Horticulture Research

Horticulture Research, 2025, **12**: uhae313

https://doi.org/10.1093/hr/uhae313

Article

A full genome assembly reveals drought stress effects on gene expression and metabolite profiles in blackcurrant (*Ribes nigrum* L.)

Freya Maria Rosemarie Ziegler [1,2,3,*], Vivien Rosenthal[2,3], Jose G. Vallarino[4], Franziska Genzel[1], Sarah Spettmann[1], Łukasz Seliga[5], Sylwia Keller-Przybyłkowicz [5], Lucas Munnes[2], Anita Sønsteby[6], Sonia Osorio[4] and Björn Usadel [1,2,3,*]

[1]CEPLAS, Institute of Bio- and Geosciences (IBG-4: Bioinformatics) & Bioeconomy Science Center (BioSC), Forschungszentrum Jülich, Wilhelm-Johnen-Straße, D-52425 Jülich, Germany
[2]Faculty of Mathematics and Natural Sciences, CEPLAS, Institute for Biological Data Science, Universitätsstr. 1, D-40225 Düsseldorf, Germany
[3]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany
[4]Departamento de Biología Molecular y Bioquímica, Campus de Teatinos, Instituto de Hortofruticultura Subtropical y Mediterránea 'La Mayora', Universidad de Málaga-Consejo Superior de Investigaciones Científicas, 29010 Málaga, Spain
[5]Department of Plant Crop breeding, INHORT, National Institute of Horticultural Research, Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland
[6]NIBIO, Norwegian Institute of Bioeconomy Research, Division of Food Production and Society Horticulture, Pb. 115, NO-1431 Ås, Norway
*Corresponding authors. E-mail: f.ziegler@fz-juelich.de; b.usadel@fz-juelich.de

Abstract

Blackcurrant (*Ribes nigrum* L., family Grossulariaceae) is a perennial shrub that is widely cultivated for its edible berries. These are rich in antioxidants, vitamin C, and anthocyanins, making them a valuable ingredient in the food and beverage industry. However, prolonged periods of drought during the fruiting season lead to drought stress, which has serious ecological and agricultural implications, inhibiting blackcurrant growth and reducing yields. To facilitate the analysis of underlying molecular processes, we present the first high-quality chromosome-scale and partially haplotype-resolved assembly of the blackcurrant genome (cv. Rosenthals Langtraubige), also the first in the family Grossulariaceae. We used this genomic reference to analyze the transcriptomic response of blackcurrant leaves and roots to drought stress, revealing differentially expressed genes with diverse functions, including those encoding the transcription factors *bZIP*, *bHLH*, *MYB*, and *WRKY*, and tyrosine kinase-like kinases such as *PERK* and *DUF26*. Gene expression was correlated with the abundance of primary metabolites, revealing 14 with significant differences between stressed leaves and controls indicating a metabolic response to drought stress. Amino acids such as proline were more abundant under stress conditions, whereas organic acids were depleted. The genomic and transcriptomic data from this study can be used to develop more robust blackcurrant cultivars that thrive under drought stress conditions.

Introduction

Blackcurrant (*Ribes nigrum* L.) is a winter-hardy shrub (family Grossulariaceae) cultivated in temperate regions mainly for its edible berries. These are associated with several health benefits reflecting the high levels of vitamin C, anthocyanins, antioxidants, and flavonoids. However, many blackcurrant cultivars produce less floral biomass and abort a greater proportion of flowers in response to drought stress, reducing the commercial berry yield [3, 4].

Drought stress affects multiple biochemical and physiological processes at the cellular and whole-plant levels, including growth, membrane integrity, pigment synthesis, osmotic adjustment, water relations, and photosynthetic activity [4]. It also promotes senescence, which is necessary for plant survival. Leaf senescence requires the combined activity of phytohormones and transcription factors such as those of the *NAC*, *AP2/ERF*, and *WRKY* families, which are modulated by ethylene [5, 6]. One of the key phytohormones mediating drought stress responses and tolerance is abscisic acid (ABA), which regulates stomatal closure and the expression of stress response genes. Roots transmit water deficit information to the leaves, where ABA accumulates in the vasculature [7]. Drought stress is also regulated by calcium signaling, reactive oxygen species (ROS), and phytohormone translocation [8]. Furthermore, in response to drought, plants synthesize and accumulate osmolytes such as soluble carbohydrates, proteins, free amino acids, glycine betaine, and proline to increase the osmotic potential of their cells [9].

The analysis of drought stress responses using omics approaches has focused on model plants and major crops such as rice [10] and maize [11], and nonmodel species have also been studied such as *Solanum lycopersicoides* [12] and grapevine (*Vitis vinifera* [13]). So far, only two studies have considered drought stress in blackcurrant. Čereković *et al.* [3] compared irrigated and nonirrigated plants of the cultivars Narve Viking and Ben Gairn after 12 days of stress and 17 days of recovery, revealing that drought stress reduced the cumulative evapotranspiration

Received: 3 June 2024; **Accepted:** 6 November 2024; **Published:** 11 November 2024; **Corrected and Typeset:** 1 February 2025

of both cultivars during flowering due to stomatal closure and a smaller leaf area. The same authors also investigated the effect of drought on gene expression in the leaves of Ben Gairn plants [14], discovering more than 2000 differentially expressed genes (DEGs) during drought treatment based on a custom *Ribes* microarray [14]. The DEGs were found to belong to several transcription factor families, including *bZIP*, *WRKY*, *MYB*, and zinc finger. Additionally, genes associated with the cell wall and cell cycle regulation, gibberellin metabolism, the cytochrome P450 family, and the 2-oxoglutarate superfamily were identified. Although this provided first insights into the molecular basis of drought stress in blackcurrant, the specific functions of many genes remain unknown.

The blackcurrant transcriptome was analyzed in more detail by Russell *et al.* [2, 15] who identified transcriptome-based markers from leaf buds and mapped them to parental genotypes of a reference mapping population using 454 sequencing. Jarret *et al.* [16] evaluated blackcurrant fruits at different developmental stages for changes in gene expression and metabolite abundance, and used RNA-Seq data to generate a *de novo* transcriptome assembly. A transcriptome assembly based on blackcurrant fruit was published by Thole *et al.* [17], and the transcriptomic response of blackcurrant cultivar Aldoniai to blackcurrant reverse virus infection was studied by Mažeikienė *et al.* [18] and Juškytė *et al.* [19]. Although the blackcurrant chloroplast genome has been published [20], we currently lack a reference genome for the genus *Ribes* and even for the family Grossulariaceae [21].

Here we present a partially haplotype-resolved chromosome-scale genome assembly of blackcurrant (*R. nigrum* L. cv Rosenthals Langtraubige) based on a hybrid approach combining Oxford Nanopore Technologies (ONT) sequencing and PacBio HiFi long-read sequencing. We used the chromosome-scale genome, comprising eight pseudo-chromosomes, to investigate drought stress effects in Rosenthals Langtraubige leaves and roots at the transcriptomic and metabolomic levels. Our data will facilitate the development and cultivation of more drought-tolerant and climate-smart blackcurrant varieties, as well as other *Ribes* species including redcurrant and gooseberry.

Results

Ribes nigrum genome sequencing and assembly

The German cultivar Rosenthals Langtraubige was selected for the genomic analysis because the plant has a strong, spreading bushy structure, and its medium-sized, early-maturing berries have high vitamin C content. We combined PacBio HiFi and ONT nanopore sequencing with the Pore-C method for nanopore chromosome conformation capture. To this aim, we used hifiasm (UL) where we supplied the PacBio HIFI data for the initial string graph assembly and the nanopore data as 'ultralong' data, which hifiasm aligns to the string graph to finally produce an assembly [22]. By linking the contig sequences with the Pore-C data, we created a genome-scale assembly with a size of 871 555 738 bp (Supplemental Table S1), 99.03% of which (867 416 594 bp) was anchored to eight pseudo-chromosomes (Fig. 1, Supplemental Fig. S1).

We found that genes were more abundant at the chromosome ends but scarcer around the centromeres, where transposable elements (TEs) and CpG methylation were prevalent and the gas chromatography (GC) content was higher, a characteristic distribution found in many plant genomes (Fig. 1). The analysis of *k*-mers revealed a genome length of 800 Mb with 0.75% predicted heterozygosity, suggesting the blackcurrant genome 'Rosenthals Langtraubige' is relatively heterozygous (Supplemental Fig. S2).

Table 1. Assembly statistics for the *R. nigrum* genome at the pseudo-chromosomal level

Genome features	*R. nigrum* genome assembly
No. of chromosomes	8
Total length (bp)	867 416 594
GC content (%)	36.97
N per 100 kbp	0.16
k-mer completeness	98.60
QV	57.50
CRE(R-AQI) (%)	0.39 (96.13)
CRE(S-AQI) (%)	0.02 (97.87)
LAI	14.77

QV = quality value based on *k*-mers; LAI = LTR assembly index; R-AQI = regional assembly quality index; S-AQI = structural assembly quality index.

Genome quality evaluation using the *k*-mer analysis tool Merqury [23] revealed a high consensus-based quality value (QV) of 57.50 and a *k*-mer completeness of 98.59%, indicating a very high-quality and accurate genome assembly. Telomeric sequences were identified at the beginning of chromosomes 2 and 4 and at the ends of chromosomes 3, 5, 6, and 7. A telomere-to-telomere sequence was successfully assembled for chromosome 8 (Supplemental Table S2). Furthermore, our assembly was almost gap free (0.16 Ns/100 kbp) and the long terminal repeat (LTR) assembly index (LAI), which assesses the completeness of LTR sequences [24], was 14.77, aligning closely with reference genome values for *Arabidopsis thaliana* (Arabidopsis) and other species (Table 1).

In accordance with these metrics, the outcomes of CRAQ quality analysis [25] demonstrated minimal clip-based regional errors (CRE = 0.39) and negligible clip-based structural errors (CSE = 0.02). The calculated average values for the regional and structural assembly quality indicators (R-AQI = 96.13 and S-AQI = 97.87) surpassed the threshold of 90, signifying a reference-level assembly quality according to Li *et al.* [25]. Furthermore, the genome was validated by benchmarking universal single-copy orthologs (BUSCO) analysis [26], resulting in a gene completeness score of 98.3% and a BUSCO duplication rate of 3.9% for the whole genome assembly. These combined results confirm the very high quality of our *de novo* genome assembly.

We compared the *R. nigrum* genome assembly with available data by anchoring 73 single nucleotide polymorphisms (SNPs) to the linkage map of the blackcurrant SCRI 9328 population [2] (Fig. 2, Supplemental Table S3). Markers in the same linkage group were likewise found on the same chromosome in our genome assembly. Furthermore, the markers were largely collinear with the assembly. However, although markers in linkage group LG7 were collinear with our assembly, we only found them at the beginning of pseudo-chromosome 7 (Fig. 2).

TEs were identified using EDTA [24] and were found to represent 73.33% of the blackcurrant genome (Table 2). Approximately half of the identified TEs were LTRs, with Gypsy elements representing 30.35% of all TEs and Copia elements accounting for 9.03% (Table 2). Terminal inverted repeat (TIR) elements included Mutator, CACTA, PIF-Harbinger, Tc1_Marnier, hAT, and polintron. Non-LTR long interspersed nuclear elements (LINEs) accounted for only 0.27% of all TEs, whereas non-TIR helitron elements made up 2.96% of all TEs (Table 2).

The hybrid hifi assembly strategy, as previously outlined, also permitted the acquisition of haplotypes, with sizes of 900 Mbp for haplotype 1 and 885 Mbp for haplotype 2, and corresponding N50 values of 61 and 51 Mbp, respectively (Supplemental Table S1). Upon analyzing both haplotypes individually by total read

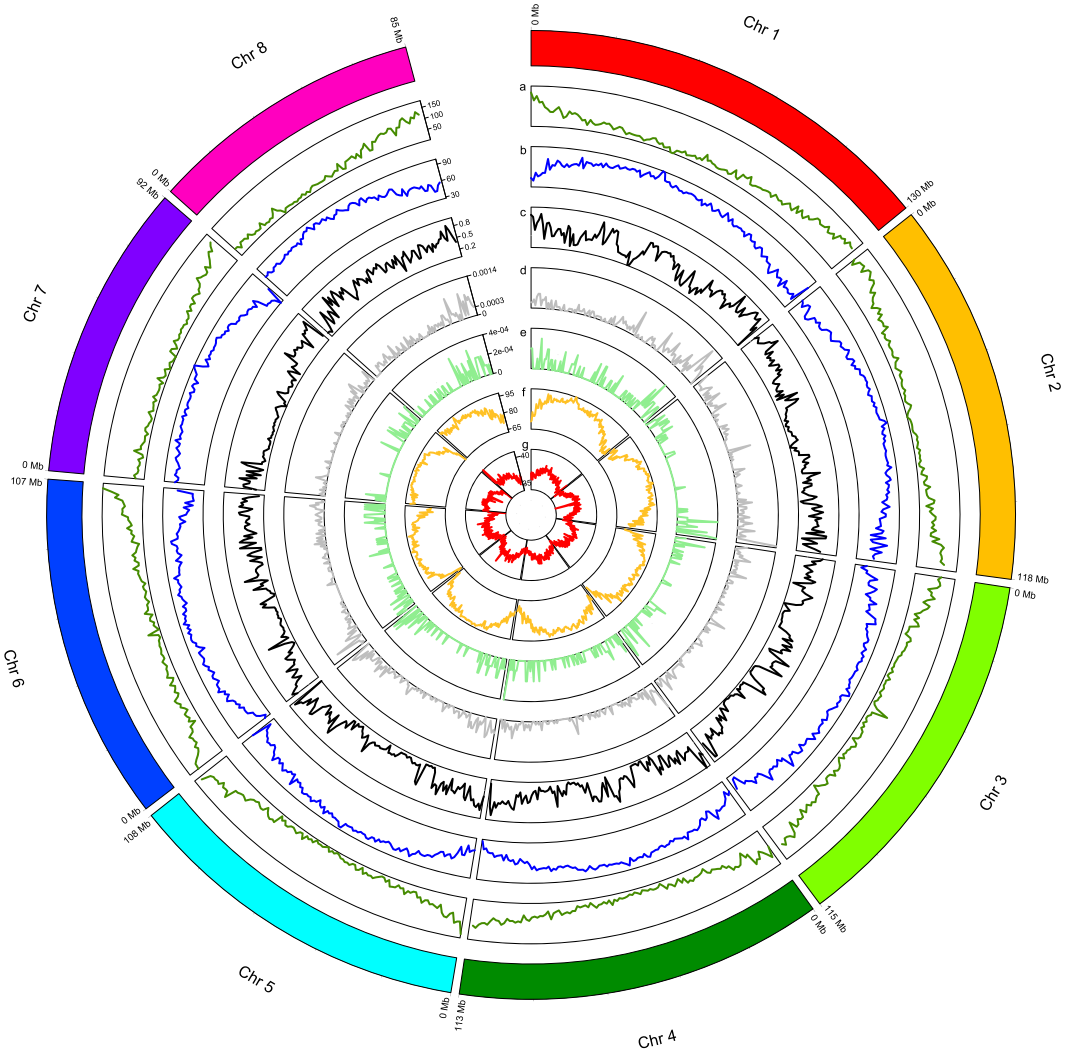

Figure 1. Chromosome-level characteristics of the *R. nigrum* genome. The genome is divided into eight chromosomes (Chr 1–8). (a) Gene density within 1-Mbp windows. (b) TE distribution within 1-Mbp windows. (c) Percentage of heterozygous SNP within 1-Mbp windows. (d) Percentage of heterozygous stop gained SNPs within 1-Mbp windows. (e) Percentage of heterozygous stop lost SNPs within 1-Mbp windows. (f) Percentage CpG methylation within 1-Mbp windows. (g) Percentage GC content within 1-Mbp windows.

Table 2. Repetitive DNA content of the scaffolded blackcurrant genome at the anchored pseudo-chromosome level

Class	Superfamily	Count	Masked sequences (bp)	Masked sequences (%)
LTR	Copia	93 001	78 309 531	9.03
	Gypsy	201 313	261 507 560	30.15
	unknown	194 422	94 960 513	10.95
TIR	CACTA	64 085	25 369 267	2.92
	Mutator	158 420	65 986 508	7.61
	PIF-Harbinger	34 197	11 702 040	1.35
	Tc1_Marnier	18 514	4 910 143	0.57
	hAT	55 797	19 738 629	2.28
	polintron	8	3314	0.00
Non-LTR	LINE	4848	2 379 098	0.27
Non-TIR	helitron	73 220	25 686 508	2.96
Repeat region		147 113	45 549 242	5.25
Total		1 044 938	636 102 353	73.33

k-mers, a *k*-mer completeness of 84.74% was achieved with a QV of 56.96 for haplotype 1, and a *k*-mer completeness of 84.84% was achieved with a QV of 58.12 for haplotype 2. This result corroborated the relatively high heterozygosity, in accordance with the 0.75% heterozygosity indicated by Genoscope analysis (Supplemental Fig. S2). Due to these results, we assessed SNPs between the two different haplotypes by calling heterozygous SNPs both using long and short read data as well as structural variants with the long read data only. This resulted in 2 176 479 SNPs and 64 644 heterozygous structural variants,

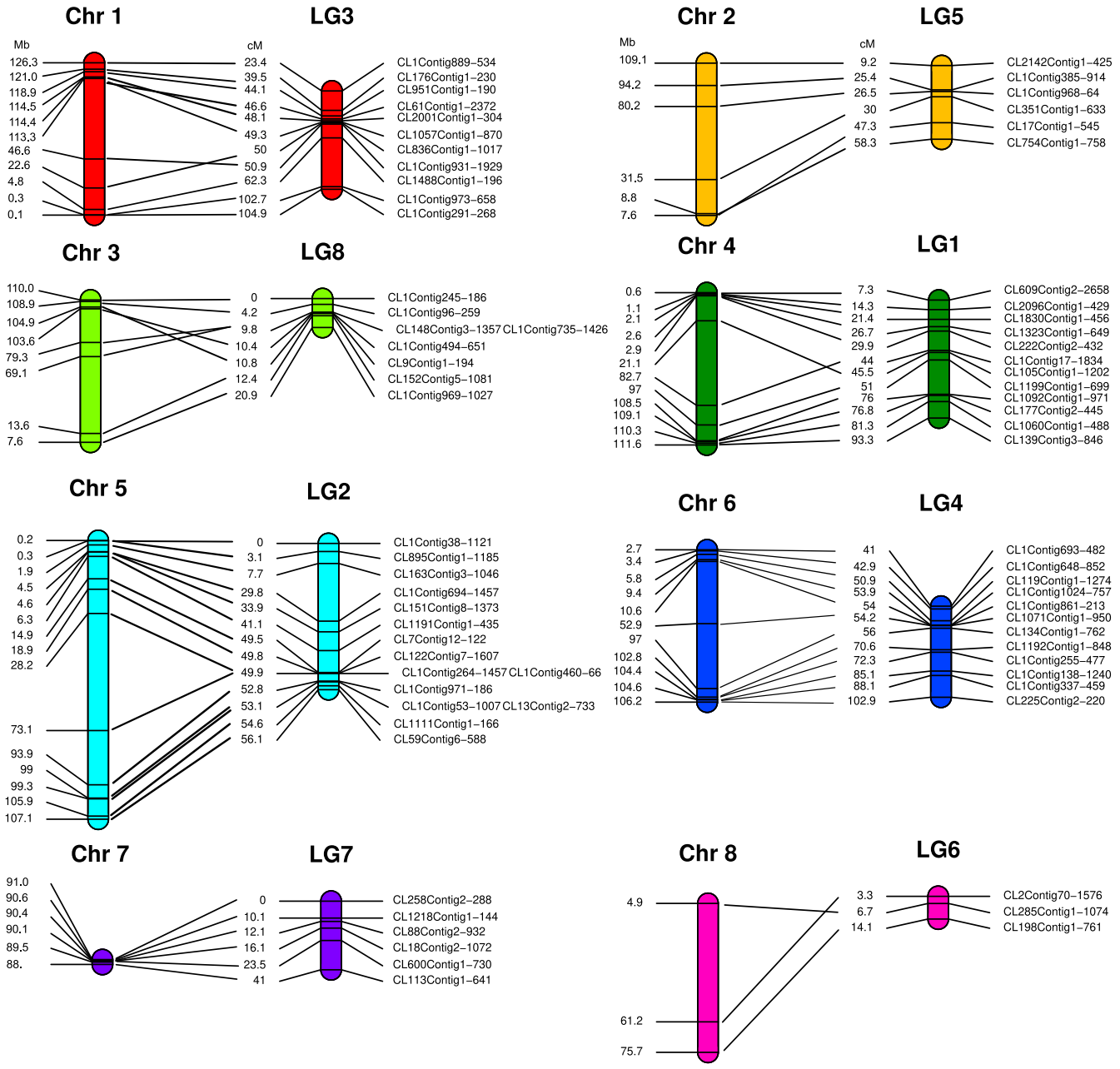

Figure 2. Assignment of 73 genetic markers to the *R. nigrum* reference genome. The 73 markers from the genetic map of *R. nigrum* population SCRI 9328 [2] were assigned to eight linkage groups (LG1–8), which were largely collinear with our eight pseudo-chromosomes (Chr 1–8).

respectively. Interestingly, we were able to detect somewhat lower heterozygosity in the inner parts of chromosome 7 (Fig. 1c). We further assessed the small variants and found that 1256 variants were identified that introduce premature stop codons, while 373 variants were found to result in the conversion of stop codons into nonstop codons (Fig. 1d, e; Supplemental Table S5). SNPs were functionally annotated, and variants causing stop codons showed a high enrichment in secondary metabolism [false discovery rate (FDR) < 0.05; Supplemental Table S6], while photosynthesis and phytohormone action are depleted (FDR < 0.001; Supplemental Table S6).

To make these data more accessible, we scaffolded the two haplotypes to the complete genome assembly using RagTag, enabling a simple chromosome-scale haplotype representation (Supplemental Fig. S3).

Ribes nigrum genome annotation

Genome annotation, using a comprehensive pipeline that integrated *ab initio* and evidence-based approaches, predicted 42 380 gene sequences, 86.20% of which were functionally annotated using Mercator [27]. The quality of the resulting gene models was validated by BUSCO analysis, yielding a completeness score of 97.6% of the proteome (Supplemental Table S4).

Drought stress-induced genes

To shed light on the response of blackcurrant leaves and roots to drought stress in mature plants, water was withheld for 4 days. The treated plants showed typical external signs of a stress response, including curled and dry leaves (Fig. 3A–C). In contrast, control plants with continuous irrigation over the same period showed no change in their appearance (Fig. 3D–F).

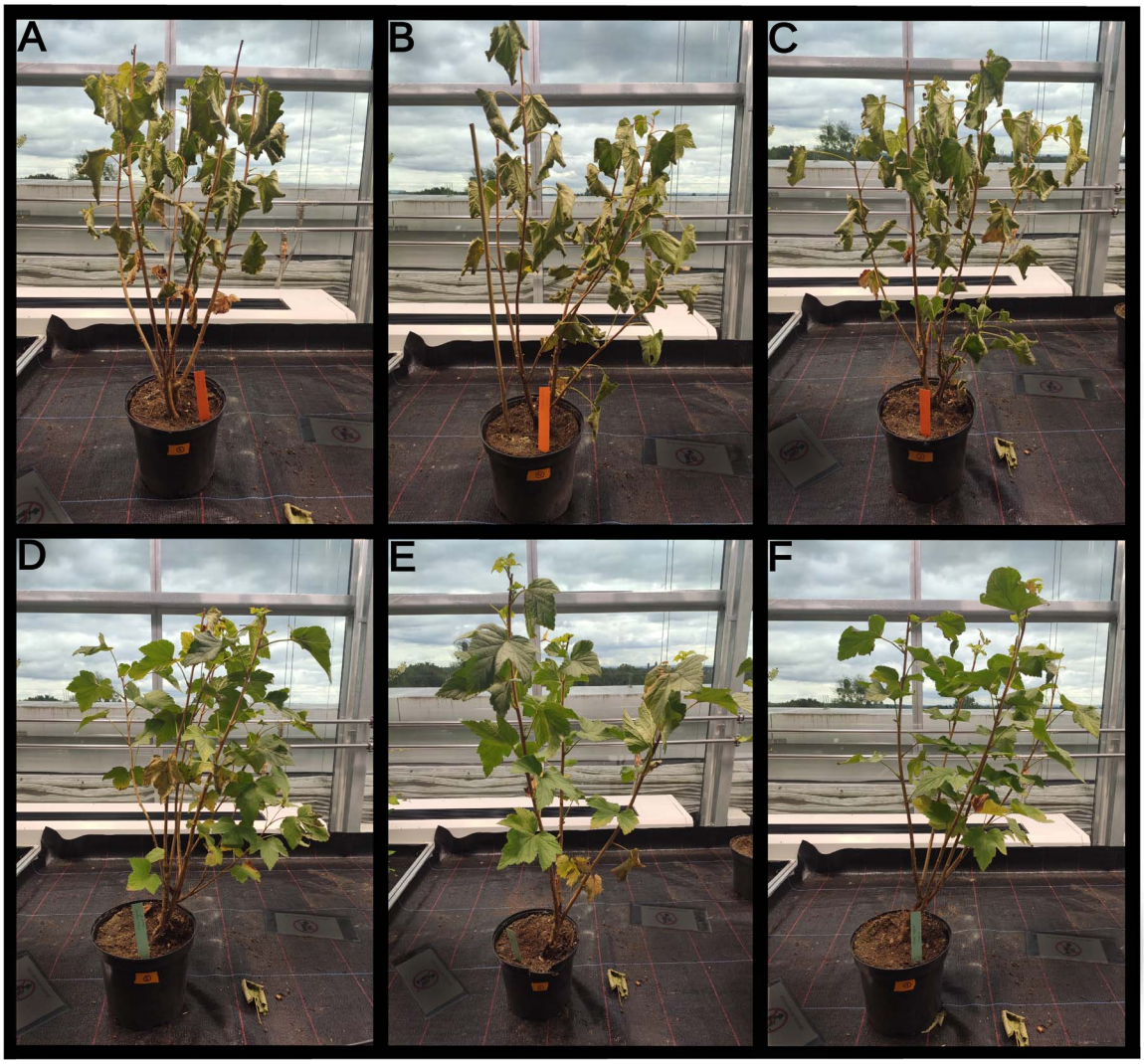

Figure 3. Drought stress treatment of *R. nigrum* cv Rosenthals Langtraubige plants. (A–C) Stressed plants after 4 days without water. (D–F) Control plants with continuous irrigation.

To identify differentially expressed transcripts in leaf and root samples of adult plants, nanopore-based RNA-Seq results were used. In the leaf samples, 540 genes were identified as DEGs (FDR < 0.1), 48 of which were upregulated and 492 downregulated, as well as 8353 DEGs in the roots even with much more stringent filtering (FDR < 0.05), 3697 of which were upregulated and 4656 downregulated.

To corroborate these data, we performed a second experiment on younger plants where we again withheld water (Supplemental Fig. S4). Here, 7891 DEGs were identified in leaf samples using the FDR threshold of <0.05. Of these, 3642 genes were upregulated, while 4249 genes were downregulated.

From the 540 DEG identified in the first experiment, 80 genes (FDR < 0.05) exhibited consistent regulation patterns.

Impact of drought stress on protein kinases

In the adult plants, we identified 243 genes encoding leucine-rich repeat (LRR) protein kinases, mainly representing families I–XV (Supplemental Table S5). For the young plants, 206 of these LRR encoding transcripts were found, showing similar expression results for most transcripts. Most LRR protein kinases were downregulated in leaves (Supplemental Table S7). However, three LRR-XV genes were slightly upregulated in leaves, and 16 LRR-XII gene transcripts were strongly downregulated ($log_2FC < -2$) in leaves (Supplemental Table S7). Here, we could find 10 of these transcripts downregulated in young plants. In the roots, most LRR protein kinases were downregulated (FDR < 0.05).

Another member of the LRR family, the receptor-like cytoplasmic kinases (RLCK) multiple transcripts were of RLCK-IXb, V, VI, VIIa, VIIb, VIII, X, XI, XII, and XV found differentially expressed. Six RLCK-V transcripts were significantly downregulated in roots of adult plants (FDR < 0.01). Three transcripts of RLCK-VI (RN5G010110.1, RN1G011660.1, RN1G053710.1) were significantly downregulated in roots and have shown a strong downregulation in leaves in adult plants ($Log_2FC < -2$). This expression pattern was also observed in leaf tissue in young plants (Supplemental Table S7). The downregulation was also identified for RN1G032180.1 (RLCK-X) and RN3G051940.2 (RLCK-XII) for both tissues in adult plants and was significant for roots. The RLCK-IXb gene RN4G001630.1 was significantly upregulated in roots (FDR < 0.05) and leaf tissue in young plants (FDR < 0.05) (Supplemental Table S7).

Several DEGs encoded components of the MAPK cascade. Three MAPK kinases (NPK/ANP) were upregulated in leaves

but significantly downregulated in roots (FDR < 0.05) (Supplemental Table S5). In contrast, the transcript encoding MAP2K (RN3G007350.1) was slightly upregulated and MAP3K-MEKK (RN5G010030.1) was strongly downregulated ($log_2FC < -2$) in leaves and RN3G007350.1 was slightly upregulated in roots (Supplemental Table S7). In young plants, these transcripts were slightly regulated but not significantly so. Ca^{2+}-dependent protein kinase (CDPK) transcripts were mostly upregulated in leaves and roots, with RN1G025100.1 significantly upregulated in roots (FDR < 0.01). In contrast, two CDPK transcripts (RN1G043640.1 and RN5G008650.1) were significantly downregulated in roots and RN1G043640.1 was downregulated in leaves. One DOMAIN OF UNKNOWN FUNCTION 26 (DUF26) transcript was significantly (FDR < 0.05) upregulated in roots and two others were significantly downregulated (Supplemental Table S7). In young plants, the transcript RN8G019550.2 (DUF26) was significantly upregulated (FDR < 0.01), which showed the same expression in leaf and root tissue of adult plants.

Five transcripts encoding proline-rich extensin-like receptor kinases (PERKs) were downregulated in leaves, including RN4G048640.1, RN7G025360.1, and RN7G034400.1, which were also significantly downregulated in roots (FDR < 0.01). In contrast, RN3G015760.1 was significantly upregulated in roots (FDR < 0.01) (Supplemental Table S7). Transcripts in young plants in leaf tissue showed the same expression pattern as roots in adult plants, and RN3G015760.1 was also significantly upregulated (FDR < 0.01).

Transcription factors and metabolism

Transcripts encoding *AP2/ERF*, *bZIP*, *CBF/DREB1*, *DREB2*, *C2H2*, *HD-ZIP*, *R2R3-MYB*, *WRKY*, *HD-ZIP*, *bHLH*, and *NAC* transcription factors were differentially expressed in both tissues of adult plants. Transcripts associated with carotenoid and ABA metabolism, including those encoding PYL/RCAR and abscisic aldehyde oxidase, were generally downregulated, and the PYL/RCAR transcripts RN3G031060.1, RN4G014070.1, and RN4G030230.1 were significantly (FDR < 0.01) downregulated in the roots (Supplemental Table S7). In young plants in leaf tissue, RN3G031060.1 showed the same significant downregulation (FDR < 0.001) as in roots in adult plants. Several transcripts encoding SnRK1 and SnRK2 SNF1-related protein kinases were modulated in both tissues in adult plants and also in leaf tissue in young plants, with significant differential expression in the roots. Several transcripts associated with cell wall regulation were identified in the leaves. Genes encoding *p-coumaroyl shikimate/quinate 3-hydroxylase* were modulated in leaves and roots (Supplemental Table S5). RN6G009060.1, encoding hydroxyproline-*O*-galactosyltransferase (GALT29), was significantly downregulated in roots and slightly downregulated in leaves. Transcripts encoding caffeoyl-CoA 3-*O*-methyltransferase were significantly downregulated in roots but were expressed at minimal levels in leaves (Supplemental Table S7).

Relative expression of drought stress induced through qRT-PCR

To validate the results of drought-stress induced genes, a total of 13 DEGs, comprising six upregulated and seven downregulated genes in leaf tissue of adult plants, were selected for qRT-PCR. The analysis was conducted using actin and eIF4A as endogenous controls. The results indicated that the selected genes exhibited comparable gene expression fold changes (FCs) between quantitative qRT-PCR and computational analyses in both leaf and root tissues of mature plants (Supplemental Fig. S5B). Additionally, qRT-PCR was conducted for the same genes, excluding 4HbD and WRKY, which could not be reliably quantified in the leaf tissue of young plants. Of these 11 genes, the expression of eight genes were consistent with those observed in the leaf tissue of drought-stressed adult plants, further validating the selected candidate genes involved in the *R. nigrum* response to drought stress (Supplemental Fig. S5A, B).

Table 3. Metabolites in blackcurrant leaf tissue that change in abundance in response to drought stress

Metabolite	Log_2FC	FDR
Amino acids		
Alanine	1.09	0.025*
Valine	2.23	0.004**
Isoleucine	2.57	0.005**
Proline	2.67	0.015*
Serine	0.92	0.056
Threonine	2.05	0.003**
GABA	1.28	0.017*
Methionine	2.31	0.015*
Glutamic acid	2.12	0.024*
Phenylalanine	2.72	0.009**
Glutamine	2.86	0.006**
Tryptophan	3.45	0.004**
Organic acids		
Quinic acid	0.82	0.017*
Citric acid	−1.39	0.047*
Sugars and derivatives		
Galactinol	0.62	0.011*

Upregulation (red) and downregulation (blue) is shown according to the log_2FC of the genes. Significance levels of adjusted *P* values (FDR) are shown in bold.

Metabolite changes due to drought stress

Next, we compared the abundance of 60 primary metabolites in response to drought stress in leaves and roots of mature plants. Several were barely affected, but 14 leaf metabolites became significantly more abundant under stress ($log_2FC > 1$, FDR < 0.05). Most of these metabolites were amino acids and their derivatives such as threonine, valine, proline, isoleucine, methionine, glutamine, glutamic acid, tryptophan, phenylalanine, alanine, and 4-aminobutanoic acid (GABA) (Table 3, Supplemental Tables S6 and S8). Sugars such as mannose, sugar derivatives such as galactinol, and the organic acid quinic acid also accumulated in response to drought, whereas citric acid was the only metabolite that was significantly depleted.

Few metabolic changes were observed in the roots. However, mannose and isoleucine accumulated in response to drought stress ($log_2FC > 1.3$), whereas β-alanine was depleted ($log_2FC < -1.3$). We also observed a slight increase in the levels of valine, isoleucine, and methionine ($log_2FC > 0.5$), and a slight decrease in the levels of citric acid ($log_2FC < -0.5$) (Supplemental Fig. S7). Although these metabolites showed minor changes according to their log_2FC values, the adjusted *P* value suggested that the changes were not significant. In the leaf tissue of young plants, similar patterns were observed for valine, isoleucine, proline, threonine, GABA, and phenylalanine compared to the leaf tissue of adult plants. Additionally, valine, proline, GABA, and phenylalanine exhibited significant upregulation (FDR > 0.05), while alanine, serine, and glutamic acid showed a slight downregulation (Supplemental Table S9).

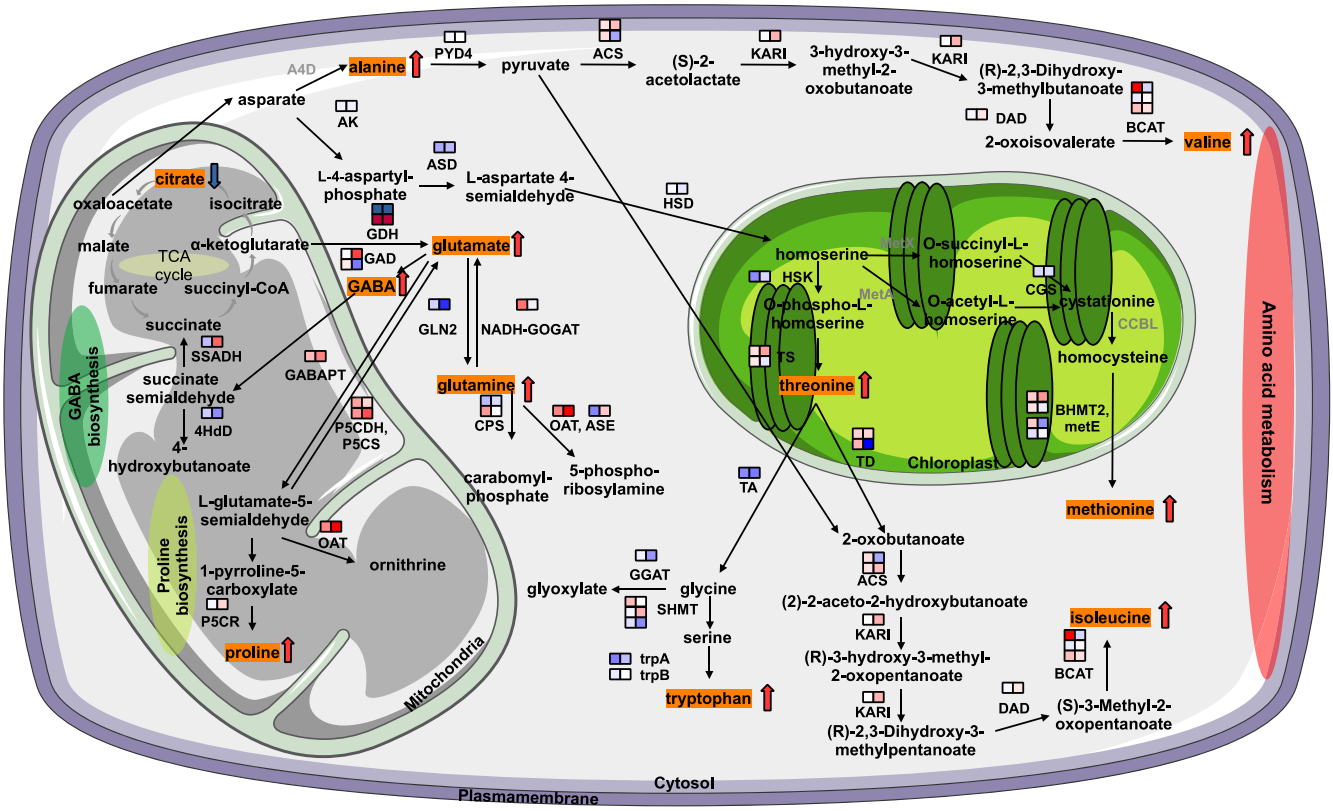

Figure 4. The connection between DEGs and metabolic pathways that significantly alter metabolite levels under drought stress in the leaves and roots of *R. nigrum*. DEGs encoding enzymes (annotated using MapMan) are aligned with specific metabolic reactions (absent = bold gray). Upregulated genes are denoted by red squares and downregulated genes by blue squares, with squares on the left and right indicating leaves and roots, respectively. Multiple rows of squares denote multiple genes with the same enzymatic functions (log_2FC of DEGs and FDRs listed in Supplemental Table S2). Metabolites differing in abundance (Table 3 and Supplemental Tables S1 and S2) under drought stress are highlighted in orange, with accumulation shown as red arrows and depletion as blue arrows. Abbreviations: 4HbD = bifunctional γ-hydroxybutyrate dehydrogenase; ACLY = ATP citrate (pro-S)-ligase; ACO = aconitase; ACS = acetolactate synthase; AK = aspartate kinase; ASAT = aspartate transaminase; ASD = aspartate-semialdehyde dehydrogenase; ASE = amidophosphoribosyltransferase; A4D = aspartate 4-decarboxylase; BCAT = branched-chain-amino-acid transaminase; BHMT2 = homocysteine S-methyltransferase; CCBL = kynurenine-oxoglutarate transaminase; CGS = cystathione-γ-synthase; CPS = large subunit of carbamoyl phosphate synthetase heterodimer; DAD = dihydroxy-acid dehydratase; GABAPT = GABA pyruvate transaminase; GAD = glutamate decarboxylase; GAT = GABA transporter; GDH = glutamate dehydrogenase; GGAT = glycine transaminase; GLN1 = cytosolic glutamine synthetase; GLN2 = plastidial glutamine synthetase; HSD = homoserine dehydrogenase; HSK = homoserine kinase; IDH3 = isocitrate dehydrogenase heterodimer subunit 1; KARI = ketol-acid reductoisomerase; metA = homoserine O-succinyltransferase; metE = 5-methyltetrahydropteroyltriglutamate homocysteine methyltransferase; metX = homoserine O-acetyltransferase; NADH-GOGAT = NADH-dependent glutamate synthase; OAT = ornithine-γ-aminotransferase; P5CDH = [1]Δ-pyrroline-carboxylate dehydrogenase; P5CR = pyrroline-carboxylate-dehydrogenase; P5CS = [1]Δ-pyrroline-carboxylate synthase; PYD4 = alanine aminotransferase; SHMT = serine hydroxymethyltransferase; SSADH = succinate-semialdehyde dehydrogenase (NAD+); TA = threonine aldolase; TD = threonine dehydratase; trpA = subunit α of tryptophan synthase complex; trpB = subunit β of tryptophan synthase complex; TS = threonine synthase.

Connecting DEGs with metabolites in response to drought stress

The integration of data from the analysis of DEGs and metabolites revealed associations involving the tricarboxylic acid (TCA) cycle, GABA biosynthesis, cysteine and methionine metabolism, proline metabolism, alanine and aspartate metabolism, as well as branched-chain amino acid metabolism. We also found associations between groups of altered metabolites in leaves (FDR < 0.05) and between groups of DEGs in both tissues (Fig. 4).

In the context of the TCA cycle and the depletion of citric acid, all related transcripts in leaves in adult plants were strongly upregulated (Supplemental Table S8) with the exception of *ACLY* (RN7G026170.1), which was downregulated in leaves and roots. In young plants, this transcript showed a similar significant downregulation (FDR < 0.01). GABA-associated genes were mainly upregulated in both tissues in adult and young plants. Notably, one gene encoding a bifunctional γ-hydroxybutyrate dehydrogenase (RN1G038820.1) was significantly downregulated in roots, while another (RN3G012700.1) showed no significant differential expression in roots. *Glutamate decarboxylase* genes were significantly upregulated (RN8G004390.1) or downregulated (RN6G037990.1) in roots but were not expressed at significant levels in leaves. The *succinate-semialdehyde dehydrogenase* gene RN8G041850.1 was significantly upregulated in roots but showed no expression in young plants in leaf tissue (Supplemental Table S8). Genes involved with alanine, aspartate, and glutamate metabolism were generally downregulated in both tissues in adult and young plants (except *ASAT*, demonstrated upregulation). Furthermore, genes encoding *CPS* and *NADH-GOGAT* were observed to be upregulated in leaves, while a threonine aldolase gene demonstrated a significant downregulation in the roots (Supplemental Table S8). All transcripts associated with valine, leucine, and isoleucine metabolism showed significantly upregulated in adult plants. One *threonine dehydratase* gene (RN6G010130.1) exhibited a marked decrease in roots (Supplemental Table S8). In young

Table 4. Differentially expressed transcripts involved in proline metabolism and their MapMan annotations

Gene ID	MapMan annotation	Enzyme Name	Leaf (log_2FC)	Leaf (FDR)	Root (log_2FC)	Root (FDR)
RN8G016110.1	[25]Δ-Pyrroline-carboxylate dehydrogenase	P5CDH	0.72	0.13	0.31	0.23
RN1G008890.1	Glutamate dehydrogenase ∗(GDH)	GDH	−3.66	0.12	−1.17	0.08
RN1G008920.1	Glutamate dehydrogenase ∗(GDH)	GDH	−5.77	0.11	−1.52	0.05
RN1G008960.1	Glutamate dehydrogenase ∗(GDH)	GDH	−2.73	0.23	−1.69	**0.03**∗
RN1G008980.1	Glutamate dehydrogenase ∗(GDH)	GDH	−1.82	0.35	−1.89	**0.03**∗
RN3G012920.1	Glutamate dehydrogenase ∗(GDH)	GDH	0.43	0.27	0.22	0.38
RN3G035860.1	Glutamate dehydrogenase ∗(GDH)	GDH	0.35	0.57	−1.68	**0.008**∗∗
RN4G025340.1	Glutamate dehydrogenase ∗(GDH)	GDH	0.51	0.82	1.12	0.08
RN2G017990.1	Ornithine aminotransferase	OAT	0.92	0.43	1.94	**0.0004**∗∗∗
RN1G015520.1	Proline dehydrogenase	PRODH	0.38	0.56	−0.26	0.57
RN1G019530.1	Pyrroline-5-carboxylate reductase	P5CR	−0.19	0.58	0.33	0.15
RN2G007260.1	Pyrroline-5-carboxylate reductase	P5CR	−0.10	0.78	0.25	0.22
RN3G051060.1	[25]Δ-Pyrroline-carboxylate synthase	P5CS	0.83	0.61	1.38	**0.007**∗∗
RN4G032560.1	[25]Δ-Pyrroline-carboxylate synthase	P5CS	0.81	0.68	1.35	**0.005**∗∗

Upregulation (red) and downregulation (blue) reflect the log_2FC values, with significance (FDR adjusted *P* values) shown in bold.

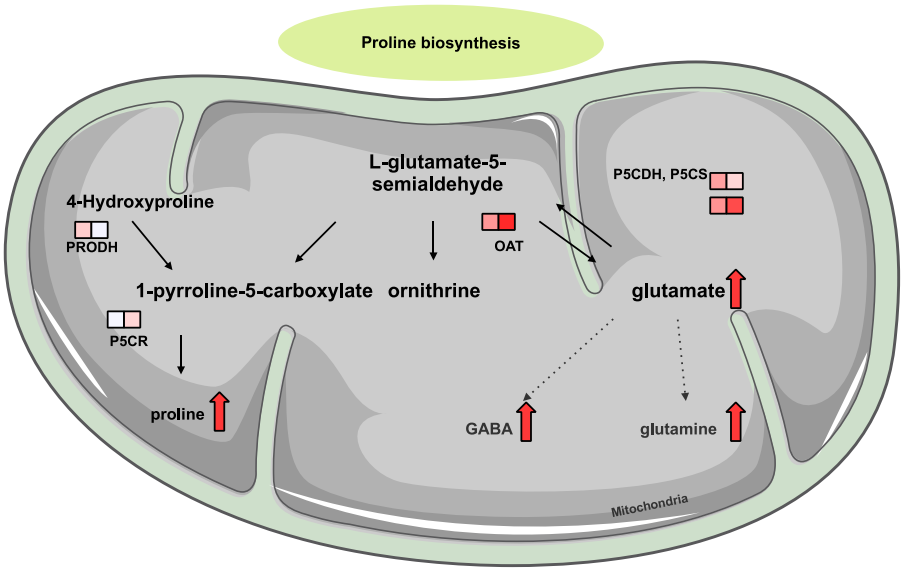

Figure 5. Proline biosynthesis metabolites with significantly altered levels under drought stress in the leaves of *R. nigrum*. DEGs (Table 4) encoding enzymes (annotated using MapMan) are aligned with specific metabolic reactions. Upregulated genes are denoted by red squares and downregulated genes by blue squares, with squares on the left and right indicating leaves and roots, respectively. Multiple rows of squares denote multiple genes with the same enzymatic functions. Metabolites that accumulate under drought stress (Table 4 and Supplemental Tables S1 and S2) are highlighted with a red arrow. Abbreviations: OAT = ornithine-γ-aminotransferase; P5CDH = [1]Δ-pyrroline-carboxylate dehydrogenase; P5CR = pyrroline-carboxylate dehydrogenase; P5CS = [1]Δ-pyrroline-carboxylate synthase; PRODH = proline dehydrogenase.

plants, most of these transcripts were downregulated except RN4G055970.1 and RN2G016250.1, which were upregulated (Supplemental Table S8).

Transcripts in adult plants involved in proline metabolism were mostly upregulated in leaves, although several glutamate dehydrogenase transcripts were downregulated and the same genes were significantly (FDR < 0.05) downregulated in roots (Table 4, Fig. 5). These results were also found in leaf tissue in young plants, except for some glutamate dehydrogenase, which have shown no expression. For transcripts of P5CDH (RN8G016110.1), OAT (RN2G017990.2), P5CR (RN2G007260.1), and P5CS (RN4G032560.1), expression in young plants was similarly regulated as in leaf tissue of young plants, while OAT, P5CR, and P5CS were significantly differentially expressed (FDR < 0.05). Transcripts encoding ornithine-γ-aminotransferase and two [1]Δ-pyrroline-carboxylate synthases were significantly upregulated in roots of adult plants (FDR < 0.01).

Discussion

Ribes nigrum genome assembly

We have generated a blackcurrant genome assembly at the pseudo-chromosomal level, the first published genome representing the genus *Ribes* or the family Grossulariaceae, and have identified multiple genes associated with drought stress as well as related changes in metabolites. The total size of the anchored assembly was ~871 Mbp, which is close to the estimated genome size of ~1 Gbp for the genus *Ribes* [21] and in line with the Genoscope estimation of 750 Mbp (Supplemental Fig. S2). We used a hybrid approach that integrates ONT long reads and precise HiFi PacBio data directly in hifiasm, as well as Pore-C ONT contact data, thus coupling long reads directly to multiway chromatin contacts without amplification [28]. This allowed the construction of pseudo-chromosomal scaffolds representing eight chromosomes, building on the complementarity of simplex ONT and precise HiFi reads [29, 30]. Our data indicate that

blackcurrant is a diploid genome with a basic chromosome number of 8, in line with other *Ribes* species [21, 31]. Given the relatively high heterozygosity of cv Rosenthals Langtraubige, we generated contiguous individual assemblies for haplotypes 1 and 2, with sizes of 900 and 885 Mbp, respectively, and scaffolded these against the haploid assembly using RagTag. While this later step introduces several switching errors, it might provide a useful representation for breeding and analysis purposes. Structural variant analysis identified a substantial number of stop codons across the genome, suggesting a high frequency of gene disruption and an elevated mutational load.

The quality of the genome assembly was checked using the LAI, which estimates completeness by calculating the percentage of intact LTR retrotransposon sequences [32]. The LAI of the blackcurrant genome was 14.77, similar to the score of 14.9 reported for Arabidopsis [24, 33]. Long-read sequencing methods like ONT and PacBio typically yield LAI scores >10, which designates a genome of reference quality [34]. This was supported by the assessment of errors using CRAQ, which revealed an R-AQI of 96.21% and an S-AQI of 98.64%, with scores >90% indicating reference quality [25]. The AQI scores for the blackcurrant genome were comparable to those of the PacBio complete long reads (CLR) rice haplotype MH63 (R-AQI = 96.25; S-AQI = 94.96) (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_001623365.2 [25]) and the PacBio CLR rice haplotype R498 (R-AQI = 96.96; S-AQI = 98.22) [25, 35] at the chromosomal level. Our results surpassed the existing ONT assemblies of *Solanum pennellii* (R-AQI ≤ 95.57; S-AQI = 91.47) and Arabidopsis KM74 (R-AQI = 85.95; S-AQI = 93.34), as well as the PacBio CLR Arabidopsis KM74 assembly [33] at the contig level (R-AQI = 95.57; S-AQI = 96.29) (Table 1 [25]). These distinctions probably reflect variations in sequencing methods and TE identification. Notably, PacBio HiFi error rates are generally lower than those of simplex ONT reads, yielding more accurate assemblies. Our R-AQI values align closely with the PacBio reference genomes of rice and Arabidopsis.

The QV of our assembly assessed using Merqury was 57.49, which corresponds to 99.999% accuracy [23]. This improves upon the PacBio MH63 gap-free rice reference genome [36], which achieved a QV of 55.05, and the chromosome-level assembly based on PacBio and Hi-C data of the berry *Fragaria pentaphylla* [37], which achieved a QV of 54.76. Furthermore, the *k*-mer completeness of our genome was 98.59, which is higher than the PacBio and Hi-C *de novo* assemblies of *F. pentaphylla* and *Arctostaphylos glauca*, with *k*-mer completeness values of 84.71 and 74.39, respectively [37, 38], and similar to the 98.87 score assigned to the PacBio haplotype-resolved assembly of *Coriaria nepalensis* [39]. To make our assembly useful for quantitative trait loci (QTL) studies, we mapped SNP marker sequences from the linkage map of population SCRI 9328 and MP7 [2]. Genetic markers in specific linkage groups were also identified on corresponding chromosomes in our assembly. This demonstrates concordance between the genetic and physical maps and shows that our assembly is accurate and can be used to compare legacy data with our genome assembly.

The high quality of our sequence assembly was also confirmed by the BUSCO completeness score of 98.3%. This surpasses the scores of chromosomal-level genomes in the Saxifragaceae family, which is part of the Saxifragaceae alliance alongside the Grossulariaceae [40]. Three Saxifragaceae genomes have been published at the chromosomal level. The largest (12 Gbp, five chromosomes) is *Paeonia ostii*, which achieved a BUSCO completeness score of 94.4%, and the smallest (412 Mbp) is *Tiarella polyphylla*, which achieved a BUSCO completeness score of 97.31% [41, 42]. The *Cercidiphyllum japonicum* genome (719 Mbp, 19 chromosomes) achieved a score of 94.1%. The *T. polyphylla* assembly is based on a combination of Illumina short reads, ONT long reads, and Hi-C data, whereas the assembly of the *C. japonicum* genome is based on PacBio long reads, Illumina short reads, and Hi-C data. The higher BUSCO completeness score of our assembly may reflect the synergistic integration of PacBio and ONT reads with ONT Pore-C data. Our high BUSCO completeness score also surpasses the 97%–98% range reported for other berry crops, such as the blueberry wild relative *Vaccinium darrowii* with a BUSCO score of 97.5% [43], and raspberry (*Rubus idaeus*) with a BUSCO score of 91.3% [44]. These genomes were assembled using PacBio, Hi-C, and Illumina data, explaining the slightly lower BUSCO scores compared to our assembly. Our assembly, anchored on eight pseudo-chromosomes, successfully mapped 99% of the genome, whereas the comparable *C. japonicum* genome only anchored 90.18% on 19 pseudo-chromosomes [45]. Our annotation allowed the prediction of 42 380 genes with a BUSCO completeness score for the transcriptome of 96.8%. In comparison, Thole *et al.* [17] assembled 145 906 transcripts (N50 = 1480 bp) with a read mapping rate of 90.4% for cv Ben Hope. Another 40 225 genes were predicted by microarray analysis, but the data are not publicly available for comparison [14].

DEGs in leaf and root tissue during drought stress

Protein kinases

The genome assembly and annotated transcriptome were used to identify DEGs in leaf and root samples from drought-stressed plants and controls. During drought stress, plants close their stomata to minimize water loss, promote deeper root growth to enhance water acquisition, and shed leaves to reduce transpiration. The perception of drought stress involves the transduction of signals via receptor-like kinases [46]. Accordingly, we identified many DEGs encoding LRR kinases, most of which were significantly downregulated in roots and leaves under drought conditions. The LRR-III gene RN5G011100.1 was downregulated in both leaf and root tissues, a finding further validated by qRT-PCR. This supports the hypothesis that LRR-III plays a role in the drought stress response of *R. nigrum*. Besides that, one gene annotated as a member of the subfamily of LRR-XV was upregulated in leaves and significantly downregulated in roots. Ksouri *et al.* [47] previously observed a similar trend in *Prunus persica* rootstocks, noting the downregulation of LRR kinase genes particularly in leaves. To date, LRR and RLK (receptor-like kinase) proteins are known to play crucial roles in the perception of environmental signals and the initiation of responses to various biotic and abiotic stresses. However, LRR-III has not been directly associated with drought stress response in plants [48].

We also identified DEGs encoding RLCKs that form complexes responsible for intracellular signaling to control plant growth, development, and stress responses [49]. RLCK-VII in particular is closely associated with plant growth and development, suggesting its significant role in signal transduction [50]. Several Arabidopsis RLCK-VII proteins are involved in brassinosteroid signaling and root hair growth as well as root development and stomatal defenses (Mara-Garcia *et al.*, 2004). Furthermore, RLCK-VIII is associated with oxidative stress signaling [51, 52]. Members of the RLCK-XII family, such as BSK1, are essential for flg22-induced maximal ROS production, a key component of immune signalin [51, 53]. Additionally, RKS1 and the NLR (nucleotide-binding LRR) protein ZAR1 work together to initiate immune responses, contributing to plant defense mechanisms in Arabidopsis [54, 55]. We

observed the modulation of RLCK genes representing classes VII, VIII, and XII in both tissues in adult and young plants. Furthermore, we detected DEGs encoding RLCK classes V, VI, X, XI, and XII that have not been extensively characterized but appear to be involved in responses to drought stress in both tissues in adult, and were strongly downregulated in the roots. Additionally, we found these genes in leaf tissue in young plants as well, validating our findings of RLCK genes under drought stress in *R. nigrum* under drought stress.

We identified six PERKs expressed in leaf and root tissues, one of which was upregulated in roots but downregulated in leaves. Arabidopsis PERK4 is predicted to be a key regulator of Ca^{2+} signaling, contributing to the production of ABA in roots [56]. Furthermore, several ZmPERK genes are induced in response to stress [57]. The cysteine-rich repeat receptor-like kinase family (CRK) is typically characterized by the presence of two DUF26 protein interaction domains involved in sensing of environmental signals and the regulation of plant development [58, 59].

Hormone-related genes

Pyrabactin resistance (PYR), pyrabactin like (PYL), and regulatory component of ABA receptors (RCAR) are major negative regulators of the ABA signaling cascade if no ABA is available [60]. We found that most PYL/RCAR components were downregulated in leaves and roots of adult plants. This result was also found in young plants in leaf tissue for one PYL/RCAR component. These proteins form a complex with SnRK1, SnRK2, and PP2C. ABA can bind to PYR/PYL/RCAR and activate the SnRK1, SnRK2, and SnRK3 kinases. Several SnRK3 (data not shown) and SnRK2 genes were upregulated in our study, and a gene encoding the regulatory FCS-like zinc finger (FLZ) of the SnRK1 complex was downregulated, suggesting the ABA pathway is activated in response to drought stress. The involvement of this pathway and the corresponding genes in abiotic stress responses has already been shown in Arabidopsis [61], barley [62], pea [63], and fava bean [64]. The same genes also regulate development, growth and stomatal movement [65].

Transcription factors

Drought stress influenced the expression of many stress-responsive genes encoding transcription factors in leaves and/or roots. The ERF-I, II, IV, IX, VI, VIII, X, and DREB1/2 genes were upregulated in both tissues in adult and young plants, and were previously shown to regulate drought stress responses [66]. CBF/DREB1 and DREB2 mediate drought tolerance in rice, soybean, and Arabidopsis [66, 67]. We likewise found that most CBF/DREB1 and DREB2 genes were upregulated under drought stress. Čereković *et al.* [14] showed that three bZIP transcription factors were modulated in *R. nigrum* leaves under drought stress, which were names according to similar Arabidopsis genes: *AT3G58120* (upregulated), as well as *AT5G24800* and *AT4G35900* (downregulated). We found that multiple bZIP genes were upregulated in both tissues. The difference between the studies may reflect the use of different cultivars and/or annotated reference genomes. The IRE1/bZIP60 gene and corresponding regulatory mediators were strongly downregulated in roots. IRE1/bZIP60 is part of the unfolded protein response.

R2R3-MYB, bHLH, WRKY, NAC, zinc finger (C2H2), and *HD-ZIP* transcription factors are also involved in drought stress responses [68–73]. R2R3-MYB and bHLH proteins have been linked to the anthocyanin pathway and ABA-dependent regulation in *R. nigrum* under drought stress [1, 14, 74, 75]. Čereković *et al.* [14] showed that MYB genes were downregulated under drought stress, which is consistent with our results for *R. nigrum* R2R3-MYB genes. Furthermore, we found that WRKY, NAC, C2H2, and HD-ZIP genes were modulated in *R. nigrum* leaves and roots of adult and young plants. The *R. nigrum* WRKY33 gene was reported to be induced in leaves under drought stress [3] and mediates the drought stress response in *Caragana korshinskii* [71]. Plants expressing CkWRKY33 were more likely to survive a period of drought and featured higher levels of soluble sugar, proline, and peroxidase activity. We found that several NAC genes were differentially expressed in *R. nigrum* leaves and with highly significant differences in the roots, in agreement with previous findings [76, 77]. NAC proteins play a key role in drought responses, activating stress-responsive genes or enhancing drought resistance in transgenic plants [76, 77]. Čereković *et al.* [3] found that several C2H2 genes were downregulated in *C. korshinskii* leaves and roots, whereas we detected both upregulated and downregulated C2H2 genes in both tissues. C2H2 transcription factors influence ROS scavenging in rice by regulating the expression of genes encoding antioxidants and the synthesis of osmoprotectants such as proline and soluble sugars [68, 78, 79]. They are also involved in ABA signaling by modulating genes associated with stomatal closure and root development in response to drought [79]. C2H2 transcription factors also regulate the MAPK signaling pathway [80–82]. Finally, we observed the modulation of HD-ZIP genes in both tissues. The corresponding transcription factors are known to regulate growth, development, drought stress responses, ABA signaling, and stomatal closure [83–86].

DEGs that influence the abundance of metabolites

We identified multiple DEGs related to ABA, GABA, and proline biosynthesis, as well as general amino acid metabolism, and we also found that several of the corresponding metabolites were affected. Interestingly, most of these metabolic differences were restricted to the leaves, whereas previous studies have shown changes in roots too, for example, in peanut and various Triticeae species (Hu *et al.*, 2016 [87, 88]). The differences between our study and the previous reports may be species dependent or may reflect the relatively short period of drought stress we applied.

The main metabolic differences we observed between stressed leaves and controls involved the accumulation of amino acids, which serve as osmolytes, antioxidants, signaling mediators, and sources of energy [89, 90]. Proline and GABA accumulated in the stressed blackcurrant leaves in adult and young plants. Proline is an osmolyte and chaperone that accumulates to high levels in many dehydrated plants [76, 91–94]. GABA increases leaf turgor and osmolyte synthesis, regulates antioxidant levels, and increases water use efficiency and drought tolerance by influencing guard cell behavior [95]. Genes associated with these pathways were also found among the DEGs (Table 2, Fig. 3). Glutamine is another osmoprotectant that accumulates under drought stress [89]. It is synthesized from aspartate and converted to glutamate, which are involved in the transduction of Ca^{2+}, ROS, and electric signals, and it is also a precursor of proline, asparagine, arginine, and GABA [74, 96–99]. These metabolites, which play a key role in nitrogen metabolism, also accumulated in our drought-stressed leaves, and the corresponding regulatory genes were also differentially expressed (Table 2 and Fig. 3).

Citric acid is an organic acid within the TCA cycle that can be utilized for amino acid or GABA biosynthesis via the production of glutamate [100]. Drought stress can disrupt the TCA cycle by reducing oxygen availability and respiration, leading to alterations in citric acid levels [101]. We found that citric acid was

depleted in stressed leaves, in agreement with previous studies (Ashraf *et al.*, 2018). Quinic acid, which was slightly elevated in the stressed leaves, is known to be involved in drought stress responses, acting as a precursor for the synthesis of phenolic compounds and osmolytes [73, 102].

The amino acids threonine, methionine, isoleucine, and valine also accumulated under drought stress. Threonine and methionine are precursors of isoleucine and valine, which accumulate as a response to osmotic stress in Arabidopsis and other plants [91, 103–105]. These amino acids regulate osmosis and reduce damage caused by ROS (Singh *et al.*, 2015; Bohnert and Jensen, 1996). The OsDIAT gene encodes a regulator of branched-chain amino acid accumulation in rice [106]. In our study, the corresponding gene RN4G055970.1 (*BCAT*) was upregulated in both tissues of adult plants and in leaf tissue of young plants, which may underpin the accumulation of isoleucine and valine (Table 2, Fig. 4).

The aromatic acids phenylalanine and tryptophan, which accumulated in our stressed plants, are precursors of secondary metabolites such as indole acetate, lignin (via the shikimate pathway), and flavonoids, all of which play important roles in drought stress tolerance [107] and plant growth [108]. Phenylalanine levels were previously shown to increase during dehydration and decrease after dehydration in *Barbacenia purpurea* [107]. Tryptophan is the main precursor of auxin, which regulates vacuolar osmotic pressure and the translocation of metabolites, thus improving osmotic balance and growth under drought stress [109]. Tryptophan has been reported to accumulate under drought stress in other plants, which is consistent with our results [110].

The accumulation of sugars in plants during drought stress is an important adaptive strategy that supports survival during drought and recovery afterward [111]. In our study, the galactinol content of the leaves increased under drought stress, which has been previously reported in other plants [112, 113]. Sucrose, glucose, and fructose also contribute to drought stress tolerance, but we did not observe the accumulation of these soluble sugars in blackcurrant. As discussed above, this may reflect the relatively short duration of stress in our experiments, which may have been insufficient for the breakdown of storage carbohydrates into monosaccharides and disaccharides. In the context of sugar metabolism, we identified an arabinogalactan protein, GALT29/GALT29A, which was significantly downregulated across all tissues in both young and adult plants under drought stress (Supplemental Fig. S9). In *A. thaliana*, AtGALT29A encodes a putative glycosyltransferase that belongs to the carbohydrate-active enzyme family GT29. AtGALT29A exhibits β-1,6-galactosyltransferase activity, elongating β-1,6-galactan side chains and forming 6-Gal branches on the β-1,3-galactan backbone of arabinogalactan proteins [114]. So far, GALT29A has not been described as being involved in the drought stress response. However, our results strongly suggest that it plays a role in both leaf and root tissues of *R. nigrum* under drought conditions.

Conclusion

We have constructed a blackcurrant (*R. nigrum*) haplotype-resolved *de novo* genome assembly at the pseudo-chromosomal level using a combination of ONT long reads, Illumina short reads, and ONT Pore-C sequencing data. The resulting high-quality assembly (confirmed by BUSCO scores, *k*-mer completeness, and AQI values) is the first for the genus *Ribes* and indeed for the family Grossulariaceae. Transcriptome analysis revealed many DEGs involved in drought stress mechanisms, which were connected to changes in metabolite levels, particularly involving the TCA cycle, amino acid metabolism, and the biosynthesis of proline and GABA. These integrated findings enhance our understanding of *R. nigrum* stress-related genes and metabolites. These findings were observed in both adult and young plants, confirming the validity of our results regarding the drought stress response of *R. nigrum* in leaf tissues. However, more specific analysis (targeted metabolomics) is needed to establish precise relationships spanning the genome, transcriptome, and metabolome. The annotated blackcurrant genome presented in this study lays the foundation for further research to determine the molecular basis of drought stress responses in this important fruit crop. Our results can also be combined with existing genetic resources such as quantitative trait loci to facilitate marker-assisted selection and genetic engineering for the development of improved blackcurrant cultivars.

Materials and methods

Plant material

For DNA sequencing, young blackcurrant (*R. nigrum* cv Rosenthals Langtraubige) leaves were collected from a single adult plant and flash frozen in liquid nitrogen. For the drought stress experiments, plants in 3-l pots were irrigated with 150 ml of water per day (control) or water was withheld for 4 days (stress treatment). Root and leaf samples (three replicates per plant) were taken the morning after the last treatment day. Blackcurrant cv Tihope plants were grown under field conditions in Poland. Samples were taken at different developmental stages: flower initiation, dormancy induction, and dormancy release. Bud samples were collected at nine different time points, whereas leaf samples were taken at three time points during flower initiation. The samples were used for RNA deep sequencing on the Illumina HiSeq 2500 platform. Blackcurrant RNA-Seq datasets of cv Ben Gairn were generated from 20 bud samples on the same sequencing platform.

Stem cuttings of *R. nigrum* cv. 'Rosenthals Langtraubige' were cultivated in 500-ml pots for 5 weeks under controlled greenhouse conditions. For the drought stress experiment, plants were initially irrigated until the soil reached full water-holding capacity. Subsequently, plants in the control group were watered to maintain 70% of pot weight based on water loss, while those in the stress treatment group received no water for a period of 9 days. Leaf samples were collected the morning following the final day of treatment. RNA was extracted as described below, and samples were used for RNA deep sequencing on the Illumina HiSeq 2500 platform.

DNA library construction and sequencing

DNA was extracted from 1 g of frozen leaf material using the NucleoBond HMW DNA Kit (Machery Nagel, Thermo Fisher Scientific, Waltham, MA, USA). Short fragments were separated using the Circulomics short read eliminator (SRE XL) Kit (Pacific Biosciences, Melno Park, CA, USA) with a threshold of >40 kbp. DNA degradation and contamination were monitored by 1% agarose gel electrophoresis. DNA purity was checked using a NanoDrop spectrophotometer (Thermo Fisher Scientific), and the DNA concentration was determined using the Qubit DNA Assay Kit and a Qubit fluorometer (Thermo Fisher Scientific). Filtered long fragments were used to prepare ONT libraries for sequencing (Qubit DNA Assay Kit and a Qubit fluorometer). We used the standard ONT protocol for the SQK-LSK112 Kit. Genomic DNA fragments were repaired and 3′-adenylated using the NEBNext FFPE DNA

Repair Mix and the NEBNext Ultra II End Repair/A-Tailing Module (New England Biolabs/NEB, Ipswich, MA, USA). Sequencing adapters provided by ONT were then ligated using the NEBNext Quick Ligation Module (NEB). After purification with AMPure XP beads (Beckmann Coulter, Brea, CA, USA), libraries were loaded onto primed 10.4.1 Spot-On Flow Cells and sequenced using a PromethION device (ONT) with 72-h runs. Cells were flushed and reloaded after 18 h. Basecalling was achieved using guppy v6.1.1 (ONT). PacBio sequencing of the ONT long-read sequences was carried out by Genohub (Brigham Young University, Provo, UT, USA). Filtered short-read sequences were sequenced by Genewiz (Leipzig, Germany).

cDNA library construction and sequencing

Total RNA was extracted from leaves and roots of adult plants of treated and control plants using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations, with minor changes. Briefly, 50 mg of frozen tissue from each plant was lysed (1.5 M NaCl, 2% CTAB, 30 mM EDTA pH 8, 100 mM Tris-Cl, 2% β-mercaptoethanol in RNase-free water) for 10 min at 70°C. Further extraction and on-column digestion with DNase I was carried out according to Qiagen recommendations. RNA quality was monitored as described above for DNA. For cDNA sequencing, 200 ng of RNA per sample was used in a strand-switching step followed by cDNA amplification by PCR using primers containing 5′ tags according to the ONT protocol for cDNA PCR sequencing (SQK-PCS111). We then added rapid sequencing adapters and loaded 25–55 ng onto primed 9.4.1 Spot-On Flow Cells for sequencing using a PromethION device.

Analysis of gene expression of selected drought-responsive genes by qRT-PCR

Based on the results of the RNA-Seq analysis, genes responsive to the drought stress treatment were selected for validation of their expression in leaves and roots of adult plants and leaves in young plants (Supplemental Table S11). Actin and eIF4A [19, 115] were selected for normalization. For cDNA synthesis, 1 μg of RNA was reverse transcribed into cDNA using iScript cDNA Synthesis kit (Bio-Rad) in a 20-μl reaction following the manufacturer's instructions. Quantitative real-time PCR of selected genes was conducted with a CFX Opus 384 Real-Time PCR System (Bio-Rad). The amplification was carried out in a reaction mixture containing 5-μl iQ SYBR Green Supermix (Bio-Rad), 1 μl of primers (mix of forward and reverse primer, 500 nM), and 4-μl cDNA (diluted 20-fold) in triplicate for each sample and gene. The PCR program started with polymerase activation at 95°C for 3 min, followed by 39 cycles of cDNA denaturation at 95°C for 10 s, annealing and amplification at 60°C for 30 s, and a melt curve analysis at 65–95°C with an increment of 0.5°C and 5 s per step at the end.

The relative gene expression and FC were determined using the threshold cycle value (C_t) and efficiency-based $2^{-\Delta\Delta C}{}_T$ method described by Ganger *et al.* [116], with a modification where $\log_2$ transformation was applied instead of $\log_{10}$. The reaction efficiencies were calculated using real-time PCR miner [117] averaged for each gene for tissue and treatment. Actin and eIF4A were employed as internal reference genes for normalization. The log2 FCs ($\log_2$FC) were visualized in a heatmap, depicting the upregulation and downregulation of the analyzed genes. Data processing and visualization were performed in R v4.2.2, utilizing the RColorBrewer v1.1–3, dplyr v1.1.4, and EnrichedHeatmap v1.28.1 packages.

Pore-C library preparation and nanopore sequencing

RE-Pore-C libraries were prepared according to the ONT Plant Pore-C protocol using the restriction enzyme DpnII. The enzyme was heat denatured after overnight incubation and cross-linked DNA clusters were ligated in proximity. The protein was then degraded to remove cross-links, releasing the chimeric Pore-C dsDNA polymers. DNA concentration and purity were determined as described above. Samples with a low $A_{260/230}$ ratio (indicating the presence of polysaccharides) were cleaned up by 1:1 phenol/chloroform extraction and ethanol precipitation. Genomic DNA fragments were damage repaired, end repaired, and A-tailed using NEBNext FFPE DNA Repair Mix and the NEBNext Ultra II End Repair/A-Tailing Module. Libraries were constructed using the ONT Ligation Sequencing Kit (SQK-LSK114) and were sequenced using two R10.4.1 PromethION flow cells. The runtime was set to 100 h in accurate speed mode (260 bps).

Genome assembly

After basecalling, any remaining adapters were removed from the genomic DNA using Porechop v0.2.4 [118] and the longest reads (> 50 kbp) were selected using Filtlong v2.9.1 (https://github.com/rrwick/Filtlong) and setting '—min_length 50 000'. The filtered ONT reads were assembled with the PacBio single-molecule real-time (SMRT) reads using hifiasm v0.19.6-r595 [22, 119] in hybrid mode where the nanopore data were supplied with setting '–ul' and an estimated genome size of 0.8 Gbp. For *k*-mer analysis, we used Jellyfish v2.2.10 (https://github.com/gmarcais/Jellyfish) and GenomeScope (http://qb.cshl.edu/genomescope/) with default parameters ($k = 23$). PurgeHaplotigs v1.1.2 [120] was used to remove duplicate sequences. The Pore-C sequences were virtually digested using pore-c-py 'digest' (https://github.com/epi2me-labs/pore-c-py) and aligned with the previously generated genome using minimap2 v2.25 [121]. Aligned files were annotated using pore-c-py 'annotate' in paired-end mode (with settings '–monomers–chromunity–chromunity_merge_distance-1–paired_end–filter_pairs–paired_end_minimum_distance 100–paired_end_maximum_distance 200') and duplicates were removed using picard v2.21.8 with setting 'MarkDuplicates' (https://github.com/broadinstitute/picard) and default parameters. The aligned Pore-C data were then used by yahs scaffolder v1.1 [122] to scaffold the contig assembly, followed by manual curation using Juicebox Assembly Tools v1.11.08 [123]. Quality was assessed using Quast v4.6.3 [124] and BUSCO v5 [26] to provide information on assembly completeness and genome quality.

Finally, the two resulting haplotypes constructed by hifiasm were scaffolded to the haploid genome representation using RagTag v2.1.0 [125] using correct and scaffold with default parameters to obtain complete chromosome scale haplotypes. Synteny and structural rearrangements were visualized using Syri v1.7.0 [126] with default parameters.

SNPs were called using Freebayes v1.3.6 [127] with DNA short reads and default parameters for diploid organisms. SNPs were then annotated using SNPEff v5.2c [128]. Furthermore, SNPs were functionally annotated using MapMan protein classes, and an enrichment analysis based on these classes was conducted using R Studio (version 2023.06.0, [129]).

Genome analysis

Methylated site probabilities from mapped PacBio reads were generated using pb-CpG-tools v2.3.2 (https://github.com/PacificBiosciences/pb-CpG-tools) with a threshold of 75% for methylated motifs and

were plotted with gene density, GC content and TE abundance, heterozygous SNP rate, stop gained, and stop lost SNP Codons in 1-Mb windows using the R circlize package v0.4.15 (Gu *et al.*, 2014). Furthermore, we used the EDTA [24] pipeline and LTR_retriever 'LAI' v2.9.8 [130], CRAQ v1.0.9α [25], and Meryl v1.4 followed by Merqury v1.3 [23] for quality validation using the Pacbio reads. We used blastn in BLAST v2.15.0+ [131] to compute the synteny of 73 SNPs common to the SCRI 9328 and MP7 linkage map [2] and our genome assembly. Quality parameters were visualized using R Studio base v2023.06.0 [129]. The scaffolded genome assembly was annotated using Helixer v0.3.1 [132] with the model 'land_plant_v0.3_a_0080.h5' and StringTie v2.2.1 [133] with RNA-Seq Illumina short-reads and cDNA nanopore RNA-Seq reads. Splice junctions were detected with Portcullis v1.2.0 (https://github.com/EI-CoreBioinformatics/portcullis) and analyzed with the Helixer and StringTie annotation results in Mikado v2.3.0 [134]. The transcriptome and proteome were extracted from the annotated genome using Gffread v0.12.4 [135]. The resulting protein sequences were used in Mercator and MapMan4 v5.0 [27] for functional gene annotation. Transcriptome quality was assessed using BUSCO v5.

Evaluation of the drought stress experiment

The cDNA long reads were mapped onto the transcriptome assembly using Minimap2 assembler v2.25 (settings '-ax splice -ts') [121]. The read count was quantified using Salmon v1.10.1 [136]. Paired-end Illumina short reads for leaf samples of young plants were pseudo-mapped and quantified using Salmon v1.10.1. DEGs were identified using edgeR v3.40.1 [137], and we calculated the *P* values and adjusted *P* values (FDR). The counts per million (CPM) data were evaluated by principal component analysis (PCA). Furthermore, overrepresentation analysis based on MapMan annotation was used to find drought stress-related pathways in leaves and roots, and candidate genes were then selected.

Metabolite analysis

Metabolites in leaf and root samples of adult plants and leaf samples of young plants were analyzed by gas chromatography-masspectrometry (GC–MS) [138, 139]. Mass spectra were compared with the Golm Metabolome database for identification [140]. Metabolomic data were $\log_2$ transformed, and the average values were calculated for control and treated tissues. The data were normalized against the average values of the control samples. To assess the significance of differences, we calculated the $\log_2$ FC ($\log_2$FC), *P* values (t-test), and adjusted *P* values (FDR, Benjamini-Hochberg). The statistical tests were used to detect significant differences between the control and stressed samples for each metabolite. The data were visualized using enhanced volcano plots.

Data integration

A database was built using MariaDB v10.6.5 to establish a connection between the transcriptome and metabolome. Protein sequences representing land plant genes with known functions were extracted from the UniProt SWISSPROT database (May 2022) and linked to MapMan protein classes. These annotated genes were supplemented with associated enzymatic reactions from the RHEA database (https://www.rhea-db.org/). The substrates and products of each reaction were linked to their CHEBI (https://www.ebi.ac.uk/chebi/) and PubChem (https://pubchem.ncbi.nlm.nih.gov/) IDs. We then used the hierarchical organization of MapMan bins, which categorize major biological processes as top-level bins and their subprocesses as child bins, to establish connections between the enzymatic reactions and their associated metabolites. This was achieved by identifying closely related DEGs that associated with these processes and their corresponding bins. By linking the DEGs and their MapMan protein functions to the corresponding metabolites, we established a comprehensive understanding of the relationship between gene expression, biological processes, and metabolites. Genes associated with metabolites were filtered by FDR < 0.05 and pathway-related genes were selected. Statistical analysis was conducted using R Studio base v2023.06.0 [129].

Acknowledgements

This study was funded by the European Union's Horizon 2020 program under grant agreement number 679303. We thank Dr. Elisa Senger and Dr. Richard M Twyman for manuscript proofreading.

Author contributions

F.M.R.Z. performed ONT sequencing, data analysis, and manuscript writing, supported and supervised by B.U. S.O. and JGV performed metabolite measurement, and V.R. helped with ONT PoreC sequencing. L.M. provided RNA-Seq data of the drought stress experiment. F.G. and S.S. performed qPCR of selected candidate genes. A.S., L.S., and S.K.P. provided RNA-Seq data for genome annotation. All authors read and approved the final manuscript.

Data availability

Processed data are available in the supplementary data including read data summaries (S12) and raw sequencing data, and the genome will be available on EMBL-EBI (PRJEB77865). In addition, we added in the supplementary data the genomes and their annotation to the PLANTdataHUB [141] https://git.nfdi4plants.org/usadellab/ribes_nigrum_genome to make them accessible before publication.

Conflict of interest statement

The authors declare no competing interest.

Supplementary Data

Supplementary data is available at *Horticulture Research* online.

References

1. Abe H, Urao T, Ito T. *et al.* Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell*. 2003;**15**:63–78
2. Russell JR, Bayer M, Booth C. *et al.* Identification, utilization and mapping of novel transcriptome-based markers from blackcurrant (*Ribes nigrum*). *BMC Plant Biol*. 2011;**11**:147
3. Čereković N, Pagter M, Kristensen HL. *et al.* Effects of drought stress during flowering of two pot-grown blackcurrant (*Ribes nigrum* L.) cultivars. *Sci Hortic*. 2013;**162**:365–73
4. Yang X, Lu M, Wang Y. *et al.* Response mechanism of plants to drought stress. *Horticulturae*. 2021;**7**:50
5. Jan S, Abbas N, Ashraf M. *et al.* Roles of potential plant hormones and transcription factors in controlling leaf senescence and drought tolerance. *Protoplasma*. 2018;**256**:313–29

6. Munné-Bosch S, Allegre L. Die and let live: leaf senescence contributes to plant survival under drought stress. *Funct Plant Biol*. 2004;**31**:203–16
7. Kuromori T, Seo M, Shinozaki K. ABA transport and plant water stress responses. *Trends Plant Sci*. 2018;**23**:513–22
8. Choi WG, Miller G, Wallace I. *et al.* Orchestrating rapid long-distance signaling in plants with Ca2+, ROS and electric signals. *Plant J*. 2017;**90**:698–707
9. Ozturk M, Unal BT, García-Carparrós P. *et al.* Osmoregulation and its actions during the drought stress in plants. *Physiol Plant*. 2020;**172**:1321–35
10. Lawas LMF, Zuther E, Jagadish SVK. *et al.* Molecular mechanisms of combined heat and drought stress resilience in cereals. *Curr Opin Plant Biol*. 2018;**45**:212–7
11. Farooqi MQU, Nawaz G, Wani SH, CHoudhary JR, Rana M, Sah RP, Afzal M, Zahra Z, Ganie SA, Razzaq A, Reyes VR, Mahmoud EA, Elansary HO, El-Abedin TKZ, Siddique KHM (2022). Recent developments in multi-omics and breeding strategies for abiotic stress tolerance in maize (Zea mays L.). *Frontiers in PLant Science* **13**:965878.
12. Powell AF, Feder A, Li J. *et al.* A *Solanum lycopersicoides* reference genome facilitates insights into tomato specialized metabolism and immunity. *Plant J*. 2022;**110**:1791–810
13. Hewitt S, Hernández-Montes E, Dhingra A. *et al.* Impact of heat stress, water stress, and their combined effects on the metabolism and transcriptome of grape berries. *Sci Rep*. 2023;**13**:9907
14. Čereković N, Jarret D, Pagter M. *et al.* The effects of drought stress on leaf gene expression during flowering in blackcurrant (*Ribes nigrum* L.). *Eur J Hortic Sci*. 2015;**80**:39–46
15. Russell JR, Hackett C, Hedley P. *et al.* The use of genotyping by sequencing in blackcurrant (*Ribes nigrum*): developing high-resolution linkage maps in species without reference genome sequences. *Mol Breed*. 2013;**33**:835–49
16. Jarret D, Morris J, Cullen DW. *et al.* A transcriptome and metabolite atlas of blackcurrant fruit development highlights hormonal regulation and reveals the role of key transcription factors. *Front Plant Sci*. 2018;**9**:1235
17. Thole V, Bassard JE, Ramírez-González R. *et al.* RNA-seq, de novo transcriptome assembly and flavonoid gene analysis in 13 wild and cultivated berry fruit species with high content of phenolics. *BMC Genet*. 2019;**20**:995
18. Mažeikienė I, Juškytė AD, Stanys V. De novo transcriptome analysis of *R. nigrum* cv. Aldoniai in response to blackcurrant reversion virus infection. *Int J Mol Sci*. 2022;**23**:3137
19. Juškytė AD, Mažeikienė I, Stanys V. Analysis of R genes related to blackcurrant reversion virus resistance in the comparative transcriptome of *Ribes nigrum* cv. Aldoniai. *Plants (Basel)*. 2022;**11**:3137
20. Sun X, Zhan Y, Li S. *et al.* Complete chloroplast genome assembly and phylogenetic analysis of blackcurrant (*Ribes nigrum*), red and white currant (*Ribes rubrum*), and gooseberry (*Ribes uva-crispa*) provide new insights into the phylogeny of Grossulariaceae. *PeerJ*. 2023;**11**:e16272
21. Chiche J, Brown SC, Leclerc JC. *et al.* Genome size, heterochromatin organisation, and ribosomal gene mapping in four species of Ribes. *Can J Bot*. 2003;**81**:191–296
22. Cheng H, Asri M, Lucas J. *et al.* Scalable telomere-to-telomere assembly for diploid and polyploid genomes with double graph. *Nat Methods*. 2024;**21**:967–70
23. Rhie A, Walenz BP, Koren S. *et al.* Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*. 2020;**21**:245
24. Ou S, Su W, Liao Y. *et al.* Benchmarking transposable element annotation methods for creation of streamlined, comprehensive pipeline. *Genome Biol*. 2019;**20**:275
25. Li K, Xu P, Wang J. *et al.* Identification of errors in draft genome assemblies at single-nucleotide resolution for quality assessment and improvement. *Nat Commun*. 2023;**14**:6556
26. Simão FA, Waterhouse RM, Ioannidis P. *et al.* BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;**31**:3210–2
27. Schwacke R, Ponce-Soto GY, Krause K. *et al.* MapMan4: a refined protein classification and annotation framework applicable to multi-omics data analysis. *Mol Plant*. 2019;**12**:879–92
28. Ulahannan N, Pendleton M, Deshpande A. *et al.* Nanopore sequencing of DNA concatemers reveals higher-order features of chromatin structure. 2019; bioRxiv preprint: 833590
29. Cheng H, Asri M, Lucas J. *et al.* Scalable telomere-to-telomere assembly for diploid and polyploid with double graph. *Nat Methods*. 2023;**21**:967–70
30. van Rengs WMJ, Schmidt MH-W, Effgen S. *et al.* A chromosome scale tomato genome built from complementary PacBio and nanopore sequences alone reveals extensive linkage drag during breeding. *Plant J*. 2022;**110**:572–88
31. Zielinski QB. Chromosome numbers and meiotic studies. *Bot Gaz*. 1953;**114**:265–74
32. Feron R, Waterhouse RM. Assessing species coverage and assembly quality of rapidly accumulating sequenced genomes. *GigaScience*. 2022;**11**:giac006
33. Michael TP, Jupe F, Bemm F. *et al.* High contiguity *Arabidopsis thaliana* genome assembly with a single nanopore flow cell. *Nat Commun*. 2018;**9**:541
34. Ou S, Chen J, Jiang N (2018). Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acid Research* **46**:e126.
35. Du H, Yu Y, Ma Y. *et al.* Sequencing and de novo assembly of a near complete indica rice genome. *Nat Commun*. 2017;**8**:15324
36. Song JM, Xie WZ, Wang S. *et al.* Two gap-free reference genomes and global view of the centromere architecture in rice. *Mol Plant*. 2021;**14**:1757–67
37. Sun R, Li S, Chang L. *et al.* Chromosome-level genome assembly of *Fragaria pentaphylla* using PacBio and hi-C technologies. *Front Genet*. 2022;**13**:873711
38. Huang Y, Escalona M, Morrison G. *et al.* Reference genome assembly of the big berry Manzanita (*Arctostaphylos glauca*). *J Hered*. 2022;**113**:188–96
39. Zhao SW, Gup JF, Kong L. *et al.* Haplotype-resolved genome assembly of *Coriaria nepalensis* a non-legume nitrogen-fixing shrub. *Sci Data*. 2023;**10**:259
40. Soltis DE, Smith SA, Cellinese N. *et al.* Angiosperm phylogeny: 17 genes, 640 taxa. *Am J Bot*. 2011;**98**:704–30
41. Liu L, Chen M, Folk RA. *et al.* Phylogenomic and syntenic data demonstrate complex evolutionary processes in early radiation of the rosids. *Mol Ecol Resour*. 2023;**23**:1673–88
42. Yuan J, Jiang S, Liu M. *et al.* Genomic basis of the giga-chromosomes and giga-genome of tree peony *Paeonia ostii*. *Nat Commun*. 2022;**13**:7328
43. Yu J, Hulse-Kemp AM, Babiker E. *et al.* High-quality reference genome and annotation aids understanding of berry development for evergreen blueberry (*Vaccinium darrowii*). *Hortic Res*. 2021;**8**:228
44. Davik J, Røen D, Lysøe E. *et al.* A chromosome -level genome sequence assembly of the red raspberry (*Rubus idaeus* L.). *PLoS One*. 2022;**17**:e0265096

45. Zhu S, Chen J, Zhao J. *et al.* Genomic insights on the contribution of balancing selection and local adaptation to the long-term survival of a widespread living fossil tree, *Cercidiphyllum japonicum*. *New Phytol*. 2020;**228**:1674–89
46. Gish LA, Clark SE. The RLK/Pelle family of kinases. *Plant J*. 2011;**66**:117–27
47. Ksouri N, Jiménez S, Wells CE. *et al.* Transcriptional responses in root and leaf of *Prunus persica* under drought stress using RNA sequencing. *Front Plant Sci*. 2016;**7**:1715
48. Wei K, Wang Y, Zhong X. *et al.* Protein kinase structure, expression and regulation in maize drought signaling. *Mol Breed*. 2014;**34**:583–602
49. Lin W, Ma X, Shan L. *et al.* Big roles of small kinases: the complex functions of receptor-like cytoplasmic kinases in plant immunity and development. *J Integr Plant Biol*. 2013;**55**:1188–97
50. He Y, Zhou J, Shan L. *et al.* Plant cell surface receptor-mediated signaling- a common theme amid diversity. *J Cell Sci*. 2018;**131**:jcs209353
51. Liang X, Zhou JM (2018). Receptor-Like Cytoplasmic Kinases: Central Players in Plant Receptor Kinase–Mediated Signaling. *Annual review of plant biology* **69**:267-99
52. Forzani C, Carreri A, de la Fuente van Bentem S. *et al.* The Arabidopsis protein kinase Pto-interacting 1–4 is a common target of the oxidative signal-inducible 1 and mitogen-activated protein kinases. *FEBS J*. 2011;**278**:1126–36
53. Majhi BB, Sobol G, Gachie S. *et al.* Brassinosteroid-signalling kinases 7 and 8 associate with the FLS2 immune receptor and are required for flg22-induced PTI responses. *Mol Plant Pathol*. 2021;**22**:786–99
54. Wang J, Hu M, Wang J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science*. 2019a;**364**:eaav5870
55. Wang J, Wang J, Hu M. *et al.* Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science*. 2019b;**364**:eaav5868
56. Bai L, Guozeng Z, Yun Z. *et al.* Plasma membrane-associated proline-rich extensin-like receptor kinase 4, a novel regulator of Ca^{2+} signaling, is required for abscisic acid responses in *Arabidopsis thaliana*. *Plant J*. 2009;**60**:314–27
57. Shahid S, Sher MA, Ahmad F. *et al.* Prediction of RNA editing sites and genome-wide characterization of *PERK* gene family in maize (*Zea mays* L.) in response to drought stress. *J King Saud Univ Sci*. 2022;**34**:102293
58. Wrzaczek M, Brosché M, Salojärvi J, Kangasjärvi S, Idänheimo N, Mersmann S, Robatzek S, Karpìński S, Karpińska B, Kangasjärvi J (2010). Transcriptional regulation of the CRK/DUF 26 group of receptor-like protein kinases by ozone and plant hormones in Arabidopsis. *BMC Plant Biology* **10**, 95.
59. Arellano-Villagómez FC, Guevara-Olvera L, Zuñiga-Mayo VM, Cerbantez-Bueno VE, Verdugo-Perales M, Medina HR, De Folter S, Gerardo Acosta-García G (2021). Arabidopsis cysteine-rich receptor-like protein kinase CRK33 affects stomatal density and drought tolerance. *Plant Signal Behaviour* **16**:1905335.
60. Kundu S, Gantait S. Abscisic acid signal crosstalk during abiotic stress response. *Plant Gene*. 2017;**11**:61–9
61. Baena-González E, Rolland F, Thevelein JM. *et al.* A central integrator of transcription networks in plant stress and energy signalling. *Nature*. 2007;**448**:938–42
62. Zhang Y, Shewry PR, Jones H. *et al.* Expression of antisense SnRK1 protein kinase sequence causes abnormal pollen development and male sterility in transgenic barley. *Plant J*. 2001;**28**:431–41
63. Radchuk R, Radchuk V, Weschke W, Borisjuk L, Weber H (2006). Repressing the expression of the SUCROSE NONFERMENTING-1-RELATED PROTEIN KINASE gene in pea embryo causes pleiotropic defects of maturation similar to an abscisic acid-insensitive phenotype. *Plant Physiology* **140**:263-78
64. Yang R, Guo Q, Gu Z. GABA shunt and polyamine degradation pathway on γ-aminobutyric acid accumulation in germinating fava bean (Vicia faba L.) under hypoxia. *Food Chem*. 2013;**136**:152–9
65. Hasan MM, Liu XD, Waseem M. *et al.* ABA activated AnRK2 kinases: an emerging role in plant growth and physiology. *Plant Signal Behav*. 2022;**17**:2071024
66. Chen K, Tang W, Zhou Y. *et al.* AP2/ERF transcription factor GmDREB1 confers drought tolerance in transgenic soybean by interacting with GmERFs. *Plant Physiol Biochem*. 2022;**170**:287–95
67. Dubouzet JG, Sakuma Y, Ito Y. *et al.* OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J*. 2003;**33**:751–63
68. Huang XY, Chao DY, Gao JP. *et al.* A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. *Genes Dev*. 2009a;**23**:1805–17
69. Joshi R, Wani SH, Singh B. *et al.* Transcription factors and plants response to drought stress: current understanding and future directions. *Front Plant Sci*. 2016;**7**:1029
70. Khoso MA, Hussain A, Ritonga FN. *et al.* WRKY transcription factors (TFs): molecular switches to regulate drought, temperature, and salinity stresses in plants. *Front Plant Sci*. 2022;**13**:3389
71. Li Z, Liang F, Zhang FN. *et al.* Enhanced tolerance to drought stress resulting from *Caragana korshinskii* CkWRKY33 in transgenic *Arabidopsis thaliana*. *BMC Genom Data*. 2021;**22**:11
72. Liu W, Tai H, Li S. *et al.* bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. *New Phytol*. 2014;**201**:1192–204
73. Qu X, Wang H, Chen M. *et al.* Drought stress-induced physiological and metabolic changes in leaves of two oil tea cultivars. *J Am Soc Hortic Sci*. 2019;**144**:439–47
74. Qiu XM, Sun YY, Ye XY. *et al.* Signaling role of glutamate in plants. *Front Plant Sci*. 2020;**10**:1743
75. Valliyodan B, Nguyen HT. Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Plant Biol*. 2006;**9**:189–95
76. Takasaki H, Maruyama K, Kidokoro S. *et al.* The abiotic stress-responsive NAC-type transcription factor OsNAC5 regulates stress-inducible genes and stress tolerance in rice. *Mol Genet Genomics*. 2010;**284**:173–83
77. Tran LS, Nakashima K, Sakuma Y. *et al.* Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early response to dehydration stress 1 promoter. *Plant Cell*. 2004;**16**:2481–98
78. Huang J, Sun S, Xu D. *et al.* Increased tolerance of rice to cold, drought and oxidative stresses mediated by the overexpression of a gene that encodes the zinc finger protein ZFP245. *Biochem Biophys Res Commun*. 2009b;**389**:556–61
79. Wang K, Ding Y, Cai C. *et al.* The role of C2H2 zinc finger proteins in plant response to abiotic stresses. *Physiol Plant*. 2018;**165**:690–700
80. He X, Wang C, Wang H. *et al.* The function of MAPK cascades in response to various stresses in horticultural plants. *Front Plant Sci*. 2020;**11**:952

81. Ku YS, Sintaha M, Cheung MY. *et al.* Plant hormone signaling crosstalks between biotic and abiotic stress responses. *Int J Mol Sci*. 2018;**19**:3206
82. Lin L, Wu J, Jiang M. *et al.* Plant mitogen-activated protein kinase cascades in environmental stresses. *Int J Mol Sci*. 2021;**22**:1543
83. Agalou A, Purwantomo S, Overnaes E. *et al.* A genome-wide survey of HD-zip genes in rice and analysis of drought-responsive family members. *Plant Mol Biol*. 2008;**66**:87–103
84. Himmelbach A, Hoffmann T, Leube M. *et al.* Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in Arabidopsis. *EMBO J*. 2002;**21**:3029–38
85. Olsson AS, Engström P, Söderman E. The homeobox genes ATHB12 and ATHB7 encode potential regulators of growth in response to water deficit in Arabidopsis. *Plant Mol Biol*. 2004;**55**:663–77
86. Perotti MF, Arce AL, Chan RL. The underground life of homeodomain-leucine zipper transcription factors. *J Exp Bot*. 2021;**72**:4005–21
87. Jiang F, Hartung W. Long-distance signaling of abscisic acid (ABA): the factors regulating the intensity of ABA signal. *J Exp Bot*. 2008;**59**:37–43
88. Ullah N, Yüce M, Gökçe ZNÖ. *et al.* Comparative metabolite profiling of drought stress in roots and leaves of seven Triticeae species. *BMC Genomics*. 2017;**18**:969
89. Hildebrandt TM, Nesi AN, Araújo WL. *et al.* Amino acid catabolism in plants. *Mol Plant*. 2015;**8**:1563–79
90. Pratelli R, Guillaume P. Regulation of amino acid metabolic enzymes and transporters in plants. *J Exp Bot*. 2014;**65**:5535–56
91. Huang T, Jander G. Abscisic acid-regulated protein degradation causes osmotic stress-induced accumulation of branched-chain amino acids in *Arabidopsis thaliana*. *Planta*. 2017;**246**:737–47
92. Rizhsky L, Liang H, Shuman J. *et al.* When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. *Plant Physiol*. 2004;**134**:1683–96
93. Szabados L, Savouré A. Proline: a multifunctional amino acid. *Trends Plant Sci*. 2010;**15**:89–97
94. Wang B, Yang X, Jia Y. *et al.* High-quality *Arabidopsis thaliana* genome assembly with nanopore and hifi long reads. *Genomics Proteomics Bioinformatics*. 2021;**20**:4–13
95. Hasan MM, Alabdallah NM, Alharbi BM. *et al.* GABA: a key player in drought stress resistance in plants. *Int J Mol Sci*. 2021;**22**:10136
96. Mousavi SAR, Chauvin A, Pascaud F. *et al.* GLUTAMATE RECEPTOR-LIKE genes mediate leaf-to-leaf wound signalling. *Nature*. 2013;**500**:422–6
97. Nguyen CT, Kurenda A, Stolz S. *et al.* Identification of cell populations necessary for leaf-to-leaf electrical signaling in a wounded plant. *PNAS*. 2018;**115**:10178–83
98. Sun C, Gao X, Zhou J. *et al.* Metabolic response of maize (Zea mays L.) plants to combined drought and salt stress. *Plant and Soil*. 2015;**388**:99–117
99. Toyota M, Spencer D, Sawai-Toyota S. *et al.* Glutamate triggers long-distance, calcium-based plant defense signaling. *Science*. 2018;**361**:1112–5
100. Degu A, Hatew B, Nunes-Nesi A. *et al.* Inhibition of aconitase in citrus fruit callus results in a metabolic shift towards amino acid biosynthesis. *Planta*. 2011;**234**:501–13
101. Araújo WL, Martins AO, Fernie AR. *et al.* 2-oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Front Plant Sci*. 2014;**5**:552
102. Wan T, Feng Y, Liang C. *et al.* Metabolomics and transcriptomics analysis of two contrasting cherry rootstocks in response to drought stress. *Biology (Basel)*. 2021;**10**:201
103. Joshi V, Joung JG, Fei Z. *et al.* Interdependence of threonine, methionine and isoleucine metabolism in plants: accumulation and transcriptional regulation under abiotic stress. *Amino Acids*. 2010;**39**:933–47
104. Sanchez DH, Siahpoosh MR, Roessner U. *et al.* Plant metabolomics reveals conserved and divergent metabolic responses to salinity. *Physiol Plant*. 2008;**132**:209–19
105. Székely G, Abrahám E, Cséplo A. *et al.* Duplicated P5CS genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. *Plant J*. 2008;**53**:11–28
106. Shim JS, Jeong HI, Bang SW. *et al.* Drought-induced branched-chain amino acid aminotransferase enhances drought tolerancein rice. *Plant Physiol*. 2018;**191**:1435–47
107. Suguiyama VF, Silva EA, Meirelles ST. *et al.* Leaf metabolite profile of the Brazilian resurrection plant *Barbacenia purpurea* hook. (Velloziaceae) shows two time-dependent responses during desiccation and recovering. *Front Plant Sci*. 2014;**5**:96
108. Kumar M, Patel MK, Kumar N. *et al.* Metabolomics and molecular approaches reveal drought stress tolerance in plants. *Int J Mol Sci*. 2021;**22**:9108
109. Wang C, Yang A, Yin H. *et al.* Influence of water stress on endogenous hormone contents and cell damage of maize seedlings. *J Integr Plant Biol*. 2008;**50**:427–34
110. Bowne JB, Erwin TA, Jutter J. *et al.* Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Mol Plant*. 2012;**5**:418–29
111. Kaur H, Manna M, Thakur T. *et al.* Imperative role of sugar signaling and transport during drought stress responses in plants. *Physiol Plant*. 2021;**171**:833–48
112. Taji T, Ohsumi C, Iuchi S. *et al.* Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J*. 2002;**29**:417–26
113. You J, Zhang Y, Liu A. *et al.* Transcriptomic and metabolomic profiling of drought-tolerant and susceptible sesame genotypes in response to drought stress. *BMC Plant Biol*. 2019;**19**:267
114. Dilokpimol A, Poulsen CP, Vereb G. *et al.* Galactosyltransferases from *Arabidopsis thaliana* in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. *BMC Plant Biol*. 2014;**14**:90
115. Hedley PE, Russel JR, Jorgensen L. *et al.* Candidate genes associated with bud dormancy release in blackcurrant (*Ribes nigrum* L.). *BMC Plant Biol*. 2010;**10**:202
116. Ganger MT, Dietz GD, Erwing SJ. A common base method for analysis of qPCR data and the application of simple blocking in qPCR experiments. *BMC Bioinformatics*. 2017;**18**:534
117. Zhao S, Fernald RD. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol*. 2005;**12**:1047–64
118. Wick RR, Judd LM, Cl G. *et al.* Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom*. 2017;**3**:e000132
119. Cheng H, Jarvis ED, Fedrigo O. *et al.* Haplotype-resolved assembly of diploid genomes without parental data. *Nat Biotechnol*. 2022;**40**:1332–5
120. Roach MJ, Schmidt SA, Bornemann AR. Purge Haplotigs_ allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*. 2018;**19**:460
121. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;**34**:3094–100

122. Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. *BMC Bioinformatics*. 2022;**39**:btac808
123. Dudchenko O, Shamim MS, Batra SS. *et al.* The Juicebox assembly tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. 2018; bioRxiv 254797
124. Gurevich A, Saveliev V, Vyahhi NT. *et al.* QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;**29**: 1072–5
125. Alonge M, Lebeigle L, Kirsche M. *et al.* Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol*. 2022;**23**:258
126. Goel M, Sun H, Jiao WB. *et al.* SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol*. 2019;**20**:277
127. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012; arXiv preprint arXiv:1207.3907
128. Cingolani P, Platts A, Wang le L. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;**2**:80–92
129. RStudio Team (2020). *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.
130. Ou S, Jiang N. Assessing genome assembly quality using the LTR assembly index (LAI). *Nucleic Acids Res*. 2018; gky730
131. Camacho C, Coulouris G, Avagyan V. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics*. 2008;**10**:421
132. Holst F, Bolger A, Günther C. *et al.* Helixer-de novo prediction of primary eukaryotic gene models combining deeep learning and a hidden markov model. 2023; bioRxiv, 2023.02.06.527280v2
133. Shumate A, Wong B, Pertea G. *et al.* Improved transcriptome assembly using a hybrid of long and short reads with StringTie. *PLoS Comput Biol*. 2022;**18**:e1009730
134. Mapleson DL, Venturini L, Kaithakottil G. *et al.* Efficient and accurate detection of splice junctions from RNAseq with portcullis. *GigaScience*. 2018;**7**:giy131
135. Pertea G, Pertea M. GFF utilities: GrrRead and GffCompare. *F1000 Res*. 2020; ISCB Comm J-304
136. Patro R, Duggal G, Love MI. *et al.* Salmon: fast and bia-aware quantification of transcript expression using dual-phase inference. *Nat Methods*. 2019;**14**:417.419
137. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;**26**:139–40
138. Osorio S, Alba R, Nikoloski Z. *et al.* Integrative comparative analyses of transcript and metabolite profiles from pepper and tomato ripening and development stages uncovers species-specific patterns of network regulatory behaviour. *Plant Physiol*. 2012;**159**:1713–29
139. Vallarino JG, de Abreu e Lima F, Soria C. *et al.* Genetic diversity of strawberry germplams using metabolomic biomarkers. *Sci Rep*. 2018;**8**:14386
140. Kopka J, Schauer N, Krueger S. *et al.* GMD@CSB.DB: the golm metabolome database. *Bioinformatics*. 2005;**21**:1635–8
141. Weil HL, Schneider K, Tschöpe M. *et al.* PLANTdataHUB: a collaborative platform for continuous FAIR data sharing in plant research. *Plant J*. 2023;**116**:974–88

Horticulture Research, 2025, **12**: uhae313
https://doi.org/10.1093/hr/uhae313
Article