BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 10(12)

In silico-prediction of downstream MYB interacting partners of MAPK3 in *Arabidopsis*

Priyanka Giri, Anil Kumar & Gohar Taj*

Department of Molecular Biology & Genetic Engineering, College of Basic Sciences & Humanities, G. B. Pant University of Agriculture & Technology, Pantnagar-263145 (Uttarakhand), India; Gohar Taj – Email: gohartajkhan@rediffmail.com; *Corresponding author

Received September 21, 2014; Accepted September 25, 2014; Published December 31, 2014

Abstract:
Protein-Protein interactions (PPIs) are vital to most biological processes thus the identification of PPIs is of primary importance. Here, we endeavor to identify the downstream interacting partners of (AtMAPK3P) in *Arabidopsis thaliana* using the docking approach. Out of 131 members of MYB transcription factors 41 members are showing interactions with AtMAPK3P while the rest are showing non-interaction. Using minimal sequence motif search as well as through docking approach several novel MYB interacting proteins were also reported in the present study which need to be confirmed by in vitro kinase assay. Together, the results obtained essentially enhance our knowledge of the MAPK interacting protein network and provide a valuable research resource for developing a nearly important link between pathogen-activated MAPK signaling pathways and downstream transcriptional programming.

Key words: Protein -Protein interactions (PPIs), Transcription factor (TF), docking, MAPK, MYB

Background:
Signal transduction pathways in plants are very well developed while at the same time they are extremely complex to reveal all the cross talks. Out of many signaling pathways involved in abiotic and biotic stress response in plants, mitogen activated protein kinase (MAPK) cascade is one of the major pathway. Mitogen-activated protein kinases (MAPKs) play central roles in the signaling network as terminal components of MAPK cascades, which are composed of at least three sequentially activated MAPK family modules MAP3Ks (for MAPK kinase kinase), MAP2Ks (for MAPK kinase), and MAPKs. The MAPKs are known to regulate a myriad of physiological and developmental responses such as cell growth, cell differentiation, hormone signaling, pathogen infection, wounding, drought, low temperature, high salinity etc. Although very little is known about MAPK's downstream targets but despite this gap in our knowledge it is clear that MAPKs interact with the transcription factors **[1].** Transcription factors are master regulators of gene expression at the transcriptional level, and controlling the activity of these factors alters the transcriptome of the plant, leading to metabolic and phenotypic changes in response to stress. The functional analysis of interactions between transcription factors and other proteins is very important for elucidating the role of these transcriptional regulators in different signaling cascades. Six major families of transcription factors involved in plant defense: basic leucine zipper containing domain proteins (bZIP), amino-acid sequence WRKYGQK (WRKY), myelocytomatosis related proteins (MYC), myeloblastosis related proteins (MYB), APETALA2/ ETHYLENE-RESPONSIVE ELEMENT BINDING FACTORS (AP2/EREBP), no apical meristem (NAM), *Arabidopsis* transcription activation factor (ATAF) and cup-shaped cotyledon (CUC) (NAC) **[2].** MYB transcription factors represent a family of proteins comprising over 131 members in *Arabidopsis* that include the conserved MYB DNA-binding domain. MYB proteins regulate transcription by binding to their target motifs as MYB/MYB homo- or heterodimers. Members of the MYB family have been characterized to various extents in a

diversity of species, and a role in germination, development, and responses to stress and hormone treatment **[3-4].** It is difficult to identify the downstream interacting partners or the substrates of MAPKs, primarily because MAPK-substrate interactions are very transient and unstable. Many experimental methods have been developed to study the protein-protein interactions including yeast two hybrid systems, affinity purification followed by mass spectrometry and the phage display libraries, but these methods have its own limitations and suffer from high false positive rate **[5].** In order to overcome these limitations *in silico* studies has been carried out. For studying protein –protein interactions *in silico* ''Docking'' strategy is extensively used in mitogen-activate protein kinase (MAPK) signaling **[6-7].** Regulation of protein activity is required for functional signaling pathways and metabolism. Protein interactions can be regulated by post-translational modifications. Protein phosphorylation is one of the most common posttranslational modifications in eukaryotic organisms and is involved in almost all cell biological processes. The phosphorylation of serine, threonine and tyrosine residues can affect protein structure, enzymatic activity and subcellular localization, interaction with other proteins as well as it is crucial in signal transduction **[8].** In eukaryotes MAPKs are catalytically inactive in their base state and require phosphorylation. The dual-specificity MAP2Ks phosphorylate MAPKs on both serine/threonine and tyrosine residues in the activation loop **[9].** Once activated, MAPKs phosphorylate many evolutionarily diverged substrates on serine or threonine residues within a minimal S/T-P motif **[10-11].** To mimic this regulation activity we phosphorylated the AtMAPK3 protein at threonine (196) and tyrosine (198) residue in TEY motif located in the activation loop (T-loop). In an effort to better understand the protein-protein interactions, we have generated a protein – protein network based on docking approach to predict the downstream interacting MYB proteins in *Arabidopsis* with AtMAPK3P.

Methodology:

The sequences for MYB transcription factors were downloaded from DATF (Database of *Arabidopsis* Transcription factor) **[12]** and the sequence of *Arabidopsis thaliana* (AtMAPK3) is downloaded from TAIR *Arabidopsis* genome database **[13]** in FASTA format. Homology modeling of the AtMAPK3 and MYB transcription factors was done with the help of MOE (Molecular Operating Environment). For constructing the structures of all, a template for homology modeling was searched with PDB search Programme of MOE. For each molecule 10 structures were generated in the database, out of which the minimized average models with maximum score, lowest E-value and with a cut off sequence identity of < 40% were selected. Structure of AtMAPK3 was phosphorylated (AtMAPK3P) with MOE as phosphorylation is essential for its enzymatic activity. The final structures were done after constructing and evaluating 3D models. Structural refinement through energy minimization model was performed using energy minimization tool keeping parameter value constant for all structures i.e Gradient: 0.5, MMFF94x Forcefield Cutoff: On=8, Off=10 Solvation: Dielectric=1, Exterior=80. The minimized structures were finally saved as *.pdb files which were validated by Ramachandran plot. After structure formation the refined structure of phosphorylated AtMAPK3P was taken as receptor and the structures of MYB transcription factor family were taken as ligand for the docking studies on the online patch dock server **[14]** which is based on shape complementarity principles and results were refined using FireDock on-line server **[15]** which rearranges the interface side chains and adjusts the relative orientation of the molecules. Taking the global energy of the interacting MYB transcription factor with MAPK3 reported in the literature MYB44 **[16]** and other interacting MYB transcription factors **[2]** as a criterion, we identified the interacting and non interacting partners of AtMAPK3P. After docking, the results were analyzed.

Results & Discussion:

Networks of protein–protein interactions provide a framework for understanding the biological processes and can give insights into the mechanism of diseases. Thus the understanding of biological mechanisms requires the knowledge of protein-protein interactions. MAPK is a conserved link between cell receptor and cell response and is mediated through gene expression which is regulated by transcription factors. To best of our knowledge no work in the literature has been done regarding the prediction of AtMAPK3P with MYB transcription factor family in *Arabidopsis*. As MYB transcription factor reported to play prominent roles in the regulation of the plant defense response **[2-16].** Therefore, the paper focuses on identifying the interacting MYB transcription factors with AtMAPK3P of *Arabidopsis thaliana* which is involved in disease signaling process through docking approach. Docking studies use geometric and steric considerations to fit the two proteins (AtMAPK3P and MYB) into a bound complex, the more stable the complex structure (less global energy) higher the probability of their interaction. The docking studies performed here, suggested that out of 131 members of MYB transcription factors, 41 members are showing interaction with AtMAPK3P while the rest are showing non-interaction **Table 1 (see supplementary material).** Sorensson *et al.* (2012) determined the primary sequence specificity of *Arabidopsis* MAPK3 and MAPK6 substrates **[17].** They indicated a minimal motif sequence L/P-P/X-S/T-P-R/K by random positional peptide library search to be the substrate for both the kinases. In an another study conducted by Hoehenwarter *et al*. (2013) they have reported other sequences surrounding the minimal motif S/T-P **Table 2 (see supplementary material).)** through the use of tandem metal oxide affinity chromatography to be the MAPK3/6 substrate **[18].** Using the minimal sequence motif identified in the above studies, we have derived a list of potential novel substrates (MYB transcription factors) from *Arabidopsis thaliana* which also showed interaction with AtMAPK3P through docking studies **Table 2.** It could be hypothesized that the amino acids surrounding the minimal S/T-P motif contribute to MAPK specificity. Those motif sequences are also considered in *Arabidopsis* in which the given hydrophobic residue is being replaced by the other hydrophobic residue and so and so for. The results of our study clearly revealed the complexity of AtMAPK3P interaction with several MYB transcription factors triggered in response to diverse upstream stimuli. Number of novel candidate AtMAPK3 substrates was predicted and need to be confirmed by in vitro kinase assay.

Conclusion:

The PPI networks can give insights into the mechanisms of diseases and provide a spectrum for the understanding of biological processes. Interaction networks can aid in designing

signal transduction pathway and help to find the disease suppressive agents as well as uncover the key genes those are responsible for senescence and defense responses against pathogens.

Acknowledgment:
Authors are grateful to sub-DIC, Bioinformatics unit at G.B. Pant University of Agriculture and Technology, Pantnagar, India for providing computational facility. This study was supported by Department of Biotechnology, Govt. of India under Programme Mode Support Project.

References:

[1] Giri P *et al. Bioinformation* 2013 **9**: 1036 [PMID: 24497732]
[2] Alves *et al. Proteomes*. 2014 **2**: 85
[3] Dubos C *et al. Trends Plant Sci.* 2010 **15**: 573 [PMID: 20674465]
[4] Prouse MB & Campbell MM, *Biochim Biophys Acta*. 2012 **1819**: 67 [PMID: 22067744]
[5] Zolghadr K *et al. Methods Mol Biol.* 2012 **812**: 275 [PMID: 22218866]
[6] Bardwell L, *Biochem Soc Trans*. 2006 **34**: 837 [PMID: 17052210]
[7] Sharrocks AD *et al. Trends Biochem Sci*. 2000 **25**: 448 [PMID: 10973059]
[8] Zulawski M *et al. Nucleic Acids Res.* 2013 **41:** D1176 PMID: 23172287
[9] Yang KY *et al. Proc Natl Acad Sci U S A*. 2001 **98:** 741 [PMID: 11209069]
[10] Clark-Lewis I *et al. J Biol Chem.* 1991 **266**: 15180 [PMID: 1907971]
[11] Alvarez E *et al. J Biol Chem*. 1991 **266:** 15277 [PMID: 1651323]
[12] (http://datf.cbi.pku.edu.cn/.)
[13] (http://www.arabidopsis.org/)
[14] http://bioinfo3d.cs.tau.ac.il/PatchDock/
[15] http://bioinfo3d.cs.tau.ac.il/FireDock/
[16] Persak H & Pitzschke A, *PLoS ONE.* 2013 **8**: e57547 [PMID: 23437396]
[17] Sorensson C *et al. Biochem* J. 2012 **446:** 271 [PMID: 22631074]
[18] Hoehenwarter W *et al. Mol Cell Proteomics*. 2013 **12:** 369 [PMID: 23172892]

Edited by P Kangueane
Citation: **Giri *et al.*** Bioinformation 10(12): 721-725 (2014)

BIOINFORMATION

open access

Supplementary material:

Table 1: MYB transcription factor family genes showing interaction and non-interaction with AtMAPK3P

Gene loci Id	41	AtMYB15, AtMYB9, AtMYB24, AtMYB17, AtMYB33, AtMYB19, AtMYB105, AtMYB45, AtMYB49, AtMYB52, AtMYB108, AtMYB13, AtMYB50, AtMYB3R3, AtMYB55, AtMYB66, AtMYB26, AtMYB112, AtMYB106, AtMYB82, AtMYB44,
Showing Interaction		AtMYB23, AtMYB78, AtMYB3, AtMYB3R2, AtMYB101, AtMYB80, AtMYB87, AtMYB77, AtMYB39, AtMYB74, AtMYB0, AtMYB61, AtMYB30, AtMYB41, AtMYB22, AtMYB93, AtMYB103, AtMYB16, AtMYB14, AtMYB11
Gene loci Id	90	AtMYB1, AtMYB10, AtMYB100, AtMYB118, AtMYB119, AtMYB12, AtMYB120, AtMYB121, AtMYB25, AtMYB27, AtMYB28, AtMYB29, AtMYB46, AtMYB47, AtMYB48, AtMYB5, AtMYB102, AtMYB104, AtMYB107, AtMYB109, AtMYB110,
Showing Non-Interaction		AtMYB111, AtMYB113, AtMYB114, AtMYB115, AtMYB116, AtMYB117, AtMYB67, AtMYB68, AtMYB69, AtMYB7, AtMYB70, AtMYB71, AtMYB72, AtMYB122, AtMYB123, AtMYB124, AtMYB125, AtMYB18, AtMYB2, AtMYB20, AtMYB21, AtMYB8, AtMYB81, AtMYB83, AtMYB84, AtMYB31, AtMYB32, AtMYB34, AtMYB35, AtMYB36, AtMYB37, AtMYB38, AtMYB4, AtMYB40, AtMYB43, AtMYB94, AtMYB95, AtMYB96, AtMYB51, AtMYB53, AtMYB54, AtMYB56, AtMYB57, AtMYB58, AtMYB59, AtMYB6, AtMYB60, AtMYB62, AtMYB63, AtMYB64, AtMYB65, AtMYB73, AtMYB75, AtMYB76, AtMYB79, AtMYB85, AtMYB86, AtMYB88, AtMYB89, AtMYB90, AtMYB91, AtMYB92, AtMYB42, AtMYB97, AtMYB98, AtMYB99, AtMYB3R1,AtMYB3R4, AtMYB3R5,

Table 2: List of potential interacting MYB transcription factors identified through both minimal sequence motif search and docking approach.

Peptide sequence	Common	Only through minimal sequence motif search
NIMGVESNVQPLTS(ph)PLSK		VQSPL (AtMYB81) VQSPL (AtMYB104) PSSPL (AtMYB97) LSSPL (AtMYB120)
NVEDLSNDFSNTFLLDEELDLEHRS(ph)PR	AtMYB41	THSPR (AtMYB41)
IHHPPS(ph)PR		AFSPR (AtMYB115) LPSPR (AtMYB3R4)
IHHPSS(ph)PR		LCSPR (AtMYB120) ASSPR (AtMYB38)
LLPLFPVTS(ph)PR		ASSPR (AtMYB38) LCSPR (AtMYB120)
TYVADVSEYLGSNS(ph)PRDPYLER	AtMYB41	THSPR (AtMYB102) MSSPR (AtMYB41)
VAPIPPS(ph)PVKVPQPVPEPVVLEPPQMFVDQR	AtMYB78	AVSPV (AtMYB78) LLSPV (AtMYB3R1) AASPV (AtMYB3R5)
ATDILQGS(ph)PVESGPTTPLPDKK	AtMYB3R3	QGSPV (AtMYB3R3)
SHLRPPGNISGSQS(ph)PVESSGLYHSK		SSSPV (AtMYB96) CMSPV (AtMYB3R4)
VHNPVVESSIQPQRS(ph)PR	AtMYB41	THSPR (AtMYB41)
YREATNLIPS(ph)PR		LPSPR (AtMYB3R4) AFSPR (AtMYB115)
HATIQPFDVLPS(ph)PTFSAAR	AtMYB26 AtMYB44	ILSPT (AtMYB26) PVSPT (AtMYB48) ALSPT (AtMYB70) LASPT (AtMYB59) PGSPT (AtMYB44) LPSPT (AtMYB120)

LLHSAYDPQNRPAIEVHLVQVQPAGISADLDSTSNDAGHSS(ph)PTRK	AtMYB74	TSSPT (AtMYB95)
		SNSPT (AtMYB2)
		TYSPT (AtMYB21)
		SSSPT (AtMYB74)
		SSSPT (AtMYB86)
EGYSQSQSRPVYGLS(ph)PTLNHR	AtMYB44	ALSPT (AtMYB70)
		ILSPT (AtMYB26)
		PVSPT (AtMYB48)
		LPSPT (AtMYB120)
		LASPT (AtMYB59)
		PGSPT (AtMYB44)
SSWTSESYQLKPQSSFSGSHPSGS(ph)PNAR	AtMYB45	YLSPN (AtMYB0)
	AtMYB0	SFSPN (AtMYB45)
		YLSPN (AtMYB66)
		YLSPN (AtMYB23)
IITDYVGS(ph)PATDPMR	AtMYB108	AVSPA (AtMYB108)
	AtMYB3R3	IASPA (AtMYB3R3)
	AtMYB30	ALSPA (AtMYB30)
LLEHFLVQEQTEGSS(ph)PSR	AtMYB3R3	FSSPS (AtMYB3R1)
		INSPS (AtMYB3R4)
		FTSPS (AtMYB43)
		ITSPS (AtMYB65)
		ASSPS (AtMYB94)
		GCSPS (AtMYB3R3)
STPVRKPHTSTADLLTWSEVPPPDSP(ph)SSASR	AtMYB19	ESPSS (AtMYB19)
VHNPVVESSIQPQRS(ph)PR		THSPR (AtMYB102)
NVEDLSNDFSNTFLLDEELDLEHRS(ph)PR		THSPR (AtMYB102)
VAPLTGPYGLAGDFAGHLPSSILS(ph)PQAQGFYYR		ALSPQ (AtMYB46)
		IPSPQ(AtMYB124)
		IPSPQ (AtMYB88)