Integrative analysis of the Regulatory Region of the FGFR3 Oncogene

Josep Bau[1], Marta Cullell[1] and Jordi Solé-Casals[2]

1- Bioinformatics and Medical Statistics Group, Department of Systems Biology, Universitat de Vic, 08500, Spain

2- Digital Technologies Group, Department of Digital Technologies and Information, Universitat de Vic, 08500, Spain

Abstract The study of transcriptional regulation often needs the integration of diverse yet independent data. In the present work, sequence conservation, prediction of transcription factor binding sites (TFBS) and gene expression analysis have been applied to the detection of putative transcription factor (TF) modules in the regulatory region of the FGFR3 oncogene. Several TFs with conserved binding sites in the FGFR3 regulatory region have shown high positive or negative correlation with FGFR3 expression both in urothelial carcinoma and in benign nevi. By means of conserved TF cluster analysis, two different TF modules have been identified in the promoter and first intron of FGFR3 gene. These modules contain activating AP2, E2F, E47 and SP1 binding sites plus motifs for EGR with possible repressor function.

Keywords: FGFR3, Bladder Cancer, Transcription Factors, Binding Sites, Conserved Sequence, Gene Expression Regulation.

Introduction

Fibroblast growth factor receptor 3 (FGFR3) is frequently activated by mutation and/or over-expressed in urothelial carcinoma (UC) and benign nevi of the skin (BN), triggering the MAPK pathway and leading or contributing to cell growth and tumorigenesis.

Although some transcription factors like SP1, AP2δ and E47 are known to promote FGFR3 transcription by binding to specific sites in the promoter and first intron [1-3], the regulatory mechanism is not yet fully understood. The discovery of additional TFs involved in FGFR3 transcriptional regulation could help explain the over-expression observed in UC and BN patients.

Several tools and algorithms already exist for scoring known transcription factor binding motifs in gene promoters. In most cases, however, a combination of dif-

ferent methodologies is required to minimize false positives. For example, TFBSs found in evolutionary conserved regions (ECRs) identified by cross-species comparative analysis have an increased likelihood of playing a role in gene regulation. In addition to this, expression correlation may also be due to a regulatory relationship. Finally, as transcription factors often work cooperatively, binding in close physical proximity, co-occurring motifs can indicate the presence of regulatory modules. The configuration of such modules can be, up to some extent, evolutionarily conserved, as regulatory elements driving shared temporal and/or spatial aspects of gene expression are most probably located on genomic elements that contain similar modules of certain TFBS [4]. Therefore, the finding of functional relationships or common expression patterns among genes that share the detected regulatory modules would further support the prediction.

Material and Methods

Regulatory region covering 2Kb 5' from the transcription start site and the 5'UTR and first intron were obtained from Ensembl for human, chimpanzee, gorilla and mouse. Conserved TFBSs were detected in ECRs by ortholog multiple local alignment with MULAN [5].
Expression data was obtained from GEO (http://www.ncbi.nlm.nih.gov/geo/) from three Bladder Cancer (Acc. GSE3167, GSE7476, GSE13507) and one Benign Nevi (Acc. GSE3189) RNA expression sets. Probes with low variability (Inter Quartil Range<0.5) were excluded and data was processed from CEL files, when available, with robust multiarray analysis (RMA).
Sample selection was conducted to include normal tissue *vs.* diseased with high FGFR3 expression levels (case > mean(ctr) + sd(ctr)). In total, 87 controls and 100 cases from 4 different experimental datasets were used.
TFs with conserved binding sites in the FGFR3 promoter were analyzed for co-expression with FGFR3 applying Pearson correlation to the expression data. Differential expression (Rank Product) was also calculated for all TFs between normal tissue samples and disease samples with increased FGFR3 expression.
Candidate TFBS modules were studied with SynoR [4] conducting human *vs.* mouse genome scans for conserved clusters of TFBS.

Results

TFBS detection in ECRs of the FGFR3 regulatory region returned 129 putative binding sites from 58 different transcription factors. Expression analysis was performed for these TFs and consistent over-expression and positive correlation with FGFR3 was found for AP2α and TCF3 while EGR1 showed clear under-expression and negative correlation (Table 1, Fig. 1).

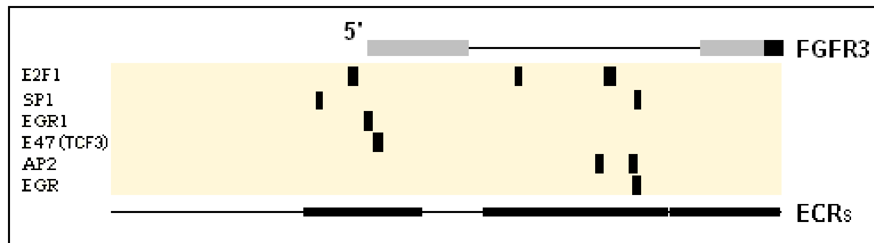

Fig. 1. Distribution of relevant TFBSs and conserved regions (ECRs) relative to FGFR3 gene features (line: intron; box: exon; grey: 5'UTR; black: CDS)

Table 1. Results for Pearson correlation with FGFR3 expression. Arrows indicate differential expression of cases with high FGFR3 expression compared to controls.

Dataset	AP2-α	AP2-γ	TCF3	TFDP1	SP1	EGR2	EGR1	N
BC-GSE3167	0.55 ↑	0.09	0.70 ↑	0.71 ↑	-	-	-0.57 ↓	18
BC-GSE7476	0.89 ↑↑	0.06 ↑	0.73 ↑	0.61 ↑	-	-0.68	-0.80 ↓↓	6
BC-GSE13507	0.69 ↑↑	0.33	0.38 ↑↑	-0.23	0.4 ↑↑	-0.55 ↓↓	-0.56 ↓↓	140
BN-GSE3189	0.82 ↑↑	0.82 ↑↑	0.63 ↑	0.79 ↑	0.5 ↑	-0.55 ↓↓	-0.69 ↓↓	23

The TCF3 gene encodes the alternative spliced TFs, E12 and E47. The latter plays major roles in embryogenesis and its phosphorylation is regulated by a MAPK. TFDP1 shows overexpression and correlation with FGFR3 in three out of four experiments (Table 1). This TF is a dimerization partner for the E2F family of TFs, which play a key role in cell cycle regulation. It is noteworthy that FGFR3 overexpression has been observed after adenoviral infection with E2F-1 [6,7].

Overlapping EGR and SP1 sites have been found in the first intron of FGFR3. EGR1 is known to interact with SP1 (MINT, Acc. 7384908) and is associated with negative regulation of transcription (GO:122). Moreover, repressor activity of Egr-1 has been reported, probably by preventing interactions between SP1 and the transcriptional machinery [8]. Interestingly, Egr-1 is upregulated by FGF1 [9], so an expression decrease should not be expected when MAPK pathway is active.

These results suggest an activating role for AP2α and E47 and a possible repression activity of EGR1.

Two evolutionary conserved modules of the detected TFs have been identified with SynoR: (1) a promoter module (PM) consisting of clusters of SP1, E2F1, EGR and E47 binding sites and (2) an intronic module (IM) with AP2, EGR and SP1 TFBSs. Instances of the PM have been found in ECRs of 634 genes and IM instances in ECRs of 799 genes. Only 10% of the IM instances were found within coding sequences while most were located in regions with possible regulatory function like promoters (35%) and introns (32%). Functional relationships and common expression patterns have also been found between FGFR3 and the genes with PM or IM detected in non-coding regions.

Conclusions

Two different conserved TF modules have been identified in the promoter and first intron of FGFR3 gene. These modules contain activating AP2, E47, E2F and SP1 binding sites and motifs for EGR with possible repressor function.

AP2α is the AP2 family member with best correlation with FGFR3 expression. Therefore, AP2α might have a stronger regulatory effect in UC and BN than AP2δ, whose FGFR3 activating function have already been reported in neuroblastoma cell lines [2].

A strong correlation with FGFR3 expression has been found for E47. This TF is functionally associated with the MAPK pathway and has been shown to induce FGFR3 expression in reporter assays [3]. It is, therefore, likely to play a role in the transcriptional regulation of FGFR3 expression.

The inverse correlation observed between the expression levels of EGR1 and FGFR3 suggests that EGR1 may be repressing FGFR3. The detection of overlapping SP1 and EGR1 sites (Fig. 1) are consistent with the repression mechanism proposed by *Tan et al.* [8] where EGR1 would interfere with the transcriptional activity of SP1 by binding to overlapping sites.

Acknowledgments: This work has been partially supported by the Generalitat de Catalunya (Spain) through the grant 2009SGR-581 and by the Ministerio de Ciencia e Innovación (Spain) under the grant AR2009-0010.

References

1. Perez-Castro AV, Wilson J, Altherr MR (1997) Genomic organization of the human fibroblast growth factor receptor 3 (FGFR3) gene and comparative sequence analysis with the mouse Fgfr3 gene. Genomics 41: 10-16
2. Tan CC, Walsh MJ, Gelb BD (2009) Fgfr3 is a transcriptional target of Ap2δ and Ash2l-containing histone methyltransferase complexes. PLoS One 4: e8535
3. Funato N, Ohtani K, Ohyama K, Kuroda T, Nakamura M (2001) Common regulation of growth arrest and differentiation of osteoblasts by helix-loop-helix factors. Mol Cell Biol 21: 7416-7428.
4. Ovcharenko I and Nobrega MA (2005) Identifying synonymous regulatory elements in vertebrate genomes. Nucleic Acids Res 33: 184-194
5. Ovcharenko I, Loots GG, Giardine BM, Hou M, Ma J, Hardison RC, Stubbs L, Miller W (2005) Mulan: multiple-sequence local alignment and visualization for studying function and evolution. Genome Res 15: 184-194
6. Jamshidi-Parsian A, Dong Y, Zheng X, Zhou HS, Zacharias W, McMasters KM (2005) Gene expression profiling of E2F-1-induced apoptosis. Gene 344: 67-77
7. Müller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E, Prosperini E, Vigo E, Oliner JD, Helin K (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. Genes Dev 15: 267: 267-285
8. Tan L, Peng H, Osaki M, Choy BK, Auron PE, Sandell LJ, Goldring MB (2003) Egr-1 Mediates Transcriptional Repression of COL2A1Promoter Activity by Interleukin-1β. J Biol Chem 278: 17688-17700
9. Delbridge GJ and Khachigian LM (1997) FGF-1–Induced Platelet-Derived Growth Factor-A Chain Gene Expression in Endothelial Cells Involves Transcriptional Activation by Early Growth Response Factor-1. Circ Res 81: 282-288.