Plant Phenomics 8 (2026) 100174

Contents lists available at ScienceDirect

Plant Phenomics

journal homepage: www.sciencedirect.com/journal/plant-phenomics

Research Article

Unraveling plant phenotype to genotype associations with daily hyperspectral traits in *Populus trichocarpa*☆

Marie C. Klein [a], Christopher YS. Wong [a,3], J. Grey Monroe [a], Jack Bailey-Bale [a], Thomas N. Buckley [a], Jin-Gui Chen [b], Mengjun Shu [b], Timothy J. Tschaplinski [b], Gerald A. Tuskan [b], Troy S. Magney [a,*,1], Gail Taylor [a,**,2]

[a] Department of Plant Sciences, University of California Davis, Davis, CA, 95616, USA
[b] Biosciences Division and the Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

ARTICLE INFO

Keywords:
Hyperspectral remote sensing
Field-based
Hyperspectral GWAS
Poplar
Vegetation indices
Heritability
Chlorophyll
Carotenoids
Chloroplasts
Genetic basis of hyperspectral traits

ABSTRACT

Hyperspectral remote sensing is a powerful, high-throughput phenotyping tool that quantifies physiologically and structurally relevant wavelengths across diverse genotypes and over varying temporal scales. In this study, we combined tower-based continuous hyperspectral sensing with genome-wide association studies to analyze 1423 wavebands (400-900 nm) and derivative vegetation indices across 505 genotypes and the genetic architecture of hyperspectral phenotypes over time in *Populus trichocarpa Torr. & Gray* grown under field conditions. Wavelengths related to chlorophyll and carotenoid absorption spectra exhibited the strongest genetic variation resulting in 98 significant SNP associations. Notably, we found substantial overlap in genetic association between the blue and red spectral regions, indicative of carotenoids and chlorophyll, respectively, and identified more than 10 candidate genes associated with chloroplast function, underpinning photosynthetic activity. Furthermore, fluctuations in associations for vegetative indices, such as the chlorophyll:carotenoid index (CCI), across the growing season reveal a temporally dynamic genetic architecture of physiological traits associated with fall senescence of this temperate tree species. Finally, we also observed correlations (spearman rho = 0.3, $p < 1x10^{-8}$) between individual wavebands or vegetative indices and growth rate, assessed as the relative change of tree height over the growing season. The growth rate prediction was substantially improved by a regularization multivariate model (spearman rho>0.5, $p < 1x10^{-16}$), reinforcing the value of hyperspectral measurements for predicting traits linked to tree productivity. These findings highlight the potential of high-throughput, rapid, hyperspectral genome wide association studies GWAS to uncover physiologically meaningful genetic variation and offer promising insights for future acceleration for plant breeding.

1. Introduction

Populus trichocarpa Torr. & Gray (poplar) has significant large-scale potential as a lignocellulosic feedstock for biofuels including Sustainable Aviation Fuels (SAFs), due to its rapid growth, genetic diversity, and established use as a model tree species [1–3]. With a fully sequenced 500 Mb genome, extensive research has elucidated the genetic basis of key phenotypes of poplar to improve traits such as growth rate, biomass yield, stress tolerance, and wood quality to increase the suitability as a bioenergy feedstock [3–10].

However, despite advances, the ability to efficiently connect genotypes to desirable phenotypic traits, such as yield and stress tolerance, is hindered by phenotyping, particularly under field conditions, which

☆ Notice: 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 and by the Genomics enabled Plant Biology for Determination of Gene Function programme by the Office of Biological and Environmental Research in the DOE Office of Science (DE-SC0020164).

* Corresponding author.
** Corresponding author.
E-mail addresses: troy.magney@umontana.edu (T.S. Magney), gail.taylor@ucl.ac.uk (G. Taylor).
[1] Current address: Department of Forest Management, W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT.
[2] Current address: Department of Genetics, Evolution and Environment, UCL, Gower Street, London WC1E 6AE
[3] Current address: Faculty of Forestry and Environmental Management, University of New Brunswick, Fredericton, NB, Canada.

https://doi.org/10.1016/j.plaphe.2026.100174
Received 17 September 2025; Received in revised form 18 January 2026; Accepted 21 January 2026
Available online 27 March 2026

Abbreviations

CCI	Chlorophyll: Carotenoid index
COV	Coefficient of variation
GWAS	Genome-wide association studies
NDVI	Normalized Difference Vegetation Index
NIR	Near- Infrared
PRI	Photochemical Reflectance Index
QTL	Quantitative Trait Locus
RGR	Relative growth rate
SD	Standard deviation
SIF	Solar-Induced fluorescence
SNP	Single nucleotide polymorphism
SNV	Standard normal variate transformation
SVI	Stem volume index

remains a major bottleneck in crop improvement [11–13]. For trees in particular, phenotyping efforts are destructive and limited in scope, constrained by time-intensive labor and the challenge of capturing dynamic plant traits over time or across large populations of tall and difficult-to-access plants. High-throughput phenotyping technologies which leverage spectroscopic techniques provide a strategy to quickly and nondestructively measure biochemical and physiological traits on a large scale, overcoming many of the phenotyping bottlenecks [14–16].

Remote sensing has been used for over 50 years for ecosystem monitoring and modeling applications [17], and recently high-throughput phenotyping to study genetic diversity has been explored consequent on user-friendly remote sensing platforms with high spatial, temporal, and spectral resolutions [18,19], but limited studies have been undertaken on tree populations.

Optical-based remote sensing measures vegetation reflectance from the visible (400-700 nm) and near-infrared (NIR, 700-1100 nm) spectral regions and variation in specific spectral regions enables inferences about vegetation traits and functions. The visible region is sensitive to chlorophyll and carotenoid pigment dynamics [20,21], whilst the NIR region is sensitive to structural dynamics such as leaf thickness and density [22]. Specific wavebands can be used as proxies of vegetation traits and functions [23] such as tree health, biomass, photosynthetic efficiency, and stress responses [24]. For example, the normalized difference vegetation index (NDVI) is sensitive to both leaf area and chlorophyll content - or total “greenness”, whilst Photochemical Reflectance Index (PRI) and Carotenoid/Chlorophyll Index (CCI) enable estimates of photosynthetic activity, light-use efficiency, and photoprotection [25,26]. Beyond hyperspectral sensing and vegetation indices, fine wavelengths (<2 nm) across the entire visible and NIR spectrum [27] are now being used to deduce novel insights into plant-based traits, alongside machine learning models that advantage of the entire spectrum to predict a suite of biochemical and photosynthetic parameters [28–30].

Recently, ultra-resolution (<1 nm spectral resolution) remote sensing systems have enabled the retrieval of solar-induced chlorophyll fluorescence (SIF), offering a physiology-based signal tied to chlorophyll fluorescence, which is linked to photosynthetic activity and non-photochemical quenching [31]. To date, Its main applications have been large-scale ecosystem monitoring via satellites across seasons and stress events [32,33]. Only recently, there have been advances in SIF applications on tower and airborne platforms in precision agriculture and high-throughput phenotyping [34–39]. Given its relatively new use in phenotyping, there is much ongoing work focused on teasing apart SIF sensitivity to physiology and structure [40–43]. Ultimately, SIF and other optical-based metrics to assess a suite of plant traits and functions in high-throughput phenotyping and screening genetic population trials for optimizing stress resilience and growth.

The integration of hyperspectral imaging with genome-wide association studies (GWAS) is an emerging field that connects high-throughput phenotyping with advanced genetic analysis [44]. GWAS is a powerful tool for identifying genetic variants associated with complex traits by analyzing genetic diversity across populations and testing for associations between genetic markers and trait variation [45]. With hyperspectral phenotyping, each spectral band or vegetation index (and its temporal profile) can be treated a phenotype. This enables genetic mapping of canopy-scale ‘endophenotypes’ that are closer to underlying physiology (e.g., pigments, water status, photosynthetic function) and can reveal loci whose effects are transient or stress-dependent over the season. A few studies have explored this combination to identify genetic regions linked to important agronomic traits, such as seed vigor, drought tolerance, yield, and photosynthetic efficiency [16,44,46–48].

However, these studies have largely focused on single time-point measurements, primarily under controlled conditions (e.g., greenhouses) and rarely using trees, without the added complexity of field environments or longitudinal measurements. Consequently, the combination of hyperspectral imaging with GWAS across continuous time periods in field conditions remains largely unexplored. In this study, we employed a tower-based hyperspectral camera covering the 400 to 900 nm spectral range to measure 595 *Populus* trees (505 genotypes). Combining field hyperspectral remote sensing with GWAS, we set out to uncover the genetic architecture of hyperspectral traits. We aimed to address the following primary questions: 1) What is the heritability of hyperspectral traits? 2) What is the genetic architecture of individual wavebands, and which wavebands exhibit the strongest signals in GWAS? 3) How does the genetic architecture of vegetative indices change over time, as evidenced by autocorrelations in GWAS signals? 4) How do these indices correlate with yield (biomass), and what is their potential relevance for breeding applications?

2. Results and discussion

2.1. Seasonal and genotypic variation in spectral reflectance and vegetation indices

Spectral reflectance exhibited strong seasonal change over time and across genotypes within the drought plot (target environment) (Fig. 1). Over time, from August to October, the visible region reflectance shows a small increase, most notably between 550 nm and 700 nm. In addition, the standard deviation gets larger over time, indicating more variation across genotypes later in the growing season. The coefficient of variation (COV, SD/mean) per day also shows seasonal changes (Fig. 1). From August to September, COV shows a distinct pattern with high variation in the visible region (400 to 700 nm) and relatively low variation in the NIR (700 to 900 nm). By October, the pattern becomes more subdued, exhibiting a more constant COV across all visible and NIR wavebands.

Vegetation indices (PRI, CCI, NDVI) and SIF showed strong seasonal and genotype variation (Fig. 2). All variables exhibited a seasonal decline from August to October. Across genotypes, variation could be observed within days and across days based on absolute values and timing of change. PRI and CCI are sensitive to carotenoid:chlorophyll ratios, where more negative values indicate more carotenoid relative to chlorophyll pools (Fig. 2c and d). NDVI, a proxy of greenness and chlorophyll, represents more general changes in the canopy, where lower values represent less green foliage (Fig. 2e). SIF is a proxy of photosynthetic activity, where lower values indicate less photosynthetic activity (Fig. 2f). We note that there are missing data gaps in the spectral data due to instrument error or sky conditions (Fig. 2a).

2.2. Genetic architecture of hyperspectral phenotypes

Heritability (H^2) estimates across the visible and near-infrared spectrum highlight key genetic influences on photosynthesis-related traits. These values reflect broad-sense heritability, encompassing

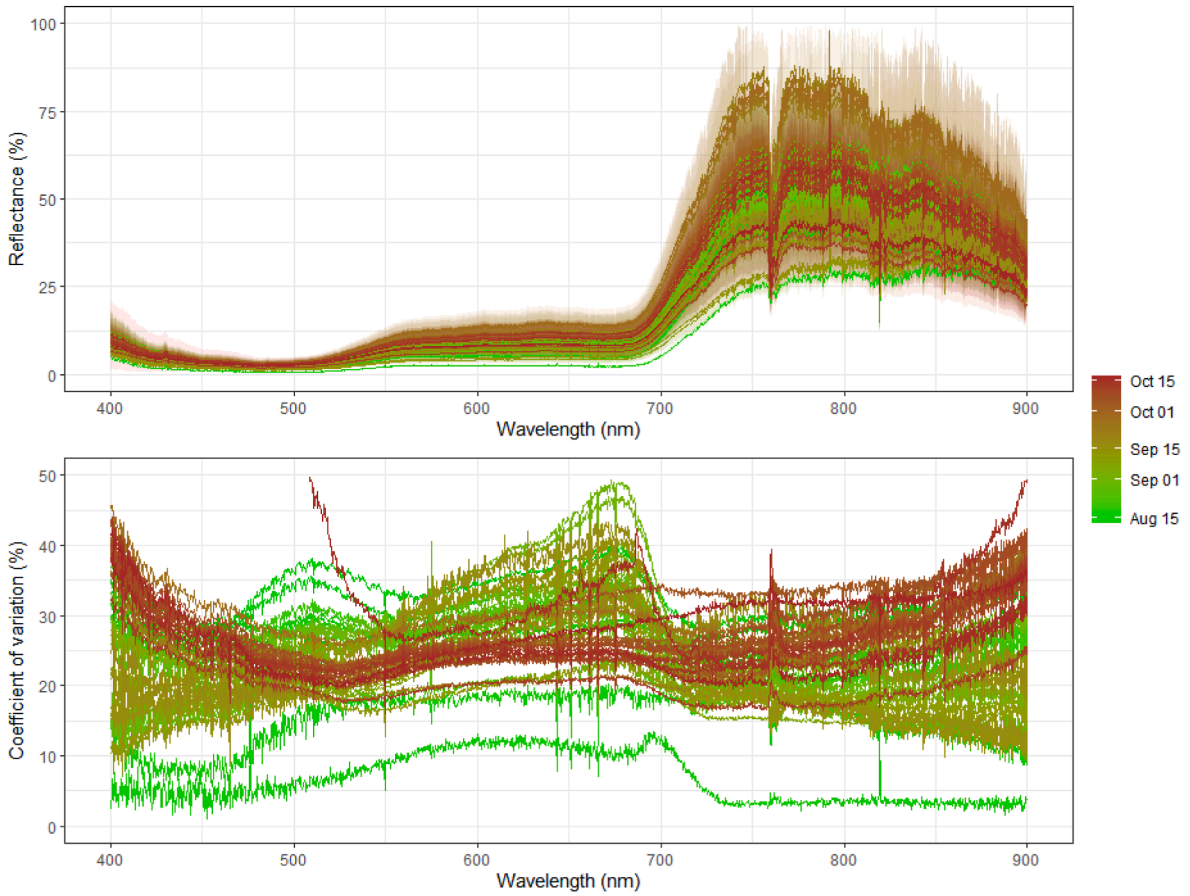

Fig. 1. Mean midday hyperspectral reflectance and coefficient of variation (standard deviation/mean) of all genotypes per date. Each line represents the date, and the shaded region represents the standard deviation of all genotypes.

genetic variance and potential genotype-by-date interactions under drought progression. The increase in heritability between 400 and 500 nm corresponds to the blue region, where chlorophyll *b* absorbs light (Fig. 3), suggesting that genetic variation influencing light capture and photosynthetic efficiency is significant in this region. The peak around 680-700 nm, corresponding to the red absorption of chlorophyll *a*, is consistent with genetic control over the primary processes of photosynthesis. In contrast, the near-infrared range (700+ nm) shows declining heritability, potentially reflecting greater environmental influence or structural variation less tied to genetic control. Among vegetative indices, CCI and NDVI exhibited the highest heritability among all four indices, with H^2 ranging from 0.2 to 0.25, indicating genetic control over chlorophyll content. CCI, which captures both carotenoid (531 nm) and chlorophyll (645 nm) absorption, had the highest heritability and may offer a more comprehensive measure of plant physiological status than NDVI, which focuses solely on chlorophyll red absorption around 670-690 nm and NIR reflectance (750-900 nm).

The left panel shows the heritability (H^2) across the visible and near-infrared spectrum (400-900 nm) for individual wavebands averaged across time points. Heritability estimates increase between 400 and 500 nm and around 600-700 nm, peaking at 700 nm, corresponding to chlorophyll *a*'s red absorption spectrum. The right panel displays the heritability of several vegetative indices, including CCI (Chlorophyll: Carotenoid Index), PRI (Photochemical Reflectance Index), NDVI (Normalized Difference Vegetation Index) and SIF (Solar Induced Fluorescence).

The genetic architecture of hyperspectral traits across different wavelengths was examined using GWASfor the genotype means of each of the 1423 bands to identify putative SNP associations. Temporal variation in GWAS helps pinpoint how genetic control shifts across developmental stages and with respect to environmental conditions, which also shift over time. We observed substantially more SNPs associated with bands centered around the blue (450–500 nm) and red (650–700 nm) regions of the spectrum (Fig. 4). This pattern persisted across different levels of significance filtering in the GWAS results, although it became more pronounced when restricted to highly significant variants ($P < 1x10^{-6}$) (Fig. 4a). After filtering for LOD score significance ($P < 1x10^{-6}$) and enrichment for significant SNPs in high LD (Materials and Methods, chi-squared test $P < 0.05$, Bonferroni-adjusted), we identified a total of 98 SNPs associated with bands in the potential chlorophyll absorption spectra (400-700 nm) (Supplemental Table S4). Notably, the distribution of these SNPs mirrored the heritability estimates from our previous analyses (Fig. 3), consistent with a genetic basis for traits associated with these specific wavelengths linked to photosynthesis and chloroplast function. SNP association density followed the pattern of H^2 across wavelengths, with more associations in high-heritability spectral regions. This supports the hypothesis of a polygenic basis for pigment-related reflectance traits. These associations may include regulatory variants influencing pigment biosynthesis or plastid development, rather than direct coding mutations. Most associations displayed moderate to small effect sizes, consistent with expectations for complex traits governed by many small-effect loci.

We found substantial overlap between SNPs associated with blue and red wavelengths (Fig. 4b and c). This overlap may point to shared genetic mechanisms affecting the broader functioning of chloroplasts, as these wavelengths relate to chlorophyll absorption. Specifically, blue and red light is primarily absorbed by chlorophyll *b* and *a*, respectively, suggesting that the SNP unique to blue or red wavelengths could be

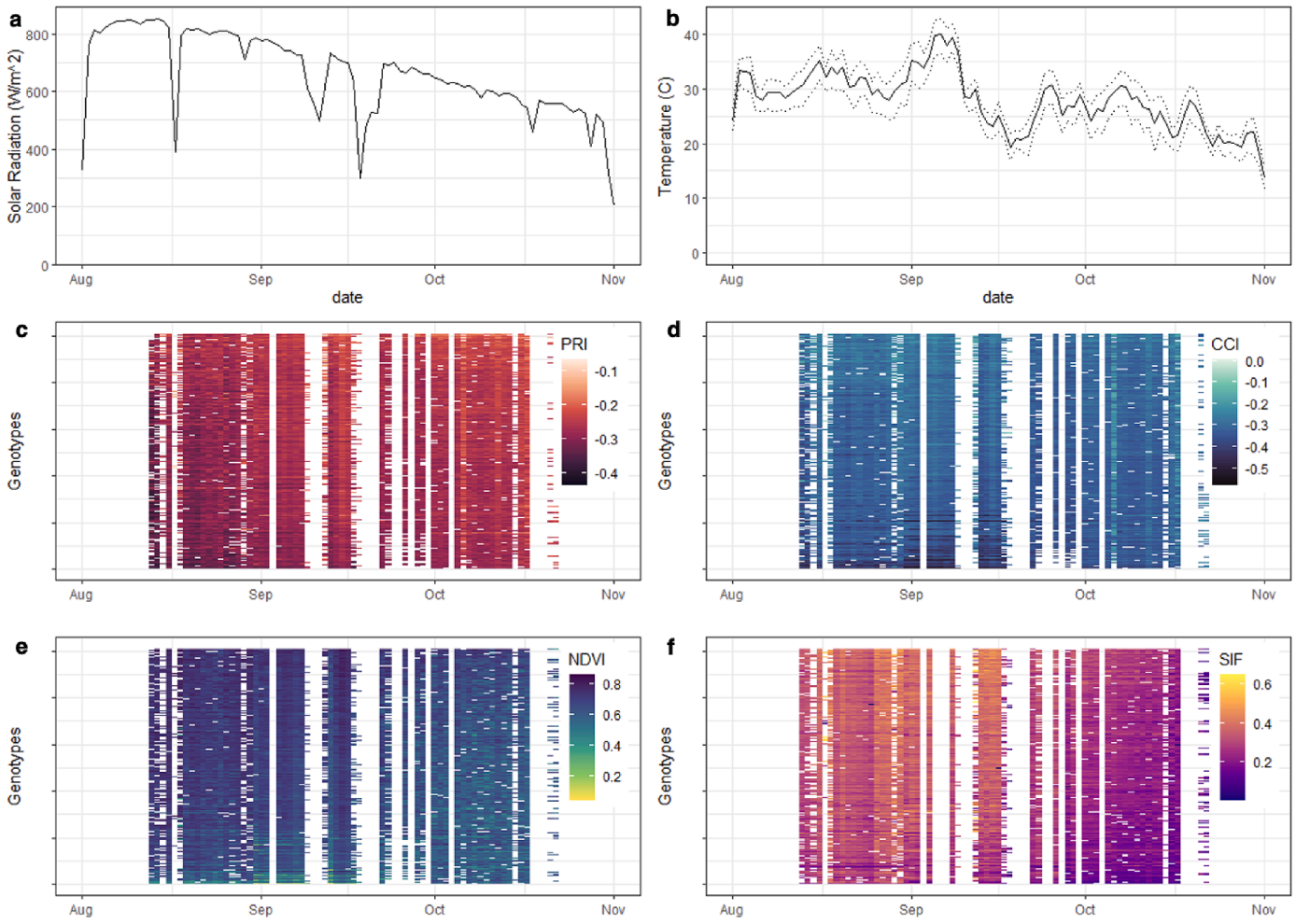

Fig. 2. Time series of midday incoming solar radiation, daily temperature (solid = mean, dotted = minimum and maximum), and genotype-specific CCI (Chlorophyll:Carotenoid Index), PRI (Photochemical Reflectance Index), NDVI (Normalized Difference Vegetation Index) and SIF (Solar Induced Fluorescence). Climate data (Sol Rad and temp) from California Irrigation Management Information System (CIMIS). 2024. Weather data from Station ID 6, Campbell Tract, Davis, CA (lat: 38.535694, long: 121.776360).

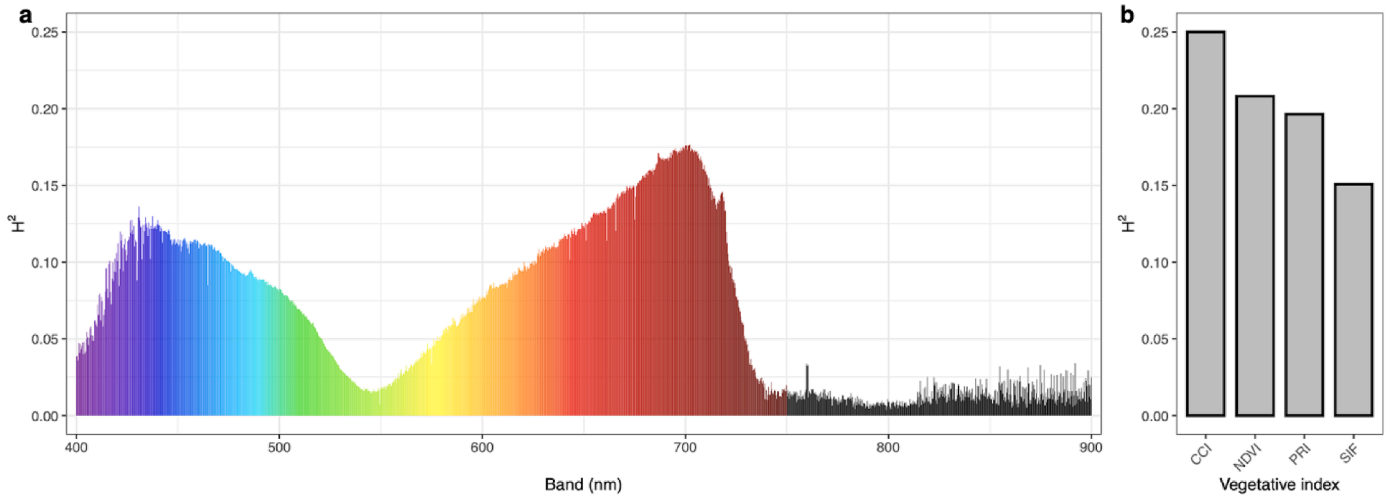

Fig. 3. Heritability (H^2) of hyperspectral bands and vegetative indices.

influencing the activity of these chlorophyll types differentially. In contrast, the SNP shared between blue and red wavelengths could represent regulatory mechanisms that affect chloroplast function more generally, potentially influencing light harvesting, photosystem efficiency, or other related chloroplast activities. Importantly, however, red and blue light is absorbed more efficiently by photosynthetic pigments than green light, resulting in their primary absorption occurring

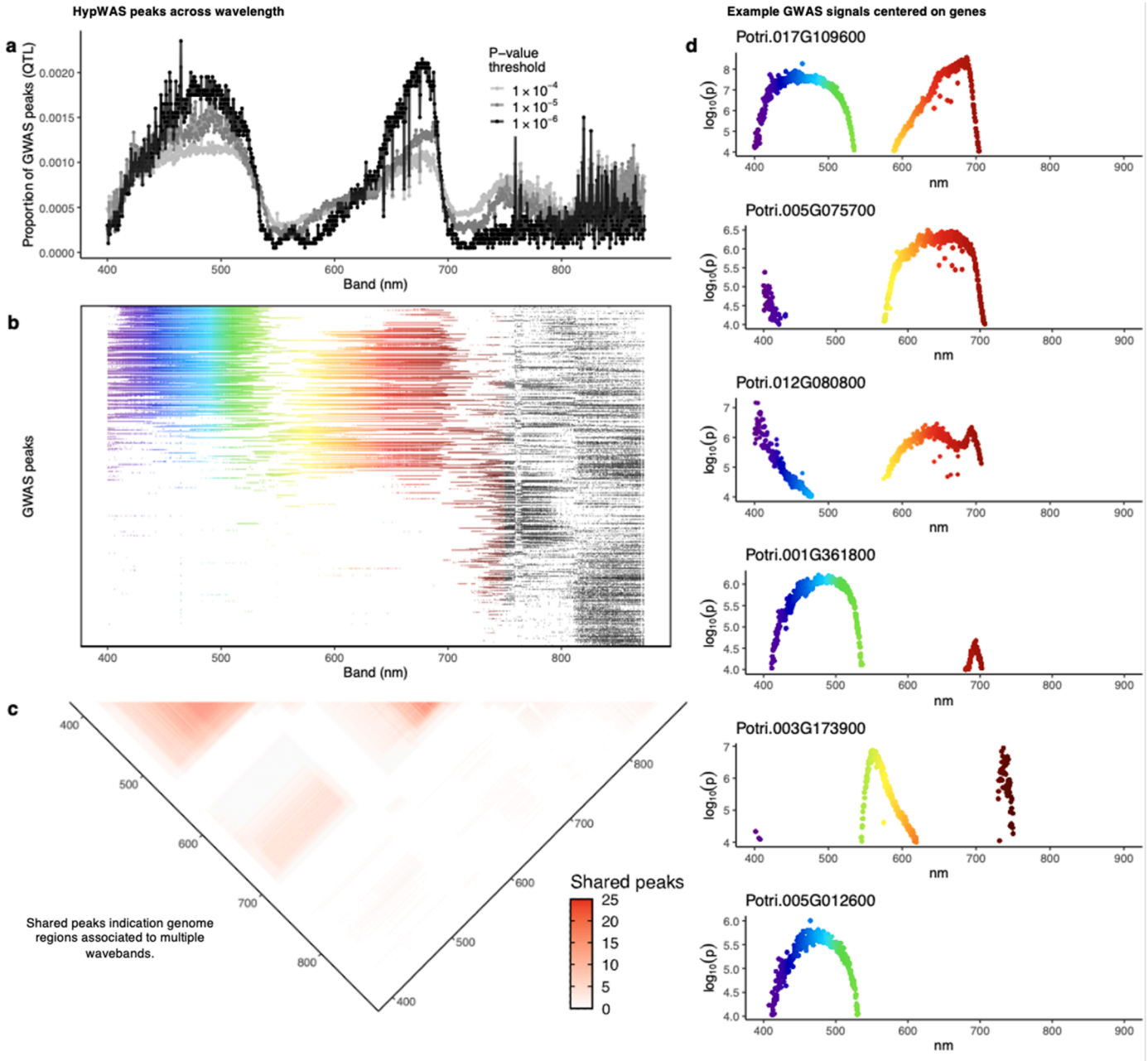

Fig. 4. Genome wide association studies (GWAS) results for hyperspectral traits. a) Relative peak number across bands following different significance thresholds indicated by the gray/black lines. **b)** GWAS peaks clustered in 50 kb detected for >20 bands after filtering, significance threshold is $p \leq 1 \times 10^{-4}$. Each row (y-axis) represents a unique GWAS peak, ordered by the median band with a detected association. The x-axis represents the bands (nm) associated with each respective GWAS peak. **c)** Number of shared peaks between bands, where the color gradient (from white to red) indicates the number of shared peaks. **d)** Example candidate genes with go-term annotation relating to chloroplast, with significant GWAS hits within their transcribed region showing bands in which GWAS, $p \leq 1 \times 10^{-5}$, and $p \leq 1 \times 10^{-6}$.

in the upper layers of leaf cells. In contrast, green light can penetrate more deeply into leaf tissues [49,50].

Wavebands from 450 to 500 nm and 650-700 nm emerged as having elevated heritability (Fig. 3a) and the greatest number of significant variants detected by GWAS (Supplemental Fig. S1, Table S4), perhaps reflecting a complex genetic architecture underlying photosynthetic or metabolic traits, especially for absorption peaks of chlorophyll *a and b and* carotenoids [51]. The strong genetic control over chlorophyll and carotenoid biosynthesis, and chlorophyll biology in general, could explain the greater number of SNPs linked to these wavelengths. Given the importance of chlorophyll content and energy transfer in plant growth and stress response, genetic variation for reflectance in these ranges may be more heritable than other bands (Fig. 3) and detectable in hyperspectral GWAS (Fig. 4).

2.3. *Candidate genes relate to chloroplast function*

To further investigate the genetic basis of associations in the hyperspectral phenotypes, we examined hyperspectral GWAS peaks by determining which variants overlapped with genes involved in chloroplast function. Interestingly, we found a tendency for genes annotated with gene ontology terms associated with chloroplasts to show associations to blue wavebands or blue and red but found few cases of chloroplast-related genes that were associated with red wave bands

alone (Fig. 4d, Supplemental Fig. 2, Supplemental Table 4). The set of associated genes falls into three primary categories: (1) regulators of chloroplast development and plastid transcription (e.g., Potri.001G361800, Potri.005G012600), (2) pigment-related metabolic enzymes (e.g., Potri.017G109600), and (3) nitrogen assimilation and stress response genes (e.g., Potri.005G075700). We identified a variant associated with both blue (450–500 nm) and red (650–700 nm) wavebands that maps to Potri.017G109600, which encodes for a tryptophan synthase. This gene is a sequence homolog to *Arabidopsis* genes with functions in the chloroplast stroma (GO:0009570) and chloroplast envelope (GO:0009941), suggesting a potential role in chloroplast metabolic processes that impact these key wavelengths for photosynthesis. Furthermore, a strong association was observed between red/violet wavebands and a variant in Potri.005G075700, which encodes an asparagine synthetase homolog. Asparagine is essential for nitrogen storage, and this gene is involved in L-asparagine biosynthesis (GO:0070981), as well as chloroplast function (GO:0009507). Nitrogen and chlorophyll often exhibit strong co-variation, as nitrogen is a key component of Rubisco, the primary enzyme facilitating carbon fixation in photosynthesis [52]. Associations spanning UV/blue and red wavebands were also detected for a variant in Potri.012G080800, a gene linked to plastid development (GO:0009536) and peroxisome function (GO:0005777). A role in photorespiration and oxidative stress management could explain the involvement of this gene in these spectral regions, as peroxisomes and chloroplasts work together to maintain the efficiency of light-energy capture and reduce oxidative damage [53,54]. A variant in Potri.001G361800 was associated with blue and green wavebands. This gene is a homolog of *Arabidopsis* transcription termination factors (mTERF family) involved in regulating plastid gene expression and chloroplast development, specifically within the thylakoid membranes (GO:0042651) [55]. These functions are crucial for optimizing light absorption in these spectral regions, highlighting the gene's potential role in maintaining photosynthetic efficiency. In the green and orange wavebands, a significant variant was found in Potri.003G173900, which encodes a sequence homolog of *Arabidopsis* HEAT SHOCK PROTEIN 60, which localizes to chloroplasts [56]. This is especially interesting, as the 531 nm wavelength is a key absorption feature associated with the de-epoxidation of xanthophyll cycle pigments, which facilitate non-photochemical quenching (NPQ) to dissipate excess energy as heat. Under stress conditions, dynamic adjustments in xanthophyll pigments enhance light absorption at this peak, leading to a characteristic decrease in reflectance at 531 nm [57]. Finally, a notable association with blue and green wavebands was identified for Potri.005G012600, a gene encoding a sequence homolog of LPA66, a chloroplast-localized protein of the pentatricopeptide repeat (PPR) family, knockouts of which have impaired photosynthesis [58]. Taken together, these candidates (Fig. 4d, Supplemental Fig. 2, Supplementary Table 5) provide novel insights into genes involved in chloroplast biology, motivating further experimental validation and inspiring a framework for using hyperspectral associations to infer gene function.

2.4. Genetic architecture of vegetative indices over time

Temporal variation in genome-wide association studies (GWAS) helps pinpoint how genetic control shifts across developmental stages and with respect to the environmental conditions which also change over time [59–62]. Here, we examined the genetic architecture of standard vegetation indices—NDVI, PRI, SIF, and CCI— to test for temporal variation in GWAS. To test this, we conducted hyperspectral GWAS for each index at daily intervals throughout the experiment, producing a total of 272 GWAS analysis (4 indices x 68 days). We compared the results by analyzing the beta coefficients and the significance of associations (p-values) for each measurement day. To assess variability over time, we calculated Pearson correlation coefficients between days for both beta coefficients and negative log10-transformed p-values. Under the null hypothesis of a stable genetic architecture, we would expect no correlation between the temporal distance of measurements and the similarity of GWAS results (Supplemental Fig. 3). However, if genetic architecture changes over time, we would expect more similarity in GWAS results for measurements taken closer together in time. Our analysis showed that GWAS results were indeed more similar for measurements taken at closer time points ($P < 1x10^{-16}$), supporting the hypothesis that the genetic architecture of these traits changes over the growing season and/or developmental stages (Fig. 5). These dynamics likely reflect shifting genetic regulation as plants transition from peak photosynthetic activity toward senescence under drought stress, indicating that the genetic control of these traits is not static but dynamic, varying across time (days and weeks) and wavelengths, highlighting the importance of considering temporal and spectral factors when designing experiments and interpreting GWAS in plant systems. Indeed, a few studies have shown that the genetic architecture of complex traits in plants changes over time, with many loci exhibiting stage-specific, transient, or even reversing effects as development progresses or environmental conditions shift [62–64]. The temporal variation in the genetic architecture of vegetative indices emphasizes the critical role of timing in phenotypic measurements, as it can significantly affect genome-trait associations. Associations found at one time-point may not be as relevant at another, especially for traits shaped by environmental conditions or developmental stages. This is particularly true for deciduous trees like poplar, where seasonal and developmental shifts strongly influence the biological processes underlying vegetative indices. In our experiment, the poplar trees transitioned from an active growing season with new leaf growth to leaf senescence (August to October). Given this developmental shift, it is logical that the genetic architecture of traits would change over time, reflecting the dynamic physiological and environmental conditions influencing gene expression and trait regulation.

Taken together, Fig. 5 highlights the dual sensitivities of PRI and CCI to diurnal and seasonal environmental changes, which can be attributed to distinct physiological processes. Diurnally, both indices reflect dynamic changes in the xanthophyll cycle, driven by environmental factors such as light intensity and associated with non-photochemical quenching [65]. This reflects immediate, transient responses to fluctuating environmental conditions. In contrast, on a seasonal timescale, PRI and CCI reflect slower changes in carotenoid and chlorophyll pools that occur over weeks and are likely tied to plant developmental processes and long-term environmental adaptation. Our interpretation suggests that carotenoid ratios, as indicated by CCI, are more genetically controlled than diurnal variations, given the slower nature of these seasonal adjustments. These gradual changes in pigment pools over the growing season reflect underlying genetic factors that influence the absolute size of the carotenoid and chlorophyll pools and their rates of change in response to long-term environmental cues. This slower, genetically driven adjustment may play a key role in plant adaptation to seasonal stressors and could have implications for understanding genetic variation in photosynthetic efficiency and stress resilience. Therefore, the seasonal dynamics of carotenoid ratios may serve as a valuable proxy for genetic strategies linked to long-term fitness and environmental adaptation in different genotypes.

2.5. Correlations to yield traits

In addition to investigating the genetic architecture of hyperspectral traits, we examined the relationship between specific wavelengths and yield—the relative growth rate (RGR) of height—to assess the adaptive significance of vegetation indices for predicting yield (Fig. 6) [46]. Vegetation indices such as NDVI are effective predictors of crop productivity, as they quantify canopy structure, greenness, and photosynthetic activity across developmental stages [66–68]. RGR exhibited a positive correlation at 500 nm, reaching a maximum of −0.3, and a negative correlation that peaked around 700 nm. CCI had the strongest

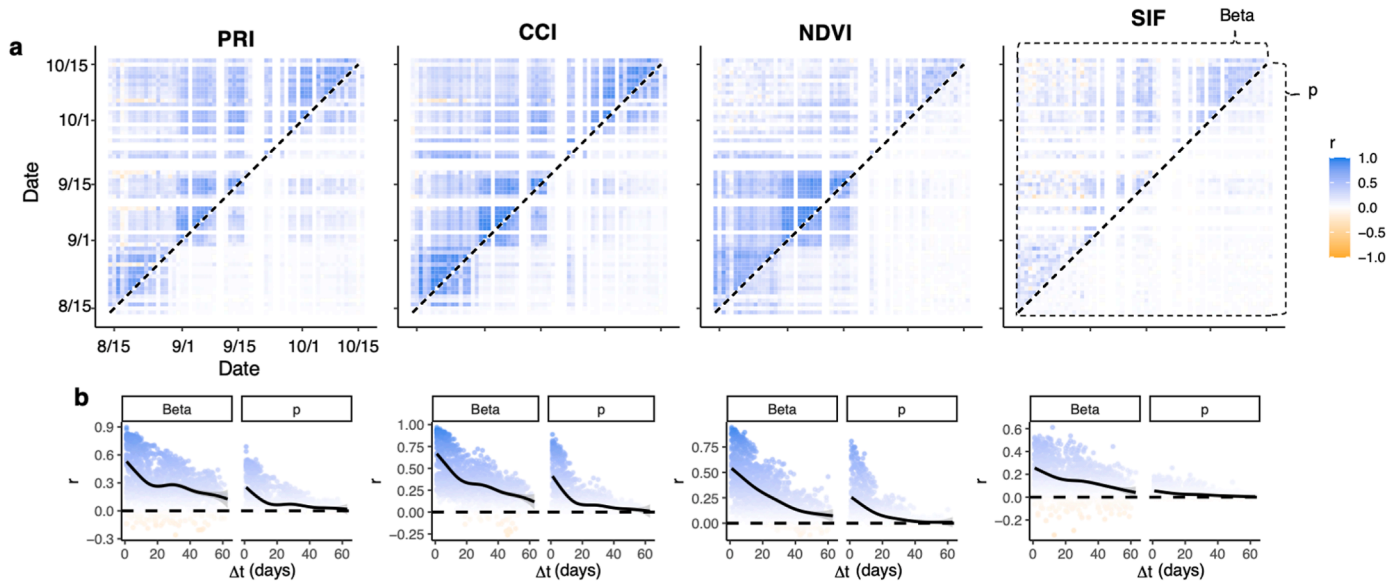

Fig. 5. Temporal Genome wise association study (GWAS) analyses for composite hyperspectral traits in *Populus trichocarpa*. **a)** Heatmaps showing Pearson correlations between timepoints for allele effects (beta coefficients, above diagonal; and p-values, below diagonal). **b)** Autocorrelations between beta and p-value correlations as a function of the difference in time among measurements.

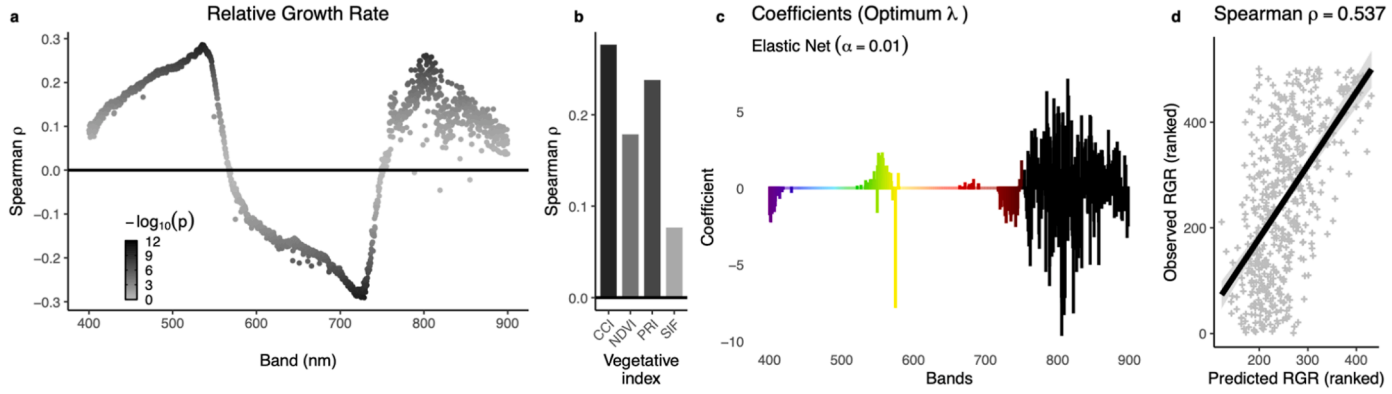

Fig. 6. Spectral and Vegetative Index Correlations with Phenotypic Traits. **a)** Spearman correlation of hyperspectral bands (400–900 nm) of Relative Growth Rate (RGR). The x-axis represents the wavelength (nm), and the y-axis represents the Spearman's rho, quantifying the strength and direction of the correlation between each spectral band and the respective trait. **b)** Correlation with vegetative indices CCI (Chlorophyll:Carotenoid Index), PRI (Photochemical Reflectance Index), NDVI (Normalized Difference Vegetation Index) and SIF (Solar Induced Fluorescence) **c)** Coefficients (slopes) from elastic net multivariate model predicting relative growth rate. **d)** Predicted versus observed rank values for relative growth rate from elastic net model.

spearman correlation with RGR ($-\log_{10}(p) > 10$), followed by NDVI and PRI, while SIF showed the weakest correlation. Overall, higher green and infrared reflectance was strongly associated with greater yield across all traits, suggesting that these wavelengths effectively indicate larger, more productive trees. NDVI has been linked to yield in other crop systems and utilized in breeding and selection for improvements in productivity and stress resilience. Lower red reflectance (higher red absorption) was linked to greater biomass, highlighting the connection between greater photosynthetic potential, likely due to more chlorophyll absorption. Red light is generally associated with increased plant biomass accumulation due to its strong stimulation of photosynthesis and influence on elongation via phytochrome signaling, whereas blue light tends to reduce overall growth by inhibiting elongation, though it plays key roles in photomorphogenesis and stomatal regulation [69,70]. Indeed, the blue reflectance in our study showed a more moderate association with biomass. These findings suggest that specific wavelengths, particularly in the green, infrared, and red regions, could serve as valuable markers for predicting yield across diverse genotypes.

To evaluate the predictive power of hyperspectral traits with a multivariate model, we used an elastic net regression, which regularizes coefficients in order to overcome multicollinearity and overfitting, especially for datasets with large number of predictors like for hyperspectral traits (see Methods for details) [71–74]. Elastic net models can reduce predictor coefficients to 0, thus serving as a form of model selection to identify informative and non-redundant predictors. Interestingly, we found that the elastic net model retained wavelengths in visible and infrared wavebands. The retention and coefficients of predictors can reflect their degree of multicollinearity and value in predicting the response (relative growth rate). Notably, this model outperformed any individual waveband or vegetative index in predictive accuracy of genotype ranks for growth rate (spearman rho >0.5, $p < 2x10^{-16}$). This result highlights the practical value of high dimensional spectral data to predict plant performance traits and inspires further applications for advanced selections. Like all traits, hyperspectral phenotypes reflect both genotype and environment effects, so loci detected here may be more meaningful under drought, and their

effects could shift under well-watered or alternative field conditions. Further work is needed to validate the prediction accuracy across other environments and genotypes to determine the reliability for breeding applications. A natural next step is to broaden this framework beyond a single season and environment. Applying GWAS across multi-year, multi-environment field trials with expanded genotype panels will be essential for testing the scalability and robustness of the genetic signals and trait–yield relationships identified here. Extending tower (and calibrated airborne/UAV) measurements across the full growing season would further capture the complete drought and phenological trajectories. From a breeding perspective, a key practical goal is to quantify how well early-season hyperspectral/SIF phenotypes predict higher-order outcomes, especially end-of-season biomass, enabling earlier and lower-cost culling of poor performers. Finally, feature-selection approaches can be used to identify a minimal set of informative wavelengths or indices, informing the design of lower-cost multispectral systems for routine screening in large breeding populations. Beyond breeding utility, hyperspectral GWAS provides a route to functional inference by linking genetic variation to specific spectral features. A clear next step is experimental validation of these interpretations using targeted genetic perturbations. Profiling hyperspectral/SIF traits in mutant or gene-edited isogenic lines for candidate loci would allow direct tests of how altering gene function shifts wavelength-specific signatures, strengthening causal links between mapped loci, spectral traits, and underlying biology.

3. Conclusions

We combined hyperspectral remote sensing and GWAS analyses to explore the genetic architecture of hyperspectral phenotypes underpinning productivity in a field common garden of diverse *Populus trichocarpa* over time. Our findings point to significant heritability and genetic control over key physiological traits, particularly those linked to chlorophyll and carotenoid dynamics, likely underpinning photosynthetic performance as demonstrated by elevated heritability in the blue and red spectral regions. Notably, we identified several candidate genes related to chloroplast function, providing novel insights into their roles in photosynthetic function linked to yield.

Temporal analyses highlighted the dynamic nature of genetic architecture, with variations in hyperspectral GWAS results correlating with developmental stages and environmental shifts. This suggests that both the genetic and physiological responses of *Populus* are influenced by seasonality, underscoring the importance of integrating time-series data in future genetic studies and providing an opportunity to explore adaptive seasonal traits that in future, are likely to be sensitive to climate change. Additionally, strong correlations between vegetation indices and yield traits emphasize the potential of hyperspectral phenotyping to predict yield productivity, contributing valuable knowledge for breeding programs focused on bioenergy crops.

Overall, this work demonstrates the utility of hyperspectral GWAS in identifying physiologically meaningful genetic variation, offering promising directions for enhancing crop traits through targeted breeding. Our results pave the way for more nuanced applications of remote sensing in plant phenotyping and underscore the importance of considering temporal dynamics when investigating the genetic basis of complex traits.

4. Methods

4.1. Site and tree details

This study was conducted on a subset of a 15-acre common garden at UC Davis (38.5°N, 121.8°W), which houses approximately 8000 trees of the bioenergy species *Populus trichocarpa* (Black Cottonwood). *P. trichocarpa* is native to the Pacific Northwest of North America, extending from Canada to Southern California (38.5-54.3°N, 116-128.7°W). The species experiences a wide range of climates across its natural range, resulting in high genetic diversity in its biological and structural development. Detailed descriptions of the field management and the climate of origin for these genotypes have been previously described [75]. Starting in April 2021, irrigation was reduced in designated drought plots to create a controlled soil moisture deficit, while control plots received full irrigation. During both 2021 and 2022, the drought plots were maintained at a soil water potential between −0.1 and −0.15 MPa throughout the dry summer months (May to October), creating a long-term contrast to the fully irrigated control plots, where soil water potential, monitored via gypsum blocks, remained between 0 and -0.03 MPa [75]. This study focused on trees within the drought site.

4.2. Tower-based hyperspectral and solar-induced fluorescence

To study the genetic architecture of photosynthetic traits over time, we employed a tower-based remote sensing system, TSWIFT, to continuously monitor individual poplar trees [37]. This system was set up on a 15 m tall tower. The tower was located North of a drought treatment block from 2022 to 08-18 to 2022-10-18. The TSWIFT system consists of an RGB camera with a co-located telescope with a 4° field of view, connected to a fiber optic cable and a FLAME and QE PRO spectrometer (Ocean Insights, FL, USA). The FLAME spectrometer collected hyperspectral data from 400 to 900 nm (1.2 nm full width at half maximum), while the QE PRO spectrometer collected ultraspectral data from 730 to 780 nm (0.3 full width at half maximum) for SIF retrievals. This system enabled spot targeting of individual trees with a field of view diameter ranging from 0.3 to 1 m, depending on the distance of the tree to the tower. Each tree was sampled using fixed pointing coordinates set at the start to target a sunlit area in the upper ~1/3 of the canopy. Given the co-located RGB camera, we were able to visually assess that the spectral sample was taken on the exact same location of every tree throughout the experiment. We chose the position in the field based on the sun angle and accessibility; therefore, all trees in that radius from the tower were selected randomly. Each tree target measurement took a maximum of 5 s, and reference sky scans occurred for every 15 target measurements (maximum 75 s). The TSWIFT system monitored 595 trees, representing 505 genotypes (some with replication) within the same block. All tower-based hyperspectral measurements were collected over the drought treatment plot to align phenotyping with the primary biological target of the experiment: canopy performance under water limitation. Because measurements were restricted to this environment, the resulting genetic associations and prediction models are interpreted as drought-context–specific.

The spectrometers were radiometrically calibrated from unitless digital numbers to radiance (mW cm^{-2} sr^{-1} nm $^{-1}$) according to Ref. [37]. This was conducted in the field by taking “sky” measurements with the TSWIFT camera pointing straight up at the sky (90°) while taking concurrent measurements from a radiometrically calibrated field spectrometer (HR-1024i, Spectra Vista Corporation, Poughkeepsie, New York, USA) pointed at a calibrated Spectralon diffuse reflectance standard (Labsphere Inc., NH, USA). The concurrent measurements took repeat measurements every minute for 1 h from 12:00 to 13:00 under clear sunny conditions. Radiometric calibration coefficients were generated for each spectrometer by relating the raw digital numbers with the calibrated spectrometer radiance via a fitted linear line with an intercept of zero. The calibration coefficients were applied to all TSWIFT measurements.

To calculate reflectance, the target radiance was divided by the closest in-time sky irradiance. The FLAME reflectance data was used to determine vegetation indices, including the normalized difference vegetation index (NDVI), photochemical reflectance index (PRI), and chlorophyll-carotenoid index (CCI). The QEP radiance data was used to determine solar-induced fluorescence (SIF).

$$\mathrm{NDVI} = \frac{(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red})}$$

$$\mathrm{PRI} = \frac{(R_{531} - R_{570})}{(R_{531} + R_{570})}$$

$$\mathrm{CCI} = \frac{(R_{531} - R_{645})}{(R_{531} + R_{645})}$$

The above equations represent reflectance (R) at respective wavebands/regions denoted in the subscript. Red is centered at 680 nm, and near-infrared (NIR) is centered at 800 nm with a ±15 and 25 nm buffer, respectively, while the PRI and CCI bands had a ±5 nm buffer. SIF was retrieved using Differential Optical Absorption Spectroscopy (DOAS) (Supplemental Tables 1 and 2) [34,76].

4.3. Data filtering and cleaning

To ensure targets were correctly positioned on trees, we used NIR reflectance data as an initial data quality filter. Here, we constrained reflectance data to >10 %, with the assumption that NIR less than 10 % reflectance was poor data quality due to partial tree targeting or weak signal-to-noise ratio. We also conducted a cloud/sky condition filtering using the reference "sky" scans. Here, we summed irradiance from 400 to 700 nm to represent a unitless value representing photosynthetic active radiation (PAR). With this parameter, we created a 7-day running mean and filtered out days that were <50% of the running mean, which we assumed were poor sky conditions such as clouds. To summarize the spectral data on a daily basis, we constrained data to solar noon from 11 to 15 h. From these data, we averaged samples to each individual target per day, as some targets have repeat samples within the solar noon period. For downstream analysis (e.g., heritability and GWAS), raw spectral measurements for each wavelength were scaled and centered. Specifically, for each target on each day, the vector of band measurements was standardized (standard normal variate transformation, SNV) to capture relative rather than absolute signal strength across bands. This normalization step adjusts for potential biases in overall reflectance intensity, which could arise from variations in field position or other sources of experimental noise.

4.4. Variance components and heritability estimation

To estimate broad-sense heritability (H^2) and variance components for hyperspectral traits, we implemented a mixed-effects modeling approach using the lme4 package in R (version 4.4.1). The analysis was carried out for each wavelength band measured by the hyperspectral imager. For each wavelength band, we fit a linear mixed-effects model to partition the variance. The model included genotype and date as random effects, allowing us to estimate the contribution of these factors to the total phenotypic variance. The full model was represented as: $band = \mu + (1|Genotype) + (1|Date) + \varepsilon$, where μ is the overall mean, $(1|Genotype)$ represents the random intercept for each genotype, capturing genetic variance ($\sigma^2_{Genotype}$), $(1|date)$ represents the random intercept for each date, capturing environmental variance due to the date effect (σ^2_{Date}) and ϵ is the residual error, capturing unexplained variance ($\sigma^2_{Residual}$). Variance components for genotype, date, and residual variance were extracted from the full mixed-effects model. These variance components were used to calculate the total phenotypic variance: $\sigma^2_{Phenotype} = \sigma^2_{Genotype} + \sigma^2_{Date} + \sigma^2_{Residual}$. We estimated the broad-sense heritability (H^2) for each wavelength band as the proportion of the total phenotypic variance attributable to genetic variance $H^2 = \frac{\sigma^2_{Genotype}}{\sigma^2_{Phenotype}}$. Formal G × E (genotype × date) variance partitioning requires denser within-date replication, a direction for future work.

4.5. GWAS genome-wide association studies with hyperspectral phenotypes

The marker phenotype association test was performed using GEMMA (Genome-wide Efficient Mixed-Model Association, version 0.98.5) to identify associations between the trait of interest and the genotypic data of the entire *P. trichocarpa* population (1523 genotypes) available through their database [4,77] and as previously described in Ref. [75,78]. We used the *Populus trichocarpa* genome version 4.1 (Phytozome genome ID: 533). GEMMA's linear mixed model incorporates a kinship matrix to account for population structure and relatedness. In this study, we performed GWASseparately for all 1423 bands (400-900 nm), as well as CCI, NDVI, PRI, and SIF. Variants (SNPs and Indels) were filtered to those with a minor allele frequency of at least 0.05. To account for linkage disequilibrium (LD) decay and ensure the selection of independent SNPs for further analysis, we applied the clumping procedure in PLINK v1.9. SNPs located within 50 kilobases (kb) of one another were grouped into clumps based on their LD with the index variant. Variants within each clump that had an r-squared (r^2) value greater than or equal to 0.05 (indicating high LD) were excluded, leaving only the most significant variant (index variant) in each region. This was done using the following PLINK command: plink –clump –clump-p1 1 –clump-p2 1 –clump-r2 0.05 –clump-kb 50. The LD thresholds were determined based on empirically measured LD across the *Populus trichocarpa* genome, determined with the command plink –r2 –ld-window-kb 100 –ld-window 500 –ld-window-r2 0. Visual inspection of LD decay showed that LD dropped to equilibrium ($r^2 = 0.05$) at distances greater than 20 kb. This approach minimized redundancy caused by variants in LD while retaining the most informative genetic markers associated with the trait.

For downstream analyses, we initially selected index variants with a significance threshold of $p \leq 1 \times 10^{-4}$. We then compared these results with more stringent filtering thresholds, specifically $p \leq 1 \times 10^{-5}$ and $p \leq 1 \times 10^{-6}$, and found that the relative enrichment of GWAS peaks across different wavelengths remained consistent across significance thresholds. For the final set of high-confidence associations, we applied a more stringent filter, retaining only variants with $p \leq 1 \times 10^{-6}$. Additionally, for each clump (i.e., group of variants in LD with the most significant variant), we calculated the proportion of variants with $p < 0.05$ and compared this to the null expectation under a uniform distribution of p-values using a Chi-squared test. To account for multiple testing, we applied Bonferroni correction to adjust the significance thresholds in the Chi-squared test. Ultimately, our high-confidence associations were those variant clusters where the most significant variant had $p \leq 1 \times 10^{-6}$, and the additional variants in LD were significantly skewed toward low p-values. Because a chi-square test was applied to each GWAS clump, we corrected for multiple testing using a Bonferroni adjustment on the clump-level analyses. However, we did not apply an additional correction across all 1423 phenotypes, as many bands are highly correlated and correcting across all traits would likely be overly conservative.

4.6. Identifying candidate genes

To identify potential candidate genes underlying the band-specific genetic association results, we focused on significant variants ($p \leq 1 \times 10^{-6}$) located within the annotated transcribed regions of *Populus trichocarpa* genes (genome annotation v4_1, JGI). To explore the functional roles of these genes, we conducted gene ontology analyses by identifying putative sequence homolog in *Arabidopsis thaliana* (TAIR10) using BLASTP with the following thresholds: E-value less than 1e-5, bitscore greater than 100, and percent identity (PIDENT) greater than 30%. The gene ontologies for the *Arabidopsis* homologs were sourced from the TAIR10 annotations (arabidopsis.org).

To further explore candidate genes, we ranked genic variants based on the median p-value across all bands where a significant association

($p < 0.0001$) was detected. To visualize the potential phenotypic effects of these variants in relation to different wavebands, we plotted the negative log10-transformed p-values (-log10(P)) for each variant across all bands.

4.7. Temporal patterns of GWAS results

For the temporal analysis, we used GWAS-estimated beta coefficients representing the estimated genetic effect size of InDels (insertion/deletion polymorphisms) and single nucleotide polymorphisms (SNPs) on SIF, NDVI, CCI, and PRI daily averages. To evaluate the temporal patterns of allele effects over time, Pearson correlations were estimated across different time points for both β and their associated p-values.

To assess the similarities in hyperspectral GWAS results between time points, we estimated correlations between the beta coefficients (β_i and β_j) and p-values (p_i and p_j) across pairs of time points. Specifically, we computed the pairwise Pearson correlation for all distinct time point (T) pairs (i,j), or: $cor\left(\beta_i, \beta_j\right), \forall\, i,j \in T \text{ with } i \neq j$ and $cor\left(p_i, p_j\right), \forall\, i,j \in T \text{ with } i \neq j$, where β_i and β_j represent the vector of β, with $-\log_{10}$(p-values) p_i and p_j at time points i and j.

Heatmaps were generated to visualize these correlations, with beta coefficient correlations displayed above the diagonal and p-value correlations below the diagonal. Autocorrelation analyses were subsequently performed to assess the relationship between the similarity of GWAS results (correlations described above) and the temporal intervals between measurements. Specifically, we calculated the autocorrelation by comparing the correlation values (either for beta coefficients or p-values) with the absolute time difference $|j = i|$, or $cor(cor(\beta_i, \beta), |j = i|)$ and $cor\left(cor\left(p_i, p_j\right), |j = i|\right)$. Under the null (no autocorrelation), we expect no correlation between difference and time and similarities in the beta coefficients or p-values. Alternatively, if GWAS results are more similar between close measurement time points, we expect a significant negative relationship between similarities in beta coefficients and p and difference in time. To validate this, we also measured similarities between beta coefficient and p-value matrices with distance matrices of time with a Mantel test in R (from *vegan* package, version 2.7-2), using 9999 permutations.

4.8. Correlations with yield traits

We used previously measured performance-related (yield) traits "Height in cm" (measured November 2021), relative growth rate, and stem volume index (measured in November 2021) [75]. The relative growth rate of height was calculated with this equation: $RGR = (ln(W2) - ln(W1))/(ln(W1))$ where *W1* represents the height taken in March of 2021, and *W2* represents the height measured in November of 2021. For each, we calculated the Spearman correlation between mean genotype values for each band, as well as vegetative indices in R (version 4.4.1) with the cor(method = " spearman") function. For a multivariate model, we used regularization to predict genotype ranks of growth rate with Elastic net regression in R using the cv.glmnet function from the glmnet package (version 4.1-8). This method combines the L1 and L2 penalties of lasso and ridge regression and is especially useful for handling multicollinearity and avoiding overfitting in high-dimensional datasets like the hyperspectral data we collected. Elastic net regularizes beta coefficients by solving to minimize both the residual sum of squares while penalizing the magnitude of coefficients

$$minimize\left(\frac{1}{2n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 + \lambda\left((1-\alpha)\frac{1}{2}\sum_{j=1}^{p}\beta_j^2 + \alpha\sum_{j=1}^{p}|\beta_j|\right)\right)$$

Elastic net models were fitted using the glmnet package. For each trait, the tuning parameter λ (lambda), which controls overall shrinkage, was selected automatically by 10-fold cross-validation using cv.glmnet; we used the cross-validated optimum (λ_min) for the final model. The mixing parameter α, which determines the relative contribution of L1 (lasso) and L2 (ridge) penalties, was set to 0.01 to provide mild L1 regularization and yield a slightly sparser, more interpretable model without compromising predictive accuracy. We also evaluated ridge regression ($\alpha = 0$) and lasso regression ($\alpha = 1$); ridge produced a fully dense model (retaining essentially all predictors), whereas lasso was overly stringent and removed most predictors. Both alternatives showed slightly lower accuracy than $\alpha = 0.01$, so we report only elastic-net results.

Author contributions

Conceptualization:
Methodology: CYSW, TM, GT.
Software: CYSW, TM.
Validation:
Formal Analysis: CYSW, MK, MS, GM.Investigation: J-GC, GT, GAT, TT, CYSW, TM. Resources: J-GC, GT, GAT, TM. Data Curation: CYSW, MK, JBB.
Writing - Original Draft: CYSW, MK.
Writing - Review & Editing: All.
Visualization: CYSW, MK, GM.
Supervision: GT, TM, TB, Project Administration: GT.
Funding Acquisition: GT, GAT, TT, CYSW.

Funding

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, world-wide 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).

MCK acknowledges the Department of Plant Sciences, UC Davis, for awarding a GSR scholarship funded by endowments, particularly the James Monroe McDonald Endowment, administered by UCANR. Research in the laboratory of GT was funded by the John B. Orr Chair Endowment, by the Genomics enabled Plant Biology for Determination of Gene Function programme by the Office of Biological and Environmental Research in the DOE Office of Science (DE-SC0020164) and as Quain Professor of Botany, University College, London.

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.

Acknowledgments

We thank all UC Davis interns, graduate students, and postdocs who helped with field collections and lab assistance on this project. MCK gratefully acknowledges the UCD Jastro-Shields Research Award for their support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphe.2026.100174.

Data availability

Hyperspectral phenotype data are publicly available through the Oak Ridge National Laboratory LabKey data portal (DOI: 10.25983/CBI/3012775). Genomic analyses used the *Populus trichocarpa* reference genome version 4.1 (Phytozome genome ID: 533).

References

[1] H.D. Bradshaw, R. Ceulemans, J. Davis, R. Stettler, Emerging model systems in plant biology: poplar (Populus) as A model forest tree, J. Plant Growth Regul. 19 (2000) 306–313.

[2] G. Taylor, M.R. Allwright, H.K. Smith, A. Polle, H. Wildhagen, M. Hertzberg, R. Bhalerao, J.J.B. Keurentjes, S. Scalabrin, D. Scaglione, et al., Bioenergy trees: genetic and genomic strategies to improve yield, in: Perennial Biomass Crops for a Resource-Constrained World, Springer International Publishing, Cham, 2016, pp. 167–190.

[3] M.R. Allwright, A. Payne, G. Emiliani, S. Milner, M. Viger, F. Rouse, J.J. B. Keurentjes, A. Bérard, H. Wildhagen, P. Faivre-Rampant, et al., Biomass traits and candidate genes for bioenergy revealed through association genetics in coppiced European Populus nigra (L.), Biotechnol. Biofuels 9 (2016) 195.

[4] G.A. Tuskan, S. DiFazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, et al., The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray), Science 313 (2006) 1596–1604.

[5] T.J. Tschaplinski, G.A. Tuskan, M.M. Sewell, G.M. Gebre, D.E. Todd, C.D. Pendley, Phenotypic variation and quantitative trait locus identification for osmotic potential in an interspecific hybrid inbred F2 poplar pedigree grown in contrasting environments, Tree Physiol. 26 (2006) 595–604.

[6] N.R. Street, O. Skogström, A. Sjödin, J. Tucker, M. Rodríguez-Acosta, P. Nilsson, S. Jansson, G. Taylor, The genetics and genomics of the drought response in populus, Plant J.: Cell Mol. Biol. 48 (2006) 321–341.

[7] A.M. Rae, M.P.C. Pinel, C. Bastien, M. Sabatti, N.R. Street, J. Tucker, C. Dixon, N. Marron, S.Y. Dillen, G. Taylor, QTL for yield in bioenergy populus: identifying G×E interactions from growth at three contrasting sites, Tree Genet. Genomes 4 (2007) 97–112.

[8] S. Berlin, U. Lagercrantz, Arnold S. von, T. Ost, A.C. Rönnberg-Wästljung, High-density linkage mapping and evolution of paralogs and orthologs in salix and populus, BMC Genom. 11 (2010) 129.

[9] P. Ranjan, T. Yin, X. Zhang, U.C. Kalluri, X. Yang, S. Jawdy, G.A. Tuskan, Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in populus, Bioenergy Res. 3 (2010) 172–182.

[10] W. Jinnan, W. Ruyue, L. Jihong, T. Yanting, G. Haoping, H. Lili, W. Dongyue, W. Xueling, Construction of a high-density genetic map using specific-length amplified fragment markers and identification of QTLs for branching angle in poplar, Mol. Genet. Genom.: MGG 297 (2022) 345–356.

[11] J.N. Cobb, G. Declerck, A. Greenberg, R. Clark, S. McCouch, Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement, Theor. Appl. Genet. 126 (2013) 867–887.

[12] F. Tardieu, L. Cabrera-Bosquet, T. Pridmore, M. Bennett, Plant phenomics, from sensors to knowledge, Curr. Biol.: CB 27 (2017) R770–R783.

[13] J. Wang, X. Li, T. Guo, M.J. Dzievit, X. Yu, P. Liu, K.P. Price, J. Yu, Genetic dissection of seasonal vegetation index dynamics in maize through aerial based high-throughput phenotyping, Plant Genome 14 (2021) e20155.

[14] L.M. York, Functional phenomics: an emerging field integrating high-throughput phenotyping, physiology, and bioinformatics, J. Exp. Bot. 70 (2019) 379–386.

[15] Y. Zhang, M. Migliavacca, J. Penuelas, W. Ju, Advances in hyperspectral remote sensing of vegetation traits and functions, Rem. Sens. Environ. 252 (2021) 112121.

[16] X. Zhi, S.R. Massey-Reed, A. Wu, A. Potgieter, A. Borrell, C. Hunt, D. Jordan, Y. Zhao, S. Chapman, G. Hammer, et al., Estimating photosynthetic attributes from high-throughput canopy hyperspectral sensing in sorghum. *Plant Phenomics* **2022**, 2022.

[17] J. Xiao, F. Chevallier, C. Gomez, L. Guanter, J.A. Hicke, A.R. Huete, K. Ichii, W. Ni, Y. Pang, A.F. Rahman, et al., Remote sensing of the terrestrial carbon cycle: a review of advances over 50 years, Rem. Sens. Environ. 233 (2019) 111383.

[18] R.T. Furbank, M. Tester, Phenomics–technologies to relieve the phenotyping bottleneck, Trends Plant Sci. 16 (2011) 635–644.

[19] F. Fiorani, U. Schurr, Future scenarios for plant phenotyping, Annu. Rev. Plant Biol. 64 (2013) 267–291.

[20] G.A. Blackburn, Hyperspectral remote sensing of plant pigments, J. Exp. Bot. 58 (2006) 855–867.

[21] S.L. Ustin, A.A. Gitelson, S. Jacquemoud, M. Schaepman, G.P. Asner, J.A. Gamon, P. Zarco-Tejada, Retrieval of foliar information about plant pigment systems from high resolution spectroscopy, Rem. Sens. Environ. 113 (2009) S67–S77.

[22] P.J. Sellers, Canopy reflectance, photosynthesis, and transpiration, II. The role of biophysics in the linearity of their interdependence, Rem. Sens. Environ. 21 (1987) 143–183.

[23] A. Bannari, D. Morin, F. Bonn, A.R. Huete, A review of vegetation indices, Remote Sens. Rev. 13 (1995) 95–120.

[24] Q. Sun, Q. Jiao, X. Qian, L. Liu, X. Liu, H. Dai, Improving the retrieval of crop canopy chlorophyll content using vegetation index combinations, Remote Sens. 13 (2021) 470.

[25] J. Peñuelas, I. Filella, J.A. Gamon, Assessment of photosynthetic radiation-use efficiency with spectral reflectance, New Phytol. 131 (1995) 291–296.

[26] J.A. Gamon, K.F. Huemmrich, C.Y.S. Wong, I. Ensminger, S. Garrity, D. Y. Hollinger, A. Noormets, J. Peñuelas, A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers, Proc. Natl. Acad. Sci. U. S. A 113 (2016) 13087–13092.

[27] A.C. Burnett, J. Anderson, K.J. Davidson, K.S. Ely, J. Lamour, Q. Li, B.D. Morrison, D. Yang, A. Rogers, S.P. Serbin, A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression, J. Exp. Bot. 72 (2021) 6175–6189.

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

[29] K. Meacham-Hensold, P. Fu, J. Wu, S. Serbin, C.M. Montes, E. Ainsworth, K. Guan, E. Dracup, T. Pederson, S. Driever, et al., Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging, in: T. Lawson (Ed.), J. Exp. Bot. 71 (2020) 2312–2328.

[30] S.P. Serbin, D.N. Dillaway, E.L. Kruger, P.A. Townsend, 2012. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature, J. Exp. Bot. 63 (2012) 489–502.

[31] A. Porcar-Castell, Z. Malenovský, T. Magney, S. Van Wittenberghe, B. Fernández-Marín, F. Maignan, Y. Zhang, K. Maseyk, J. Atherton, L.P. Albert, et al., Chlorophyll a fluorescence illuminates a path connecting plant molecular biology to Earth-system science, Nat. Plants 7 (2021) 998–1009.

[32] Y. Sun, L. Gu, J. Wen, C. van der Tol, A. Porcar-Castell, J. Joiner, C.Y. Chang, T. Magney, L. Wang, L. Hu, et al., From remotely sensed solar-induced chlorophyll fluorescence to ecosystem structure, function, and service: part I-Harnessing theory, Glob. Change Biol. 29 (2023) 2926–2952.

[33] Y. Sun, J. Wen, L. Gu, J. Joiner, C.Y. Chang, C. van der Tol, A. Porcar-Castell, T. Magney, L. Wang, L. Hu, et al., From remotely-sensed solar-induced chlorophyll fluorescence to ecosystem structure, function, and service: part II-Harnessing data, Glob. Change Biol. 29 (2023) 2893–2925.

[34] K. Grossmann, C. Frankenberg, T.S. Magney, S.C. Hurlock, U. Seibt, J. Stutz, PhotoSpec: a new instrument to measure spatially distributed red and far-red solar-induced chlorophyll fluorescence, Rem. Sens. Environ. 216 (2018) 311–327.

[35] C. Camino, V. Gonzalez-Dugo, P. Hernandez, P.J. Zarco-Tejada, Radiative transfer vcmax estimation from hyperspectral imagery and SIF retrievals to assess photosynthetic performance in rainfed and irrigated plant phenotyping trials, Rem. Sens. Environ. 231 (2019) 111186.

[36] C.Y. Chang, R. Zhou, O. Kira, S. Marri, J. Skovira, L. Gu, Y. Sun, An unmanned aerial system (UAS) for concurrent measurements of solar-induced chlorophyll fluorescence and hyperspectral reflectance toward improving crop monitoring, Agric. For. Meteorol. 294 (2020) 108145.

[37] C.Y. Wong, M.E. Gilbert, M.A. Pierce, T.A. Parker, A. Palkovic, P. Gepts, T. S. Magney, T.N. Buckley, Hyperspectral remote sensing for phenotyping the physiological drought response of common and tepary bean, Plant Phenom. 5 (2023) 21. Washington, D.C.

[38] Z.A. Pierrat, T.S. Magney, R. Cheng, A.J. Maguire, C.Y.S. Wong, M.F. Nehemy, M. Rao, S.E. Nelson, A.F. Williams, J.A.H. Grosvenor, et al., The biological basis for using optical signals to track evergreen needleleaf photosynthesis, Bioscience 74 (2024) 130–145.

[39] Z.A. Pierrat, T.S. Magney, W.P. Richardson, B.R.K. Runkle, J.L. Diehl, X. Yang, W. Woodgate, W.K. Smith, M.R. Johnston, Y.R.S. Ginting, et al., Proximal remote sensing: an essential tool for bridging the gap between high-resolution ecosystem monitoring and global ecology, New Phytol. 246 (2025) 419–436.

[40] L.T. Helm, H. Shi, M.T. Lerdau, X. Yang, Solar-induced chlorophyll fluorescence and short-term photosynthetic response to drought, Ecol. Appl.: Publ. Ecol. Soc. Am. 30 (2020) e02101.

[41] T.S. Magney, M.L. Barnes, X. Yang, On the covariation of chlorophyll fluorescence and photosynthesis across scales, Geophys. Res. Lett. 47 (2020).

[42] C.Y. Chang, J. Wen, J. Han, O. Kira, J. LeVonne, J. Melkonian, S.J. Riha, J. Skovira, S. Ng, L. Gu, et al., Unpacking the drivers of diurnal dynamics of sun-induced chlorophyll fluorescence (SIF): Canopy structure, plant physiology, instrument configuration and retrieval methods, Rem. Sens. Environ. 265 (2021) 112672.

[43] J. Zhu, Y. Yin, J. Lu, T.A. Warner, X. Xu, M. Lyu, X. Wang, C. Guo, T. Cheng, Y. Zhu, et al., The relationship between wheat yield and sun-induced chlorophyll fluorescence from continuous measurements over the growing season, Rem. Sens. Environ. 298 (2023) 113791.

[44] M. Yoosefzadeh-Najafabadi, S. Torabi, D. Tulpan, I. Rajcan, M. Eskandari, Genome-wide association studies of soybean yield-related hyperspectral reflectance bands using machine learning-mediated data integration methods, Front. Plant Sci. 12 (2021) 777028.

[45] S. Liu, Z. Xiong, Z. Zhang, Y. Wei, D. Xiong, F. Wang, J. Huang, Exploration of chlorophyll fluorescence characteristics gene regulatory in rice (Oryza sativa L.): a genome-wide association study, Front. Plant Sci. 14 (2023) 1234866.

[46] H. Feng, Z. Guo, W. Yang, C. Huang, G. Chen, W. Fang, X. Xiong, H. Zhang, G. Wang, L. Xiong, et al., An integrated hyperspectral imaging and genome-wide association analysis platform provides spectral and genetic insights into the natural variation in rice, Sci. Rep. 7 (2017) 4401.

[47] X. Wu, H. Feng, D. Wu, S. Yan, P. Zhang, W. Wang, J. Zhang, J. Ye, G. Dai, Y. Fan, et al., Using high-throughput multiple optical phenotyping to decipher the genetic architecture of maize drought tolerance, Genome Biol. 22 (2021) 185.

[48] Z. Xiong, S. Liu, J. Tan, Z. Huang, X. Li, G. Zhuang, Z. Fang, T. Chen, L. Zhang, Combining hyperspectral techniques and genome-wide association studies to predict peanut seed vigor and explore associated genetic loci, Int. J. Mol. Sci. 25 (2024) 8414.

[49] C.R. Brodersen, T.C. Vogelmann, Do changes in light direction affect absorption profiles in leaves? Funct. Plant Biol.: FPB 37 (2010) 403.
[50] J. Liu, M.W. van Iersel, Photosynthetic physiology of blue, green, and red light: light intensity effects and underlying mechanisms, Front. Plant Sci. 12 (2021) 619987.
[51] E.L. Ashenafi, M.C. Nyman, J.T. Shelley, N.S. Mattson, Spectral properties and stability of selected carotenoid and chlorophyll compounds in different solvent systems, Food Chem. Adv. 2 (2023) 100178.
[52] J.R. Evans, Nitrogen and photosynthesis in the flag leaf of wheat (Triticum aestivum L.), Plant Physiol. 72 (1983) 297–302.
[53] A.U. Igamberdiev, P.J. Lea, The role of peroxisomes in the integration of metabolism and evolutionary diversity of photosynthetic organisms, Phytochemistry 60 (2002) 651–674.
[54] K. Sørhagen, M. Laxa, C. Peterhänsel, S. Reumann, The emerging role of photorespiration and non-photorespiratory peroxisomal metabolism in pathogen defence, Plant Biol. 15 (2013) 723–736.
[55] E. Babiychuk, K. Vandepoele, J. Wissing, M. Garcia-Diaz, R. De Rycke, H. Akbari, J. Joubès, T. Beeckman, L. Jänsch, M. Frentzen, et al., Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family, Proc. Natl. Acad. Sci. U. S. A 108 (2011) 6674–6679.
[56] Y.-W. Hsu, C.-T. Juan, C.-M. Wang, G.-Y. Jauh, Mitochondrial heat shock protein 60s interact with what's this factor 9 to regulate RNA splicing of ccmFC and rpl2, Plant Cell Physiol. 60 (2019) 116–125.
[57] T.S. Magney, Hyperspectral reflectance integrates key traits for predicting leaf metabolism, New Phytol. (2024).
[58] W. Cai, D. Ji, L. Peng, J. Guo, J. Ma, M. Zou, C. Lu, L. Zhang, LPA66 is required for editing psbF chloroplast transcripts in arabidopsis, Plant Physiol. 150 (2009) 1260–1271.
[59] L. Busemeyer, A. Ruckelshausen, K. Möller, A.E. Melchinger, K.V. Alheit, H. P. Maurer, V. Hahn, E.A. Weissmann, J.C. Reif, T. Würschum, Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation, Sci. Rep. 3 (2013) 2442.
[60] J.A. Bac-Molenaar, D. Vreugdenhil, C. Granier, J.J.B. Keurentjes, Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci, J. Exp. Bot. 66 (2015) 5567–5580.
[61] K. Neumann, Y. Zhao, J. Chu, J. Keilwagen, J.C. Reif, B. Kilian, A. Graner, Genetic architecture and temporal patterns of biomass accumulation in spring barley revealed by image analysis, BMC Plant Biol. 17 (2017) 137.
[62] D. Knoch, A. Abbadi, F. Grandke, R.C. Meyer, B. Samans, C.R. Werner, R. J. Snowdon, T. Altmann, Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola, Plant Biotechnol. J. 18 (2020) 68–82.
[63] T. Würschum, W. Liu, L. Busemeyer, M.R. Tucker, J.C. Reif, E.A. Weissmann, V. Hahn, A. Ruckelshausen, H.P. Maurer, Mapping dynamic QTL for plant height in triticale, BMC Genet. 15 (2014) 59.
[64] Y. Fu, H. Zhao, J. Huang, H. Zhu, X. Luan, S. Bu, Z. Liu, X. Wang, Z. Peng, L. Meng, et al., Dynamic analysis of QTLs on plant height with single segment substitution lines in rice, Sci. Rep. 12 (2022) 5465.
[65] C.Y.S. Wong, J.A. Gamon, Three causes of variation in the photochemical reflectance index (PRI) in evergreen conifers, New Phytol. 206 (2015) 187–195.
[66] N. Tilly, H. Aasen, G. Bareth, Fusion of plant height and vegetation indices for the estimation of barley biomass, Remote Sens. 7 (2015) 11449–11480.
[67] S. Guan, K. Fukami, H. Matsunaka, M. Okami, R. Tanaka, H. Nakano, T. Sakai, K. Nakano, H. Ohdan, K. Takahashi, Assessing correlation of high-resolution NDVI with fertilizer application level and yield of rice and wheat crops using small UAVs, Remote Sens. 11 (2019) 112.
[68] E. Benami, Z. Jin, M.R. Carter, A. Ghosh, R.J. Hijmans, A. Hobbs, B. Kenduiywo, D. B. Lobell, Uniting remote sensing, crop modelling and economics for agricultural risk management, Nat. Rev. Earth Environ. 2 (2021) 140–159.
[69] C. Lin, Plant blue-light receptors, Trends Plant Sci. 5 (2000) 337–342.
[70] C. Kami, S. Lorrain, P. Hornitschek, C. Fankhauser, Light-regulated plant growth and development, Curr. Top. Dev. Biol. 91 (2010) 29–66.
[71] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. R. Stat. Soc., Ser. B Stat. Methodol. 67 (2005) 301–320.
[72] W. Liu, Q. Li, An efficient elastic net with regression coefficients method for variable selection of spectrum data, PLoS One 12 (2017) e0171122.
[73] A. Marzougui, Y. Ma, R.J. McGee, L.R. Khot, S. Sankaran, Generalized linear model with elastic net regularization and convolutional neural network for evaluating aphanomyces root rot severity in lentil, in: Plant Phenomics, 2020, 2020 2393062. Washington, D.C.
[74] P. Mangalraj, B.-K. Cho, Recent trends and advances in hyperspectral imaging techniques to estimate solar induced fluorescence for plant phenotyping, Ecol. Indic. 137 (2022) 108721.
[75] G. Taylor, J.H. Bailey-Bale, M.C. Klein, S. Milner, J.-G. Chen, W. Muchero, P. Freer-Smith, T.J. Tschaplinski, J. Tuskan, Harnessing the power of poplar tree natural genetic variation for the development of future sustainable biofuels and bioproducts: a droughted marginal-land experiment for multi-disciplinary investigations, bioRxiv 2024 (2024), 01.11.575272.
[76] U. Platt, J. Stutz, Evaluation of DOAS spectra, sensitivity and detection limits, in: Physics of Earth and Space Environments, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 287–328.
[77] L.M. Evans, G.T. Slavov, E. Rodgers-Melnick, J. Martin, P. Ranjan, W. Muchero, A. M. Brunner, W. Schackwitz, L. Gunter, J.-G. Chen, et al., Population genomics of Populus trichocarpa identifies signatures of selection and adaptive trait associations, Nat. Genet. 46 (2014) 1089–1096.
[78] M.C. Klein, Z. Meng, J. Bailey-Bale, S. Milner, P. Shi, W. Muchero, J.-G. Chen, T. J. Tschaplinski, D. Jacobson, J. Lagergren, et al., Climate adaptation in*P. trichocarpa*: key adaptive loci identified for stomata and leaf traits, bioRxiv (2024) 603099, 2024.07.11.