World Academy of Science, Engineering and Technology
International Journal of Bioengineering and Life Sciences
Vol:8, No:12, 2014

Identification of Conserved Domains and Motifs for GRF Gene Family

Jafar Ahmadi, Nafiseh Noormohammadi, Sedigheh Fabriki Ourang

Open Science Index, Bioengineering and Life Sciences Vol:8, No:12, 2014 publications.waset.org/9999904.pdf

***Abstract*—**GRF, Growth regulating factor, genes encode a novel class of plant-specific transcription factors. The GRF proteins play a role in the regulation of cell numbers in young and growing tissues and may act as transcription activations in growth and development of plants. Identification of GRF genes and their expression are important in plants to performance of the growth and development of various organs. In this study, to better understanding the structural and functional differences of GRFs family, 45 GRF proteins sequences in *A. thaliana, Z. mays, O. sativa, B. napus, B. rapa, H. vulgare* and *S. bicolor*, have been collected and analyzed through bioinformatics data mining. As a result, in secondary structure of GRFs, the number of alpha helices was more than beta sheets and in all of them QLQ domains were completely in the biggest alpha helix. In all GRFs, QLQ and WRC domains were completely protected except in AtGRF9. These proteins have no trans-membrane domain and due to have nuclear localization signals act in nuclear and they are component of unstable proteins in the test tube.

***Keywords*—**Domain, Gene Family, GRF, Motif.

I. Introduction

GRF Growth regulating factor, proteins play a role in the regulation of cell numbers in young and growing tissues and may act as transcription activations in growth and development of plants. The GRF family proteins contain the evolutionally conserved QLQ and WRC domains in their N-terminal region [1], [2]. The QLQ domain of GRF proteins is characterized by the conserved Gln-Leu-Gln residues, which has Phe in place of Leo [1]. The QLQ domain of GRFs shows similarity to the N-terminal region of the yeast SWI2/ SNF2 protein, another component of the SWI/SNF chromatin-remodeling complex in yeast. It suggests that QLQ domain may be involved in protein–protein interaction [3]. Another feature of QLQ domain is the absolute conservation of bulky aromatic/hydrophobic and acidic amino acid residues as Phe, Trp, Tyr, Leu, Glu or their equivalents in terms of chemical and radial properties, The Pro residue is also absolutely conserved [1]. This finding indicates that this amino acid residue is critical for the function of QLQ domain, probably or protein-protein interaction [2]. The WRC domain is also protected by the conserved residue (Trp, Arg, Cys) and contains two distinctive structural feature, namely many basic amino acids (Arg and Lys) and the conserved spacing of three Cys and one His residues, the putative zinc finger (C3H) motif. The basic amino acids are highly conserved in all the GRF horologes in other seed plants. The WRC domain also contains a functional nuclear localization signal [1]. It has been suggested that WRC domain functions in DNA binding [1], [2]. In the GRFs has been confirmed another conserved motif like GGPL, TQL, FFD, that presence of this motifs are very effective in the GRF proteins function in plant specific tissues and organs. These motifs are different in terms of level of protection, existence and probably this is one of the reasons for difference of the amount and type of the GRF protein effects in different plants.

In this study, With regard to the critical role of transcription factors on important biological processes, it is important and useful to know the biochemical and structural properties of GRF proteins.

Jafar Ahmadi is Associate Professor of Imam Khomeini International University, Qazvin, Iran (Corresponding author Address: Department of Plant Breeding, Engineering Faculty, Imam Khomeini International University, Qazvin, Iran, 34149-16818 (Tel: +982833901139; Fax: +982833780073; e-mail: njahmadi910@yahoo.com).

Nafiseh Noormohammadi was M.Sc. Student of Imam Khomeini International University, Qazvin, Iran.

Sedigheh Fabriki Ourang is Assistant Professor of Imam Khomeini International University, Qazvin, Iran (e-mail: s.ourang910@gmail.cm).

II. Material and Methods

A. The Collection of Protein Sequences

45 GRF protein sequences in *Arabidopsis thaliana* (9 proteins), *Zea mays* (14 proteins), *Oryza sativa* (12 proteins), *Brassica napus* (4 proteins), *Brassica rapa* (4 proteins), *Hordeum vulgare* (1 protein), *Sorghum bicolor* (1 protein) were identified and collected after searching in NCBI (http://www.ncbi.nlm.nih.gov).

B. Prediction of Conserved Domains

All tables and figures you insert in your document are only to help you gauge the size of your paper, for the convenience of the referees, and to make it easy for you to distribute preprints.

C. Alignment and Phylogenetic Tree

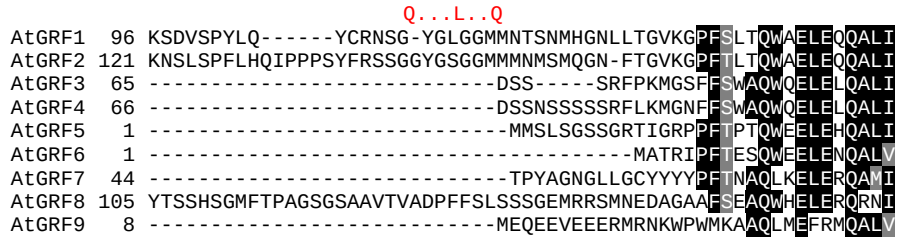
Sequence alignments to find similar and conserved regions were created by Clustal W (http://www.ch.embnet.org/software/Clustal W) and Clustal X softwares. Phylogenetic tree was constructed by MEGA4.1 software using 1000 bootstrap replicates. Rate of similarity between proteins sequences was calculated by clustalO program (http://www.uniprot.org/align).

III. Results and Discussion

Alignments confirmed conservation of QLQ and WRC domains and zinc finger motif (C3H) completely in all sequences except AtGRF9 which has Phe, a conservative substitution of Leo, in place of Leo in QLQ domain (Fig. 1). As well as predicted protein motifs that were previously

identified, TQL, FFD, GGPL and C3H are identified as motifs that they are involved in structure and act of GRF protein sequences. In AtGRF proteins, GGPL and TQL that had previously been identified, we have also confirmed the existence of these motifs. According to Kim's report [1], TQL motif was already identified in AtGRF1, AtGRF2, AtGRF3, AtGRF4, but we identified this motif only in AtGRF1 and AtGRF2. We also identified FFD motif in AtGRF3 and AtGRF4. Also according to Zhang's report [4], GGPL and TQL motifs were already identified in ZmGRF proteins. Moreover, we identified GGPL motif in ZmGRF8 and ZmGRF13. we found that C3H motif, it was not completely protected so that in these proteins Tyr and Arg residue, respectively, in ZmGRF2 and ZmGRF4 have been replaced instead of His residue. Choi et al. [5] previously reported TQL motifs in OsGRF proteins. We also identified FFD and GGPL motifs in OsGRF proteins. All of BnGRF and BrGRF proteins were shared TQL and GGPL motifs. We also identified TQL and FFD motifs in HvGRF1 and SbGRF1.

There were seven subfamilies in phylogenic tree (Fig. 2). Subfamily A was divided into two sub-subfamilies, which ZmGRF1, ZmGRF5, ZmGRF6, OsGRF3, OsGRF4 and HvGRF1 were into A1 sub-subfamily. Since these proteins in addition to QLQ and WRC domains were shared TQL and FFD motifs in their sequences, for this reason they have high similarity and have been located in a same subfamily. Probably this sequences similarity has caused that these proteins have a similar function. ZmGRF12, ZmGRF14 and OsGRF5 were into A2 sub-subfamily. These proteins were already into a subfamily in Zhang's report [6]. Also as Zhang previously reported, in C and B subfamilies were often presented AtGRF, ZmGRF, OsGRF proteins [6]. All BnGRF and BrGRF proteins, AtGRF1 and AtGRF2 were into D subfamily. All of these proteins are shared in GGPL motif. Proteins in *B. rapa* and *B. napus* had high similarity together; therefore proteins in this subfamily have similar operations. ZmGRF2, OsGRF7 and OsGRF8 were in E subfamily. OsGRF6 was independently in F subfamily. In G subfamily were ZmGRF8, ZmGRF13 and OsGRF9 that they are shared in GGPL motif in addition to other conserved domains and motifs.



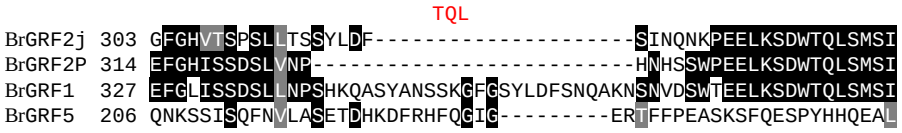
(A)

(B)

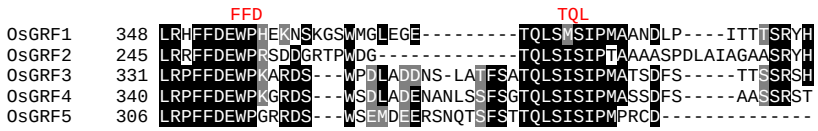
(C)

(D)

Open Science Index, Bioengineering and Life Sciences Vol:8, No:12, 2014 publications.waset.org/9999904.pdf



(E)



(F)

Fig. 1 Multiple sequence alignment of 45 GRF proteins in *A. thaliana*, O. *sativa*, *Z. mays*, *B. napus*, *B. rapa*, *H. vulgare*, and *S. bicolore*. (A) QLQ domain in *Arabidopsis thaliana*, (B) WRC domain and C3H motif in *Zea mays*, (C) GGPL motif in *Oriza sativa*, (D, E) TQL motif in *Brassica napus* and *Brassica rapa*, (F) FFD motif in *Oriza sativa*

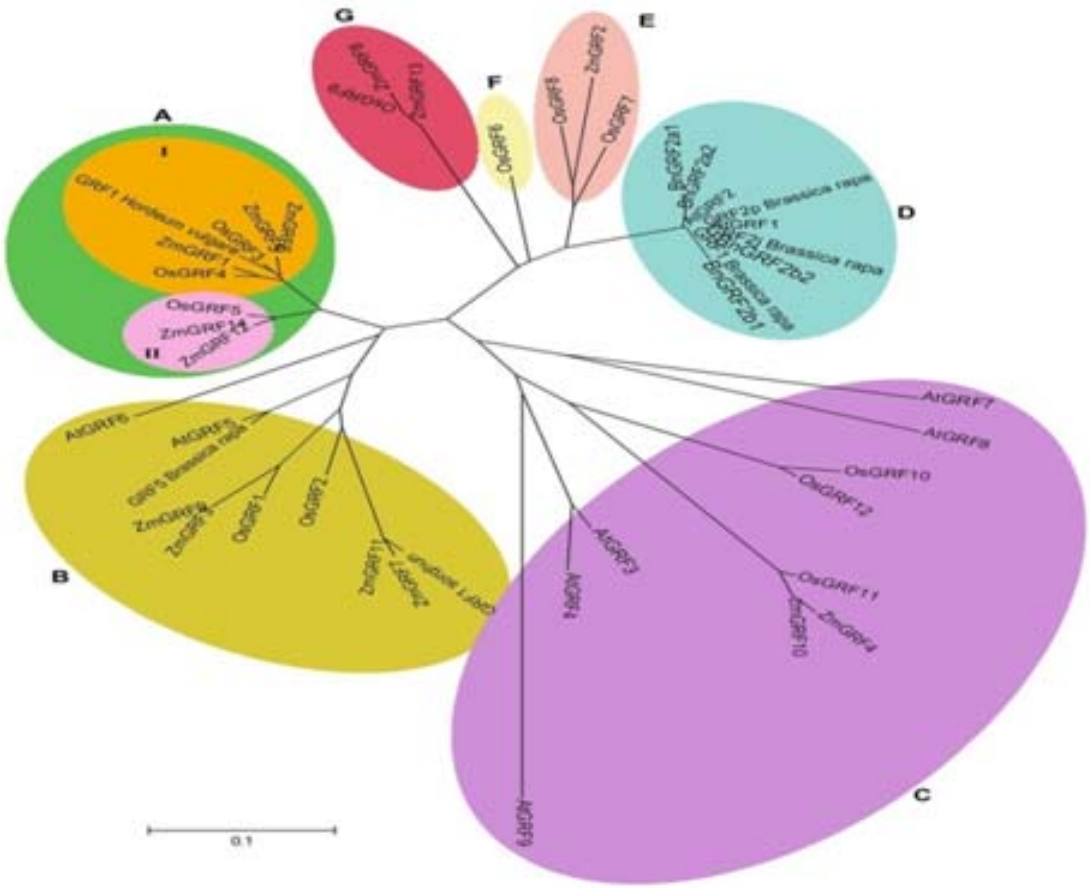


Fig. 2 Unrooted phylogenetic tree of 45 GRF proteins in *A. thaliana*, *O. sativa*, *Z. mays*, *B. napus*, *B. rapa*, *H. vulgare*, and *S. bicolore*. The deduced full-length amino acid sequences were used to create alignments with the Clustal W and Clustal X programs

IV. Conclusion

In all proteins in this study QLQ and WRC domains were completely protected except in AtGRF9. Also C3H, TQL, FFD and GGPL motifs were involved different in the structure and act of GRF proteins. This study provided useful information on the molecular and structural properties of GRF proteins for future performance reviews and it may increase general understanding of the exact mechanism of regulating the growth of other plants and obtain better results in subsequent studies.

Acknowledgment

This research was supported by a research grant in genomics laboratory at Imam Khomeini International University. We would like to thank the authority of Imam Khomeini International University (IKIU) for providing us with a good environment and facilities to complete this project

Open Science Index, Bioengineering and Life Sciences Vol:8, No:12, 2014 publications.waset.org/9999904.pdf

REFERENCES

[1] J. H. Kim, D. Choi, H. Kende, "The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis" Plant Journal. 2003, 36: 94-104.
[2] E. Van der Knaap, J.H. Kim, H. Kende, "A novel gibberellin-induced gene from rice and its potential regulatory role in stem growth" Plant Physiol. 2000, 122: 695–704.
[3] I. Treich, B. R. Cairns, T. delos Santos, E. Brewster, M. Carlson, "SNF11, a new component of the yeast SNF-SWI complex that interacts with a conserved region of SNF2" Mol. Cell. Biol. 1995, 15: 4240–4248.
[4] W. Zhang, T. Zhang, Y. Wu, J. Jiang, "Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis" Plant Cell. 2012, 15: 2719–2731. doi: 10.1105/tpc.112.098061.
[5] D. Choi, J. H. Kim, H. Kende "Whole genome analysis of the OsGRF gene family encoding plant specific putative transcription activators in rice (Oryza sativa L.)" Plant Cell Physiol. 2004, 45: 897-904.
[6] J. Kyte and R. F. Doolittle "A simple method for displaying the hydropathic character of a protein" Journal Molecular Biology. 1982, 157: 105–132.

Open Science Index, Bioengineering and Life Sciences Vol:8, No:12, 2014 publications.waset.org/9999904.pdf