Research Press

ARTICLE

Prediction of bioactive compound activity against wood contaminant fungi using artificial neural networks

Henrique Vicente, José C. Roseiro, José M. Arteiro, José Neves, and A. Teresa Caldeira

Abstract: Biopesticides based on natural endophytic bacteria to control plant diseases are an ecological alternative to chemical treatments. *Bacillus* species produce a wide variety of metabolites with biological activity like iturinic lipopeptides. This work addresses the production of biopesticides based on natural endophytic bacteria isolated from *Quercus suber* L. Artificial neural networks were used to maximize the percentage of inhibition triggered by the antifungal activity of bioactive compounds produced by *Bacillus amyloliquefaciens*. The active compounds, produced in liquid cultures, inhibited the growth of 15 fungi and exhibited a broader spectrum of antifungal activity against surface contaminant fungi, blue stain fungi, and phytopathogenic fungi. A 19-7-6-1 neural network was selected to predict the percentage of inhibition produced by antifungal bioactive compounds. A good match among the observed and predicted values was obtained with the R^2 values varying between 0.9965–0.9971 and 0.9974–0.9989 for training and test sets. The 19-7-6-1 neural network was used to establish the dilution rates that maximize the production of antifungal bioactive compounds, namely, 0.25 h^{-1} for surface contaminant fungi, 0.45 h^{-1} for blue stain fungi, and between 0.30 and 0.40 h^{-1} for phytopathogenic fungi. Artificial neural networks show great potential in the modelling and optimization of these bioprocesses.

Résumé : Les biopesticides à base de bactéries endophytes naturelles pour lutter contre les maladies des plantes constituent une alternative écologique aux traitements chimiques. Les espèces de *Bacillus* produisent une grande variété de métabolites biologiquement actifs tels que les lipopeptides ituriniques. Cette étude porte sur la production de biopesticides par des bactéries endophytes naturelles isolées du *Quercus suber* L. Des réseaux neuronaux artificiels ont été utilisés pour maximiser le pourcentage d'inhibition provoquée par l'activité antifongique des composés bioactifs produits par *Bacillus amyloliquefaciens*. Les composés actifs, produits en culture liquide, ont inhibé la croissance de 15 champignons et avaient un spectre d'activé antifongique plus large contre les contaminants fongiques de surface, les champignons de bleuissement et les champignons phytopathogènes. Un réseau neuronal 19-7-6-1 a été choisi pour prédire le pourcentage d'inhibition produit par les composés bioactifs antifongiques. Une bonne concordance entre les valeurs observées et prédites a été obtenue; les valeurs de R^2 variaient de 0,9965 à 0,9971 et de 0,9974 à 0,9989 pour les bases d'apprentissage et de test. Le réseau neuronal 19-7-6-1 a été utilisé pour établir les taux de dilution qui maximisent la production des composés bioactifs antifongiques, nommément 0,25 h^{-1} pour les contaminants fongiques de surface, 0,45 h^{-1} pour les champignons de bleuissement et entre 0,30 et 0,40 h^{-1} pour les champignons phytopathogènes. Les réseaux neuronaux artificiels ont un potentiel élevé pour modéliser et optimiser ces processus biologiques. [Traduit par la Rédaction]

Introduction

The alarming problem of some fungal diseases in forest systems increases the urgency of discovering the discovery of new compounds with a higher activity against pathogenic agents that can cause common diseases in forest systems. Microbial populations possess enormous metabolic diversity and can activate protective mechanisms that allow them to withstand the pressures imposed either by the natural environment or human intervention (Aminov 2009). Indeed, the integration of various disease-control strategies, including biological control, should be considered to improve the efficacy and reduce fungicide levels in the environment.

Sapstain and mould growth on lumber are serious problems in the forest products industry. Structural damage is minimal to timber, but the discolorations are objectionable to buyers and highly detrimental to the pulp and paper industry. Recently, there have been many reports about the application of antagonistic fungi to control plant diseases such as the use of the *Trichoderma* species (Soytong et al. 2005). The potential of the *Trichoderma* species as a biocontrol agent of plant pathogens was first recognized

in the early 1930s (Weindling 1932). Subsequently, this species was applied successfully as a biocontrol agent against several plant diseases in commercial agriculture, as a competitive antagonist against important pathogens of forest trees, and as a biocontrol agent against wood decay fungi in urban trees (Schubert et al. 2008). Marine fungi, namely *Aspergillus japonicus*, also have proved their potential to complete in vitro strong inhibition of the growth of the pathogenic fungus *Fusarium solani* as they are biologically based and environmentally safe alternatives (El-Kassas 2008; El-Kassas and Khairy 2009). On the other hand, Cartapip (a commercial product consisting of inoculum from a melanin-deficient strain of *Ophiostoma piliferum*) is used for the prevention of sapstain on wood as well as in biopulping, where *O. piliferum* reduces the pitch content in wood chips before pulping (Beer et al. 2003). The *Bacillus* species are well-known producers of metabolites with antimicrobial properties. Usually, three different classes of bioactive peptides can be distinguished: antifungal peptides, such as bacilysin and rhizocticin; antifungal lipopeptides, such as surfactins, iturins and fengycins; and antimicrobial polypeptides,

such as subtilin (Pabel et al. 2003). The species *Bacillus amyloliquefaciens* has been reported to produce lipopeptides with antifungal proprieties. The presence of these compounds and (or) their conjugation confers antifungal activity to the cultures. The biological activity of lipopeptides is the result of the peptide cycle and the lipid chain, leading to differential inhibitory effects against various fungal species (Kowall et al. 1998; Moyne et al. 2001). Furthermore, the production of antifungal bioactive compounds by *Bacillus* in liquid cultures depends on the dilution rate used, and the different sensitivities of the compounds produced are related to the type of fungi (Caldeira et al. 2011*a*).

Artificial neural networks (ANNs) are accepted as reliable data modelling tools that are able to capture and represent complex relationships between inputs and outputs. ANNs are computational tools inspired by the architecture and internal operational features of the human brain. ANNs can learn from example, are fault-tolerant (in the sense that they are able to handle noisy and incomplete data), are able to deal with nonlinear problems, and, once trained, can perform prediction and generalization at high speed (Galushkin 2007; Haykin 2008). In recent years, ANNs have been applied to model and to optimize the production of biocompounds by *Bacillus* (Silva et al. 2008; Sathish and Prakasham 2010; Caldeira et al. 2011*b*; Zafar et al. 2012). However, the establishment of the dilution rate, which maximizes the activity of compounds against several types of fungi, is a complex and highly nonlinear problem for which, to our knowledge, no methods have been reported in the literature for a direct and accurate prediction. The objective of the current study was to use artificial-intelligence-based tools, namely ANNs, to solve this problem. The ANN-based approach was used to establish the dilution rate for *B. amyloliquefaciens* CCMI 1051 cultures to maximize the production of antifungal bioactive compounds against surface contaminant fungi (SCF), blue stain fungi (BSF), and phytopathogenic fungi (PhF).

Materials and methods

Microorganisms

Bacillus amyloliquefaciens CCMI 1051 was isolated from healthy *Quercus suber* L. (Caldeira et al. 2006). The strain was identified by morphological, physiological, and biochemical characteristics based on *Bergey's Manual of Systematic Bacteriology* and by 16S rDNA sequence analysis (Accession No. AY785773). *Bacillus amyloliquefaciens* CCMI 1051 was maintained on nutrient agar (Difco) slants and stored at 4 °C.

Fungi used in antifungal activity assays were grown on malt extract agar (MEA) (Merck, Darmstadt, Germany). The contaminant fungi were *Aspergillus niger* CCMI 296, *Penicillium expansum* CCMI 625, *Trichoderma pseudokoningii* CCMI 304, *Trichoderma harzianum* CCMI 783, *Trichoderma harzianum* CCMI 822, and *Trichoderma koningii* CCMI 868; the blue stain fungi were *Cladosporium cladosporioides* CCMI 680, *Cladosporium resinae* CCMI 262, and *Cladosporium resinae* CCMI 667; and the phytopathogenic fungi were *Botrytis cinerea* CCMI 899, *Cephalosporium* sp. F25, *Fusarium oxysporum* CCMI 898, *Fusarium solani* F4, and *Cladosporium cucumerinum* CCMI 206. *Cephalosporium* sp. F25 and *F. solani* F4 were isolated from cork and were provided by the National Agronomic Institute Collection. *Rhizopus oryzae* L122 provided by the Wood Fungi Collection of the Civil Engineering National Laboratory was also used. The remaining fungi were provided by the Culture Collection of Industrial Microorganisms (CCMI) of the Portuguese Laboratory of Industrial Microbiology, Lisbon, Portugal.

Growth conditions

Bacillus sp. CCMI 1051 was grown in continuous culture at 30 °C using a bench-scale bioreactor (Infors HT CH-4103, Bottmingen, Switzerland) with a working volume of 1 L. The chemically defined medium used contained (per litre): $(NH_4)_2SO_4$, 1.5 g; KH_2PO_4, 1.7 g; $Na_2HPO_4 \cdot 2H_2O$, 1.7 g; $MgSO_4 \cdot 7H_2O$, 0.2 g; yeast extract (Difco),

0.1 g; glucose, 2 g; and 2 mL of a mineral-rich solution (the Vishniac solution, sterilized separately) (Vishniac and Santer 1957). The medium components were dissolved in 80% of the total distilled water volume and glucose in the remaining 20% of the final volume of distilled water and thensterilized separately to avoid Maillard reactions. Solutions were sterilized in an autoclave at 121 °C and 1 bar (1 bar = 100 kPa) for 20 min. Control of foam was performed by the addition of 2 mL/L of a sterilized aqueous solution of polypropylene glycol 2000 (2% (v/v)) into the medium.

After a period of batch cultivation, a variable speed peristaltic pump (Master Flex, Cole-Parmer Instrument Co., Vernon Hills, Illinois) was used to feed sterile medium into the bioreactor at dilution rates between 0.15 and 0.75 h⁻¹. The culture volume in the vessel was kept constant using an overflow tube. Metabolic steady states were achieved when cell and glucose concentrations were constant for more than four to five retention times.

Antifungal assays

Fungal spore suspension was prepared by suspending loopfuls of hyphae and spores from a MEA slant (7 day cultures at 25 °C) in 5 mL of a 0.85% NaCl solution. The suspension was filtered through sterilized cotton. A 10^8 CFU/mL spore suspension was obtained through dilutions and adjusted to 10^5 CFU/mL using a serial dilution operation. Malt extract (2.5 mL), 0.25 mL of 10^5 CFU/mL fungi, and 2.5 mL of *Bacillus* sp. CCMI 1051 cell-free supernatants were incubated for 24 h at 25 °C. After this period, 1 mL of each interaction mixture was plated, by incorporation in 20 mL of CRB, and the petri dishes were incubated at 25 °C for 24–48 h. The relative inhibition against a control test (%) was determined in quintuplicate by counting the number of colonies.

Artificial neural networks

In the present study, the most common neural network type, the multilayer perceptron (Haykin 2008), was adopted. This type of network is formed by three or more layers of basic computing units designated by artificial neurons or nodes. It includes an input layer, an output layer, and a number of hidden layers with a certain number of active neurons connected by feed-forward links, to which are associated modifiable weights. In addition, there are also biases, which are connected to neurons in the hidden and output layers. The input layer is where the information is introduced into the network and then the information is passed to the hidden layers before reaching the output layer. The number of nodes in the input layer denotes the number of independent variables and the number of nodes in the output layer stands for the number of dependent variables (Galushkin 2007; Haykin 2008).

Network design, i.e., the establishment of network topology, is a trial-and-error process and may affect the accuracy of the resulting trained network. A number of automated techniques have been proposed to search for a "good" network topology. These typically use a hill-climbing approach that starts with an initial topology that is selectively modified to improve performance, i.e., to minimize an internal measure of error. The internal error metric most commonly used is the mean square error (Cortez et al. 2004; Han and Kamber 2006).

The training phase is the process whereby the weights are adjusted to find the optimal set of weights that would produce the right output for any input in the ideal case. There are two different training methodologies, i.e., supervised and unsupervised. The former requires input and output parameters and the latter only needs the inputs. In the present study, a supervised training method was used. Thus, the available data were randomly divided into two mutually exclusive partitions: the training set, with two-thirds of the available data, used to establish the ANN model, and the test set, with the remaining one-third of data, used after training, to evaluate the model performance (Souza et al. 2002). To ensure statistical significance of the attained results and no

dependence from the random partition of the available data into training and test sets, 20 runs were applied in all tests.

In the training phase, the back-propagation algorithm (BP) (Rumelhart et al. 1986) was applied. This is the most widely used training algorithm for the multilayered perceptron, and evolves in two phases. The first phase is concerned with the forward stage, where the information is propagated from the input to the output layer. The second phase is the backward stage where the error, defined as the difference between the observed value and the desired nominal value in the output layer, is propagated backwards to adjust the weightings and bias values. In the forward phase, the weighted sum of input components, u_j, is calculated as

$$(1) \qquad u_j = \sum_{i=1}^{n} w_{ij} x_i + \text{bias}_j$$

where w_{ij} stands for the weight linking the jth and the ith neurons in the preceding layer, x_i represents the output of the ith neuron in the preceding layer, and bias_j denotes the weight between the jth neuron and the bias neuron in the preceding layer. The output of the jth neuron in any layer, y_j, is calculated as

$$(2) \qquad y_j = f(u_j)$$

where f represents the activation function. Despite the various activation functions that can be used in the multilayered perceptrons, the sigmoid activation function, given as follows, is widely used (Han and Kamber 2006).

$$(3) \qquad \varphi(u_j) = \frac{1}{1 + e^{-u_j}}$$

The software used to implement the ANNs was the Waikato Environment for Knowledge Analysis (WEKA), keeping the default software parameters (Hall et al. 2009). In all experiments, the sigmoid activation function was applied.

Modelling performance criteria

To evaluate the performance of the different ANN models, mean absolute deviation (MAD) and mean square error (MSE) were used.

$$(4) \qquad \text{MAD} = \frac{\sum_{i=1}^{N} |Y_i' - Y_i|}{N}$$

$$(5) \qquad \text{MSE} = \frac{\sum_{i=1}^{N} (Y_i' - Y_i)^2}{N}$$

where Y is an experimental value, Y' represents a predicted value, and N denotes the number of observations. If the goal is a model that avoids large errors, MSE should be minimized, since this error metric amplifies large deviations. Conversely, if some large deviations can be allowed, the amplification effect is not necessary and MAD should be used, since it attenuates large errors. These two measures of goodness of fit do not provide any information on the nature of the errors. To investigate this point, the average of all individual errors, named bias, was calculated using

$$(6) \qquad \text{bias} = \frac{\sum_{i=1}^{N} (Y_i' - Y_i)}{N}$$

Fig. 1. Inhibition halos triggered by antifungal activity of bioactive compounds produced by *Bacillus amiloliquefaciens* at different dilution rate (*D*) culture broths on a malt extract agar (MEA) petri dish with *Trichoderma harzianum* CCMI 783.
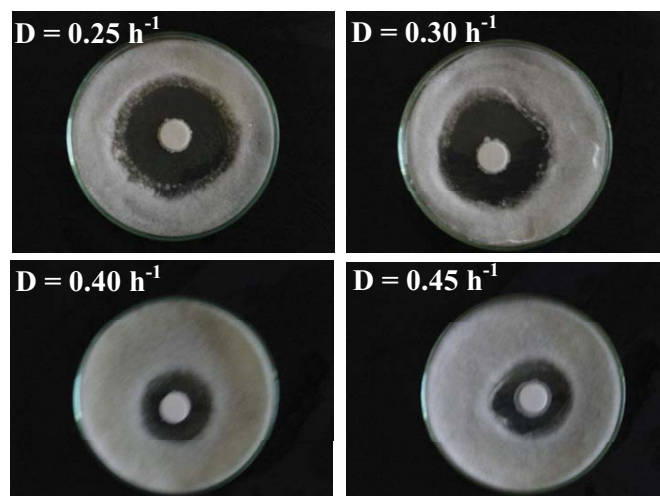
This metric is helpful, since it indicates whether the model over-estimates (bias greater than zero) or underestimates the output variables (bias less than zero).

To obtain the percentage of variability that can be explained by the selected ANN model, the coefficient of determination (R^2) was calculated using the following equation:

$$(7) \qquad R^2 = \left[\frac{N\Sigma_{i=1}^{N} Y_i' Y_i - (\Sigma_{i=1}^{N} Y_i')(\Sigma_{i=1}^{N} Y_i)}{\sqrt{\left[N\Sigma_{i=1}^{N} Y_i'^2 - (\Sigma_{i=1}^{N} Y_i')^2 \right] \times \left[N\Sigma_{i=1}^{N} Y_i^2 - (\Sigma_{i=1}^{N} - Y_i)^2 \right]}} \right]^2$$

Results

A strain of *B. amyloliquefaciens* isolated from healthy *Q. suber* was used to produce antifungal bioactive compounds against SCF, BSF, and PhF. The operating conditions that maximize the production of these compounds was previously reported (Caldeira et al. 2011*b*). The cultures of this strain were performed in a chemostat culture with different dilution rates ranging from 0.15 to 0.45 h⁻¹. Compounds produced by *B. amyloliquefaciens* inhibited the growth of the 15 strains of fungi (presented in the Microorganisms section) and exhibit a broader spectrum of activity, highly influenced by culture conditions. Figure 1 shows the inhibition halos triggered by antifungal activity of the bioactive compounds produced by *B. amyloliquefaciens*, at different dilution rates of culture broth, on an MEA petri dish with *T. harzianum* CCMI 783. The culture broth, performed in triplicate, produced distinct inhibition halos according to the dilution rate used. As can be noted from Fig. 1, the higher dilution rates (0.40 and 0.45 h⁻¹) reveal a weak diffusion into agar and a more limited antifungal activity.

Database

The data from the antifungal activity of biocompounds produced at different dilution rates by *B. amyloliquefaciens* against the 15 contaminant fungi were used to construct a database (Table 1). This database contained a total of 60 records with four fields; namely, the type of fungi (SCF, BSF, or PhF), fungal strain, dilution rate, and percentage of inhibition. Table 2 shows the statistical characterization of the numerical fields included in the database. A perusal of Table 2 reveals that the inhibition shows a large dispersion of the data with a coefficient of variation of 54.2%. This coefficient is a measure of dispersion of data and it is calculated as (SD/mean)100.

Table 1. Antifungal activity of *Bacillus amyloliquefaciens* CCMI 1051 cell-free supernatants at dilution rates ranging between 0.15 and 0.45 h⁻¹.

Dilution rate (h⁻¹)	Inhibition (%)*														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0.15	34.3±5.3	77.3±6.9	98.3±6.8	28.0±3.2	62.0±2.8	42.7±2.1	12.9±1.4	0.0±0.0	64.5±3.0	97.4±8.8	58.0±2.1	52.0±6.5	86.0±8.7	15.7±3.1	70.8±3.5
0.25	29.3±1.8	79.2±6.0	93.4±6.0	5.5±0.2	81.3±7.2	80.8±2.8	19.3±1.9	0.0±0.0	35.7±5.2	97.4±8.8	52.8±8.5	57.9±1.5	96.6±1.0	37.8±3.2	74.4±6.8
0.35	27.1±4.5	71.3±5.1	82.9±7.9	0.7±0.2	62.5±3.1	58.5±3.1	0.0±0.0	0.0±0.0	47.8±5.3	83.9±3.7	58.1±8.3	59.1±4.6	86.5±7.4	41.3±7.2	98.6±3.2
0.45	35.2±5.2	59.8±8.3	87.4±0.2	0.3±0.1	52.2±2.5	61.8±7.3	2.8±0.5	33.5±4.2	86.1±3.5	91.9±3.4	73.4±8.2	56.3±4.0	91.8±4.7	46.6±6.1	73.0±5.6

Note: *Aspergillus niger* (A), *Penicillium expansum* (B), *Trichoderma koningii* (C), *Trichoderma pseudokoningii* (D), *Trichoderma harzianum* sp1 (E), *Trichoderma harzianum* sp2 (F), *Rhizopus* sp (G), *Cladosporium cladosporioides* (H), *Cladosporium resinae* sp1 (I), *Cladosporium resinae* sp2 (J), *Cladosporium* sp (K), *Botrytis cinerea* (L), *Cephalosporium* sp (M), *Fusarium oxysporum* (N), and *Fusarium solani* (O).

*The relative inhibition against a control test was represented as % growth by quintuplicate colony counting in the presence of *Bacillus amyloliquefaciens* CCMI 1051 cell-free supernatants as compared with a control culture.

Table 2. Statistical characterization of the numerical variables used in the study.

Variable	Unit	No. of records	Min.	Max.	Mean	SD	Coefficient of variation (%)
Dilution rate	h⁻¹	60	0.15	0.45	0.299	0.114	38.1
Inhibition	%	60	0.0	98.6	55.46	30.09	54.2

ANN model

To obtain the best prediction for the percentage of inhibition produced by antifungal bioactive compounds, different network topologies were elaborated and evaluated. The optimum number of hidden layers and the optimum number of nodes in each of these was found by trial and error. Table 3 presents the values of MAD, MSE, and bias for some of the topologies considered. A glance at Table 3 shows that the 19-7-6-1 ANN topology minimizes MAD and MSE and exhibits a bias value closer to zero for the training set and for the test set. Thus, the ANN depicted in Fig. 2, formed by an input layer with 19 nodes, two hidden layers with seven and six nodes, respectively, and a one-node output layer was chosen to model the data. Figure 3 shows the mark point on a graph for experimental and predicted values of the percentage of inhibition triggered by the antifungal activity of bioactive compounds produced by *B. amyloliquefaciens* for the training and test sets. Table 4 presents the values of the coefficient of determination (R^2) for the evaluation of model performance. A good match between the observed and predicted values can be observed with the R^2 values varying in the range 0.9965–0.9971 for the training set and 0.9974–0.9989 for the test set. The agreement between the predicted and experimental values, R^2, MAD, MSE, and bias seems to suggest a good fit of the model to the data set, and demonstrates the potential of the ANNs in addressing the antifungal activity of compounds produced by *Bacillus* at different dilution rates.

In addition, Fig. 4 shows the plots for residuals versus predicted values of the percentage of inhibition triggered by the antifungal activity of the bioactive compounds produced by *Bacillus* for the training and test sets. The observed relationship between residuals and predicted values for the output variables for the training and test sets shows complete independence and random distribution. In fact, the determination coefficients are negligible (varying in the range 0.011–0.050 for the training set and 3.0×10^{-6} to 0.054 for the test set). Figure 4 shows that the points on the chart are well-distributed on both sides of the horizontal line of the zero ordinate, corresponding to the correct prediction. Plots of the residuals versus predicted values can be more informative regarding model fitting to a data set. If the residuals appear to behave randomly, it suggests that the model fits the data quite well. On the other hand, if nonrandom distribution is evident in the residuals, the model does not fit the data adequately (McBride 2005).

After the training and validation stages, the selected 19-7-6-1 ANN model was used to establish the effect of dilution rate on antifungal inhibition. Thus, the selected ANN was brought into play to predict the responses for different initial conditions, which exhaustively covered the entire experimental space. In Fig. 5, the relationships between dilution rate and the antifungal inhibition for the different fungi are depicted. Response contours for antifungal inhibition against fungi (Fig. 5) showed a dependence upon dilution rate, illustrating different behaviours for different types of fungi. Surface contaminant fungi showed a profile converging to a common maximum value of inhibition corresponding to a dilution rate of 0.25 h⁻¹. Blue strain fungi showed a sigmoidlike profile, converging to a common maximum value of fungal inhibition at a dilution rate of 0.45 h⁻¹. In contrast, the results obtained indicated that the maximum PhF inhibition was achieved with dilution rates between 0.3 and 0.4 h⁻¹.

Table 3. Mean absolute deviation (MAD), mean squared error (MSE), and bias for some of the artificial neural network (ANN) topologies tested.

ANN topology	MAD*		MSE*		bias*	
	Training set	Test set	Training set	Test set	Training set	Test set
19-5-1	0.0486	0.0322	0.0047	0.0015	0.0132	0.0125
19-12-1	0.0725	0.1179	0.0080	0.0203	0.0126	−0.0134
19-3-3-1	0.0527	0.0522	0.0041	0.0042	0.0232	−0.0018
19-4-3-1	0.0733	0.1054	0.0081	0.0162	0.0207	0.0011
19-7-6-1	0.0138	0.0069	2.74×10^{-4}	7.37×10^{-5}	0.0015	$-1.38,24.10^{-4}$
19-12-5-1	0.0908	0.1243	0.0137	0.0259	0.0208	−0.0070

*$MAD = \frac{\sum_{i=1}^{N} Y_i' - Y_i}{N}$, $MSE = \frac{\sum_{i=1}^{N}(Y_i' - Y_i)^2}{N}$, $bias = \frac{\sum_{i=1}^{N}(Y_i' - Y_i)}{N}$, where Y, Y', and N denote, respectively, an experimental value, a predicted value, and the number of observations.
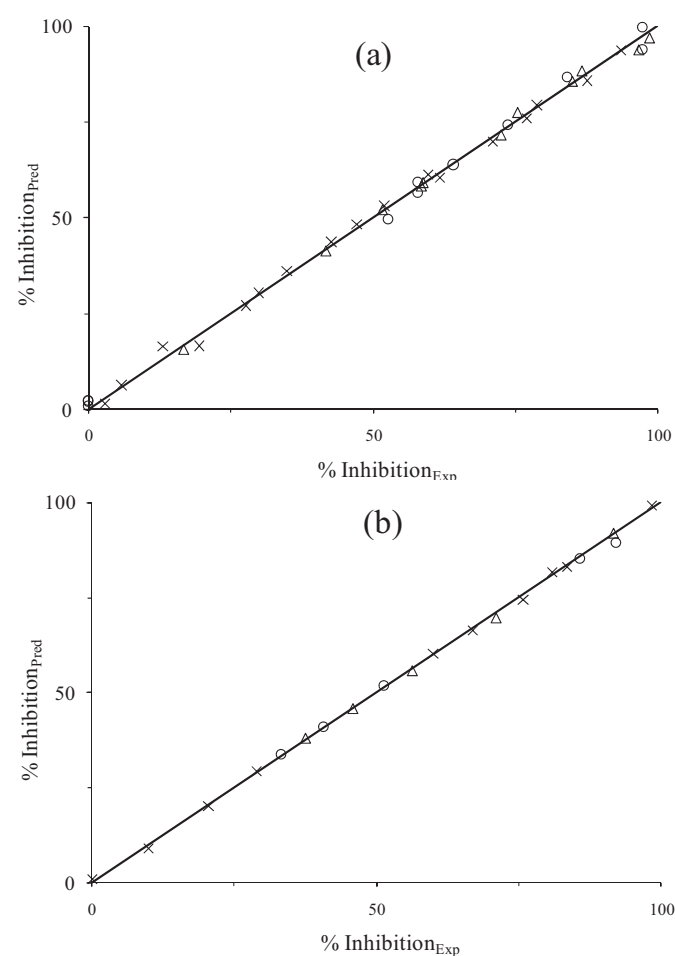
Fig. 2. The artificial neural network (ANN) structure for modelling antifungal activity against surface contaminant fungi, blue stain fungi, and phytopathogenic fungi.

Fig. 3. Plot of the predicted response by the artificial neural network (ANN) model and experimental values for surface contaminant fungi (×), blue stain fungi (○), and phytopathogenic fungi (Δ) for training (*a*) and test (*b*) sets.
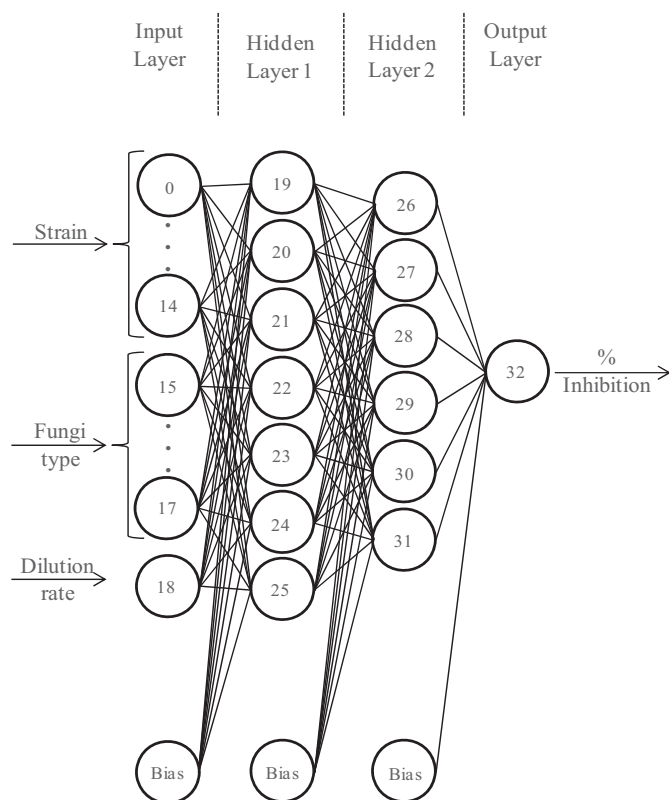
Sensitivity analysis of the ANN model

Typically, the efforts in data acquisition will be focused on the more relevant variables for model accuracy and dropping or ignoring those that matter least. Sensitivity analysis is a simple procedure that is applied after the modelling phase and analyzes the model responses when the inputs are changed. Sensitivity according to variance (Kewley et al. 2000) was used to compute the relative importance of the input variables for the selected model. The results are presented in Fig. 6 and reveal that the most informative variable is the fungal strain. However, despite the importance of this variable, the other variables (i.e., fungi type and dilution rate) also have a contribution to the final performance of the model.

Discussion

Glucose-limited chemostat cultures of *B. amyloliquefaciens* CCMI 1051 growing at different dilution rates showed different antifungal properties, indicating changes in metabolite concen-
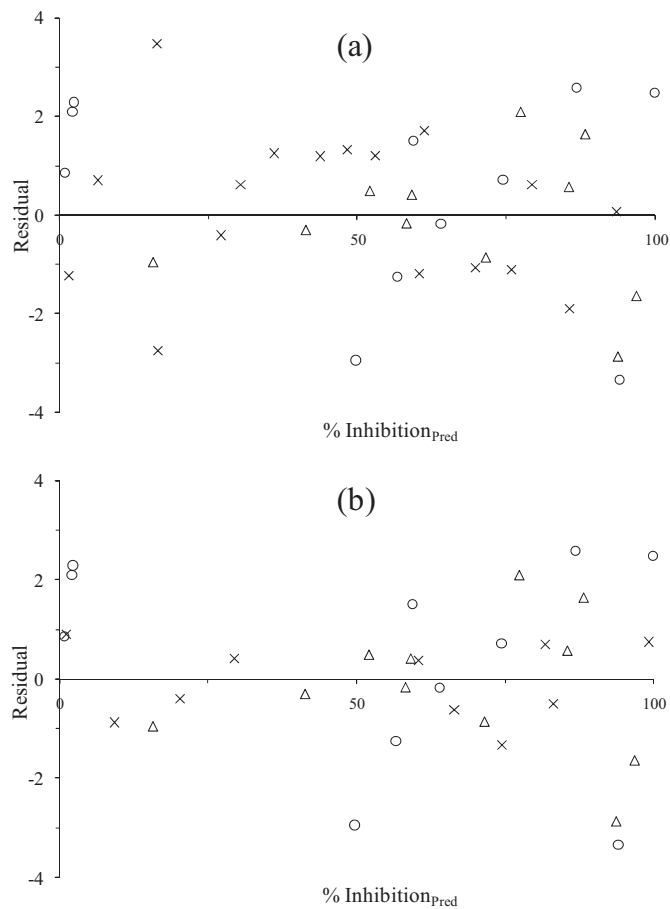
tration and thus the possibility to obtain a more intense extract activity. Continuous cultivation allows changes in the specific growth rate to be studied and allows the detection of physiological effects (Nobre et al. 2009). This strain showed the production of several compounds between 1000 and 1100 Da, comparable to that of iturin and surfactin compounds, and it also produces antifungal lipopeptides between 1436 and 1478 Da (Caldeira et al. 2011*a*).

The antifungal activity data for the biocompounds produced at different dilution rates by *B. amyloliquefaciens* against the contaminant fungi show a large dispersion of the data. These results seem to suggest that the bioactive culture broth produced at different dilution rates produces different sensitivities against the same

Table 4. Comparison between measured and evaluated responses of the selected artificial neural network (ANN) model for surface contaminant fungi (SCF), blue stain fungi (BSF), and phytopatogenic fungi (PhF) for the training and test sets.

	Response	
Fungi type	Training set	Test set
SCF	0.9971	0.9989
BBSF	0.9965	0.9982
PhF	0.9969	0.9974

Fig. 4. Plot of residuals versus the predicted response by the artificial neural network (ANN) model for surface contaminant fungi (×), blue stain fungi (○), and phytopathogenic fungi (Δ) for training (*a*) and test (*b*) sets.

Fig. 5. Response contours for antifungal inhibition, obtained with the 19-6-7-1 artificial neural network (ANN), against fungi upon dilution rate. Surface contaminant fungi (*a*), blue strain fungi (*b*), and phytopathogenic fungi (*c*). *Aspergillus niger* (A), *Penicillium expansum* (B), *Trichoderma koningii* (C), *Trichoderma pseudokoningii* (D), *Trichoderma harzianum* sp1 (E), *Trichoderma harzianum* sp2 (F), *Rhizopus* sp (G), *Cladosporium cladosporioides* (H), *Cladosporium resinae* sp1 (I), *Cladosporium resinae* sp2 (J), *Cladosporium* sp (K), *Botrytis cinerea* (L), *Cephalosporium* sp (M), *Fusarium oxysporum* (N), and *Fusarium solani* (O).
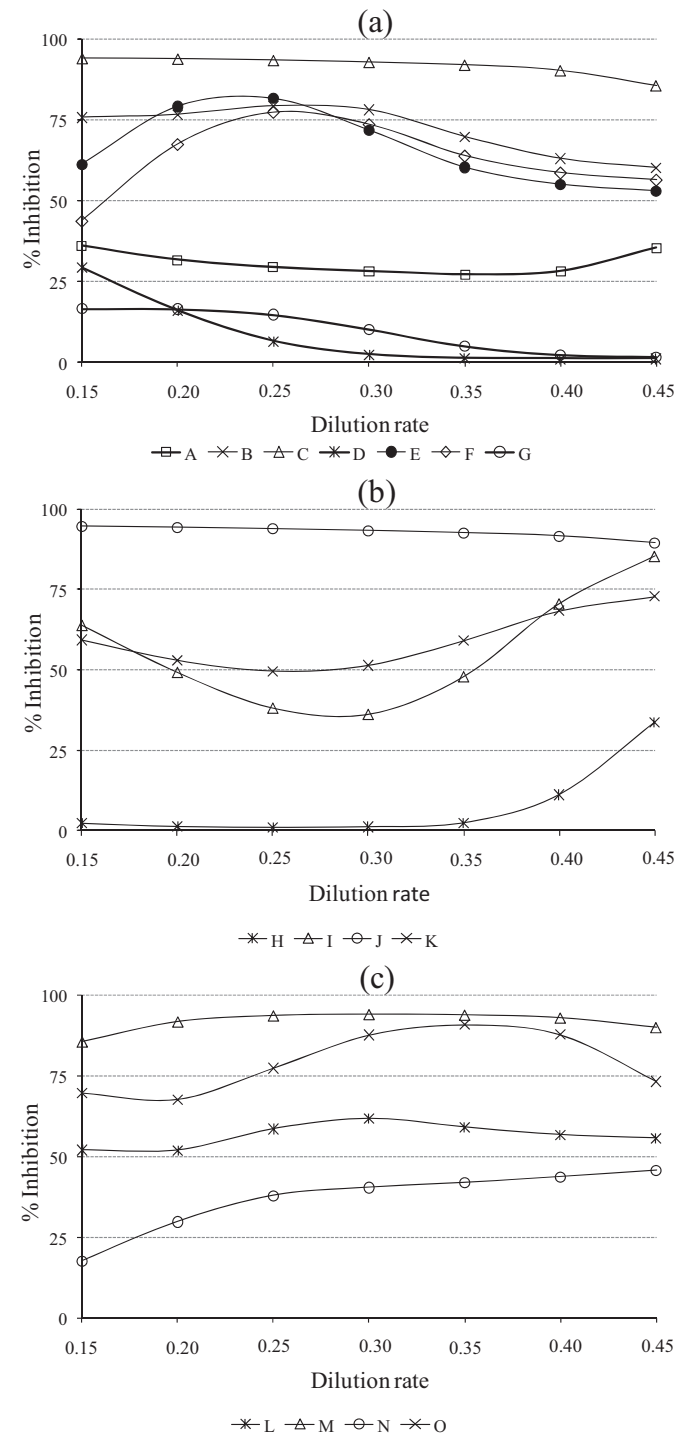

fungi (Figs. 1 and 5). In addition to this, the variation in antifungal activity may reflect differences in the sites of action or the abilities of the fungi to detoxify the metabolites (Fiddaman and Rossal 1994). Such variability may be attributed to different inhibitory effects against various fungal species. Indeed, this fact may be explained by the different responses of each species in the presence of the components of the metabolite pool (Caldeira et al. 2007). This synergistic effect was also described in cultures of *Bacillus subtilis* against *Botrytis cinerea* because of the simultaneous presence of surfactin and plipastatin (Tsuge et al. 1996), and in cultures of *B. amyloliquefaciens* against *Fusarium oxysporum* because of the presence of bacilomicin D and fengycin (Koumoutsi et al. 2004).

Bacillus amyloliquefaciens CCMI 1051 can be a good biological control agent, which may contribute to the alleviation of the exces-

Fig. 6. Relative importance of input variables for the selected artificial neural network (ANN) to model antifungal activity against surface contaminant fungi, blue stain fungi, and phytopathogenic fungi.
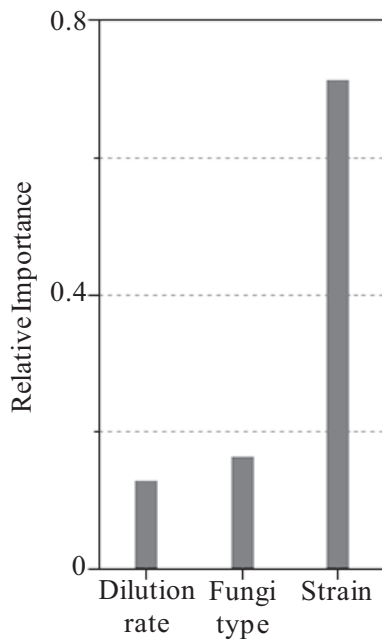
sive use of chemical pesticides, and in the end, the reduction of environmental pollution; in fact, these compounds, made of amino acids and fatty acid, are easily biodegradable in soils. Furthermore, multiple strains of *Bacillus* spp. have demonstrated the ability to stimulate plant defence responses because of the production of lipopeptides (Ongena et al. 2008).

Glucose-limited chemostat cultures of *B. amyloliquefaciens* CCMI 1051 growing at different dilution rates showed different antifungal properties. The agreement between the predicted and experimental values demonstrates the potential of the ANN model to address the antifungal activity of compounds, produced by *Bacillus* culture, at different dilution rates. Thus, ANNs can predict responses for different initial conditions, which exhaustively cover the entire experimental space, to maximize the production of antifungal bioactive compounds against SCF, BSF, and PhF. In fact, treatments based on the use of toxic chemicals that may accumulate in animal tissues (Greaves 1970) are subject to increasing restrictions. Thus, biological control (Leifert et al. 1995; Emmert and Handelsman 1999; Feio et al. 2004) is being seriously considered for wood protection, not only against sapstain but also against decay fungi (Bruce 1998; Caldeira et al. 2008). The prediction of bioactive compound activity against wood contaminant fungi using ANNs can become an important tool in the development of an integrated strategy for the management of fungal pathogens and undesirable saprobes.

References

Aminov, R.I. 2009. The role of antibiotics and antibiotic resistance in nature. Environ. Microbiol. **11**: 2970–2988. doi:10.1111/j.1462-2920.2009.01972.x. PMID:19601960.

Beer, Z., Wingfieldb, B., and Wingfield, M. 2003. The occurrence of *Ophiostoma piliferum*-like fungi on pulpwood chips and other wood sources in South Africa. S. Afr. J. Sci. **99**: 34–36.

Bruce, A. 1998. Biological control of wood decay. *In* Forest products biotechnology. *Edited by* A. Bruce and J.W. Palfreyman. Taylor & Francis, London. pp. 251–266.

Caldeira, A.T., Savluchinske-Feio, S., Arteiro, J.M., and Roseiro, J.C. 2006. Antimicrobial activity of steady state cultures of *Bacillus* sp. CCMI 1051 against wood contaminant fungi. Biochem. Eng. J. **30**: 231–236. doi:10.1016/j.bej.2006.04.005.

Caldeira, A.T., Savluchinske-Feio, S., Roseiro, J.C., and Arteiro, J.M. 2007. *Bacillus amyloliquefaciens* CCMI 1051 in vitro activity against wood contaminant fungi. Ann. Microbiol. **57**: 29–34. doi:10.1007/BF03175046.

Caldeira, A.T., Savluchinske-Feio, S., Arteiro, J.M., Coelho, A., and Roseiro, J.C. 2008. Environmental dynamics of *Bacillus amyloliquefaciens* CCMI 1051 antifungal activity under different nitrogen patterns. J. Appl. Microbiol. **104**: 808–816. doi:10.1111/j.1365-2672.2007.03601.x. PMID:17953685.

Caldeira, A.T., Arteiro, J.M., Coelho, A.V., and Roseiro, J.C. 2011*a*. Combined use of LC–ESI-MS and antifungal tests for rapid identification of bioactive lipopeptides produced by *Bacillus amyloliquefaciens* CCMI 1051. Process Biochem. **46**: 1738–1746. doi:10.1016/j.procbio.2011.05.016.

Caldeira, A.T., Arteiro, J.M., Roseiro, J.C., Neves, J., and Vicente, H. 2011*b*. An artificial intelligence approach to *Bacillus amyloliquefaciens* CCMI 1051 cultures: application to the production of anti-fungal compounds. Bioresour. Technol. **102**: 1496–1502. doi:10.1016/j.biortech.2010.07.080. PMID:20801027.

Cortez, P., Rocha, M., and Neves, J. 2004. Evolving time series forecasting ARMA models. J. Heuristics, **10**: 415–429. doi:10.1023/B:HEUR.0000034714.09838.1e.

El-Kassas, H. 2008. Decolorization and detoxification of Direct Fast Red 8B by a marine Fungus *Aspergillus japonicus* HK. World Appl. Sci. J. **5**: 460–468.

El-Kassas, H., and Khairy, H. 2009. A trial for biological control of a pathogenic fungus (*Fusarium solani*) by some marine microorganisms. Am.-Eurasian J. Agric. Environ. Sci. **5**: 434–440.

Emmert, E.A.B., and Handelsman, J. 1999. Biocontrol of plant disease: a (gram-) positive perspective. FEMS Microbiol. Lett. **171**: 1–9. doi:10.1111/j.1574-6968.1999.tb13405.x. PMID:9987836.

Feio, S., Barbosa, A., Cabrita, M., Nunes, L., Esteves, A., Roseiro, J.C., and Curto, M.M. 2004. Antifungal activity of *Bacillus subtilis* 355 against wood-surface contaminant fungi. J. Ind. Microbiol. Biotechnol. **31**: 199–203. PMID:15197600.

Fiddaman, P.J., and Rossal, S. 1994. Effect of substrate on the production of antifungal volatiles from *Bacillus subtilis*. J. Appl. Bacteriol. **76**: 395–405. doi:10.1111/j.1365-2672.1994.tb01646.x. PMID:8200865.

Galushkin, A.I. 2007. Neural networks theory. Springer, New York.

Greaves, H. 1970. The effects of selected bacteria and Actnomycetes on the decay capacity of some wood-rotting fungi. Mater. Organismen, **5**: 265–279.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. 2009. The WEKA data mining software: an update. SIGKDD Explor. **11**: 10–18. doi:10.1145/1656274.1656278.

Han, J., and Kamber, M. 2006. Data mining: concepts and techniques. 2nd ed. Morgan Kauffmann Publishers, San Francisco, Calif.

Haykin, S. 2008. Neural networks and learning machines. 3rd ed. Prentice Hall, New York.

Kewley, R., Embrechts, M., and Breneman, C. 2000. Data strip mining for the virtual design of pharmaceuticals with neural networks. IEEE Trans. Neural Netw. **11**: 668–679. doi:10.1109/72.846738. PMID:18249794.

Koumoutsi, A., Chen, X.-H., Henne, A., Liesegang, H., Hitzeroth, G., Franke, P., Vater, J., and Borriss, R. 2004. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. J. Bacteriol. **186**: 1084–1096. doi:10.1128/JB.186.4.1084-1096.2004. PMID:14762003.

Kowall, M., Vater, J., Kluge, B., Stein, T., Franke, P., and Ziessow, D. 1998. Separation and characterization of surfactin isoforms produced by *Bacillus subtilis* OKB 105. J. Colloid Interface Sci. **204**: 1–8. doi:10.1006/jcis.1998.5558. PMID:9665760.

Leifert, C., Li, H., Chidburee, S., Hampson, S., Workman, S., Sigee, D., Epton, H., and Harbour, A. 1995. Antibiotic production and biocontrol activity by *Bacillus subtilis* CL27 and *Bacillus pumilus* CL45. J. Appl. Bacteriol. **78**: 97–108. doi:10.1111/j.1365-2672.1995.tb02829.x. PMID:7698955.

McBride, G.B. 2005. Using statistical methods for water quality management issues, problems and solutions. John Wiley & Sons, Hoboken, N.J.

Moyne, A.L., Shelby, R., Cleveland, T.E., and Tuzun, S. 2001. Bacillomycin D: an iturin with antifungal activity against *Aspergillus flavus*. J. Appl. Microbiol. **90**: 622–629. doi:10.1046/j.1365-2672.2001.01290.x. PMID:11309075.

Nobre, A., Duarte, L.C., Roseiro, J.C., and Gírio, F.M. 2009. A physiological and enzimatic study of *Debaryomyces hansenii* growth on xylose and oxygen-limited chemostats. Appl. Microbiol. Biotechnol. **59**: 509–516. doi:10.1007/s00253-002-1050-4. PMID:12172618.

Ongena, M., Jourdan, E., Adam, A., Paquot, M., Brans, A., Joris, B., Arpigny, J.L., and Thonart, P. 2008. Surfactin and fengycin lipopeptides of *Bacillus subtilis* as elicitors of induced systemic resistance in plants. Environ. Microbiol. **10**: 1823–1830. doi:10.1111/j.1462-2920.2008.01603.x. PMID:18397310.

Pabel, C.T., Vater, J., Wilde, C., Franke, P., Hofemeister, J., Adler, B., Bringmann, G., Hacker, J., and Hentschel, U. 2003. Antimicrobial activities and matrix-assisted laser desorption/ionization mass spectrometry of *Bacillus* isolates from the marine sponge *Aplysina aerophoba*. Mar. Biotechnol. **5**: 424–434. doi:10.1007/s10126-002-0088-8. PMID:14730425.

Rumelhart, D., Hinton, G., and Williams, R. 1986. Learning internal representation by error propagation. *In* Parallel distributed processing. *Edited by* David E. Rumelhart and James L. McCleland. MIT Press, Cambridge, Massachusetts. pp. 318–362.

Sathish, T., and Prakasham, R.S. 2010. Enrichment of glutaminase production by *Bacillus subtilis* RSP-GLU in submerged cultivation based on neural network – genetic algorithm approach. J. Chem. Technol. Biotechnol. **85**: 50–58. doi:10.1002/jctb.2267.

Schubert, M., Fink, S., and Schwarze, F. 2008. Evaluation of *Trichoderma* spp. as a biocontrol agent against wood decay fungi in urban trees. Biol. Control, **45**: 111–123. doi:10.1016/j.biocontrol.2008.01.001.

Silva, R.G., Pinotti, L.M., Cruz, A.J.G., Giordano, R.L.C., and Giordano, R.C. 2008. Artificial neural networks to infer biomass and product concentration during the production of penicillin G acylase from *Bacillus megaterium*. J. Chem. Technol. Biotechnol. **83**: 739–749. doi:10.1002/jctb.1864.

Souza, J., Matwin, S., and Japkowicz, N. 2002. Evaluating data mining models: a pattern language. *In* Proceedings of the 9th Conference on Pattern Language of Programs, Urbana, Illinois, 8–12 September 2002. pp. 1–23.

Soytong, K., Srinon, W., Ratanacherdchai, K., Kanokmedhakul, S., and Kanokmedhakul, K. 2005. Application of antagonistic fungi to control anthracnose disease of grape. Agric. Technol. **1**: 33–42.

Tsuge, K., Ano, T., and Shoda, M. 1996. Isolation of a gene essential for biosynthesis of the lipopeptide antibiotics plipastatin B1 and surfactin in *Bacillus subtilis* YB8. Arch. Microbiol. **165**: 243–251. doi:10.1007/s002030050322. PMID:8639027.

Vishniac, W., and Santer, M. 1957. The Thiobacilli. Bacteriol. Rev. **21**: 195–213. PMID:13471458.

Weindling, R. 1932. *Trichoderma lignorum* as a parasite of other soil fungi. Phytopathology, **22**: 837–845.

Zafar, M., Kumar, S., and Dhiman, A.K. 2012. Artificial intelligence based modeling and optimization of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) production process by using *Azohydromonas lata* MTCC 2311 from cane molasses supplemented with volatile fatty acids: a genetic algorithm paradigm. Bioresour. Technol. **104**: 631–641. doi:10.1016/j.biortech.2011.10.024. PMID:22074908.