Food Chemistry 227 (2017) 264–270

Contents lists available at ScienceDirect

Food Chemistry

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

Artificial neural network – Genetic algorithm to optimize wheat germ fermentation condition: Application to the production of two anti-tumor benzoquinones

Zi-Yi Zheng, Xiao-Na Guo, Ke-Xue Zhu *, Wei Peng, Hui-Ming Zhou *

State Key Laboratory of Food Science and Technology, Collaborative Innovation Center for Food Safety and Quality Control, School of Food Science and Technology, Jiangnan University, 1800 Lihu Avenue, Wuxi 214122, Jiangsu Province, People's Republic of China

ARTICLE INFO

Article history:
Received 20 June 2016
Received in revised form 23 December 2016
Accepted 15 January 2017
Available online 18 January 2017

Keywords:
Wheat germ
Benzoquinones
Fermentation
Artificial neural network
Second-order interactions

ABSTRACT

Methoxy-ρ-benzoquinone (MBQ) and 2, 6-dimethoxy-ρ-benzoquinone (DMBQ) are two potential anti-cancer compounds in fermented wheat germ. In present study, modeling and optimization of added macronutrients, microelements, vitamins for producing MBQ and DMBQ was investigated using artificial neural network (ANN) combined with genetic algorithm (GA). A configuration of 16-11-1 ANN model with Levenberg-Marquardt training algorithm was applied for modeling the complicated nonlinear interactions among 16 nutrients in fermentation process. Under the guidance of optimized scheme, the total contents of MBQ and DMBQ was improved by 117% compared with that in the control group. Further, by evaluating the relative importance of each nutrient in terms of the two benzoquinones' yield, macronutrients and microelements were found to have a greater influence than most of vitamins. It was also observed that a number of interactions between nutrients affected the yield of MBQ and DMBQ remarkably.

1. Introduction

Wheat germ, which is the by-product of flour-milling industry, contains various bioactive compounds which are beneficial to human physiology (Rizzello, Cassone, Coda, & Gobbetti, 2011). In the past few decades, several studies demonstrated the safety of fermented wheat germ extract for its intended use as a dietary supplement ingredient (Roberta et al., 2002). According to the findings of literatures, fermented (by *Saccharomyces cerevisiae*) wheat germ extract exhibited cytotoxic activity towards cancer cell lines and positive immunological effects (Jakab et al., 2003). It is currently assumed that two methoxy-substituted benzoquinones, i.e., methoxy-ρ-benzoquinone (MBQ) and 2,6-dimethoxy-ρ-benzoquinone (DMBQ) were responsible for the biological properties (Mueller, Jordan, & Voigt, 2011). The two benzoquinones were derived from hydroquinones substituted by β-1,6-linked oligosaccharides in wheat germ (Zhokhov, Broberg, Kenne, & Jastrebova, 2010). In the fermentation process, β-glucosidic linkages of hydroquinones glucosides were split by β-glucosidase (EC 3.2.1.21) of *Saccharomyces cerevisiae* to form methoxy-ρ-hydroquinone and 2, 6-dimethoxy-ρ-hydroquinone, as shown in Fig. S1. Further, the two hydroquinones were catalyzed by wheat germ peroxidase (WGP) to form MBQ and DMBQ (Garcia, Rakotozafy, Telef, Potus, & Nicolas, 2002). From the above description, the production of high-price bioactives (MBQ and DMBQ) from fermented wheat germ requires high activities of β-glucosidase and peroxidase. However, the activities of β-glucosidase and peroxidase in fermentation broth were low and unsatisfactory for the bioconversion of the hydroquinone glucosides (Rizzello et al., 2013).

Exogenous β-glucosidase and peroxidase could be added into fermentation broth to enhance the yields of MBQ and DMBQ. However, high costs and loss of activities during the fermentation process are two major obstacles of exogenous enzymes. It is well known that nutrients, i.e. macronutrients, microelements and vitamins are essential for microbial metabolism. Further, some macronutrients and microelements have been proven to modulate enzyme activities due to their location in the active site of enzyme (Jeng et al., 2011). Therefore, addition of appropriate amount of nutrients might be speculated to increase the yields of MBQ and DMBQ.

There is a crucial need to find a predictive model to illuminate the complicated law between multiple nutrients and two benzoquinones in wheat germ fermentation process. Traditional modeling and optimization approaches for multiple variables such as response surface methodology, present restrictions for modeling highly complex biological systems (Rafigh, Yazdi, Vossoughi,

* Corresponding authors.
E-mail addresses: 18206188734@163.com (Z.-Y. Zheng), xiaonaguo@jiangnan.edu.cn (X.-N. Guo), kxzhu@jiangnan.edu.cn (K.-X. Zhu), pengwei@jiangnan.edu.cn (W. Peng), hmzhou@jiangnan.edu.cn (H.-M. Zhou).

http://dx.doi.org/10.1016/j.foodchem.2017.01.077
0308-8146/

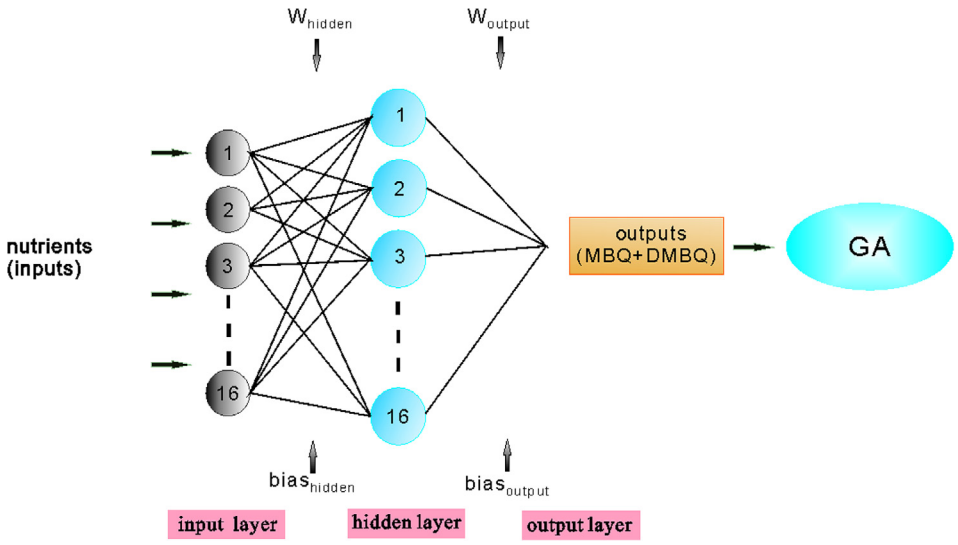

Fig. 1. Schematic diagram of the combined back-propagation artificial neural network modeling and genetic algorithm. The connections between nodes are associated with weights and biases.

Safekordi, & Ardjmand, 2014). In the last decade, artificial neural network (ANN), which exhibited high accuracy and generalization ability in modeling, has been widely applied to model the non-linear biological systems (Cimpoiu, Cristea, Hosu, Sandru, & Seserman, 2011; Hosu, Cristea, & Cimpoiu, 2014; Vani, Sukumaran, & Savithri, 2015). A common artificial neural network is a connected parallel architecture consisting of an input layer composed of neurons (corresponding to the input variables), a hidden layers composed of neurons and an output layer composed of neurons (corresponding to the output variables), as schematized in Fig. 1. Genetic algorithm (GA) is a stochastic global optimizing algorithm, which is based on the laws of biological evolution (Kumar, Pathak, & Guria, 2015). GA has been shown the ability to solve smooth or non-smooth optimization and this algorithm does not require differentiable or continuous functions. Applying of GA has been proved to be effective in improving the productivity of fermentation (Camacho-Rodríguez, Cerón-García, Fernández-Sevilla, & Molina-Grima, 2015; García-Camacho, Gallardo-Rodríguez, Sánchez-Mirón, Chisti, & Molina-Grima, 2011; Kumar et al., 2015).

The fermentation parameters, i.e. agitation speed, initial pH, fermentation temperature and fermentation time have been optimized in order to maximize the total contents of MBQ and DMBQ in our previous study (Zheng, Guo, Zhu, Peng, & Zhou, 2016). In present work, ANN was assessed as a predictive model between the total contents of MBQ and DMBQ in fermented wheat germ and 16 nutrients (macronutrients, micronutrients and vitamins). Then two approaches of evaluating the relative importance of each nutrient for the total contents of MBQ and DMBQ was performed. Moreover, to evaluate nutrients' second order interactions for the total contents of MBQ and DMBQ, a novel method was established based on sixteen newly established ANN models with partial derivatives as the outputs. Finally, an optimal formulation of exogenous nutrients to achieve the maximum total contents of MBQ and DMBQ was worked out by the combined ANN-GA method.

2. Materials and methods

2.1. Materials

The raw wheat germ was obtained from Yihai Kerry Food Industry Co, Ltd (Kunshan, China). It was stored at −18 °C in a freezer before experiments. Calcium chloride ($CaCl_2$), Magnesium sulfate heptahydrate ($MgSO_4{\cdot}7H_2O$), iron vitriol, ($FeSO_4{\cdot}7H_2O$), copper sulfate pentahydrate ($CuSO_4{\cdot}5H_2O$), zinc sulfate heptahydrate ($ZnSO_4{\cdot}7H_2O$), Manganese sulfate monohydrate ($MnSO_4{\cdot}H_2O$), selenium (Se), thiamine hydrochloride (B_1), riboflavin (B_2), nicotinic acid (B_3), calcium d-pantothenate (B_5), pyridoxine hydrochloride (B_6), Cyanocobalamin (B_{12}), 2-methyl-1,4-naphthoquinone (K_3), 1-naphthalenol, 4-amino-2-methyl-,hydrochloride (1:1) (K_5), and d-ascorbic acid (C) were purchased from Sinopharm Chemical Reagent Co. Ltd (Shanghai, China). Methoxy-ρ-benzoquinone, and 2,6-ρ-dimethoxybenzoquinone were purchased from Meilian biotechnology Co., Ltd (Shanghai, China).

2.2. Media preparation and fermentation

The strain *S. cerevisiae* (RC212, Huankai Biological Technology Co., Ltd, Guang Zhou, China) were cultivated on a YPD (1% yeast extract, 1% peptone, 2% glucose, 2% agar) medium in a 500 ml Erlenmeyer flask under aerobic condition with chloramphenicol (0.01%) to avoid bacterial growth. The Erlenmeyer flask was incubated in a controlled incubator shaking at 30 °C. Then *S. cerevisiae* was transferred into liquid YPD overnight under aerobic condition with constant shaking at 30 °C in Erlenmeyer flasks. For all experiments, *S. cerevisiae* was collected at the beginning of the exponential growth phase (OD_{600nm} = 1.1). 5 ml *S. cerevisiae* suspension was inoculated in a 500 ml Erlenmeyer flask with 200 ml tap water and 10 g wheat germ.

Sixteen nutrients including macronutrients (Ca, Mg), micronutrients (Fe, Cu, Zn, Mn, Se), vitamins (B_1, B_2, B_3, B_5, B_6, B_{12}, K_3, K_5, C) were added into medium before fermentation. The concentration range of each component was determined by single – factor experiments and could be seen in Supplementary Figs. S4–S19. Erlenmeyer flask was held on an orbital shaker incubator at 142 rpm, 31.6 °C. After 39.8 h of fermentation, each sample was freeze-dried and then kept in the refrigerator. The agitation speed, fermentation temperature and time were optimized previously.

2.3. Determination of MBQ and DMBQ

Ten grams of lyophilized sample were dissolved in 250 ml of double distilled water and extracted three times by shaking with 100 ml $CHCl_3$. $CHCl_3$ layers were collected and washed three times

with distilled water. The filtrate was evaporated to dry condition by vacuum evaporator at 30 °C. The dry material was dissolved into mobile phase then filtered through 0.22 μm PTFE filter aid. 20 μl filtrate was injected into the HPLC column and measured, using a previous method with modifications (Rizzello et al., 2013). The HPLC system was equipped with C-18 (5 μm, 250 × 4.6 mm) column (Macherey-Nagel, Germany) and a detector UV 900 operating at 275 nm. The mobile phase was 20% acetonitrile: 80% water (v/v) mixture. Flow rate was 0.7 ml/min. MBQ and DMBQ were dissolved in mobile phase as references. Identification of peaks was confirmed by retention time. MBQ and DMBQ concentration were extrapolated from pure standards. All quantifications were determined in triplicate.

2.4. Modeling of artificial neural network

The input variables of the ANN consisted of a matrix whose column vectors were the concentrations of nutrients. Before developing an ANN model, the input and output data were scaled by 'mapminmax' function. The data points (330) were randomly divided into three subsets, i.e. training set (60%), validation set (20%) and testing set (20%). The training set was used to adjust the weights and biases of network. The second subset (validation set) was used to minimize overfitting, in order to make the model more reliable and robust. The testing set was used to confirm the actual predictive power of the network. The output function can be represented according to the following expression:

$$f_2(\mathrm{x}, \mathrm{w}_{\mathrm{hidden}}, \mathrm{w}_{\mathrm{output}}) = \mathrm{w}_{\mathrm{output}} \cdot f_1(\mathrm{w}_{\mathrm{hidden}} \cdot \mathrm{x} + \mathrm{bias}_{\mathrm{hidden}}) + \mathrm{bias}_{\mathrm{output}}$$

where $\mathrm{w}_{\mathrm{hidden}}$, $\mathrm{w}_{\mathrm{output}}$, $\mathrm{bias}_{\mathrm{hidden}}$, $\mathrm{bias}_{\mathrm{output}}$ are the parameter matrices associated with connections between the nodes of adjacent layers, and f_1 and f_2 are the activation function, as shown in Fig. 2. The activation function between the input layer and hidden layer was 'tansig', and 'purelin' was used as activation function between the hidden layer and the output layer. The learning rate was set as 0.01. The epoch size was adjusted in the range of 200–5000.

The ANN model was trained with nine back-propagation training algorithms one after another, i.e. Levenberg-Marquardt (trainlm), Bayesian-regularization (trainbr), conjugate gradient with Fletcher-Reeves (traincgf), conjugate gradient with Powell-Beale (traincgb), conjugate gradient with Polak-Ribiere (traincgp), gradient descent with momentum (traindm), gradient descent with adaptive learning rate (trainda), gradient descent with momentum and adaptive learning rate (traingdx), scaled conjugate gradient (trainscg). During the training process, the biases and weights of the neural network were systematically updated to obtain the lowest mean square error (MSE) between the predicted values and experimental values. The MATLAB 7.0 Neural Network Toolbox was used to train the ANN model.

2.5. Methods of ranking components' relative importance

2.5.1. Garson's algorithm

Garson's algorithm has been widely used to evaluate the relative importance of each independent variable for a dependent variable in an ANN model (Giam & Olden, 2015; Grahovac, Jokić, Dodić, Vučurović, & Dodić, 2016). Based on the Garson's algorithm, the product sum of connecting weights obtained from trained ANN model was used to calculate the relative importance of each nutrient for the total contents of MBQ and DMBQ. The relative importance (R_{ik}) of input variables is defined by:

$$\mathrm{R}_{\mathrm{ik}} = \frac{\sum_{\mathrm{J}=1}^{\mathrm{M}}(|\mathrm{W}_{\mathrm{ij}}\mathrm{W}_{\mathrm{jk}}|/\sum_{\mathrm{r}=1}^{\mathrm{N}}|\mathrm{S}_{\mathrm{rj}}|)}{\sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{J}=1}^{\mathrm{M}}(|\mathrm{W}_{\mathrm{ij}}\mathrm{W}_{\mathrm{jk}}|/\sum_{\mathrm{r}=1}^{\mathrm{N}}|\mathrm{S}_{\mathrm{rj}}|)} \quad (\mathrm{i}=1,\ldots,\mathrm{N};\ \mathrm{r}=1,\ldots,\mathrm{N};\ \mathrm{j}=1,\ldots,\mathrm{M};\ \mathrm{k}=1,\ldots,\mathrm{L})$$

where R_{ik} is the relative importance of input variable i to output k; N is the number of neurons in input layer; M is the number of neurons in hidden layer and K is the number of neurons in output layer; W_{ij} is the connected weight between input node i and hidden node j; W_{jk} is the weight between hidden node j and output node k; S_{rj} represents the sum of $\mathrm{W}_{\mathrm{ij}}\mathrm{W}_{\mathrm{jk}}$ in j neuron of hidden layer.

2.5.2. Partial derivatives method

The method of partial derivatives is considered to be another efficient method to rank the relative importance of each independent variable for a dependent variable applied in an ANN model (Gevrey, Dimopoulos, & Lek, 2003). The partial derivatives d_{ik} of the output y_k with respect to input x_i could be computed by the following equation:

$$d_{ik} = \frac{\partial y_k}{\partial x_i} = f'_1 \sum_{j=1}^{L} W_{jk} f'_2 W_{ij}$$

where f'_1 is the derivatives of activation function between hidden nodes and output nodes; f'_2 is the derivatives of activation function between hidden nodes and input nodes; W_{ij} is the connected weight between input node i and hidden node j; W_{jk} is the weight between hidden node j and output node k. Thus, a set of graphs of partial derivatives with respect to each corresponding input variable were plotted (Figs. S4–S19). Further, in order to rank the relative importance of each input to the output, the sum of the square partial derivatives (SSD) with respect to an input variable was calculated as follow:

$$SSD_{\mathrm{i}} = \sum_{i=l}^{N} (d_{ik})^2$$

The highest SSD of the input variable is considered to influence the output variable most.

2.6. Factorial design methodology

The fractional factorial design (FFD) is a common method to evaluate the relative influence of second-order interactions (Panić et al., 2015). In present work, the fractional factorial design was applied to verify the feasibility of the proposed method which was applied to evaluate second-order interactions (in Section 3.4). Factors were coded at three levels, i.e., upper level, central point and lower level. The full experimental designs were listed in Tables S1 and S2. The statistical analysis was performed by Minitab 17.0.

2.7. Genetic algorithm optimization

Genetic algorithm was employed to optimize the concentrations of 16 nutrients in order to proved a guide line for maximizing the total contents of MBQ and DMBQ. The constructed ANN was used as the fitness function in GA. Each variable was coded to form gene, and 16 genes formed a chromosome which represented an individual. The genetic algorithm comprised a cycle of four steps. First, a population of random individuals was generated. Secondly, evaluation of these individuals was performed based on fitness function. Thirdly, individuals of higher fitness values would have a greater probability of producing offspring. After selection, mutation and crossover, the offspring turned into new parents. The above steps were repeated until the optimum individual stopped changing. After decoding, the maximum value of MBQ + DMBQ and the optimized concentration value for each nutrient were obtained. The Genetic Algorithm Toolbox, developed by the University of Sheffield, was applied to obtain the optimal solution.

3. Results and discussion

3.1. Evaluation of ANN

The back-propagation network topology might be most frequently applied network in modeling complex fermentation process (Ma et al., 2011; Sridevi, Sivaraman, & Mullai, 2014; Zhang, Xie, Yu, & Li, 2014). The present work attempted to develop a reliable back-propagation ANN to model the wheat germ fermentation process with addition of 16 nutrients for the production of MBQ and DMBQ. First of all, optimization of the number of neurons in the hidden layer is important in order to obtain good prediction. The upper limit of the number of hidden layer nodes was computed based on a widely accepted empirical rule:

$$N \leqslant \frac{N^{tr}}{N^{inputs}+1}$$

where N is the maximum number of hidden layer nodes, N^{TR} is the number of samples in training set, and N^{inputs} is the number of input nodes (Nagata & Chu, 2003). The equation provided a value of 11 for the upper limit of nodes number in hidden layer. Therefore, the number of neurons in hidden layer was tested from 5 to 11. Results showed that an ANN model with 11 neurons in the hidden layer gave lowest mean square error (MSE) for validation dataset and test dataset. Thus, final architecture of ANN consisted of 11 neurons in the hidden layer.

On the other hand, it is important to select a proper training algorithm for the ANN. Hence, in present work, nine training algorithms were tested for the ANN model, and the results are exhibited in Table 1. As can be seen, Levenberg-Marquardt algorithm could be considered as the most appropriate training algorithm on account of smallest mean square error (MSE) for the test dataset. This is more clearly illustrated in Fig. 2, where the experimented values of the total contents of MBQ and DMBQ were compared with the predicted values. As shown in Fig. 2, the predicted values and experimented values in training set, validation set and test set were uniformly distributed around the regression line (y = x, i.e. predicted value = experimented value). The coefficients of determination (R^2) for the training set (R^2 = 0.971) and validation set (R^2 = 0.931) were better than that for the test set (R^2 = 0.928). The coefficients of determination (R^2 = 0.928) for test set, which demonstrated a good linear fit between the experimented values and the predicted values, confirmed the ability of the trained ANN to predict new data precisely.

3.2. Relative importance of input variables

The relative importance of nutrients to the total contents of MBQ and DMBQ was calculated based on Garson' algorithm and PaD method. The relative importance of nutrients to the total contents of MBQ and DMBQ was converted into relative terms for comparison purpose. As observed in Fig. S2, for Garson' algorithm and PaD, interestingly, results of relative importance of 16 nutrients contributing to the total contents of MBQ and DMBQ were similar. Two methods both showed vitamin C as the most important nutrient to the total contents of MBQ and DMBQ. Except vitamin C, the relative importance ranged from 0.62% for B_{12} to 9.63% for $ZnSO_4{\cdot}7H_2O$ based on PaD method, and 0.88% for K_3 to 13.6% for $ZnSO_4{\cdot}7H_2O$ based on Garson's algorithm. Despite vitamin C and B_1, it was observed that other vitamins have a relatively small impact on the total contents of MBQ and DMBQ, compared with macronutrients (Ca, Mg) and micronutrients (Fe, Cu, Zn, Mn).

3.3. Effect of each nutrient on the total contents of MBQ and DMBQ

Due to the Black-Box structure of ANN, the relation between each component and total contents of MBQ and DMBQ was unknowable in the present ANN model. However, profiles of the first-order partial derivative with respect to each input variable (Figs. S4–S19) could provide more information about the relation between each nutrient and the total contents of MBQ and DMBQ (Nourani & Sayyah Fard, 2012). In Figs. S7, S10, S11 and S13, it can be seen that most of partial derivatives with respect to

Table 1
Evaluation of nine back-propagation training algorithms.

Back-propagation algorithms	Function	MSE (training)	MSE (validation)	MSE (test)
Levenberg-Marquardt	**trainlm**	**0.018**	**0.0387**	**0.0396**
Bayesian-regularization	trainbr	0.029	0.089	0.103
Conjugate gradient with Fletcher-Reeves	traincgf	0.030	0.064	0.088
Conjugate gradient with Powell-Beale	traincgb	0.034	0.071	0.088
Conjugate gradient with Polak-Ribiere	traincgp	0.032	0.074	0.089
Gradient descent with momentum	traindm	0.044	0.073	0.082
Gradient descent with adaptive learning rate	trainda	0.041	0.074	0.105
Gradient descent with momentum and adaptive learning rate	traingdx	0.089	0.063	0.111
Scaled conjugate gradient	trainscg	0.029	0.070	0.079

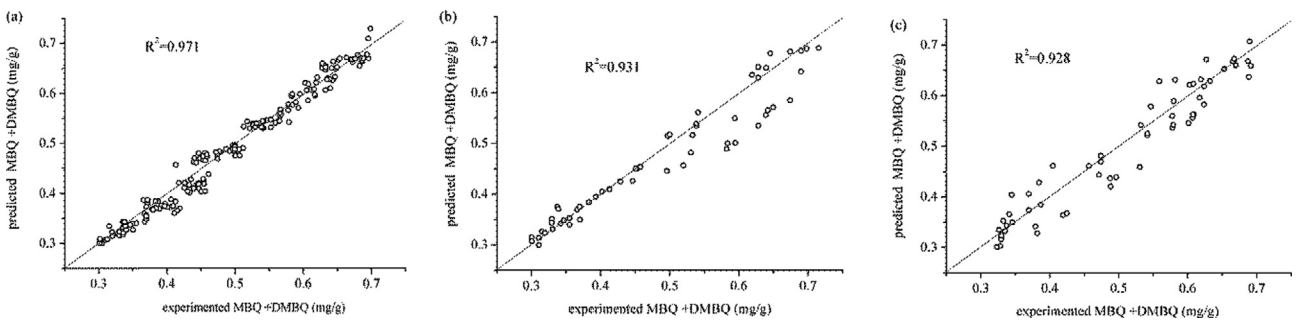

Fig. 2. The predicted total contents of MBQ and DMBQ versus experimental values: (a) training dataset, (b) validation dataset, (c) test dataset.

$CuSO_4 \cdot 5H_2O$, Se, B_1 and B_3 were positive. One example of an interpretation is that, if the partial derivative is positive at one data point, change of input variable and output has the identical tendency at this data point. On the contrary, if the partial derivative is negative at this data point, the output will increase as the input variable decreases. Nevertheless, it can be seen that the partial derivatives in most of sixteen graphs (Figs. S4–S19) were highly dispersive, meaning that the partial derivatives depended on not only x value in one graph but also concentrations of other fifteen components at this data point. In other words, complicated interactions among sixteen nutrients actually existed in bioconversion of hydroquinone glucosides during the wheat germ fermentation process (Vasilakos, Kalabokidis, Hatzopoulos, & Matsinos, 2009).

Thus, a comprehensive statistical analysis of partial derivative values with respect to each input variable was performed. Based on the changed box-plots, five percentiles (10th, 25th, 50th, 75th and 90th) of partial derivative values for each input variable were calculated and shown in Fig. 3. As can be seen, 75th percentile value of partial derivatives with respect to $FeSO_4 \cdot 7H_2O$ was negative, meaning that with the probability of more than 75%, the total contents of MBQ and DMBQ decrease with increasing concentration of $FeSO_4 \cdot 7H_2O$ (Nourani & Sayyah Fard, 2012). Same trends were found in $MnSO_4 \cdot H_2O$, B_5, C. Regarding $CuSO_4 \cdot 5H_2O$, $ZnSO_4 \cdot 7H_2O$ and B_3, over 75% of the partial derivatives were positive, and more than 90% of the partial derivatives were positive with respect to Se and B_1. Moreover, the partial derivative values with respect to B_{12}, K_3 and K_5 were clustering around the zero line, indicating that B_{12}, K_3 and K_5 might have relatively small impact on the total contents of MBQ and DMBQ. These results were consistent with the results of relative importance of 16 nutrients (Fig. S2).

3.4. Analysis of nutrients' second-order interactions

As discussed in Section 3.3, based on highly dispersivity of partial derivatives, it could be concluded that the effect of one nutrient to the total contents of MBQ and DMBQ also depended on the concentrations of other nutrients. In other words, complex interactions existed among nutrients for the total contents of MBQ and DMBQ. In previous studies, multiple-variable interactions analysis was mostly performed based on a multiple linear regression (MLR) model, and the linear relationship between two variables could be calculated by *Pearson product moment correlations* and *Spearman rank correlations* (García-Camacho et al., 2011). Therefore, a multiple linear regression (MLR) equation was established between MBQ + DMBQ and the concentrations of 16 nutrients using SPSS 19 software. The linear fitting equation was:

$$y = a_0 + \sum_{i=1}^{16} a_i x_i$$

where a_0 and a_i are the regression coefficients and x_i are concentrations of 16 components. However, the coefficients of determination R^2 (64.5%), which is an important indicator to check the validity of the MLR model, indicated that the established model explained less than 65% of the variability. For this reason, the MLR model was not appropriate for modeling the present fermentation process.

To the best of our knowledge, there are no literatures about analysis of the second-order interactions based on an ANN model. Herein, we tried to propose a method based on the authenticated ANN model to investigate nutrients' second-order interactions for the total contents of MBQ and DMBQ. According the principle of conventional full factorial design, which provided a unique estimation of main effects of factors and the second-order interactions, if the interaction between x_i and x_j is negligible, the relative importance of x_j for slope ($\partial y/\partial x_i$) should be small (Panić et al., 2015; Tarley et al., 2009). If sixteen new ANN models were constructed, using partial derivatives with respect to one variable ($\partial y/\partial x_i$) as output values and concentrations of sixteen nutrients as input values, the relative importance of sixteen nutrients contributing to partial derivatives with respect to one variable could be evaluated based on Garson's algorithm, as described in Section 2.5.1.

Therefore, sixteen new ANN models which used $\partial y/\partial x_i$ as outputs were established and trained as described in Section 3.1. After presenting good fitting ability and generalization capacity, sixteen weights matrixes were obtained from the trained ANN model. By utilization of Garson's algorithm, sixteen profiles of relative importance for $\partial y/\partial x_i$ were obtained. R_{ij} was defined as the relative importance of x_j for $\partial y/\partial x_i$. Then the sixteen values of relative importance in each profile were recalculated, in order to make R_{ii} in ith profile consistent with the relative importance value of each

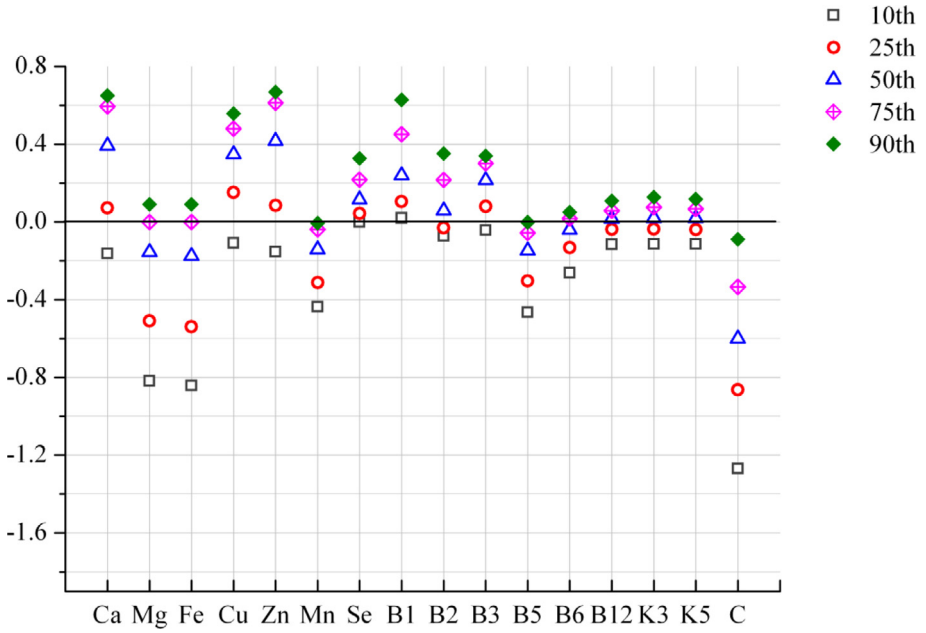

Fig. 3. Five percentiles (the 10th, 25th, 50th, 75th and 90th) of partial derivative values with respected to the input variables.

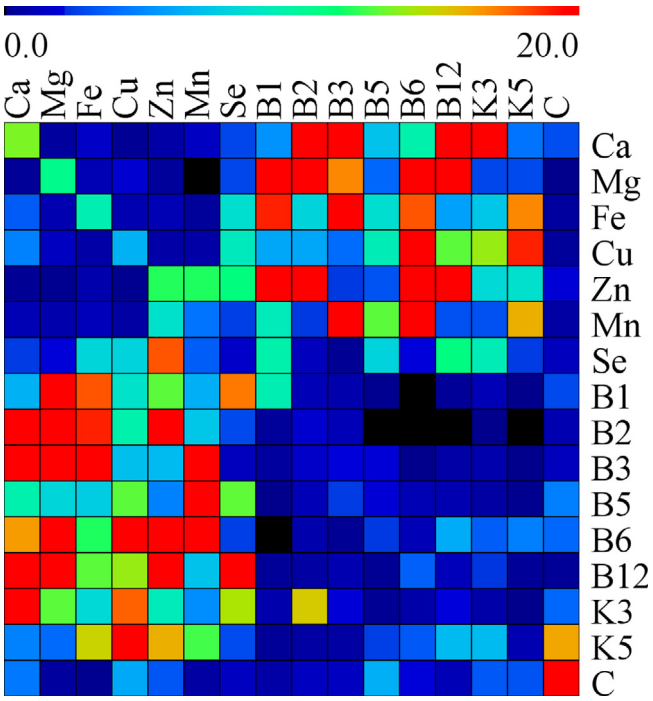

Fig. 4. Heat – map for expression of the second-order interactions of nutrients influencing the total yields of MBQ and DMBQ.

nutrient for MBQ + DMBQ (R_i) in Fig. S2. Finally a total of 16 × 16 R_{ij} were expressed as the form of heat – map (Fig. 4).

As observed in Fig. 4, B_2 and B_{12} had strong interaction with Ca and Mg for the total yields of MBQ and DMBQ. It was noticed that most of vitamins (B_1, B_2, B_3, B_6, B_{12}, K_3, K_5, C) showed weak interactions with other vitamins, which was reflected by that the area in the lower right of Fig. 4 was almost light blue. Similar weak interactions were observed with macronutrient – macronutrient. As seen in the heat – map, 15 of the most important factors and interactions could be ranked as: Mg × B_6 > Mg × B_2 > Fe × B_5 > Ca × B_3 > Zn × B_3 > Mg × B_1 > Ca × B_2 > Ca × B_{12} > Cu × B_6 > Cu × K_5 > Mg × B_{12} > Ca × K_3 > Mg × K_3 > C > Zn × B_3.

As shown in Fig. 4, most of the added vitamins were less important for the total contents of MBQ and DMBQ, in comparison with macronutrients and micronutrients. However, vitamins might exert important influences to the total contents of MBQ and DMBQ through interactions with macronutrients and micronutrients. It was indicated that adding appropriate vitamins was essential to enhance the yields of MBQ and DMBQ. Further validation experiments for the proposed method were performed using conventional fractional factorial.

To avoid possible non-linear response caused by the increased number of the independent variables, fractional factorial design with three factors as independent variables was performed in order to evaluate relative importance of nutrients and their second-order interactions for the total contents of MBQ and DMBQ in present study. Fig. S20 showed the information regarding the effects of Ca, Mg, B_2 and their second-order interactions on the total contents of MBQ and DMBQ. The length of bars in the chart represented relative importance of the factors on the total contents of MBQ and DMBQ. The relative importance could be ranked as Ca × B_2 > Ca > Mg × B_2 > Mg > Ca × Mg > B_2, which was similar with the results in Fig. 4. Another validation experiment was performed to compute the relative importance of Zn, B_6 and B_{12} (Fig. S21). The relative importance of the factors and interactions could be ranked as Zn × B_6 > Zn > Zn × B_{12} > B_6 × B_{12} > B_{12} > B_6, and the sequence was similar with the results exhibited in Fig. 4. The slight difference of the results between fractional factorial design and the proposed method might be attributed to that fractional factorial design was based on assuming a linear response generated by all the factors, which was different with ANN (Panić et al., 2015). The proposed method for evaluating the second-order interactions may be suitable in present work than the conventional fractional factorial design.

Table 2
The best formulation of nutrients in the fermentation medium obtained by ANN-GA.

Component	Concentration (mg/L water)
$CaCl_2$	782.4
$MgSO_4{\cdot}7H_2O$	18.8
$FeSO_4{\cdot}7H_2O$	40.9
$CuSO_4{\cdot}5H_2O$	84.2
$ZnSO_4{\cdot}7H_2O$	72.3
$MnSO_4{\cdot}H_2O$	2.3
Se	116
B_1	61.1
B_2	33.6
B_3	530.2
B_5	64.5
B_6	66.1
B_{12}	8.4
K_3	91
K_5	0.48
C	0.04

3.5. Optimization using the hybrid ANN-GA

After the ANN model was established, an effort was made to optimize the input variables for the goal of maximizing the total contents of MBQ and DMBQ, as shown in Fig. 1. The trained ANN model was applied as the fitness function. The Genetic Algorithm Toolbox, developed by the University of Sheffield, was applied to obtain the optimal solution. The operation parameters of genetic algorithm were assigned as follows: number of individuals: 30, maximum number of generations: 100, number of variables: 16, crossover probability: 90%, mutation probability: 0.01. The optimization process was run several times with various initial populations to avoid local optimum. Based on the combined ANN-GA method, the maximum predicted value of total contents of MBQ and DMBQ was 0.939 mg/g and the best formulation of nutrients was shown in Table 2. Further, the optimal solution was validated experimentally. The experimental value of MBQ + DMBQ based on the optimized formulation of nutrients was 0.913 ± 0.021 mg/g, which was close to the predicted value. The experimental value was 31% higher than the best value in non-optimized experimental data and 117% higher than that in control.

4. Conclusion

A hybrid ANN-GA was utilized to search a optimal formulation of sixteen nutrients in wheat germ fermentation for maximizing the total contents of MBQ and DMBQ. The bach-propagation neural network with 16-11-1 topology presented satisfactory predictive ability and generalization capacity for modeling the complicated fermentation process with multi-factors interaction. The obtained optimal formulation of sixteen additives was verified by experiment and the total contents of MBQ and DMBQ were 117% higher than the contents in the control group. Through analysis of relative importance, both Garson's algorithm and PaD methods showed feasibility for evaluating the relative importance of sixteen nutrients and the results obtained from the two methods were nearly the same. Further, a methodology for analysing the second-order

interactions was proposed and the preliminary validation experiments were performed by fractional factorial design. The application of this methodology could provide valuable information of the second-order interactions in fermentation process which have been modeled by the artificial neural network.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was co-financed by the National Natural Science Foundation of China (Grant No. 31101384), Qing Lan Project, the National Key Technology R&D Program (2013AA102201), China Postdoctoral Science Foundation (Grant no. 2014M560396), the Key Technologies R&D Program of Jiangsu (BE2014325), Jiangsu Planned Projects for Postdoctoral Research Funds (Grant no. 1402072C) and the Jiangsu province 'Collaborative Innovation Center for Modern Grain Circulation and Safety' industry development program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2017.01.077.

References

Camacho-Rodríguez, J., Cerón-García, M. C., Fernández-Sevilla, J. M., & Molina-Grima, E. (2015). Genetic algorithm for the medium optimization of the microalga Nannochloropsis gaditana cultured to aquaculture. *Bioresource Technology, 177*, 102–109.

Cimpoiu, C., Cristea, V.-M., Hosu, A., Sandru, M., & Seserman, L. (2011). Antioxidant activity prediction and classification of some teas using artificial neural networks. *Food Chemistry, 127*(3), 1323–1328.

Garcia, R., Rakotozafy, L., Telef, N., Potus, J., & Nicolas, J. (2002). Oxidation of ferulic acid or arabinose-esterified ferulic acid by wheat germ peroxidase. *Journal of Agricultural and Food Chemistry, 50*(11), 3290–3298.

García-Camacho, F., Gallardo-Rodríguez, J. J., Sánchez-Mirón, A., Chisti, Y., & Molina-Grima, E. (2011). Genetic algorithm-based medium optimization for a toxic dinoflagellate microalga. *Harmful Algae, 10*(6), 697–701.

Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling, 160*(3), 249–264.

Giam, X., & Olden, J. D. (2015). A new R2-based metric to shed greater insight on variable importance in artificial neural networks. *Ecological Modelling, 313*, 307–313.

Grahovac, J., Jokić, A., Dodić, J., Vučurović, D., & Dodić, S. (2016). Modelling and prediction of bioethanol production from intermediates and byproduct of sugar beet processing using neural networks. *Renewable Energy, 85*, 953–958.

Hosu, A., Cristea, V.-M., & Cimpoiu, C. (2014). Analysis of total phenolic, flavonoids, anthocyanins and tannins content in Romanian red wines: Prediction of antioxidant activities and classification of wines using artificial neural networks. *Food Chemistry, 150*, 113–118.

Jakab, F., Shoenfeld, Y., Balogh, Á., Nichelatti, M., Hoffmann, A., Kahán, Z., ... Hidvégi, M. (2003). A medical nutriment has supportive value in the treatment of colorectal cancer. *British Journal of Cancer, 89*(3), 465–469.

Jeng, W.-Y., Wang, N.-C., Lin, M.-H., Lin, C.-T., Liaw, Y.-C., Chang, W.-J., ... Wang, A. H. J. (2011). Structural and functional analysis of three β-glucosidases from bacterium Clostridium cellulovorans, fungus Trichoderma reesei and termite Neotermes koshunensis. *Journal of Structural Biology, 173*(1), 46–56.

Kumar, A., Pathak, A. K., & Guria, C. (2015). NPK-10:26:26 complex fertilizer assisted optimal cultivation of Dunaliella tertiolecta using response surface methodology and genetic algorithm. *Bioresource Technology, 194*, 117–129.

Ma, Y., Huang, M., Wan, J., Wang, Y., Sun, X., & Zhang, H. (2011). Prediction model of DnBP degradation based on BP neural network in AAO system. *Bioresource Technology, 102*(6), 4410–4415.

Mueller, T., Jordan, K., & Voigt, W. (2011). Promising cytotoxic activity profile of fermented wheat germ extract (Avemar(R)) in human cancer cell lines. *Journal of Experimental and Clinical Cancer Research, 42*, 42.

Nagata, Y., & Chu, K. H. (2003). Optimization of a fermentation medium using neural networks and genetic algorithms. *Biotechnology Letters, 25*(21), 1837–1842.

Nourani, V., & Sayyah Fard, M. (2012). Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Advances in Engineering Software, 47*(1), 127–146.

Panić, S., Rakić, D., Guzsvány, V., Kiss, E., Boskovic, G., Kónya, Z., & Kukovecz, Á. (2015). Optimization of thiamethoxam adsorption parameters using multi-walled carbon nanotubes by means of fractional factorial design. *Chemosphere, 141*, 87–93.

Rafigh, S. M., Yazdi, A. V., Vossoughi, M., Safekordi, A. A., & Ardjmand, M. (2014). Optimization of culture medium and modeling of curdlan production from Paenibacillus polymyxa by RSM and ANN. *International Journal of Biological Macromolecules, 70*, 463–473.

Rizzello, C. G., Cassone, A., Coda, R., & Gobbetti, M. (2011). Antifungal activity of sourdough fermented wheat germ used as an ingredient for bread making. *Food Chemistry, 127*(3), 952–959.

Rizzello, C., Mueller, T., Coda, R., Reipsch, F., Nionelli, L., Curiel, J., & Gobbetti, M. (2013). Synthesis of 2-methoxy benzoquinone and 2,6-dimethoxybenzoquinone by selected lactic acid bacteria during sourdough fermentation of wheat germ. *Microbial Cell Factories, 12*(1), 1–9.

Roberta, F.-B., Mate, H., Shoenfeld, Y., Gabriela, I., Dmytro, D., Rita, T.-F., ... Eva, M. (2002). Fermented wheat germ extract induces apoptosis and downregulation of major histocompatibility complex class I proteins in tumor T and B cell lines. *International Journal of Oncology, 20*(3), 563–570.

Sridevi, K., Sivaraman, E., & Mullai, P. (2014). Back propagation neural network modelling of biodegradation and fermentative biohydrogen production using distillery wastewater in a hybrid upflow anaerobic sludge blanket reactor. *Bioresource Technology, 165*, 233–240.

Tarley, C. R. T., Silveira, G., dos Santos, W. N. L., Matos, G. D., da Silva, E. G. P., Bezerra, M. A., ... Ferreira, S. L. C. (2009). Chemometric tools in electroanalytical chemistry: Methods for optimization based on factorial design and response surface methodology. *Microchemical Journal, 92*(1), 58–67.

Vani, S., Sukumaran, R. K., & Savithri, S. (2015). Prediction of sugar yields during hydrolysis of lignocellulosic biomass using artificial neural network modeling. *Bioresource Technology, 188*, 128–135.

Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., & Matsinos, I. (2009). Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Natural Hazards, 50*(1), 125–143.

Zhang, R., Xie, W.-M., Yu, H.-Q., & Li, W.-W. (2014). Optimizing municipal wastewater treatment plants using an improved multi-objective optimization method. *Bioresource Technology, 157*, 161–165.

Zheng, Z., Guo, X., Zhu, K., Peng, W., & Zhou, H. (2016). The optimization of the fermentation process of wheat germ for flavonoids and two benzoquinones using EKF-ANN and NSGA-II. *RSC Advances, 6*(59), 53821–53829.

Zhokhov, S. S., Broberg, A., Kenne, L., & Jastrebova, J. (2010). Content of antioxidant hydroquinones substituted by β-1,6-linked oligosaccharides in wheat milled fractions, flours and breads. *Food Chemistry, 121*(3), 645–652.