



Integrative Metabolomics and Predictive Modelling of Doubled Haploid Induction in Wheat (*Triticum aestivum*)

Teresa Perez-Piñar^{1,2} · J. Eduardo Meyers-Angulo² · Armin Meister³ · Sandra Bössow⁴ · Heike Gnad^{4,5} · Hans-Peter Mock¹

Received: 9 July 2025 / Accepted: 3 March 2026 / Published online: 21 April 2026
© The Author(s) 2026

Abstract

Doubled haploid plants can be produced from immature female or male germ cells. The male germ line is mostly preferred since the number of microspores exceeds by far the number of egg cells. Applying this method, microspores need to be switched from the gametophytic to the sporophytic development. In contrast to conventional breeding programs, the time to obtain homozygosity with this method is shortened from approx. 6 years to approx. 1 year. However, the efficiency of anther culture in wheat has been reported to be genotype dependent and to result in a high proportion of albinism. Here, we studied the metabolic changes between *Triticum aestivum* DH lines with different regeneration rates, coming from a cross of the wheat cultivars Svilen and Berengar, contrasting in their ability to regenerate green plants. The metabolic analysis showed clear differences between the lines with high and low regeneration capability. The metabolic profiles found in terms of compounds were very similar in the lines, the differences resided mainly in their concentrations. Amino acids and organic acids were more abundant in the lines with higher rate of regeneration. In contrast, among sugars, galactose was found to be more abundant in the low regenerating lines. Based on the metabolic data, a predictive model was developed for the green plant regeneration rate. For the first time, a parsimonious linear model including isoleucine and phenylalanine accurately predicted regenerative competence, highlighting these amino acids as key metabolic indicators.

Highlight

We describe the metabolic differences during androgenesis up to green plant regeneration within a segregating wheat population produced via interspecific crosses with maize (*Zea mays*). The progeny of fixed DH lines differed in their tissue culture ability during anther culture. Based on selected metabolic data we developed a model to predict the green plant regeneration rate from the uninucleate stage of microspores.

Keywords Microspores · Green plant regeneration rate · Metabolites · Prediction · Androgenesis · Wheat

✉ Teresa Perez-Piñar
tperez@uax.es

J. Eduardo Meyers-Angulo
jmeyeang@uax.es

Armin Meister
meister@ipk-gatersleben.de

Sandra Bössow
boessow@saaten-union-biotec.com

Heike Gnad
heike.gnad@sgs.com

Hans-Peter Mock
hpmock@t-online.de

¹ Department of Physiology and Cell Biology, Applied Biochemistry, Leibniz Institute of Plant Genetics and Crop Plant Research, Corrensstraße 3, 06466 Seeland/OT Gatersleben, Germany

² Present address: Facultad de Ciencias Biomédicas y de la Salud, Universidad Alfonso X el Sabio, Avenida de la Universidad 1, Villanueva de la Cañada, Madrid 28691, Spain

³ Department of Breeding Research, Quantitative Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research, Corrensstraße 3, 06466 Seeland, Gatersleben, OT, Germany

⁴ Saaten-Union Biotec GmbH, Am Schwabeplan 6, 06466 Seeland/OT Gatersleben, Germany

⁵ Present address: TraitGenetics Section, SGS Institut Fresenius GmbH, Am Schwabeplan 1b, 06466 Seeland/OT Gatersleben, Germany

Abbreviations

DH	Doubled haploid
GC-MS	Gas chromatography-mass spectrometry
GPR	Green plant regeneration
TCA	Tricarboxylic acid

Introduction

Microspore embryogenesis, also called androgenesis, is a cellular reprogramming process in which a microspore is switched from a gametophytic to a sporophytic development, usually after the application of a stress pre-treatment (Zheng et al. 2003). This process can be induced to obtain doubled haploid (DH) plants (Mishra and Goswami, 2014). In conventional breeding programs up to 6 cycles of self-pollination are usually required to achieve a desired genotype with high percentage of homozygosity (De La Fuente et al. 2013). DH plants represent an alternative (Murovec and Bohanec 2012) that widely reduces the time needed to obtain homozygous lines and thus, cultivars.

The main advantage of androgenesis resides in the high number of microspores that can be obtained from a single plant, which have the potential to regenerate into a new plant. On the other hand, genotype dependency and albinism represent the prevailing disadvantages (Lantos et al. 2013).

The microspore embryogenesis is typically initiated with the application of a stress pretreatment (Fig. 1). The stress is applied when the microspore is in the mid- to late- uninucleate stage of the microsporogenesis in the anther (also called the vacuolated stage). With this pretreatment, the

nucleus is displaced to the center of the cell, dividing the vacuole into several small vacuoles, giving the microspore an internal appearance of a star-like pattern (Maraschin et al. 2005). Subsequent cell divisions lead to embryogenesis or organogenesis, and when successful, a green plant will be regenerated. This approach has been widely applied to obtain doubled haploid plants (Kambale et al. 2023; Yang et al. 2024).

The efficiency of the microspore embryogenesis during in-vitro culture depends on several factors such as the developmental stage of the microspore (Hoekstra et al. 1992), the stress pretreatment used to induce embryogenesis (Žur et al. 2009), and the media composition (Saharan et al. 2004). Different strategies have been tested to overcome the poor response in recalcitrant genotypes. For example, ovary-conditioned medium has been shown to improve the efficiency in recalcitrant genotypes of wheat (Zheng et al. 2002). Addition of indoleacetic acid has improved the efficiency of microspore embryogenesis in *Brassica campestris* (Wang et al. 2023). Plant growth regulators were studied to improve the regeneration rates in rice and pepper (Cheng et al. 2013; Khalequzzaman et al. 2005). Among them, auxin has been found to be involved in microspore embryogenesis (reviewed in Testillano 2019).

Even after optimization of a protocol, the influence of the individual genotypes remains one of the biggest challenges to handle. Another bottleneck strongly influenced by the genotype is albinism, which is characterized by a deficiency in chloroplasts. It has been studied in several species such as wheat, barley, oat or rye (reviewed in Makowska and Oleszczuk 2014). Such plants cannot survive to maturity and reduce total green-plant yields (Žur et al. 2022). On

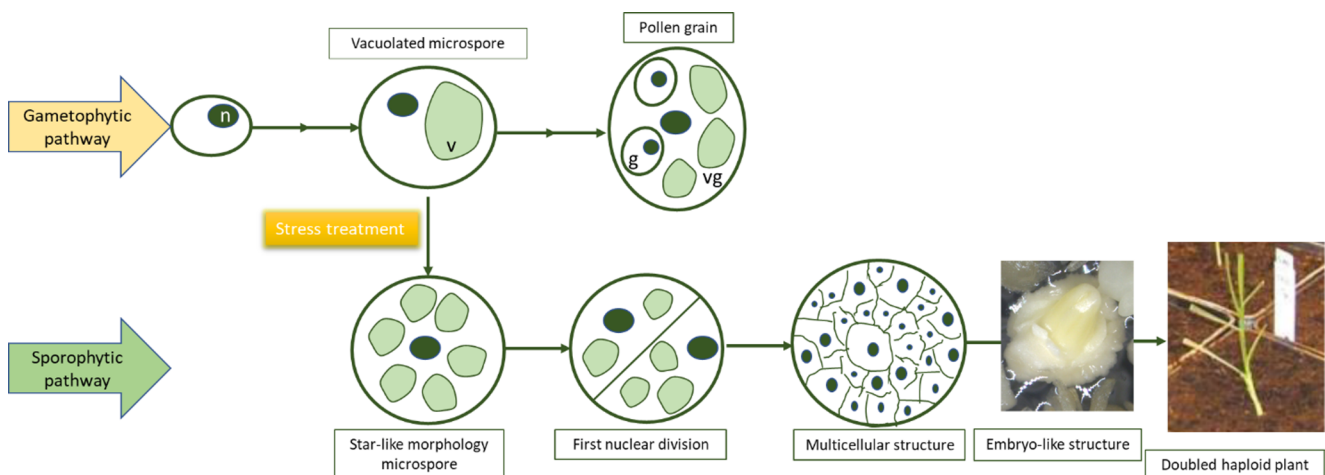


Fig. 1 Gametophytic and microspore embryogenesis pathways in bread wheat. In the gametophytic pathway, different steps of the development of a pollen grain are represented. From a mother cell, later a “Vacuolated microspore” appears. Finally, the pollen grain is formed with two spermatid nuclei. The microspore embryogenesis pathway starts after the application of a stress treatment. After that, the cell

usually appears with a star-like morphology. It divides forming first multicellular structures, and later embryo-like structures. Cellular division and differentiation continue until a green plant develops. Abbreviations: n-nucleus, v-vacuole, g-generative cell, vg-vegetative cell. Pictures correspond to a *T. aestivum* embryo like structure (2–3 mm, approx.) and a doubled haploid plant (10 cm approx.)

the other hand, genotype dependence has been studied in several species like *Triticum turgidum* (Arzani and Mirodjagh 1999), *Aegilops speltoides* (Zale et al. 2004) or *Oryza sativa* (Hoque and Mansfield 2004).

Genetic control of plant regeneration has been shown to involve multiple genes in the study of Mano et al. (1996). Some of the genes control the callus growth, and others, the shoot regeneration in barley. In bread wheat, genetic markers associated with microspore embryogenesis were described in cultivars with different regeneration rates (Sánchez-Díaz et al. 2013; Schlegel et al. 2000). Furthermore, epigenetic modifications are involved in microspore embryogenesis (reviewed in Testillano 2019). The use of demethylating agents together with a stress treatment has been used to induce microspore embryogenesis in triticale (Krzewska et al. 2021). Epigenetic mechanisms that control microspore embryogenesis efficiency have been reported in bread wheat (Valero-Rubira et al. 2023).

Comparable genotype effects on regeneration efficiency are also known in somatic embryogenesis of other crops (Raza et al., 2019; Carra et al., 2024). In coffee, embryogenic capacity was shown to be stable across early generations (Molina et al. 2002). The authors suggested the predictability of the embryogenic capacity of the F5 lines based on that of the F3 and F4 ancestors' lines.

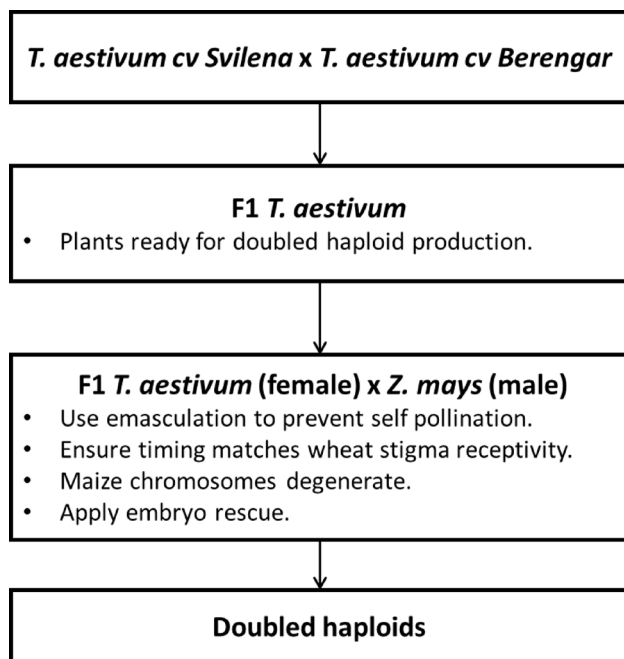


Fig. 2 Doubled haploid production in *T. aestivum*. The crossing strategy is illustrated. Two parental wheat varieties were crossed to generate the F1 generation (*T. aestivum* cv. Svilenia and *T. aestivum* cv. Berengar). F1 plants were pollinated with *Zea mays* (maize). After fertilization, maize chromosomes degenerate resulting in wheat haploid lines. The haploids are subsequently treated to double the chromosome number, producing homozygous doubled haploid plants

Other bottlenecks in microspore and somatic embryogenesis include oxidative stress, abnormal embryo development, inefficient chromosome doubling or poor reproducibility (Cheng et al. 2025; Fomicheva et al. 2024; Wang et al. 2023). However, advances have also been identified, such as a modified procedure to reduce oxidative stress and improve the effectiveness of the process in wheat (Dubas et al. 2024).

Despite advances in knowledge of both androgenesis and somatic embryogenesis, no metabolic predictive models for the regeneration efficiency have been reported to date. Here, we address this gap by examining metabolic differences during androgenesis between *Triticum aestivum* DH lines differing in their green plant regeneration (GPR) efficiency. We developed a model allowing to predict the green plant regeneration rate of androgenic wheat DH lines based on few metabolites out of the metabolic dataset.

Materials and Methods

Plant Material

A population of 137 doubled haploid lines was generated using the highly responsive to tissue culture winter wheat cultivar Svilenia (Schlegel et al. 2000) and the tissue culture recalcitrant cultivar Berengar, using interspecific crosses of F1-plants with maize (Fig. 2). After fertilization, maize chromosomes degenerate resulting in wheat haploid lines. The haploids are subsequently treated to double the chromosome number, producing homozygous doubled haploid plants.

This egg-cell-based method was used to avoid any bias in our further analysis studying the microspore-based anther culture. The population of 137 doubled haploid lines were regenerated and phenotyped for their responsiveness in anther culture following the protocol described in Rubtsova et al. (2013). The green plant regeneration (GPR) rate was defined as the number of green plants, which were regenerated after previous induction of microspores per spike. The eight highest and lowest responding doubled haploid lines each were selected, i.e. 16 DH-lines and their parental lines were used for further analysis. Metabolomics and predictive modelling were conducted on the subset of selected genotypes with complete metabolite measurements at each time point (a minimum of 13 lines were included in all analyses; Supplementary_data_S1 and Supplementary_data_S2).

Time Course of Androgenesis up to Green Plant Regeneration and Microspore Isolation for Metabolic Analysis

Donor plants of the 16 pre-selected DH-lines and their 2 parental lines were grown as described in Rubtsova et al. (2013). For anther culture, the pre-treatment of spikes for 10 days, in which the temperature was set to 4 °C and the following in-vitro culture of anthers were performed in the dark at 22 °C as described in Rubtsova et al. (2013).

Microspores were isolated in three stages, as described in Seifert et al. (2016). In the first stage (T1) untreated microspores were at the late vacuolated uninucleate stage and they were directly isolated from freshly harvested spikes. In the second time point (T2), microspores were slightly enlarged in the anthers and with a visible star-like structure after a cold stress pre-treatment of 10 days. In the third stage (T3), after 4–8 days in in-vitro culture, enlarged microspores with first visible karyokinesis were isolated directly out of in-vitro anther culture. The developmental stage T3 of the microspores was checked daily by squashing anthers out of in-vitro anther cultures using a 20x magnification and determining the microspore stage. When 5–10% of microspores showed visible karyokinesis, all anthers of the petri dish were used for isolation (Fig. 1).

To evaluate the green plant regeneration rate for each DH-line and the parental lines, up to 12 spikes per line were not used for the isolation of microspores for metabolic analysis, but used to regenerate in-vitro plants via anther culture. From these spikes, induced somatic embryos were transferred to regeneration media and the number of green and albino plants was counted as described in Lantos et al. (2013).

Microspores for the 3 different time point T1, T2, and T3 were isolated exactly as described in Perez-Piñar et al. (2024).

For each of the three time points, at least three biological replicates were analyzed, each represented by three technical replicates.

Gas-chromatography Mass Spectrometry (GC-MS) Analysis

An aliquot of the lysate corresponding to 6000 microspores was vacuum-dried. Samples from at least 6 good responding and 6 bad responding DH lines as well as their 2 parental lines were extracted from the three developmental stages T1, T2, T3, derivatized, measured and results analyzed according to Perez-Piñar et al. (2024).

Amino Acid Analysis

A second aliquot of the lysate of all samples corresponding to 2000 microspores was vacuum-dried and processed for amino acid analysis by UPLC as described in Perez-Piñar et al. (2024).

Correlation Analysis

The Spearman correlation was calculated between the concentration of each metabolite and the average GPR rate (averaged from at least three biological replicates from each line), and also with the albino plant regeneration rates, using SigmaPlot v13.0 (Systat software Inc, San Jose, CA, USA). For amino acids two values were obtained, one from the untargeted (GC-MS) and the other one from the targeted (UPLC) analysis and both were included in the correlation study.

Statistical Analysis of Metabolite Dynamics

To analyze metabolite changes across microspore developmental stages (T1, T2 and T3) in relation to green plant regeneration capacity, doubled haploid lines were classified into two contrasting groups (high-GPR and Low-GPR). Metabolite concentrations from replicates were averaged per genotype and time point. For each metabolite, two complementary statistical analyses were performed:

- (i) High-GPR vs. Low-GPR comparisons. For each time point (T1, T2 and T3), metabolite levels were compared between the two groups using the Mann-Whitney U test (two sided). Multiple testing correction was performed within each time point using the Benjamini-Hochberg false discovery rate (FDR) procedure.
- (ii) Time points comparisons (T1, T2 and T3). Within each GPR group, temporal changes were analyzed using the Kruskal-Wallis one way ANOVA on ranks. P-values were corrected for multiple testing using the Benjamini-Hochberg FDR method.

In both analyses, metabolites with FDR-adjusted q-values < 0.05 were considered significantly different.

A heatmap was developed representing the z-score-normalized concentrations of all quantified metabolites across the three developmental stages (T1, T2 and T3), and the two contrasting GPR groups (high vs. low).

Principal Component Analysis plot was performed to scaled data (z-score), to provide an unbiased overview of the data.

All analyses and graphics were developed in R using base functions (`wilcox.test` and `Kruskal.test`), and the `dplyr` and `readxl` packages for data processing.

Predictive Model Development

For predictive modelling, the experimental unit was the DH line at stage T1. The dataset comprised 13 DH lines with complete T1 data and replicate measurements averaged. Several modeling strategies were evaluated, including classical linear regression, penalized regression models (Ridge and LASSO), PCA-based dimensionality reduction, and variable-selection approaches. For classical linear regression, a forward stepwise regression analysis was first carried out with Sigmaplot v15.0 (Systat, Software Inc., San José, CA, USA). Remaining analyses were performed in R software (R Core Team 2024). Model performance was assessed through cross-validation schemes appropriate to the dataset structure, primarily leave-one-out cross-validation (LOOCV) at the line level. Model selection was guided by cross-validation performance (predictive ability), parsimony, and multicollinearity diagnostics (VIF). Variance inflation factors (VIF) values above 10 were interpreted as indicative of severe multicollinearity.

Based on this selection workflow, the final predictive model chosen was a parsimonious linear regression including two metabolites (Isoleucine_UPLC and Phenylalanine_UPLC). This model provided the best balance between interpretability, coefficient stability, and cross-validated predictive performance, while avoiding the multicollinearity and overfitting observed in more complex models.

Results

Effect of Genotype on Green Plant Regeneration

A DH-mapping population for testing the GPR rate was generated by crossing cultivars Svilená and Berengar, with high and low GPR rates, respectively, and using the corresponding F1 plants for interspecific crosses with maize. With this approach, any bias associated with genotype dependence for androgenesis could be excluded during selection. The resulting population contained 137 DH lines ranging from a GPR of 0.00 green regenerated plants per spike in line 001 or 0.22 in Berengar, to 37.00 green regenerated plants per spike in line 081 or 34.00 in Svilená (Supplementary_data_S3), showing clearly that not only the genotype, but also the phenotype was different between the lines, and that green plant regeneration is a quantitative trait within this population. In addition to the green plant regeneration, the albino plant regeneration was determined. Albino plant

regeneration was low for all the lines, ranging from 0.00 albinos per spike in several lines (e.g. line 001) to 1.96 albinos per spike in line 069.

For further analysis of plant regeneration capacity, two contrasting groups were formed: the 8 DH lines with lowest GPR and the 8 lines with highest GPR, and both parents (Fig. 3).

Untargeted and Targeted Analysis of Metabolites in DH-lines with High and Low GPR Rates

Primary metabolism was studied in enriched populations of microspores from 3 different time points (Fig. 1) using contrasting DH lines and their parents differing in their regeneration rate in order to find metabolic markers for the GPR. Relative concentrations of primary metabolites were measured by an untargeted analysis using GC-MS as described in the Materials and Methods section. More than 80 compounds were detected, with relative quantification possible for 60 of them (Fig. 4). The metabolites detected could be grouped into amino acids, organic acids, sugars, fatty acids and unknown compounds, in which annotation was not possible with the spectrum library used. Most of them showed significant temporal variation among T1, T2 and T3 stages (Supplementary_data_S1).

The overall metabolite profile showed an increase in several amino acids at the T3 stage (Fig. 4), notably asparagine, ornithine, phenylalanine (Fig. 5) or gamma-amino butyric acid (GABA). Others like alanine, glutamine and glycine were found with a tendency to increase in T2 and to decrease in T3. Organic acids and sugars showed a tendency to increase in T2 and to decrease in T3 stage, except succinic acid, which decreased across the stages. Within the sugars, glucose and sucrose showed the highest concentrations. Slight changes were found in fatty acids. In other compounds, such as putrescine, an increase in T2 and a later decrease in T3 were observed. Finally, within the unknown compounds, unknown #4 and unknown named D were seen to increase in T3 stage.

The amino acid with the highest concentration in T1 was alanine in both high- and low-GPR lines, with 84 pmol per 100 microspores and 32 pmol per 100 microspores on average, respectively. In T2, alanine continued to be the most abundant amino acid, with 170 pmol per 100 microspores in the high-GPR lines and 101 pmol per 100 microspores on average in the low-GPR lines. In T3, the most abundant amino acids were glutamine and asparagine, reaching more than 170 pmol per 100 microspores on average in both amino acids in high-GPR lines, and more than 120 pmol per 100 microspores on average in both in low-GPR lines (Supplementary_data_S1).

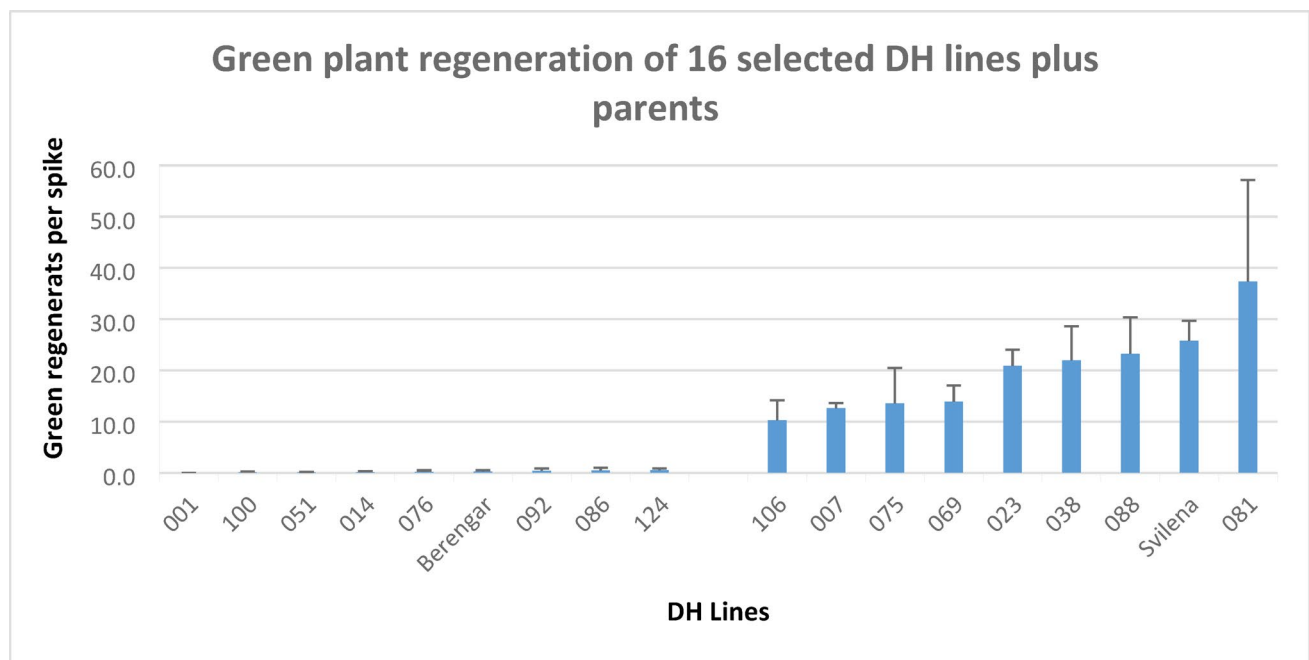


Fig. 3 Green plant regeneration rates of 16 selected DH lines coming from the cross of cultivars Svilena and Berengar. The green plant regeneration (GPR) rate was defined as the number of green plants, which were regenerated after previous induction of microspores per spike. Up to 12 spikes per line were used to regenerate in-vitro plants via anther culture. From these spikes, induced somatic embryos were transferred to regeneration media and the number of green and albino

plants was counted as described in Lantos et al. (2013). Mann Whitney U test between good and bad responsive DH lines showed significant differences ($p < 0.001$). Units are given in number of green regenerants per spike. The error bars represent the standard deviation across three independent GPR rates, which reflects the natural variation observed in some lines

Both untargeted and targeted analyses showed similar trends across stages in both groups of lines, although concentration levels differed between them (Fig. 4).

Effect of Metabolites on GPR

In order to study the metabolic implications for the GPR and to detect potential metabolite markers, firstly, comparisons were made between high-GPR and low-GPR lines for each metabolite and time point. Secondly, a correlation analysis was performed between the concentrations of the metabolites detected and the GPR rate.

A large fraction of metabolites differed significantly between high- and low-GPR DH lines, explaining a high percentage of the variance (Fig. 6). Out of the 60 metabolites, 31 showed significant differences between the GPR groups at T1 stage, 19 at T2, and 13 at T3 (Supplementary_data_S1). Amino acids, TCA intermediates, sucrose and putrescine consistently showed significantly higher concentrations in high-GPR lines across stages (Fig. 4). Within the sugars, galactose presented a tendency of lower concentrations in the high-GPR lines than the low-GPR ones. The remaining sugars and fatty acids showed similar concentrations in all genotypes.

Regarding correlation (Table 1), in general, amino acids showed positive correlation with the GPR rate in the three studied stages. Phenylalanine presented the highest correlation, with 0.792 in T1 and 0.693 in T2, and histidine was the highest correlated amino acid in T3, with 0.479. Organic acids of the TCA cycle also showed positive correlation. Fumaric acid and malic acid had the highest correlation values in T1, with 0.698 and 0.691, respectively. In T2, the highest correlation was found in malic acid, with 0.487. In T3 we found citric acid, with 0.413. The remaining organic acids were detected with lower correlations. Concerning sugars, either no correlation or negative correlation was shown, being the highest negative detected in galactose, with -0.109 in T1 stage, -0.498 in T2 and -0.367 in T3. Sucrose was found with positive correlation of 0.5. Other compounds had low values. Finally, fatty acids had either small negative or no correlation, only a small increase in correlation values was seen in T2 stage.

GPR Markers and Predictive Model

The prediction of the GPR rate (the dependent variable) was achieved by using the values of the metabolites previously detected and identified (independent variables or predictors) at the T1 stage.

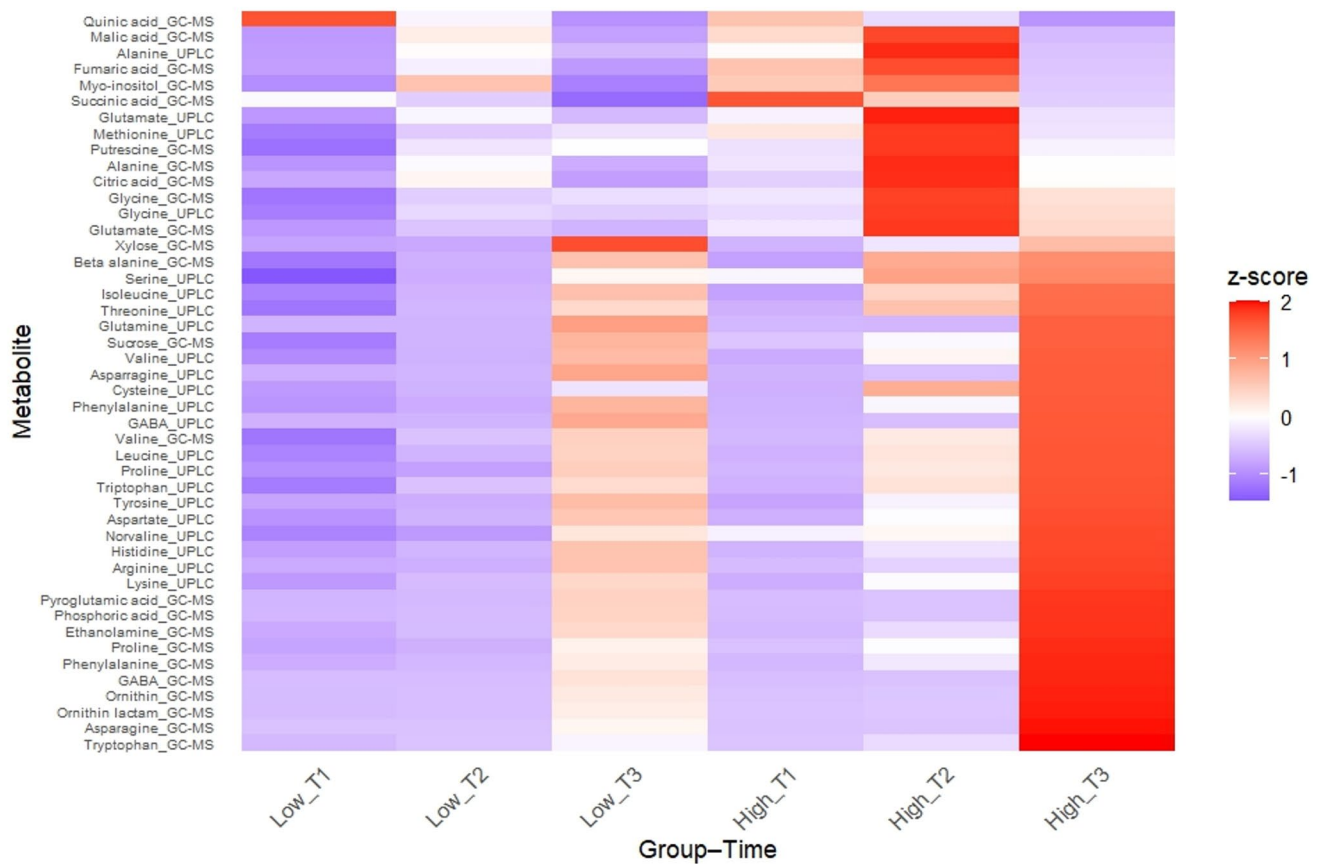


Fig. 4 Heatmap of relative concentrations of metabolites of DH lines differing in their regeneration capacity. The three stages sampled were: T1, uninucleate microspores freshly harvested; T2, uninucleate microspores following a 10 days pre-treatment of low temperature; and T3, microspores showing first nuclear divisions after 4–8 days in culture. Samples from at least 6 good responding and 6 bad responding DH lines as well as their 2 parental lines, each with 6000 microspores, were extracted from the three developmental stages T1, T2, T3, deriva-

tized, measured and results analyzed according to Perez-Piñar et al. (2024). Detection was performed with GC-MS and UPLC. Units: relative units (z-score normalized concentrations). The differences between lines with high and low GPR rate were statistically evaluated by the Mann-Whitney U test. The differences across the stages (T1, T2 and T3) were evaluated by applying the Kruskal-Wallis ANOVA on ranks test, * $p < 0.05$

A total of six predictive models (M1–M6) were generated and evaluated to determine the most suitable approach for predicting the trait of interest from the metabolomic dataset (Supplementary_data_S2). The models differed in the number of predictors, the type of regression method, and the validation schemes applied.

Two linear regression models were selected for detailed evaluation due to their relevance and performance: the full seven-metabolite model (M1) and the parsimonious two-metabolite model (M2):

M1

$$GPR = 1.648 + 5.621 (Phenylalanine_{UPLC}) + 0.569 (Serine_{UPLC}) + (-2.1) (Aspartate_{UPLC}) + (-0.063) (Galactose_{GC_{MS}}) + 0.391 (Alanine_{GC_{MS}}) + (-0.879) (Ethanolamine_{GC_{MS}}) + 0.898 (GABA_{GC_{MS}})$$

M2

$$GPR = -10.910 + 7.750 (Isoleucine_{UPLC}) + 7.186 (Phenylalanine_{UPLC})$$

The selected model (M2) was a parsimonious linear regression including only two metabolites, Isoleucine_UPLC and Phenylalanine_UPLC, which were consistently identified as relevant predictors both in the reduced model and through LASSO variable selection. Despite its simplicity, the model showed a solid in-sample fit ($R^2 = 0.811$) and acceptable performance under leave-one-out cross-validation (LOOCV $R^2 = 0.667$; RMSE=7.91). Multicollinearity was negligible, with both predictors displaying low VIF values (VIF_Isoleucine=5.63; VIF_Phenylalanine=5.63), resulting in stable and interpretable coefficient estimates. Compared with the full model, M2 offered improved robustness, clearer

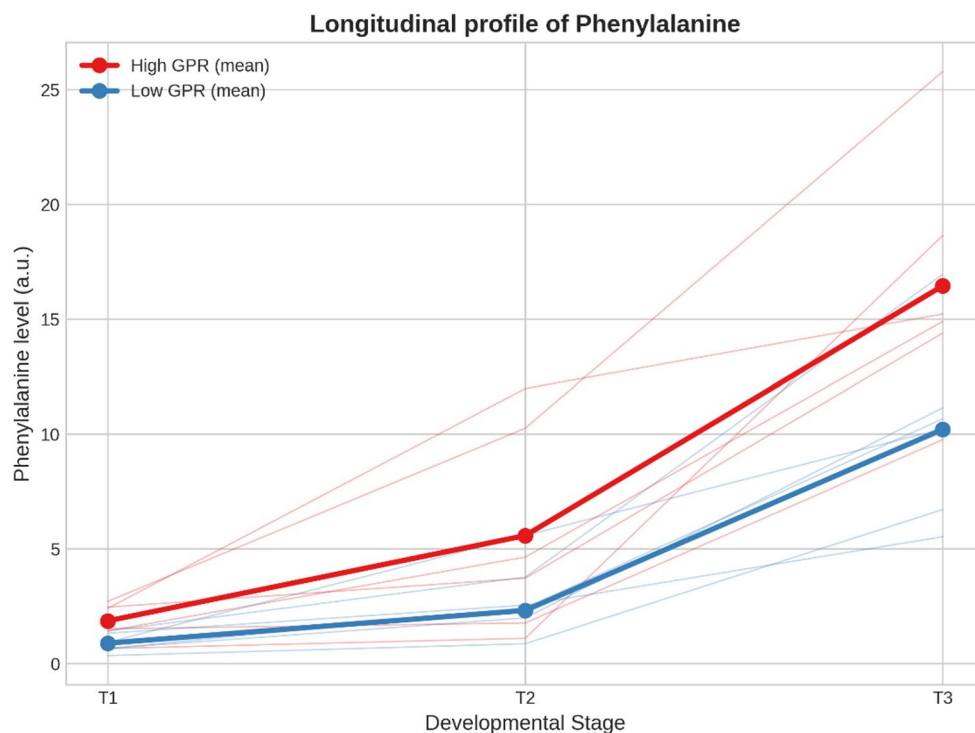


Fig. 5 Phenylalanine concentrations in microspores of DH lines in T1, T2 and T3, in lines with highest GPR and in lines with lowest GPR, including the parents. The three stages sampled were: T1, uninucleate microspores prior to low temperature treatment; T2, uninucleate microspores following a 10 days-exposure to low temperature stress; and T3, microspores showing first nuclear divisions after 4–8 days in culture. Samples from at least 6 good responding and 6 bad responding DH lines as well as their 2 parental lines, each with 2000 microspores, from the three developmental stages T1, T2, T3, were derivatized, and

analyzed as described in Perez-Piñar et al. (2024). Detection was performed with UPLC. Units are given in arbitrary units (z-scored concentrations). The differences between lines with highest and lowest GPR were statistically evaluated by the Mann-Whitney U test. The differences across the T1, T2 and T3 stages were evaluated by applying Kruskal-Wallis ANOVA on ranks test, * $p < 0.05$. Spaghetti plot shows individual genotype trajectories (thin lines) and group means (thick lines) for phenylalanine levels across the stages

biological interpretability, and higher generalizability, making it the most reliable and consistent predictive model in this study. Also only analysis of amino acids is required.

In order to increase the prediction efficiency, another model evaluated (M1) included seven metabolites that had been previously selected through a forward step-wise regression procedure (Supplementary_data_S4). For consistency with the rest of the analyses, this model was reconstructed and refitted in R. Although it showed a high degree of explained variance ($R^2 = 0.982$; LOOCV $R^2 = 0.855$; RMSE_LOOCV = 5.15 GPR units), diagnostic checks revealed marked multicollinearity among predictors and unstable coefficient estimates, with several metabolites VIF values exceeding 10 (VIF_Alanine_GC_MS = 18.64; VIF_Ethanolamine_GC_MS = 17.10; VIF_Aspartate_UPLC = 12.77). These results limit its interpretability and reduce its reliability for prediction.

Discussion

In our study, we showed a strong correlation of the green plant regeneration rate and the metabolic patterns of the microspores in the T1 stage. We identified the amino acid, organic acid and sugar metabolism as key pathways associated with the green plant regeneration rate. In particular, several potential GPR markers were found, and used in a model, which serves as a tool for the early prediction of the GPR.

Effect of Metabolites on GPR: High Concentrations of Amino Acids and Organic Acids are Markers for a successful Regeneration

Almost all the metabolites showed a similar profile in good and bad responding DH-lines across the three stages (T1, T2 and T3), in line with results from our previous study of Svilen alone (Perez-Piñar et al. 2024). The main differences between both groups were quantitative rather than qualitative, with amino acids and organic acids consistently

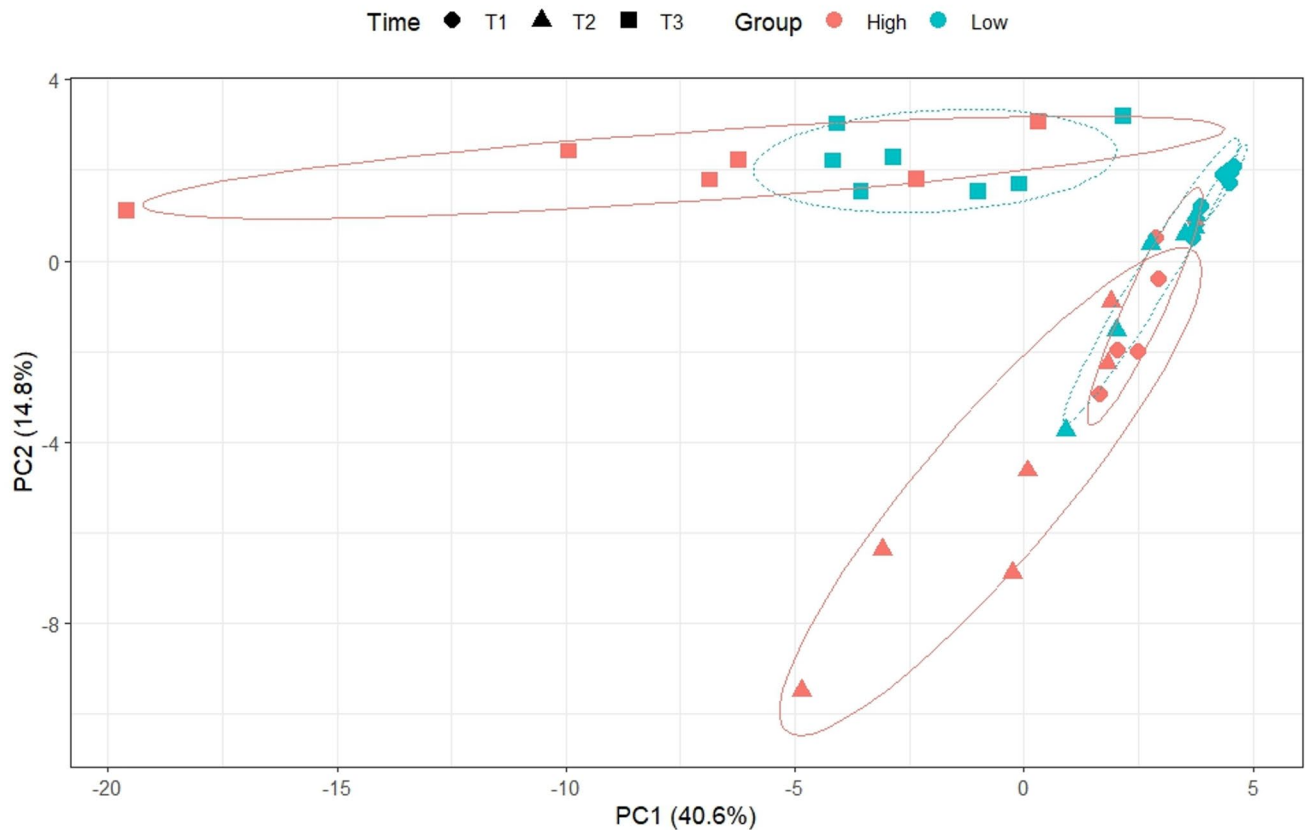


Fig. 6 Principal Component Analysis of metabolite profiles in high- and low-GPR lines, across stages (T1, T2 and T3). The three stages sampled were: T1, uninucleate microspores freshly harvested; T2, uninucleate microspores following a 10 days pre-treatment of low temperature; and T3, microspores showing first nuclear divisions after 4–8 days in culture. Samples from at least 6 good responding and 6 bad responding DH lines as well as their 2 parental lines, each with 6000 microspores, were extracted from the three developmental stages T1,

T2, T3, derivatized, measured and results analyzed according to Perez-Piñar et al. (2024). Detection was performed with GC-MS. Units: relative units (z-score normalized concentrations). Each point represents one line at a given time point (T1, T2 and T3). Colors indicate GPR group (high in red and low in blue), and shapes indicate sampling time (T1, circle; T2, triangle; T3, square). The PC1 explains 40.6% of variance, whilst PC2 explains the 14.8%. Elipses represent 68% confidence regions for each group

showing higher concentrations in good responding lines across all stages. In sugars, fatty acids and other compounds the differences between the two groups were not so obvious, showing only minor differences.

The involvement of amino acids in increasing the regeneration capacity in wheat embryos has been shown only by external application of these metabolites (Duran et al. 2013; Hashemi et al. 2023). They are involved in protein, nitrogen and carbohydrate metabolism, and also act as precursors of secondary metabolites. Here, the high-GPR lines could have enough of these resources to overcome the stress and later steps.

Regarding sugars, genes related to carbohydrate metabolism have been associated with heat-treated non-dividing microspores of *Brassica napus* (Malik et al. 2008). In our study this was shown in few sugars, such as galactose or fructose. Particularly for galactose, the low regenerating lines showed higher concentrations after the pretreatment compared to the responsive genotypes. Additionally,

galactose is known to be part of plant cell wall polymers. Modifications in cell wall protein pectin have been shown to contribute to freezing tolerance. Particularly, an increase in β -1,4-galactan in cell walls of *Arabidopsis* has been shown in cold acclimation (Takahashi et al. 2024). Therefore, the high concentrations of galactose in the lower level regenerating microspores may reflect an altered cell wall metabolism in those lines. We suggest the further study of the enzyme β -1,4-galactan synthase, but also glucose epimerase, which interconverts UDP-glucose and UDP-galactose, since a difference either in abundance or in activity may also explain the result shown here. This suggestion would be supported by the higher glucose concentration of the lines with good regeneration rate found.

In this study, lines with high GPR rate were metabolically better prepared in each phase to continue with the microspore embryogenesis, since amino acids and organic acids are significant metabolites of microspore embryogenesis. On the other hand, the DH lines with low GPR rate, may

Table 1 Correlation coefficients between metabolites and GPR rates and albino regeneration rates. Spearman correlation coefficient of metabolic GC-MS relative concentrations and absolute concentrations obtained by UPLC measurements and both green plant regeneration and albinos plant regeneration rates of the studied DH lines. Non-significant correlations ($p > 0.05$) are shown with red characters. Heat map applied to the values, going from green in value 1 to yellow in value 0, and to red in value -1

Alanine	GPR T1	GPR T2	GPR T3
	0.597	0.55	0.0939
Arginine	0.503	0.479	0.467
Asparagine	0.251	0.533	0.169
Aspartic acid	0.327	0.412	0.198
Beta alanine	0.226	0.343	0.252
Cystein	0.124	0.412	0.143
GABA	0.44	0.669	−0.00867
Glutamine	0.424	0.393	0.214
Glutamic acid	0.629	0.566	0.23
Glycine	0.608	0.635	0.343
Histidine	0.432	0.678	0.479
Isoleucine	0.721	0.606	0.27
Leucine	0.703	0.599	0.408
Lysine	0.613	0.494	0.456
Metionine	0.625	0.514	0.0581
Ornithin	0.449	0.455	0.276
Ornithin lactam	0.518	0.47	0.424
Phenylalanine	0.792	0.693	0.431
Proline	0.725	0.603	0.335
Serine	0.654	0.662	0.41
Threonine	0.464	0.614	0.31
Tryptophan	0.488	0.0707	0.27
Tyrosine	0.441	0.658	0.42
Valine	0.598	0.338	0.21
Citric acid	0.508	0.277	0.413
Fumaric acid	0.698	0.373	0.269
Malic acid	0.691	0.487	0.203
Phosphoric acid	0.439	0.422	0.384
Propanoic acid	0.213	0.231	0.26
Pyroglutamic acid	0.483	0.243	0.252
Quinic acid	0.0292	0.089	0.00897
Succinic acid	0.344	0.166	0.372
Threonic acid	0.0314	0.199	−0.209
Erythritol	0.057	−0.144	−0.144
Fructose	−0.114	−0.0689	−0.236
Galactose	−0.109	−0.498	−0.367
Glucose	0.0548	0.201	−0.0233
Glycerol	0.0755	0.0904	0.223
Mannose	0.0563	−0.106	−0.115
Myo-inositol	0.39	0.0442	−0.0286
Sucrose	0.5	0.181	0.272
Xylose	0.28	0.0388	−0.25
Ethanolamine	0.197	0.355	0.454
Putrescine	0.335	0.217	−0.0527
Urea	0.295	0.183	0.206
Unknown4	0.338	0.164	0.209
Unknown23	0.41	0.0809	0.156
Unknown27	0.193	0.288	0.255
Unknown36	0.301	−0.14	−0.197
UnknownD	0.0384	0.0306	−0.132
Lauric acid	−0.115	0.163	0.0636
Myristic acid	−0.0764	0.151	−0.0563

Table 1 (continued)

Alanine	GPR T1	GPR T2	GPR T3
	0.597	0.55	0.0939
Palmitic acid	−0.0866	0.374	0.139
Stearic acid	0.0278	0.305	−0.0305
Unknown19	0.247	0.14	−0.135
Unknown20	0.246	0.248	0.0838
Unknown22	0.0697	0.193	0.0181
Unknown27	−0.0771	0.0651	0.1
Unknown32	−0.206	−0.0285	−0.124
Unknown34	0.0855	0.0413	0.159
Unknown41	−0.197	0.17	0.225
Legend	1	0	−1

not efficiently accomplish the reprogramming and embryogenesis steps, because the levels of these metabolites were not high enough.

Since these differences are focused on individual metabolites at each stage of the process, a predictive model can provide information about metabolic pathways involved in the process.

Future work including secondary metabolites and hormonal profiling may provide additional insights into the regulation of green plant regeneration. In addition, linking these metabolic markers to genomic loci through QTL mapping would further strengthen the understanding of the genetic basis of green plant regeneration.

GPR Rate can be Predicted before the Microspore Embryogenesis Process Starts

Metabolic markers can be used to estimate the performance of a plant (Fernandez et al. 2016) also under stress conditions (Cardoso et al. 2023; Mashabela et al. 2023). Diverse conditions in different studies exhibit metabolic markers, such as trehalose in drought-tolerant rice (Quistián-Martínez et al. 2011), or allantoin in salt-stress tolerant rice (Nam et al. 2015). Among others, *BABY BOOM1* in *Brassica napus* can be used as a genetic marker of the microspore embryogenesis efficacy, but only after the induction of the microspore embryogenesis (Malik et al. 2007). In our study, we found different potential metabolite markers for the regeneration rate of the DH lines in the vacuolated stage, prior to the induction phase.

The comparison between the six exploratory models revealed that the model with seven metabolites (M1), suffered from substantial multicollinearity, limiting its practical usefulness. However, it showed excellent apparent fit. Such behavior is consistent with the challenges typically associated with high-dimensional metabolomic datasets, where correlated predictors can overestimate parameters and can increase the risk of overfitting, compromising model generalization (Otse et al. 2025). In contrast, the parsimonious

model with two metabolites (M2) showed a more favourable balance between predictive accuracy, and robustness. Its lower VIF values together with consistent support from the statistic procedure applied (LASSO), highlight the importance of isoleucine and phenylalanine as reliable predictors. Regarding metabolic prediction, another prediction equation was reported using several plant growth regulators in triticale, such as zeatin or ABA (Žur et al. 2015), however, these phytohormones are difficult to analyze and quantify. This suggests the necessity of approaches based on metabolites which can be measured easily. In addition, the superior stability of our developed model M2 compared to model M1 suggests that simpler models may outperform more complex formulations when sample size is limited, reinforcing the value of parsimony in metabolomic prediction. This is confirmed by other studies where more simple models are preferred (Ahlmann-Eltze et al. 2025; Ramapiandra et al. 2023).

The validation analysis of our model allowed the differentiation of both contrasting lines by their green plant regeneration rate, which is crucial to confirm that the identified metabolic predictors are not artefacts of model overfitting but reflect biologically consistent patterns across independent datasets.

Overall, our results support the use of a reduced model as a biologically meaningful and statistically reliable approach for predicting the trait of interest.

Despite the high accuracy found in this study, further validation of the predictive model is needed, to exclude possible limitations in the prediction of the middle levels of green plant regeneration, potential bias in the data and to establish causality. This could be achieved by using a broader genetical base and more replicates in further studies, and by adding selected markers in the culture during the microspore embryogenesis process.

The model developed here was robust and allowed the prediction of the potential of green plant regeneration rate of microspores from DH lines before applying the stress pretreatment. This indicates that it is possible to select early the

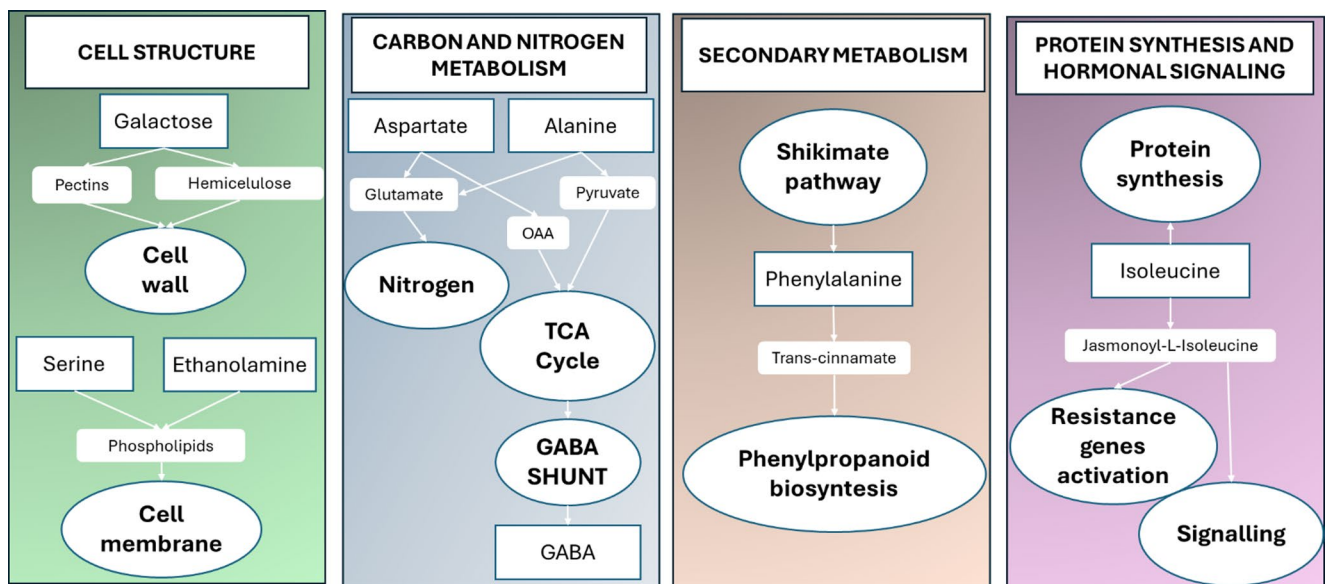


Fig. 7 Metabolic modules associated with regeneration competence. Metabolic pathways related to the biomarkers detected are represented in four modules. (1) Cell structure which includes galactose involved in cell wall formation, serine and ethanolamine involved in plasma membrane synthesis and remodeling. (2) Carbon and nitrogen

metabolism which includes aspartate, alanine and GABA, involving the TCA cycle, the GABA shunt and nitrogen metabolism. (3) Secondary metabolism which includes phenylalanine as a central point between primary and secondary metabolism. (4) Protein synthesis and signalling, via jasmonoyl-isoleucine

lines with high GPR rate. This knowledge would be highly desirable for both breeders and researchers: breeders could adapt their efforts in doubled haploid production using less input for highly responsive starting material (F1 or later segregating generations) and more input for less responsive starting material. On the other hand, researchers could focus on highly responsive lines or crosses for e.g. any proof of concept studies where DH lines are involved. From a biological point of view, it will be interesting to see if other species the same metabolites will be guiding the model.

Successful GPR is Associated with Coordinated Carbon, Nitrogen, Phenylpropanoid and Cell Wall Metabolism

The observed multicollinearity among the seven biomarkers in our first GPR model may reflect their involvement in shared metabolic pathways, as has been reported in metabolomics studies where co-varying metabolites cluster within common biochemical routes (Ash et al. 2019; Fukushima et al. 2011). Therefore, it allows to identify meaningful metabolic implication in regeneration. These pathways can be grouped into 4 modules (Fig. 7).

The first module is defined by cell structures generation, which connects galactose, serine, and ethanolamine. Galactose is essential for cell wall biosynthesis and assembly, cell expansion, differentiation and regeneration (Ma and Johnson 2023). Our findings may reflect different cell wall metabolism between the lines with high and low

regeneration efficiency. This would support the up-regulated genes coding for cell wall metabolism proteins found in *Brassica napus* in the study of Malik et al. (2008). On the other hand, serine and ethanolamine are components of the plasma membrane, being part of phospholipids. Salt stress has been reported to increase the proportion of phosphatidylethanolamine and phosphatidylserine in *Mesembryanthemum crystallinum* L (Guo et al. 2022). Another study has found a positive correlation between phosphatidylserine and plant salt tolerance in *Salicornia europaea* (Lv et al. 2021).

In sum, readiness in cell protective structures seems to be essential to overcome the stress pretreatment and regeneration successfully.

The second module involves carbon and nitrogen metabolism and the GABA shunt. Here, aspartate and alanine can enter the TCA cycle, providing resources for energy production. This correlates with our previous published results in *Triticum aestivum* cv. Svlena (Perez-Pinar et al., 2024). In addition, increased ATP synthase protein was found in responsive DH lines of triticale after cold stress treatment (Krzewska et al. 2017). This finding also correlates with the mobilization of carbohydrate reserve to meet the energy demanding process of reprogramming in triticale (Krzewska et al. 2021). Energy may be required for the active protein synthesis as reported in a responsive DH line of *Brassica napus* compared to a non-responsive cultivar (Malik et al. 2008). Therefore, further analysis of ATP metabolism would be of high interest. In addition, both alanine and aspartate are linked to nitrogen metabolism through glutamate production

and are necessary for nitrogen distribution during different stages of development (Li et al. 2025). Addition of aspartate has been reported to enhance heat tolerance in perennial ryegrass (Lei et al. 2022), confirming its importance as supporting metabolite during the stress pretreatment.

Moreover, the GABA shunt links with the TCA cycle via succinate. GABA has been found to enhance drought stress tolerance in *Brassica napus* as a resource of succinate to support mitochondrial metabolism (Palabıyık et al. 2024). This metabolite has different roles in *Arabidopsis*, such as ROS scavenging, maintaining the carbon and nitrogen balance and metabolic reprogramming (Xu et al. 2024).

In summary, our results in this second module suggest a carbon flux and energy production as key pathways related to better regeneration capacity.

The third module involves phenylalanine and the phenylpropanoid pathway. Phenylalanine links primary metabolism via the shikimate pathway, to phenylpropanoids. These compounds are well known to promote growth, cell wall remodeling and stress responses (Vogt 2010). This suggests that good responsive lines have enough phenylalanine to better overcome next steps of the process by redirecting carbon to the secondary metabolites needed, such as lignin, flavonoids, and others.

The last module involves isoleucine, which is an essential branched-chain amino acid required for protein synthesis during proliferation. It has a regulatory function in the context of regenerative competence, mediated through the Jasmonoyl-L-Isoleucine (JA-Ile) pathway (Wasternack and Hause 2013) (Fig. 7). In this pathway, JA-Ile initiates the Jasmonate signaling pathway, which has been associated with stress responses and defense mechanisms in plants (Kondo et al. 2021). In addition, the JA-Ile transduction pathway plays critical roles in somatic embryogenesis, noting a difference in regenerative capacities between mature seeds and highly regenerative immature embryos (Awere et al. 2024). Defects in cell proliferation and expansion in partial deficiency of isoleucine in *Arabidopsis* highlights its specific role in development (Yu et al. 2013).

In summary, successful regeneration in bread wheat relies on the readiness of structural, energetic and regulatory pathways, which ensures stress resistance and efficient cell proliferation. Taking together, these modules offer a framework of the variability in regeneration efficiency and also offer potential targets to enhance plant regenerative capacity.

Phenylalanine and Isoleucine as only Indicators of Embryogenic Potential

From the parsimonious predictive model, both phenylalanine and isoleucine emerged as robust unique predictors of competence.

Therefore, phenylalanine stands as a critical metabolic checkpoint. Its finding in both M1 and M2 predictive models strengthens its biological significance. Additionally, it arises as a robust biomarker for green plant regeneration in *Triticum aestivum*, being detected even when the model is simplified.

Conclusion

In this study, we assessed the metabolic differences in different time points during androgenesis between wheat DH lines with high and low green plant regeneration rates. Metabolites linked to cell protection, carbon and nitrogen metabolism, phenylpropanoids production and hormonal signaling, appear to provide the microspore with the resources needed for successful androgenesis and regeneration. For the first time, prediction of green plant regeneration capacity through microspore embryogenesis in wheat DH lines has been shown with only two metabolites.

Successful biotechnology and research applications demand an effective plant regeneration. Our modelling approach might help to improve the efficiency of anther culture, since it allows to selectively control the input of resources into DH production. However, further analysis must be done in order to continue clarifying the reason for the differences in metabolite content between the DH lines. Additionally, it will be necessary to complete the validation of the model presented here with more lines and even other species. Inclusion of other species will also help to elucidate the biological mechanisms underlying the plant regenerating capacity.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00344-026-12149-4>.

Acknowledgements This project was financially supported by the German Bundesministerium für Bildung und Forschung (FKZ 031A108A and FKZ 031A108B).

Author Contributions **TPP**: formal analysis, investigation, methodology, writing; **JEMA**: formal analysis; **AM**: formal analysis, supervision; **SB**: methodology; **HG**: Project administration, conceptualization, supervision; **HPM**: Project administration, conceptualization, supervision.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest Authors declare no competing interests in this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahlmann-Eltze C, Huber W, Anders S (2025) Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. *Nat Methods* 22:1657–1661. <https://doi.org/10.1038/s41592-025-02772-6>
- Arzani A, Mirodjagh SS (1999) Response of durum wheat cultivars to immature embryo culture, callus induction and *in-vitro* salt stress. *Plant Cell Tissue Organ Cult* 58:67–72. <https://doi.org/10.1023/A:1006309718575>
- Ash JR, Kuenemann MA, Rotroff D et al (2019) Cheminformatics approach to exploring and modeling trait-associated metabolite profiles. *J Cheminform* 11:43. <https://doi.org/10.1186/s13321-019-0366-3>
- Awere CO, Rakkammal K, Çolak AB, Bayrak M, Fredrick O, Anadebe VC, Ramesh M (2024) Enhanced and predictive modelling of direct shoot regeneration of *Evolvulus alsinoides* (L.) using ANN machine learning model and genetic stability studies. *Curr Plant Biol* 40:100423. <https://doi.org/10.1016/j.cpb.2024.100423>
- Cardoso LL, Freire FBS, Daloso DM (2023) Plant Metabolic Networks Under Stress: a Multi-species/Stress Condition Meta-analysis. *J Soil Sci Plant Nutr* 23:4–21. <https://doi.org/10.1007/s42729-022-01032-2>
- Carra A, Wijerathna-Yapa A, Pathirana R, Carimi F (2024) Development and Applications of Somatic Embryogenesis in Grapevine (*Vitis* spp). *Plants* 13:3131. <https://doi.org/10.3390/plants13223131>
- Cheng Y, Ma R, Jiao Y, Qiao N, Li T (2013) Impact of genotype, plant growth regulators and activated charcoal on embryogenesis induction in microspore culture of pepper (*Capsicum annuum* L.). *S Afr J Bot* 88:306–309. <https://doi.org/10.1016/j.sajb.2013.08.012>
- Cheng Z, Guimei G, Shuwei Z, He T, Feng S, Liu C, Wang Y, Zhou L (2025) Workflow for efficiency isolating microspore cultures of different rice genotypes by optimizing the callus induction medium. *Front Plant Sci* 16. <https://doi.org/10.3389/fpls.2025.1662463>
- De La Fuente GN, Frei UK, Lübberstedt T (2013) Accelerating plant breeding. *Trends Plant Sci* 18:667–672. <https://doi.org/10.1016/j.tplants.2013.09.001>
- Dubas E, Krzewska M, Surówka E, Kopeć P, Springer A, Janowiak F, Weigt D, Mikołajczyk SK, Telk A, Żur I (2024) New Prospects for Improving Microspore Embryogenesis Induction in Highly Recalcitrant Winter Wheat Lines. *Plants* 13(3):363. <https://doi.org/10.3390/plants13030363>
- Duran RE, Coskun Y, Demirci T (2013) Comparison of amino acids for their efficiency on regeneration in wheat embryo culture. *Asian J Plant Sci Res* 3:115–119
- Fernandez O, Urrutia M, Bernillon S et al (2016) Fortune telling: metabolic markers of plant performance. *Metabolomics* 12:158. <https://doi.org/10.1007/s11306-016-1099-1>
- Fomicheva M, Kulakov Y, Alyokhina K, Domblides E (2024) Spontaneous and Chemically Induced Genome Doubling and Polyploidization in Vegetable Crops. *Horticultrae* 10(6):551. <https://doi.org/10.3390/horticultrae10060551>
- Fukushima A, Kusano M, Redestig H et al (2011) Metabolomic correlation-network modules in *Arabidopsis* based on a graph-clustering approach. *BMC Syst Biol* 5:1. <https://doi.org/10.1186/1752-0509-5-1>
- Guo Q, Liu L, Rupasinghe TW, Roessner U, Barkla BJ (2022) Salt stress alters membrane lipid content and lipid biosynthesis pathways in the plasma membrane and tonoplast. *Plant Physiol* 189(2):805–826. <https://doi.org/10.1093/plphys/kiac123>
- Hashemi M, Moieni A, Sabet MS (2023) Improving the isolated microspore culture in eggplant (*Solanum melongena* L.) with amino acid nutrition. *PLoS ONE* 18(6):e0286809. <https://doi.org/10.1371/journal.pone.0286809>
- Hoekstra S, van Zijderveld MH, Louwerse JD, Heidekamp F, van der Mark F (1992) Anther and microspore culture of *Hordeum vulgare* L. cv. Igri. *Plant Sci* 86:89–96. [https://doi.org/10.1016/0168-9452\(92\)90182-L](https://doi.org/10.1016/0168-9452(92)90182-L)
- Hoque ME, Mansfield JW (2004) Effect of genotype and explant age on callus induction and subsequent plant regeneration from root-derived callus of indica rice genotypes. *Plant Cell Tissue Organ Cult* 78:217–223. <https://doi.org/10.1023/B:TICU.0000025640.75168.2d>
- Kambale R, Ramasamy G, Balasubramanian R, Thiruvengatasamy S, Sekar V (2023) Androgenesis in tomato (*Solanum lycopersicum* L.)-Effect of genotypes, microspore development stage, pre-treatments and media composition on induction of haploids. *J Phytol* 15:80–93. <https://doi.org/10.25081/jp.2023.v15.8366>
- Khalequzzaman M, Haq N, Hoque E, Aditya TL (2005) Regeneration efficiency and genotypic effect of 15 indica type Bangaldeshi rice (*Oryza sativa* L.) landraces. *Plant Tissue Cult Biotechnol* 15:33–42
- Kondo S, Tomiyama H, Opio P et al (2021) l-Isoleucine (Ile) Promotes Anthocyanin Accumulation in Apples. *J Plant Growth Regul* 40:541–549. <https://doi.org/10.1007/s00344-020-10117-0>
- Krzewska M, Gołębiowska-Pikania G, Dubas E, Gawin M, Żur I (2017) Identification of proteins related to microspore embryogenesis responsiveness in anther cultures of winter triticale (*×Triticosecale* Wittm). *Euphytica* 213:192. <https://doi.org/10.1007/s10681-017-1978-1>
- Krzewska M, Dubas E, Gołębiowska G et al (2021) Comparative proteomic analysis provides new insights into regulation of microspore embryogenesis induction in winter triticale (*×Triticosecale* Wittm.) after 5-azacytidine treatment. *Sci Rep* 11:22215. <https://doi.org/10.1038/s41598-021-01671-y>
- Lantos C, Weyen J, Orsini JM et al (2013) Efficient application of *in-vitro* anther culture for different European winter wheat (*Triticum aestivum* L.) breeding programmes. *Plant Breed* 132:149–154. <https://doi.org/10.1111/pbr.12032>
- Lei S, Rossi S, Huang B (2022) Metabolic and Physiological Regulation of Aspartic Acid-Mediated Enhancement of Heat Stress Tolerance in Perennial Ryegrass. *Plants* 11(2):199. <https://doi.org/10.3390/plants11020199>
- Li S, Ahmed W, Jiang T et al (2025) Amino acid metabolism pathways as key regulators of nitrogen distribution in tobacco: insights from transcriptome and WGCNA analyses. *BMC Plant Biol* 25:393. <https://doi.org/10.1186/s12870-025-06390-4>
- Lv S, Tai F, Guo J, Jiang P, Lin K, Wang D, Zhang X, Li Y (2021) Phosphatidylserine Synthase from *Salicornia europaea* Is Involved in Plant Salt Tolerance by Regulating Plasma Membrane Stability.

- Plant Cell Physiol 62(1):66–79. <https://doi.org/10.1093/pcp/pca141>
- Ma Y, Johnson K (2023) Arabinogalactan proteins - Multifunctional glycoproteins of the plant cell wall. Cell Surf 17(9):100102. <https://doi.org/10.1016/j.tcs.2023.100102>
- Makowska K, Oleszczuk S (2014) Albinism in barley androgenesis. Plant Cell Rep 33:385–392. <https://doi.org/10.1007/s00299-013-1543-x>
- Malik MR, Wang F, Dirpaul JM, Zhou N, Polowick PL, Ferrie AMR, Krochko JE (2007) Transcript profiling and identification of molecular markers for early microspore embryogenesis in *Brassica napus*. Plant Physiol 144:134–154. <https://doi.org/10.1104/p.106.092932>
- Malik MR, Wang F, Dirpaul JM, Zhou N, Hammerlindl J, Keller W, Abrams SR, Ferrie AMR, Krochko JE (2008) Isolation of an embryogenic line from non-embryogenic *Brassica napus* cv. Westar through microspore embryogenesis. J Exp Bot 59:2857–2873. <https://doi.org/10.1093/jxb/ern149>
- Mano Y, Takahashi H, Sato K, Takeda K (1996) Mapping genes for callus growth and shoot regeneration in barley (*Hordeum vulgare* L.). Breed Sci 46:137–142. <https://doi.org/10.1270/jsbbs1951.46.137>
- Maraschin SF, de Priester W, Spaik HP, Wang M (2005) Androgenic switch: an example of plant embryogenesis from the male gametophyte perspective. J Exp Bot 56:1711–1726. <https://doi.org/10.1093/jxb/eri190>
- Mashabela MD, Masamba P, Kappo AP (2023) Applications of Metabolomics for the Elucidation of Abiotic Stress Tolerance in Plants: A Special Focus on Osmotic Stress and Heavy Metal Toxicity. Plants 12(2):269. <https://doi.org/10.3390/plants12020269>
- Mishra VK, Goswami R (2014) Haploid production in higher Plant. Int J ChemBiolSci 1:25–45
- Molina MD, Aponte EM, Cortina H, Moreno G (2002) The effect of genotype and explant age on somatic embryogenesis of coffee. Plant Cell Tissue Organ Cult 71:117–123. <https://doi.org/10.1023/A:1019965621041>
- Murovec J, Bohanec B (2012) Haploids and Doubled Haploids in Plant Breeding. In: *Plant Breeding*. Dr. Ibrokhim Abdurakhmonov (Ed.), InTech: 87–107
- Nam HM, Bang E, Kwon YT, Kim Y, Kim HE, Cho K et al (2015) Metabolite profiling of diverse rice germplasm and identification of conserved metabolic markers of rice roots in response to long-term mild salinity stress. Int JMol Sci 16:21959–21974. <https://doi.org/10.3390/ijms160921959>
- Otse EJ, Obunadike GN, Abubakar A (2025) Linear Regression Approach to Solving Multicollinearity and Overfitting in Predictive Analysis. J Sci Res Rev 2(1):108–117. <https://doi.org/10.7082/josrar.2025.v2i1.35>
- Palabiyik Ş, Çetinkaya İ, Öztürk TA et al (2024) Flagellin Induced GABA-shunt improves Drought stress tolerance in *Brassica napus* L. BMC Plant Biol 24:864. <https://doi.org/10.1186/s12870-024-05503-9>
- Perez-Piñar T, Hartmann A, Bössow S, Gnad H, Mock HP (2024) Metabolic changes during wheat microspore embryogenesis induction using the highly responsive cultivar Svilen. J Plant Physiol 294:154193. <https://doi.org/10.1016/j.jplph.2024.154193>
- Quistián-Martínez D, Estrada-Luna AA, Altamirano-Hernández J, Peña-Cabriaes JJ, Oca-Luna RM, Cabrera-Ponce JL (2011) Use of trehalose metabolism as a biochemical marker in rice breeding. Mol Breed 30:469–477. <https://doi.org/10.1007/s11032-011-9636-0>
- R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Ramampandra EC, Scheidegger A, Wydler J, Schuwirth N (2023) A comparison of machine learning and statistical species distribution models: Quantifying overfitting supports model interpretation. Ecol Model 481:110353. <https://doi.org/10.1016/j.ecolmodel.2023.110353>
- Raza G, Singh MB, Bhalla PL (2019) Somatic Embryogenesis and Plant Regeneration from Commercial Soybean Cultivars. Plants 9(1):38. <https://doi.org/10.3390/plants9010038>
- Rubtsova M, Gnad H, Melzer M, Weyen J, Gils M (2013) The auxins centrophenoxine and 2,4-D differ in their effects on non-directly induced chromosome doubling in anther culture of wheat (*T. aestivum* L.). Plant Biotechnol Rep 7:247–255. <https://doi.org/10.1007/s11816-012-0256-x>
- Saharan V, Yadav RC, Yadav RN, Chapagain BP (2004) High frequency plant regeneration from desiccated calli of indica rice (*Oryza sativa* L.). Afr J Biotechnol 3:256–259. <https://doi.org/10.5897/AJB2004.000-2047>
- Sánchez-Díaz RA, Castillo AM, Vallés MP (2013) Microspore embryogenesis in wheat: new marker genes for early, middle and late stages of embryo development. Plant Reprod 26:287–296. <https://doi.org/10.1007/s00497-013-0225-8>
- Schlegel R, Belchev I, Kostov K, Atanassova M (2000) Inheritance of high anther culture response in hexaploid wheat, *Triticum aestivum* L. var. Svilen. Bulg J Agric Sci 6(3):261–270
- Seifert F, Bössow S, Kümlehn J, Gnad H, Scholten S (2016) Analysis of wheat microspore embryogenesis induction by transcriptome and small RNA sequencing using the highly responsive cultivar Svilen. BMC Plant Biol 16(1):1–16. <https://doi.org/10.1186/s12870-016-0782-8>
- Takahashi D et al (2024) Structural changes in cell wall pectic polymers contribute to freezing tolerance induced by cold acclimation in plants. Curr Biol 34(5):958–968. <https://doi.org/10.1016/j.cub.2024.01.045>
- Testillano P (2019) Microspore embryogenesis: targeting the determinant factors of stress-induced cell reprogramming for crop improvement. J Exp Bot 70(11):2965–2978. <https://doi.org/10.1093/jxb/ery464>
- Valero-Rubira I, Castillo AM, Burrell MÁ, Vallés MP (2023) Microspore embryogenesis induction by mannitol and TSA results in a complex regulation of epigenetic dynamics and gene expression in bread wheat. Front Plant Sci 13:1058421. <https://doi.org/10.3389/fpls.2022.1058421>
- Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3(1):2–20. <https://doi.org/10.1093/mp/ssp106>
- Wang C, Zhang P, He Y et al (2023) Exogenous spraying of IAA improved the efficiency of microspore embryogenesis in Wucai (*Brassica campestris* L.) by affecting the balance of endogenous hormones, energy metabolism, and cell wall degradation. BMC Genomics 24:380. <https://doi.org/10.1186/s12864-023-09483-2>
- Wasternack C, Hause B (2013) Jasmonates: Biosynthesis, perception, signal transduction and action in plant stress response, growth and development. Ann Bot 111(6):1021–1058. <https://doi.org/10.1093/aob/mct067>
- Xu B, Feng X, Piechatzek A, Zhang S, Konrad KR, Kromdijk J, Hedrich R, Gilliam M (2024) The GABA shunt contributes to ROS homeostasis in guard cells of Arabidopsis. New Phytol 241(1):73–81. <https://doi.org/10.1111/nph.19390>
- Yang F, Liu X, Qiao Y, Tang X, Luo P (2024) Microspore embryogenesis: *in-vitro* cultivation induced cell reprogramming for plant breeding. Seed Biol 3:e021. <https://doi.org/10.48130/seedbio-0024-0019>
- Yu H, Zhang F, Wang G, Liu Y, Liu D (2013) Partial deficiency of isoleucine impairs root development and alters transcript levels of the genes involved in branched-chain amino acid and glucosinolate metabolism in *Arabidopsis*. J Exp Bot 64(2):599–612. <https://doi.org/10.1093/jxb/ers352>
- Zale JM, Borchardt-Wier H, Kidwell KK, Steber CM (2004) Callus induction and plant regeneration from mature embryos of

- a diverse set of wheat genotypes. *Plant Cell Tissue Organ Cult* 76:277–281. <https://doi.org/10.1023/B:TICU.0000009248.32457.4c>
- Zheng M, Weng Y, Liu W, Konzak C (2002) The effect of ovary-conditioned medium on microspore embryogenesis in common wheat (*Triticum aestivum* L.). *Plant Cell Rep* 20:802–807. <https://doi.org/10.1007/s00299-001-0411-2>
- Zheng MY, Liu W, Weng Y, Polle E, Konzak CF (2003) Production of doubled haploids in wheat (*Triticum aestivum* L.) through microspore embryogenesis triggered by inducer chemicals. In *Doubled Haploid Production in Crop Plants*. Maluszynski M, Kasha K J, Forster B P, Szarejko I (eds). Springer: 83–94
- Żur I, Dubas E, Golemieć E, Szechyńska-Hebda M, Gołebiowska G, Wedzony M (2009) Stress-related variation in antioxidative enzymes activity and cell metabolism efficiency associated with embryogenesis induction in isolated microspore culture of triticale ($\times$ *Triticosecale* Wittm.). *Plant Cell Rep* 28:1279–1287. <https://doi.org/10.1007/s00299-009-0730-2>
- Żur I, Dubas E, Krzewska M, Waligórski P, Dziurka M, Janowiak F (2015) Hormonal requirements for effective induction of microspore embryogenesis in triticale ($\times$ *Triticosecale* Wittm.) anther cultures. *Plant Cell Rep* 34:47–62. <https://doi.org/10.1007/s00299-014-1686-4>
- Żur I, Adamus A, Cegielska-Taras T et al (2022) Doubled Haploids: Contributions of Poland's Academies in Recognizing the Mechanism of Gametophyte Cell Reprogramming and Their Utilization in Breeding of Agricultural and Vegetable Species. *Acta Soc Bot Pol* 91(1). <https://doi.org/10.5586/asbp.9128>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.