Combining Digital Image Processing and Machine Learning is Useful for the Early Detection of Salinity and Drought Stresses in Cucumber

Parvin Mohammadi[1] (✉) and Keyvan Asefpour Vakilian[2]

[1] Department of Agrotechnology, College of Abouraihan, University of Tehran, Tehran, Iran
p.mohammadi68@ut.ac.ir

[2] Department of Biosystems Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

Abstract. Timely detection of plant abiotic stresses and their type and severity can be beneficial in order to prevent the loss of yield in crop production systems. This study introduces an image processing-based method to detect the type and severity of salinity and drought stress, as crucial abiotic stresses, in cucumber plants. Plants were cultivated in the greenhouse environment. Plant morphological features in the form of textural features were measured from the images captured from the plants. Sampling was performed five times at 3-day intervals beginning with applying the abiotic stresses. Measurements were conducted by transferring three leaves randomly selected from each pot to a chamber with artificial lighting equipped with a camera for image acquisition. The artificial neural network (ANN) was used based on the image textural features of the leaves as inputs and stress type and severity as output. The parameters of the network were optimized using a sophisticated optimization algorithm to achieve the most efficient machine. As a robust evolutionary method, the genetic algorithm was used to optimize the architecture of the ANNs. The results revealed that the image textural features for training ANNs optimized using the genetic algorithm could predict the type and severity of the plant abiotic stresses with MSE and R^2 values of 0.092 and 0.74, respectively. Since the machine was able to perform a reliable stress prediction within a short period after applying the stress, the proposed method can be used for the early detection of salinity and drought stresses in cucumber plants.

Keywords: Artificial neural networks · Severe stress conditions · Image textural features · Genetic algorithm · Optimization

1 Introduction

While the morphological, physiological, and biochemical attributes of plants undergo significant alterations due to biotic and abiotic stresses, these alterations are not exclusive to particular stressors [1]. To illustrate, distinct stress origins can lead to comparable alterations in characteristics such as plant height, as well as the weights of roots and

E. Cavallo et al. (Eds.): ANKAgEng 2023, LNCE 458, pp. 373–382, 2024.
https://doi.org/10.1007/978-3-031-51579-8_33

shoots, as morphological attributes of plants. To enhance the comprehension of plant responses within stress-related investigations, it is preferable to focus on traits at a molecular level that are almost exclusively associated with specific stress conditions. However, these methods are costly and time-consuming. Finding a reliable solution to determine the severity and type of the stresses based on the morphological traits of plants can be beneficial for farmers and plant pathologists since they are cost- and time-effective [2].

Abiotic stresses, including light, temperature, deficiency and excess of nutrients, drought, salinity, elevated carbon dioxide levels, and contamination by heavy metals, constitute some of the prevalent challenges encountered by crops [3]. Various types of stresses, such as salt and drought, are environmentally undesirable and can significantly impact successful crop production [4]. Elevated levels of salt and drought stress have a notable influence on the physiological and biochemical processes of numerous plant species [5]. The response of plants to drought stress can be assessed based on physiological indicators like relative water content, stomatal responses, active oxygen species, and the activation of antioxidative enzymes [6]. Salt stress can impede plant growth and development through mechanisms involving osmotic stress, ion toxicity, disrupted nutrient balance, and oxidative stress [7]. Cucumber (*Cucumis sativus* L.), particularly during its initial growth stages, is highly susceptible to both drought and salt stress [8]. Severe instances of drought and salinity stresses have adverse consequences on the growth, photosynthesis, biochemical composition, and textural attributes of cucumber fruits [9, 10].

Supervised and unsupervised machine learning algorithms are being used in plant science to model multivariate plant growth problems [11]. With the help of training data in the database, the machine learns the complex non-linear patterns between inputs and output [12–14]. When presented with a dataset that encompasses morphological characteristics of cucumber plants subjected to salt and drought stresses, a crucial question arises: Which specific features should be employed to train machine learning methods in order to ensure the accurate identification of stress type and severity? Consequently, this study undertakes the imposition of varying levels of salt and drought stresses on cucumber plants, both individually and concurrently. Throughout the experimental process, diverse attributes of the plants based on the images captured from the plant leaves are measured, contributing to the creation of an informative database that holds utility for the training of machine learning models. Subsequently, the artificial neural network (ANN) algorithm is employed, and its performance is amplified through the utilization of a metaheuristic optimization technique, namely the genetic algorithm (GA), to identify the category and intensity of stress based on the provided database.

The primary objectives of this research encompass: (a) assessing the performance of image-processing-based morphological traits in identifying salt and drought stresses in cucumber plants, and (b) determining a machine learning algorithm, honed through intricate evolutionary processes, tailored specifically for the accurate detection of stress conditions. A search in bibliographic databases reveals that a combination of image processing techniques and machine learning methods optimized with metaheuristic algorithms has not been studied so far. Therefore, the novelty of this work involves developing sophisticated machines with performances improved with novel optimization methods to predict stress conditions in plants by having their images captured by RGB cameras.

2 Material and Methods

Figure 1 demonstrates the flowchart of the present study. Morphological attributes, encompassing energy, entropy, and local homogeneity, were extracted from images acquired from plant leaves using an image processing module. These attributes, as well as the sampling time, designated as machine inputs, were subsequently employed for the training of ANNs, with their parameters, including the number of hidden (learner) layers, the number of neurons in each layer, and their weight and bias parameters, fine-tuned through GA. The machine-generated outputs corresponded to the identification of stress type and quantification of its severity.

2.1 Plant Material and Experimental Design

Cucumber seeds were procured from a local vendor and subsequently subjected to surface sterilization using NaClO for a duration of 1 min, followed by thorough rinsing with distilled water. The seeds were initially cultivated in plug trays utilizing a growth medium comprising a blend of vermiculite, peat, and perlite in a ratio of 3:2:1. Upon reaching the 21st day of growth, the seedlings were transplanted into 5-L pots containing the same growth medium as that of the plug trays. Nutrient fertilization management, encompassing both macro- and micro-nutrients, was executed in accordance with the guidelines outlined in the literature. To sustain optimal growth conditions, the growth room's relative humidity was maintained at $70 \pm 5\%$, accompanied by a light intensity of 120 μmolm^{-2} s^{-1}.

A consistent and uniform irrigation regimen was upheld across all pots for a duration of 60 days. Subsequent to this period, the control group of plants received daily supplementation with tap water to sustain soil water potential at a level proximate to the field capacity. Contrastingly, plants subjected to drought treatment underwent water deprivation until the soil water potential diverged from the field capacity. Monitoring of soil moisture was carried out on a daily basis utilizing a time-domain reflectometry (TDR) device (PMS-714, LUTRON, Taiwan).

To explore various levels of drought stress severity, the soil moisture content was meticulously regulated to remain nonlethal and surpass the wilting point. Specifically, soil moisture was upheld at 100% (W_0, control), 80% (W_1), 60% (W_2), and 40% (W_3) of the field capacity. For plants undergoing salinity treatment, irrigation was administered using water containing distinct concentrations of NaCl: 0 (S_0, control), 20 mM (S_1), 40 mM (S_2), and 60 mM (S_3). Notably, concentrations exceeding 100 mM were deemed potentially lethal to young plants. Various studies have recommended similar drought and salinity levels to apply stresses to cucumber plants. It should be noted that both abiotic and biotic stresses can exert changes in various characteristics of cucumber plants [15].

Plant attributes were assessed five times, each separated by 3-day intervals, commencing from the initiation of the stress application. Experiments were done with four replications for each treatment. The total number of pots was $4 \times 4 \times 5 \times 3 = 240$.

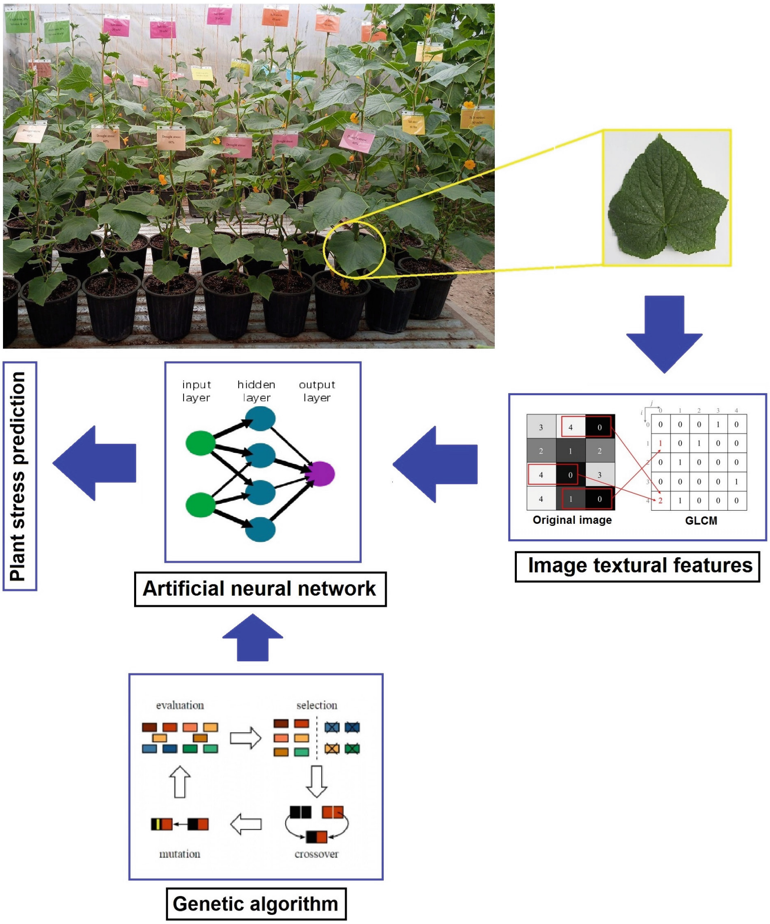

Fig. 1. Flowchart of the present study

2.2 Morphological Measurements

The digital image processing technique was employed to extract morphological attributes from leaf images. The image capture setup encompassed three components: a dark chamber, a CCD digital camera, and a 200-LED lighting array featuring a 70° viewing angle. This arrangement enhanced light uniformity within the area of interest. From each captured image, the leaf region of interest was isolated from the background through Canny edge detection, a process known as image segmentation [16]. Grayscale transformation was applied to convert the images into grayscale format by combining the weighted sums of the red (R), green (G), and blue (B) color components.

To capture the spatial correlation of gray-level values, a Gray-level Co-occurrence Matrix (GLCM) was utilized [17, 18]. Each element (i, j) within the GLCM denoted the frequency of occurrences where a pixel with value i was found horizontally adjacent to a pixel with value j. While a comprehensive set of 21 textural parameters was previously identified, recent studies have indicated that the use of only three textural variables, adopted in this study, can effectively discern plant health: entropy (Eq. 1), energy (Eq. 2), and local homogeneity (Eq. 3), as established by Story et al. [19]

$$\text{Entropy} = -\sum_i \sum_j p(i,j)\log(p(i,j)) \tag{1}$$

$$\text{Energy} = \sum_i \sum_j p(i,j)^2 \tag{2}$$

$$\text{Local homogeneity} = \sum_i \sum_j \frac{p(i,j)}{1+(i-j)^2} \tag{3}$$

where $p(i, j)$ is the (i, j)-th element of the GLCM. The program for image processing was written with MATLAB R2018b programming environment.

2.3 Machine Learning Method

The obtained measurements of the various attributes under different stress conditions were compiled to form a database tailored for utilization in machine learning applications. Given that statistical analysis revealed insignificant effects of replication on the outcomes, mean values were employed in the preparation of the database. While statistical regression models offer mathematical equations for estimating the dependent variable based on input features, it's generally advisable to limit the number of input features to 2 due to the complexity of parameter estimation in high-dimensional problems. In contrast, machine learning techniques are capable of assimilating databases including hundreds of input features alongside corresponding dependent variables. In this study, ANN was employed to predict plant stress utilizing attribute values as inputs. For optimization purposes, GA, which is an evolutionary technique was employed to optimize the architecture of ANN during the training.

GA is a robust metaheuristic optimization approach that operates through a population comprising a collection of chromosomes, each representing a potential solution to the problem. Through an iterative process, termed generations, individuals yielding more favorable values for the objective function are allowed to persist and undergo crossover operations in subsequent iterations. In simpler terms, pairs of chromosomes acting as parents generate more dependable offspring in each generation by discarding weaker solutions associated with undesirable values of the objective function. Table 1 presents the parameters that were taken into consideration for the optimization methods in this study. The primary aim of these methods was to minimize the error associated with predicting stress levels utilizing the ANNs. In other words, the objective (fitness) function of GA during the optimization was to minimize the error of prediction performed by ANN. The optimization process ceased upon reaching the predefined maximum number of iterations. The parameters of the network that were optimized by GA to reach the lowest prediction error included the number of hidden (learner) layers, the number of neurons in each layer, and their weight and bias parameters.

In this study, a program was developed using the MATLAB programming environment to harness machine learning techniques based on the textural features extracted

Table 1. Parameters of the optimization algorithms

Optimization method	Parameter	Value
GA	Population size	100
	Maximum number of iterations	500
	Mutation rate	0.1
	Crossover percentage	0.5

from leaf images, as well as sampling time, as inputs and stress type and severity as outputs. The program encompassed the formulation of the ANN regression model designed to predict the model's output while utilizing the four input features (three textural features and one sapling time). The configuration of the ANN architecture, including the number of hidden (learner) layers (ranging from 1 to 3), the number of neurons within each layer (ranging from 5 to 15), and their corresponding weight and bias parameters, were systematically optimized employing advanced optimization algorithms. The aim was to attain the most efficient machine-learning model possible. The performance of the models was evaluated based on the mean squared error (MSE) (Eq. 4) and coefficient of determination (R^2) (Eq. 5)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_\mathrm{p} - x_\mathrm{o})^2 \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_\mathrm{p} - x_\mathrm{o})^2}{\sum_{i=1}^{n}(x_\mathrm{o} - \overline{x}_\mathrm{o})^2} \tag{5}$$

where x_o is the severity level of the stress applied to the plants, x_p is the predicted value using ANNs, and n is the number of samples. The lower the MSE and higher the R^2, the better the performance of the machine learning model is.

3 Results and Discussion

Prior research has established the noteworthy impacts of both salinity and drought stress on the morphological attributes of cucumber plants. Nonetheless, accurately pinpointing the occurrence of these stresses poses a challenge due to the intricate and multifaceted nature of plant responses. These responses often exhibit delays in reacting to the underlying stress factors, rendering the process of pinpointing the exact source of stress highly intricate.

In control plants grown under optimal conditions, their leaves displayed robust health and vivid colors, resulting in elevated entropy levels. Conversely, plants subjected to salinity and drought treatments exhibited reduced surface structural complexity, leading to diminished entropy in their leaf images. As time progressed, the leaves of control plants adopted a deeper green hue, causing a decline in image energy levels, as observed in previous studies [19]. On the contrary, the appearance of a yellowish tinge in treated plants contributed to an increase in image energy levels. Furthermore, the local homogeneity of images originating from control plants experienced a decrease as these plants evolved, displaying a spectrum of green shades and becoming more colorful during their growth. In contrast, treated plants, characterized by a uniform coloration, exhibited higher levels of homogeneity in their images. Similar results and findings are reported by Asefpour Vakilian and Massah [17, 20].

In alignment with the objectives outlined in this study, the utilization of machine learning techniques, in conjunction with morphological traits of cucumber plants, is pursued to enable the targeted identification of stress severity caused by abiotic factors. This approach aims to achieve accurate and specific detection of stress levels, particularly during the initial stages of stress imposition.

Table 2 presents the evaluation of morphological variables in their capacity to predict plant stress, as investigated within the scope of this study. The results indicate that the employed machine learning approach was proficient in identifying stress instances under conditions of severe salinity and drought. This capability holds particular significance for farmers, as the most substantial reduction in agricultural yield is observed when crops face severe abiotic stresses during their growth cycle.

Table 2. Performance evaluation of ANN and GA-ANN in the prediction of plant stress

	Training		Validation		Test	
ANN architecture	MSE	R^2	MSE	R^2	MSE	R^2
4-5-1	0.397	0.39	0.460	0.34	0.587	0.25
4-6-1	0.316	0.46	0.373	0.41	0.385	0.40
4-7-1	0.218	0.56	0.316	0.46	0.422	0.37
4-8-1	0.338	0.44	0.487	0.32	0.409	0.38
4-9-1	0.255	0.52	0.306	0.47	0.361	0.42
4-10-1	0.227	0.55	0.201	0.58	0.265	0.51
4-11-1	0.169	0.62	0.275	0.50	0.361	0.42
4-12-1	0.338	0.44	0.338	0.44	0.422	0.37
4-13-1	0.350	0.43	0.397	0.39	0.305	0.47
4-14-1	0.316	0.46	0.373	0.41	0.447	0.35
4-15-1	0.193	0.59	0.275	0.50	0.245	0.53
GA-ANN	0.087	0.75	0.083	0.76	0.092	0.74

Notably, the highest R^2 value for the test data achieved by the ANN machine in predicting both stresses simultaneously reached 0.74, achieved through an architecture featuring three hidden layers and optimized by GA. However, as Table 2 reveals, in the case of not using GA to optimize the structure of ANN, the R^2 values for test data did not exceed 0.53. Proper determination of the weight and bias of the learner neurons and the number of hidden layer neurons is important since they affect not only the convergence of the network, but also the prediction performance. In order to understand the effect of the ANN parameters on the prediction, Table 2 shows the changes in the prediction performance due to the change of learner neurons in the hidden layer.

The morphological variables utilized in this study were image textural features, assessed through probability-density functions on GLCM. These features were extracted from leaf images obtained through an image acquisition system and subsequently transferred to a computer for in-depth analysis.

Furthermore, the utilization of GA as a metaheuristic optimization technique proved instrumental in enhancing the predictive capability of the machine learning model for plant stress detection. This improvement was evident from the significant decrease in MSE values, while a remarkable increase in R^2 values. The successful integration of GA optimization resulted in the development of an efficient machine learning approach, proficient in utilizing variables derived from the image processing technique to accurately predict and identify severe levels of abiotic stresses in cucumber plants. It should be noted that considering the size of the dataset investigated in this work, it was not possible to utilize deep learning for the prediction of stresses. Therefore, ANN, as a base regression model optimized by GA was used for the prediction.

In the context of other morphological attributes, such as plant height, shoot weight, and root weight, that were not considered in this study; while these traits generally showed susceptibility to increasing stress severity, they did not demonstrate specificity in terms of stress type. This aligns with findings from earlier studies [1, 21, 22] which highlighted that these traits' responses to stress were not unique to stress types and did not offer reliable discrimination between different types of stressors. Therefore, the use of image textural features as the only morphological features of the plants exerted acceptable results in predicting the outputs, i.e., type and severity of the stress.

4 Conclusions

An effort to introduce a promising image processing-based method to detect the type and severity of two main abiotic stresses, i.e., salinity and drought, in cucumber plants is reported in this paper. Treatments were selected in levels to apply a range of mild to rather severe stress conditions. The performance assessment of image-processing-based morphological traits showed that training these features to a machine such as ANN can identify the stress in the plants with acceptable efficiency. ANN, equipped with GA as an intricate evolutionary optimization technique can increase the MSE and R^2 values of the prediction up to 0.092 and 0.74, respectively. Due to the fact that the machine was capable of performing a reliable stress prediction within a short period after applying the stress, the proposed technique can be utilized for the early detection of abiotic stresses in cucumber plants.

Although physiological and biochemical features of plants such as enzymatic activities and microRNA regulation might provide us with higher prediction performance, they all require expensive laboratory equipment and time-consuming protocols [23].

Declaration of Competing Interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Asefpour Vakilian K (2019) Gold nanoparticles-based biosensor can detect drought stress in tomato by ultrasensitive and specific determination of miRNAs. Plant Physiol Biochem 145:195–204
2. Javidan SM, Banakar A, Asefpour Vakilian, K, Ampatzidis Y (2023) Tomato leaf diseases classification using image processing and weighted ensemble learning. Agronomy J (in Press)
3. Gong Z et al (2020) Plant abiotic stress response and nutrient use efficiency. Sci China Life Sci 63:635–674
4. Chandra S, Roychoudhury A (2020) Penconazole, paclobutrazol, and triacontanol in overcoming environmental stress in plants. In: Roychoudhury A, Tripathi DK (eds) Protective chemical agents in the amelioration of plant abiotic stress: biochemical and molecular perspectives. John Wiley & Sons, Hoboken, pp 510–534
5. Massonnet C, Costes E, Rambal S, Dreyer E, Regnard JL (2007) Stomatal regulation of photosynthesis in apple leaves: evidence for different water-use strategies between two cultivars. Ann Bot 100(6):1347–1356
6. Batool T et al (2020) Plant growth promoting rhizobacteria alleviates drought stress in potato in response to suppressive oxidative stress and antioxidant enzymes activities. Sci Rep 10:16975
7. Behera LM, Hembram P (2021) Advances on plant salinity stress responses in the post-genomic era: a review. J Crop Sci Biotechnol 24:117–126
8. Du C, Li H, Liu C, Fan H (2021) Understanding of the postgerminative development response to salinity and drought stresses in cucumber seeds by integrated proteomics and transcriptomics analysis. J Proteomics 232:104062
9. Ouzounidou G, Giannakoula A, Ilias I, Zamanidis P (2016) Alleviation of drought and salinity stresses on growth, physiology, biochemistry and quality of two *Cucumis sativus* L. cultivars by Si application. Br J Bot 39:531–539
10. Alsaeedi A, El-Ramady H, Alshaal T, El-Garawany M, Elhawat N, Al-Otaibi A (2019) Silica nanoparticles boost growth and productivity of cucumber under water deficit and salinity stresses by balancing nutrients uptake. Plant Physiol Biochem 139:1–10
11. Javidan SM, Banakar A, Asefpour Vakilian K, Ampatzidis Y (2023) Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. Smart Agric Technol 3:100081
12. Hashemi A, Asefpour Vakilian K, Khazaei J, Massah J (2014) An artificial neural network modeling for force control system of a robotic pruning machine. J Inf Organ Sci 38(1):35–41
13. Sarlaki E, Sharif Paghaleh A, Kianmehr MH, Asefpour Vakilian K (2021) Valorization of lignite wastes into humic acids: process optimization, energy efficiency and structural features analysis. Renew Energy 163:105–122
14. Esmaili M et al (2021) Assessment of adaptive neuro-fuzzy inference system (ANFIS) to predict production and water productivity of lettuce in response to different light intensities and CO_2 concentrations. Agric Water Manag 258:107201
15. Liao L, Hu Z, Liu S, Yang Y, Zhou Y (2021) Characterization of Germin-like proteins (GLPs) and their expression in response to abiotic and biotic stresses in cucumber. Horticulturae 7(10):412
16. Asefpour Vakilian K, Massah J (2013) An artificial neural network approach to identify fungal diseases of cucumber (*Cucumis sativus* L.) plants using digital image processing. Arch Phytopathol Plant Prot 46(13):1580–1588

17. Asefpour Vakilian K, Massah J (2017) A farmer-assistant robot for nitrogen fertilizing management of greenhouse crops. Comput Electron Agric 139:153–163
18. Javidan SM, Banakar A, Asefpour Vakilian K, Ampatzidis Y (2022) A feature selection method using slime mould optimization algorithm in order to diagnose plant leaf diseases. In: 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), pp. 1–5, Behshahr, Iran
19. Story D, Kacira M, Kubota C, Akoglu A, An L (2010) Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. Comput Electron Agric 74(2):238–243
20. Asefpour Vakilian K, Massah J (2012) Design, development and performance evaluation of a robot to early detection of nitrogen deficiency in greenhouse cucumber (*Cucumis sativus*) with machine vision. Int J Agric Res Rev 2:448–454
21. Schwarz D, Rouphael Y, Colla G, Venema JH (2010) Grafting as a tool to improve tolerance of vegetables to abiotic stresses: thermal stress, water stress and organic pollutants. Sci Hortic 127(2):162–171
22. Khan MM, Al-Mas'oudi RS, Al-Said F, Khan I (2013) Salinity effects on growth, electrolyte leakage, chlorophyll content and lipid peroxidation in cucumber (*Cucumis sativus* L.). In: International conference on food and agricultural sciences. IACSIT Press. Malaysia, pp 28–32
23. Lu XY, Huang XL (2008) Plant miRNAs and abiotic stress responses. Biochem Biophys Res Commun 368(3):458–462