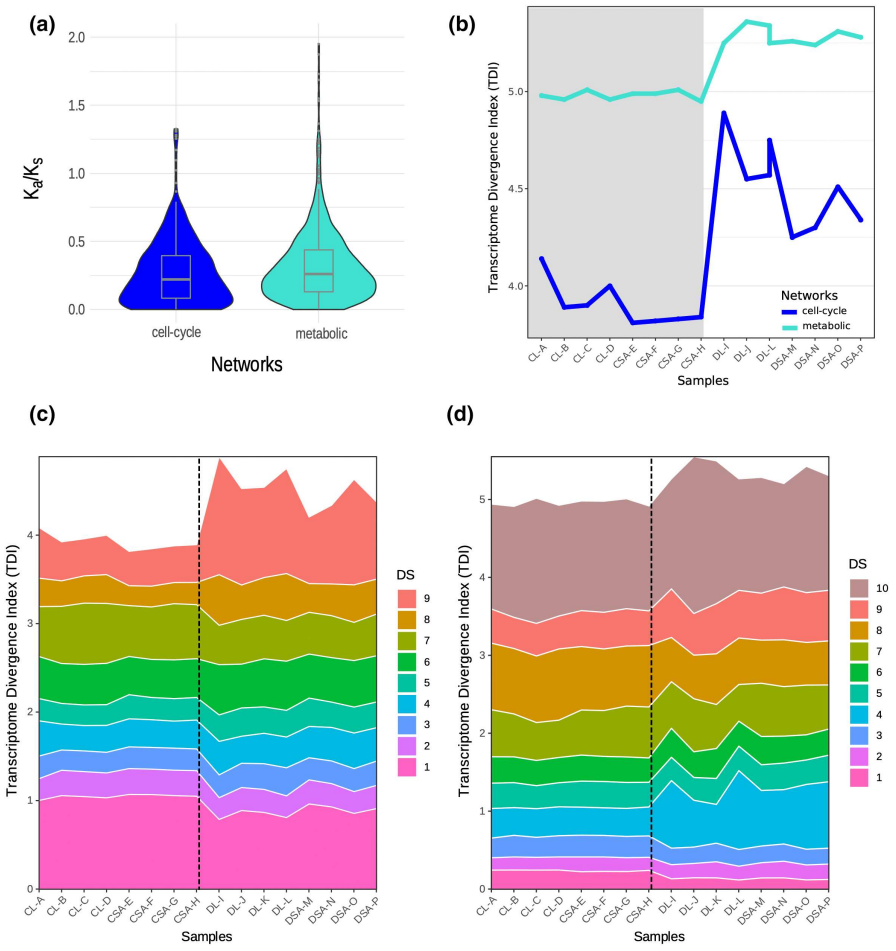

FIGURE 5 Transcriptome divergence index (TDI) profiles of cell cycle and metabolic networks. (a) Distribution of Ka/Ks ratio of genes in two networks respectively. (b) TDI profiles of two networks in control samples (grey background) and drought samples (white background). (c) TDI contributions split according to different DS in cell cycle network. A vertical dotted line separates control and drought samples. (d) TDI contributions split according to different DS in the metabolic network.

3.8 | Comparative analysis reveals stronger positive selection signals in metabolic GCN genes compared to cell cycle GCN

We also study the selective forces acting on the identified drought-responsive GCNs at the population level. We first investigate the effects of these drought-responsive genes on population structure using whole-genome sequences of six *S. chilense* populations (three central populations: C_LA1963, C_LA3111, C_LA2931; two south coast populations: SC_LA2932, SC_LA4107; one south highland population: SH_LA4330; five plants each) analysed in Wei et al. (2023). Here, we aligned to the reference genome of *S. chilense*, and we identified 45,208,263 high-quality SNPs, of which 111,606 SNPs are found in genes of the cell cycle GCN and 167,334 SNPs in genes of the metabolic GCN. We confirmed that the overall population structure between the whole-genome data and drought-responsive genes follows the results in Wei et al. (2023) (Figure S8a,c). However, the structure exhibited by drought genes shows (1) stronger differentiation among populations for the central region and SH_LA4330 and (2) weaker differentiation between the two south coastal populations (SC_LA2932 and SC_LA4107) (Figure S8b,d). Moreover, the two SC populations showed clearer structure and a higher proportion of variance elucidated in PCA conducted with whole-genome SNPs aligned to the *S. chilense* reference (Figure S8a), contrasting with findings from alignment to the *S. pennellii* reference (see Figure 1c in Wei et al., 2023). This discrepancy likely stems from the enhanced precision in read mapping and SNP calling between populations when employing the *S. chilense* reference genome.

Focusing on the genetic diversity of population C_LA1963, we find that the mean nucleotide diversity (π) per gene does not differ between the two GCNs (Figure S9a; Table S5; Kolmogorov–Smirnov test, $p = .15$) and the π values of the promoter regions (2 kb upstream of the transcription initiation site) are significantly higher than those of the coding regions (Figure S9a; Table S5; Kolmogorov–Smirnov test, $p = .03$) for both GCNs. This result suggests that the selective constraints in promoter regions may be more relaxed, which could partly explain why certain transcription factors (TF) can bind to multiple genes in the GCNs (Table S2). TFs are indeed conserved at the coding sequence level, especially at the functional domains. Still, a higher amount of polymorphism of TF binding sites in the promoter can indicate complex and diverse regulation, for example, in response to stressful conditions (Sato et al., 2016; Spivakov, 2014). Also, considering that the affinity of the TFs for promotors is associated with smaller regions (boxes), we perform an additional analysis of the nucleotide diversity by bins (100 bp), which shows no difference between the two GCNs (Figure S9a; Table S5). Furthermore, the genes for the metabolic GCN show lower Tajima's D values than

those of the cell cycle GCN (Figure S9b; Table S5; Kolmogorov–Smirnov test, $p = .04$), suggesting possible recent positive selection pressure in the metabolic GCN.

We further find significant but opposite correlations between π or Tajima's D and the contributions of the different DS for the two GCNs (Figure S10; Table S6). The cell cycle GCN shows a positive correlation of the different DS with π and Tajima's D (Figures S10a,c), indicating that DS with high contribution (see DS1 in Figure 5c) show high nucleotide diversity. In contrast, a negative correlation is observed between the contribution of each DS and π or Tajima's D in the metabolic network (Figures S10b,d). Hence, DS with high contribution (see DS10 in Figure 5d) shows low nucleotide diversity and low Tajima's D in the metabolic network. So, it appears likely that the metabolic genes may be under positive selection, underpinning the recent evolution of the drought response transcriptome.

3.9 | The high correlation between the timing of adaptation and gene connectivity suggests strong network re-wiring during colonization to dry habitats

To investigate drought-responsive genes that have potentially undergone a shift in selection regime in *S. chilense*, we identify candidate genes under positive selection and estimate the age of selective sweep using the whole-genome SNPs of the population C_LA1963 and the *S. chilense* reference genome (see Methods). In total, 284 candidate genes from 492 selective sweep regions are obtained using the same pipeline previously used (Wei et al., 2023). The number of candidate genes is more than three times that of the previous studies (86 sweep genes) based on the *S. pennellii* reference genome, and these selective sweeps cover 79% of the putatively selected regions in C_LA1963 previously obtained in Wei et al. (2023). We find 17 and 31 drought-responsive genes in the cell cycle and metabolic networks, respectively, in the list of candidate genes under positive selection (Dataset S7).

To address the role that (putatively) positively selected genes may play within the drought-responsive networks, we compare the connectivity of these genes in the two GCNs. In the metabolic GCN, the connectivity of positively selected genes (mean connectivity 0.50 ± 0.11) is significantly higher than other drought-responsive genes (mean connectivity 0.43 ± 0.13) (Figure S11a; Kolmogorov–Smirnov test, $p = .028$), but no such difference is observed for the cell cycle network (Figure S11a; Kolmogorov–Smirnov test, $p = .58$). The connectivity of positively selected genes of the metabolic network (0.50 ± 0.11) is much higher than those from the cell cycle network (0.43 ± 0.083) (Figure 6a; Kolmogorov–Smirnov test, $p = .0028$). These results suggest that highly connected (likely more pleiotropic) genes in the metabolic GCN may have facilitated the species colonization into dry habitats (Hämälä et al., 2020) and that the two networks underwent different evolutionary selective pressures during the species divergence of *S. chilense*.

Finally, we compare the age of the selective sweep at the candidate genes of the two GCNs. We find that sweep ages at the cell cycle genes (34.74 ± 13.11 kya) are significantly younger than at sweeps in the metabolic network (39.10 ± 12.61 kya; Figure 6b; Kolmogorov–Smirnov test, $p = .016$). This supports drought-responsive GCNs as an important mechanism underlying the recent local adaptation or (re) colonization (Raduski & Igić, 2021; Wei et al., 2023). Interestingly, we find a significant positive correlation between the age of the sweep and gene connectivity for both GCNs (Figure 6c–e). In other words, selective sweeps appear first at more connected genes and, subsequently, at less connected genes, during the history of colonization/adaptation of arid habitats. To our knowledge, this is the first report of a correlation between the age of a selective sweep and the gene connectivity in a network. We support this conclusion by revealing a positive correlation between the tMRCA of drought-responsive genes and gene connectivity (Pearson's cor = .72, $p = 4.28\text{e-}6$). In contrast, this correlation is low for drought-responsive genes outside sweep regions (Pearson's cor = .32, $p = .14$). Our results support the hypothesis of polygenic adaptation in GCNs where the positive selection acts first on core genes (with high connectivity and more pleiotropic effects) of networks and subsequently on the peripheral genes (less connectivity and less pleiotropic effects).

4 | DISCUSSION

This study identifies two drought-responsive GCNs by analysing gene expression profiles of plants growing under control and drought conditions. Two GCNs involved in cell cycle and metabolic biological processes are detected, and their structural relevance is supported by TF/TFBS predictions. These networks represent two different strategies for drought response (Danilevskaya et al., 2019; Farooq et al., 2009). We then demonstrate that the cell cycle network is evolutionary older, and more conserved than the metabolic network. Despite the ancient history of these two GCNs, we further show that both GCNs contribute to different extents to contemporary adaptation processes to arid habitats. The joint analyses of genomic and transcriptomic data indicate that (1) at the transcriptome level, metabolic GCNs present higher evolvability, especially with younger selection events linked to response to drought environment, (2) cell cycle GCN is less evolvable and (3) both networks still present signals of evolution under positive selection in core elements of the GCN, while peripheral genes of the network can be involved in adaptation at later (more recent) stages of the colonization processes.

4.1 | Drought tolerance is mediated by the regulation of cell proliferation and metabolism

When roughly defining organ development into cell proliferation and differentiation, water deficit is a limiting factor for both processes (Alves & Setter, 2004; Verelst et al., 2013). Drought stress reduces the cell cycle's activity and thus slows plants' growth and development. The down-regulated genes we find in the cell cycle network likely indicate that the expression of genes related to the cell cycle

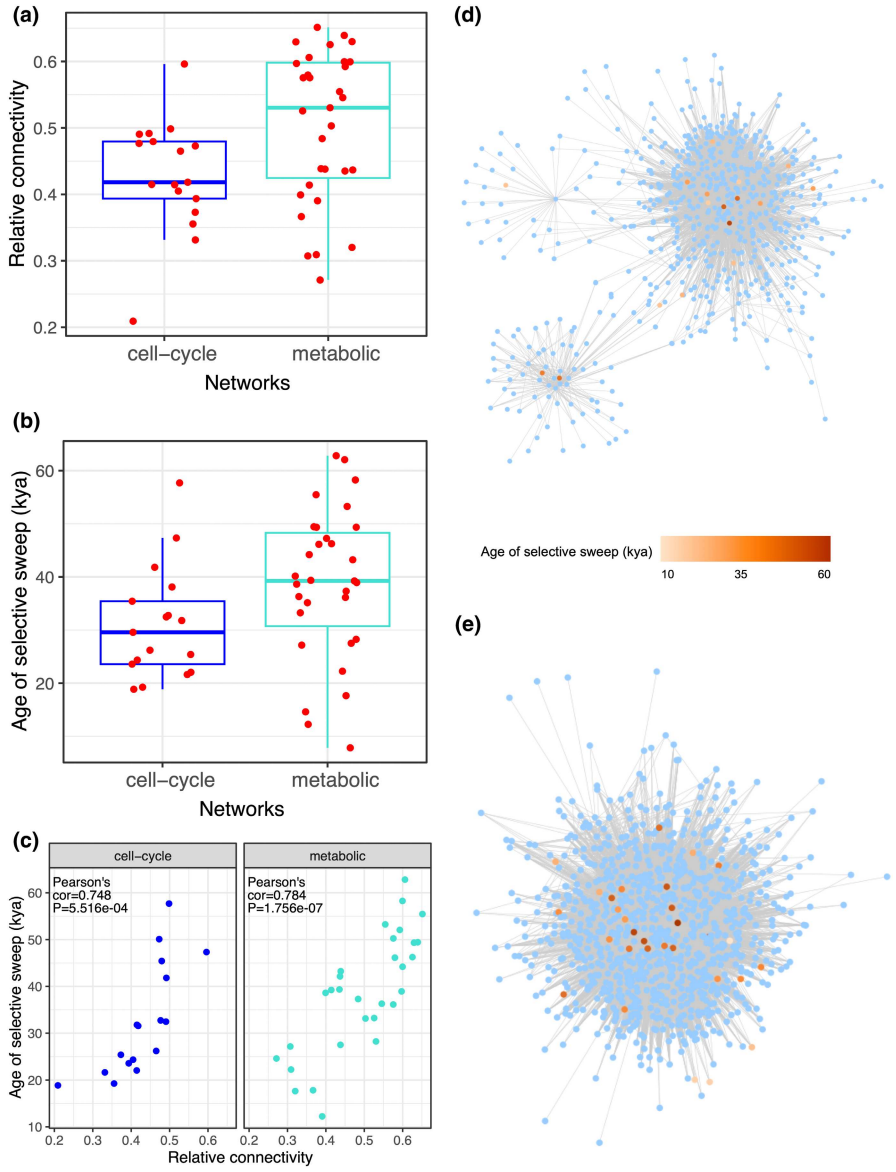

FIGURE 6 Drought-responsive genes under positive selection in population C_LA1963. (a) The connectivity of drought-responsive genes under positive selection in the two networks (b) The age of selective sweep of drought-responsive genes under positive selection in the two networks. (c) The correlations between connectivity and the sweep age of drought-responsive genes under positive selection in the two networks. (d) The visualization of the cell cycle network includes 643 genes with a weighted correlation greater than 0.65 between them. (e) The visualization of the metabolic network includes 937 genes with a weighted correlation greater than 0.65 between them. Red-to-orange coloured dots denote drought-responsive genes under positive selection, and the red-to-orange scale denotes the ages of selective sweeps. The location of the dots closer to the centre of the networks indicates genes exhibiting higher connectivity.

is suppressed by drought stress to possibly restrict cell division in *S. chilense*. The reduction in cell number due to mild drought stress is also found in *A. thaliana* (Skirycz & Inzé, 2010). This means that the cell cycle response to drought may directly and indirectly impact the general physiology of plants. Conversely, the changes in fundamental metabolic activity may be a faster and a flexible drought-responsive strategy (Harb et al., 2010). Plant water shortage is first reflected in changes in metabolic processes, such as accelerating the catabolism of macromolecules to regulate the penetration of tissues, maintaining physiological water balance or slowing down metabolism to reduce energy and water consumption (Gupta et al., 2020; Reddy et al., 2004). In addition, several signalling pathways related to the metabolic gene network have also been demonstrated to be involved in the response to drought stress. For example, the abscisic acid (ABA) signalling pathway regulates the response to dehydration and optimizes water utilization (Harb et al., 2010; Wilkinson & Davies, 2010). Although these two GCNs correspond to two different drought response strategies, they are not isolated but interact with one another in a time-dependent manner. Water deprivation and heat first change the metabolic processes, leading to stomata closure, which leads to the cell cycle network being affected under prolonged water deficit. In return, the increased or decreased cell cycle gene expression affects the further physiology and metabolism of the plant (Gupta et al., 2020). Indeed, drought-responsive strategies regulating the cell cycle appear to be activated later than metabolism processes, as glucose metabolism rapidly follows drought stress. In contrast, the accumulation of amino acids, a crucial part of the cell cycle response, is later initiated in response to drought (Fàbregas & Fernie, 2019).

4.2 | Rewiring of ancient GCNs drives recent adaptation to dry environments

The phylostratigraphic analysis supports that the majority of drought-responsive genes in *S. chilense* evolved during the

early-to-intermediate stages of plant evolutionary history, which is in agreement with the time of origin of multiple abiotic response genes studied in *Arabidopsis thaliana* (Mustafin et al., 2019). The emergence of a majority of drought-responsive genes, therefore, coincides with the periods of divergence among major plant groups (land and flowering plants), which are marked by frequent whole-genome duplication events that trigger gene family expansions, gene neo- and sub-functionalization and genome reorganization processes (Clark & Donoghue, 2018; Wang et al., 2012). These genomic processes likely contributed to the enrichment of drought-responsive GCNs. For instance, fundamental morphological traits involved in drought responses, such as stomata, are present in the ancestral land plants. However, stomatal genes existed prior to the divergence of land plants and underwent multiple duplications during evolution. Additionally, the stomata response to environmental cues, such as humidity, light, CO_2 and phytohormones, such as ABA, is widely distributed and may be ancestral to land plants (Clark et al., 2022). Therefore, we propose that our two drought-responsive networks were primarily established during or shortly after the divergence of land plants and have subsequently undergone expansion.

Previous studies show that TAI and TDI profiles across embryogenesis, seed germination and transition to flowering in *A. thaliana* exhibit an 'hourglass pattern' (older and conserved transcriptomes are preferentially active at the mid-development stages) (Drost et al., 2016; Quint et al., 2012). However, our TAI/TDI profiles for the two developmental stages remain stable for a given condition (Figures 4c and 5b). The similarity of TAI/TDI values between developmental stages (Figures 4c and 5b) may be certainly because our analyses focused on two modules (co-expressed genes) highly correlated with the differential expression between drought and control conditions (Figure 2b; Table S1). Therefore, developmental stage-specific response genes are likely underrepresented in the two analysed networks. We speculate that although abiotic stress response regulatory networks are primarily composed of highly ancient and conserved elements across species (Chen & Zhu, 2004), networks can change expression patterns to respond rapidly to environmental changes and thus allow the exploration of new ecological niches.

Extensive network re-wiring in relatively recent and short timeframes has been found in maize and tomato in response to domestication (Koenig et al., 2013; Swanson-Wagner et al., 2012). It is, therefore, not surprising to find signs of adaptive variation in core elements of rather conserved regulatory networks related to adaptation processes to arid habitats. While the genetic (and morphological) divergence of the *S. chilense* marginal southern populations, southern coastal and highland, is recent but substantial (Raduski & Igić, 2021), drought adaptation is an ancestral and widespread characteristic of *S. chilense* populations. It is congruent with theoretical results showing that gene networks with higher mutation sensitivity can facilitate local adaptation, increase gene expression variance and underlie accelerated range expansion processes across abiotic environmental gradients (Deshpande & Fronhofer, 2022). Our results indeed indicate that changes at central genes (with higher connectivity) can be responsible for the short-term response to selection (Jovelin & Phillips, 2009; Luisi et al., 2015) and promote re-wiring of the gene network (Koubkova-Yu et al., 2018), a critical process for adaptation to contrasting habitats and lineage divergence. Complementarily, our empirical approach shows two regulatory networks with different evolutionary trends and exhibiting different gene expression responses. One GCN would exhibit a faster and more variable response (metabolic), while the other would exhibit a later (delayed) but more constitutive response (cell cycle) to drought. Despite the differences in gene age and variation between the networks, our results show that both GCNs have undergone sufficient changes leading to their re-wiring during the divergent process of *S. chilense* in dry habitats in South Peru and North Chile. Nevertheless, genes in the metabolic network show more recent evolution, with new gene members appearing in *S. chilense*, concomitantly with more variable expression in the drought transcriptome.

4.3 | Limitations and further work

A limitation of our gene expression study is that our transcriptomic analyses are based on few individuals from a single location (near the putative region of origin of the species; Wei et al., 2023), while variability in gene expression and phenotypic response has been observed between different populations (Fischer et al., 2013; Mboup et al., 2012; Nosenko et al., 2016). Further expression studies, including multiple plants from various locations, would be helpful to verify that the identified GCNs are also present and expressed in other populations and study the possible variation in southern (divergent) populations. However, as most genes we find in the two CGNs also show expression differences in response to drought in other tomato species, our results are likely representative of the CGN evolution across the whole range of *S. chilense*. Finally, we refer the reader to our previous study (Wei et al., 2023), where we discuss possible limitations of our population genetics analyses due to using plant material from the TGRC (Tomato Genetics Resource Center, UC Davis, USA). To conclude, we suggest that conducting sampling and experimental work in the field is necessary to improve the resolution of transcriptome and genomic studies, and better assess phenotypic differences between organs and stages of development.

AUTHOR CONTRIBUTIONS

KW, GAS-A and AT planned and designed the study. SS, NS and HN performed all drought experiments and obtained the transcriptome data. KW performed data analyses and XZ and SS supported the gene expression analyses. KW wrote the first draft of the manuscript and SS, GAS-A, AT, NS and HN edited and improved the manuscript. All authors approved the final manuscript. KW and SS contributed equally to this work.

ACKNOWLEDGEMENTS

KW was funded by the Chinese Scholarship Council. SS was funded by the Fulbright Visiting Scholar Program sponsored by the U.S.

Department of State and by the German Academic Exchange Service (DAAD) within the framework of Research Stays for University Academics and Scientists. GAS-A was funded by the Technical University of Munich. AT acknowledges funding from DFG (Deutscheforschungsgemeinschaft) Grant Number: 317616126 (TE809/7-1). NS acknowledges funding from the U.S. National Science Foundation (NSF/IOS-1238243). This work was supported by grants from JSPS KAKENHI (JP19K23742, JP20K06682 and JP20KK0340 to H.N.), and NSF PGRP IOS-1856749 and IOS-211980 to NS. The laboratory work of H.N. was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (JP19H05672). We thank the TUM GHL in Dürnast for providing plant growth facilities and the Tomato Genetics Resource Center (TGRC) of the University of California, Davis, for generously providing us with the seeds of the accession included in this study. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The raw-sequencing RNA data are available in PRJDB15063 (*S. chilense*), PRJEB5809 (*S. pennellii*) and PRJNA812356 (*S. lycopersicum*). The raw pair-end whole-genome sequencing data are available at the European Nucleotide Archive (ENA) under project accession no. PRJEB47577. The nucleotide variants derived from the WGS data are available in VCF format in our GitLab repository: https://gitlab.lrz.de/population_genetics/s.chilense-drought-transcriptome/-/blob/main/chilense_snps.vcf.gz. All codes used in this study and other previously published genomic data are available at the sources referenced. The code for implementing the analyses used in this paper can be found on our GitLab repository: https://gitlab.lrz.de/population_genetics/s.chilense-drought-transcriptome.

ORCID

Aurélien Tellier https://orcid.org/0000-0002-8895-0785
Gustavo A. Silva-Arias https://orcid.org/0000-0002-7114-9916

REFERENCES

Albers, P. K., & McVean, G. (2020). Dating genomic variants and shared ancestry in population-scale sequencing data. *PLoS Biology*, *18*, e3000586.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

Alves, A. A., & Setter, T. L. (2004). Response of cassava leaf area expansion to water deficit: Cell proliferation, cell expansion and delayed development. *Annals of Botany*, *94*, 605–613.

Anderson, L., & Stack, S. (2002). Meiotic recombination in plants. *Current Genomics*, *3*, 507–525.

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, *43*, W39–W49.

Barrera-Ayala, D., Tapia, G., & Ferrio, J. P. (2023). Leaf carbon and water isotopes correlate with leaf hydraulic traits in three *Solanum* species (*S. peruvianum*, *S. lycopersicum* and *S. chilense*). *Agriculture*, *13*, 525.

Basu, S., Ramegowda, V., Kumar, A., & Pereira, A. (2016). Plant adaptation to drought stress. *F1000Research*, *5*, 1554.

Blanchard-Gros, R., Bigot, S., Martinez, J.-P., Lutts, S., Guerriero, G., & Quinet, M. (2021). Comparison of drought and heat resistance strategies among six populations of *Solanum chilense* and two cultivars of *Solanum lycopersicum*. *Plants*, *10*, 1720.

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sørensen, I., & Lichtenstein, G. (2014). The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics*, *46*, 1034–1038.

Böndel, K. B. (2014). *Evolution of the wild tomato species Solanum chilense: Demography and natural selection*. Dissertation. Ludwig-Maximilians-Universität.

Böndel, K. B., Lainer, H., Nosenko, T., Mboup, M., Tellier, A., & Stephan, W. (2015). North–south colonization associated with local adaptation of the wild tomato species *Solanum chilense*. *Molecular Biology and Evolution*, *32*, 2932–2943.

Böndel, K. B., Nosenko, T., & Stephan, W. (2018). Signatures of natural selection in abiotic stress-responsive genes of *Solanum chilense*. *Royal Society Open Science*, *5*, 171198.

Bowles, A., Paps, J., & Bechtold, U. (2021). Evolutionary origins of drought tolerance in spermatophytes. *Frontiers in Plant Science*, *12*, 655924.

Bushnell, B. (2014). BBTools software package, 578–579. http://sourceforge.net/projects/bbmap

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421.

Chen, W. J., & Zhu, T. (2004). Networks of transcription factors with roles in environmental stress response. *Trends in Plant Science*, *9*, 591–596.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., de Noblet, N., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., … Valentini, R. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, *437*, 529–533.

Clark, J. W., & Donoghue, P. C. (2018). Whole-genome duplication and plant macroevolution. *Trends in Plant Science*, *23*, 933–945.

Clark, J. W., Harris, B. J., Hetherington, A. J., Hurtado-Castano, N., Brench, R. A., Casson, S., Williams, T. A., Gray, J. E., & Hetherington, A. M. (2022). The origin and evolution of stomata. *Current Biology*, *32*, R539–R553.

Comeron, J. M. (1995). A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. *Journal of Molecular Evolution*, *41*, 1152–1159.

Crow, M., Suresh, H., Lee, J., & Gillis, J. (2022). Coexpression reveals conserved gene programs that co-vary with cell type across kingdoms. *Nucleic Acids Research*, *50*, 4302–4314.

Danilevskaya, O. N., Yu, G., Meng, X., Xu, J., Stephenson, E., Estrada, S., Chilakamarri, S., Zastrow-Hayes, G., & Thatcher, S. (2019). Developmental and transcriptional responses of maize to drought stress under field conditions. *Plant Direct*, *3*, e00129.

De Vries, J., & Archibald, J. M. (2018). Plant evolution: Landmarks on the path to terrestrial life. *New Phytologist*, *217*, 1428–1434.

Deshpande, J. N., & Fronhofer, E. A. (2022). Genetic architecture of dispersal and local adaptation drives accelerating range expansions. *Proceedings of the National Academy of Sciences*, *119*, e2121858119.

Domazet-Lošo, T., Brajković, J., & Tautz, D. (2007). A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends in Genetics*, *23*, 533–539.

Domazet-Lošo, T., & Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature*, *468*, 815–818.

Drost, H.-G., Bellstädt, J., Ó'Maoiléidigh, D. S., Silva, A. T., Gabel, A., Weinholdt, C., Ryan, P. T., Dekkers, B. J., Bentsink, L., & Hilhorst, H. W. (2016). Post-embryonic hourglass patterns mark ontogenetic

transitions in plant development. *Molecular Biology and Evolution*, *33*, 1158–1163.
Drost, H.-G., Gabel, A., Grosse, I., & Quint, M. (2015). Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. *Molecular Biology and Evolution*, *32*, 1221–1231.
Erwin, D. H. (2020). Chapter thirteen - evolutionary dynamics of gene regulation. In I. S. Peter (Ed.), *Current topics in developmental* (pp. 407–431). Academic Press.
Fàbregas, N., & Fernie, A. R. (2019). The metabolic response to drought. *Journal of Experimental Botany*, *70*, 1077–1085.
Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., & Basra, S. M. A. (2009). Plant drought stress: Effects, mechanisms and management. *Agronomy for Sustainable Development*, *29*, 185–212.
Ficklin, S. P., & Feltus, F. A. (2011). Gene coexpression network alignment and conservation of gene modules between two grass species: Maize and rice. *Plant Physiology*, *156*, 1244–1256.
Fischer, I., Camus-Kulandaivelu, L., Allal, F., & Stephan, W. (2011). Adaptation to drought in two wild tomato species: The evolution of the Asr gene family. *New Phytologist*, *190*, 1032–1044.
Fischer, I., Steige, K. A., Stephan, W., & Mboup, M. (2013). Sequence evolution and expression regulation of stress-responsive genes in natural populations of wild tomato. *PLoS One*, *8*, e78182.
Flowers, J. M., Sezgin, E., Kumagai, S., Duvernell, D. D., Matzkin, L. M., Schmidt, P. S., & Eanes, W. F. (2007). Adaptive evolution of metabolic pathways in Drosophila. *Molecular Biology and Evolution*, *24*, 1347–1354.
Gao, M.-F., Peng, H.-Z., Li, S.-S., Wang, X.-L., Gao, L., Wang, M.-H., Zhao, P.-F., & Zhao, L.-X. (2015). Insight into flower diversity in *Solanum lycopersicum* and *Solanum chilense* using comparative biological approaches. *Canadian Journal of Plant Science*, *95*, 467–478.
Gehan, M. A., Greenham, K., Mockler, T. C., & McClung, C. R. (2015). Transcriptional networks—Crops, clocks, and abiotic stress. *Current Opinion in Plant Biology*, *24*, 39–46.
Gerstein, M. B., Rozowsky, J., Yan, K.-K., Wang, D., Cheng, C., Brown, J. B., Davis, C. A., Hillier, L., Sisu, C., & Li, J. J. (2014). Comparative analysis of the transcriptome across distant species. *Nature*, *512*, 445–448.
Guo, A.-Y., Chen, X., Gao, G., Zhang, H., Zhu, Q.-H., Liu, X.-C., Zhong, Y.-F., Gu, X., He, K., & Luo, J. (2007). PlantTFDB: A comprehensive plant transcription factor database. *Nucleic Acids Research*, *36*, D966–D969.
Gupta, A., Rico-Medina, A., & Caño-Delgado, A. I. (2020). The physiology of plant responses to drought. *Science*, *368*, 266–269.
Haak, D. C., Kostyun, J. L., & Moyle, L. C. (2014). Merging ecology and genomics to dissect diversity in wild tomatoes and their relatives. In C. R. Landry & N. Aubin-Horth (Eds.), *Ecological genomics: Ecology and the evolution of genes and genomes* (pp. 273–298). Springer Netherlands.
Hämälä, T., Gorton, A. J., Moeller, D. A., & Tiffin, P. (2020). Pleiotropy facilitates local adaptation to distant optima in common ragweed (Ambrosia artemisiifolia). *PLoS Genetics*, *16*, e1008707.
Harb, A., Krishnan, A., Ambavaram, M. M., & Pereira, A. (2010). Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiology*, *154*, 1254–1271.
Jill Harrison, C. (2017). Development and genetics in the evolution of land plant body plans. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*, 20150490.
Josephs, E. B., Wright, S. I., Stinchcombe, J. R., & Schoen, D. J. (2017). The relationship between selection, network connectivity, and regulatory variation within a population of Capsella grandiflora. *Genome Biology and Evolution*, *9*, 1099–1109.
Jovelin, R., & Phillips, P. C. (2009). Evolutionary rates and centrality in the yeast gene regulatory network. *Genome Biology*, *10*, R35.
Juenger, T. E. (2013). Natural variation and genetic constraints on drought tolerance. *Current Opinion in Plant Biology*, *16*, 274–281.
Kahlon, P. S., Förner, A., Muser, M., Oubounyt, M., Gigl, M., Hammerl, R., Baumbach, J., Hückelhoven, R., Dawid, C., & Stam, R. (2023). Laminarin-triggered defence responses are geographically dependent in natural populations of *Solanum chilense*. *Journal of Experimental Botany*, *74*, 3240–3254.
Kashyap, S., Kumari, N., Mishra, P., Moharana, D. P., Aamir, M., Singh, B., & Prasanna, H. (2020). Transcriptional regulation-mediating ROS homeostasis and physio-biochemical changes in wild tomato (*Solanum chilense*) and cultivated tomato (*Solanum lycopersicum*) under high salinity. *Saudi Journal of Biological Sciences*, *27*, 1999–2009.
Kashyap, S. P., Prasanna, H. C., Kumari, N., Mishra, P., & Singh, B. (2020). Understanding salt tolerance mechanism using transcriptome profiling and de novo assembly of wild tomato *Solanum chilense*. *Scientific Reports*, *10*, 15835.
Kim, P. M., Korbel, J. O., & Gerstein, M. B. (2007). Positive selection at the protein network periphery: Evaluation in terms of structural constraints and cellular context. *Proceedings of the National Academy of Sciences*, *104*, 20274–20279.
Koenig, D., Jiménez-Gómez, J. M., Kimura, S., Fulop, D., Chitwood, D. H., Headland, L. R., Kumar, R., Covington, M. F., Devisetty, U. K., & Tat, A. V. (2013). Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proceedings of the National Academy of Sciences*, *110*, E2655–E2662.
Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*, 356.
Koubkova-Yu, T. C.-T., Chao, J.-C., & Leu, J.-Y. (2018). Heterologous Hsp90 promotes phenotypic diversity through network evolution. *PLoS Biology*, *16*, e2006450.
Kumar, R., Ichihashi, Y., Kimura, S., Chitwood, D. H., Headland, L. R., Peng, J., Maloof, J. N., & Sinha, N. R. (2012). A high-throughput method for Illumina RNA-seq library preparation. *Frontiers in Plant Science*, *3*, 202.
Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*, 559.
Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930.
Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 1–21.
Luisi, P., Alvarez-Ponce, D., Pybus, M., Fares, M. A., Bertranpetit, J., & Laayouni, H. (2015). Recent positive selection has acted on genes encoding proteins with more interactions within the whole human interactome. *Genome Biology and Evolution*, *7*, 1141–1154.
Mähler, N., Wang, J., Terebieniec, B. K., Ingvarsson, P. K., Street, N. R., & Hvidsten, T. R. (2017). Gene co-expression network connectivity is an important determinant of selective constraint. *PLoS Genetics*, *13*, e1006402.
Martínez, J. P., Antúnez, A., Araya, H., Pertuzé, R., Fuentes, L., Lizana, X. C., & Lutts, S. (2014). Salt stress differently affects growth, water status and antioxidant enzyme activities in *Solanum lycopersicum* and its wild relative *Solanum chilense*. *Australian Journal of Botany*, *62*, 359–368.
Masalia, R. R., Bewick, A. J., & Burke, J. M. (2017). Connectivity in gene coexpression networks negatively correlates with rates of molecular evolution in flowering plants. *PLoS One*, *12*, e0182289.
Mboup, M., Fischer, I., Lainer, H., & Stephan, W. (2012). Trans-species polymorphism and allele-specific expression in the CBF gene family of wild tomatoes. *Molecular Biology and Evolution*, *29*, 3641–3652.
Mi, H., Muruganujan, A., & Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*, *41*, 377–386.

Molitor, C., Kurowski, T. J., Fidalgo de Almeida, P. M., Kevei, Z., Spindlow, D. J., Chacko Kaitholil, S. R., Iheanyichi, J. U., Prasanna, H., Thompson, A. J., & Mohareb, F. R. (2024). A chromosome-level genome assembly of *Solanum chilense*, a tomato wild relative associated with resistance to salinity and drought. *Frontiers in Plant Science*, *15*, 1342739.

Mustafin, Z. S., Zamyatin, V. I., Konstantinov, D. K., Doroshkov, A. V., Lashin, S. A., & Afonnikov, D. A. (2019). Phylostratigraphic analysis shows the earliest origination of the abiotic stress associated genes in *A. thaliana*. *Genes*, *10*, 963.

Nakazato, T., Warren, D. L., & Moyle, L. C. (2010). Ecological and geographic modes of species divergence in wild tomatoes. *American Journal of Botany*, *97*, 680–693.

Nicolas, P., Shinozaki, Y., Powell, A., Philippe, G., Snyder, S. I., Bao, K., Zheng, Y., Xu, Y., Courtney, L., & Vrebalov, J. (2022). Spatiotemporal dynamics of the tomato fruit transcriptome under prolonged water stress. *Plant Physiology*, *190*, 2557–2578.

Nieri, D., Di Donato, A., & Ercolano, M. R. (2017). Analysis of tomato meiotic recombination profile reveals preferential chromosome positions for NB-LRR genes. *Euphytica*, *213*, 206.

Nosenko, T., Böndel, K. B., Kumpfmüller, G., & Stephan, W. (2016). Adaptation to low temperatures in the wild tomato species *Solanum chilense*. *Molecular Ecology*, *25*, 2853–2869.

Papakostas, S., Vøllestad, L. A., Bruneaux, M., Aykanat, T., Vanoverbeke, J., Ning, M., Primmer, C. R., & Leder, E. H. (2014). Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. *Nature Communications*, *5*, 4071.

Pertea, G., & Pertea, M. (2020). GFF utilities: GffRead and GffCompare. *F1000Res*, *9*, 304.

Piasecka, B., Lichocki, P., Moretti, S., Bergmann, S., & Robinson-Rechavi, M. (2013). The hourglass and the early conservation models—co-existing patterns of developmental constraints in vertebrates. *PLoS Genetics*, *9*, e1003476.

Quint, M., Drost, H.-G., Gabel, A., Ullrich, K. K., Bönn, M., & Grosse, I. (2012). A transcriptomic hourglass in plant embryogenesis. *Nature*, *490*, 98–101.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raduski, A. R., & Igić, B. (2021). Biosystematic studies on the status of *Solanum chilense*. *American Journal of Botany*, *108*, 520–537.

Reddy, A. R., Chaitanya, K. V., & Vivekanandan, M. (2004). Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. *Journal of Plant Physiology*, *161*, 1189–1202.

Rodrigues, J., Inzé, D., Nelissen, H., & Saibo, N. J. (2019). Source–sink regulation in crops under water deficit. *Trends in Plant Science*, *24*, 652–663.

Roselius, K., Stephan, W., & Städler, T. (2005). The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics*, *171*, 753–763.

Sato, M. P., Makino, T., & Kawata, M. (2016). Natural selection in a population of Drosophila melanogaster explained by changes in gene expression caused by sequence variation in core promoter regions. *BMC Evolutionary Biology*, *16*, 35.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498–2504.

Shen, W., & Ren, H. (2021). TaxonKit: A practical and efficient NCBI taxonomy toolkit. *Journal of Genetics and Genomics*, *48*, 844–850.

Silva Arias, G. A., Gagnon, E., Hembrom, S., Fastner, A., Khan, M. R., Stam, R., & Tellier, A. (2023). Contrasting patterns of presence-absence variation of NLRS within *S. chilense* are mainly shaped by past demographic history. *bioRxiv*. 2023.2010. 2013.562278.

Skirycz, A., & Inzé, D. (2010). More from less: Plant growth under limited water. *Current Opinion in Biotechnology*, *21*, 197–203.

Spivakov, M. (2014). Spurious transcription factor binding: Non-functional or genetically redundant? *BioEssays*, *36*, 798–806.

Städler, T., Arunyawat, U., & Stephan, W. (2008). Population genetics of speciation in two closely related wild tomatoes (Solanum section Lycopersicon). *Genetics*, *178*, 339–350.

Stuart, J. M., Segal, E., Koller, D., & Kim, S. K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science*, *302*, 249–255.

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, *34*, W609–W612.

Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M. B., Ross-Ibarra, J., Myers, C. L., Tiffin, P., & Springer, N. M. (2012). Reshaping of the maize transcriptome by domestication. *Proceedings of the National Academy of Sciences*, *109*, 11878–11883.

Tapia, G., Méndez, J., & Inostroza, L. (2016). Different combinations of morpho-physiological traits are responsible for tolerance to drought in wild tomatoes *Solanum chilense* and *Solanum peruvianum*. *Plant Biology*, *18*, 406–416.

Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J., & Gao, G. (2020). PlantRegMap: Charting functional regulatory maps in plants. *Nucleic Acids Research*, *48*, D1104–D1113.

Tournebize, R., Poncet, V., Jakobsson, M., Vigouroux, Y., & Manel, S. (2019). McSwan: A joint site frequency spectrum method to detect and date selective sweeps across multiple population genomes. *Molecular Ecology Resources*, *19*, 283–295.

Townsley, B. T., Covington, M. F., Ichihashi, Y., Zumstein, K., & Sinha, N. R. (2015). BrAD-seq: Breath adapter directional sequencing: A streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. *Frontiers in Plant Science*, *6*, 366.

True, J. R., & Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development*, *3*, 109–119.

Verelst, W., Bertolini, E., De Bodt, S., Vandepoele, K., Demeulenaere, M., Pè, M. E., & Inzé, D. (2013). Molecular and physiological analysis of growth-limiting drought stress in Brachypodium distachyon leaves. *Molecular Plant*, *6*, 311–322.

Wagner, G. P., Kin, K., & Lynch, V. J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, *131*, 281–285.

Wang, X., Feng, H., Chang, Y., Ma, C., Wang, L., Hao, X., Al, L., Cheng, H., Wang, L., & Cui, P. (2020). Population sequencing enhances understanding of tea plant evolution. *Nature Communications*, *11*, 4447.

Wang, Y., Wang, X., & Paterson, A. H. (2012). Genome and gene duplications and gene expression divergence: A view from plants. *Annals of the New York Academy of Sciences*, *1256*, 1–14.

Wei, K., Silva-Arias, G. A., & Tellier, A. (2023). Selective sweeps linked to the colonization of novel habitats and climatic changes in a wild tomato species. *New Phytologist*, *237*, 1908–1921.

Wilkinson, S., & Davies, W. J. (2010). Drought, ozone, ABA and ethylene: New insights from cell to plant to community. *Plant, Cell & Environment*, *33*, 510–525.

Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, *88*, 76–82.

Yang, J., Ren, Y., Zhang, D., Chen, X., Huang, J., Xu, Y., Aucapiña, C. B., Zhang, Y., & Miao, Y. (2021). Transcriptome-based WGCNA analysis reveals regulated metabolite fluxes between floral color and scent in *Narcissus tazetta* flower. *International Journal of Molecular Sciences*, *22*, 8249.

Yang, L., Bu, S., Zhao, S., Wang, N., Xiao, J., He, F., & Gao, X. (2022). Transcriptome and physiological analysis of increase in drought stress tolerance by melatonin in tomato. *PLoS One*, *17*, e0267594.

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*, *16*, 284–287.

Zarrineh, P., Sánchez-Rodríguez, A., Hosseinkhan, N., Narimani, Z., Marchal, K., & Masoudi-Nejad, A. (2014). Genome-scale co-expression network comparison across Escherichia coli and salmonella enterica serovar typhimurium reveals significant conservation at the regulon level of local regulators despite their dissimilar lifestyles. *PLoS One*, *9*, e102871.

Zinkgraf, M., Zhao, S. T., Canning, C., Gerttula, S., Lu, M. Z., Filkov, V., & Groover, A. (2020). Evolutionary network genomics of wood formation in a phylogenetic survey of angiosperm forest trees. *New Phytologist*, *228*, 1811–1823.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wei, K., Sharifova, S., Zhao, X., Sinha, N., Nakayama, H., Tellier, A., & Silva-Arias, G. A. (2024). Evolution of gene networks underlying adaptation to drought stress in the wild tomato *Solanum chilense*. *Molecular Ecology*, *00*, e17536. https://doi.org/10.1111/mec.17536