Plant Cell, Tissue and Organ Culture (PCTOC) (2026) 165:89
https://doi.org/10.1007/s11240-026-03509-9

ORIGINAL ARTICLE

Integrating machine learning and SHAP analysis to boost callus growth and rutin biosynthesis in *Capparis spinosa* L.

Marouane Mohaddab[1,2] · Younes EL Goumi[3] · Mohammed Elakrouch[4] · Soufiane Hasni[5] · Clément Burgeon[1] · Manon Genva[1] · Malika Fakiri[2] · Marie-Laure Fauconnier[1]

Received: 19 January 2026 / Accepted: 8 June 2026
© The Author(s), under exclusive licence to Springer Nature B.V. 2026

Abstract

Capparis spinosa L. is a Mediterranean medicinal species of high economic value, yet its large-scale propagation and metabolite production remain constrained by conventional approaches. This study evaluated the effects of four hormonal combinations, 6-benzylaminopurine (BAP) + 1-naphthaleneacetic acid (NAA), BAP + 2,4-dichlorophenoxyacetic acid (2,4-D), kinetin (KIN) + 2,4-D, and KIN + NAA, on callus fresh weight gain (FWG) from leaf explants using a pairwise combinatorial experimental design. The resulting dataset was analyzed using three machine learning (ML) algorithms, Random Forest (RF), Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost), as well as a second-degree polynomial regression model (PL2). Among these models, RF achieved the best performance, with a cross-validation coefficient of determination (R^2CV) of 0.89, a coefficient of determination (R^2) of 0.86, and a root mean square error (RMSE) of 263.33. SHapley Additive exPlanations (SHAP) analysis revealed 2,4-D as the most influential predictor of callogenesis, followed by a positive contribution from BAP, while KIN and NAA had minimal or negative effects on FWG prediction. Experimental validation was restricted to the five highest-ranked RF-predicted combinations, demonstrating good agreement between predicted and observed values. In addition, rutin, the major bioactive flavonoid of *C. spinosa*, was identified by LC-QTOF-MS/MS and relatively quantified by LC-TQ-MS/MS under BAP and 2,4-D combinations. A stacked RF model integrating FWG predictions was further developed to estimate rutin accumulation, achieving satisfactory performance with an R^2 of 0.82, an R^2CV of 0.84, and an RMSE of 0.41. Maximum accumulation was observed at moderate hormone concentrations. Overall, this integrative ML and SHAP-based approach provides an interpretable and scalable strategy for optimizing callus culture and enhancing metabolite production, offering a sustainable alternative to wild plant harvesting.

Keywords *Capparis spinosa* L. · Plant tissue culture · Callogenesis · Machine learning · Rutin

Communicated by Melekşen Akın

✉ Marouane Mohaddab
mmohaddab@uliege.be

Younes EL Goumi
y.elgoumi@usms.ma

Mohammed Elakrouch
Mohammed.Elakrouch@uliege.be

Soufiane Hasni
Soufiane.hasni@etu.uae.ac.ma

Clément Burgeon
cburgeon@uliege.be

Manon Genva
M.Genva@uliege.be

Malika Fakiri
malika.fakiri@uhp.ac.ma

Marie-Laure Fauconnier
marie-laure.fauconnier@uliege.be

1 Laboratory of Chemistry of Natural Molecules, Gembloux Agro-Bio Tech, University of Liege, Gembloux 5030, Belgium

2 Laboratory of Agrifood and Health, Faculty of Sciences and Techniques, Hassan First University of Settat, BP 577, Settat 26000, Morocco

3 Polyvalent Team in R&D, Higher School of Technology of Fkih Ben Salah, Sultan Moulay Slimane University, Beni Mellal, Morocco

4 Laboratory of Integrated and Urban Plant Pathology (LIUPP), Gembloux Agro-Bio Tech, University of Liege, Gembloux 5030, Belgium

5 Biology Department, Faculty of Sciences and Technology of Tangier, Abdelmalek Essaadi University, Tangier, Morocco

Published online: 15 June 2026

Introduction

Capparis spinosa L., a native species widely distributed across arid and semi-arid Mediterranean regions, has attracted increasing scientific attention due to its medicinal, nutritional, and agroecological values. The species is extensively harvested and commercially exploited across several Mediterranean countries, including Morocco, Spain, Italy, Greece, and Turkey (Sozzi and Vicente 2006; Grimalt et al. 2018, 2022; Hazrati et al. 2025). *C. spinosa* is characterized by a rich phytochemical profile, encompassing glucosinolates, phenolic acids, alkaloids, and flavonoids, which are associated with a wide range of biological activities such as antioxidants, antidiabetic, anti-inflammatory, and anticancer effects (Moghadamnia et al. 2019; Esmaeilpour et al. 2020; Rahimi et al. 2020). These properties highlight its considerable potential for applications in the pharmaceutical, nutraceutical, and cosmetic industries (Karous et al. 2021; Kirkan et al. 2021; Annaz et al. 2022). Among these compounds, rutin is the most abundant molecule in the leaves of *C. spinosa*, with a content of approximately 0.61 mg/g, making it a key target for biotechnological production strategies (Ramezani et al. 2008). Its biosynthesis is integrated into the phenylpropanoid and flavonoid pathways derived from phenylalanine. The main enzymatic steps involve phenylalanine ammonia-lyase (PAL), 4-coumarate-CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone-3-hydroxylase (F3H), flavonoid-3'-hydroxylase (F3'H), and flavonol synthase (FLS). The pathway proceeds through successive glycosylation reactions catalyzed by flavonoid-3-*O*-glucosyltransferase (UFGT) and flavonol-3-*O*-glucoside L-rhamnosyltransferase (RT), the latter representing the committed final step in the biosynthesis of rutin (Kianersi et al. 2020). Furthermore, plant growth regulators (PGRs) may modulate the expression of genes involved in key metabolic pathways, thereby inducing significant variations in the accumulation of secondary metabolites (Kumari et al. 2024; Li et al. 2025).

The exploitation of this plant supports the livelihoods of rural and marginalized communities while contributing to local economies (Sozzi et al. 2012). However, commercial harvesting relies mainly on wild populations, increasing pressure on natural resources and threatening long-term sustainability (Zhang and Ma 2018). Due to its tolerance to water deficit and adverse soils, *C. spinosa* also represents an ideal candidate for climate-resilient agriculture (Yousefi et al. 2025). Despite these advantages, traditional propagation methods, whether sexual or vegetative, face significant limitations, including low efficiency, slow growth, and the potential loss of desirable traits (Gristina et al. 2014). Pretreatments such as scarification, sulfuric acid exposure, or hormone application partially improve germination and rooting but remain insufficient for rapid large-scale multiplication (Labbafi et al. 2018; Nowruzian and Aalami 2023).

In vitro culture techniques, particularly callus induction, offer an effective solution (Efferth 2019). They enable the regeneration of uniform, pathogen-free plants (Nimavat and Parikh 2024). They can constitute an effective system for enhancing the accumulation of secondary metabolites, thereby contributing to the conservation of natural populations and meeting the growing demand for phytotherapeutic resources (Yin et al. 2014; Mohaddab et al. 2022). Although research on *C. spinosa* remains limited (Caglar et al. 2005; Movafeghi et al. 2008; Al-Safadi and Elias 2011; Carra et al. 2012; Fahmideh et al. 2019; Gianguzzi et al. 2019; Mohaddab et al. 2024), callus tissue, being totipotent, serves as an ideal platform for targeted secondary metabolite production. Hagaggi et al. (2024) reported that callus extracts contain higher levels of phenolic and flavonoid compounds, a greater diversity of volatile metabolites, and superior antioxidant and antibacterial activities compared to leaf extracts. Thus, strict control of culture conditions enhances the accumulation of bioactive flavonoids, offering direct potential for industrial and pharmaceutical applications (Wijerathna-Yapa et al. 2025).

Callus formation and proliferation depend on intrinsic factors, including explant type, age, and medium composition, as well as extrinsic factors such as light, temperature, and humidity (Pasternak and Steinmacher 2024). PGRs, particularly auxins and cytokinins, are essential for morphogenesis, biomass accumulation, and secondary metabolite production (Ikeuchi et al. 2013; Bravo-Vázquez et al. 2023). Commonly used phytohormones include 2,4-D and NAA, applied alone or in combination with cytokinins such as BAP or KIN (Koufan et al. 2022; Higazy et al. 2023). Studies indicate that 2,4-D at 1.0 mg/L combined with BAP at 1.5 mg/L effectively induces callus from *C. spinosa* leaf explants (Wang et al. 2007; Liu et al. 2011; Yin et al. 2014). Strategic manipulation of PGRs during callogenesis can therefore enhance both callus regeneration and flavonoid production, particularly rutin (Sobhy et al. 2025).

In vitro plant development is a complex process governed by multiple interconnected variables (Ma et al. 2018; Svolacchia and Sabatini 2023). Studying these factors in isolation is insufficient, as their interactions strongly influence outcomes. Classical statistical approaches are limited when handling nonlinear and multidimensional datasets and often exhibit high error rates (Aasim et al. 2023; Katırcı et al. 2024). Mechanistic models, although grounded in biological knowledge, require precise measurement of numerous parameters that are difficult to standardize experimentally (Noordijk et al. 2024), reducing their predictive power and practical usefulness for rapid protocol optimization (Baker et al. 2018). Artificial intelligence (AI), especially

supervised learning or ML, offers a means to manage this complexity (Peng and Rajjou 2024). ML applications to callogenesis and organogenesis represent an emerging and promising field (Hesami and Jones 2021). Algorithms such as RF, GB, and XGBoost have been successfully applied to predict callus induction, shoot regeneration, biomass yield, and secondary metabolite production in various species (Hesami and Jones 2021; Alcalde et al. 2022; Eren et al. 2023; Demirel et al. 2023; Sarabandi et al. 2024; Bozkurt et al. 2024; Tan et al. 2024; Zarbakhsh et al. 2024; Koçak et al. 2025; Sarmah et al. 2025). Despite their predictive performance, these models often remain "black boxes," where relationships between biological variables and observed responses are opaque (Petch et al. 2022). SHAP values provide a solution to this opacity. Derived from game theory, they quantify the contribution of each variable to the final prediction by considering all possible factor combinations (Lundberg and Lee 2017; Lundberg et al. 2020). By transforming a black-box model into a "white-box" model, SHAP values allow identification of key variables and their interactions, offering biologically meaningful interpretation (Ekanayake et al. 2022).

Although ML methods have been progressively integrated into plant tissue culture research across various species (Hesami and Jones 2021; Sarabandi et al. 2024; Bozkurt et al. 2024), their application to *C. spinosa* remains unexplored. Moreover, the use of SHAP-based interpretability frameworks to analyze key hormonal regulators involved in both callus induction and secondary metabolite biosynthesis has not yet been reported for this plant. To our knowledge, this study is the first to integrate ML to simultaneously model callus growth and rutin accumulation in *C. spinosa*, thereby proposing a novel and reproducible framework for the biotechnological optimization of this medicinal plant.

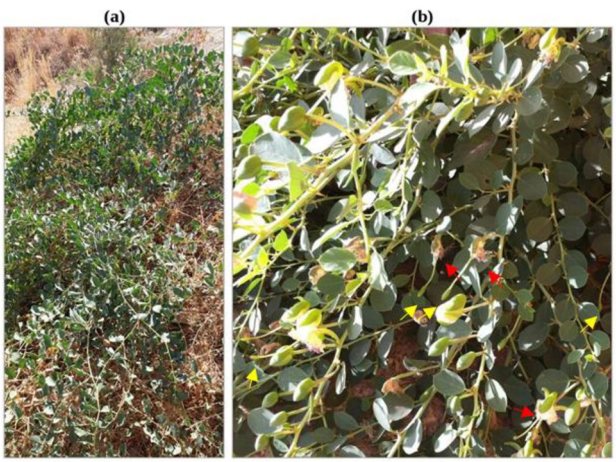

Fig. 1 Wild *Capparis spinosa* L. from Taounate, Morocco. (**a**) Natural habitat. (**b**) Mature fruiting plant showing explant source leaves. Red arrow: pollinated flowers. Yellow arrow: fruits

This study examined the effects of growth regulators, individually and in combination, on callus induction and proliferation, as well as on rutin accumulation. The collected data were evaluated using ML models, complemented by SHAP analyses, to quantify the relative influence of each regulator and identify the key factors modulating both callus development and rutin accumulation.

Materials and methods

Collection of plant material

Fresh leaves of *C. spinosa* L. were collected in April 2022 from mature wild plants at the fruiting stage in the Taounate region (Morocco; latitude: 34°26'31.187" N, longitude: 4°42'11.97" W, altitude: 435 m). The plant material was identified and authenticated based on morphological characteristics by the National Agency for Medicinal and Aromatic Plants in Taounate, Morocco. The leaves were immediately processed on the same day for subsequent experiments (Fig. 1).

Surface sterilization

Surface sterilization of the harvested leaves was performed before culture establishment. The procedure followed the protocol described by Slimani et al. (2021), with minor modifications, to ensure the absence of microbial contamination prior to inoculation onto the culture medium. The leaves were first washed under running tap water for 15 min, then immersed in 70% ethanol for 30 s. Surface disinfection was subsequently carried out using a 2.5% (w/v) sodium hypochlorite (NaOCl) solution supplemented with two drops of Tween-20 per 100 mL, under continuous agitation for 15 min. The explants were then rinsed three times with sterile distilled water (5 min each) under laminar airflow conditions. All subsequent procedures were performed under a laminar flow hood to prevent microbial contamination.

Culture medium preparation

Callus induction was carried out on Murashige and Skoog (MS) medium (Murashige and Skoog 1962), supplemented with 0.8% agar, 3% sucrose, and MES buffer (M0255, Duchefa). The leaf explants, approximately 1 cm in diameter with an average weight of 32.90 ± 1.42 mg, were placed abaxial side down on 20 mL of medium in each Petri dish, with five explants per dish. A total of 795 explants were used in the experiment. The pH of the medium was adjusted to 5.7 ± 0.1 before autoclaving at 121 °C for 20 min under 1 bar of pressure.

Plant growth regulator combination design

The basal medium was supplemented with various concentrations of exogenous PGRs to evaluate their effects on callus induction. The PGRs used included two cytokinins, BAP and KIN, as well as two auxins, 2,4-D and NAA. A combinatorial experimental design was applied using four combinations: BAP+NAA, BAP+2,4-D, KIN+NAA, and KIN+2,4-D. For each combination, four concentrations (0.0, 0.5, 1.0, and 1.5 mg/L) were tested for each regulator, generating 16 concentration pairs per combination, resulting in a total of 64 treatments prior to removal of duplicates. In addition, each plant growth regulator was evaluated individually at a concentration of 2.0 mg/L. All treatments were conducted in three independent replicates, bringing the total number of experimental units to 159.

In vitro culture conditions

All cultures were maintained in a phytotron under a 16-hour light/8-hour dark photoperiod at 24 ± 2 °C for 30 days. Illumination was provided by Sylvania fluorescent tubes delivering a light intensity of 40 µmol m^{-2} s^{-1}.

Measurement of callus growth

After 30 days, FWG was determined by weighing the callus from each Petri dish immediately after removal from the culture medium. The resulting data were subsequently compiled into a dataset for ML analysis.

FWG was calculated as:

$$\text{FWG (mg)} = \text{Final fresh weight} - \text{Initial weight of explant}$$

Identification and relative quantification of rutin in callus cultures using LC–MS/MS

The identification and relative quantification of rutin were performed using LC-QTOF for qualitative structural elucidation coupled with LC-TQ-MS/MS in negative ionization mode based on the internal standard method and Multiple Reaction Monitoring (MRM) mode. Approximately 150 mg of oven-dried callus material, dried at 30 °C to constant weight, was extracted with 2 mL of methanol-water (80:20, v/v). The extracts were spiked with 5 µL of an internal standard solution of flavone (molecular weight = 220 g/mol) dissolved in DMSO at a concentration of 2 mg/mL. Samples were sonicated in an ultrasonic bath for 20 min and centrifuged at 3000 × g for 10 min. The resulting supernatants were filtered through 0.22 µm syringe filters prior to injection.

Structural identification of rutin was carried out using an Agilent 1290 Infinity II LC system coupled to an Agilent 6530 quadrupole time-of-flight (Q-TOF) mass spectrometer (Agilent Technologies, Santa Clara, CA, USA), equipped with an electrospray ionization (ESI) source operating in negative ionization mode. Chromatographic separation was achieved on an Acquity® BEH C18 column (2.1 × 100 mm, 1.8 µm particle size, Waters Co., Milford, MA, USA) with an injection volume of 10 µL. A 16-minute gradient elution was applied according to Belkessam et al. (2025), at a flow rate of 0.450 mL/min. The mobile phase consisted of solvent A (Milli-Q water with 0.1% formic acid) and solvent B (MS-grade acetonitrile with 0.1% formic acid). The gradient started at 1% B for 1 min, increased linearly to 100% B over 11 min, was maintained isocratically at 100% B for 1.5 min, then returned to 1% B over 2.5 min for column re-equilibration. Column and autosampler temperatures were set at 30 °C and 15 °C, respectively. Full scan MS1 data were acquired over an m/z range of 100–1000, with a mass accuracy tolerance below 16 ppm. MS/MS fragmentation was performed at a collision energy of 35 eV, isolation width of 1.3 m/z, and fragment mass range of 100–650 m/z. Rutin identification was confirmed by comparison with MassBank and PubChem databases and published literature.

For relative quantification, MRM transitions were recorded using an Agilent 1290 Infinity II LC system coupled to an Agilent 6475 triple quadrupole mass spectrometer (Agilent Technologies, Santa Clara, CA, USA), under identical chromatographic conditions. Ion source parameters were as follows: drying gas temperature 300 °C, drying gas flow 13 L/min, nebulizer pressure 35 psi, sheath gas temperature 250 °C, sheath gas flow 11 L/min, and capillary voltage 4000 V. MRM transitions were acquired in negative ionization mode with an acquisition time of 5 ms, fragmentor voltage of 150 V, and collision energy of 30 V. Rutin was monitored using the transition 609 → 300. Relative quantification was calculated as follows:

$$\text{Rutin content (A.U./}mg\text{ DW)} = \frac{\text{MRM peak area of rutin}}{\text{MRM flavone peak area x Sample dry weight}}$$

where A.U. = Arbitrary Units and DW = Dry Weight. All data were acquired and processed using MassHunter® Workstation software (version 10.0, Agilent Technologies, Santa Clara, CA, USA).

Modeling procedures

Callogenesis was induced using four PGR combinations: BAP/NAA, BAP/2,4-D, KIN/NAA, and KIN/2,4-D at varying concentrations to evaluate their effects on FWG and

rutin accumulation. To ensure data quality, the data were normalized, outliers were identified and removed using the *Z*-score method, and the relationships between parameters were examined using a Pearson correlation test. The final dataset was partitioned into a training set (75%) and an independent test set (25%), while maintaining a balanced distribution of experimental conditions (Fig. 2A).

Four ML models, RF, GB, XGBoost, and PL2, were evaluated to predict FWG. Model performance was primarily evaluated using k-fold cross-validation, based on the R^2CV. This approach optimizes the efficient use of available data by simultaneously utilizing training and validation subsets, while providing a robust estimate of the models' generalization ability. In this context, cross-validation serves to reduce the risk of overfitting and ensures that performance metrics reliably reflect the models' true predictive ability under varying experimental conditions. The hyperparameters used for each model were defined as follows: the RF model was configured with n_estimators = 100, max_depth = 10, and random_state = 42; the GB model with n_estimators = 100, max_depth = 3, learning_rate = 0.1, and random_state = 42; the XGBoost model with n_estimators = 100, max_depth = 3, learning_rate = 0.1, subsample = 0.8, colsample_bytree = 0.8, and random_state = 42; and the polynomial regression model with a degree of 2. Following the training phase, the models were evaluated on an independent test set to verify their predictive performance. Statistical metrics were used to quantify prediction accuracy, including the R^2 and the RMSE, complemented by residual analysis and evaluation of their distribution. Final model selection was performed using Friedman and Wilcoxon signed-rank tests with Bonferroni correction.

SHAP analysis was applied as a model interpretation framework to quantify the individual contribution of each

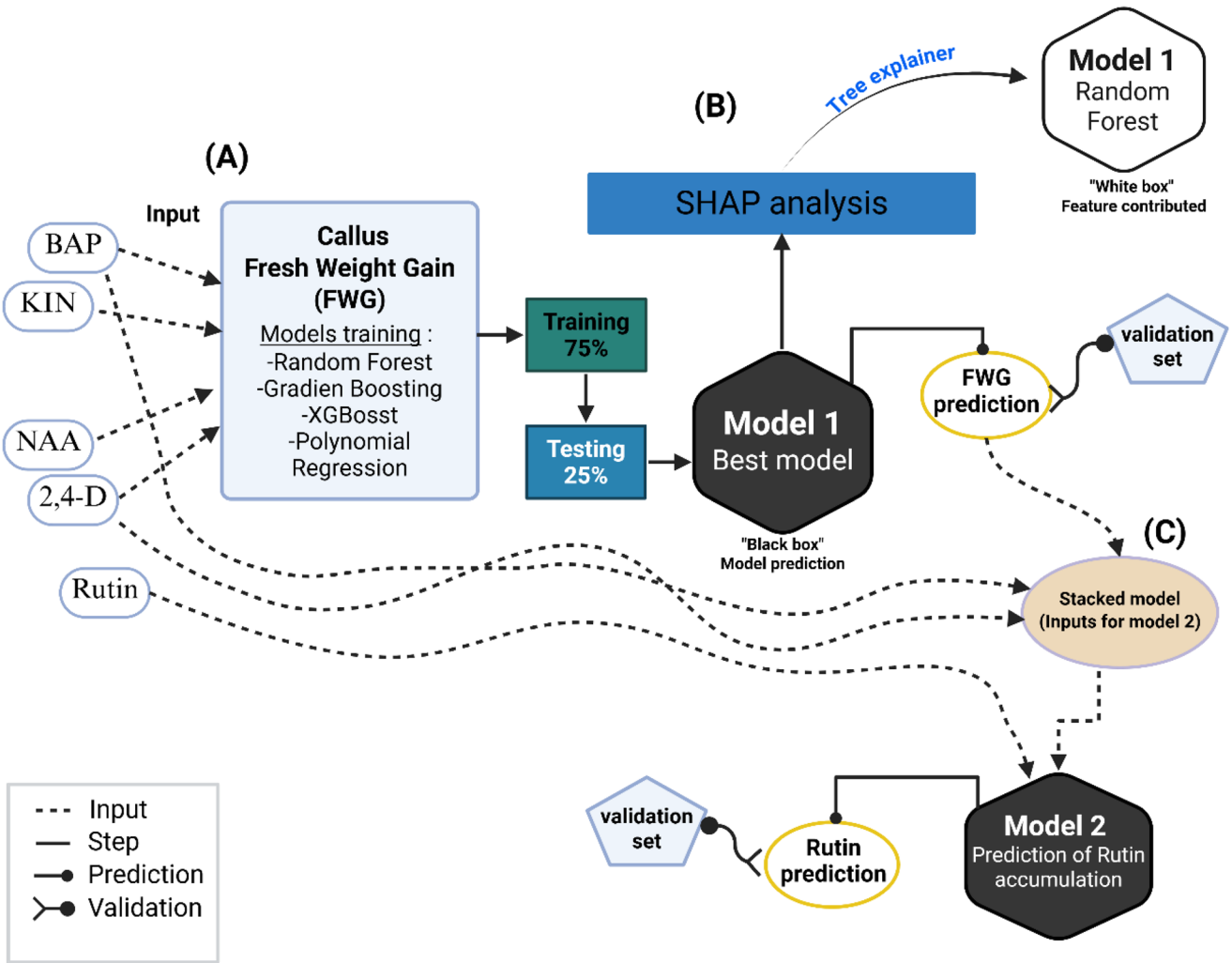

Fig. 2 A schematic representation of the stepwise computational approach used in this study. (**A**) Modeling and prediction of callus fresh weight gain (FWG) (output) based on 6-benzylaminopurine (BAP), kinetin (KIN), 1-naphthaleneacetic acid (NAA), and 2,4-dichlorophenoxyacetic acid (2,4-D) as inputs using four models. (**B**) SHAP value analysis of the best-selected model to identify and quantify the key model features. (**C**) Stacking of the best model and its integration as input for the accumulation of rutin (Model 2)

PGR to FWG predictions generated by the best-performing ML model (Fig. 2B). This approach decomposes each model prediction into additive contributions from the different features, while accounting for all possible interactions between variables. It provides both local explanations for individual predictions and a global perspective on feature importance across the entire dataset (Lundberg and Lee 2017; Lundberg et al. 2020).

Predictions from model 1 were then incorporated into a stacked model designed to predict rutin accumulation (model 2) (Fig. 2C). This model used both the initial PGR concentrations and the FWG predictions as input variables, directly linking callus growth to secondary metabolite accumulation. To evaluate potential error propagation inherent to stacked modeling architectures, model 2 was trained and assessed using experimentally measured FWG values and RF-predicted FWG values. To validate the efficiency and reliability of this optimized model, the predicted conditions for maximizing FWG and rutin accumulation were independently tested in triplicate, following the same protocols used in the initial experiment.

Statistical analysis

Statistical analysis was performed on experimental data using one-way analysis of variance (ANOVA). When significant differences were detected, pairwise comparisons were conducted using Tukey's Honestly Significant Difference (HSD) post-hoc test, with $p \leq 0.05$ as the significance level. Data are presented as mean ± standard deviation (SD), based on three independent replicates per treatment.

ML model performance was evaluated using k-fold cross-validation (k = 5), based on the R^2CV and RMSE. To formally compare predictive performance across models, the Friedman test was applied, followed by pairwise Wilcoxon signed-rank tests with Bonferroni correction for multiple comparisons. Residual and distribution analysis was performed to assess model reliability. Pearson correlation analysis was conducted to examine relationships between input variables prior to model training. All statistical and ml computations were performed using Python (version 3.12).

Results

Influence of plant growth regulators on *C. spinosa* callus induction

This study investigated the influence of phytohormones combinations at different concentrations (0, 0.5, 1 and 1.5 mg/L) and hormonal combinations BAP/NAA, BAP/2,4-D, KIN/NAA, and KIN/2,4-D on the FWG of callus derived from *C. spinosa* leaf explants after 30 days of culture (Fig. 3; Supplementary Table A). As expected, no callus formation was observed in hormone-free control medium. The analysis of variance (ANOVA) revealed a significant overall effect of phytohormone treatments on FWG (Supplementary Table B), indicating that the different hormonal combinations induce statistically significant ($p < 0.001$) variations in biomass.

Among the tested hormone combinations, 2,4-D and BAP showed a significant effect on callus induction and growth (Supplementary Figure A). The highest FWG (3639.5 mg ± 208.0) was recorded with a combination of 1 mg/L BAP with 1 mg/L 2,4-D. Similarly high values were observed with 1 mg/L BAP + 0.5 mg/L 2,4-D (3459.33 mg ± 84.68). Additionally, 2,4-D applied alone also promoted substantial callus proliferation, with FWG values of 3438.5 mg ± 493.5 and 3394 mg ± 31 at 0.5 and 1 mg/L, respectively.

In contrast, the (BAP/NAA) and (KIN/NAA) combinations produced significantly lower FWG (Supplementary Figure A), indicating less efficiency in promoting callus formation and proliferation. The maximum FWG in these groups was observed with 0.5 mg/L BAP combined with 1.5 mg/L NAA (1874.45 mg ± 25.66) and with 0.5 mg/L

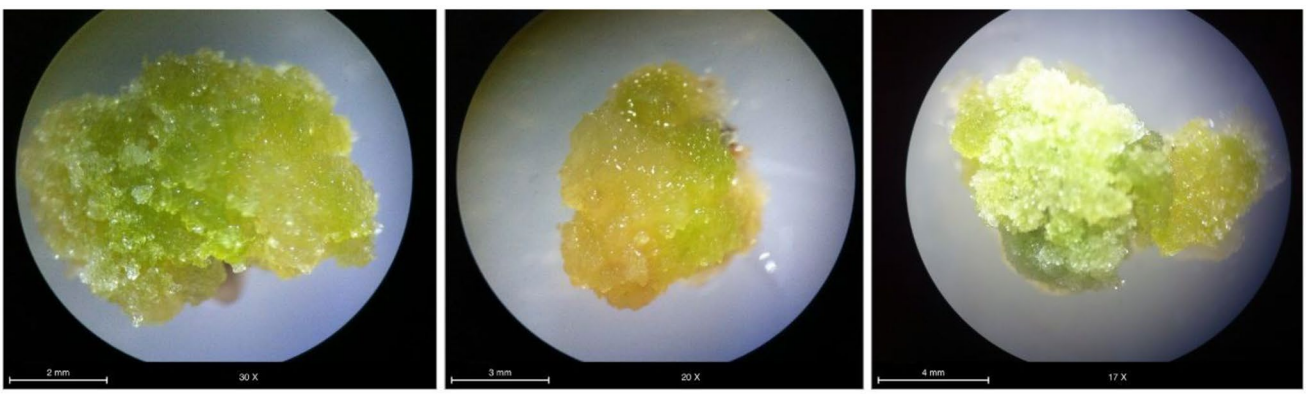

Fig. 3 Stereomicroscopic observations of *Capparis spinosa* callus derived from leaf explants

Table 1 Machine learning model performance for predicting callus fresh weight gain

ML Models	R^2	R^2 CV	RMSE
RF	0.86	0.89	263.33
GB	0.83	0.86	285.59
XGBoost	0.85	0.86	274.46
PL2	0.56	0.63	471.98

R²: coefficient of determination, R²CV: cross-validated R², RMSE: root mean square error, RF: Random Forest, GB: Gradient Boosting, XGBoost: Extreme Gradient Boosting, PL2: second-degree polynomial regression

KIN combined with 1 mg/L NAA (1350.5 mg±215.5). To further investigate these complex hormonal interactions and their effects on callus induction, ML was employed to analyze the dataset.

Model performance in predicting callus induction

The callogenesis of *C. spinosa* was quantitatively modeled using a set of explanatory variables representing four exogenously applied PGRs: BAP, KIN, NAA, and 2,4-D. The target variable for modeling was the FWG of callus tissue. To capture the complex relationships between hormonal treatments (inputs) and FWG (outputs), four ML models were applied and compared: RF, GB, XGBoost, and a traditional second-degree Polynomial Regression.

The ensemble learning models, RF, GB, and XGBoost, demonstrated superior and comparable performance (Table 1). RF achieved the best predictive performance with an R² of 0.86, R²CV of 0.89, and a low RMSE of 263.33, indicating satisfactory generalization capacity. XGBoost and GB also performed well, with R² values of 0.85 and 0.83, R²CV of 0.86 for both, and RMSE values of 274.46 and 285.59, respectively. These results highlight the ability of ensemble methods to capture complex, non-linear relationships in FWG data effectively. In contrast, the PL2 model exhibited substantially lower performance, with an R² of 0.56, R²CV of 0.63, and a markedly higher RMSE of 471.98. This indicates that simple polynomial relationships are insufficient to model the underlying biological complexity, resulting in larger prediction errors.

The combined analysis of numerical and graphical results highlights the superiority of ensemble tree-based models (RF, GB, and XGBoost) for predicting callus FWG. As shown in Fig. 4, these models exhibit regression slopes close to unity (RF: 0.98x+42.04; GB: 0.94x+83.13; XGBoost: 0.96x+71.62) and a tight clustering of points around the ideal y=x line, indicating precise calibration.

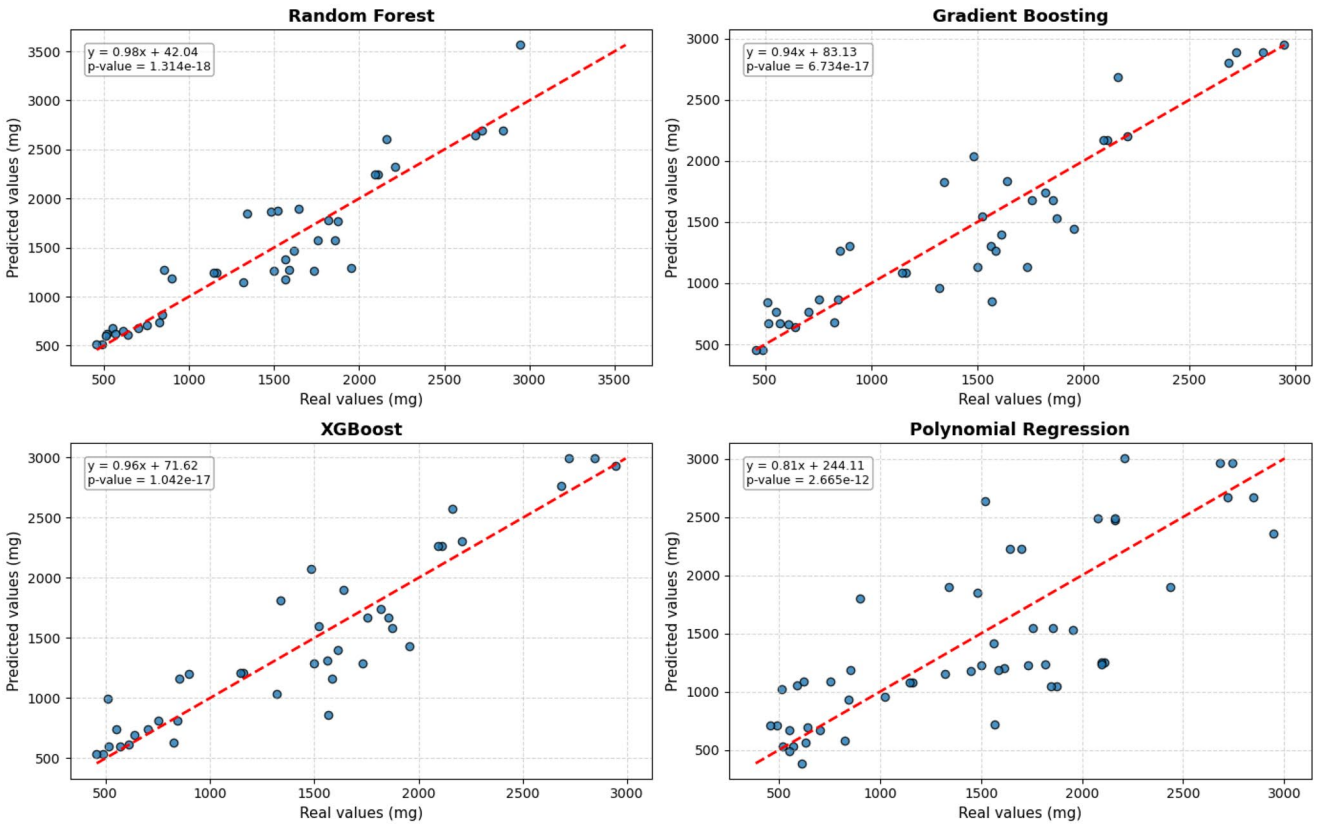

Fig. 4 Scatter plots presenting real vs. predictions of callus fresh weight gain for four models: Random Forest, Gradient Boosting, XGBoost, and second-degree polynomial regression

These findings are consistent with the high R^2 and low RMSE reported in Table 1, confirming the robustness and generalizability of these models across the full FWG range. Residual analysis further corroborated these observations, as the RF, GB, and XGBoost models exhibited residuals randomly distributed around zero with approximately normal distributions (Supplementary Figure B (a – f)), indicating the absence of notable systematic bias. In contrast, PL2 showed a markedly lower slope (0.81x + 244.11) and substantial dispersion, reflecting systematic underestimation of high FWG values and poor ability to capture the non-linearities inherent in biological data, a pattern further confirmed by its more dispersed and asymmetric residuals (Supplementary Figure B (g and h)). Overall, the Friedman test revealed highly significant differences among the tested models for both R^2CV and RMSE scores ($p < 0.001$, Supplementary Table C). Pairwise Wilcoxon signed-rank tests with Bonferroni correction showed that RF significantly outperformed GB, XGBoost, and PL2. No significant difference was observed between GB and XGBoost ($p < 0.001$, Supplementary Table D).

Therefore, RF was selected for subsequent analyses, including SHAP-based interpretation of variable importance. Figure 5 illustrates the overall importance of the variables and reveal that 2,4-D is the most influential predictor of callus FWG in the RF model, with the highest mean absolute SHAP value (~ 700), far exceeding those of BAP, KIN, and NAA (Fig. 5A). Directional SHAP analysis (Fig. 5B) indicates that 2,4-D exhibits the strongest positive association with FWG predictions (+ 12.19), suggesting a favorable association under the tested experimental conditions. BAP also contributes positively to the predictions, but to a more moderate extent (+ 6.39). In contrast, KIN is associated with a decrease in FWG predictions (− 6.95), whereas NAA showed a marginal negative contribution (− 2.01).

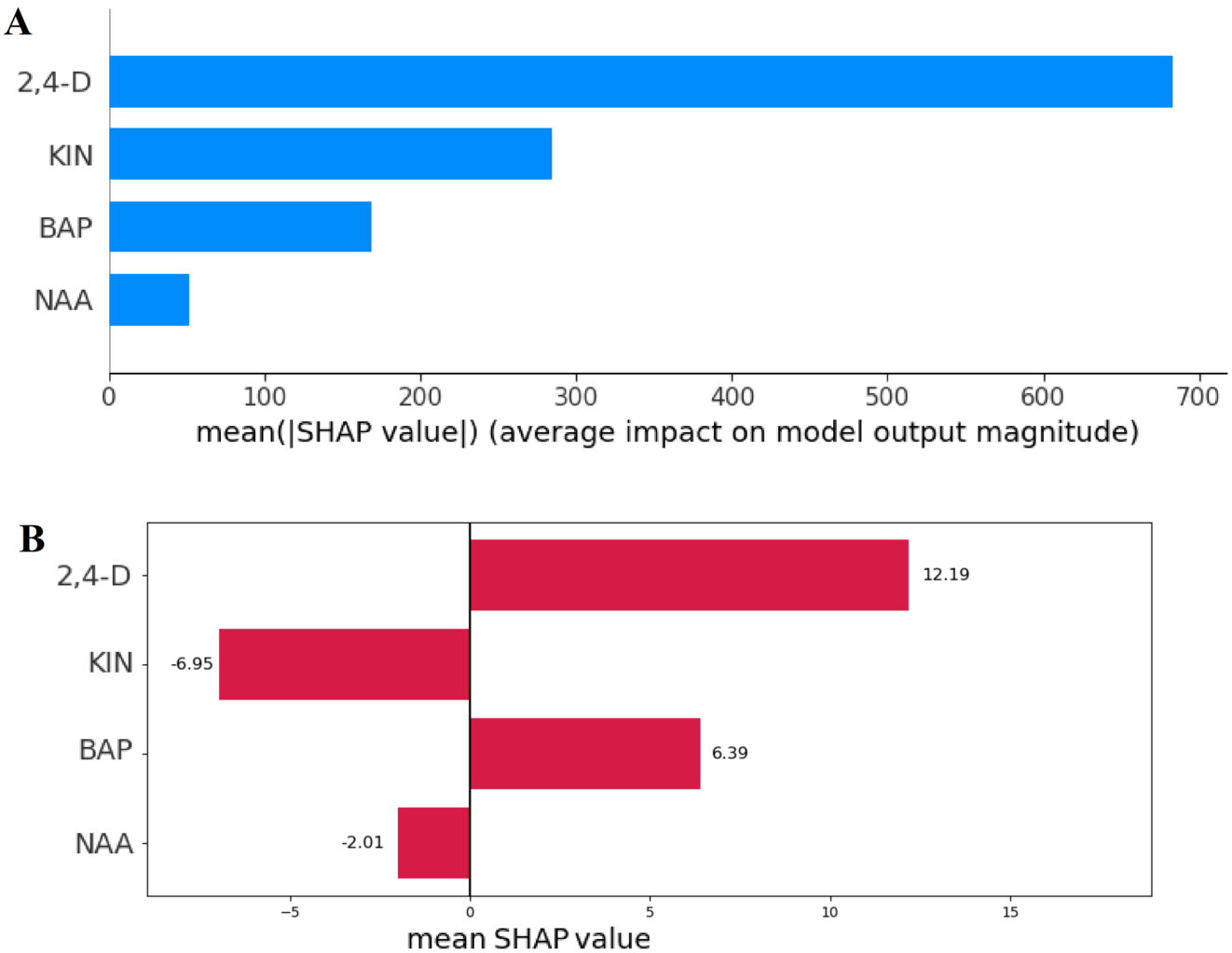

Fig. 5 Global importance (**A**) and directional effects (**B**) of hormones on predicted callus fresh weight gain using SHAP values in the Random Forest model. 6-benzylaminopurine (BAP), kinetin (KIN), 1-naphthaleneacetic acid (NAA), and 2,4-dichlorophenoxyacetic acid (2,4-D)

Experimental validation of callus fresh weight gain

Laboratory experiments were conducted to validate the five highest-ranked phytohormone combinations predicted by the RF model. The validation results demonstrated satisfactory agreement between predicted and observed callus FWG values (Fig. 6), confirming the reliability of the RF model for predicting optimal hormonal combinations for callus induction in *C. spinosa.*

Callus culture as a platform for rutin accumulation

Hormonal regulation is a key factor in controlling the accumulation of secondary metabolites in *C. spinosa* callus cultures. In this study, rutin, the major bioactive flavonoid of this species, was first identified using LC-QTOF-MS/MS, while its relative concentration was subsequently quantified using LC-TQ-MS/MS in MRM mode (609 → 300). The BAP/2,4-D hormonal combination was selected for these experiments due to its favorable influence on FWG (Fig. 5), thereby providing optimal conditions for evaluating rutin accumulation in callus tissues. The effect of hormonal treatments on rutin content was further investigated using an RF model. Figure 7 presents a comparison between experimentally determined and predicted rutin concentrations. The data indicate that rutin accumulation is strongly dependent on the relative concentrations of BAP and 2,4-D. The highest rutin levels were observed under moderate BAP concentrations (0.5–1.0 mg/L) combined with low 2,4-D concentrations (0.0–1.0 mg/L). In contrast, the observed reduction in rutin content at high BAP (≥1.5 mg/L) and 2,4-D (≥1.5 mg/L) concentrations may be explained by the fact that elevated auxin and cytokinin levels preferentially promote cell division and biomass accumulation rather than secondary metabolite accumulation, consistent with the high FWG values observed under these conditions (Fig. 6). Overall, the RF model effectively captured these trends, as well as the non-linear relationship between PGRs and rutin accumulation, with predicted values closely matching the experimental observations.

The RF model applied to rutin accumulation prediction achieved an R^2 of 0.821, an R^2CV of 0.837, and an RMSE of 0.408, with a regression slope of 0.79 and an intercept of 0.25 (Fig. 8). The associated *p*-value was significant ($p<0.05$), indicating satisfactory predictive performance. Residual analysis further supported these results, as the model exhibited residuals randomly distributed around zero and approximately following a normal distribution (Supplementary Figure C), confirming the absence of systematic bias and the adequacy of model fit. These findings indicate that the model reliably reproduces overall trends, even within a complex biological system, despite the relatively limited dataset size. The marginal differences in R^2 between experimentally measured FWG values and RF-predicted FWG values confirmed that error propagation had a limited impact on rutin accumulation predictions. Ensemble tree-based algorithms are distinguished by their flexibility, ability to model higher-order nonlinear interactions, capacity to handle heteroscedasticity, and robustness against overfitting. These properties make them particularly well suited for predicting and understanding the accumulation of secondary metabolites, such as rutin, in plant tissue cultures.

Experimental validation of rutin accumulation in callus cultures

The predictions of rutin accumulation obtained by the RF model (Fig. 9), based on the BAP/2, 4-D combinations, show overall agreement with the experimental data, particularly for the condition BAP=2.0+2,4-D=2.0, highlighting the robustness of the model in anticipating metabolic responses to different hormonal treatments. The slight overestimation observed for certain combinations (BAP=1.5+2,4-D=2.0 and BAP=2.0+2,4-D=1.0) does not alter the overall

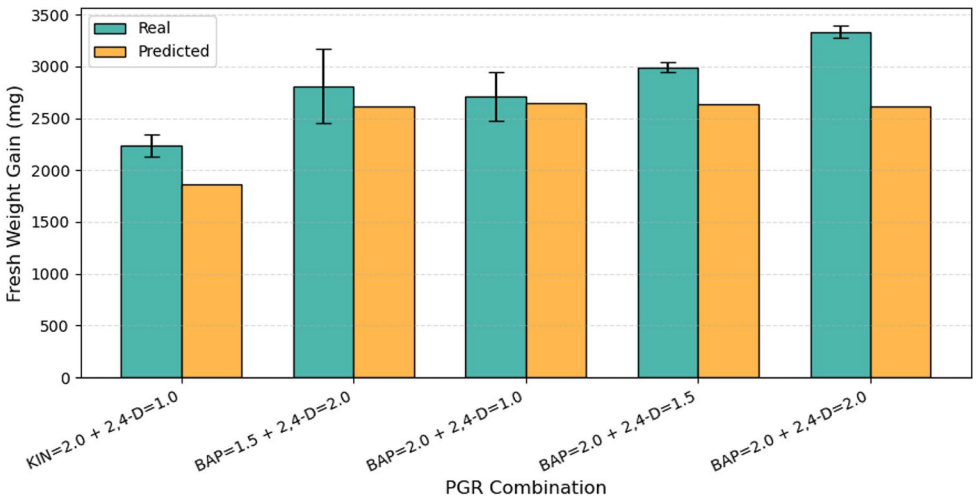

Fig. 6 Real vs. five highest-ranked predicted hormonal combinations for maximization of callus fresh weight gain using the Random Forest model

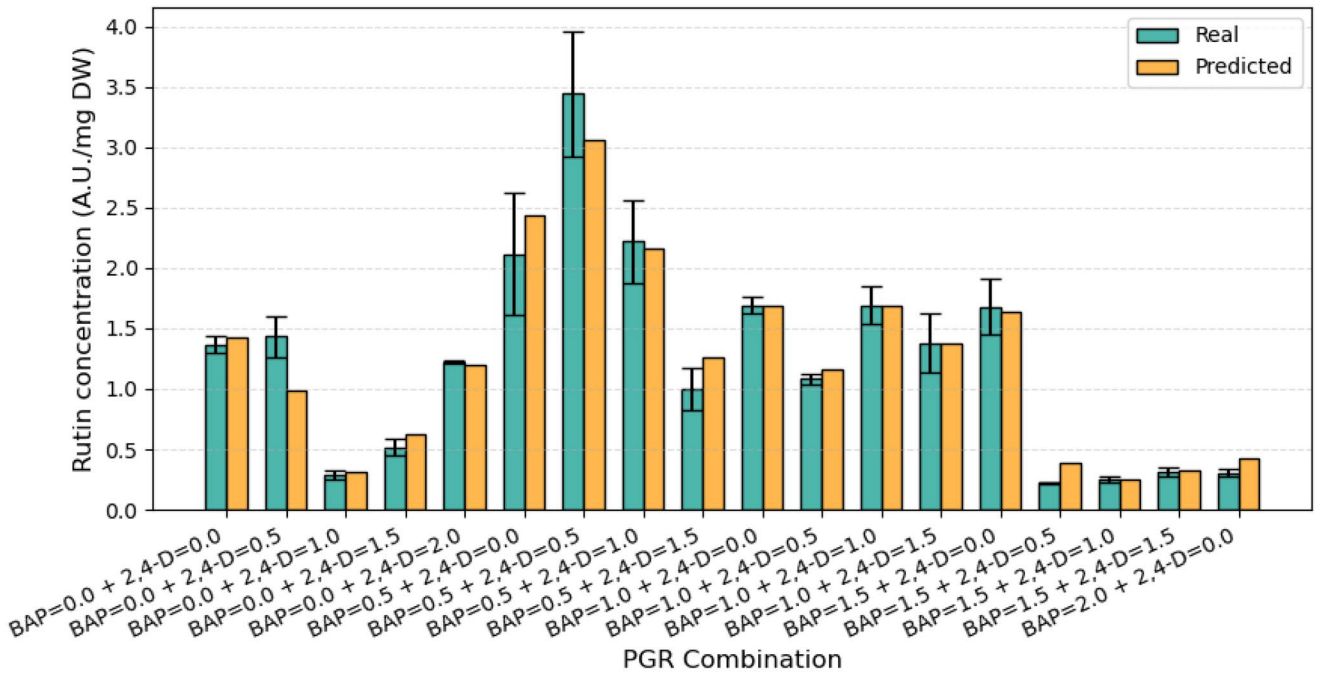

Fig. 7 Comparison of real (green) and predicted (orange) rutin accumulation under different hormonal combinations of 6-benzylaminopurine and 2,4-dichlorophenoxyacetic acid using the Random Forest model

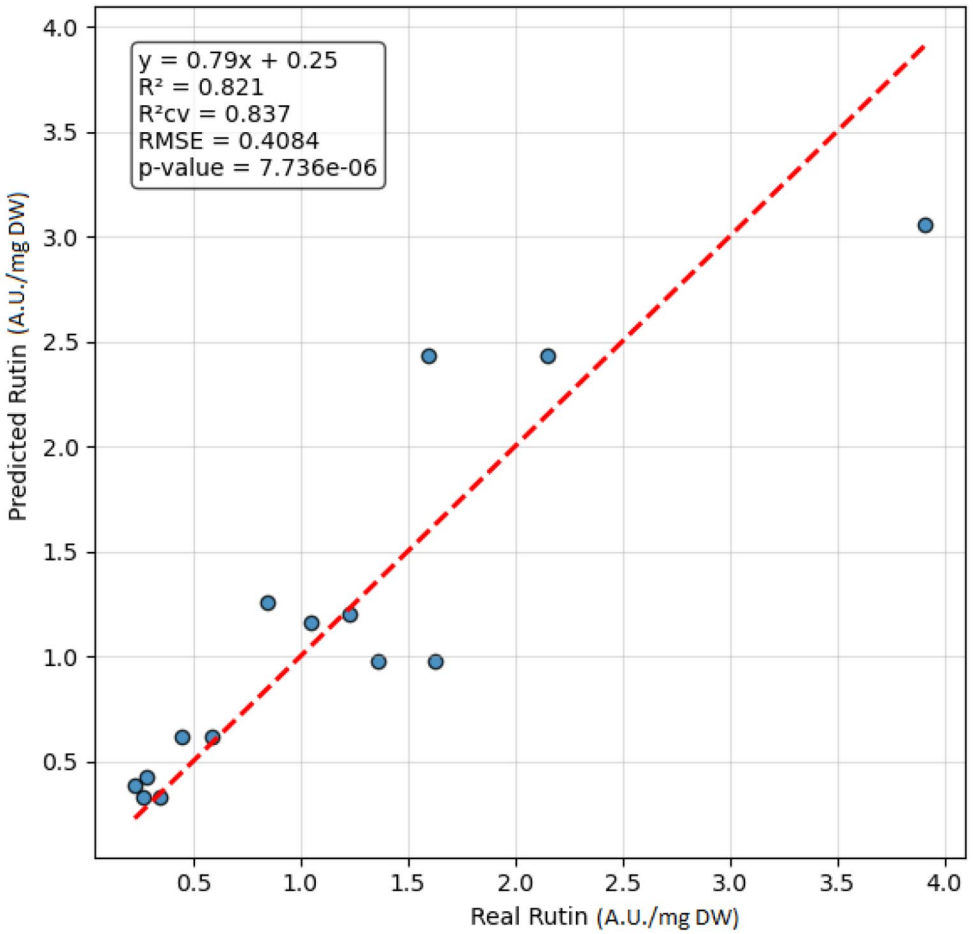

Fig. 8 Scatter plots presenting real vs. predictions rutin accumulation values obtained by the Random Forest model

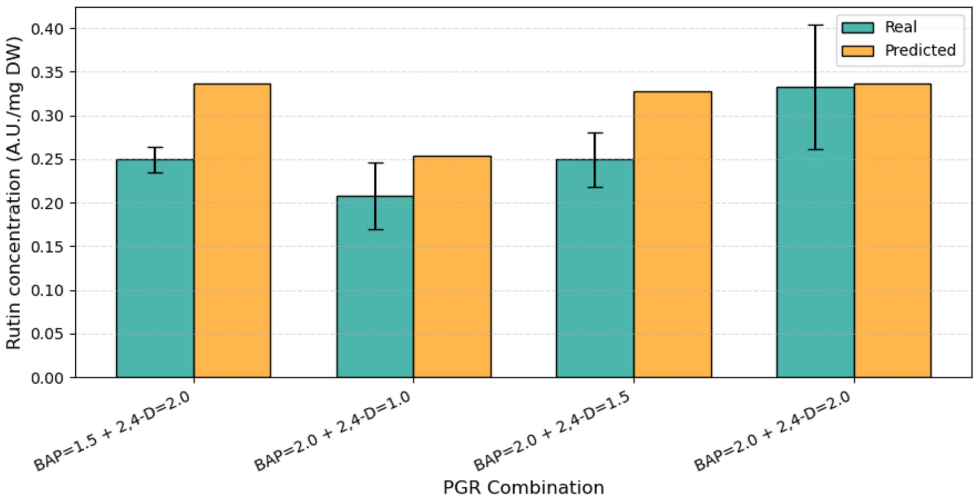

Fig. 9 Real vs. predicted rutin accumulation using the Random Forest model

consistency of the predictions. These results demonstrate that the model represents a reliable tool for estimating rutin accumulation in *C. spinosa* callus cultures and provide effective support for optimizing culture conditions to maximize secondary metabolite accumulation.

Discussion

Callus induction is a central step in plant tissue culture, representing a cellular reprogramming process in which differentiated cells revert to an undifferentiated and proliferative state (Lee et al. 2024). This process is primarily regulated by the exogenous application of PGRs, which provide the plasticity required for subsequent morphogenetic events, such as organogenesis and somatic embryogenesis (Pasternak and Steinmacher 2024). Callus also serves as a reservoir for secondary metabolite accumulation, including rutin, highlighting the importance of understanding auxin-cytokinin effects on both callogenesis and metabolic activity.

In this study, the effects of different concentrations of BAP, KIN, NAA, and 2,4-D on FWG and rutin accumulation in *C. spinosa* were evaluated using RF, GB, XGBoost, and PL2. ML algorithms, particularly ensemble tree-based models, have been increasingly applied in plant tissue culture, although comparative studies remain limited.

These findings highlight that classic statistical methods, based on linear relationships or limited curvature, struggle to capture the complex interactions, threshold effects, and saturation phenomena characteristic of plant hormonal responses (Sadat-Hosseini et al. 2022). Callus formation in vitro involves highly non-linear and synergistic interactions among PGRs, requiring more flexible modeling approaches (Fallah Ziarani et al. 2022). In this context, ensemble tree-based methods such as RF and XGBoost are particularly well suited to this challenge (Munasinghe et al. 2020). Their key advantages include the ability to model high-order interactions, handle heteroscedasticity, and remain robust to overfitting (Ali and Aasim 2024; Kaushik et al. 2025). However, several limitations of these approaches deserve recognition, as they have the potential to affect the model's overall performance. In particular, the relatively limited size of the dataset, a common feature of in vitro studies (Hesami and Jones 2020), restricts the model's generalizability and may reduce the reliability of predictions outside the tested hormonal range. Furthermore, the models developed in this study were trained exclusively on plant PGR concentrations, while other influential factors, such as temperature, light intensity, explant age, and culture medium composition, were held constant and therefore not included as input variables. Although SHAP analysis significantly improves the model's interpretability, the RF model remains partially opaque compared to mechanistic biological models.

The demonstrated accuracy of these models highlights the value of integrating ML into plant tissue culture research, not only to improve predictive ability but also to enhance our understanding of the complex mechanisms driving hormonal responses.

To evaluate the relative contributions of individual growth regulators to callus development, we focused the SHAP analysis on the RF model, selected for its superior performance in terms of predictive accuracy. This approach allows for a quantitative assessment of the effect of each factor (BAP, KIN, NAA, and 2,4-D) on FWG predictions for *C. spinosa* callus tissue, while taking into account potential interactions between variables.

Recent studies comparing ML models for predicting callogenesis in saffron (*Crocus sativus*) have reported that boosting-based algorithms, particularly XGBoost and Gradient Boosting, achieve high predictive accuracy ($R^2 > 0.95$), outperforming artificial neural networks and RF in modeling in vitro culture systems (Sarabandi et al. 2024).

In the present study, the predictive performance of several approaches, RF, GB, XGBoost, and PL2, was evaluated for FWG in *C. spinosa*. Tree-based ensemble models consistently outperformed polynomial regression, with RF exhibiting the best overall predictive performance for callogenesis ($R^2 = 0.86$). Although this performance is slightly lower than the R^2 values exceeding 0.95 reported for *C. sativus*, this discrepancy should be interpreted cautiously due to differences in input variables, experimental protocols, and species-specific biological complexity, which limit direct comparability between studies. Nevertheless, the results consistently highlight the central role of auxins in callus induction (Sarabandi et al. 2024). Although this study represents the first direct comparison of these algorithms in the context of *C. spinosa in vitro* culture, the obtained performance levels are consistent with those reported in comparable plant tissue culture studies conducted under similar experimental conditions (Özcan et al. 2023; Bozkurt et al. 2024; Sarmah et al. 2025). Furthermore, the robustness of RF has been demonstrated in related biological systems, including *Nilgirianthus ciliatus*, where R^2 values of up to 0.97 were reported for callus growth and metabolite production (Jeevan Ram et al. 2025), further supporting its suitability for such predictive applications.

The utility of SHAP values as an interpretive tool for identifying and quantifying the importance of variables has been demonstrated in several recent studies (Ekanayake et al. 2022; Antonini et al. 2024). SHAP constitutes a mathematically rigorous, consistent, and objective approach for decomposing a model's predictions into contributions attributed to each of the explanatory variables (Lundberg and Lee 2017). This method enables both a precise local interpretation of individual predictions and a global analysis of the importance of the model's variables (Lundberg et al. 2020). Using SHAP, key variables such as 2,4-D, NAA, sucrose concentration, and culture duration, have been identified as determinants in various experimental contexts (Sarabandi et al. 2024). In a similar approach, Kumar and Kumar (2026) used SHAP-based interpretability to evaluate the effects of PGRs on callogenesis in *Clerodendrum phlomidis* using an XGBoost model. Their results indicated that fresh callus weight is a major response variable, with the model primarily highlighting the influence of auxins, particularly NAA and indole-3-acetic acid (IAA). Moreover, cytokinins such as BAP exhibited a positive effect on induction time, whereas KIN showed a relatively weaker contribution to both studied parameters. This is consistent with the present study, where SHAP analysis highlighted 2,4-D as the most influential factor in predicting FWG, followed by the effect of BAP, while KIN exhibited a negative contribution on callus formation in *C. spinosa*. The identification of these factors is of particular importance when the accumulation of secondary metabolites is the primary objective of callus cultures.

To our knowledge, this study provides the first systematic assessment of the impact of various PGRs on both callogenesis and rutin accumulation in *C. spinosa*. To directly link callus growth with metabolite accumulation, we employed a stacked RF model, using callogenesis outputs as inputs to predict rutin accumulation. This approach effectively modeled the nonlinear relationship between cell proliferation and rutin accumulation, yielding satisfactory predictive performance ($R^2 = 0.82$) and minimal deviation between observed and predicted values, confirming method robustness.

Direct comparisons with previous studies remain limited due to differences in explant sources, methodologies, and experimental objectives. However, qualitative trends are consistent across studies. Yin et al. (2014) and Duran and Issah (2022) reported similar callus responses to BAP, NAA, and 2,4-D combinations, although values are not directly comparable. In the present study, maximum callus formation was observed with 2,4-D, consistent with reports in *C. spinosa* (Sobhy et al. 2025), and other species where auxin application significantly promotes callogenesis (Slimani et al. 2021; Teoh et al. 2023; Ranade et al. 2023; Mahood et al. 2024; Abdelazeez et al. 2025).

Overall, experimental validation confirmed the reliability of the RF model, with satisfactory agreement between predicted and observed values, despite slight overestimations in some cases. These results highlight the potential of ML approaches to optimize callus culture and the accumulation of secondary metabolites. A similar trend was reported by Jeevan Ram et al. (2025) in *Nilgirianthus ciliatus*, where ML models, including the multilayer perceptron, LightGBM, and RF, showed high concordance between predicted and observed values for key callus growth parameters and metabolic parameters. Future studies incorporating a wider range of physiological and environmental variables, as well as larger datasets, may further improve the model's generalizability. Investigations might also focus on the combination of 2,4-D with other plant growth regulators, or extend to cell suspension cultures and bioreactor production systems, which represent promising avenues for future research.

Conclusion

This research provides the first comprehensive evaluation of the combined effects of BAP, KIN, NAA, and 2,4-D on callogenesis and rutin accumulation in *C. spinosa* using an integrated ML and explainable-AI framework. Ensemble tree-based models, particularly RF, accurately predicted callus biomass and metabolite accumulation,

outperforming classical statistical approaches. SHAP analysis identified 2,4-D as the most influential predictor, while BAP showed a positive additive contribution to callus growth, and KIN and NAA contributed weak or inhibitory effects. Experimental validation confirmed the reliability of the predicted optimal hormonal combinations, demonstrating satisfactory agreement between predicted and observed FWG and rutin levels. Moderate concentrations of BAP and 2,4-D favored rutin accumulation, underscoring the importance of fine hormonal tuning for metabolic optimization.

Overall, this study highlights the relevance of ML-driven modeling for deciphering complex hormonal interactions and accelerating protocol optimization in plant tissue culture. The established models constitute a robust foundation for scaling up callus-based production systems, including suspension cultures and bioreactors, aimed at sustainable, high-yield accumulation of rutin and other valuable metabolites in *C. spinosa*. Future investigations integrating larger datasets, molecular markers, and elicitation strategies will further strengthen predictive performance and advance the biotechnological exploitation of this species.

Author contribution Marouane Mohaddab: Writing – review & editing, Writing – original draft, Conceptualization, Methodology, Data curation, Formal analysis. Younes EL Goumi: Supervision, Writing – review and editing, Methodology. Mohammed Elakrouch: Formal analysis, Software, Methodology. Soufiane Hasni: Formal analysis, Software, Data curation. Clément Burgeon: Writing – review & editing, Software, Data curation. Manon Genva: Writing – review & editing, Software. Malika Fakiri: Writing – review & editing, Validation, Project administration Marie-Laure Fauconnier: Writing – review & editing, Validation, Project administration.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval Not applicable.

Conflict of interest The authors declare that there is no conflict of interest.

References

Aasim M, Ali SA, Altaf MT et al (2023) Artificial neural network and decision tree facilitated prediction and validation of cytokinin-auxin induced in vitro organogenesis of sorghum (Sorghum bicolor L). Plant Cell Tiss Organ Cult 153:611–624. https://doi.org/10.1007/s11240-023-02498-3

Abdelazeez WMA, Abouelidis GR, Suliman AA, Mohammed DM (2025) Production of secondary metabolites in callus cultures of Scutellaria baicalensis L. and assessment of their anti-inflammatory and antioxidant efficacy in ulcerative colitis rats. Plant Cell Tiss Organ Cult 160:80. https://doi.org/10.1007/s11240-025-02996-6

Al-Safadi B, Elias R (2011) Improvement of caper (Capparis spinosa L.) propagation using in vitro culture and gamma irradiation. Sci Hortic 127:290–297. https://doi.org/10.1016/j.scienta.2010.10.014

Alcalde MA, Müller M, Munné-Bosch S et al (2022) Using machine learning to link the influence of transferred Agrobacterium rhizogenes genes to the hormone profile and morphological traits in Centella asiatica hairy roots. Front Plant Sci 13:1001023. https://doi.org/10.3389/fpls.2022.1001023

Ali SA, Aasim M (2024) Response surface methodology and artificial intelligence modeling for in vitro regeneration of Brazilian micro sword (Lilaeopsis brasiliensis). Plant Cell Tiss Organ Cult 157:10. https://doi.org/10.1007/s11240-024-02734-4

Annaz H, Sane Y, Bitchagno GTM et al (2022) Caper (*Capparis spinosa* L.): an updated review on its phytochemistry, nutritional value, traditional uses, and therapeutic potential. Front Pharmacol 13:878749. https://doi.org/10.3389/fphar.2022.878749

Antonini AS, Tanzola J, Asiain L et al (2024) Machine Learning model interpretability using SHAP values: Application to Igneous Rock Classification task. Appl Comput Geosci 23:100178. https://doi.org/10.1016/j.acags.2024.100178

Baker RE, Peña J-M, Jayamohan J, Jérusalem A (2018) Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol Lett 14:20170660. https://doi.org/10.1098/rsbl.2017.0660

Belkessam M, Genva M, Kouki A et al (2025) Phytochemical profile and cosmeceutical potential of leaf extracts of two species of the Anacardiaceae family from the Mediterranean scrubland: Pistacia lentiscus L. and Pistacia atlantica Desf. Processes 13:3712. https://doi.org/10.3390/pr13113712

Bozkurt T, İnan S, Dündar İ et al (2024) Optimizing the in vitro propagation of tea plants: a comparative analysis of machine learning models. Horticulturae 10:721. https://doi.org/10.3390/horticulturae10070721

Bravo-Vázquez LA, Angulo–Bejarano PI, Bandyopadhyay A et al (2023) Regulatory roles of noncoding RNAs in callus induction and plant cell dedifferentiation. Plant Cell Rep 42:689–705. https://doi.org/10.1007/s00299-023-02992-0

Caglar G, Caglar S, Ergin O, Yarim M (2005) The influence of growth regulators on shoot proliferation and rooting of in vitro propagated caper. J Environ Biol 26:479–485

Carra A, Del Signore MB, Sottile F et al (2012) Potential use of new diphenylurea derivatives in micropropagation of Capparis spinosa L. Plant Growth Regul 66:229–237. https://doi.org/10.1007/s10725-011-9645-3

Demirel F, Uğur R, Popescu GC et al (2023) Usage of machine learning algorithms for establishing an effective protocol for the in vitro micropropagation ability of black chokeberry (Aronia melanocarpa (Michx.) Elliott). Horticulturae 9:1112. https://doi.org/10.3390/horticulturae9101112

Duran RE, Issah H (2022) The impact of strigolactone GR24 on *Capparis spinosa* L. callus production and phenolic compound content. Plant Cell Tiss Organ Cult 149:197–204. https://doi.org/10.1007/s11240-021-02212-1

Efferth T (2019) Biotechnology applications of plant callus cultures. Engineering 5:50–59. https://doi.org/10.1016/j.eng.2018.11.006

Ekanayake IU, Meddage DPP, Rathnayake U (2022) A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Stud Constr Mater 16:e01059. https://doi.org/10.1016/j.cscm.2022.e01059

Eren B, Türkoğlu A, Haliloğlu K et al (2023) Investigation of the influence of polyamines on mature embryo culture and DNA methylation of wheat (Triticum aestivum L.) using the machine learning

algorithm method. Plants 12:3261. https://doi.org/10.3390/plants12183261
Esmaeilpour HR, Boostani R, Shoeibi A et al (2020) The potential effects of Caper (Capparis spinosa L.) in the treatment of diabetic neuropathy. Tradit Med Res 5:442–448. https://doi.org/10.12032/TMR20200223167
Fahmideh L, Sheikhi M, Benakashani F, Solouki M (2019) Callus induction and organogenesis from various explants of plant Capparis spinosa L. under in vitro conditions. J Plant Prod Res 26:75–88. https://doi.org/10.22069/jopp.2019.14326.2284
Fallah Ziarani M, Tohidfar M, Navvabi M (2022) Modeling and optimizing in vitro percentage and speed callus induction of carrot via multilayer perceptron-single point discrete GA and radial basis function. BMC Biotechnol 22:19. https://doi.org/10.1186/s12896-022-00764-4
Gianguzzi V, Inglese P, Barone E, Sottile F (2019) In vitro regeneration of Capparis spinosa L. by using a temporary immersion system. Plants 8:177. https://doi.org/10.3390/plants8060177
Grimalt M, Hernández F, Legua P et al (2018) Physicochemical composition and antioxidant activity of three Spanish caper (Capparis spinosa L.) fruit cultivars in three stages of development. Sci hortic 240:509–515. https://doi.org/10.1016/j.scienta.2018.06.061
Grimalt M, Hernández F, Legua P et al (2022) Antioxidant activity and the physicochemical composition of young caper shoots (Capparis spinosa L.) of different Spanish cultivars. Sci Hortic 293:110646. https://doi.org/10.1016/j.scienta.2021.110646
Gristina AS, Fici S, Siragusa M et al (2014) Hybridization in Capparis spinosa L.: molecular and morphological evidence from a Mediterranean island complex. Flora 209:733–741. https://doi.org/10.1016/j.flora.2014.09.002
Hagaggi NShA, Abdul-Raouf UM, Radwan TAA (2024) Variation of antibacterial and antioxidant secondary metabolites and volatiles in leaf and callus extracts of Phulai (Acacia Modesta Wall). BMC Plant Biol 24:93. https://doi.org/10.1186/s12870-024-04747-9
Hazrati S, Mousavi Z, Mollaei S et al (2025) Effect of Phenological variation on the phytochemical composition and antioxidant activity of different organs of Capparis spinosa L. Horticulturae 11:702. https://doi.org/10.3390/horticulturae11060702
Hesami M, Jones AMP (2020) Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. Appl Microbiol Biotechnol 104:9449–9485. https://doi.org/10.1007/s00253-020-10888-2
Hesami M, Jones AMP (2021) Modeling and optimizing callus growth and development in Cannabis sativa using random forest and support vector machine in combination with a genetic algorithm. Appl Microbiol Biotechnol 105:5201–5212. https://doi.org/10.1007/s00253-021-11375-y
Higazy AE, El-Mahrouk ME, El-Banna AN et al (2023) Production of black cumin via somatic embryogenesis, chemical profile of active compounds in callus cultures and somatic embryos at different auxin supplementations. Agronomy 13:2633. https://doi.org/10.3390/agronomy13102633
Ikeuchi M, Sugimoto K, Iwase A (2013) Plant callus: mechanisms of induction and repression. Plant cell 25:3159–3173. https://doi.org/10.1105/tpc.113.116053
Jeevan Ram PS, Singh S, Ali SA et al (2025) Machine learning-driven screening of promising mutants in Nilgirianthus ciliatus callus cultures for enhanced phytochemicals production. Physiol Plant 177:e70536. https://doi.org/10.1111/ppl.70536
Karous O, Ben Haj Jilani I, Ghrabi-Gammar Z (2021) Ethnobotanical study on plant used by semi-nomad descendants' community in Ouled Dabbeb—Southern Tunisia. Plants 10:642. https://doi.org/10.3390/plants10040642
Katırcı R, Aasim M, Deveci G, Mustafa Z (2024) Comparing quantum machine learning and classical machine learning for in vitro regeneration of cowpea (Vigna unguiculata). Plant Cell Tiss Organ Cult 159:32. https://doi.org/10.1007/s11240-024-02880-9
Kaushik P, Rani M, Khurana N et al (2025) Revolutionizing Plant Tissue Culture: harnessing artificial intelligence for precision propagation and optimization. NPJ 15:e040624230642. https://doi.org/10.2174/0122103155302871240527094915
Kianersi F, Abdollahi MR, Mirzaie-asl A et al (2020) Biosynthesis of rutin changes in Capparis spinosa due to altered expression of its pathway genes under elicitors' supplementation. Plant Cell Tiss Organ Cult 141:619–631. https://doi.org/10.1007/s11240-020-01823-4
Kirkan B, Ceylan O, Sarikürkcü C, Bektas T (2021) Phenolic profile, antioxidant and enzyme inhibitory activity of the ethyl acetate, methanol and water extracts of Capparis spinosa L. Int J Second Metabolite 8:337–351. https://doi.org/10.21448/ijsm.981149
Koçak M, Yılmaz MC, Kuzğun C et al (2025) Synthetic seed production in Crataegus monogyna L. and prediction of regeneration of synthetic seeds with machine learning algorithms. Plant Cell Tiss Organ Cult 161:44. https://doi.org/10.1007/s11240-025-03049-8
Koufan M, Belkoura I, Mazri MA (2022) In vitro propagation of Caper (Capparis spinosa L.): a review. Horticulturae 8:737. https://doi.org/10.3390/horticulturae8080737
Kumar RP, Kumar SR (2026) Machine learning-driven optimization of Clerodendrum phlomidis L. f. (sage glory bower) tissue culture for enhanced callogenesis. Plant Cell Tiss Organ Cult 165:43. https://doi.org/10.1007/s11240-026-03463-6
Kumari S, Nazir F, Maheshwari C et al (2024) Plant hormones and secondary metabolites under environmental stresses: enlightening defense molecules. Plant Physiol Biochem 206:108238. https://doi.org/10.1016/j.plaphy.2023.108238
Labbafi M, Mehrafarin A, Badi H et al (2018) Improve germination of caper (Capparis spinosa L.) seeds by different induction treatments of seed dormancy breaking. Trakia J Sci 16:71–74. https://doi.org/10.15547/tjs.2018.01.010
Lee KY, Nam D-H, Jeon Y et al (2024) Exploring the production of secondary metabolites from a Halophyte Tetragonia tetragonoides through callus culture. Horticulturae 10:244. https://doi.org/10.3390/horticulturae10030244
Li W, Lin S, Wang R et al (2025) Regulation of plant hormones on the secondary metabolism of medicinal plants. Med plants Biol 4:e0016. https://doi.org/10.48130/mpb-0025-0016
Liu W, He Y, Xiang J et al (2011) The physiological response of suspension cell of Capparis spinosa L. to drought stress. J Med Plants Res 5:5899–5906
Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30:4765–4774. https://doi.org/10.48550/arXiv.1705.07874
Lundberg SM, Erion G, Chen H et al (2020) From local explanations to global understanding with explainable ai for trees. Nat Mach Intell 2:56–67. https://doi.org/10.1038/s42256-019-0138-9
Ma Q, Grones P, Robert S (2018) Auxin signaling: a big question to be addressed by small molecules. J Exp Bot 69:313–328. https://doi.org/10.1093/jxb/erx375
Mahood HE, Sarropoulou V, Tzatzani T-T et al (2024) Biomass and polyphenolic enhancement in date palm (Phoenix dactylifera L.) callus cultures through Fusarium oxysporum elicitation. Plant Cell Tiss Organ Cult 159:76. https://doi.org/10.1007/s11240-024-02934-y
Moghadamnia Y, Kani SNM, Ghasemi-Kasman M et al (2019) The anti-cancer effects of *Capparis spinosa* hydroalcoholic extract. Avicenna J med Biotech 11:43–47. https://doi.org/10.1016/j.avb.2019.01.001
Mohaddab M, El Goumi Y, Gallo M et al (2022) Biotechnology and in vitro culture as an alternative system for secondary metabolite production. Molecules 27:8093. https://doi.org/10.3390/molecules27228093

Mohaddab M, Genva M, Malika F et al (2024) *Capparis spinosa*: a rich source of phenolic compounds — a comprehensive review of its phytochemistry, health benefits, and biotechnological applications. Biocatal Agric Biotechnol 55:103409. https://doi.org/10.1016/j.bcab.2024.103409

Movafeghi A, HABIBI G, ALI AM (2008) Plant regeneration of Capparis spinosa L. using hypocotyl explants. Iran J Biol 21:289–297

Munasinghe SP, Somaratne S, Weerakoon SR, Ranasinghe C (2020) Prediction of chemical composition for callus production in Gyrinops walla Gaetner through machine learning. Inf Process Agric 7:511–522. https://doi.org/10.1016/j.inpa.2019.09.006

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Nimavat N, Parikh P (2024) Innovations in date palm (Phoenix dactylifera L.) micropropagation: detailed review of in vitro culture methods and plant growth regulator applications. Plant Cell Tiss Organ Cult 159:6. https://doi.org/10.1007/s11240-024-02866-7

Noordijk B, Garcia Gomez ML, ten Tusscher KHWJ et al (2024) The rise of scientific machine learning: a perspective on combining mechanistic modelling with machine learning for systems biology. Front Syst Biol 4:1407994. https://doi.org/10.3389/fsysb.2024.1407994

Nowruzian A, Aalami A (2023) Effect of different treatments on dormancy breaking for seed germination enhancement and metabolite analysis of Capparis spinosa L. J Med plants By-Prod 12:387–396. https://doi.org/10.22092/jmpb.2022.357295.1443

Özcan E, Atar HH, Ali SA, Aasim M (2023) Artificial neural network and decision tree–based models for prediction and validation of in vitro organogenesis of two hydrophytes—Hemianthus callitrichoides and Riccia fluitans. Vitro Cell Dev Biol Plant 59:547–562. https://doi.org/10.1007/s11627-023-10367-z

Pasternak TP, Steinmacher D (2024) Plant growth regulation in cell and tissue culture in vitro. Plants 13:327. https://doi.org/10.3390/plants13020327

Peng S, Rajjou L (2024) Advancing plant biology through deep learning-powered natural language processing. Plant Cell Rep 43:208. https://doi.org/10.1007/s00299-024-03294-9

Petch J, Di S, Nelson W (2022) Opening the black box: the promise and limitations of explainable machine learning in cardiology. Can J Cardiol 38:204–213. https://doi.org/10.1016/j.cjca.2021.09.004

Rahimi VB, Rajabian A, Rajabi H et al (2020) The effects of hydro-ethanolic extract of *Capparis spinosa* (*C. spinosa*) on lipopolysaccharide (LPS)-induced inflammation and cognitive impairment: evidence from *in vivo* and *in vitro* studies. J Ethnopharmacol 256:112706. https://doi.org/10.1016/j.jep.2020.112706

Ramezani Z, Aghel N, Keyghobadi H (2008) Rutin from different parts of *Capparis spinosa* growing wild in Khuzestan/Iran. Pak J Biol Sci 11:768–772. https://doi.org/10.3923/pjbs.2008.768.772

Ranade R, Joshi N, Kudale S (2023) Comparative secondary metabolite expression in callus cultures and mother plant in Barleria prionitis L. Plant Cell Tiss Organ Cult 155:653–663. https://doi.org/10.1007/s11240-023-02585-5

Sadat-Hosseini M, Arab MM, Soltani M et al (2022) Predictive modeling of Persian walnut (Juglans regia L.) in vitro proliferation media using machine learning approaches: a comparative study of ANN, KNN and GEP models. Plant Methods 18:48. https://doi.org/10.1186/s13007-022-00871-5

Sarabandi M, Singh RK, Kalantari S et al (2024) Unveiling machine learning's impact on *in vitro* callogenesis optimization in *Crocus sativus* L. S Afr J Bot 168:1–8. https://doi.org/10.1016/j.sajb.2024.02.053

Sarmah P, Borah T, Mudoi KD et al (2025) In vitro micropropagation of the critically endangered Aquilaria malaccensis Lam. and comparing the effect of PGRs through machine learning models. Plant Cell Tiss Organ Cult 162:6. https://doi.org/10.1007/s11240-025-03115-1

Slimani C, El Goumi Y, Rais C et al (2021) Optimization of callogenesis/caulogenesis induction protocol in saffron plant (*Crocus sativus* L.) using response surface methodology. Biointerface Res Appl Chem 12:4731–4746. https://doi.org/10.33263/BRIAC124.47314746

Sobhy FS, Ahmed S, Abdelrahim E et al (2025) Chemo preventive potential of rutin-rich extract obtained from Capparis spinosa culture against HepG2 cancer cell line activity. Egypt J Chem 68:855–866

Sozzi GO, Vicente AR (2006) Capers and caperberries. In: Peter KV (ed) Handbook of herbs and spices. Elsevier, Cambridge, pp 230–256

Sozzi GO, Peter KV, Babu KN, Divakaran M (2012) Capers and caperberries. In: Peter KV (ed) Handbook of herbs and spices, 2nd edn. Woodhead Publishing/Elsevier, Cambridge, pp 193–224

Svolacchia N, Sabatini S (2023) Cytokinins. Curr Biol 33:R10–R13. https://doi.org/10.1016/j.cub.2022.11.022

Tan WH, Tong CY, Chua MX, Derek CJC (2024) Modelling and optimizing secondary metabolites production in Spirodela polyrhiza using machine learning. J Clean Prod 478:143986. https://doi.org/10.1016/j.jclepro.2024.143986

Teoh SC, Subramaniam S, Chew BL (2023) The effects of 2, 4-dichlorophenoxyacetic acid on the induction of callus from cotyledon and hypocotyl explants of butterfly pea (*Clitoria ternatea*). Malays Appl Biol 52:61–72. https://doi.org/10.55230/mabjournal.v52i1.2444

Wang YT, Gan L, Liu W et al (2007) Research on the callus inducing and the cell growth and metabolism characteristics of Capparis spinosa L. Prog Mod Biomed 7:1779–1783

Wijerathna-Yapa A, Hiti-Bandaralage J, Pathirana R (2025) Harnessing metabolites from plant cell tissue and organ culture for sustainable biotechnology. Plant Cell Tiss Organ Cult 162:55. https://doi.org/10.1007/s11240-025-03180-6

Yin Y, He Y, Liu W et al (2014) The durative use of suspension cells and callus for volatile oil by comparative with seeds and fruits in *Capparis spinosa* L. PLoS ONE 9:e113668. https://doi.org/10.1371/journal.pone.0113668

Yousefi E, Abedi M, Aghajanzadeh TA, Moreno DA (2025) Caper bush (Capparis spinosa L.) bioactive compounds and antioxidant capacity as affected by adaptation to harsh soils. Sci Rep 15:11893. https://doi.org/10.1038/s41598-025-97298-4

Zarbakhsh S, Shahsavar AR, Soltani M (2024) Optimizing PGRs for in vitro shoot proliferation of pomegranate with Bayesian-tuned ensemble stacking regression and NSGA-II: a comparative evaluation of machine learning models. Plant Methods 20:1–24. https://doi.org/10.1186/s13007-024-01211-5

Zhang H, Ma ZF (2018) Phytochemical and pharmacological properties of Capparis spinosa as a medicinal plant. Nutrients 10:116. https://doi.org/10.3390/nu10020116

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.