

RESEARCH

Open Access



# Optimization of preparation conditions for *Salsola laricifolia* protoplasts using response surface methodology and artificial neural network modeling

Hao Guo<sup>1,2,3†</sup>, Yuxin Xi<sup>1,4,5†</sup>, Kuerban Guzailinuer<sup>1,5,6</sup> and Zhibin Wen<sup>1,7\*</sup>

## Abstract

**Background** *Salsola laricifolia* is a typical C<sub>3</sub>–C<sub>4</sub> typical desert plant, belonging to the family Amaranthaceae. An efficient single-cell system is crucial to study the gene function of this plant. In this study, we optimized the experimental conditions by using Box-Behnken experimental design and Response Surface Methodology (RSM)–Artificial Neural Network (ANN) model based on the previous studies.

**Results** Among the 17 experiment groups designed by Box-Behnken experimental design, the maximum yield ( $1.566 \times 10^6/100$  mg) and the maximum number of viable cells ( $1.367 \times 10^6/100$  mg) were obtained in group 12, and the maximum viability (90.81%) was obtained in group 5. Based on these results, both the RSM and ANN models were employed for evaluating the impact of experimental factors. By RSM model, cellulase R-10 content was the most influential factor on protoplast yield, followed by macerozyme R-10 content and mannitol concentration. For protoplast viability, the macerozyme R-10 content had the highest influence, followed by cellulase R-10 content and mannitol concentration. The RSM model performed better than the ANN model in predicting yield and viability. However, the ANN model showed significant improvement in predicting the number of viable cells. After comprehensive evaluation of the protoplast yield, the viability and number of viable cells, the optimal results was predicted by ANN yield model and tested. The amount of protoplast yield was  $1.550 \times 10^6/100$  mg, with viability of 90.65% and the number of viable cells of  $1.405 \times 10^6/100$  mg. The corresponding conditions were 1.98% cellulase R-10, 1.00% macerozyme R-10, and 0.50 mol L<sup>-1</sup> mannitol. Using the obtained protoplasts, the reference genes (*18SrRNA*, *β-actin* and *EF1-α*) were screened for expression, and transformed with PEG-mediated *pBI121-SaNADP-ME2-GFP* plasmid vector. There was no significant difference in the expression of *β-actin* and *EF1-α* before and after treatment, suggesting that they can be used as internal reference genes in protoplast experiments. And *SaNADP-ME2* localized in chloroplasts.

**Conclusion** The current study validated and evaluated the effectiveness and results of RSM and ANN in optimizing the conditions for protoplast preparation using *S. laricifolia* as materials. These two methods can be used independently of experimental materials, making them suitable for isolating protoplasts from other plant materials. The selection of the number of viable cells as an evaluation index for protoplast experiments is based on its ability to consider

<sup>†</sup>Guo Hao and Xi Yuxin have contributed equally to this work.

\*Correspondence:

Zhibin Wen

zhibinwen@ms.xjb.ac.cn

Full list of author information is available at the end of the article



both protoplast yield and viability. The findings of this study provide an efficient single-cell system for future genetic experiments in *S. laricifolia* and can serve as a reference method for preparing protoplasts from other materials.

**Keywords** *Salsola laricifolia*, Protoplasts, Response surface methodology, Artificial neural network, Number of viable cells

## Background

Protoplast is a general term for a variety of substances in plant cells other than the cell wall removed, which has the totipotency as well as a part of the viability of ordinary plant cells [1], and have been extensively utilized in various experiments, such as gene promoter screening, verification of exogenous gene function, subcellular localization of proteins, protein interactions and multiplex genome editing [2–5]. Protoplast systems of model plants, such as *Arabidopsis thaliana* and *Nicotiana tabacum*, have been extensively used in gene function studies. However, it is important to note that due to the genetic variations among different plant protoplasts, these systems may provide inaccurate expression information for exogenous genes [6]. Therefore, in recent years, many non-model plants have also developed their own protoplast preparation systems to verify and explore gene functions. Some examples include *Apium graveolens*, *Camellia sinensis*, *Gossypium spp*, *Panicum virgatum*, *Populus tomentosa* [7–12].

Protoplast preparation was influenced by plant materials, sampling locations, and growth time [13, 14]. The amount of protoplast yield in *Oryza sativa* callus was  $2.500 \times 10^6$ /g FW through a 3 h enzymatic digestion. The digestion process involved the use of 1.50% cellulase, 1.00% macerozyme, and 0.60 mol L<sup>-1</sup> mannitol [3]. In the case of *Colobanthus quitensis* leaves, enzymatic digestion for 3 h was carried out. The digestion solution contained 3.00% cellulase RS, 1.20% macerozyme R-10, 1.50% viscozyme, and 0.50 mol L<sup>-1</sup> mannitol, resulting in a protoplast yield of  $8.7 \times 10^5$ /g FW [15]. Protoplasts were prepared using *Camellia oleifera* leaves and subjected to a 10 h digestion with 1.50% cellulase R-10, 0.50% macerozyme R-10, and 0.25% snailase, resulting in a maximum yield of  $3.500 \times 10^7$ /g FW [16]. However, the digestion time used for protoplast preparation of *C. oleifera* petals was reduced to 8 h, and the digestion solution contained 3.00% cellulase R-10, 1.00% macerozyme R-10, and 0.50 mol L<sup>-1</sup> mannitol, resulting in a maximum yield of  $1.42 \times 10^6$ /100 mg [10]. Furthermore, the yield of protoplast preparation in *Zea mays* root was significantly increased by incorporating a buffer pretreatment step [17].

RSM (Response Surface Methodology) and ANN (Artificial Neural Network) models are commonly employed for data analysis and prediction in optimization

experiments across various domains, such as food [18, 19] and materials [20, 21]. The RSM model utilizes a second-order polynomial model:

$$g^0(x) = a + \sum b_i x_i + \sum c_i x_i^2 \quad i = 1, 2, 3, \dots n(\text{Number of basic variables})$$

to study the relationship between one or more response variables and a number of independent variables, allowing for intuitive analysis of the influence of experimental factors on the results [22, 23]. Central composite, Doehlert and Box-Behnken are three main response surface design methods [24, 25] where Box-Behnken design is based on the multivariate optimization technique of three-level incomplete factorial design, which, due to its absence of axial points, helps to reduce the cost of the experiment and the number of experimental runs number of times, allowing experiments to be done more efficiently and economically [26, 27]. The ANN model is a machine learning technique that utilizes biological information to construct mathematical models for predicting the output of new datasets. Unlike RSM model, the ANN model does not require special experimental requirements or additional workload for its application [28]. Additionally, the ANN model often provides optimality-seeking results that are more aligned with experimental needs [29, 30]. One of the most commonly used ANN models is the Back-Propagation Neural Network (BPNN), which adjusts the thresholds and weights in the opposite direction of the conventional method during the learning process [31, 32]. The simultaneous use of RSM and ANN models can complement each other's strengths to better analyze and predict data [33, 34].

*Salsola laricifolia* belongs to the family Amaranthaceae [35]. The genus *Salsola* encompasses a diverse range of photosynthetic evolutionary types [36–38]. *S. laricifolia* is classified as a type I C<sub>3</sub>-C<sub>4</sub> intermediate plant [39]. Studying the physiological and biochemical properties of this plant type, as well as its gene functions, is crucial for understanding the evolution of the C<sub>4</sub> photosynthetic pathway and the mechanism of photorespiration reduction [40–42]. Xi et al. [43] explored the preparation conditions of protoplasts of *S. laricifolia* using the orthogonal method, and have preliminarily determined the influencing factors and experiment ranges. Although obtaining a maximum yield of protoplasts of  $1.210 \times 10^6$ /100 mg

and viability of 85.00%, but the results was not stable, and could not satisfy the experimental needs. Therefore, in this study we employed the Box-Behnken experimental design and RSM-ANN model to optimize the preparation system of *S. laricifolia* protoplasts. Additionally, reference genes were chosen and the subcellular localization of the NADP-malic enzyme gene 2 (*Sa*NADP-ME2) was determined for the prepared protoplasts. These steps ensured the continuation of gene research.

## Materials and methods

### Experimental materials and growth conditions

*S. laricifolia* seeds were collected from Bole, Xinjiang, China, in October 2022 and stored in a refrigerator at 4 °C. The seeds were soaked in sterile water for 1.5 h, followed by washing with a sodium hypochlorite solution (sterile water: sodium hypochlorite = 7: 3) for 7 min. During the washing process, the seeds were shaken well and rinsed 10–15 times with sterile water. After that, the seeds were treated with thiram (1 g of thiram dissolved in 20 mL of ddH<sub>2</sub>O) for 30 s. Subsequently, the seeds were placed on a culture medium (1/2 MS+15 g L<sup>-1</sup> sucrose+8 g L<sup>-1</sup> agar) and incubated in an artificial climate chamber. The chamber was maintained at a temperature of 25 °C for 14 h during the daytime and 18 °C for 10 h during the nighttime. The light intensity was set at 6000 Lux. After 3 days of germination, healthy and uniform size of seedlings were selected and transplanted into plastic pots with dimensions of 8.50 cm (height) × 10.00 cm (inner diameter) for hydroponic cultivation. The culture solution used was Hoagland's nutrient solution and changed every 3 days. The protoplast preparation experiment was conducted after 25 days.

### Reagent preparation

The reagents including mannitol, cellulase R-10, macerozyme R-10 (Solarbio, Beijing, China), MES, BSA (Sangon Biotech, Shanghai, China) were used. The formulations of pretreatment solution, enzyme digestion solution, W5 solution, WI solution, MMG solution and PEG solution were formulated as shown in Additional file 1: Table S1, and all solutions were sterilized by 0.20 µm filter. The pretreatment and enzymatic digestion solutions were prepared for use now, and when prepared, MES and mannitol were added first, and then mixed in a 55 °C water bath for 10 min, and then BSA, CaCl<sub>2</sub> and KCl were added sequentially after cooling down to room temperature; At last, cellulase and macerozyme were added. The PEG solution was required to be prepared for use now, and the WI and MMG solutions could be stored in the refrigerator at 4 °C for 1 week after preparation.

## Experimental methods

### Isolation and purification of protoplasts

The steps of protoplast preparation were referred to Wang et al. [44] and Xi et al. [43] with some modifications. Using 100 mg leaves were longitudinally dissected from the middle using a sterile scalpel in an ultra-clean bench, then cut into 2–3 mm segments and placed in a pretreatment solution for 30 min. Subsequently, the segments were transferred to a 5.00 mL centrifuge tube containing 1.00 mL of enzyme digestion solution (previously heated in a 50 °C water bath for 1 min). After 4 h of enzyme digestion at 26 °C for 40 r min<sup>-1</sup> avoiding light, the enzyme was gently inverted several times to fully release the protoplasts and placed on ice for spare time.

The W5 solution was pre-wetted on a 200 mesh sieve, after which the digest was filtered. The enzyme reaction was terminated by adding an equal volume of W5 solution to the enzyme filtrate (the pipette gun tip was cut with a 0.50 cm tip and the incision sharp filaments were removed over a flame), centrifuge the supernatant at 300 rpm min<sup>-1</sup> for 2 min at 4 °C and discard it, add 1 mL of W5 solution into the precipitate to resuspend the protoplasts at the bottom, and then discard the supernatant after 10 min of static time on the ice; Add 2.00 mL of W5 solution to resuspend the bottom protoplasts, leave on ice for 40 min, discard the supernatant, add 200 µL of W5 to obtain the suspension of *S. laricifolia* protoplasts, and put it in the refrigerator at 4 °C for spare. Protoplast yield was calculated using hemocytometer counting method. Evans blue staining solution (0.25%, W/V) was used to dip the protoplast suspension (ratio 2: 5) for 5 min to calculate protoplast viability.

### Box-Behnken experimental design

A Box-Behnken design of experiments using Design-Expert V8.0.6.1 software was used with protoplast yield × 10<sup>6</sup>/100 mg (Y<sub>1</sub>) and viability % (Y<sub>2</sub>) as response values, cellulase R-10 concentration % (W/V) as A, macerozyme R-10 concentration % (W/V) as B and mannitol concentration (mol L<sup>-1</sup>) as C. The factors in this experiment were designed as three levels, A (1.00, 2.00, 3.00), B (0.50, 0.75, 1.00), and C (0.50, 0.60, 0.70). The number of viable cells was calculated by multiplying the protoplast yield with the viability results obtained.

### Comparison of RSM-ANN model

The results obtained from Box-Behnken design were analyzed by RSM and ANN models using Design Expert and Matlab R2023a, respectively, to evaluate the effect of model fitting and obtain the optimal values. The results of regression equations Y<sub>1</sub> and Y<sub>2</sub> obtained from the RSM model were plotted on the three-dimensional response

surface plots using Origin 2021. By examining the interaction density and steepness of the curves in the plots, the magnitude of the effect of experimental factors on the response values was assessed. Typically, a steeper response surface curve indicates a greater impact of the factor interaction on the response value, while a less steep curve suggests a smaller effect [45–47]. The ANN model was analyzed using BPNN. It comprised an input layer of three neurons, an output layer of one neuron, and a variable number of hidden neurons layer, and was trained using different learning algorithms to select the set with the best fitted data, and the results of the experiments were used 70.00% for training, 15.00% for validation, and 15.00% for testing. The experimental optimums were solved by the simulation model of the BPNN using the GA in Matlab [48, 49].

After obtaining the models, their predictions were compared with the actual results, and several metrics were calculated to evaluate the goodness of fit of each model. These metrics included the correlation coefficient (R), the coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and the mean absolute percentage error (MAPE). A higher value of R and  $R^2$  indicates a stronger correlation and better fitting ability of the model [50, 51]. Smaller values of RMSE and MAPE indicate a lower deviation and relative error between the predicted and actual values, indicating the higher accuracy in the model's predictions [33, 52].

#### **RNA extraction and gene expression analysis**

After enzymatic digestion of protoplasts, the protoplasts were centrifuged at 300 rpm  $\text{min}^{-1}$  for 2 min at 4 °C. RNA was extracted immediately after removing the supernatant using the MiniBEST Plant RNA Extraction Kit (Takara, Japan). The entire extraction process was performed on ice. Three sets of replicates were performed for each experiment, and each set of replicates included three technical replicates. The quality of the total RNA was assessed using the NanoDrop 2000 Ultra-Micro Spectrophotometer (Thermo, America). Subsequently, the total RNA was reverse transcribed into cDNA using the PrimeScript™ RT reagent kit with gDNA-Eraser kit (Takara, Japan). The resulting cDNA products were either used directly for quantitative Real-time Polymerase Chain Reaction (qPCR) or stored at – 20 °C.

Referring to the study of Wen and Zhang [53], three housekeeping genes (*18S rRNA*,  *$\beta$ -actin* and *EF1- $\alpha$* ) were selected from *S. laricifolia*, and the comparison in expression of these genes between protoplasts and untreated leaves were analyzed by using qPCR. The qPCR primers for each housekeeping gene were in Additional file 1: Table S2, and the reaction systems were provided in Additional file materials. The relative

expression of each housekeeping gene was calculated using the formula  $Q = E^{-\Delta\Delta Ct}$ , where E represents the gene amplification efficiency, typically assumed to be two (100.00% efficiency).  $\Delta Ct$  is calculated as  $Ct(\text{min}) - Ct(\text{sample})$ , where  $Ct(\text{min})$  is the lowest  $Ct$  value among all samples, and  $Ct(\text{sample})$  is the  $Ct$  value of each sample [54, 55]. The data were subjected to two-sample anova and plotted using Origin.

#### **Exogenous gene transformation**

The *pBI121-SaNADP-ME2* and *pBI121-GFP* null-loaded *E. coli* strains were obtained from our laboratory and contained kanamycin resistance genes and *GFP* reporter gene. Refer to Ren et al. [56] for the procedure of protoplast transformation and improve it. Protoplasts were used as receptors and were first resuspended in 1 mL of MMG solution, then incubated on ice for 30 min and centrifuged at 300 rpm  $\text{min}^{-1}$  for 2 min to remove the supernatant. 600  $\mu\text{L}$  of protoplast solution was pipetted into a 5 mL centrifuge tube. 60  $\mu\text{L}$  of *pBI121-SaNADP-ME2-GFP* plasmid vector was added to the bottom of the tube and the tube was gently flicked to mix the contents. The tube was then inverted several times to mix the 660  $\mu\text{L}$  of PEG solution. The mixture was incubated in a dark environment at room temperature for 30 min, completing the transformation process. The transformation reaction was stopped by adding a twofold volume of W5 solution. After centrifugation at 300 rpm  $\text{min}^{-1}$  for 2 min, the supernatant was discarded. The protoplasts were then resuspended in 3 mL of WI solution and placed in cell culture chambers (Labsselect, Beijing, China). They were incubated in the dark at room temperature for 16–24 h. A suitable amount of the transformed protoplasts was used for further analysis. Protoplasts transformed with *pBI121-GFP* empty load were used as a control. The images were captured using a laser confocal scanning microscope (Zeiss LSM 800, Jena, Germany) and processed with Zen 2012 software. GFP and chlorophyll were excited using 488 nm and 633 nm laser lines, respectively, to observe GFP expression.

## **Results**

### **Analysis of protoplasmic preparation results**

Seventeen groups of experiments were conducted using the Box-Behnken design, and the results obtained were presented in Table 1. Among the 17 groups, group 12 exhibited the highest yield with a value of  $1.566 \times 10^6/100 \text{ mg}$ . Group 5 showed the highest viability at 90.81%, while group 12 had the maximum number of viable cells, reaching  $1.367 \times 10^6/100 \text{ mg}$ .

**Table 1** Experimental results of three-factor, two-response-value Box-Behnken design test based on the preparation of protoplasts of *Salsola laricifolia*

Group	Cellulase R-10 content (%)	macerozyme R-10 content (%)	Mannitol concentration (mol·L <sup>-1</sup> )	Protoplast yield (10 <sup>6</sup> /100 mg)	Protoplast viability (%)	Number of viable cells (yield*vitality) (10 <sup>6</sup> /100 mg)
1	1.00	0.50	0.60	0.575	85.78	0.493
2	3.00	0.50	0.60	0.256	85.58	0.219
3	1.00	1.00	0.60	0.683	88.02	0.601
4	3.00	1.00	0.60	0.800	78.69	0.630
5	1.00	0.75	0.50	0.563	90.81	0.511
6	3.00	0.75	0.50	0.383	84.65	0.324
7	1.00	0.75	0.70	0.583	87.62	0.511
8	3.00	0.75	0.70	0.615	84.27	0.518
9	2.00	0.50	0.50	1.125	90.10	1.013
10	2.00	1.00	0.50	1.460	90.62	1.323
11	2.00	0.50	0.70	1.320	89.23	1.178
12	2.00	1.00	0.70	1.566	87.27	1.367
13	2.00	0.75	0.60	0.955	89.63	0.856
14	2.00	0.75	0.60	0.945	89.42	0.845
15	2.00	0.75	0.60	1.035	88.25	0.913
16	2.00	0.75	0.60	1.090	88.57	0.965
17	2.00	0.75	0.60	1.010	87.21	0.881

**Table 2** Response surface model evaluation indexes

Evaluation metrics	Protoplast yield model	Protoplast viability model
F-value	105.270	21.080
P-value	< 0.0001	0.0003
R <sup>2</sup>	0.993	0.964
R <sup>2</sup> of experiment value	0.983	0.919
R <sup>2</sup> of predicted value	0.976	0.815
Coefficient of variation (C.V. %)	5.46	0.98
signal-to-noise ratio	35.072	18.012

**RSM results analysis**

**RSM regression model analysis**

The test results were analyzed using RSM to establish the response values Y<sub>1</sub> and Y<sub>2</sub> based on the experimental factors A, B, and C. Quadratic multinomial regression equations was obtained. The model evaluation indexes were in Table 2, and the detailed results of the model were provided in Additional file 1: Tables S3, S4.

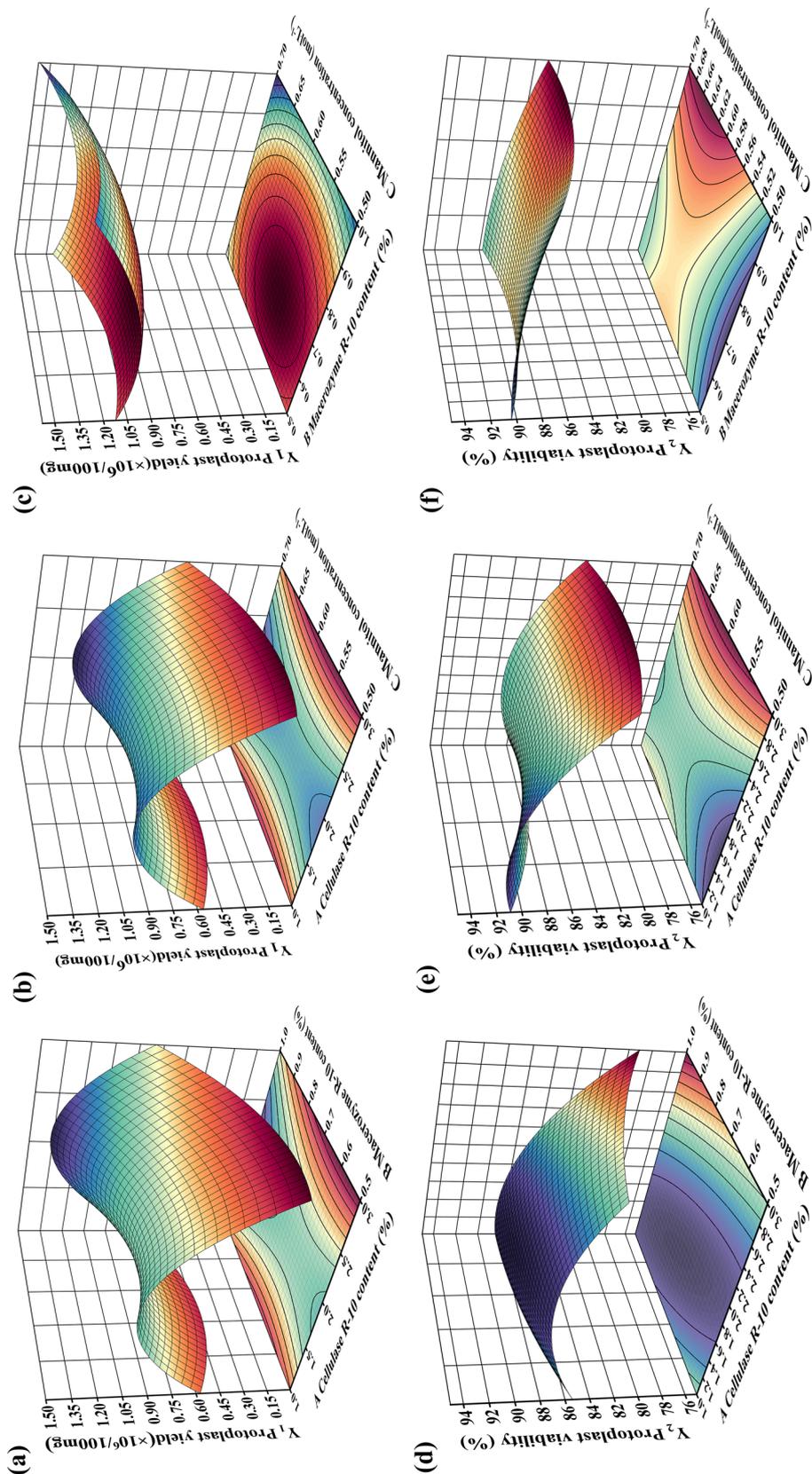
$$Y_1 = 6.119 + 1.815A - 4.626B - 18.608C + 0.436AB + 0.580AC - 0.887BC - 0.633A^2 + 3.269B^2 + 15.650C^2$$

$$Y_2 = 18.883 + 13.360A + 49.779B - 186.192C - 9.063AB + 7.010AC - 24.790BC - 3.289A^2 - 13.194B^2 + 150.865C^2$$

The yield model (Y<sub>1</sub>) had an F-value of 105.270, indicating highly significant level differences (p < 0.0001). The R<sup>2</sup> value was 0.993, and the signal-to-noise ratio was 35.072, which was greater than 10, indicating that the model was sufficiently accurate and not easily affected by external disturbances. Additionally, the data had an acceptable level of variability, as evidenced by the range of 4% to 10% for the coefficient of variation (C.V. %). The viability model (Y<sub>2</sub>) had an F-value of 21.080 (p = 0.0003), which was less than 0.010, indicating a significant model variance. The R<sup>2</sup> value was 0.964, and the signal-to-noise ratio was 18.012, indicating a high level of accuracy and resistance to external disturbances. Additionally, the C.V.% value was less than 4.00, suggesting a low degree of data variability in the model.

**RSM model three-dimensional result graph analysis**

From Fig. 1a, b and c, it can be observed that the interaction effects of cellulase R-10 content (A) and macerozyme R-10 content (B) were the most significant. The influence of A on protoplast yield (Y<sub>1</sub>) was greater than that of B, which was manifested by the fact that the surface intersecting AB was the steepest, and the contours intersecting with the A-axis were more densely packed



**Fig. 1** Three-dimensional response surface plots based on the RSM model resolving the interaction of various factors of the experiment on the results of protoplast preparation of *Salsola laricifolia*. The plots labeled **a**, **b**, and **c** represent the protoplast yield (Y<sub>1</sub>) and show how it is influenced by the interaction between two experimental factors. Similarly, the plots labeled **d**, **e**, and **f** represent protoplast viability (Y<sub>2</sub>) and how it is influenced by the interaction between two experimental factors. X-axis, Y-axis are the experimental factors including cellulase R:10 content, macerozyme R:10 content, and mannitol concentration, denoted by A, B, and C, respectively

than those of the B-axis (Fig. 1a), the interaction between A and mannitol concentration (C) had more significant effect on  $Y_1$ . A had a greater effect on  $Y_1$  than C (Fig. 1b), where the surface was smoother and the contours intersecting the A-axis were denser than those of the C-axis. The interaction between B and mannitol concentration (C) had a significant effect, and B had a greater effect on  $Y_1$  than C (Fig. 1c), where the surface was smoothest and the contours intersecting the B-axis were denser than those of the C-axis. Therefore, it can be concluded that the effects of the experimental factors A, B, and C on  $Y_1$  following the order of  $A > B > C$ . Additionally, the effect of the interaction between these factors on  $Y_1$  followed the order of  $AB > AC > BC$ .

Similarly, as shown in Figs. 1d, 1e, and 1f, the experimental factors on the protoplast preparation viability ( $Y_2$ ) resulted in the order of the influence of  $B > A > C$ , the interaction between the factors on the degree of influence of  $Y_2$  in the order of  $AB > BC > AC$ .

#### ANN result analysis

The mean square values, error values, and simulation effects of the ANN model were shown in Fig. 2. The yield model was optimized using the Levenberg–Marquardt algorithm with 18 hidden layers. The vitality model was optimized using the scaled conjugate gradient method with 17 hidden layers. The viable cell model was optimized using the Levenberg–Marquardt algorithm with 14 hidden layers.

The correlation coefficient of the yield model obtained after training the ANN model was 0.96736. The end-of-training condition was met after 5 iterations, with a mean square error (MSE) of 0.045 (Figs. 2, 3a). The correlation coefficient of the ANN-viability model was 0.977, and the MSE of the ANN-viability model was  $8.5388 \times 10^{-5}$  after 10 iterations, satisfying the end-of-training condition (Figs. 2, 3b). The correlation coefficient of the ANN-vital cell model was 0.995, and its mean square error was 0.003 after four iterations, also satisfying the end-of-training condition (Figs. 2, 3c).

#### Comparison between RSM and ANN model fitting and prediction performance

##### Fitting effect of RSM, ANN

According to Fig. 4, the RSM yield and viability model had a better fit compared to the ANN yield and viability model. The R and  $R^2$  values of the RSM model were 0.996, 0.992, and 0.982, 0.964 respectively, which were closer to 1 than the values of the ANN model. Additionally, the RSM model had smaller RMSE and MAPE values compared to the ANN model. However, the ANN model showed improved accuracy in predicting

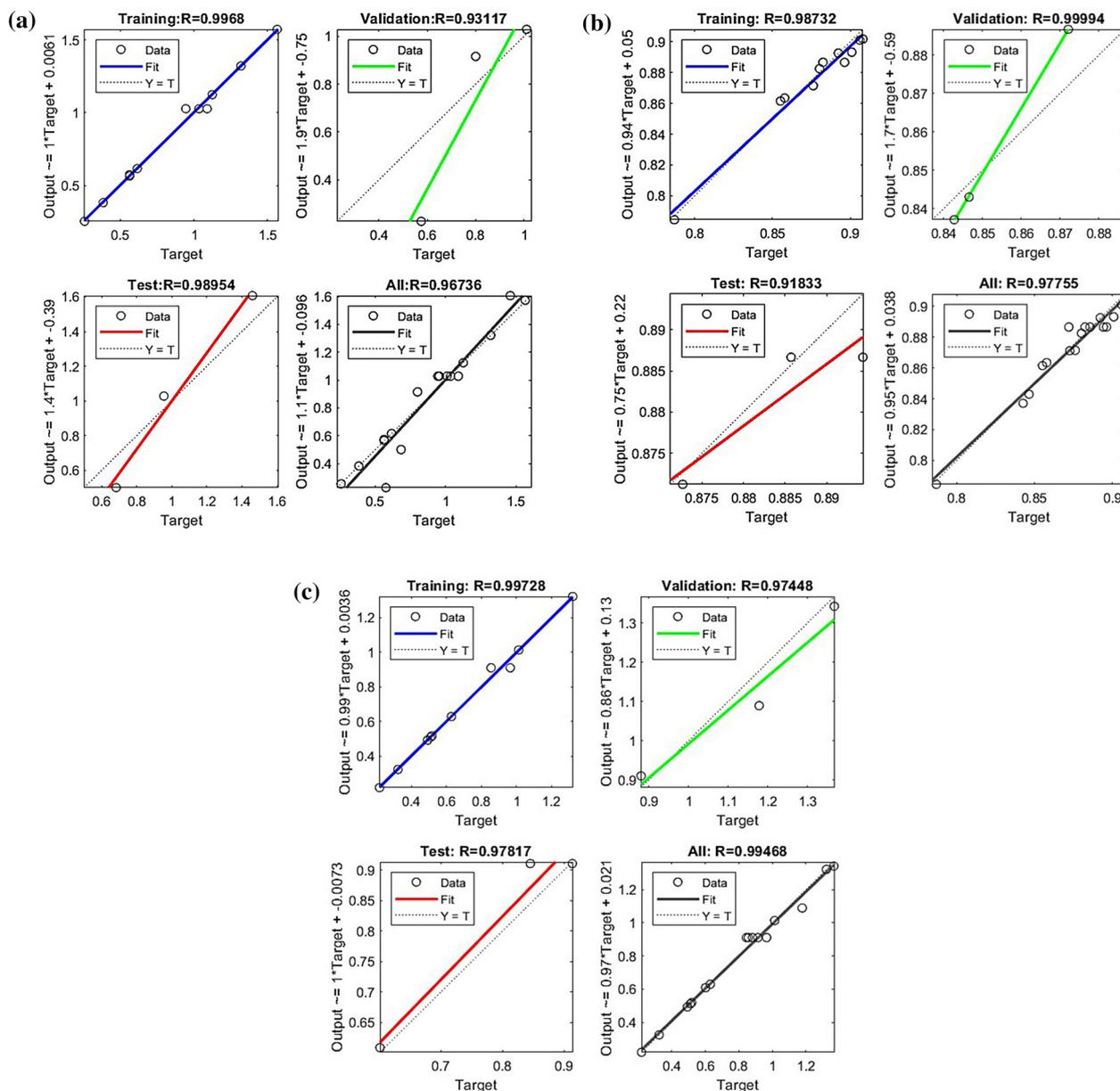
the number of live cells, with its R and  $R^2$  values being second only to the best-fitting RSM yield model (Fig. 4, Table 3).

When analyzing only yield and viability, protoplast yield fluctuated greatly by changes in experimental conditions, while viability fluctuated less. The coefficients of variation (C.V.%) of the actual and predicted values of protoplast yield were 42.203, 41.943 and 48.438, which were much larger than those of the actual and predicted values of viability, which were 3.428, 3.366, and 3.354 (Fig. 4). Consequently, the RMSE and MAPE of the predicted values for yield were higher than those for viability predictions in Table 3, indicating a greater bias in yield prediction.

##### RSM, ANN model optimization and verification

The RSM model was used to solve for the optimal values under the conditions of simultaneous consideration of the yield and viability. The results revealed that under the conditions of a cellulase R-10 content of 1.86%, macerozyme R-10 content of 1.00% and mannitol concentration of  $0.50 \text{ mol L}^{-1}$ , the theoretical values for protoplast yield were  $1.464 \times 10^6/100 \text{ mg}$  and viability was 90.81% (Table 4). GA was utilized for optimization of ANN-yield model, ANN-viability model and ANN-viable cell number model respectively. ANN-yield model predicted a theoretical maximum protoplast yield of  $1.625 \times 10^6/100 \text{ mg}$ . This prediction was based on the following experimental conditions: 1.98% cellulase R-10 content, 1.00% macerozyme R-10 content, and  $0.50 \text{ mol L}^{-1}$  mannitol concentration. Similarly, the viability model predicted a theoretical maximum viability of 96.29 based on experimental conditions including 1.65% cellulase R-10 content, 0.98% macerozyme R-10 content, and  $0.53 \text{ mol L}^{-1}$  mannitol concentration. Additionally, the ANN viable cell number model predicted a theoretical maximum viable cell number of  $1.342 \times 10^6/100 \text{ mg}$ , which was based on 2.00% cellulase R-10 content, 1.00% macerozyme R-10 content, and  $0.70 \text{ mol L}^{-1}$  mannitol concentration (Table 4). To validate these predictions, the predicted values, experimentally validated values, and relative errors of the results were calculated under the respective optimal conditions of the RSM and ANN models, and Table 4 was obtained.

From Table 4, the relative errors between the prediction results and the actual results using RSM and ANN models were less than 5.00%, indicating a high degree of confidence in the prediction results of both models [57]. Considering the protoplast yield, viability and viable cell number, the ANN yield model predicted the optimal results, and the validation resulted in a protoplast yield of  $1.550 \times 10^6/100 \text{ mg}$ , the viability of 90.65% and a viable cell number of  $1.405 \times 10^6/100 \text{ mg}$ , which corresponded to the



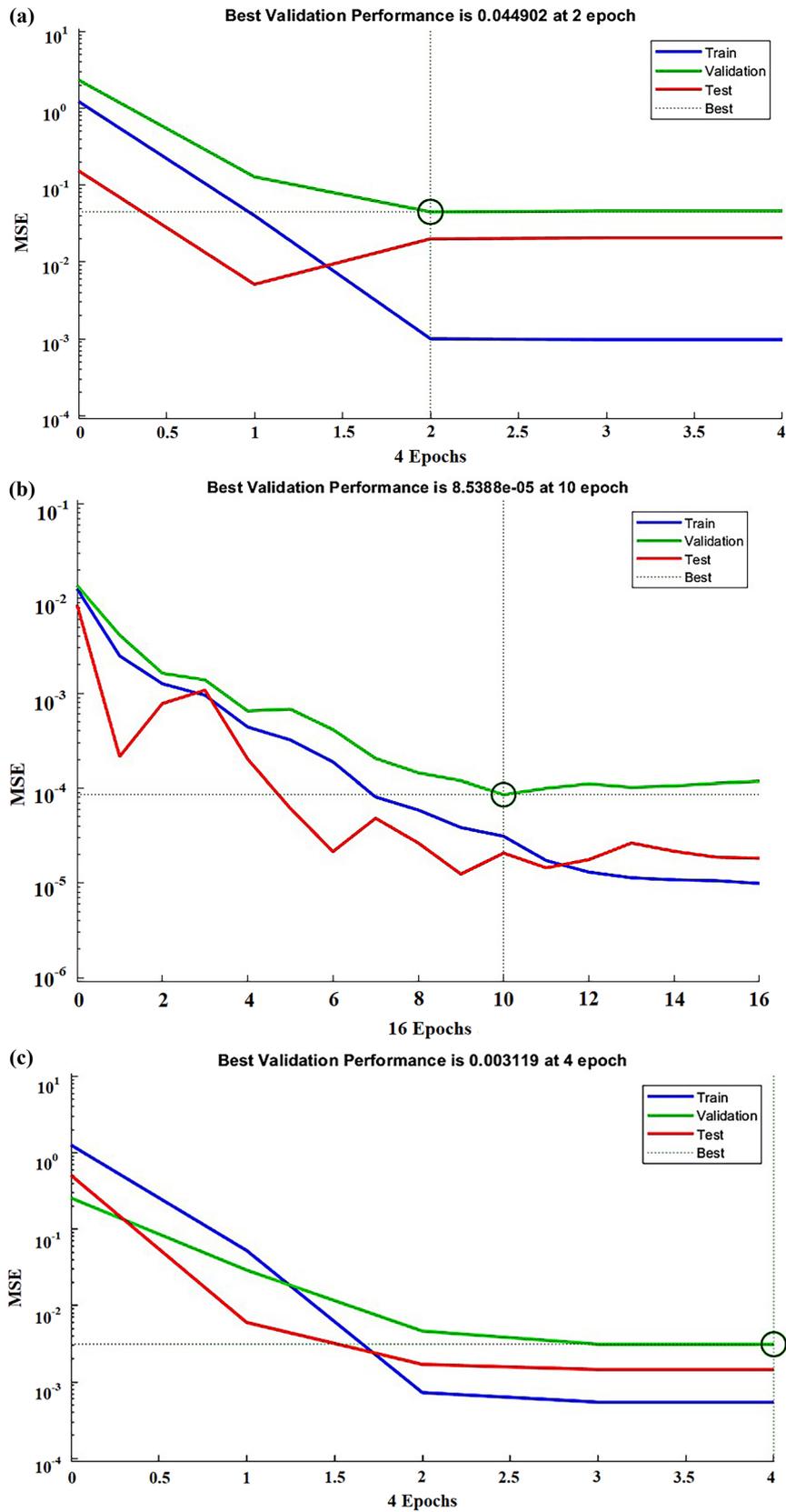
**Fig. 2** Artificial neural network (ANN) model simulation effect. It displays the training, verification, test and all data fitting for each model. The fitting effect is represented by correlation coefficient (R). **a** represents the yield model, **b** represents the viability model, and **c** represents the viable cell model

preparation conditions of 1.98% cellulase R-10 content, 1.00% macerozyme R-10 content and 0.50 mol L<sup>-1</sup> mannitol concentration. When compared to the RSM model

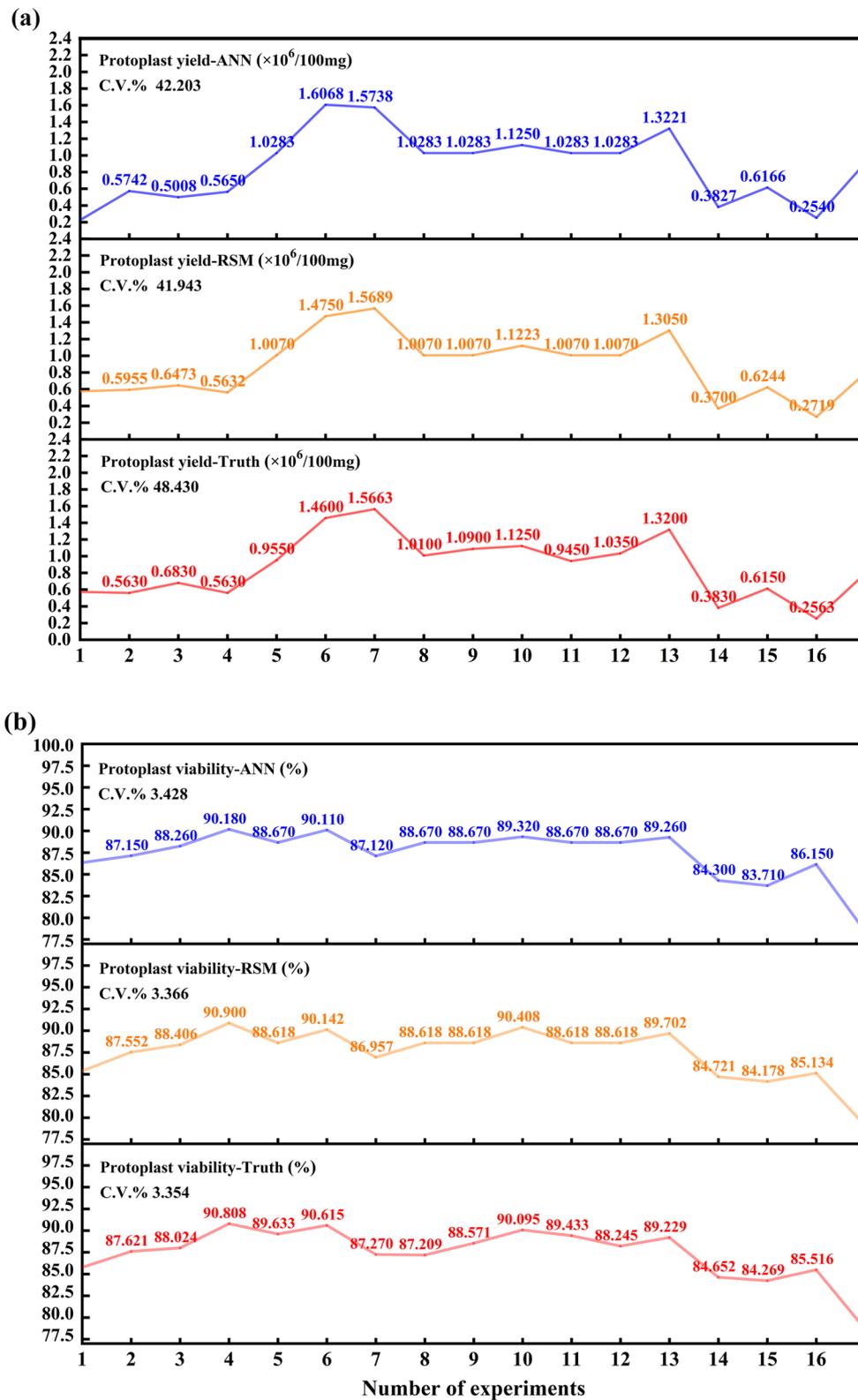
prediction validation results (1.470×10<sup>6</sup>/100 mg, 89.73%, 1.319×10<sup>6</sup>/100 mg), the indicators showed improvements of 5.44%, 1.03%, and 6.52% respectively. Figure 5 demonstrates

(See figure on next page.)

**Fig. 3** Artificial neural network (ANN) model mean square error (MSE) effect plot. The top shows the number of ANN model epochs when the minimum MSE is obtained for validation data. **a** represents the yield model, **b** represents the vigor model, and **c** represents the viable cell model



**Fig. 3** (See legend on previous page.)



**Fig. 4** Comparison of experimental values with predicted values of response surface methodology (RSM), artificial neural network (ANN) models. **a** for protoplast yield, **b** for protoplast viability

**Table 3** Evaluation of response surface methodology (RSM), artificial neural network (ANN) model based on statistical indexes

Statistical indexes	RSM-yield	RSM-viability	ANN-yield	ANN-viability	ANN-number of viable cells
R	0.996	0.982	0.967	0.977	0.994
R <sup>2</sup>	0.992	0.964	0.935	0.955	0.989
RMSE	0.034	0.005	0.110	0.006	0.035
MAPE/%	2.825	0.478	8.274	0.595	2.009

R correlation coefficient, R<sup>2</sup> coefficient of determination, RMSE root mean square error, MAPE mean absolute percentage error

the results of the protoplast preparation under these conditions.

### Reference gene stability analysis

The  $A_{260}/A_{280}$  and  $A_{260}/A_{230}$  ratios of RNA in both protoplasts and untreated leaves of *S. laricifolia* were approximately 2.00, which satisfied the requirements for subsequent experiments. The relative expression of the three housekeeping genes in both protoplasts and leaves was shown in Fig. 6. There was no significant difference in the expression of  $\beta$ -actin and *EF1- $\alpha$*  pre-and post-treatments. However, there was a significant difference in the expression of *18SrRNA* ( $p < 0.05$ ), indicating that it could not be used as a reference gene in the protoplast experiment.

### Results of protoplast transformation

The results of the transformation of *S. laricifolia* protoplasts using the *pBII121-SaNADP-ME2* plasmid vector were shown in Fig. 7, where fluorescence signals of GFP and chlorophyll were detected respectively at 500–530 nm and 650–750 nm after excitation at 488 nm and 633 nm. The protoplasts transformed with *pBII121-SaNADP-ME2* plasmid vector could observe green fluorescence of GFP protein in chloroplasts (Figs. 7a, b, c, d), while protoplasts imported with empty plasmid only emitted green fluorescence in cytoplasm, with no obvious fluorescence in chloroplasts (Figs. 7e, f, g, h). The protoplast cells that were successfully transformed appeared transparent and had a regular shape without any breakage.

### Discussion

In this study, we utilized RSM and ANN models to predict the best preparation conditions for *S. laricifolia* protoplasts. Our findings revealed that the Box-Behnken design test provided the necessary data for our analysis. The RSM model demonstrated a better fit compared to the ANN model (Table 3). However, the final protoplast preparation conditions were determined using the ANN-yield model (Table 4). Because the amount of protoplast yield was fluctuated more when changing experimental conditions compared to protoplast viability, leading

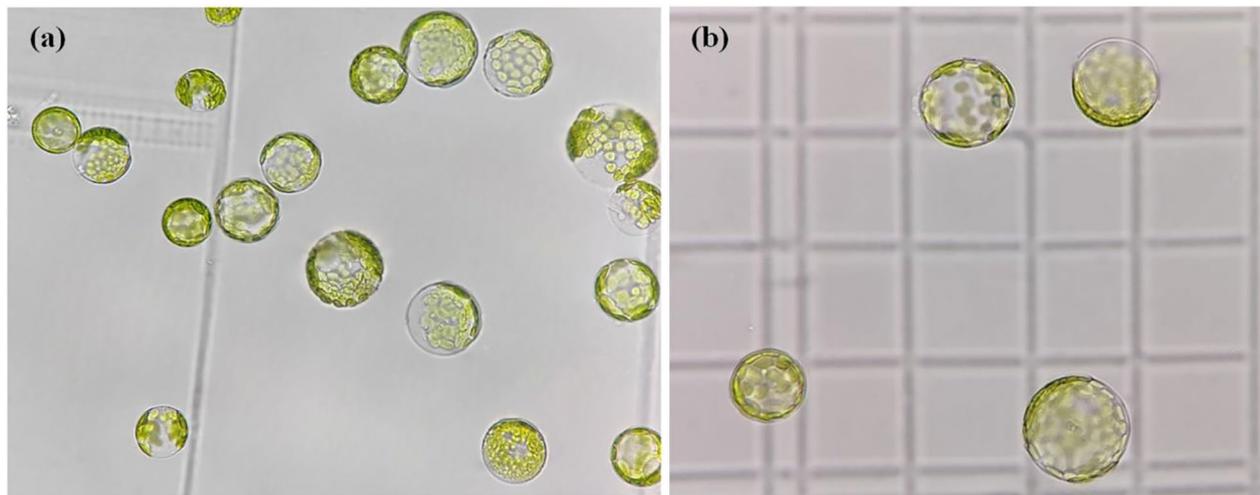
to the fact that protoplast yield tends to be more determinative of the number of viable cells (Table 3, Fig. 4). Although the ANN-yield model produced a viability value of 90.65%, slightly lower than the predicted value of the ANN-viability model (90.81%), the final number of viable cells obtained was  $1.405 \times 10^6/100$  mg, which was the highest value (Table 4). As a result, the preparation conditions predicted by the ANN-viability model were chosen as the final conditions. Compared to the previous experimental results of *S. laricifolia* protoplasts [41], the protoplast yield increased by 28.10%, the protoplast viability increased by 6.65%, and the number of viable cells increased by 36.61%. Furthermore, when compared to the best results before the ANN treatment (Table 1), there was an additional increase of 3.87% in the protoplast viability and 2.78% in the number of viable cells.

After modeling the ANN model using the number of viable cells, the fitting effect of the ANN model was significantly improved, with R, and R<sup>2</sup> were next only to the best-fitting RSM-yield model, and the relative error between the experimental validation results and the predicted results was also less than 5.00% (Tables 3, 4). Therefore, it is recommended to consider the number of living cells when using the ANN model to optimize the results of protoplast preparation. This consideration, along with the analysis of protoplast yield and viability, can help obtain predicted results that align more closely with experimental expectations. With the further application of deep learning in cell research, studies have been conducted to identify dead/surviving cells directly using convolutional neural networks [58], which makes it possible to count the number of living cells more easily. Additionally, the fitting ability of RSM and ANN models varies depending on the predicted objects [33, 52, 59–61]. Therefore, when different materials are used and steps are taken to prepare protoplasts, they need to be evaluated in detail with specific evaluation indexes and experimental data.

In this study, the expression differences of three housekeeping genes (*18sRNA*, *EF1- $\alpha$* , and  $\beta$ -actin) before and after the preparation of pine leaf pigweed protoplasts were shown in Fig. 6. We optimized the protoplast preparation system and identified suitable internal reference

**Table 4** The response surface methodology (RSM) yield-vigor model, the artificial neural network(ANN) yield model, the vigor model, and the viable cell number model predictions were validated, and the results were considered credible with a relative error of less than 5%

Model	Cellulase R-10 content (%)	Macerozyme R-10 content (%)	Mannitol concentration(mol L <sup>-1</sup> )	Protoplast yield(X 10 <sup>6</sup> /100 mg)		Protoplast viability (%)		Number of viable cells(X 10 <sup>6</sup> /100 mg)		Relative error (%)		
				Predicted value	Experiment value	Predicted value	Experiment value	Predicted value	Experiment value	Protoplast yield	Protoplast viability	Protoplast yield
RSM yield-viability	1.86	1.00	0.50	1.464	1.470	90.81	89.73	1.329	1.319	0.41	1.20	0.76
ANN-yield	1.98	1.00	0.50	1.625	1.550	90.16	90.65	1.337	1.405	4.84	0.54	4.83
ANN-viability	1.00	0.75	0.50	0.565	0.563	90.18	90.81	0.510	0.508	0.36	0.69	0.39
ANN-number of viable cells	2.00	1.00	0.70	1.574	1.566	87.12	87.27	1.342	1.367	0.51	0.17	1.83



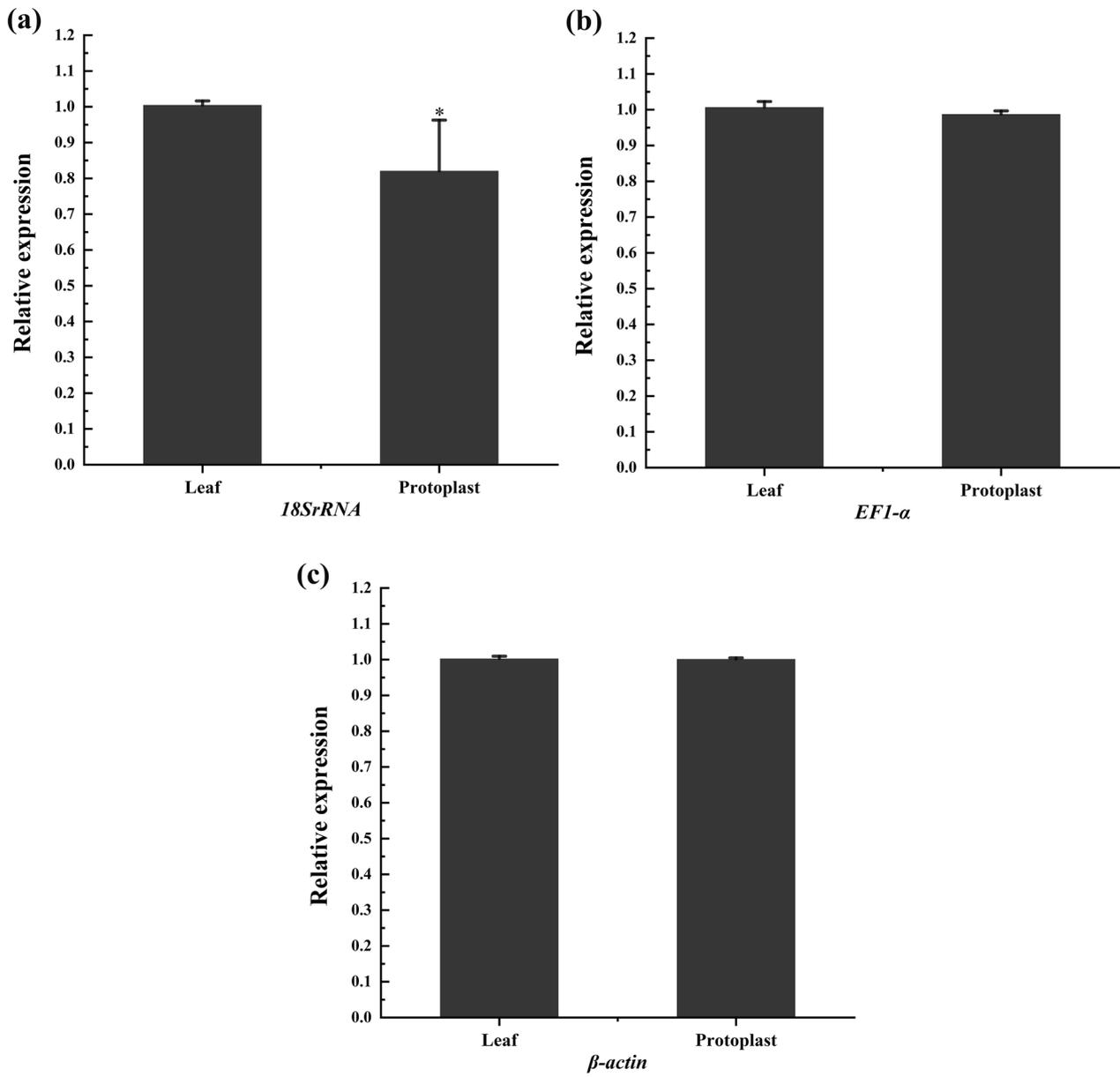
**Fig. 5** Results of the preparation of protoplasts of *Salsola laricifolia* under the optimal preparation conditions. To enhance visualization of protoplasts, **a** and **b** were captured using blood counting plates (25 × 16) as a background under a microscope (10 × 40). Additionally, the protoplast concentrations in **b** were diluted by a factor of two

genes to avoid any potential alteration of their expression in the pigweed protoplasts, which could affect the results of subsequent experiments. Housekeeping genes are essential for maintaining minimum cellular functions and are generally considered to be stably expressed [62–65]. However, their expression may be altered to varying degrees under different adverse conditions. Unlike *β-actin* and *EF1-α*, the expression of *18sRNA* selected in this experiment were found to be unstable and unsuitable for use as endogenous genes after the preparation process of *S. laricifolia* protoplasts (Fig. 5), this aligns with the findings of a previous study that investigated the suitability of these genes as endogenous genes in *S. laricifolia* leaves under drought stress [53].

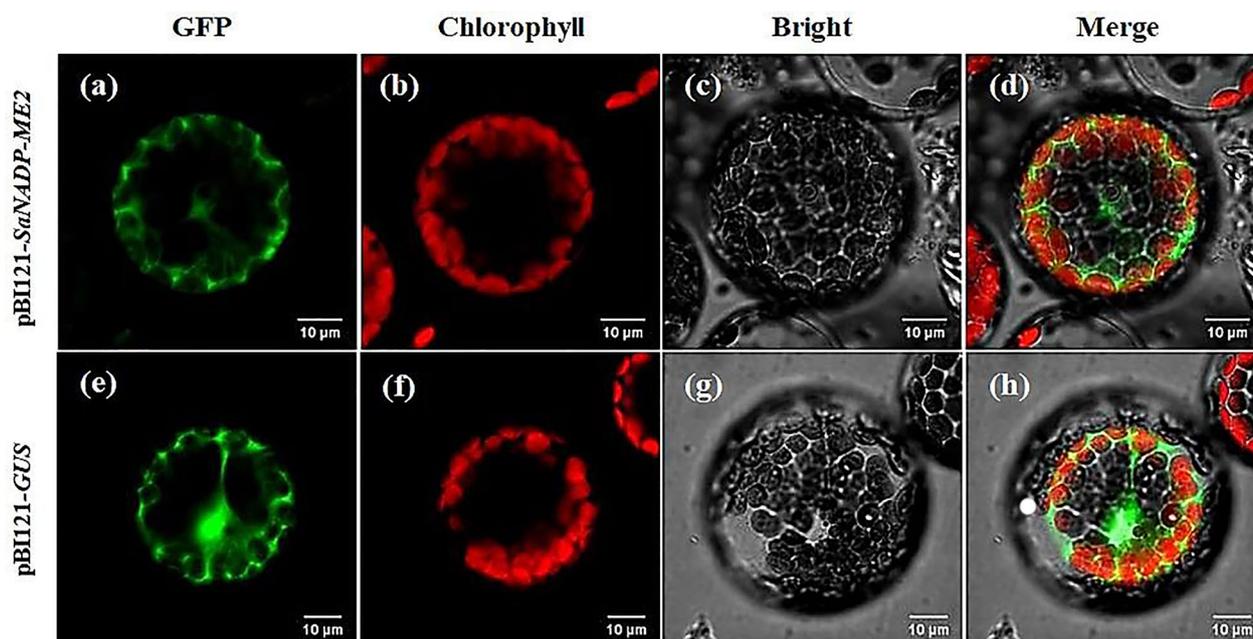
## Conclusions

In this experiment, Box-Behnken design method with RSM-ANN model was used to optimize the preparation conditions of protoplasts of *S. laricifolia*, these methods are not commonly used in previously similar studies. To validate the methodology, a series of

experiments were conducted, including transient gene conversion and internal reference gene analysis. After considering the protoplast yield, viability and number of viable cells, the ANN yield model predicted the best results, and the experimental validation protoplast yield of  $1.550 \times 10^6/100$  mg, viability of 90.65%, and a number of viable cells of  $1.405 \times 10^6/100$  mg. Corresponding preparation conditions were 1.98% cellulase R-10, 1.00% macerozyme R-10,  $0.50 \text{ mol L}^{-1}$  mannitol. Furthermore, *β-actin* and *EF1-α* were identified as internal reference genes in protoplast experiment. The experiments also demonstrated that, in addition to protoplast yield and viability, the number of viable cells (yield × viability) can serve as an evaluation index for predicting protoplast experiments using RSM and ANN. This index not only considers the viability and yield of protoplasts simultaneously but also aligns better with the requirements of subsequent genetic experiments.



**Fig. 6** Relative expression of the three housekeeping genes (a, b, c represent *18SRNA*, *EF1-α*, and *β-actin*, respectively) in the protoplasts and leaves of *Salsola laricifolia*. \* represents significant differences ( $p < 0.05$ )



**Fig. 7** Transformation results of protoplasts under laser confocal scanning microscope, green and red fluorescence indicate the localization of GFP protein and chlorophyll, respectively. **a, b, c, d:** sequentially, *pBI121-SαNADP-ME2* plasmid transformed protoplasts in GFP fluorescence, chloroplast autofluorescence, white light and three images superimposed; **e, f, g, h:** sequentially, *pBI121-GFP* empty plasmid transformed protoplasts in GFP fluorescence, chloroplast autofluorescence, white light and three images superimposed

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13007-024-01180-9>.

**Additional file 1: Table S1.** Components of experimental reagents. **Table S2.** Primer sequences of three housekeeping genes in *Salsola laricifolia*. **Table S3.** Regression analysis of response surface yield and vitality models, \*, \*\*, and \*\*\* indicate significant differences at the  $p < 0.05$ ,  $p < 0.01$ , and  $p < 0.001$  levels, respectively. **Table S4.** Fit statistics of Box-Behnken design-response surface methodology. *Salsola laricifolia* housekeeping gene reaction system in the supplementary Information.

## Acknowledgements

We are grateful to Tong Liu, Shihezi University, for providing support for this research.

## Author contributions

Conceptualization: ZW, HG, YX. Methodology: HG, KG. Software: HG. Validation: YX, KG. Investigation: HG, YX, ZW. Data Curation: HG, KG, YX. Writing—original draft: HG. Writing review and editing: HG, ZW, YX, KG. Supervision: KG, ZW. Funding acquisition: ZW.

## Funding

The research was supported by the National Natural Science Foundation of China (No. 31970354), Tianshan Talent Training Program (No. 2023TSY-CCX0090), Key Science and Technology Project of Xinjiang Production and Construction Corps (No. 2022AB010).

## Availability of data and materials

Not applicable.

## Declarations

### Ethical approval and consent to participate

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>State Key Laboratory of Desert and Oasis Ecology, Key Laboratory of Ecological Safety and Sustainable Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China. <sup>2</sup>College of Life Sciences, Shihezi University, Shihezi 832003, China. <sup>3</sup>Xinjiang Production and Construction Corps Key Laboratory of Oasis Town and Mountain-Basin System Ecology, Shihezi 832003, China. <sup>4</sup>Xinjiang Key Lab of Conservation and Utilization of Plant Gene Resources, Urumqi 830011, China. <sup>5</sup>University of Chinese Academy of Sciences, Beijing 100049, China. <sup>6</sup>Sino-Tajikistan Joint Laboratory for Conservation and Utilization of Biological Resources, Urumqi 830011, China. <sup>7</sup>The Specimen Museum of Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China.

Received: 12 December 2023 Accepted: 25 March 2024

Published online: 07 April 2024

## References

- Lei R, Qiao W, Hu F, Jiang H, Zhu S. A simple and effective method to encapsulate tobacco mesophyll protoplasts to maintain cell viability. *MethodsX*. 2015;2:24–32.
- Moon KB, Park JS, Park SJ, Lee HJ, Cho HS, Min SR, Park YI, Jeon JH, Kim HS. A more accessible, time-saving, and efficient method for in vitro plant regeneration from potato protoplasts. *Plants-Basel*. 2021;10:781.
- Poddar S, Tanaka J, Cate JHD, Staskawicz B, Cho MJ. Efficient isolation of protoplasts from rice calli with pause points and its application in

- transient gene expression and genome editing assays. *Plant Methods*. 2020;16:151–64.
4. Rather GA, Ayzenshtat D, Teper-Bamnlker P, Kumar M, Forotan Z, Eshel D, Bocobza S. Advances in protoplast transfection promote efficient CRISPR/Cas9-mediated genome editing in tetraploid potato. *Planta*. 2022;256:14.
  5. Xia K, Zhang D, Xu X, Liu G, Yang Y, Chen Z, Wang X, Zhang GQ, Sun HX, Gu Y. Protoplast technology enables the identification of efficient multiplex genome editing tools in *Phalaenopsis*. *Plant Sci*. 2022;322:111368.
  6. Ambastha V, Chauhan G, Tiwari BS, Tripathy BC. Execution of programmed cell death by singlet oxygen generated inside the chloroplasts of *Arabidopsis thaliana*. *Protoplasma*. 2020;257:841–51.
  7. Burris KP, Dlugosz EM, Collins AG, Stewart CN Jr, Lenaghan SC. Development of a rapid, low-cost protoplast transfection system for switchgrass (*Panicum virgatum* L). *Plant Cell Rep*. 2016;35:693–704.
  8. Du J, Zhang H, Li W, Li X, Wang Z, Zhang Y, Xiong A, Li M. Optimization of protoplast preparation system from leaves and establishment of a transient transformation system in *Apium graveolens*. *Agronomy-Basel*. 2023;13:2154.
  9. Jiang H, Cao Q, Wang X, Lv W, Wang Y. Pectate lyase genes abundantly expressed during the infection regulate morphological development of *Colletotrichum camelliae* and *CcPEL16* is required for full virulence to tea plants. *Mosphere*. 2023;8:677.
  10. Lin Z, Huang LJ, Yu P, Chen J, Du S, Qin G, Zhang L, Li N, Yuan D. Development of a protoplast isolation system for functional gene expression and characterization using petals of *Camellia Oleifera*. *Plant Physiol Biochem*. 2023;201:107885.
  11. Wang H-L, Wang T, Yang Q, Yin W, Xia X, Guo H, Li Z. Preparation and transfection of *Populus tomentosa* mesophyll protoplasts. *Bio-Protoc*. 2021;11: e4220.
  12. Xu XF, Zhu HY, Ren YF, Feng C, Ye ZH, Cai HM, Wan XC, Peng CY. Efficient isolation and purification of tissue-specific protoplasts from tea plants (*Camellia sinensis* (L) O Kuntze). *Plant Methods*. 2021;17:84.
  13. Eeckhaut T, Lakshmanan PS, Deryckere D, Van Bockstaele E, Van Huylenbroeck J. Progress in plant protoplast research. *Planta*. 2013;238:991–1003.
  14. Naing AH, Adedeji OS, Kim CK. Protoplast technology in ornamental plants: current progress and potential applications on genetic improvement. *Sci Hortic*. 2021;283: 110043.
  15. Cha OK, Lee J, Lee HS, Lee H. Optimized protoplast isolation and establishment of transient gene expression system for the antarctic flowering plant *Colobanthus quitensis* (Kunth) Bartl. *Plant Cell Tissue Organ Culture (PCTOC)*. 2019;138:603–7.
  16. Li S, Zhao R, Ye T, Guan R, Linjie X, Ma X, Zhang J, Xiao S, Yuan D. Isolation, purification and PEG-mediated transient expression of mesophyll protoplasts in *Camellia oleifera*. *Plant Methods*. 2022;18:141.
  17. Ortiz-Ramirez C, Arevalo ED, Xu X, Jackson DP, Birnbaum KD. An efficient cell sorting protocol for maize protoplasts. *Current Protocols Plant Biol*. 2018;3:e20072–e20072.
  18. Kumar S, Pipliya S, Srivastav PP. Effect of cold plasma processing on physicochemical and nutritional quality attributes of kiwifruit juice. *J Food Sci*. 2023;88:1533–52.
  19. Pradhan D, Abdullah S, Pradhan RC. Chironji (*Buchanania lanzan*) fruit juice extraction using cellulase enzyme: modelling and optimization of process by artificial neural network and response surface methodology. *J Food Sci Technol-Mysore*. 2021;58:1051–60.
  20. Belaadi A, Boumaaza M, Alshahrani H, Bourchak M. Effect of jute fiber length on drilling performance of biocomposites: optimization comparison between RSM, ANN, and genetic algorithm. *Int J Adv Manuf Technol*. 2023;124:3579–99.
  21. Khetib Y, Sedraoui K, Gari A. Improving thermal conductivity of a ferrofluid-based nanofluid using Fe<sub>3</sub>O<sub>4</sub>-challenging of RSM and ANN methodologies. *Chem Eng Commun*. 2022;209:1070–81.
  22. Oyedoh EA, Erumi GA, Akhabue CE. Comparative study of response surface methodology and artificial neural network for modeling and optimization of extraction process parameters on *Tetrapleura tetraptera*. *J Appl Sci Environ Manag*. 2020;24:313–21.
  23. Su W, Yang H, Tan Y, Zhou Z, Qin Y, Li X. Optimization of protoplast formation of *Auricularia cornea* by response surface methodology. *J Southern Agric*. 2020;51:169–75.
  24. Cazetta AL, Spessato L, Almeida VC. The use of chemometric tools for screening and optimization of variables in the preparation and application of carbon-based materials. *J Taiwan Inst Chem Eng*. 2021;121:321–36.
  25. Ferreira Mata Cerqueira UM, Bezerra MA, Costa Ferreira SL, Araujo Rdd, da Silva BN, Novaes CG. Doehlert design in the optimization of procedures aiming food analysis—a review. *Food Chem*. 2021;364:130429.
  26. Elleuch L, Messaoud M, Djebali K, Attafi M, Cherni Y, Kasmi M, Elaoud A, Trabelsi I, Chatti A. A new insight into highly contaminated landfill leachate treatment using Kefir grains pre-treatment combined with Ag-doped TiO<sub>2</sub> photocatalytic process. *J Hazard Mater*. 2020;382: 121119.
  27. Gomez-Obando VA, Garcia-Mora AM, Basante JS, Hidalgo A, Galeano LA. CWPO degradation of methyl orange at circumneutral pH: multi-response statistical optimization, main intermediates and by-products. *Front Chem*. 2019;7:00772.
  28. Kumar AN, Geetha G. Recognizing ancient south indian language using opposition based grey wolf optimization. *Intell Automation Soft Comput*. 2023;35:2619–37.
  29. Amdoun R, Benyoussef EH, Benamghar A, Khelifi L. Prediction of hyoscyamine content in *Datura stramonium* L. hairy roots using different modeling approaches: response surface methodology (RSM), artificial neural network (ANN) and kriging. *Biochem Eng J*. 2019;144:8–17.
  30. Belaadi A, Boumaaza M, Alshahrani H, Bourchak M, Satha H. Improving the mechanical performance of biocomposite plaster/*Washingtonia filifera*: optimization comparison between ANN and RSM approaches. *J Natural Fibers*. 2023;20:2170945.
  31. Rafique D, Mussolin M, Forzati M, Martensson J, Chugtai MN, Ellis AD. Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm. *Opt Express*. 2011;19:9453–60.
  32. Silaban H, Zarlis M, Sawaluddin, Iop. Analysis of accuracy and epoch on back-propagation BFGS Quasi-Newton. In: International conference on information and communication technology (IconICT), Medan, Indonesia. 2017.
  33. Elijah OC, Philomena KI, Joseph TN, Charles ON, Paschal EO. Evaluation of optimization techniques in predicting optimum moisture content reduction in drying potato slices. *Artif Intell Agric*. 2020;4:39–47.
  34. Ishola NB, McKenna TFL. Influence of process parameters on the gas phase polymerization of ethylene: RSM or ANN statistical methods? *Macromol Theory Simul*. 2021;30:59.
  35. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*. 2016; 181: 1–20.
  36. Murshid SSA, Atoum D, Abou-Hussein DR, Abdallah HM, Hareeri RH, Almukadi H, Edrada-Ebel R. Genus *Salsola*: chemistry, biological activities and future prospective - A review. *Plants-Basel*. 2022; 11: 714.
  37. Tefarikis DT, Morales-Briones DF, Yang Y, Edwards G, Kadereit G. On the hybrid origin of the C<sub>2</sub> *Salsola divaricata* agg. (Amaranthaceae) from C<sub>3</sub> and C<sub>4</sub> parental lineages. *New Phytologist*. 2022;234:1876–90.
  38. Wen ZB, Wang YL, Xia CL, Zhang YF, Zhang HX. Chloroplastic *SaNADP-ME4* of C<sub>3</sub>–C<sub>4</sub> woody desert species *Salsola laricifolia* confers drought and salt stress resistance to *Arabidopsis*. *Plants-Basel*. 2021;10:1827.
  39. Wen ZB, Zhang ML. *Salsola laricifolia*, another C<sub>3</sub>–C<sub>4</sub> intermediate species in tribe *Salsoleae* s.l. (Chenopodiaceae). *Photosynth Res*. 2015;123:33–43.
  40. Bianconi ME, Sotelo G, Curran EV, Milenkovic V, Samaritani E, Dunning LT, Bertolino LT, Osborne CP, Christin PA. Upregulation of C<sub>4</sub> characteristics does not consistently improve photosynthetic performance in intraspecific hybrids of a grass. *Plant, Cell Environ*. 2022;45:1398–411.
  41. Jacob J, Bossard N, Bariac T, Terwilliger V, Biron P, Richard P, Le Milbeau C, Verges E. Hydrogen isotopic fractionations during syntheses of lipid biomarkers in the seeds of broomcorn millet (*Panicum miliaceum* L.) under controlled environmental conditions. *Organic Geochem*. 2021;154:104221.
  42. Yorimitsu Y, Kadosono A, Hatakeyama Y, Yabiku T, Ueno O. Transition from C<sub>3</sub> to proto-Kranz to C<sub>3</sub>–C<sub>4</sub> intermediate type in the genus *Chenopodium* (Chenopodiaceae). *J Plant Res*. 2019;132:839–55.
  43. Xi YX, Zhang YH, Wen ZB. Preparation and optimization of leaf protoplasts of *Salsola laricifolia*. *Arid Zone Res*. 2023;40:655–62.
  44. Wang JJ, Wang Y, Lu TF, Yang X, Liu J, Dong Y, Wang YZ. An efficient and universal protoplast isolation protocol suitable for transient gene expression analysis and single-cell RNA sequencing. *Int J Mol Sci*. 2022;23:3419.

45. Li J, Zhang P, Hu J, Zhang Y. Study of the synergistic effect of induction heating parameters on heating efficiency using taguchi method and response surface method. *Appl Sci-Basel*. 2023;13:555.
46. Li Q, Yu H, Zhang Z, Sun Z, Zhang Y, Zhang D, Wang X, Ma Z, Yan Y. Optimization of cotton mesophyll protoplast transient expression system. *Scientia Agricultura Sinica*. 2021;54:4514–24.
47. Pandian PS, Selvan SS, Subathira A, Saravanan S. Optimization of aqueous two phase extraction of proteins from *Litopenaeus Vannamei* waste by response surface methodology coupled multi-objective genetic algorithm. *Chem Prod Process Model*. 2020;15:34.
48. Ajala SO, Alexander ML. Multi-objective optimization studies of microalgae dewatering by utilizing bio-based alkali: a case study of response surface methodology (RSM) and genetic algorithm (GA). *Sn Appl Sci*. 2020;2:287.
49. Bello VE, Olafadehan OA. Comparative investigation of RSM and ANN for multi-response modeling and optimization studies of derived chitosan from *Archachatina marginata* shell. *Alex Eng J*. 2021;60:3869–99.
50. Hlangwani E, Doorsamy W, Adebiji JA, Fajimi LI, Adebo OA. A modeling method for the development of a bioprocess to optimally produce *umqombothi* (a South African traditional beer). *Sci Rep*. 2021;11:20626.
51. Kannaiyan M, Govindan K, Raghuvaran JGT. Prediction of specific wear rate for LM25/ZrO<sub>2</sub> composites using Levenberg-Marquardt back-propagation algorithm. *J Market Res*. 2020;9:530–8.
52. Nazerian M, Naderi F, Partovinia A, Papadopoulos AN, Younesi-Kordkheili H. Modeling the bending strength of MDF faced, polyurethane foam-cored sandwich panels using response surface methodology (RSM) and artificial neural network (ANN). *Forests*. 2021;12:1514.
53. Wen ZB, Zhang ML. Reference Gene Selection for real-time quantitative PCR in *Salsola laricifolia* under soil drought stress. *Plant Physiol J*. 2015;51:2031–8.
54. Nevone A, Lattarulo F, Russo M, Panno G, Milani P, Basset M. A strategy for the selection of RT-qPCR reference genes based on publicly available transcriptomic datasets. *Biomedicines*. 2023;11:1079.
55. Tang J, Li E, Liu J, Zhang Z, Hua B, Jiang J, Miao M. Selection of reliable reference genes for gene expression normalization in *Sagittaria trifolia*. *Genes*. 2023;14:1321.
56. Ren R, Gao J, Lu C, Wei Y, Jin J, Wong SM, Zhu G, Yang F. Highly efficient protoplast isolation and transient expression system for functional characterization of flowering related genes in *Cymbidium Orchids*. *Int J Mol Sci*. 2020;21:2264.
57. Mohammed M, Alqahtani NK. Design and validation of automated sensor-based artificial ripening system combined with ultrasound pretreatment for date fruits. *Agronomy-Basel*. 2022;12:2805.
58. Eren F, Aslan M, Kanarya D, Uysalli Y, Aydin M, Kiraz B, Aydin O, Kiraz A. DeepCAN: a modular deep learning system for automated cell counting and viability analysis. *IEEE J Biomed Health Inform*. 2022;26:5575–83.
59. Alrugaibah M, Yagiz Y, Gu L. Use natural deep eutectic solvents as efficient green reagents to extract procyanidins and anthocyanins from cranberry pomace and predictive modeling by RSM and artificial neural networking. *Sep Purif Technol*. 2021;255:117720.
60. Avinash A, Murugesan A. Prediction capabilities of mathematical models in producing a renewable fuel from waste cooking oil for sustainable energy and clean environment. *Fuel*. 2018;216:322–9.
61. Zhang X, Zhang Q, Li Y, Zhang H. Modeling and optimization of photo-fermentation biohydrogen production from co-substrates basing on response surface methodology and artificial neural network integrated genetic algorithm. *Biores Technol*. 2023;374:128789.
62. Faheem M, Khaliq S. Validation of housekeeping genes for gene expression profiling in fish: a necessity. *Crit Rev Eukaryot Gene Expr*. 2019;29:565–79.
63. Joshi CJ, Ke W, Drangowska-Way A, O'Rourke EJ, Lewis NE. What are housekeeping genes? *PLoS Comput Biol*. 2022;18:e1010295.
64. Li R, Shen Y. An old method facing a new challenge: re-visiting housekeeping proteins as internal reference control for neuroscience research. *Life Sci*. 2013;92:747–51.
65. Wei K, Ma L. Concept development of housekeeping genes in the high-throughput sequencing era. *Yichuan Hereditas*. 2017;39:127–34.

## Publisher's Note

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