Implications of RBBP6 in various types of Cancer

Khan, F [1]; Allam, M [2]; Tincho, MB [1] and Pretorius, A [1]

[1] University of the Western Cape, Cape Town South Africa

[2] South African National Bioinformatics Institute, Cape Town South Africa

Abstract. *Background:* The 250 kDa retinoblastoma binding protein 6 (RBBP6) was cloned over a decade ago and was found to bind both the p53 and Rb1 tumor suppressor proteins. In addition, RBBP6 has been associated with multiple biological functions, such as mitosis, mRNA processing, translation and ubiquitination.

Objectives: In the current study, an in silico approach was used to identify RBBP6 binding partners. The information will be used to investigate the relationship between RBBP6 and its binding partners as well as to further probe the biological function of RBBP6 and its binding partners in various cancer types.

Materials & Methods: Putative genes were identified as possible binding partners for RBBP6. These genes were subjected to expression profiling to determine differential expression. Promoter analysis were done on the promoters of the genes encoding RBBP6 and each of its binding partners to investigate the underlying regulatory elements involved in the progression of cancer.

Results: In the current study 20 binding partners were identified for RBBP6. Expression profiling revealed 14 types of cancer in which RBBP6 and its binding partners show differential expression. Functional annotation indicated that RBBP6 and its binding partners are involved in similar biological processes. De novo motif discovery revealed 10 unique signatures present in the promoters of RBBP6 and its binding partners.

Discussion: From previously explained results it is clear that RBBP6 and its binding partners are involved in similar processes and this fact is indicated by all the genes being involved. All the genes also share similar and unique regulatory elements within their promoters validating the possible co-expression relationship between RBBP6 and its binding partners.

Conclusions: The study showed that RBBP6 and its binding partners are involved in the same biological processes and that they share underlying regulatory elements within their promoters. Because they are 2-fold differentially expressed in a number of cancers and they share common regulatory elements it could be inference that they are highly involved in the progression of cancer.

1 Introduction

One of the most limiting aspects of biological research in the post-genomic era is the capability to integrate massive datasets on gene structure and function for producing useful biological knowledge [1]. The identification of stable and reliable human gene to gene co-expression networks is essential to unravel the interactions and functional correlations between human genes at an omic scale [2]. Exploration and analysis of gene expression data using genome-wide microarrays is a technique often used in genomic studies to find co-expression patterns and locate groups of co-transcribed genes. This kind of studies has been used in model organisms, like yeast to discover gene functions, to define biological processes and to find related transcription factors and their products [3]. The main features of expression patterns that give a wide utility in bioinformatic studies are: the functional information associated with the gene [4], the high conservation of gene co-expression groups along evolution [5] and the high correlation of these groups with biomolecular pathways or reactions [6].

In humans, retinoblastoma binding protein (RBBP6) has been shown to play an important role in apoptosis and cell cycle regulation through its interaction with the tumor suppressor genes p53 and Rb1 [7]. It is an evolutionarily conserved 250-kDa multi-domain protein that has been found in a wide variety of eukaryotic organisms ranging from spore forming unicellular organisms, to plants and to vertebrates [8]. RBBP6 was first identified in mouse testis and was implicated in controlling cell proliferation and differentiation [9]. Two research groups identified the partial cDNAs of the mouse RBBP6 gene and named them PACT [10] and P2P-R [9]. The gene is known to possess six different domains that have been characterized and linked with different types of cancer such as breast cancer [11]. These domains include the domain with no name (DWNN), zinc knuckle, RING finger, SR, Rb, and a p53-binding domain which are present in the RBBP6 in vertebrates. In other species homologues of the gene are found however, they lack the p53- and Rb-binding domain [8]. Its involvement in wide range of biological processes such mitosis, mRNA processing, ubiquitination, and translation, further gives rise to the in depth investigation of RBBP6 [7, 9, 12, 13]. Differential expression level of RBBP6 has been implicated in many cancer types including but not limited to colorectal, breast, cervical, ovarian, and prostate cancer when compared to normal samples [11]. Three splice variant namely P2P-R, PACT, and RBQ-1 have been found in RBBP6 in mice [10, 12, 14]. The P2P-R was shown to localize to the nucleolus of interphase cells and the periphery of chromosomes in cells undergoing mitosis [15].

Due to the ability to regulate p53 pathway and to prevent tumorigenesis and the consequent potential role of RBBP6 as a target for cancer therapy, plenty of studies have been performed on human and mouse cells [8]. However, little research has been carried out to identify the interaction, co-expression analysis and protein modification of this gene and its protein products on so many other genes and proteins that are co-expressed. The importance of identifying biological networks and predicting molecular interactions has been emphasized by several studies [16, 17]. Such studies emphasize that when knowledge about DNA variation within populations is interfaced with data on gene expression, protein interactions and DNA-protein binding, biological

networks can be constructed that are predictive of the physiological molecular interactions and disease susceptibility [18]. In 2011, using bioinformatics and molecular approach we identified the regulatory role of the two promoter sequence in RBBP6 on apoptosis [19]. Motadi et al., 2011, using the bioinformatics tool, GeneNetwork, established that large genetically defined transcription networks that include P2P-R exist in fat cells and eye tissues [17]. Generally the genetic architecture of common human diseases is characterized by interactions between genes, i.e., epistasis [22]. Accordingly, the focus of recent research has shifted from identifying single locus susceptibility [20, 21] to quantifying interaction effects between multiple candidate loci throughout the human genome [22]. The aim of this study is to characterize RBBP6 in relation to genes that are co-expressed with it, identifying the pathways and their role in regulating cell growth, using the bioinformatics approaches.

2 Methodology

2.1 Co-expression analysis

For the purpose of co-expression analysis a combination of the two approaches were used to identify genes that were co-expressed with RBBP6, namely a guide-gene approach and a non-targeted approach described by Aoki and his colleges (Approaches for Extracting Practical Information from Gene Co-expression Networks in Plant Biology). The results obtained are then visualized and evaluated based on their hypothesis.

2.2 RBBP6 protein-protein interaction network

Protein coding sequence of RBBP6 was used as a query to search for known and putative protein-protein interactions between RBBP6 and its co-expressed genes using the STRING database version 9 [23]. STRING uses a scoring system that is intended to reflect the evidence of predicted interactions. To produce each of the interaction networks, parameters were judiciously chosen as a follows: (i) a confidence level of 0.7, (ii) a network depth of 4 and (iii) restricting to show only the top 20 interactions [23].

2.3 Expression profiling of RBBP6 and its binding partners across cancer

Microarray expression data of RBBP6 and its binding partners were downloaded from Gene Expression Atlas (GEA) (http://www.ebi.ac.uk/gxa/). Information regarding the differential regulation of all the genes were extracted. And were used to further probe the implication of RBBP6 in an array of cancers.

2.4 GO term enrichment and functional annotation of RBBP6 and its binding partners

GO term enrichment was performed to determine which functional annotations were associated with both RBBP6 and its binding partners. By determining which GO terms were enriched for the genes, a link could be made to, their involvement in biological processes, molecular functions or cellular components. The information obtained will thus give a clear understanding of the functional relatedness of RBBP6 and its binding partners.

2.5 Promoter Sequence extraction

Promoter sequences of RBBP6 and its binding partners were extracted from the human promoter database (Cold spring Harbour) available at (http://rulai.cshl.edu/cgi-bin/CSHLmpd2/hspd.pl.). The HUGO Gene Nomenclature Committee (HGNC) identifiers of the genes were used as a query to extract the promoter sequences. Promoter sequences flanking from 1000bp upstream and 200bp downstream were extracted for RBBP6 and its binding partners. The promoter sequences were exported in FASTA format for further analyses.

2.6 Promoter Content Analysis

The Promoter sequences extracted from the human transcription factor database were used as the test dataset for Transcription Factor Binding Sites (TRANSFAC) analyses [24]. The selected promoters were compared against the background dataset available in TRANSFAC which contains all known human promoters. An enrichment analysis of all the transcription factors present in our dataset in comparison to the background set was generated. Only transcription factors with a p-value of 0.05 or above associated with the prediction were selected for further analyses [24].

2.7 De novo motif discovery

Using the command line, the promoter proximal regions were extracted from the promoter sequences file. Lists of promoter proximal regions were assigned as the unaligned sequences. Then the file with the unaligned sequences was used as an input for the motif-based sequence analysis tools (MEME) [25]. The MEME result files then used as input for average motif affinity (AMA), Gene Ontology for Motifs (GOMO) and motif comparison tool (TOMTOM) [26].

3 Results

The STRING database search revealed protein-protein interactions between RBBP6 and 20 proteins with a score of at least 0.7 (Figure 1). Six of these proteins are DNA-binding proteins (SON, GTF2H3, YBX1, ZDHHC17, CIZ1 and BLZF1) and two are RNA-binding proteins (BICC1 and PUM2). The interaction network also includes onco-proteins (MDM2, TP53 and Rb1), cytokines (AZ12), cytokines signaling suppressors (ASB5), signal transduction molecules (GNLA, GNL3L, ASAP2 and TOM1L1), growth factor (LTBP4) and transcription regulators (DMXL1 and SCYL1).

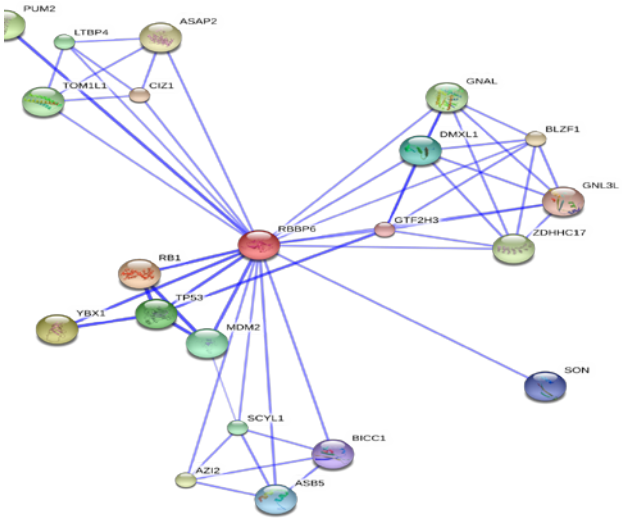

Fig. 1. Graphical representation of RBBP6 and its binding partners identified using STRING database version 9. Interactions are based on a number of criteria including co-expression and experimental validation.

Fourteen types of cancer were identified as having RBBP6 and its binding partners differentially regulated during the progression of the disease in comparison to normal conditions (Figure 2). The cancer types were ranked by assigning a value to each of the genes according to category, that is up regulated genes were assigned the value of 2, down regulated genes were assigned the value of 1 and genes that showed no change under the influence of each of the cancer types were assigned the value 0. These results showed highly similar expression patterns between RBBP6 and its binding partners and various cancer types.

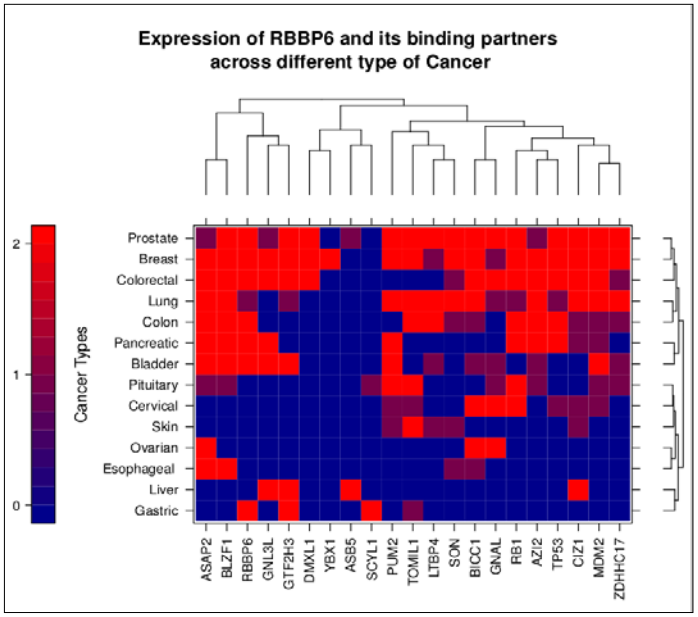

Fig. 2. Expression profile of RBBP6 and its binding partners under the influence of 14 different types of cancer. The key represents up-regulated (red), down-regulated (purple) and genes that show no change in expression (blue).

GO annotations were extracted for three categories namely, biological processes, molecular function and cellular components. Under the category ”biological processes” ten GO terms were enriched for both RBBP6 and its binding partners. These GO terms include ”cel lular co mponent o rganization”, ”cel lular p rotein metabolic p rocess”, ” macromolecule modification”, ”r egulation o f cel lular co mponent o rganization”, ”growth”, ”response to DNA damage stimuli's”, ”regulation of growth”, ”negative regulation o f apoptosis”, ”n egative regulation of cel l d eath” and ”regulation o f transport” . F or t he cat egory ”cel lular co mponents” G O t erms ”i ntracellular p art”, ”intracellular membrane bounded organelle”, ”nucleus”, ”membrane enclosed lumen” and ”intracellular organelle part” were enriched for both RBBP6 and its binding partners. For the category ”molecular function” only three GO terms were enriched for both sets of genes these included ”binding”, ”protein binding” and ”nucleic acid binding”. These results are shown in Figures 3.

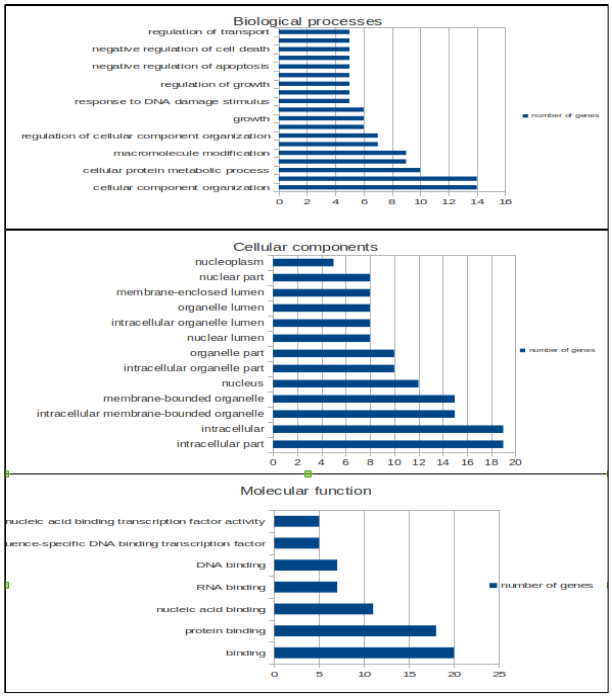

Fig. 3. Bar chart representation of GO terms enriched for RBBP6 and its binding partners in comparison to the entire human genome under the categories: Biological processes, Cellular Components and Molecular Function

MEME motifs are represented by position-specific probability matrices that specify the probability of each possible letter appearing at each possible position in an occurrence of the motif. These are displayed as "sequence LOGOS", containing stacks of letters at each position in the motif. The total height of the stack is the "information content" of that position in the motif in bits. The height of the individual letters in a stack is the probability of the letter at that position multiplied by the total information content of the stack. Ten motifs were identified as being significant amongst RBBP6 and its binding partners as shown in Figure 4. A similarity matrix for all the motifs identified for RBBP6 and its binding partners was generated and is represented in Figure 5. This matrix shows the correlation between the input genes and the identified motifs.

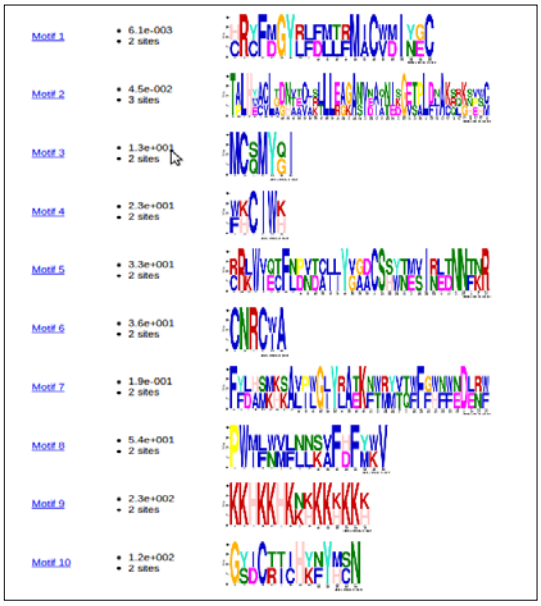

Fig. 4. List of identified motifs shared amongst RBBP6 and its binding partners. Where column 1 contains the motif name, column2 contains the E- value associated with the motif

Motif	Width	Best possible match	Similarity									
			1	2	3	4	5	6	7	8	9	10
1	22	CRCFDGYRFFMLFMICWDINGC	-	0.18	0.33	0.31	0.21	0.39	0.22	0.36	0.10	0.23
2	47	TALIWACLADNVACIKLLLEAGANVDAINIDGETPLDIAKQQKNEWC	0.18	-	0.29	0.44	0.22	0.36	0.14	0.24	0.21	0.26
3	7	MCQMYGI	0.33	0.29	-	0.20	0.30	0.17	0.45	0.40	0.08	0.33
4	6	WICIWI	0.31	0.44	0.20	-	0.41	0.34	0.35	0.41	0.22	0.35
5	35	CRKWIQCFNDNDCIIYGGDCSIWNMVIREDNNFNR	0.21	0.22	0.30	0.41	-	0.32	0.16	0.24	0.17	0.25
6	6	CNRCWA	0.39	0.36	0.17	0.34	0.32	-	0.29	0.26	0.13	0.32
7	36	FFDIMMIKAVIWGIYRAEKFWMMVQWFFWFWEDERW	0.22	0.14	0.45	0.35	0.16	0.29	-	0.29	0.17	0.27
8	17	FWMFWMFNNKVFDFMWV	0.36	0.24	0.40	0.41	0.24	0.26	0.29	-	0.06	0.29
9	15	KKIKKIKNIKKIKKI	0.10	0.21	0.08	0.22	0.17	0.13	0.17	0.06	-	0.13
10	14	GYDCRICIYFYICN	0.23	0.26	0.33	0.35	0.25	0.32	0.27	0.29	0.13	-

Fig. 5. Similarity matrix for all the motifs identified for RBBP6 and its binding partners. This matrix shows the correlation between the the genes that were used to identify the motifs

4 Discussion

The data reported in this paper establish that the function of gene products can be predicted by determining their association with specific genetically defined biological networks, using systems genetics approaches and tools, such as STRING, Transfac and MEME. The generation of gene networks has previously been used to predict a molecular interaction between RBBP6 and Pum2 via the gene network software. The gene transcript which is most consistently co-expressed with RBBP6 in three tissues

of the BXD mouse genetic reference panel. That prediction was confirmed by molecular studies showing that Pum2 binds to a specific sequence in the 5 'UTR of RBBP6 mRNA to modulate RBBP6 expression [20]. The current study expands the use of generating gene networks in an attempt to identify biological networks in which RBBP6 is a member, leading to the prediction of new function(s). Data derived using these systems approaches suggest that RBBP6 is an important member of large genetically defined transcription networks in various cancers, thereby having the potential to impact the expression of many hundreds of genes. Molecular studies by Peidis and colleagues (2010) confirm the prediction showing that RBBP6 can indeed function as a transcriptional co- repressor.

GO annotation and functional enrichment indicate that both RBBP6 and its binding partners share commonly enriched terms across the three categories namely biological processes, molecular function and cellular components. Go terms such as negative regulation of cell death, negative regulation of apoptosis and regulation of growth which are commonly enriched amongst RBBP6 and its binding partners suggests that these genes may share an underlying regulatory mechanism. This finding shows that RBBP6 and its binding partners are tightly co-expressed and validates our previous findings. Under the category of cellular components, GO annotations such as intracellular, nucleus, intracellular membrane bound organelle were most prevalent. This finding are to be expected as both RBBP6 and its binding partners are localized to the intracellular membrane region. Furthermore the link between RBBP6, its binding partners and the enriched biological processes hence suggest RBBP6's implication in disease progression.

A total of 10 motifs were identified amongst the promoter proximal regions of RBBP6 and its binding partners. The associated E-value for each motif is statistically significant (E-value < 1) indicating that these motif predictions are highly accurate. RBBP6 and its binding partners were used as input to create a similarity matrix. The similarity matrix produced showed a direct correlation between the input genes and the motifs identified indicating that all the genes have a positive correlation to RBBP6 and its binding partners. This result suggests a strong link between co-expression of genes and transcriptional regulation. The motifs identified were seen to be involved in similar biological process as previously seen between RBBP6 and its binding partners. These findings further suggest a concrete link between co-expression, correlation, transcriptional regulation and disease progression.

Future prospects include experimentally validating the findings from the in-silico studies, in particular generating comprehensive expression profiles using RT-PCR. Although much work has been covered within the realms of this study a lot still needs to be done to fully understand the mechanism by which genes are regulated with regard to disease progression. Furthermore a more comprehensive look at theregulatory aspects of both RBBP6 and its binding partners promoters may lead to a more indepth understanding of the underlying mechanisms involved in controlling the progression of disease. Furthermore it will also allow for the further characterization of the cancer progression pathway.

5 References

1. Hebb A, Moore C, Bhan V, et al. Expression of the inhibitor ofapoptosis protein family in multiple sclerosis reveals a potentialimmunomodulatory role during autoimmune mediated demyeli-nation.Mult Scler. 2008;14:577–594
2. Blaydes, J. P. (2010) Cooperation between MDM2 and MDMX in the regulation of p53. In: Hupp T, Ayed A (eds). p53: Landes bioscience.
3. van Noort,V.et al. (2004) The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model. EMBO rep. ,5, 280–284.
4. Lee,H.K.et al.(2004) Coexpression analysis of human genes across many microarraydata sets.Genome Res.,14, 1085–1094
5. Chibi, M., Meyer, M., Skepu, A., Rees, D. J. G., Moolman-Smook, J. C., Pugh, D. J. R. (2008). RbBP6 interacts with multi-functional protein YB-1 through its RING finger domain, leading to ubiquitination and proteasomal degradation of YB-1. *Journal of Molecular Biology*, 384:908-916.
6. Kim KC, Geng L, Huang S (2003) Inactivation of a histone methyltransferase by mutations in human cancers. Cancer Res 63:7619–7623.
7. Pugh, D. J. R., Eiso, A. B., Faro, A., Lutya, P. T., Hoffmann, E., Rees, D. J. G. (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquintin-like pathway. *BMC Structural Biology*, 6: 1-12.
8. Huang da, W, Sherman, BT and Lempicki, RA. 2009a. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37: 1–13.
9. Gao S, Witte MM, Scott RE: P2P-R protein localizes to the nucleolus of interphase cells and the periphery of chromosomes in mitotic cells which show maximum P2P-R immunoreactivity. *J Cell Physiol* 2002, 191:145-154.
10. Witte MM, Scott RE: The proliferation potential protein-related (P2P-R) gene with domains encoding heterogeneous nuclear ribonucleoprotein association and Rb1 binding shows repressed expression during terminal differentiation. *Proc Natl Acad Sci U S A* 1997, 94:1212-1217
11. Simons A, Melamed-Bessudo C, Wolkowicz R, Sperling J, Sperling R, Eisenbach L, Rotter V: PACT: cloning and characterization of a cellular p53 binding protein that interacts with Rb. *Oncogene* 1997, 14:145-155.
12. Stuart,J.M.et al.(2003) A gene-coexpression network for global discovery ofconserved genetic modules.Science,302, 249–255
13. Bykov, V. J. N., Issaeva, N., Shilov, A., Hultcrantz, M., Pugacheva, E., Chumakov, P., Bergman, J., Wiman, K. G., Selivanova, G. (2002). Restoration of the tumor suppressor functions to mutant p53 by a low-molecular-weight compound. *Nature medicine*, 8 (3): 282-288.
14. Mbita, Z., Meyer, M., Skepu, A., Hosie, M., Rees, J., Dlamini, Z. (2012). De-regulation of the RBBP6 isoform 3/DWNN in human cancer. Molecular cell Biochemistry, 362: 249-262.
15. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: Model validation, design issues and standard error application. Genome Biol. 2001;2:RESEARCH0032
16. Gao S, Scott RE: P2P-R protein overexpression restricts mitotic progression at prometaphase and promotes mitotic apoptosis. *J Cell Physiol* 2002, 193:199-207.
17. Harris C. C. (1996). Structure and Function of the p53 Tumor Suppressor Gene: Clues for Rational Cancer Therapeutic Strategies. Journal of the National Cancer Institute, 88 (20): 1442-1455.

18. Pretorius,A, Kaur M, Wamalwa M, February MF, Essack M, Bajic VB, Rees DJG (2011). Functional analyses and characterization of the human RBBP6 promoters on a combination of molecular biology and insilico approaches provide additional evidence for RBBP6 role in apoptosis. GSTF international journal on Bioinformatics and Biotechnology. Vol 1, No 1
19. Motadi, C. R., Bhoda, K. D., Dlamini, Z. (2011). Expression and function of retinoblastoma binding protein 6 (RbBP6) in human cancer. Immunobiology, 216: 1065-1073.
20. Scott, R .E, W hite-Grindley E, Ruley H E, Chesler EJ , W illiams R W (2005): P2P-R expression i s g enetically c oregulated w ith c omponents of t he translation machinery and with PUM2, a translational repressor that associates with P2P-R mRNA. J Cell Physiol , 204:99-105.
21. Hashimoto S, Onodera Y, Hashimoto A, Tanaka M, Hamaguchi M, et al. (2004) Requirement for Arf6 in breast cancer invasive activities. PNAS vol. 101(17): 6647–6652.
22. Wang Y, Wu C, Ji Z, Wang B, Liang Y (2011) Non-Parametric Change-Point Method for Differential Gene Expression Detection. PLoS ONE 6(5): e20060.
23. Mering, C . V . (2003). S TRING: a d atabase o f p redicted functional as sociations b etween proteins. Nucleic Acids Research, 31(1):258–261.
24. Matys, V . et al., 2006. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic acids research, 34(Database issue), pp.D108–10.
25. Bailey, T.L. et al., 2006. MEME : discovering and analyzing DNA and protein sequence motifs. , 34, pp.369–373.
26. Bailey, T .L. e t a l., 20 09. M EME S U ITE : t ools f or m otif di scovery a nd s earching. , 37(May),pp.202–208.