

OPEN

# Machine learning improves our knowledge about miRNA functions towards plant abiotic stresses

Keyvan Asefpour Vakilian <sup>1,2</sup>

During the last two decades, human has increased his knowledge about the role of miRNAs and their target genes in plant stress response. Biotic and abiotic stresses result in simultaneous tissue-specific up/down-regulation of several miRNAs. In this study, for the first time, feature selection algorithms have been used to investigate the contribution of individual plant miRNAs in *Arabidopsis thaliana* response towards different levels of several abiotic stresses including drought, salinity, cold, and heat. Results of information theory-based feature selection revealed that miRNA-169, miRNA-159, miRNA-396, and miRNA-393 had the highest contributions to plant response towards drought, salinity, cold, and heat, respectively. Furthermore, regression models, i.e., decision tree (DT), support vector machines (SVMs), and Naïve Bayes (NB) were used to predict the plant stress by having the plant miRNAs' concentration. SVM with Gaussian kernel was capable of predicting plant stress ( $R^2 = 0.96$ ) considering miRNA concentrations as input features. Findings of this study prove the performance of machine learning as a promising tool to investigate some aspects of miRNAs' contribution to plant stress responses that have been undiscovered until today.

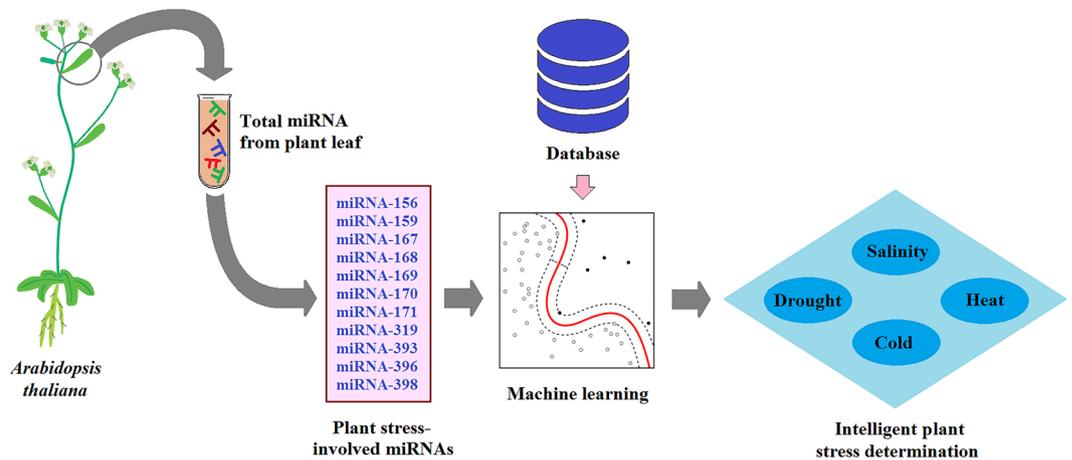
microRNAs (miRNAs) are small single-stranded RNAs with low protein-coding potential<sup>1</sup>. Although plant miRNAs target only a small number of mRNAs (<1%)<sup>2</sup>, the role of miRNA-controlled gene regulation in plants cannot be neglected because most of the target mRNAs participate in most plant developmental processes<sup>3,4</sup>. Furthermore, there are evidences showing the relationships between plant stress responses and changes in miRNAs' expression<sup>5,6</sup>. miRNAs are known as negative post-transcription regulators since they exert specific binding to their target mRNAs or repressing target mRNA translation<sup>7-9</sup>.

Among the major plant abiotic stress sources, drought, salinity, cold, heat, ultraviolet irradiation, carbon dioxide, and heavy metal pollution have significant effects on plant morphological, physiological, and biochemical characteristics<sup>10,11</sup>. To adapt and survive under stress conditions, plants exert miRNA up/down-regulation which results in gene expression reprogramming to restore cellular homeostasis<sup>12,13</sup>. Plant miRNA expression towards stress is generally spatial (plant tissue) and temporal (developmental/growth stage) specific<sup>4,14</sup>.

With the identification of stress-responsive miRNAs, useful information on their role in improving the stress tolerance mechanism of plants can be obtained. A search in bibliographic resources reveals that hundreds of research studies have been dedicated to the changes in plant miRNA expression in response to biotic and abiotic stresses. A large part of these studies has focused on *Arabidopsis thaliana*, *Brachypodium distachyon*, *Glycine max*, *Hordeum vulgare*, *Medicago truncatula*, *Manihot esculenta*, *Phaseolus vulgaris*, *Populus euphratica*, *Populus trichocarpa*, *Populus tremula*, *Triticum turgidum*, *Oryza sativa*, *Vigna unguiculata*, and *Zea mays*<sup>15</sup>. Studies have shown that a miRNA most probably functions in several stresses in one hand. miRNA-167, miRNA-169, miRNA-171, miRNA-319, miRNA-393, miRNA-394, and miRNA-396 are some examples of miRNAs that function in many abiotic stress-related processes<sup>16-20</sup>. On the other hand, a stress can involve changes in the expression of several miRNAs. As an example, nitrogen deficiency can result in an overexpression of miRNA-156, miRNA-160, miRNA-171, miRNA-780, miRNA-826, miRNA-842, and miRNA-846<sup>1</sup>.

Some of plant abiotic stresses such as drought, salinity, cold, and heat are of major constraints to agricultural productivity worldwide<sup>21</sup>. The study of miRNAs involving these stresses and their contribution to plant response is important since it can provide us with valuable information on plant stress physiology. However, only involved miRNAs, their expression in stress conditions, and their target genes are already identified in previous studies and their contribution to plant response towards different levels of plant stress is still a matter of question. Therefore,

<sup>1</sup>Department of Agrotechnology, College of Abouraihan, University of Tehran, Tehran, Iran. <sup>2</sup>Private Laboratory of Biosensor Applications, Hamadan, Iran. email: [keyvan.asefpour@ut.ac.ir](mailto:keyvan.asefpour@ut.ac.ir)



**Figure 1.** The proposed model to link the plant miRNA concentration to the stress using machine learning.

it seems that investigating the contribution (in other words, importance) of each miRNA in plant stress response can be interesting. The preparation of a database based on the observations of miRNA expressions at different levels of plant stress can be the first step. Methods such as northern blot and polymerase chain reaction (PCR) which have been widely used to measure miRNA expressions suffer from weak analytical characteristics, e.g., limit of detection, response linear range, and precision<sup>22</sup>. Therefore, it seems that using biosensors equipped with gold nanoparticles which work on the basis of nanoparticle aggregation is a reliable method to gather required information for such databases<sup>23,24</sup>. Afterwards, using feature selection algorithms to rank the miRNAs will be a possible solution in miRNA contribution investigations.

Feature selection is one of the fundamental problems in the fields of machine learning and pattern recognition<sup>25,26</sup>. Several approaches have been employed in feature selection, such as: embedded<sup>27</sup>, wrapper<sup>28</sup>, and filter<sup>29</sup> methods. These methods use various evaluation criteria for scoring the input features. Among these criteria, information theory-based measurements achieve excellent performance according to their robust algorithm<sup>30</sup>. In contrast with conventional feature selection methods which discard features that are highly correlated to other features but relevant to the target class<sup>31</sup>, information-based feature selecting methods such as cooperative game theory and minimum redundancy - maximum relevance do not ignore features which have strong discriminatory power as a group but are weak as individuals<sup>32</sup>. In cooperative game theory, features that make a big difference as group are usually selected. However, it is possible that they individually perform poorly<sup>33</sup>. In fact, the accuracy of target prediction is assumed as a game in which the features (miRNAs in this study) are the players and a team of players should be selected who can achieve the best results (better prediction of plant stress). In this method, a score is assigned to each feature and the features which are identified with low scores can be eliminated in the measurements.

Machine learning can also be used to predict plant stress by having the plant miRNA expressions. In this situation, machine learns the complex non-linear patterns between inputs (miRNA expressions) and output (plant stress) using the training data in the database and predicts the stress condition of unknown plant samples. Although several learning algorithms including supervised, unsupervised, reinforcement, sparse dictionary, and rule-based learnings have been extensively utilized in previous studies, supervised learning is a reliable and efficient method for life science problems<sup>34,35</sup>. Decision tree (DT), support vector machines (SVMs), least-square support vector machines (LSSVMs), and Naïve Bayes (NB) can be used as supervised learning methods to find the patterns in a database<sup>36,37</sup>. Acceptable performance of the machine (which is indicated by performance evaluation criteria) will show that the expressions of the selected miRNAs which have been used to train the machine are the most important miRNAs in plant response towards stress.

The objectives of this study are: (a) to measure the effects of different levels of abiotic stress including drought, salinity, cold, and heat on the expression of already-known *Arabidopsis thaliana* miRNAs using a gold nanoparticle-based optical biosensor; (b) to investigate the contribution of miRNAs to the plant response towards studied abiotic stresses using information theory-based feature selection algorithms; and (c) to use supervised regression models to predict the plant stress by having the plant miRNA expressions.

## Results and Discussion

Several studies have shown that 11 miRNAs exert tissue-specific expression towards major abiotic stresses in *Arabidopsis thaliana*<sup>1</sup>. This study tries to demonstrate a model in which, machine learning links the plant leaf miRNA expression to the stress. In this situation, a successful learning-based model will be capable of precise plant stress determination by having the concentration of stress-involved miRNAs. Figure 1 depicts the model proposed in this study. Furthermore, features selection algorithms can reveal the contribution of each miRNA to the plant stresses; the information which may require rather complex and expensive laboratory efforts to obtain.

**The effects of different levels of abiotic stress on miRNA concentrations.** The miRNA concentrations towards different levels of abiotic stress is brought in Table 1. This information was necessary to construct the database required for machine learning algorithms. Similar to the results obtained in previous investigations,

Treatment	miRNA concentration										
	miRNA-156	miRNA-159	miRNA-167	miRNA-168	miRNA-169	miRNA-170	miRNA-171	miRNA-319	miRNA-393	miRNA-396	miRNA-398
Control	125 ± 6	45 ± 3	79 ± 7	16 ± 4	84 ± 6	322 ± 15	8 ± 1	578 ± 6	65 ± 6	844 ± 29	13 ± 3
W <sub>1</sub>	462 ± 23 <sup>c</sup>	74 ± 4 <sup>d</sup>	545 ± 28 <sup>d</sup>	25 ± 4 <sup>d</sup>	63 ± 4 <sup>a</sup>	365 ± 21 <sup>d</sup>	13 ± 2 <sup>d</sup>	594 ± 9 <sup>d</sup>	83 ± 5 <sup>d</sup>	1342 ± 18 <sup>d</sup>	14 ± 4 <sup>a</sup>
W <sub>2</sub>	670 ± 17 <sup>b</sup>	115 ± 7 <sup>c</sup>	963 ± 47 <sup>c</sup>	41 ± 4 <sup>c</sup>	61 ± 5 <sup>a</sup>	392 ± 17 <sup>c</sup>	19 ± 2 <sup>c</sup>	612 ± 10 <sup>c</sup>	102 ± 9 <sup>c</sup>	1526 ± 17 <sup>c</sup>	12 ± 2 <sup>a</sup>
W <sub>3</sub>	831 ± 25 <sup>a</sup>	143 ± 11 <sup>b</sup>	1328 ± 42 <sup>b</sup>	64 ± 6 <sup>b</sup>	45 ± 3 <sup>b</sup>	440 ± 23 <sup>b</sup>	25 ± 3 <sup>b</sup>	625 ± 8 <sup>b</sup>	136 ± 8 <sup>b</sup>	1942 ± 25 <sup>b</sup>	12 ± 3 <sup>a</sup>
W <sub>4</sub>	843 ± 32 <sup>a</sup>	234 ± 14 <sup>a</sup>	1743 ± 53 <sup>a</sup>	93 ± 8 <sup>a</sup>	29 ± 4 <sup>c</sup>	487 ± 20 <sup>a</sup>	37 ± 4 <sup>a</sup>	647 ± 8 <sup>a</sup>	175 ± 11 <sup>a</sup>	2416 ± 32 <sup>a</sup>	14 ± 3 <sup>a</sup>
S <sub>1</sub>	174 ± 10 <sup>d</sup>	94 ± 8 <sup>c</sup>	1005 ± 65 <sup>c</sup>	44 ± 4 <sup>c</sup>	124 ± 8 <sup>d</sup>	349 ± 21 <sup>c</sup>	15 ± 2 <sup>b</sup>	603 ± 7 <sup>c</sup>	97 ± 4 <sup>d</sup>	974 ± 15 <sup>d</sup>	15 ± 3 <sup>a</sup>
S <sub>2</sub>	301 ± 21 <sup>c</sup>	263 ± 17 <sup>b</sup>	1655 ± 85 <sup>b</sup>	57 ± 4 <sup>b</sup>	194 ± 9 <sup>c</sup>	363 ± 11 <sup>b</sup>	21 ± 2 <sup>a</sup>	600 ± 10 <sup>c</sup>	132 ± 8 <sup>c</sup>	1081 ± 13 <sup>c</sup>	11 ± 2 <sup>b</sup>
S <sub>3</sub>	372 ± 22 <sup>b</sup>	554 ± 28 <sup>a</sup>	1664 ± 93 <sup>b</sup>	83 ± 7 <sup>a</sup>	247 ± 11 <sup>b</sup>	361 ± 14 <sup>b</sup>	20 ± 3 <sup>a</sup>	671 ± 9 <sup>b</sup>	176 ± 7 <sup>b</sup>	1211 ± 19 <sup>b</sup>	10 ± 3 <sup>b</sup>
S <sub>4</sub>	435 ± 25 <sup>a</sup>	546 ± 22 <sup>a</sup>	2155 ± 53 <sup>a</sup>	81 ± 5 <sup>a</sup>	283 ± 10 <sup>a</sup>	401 ± 14 <sup>a</sup>	21 ± 4 <sup>a</sup>	748 ± 11 <sup>a</sup>	213 ± 9 <sup>a</sup>	1307 ± 21 <sup>a</sup>	7 ± 2 <sup>c</sup>
C <sub>1</sub>	132 ± 8 <sup>a</sup>	45 ± 3 <sup>a</sup>	73 ± 6 <sup>a</sup>	15 ± 4 <sup>d</sup>	98 ± 8 <sup>d</sup>	338 ± 21 <sup>c</sup>	14 ± 2 <sup>d</sup>	734 ± 8 <sup>d</sup>	79 ± 5 <sup>c</sup>	1138 ± 23 <sup>c</sup>	13 ± 4 <sup>a</sup>
C <sub>2</sub>	129 ± 7 <sup>a</sup>	43 ± 2 <sup>a</sup>	78 ± 8 <sup>a</sup>	23 ± 3 <sup>c</sup>	137 ± 8 <sup>c</sup>	335 ± 11 <sup>c</sup>	23 ± 2 <sup>c</sup>	957 ± 11 <sup>c</sup>	93 ± 4 <sup>b</sup>	1472 ± 19 <sup>b</sup>	13 ± 4 <sup>a</sup>
C <sub>3</sub>	122 ± 10 <sup>a</sup>	45 ± 3 <sup>a</sup>	75 ± 9 <sup>a</sup>	35 ± 5 <sup>b</sup>	199 ± 9 <sup>b</sup>	378 ± 14 <sup>b</sup>	35 ± 3 <sup>b</sup>	1143 ± 9 <sup>b</sup>	124 ± 7 <sup>a</sup>	2131 ± 30 <sup>a</sup>	9 ± 2 <sup>b</sup>
C <sub>4</sub>	126 ± 8 <sup>a</sup>	46 ± 3 <sup>a</sup>	71 ± 8 <sup>a</sup>	52 ± 5 <sup>a</sup>	244 ± 8 <sup>a</sup>	422 ± 14 <sup>a</sup>	45 ± 4 <sup>a</sup>	1404 ± 9 <sup>a</sup>	121 ± 9 <sup>a</sup>	2142 ± 27 <sup>a</sup>	3 ± 3 <sup>c</sup>
H <sub>1</sub>	273 ± 22 <sup>d</sup>	112 ± 11 <sup>c</sup>	238 ± 23 <sup>d</sup>	17 ± 2 <sup>a</sup>	57 ± 4 <sup>a</sup>	327 ± 17 <sup>a</sup>	12 ± 1 <sup>d</sup>	585 ± 7 <sup>d</sup>	83 ± 4 <sup>d</sup>	1005 ± 18 <sup>d</sup>	13 ± 2 <sup>a</sup>
H <sub>2</sub>	429 ± 27 <sup>c</sup>	218 ± 21 <sup>b</sup>	524 ± 33 <sup>c</sup>	15 ± 3 <sup>a</sup>	32 ± 5 <sup>b</sup>	319 ± 25 <sup>a</sup>	37 ± 1 <sup>c</sup>	607 ± 9 <sup>c</sup>	91 ± 4 <sup>c</sup>	1321 ± 23 <sup>c</sup>	12 ± 3 <sup>a</sup>
H <sub>3</sub>	526 ± 28 <sup>b</sup>	233 ± 15 <sup>b</sup>	745 ± 30 <sup>b</sup>	16 ± 4 <sup>a</sup>	28 ± 6 <sup>b</sup>	328 ± 18 <sup>a</sup>	55 ± 5 <sup>b</sup>	620 ± 7 <sup>b</sup>	102 ± 10 <sup>b</sup>	1584 ± 22 <sup>b</sup>	14 ± 4 <sup>a</sup>
H <sub>4</sub>	641 ± 31 <sup>a</sup>	342 ± 14 <sup>a</sup>	1254 ± 42 <sup>a</sup>	15 ± 3 <sup>a</sup>	19 ± 5 <sup>c</sup>	325 ± 21 <sup>a</sup>	94 ± 7 <sup>a</sup>	642 ± 8 <sup>a</sup>	117 ± 10 <sup>a</sup>	1859 ± 24 <sup>a</sup>	12 ± 4 <sup>a</sup>

**Table 1.** Plant leaf miRNA concentrations (fM) under stress conditions. Treatments are demonstrated in the text. Data are shown as means ± SD ( $n=5$ ). Different small letter within the same column for each treatment indicate significant differences between the stress levels based on LSD test ( $P < 0.01$ ).

Table 1 shows that the studied miRNAs were significantly affected by the plant major abiotic stresses. The miRNA concentration determination was carried out using a gold nanoparticles-based optical biosensor because the biosensor response towards analyte is generally more sensitive and specific than that obtained by conventional analytical methods, e.g., qRT-PCR, northern blot, and microarrays<sup>22,38,39</sup>. Although an optical biosensor is developed in this study, amperometric<sup>40</sup>, impedimetric<sup>41</sup>, and fluorescence-based<sup>42</sup> biosensors are also introduced for miRNA measurements. However, on one side, electrochemical methods require extensive electrode pre-treatments and costly equipment. On the other side, fluorescence-based methods are not sensitive which decreases their application as a promising technique for miRNA determination<sup>43</sup>. A signal-to-noise ratio of 3:1 was considered to calculate the limit of detection (LoD) of the developed biosensor in this study<sup>44</sup>. It was found that the LoD of the biosensor is equal to 0.5 fM. As another important analytical parameter, the resolution of the biosensor was 1 fM.

As shown in Table 1, some miRNAs have been induced during the stress condition whilst some other miRNAs have been inhibited or unaltered. As expected, the concentration of studied miRNAs has been changed towards at least one stress. The stress levels in this study are selected in a manner to divulge mild, moderate and severe stress conditions in the *Arabidopsis thaliana* plants. According to the table, some miRNAs, i.e., miRNA-171 (with concentrations lower than 100 fM even in induced form) and miRNA-398 (with concentrations lower than 20 fM) were found in very small amounts during the experiments which were induced and inhibited during the stress conditions, respectively. The small concentrations of these two miRNAs should not be interpreted their weak role in plant physiology and biochemistry. The target genes of miRNA-171 in *Arabidopsis thaliana* are *SCL6-II* and *SCL6*, the genes that function in plant root and leaf development, photochrome signalling, lateral organ polarity, meristem formation, vascular development, and stress response<sup>45–47</sup>. miRNA-398 is also an important miRNA which targets *CSD*, *COC5b-1*, and *CCS1* genes playing remarkable roles in Cu homeostasis, heavy metal tolerance, and oxidative stress<sup>48</sup>. There are several stress conditions that the studied miRNAs did not show any significant alteration towards them ( $P < 0.01$ ) (Table 1): miRNA-398 in drought, miRNA-156, miRNA-159, and miRNA-167 in cold, and miRNA-168, miRNA-170, and miRNA-398 in heat conditions.

Although similar results have been reported in previous studies<sup>1</sup>, they have generally considered the effects of severe stress conditions. However, findings of this study revealed that even slight to moderate stress can significantly affect the miRNA concentration. This is interesting since the changes in miRNA concentration detected by an optical biosensor similar to the biosensor developed in this study can be useful for early detection of stress. In this situation, we should know that the result is not specific to a certain stress because as shown in Table 1, a miRNA can be affected by several stress sources.

According to the results, miRNA-167 exerted the highest change in its concentration towards stress among the studied miRNAs. A ca. 27-fold increment in this miRNA was observed during severe salinity stress. The function of miRNA-167 was not very clear in plant stress response until it was recently showed that the transgenic tomato plants overexpressing miRNA-167 exhibited reductions in leaf size and internode length as well as shortened petals, stamens, and styles<sup>49</sup>. In another study, the differential expression of the cassava miRNA-167 target genes (*MesARF6/8*) were observed to be associated with changes in the leaf shape and stomatal closure towards drought stress<sup>50</sup>.

**The contribution of miRNAs to the plant response towards abiotic stresses.** The mean values of miRNA concentrations towards different levels of plant stress were used to construct a database suitable for machine learning. Table 2 shows the contribution of each miRNA to plant stress response which was obtained

miRNA	Stress			
	Drought	Salinity	Cold	Heat
miRNA-156	<b>2</b>	7	9	<b>3</b>
miRNA-159	<i>10</i>	<b>1</b>	<i>11</i>	8
miRNA-167	7	9	<i>10</i>	6
miRNA-168	6	6	5	9
miRNA-169	<b>1</b>	5	<b>2</b>	7
miRNA-170	<b>3</b>	<i>10</i>	7	<i>11</i>
miRNA-171	8	<i>11</i>	<b>4</b>	<b>5</b>
miRNA-319	9	<b>4</b>	6	<b>4</b>
miRNA-393	<b>5</b>	<b>3</b>	8	<b>1</b>
miRNA-396	<b>4</b>	<b>2</b>	<b>1</b>	<b>2</b>
miRNA-398	<i>11</i>	8	3	<i>10</i>

**Table 2.** The importance of miRNAs in plant stress response. The numbers show the importance of the miRNAs. The lower the number, the higher the importance. The five most important miRNAs in each stress condition are shown in bold whilst the five least important miRNAs are in italic.

by implementing the feature selection algorithm on the constructed database. The concept of miRNA's contribution (and importance) to a plant stress has been introduced in this study for the first time. It shows the possibility of proper prediction of a plant stress by having the concentration of a specific miRNA. As shown in Table 2, miRNA-169 concentration had the highest contribution to drought stress. This reveals that among the abiotic stress-involved miRNAs, drought condition has the highest correlation with miRNA-169 concentration in *Arabidopsis thaliana*. Accordingly, miRNA-159, miRNA-396, and miRNA-393 had the highest contributions to the salinity, cold, and heat stresses, respectively. The five most important miRNAs in each stress condition are shown in bold. Studies have shown that even large-scale data in standard databases can be classified with acceptable performance having the five features with the highest scores<sup>32</sup>.

According to Table 2, miRNA-169, miRNA-393, and miRNA-396 have important contributions to at least three studied abiotic stresses. This means that they are good indicators of a wide range of plant stress: from slight to severe stress conditions. These miRNAs have been extensively investigated in plant physiology since there are several evidences of significant changes in their expression towards stress not only in *Arabidopsis thaliana*, but in many plants such as *Phaseolus vulgaris*, *Populus euphratica*, *Populus trichocarpa*, *Populus tremula*, *Triticum turgidum*, *Oryza sativa*, *Vigna unguiculata*, *Medicago truncatula*, and *Zea mays*<sup>51–54</sup>.

miRNA-169 targets *HAP2*, a gene that functions in lots of biotic and abiotic stresses. As the largest miRNA family in *Arabidopsis thaliana*, miRNA-169 has 14 members and can be divided into four groups based on mature miRNA sequences: miRNA-169a, miRNA-169b/c, miRNA-169d/e/f/g and miRNA-169h/i/j/k/l/m/n<sup>55</sup>. Several studies have been carried out to determine whether the different biogenesis in the miRNA-169 family affects the properties of the small RNAs. Combiere and coworkers showed that miRNA-169 involves the symbiotic nodule development in *Medicago truncatula*<sup>56</sup>. In general, the expression of miRNA-169 is significantly down-regulated by nitrogen deficiency<sup>57</sup>. It has been shown that transgenic plants are more sensitive to drought stress compared to wild type plants since they exert more overexpression of miRNA-169<sup>58</sup>. Although plants have different signalling pathways for detecting and responding to dehydration shock versus drought stress<sup>59</sup>, the results of this study along with previous investigations show the important role of miRNA-169 in these pathways. As an example, an abscisic acid-dependent pathway is reported for plants subjected to gradual water stress by withholding irrigation in which, miR169 transcripts are down-regulated during the stress<sup>60</sup>. Furthermore, it has been recently shown that among the plant miRNAs, miRNA-169 is the only miRNA that is inhibited by titanium dioxide nanoparticles with a dose-dependent pattern<sup>61</sup>. During the last two decades, titanium dioxide has become a potentially dangerous contaminant to the environment which undesirably affects the plant growth and development<sup>62</sup>.

One of the conserved miRNA families in plants, the *MIR393* genes, have been found in different plant species<sup>63</sup>. miRNA-393 targets a *TIR1/AFB2* auxin receptor<sup>4</sup> and manipulates the auxin responses<sup>64</sup>, such as controlling the root architecture<sup>65</sup>, regulating leaf development<sup>66</sup>, and maintenance of normal plant growth<sup>67</sup>. It has been found that the overexpression of a cleavage resistant form of *TIR1* leads to an increase in salt tolerance<sup>68</sup>. During the stress, up-regulated miRNA-393 contributes to the repression of auxin signalling by keeping *TIR1* levels low, thereby increasing *AUX/IAA-ARF* heterodimerization<sup>69</sup>. Besides, miRNA-160 and miRNA-167, which are also up-regulated as a result of the stress, down-regulate *ARF* levels and consequently, *ARF* mediated gene expression. Therefore, overall *ARF*-mediated gene expression is suppressed by miRNA-393, miRNA-160 and miRNA-167, leading to the attenuation of plant growth and development under stress, and possibly promoting plant stress tolerance as well.

Finally, miRNA-396, as an important contributor to plant stress, targets four classes of stress resistance protein: pathogen-related, nucleotide binding site resistance protein-like, dirigent-like, and ribonuclease-like proteins<sup>70</sup>. Growth regulating factors targeted by miRNA-396 are cell cycle regulators, which control plant growth and differentiation<sup>71</sup>. miRNA-396 contributes to leaf and flower shape control and axillary meristem maintenance as it balances differentiation and proliferation of cell masses and hence morphogenesis. Stress induction of miRNA-396 represses cell multiplication under unfavourable conditions<sup>20</sup>.

**Supervised prediction of plant stress by having the plant miRNAs' concentration.** The results of the contribution of miRNAs to the plant stress response revealed that stress-involved miRNAs were up-regulated or down-regulated towards all the studied stresses which means that the behaviour of these miRNAs is not specific to the stress. This non-specificity limits the performance of predicting the plant stress by having the concentration of one single miRNA even with the most sophisticated computing tools. Some of the research studies have shown that there are several miRNAs that function as specific indicators of plant biotic and abiotic stresses in some plants<sup>72</sup>. However, in the present study, the concentration of the miRNAs did not exert a selective and specific change towards the stresses.

To untie this knot and to introduce a reliable method for the prediction of plant stress by having the plant miRNAs' concentration, the most important contributing miRNAs in plant stresses which were obtained by the feature selection algorithm in the previous subsection were considered to train the machine learning algorithms. Therefore, the concentration of miRNA-169, miRNA-393, and miRNA-396 in different stress levels were used to train several machines including DT, SVM, LSSVM, and NB. The detail, types, and parameters of these regression-based machine learning algorithms are brought in Methods. The reason why I emphasis on the term "regression" is that in contrast with the feature selection method used in this study, the outputs of machine learning algorithms are continuous (and not discrete). As an example, for a sample in which the concentration of miRNA-169, miRNA-393, and miRNA-396 are 283, 213, and 1307 fM, respectively, a well-trained machine should have an output equal to 80 for the salinity stress which shows that the sample has irrigated with water containing 80 mM of NaCl (see Table 1). Predicting values lower or higher than 80 shows the undesirability of the machine and  $R^2$  of the machine decreases. It should be noted that the machines predict the type of the stress (i.e., salinity, drought, heat, and cold) along with the stress level. In this situation, the researcher does not need to know the type of stress before predicting the stress level by having the miRNA concentrations using the machines since both the type and level of the stress are predicted by the machines simultaneously. Results showed that all the machines were capable of accurate prediction of the stress type. However, some machines were more accurate in predicting the stress level while some of them were less accurate.

The results of model performance evaluation in the prediction of plant stress is shown in Table 3. According to the table, the performance of SVM was better than the performance of other machine learning methods in prediction of plant stress. SVM was able to predict the output with  $R^2 = 0.96$  which means that if we measure the concentrations of miRNA-169, miRNA-393, and miRNA-396 in *Arabidopsis thaliana* plant leaf, there is a good chance that we will be able to predict the plant stress using SVM. It should be noted that in this study and possible similar investigations in the future including the relationships between miRNA concentrations and plant stress, it is not possible to use artificial neural networks (ANNs) as a learning algorithm since ANNs require a lot of training data for the optimization of sigmoid functions belonging to the hidden layer's neurons<sup>73</sup>. Therefore, in this study, where the number of training samples was small, the optimization process cannot be properly carried out even by using back-propagation algorithms. Furthermore, this small number of training data may result in over-fitting and local minima in ANNs. These phenomena can cause an unrealistic increase in the  $R^2$  of the model.

A comparison among machine learning results shows that CART and CHAID had better performance compared to other DT algorithms (Table 3). Results of SVM and LSSVM methods showed that Gaussian kernel was more accurate than other kernels. The suitable kernel used in SVM method depends on the type of samples' scattering in the feature space. Former studies have shown that Gaussian kernel is useful in modelling many biological and biotechnological phenomena<sup>74</sup>. NB is another learning algorithm which is based on logistic regression. Logistic regression is mainly used in cases where the output can be expressed as Boolean values. Table 3 shows that this type of model did not have acceptable results.

## Conclusions

This study is the first report of using machine learning to investigate the contribution of miRNAs in plant stress response. Although the response of *Arabidopsis thaliana* miRNAs towards abiotic stresses such as drought, salinity, cold, and heat is not specific, machine learning can be a useful technique to predict plant stress by having the concentration of stress-involved miRNAs. To do this, laboratory data of miRNA concentrations in different levels of plant stress were extracted using an optical nanoparticles-based biosensor to construct a database required by the machine learning algorithms. Then, feature selection algorithm showed that miRNA-169, miRNA-159, miRNA-396, and miRNA-393 have the highest contributions to plant response towards drought, salinity, cold, and heat, respectively. Furthermore, miRNA-169, miRNA-393, and miRNA-396 were considered as the input variables of machine learning algorithms to predict plant stress since they had the highest contributions in all the studied stresses. The results of this study confirm the hypothesis describing machine learning as an efficient technique to improve our knowledge about the relationships between plant stress and miRNAs' expression.

## Methods

**Plant materials and growth condition.** Double-knockout mutant (*acl5/spms*) seeds of *Arabidopsis thaliana* ecotype Columbia (Col-0) were surface-sterilized by treating with 70% ethanol for 5 min, then in a solution of 1% sodium hypochloride and 0.1% Tween 20 for 15 min, followed by extensive washing with sterile distilled water. Seeds were then sown in moistened peat pellets, stratified at 4 °C for 2 d, and then transferred to a growth room. Macro and micro nutrient fertilization management of the plants was according to Cipollini<sup>75</sup>. The relative humidity and temperature of the growth room was adjusted at  $70 \pm 5\%$  and 22 °C under a light intensity of  $100 \mu\text{mol m}^{-2} \text{s}^{-1}$  with a photocycle of 16/8 h (light/dark). Complete and equally irrigation of all plants was conducted with 100% of field capacity.

Model	Parameters	R <sup>2</sup>
DT	type = ID3	0.85
	type = C4.5	0.79
	type = CART	0.89
	type = CHAID	0.88
	type = MARS	0.64
SVM	kernel type = linear, $\gamma = 0.01$	0.76
	kernel type = linear, $\gamma = 0.05$	0.77
	kernel type = linear, $\gamma = 0.10$	0.72
	kernel type = polynomial, $\gamma = 0.01$	0.80
	kernel type = polynomial, $\gamma = 0.05$	0.83
	kernel type = polynomial, $\gamma = 0.10$	0.82
	kernel type = Gaussian, $\gamma = 0.01$	0.87
	kernel type = Gaussian, $\gamma = 0.05$	0.96
	kernel type = Gaussian, $\gamma = 0.10$	0.92
	kernel type = sigmoid, $\gamma = 0.01$	0.86
	kernel type = sigmoid, $\gamma = 0.05$	0.88
	kernel type = sigmoid, $\gamma = 0.10$	0.80
LSSVM	kernel type = linear, $\gamma = 0.01$	0.72
	kernel type = linear, $\gamma = 0.05$	0.81
	kernel type = linear, $\gamma = 0.10$	0.81
	kernel type = polynomial, $\gamma = 0.01$	0.79
	kernel type = polynomial, $\gamma = 0.05$	0.80
	kernel type = polynomial, $\gamma = 0.10$	0.82
	kernel type = Gaussian, $\gamma = 0.01$	0.87
	kernel type = Gaussian, $\gamma = 0.05$	0.90
	kernel type = Gaussian, $\gamma = 0.10$	0.83
	kernel type = sigmoid, $\gamma = 0.01$	0.74
	kernel type = sigmoid, $\gamma = 0.05$	0.79
	kernel type = sigmoid, $\gamma = 0.10$	0.71
NB	—	0.63

**Table 3.** Performance evaluation of regression models in the prediction of plant stress.

**Stress treatments.** Twenty-five-day-old (sixth true leaf stage) seedlings were used for stress investigations. Control seedlings were kept in the condition described above. Each stress was conducted individually at four levels with five replications for 15 days.

Drought treated plants were stressed by withholding water until soil water potential became different with field capacity<sup>76</sup>. Soil moisture was measured daily by a time-domain reflectometry (TDR) device (PMS-714, LUTRON, Taiwan). The soil moisture level was maintained at a level that was nonlethal and above the wilting point, at 85% ( $W_1$ ), 70% ( $W_2$ ), 55% ( $W_3$ ), and 40% ( $W_4$ ) of field capacity to study different severities of drought stress. Salinity treated plants were irrigated by water which contained different concentrations of NaCl, i.e., 20 mM ( $S_1$ ), 40 mM ( $S_2$ ), 60 mM ( $S_3$ ), and 80 mM ( $S_4$ ). Higher concentrations of NaCl, more than 100 mM, may result in lethal damage to the young plants<sup>77</sup>. For cold treatment, the temperature of growth room was adjusted at 16 °C ( $C_1$ ), 12 °C ( $C_2$ ), 8 °C ( $C_3$ ), and 4 °C ( $C_4$ ) whilst seedlings of the same growth stage were kept at 28 °C ( $H_1$ ), 32 °C ( $H_2$ ), 36 °C ( $H_3$ ), and 40 °C ( $H_4$ ) for heat treatment. The threshold values to select the stress temperatures was according to Kaplan and coworkers<sup>78</sup>.

**miRNA concentration determination.** Total RNA was isolated from 50 mg of the uppermost leaf of forty-day-old plants using Total RNA Purification Kit (Norgen Biotek, ON, Canada) according to the manufacturer's instructions based on Yamaguchi and coworkers<sup>79</sup>. The concentration of plant stress-involved miRNAs in the isolated samples was measured using a gold nanoparticles (AuNPs)-based biosensor. The list of the miRNAs and their sequence is brought in Table S1 in Supporting information. Biosensor preparation was carried out in three steps according to Hakimian and coworkers<sup>39</sup> and Asefpour Vakilian<sup>80</sup> which is briefly described below:

Step 1: 100  $\mu$ L of polyethylenimine (PEI) (42 mM) was added to 3 mL of HAuCl<sub>4</sub> (1.5 M) under vigorous stirring at constant pH of 7.4 which was adjusted by adding HCl to the solution. Afterwards, the temperature of solution was increased so its colour changed from yellow to red as an indication of the reduction process<sup>39</sup>. PEI-AuNPs were then dialyzed against deionized water using a 3.5 kDa molecular weight cut-off membrane. The resulting red solution was stored at 4 °C before use. Five microliters of the sample were incubated with 40  $\mu$ L of synthesized PEI-AuNP for 30 min at room temperature.

Step 2: 1.5 mL of sodium citrate 1% was added to 21 mL of boiling HAuCl<sub>4</sub> solution (0.8 mM), whilst vigorously stirring until its colour changed from pale yellow to deep red. The solution, was then stirred for an additional 15 min and gradually cooled down to room temperature. After that, 400  $\mu$ L of the solution was mixed with

2  $\mu$ L of Tween-20 and 400  $\mu$ L of each thiolated probe (1  $\mu$ M). Probes and their sequence are brought in Table S2 in Supporting information. The product was left for 48 h in room temperature and then, centrifuged for 23 min at 10,000 rpm. Finally, the supernatant was removed and the oily red precipitate re-dispersed in 200  $\mu$ L of deionized water.

Step 3: by mixing 5  $\mu$ L of the products from the steps 1 and 2, probe-target hybridization resulted in the reduction of distance between nanoparticles and an interparticle cross-linking aggregation happened. The higher the target concentration, the greater the aggregation is. As an indicator of reaction, the colour of mixture changed from red-pink to pink and the absorption intensity decreases at 530 nm<sup>39</sup>. Since the UV-Vis absorptions at 530 and 750 nm indicate the quantity of dispersed and aggregated AuNPs, respectively<sup>81</sup>, the absorbance ratio of 750/530 nm was considered as an indicator of probe-target hybridization, and consequently, the concentration of target miRNA. In this study, UV-Vis absorption spectra of the aggregated particles were recorded after 15 min of reaction using a multi-mode microplate reader (SpectraMax iD5, Molecular Devices, USA).

**Statistical analysis.** The data obtained from each stress source investigation were individually subjected to analysis of variance (ANOVA) using LSD test at the significance level of 0.01 in SAS 9.0 programming environment.

**Database preparation.** The measured concentrations of the studied miRNAs at different stress levels were used to construct a database suitable for machine learning purposes. Due to the fact that statistical analysis showed that replication does not have significant effects on the results, mean values were used for database preparation. Since the effects of four levels of four stress conditions were studied on the concentration of 11 stress-involved miRNAs, a total of  $4 \times 4$  patterns along with one control pattern were used to construct the database. Each pattern consisted of the plant stress level and the corresponding miRNA concentrations.

**Feature selection algorithms.** Since one of the research objectives was to investigate the contribution of miRNAs to the plant response towards studied abiotic stresses, miRNA concentrations were considered as the inputs of information theory-based feature selection algorithms whilst stress levels as discrete classes were considered as outputs. Cooperative game theory was used to score the miRNAs based on their contributions to the plant stress response. Intrinsic correlative structures among variables results in different importance of every individual. Cooperative game theory focuses on evaluating the importance (in other words, power) of each feature (input) using the Banzhaf power index<sup>32,82,83</sup>.

In brief, the Banzhaf index can be described as follows<sup>32</sup>: A winning coalition is one for which  $v(S) = 1$  and a losing coalition is one for which  $v(S) = 0$ . Each coalition  $SU\{i\}$  that wins when  $S$  loses is called a swing for player  $i$ , because the membership of player  $i$  in the coalition is crucial to the coalition winning. Let  $\sigma_i(N,v)$  be the number of swings for  $i$ , and let  $\sigma_o(N,v)$  be the total number of swings of all players in the game. Then, the normalized Banzhaf index is  $b_i(N,v) = \sigma_i(N,v) / \sigma_o(N,v)$ . The idea is that every subset of features can be regarded as a candidate subset for the final selected optimal subset<sup>32</sup>. Thus, the power of each feature can be measured by averaging the contributions that it makes to each of the subset which it belongs to. The algorithm of the feature selection method was implemented using a code written in MATLAB R2016b programming environment (Mathworks, MA, USA).

**Learning-based regression models.** Although statistical regression models provide reliable mathematical equations to calculate the dependent variable by having the input features, the number of input features should not generally exceed 2 since finding model parameters is rather difficult in high-dimension problems<sup>74,84</sup>. In contrast, machine learning methods can learn a database including hundreds of input features and corresponding dependent variables<sup>85</sup>. In this study, learning-based regression models including DT, SVM, LSSVM, and NB were used to predict plant stress by having the miRNA concentrations as inputs. Since the number of patterns in this study is limited, node-based algorithms, e.g., artificial neural networks (ANN) seem to be inefficient in modelling the database and therefore, they were not used.

Different types of DT modelling including iterative dichotomiser 3 (ID3), statistical model (C4.5), classification and regression tree (CART), chi-squared automatic interaction detection (CHAID), and multivariate adaptive regression splines (MARS) were used in the modelling. SVM and LSSVM models have two parameters including kernel type and kernel parameter which affect the performance of the model<sup>86</sup>. Three kernel types including linear ( $f = \gamma xx_o$ ), polynomial ( $f = (\gamma xx_o)^3$ ), Gaussian ( $f = \exp(-\gamma (x - x_o)^2)$ ) and sigmoid ( $f = \tanh(\gamma xx_o)$ ) were considered for modelling where  $f$  is the kernel function,  $\gamma$  is the kernel parameter,  $x$  is a train or test sample in the modelling hyperplane, and  $x_o$  is the origin point in the hyperplane<sup>87</sup>. The machine learning methods were implemented using a code written in MATLAB R2016b programming environment (Mathworks, MA, USA).

To investigate the performance of the machine learning methods, 3-fold cross-validation was used for training and testing. The performance of the models was evaluated based on the coefficient of determination ( $R^2$ ). The higher the  $R^2$ , the better performance of the machine learning model is.

Received: 10 December 2019; Accepted: 6 February 2020;

Published online: 20 February 2020

## References

- Wang, J., Meng, X., Dobrovolskaya, O. B., Orlov, Y. L. & Chen, M. Non-coding RNAs and their roles in stress response in plants. *Genom. Proteom. Bioinform.* **15**(5), 301–312 (2017).
- Li, Y. F. *et al.* Transcriptome-wide identification of microRNA targets in rice. *Plant J.* **62**(5), 742–759 (2010).
- Sunkar, R. MicroRNAs with macro-effects on plant stress responses. In *Seminars in cell & developmental biology*, **21**, 8, 805–811 Academic Press (2010).

4. Sunkar, R., Li, Y. F. & Jagadeeswaran, G. Functions of microRNAs in plant stress responses. *Trends Plant Sci.* **17**(4), 196–203 (2012).
5. Kumar, V., Khare, T., Shriram, V. & Wani, S. H. Plant small RNAs: the essential epigenetic regulators of gene expression for salt-stress responses and tolerance. *Plant Cell Rep.* **37**(1), 61–75 (2018).
6. Hou, J. *et al.* Non-coding RNAs and transposable elements in plant genomes: emergence, regulatory mechanisms and roles in plant development and stress responses. *Planta* **250**(1), 23–40 (2019).
7. Wu, L. *et al.* Rice microRNA effector complexes and targets. *Plant Cell* **21**(11), 3421–3435 (2009).
8. Zeng, H. *et al.* Role of microRNAs in plant responses to nutrient stress. *Plant Soil* **374**(1–2), 1005–1021 (2014).
9. Shriram, V., Kumar, V., Devarumath, R. M., Khare, T. S. & Wani, S. H. MicroRNAs as potential targets for abiotic stress tolerance in plants. *Front. Plant Sci.* **7**, 817 (2016).
10. Pessarakli, M. *Handbook of plant and crop stress*. 3rd ed. (CRC Press 2010).
11. Sewelam, N., Kazan, K. & Schenk, P. M. Global plant stress signaling: reactive oxygen species at the cross-road. *Front. Plant Sci.* **7**, 187 (2016).
12. Pandey, P., Wang, M., Baldwin, I. T., Pandey, S. P. & Groten, K. Complex regulation of microRNAs in roots of competitively-grown isogenic *Nicotiana attenuata* plants with different capacities to interact with arbuscular mycorrhizal fungi. *BMC Genomics* **19**(1), 937 (2018).
13. Ahmed, W. *et al.* Non-coding RNAs: Functional roles in the regulation of stress response in Brassica crops. *Genomics* <https://doi.org/10.1016/j.ygeno.2019.08.011> (2019).
14. Islam, W., Noman, A., Qasim, M. & Wang, L. Plant responses to pathogen attack: small RNAs in focus. *Int. J. Mol. Sci.* **19**(2), 515 (2018).
15. Noman, A. & Aqeel, M. miRNA-based heavy metal homeostasis and plant growth. *Environ. Sci. Pollut. R.* **24**(11), 10068–10082 (2017).
16. Wang, B. *et al.* MicroRNAs involving in cold, wounding and salt stresses in *Triticum aestivum* L. *Plant Physiol. Bioch.* **80**, 90–96 (2014).
17. Xie, F., Wang, Q., Sun, R. & Zhang, B. Deep sequencing reveals important roles of microRNAs in response to drought and salinity stress in cotton. *J. Exp. Bot.* **66**(3), 789–804 (2014).
18. Gao, S. *et al.* A cotton miRNA is involved in regulation of plant response to salt stress. *Sci. Rep.* **6**, 19736 (2016).
19. Ghani, A., Din, M. & Barozai, M. Y. K. Convergence and divergence studies of plant precursor microRNAs. *Pakistan J. Bot.* **50**(3), 1085–1091 (2018).
20. Patel, P., Yadav, K., Ganapathi, T. R. & Penna, S. Plant miRNAome: Cross Talk in Abiotic Stressful Times. In *Genetic Enhancement of Crops for Tolerance to Abiotic Stress: Mechanisms and Approaches*, I, pp. 25–52 (Springer, 2019).
21. Mahajan, S. & Tuteja, N. Cold, salinity and drought stresses: an overview. *Arch. Biochem. Biophys.* **444**(2), 139–158 (2005).
22. Petralia, S. *et al.* An innovative chemical strategy for PCR-free genetic detection of pathogens by an integrated electrochemical biosensor. *Analyst* **142**(12), 2090–2093 (2017).
23. Thanh, N. T. K. & Rosenzweig, Z. Development of an aggregation-based immunoassay for anti-protein A using gold nanoparticles. *Anal. Chem.* **74**(7), 1624–1628 (2002).
24. Oh, J. H. & Lee, J. S. Designed hybridization properties of DNA–gold nanoparticle conjugates for the ultrasensitive detection of a single-base mutation in the breast cancer gene BRCA1. *Anal. Chem.* **83**(19), 7364–7370 (2011).
25. Bolón-Canedo, V., Sánchez-Marroño, N. & Alonso-Betanzos, A. Recent advances and emerging challenges of feature selection in the context of big data. *Knowl.-Based Syst.* **86**, 33–45 (2015).
26. Cai, J., Luo, J., Wang, S. & Yang, S. Feature selection in machine learning: A new perspective. *Neurocomputing* **300**, 70–79 (2018).
27. Maldonado, S. & López, J. Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification. *Appl. Soft Comput.* **67**, 94–105 (2018).
28. González, J., Ortega, J., Damas, M., Martín-Smith, P. & Gan, J. Q. A new multi-objective wrapper method for feature selection—Accuracy and stability analysis for BCI. *Neurocomputing* **333**, 407–418 (2019).
29. Hancer, E., Xue, B. & Zhang, M. Differential evolution for filter feature selection based on information theory and feature ranking. *Knowl.-Based Syst.* **140**, 103–119 (2018).
30. Bannasar, M., Hicks, Y. & Setchi, R. Feature selection using joint mutual information maximisation. *Expert Syst. Appl.* **42**(22), 8520–8532 (2015).
31. Vergara, J. R. & Estévez, P. A. A review of feature selection methods based on mutual information. *Neural Comput. Appl.* **24**(1), 175–186 (2014).
32. Sun, X. *et al.* Feature evaluation and selection with cooperative game theory. *Pattern Recogn.* **45**(8), 2992–3002 (2012).
33. Wang, Z., Wu, D., Chen, J., Ghoneim, A. & Hossain, M. A. A triaxial accelerometer-based human activity recognition via EEMD-based features and game-theory-based feature selection. *IEEE Sens. J.* **16**(9), 3198–3207 (2016).
34. Massah, J. & Asefpour Vakilian, K. An intelligent portable biosensor for fast and accurate nitrate determination using cyclic voltammetry. *Biosyst. Eng.* **177**, 49–58 (2019).
35. Massah, J., Asefpour Vakilian, K. & Torktaz, S. Supervised Machine Learning Algorithms Can Predict Penetration Resistance in Mineral-fertilized Soils. *Commun. Soil Sci. Plant* **50**(17), 2169–2177 (2019).
36. Khan, A., Baharudin, B., Lee, L. H. & Khan, K. A review of machine learning algorithms for text-documents classification. *J. Adv. Inform. Tech.* **1**(1), 4–20 (2010).
37. Hashemi, A., Asefpour Vakilian, K., Khazaei, J. & Massah, J. An artificial neural network modeling for force control system of a robotic pruning machine. *J. Inform. Organ. Sci.* **38**(1), 35–41 (2014).
38. Konishi, H. *et al.* Detection of gastric cancer-associated microRNAs on microRNA microarray comparing pre-and post-operative plasma. *Brit. J. Cancer* **106**(4), 740 (2012).
39. Hakimian, F., Ghourchian, H., Sadat Hashemi, A., Arastoo, M. R. & Rad, M. B. Ultrasensitive optical biosensor for detection of miRNA-155 using positively charged Au nanoparticles. *Sci. Rep.* **8**(1), 2943 (2018).
40. Cheng, F. F. *et al.* Bimetallic Pd–Pt supported graphene promoted enzymatic redox cycling for ultrasensitive electrochemical quantification of microRNA from cell lysates. *Analyst* **139**(16), 3860–3865 (2014).
41. Congur, G., Eksin, E. & Erdem, A. Impedimetric detection of microRNA at graphene oxide modified sensors. *Electrochim. Acta* **172**, 20–27 (2015).
42. Almlie, C. K., Larkey, N. E. & Burrows, S. M. Fluorescent microRNA biosensors: a comparison of signal generation to quenching. *Anal. Methods* **7**(17), 7296–7310 (2015).
43. Kilic, T., Erdem, A., Ozsoz, M. & Carrara, S. microRNA biosensors: opportunities and challenges among conventional and commercially available techniques. *Biosens. Bioelectron.* **99**, 525–546 (2018).
44. Shrivastava, A. & Gupta, V. B. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chron. Young Sci.* **2**(1), 21 (2011).
45. Lee, M. H. *et al.* Large-scale analysis of the GRAS gene family in Arabidopsis thaliana. *Plant Mol. Biol.* **67**(6), 659–670 (2008).
46. Wang, L., Mai, Y. X., Zhang, Y. C., Luo, Q. & Yang, H. Q. MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in Arabidopsis. *Mol. Plant* **3**(5), 794–806 (2010).
47. Zhu, X. *et al.* Discovery of conservation and diversification of miR171 genes by phylogenetic analysis based on global genomes. *Plant Genome* **8**(2), 1–11 (2015).

48. Noman, A. *et al.* Crosstalk Between Plant miRNA and Heavy Metal Toxicity. In *Plant Metallomics and Functional Omics*, pp. 145–168 (Springer, 2019).
49. Liu, N. *et al.* Down-regulation of AUXIN RESPONSE FACTORS 6 and 8 by microRNA 167 leads to floral development defects and female sterility in tomato. *J. Exp. Bot.* **65**(9), 2507–2520 (2014).
50. Phookaew, P., Netrphan, S., Sojikul, P. & Narangajavana, J. Involvement of miR164-and miR167-mediated target gene expressions in responses to water deficit in cassava. *Biol. Plantarum* **58**(3), 469–478 (2014).
51. Lu, S., Sun, Y. H. & Chiang, V. L. Stress-responsive microRNAs in Populus. *Plant J.* **55**(1), 131–151 (2008).
52. Wang, T., Chen, L., Zhao, M., Tian, Q. & Zhang, W. H. Identification of drought-responsive microRNAs in *Medicago truncatula* by genome-wide high-throughput sequencing. *BMC Genomics* **12**(1), 367 (2011).
53. Ding, Y., Tao, Y. & Zhu, C. Emerging roles of microRNAs in the mediation of drought stress response in plants. *J. Exp. Bot.* **64**(11), 3077–3086 (2013).
54. Budak, H., Kantar, M., Bulut, R. & Akpinar, B. A. Stress responsive miRNAs and isomiRs in cereals. *Plant Sci.* **235**, 1–13 (2015).
55. Du, Q., Zhao, M., Gao, W., Sun, S. & Li, W. X. micro RNA/micro RNA\* complementarity is important for the regulation pattern of NFYA 5 by miR169 under dehydration shock in Arabidopsis. *Plant J.* **91**(1), 22–33 (2017).
56. Combier, J. P. *et al.* MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Gene Dev.* **20**(22), 3084–3088 (2006).
57. Zhao, M., Ding, H., Zhu, J. K., Zhang, F. & Li, W. X. Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. *New Phytol.* **190**(4), 906–915 (2011).
58. Li, W. X. *et al.* The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell* **20**(8), 2238–2251 (2008).
59. Talamè, V., Ozturk, N. Z., Bohnert, H. J. & Tuberosa, R. Barley transcript profiles under dehydration shock and drought stress treatments: a comparative analysis. *J. Exp. Bot.* **58**(2), 229–240 (2007).
60. Ni, Z., Hu, Z., Jiang, Q. & Zhang, H. GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. *Plant Mol. Biol.* **82**, 113–129 (2013).
61. Boykov, I. N., Shuford, E. & Zhang, B. Nanoparticle titanium dioxide affects the growth and microRNA expression of switchgrass (*Panicum virgatum*). *Genomics* **111**(3), 450–456 (2019).
62. Frazier, T. P., Burklew, C. E. & Zhang, B. Titanium dioxide nanoparticles affect the growth and microRNA expression of tobacco (*Nicotiana tabacum*). *Funct. Integr. genomic.* **14**(1), 75–83 (2014).
63. Navarro, L. *et al.* A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* **312**(5772), 436–439 (2006).
64. Windels, D. *et al.* miR393 is required for production of proper auxin signalling outputs. *PLoS One* **9**(4), e95972 (2014).
65. Vidal, E. A. *et al.* Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in *Arabidopsis thaliana*. *P. Natl. Acad. Sci. USA* **107**(9), 4477–4482 (2010).
66. Si-Ammour, A. *et al.* miR393 and secondary siRNAs regulate expression of the TIR1/AFB2 auxin receptor clade and auxin-related development of Arabidopsis leaves. *Plant Physiol.* **157**(2), 683–691 (2011).
67. Chen, Z. H. *et al.* Regulation of auxin response by miR393-targeted transport inhibitor response protein1 is involved in normal development in Arabidopsis. *Plant Mol. Biol.* **77**(6), 619–629 (2011).
68. Chen, Z. *et al.* Overexpression of a miR393-resistant form of transport inhibitor response protein 1 (*mTIR1*) enhances salt tolerance by increased osmoregulation and Na<sup>+</sup> exclusion in *Arabidopsis thaliana*. *Plant Cell Physiol.* **56**(1), 73–83 (2015).
69. Wang, R. & Estelle, M. Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. *Curr. Opin. Plant Biol.* **21**, 51–58 (2014).
70. Zhang, B. *et al.* Identification of cotton microRNAs and their targets. *Gene* **397**(1–2), 26–37 (2007).
71. Liu, D., Song, Y., Chen, Z. & Yu, D. Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. *Physiol. Plantarum* **136**(2), 223–236 (2009).
72. Din, M. & Barozai, M. Y. K. Profiling microRNAs and their targets in an important fleshy fruit: tomato (*Solanum lycopersicum*). *Gene* **535**(2), 198–203 (2014).
73. Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. *Neural network design*, Vol. 20 (PWS publishing company, 1996).
74. Asefpour Vakilian, K. & Massah, J. A portable nitrate biosensing device using electrochemistry and spectroscopy. *IEEE Sens. J.* **18**(8), 3080–3089 (2018).
75. Cipollini, D. Constitutive expression of methyl jasmonate-inducible responses delays reproduction and constrains fitness responses to nutrients in *Arabidopsis thaliana*. *Evol. Ecol.* **24**(1), 59–68 (2010).
76. Harb, A., Krishnan, A., Ambavaram, M. M. & Pereira, A. Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiol.* **154**(3), 1254–1271 (2010).
77. Sun, J. *et al.* The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in Arabidopsis. *Plant Cell Physiol.* **48**(8), 1148–1158 (2007).
78. Kaplan, F. *et al.* Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiol.* **136**(4), 4159–4168 (2004).
79. Yamaguchi, K. *et al.* A protective role for the polyamine spermine against drought stress in Arabidopsis. *Biochem. Biophys. Res. Co.* **352**(2), 486–490 (2007).
80. Asefpour Vakilian, K. Gold nanoparticles-based biosensor can detect drought stress in tomato by ultrasensitive and specific determination of miRNAs. *Plant Physiol. Bioch.* **145**, 195–204 (2019).
81. Chen, S. J. *et al.* Colorimetric determination of urinary adenosine using aptamer-modified gold nanoparticles. *Biosens. Bioelectron.* **23**(11), 1749–1753 (2008).
82. Bachrach, Y. *et al.* Approximating power indices: theoretical and empirical analysis. *Auton. Agent. Multi-Ag.* **20**(2), 105–122 (2010).
83. Sun, J., Zhong, G., Huang, K. & Dong, J. Banzhaf random forests: Cooperative game theory based random forests with consistency. *Neural Networks* **106**, 20–29 (2018).
84. Obermeyer, Z. & Emanuel, E. J. Predicting the future—big data, machine learning, and clinical medicine. *New Engl. J. Med.* **375**(13), 1216 (2016).
85. Parmar, C., Grossmann, P., Bussink, J., Lambin, P. & Aerts, H. J. Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **5**, 13087 (2015).
86. Durgesh, K. S. & Lekha, B. Data classification using support vector machine. *J. Theor. Appl. Inform. Tech.* **12**(1), 1–7 (2010).
87. Basak, D., Pal, S. & Patranabis, D. C. Support vector regression. *Neu. Inf. Pro.* **11**(10), 203–224 (2007).

## Author contributions

The experiments and paper preparation were carried out by the sole author.

## Competing interests

The author declares no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-59981-6>.

**Correspondence** and requests for materials should be addressed to K.A.V.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020