



Contents lists available at BioMedSciDirect Publications

## International Journal of Biological & Medical Research

Journal homepage: www.biomedscidirect.com



### Original Article

# In Silico approach to identify transcription factor binding sites and Cis-regulatory elements in tubulin gene promoter

Hira Mubeen, Shahid Raza

University of South Asia, Lahore, Pakistan

#### ARTICLE INFO

##### Keywords:

gene regulations  
motif  
core promoter  
TFB's.

#### ABSTRACT

Transcription factor (TF) binding sites or motifs (TFBMs) are functional cis-regulatory DNA sequences that play an essential role in gene transcriptional regulation. TFs play important roles in cellular physiology, developmental processes and responses to environmental stimuli. The importance of promoter analysis relies on how abundantly TFBMs and TFs are integrated into reference databases. The primary components of promoter are cis acting regulatory regions. Various tools and software's are being used to analyze and predict regulatory motif within promoter region of a gene. PlantPAN serves as a systematic platform for plant promoter analysis and reconstructs transcriptional regulatory networks. In the present study, we identified various motif including: light responsive elements, cis regulatory elements, core promoter elements and transcription factor binding motifs on the enhancer sequences of tubulin gene participating in variety of functions occurring in cellular processes.

© Copyright 2010 BioMedSciDirect Publications IJBMR - ISSN: 0976:6685. All rights reserved.

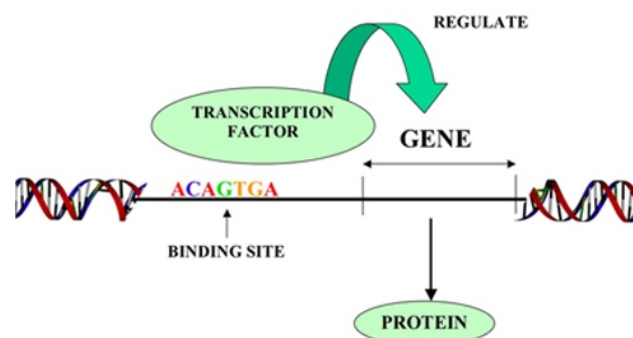
### 1. Introduction

Gene expression regulation is essential for all cellular processes with proper transcriptional control. Gene expression is the process in which gene is decoded to form an mRNA sequence to produce the particular protein sequence. For this, a molecule called a transcription factor is needed. Transcription factors are key regulators of gene expression and play a critical role in the life cycle of both animals and plants. They can act as sequence specific DNA binding proteins which will bind with promoter region of gene to make it express. However, it is difficult to discover many cis-regulatory elements with existing software's because many transcription factors are still unknown. [1,2]

Studies on transcription factors (TFs) and their corresponding cis-acting elements in promoters have attracted much attention due to their role in gene expression and its regulation. However, defining all functional binding sites within an identified promoter is difficult and the existence of some additional binding sites should be assumed. TFs regulate target genes through the recognition of specific cis-regulatory sequences in promoter regions. In plants, TFs play important roles in cellular physiology, developmental processes and responses to environmental stimuli. Also, their role in

transcription regulation can never be ignored. Bioinformatics prediction of gene regulation on the basis of the presence of TF binding sites in the promoters is of priority concern in systems biology.

For this purpose, several public web-based resources were promptly established. For instance, TRANSFAC is the commercial database manually collecting experimentally verified TFs, transcription factor binding sites (TFBSs) and matrix-based target prediction profiles [3,4]. Two other well-known TFBS repositories, PLACE and JASPAR, were created to facilitate the identification of TFBSs in input sequences [5,6]. However, only information on TFBSs is provided rather than their corresponding



\* Corresponding Author : **Hira Mubeen**

University of South Asia,  
Lahore, Pakistan

**Fig 1: A transcription factor molecule binds to DNA at its binding site and regulates the production of a protein from its gene.**

### Tubulin

Tubulin is a globular protein that constitutes the building block of microtubules, a major element of the cytoskeleton. Tubulin heterodimers are composed of two major classes, alpha tubulin and beta tubulin. Since the initial cloning of chicken and tubulin cDNA during the 1980s, tubulin genes of a wide variety of organisms have been isolated and characterized. Identification of functional transcripts has helped classifying and tubulin into classes named isotypes. In addition, the complexity of the tubulin population is amplified by various post-translational modifications. These post-translational modifications allowed subdividing isotypes into several isoforms.<sup>[7,8,9,10]</sup>

### Cis-Regulatory Module

The TFs bind to sequences called as transcription factor binding sites (TFBS) in the regulatory regions of the gene called enhancers which are organized in the form of modules, called as Cis-Regulatory Module (CRM). CRM sare regulatory sequences located few kilo bases away from gene of interest and bind to specific TFs at specific developmental stage to result in specific cell specification. [11]Overall, gene expression is regulated by the combination of all CRMs acting on genes throughout the organism's life. Studies shown, there exists as many as 10-fold more CRMs than genes. [12]Interaction between the TFs and CRMs form a development transcriptional regulatory network, encoding the specification and differentiation programmes of various cell types that are expressed at a particular stage in the development and finally lead to a full grown organism.

### PlantPAN

Provides an informative resource for detecting transcription factor binding sites (TFBSs), corresponding transcription factors (TFs), and other important regulatory elements in a promoter or a set of plant promoters. Additionally, TFBSs, CpG islands, and tandem repeats in the conserve regions between similar gene promoters are also identified. The current PlantPAN release (version 2.0) contains 16 960 TFs and 1143 TF binding site matrices among 76 plant species. In addition to updating of the annotation information, adding experimentally verified TF matrices, and making improvements in the visualization of transcriptional regulatory networks, several new features and functions are incorporated.

In the present study, we identified putative transcription factor binding motifs on the enhancer sequences of tubulin gene participating in variety of functions occurring in cellular processes. These mapped sequences are of great importance in finding out other regions within the genome having the similar or identical set of motifs. The importance of having these motifs mapped is the knowledge about which transcription factor will bind to particular sequence. The interconnection between the sequences and transcription factor regulate the gene expression.

## Material and Methods

### TFB Prediction

Transcription factor (TF) binding site prediction remains a challenge in gene regulatory research due to degeneracy and potential variability in binding sites in the genome.

#### PlantPAN

Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups.<http://plantpan2.itsps.ncku.edu.tw/>.

#### PLACE

Database of motifs found in plant cis-acting regulatory DNA elements, all from previously published reports. It covers vascular plants only.<http://www.dna.affrc.go.jp/htdocs/PLACE/>.

#### Motif Analysis

Motif analysis involves identification of common patterns from promoter region without knowing their binding sites.

#### PlantCARE

PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>.

## Results

### Analysis of Transcription Factor Binding Sites (TFB's)

PlantPAN software ([www.plantpan.mbc.nctu.edu.tw/](http://www.plantpan.mbc.nctu.edu.tw/)) identifies the transcription factors that are key regulators of gene expression. The putative transcription factor binding sites (TFBs) in Beta Tubulin (BT)genepromoters are illustrated below in table 1.

**Table 1: Transcription factor binding sites**

Transcription Factor	Species	Sequence	Site
Athb-1	Arabidopsis	gtgcATAATaagga	117
ANT	Arabidopsis	taacgCCCGatt	221
AG	Arabidopsis	aactgaTTTGGtga	887
AGL3	Arabidopsis	acgttaTATGGtcct	881
Athb-1	Arabidopsis	tttcATAATcaag	174

**Table 2: Shows light responsive elements**

Motif	Species	Sequence
GAG-motif	<i>Arabidopsis thaliana</i>	AGAGAGT
I-box	<i>Solanum tuberosum</i>	TATTATCTAGA

**Table 3: Shows core promoter elements**

Motif	Species	Sequence
TATA-box	<i>Arabidopsis thaliana</i>	TATAAA
TATA-box	<i>Brassica napus</i>	ATATAT

**Table 4: Shows cis-regulatory elements**

Motif	Species	Sequence
5UTR Py-rich stretch	<i>Lycopersicon esculentum</i>	TTTCTTCTCT
ACE	<i>Petroselinum crispum</i>	CTAACGTATT
TC-rich repeats	<i>Nicotianatabacum</i>	ATTTTCTTCA

## Discussion

Many cis-regulatory motifs were detected in the beta tubulin gene promoter. TC-rich repeats are important cis-acting element involved in defense response. Beta tubulin gene promoter was found to be enriched with conserved Transcription binding site motifs AGL3, AG, Athb-1, ANT. These motifs are located in sense (+) and (-) strands of promoter sequences. Major transcription binding sites are AGL3, AG, Athb-1, ANT. The transcription factor AGL3 is expressed in all above-ground vegetative organs; AGL3 may be involved the transcriptional regulation of genes. The prediction of regulatory motifs, TFBS form an essential link in comparative genomics. These sequences are evolutionary conserved, and eventually we can find out the orthologous of these genes in higher and complex organism which help in understanding molecular mechanisms. But some of the hurdles to predictions are: these modules are located far away from the genes they regulate. Next, the presence of multiple transcription factor binding sites for various TFs leads to combinatorial control of gene regulation, thus making it difficult to associate with one gene. [13,14]

## Conclusion

A large proportion of gene regulations happen at the transcriptional level through the binding of transcription factors to short regulatory DNA sequences in the upstream regions of TF target genes, i.e., promoter. The problem of computationally identifying TF binding sites and motifs (TFBMs) remains a challenge. Hence it is suggested to develop and maintain more advanced tools to analyze and identify the cis regulatory elements dispersed over promoter regions of genes

## References

1. H. Leung, F. Chin, S.M. Yiu, R. Rosenfeld and W.W. Tsang, JCB, 12(6), 686-701 (2005).
2. M. Li, B. Ma, and L. Wang, Journal of Computer and System Sciences, 65, 73-96 (2002).
3. Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K. et al. (2006) TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. Nucleic Acids Res., 34, D108-D110.

4. Wingender, E. (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Bioinformatics, 9, 326-333.
5. Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H. et al. (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res., 42, D142-D147.
6. Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res., 27, 297-300.
7. Cleveland L, MacDonald C, Rutter K (1980). Number and evolutionary conservation of  $\alpha$ - and beta-tubulin and cytoplasmic beta- and gamma-actin genes using specific cloned cDNA probes. Cell. 20: 95-105.
8. Cleveland L, Sherline K (1981). Unpolymerized tubulin modulates the level of tubulin mRNAs. Cell. 25: 537-546.
9. Luduena (1998). Multiple forms of tubulin: different gene products and covalent modifications. Int. Rev. Cytol. 178: 207-275.
10. MacRae (1997). Tubulin post-translational modifications--enzymes and their mechanisms of action. Eur. J. Biochem. 244: 265-278.
11. Girardot C. Deciphering enhancer activity in Drosophila based on transcription factor occupancy and chromatin state chromatin state characterization (Doctoral dissertation, Université Pierre et Marie Curie-Paris VI). 2012.
12. Bryantsev AL, Cripps RM. Cardiac gene regulatory networks in Drosophila. See comment in PubMed Commons below BiochimBiophysActa. 2009; 1789: 343-353.
13. Berman BP, Pfeiffer BD, Lavery TR, Salzberg SL, Rubin GM, Eisen MB, et al. Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in Drosophila melanogaster and Drosophila pseudoobscura. Genome Biol. 2004; 5: R61.
14. Berman BP, Nibu Y, Pfeiffer BD, Tomancak P, Celniker SE, Levine M, et al. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the Drosophila genome. See comment in PubMed Commons below Proc Natl AcadSci U S A. 2002; 99: 757-762.