Environmental and Experimental Botany 237 (2025) 106218

Contents lists available at ScienceDirect

Environmental and Experimental Botany

journal homepage: www.elsevier.com/locate/envexpbot

Leveraging hyperspectral phenotyping for accurate, non-destructive prediction of metabolite profiles in poplar under drought stress☆

Mengjun Shu[a,b], Antoine L. Harfouche[c], Martin Trtílek[d], Klára Panzarová[d], Omar F. Alasia[e], John H. Lagergren[a,b], Audrey Labbé[a,b], Nancy L. Engle[a,b], Miranda M. Clark[a,b], Jin-Gui Chen[a,b], Gerald A. Tuskan[a,b], Timothy J. Tschaplinski[a,b,*]

[a] *Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA*
[b] *Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, USA*
[c] *Department for Innovation in Biological, Agro-food and Forest Systems, University of Tuscia, Viterbo, Italy*
[d] *PSI (Photon Systems Instruments),spol. s.r.o., Prumyslova 470, Drásov CZ-664 24, Czech Republic*
[e] *Franco Alasia Vivai, Savigliano, Italy*

ARTICLE INFO

Keywords:
Hyperspectral imaging
VNIR
SWIR
Untargeted metabolomics
Populus
Drought stress
LASSO regression
Stress biomarkers
Non-destructive phenotyping
Biochemical trait prediction

ABSTRACT

Accurately predicting drought tolerance in woody perennial bioenergy crops is critical for sustainable biomass production under fluctuating precipitation. Hyperspectral imaging (HSI) in the visible-near-infrared (VNIR) and shortwave-infrared (SWIR) ranges offers a promising approach for predicting plant biochemical traits, yet its application in metabolite profiling remains underexplored. We integrated VNIR+SWIR HSI with untargeted metabolomics to investigate drought-induced metabolic shifts in *Populus* leaves from eight *Populus* genotypes. Metabolite profiling identified 127 compounds, with 73 showing significant drought responses spanning amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and unclassified metabolites. Spectral analysis revealed consistently higher reflectance across VNIR and SWIR wavelengths in drought-stressed plants, corresponding with increased accumulation of AA and reduced CHO and PG levels. Least absolute shrinkage and selection operator (LASSO) regression modeling identified robust spectral predictors of metabolite concentrations, associating VNIR wavelengths (500–700 nm) predominantly with AA and P, whereas SWIR wavelengths (1680–1700 nm) reliably predicted CHO, OA, and T. Several stable spectral-metabolite associations persisted across the two watering regimes (drought vs. well-watered), highlighting their potential as spectral biomarkers for non-destructive stress monitoring. Minimal genotype-specific variation suggests that observed spectral and metabolic responses were driven primarily by environmental factors, likely reflecting limited genetic diversity among the commercial *Populus* genotypes examined. This work establishes VNIR+SWIR hyperspectral imaging as a powerful, non-destructive phenotyping tool for precision monitoring and targeted improvement of drought resilience in bioenergy crops.

1. Introduction

The increasing severity and frequency of drought events pose critical threats to global agricultural productivity and ecosystem function (Lesk et al., 2016; Masson-Delmotte et al., 2021; Naik et al., 2010; Raina et al., 2024). Enhancing plant resilience to drought stress is imperative to address these challenges, particularly in the context of sustainable biomass production. Perennial bioenergy crops offer distinct environmental benefits over annual crops, including enhanced carbon sequestration, reduced soil erosion, and improved nutrient retention

---

☆ Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05–00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

* Correspondence to: Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA.
E-mail address: tschaplinstj@ornl.gov (T.J. Tschaplinski).

https://doi.org/10.1016/j.envexpbot.2025.106218
Received 25 April 2025; Received in revised form 29 July 2025; Accepted 13 August 2025
Available online 14 August 2025
0098-8472/© 2025 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

(Anderson-Teixeira et al., 2013; Oates et al., 2016; Pugesgaard et al., 2015; Schmer et al., 2014). Poplar (*Populus* spp.), a fast-growing and genetically diverse genus, has emerged as a model system for studying plant responses to abiotic stress and a leading candidate for bioenergy applications (Clifton-Brown et al., 2019; Taylor et al., 2019; Tuskan et al., 2006). Its extensive genomic resources, metabolic plasticity, and adaptability to diverse environments make poplar an ideal species for dissecting the biochemical and physiological mechanisms underlying drought resilience (Jansson and Douglas, 2007; Sun et al., 2024; Viger et al., 2016).

Metabolite profiling has long been a cornerstone of plant stress biology, revealing pivotal roles of metabolites as osmoprotectants, antioxidants, and signaling molecules in mediating drought responses (Kumar et al., 2021; Obata and Fernie, 2012; Ramachandra Reddy et al., 2004; Tschaplinski and Tuskan, 1994). However, traditional approaches typically require destructive sampling and extensive labor, limiting their scalability and real-time applicability, and relevance for practical breeding and management decisions (Fernie and Schauer, 2009; Hall et al., 2002). Current field-based phenotyping methods for poplar genotypes typically involve time-consuming and costly visual assessments. Developing high-throughput phenotyping methods capable of rapidly screening large poplar germplasm collections under environmental stresses could significantly improve genetic selection for drought resilience. Recent advancements in high-throughput phenotyping, particularly hyperspectral imaging (HSI), offer a transformative approach for non-destructive monitoring of plant physiological and biochemical traits (Abebe et al., 2023; Araus and Cairns, 2014; Fahlgren et al., 2015; Homolová et al., 2013; Lowe et al., 2017; Moncholi-Estornell et al., 2022). By capturing detailed reflectance spectra across a broad range of wavelengths, HSI enables the detection of subtle biochemical and structural changes in plant tissues associated with drought stress (Abdelhakim et al., 2024; Montes et al., 2022; Serbin et al., 2016; Yendrek et al., 2017). Integrating HSI with metabolite profiling could provide a real-time, non-destructive method for monitoring metabolic shifts, offering new insights into plant stress responses.

Despite advances in high-throughput phenotyping, linking spectral data to metabolite composition remains a key challenge. The complexity of metabolic networks and the variability in spectral reflectance across environmental conditions make direct associations difficult to establish. Least absolute shrinkage and selection operator (LASSO) regression provides a powerful statistical framework to address this challenge by identifying key spectral predictors while minimizing overfitting (Hastie, 2009; Tibshirani, 1996; Wold et al., 2001). While HSI and machine learning approaches have been applied to plant trait characterization, their potential for metabolite prediction remains largely unexplored (Meacham-Hensold et al., 2020; Ting et al., 2023; Vergara-Diaz et al., 2020).

We hypothesize that specific metabolite classes and even individual metabolites exhibit distinct spectral signatures due to differential absorption and reflectance properties across the electromagnetic spectrum. In particular, based on work Curran (1989) and Tschaplinski and Tuskan (1994), we propose that drought-induced osmotic adjustment alters amino-acid and carbohydrate pools, which in turn create diagnostic shifts in visible-near-infrared (VNIR, 500–900 nm) and shortwave-infrared (SWIR, 950–1700 nm) reflectance via changes in N-rich pigments and O-H bond absorptions. By integrating VNIR and SWIR HSI, we aim to identify robust spectral-metabolite associations to enable non-destructive drought monitoring in *Populus*.

This study combines VNIR+SWIR hyperspectral imaging with untargeted metabolite profiling to investigate metabolic shifts in leaves of drought-tolerant and drought-susceptible *Populus* genotypes under drought (D) and well-watered (W) conditions. Using LASSO regression, we analyzed 637 spectral profiles and quantified 127 metabolites spanning eight biochemical classes, assessing the relationships between spectral reflectance and metabolic composition. Our primary objective was to identify spectral signatures associated with drought-responsive metabolites and to evaluate the predictive capability of hyperspectral imaging for non-destructive metabolic phenotyping. This work establishes a foundation for deploying HSI in bioenergy crop breeding, bridging the gap between genotype and phenotype with precision phenotyping tools.

2. Material and methods

2.1. Plant material, growth conditions, and drought experiment

The greenhouse experiment was conducted at the Photon Systems Instruments (PSI) Research Center phenomics facility in Drásov, Czech Republic, using dormant poplar cuttings derived from eight commercial hybrid poplar genotypes. These genotypes were selected based on their known growth characteristics, drought tolerance, and suitability for biomass production, with each genotype represented by 10 cuttings, yielding a total of 80 plants. The drought-tolerant hybrids ‘AF18’ and ‘AF34’ (*Populus deltoides* × *P. nigra*) are characterized by vigorous growth, tolerance to fungal diseases, and adaptability to short- and medium-rotation forestry. Drought-susceptible hybrids ‘AF16’, ‘AF17’, and ‘AF24’ (*P. deltoides* × *P. nigra*), as well as ‘AF28’ (*P. deltoides* × (*P. deltoides* × *P. nigra*)), were chosen for their favorable stem form, growth rates, and disease resistance despite lower drought tolerance. Additionally, ‘I-214’ and ‘Boccalari’, traditional *P.* × *canadensis* genotypes widely used in poplar cultivation, were included despite their sensitivity to drought and various biotic stresses.

Prior to planting, poplar cuttings were treated with a 1 % fungicide solution, and their distal ends lightly coated with rooting hormone. Cuttings were then planted in 250-mL pots containing a standardized substrate mixture consisting of Klassman substrate 2, supplemented with 15 % perlite and dolomite limestone. Plants were cultivated under controlled environmental conditions (16-h photoperiod/8-h dark regime; light intensity: 300 μmoles·m^{-2}·s^{-1} white light [54 %] and 4.5 μmoles·m^{-2}·s^{-1} far-red light [19 %]; temperature: 19–21°C; relative humidity: 60 %) for a total of 11 weeks. After an initial seven-week establishment period, five plants from each genotype were randomly assigned to drought (D) treatment, and the remaining five served as well-watered controls (W). For the W group, substrate moisture levels were maintained at the second-highest pot weight recorded during the experiment through daily watering. For the D treatment, watering was progressively withheld to induce stress conditions by weighing each pot and restricting watering to gradually reduce substrate moisture levels (Balasubramanian et al., 2024; Mateus et al., 2021). This approach ensured consistent drought stress while allowing for physiological monitoring of the plants stress responses. At the end of the full 11-week experimental period (7-week establishment plus 4-week treatment), 55 viable plants remained (D: $n = 28$, W: $n = 27$) for subsequent spectral and metabolite analyses. These 55 plants represented all eight original genotypes: five *P. deltoides* × *P. nigra* (‘AF16’, ‘AF17’, ‘AF18’, ‘AF24’, ‘AF34’), one *P. deltoides* × (*P. deltoides* × *P. nigra*) (‘AF28’), and two *P.* × *canadensis* (‘I-214’, ‘Boccalari’), with 7–8 ramets per genotype.

2.2. Hyperspectral, thermal infrared, kinetic chlorophyll fluorescence, and RGB imaging in a high-throughput phenotyping facility

Following the seven-week establishment period, plants underwent automated phenotyping every 2–3 days for a total of 11 rounds to capture detailed physiological and morphological traits. Imaging was conducted using the high-throughput PlantScreen™ Modular System (Photon Systems Instruments, PSI, Czech Republic), which integrates kinetic chlorophyll fluorescence, thermal infrared (IR), hyperspectral, and red–green–blue (RGB) imaging sensors.

Photosynthesis-related traits were measured regularly at 10:00 am using kinetic chlorophyll fluorescence imaging and a short light protocol as described in Abdelhakim et al. (2024). Temperature profiles of plants were recorded via thermal imaging, and image acquisition and segmentation were processed according to established methods (Abdelhakim et al., 2021; Findurová et al., 2023; Zagorščak et al., 2025). Morphological and growth dynamics were assessed using RGB imaging from the top and multiple-angle side views, (0°, 120°, 240°), with projected leaf area estimated by summing non-overlapping plant pixel counts across views. Image processing followed the methods of Awlia et al. (2016) and Abdelhakim et al. (2021).

Hyperspectral reflectance data were collected using visible-near-infrared (VNIR, 420–900 nm) and shortwave infrared (SWIR, 950–1700 nm) sensors integrated into the PlantScreen™ system as described in Abdelhakim et al. (2024). These sensors cover the spectral range from the green–red portion of visible wavelengths through the entire near-infrared and initial shortwave-infrared regions, capturing detailed spectral reflectance information with spectral resolutions of ~1.14 nm and ~2.54 nm, and with a 1:1 binning setting forming 480- and 318-band hypercubes for VNIR and SWIR, respectively. The phenotyping workflow consisted of two stages: the first stage included automatic height measurement, light adaptation for chlorophyll fluorescence imaging, and IR imaging. The second stage involved RGB imaging (top and side view), hyperspectral reflectance measurements, and plant watering and weighing.

Raw image data acquired from chlorophyll fluorescence, RGB, IR, and hyperspectral imaging sensors were automatically processed, and phenotypic parameters were extracted using the PlantScreen™ Analyzer software (PSI, Czech Republic). For predicting metabolite concentrations, only hyperspectral data obtained during the final measurement (week 11) were used, as leaf samples for metabolite analysis were collected immediately following this final hyperspectral measurement. VNIR and SWIR hyperspectral data were normalized using white and dark references prior to subsequent analysis to account for sensor variability and environmental factors. To minimize spectral noise, bands below 500 nm (VNIR) and 950 nm (SWIR) were excluded. Subsequent analyses focused on VNIR wavelengths (500–700 nm) and SWIR wavelengths (1680–1700 nm). Pre-computed plant masks from the PlantScreen™ Analyzer were used to isolate plant tissue from background pixels for each spectral band. Hyperspectral signatures were derived by computing the mean, standard deviation, median, minimum, and maximum values for each spectral band.

The data of all phenotyping parameters were collected from 55 biologically independent plants (28 drought-treated, 27 well-watered), with one full image set per plant per timepoint. While direct leaf relative water content and water potential were not measured, drought was induced using a pot-weight–based watering protocol (see 2.1), which was accompanied by consistent declines in chlorophyll fluorescence and reproducible metabolic and spectral shifts.

2.3. *Growth assessment and statistical analysis*

Morphological measurements, including total leaf area, were captured for all plants of each genotype using RGB imaging from three camera angles previously specified. Growth was defined as the difference in total leaf area between the initial and final measurement rounds. To assess treatment-specific effects, average growth per plant was calculated from the three camera angles under D and W conditions. A two-way ANOVA was performed to examine the main effects of treatment (D vs. W), genotype, and their interaction on total leaf area growth. In addition, genotype-specific comparisons were performed using Wilcoxon rank-sum tests to evaluate the differences in growth between D and W treatments for each genotype individually. *P*-values from the Wilcoxon rank-sum tests were adjusted using the Benjamini-Hochberg correction to control for the false discovery rate (FDR).

To validate that drought treatment induced physiological stress, four light-adapted chlorophyll fluorescence parameters—maximum PSII efficiency (Fv.Fm_Lss), operating PSII efficiency (QY_Lss), photochemical quenching (qP_Lss), and the fraction of open PSII centers (qL_Lss)—were analyzed. Raw fluorescence data were extracted from the PlantScreen™ Analyzer. Only measurements from the final imaging round, corresponding to the end of the 4-week drought period, were used for comparison. Fluorescence data were averaged by genotype and treatment group, and two-way ANOVA was used to assess the effects of drought treatment and genotype for each parameter. Statistical significance from the ANOVA was used to annotate treatment comparisons in Supplementary Fig. S1, with *p*-values indicating differences between drought and control groups for each fluorescence parameter.

All statistical analyses were performed in R (v4.2.0) using the aov function for ANOVA and wilcox.test for non-parametric comparisons (R Core Team, 2013). Data visualizations were generated using the ggplot2 package (Wickham, 2011).

2.4. *Leaf metabolites profiling*

Untargeted metabolomics was performed on all 55 viable plants (D: n = 28, W: n = 27) using gas chromatography-mass spectrometry (GC-MS), with one fully expanded leaf collected per plant. Leaf metabolites were extracted from frozen, finely powdered leaf tissue, with approximately 50 mg weighed into a centrifuge tube per sample. Each sample was extracted twice overnight using 2.5 mL of 80 % ethanol. For internal standardization, 75 μL of sorbitol (1 mg·mL^{-1}) was added to the first ethanol extract. Following centrifugation, the extracts were combined, thoroughly mixed, and a 1 mL aliquot was evaporated to dryness under a gentle nitrogen stream.

To prepare samples for GC-MS, the dried extracts were silylated to produce trimethylsilyl (TMS) derivatives by dissolving them in 500 μL silylation grade acetonitrile, followed by addition of 500 μL of N-methyl-N-trimethylsilyltriflouroacetamide (MSTFA) with 1 % trimethylchlorosilane (TMCS), and heated at 70 °C for 1 h. After 2 days, a 1 μL aliquot was injected into an Agilent 7890 A GC coupled with a 5975 C Inert XL MS. The instrument was configured as previously described (Abraham et al., 2016). Briefly, the GC inlet was configured in splitless mode, and the MS was operated in electron impact (70 eV) ionization mode with a scan range of 50–650 Da.

Metabolite peaks were extracted using key mass-to-charge ratios (*m/z*) to minimize interference from co-eluting peaks, scaled to total ion chromatogram using previously calculated correction factors, and normalized to internal standard recovered, analyzed volume, injection volume, and leaf mass extracted. A large user-created database (~2800 spectra) of electron-impact ionization fragmentation patterns of TMS-derivatized compounds, as well as the Wiley Registry 12th Edition combined with NIST 2020 mass spectral library, were used to identify metabolites.

All leaf samples were processed and analyzed individually; no pooling was performed. The 127 metabolites reported represent the total number of unique compounds detected across all 55 samples, including both D and W treatments.

2.5. *LASSO regression analysis*

To predict metabolite levels from hyperspectral data, LASSO regression models were applied, because the method shrinks many coefficients exactly to zero, yielding a concise, interpretable subset of informative wavelengths—an advantage in high-dimensional hyperspectral data. Reflectance spectra from VNIR, SWIR, and combined VNIR+SWIR datasets were used as predictors, while metabolite levels served as response variables. Four distinct scenarios were evaluated: raw metabolite data with and without treatment information, and log-transformed data with and without treatment information. Each of these scenarios was applied across the VNIR, SWIR, and combined VNIR+SWIR spectral datasets, resulting in a total of twelve LASSO model configurations.

Spectral data consisted of median reflectance values for each cutting, which were normalized to ensure comparability. Metabolite data were processed with low-variance metabolites filtered out prior to analysis.

To ensure robust model performance, we utilized 5-fold cross-validation, repeated 20 times for each scenario. This cross-validation approach allowed for optimal selection of the regularization parameter (λ) to prevent overfitting. The performance of each model was assessed using the adjusted R^2 to quantify the proportion of explained variance and the root mean square error (RMSE) to evaluate prediction accuracy. Additionally, the standardized RMSE was obtained by normalizing the RMSE against the range of the actual metabolite values, enabling comparison across metabolites with different scales.

Spectra selected as predictors were identified based on their non-zero coefficients in the LASSO models, with frequency tracking across replications to provide insights into the consistency and importance of specific wavelengths. Detection rates for each spectrum were calculated to highlight key regions contributing to metabolite prediction accuracy.

3. Results

3.1. Physiological, spectral, and metabolite data overview

This study evaluated 55 poplar plants, including five *Populus deltoides* × *P. nigra* genotypes, one *P. deltoides* × (*P. deltoides* × *P. nigra*) genotype, and two *P.* × *canadensis* genotypes, grown for 11 weeks in a high-throughput phenomics greenhouse. Twenty-eight plants were subjected to drought treatment (D), while 27 served as well-watered controls (W), with 3–4 biological replicates per genotype per treatment. Plant growth was monitored via total leaf area, quantified using RGB images captured from three angles per plant.

Total leaf area was significantly greater in W plants than in D plants (P-adj$\leq$0.001; Fig. 1a). Among the eight commercial genotypes, genotypes, six exhibited statistically significant growth reductions under drought (P-adj < 0.01), with reductions ranging from 59 % to 91 %. Specifically, growth was reduced by 91 % in AF16, 65 % in AF17, 76 % in AF24, 66 % in AF28, 74 % in Boccalari, and 59 % in I-214. The drought-tolerant hybrids AF18 and AF34 showed no or only minor reductions in growth: AF18 displayed a smaller, although statistically significant, reduction (P-adj $= 0.032$), while AF34 exhibited no significant difference (P-adj $= 0.272$). These results confirm that drought stress was effectively imposed and highlight strong genotype-specific differences in growth responses to water deficit.

Chlorophyll fluorescence imaging further corroborated drought-

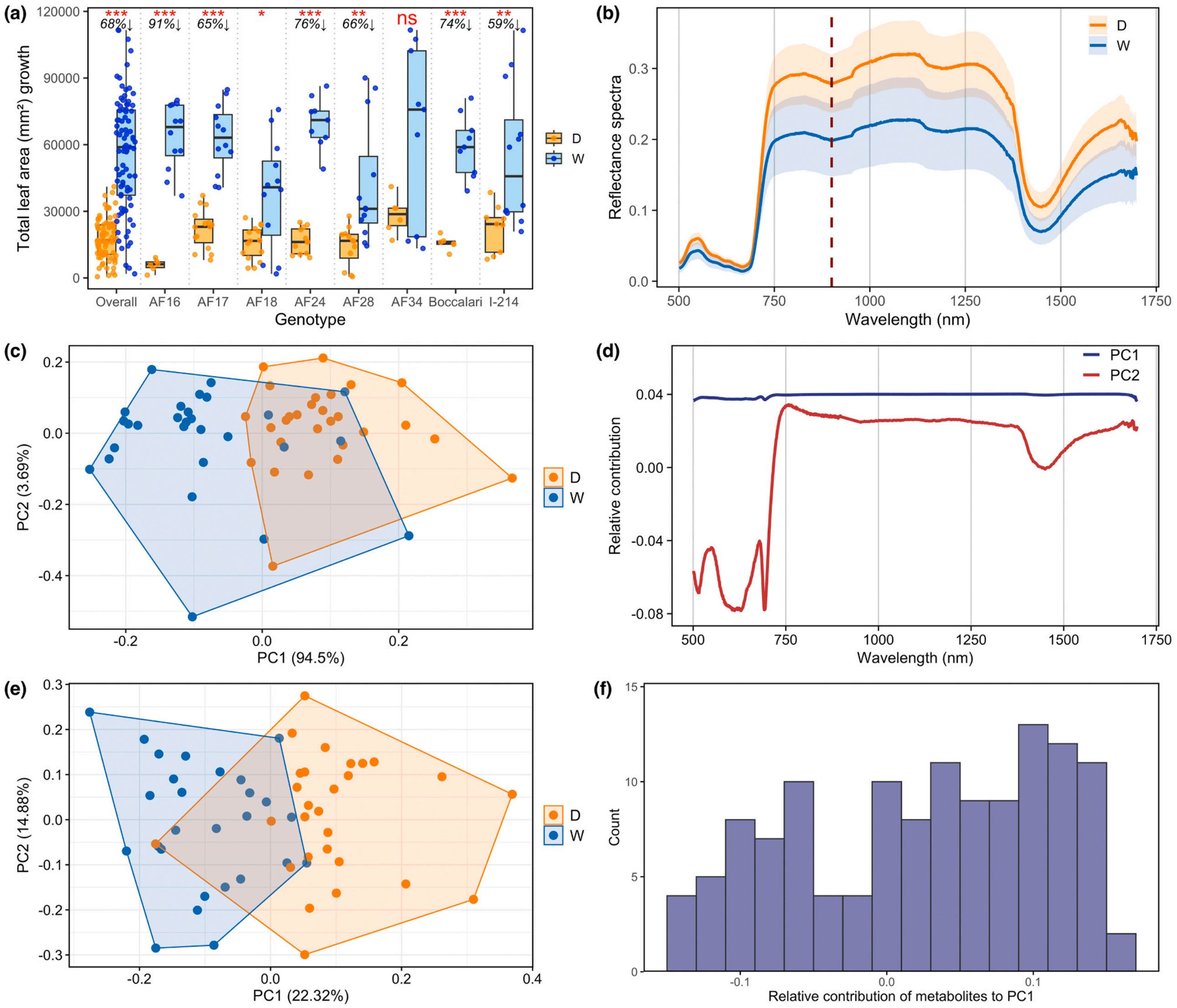

Fig. 1. (a) Boxplot of total leaf area (mm^2) growth across all samples and per genotype, comparing drought-treated (D) and well-watered (W) groups. Horizontal lines within each box indicate the median, the boxes represent the interquartile range, and symbols (*, **, ***, ns) denote ANOVA-adjusted P-values < 0.05, < 0.01, < 0.001, and not significant, respectively. (b) Mean visible and near-infrared (VNIR)+shortwave-infrared (SWIR) reflectance spectra with standard deviations for D and W groups. (c-d) Principal component analysis (PCA) of VNIR+SWIR reflectance spectra, showing D and W clusters and corresponding waveband contributions in the first two components (PC1 and PC2). (e-f) PCA of metabolite profiles, highlighting D and W clusters and metabolite contributions in the first component (PC1).

induced physiological stress (Supplementary Fig. S1). The maximum efficiency of PSII photochemistry in light-adapted samples under steady-state illumination (Fv/Fm_Lss) was significantly reduced in drought-treated plants ($p < 0.001$), indicating stress-induced impairment of photosystem II. Similarly, the fraction of open PSII centers (qL_Lss) was significantly lower under drought conditions ($p < 0.001$), reflecting a decline in electron transport capacity. In contrast, the quantum yield or operating efficiency of photosystem II under light steady-state (QY_Lss) showed no significant difference between treatments, suggesting that drought primarily affected maximal rather than steady-state PSII function. Interestingly, photochemical quenching (qP_Lss) was significantly increased under drought. Together, these fluorescence parameters confirm the presence of drought-induced photosynthetic stress. The observed reductions in Fv/Fm_Lss, qL_Lss, and the compensatory increase in qP_Lss are consistent with pigment degradation and impaired electron transport, which likely contributed to the elevated VNIR and SWIR reflectance observed in drought-treated plants (Fig. 1b). This observation supports a physiological basis for the optical signals distinguishing drought from well-watered treatments.

Hyperspectral reflectance data were collected in the VNIR (500–900 nm) and SWIR (950–1700 nm) regions, yielding 340 and 297 spectral measurements per sample, respectively. For each wavelength, reflectance spectra were averaged across all samples within each treatment (D or W), with standard deviations used to assess within-group variability. D plants consistently exhibited higher reflectance than W plants across both VNIR and SWIR regions (Fig. 1b), with similar standard deviations between treatments. This suggests that the spectral shifts observed were primarily driven by treatment effects rather than differences in spectral variability among replicates.

Principal component analysis (PCA) of spectral data further supported treatment-related differentiation. The first PC (PC1) explained 94.5 % of the total variance, with eigenvalues distributed evenly across VNIR and SWIR wavelengths, suggesting a general spectral response to drought stress rather than specific wavelength-driven effects (Fig. 1c-d). The second PC (PC2) captured an additional 3.7 % of the variance. While PC1 highlighted a trend towards separation between the D and W samples, overlap along this axis indicated substantial shared spectral variation between treatments.

Untargeted metabolite profiling using GC-MS identified 127 metabolites in poplar leaves across the 55 plants. PCA of metabolite data revealed treatment-related clustering, though with greater overlap than spectral data (Fig. 1e). Unlike spectral data, where PC1 explained a dominant 94.5 % of the variance, the first two metabolite PCA components explained only 22.3 % (PC1) and 14.9 % (PC2), capturing a combined 37.2 % of the total variance. This lower explained variance suggests that metabolite profiles are influenced by a broader range of factors beyond treatment effects alone. Contributions of individual metabolites to PC1 varied significantly, with eigenvalues ranging from −0.14–0.15 (Fig. 1f), reflecting the complex and diverse responses of metabolites to environmental conditions. This broad distribution of metabolite contributions may suggest that multiple metabolic pathways are involved, rather than a response dominated by a few key metabolites.

3.2. Metabolite changes in response to drought

To further assess metabolite responses to drought, ANOVA with FDR adjustment was applied to log-transformed metabolite data. Interaction effects between treatment and genotype were tested but were not statistically significant for all metabolites (FDR>0.05), indicating consistent treatment responses across genotypes. Therefore, our analysis focused on overall treatment effects (D vs. W).

Out of the 127 metabolites detected through untargeted GC-MS profiling, 73 exhibited significant treatment effects (FDR < 0.05), spanning amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and other secondary metabolites (Fig. 2a). Among them, 46 metabolites were elevated under drought conditions, whereas 27 decreased in D-treated plants (Fig. 2b, c; Table 1).

Across metabolite classes, AA exhibited the most pronounced increase under drought stress, with 12 metabolites increased and only one showing no change (Fig. 2a). CHO also responded strongly, with eight metabolites increasing and only one decreasing. PG and OA exhibited a more balanced response, with a similar proportion of increased and decreased metabolites, reflecting dynamic adjustments in primary and secondary metabolism. Notably, most FA and P remained unchanged, suggesting these classes were less responsive to drought stress. These results indicate differential metabolic reprogramming across classes, with AA, CHO, OA, and PG exhibiting the most substantial shifts in response to drought.

To further examine metabolite-specific responses to drought, we analyzed the $\log_2$ fold changes of individual metabolites between D and W conditions (Fig. 2b, c). Log_2 fold change values are calculated as $\log_2$(metabolite levels in D + 1) - $\log_2$(metabolite levels in W + 1), reflecting relative changes under drought stress. Among the significantly increased metabolites, AA accounted for a substantial proportion of elevated metabolites, with citrulline (3.09 $\log_2$ fold change), glutamic acid (3.02), and asparagine (2.41) displaying the largest increases. These AA are known to be integral to osmotic adjustment and nitrogen cycling, which are well-documented adaptive responses to drought stress (Abbas et al., 2023; Song et al., 2020; Zulfiqar et al., 2019). Other significantly increased AA included tryptophan (1.52), aspartic acid (1.50), glutamine (1.50), and glycine (1.49). Similarly, CHO metabolites such as raffinose (1.15), arabitol (0.82), and ribitol (0.81) increased, consistent with their known roles in osmoprotection (Vinson et al., 2020). In the PG class, p-coumaric acid 4-O-glucoside (0.93), purpurein (0.89), and grandidentatin (0.83) showed significant increases, likely supporting antioxidant defense. Additionally, terpenes (T), including α-tocopherol (1.38) and phytol (0.41), were significantly elevated, indicating their role in membrane stability and antioxidant protection during drought stress.

In contrast, 27 metabolites were significantly reduced under D conditions, with PG and OA metabolites being the most affected groups. Among PG metabolites, α-salicyloylsalicin (-1.56) and salicortin (-1.38) exhibited the largest decreases, suggesting a shift in secondary metabolism investment under unfavorable conditions. Similarly, OA metabolites, including threonic acid (-1.30) and shikimic acid (-1.16), declined under drought stress, indicating potential suppression of biosynthetic pathways involved in plant growth and defense mechanisms. Glucose (-0.33) was the only CHO metabolite significantly reduced under drought, likely reflecting increased energy demand and carbohydrate mobilization to support osmotic adjustment and other stress responses. This reduction in carbohydrate levels may be closely linked to impaired photosynthetic performance, as indicated by significant declines in Fv/Fm_Lss and qL_Lss (Supplementary Fig. S1). Reduced carbon assimilation and compromised electron transport efficiency under drought likely constrained primary metabolism, thereby limiting the availability of precursors for the synthesis of OA and secondary metabolites such as PG.

To investigate potential genotype-specific metabolic responses, PCA was conducted using the 73 metabolites with significant treatment effects (Fig. 2d). The analysis aimed to determine whether drought-tolerant genotypes (AF18 and AF34) exhibited distinct clustering compared to drought-susceptible genotypes. However, no distinct genotype-related clustering was observed, suggesting that variation in metabolite profiles was primarily driven by treatment effects rather than genetic differences among the tested genotypes. This indicates that while drought triggered broad biochemical reprogramming, genotype-specific responses were not clearly distinguished. This may be due to the limited genetic diversity among the tested commercial genotypes, which could have reduced the likelihood of detecting strong genotype-driven metabolic differentiation.

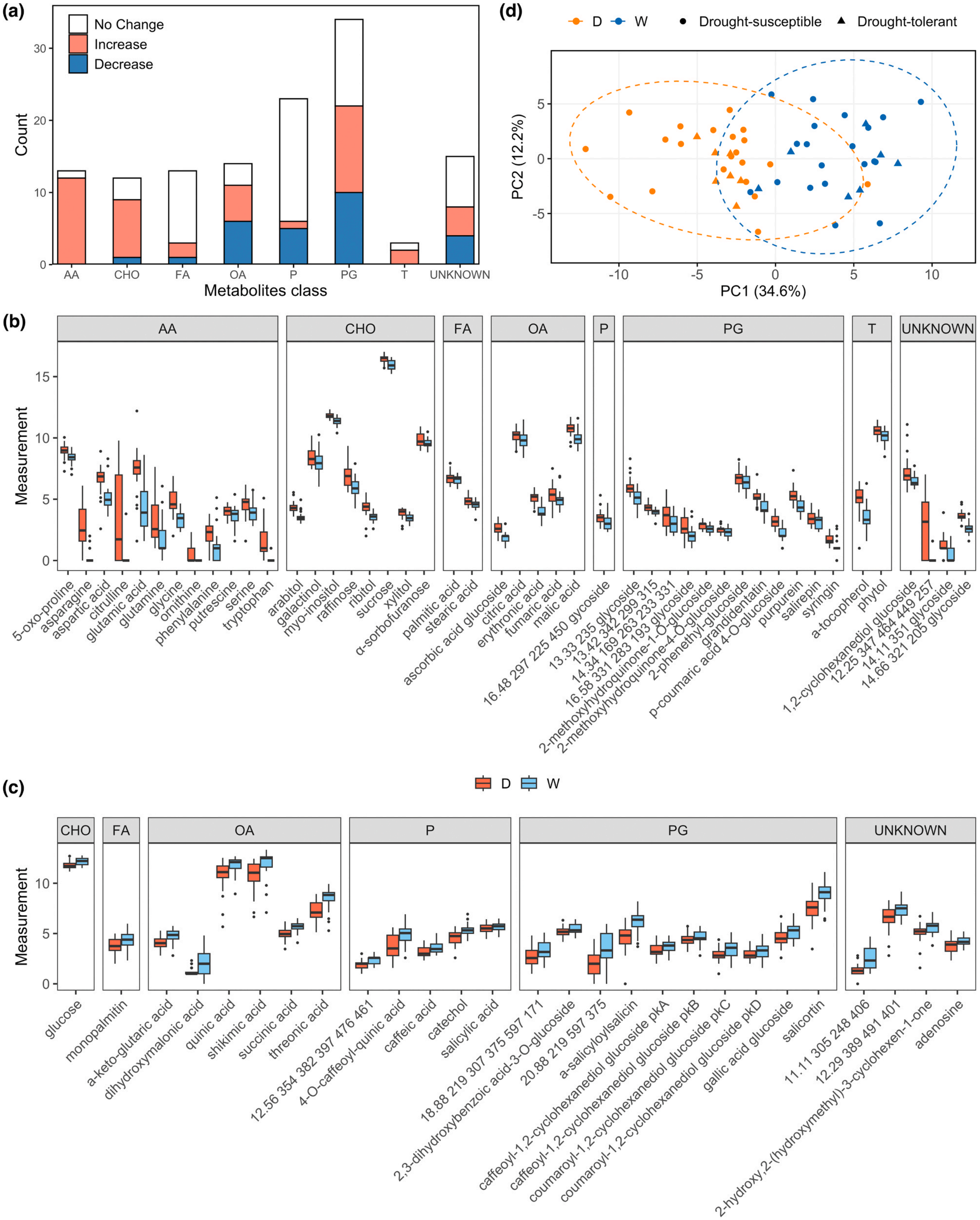

Fig. 2. (a) Bar plot showing the number of metabolites per class that exhibited a significant increase, decrease, or no change when comparing drought-treated (D) and well-watered (W) groups. A total of 73 metabolites with significant treatment effects were identified, including amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and unclassified compounds (UNKNOWN). (b) Box plot of the 46 metabolites that showed a significant increase in D, categorized by metabolite class. (c) Box plot of the 27 metabolites that showed a significant decrease in D, categorized by metabolite class. (d) The first two principal components (PC1 and PC2) of the 73 metabolites with significant treatment effects, with color indicating D (orange) and W (light blue) clusters and shape distinguishing drought-tolerant (circle) and drought-susceptible (triangle) genotypes.

Table 1
Metabolite changes in response to drought identified through ANOVA with false discovery rate (FDR) adjustment. Metabolites are classified into biochemical groups, including amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and unclassified compounds (UNKNOWN). Column D-W represents the $\log_2$ fold change of metabolite levels in drought-treated (D) plants compared to well-watered (W) plants, where positive values indicate increased abundance under drought and negative values indicate a decrease. FDR column reports the adjusted p-value with significant values ($\text{FDR} < 0.05$).

Metabolites	D-W	Class	FDR
citrulline	3.09	AA	7.11E−06
glutamic acid	3.02	AA	2.21E−11
12.25 347 464 449 257	2.83	UNKNOWN	1.31E−09
asparagine	2.41	AA	3.53E−12
a-tocopherol	1.52	S	2.67E−08
aspartic acid	1.51	AA	1.01E−08
tryptophan	1.50	AA	1.04E−07
glutamine	1.50	AA	4.17E−04
glycine	1.49	AA	1.31E−09
raffinose	1.15	CHO	3.27E−08
14.66 321 205 glycoside	1.09	UNKNOWN	1.19E−15
erythronic acid	1.09	OA	7.23E−13
phenylalanine	1.07	AA	3.17E−04
p-coumaric acid 4-O-glucoside	0.93	PG	1.71E−09
13.33 235 glycoside	0.92	PG	3.02E−06
purpurein	0.89	PG	3.84E−07
grandidentatin	0.83	PG	7.47E−07
ascorbic acid glucoside	0.83	OA	5.70E−11
arabitol	0.82	CHO	1.75E−11
ribitol	0.81	CHO	1.25E−07
malic acid	0.76	OA	1.31E−09
serine	0.73	AA	1.24E−03
1,2-cyclohexanediol glucoside	0.71	UNKNOWN	1.41E−03
16.58 331 283 193 glycoside	0.70	PG	8.21E−05
14.34 169 263 233 331	0.69	PG	5.53E−04
14.11 351 glycoside	0.63	UNKNOWN	1.85E−03
5-oxo-proline	0.62	AA	6.11E−06
ornithine	0.54	AA	1.69E−03
syringin	0.53	PG	3.46E−04
galactinol	0.48	CHO	2.96E−02
sucrose	0.48	CHO	1.40E−06
myo-inositol	0.45	CHO	3.84E−07
16.48 297 225 450 gycoside	0.45	P	3.45E−03
phytol	0.45	S	3.17E−04
fumaric acid	0.42	OA	1.53E−03
putrescine	0.42	AA	1.20E−02
2-methoxyhydroquinone−1-O-glucoside	0.39	PG	2.65E−05
salirepin	0.38	PG	1.15E−04
xylitol	0.38	CHO	2.98E−04
2-phenethyl-glucoside	0.38	PG	1.07E−02
13.42 342 299 315	0.37	PG	6.67E−07
citric acid	0.33	OA	5.81E−03
stearic acid	0.28	FA	6.31E−04
α-sorbofuranose	0.26	CHO	4.10E−02
2-methoxyhydroquinone−4-O-glucoside	0.17	PG	3.47E−02
palmitic acid	0.17	FA	3.27E−02
salicylic acid	−0.18	P	3.25E−02
caffeoyl−1,2-cyclohexanediol glucoside pkB	−0.23	PG	3.25E−02
coumaroyl−1,2-cyclohexanediol glucoside pkD	−0.31	PG	4.28E−02
2,3-dihydroxybenzoic acid−3-O-glucoside	−0.31	PG	2.95E−02
glucose	−0.33	CHO	2.87E−03
caffeic acid	−0.39	P	2.87E−03
caffeoyl−1,2-cyclohexanediol glucoside pkA	−0.45	PG	5.31E−04
12.56 354 382 397 476 461	−0.48	P	6.67E−07
adenosine	−0.50	UNKNOWN	2.59E−03
6-hydroxy−2-cyclohexenone alcohol	−0.57	UNKNOWN	1.70E−02
coumaroyl−1,2-cyclohexanediol glucoside pkC	−0.60	PG	1.69E−03
gallic acid glucoside	−0.61	PG	1.26E−04
monopalmitin	−0.61	FA	8.92E−04
18.88 219 307 375 597 171	−0.63	PG	1.44E−03
succinic acid	−0.68	OA	5.31E−06
a-keto-glutaric acid	−0.71	OA	4.76E−05
catechol	−0.78	P	4.76E−05
12.29 389 491 401	−0.92	UNKNOWN	2.11E−03
dihydroxymalonic acid	−0.95	OA	1.31E−05
quinic acid	−1.03	OA	2.24E−03
11.11 305 248 406	−1.12	UNKNOWN	2.06E−06
shikimic acid	−1.16	OA	5.68E−03
4-O-caffeoyl-quinic acid	−1.22	P	3.51E−06
threonic acid	−1.30	OA	1.09E−05
salicortin	−1.38	PG	6.67E−07
20.88 219 597 375	−1.39	PG	5.31E−06
a-salicyloylsalicin	−1.56	PG	2.35E−08

3.3. Metabolite-spectrum correlations

To investigate the relationships between leaf reflectance spectra and metabolite profiles, Pearson correlation analyses were performed across the full spectral range (VNIR and SWIR; 500–1700 nm) and all 127 identified metabolites. Analyses were performed for the full dataset and separately for D and W treatments.

In the full dataset, significant correlations ($\text{FDR} \leq 0.05$) were observed for 78 metabolites across various spectral regions, as visualized in a heatmap (Fig. 3a). Correlation coefficients (r) ranged from 0.31 to 0.79 in absolute value, with the most pronounced associations observed across both VNIR and SWIR regions. These metabolites spanned multiple biochemical classes, including amino acids (AA, 7 metabolites), carbohydrates (CHO, 6), fatty acids and alcohols (FA, 3), organic acids (OA, 12), phenolic metabolites (P, 12), phenolic glycosides (PG, 22), terpenes (T, 3), and 13 unclassified compounds. Positive correlations were predominant in AA, CHO, and T classes, whereas negative correlations primarily occurred with FA, OA, and P compounds. The PG class exhibited an equal occurrence of positive and negative correlations. Some metabolites displayed broad spectral associations, while others only exhibited correlations within specific spectral regions, suggesting wavelength-dependent sensitivity.

To identify the strongest metabolite-spectral associations, we examined correlations exceeding $|r| > 0.5$ ($\text{FDR} \leq 0.05$). Out of the 78 metabolites, 33 showed correlations with spectral data with $|r| > 0.5$ (Fig. 3b). Several metabolites exhibited strong correlations within specific spectral regions, while others showed consistent associations across the entire spectrum (e.g., shikimic acid and threonic acid at 500–1700 nm). Notably, arabitol exhibited the strongest positive correlation (average $r = 0.72$), while shikimic acid exhibited the strongest negative correlation (average $r = -0.72$) across the entire spectral range.

When analyzed separately by treatment, fewer metabolites showed significant correlations, although some had higher individual r values. Under D, significant correlations ($|r| > 0.5$; $\text{FDR} \leq 0.05$) ranged from 0.57 to 0.89 (Fig. 3c). Arabitol, fructose, and α-sorbofuranose, ascorbic acid, and maleic acid, as well as UNIDENTIFIED compounds 13.33 235 glycoside and 14.11 351 glycoside), 1,2-cyclohexanediol glucoside, and ethyl-phosphate (an extraction artefact correlated with phosphate) exhibited positive correlations. Only one partially-identified PG metabolite, 20.88 219 597 375, a coumaroyl-conjugate, showed a significant negative correlation.

Under W, more significant correlations were observed compared to the D treatment (Fig. 3d). Positively correlated metabolites included citrulline, xylitol, glycerol 1/3-phosphate, 16.48 297 225 450 glycoside, 2-phenethyl-glucoside, and coniferin, 13.33 235 glycoside and 14.11 351 glycoside, and ethyl-phosphate (extraction artefact correlated with phosphate). Negatively correlated metabolites included glutamine, quinic acid, shikimic acid, and threonic acid, as well as 2-methoxyhydroquinone-4-O-glucoside, salicin, and 2-hydroxy,2-(hydroxymethyl)-3-cyclohexen-1-one.

Despite significant metabolite-spectrum associations under both conditions, no clear treatment-specific spectral response patterns were identified. However, certain metabolites, including 13.33 235 glycoside,

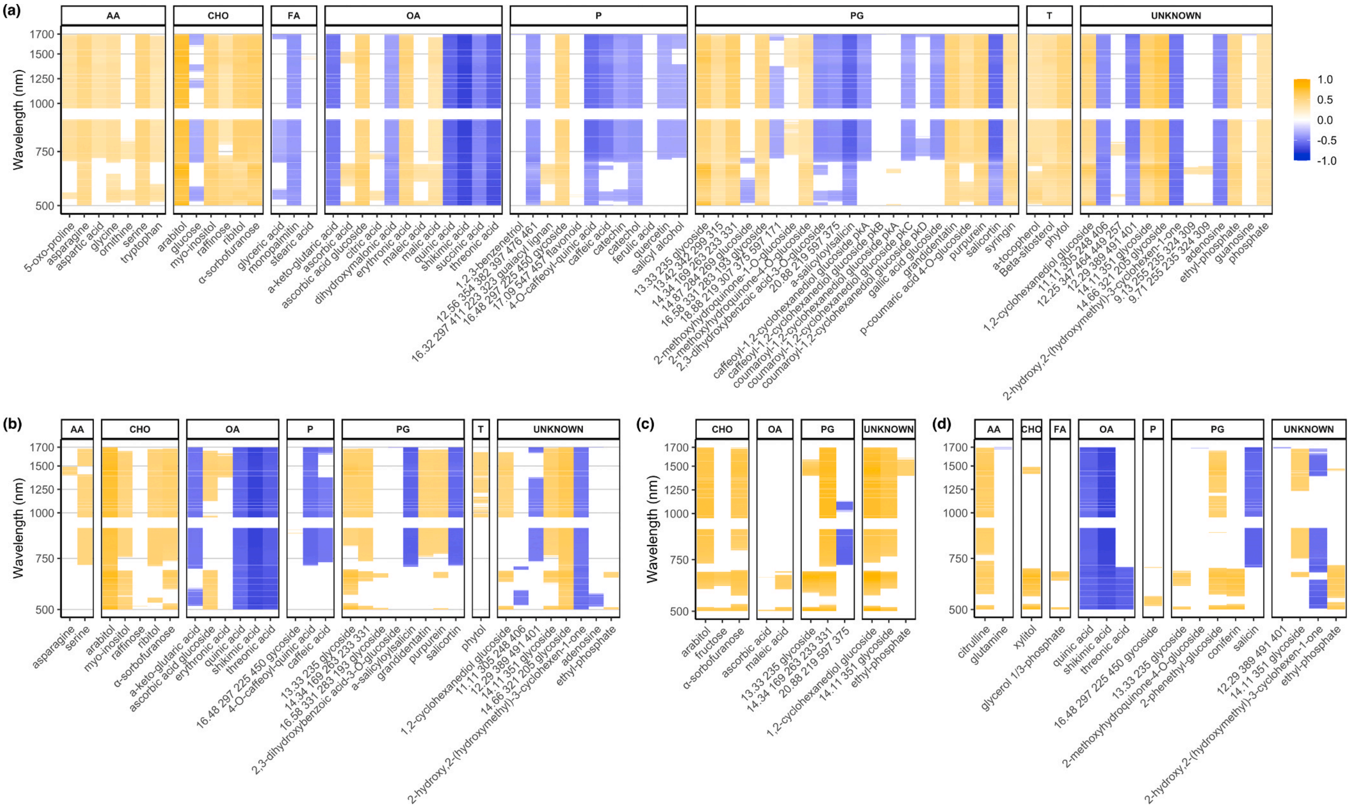

Fig. 3. (a) Heat map showing significant correlations (FDR < 0.05) between 78 metabolites and spectral data (500–1700 nm) across the entire drought-treated (D) and well-watered (W) datasets. Correlation coefficients (r) range from 0.31 to 0.79 (absolute value). (b) Heat map of 33 metabolites showing significant correlations (FDR < 0.05) with $|r| > 0.5$ for the entire D and W datasets. (c) Heat map of 11 metabolites showing significant correlations (FDR < 0.05) with $|r| > 0.5$ for the D treatment. (d) Heat map of 17 metabolites showing significant correlations (FDR < 0.05) with $|r| > 0.5$ for the W treatment.

a likely PG, 14.11 351 glycoside, and ethyl-phosphate, showed consistent positive correlations under both D and W treatments, suggesting a stable influence on leaf reflectance. In contrast, arabitol and α-sorbofuranose showed strong correlations exclusively under drought, while citrulline and shikimic acid were more strongly associated with spectral variation under W conditions. These findings highlight that while some metabolites serve as general indicators of reflectance variability, others may provide condition-specific insights into physiological or biochemical status.

3.4. Prediction accuracy and detection rates of metabolite classes using LASSO models

To assess the potential of hyperspectral reflectance data in predicting metabolite concentrations, LASSO regression models were applied to spectral data from the VNIR (500–900 nm), SWIR (950–1700 nm), and combined VNIR+SWIR regions. Model performance was evaluated using adjusted R^2 (adj-R^2) and standardized root mean square error (sRMSE), with 5-fold cross-validation repeated 20 times to ensure robust assessment. LASSO models were constructed under four analytical conditions: raw metabolite data with treatment (RT), raw data without treatment (R), $\log_2$-transformed data with treatment (LT), and $\log_2$-transformed data without treatment (L).

Across all models, 51 metabolites achieved adj-$R^2 > 0.3$, indicating moderate to strong prediction accuracy, and 15 of them exceeded adj-$R^2 = 0.5$ (Table 2, Fig. 4a). Among the best-performing models, top predictions included 1,2-cyclohexanediol glucoside (adj-$R^2 = 0.77$, VNIR), 13.33 235 glycoside (adj-$R^2 = 0.72$, VNIR), arabitol (adj-$R^2 = 0.70$, VNIR+SWIR), and ascorbic acid (adj-$R^2 = 0.66$, SWIR). These models also showed low sRMSE values (0.07–0.17), indicating high predictive reliability.

Among spectral sources, VNIR alone provided the highest adj-R^2 for 24 metabolites, SWIR for 18, and VNIR+SWIR for 9 metabolites (Table 2, Fig. 4b). Including treatment as a predictor improved adj-R^2 for 34 of the 51 metabolites, while $\log_2$-transformation enhanced prediction for 26 metabolites. These results highlight the predictive strength of VNIR spectra and the influence of data preprocessing choices on model performance.

Among the 51 well-predicted metabolites (adj-$R^2 > 0.3$), 43 also displayed significant changes under drought stress—30 of which increased and 13 decreased (Table 2). Notably, many of these drought-responsive metabolites, which also exhibited strong spectral correlations, are involved in osmotic adjustment or stress mitigation. For example, arabitol, a known sugar alcohol osmolyte, showed both high prediction accuracy (adj-$R^2 = 0.70$) and high average spectral correlation ($r = 0.72$), highlighting its clear optical signature. Similarly, shikimic acid, one of the most significantly reduced metabolites under drought, showed a strong negative correlation with spectral data ($r = -0.72$). These results demonstrate that spectral shifts reflect underlying biochemical changes relevant to drought physiology.

To further elucidate the spectral regions most relevant for metabolite prediction, detection rates were calculated across all wavelengths for the 51 metabolites with adj-$R^2 > 0.3$ (Fig. 4c). In the SWIR region, the highest detection frequencies occurred at 1695 nm (49.5 %) and 1698 nm (36.3 %), while 502 nm (30.9 %) and 698 nm (17.1 %) were most frequently selected in VNIR. These wavelengths represent key spectral regions that consistently contributed to metabolite predictions across multiple classes. Additional high-detection regions included the 600–700 nm range in VNIR and 1400–1500 nm in SWIR, suggesting their importance in capturing biochemical variability across different metabolite groups.

Detection rates also varied by metabolite class (Fig. 4d). The

Table 2

Performance summary of least absolute shrinkage and selection operator (LASSO) regression models for predicting metabolite concentrations from hyperspectral reflectance data. Metabolites are classified into biochemical groups, including amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and unclassified compounds (UNKNOWN). The highest adjusted R^2 (adj-R^2) value among the four LASSO model scenarios and three spectral sources is reported for each metabolite. Model performance metrics include adjusted R^2 (adj-R^2) and standardized root mean square error (sRMSE). LASSO models were run using three spectral sources: visible and near-infrared (VNIR), shortwave-infrared (SWIR), and VNIR+SWIR across four data transformation scenarios: raw metabolite data with treatment (RT), without treatment (R), $\log_2$-transformed data with treatment (LT), and without treatment (L). Drought response column indicates whether a metabolite significantly increased, decreased, or remained unchanged under drought conditions. Correlation column represents the average spectral correlation between metabolite abundance and hyperspectral reflectance.

Metabolites	Class	adj-R^2	sRMSE	Model	Source	Drought response	Correlation
1,2-cyclohexanediol glucoside	UNKNOWN	0.77	0.07	R	VNIR	Increase	0.48
13.33 235 glycoside	PG	0.72	0.07	RT	VNIR	Increase	0.55
arabitol	CHO	0.70	0.12	LT	VNIR+SWIR	Increase	0.72
ascorbic acid	OA	0.66	0.12	R	SWIR	No change	0.38
14.11 351 glycoside	UNKNOWN	0.64	0.10	R	VNIR	Increase	0.53
14.66 321 205 glycoside	UNKNOWN	0.62	0.11	RT	VNIR	Increase	0.62
erythronic acid	OA	0.59	0.15	LT	SWIR	Increase	0.48
14.34 169 263 233 331	PG	0.58	0.10	R	VNIR+SWIR	Increase	0.53
maleic acid	OA	0.58	0.11	RT	VNIR	No change	0.35
2-hydroxy,2-(hydroxymethyl)−3-cyclohexen−1-one	UNKNOWN	0.58	0.09	L	VNIR	Decrease	−0.54
tryptophan	AA	0.56	0.15	L	VNIR	Increase	0.33
shikimic acid	OA	0.54	0.16	R	SWIR	Decrease	−0.72
a-salicyloylsalicin	PG	0.52	0.11	L	VNIR	Decrease	−0.55
monogalactosylglycerol	FA	0.51	0.17	RT	VNIR+SWIR	No change	
a ±-sorbofuranose	CHO	0.50	0.13	RT	VNIR	Increase	0.55
myo-inositol	CHO	0.48	0.14	RT	SWIR	Increase	0.51
alpha-linolenic acid	FA	0.47	0.18	R	VNIR+SWIR	No change	
ribitol	CHO	0.47	0.12	RT	VNIR	Increase	0.50
12.29 389 491 401	UNKNOWN	0.47	0.12	L	VNIR+SWIR	Decrease	−0.48
threonic acid	OA	0.46	0.17	LT	VNIR	Decrease	−0.63
raffinose	CHO	0.46	0.12	RT	VNIR	Increase	0.37
phytol	S	0.45	0.13	RT	SWIR	Increase	0.48
ascorbic acid glucoside	OA	0.45	0.17	RT	VNIR	Increase	0.51
asparagine	AA	0.45	0.20	L	VNIR	Increase	0.46
salicortin	PG	0.44	0.14	L	SWIR	Decrease	−0.53
glycine	AA	0.44	0.16	LT	VNIR	Increase	0.38
quinic acid	OA	0.43	0.19	R	SWIR	Decrease	−0.59
sucrose	CHO	0.42	0.17	LT	SWIR	Increase	
adenosine	UNKNOWN	0.40	0.16	R	SWIR	Decrease	−0.47
purpurein	PG	0.40	0.13	LT	VNIR	Increase	0.50
4-O-caffeoyl-quinic acid	P	0.40	0.15	LT	VNIR	Decrease	−0.53
p-coumaric acid 4-O-glucoside	PG	0.40	0.16	LT	SWIR	Increase	0.39
5-oxo-proline	AA	0.39	0.10	R	VNIR+SWIR	Increase	0.38
12.25 347 464 449 257	UNKNOWN	0.38	0.24	LT	VNIR	Increase	0.31
caffeic acid	P	0.38	0.14	LT	VNIR+SWIR	Decrease	−0.48
grandidentatin	PG	0.37	0.13	LT	VNIR	Increase	0.49
11.11 305 248 406	UNKNOWN	0.36	0.17	LT	VNIR	Decrease	−0.47
glutamic acid	AA	0.34	0.18	LT	SWIR	Increase	
gallic acid glucoside	PG	0.33	0.14	R	VNIR	Decrease	−0.42
malic acid	OA	0.32	0.14	RT	VNIR+SWIR	Increase	0.34
caffeoyl−1,2-cyclohexanediol glucoside pkB	PG	0.32	0.14	RT	VNIR+SWIR	Decrease	−0.31
serine	AA	0.32	0.14	L	VNIR	Increase	0.49
glycerol	FA	0.32	0.18	RT	SWIR	No change	
aspartic acid	AA	0.32	0.17	LT	SWIR	Increase	0.35
guanosine	UNKNOWN	0.32	0.14	LT	SWIR	No change	−0.32
a-keto-glutaric acid	OA	0.31	0.19	LT	SWIR	Decrease	−0.53
a-tocopherol	S	0.31	0.19	LT	SWIR	Increase	0.36
linoleic acid	FA	0.31	0.16	RT	SWIR	No change	
monopalmitin	FA	0.31	0.15	LT	VNIR	Decrease	−0.41
citrulline	AA	0.31	0.22	LT	VNIR	Increase	
ethyl-phosphate	UNKNOWN	0.30	0.18	RT	SWIR	No change	0.45

1695 nm region in SWIR exhibited the highest class-specific detection rates for OA (70.8 %), T (70.0 %), CHO (60.0 %), P (50.0 %), AA (45.3 %), PG (44.7 %). Additionally, 1698 nm was a major detection region for OA (52.5 %), P (50.0 %), and AA (45.0 %). The strong spectral response of 1693 nm in SWIR is for AA (47.5 %), T (43.7 %), and PG (39.7 %). P (55.0 %) is the only category that showed a high detection rate at 1670 nm.

In the VNIR region, 698 nm exhibited the highest detection rates for AA (43 %) and T (42 %), reinforcing its potential role in nitrogen-related and membrane-associated metabolic processes. The 502 nm band was particularly important for detecting phenolic metabolites (66 %), while 697 nm showed the highest detection frequency for P (76 %), indicating this region's strong responsiveness to aromatic ring-containing molecular structures. The highest detection for unidentified metabolites occurred at 504 nm (43 %) in VNIR and 1695 nm (41 %) in SWIR, suggesting that these wavelengths are associated with strong, but as yet unclassified biochemical signals.

These results confirm that metabolite predictability from hyperspectral data is not only robust, but also closely aligned with drought-induced biochemical responses. Metabolites that accumulated under water deficit (e.g., arabitol) or declined (e.g., shikimic acid) were also those with the clearest and most stable optical signatures, indicating a direct biochemical basis for the spectral models. The VNIR region (500–700 nm) exhibited strong associations with AA and P, while the

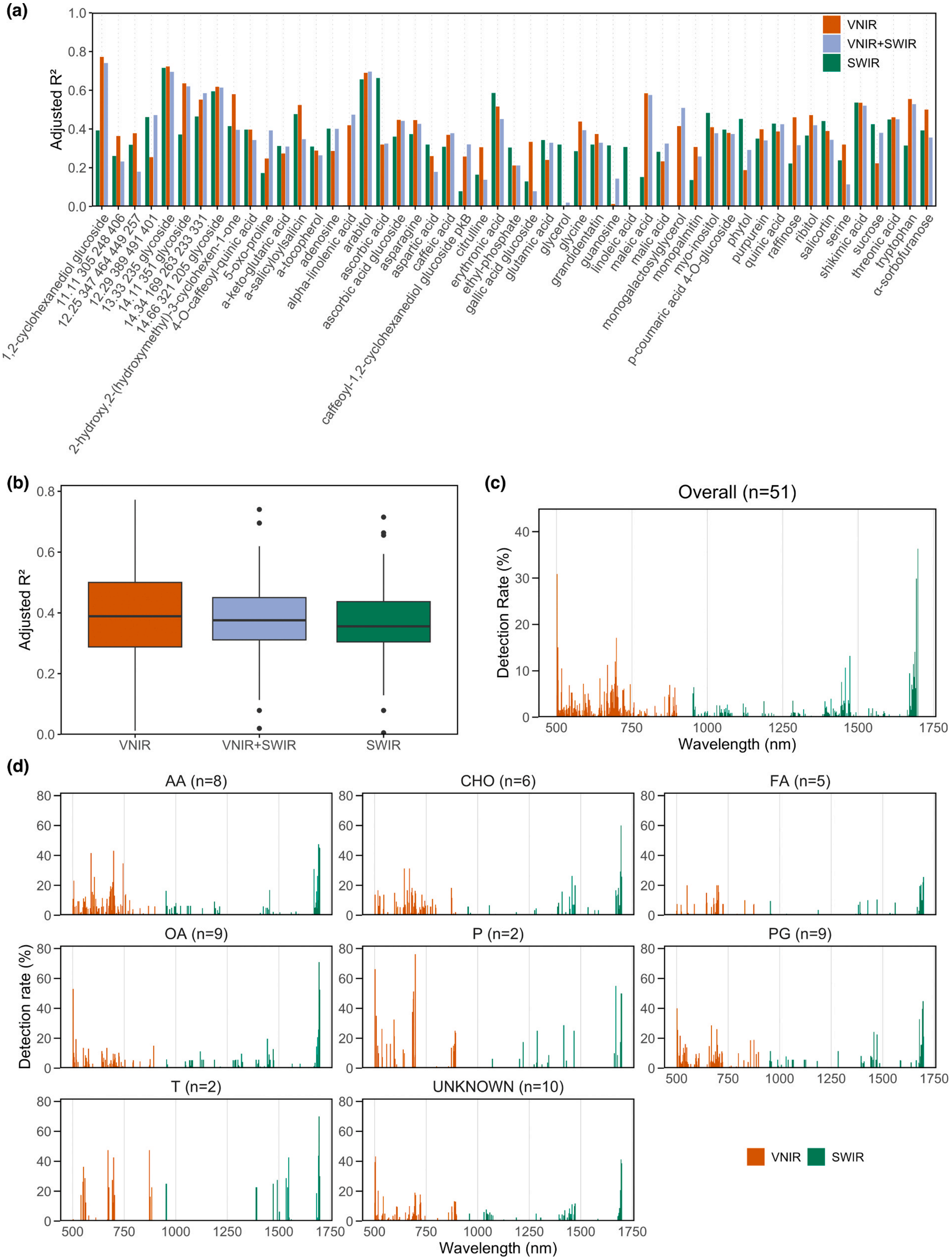

Fig. 4. (a) Bar plot showing the adjusted R^2 (adj-R^2) values from least absolute shrinkage and selection operator (LASSO) regression models for the 51 metabolites belonging to amino acids (AA), carbohydrates (CHO), phenolic glycosides (PG), organic acids (OA), fatty acids and alcohols (FA), terpenes (T), phenolic metabolites (P), and unclassified compounds (UNKNOWN) with adj-R^2 > 0.3, using visible and near-infrared (VNIR), shortwave-infrared (SWIR), and VNIR+SWIR spectral sources. (b) Box plot comparing the adj-R^2 values across VNIR, SWIR, and VNIR+SWIR for the 51 metabolites with adj-R^2 > 0.3, highlighting the distribution of prediction accuracy across spectral sources. (c) Overall detection rate of individual wavelengths across the 51 metabolites with adj-R^2 > 0.3, representing the frequency of wavelength selection in LASSO models over 20 iterations. (d) Class-specific detection rates, showing the proportion of metabolites within each biochemical class that were detected by key VNIR and SWIR wavelengths.

SWIR region (1680–1700 nm) demonstrated high predictive accuracy for OA, CHO, and T. The overlap between spectral detection and drought-induced metabolic responses underscores the feasibility of HSI as a non-destructive and accurate approach for monitoring plant physiological changes. By identifying key spectral wavelengths linked to biochemical composition, this study provides a framework for future targeted spectral modeling approaches, enabling efficient metabolic phenotyping under various environmental stress conditions.

4. Discussion

This study highlights the potential of VNIR+SWIR hyperspectral imaging (HSI) for detecting drought-induced metabolic shifts in *Populus*. By integrating spectral analysis, untargeted metabolomics, and predictive modeling, we established an accurate, non-destructive framework for monitoring drought stress in a bioenergy crop context. Clear spectral differences were observed between drought-treated (D) and well-watered (W) plants, with higher reflectance across VNIR and SWIR wavelengths under drought. These spectral shifts coincided with drought-responsive metabolic changes—most notably, increased amino acid (AA) accumulation and reduced carbohydrate (CHO) and phenolic glycoside (PG) concentrations.

4.1. Spectral and metabolic responses to drought

Drought stress induces well-documented structural and biochemical changes in plants that alter leaf optical properties and metabolism. In this study, D-treated plants exhibited higher reflectance in both VNIR and SWIR regions, consistent with previous findings that link water deficit to changes in leaf surface structure, internal anatomy, and pigment concentration (Serbin et al., 2012; Yendrek et al., 2017). These changes—such as reduced leaf hydration, cell wall thickening, and pigment degradation—are known to increase reflectance under stress (Ge et al., 2016). Furthermore, drought-treated plants exhibited significant reductions in fluorescence parameters such as Fv/Fm_Lss and qL_Lss, alongside increases in qP_Lss (Fig. S1), confirming the presence of physiological stress impairing photosynthetic function. The observed increase in qP_Lss has been documented in other plant species and is interpreted as a compensatory mechanism to sustain photochemical activity during stress conditions (Yin et al., 2024). The declines in fluorescence parameters are consistent with pigment loss and alterations in mesophyll structure, which together support the mechanistic basis for elevated VNIR/SWIR reflectance observed under drought stress.

Among the 127 detected metabolites, 73 exhibited significant treatment effects, with 46 increasing and 27 decreasing under drought. The most pronounced changes were observed in AA, CHO, and PG, reflecting a trade-off between osmotic adjustment, stress signaling, and energy storage. AA showed the most substantial increase, with citrulline, glutamic acid, and asparagine displaying the highest accumulation under drought. These compounds are well-known for their roles in osmoprotection, nitrogen metabolism, and drought tolerance, as they help maintain cellular turgor pressure and facilitate stress adaptation (Abbas et al., 2023; Song et al., 2020; Zulfiqar et al., 2019). In contrast, CHO and PG levels were generally reduced, suggesting a shift in metabolic allocation away from energy storage and secondary metabolite synthesis under stress. The decline in CHO such as glucose and raffinose likely reflects increased metabolic demand for osmoprotectants and enhanced carbohydrate catabolism to sustain cellular function during drought (Vinson et al., 2020). Similarly, reduced PG levels, including salicortin and α-salicyloylsalicin, suggest a lower investment in antioxidant defense pathways, possibly due to resource reallocation toward immediate stress mitigation.

Despite these pronounced metabolic shifts, genotype-specific clustering was not observed in either spectral or metabolite data, indicating that drought responses were primarily driven by treatment effects rather than genetic differences and/or the limited genetic diversity contained in the tested plant materials. One possible explanation is that the tested commercial clones share a narrow genetic base and similar physiological traits, limiting detectable genotype-specific variation. Future research incorporating a broader range of genetic materials will be essential to determine the extent of genotype-specific metabolic plasticity in response to drought. The present study underscores the potential of our approach, and it also sets the stage for larger-scale investigations encompassing broader poplar genetic material, and potentially in other environmental stress areas as well.

4.2. Linking spectral and metabolic data for non-destructive monitoring

Accurate, non-destructive prediction of plant biochemical composition from hyperspectral data represents a transformative approach for not only high-throughput phenotyping programs but also for precision forestry research. In this study, we demonstrated that hyperspectral reflectance data from VNIR (500–700 nm) and SWIR (1680–1700 nm) spectral regions, combined with LASSO regression modeling, accurately predicted the concentrations of a wide range of metabolites in *Populus* leaves. Among the 127 metabolites analyzed, 51 exhibited moderate to strong predictive accuracies (adj-$R^2 > 0.3$), including 15 with high accuracy (adj-$R^2 > 0.5$)—notably, 1,2-cyclohexanediol glucoside, arabitol, and ascorbic acid.

The strong predictive performance of VNIR spectra (500–700 nm) for specific AA and P aligns well with prior findings highlighting the sensitivity of VNIR reflectance to nitrogen-rich cellular constituents, including proteins and chlorophyll-associated nitrogenous compounds (Blackburn, 2007; Serbin et al., 2016). Asparagine and serine exhibited high spectral predictability, consistent with their central roles in osmotic adjustment and nitrogen remobilization under drought stress (Abbas et al., 2023; Hildebrandt et al., 2015). Additionally, phenolic compounds exhibited notable spectral associations at 502 nm and 698 nm, possibly reflecting changes in aromatic ring molecular structures and leaf pigments sensitive to drought-induced stress (Meacham-Hensold et al., 2019; Ustin et al., 2009).

In contrast, SWIR wavelengths, especially in the 1680–1700 nm range, provided strong predictive power for organic acids (e.g., ascorbic acid, shikimic acid), carbohydrates (e.g., arabitol, sucrose), and terpenes (e.g., α-tocopherol, phytol). This aligns with existing literature demonstrating that SWIR reflectance captures biochemical information linked to carbon-oxygen and carbon-hydrogen bonds common in carbohydrate and lipid structures (Curran, 1989; Ely et al., 2019; Kokaly and Skidmore, 2015). OA such as shikimic acid and threonic acid, critical intermediates in plant primary metabolism, showed strong negative spectral correlations, underscoring the sensitivity of SWIR wavelengths to metabolic alterations driven by drought stress (Serbin et al., 2016; Ustin et al., 2004). Specific CHO metabolites such as arabitol—an established osmolyte—demonstrated robust spectral predictability (adj-$R^2 = 0.70$) and a strong spectral correlation (average $r = 0.72$), confirming that our model captures physiologically meaningful carbohydrate-based osmoprotection under stress conditions (Silva-Perez et al., 2018; Vinson et al., 2020).

Incorporating drought treatment as a model covariate improved prediction accuracy for 34 metabolites, underscoring how environmental context enhances spectral-metabolite separability. Furthermore, applying $\log_2$-transformation to metabolite concentrations improved model accuracy for 26 metabolites. These findings highlight the value of considering environmental context and data normalization to strengthen spectral-metabolite predictive relationships, consistent with previous hyperspectral phenotyping studies (Burnett et al., 2021; Serbin et al., 2019).

4.3. Stable spectral-metabolite associations and their applications

A key advancement of this study is the identification of metabolites that exhibited stable spectral associations across both drought and well-watered conditions, highlighting their potential as robust, non-destructive indicators of metabolic status. Metabolites such as 13.33 235 glycoside (a likely PG), 14.11 351 glycoside (UNKNOWN), and ethyl-phosphate (reflecting phosphate level) consistently displayed strong positive correlations with spectral reflectance under both drought and well-watered conditions. The persistence of these spectral signatures across distinct moisture regimes implies their connection to inherent biochemical traits that are minimally influenced by environmental fluctuations. Such metabolites may thus serve as reliable spectral biomarkers for general growth assessment and metabolic status monitoring in precision agriculture and forestry (Serbin et al., 2019; Silva-Perez et al., 2018).

Conversely, several metabolites demonstrated pronounced treatment-specific spectral associations, highlighting their dynamic involvement in stress-specific metabolic adjustments. Under drought stress, carbohydrates including arabitol and α-sorbofuranose showed strong correlations with reflectance spectra, reflecting their critical roles in osmoprotection and stress adaptation (Vinson et al., 2020). The robust spectral association of these CHO specifically under drought conditions indicates their potential as diagnostic spectral markers for early drought detection. Conversely, AA, including citrulline, and OA, including shikimic acid, displayed stronger spectral associations under W conditions, suggesting active contributions to growth and homeostasis in the absence of stress. Such condition-dependent spectral-metabolite patterns are consistent with previously documented adaptive metabolic strategies in plants responding to fluctuating environmental pressures (Araus et al., 2018; Ely et al., 2019).

The combination of stable and condition-specific spectral-metabolite associations uncovered in this study provides a comprehensive spectral framework for plant stress monitoring. Stable spectral markers offer broad applicability for monitoring baseline metabolic health, while treatment-specific associations enhance precision in stress diagnosis and facilitate targeted physiological assessments (Burnett et al., 2021; Meacham-Hensold et al., 2019). In parallel, future work could benchmark alternative regression or machine-learning algorithms (ridge or elastic-net regression, random forests, neural networks) to determine whether modest gains in predictive accuracy justify sacrificing the sparsity and interpretability provided by the LASSO models used here. Leveraging these insights, future phenotyping efforts could develop targeted hyperspectral indices and sensor-based analytical platforms capable of discriminating specific metabolic responses to drought. Such targeted approaches would significantly improve the effectiveness of plant phenotyping platforms, aiding the selection and breeding of stress-resilient genotypes and optimizing management strategies in precision forestry.

5. Conclusions

Our results demonstrate that HSI in VNIR and SWIR regions when integrated with untargeted metabolite profiling and statistical modeling, effectively captures drought-induced metabolic responses in *Populus*. We identified specific spectral-metabolite associations, highlighting VNIR wavelengths (500–700 nm) as predictors of nitrogen-rich metabolites, particularly AA, and SWIR wavelengths (1680–1700 nm) as robust indicators for OA, CHO, and T. This study demonstrates the utility of HSI as a non-destructive, high-resolution phenotyping tool capable of accurately capturing biochemical shifts under drought stress in poplar.

Furthermore, by distinguishing metabolites exhibiting stable spectral correlations from those with environment-specific responses, this work advances our understanding of how spectral signatures relate to biochemical pathways underlying drought adaptation. These stable spectral biomarkers could enable real-time and accurate monitoring of plant health, whereas metabolites with environment-specific spectral patterns offer potential for real-time stress diagnostics in the field.

Future research should integrate more diverse genotypes and tailored multi-omics analyses, including genomics, transcriptomics, proteomics, and other multidimensional data to refine predictive models, ultimately advancing HSI as a real-time precision phenotyping tool for developing drought-resilient crops under changing climatic conditions. This approach should therefore be a successful strategy for high-throughput phenotyping of large breeding populations, thereby facilitating informed and expedited decision-making within poplar breeding programs. With high-quality data available to poplar breeders at an accelerated pace, this study offers a competitive advantage over conventional evaluation techniques. The success of coupling spectral phenomics, non-targeted metabolomics, and predictive modeling in poplar also suggests its application to other plant species, or at least forest crops.

Author contributions

M.S. conducted all data analyses, interpreted the results, prepared the figures, and wrote the manuscript. A.L.H. edited the manuscript. M. T., K.P, O.F.A, and M.M.C. conducted the drought experiment, performed imaging and phenotyping. J.H.L. cleaned the original hyperspectral data. A.L. and N.L.E. conducted metabolite profiling. J.-G.C. and G.A.T. provided general guidance on experimental design and edited the manuscript. T.J.T. conceived and supervised the project, conducted metabolomic analyses, and served as the corresponding author. All authors reviewed and approved the final manuscript.

CRediT authorship contribution statement

Antoine L. Harfouche: Writing – review & editing, Resources, Data curation. **Martin Trtílek:** Writing – review & editing, Data curation. **Gerald A. Tuskan:** Writing – review & editing, Conceptualization. **Timothy J. Tschaplinski:** Writing – review & editing, Funding acquisition, Conceptualization. **Mengjun Shu:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Miranda M. Clark:** Data curation. **Jin-Gui Chen:** Writing – review & editing. **Audrey Labbé:** Data curation. **Nancy L. Engle:** Methodology, Data curation. **John H. Lagergren:** Writing – review & editing, Methodology, Formal analysis. **Klára Panzarová:** Writing – review & editing, Formal analysis, Data curation. **Omar F. Alasia:** Resources, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This material is based upon work supported by the Center for Bioenergy Innovation (CBI), U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number ERKP886.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.envexpbot.2025.106218.

References

Abbas, K., Li, J., Gong, B., Lu, Y., Wu, X., Lü, G., Gao, H., 2023. Drought stress tolerance in vegetables: the functional role of structural features, key gene pathways, and exogenous hormones. Int. J. Mol. Sci. 24, 13876. https://doi.org/10.3390/ijms241813876.

Abdelhakim, L.O.A., Pleskačová, B., Rodriguez-Granados, N.Y., Sasidharan, R., Perez-Borroto, L.S., Sonnewald, S., Gruden, K., Vothknecht, U.C., Teige, M., Panzarová, K., 2024. High throughput image-based phenotyping for determining morphological and physiological responses to single and combined stresses in potato. J. Vis. Exp., e66255 https://doi.org/10.3791/66255.

Abdelhakim, L.O.A., Rosenqvist, E., Wollenweber, B., Spyroglou, I., Ottosen, C.-O., Panzarová, K., 2021. Investigating combined drought- and heat stress effects in wheat under controlled conditions by dynamic Image-Based phenotyping. Agronomy 11, 364. https://doi.org/10.3390/agronomy11020364.

Abebe, A.M., Kim, Y., Kim, J., Kim, S.L., Baek, J., 2023. Image-based high-throughput phenotyping in horticultural crops. Plants 12, 2061.

Abraham, P., Yin, H., Borland, A.M., Weighill, D., Lim, S.D., Cestari De Paoli, H., Engle, N.L., Agh, R., Weston, D.J., Wullschleger, S.D., Tschaplinski, T., Jacobson, D., Cushman, J.C., Hettich, R.L., Tuskan, G.A., Yang, X., 2016. Transcript, protein and metabolite temporal dynamics in the CAM plant *Agave*. Nature Plant 2, 16178. https://doi.org/10.1038/nplants.2016.178.

Anderson-Teixeira, K.J., Masters, M.D., Black, C.K., Zeri, M., Hussain, M.Z., Bernacchi, C. J., DeLucia, E.H., 2013. Altered belowground carbon cycling following land-use change to perennial bioenergy crops. Ecosystems 16, 508–520. https://doi.org/10.1007/s10021-012-9628-x.

Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19, 52–61.

Araus, J.L., Kefauver, S.C., Zaman-Allah, M., Olsen, M.S., Cairns, J.E., 2018. Translating high-throughput phenotyping into genetic gain. Trends Plant Sci. 23, 451–466. https://doi.org/10.1016/j.tplants.2018.02.001.

Awlia, M., Nigro, A., Fajkus, J., Schmoeckel, S.M., Negrão, S., Santelia, D., Trtílek, M., Tester, M., Julkowska, M.M., Panzarová, K., 2016. High-throughput non-destructive phenotyping of traits that contribute to salinity tolerance in arabidopsis thaliana. Front. Plant Sci. 7, 1414.

Balasubramanian, V.K., Veličković, D., Rubio Wilhelmi, M.D.M., Anderton, C.R., Stewart, C.N., DiFazio, S., Blumwald, E., Ahkami, A.H., 2024. Spatiotemporal metabolic responses to water deficit stress in distinct leaf cell-types of poplar. Front. Plant Sci. 15. https://doi.org/10.3389/fpls.2024.1346853.

Blackburn, G.A., 2007. Hyperspectral remote sensing of plant pigments. J. Exp. Bot. 58, 855–867. https://doi.org/10.1093/jxb/erl123.

Burnett, A.C., Anderson, J., Davidson, K.J., Ely, K.S., Lamour, J., Li, Q., Morrison, B.D., Yang, D., Rogers, A., Serbin, S.P., 2021. A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. J. Exp. Bot. 72, 6175–6189. https://doi.org/10.1093/jxb/erab295.

Clifton-Brown, J., Harfouche, A., Casler, M.D., Dylan Jones, H., Macalpine, W.J., Murphy-Bokern, D., Smart, L.B., Adler, A., Ashman, C., Awty-Carroll, D., Bastien, C., Bopper, S., Botnari, V., Brancourt-Hulmel, M., Chen, Z., Clark, L.V., Cosentino, S., Dalton, S., Davey, C., Dolstra, O., Donnison, I., Flavell, R., Greef, J., Hanley, S., Hastings, A., Hertzberg, M., Hsu, T.-W., Huang, L.S., Iurato, A., Jensen, E., Jin, X., Jørgensen, U., Kiesel, A., Kim, D.-S., Liu, J., McCalmont, J.P., McMahon, B.G., Mos, M., Robson, P., Sacks, E.J., Sandu, A., Scalici, G., Schwarz, K., Scordia, D., Shafiei, R., Shield, I., Slavov, G., Stanton, B.J., Swaminathan, K., Taylor, G., Torres, A.F., Trindade, L.M., Tschaplinski, T., Tuskan, G.A., Yamada, T., Yeon Yu, C., Zalesny Jr, R.S., Zong, J., Lewandowski, I., 2019. Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, miscanthus, willow and poplar. GCB Bioenergy 11, 118–151. https://doi.org/10.1111/gcbb.12566.

R. Core Team, 2013. R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. 30, 271–278. https://doi.org/10.1016/0034-4257(89)90069-2.

Ely, K.S., Burnett, A.C., Lieberman-Cribbin, W., Serbin, S.P., Rogers, A., 2019. Spectroscopy can predict key leaf traits associated with source–sink balance and carbon–nitrogen status. J. Exp. Bot. 70, 1789–1799.

Fahlgren, N., Gehan, M.A., Baxter, I., 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr. Opin. Plant Biol. 24, 93–99.

Fernie, A.R., Schauer, N., 2009. Metabolomics-assisted breeding: a viable option for crop improvement? Trends Genet. 25, 39–48. https://doi.org/10.1016/j.tig.2008.10.010.

Findurová, H., Veselá, B., Panzarová, K., Pytela, J., Trtílek, M., Klem, K., 2023. Phenotyping drought tolerance and yield performance of barley using a combination of imaging methods. Environ. Exp. Bot. 209, 105314. https://doi.org/10.1016/j.envexpbot.2023.105314.

Ge, Y., Bai, G., Stoerger, V., Schnable, J.C., 2016. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. Comput. Electron. Agric. 127, 625–632. https://doi.org/10.1016/j.compag.2016.07.028.

Hall, R., Beale, M., Fiehn, O., Hardy, N., Sumner, L., Bino, R., 2002. Plant metabolomics. Plant Cell 14, 1437–1440. https://doi.org/10.1105/tpc.140720.

Hastie, T., 2009. The elements of statistical learning: data mining, inference, and prediction.

Hildebrandt, T.M., Nunes Nesi, A., Araújo, W.L., Braun, H.-P., 2015. Amino acid catabolism in plants. Mol. Plant 8, 1563–1579. https://doi.org/10.1016/j.molp.2015.09.005.

Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G., Schaepman, M.E., 2013. Review of optical-based remote sensing for plant trait mapping. Ecol. Complex. 15, 1–16. https://doi.org/10.1016/j.ecocom.2013.06.003.

Jansson, S., Douglas, C.J., 2007. Populus: a model system for plant biology. Annu Rev. Plant Biol. 58, 435–458. https://doi.org/10.1146/annurev.arplant.58.032806.103956.

Kokaly, R.F., Skidmore, A.K., 2015. Plant phenolics and absorption features in vegetation reflectance spectra near 1.66μm. Int. J. Appl. Earth Obs. Geoinf. Spec. Issue Adv. Remote Sens. Veg. Funct. Andtraits 43, 55–83. https://doi.org/10.1016/j.jag.2015.01.010.

Kumar, M., Kumar Patel, M., Kumar, N., Bajpai, A.B., Siddique, K.H.M., 2021. Metabolomics and molecular approaches reveal drought stress tolerance in plants. Int J. Mol. Sci. 22, 9108. https://doi.org/10.3390/ijms22179108.

Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529, 84–87. https://doi.org/10.1038/nature16467.

Lowe, A., Harrison, N., French, A.P., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods 13, 80. https://doi.org/10.1186/s13007-017-0233-z.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., 2021. Climate change 2021: the physical science basis. Contrib. Work. Group I sixth Assess. Rep. Intergov. Panel Clim. Change 2, 2391.

Mateus, N., de, S., Florentino, A.L., Santos, E.F., Ferraz, A., de, V., Goncalves, J.L., de, M., Lavres, J., 2021. Partial substitution of k by na alleviates drought stress and increases water use efficiency in eucalyptus species seedlings. Front. Plant Sci. 12. https://doi.org/10.3389/fpls.2021.632342.

Meacham-Hensold, K., Fu, P., Wu, J., Serbin, S., Montes, C.M., Ainsworth, E., Guan, K., Dracup, E., Pederson, T., Driever, S., Bernacchi, C., 2020. Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging. J. Exp. Bot. 71, 2312–2328. https://doi.org/10.1093/jxb/eraa068.

Meacham-Hensold, K., Montes, C.M., Wu, J., Guan, K., Fu, P., Ainsworth, E.A., Pederson, T., Moore, C.E., Brown, K.L., Raines, C., Bernacchi, C.J., 2019. High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. Remote Sens. Environ. 231, 111176. https://doi.org/10.1016/j.rse.2019.04.029.

Moncholi-Estornell, A., Van Wittenberghe, S., Cendrero-Mateo, M.P., Alonso, L., Malenovský, Z., Moreno, J., 2022. Impact of structural, photochemical and instrumental effects on leaf and canopy reflectance variability in the 500–600 nm range. Remote Sens. 14, 56. https://doi.org/10.3390/rs14010056.

Montes, C.M., Fox, C., Sanz-Sáez, Á., Serbin, S.P., Kumagai, E., Krause, M.D., Xavier, A., Specht, J.E., Beavis, W.D., Bernacchi, C.J., Diers, B.W., Ainsworth, E.A., 2022. High-throughput characterization, correlation, and mapping of leaf photosynthetic and functional traits in the soybean (Glycine max) nested association mapping population. Genetics 221, iyac065. https://doi.org/10.1093/genetics/iyac065.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: a comprehensive review. Renew. Sustain. Energy Rev. 14, 578–597. https://doi.org/10.1016/j.rser.2009.10.003.

Oates, L.G., Duncan, D.S., Gelfand, I., Millar, N., Robertson, G.P., Jackson, R.D., 2016. Nitrous oxide emissions during establishment of eight alternative cellulosic bioenergy cropping systems in the north central United States. GCB Bioenergy 8, 539–549. https://doi.org/10.1111/gcbb.12268.

Obata, T., Fernie, A.R., 2012. The use of metabolomics to dissect plant responses to abiotic stresses. Cell. Mol. Life Sci. 69, 3225–3243. https://doi.org/10.1007/s00018-012-1091-5.

Pugesgaard, S., Schelde, K., Larsen, S.U., Lærke, P.E., Jørgensen, U., 2015. Comparing annual and perennial crops for bioenergy production – influence on nitrate leaching and energy balance. GCB Bioenergy 7, 1136–1149. https://doi.org/10.1111/gcbb.12215.

Raina, N., Chuetor, S., Elalami, D., Tayibi, S., Barakat, A., 2024. Biomass valorization for bioenergy production: current techniques, challenges, and pathways to solutions for sustainable bioeconomy. Bioenerg. Res. 17, 1999–2028. https://doi.org/10.1007/s12155-024-10792-x.

Ramachandra Reddy, A., Chaitanya, K.V., Vivekanandan, M., 2004. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. J. Plant Physiol. 161, 1189–1202. https://doi.org/10.1016/j.jplph.2004.01.013.

Schmer, M.R., Vogel, K.P., Varvel, G.E., Follett, R.F., Mitchell, R.B., Jin, V.L., 2014. Energy potential and greenhouse gas emissions from bioenergy cropping systems on marginally productive cropland. PLOS ONE 9, e89501. https://doi.org/10.1371/journal.pone.0089501.

Serbin, S.P., Dillaway, D.N., Kruger, E.L., Townsend, P.A., 2012. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. J. Exp. Bot. 63, 489–502. https://doi.org/10.1093/jxb/err294.

Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2016. Spectroscopic determination of leaf morphological and biochemical traits for Northern temperate and boreal tree species. Ecol. Appl. 24, 1651–1669. https://doi.org/10.1890/13-2110.1.

Serbin, S.P., Wu, J., Ely, K.S., Kruger, E.L., Townsend, P.A., Meng, R., Wolfe, B.T., Chlus, A., Wang, Z., Rogers, A., 2019. From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. N. Phytol. 224, 1557–1568. https://doi.org/10.1111/nph.16123.

Silva-Perez, V., Molero, G., Serbin, S.P., Condon, A.G., Reynolds, M.P., Furbank, R.T., Evans, J.R., 2018. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. J. Exp. Bot. 69, 483–496. https://doi.org/10.1093/jxb/erx421.

Song, Q., Joshi, M., DiPiazza, J., Joshi, V., 2020. Functional relevance of citrulline in the vegetative tissues of watermelon during abiotic stresses. Front. Plant Sci. 11, 512.

Sun, J., Xu, J., Qiu, C., Zhai, J., Zhang, S., Zhang, X., Wu, Z., Li, Z., 2024. The chromosome-scale genome and population genomics reveal the adaptative evolution of populus pruinosa to desertification environment. Hortic. Res. 11, uhae034. https://doi.org/10.1093/hr/uhae034.

Taylor, G., Donnison, I.S., Murphy-Bokern, D., Morgante, M., Bogeat-Triboulot, M.-B., Bhalerao, R., Hertzberg, M., Polle, A., Harfouche, A., Alasia, F., Petoussi, V., Trebbi, D., Schwarz, K., Keurentjes, J.J.B., Centritto, M., Genty, B., Flexas, J., Grill, E., Salvi, S., Davies, W.J., 2019. Sustainable bioenergy for climate mitigation: developing drought-tolerant trees and grasses. Ann. Bot. 124, 513–520. https://doi.org/10.1093/aob/mcz146.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 58, 267–288.

Ting, T.-C., Souza, A.C., Imel, R.K., Guadagno, C.R., Hoagland, C., Yang, Y., Wang, D.R., 2023. Quantifying physiological trait variation with automated hyperspectral imaging in rice. Front. Plant Sci. 14, 1229161.

Tschaplinski, T., Tuskan, G., 1994. Water-stress tolerance of black and eastern cottonwood clones and four hybrid progeny. Ii. metabolites and inorganic ions that constitute osmotic adjustment. Can. J. For. Res. 24, 681–687.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.-L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Déjardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjärvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J.-C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouzé, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.-J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., Rokhsar, D., 2006. The genome of black cottonwood, populus trichocarpa (Torr. & Gray). Science 313, 1596–1604. https://doi.org/10.1126/science.1128691.

Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A., Zarco-Tejada, P., 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sens. Environ. Imaging Spectrosc. Spec. Issue 113, S67–S77. https://doi.org/10.1016/j.rse.2008.10.019.

Ustin, S.L., Roberts, D.A., Gamon, J.A., Asner, G.P., Green, R.O., 2004. Using imaging spectroscopy to study ecosystem processes and properties. BioScience 54, 523–534. https://doi.org/10.1641/0006-3568(2004)054[0523:UISTSE]2.0.CO;2.

Vergara-Diaz, O., Vatter, T., Kefauver, S.C., Obata, T., Fernie, A.R., Araus, J.L., 2020. Assessing durum wheat ear and leaf metabolomes in the field through hyperspectral data. Plant J. 102, 615–630. https://doi.org/10.1111/tpj.14636.

Viger, M., Smith, H.K., Cohen, D., Dewoody, J., Trewin, H., Steenackers, M., Bastien, C., Taylor, G., 2016. Adaptive mechanisms and genomic plasticity for drought tolerance identified in european black poplar (Populus nigra L. Tree Physiol. 36, 909–928. https://doi.org/10.1093/treephys/tpw017.

Vinson, C.C., Mota, A.P.Z., Porto, B.N., Oliveira, T.N., Sampaio, I., Lacerda, A.L., Danchin, E.G.J., Guimaraes, P.M., Williams, T.C.R., Brasileiro, A.C.M., 2020. Characterization of raffinose metabolism genes uncovers a wild arachis galactinol synthase conferring tolerance to abiotic stresses. Sci. Rep. 10, 15258. https://doi.org/10.1038/s41598-020-72191-4.

Wickham, H., 2011. ggplot2. Wiley Interdiscip. Rev. Comput. Stat. 3, 180–185.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130.

Yendrek, C.R., Tomaz, T., Montes, C.M., Cao, Y., Morse, A.M., Brown, P.J., McIntyre, L. M., Leakey, A.D.B., Ainsworth, E.A., 2017. High-Throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. Plant Physiol. 173, 614–626. https://doi.org/10.1104/pp.16.01447.

Yin, X., Wang, H., Wang, C., Wang, S., Ren, X., Zong, Z., 2024. Effect of watering gradient on chlorophyll fluorescence parameters of prunus domestica trees. Agronomy 14, 2373. https://doi.org/10.3390/agronomy14102373.

Zagorščak, M., Abdelhakim, L., Rodriguez-Granados, N.Y., Široká, J., Ghatak, A., Bleker, C., Blejec, A., Zrimec, J., Novák, O., Pěnčík, A., Baebler, Š., Perez Borroto, L., Schuy, C., Županič, A., Afjehi-Sadat, L., Wurzinger, B., Weckwerth, W., Pompe Novak, M., Knight, M.R., Strnad, M., Bachem, C., Chaturvedi, P., Sonnewald, S., Sasidharan, R., Panzarová, K., Gruden, K., Teige, M., 2025. Integration of multi-omics data and deep phenotyping provides insights into responses to single and combined abiotic stress in potato. Plant Physiol., kiaf126 https://doi.org/10.1093/plphys/kiaf126.

Zulfiqar, F., Akram, N.A., Ashraf, M., 2019. Osmoprotection in plants under abiotic stresses: new insights into a classical phenomenon. Planta 251, 3. https://doi.org/10.1007/s00425-019-03293-1.