Research Article

Molecular characterization of two AP2/ERF transcription factor genes from Egyptian tomato cultivar (Edkawy)

Eslam A. Heb El-din1, Fawzy A. El-Feky1, Ali M. El-Refy1, Ahmed I. Ismail1 and Kareem A. Mosa1, 2*

1Department of Biotechnology, Faculty of Agriculture, Al-Azhar University, Cairo, Egypt
2Department of Applied Biology, Faculty of Science, Sharjah University, P.O. Box 27272, Sharjah, United Arab Emirates

Article history
Received: 18 September 2016
Accepted: 08 November 2016
Published: 01 January 2017

Abstract

The tomato is ranked first amongst vegetable crops in Egypt in relation to surface area and production. The Egyptian tomato cultivar Edkawy has shown abiotic stress tolerance characteristics. However, there is not much information about the molecular characterization of this cultivar. Furthermore, information regarding the identification of abiotic stress tolerance genes from the Edkawy tomato cultivar is lacking. Here, we investigated the ability of the Edkawy cultivar to tolerate drought stress. Two varieties were used as a control in this study; Peto86 (sensitive variety) and Strain B (tolerant variety). Edkawy, Peto86 and Strain B varieties were exposed to drought stress by reducing the water supply gradually. Interestingly, Edkawy demonstrated a remarkable tolerance phenotype to drought stress. Furthermore, we identified and isolated two members of the AP2/ERF transcription factor family from Edkawy which are associated with abiotic stress, particularly drought, i.e. ERF1 and ERF5. Protein prediction, validation and active site prediction of ERF1 and ERF5 were also determined. In addition to the domain obtained by the pfam online tool, the interaction between Edkawy ERFs proteins and other proteins in the Solanaceae family was obtained. Furthermore, subcellular localization was determined by the ngLOC and Plant-mPlLoc online tools. Characterization of the Edkawy tomato cultivar and isolation and identification of such transcription factors will help in the engineering of tomato plants with abiotic stress tolerance.

Keywords
Edkawy; molecular; tomato; transcription factor genes

Introduction

The tomato (Solanum lycopersicum L., formerly Lycopersicon esculentum L.) is one of the most economically important and widely grown plants in the Solanaceae family and is one of the most important vegetable crops grown widely all over the world. It is a self-pollinated crop with 2n = 24 chromosomes (Peterson et al. 1996). The Peru Equador region is considered to be the center of its origin (Rick, 1969). In Egypt, the tomato is ranked first amongst vegetable crops in terms of surface area and production (FAOSTAT, 2012). It is very important to understand the molecular mechanisms of tomato responses to abiotic stresses, in order to improve stress tolerance and productivity. Abiotic stresses such as drought,
salinity, cold, heat and mechanical wounding regulate many genes, and this often occurs at the transcriptional level in which several genes are activated in response to different abiotic stresses (Kavar et al., 2007). In the promoter regions, the transcription factors interact with cis-elements of several stress-related genes which lead to the upregulation of many downstream genes causing abiotic stress tolerance (Agarwal and Jha., 2010).

ERFs (ethylene responsive factors) also belong to the AP2-EREBP transcription factors family and have been found to be involved in growth, development, metabolic regulation and biotic and abiotic stress responses (Hussain et al., 2011). AP2/EREBP (ERF) family members have been reported to function through both ABA-dependent and –independent pathways (Yamaguchi-shinozaki and Shinozaki, 1994; Kizis and Investigacio, 2002). Ethylene response factor (ERF) proteins have been shown to interact with DRE/CFB genes in enhancing plant stress responses, however, the regulatory mechanism is not well known (Zhang and Huang, 2010).

The Edkawy tomato cultivar has been selected conventionally by Egyptian farmers in the Edko region, El-Beherah Governorate, Egypt (El-Awady et al., 2014). It is considered as one of the most important Egyptian tomato cultivars which have been developed using conventional methods. Studies have shown the ability of this cultivar against salt stress (Sarg et al., 1993; Saker and Rady, 1999). However, to the best of our knowledge, there is no available information regarding the identification of abiotic stress tolerance genes in the Edkawy tomato cultivar. Our objectives of this work were to investigate the ability of the Edkawy cultivar to tolerate drought stress. We also focused our efforts in identifying candidates belonging to the AP2/ERF transcription factors for their important role on abiotic stress responses as explained above. Characterization of the Edkawy tomato cultivar and isolation and identification of such transcription factors will help in the engineering of tomato plants with abiotic stress tolerance.

Materials and methods

**Plant materials and drought stress treatments**

Seeds of the varieties Peto86, Strain B and Egyptian tomato cultivar Edkawy were kindly provided by the Desert Research Center (DRC), Egypt. The seeds of the varieties were grown in a nursery using cells of compressed foam trays, and totally covered with plastic bags for 7 days to increase tray temperatures which break seed dormancy until germination. The plastic bags were then removed and the seedling would start to grow. Tomato seedlings of Edkawy, Peto86 and Strain B were exposed to drought stress conditions. For drought stress treatment, seedlings were transplanted to pots filled with Peat Moss and Vermiculite. Four-leaf stage tomato seedlings were deprived of water for 7, 10 and 15 days.

**DNA extraction**

Genomic DNA was isolated from seedlings (25 days old) of the Edkawy cultivar. Fresh leaf samples (100 - 250 mg) were taken and immersed immediately in liquid nitrogen for genomic DNA extraction. Total genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen). The DNA extraction protocol was conducted according to the manufacturer’s manual. The DNA obtained was run in 1% agarose gel electrophoresis to determine DNA quality.

**ERF1 and ERF5 primer design, PCR amplification, and fragment visualization**

Tomato (Solanum lycopersicum) ERF1 and ERF5 primers were designed from the tomato Solanum lycopersicum sequences available on NCBI database using Primer3Plus (http://primer3plus.com/web_3.0.0/primer3web_in.put.htm online program. The Primer3Plus program predicted forward and reverse primers as ERF1 Forward ATGTCAAGCCCTACTAGAGA and reverse TTCCTATGATGAAATGCTAATTAAGG and also ERF5 Forward ATGGGTITCCTCCAGAGAC and reverse AAATTATATCATCAAGAGCTGAGAT. The FASTA format file which was extracted from the flat file of the ERF1 and ERF5 genes from tomato S. lycopersicum was used as a query sequence to hit the database of GenBank with a nonredundant database. PCR was performed using the GeneAmp PCR system and Veriti Applied Biosystems which was programmed as follows: initial denaturation at 95°C for 5 minutes, 35 cycles each containing (denaturation at 95°C for 45 sec., annealing gradient temperature at 40°C - 65°C for 40 sec, extension at 72°C for 45 sec.). A final extension time at 72°C for 7 min was also applied. ERF1 was amplified at 53°C, while ERF5 was amplified at 65°C annealing temperature. 1% Agarose gel electrophoresis was used to visualize the amplified DNA fragments. The gel was stained with ethidium bromide. Only 5 µl of each PCR product was directly loaded on the gel. DNA ladder was also loaded at the gel for fragment size comparison.

**DNA Sequencing and database submission**

Due to our limited resources for genome sequencing, the amplified ERF1 and ERF5 genes fragments were outsourced for sequencing (Helmy et al., 2016) by the University of Potsdam, Institute of Biochemistry and Biology, (Potsdam, Germany) through Sigma Scientific Service (Giza, Egypt) using the ABI sequencer. The nucleotide sequence of the resulting fragments was submitted to GenBank using the BankIt tool (http://www.ncbi.nlm.nih.gov/WebSub/?form=authandtool=genbank), and subjected to alignment (http://ncbi.nlm.nih.gov/BLAST) with
available data in the NCBI databases (nonredundant nucleotide database).  

**ERF1 and ERF5 genomic sequence fragment analysis and protein prediction features**

The PCR product sequence was submitted for structural analysis. Homology search was done using BLASTN software available online at (http://blast.ncbi.nlm.nih.gov/Blast.cgi) through Entrez to achieve alignment between our sequenced ERF1 and ERF5 and other genes in database. Fasta format files for S. lycopersicum genes were retrieved from the database and subjected to generate multiple sequence alignments using the CLC Sequence Viewer 6 program. The translation of the obtained genomic DNA sequence of ERF1 and ERF5 was obtained using Expasy online software (http://web.expasy.org/translate/). The 3D structure prediction of ERF1 and ERF5 was done by the Swiss model server (http://swissmodel.expasy.org). The prediction of three different models of ERF1 for six templates was generated using the mentioned protocol and the model showing the best overall stereo chemical quality was selected for further quality assessment. Similarly, the prediction of three different models of ERF5 for six templates was also generated. Additionally, PROCHECK suite analysis of ERF1 and ERF5 was carried out to investigate the details of the stereo chemical quality of the protein structure. The Ramachandran plot acquired helped to understand and analyze the highlighted region of the protein, which appeared to have unusual geometry and provided an overall assessment of the ERF1 and ERF5 structures (http://services.mbi.ucla.edu/SAVES/). Furthermore, Pfam domain search was utilized to predict and determine the domain for ERF1 and ERF5 proteins by the available online software (http://pfam.xfam.org/search/sequence). Other bioinformatics analyses, such as protein-protein interactions were performed using String online software (http://string-db.org/) while subcellular localization was performed using the ngLOC tool (http://genome.unmc.edu/ngLOC/).

**Results and Discussion**

**Drought stress analysis**

To examine the response of Edkawy to drought stress, Peto86 and Strain B varieties were chosen as a control. Peto86 is a sensitive variety to abiotic stresses (such as salinity, heat and drought) (Abdelmageed and Gruda, 2009; Amjad et al., 2013), while the Strain B variety is tolerant to abiotic stresses (mainly heat and drought with moderate tolerance to salinity) (Wahb-Allah, Alsadon, and Ibrahim, 2011; Alsadon, Sadder, and Wahb-Allah, 2013; Abdelmageed and Gruda, 2009). Thirty five days old seedlings were grown in a nursery, and then grown under water-deficit stress for 7, 10, and 15 days. After 7 days of drought stress, leaves of Peto86 plants started to wilt, roll and exhibited chlorosis, while in Strain B plants only a few leaves rolled (Figure 1). However, Edkawy plants were still robust without leaf rolling. After 10 days of drought stress, leaves of Peto86 plants were totally wilted, rolled and showed chlorosis symptoms. Additionally, some of the Strain B leaves were rolled (Figure 1). In comparison, only few of the Edkawy leaves were rolled. Similarly, Edkawy tomato cultivar plants were more tolerant to drought after 15 days of stress in comparison with Peto86 and Strain B plants.

**PCR amplification and sequencing of ERF1 and ERF5 fragments**

Total DNA was successfully extracted from the Egyptian tomato Edkawy cultivar using the DNeasy Plant Mini Kit (QIAGen). Gene specific primers were used to amplify the ERF1 and ERF5 fragments using gradient PCR (Figure 2). Amplified ERF1 and ERF5 products were sequenced and the nucleotide sequences for resulting fragments were blasted against the NCBI database to confirm its identity. Partial sequences were submitted to the GenBank and released with an accession number (KP780206) for the ERF1 gene and an accession number (KP835548) for the ERF5 gene respectively.

**ERF1 and ERF5 genomic sequence fragment analysis and protein features prediction**

**Structural analysis of nucleotide sequence**

As show in (Table 1), the nucleotide sequence of our Edkawy partial amplified ERF1 fragment showed 100% query coverage and 100% identity with 1125 bp reported in (NM_001247912.2) and (HG975517.1) with e-value (0.0). Additionally, this amplified fragment of the ERF1 nucleotide sequence showed 99% similarity with the Solanum lycopersicum ethylene response factor (ERF1) gene, promoter region (EU395634.1) and Lycopersicon esculentum ethylene-responsive factor 1 (ERF1) mRNA, complete CDs (AY044236.1) with e-value (0.0). Our next step was to confirm the identity of our identified Edkawy ERF1 gene with similar genes in the GenBank. Therefore, the CLC Sequence Viewer 6 program was used to access and search all sequences in the GenBank and graphically view the output. Our results showed 100% identity between the Edkawy ERF1 gene and (NM_001247912.2, HG975517.1, and AK323010.1). As a result, we confirmed that Edkawy has the same sequence as the ERF1 gene while we found only one SNP between the Edkawy sequence and EU395634.1 (Figure 3A).

As shown in (Table 2), the nucleotide sequence of our Edkawy partial amplified ERF5 fragment showed 99% query coverage and 99% identity with 1347 bp reported in (AY59315.1) with e-value (0.0). Moreover, this amplified fragment of the
ERF5 nucleotide sequence showed 99% similarity with the *Solanum lycopersicum* ethylene response factor (ERF5) gene, promoter region (EU395634.1) and *Lycopersicon esculentum* ethylene-responsive factor 1 (ERF1) mRNA, complete cds (AY044236.1) with e-value (0.0). Similarly, the CLC Sequence Viewer 6 program was used and our results showed 100% identity between our ERF5 gene and (HG975515.1) to confirm that Edkawy has the same sequence as ERF5 while we found only one SNP between our sequence and AY559315.1 (Figure 3B).

A phylogenetic tree was constructed using the tool on NCBI, with the horizontal lines representing the evolutionary relationships between our translated protein and other proteins in GenBank. The results revealed that our predicted proteins of ERF1 and ERF5 and the proteins from *Solanum lycopersicum* have a closer common ancestry with each other than they do...
with any other species (Figure 4A and 4B). Additionally, the figures represent the already proven relationship between proteins in this family which play their major role as ethylene-responsive transcription factor proteins.

**Protein prediction, validation and active site prediction of ERF1 and ERF5**

The 3D structure prediction was determined for ERF1 and ERF5 proteins based on homology modeling. The confirmation, validation and active site prediction of ERF1 and ERF5 were performed by Ramachandran plot statistics (%) of residues in favored, allowed and disallowed regions) which quantified the residues as shown in (Figure 5A). The red shaded regions (A, B, L represented on the plot in Figure 5A) depict the favored region having 46 residues with the most favorable combination of phi/psi values. Additional allowed regions with 5 residues were placed in a, b, l and p, generously allowed regions showed zero residues (~a, ~b, ~l, and ~p) and without disallowed regions. The maximum likelihood of finding the residues of protein in the favored region was 90.2%, whereas the likelihood in the additional allowed regions was 9.8%. The percentage of residues in the disallowed regions was zero as suggested by the Ramachandran plot for (ERF1 model 3). The Ramachandran plot statistics (% of residues in the favored, allowed and disallowed regions) quantified the residues as shown in (Figure 5B), with the red shaded regions (A, B, L represented on the plot) depicting the favored region having 40 residues with the most favorable combination of phi/psi values. Additional allowed regions with 12 residues were placed in a, b, l and p, the generously allowed region showed 1 residue (~a, ~b, ~l, and ~p) but without disallowed regions. The maximum likelihood of finding the residues of protein in the favored region was 75.5%, whereas the likelihood in the additional allowed regions was 22.6%. The percentage in the generously allowed region was 1.9% while residues in the disallowed regions were 0.0% as suggested by the Ramachandra plot for (ERF5 model 1).
Domain search in Pfam and protein-protein Interactions analysis

The Pfam domain search was conducted in order to identify the amino acid residue domain that can bind to DNA and is found in transcription factor proteins. Our results showed that all members in the ERF1 and ERF5 proteins contain the AP2 domain and the ERF1 and ERF5 domain alignment regions (Table 3).

According to our sequence obtained from the Edkawy Egyptian tomato cultivar, we found several interactions belonging to ERF1 and ERF5 proteins (Table 4). The interactome network of ERF1 analysis revealed that the EIL3, EIL1, EIL4, EIN2 subunits demonstrated strong interaction with ERF1 folding (Figure 6A). Additionally, we found several interactions belonging to the ERF5 protein. The interactome network of ERF5 analysis revealed that the UB13, Solyc04g011500.2.1, Solyc10g006660.2.1, CZFP1 and Solyc08g078190.1.1 subunits demonstrated robust interaction with ERF5 (Figure 6B).

Subcellular localization prediction

The ngLOC Subcellular localization Predictor tools program, used to determine and analyze the subcellular localization showed that ERF1 protein is localized in the nucleus, chloroplast and cytoplasm. In contrast, ERF5 protein is predicted to be localized in the nucleus, chloroplast and plasma membrane (Table 5). Therefore, both ERF1 and ERF5 are localized in the nucleus and chloroplast with high accuracy. Additionally, it is also found in the cytoplasm or plasma membrane. We also confirmed this result by the Plant-mPLoc Predicting subcellular localization of plant proteins tools program and found predicted locations for ERF1 and ERF5 (Table 5).
Figure 4. Phylogenetic tree of Edkawy ERFs. (A) Ethylene-responsive transcription factor 1 subfamily phylogenetic tree. (B) Ethylene-responsive transcription factor 5 subfamily phylogenetic tree. Edkawy ERF1 and ERF5 are highlighted in yellow.

Table 4. ERF1 and ERF5 interactome analysis

<table>
<thead>
<tr>
<th>Protein</th>
<th>Interactom ID</th>
<th>Interactome name and function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERF1</td>
<td>EIL3</td>
<td>Ethylene_insens-like_DNA-bd</td>
</tr>
<tr>
<td>ERF1</td>
<td>EIL1</td>
<td>Ethylene_insens-like_DNA-bd 1</td>
</tr>
<tr>
<td>ERF1</td>
<td>EIL4</td>
<td>EIN3-like protein</td>
</tr>
<tr>
<td>ERF1</td>
<td>EIL2</td>
<td>Ethylene_insens-like_DNA-bd 2</td>
</tr>
<tr>
<td>ERF1</td>
<td>Solyc04g054840.1.1</td>
<td>putative ETHYLENE INSENSITIVE 3-like 4 protein-like</td>
</tr>
<tr>
<td>ERF1</td>
<td>Solyc03g096630.1.1</td>
<td>putative ETHYLENE INSENSITIVE 3-like 4 protein-like</td>
</tr>
<tr>
<td>ERF1</td>
<td>Solyc00g154980.1.1</td>
<td>putative ETHYLENE INSENSITIVE 3-like 4 protein-like</td>
</tr>
<tr>
<td>ERF1</td>
<td>EIN2</td>
<td>Ethylene signaling protein</td>
</tr>
<tr>
<td>ERF5</td>
<td>UBI3</td>
<td>Ubiquitin-40S ribosomal protein S27a Ubiquitin 40S ribosomal protein</td>
</tr>
<tr>
<td>ERF5</td>
<td>Solyc04g011500.2.1</td>
<td>Actin-41; Actins are highly conserved proteins that are involved in various types of cell</td>
</tr>
<tr>
<td>ERF5</td>
<td>Solyc10g006700.1.1</td>
<td>calcium-binding protein PBP1-like</td>
</tr>
<tr>
<td>ERF5</td>
<td>Solyc10g006660.2.1</td>
<td>calcium-binding protein PBP1-like</td>
</tr>
<tr>
<td>ERF5</td>
<td>CZFP1</td>
<td>C2H2-type zinc finger protein; Cold zinc finger protein 1</td>
</tr>
<tr>
<td>ERF5</td>
<td>Solyc08g078190.1.1</td>
<td>ethylene-responsive transcription factor 5-like</td>
</tr>
</tbody>
</table>
Table 5. ERF1 and ERF5 Subcellular localization

<table>
<thead>
<tr>
<th>Protein</th>
<th>Subcellular localization</th>
<th>Prediction accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERF1</td>
<td>Nucleus</td>
<td>42.92</td>
</tr>
<tr>
<td></td>
<td>Chloroplast</td>
<td>17.99</td>
</tr>
<tr>
<td></td>
<td>Cytoplasm</td>
<td>13.8</td>
</tr>
<tr>
<td>ERF5</td>
<td>Nucleus</td>
<td>55.37</td>
</tr>
<tr>
<td></td>
<td>Cytoplasm</td>
<td>21.75</td>
</tr>
<tr>
<td></td>
<td>Chloroplast</td>
<td>8.953</td>
</tr>
</tbody>
</table>

Figure 5. 3D structure and Ramachandran plot of Edkawy ERFs. (A) ERF1 model 3 residues obtained through the PROCHECK program. (B) ERF5 model 1 residues obtained through the PROCHECK program. The plots were subdivided into the most favored regions (A,B,L,red), additional allowed regions (a,b,l,p; yellow), generously allowed regions (~a,~b,~l,~p; beige) and disallowed regions.

Figure 6. Interactome of ERF1 (A) and ERF5 (B). ERF1 and ERF5 interactomes were obtained using STRING database.
Conclusion

Two members of the AP2/ERF family from the Egyptian tomato domestic cultivar (Edkaway) were identified and isolated; ERF1 and ERF5. It has been reported that overexpression of the tomato TERF1 gene in transgenic tobacco showed increased tolerance to drought, salt, and osmotic stresses (Huang et al., 2004; Zhang et al., 2005). Additionally, overexpression of tomato ERF1 has shown to result in an increased tolerance to drought and high-salt in transgenic rice (Gao et al., 2008). When we predicted the protein structure and obtained the interaction between our ERF1 and other ERF proteins in the Solanaceae family, we found high similarity to TERF1 which suggested similar protein function. Similarly, Pan et al. (2012) reported that overexpression of SIERF5 in transgenic tomato plants resulted in high tolerance to drought and salt stress and increased levels of relative water content. When we predicted the protein structure and obtained the interaction between our Edkaway ERF5 and other ERF proteins in the Solanaceae family, we found one SNP between our ERF5 and SIERF5 proteins. Thus, this would suggest that we need to determine the gene expression and protein function for Edkaway ERF5 in future. Characterization of the Edkaway tomato cultivar and isolation and identification of such transcription factors will help in the engineering of tomato plants with abiotic stress tolerance.

Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

The authors would like to acknowledge the support from the Biotechnology Department staff, Al-Azhar University, Egypt. We are grateful to Mr. Ahmed Gameel for his help with the bioinformatics section of the manuscript.

References


Article correspondence to the corresponding author. E-mail: amlahmede@mansoura.edu.eg


Huang, Z., Z. Zhang, X. Zhang, H. Zhang, D. Huang, and R. Huang. 2004. Tomato TERF1 modulates ethylene response and enhances osmotic stress tolerance by activating expression of downstream genes. Festschrift for his help with the bioinformatics section of the University, Egypt. We are grateful to Mr. Ahmed Gameel from the Biotechnology Department staff, Al-Azhar University, Egypt. We are grateful to Mr. Ahmed Gameel for his help with the bioinformatics section of the manuscript.
