Transcriptional regulation of human *DUSP4* gene by cancer-related transcription factors

Tatiana Varela[1,2], Natércia Conceição[1,3,4,*], Vincent Laizé[1], M. Leonor Cancela[1,3,4]

[1] Centre of Marine Sciences, University of Algarve, Faro, Portugal

[2] PhD in Biomedical Sciences, Faculty of Medicine and Biomedical Sciences, University of Algarve, Faro, Portugal

[3] Faculty of Medicine and Biomedical Sciences, University of Algarve, Faro, Portugal

[4] Algarve Biomedical Center, University of Algarve, Faro, Portugal

*Corresponding author

Faculty of Medicine and Biomedical Sciences, University of Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

Email: nconcei@ualg.pt

Tel.: 00 351 289 800 057

Keywords: DUSP4, Transcriptional regulation, CTCF, FOXA1, STAT3, YY1

Running title: *DUSP4* transcriptional regulation

Conflict of interest: The authors declare that they have no conflict of interest.

Authors Contribution: VL and MLC planned the study. TV and NC performed the experiments. TV, NC and VL wrote the manuscript. All authors analysed the data and revised the manuscript.

Funding: This study received national funds from the Portuguese Foundation for Science and Technology (FCT) through the project UIDB/04326/2020 (CCMAR).

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

ABSTRACT

Dual specificity phosphatase 4 (DUSP4), a member of the dual specificity phosphatase family, is responsible for the dephosphorylation and inactivation of ERK, JNK and p38, which are mitogen-activated protein kinases involved in cell proliferation, differentiation and apoptosis, but also in inflammation processes. Given its importance for cellular signalling, DUSP4 is subjected to a tight regulation and there is growing evidence that its expression is dysregulated in several tumors. However, the mechanisms underlying *DUSP4* transcriptional regulation remain poorly understood. Here, we analysed the regulation of the human *DUSP4* promoters 1 and 2, located upstream of exons 1 and 2, respectively, by the cancer-related transcription factors (TFs) STAT3, FOXA1, CTCF and YY1. The presence of binding sites for these TFs was predicted in both promoters through the *in silico* analysis of *DUSP4*, and their functionality was assessed through luciferase activity assays. Regulatory activity of the TFs tested was found to be promoter-specific. While CTCF stimulated the activity of promoter 2 that controls the transcription of variants 2 and X1, STAT3 stimulated the activity of promoter 1 that controls the transcription of variant 1. YY1 positively regulated both promoters, although to different extents. Through site-directed mutagenesis, the functionality of YY1 binding sites present in promoter 2 was confirmed. This study provides novel insights into the transcriptional regulation of *DUSP4*, contributing to a better comprehension of the mechanisms of its dysregulation observed in several types of cancer.

1. INTRODUCTION

Dual specificity phosphatases (DUSPs) are well-established negative regulators of mitogen-activated protein kinases (MAPKs), being responsible for their inactivation through the dephosphorylation of both the phosphoserine/threonine and phosphotyrosine residues of the conserved T-X-Y motif (Bermudez et al., 2010). These non-transmembrane phosphatases are involved in many physiological mechanisms including metabolic homeostasis and immunity. DUSP family comprises 25 members with different substrate specificities that are divided in two sub-families: MAP kinase phosphatases (MKPs) that interact with MAPKs through the kinase-interacting motif (KIM) and atypical DUSPs that do not contain the KIM domain (Pulido & Lang, 2019). MAPKs are essential components of signal transduction pathways. They are evolutionary conserved and involved in various cellular functions, including stress response, cell proliferation and differentiation (Kondoh & Nishida, 2007). Given their critical role, the magnitude and duration of MAPK activity must be tightly regulated to guarantee appropriate cellular responses (Liu et al., 2006) and abnormal MAPK signalling has been associated with human disorders such as cancer (Keyse, 2008).

Given their MAPK-regulating activity, DUSPs were hypothesized to be key players in cancer induction and progression (Bermudez et al., 2010). In this regard, the dual specificity phosphatase 4 (DUSP4) – a specific inhibitor of ERK, JNK and p38 MAPKs (Sim et al., 2015) – has been associated with mechanisms of diseases including cancer (Owens & Keyse, 2007), and there are growing evidences that its expression may be abnormally regulated in some tumors, e.g. up-regulated in colorectal cancer (Gröschl et al., 2013; Saigusa et al., 2013; Sim et al., 2015; Varela et al., 2020), breast cancers (H. Y. Wang et al., 2003) and pancreatic tumors (Yip-Schneider et al., 2001) and down-regulated in breast carcinomas (Armes et al., 2004), glioblastomas (Waha et al., 2010) and ovarian carcinomas (Sieben et al., 2005). Three alternative spliced transcripts have been described for *DUSP4*, and while transcript variant 1

initiates in exon 1, both transcript variants 2 and X1 start in exon 2. Variant X1 is similar to variant 2 in the 5'-UTR but retains intron 2 (Varela et al., 2020). As for other DUSPs, the transcription of *DUSP4* is tightly regulated, although very little is known about the transcription factors involved in this regulation (Bermudez et al., 2010). In this regard, while *DUSP4* transcripts are present at a low level in resting, non-stimulated cells, gene transcription is rapidly induced following the activation of the target MAPK pathways by mitogenic or stress stimulation (Bermudez et al., 2010). Among cancer-related transcription factors, only p53 (Shen et al., 2006), E2F1 (J. Wang et al., 2007), Hoxa10 (H. Wang et al., 2007) and ETS1 (Plotnik et al., 2014) were reported to regulate *DUSP4* transcription.

Given the scarce information on the transcriptional regulation of *DUSP4* in a cancer context, this work aims at collecting basic knowledge on its promoter activity and regulation by cancer-related transcription factors STAT3, FOXA1, CTCF and YY1 to better understand the role of DUSP4 in mechanisms regulating cancer induction and progression.

2. MATERIALS AND METHODS

2.1. *In silico* analysis of *DUSP4* promoter

The sequences of the regions upstream of exon 1 (~1.5 kb) and exon 2 (~1.4 kb) in human *DUSP4* gene (referred thereafter as promoters 1 and 2, respectively) were retrieved from the Ensembl database (www.ensembl.org, accession number ENSG00000120875). These sequences were analysed for the presence of transcription factor binding sites (TFBSs) using (1) ConTra v3 prediction tool (www.bioit2.irc.ugent.be) with the stringency criteria core=0.95 and similarity matrix=0.85, (2) JASPAR (www.jaspar.genereg.net) with a score threshold of 80% and (3) LASAGNA-Search 2.0 (www.biogrid-lasagna.engr.uconn.edu) with a cut-off $p<0.001$. Promoter sequences were also scanned for CTCF binding sites with a score threshold

set to 70% in JASPAR and a cut-off p-value set to $p<0.05$ in LASAGNA-Search 2.0. Only binding sites predicted by at least two of the above prediction tools were considered.

TFBSs present in *DUSP4* promoters were also retrieved from Encyclopedia of DNA Elements (ENCODE) ChIP-seq data, mapped on UCSC Genome Browser (genome.ucsc.edu), that provide evidence for regions of the genome that bind a specific TF in a given cell line or tissue sample.

2.2. Amplification of *DUSP4* promoter fragments

Promoter fragments (three for each promoter) were amplified from human genomic DNA using specific primers (Table 1) and KAPA Hifi DNA polymerase (KAPA Biosystems) following manufacturer instructions. PCR conditions were as follows: initial denaturation step at 95ºC for 5 min; 35 cycles of amplification (denaturation at 98ºC for 20 s; annealing at 56ºC for 15 s; extension at 72ºC for 3 min); and a final extension at 72ºC for 7 min. Promoter fragments (referred thereafter as F1.1-3 and F2.1-3 for promoters 1 and 2, respectively) were inserted into pCRII-TOPO plasmid (Invitrogen) and sequenced.

2.3. Preparation of promoter-luciferase reporter constructs

DUSP4 promoter fragments inserted into pCRII-TOPO were sub-cloned into pGL3-Basic vector (Promega) upstream of the firefly luciferase gene as follows: F1.1, F1.2 and F2.3 were excised from pCRII-TOPO using *Hin*dIII and *Xho*I endonucleases and F1.3 using *Kpn*I and *Xho*I, while F2.1 and F2.2 were PCR amplified from pCRII-TOPO using primers containing *Bgl*II or *Hin*dIII restriction sites (Table 1) then digested by both endonucleases. Promoter fragments were cloned into pGL3-Basic vector previously digested with the appropriate endonucleases, then sequenced on both strands to confirm their correct insertion.

2.4. Site-directed mutagenesis of putative YY1 binding sites

YY1 binding sites at position -1267/-1251 and -293/-288 in *DUSP4* promoter 2 were mutated using F2.1 and F2.3 luciferase reporter constructs, respectively, as DNA template and the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) following manufacturer instructions. Binding site-specific primers are listed in Table 1.

2.5. Transcription factor expressing vectors

Vectors expressing the human STAT3 (pcDNA3.1-hSTAT3), human CTCF (pcDNA3-hCTCF), human YY1 (pCMV-hYY1) and mouse FOXA1 (pcDNA3-mFOXA1) were kindly provided by Dr. Xinliang Mao (Soochow University, China), Dr. Dolores Delgado (Universidad de Cantabria, Spain), Dr. Yang Shi (Harvard medical school, MA, USA) and Dr. Jérôme Eeckhoute (University of Lille, France), respectively.

2.6. Cell culture and transient transfection assays

Human embryonic kidney 293 cells (HEK-293, ATCC number CRL-1573) were grown in DMEM (Dulbecco's modified eagle medium, Invitrogen) supplemented with 10% (v/v) of fetal bovine serum (FBS, Invitrogen), 1% (v/v) of L-glutamine (200 mM; Invitrogen) and 1% (v/v) of penicillin-streptomycin (10,000 U/ml; Invitrogen). Cell cultures were maintained at 37ºC in a humidified atmosphere with 5% of CO_2 and sub-cultured every three days to a lower cell density. The day prior to their transfection, cells were seeded in a 24-well plate at 5×10^4 cells per well. Cultures at 50–60% confluence were transfected with 250 ng of promoter-luciferase constructs (prepared with the Promega endotoxin-free PureYield plasmid DNA miniprep kit) and 5 ng of pRL-TK vector (plasmid encoding *Renilla* luciferase; Promega) using 1 μl of XtremeGENE HP DNA Transfection reagent (Roche) and DMEM without FBS and antibiotics. For co-transfections, 25 ng of each transcription factor-expressing vector were used. pGL3-

Basic and pGL3-control vectors were also separately transfected to serve as negative and positive controls, respectively. At two days post-transfection cells were lysed, and firefly and *Renilla* luciferase activities were measured using the Dual-Luciferase Reporter Assay kit (Promega) in a multiplate reader (BioTek Synergy 4). Construct-specific luciferase activity was calculated from the ratio between firefly and *Renilla* luciferase activities. All experiments were carried out at least three times and in duplicates. Results are shown as the mean of luciferase activity ± standard deviation (SD). Statistical analysis was performed using Prism 5 (GraphPad Software) and significant differences were determined through one-way ANOVA followed by Tukey's post-test for $p<0.05$.

3. RESULTS AND DISCUSSION

3.1. Functional analysis of *DUSP4* promoter regions

Based on the information retrieved from transcript sequence databases, human *DUSP4* contains two transcription start sites, thus two promoter regions. The first promoter region is situated upstream of exon 1 and controls variant 1 (NM_001394) transcription, while the second promoter is placed upstream of exon 2 and controls the transcription of variants 2 (NM_057158) and X1 (XM_011544428). Mining of the Ensembl database revealed that several *DUSP* genes (e.g *DUSP3*, *DUSP8*, *DUSP9*, *DUSP13*, *DUSP14*, *DUSP16*, *DUSP22* and *DUSP26*) also have more than one transcription start site and therefore different promoters. The presence of 2 promoters was confirmed in *DUSP22* by Stelzer et al. (2015).

The transcriptional activity of the two promoter regions of human *DUSP4* was determined using promoter reporter constructs (i.e. promoter fragments cloned upstream of the firefly luciferase cDNA) transiently transfected into HEK-293 cells. The relative luciferase activity of the promoter constructs was significantly higher than the activity of pGL3-Basic (Figure 1), indicating that both promoter regions are functional. Basal activity was higher in promoter 1

than that in promoter 2 (approximately 10-fold for the longest fragments), possibly suggesting that transcript variant 1 may be more expressed than variants 2 and X1 in HEK-293. In this regard, GTEx RNA-seq data from UCSC Genome Browser (genome.ucsc.edu) showed that variant 1 is indeed always more expressed (data not shown), regardless the tissue type.

Relative luciferase activity of F1.1 (fragment -1454/+248 of promoter 1) was approximately 150-fold the activity of pGL3-Basic and deletion of the sequence -1454/-743 doubled this activity in F1.2 (fragment -742/+248; $p<0.001$), suggesting the presence of negative regulatory elements in the deleted sequence. On the contrary, removal of the -742/-282 fragment resulted in a significant decrease ($p<0.01$) of F1.3 (fragment -281/+248) luciferase activity to approximately half the activity of F1.2, suggesting the presence of positive regulatory elements in this sequence. A similar pattern was observed for the basal activity of promoter 2 constructs. Relative luciferase activity of F2.1 (fragment -1298/+304) was approximately 17-fold the activity of pGL3-Basic and deletion of sequence -1298/-1062 doubled this activity in F2.2 (fragment -1061/+304; $p<0.01$) suggesting the presence of negative regulatory elements. The elimination of the sequence -1061 to -544 resulted in a significant decrease ($p<0.001$) of the luciferase activity of F2.3 (fragment -543/+304), approximately 3-fold the activity of F2.2, suggesting the presence of positive regulatory elements in the eliminated sequence.

3.2. *In silico* analysis of *DUSP4* promoter regions

Promoter regions 1 (1.5 kb) and 2 (1.4 kb) were searched for the presence of transcription factors binding sites (TFBSs) using three different *in silico* prediction tools. Since *DUSP4* expression was shown to be dysregulated in several cancers, we focused on transcription factors which altered expression has been associated with cancer development: STAT3 (Signal transducer and activator of transcription 3), FOXA1 (Forkhead box A1), CTCF (CCCTC-Binding Factor) and YY1 (Yin Yang 1). Putative binding sites for each of these transcription

factors was mapped on both promoter regions (Figure 2) And additional information on respective DNA sequences is provided in Supplementary Figure S1. Six and three binding sites for STAT3 (consensus motif GGAA) were predicted in promoter 1 and 2, respectively, Two and one binding sites for FOXA1 (consensus motif AAACA) were predicted in promoter 1 and 2, respectively. Nine and four binding sites for YY1 (consensus motifs CCAT and ACAT; Yant et al. (1995)) were predicted in promoter 1 and 2, respectively. The core sequence of CTCF binding sites [C(A/C)(C/T)CT] was similar but not identical and the remaining sequence of the consensus motif was also variable, a divergence already reported for CTCF binding sites by Filippova et al. (1996). Interestingly, promoter 1 has a TATA box at position -45, while promoter 2 is apparently a TATA-less promoter, a feature that may indicate higher transcriptional activity of promoter 1 (Hieb et al., 2014; Yang et al., 2014) and is in agreement with *DUSP4* promoter 1 having higher basal activity than promoter 2.

3.3. Regulation of *DUSP4* promoter activity by selected transcription factors

The ability of STAT3, FOXA1, CTCF and YY1 to bind to the sites predicted in *DUSP4* promoters (1 and 2) and stimulate/inhibit transcription was evaluated by co-transfections of promoter constructs with vectors expressing the transcription factors in study. In this set of experiments, relative luciferase activity is presented as fold changes, i.e. the ratio between the luciferase activity of the promoter constructs when co-transfected with TF expression vectors and when co-transfected with the empty expression vector. The (over)expression of the different transcription factors in the co-transfection experiments was confirmed by RT-qPCR (Supplementary Figure S2).

3.3.1. CTCF enhances *DUSP4* transcription through promoter 2

Upon co-transfection of the promoter constructs and the expression vector carrying the human *CTCF*, luciferase activity remained unchanged for promoter 1 constructs, indicating that the two binding sites predicted within this region are most likely not functional (Figure 3), while it was significantly increased for promoter 2 constructs. CTCF stimulated luciferase activity by 4.5-fold in F2.1 and by 3.1-fold in F2.2, where the region -1298/-1062 has been deleted, suggesting that the binding site predicted at position -1180/-1162 may be functional. The fact that the stimulation of luciferase activity by CTCF was not totally abolished in F2.2 indicated that additional binding sites are present in this fragment. Since it remained unchanged in F2.3 upon the elimination of the region -1061/-544, this suggested that the binding site predicted at position -163/-145 may be functional, but not the one at position -829/-811. CHIP-seq data retrieved from UCSC Genome Browser indicates that CTCF binds to DNA motifs within the region -1453/-1090 of the *DUSP4* promoter 2 (Figure 4), a region that comprises the binding site -1180/-1162 that we propose to be functional, thus supporting our finding. In the future, the functionality of the two binding sites in promoter 2 should be confirmed, e.g. by assessing luciferase activity of promoter constructs where the core of each binding site is mutated. CTCF is a transcription factor highly conserved throughout evolution and ubiquitously expressed in many cell types, with a high variety of functions (Filippova et al., 1996; Kim et al., 2015). It can function either as a repressor or an activator of transcription of genes associated with tumor development, in particular genes involved in growth, proliferation, differentiation, and apoptosis (Docquier et al., 2005; Torrano et al., 2005). We propose that the transcriptional activation of *DUSP4* promoter 2 and the stimulation of variants 2 and X1 expression by cancer-related transcription factor CTCF may be part of the mechanisms underlying the function of DUSP4 in cancer.

3.3.2. FOXA1 does not regulate *DUSP4* transcription

Upon co-transfection of the promoter constructs and the expression vector carrying the mouse *FOXA1*, luciferase activity remained unchanged for both promoters, which indicates that the binding sites predicted within these two regions are most likely not functional (Figure 5). FOXA1 is involved in several biological processes, such as cell proliferation, differentiation and apoptosis (Bernardo et al., 2013) and was recently shown to repress a subset of genes in luminal breast cancer cell lines (Bernardo et al., 2013). Several studies indicated that it can act either as an activator or a repressor, being abnormally expressed in several types of cancer (Bernardo & Keri, 2012). Our data indicates that FOXA1 does not regulate *DUSP4* transcription through the promoter regions analysed in this work. However, we cannot exclude a regulation of DUSP4 transcription by FOXA1 through binding site(s) located upstream the regions studied here or that mouse transcription factor did not bind efficiently the binding sites in human promoter, although cross-compatibility was reported for other human genes (Besnard et al., 2005)

3.3.3. STAT3 enhances *DUSP4* transcription through promoter 1

Upon co-transfection of the promoter constructs and the expression vector carrying the human *STAT3*, luciferase activity remained unchanged for promoter 2, indicating that the three binding sites predicted within this region are most likely not functional (Figure 6), while it was significantly increase for promoter 1 constructs. STAT3 enhanced luciferase activity by 1.6-fold in F1.1 and the elimination of the sequence -1454/-743 totally abolished STAT3 stimulation, indicating that one or several of the three binding sites in this sequence are functional and account for the entire regulatory effect of STAT3 on *DUSP4* expression driven by the promoter 1. It also suggests that the three binding sites predicted in the region -742/+248 are not functional or were wrongly predicted. To identify the binding site(s) accounting for the

regulatory activity of STAT3 in the region -1454/-743, additional deletion constructs should be prepared and/or the core of each binding site should be altered through site-directed mutagenesis. STAT3 is involved in several cellular functions and growing evidences suggest a critical role in the development of diverse cancer types (Corvinus et al., 2005). The aberrant activity of STAT3 observed in several cancers promoted carcinogenesis through the up-regulation of genes involved in cell cycle progression and the prevention/promotion of apoptosis (Xiong et al., 2014). Our results indicate that STAT3 positively regulates the transcription of *DUSP4* through the promoter driving the expression of variant 1 and may be implicated in the processes underlying the role of DUSP4 in cancer.

3.3.4. YY1 enhances *DUSP4* transcription through both promoter 1 and promoter 2

Upon co-transfection of the promoter constructs and the expression vector carrying the human *YY1*, luciferase activity was significantly increase for promoter 1 and 2 constructs (Figure 7). YY1 enhanced luciferase activity by 2-fold in F1.1 and the removal of the sequence -1454/-743 totally abolished YY1 stimulation indicating that one or several of the five binding sites predicted in this region are functional and account for the entire regulatory activity of YY1. It also suggests that the four binding sites predicted in the region -742/+248 are not functional or were wrongly predicted. YY1 also increased luciferase activity in F2.1 but to a much higher extent (11-fold) than in F1.1. The elimination of the sequence -1298/-1062 significantly reduced the stimulation of luciferase activity by YY1, suggesting the functionality of the binding site predicted at position -1265/-1255. Activation by YY1 was however not abolished in F2.2 (still 5-fold), indicating that additional functional binding sites are present in this fragment. The removal of the region -1061/-544 did not affect the luciferase activity indicating that binding sites predicted at position -1003/-998 and -738/-726 are not functional or were wrongly predicted. On the contrary, the increase in luciferase activity (5-fold) still present in

F2.3 indicated that the binding site predicted at position -293/-288 is likely functional. CHIP-seq data from UCSC Genome Browser indicates that YY1 binds to DNA motifs within the region -1862/-1181 in *DUSP4* promoter 1 (Figure 4) that comprises of the three binding sites potentially functional (-1218/-1207, -1322/-1311 and -1412/-1407; Figure 7) and within regions -1814/-1174 and -444/-60 in promoter 2 (Figure 4) that comprises the two binding sites potentially functional (-1265/-1255 and -293/-288; Figure 7). Because it resulted in the largest effect observed so far, the transcriptional activation of promoter 2 by YY1 was further studied through site-directed mutagenesis of the predicted binding sites at position -1265/-1255 and -293/-288. The removal of these binding sites significantly reduced the stimulation of luciferase activity by YY1, confirming their functionality and further supporting *DUSP4* transcriptional regulation by YY1 through promoter 2. The functionality of YY1 binding sites in promoter 1 should be evaluated following the same approach but the evaluation of additional promoter deletion constructs may also help to determine which of the binding sites in the region -1454/-743 are functional.

YY1 is a multifunctional and ubiquitously expressed transcription factor that can act either as a repressor or activator of gene expression, depending on the stimuli and interaction with other factors (Zhang et al., 2011). Several studies suggested that YY1 is overexpressed in several cancer types, being associated with the malignant processes where is known to activate cancer-related genes involved in proliferation, differentiation and apoptosis (Castellano et al., 2009). Our data indicate that YY1 is a positive regulator of *DUSP4* expression through binding in promoters 1 and 2, driving the transcription of variant 1 and variants 2 and X1, respectively. Although this remains to be proven, the over-expression of YY1 reported in several cancers may be at the origin of the up-regulation of *DUSP4* in cancer tissues. Interestingly, YY1 activated both promoters (although promoter 2 in a much higher extent) and therefore may have the ability to activate the transcription of all the three *DUSP4* transcript variants.

3. CONCLUSIONS

In the scope of this work, novel data was collected to improve our knowledge on the regulation of *DUSP4* expression by cancer-related transcriptional factors. Both promoters were found to be functional in the HEK-293 cells, although to different extents, and to be regulated by CTCF, STAT3 and YY1 in a promoter-specific manner, i.e. the two promoters that drive the expression of the different transcripts are regulated by different transcription factors to different extents. CTCF specifically activated the promoter 2 that drives the expression of variants 2 and X1, while STAT3 positively regulated the activity of promoter 1 that drives the transcription of the variant 1. YY1 enhanced the activity of both promoters, although to a larger extent for promoter 2. To support our *in vitro* data, CHIP-seq data from the UCSC Genome Browser showed a physical interaction between *DUSP4* promoter and YY1 and CTCF in regions carrying binding sites that we proposed to be functional. Current knowledge on the transcriptional regulation of human *DUSP4* provided by our work and previous studies (regarding regulation by p53, E2F-1 and HoxA10) has been summarized in Figure 8. The functionality of the TFBSs identified in this work should be further demonstrated by mutating the core sequence of each candidate binding site in the reporter constructs as it has been successfully done for YY1. To further confirm the physical interaction between the TFs and the putative sequence identified in *DUSP4* promoters, electrophoretic mobility shift assay (EMSA) and chromatin immunoprecipitation (ChIP) should also be performed. Data reported here provides new information about *DUSP4* transcriptional regulation by cancer-related TFs.

ACKNOWLEDGEMENTS

Authors acknowledge Débora Varela (BIOSKEL laboratory, CCMAR, Faro, Portugal) for her assistance during transfection experiments. Tatiana Varela is supported by the PhD fellowship SFRH/BD/144230/2019 from the FCT.

REFERENCES

Armes, J. E., Hammet, F., de Silva, M., Ciciulla, J., Ramus, S. J., Soo, W.-K., Mahoney, A., Yarovaya, N., Henderson, M. a, Gish, K., Hutchins, A.-M., Price, G. R., & Venter, D. J. (2004). Candidate tumor-suppressor genes on chromosome arm 8p in early-onset and high-grade breast cancers. *Oncogene*, *23*(33), 5697–5702. https://doi.org/10.1038/sj.onc.1207740

Bermudez, O., Pagès, G., & Gimond, C. (2010). The dual-specificity MAP kinase phosphatases: Critical roles in development and cancer. *American Journal of Physiology: Cell Physiology*, *299*, 189–202. https://doi.org/10.1152/ajpcell.00347.2009

Bernardo, G. M., Bebek, G., Ginther, C. L., Sizemore, S. T., Lozada, K. L., Miedler, J. D., Anderson, L. A., Godwin, A. K., Abdul-Karim, F. W., Slamon, D. J., & Keri, R. A. (2013). FOXA1 represses the molecular phenotype of basal breast cancer cells. *Oncogene*, *32*(5), 554–563. https://doi.org/10.1038/onc.2012.62

Bernardo, G. M., & Keri, R. A. (2012). FOXA1: A transcription factor with parallel functions in development and cancer. *Bioscience Reports*, *32*(2), 113–130. https://doi.org/10.1042/BSR20110046

Besnard, V., Wert, S. E., Kaestner, K. H., & Whitsett, J. A. (2005). Stage-specific regulation of respiratory epithelial cell differentiation by Foxa1. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, *289*, 750–759. https://doi.org/10.1152/ajplung.00151.2005

Castellano, G., Torrisi, E., Ligresti, G., Malaponte, G., Militello, L., Russo, A. E., McCubrey, J. A., Canevari, S., & Libra, M. (2009). The involvement of the transcription factor Yin Yang 1 in cancer development and progression. *Cell Cycle*, *8*(9), 1367–1372. https://doi.org/10.4161/cc.8.9.8314

Corvinus, F. M., Orth, C., Moriggl, R., Tsareva, S. A., Wagner, S., Pfitzner, E. B., Baus, D., Kaufman, R., Huber, L. A., Zatloukal, K., Beug, H., Öhlschläger, P., Schütz, A., Halbhuber, K.-J., & Friedrich, K. (2005). Persistent STAT3 activation in colon cancer is associated with enhanced cell proliferation and tumor growth. *Neoplasia*, *7*(6), 545–555. https://doi.org/10.1593/neo.04571

Docquier, F., Farrar, D., D'Arcy, V., Chernukhin, I., Robinson, A. F., Loukinov, D., Vatolin, S., Pack, S., Mackay, A., Harris, R. A., Dorricott, H., O'Hare, M. J., Lobanenkov, V., & Klenova, E. (2005). Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. *Cancer Research*, *65*(12), 5112–5122.

https://doi.org/10.1158/0008-5472.CAN-03-3498

Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J., & Lobanenkov, V. V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology*, *16*(6), 2802–2813. https://doi.org/10.1128/mcb.16.6.2802

Gröschl, B., Bettstetter, M., Giedl, C., Woenckhaus, M., Edmonston, T., Hofstädter, F., & Dietmaier, W. (2013). Expression of the MAP kinase phosphatase DUSP4 is associated with microsatellite instability in colorectal cancer (CRC) and causes increased cell proliferation. *International Journal of Cancer Expression*, *132*(7), 1537–1546. https://doi.org/10.1002/ijc.27834

Hieb, A. R., Gansen, A., Böhm, V., & Langowski, J. (2014). The conformational state of the nucleosome entry-exit site modulates TATA box-specific TBP binding. *Nucleic Acids Research*, *42*(12), 7561–7576. https://doi.org/10.1093/nar/gku423

Keyse, S. M. (2008). Dual-specificity MAP kinase phosphatases (MKPs) and cancer. *Cancer Metastasis Reviews*, *27*(2), 253–261. https://doi.org/10.1007/s10555-008-9123-1

Kim, S., Yu, N. K., & Kaang, B. K. (2015). CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental & Molecular Medicine*, *47*(6), e166. https://doi.org/10.1038/emm.2015.33

Kondoh, K., & Nishida, E. (2007). Regulation of MAP kinases by MAP kinase phosphatases. *Biochimica et Biophysica Acta-Molecular Cell Research*, *1773*(8), 1227–1237. https://doi.org/10.1016/j.bbamcr.2006.12.002

Liu, S., Sun, J.-P., Zhou, B., & Zhang, Z.-Y. (2006). Structural basis of docking interactions between ERK2 and MAP kinase phosphatase 3. *Proceedings of the National Academy of Sciences USA*, *103*(14), 5326–5331. https://doi.org/10.1073/pnas.0510506103

Owens, D. M., & Keyse, S. M. (2007). Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases. *Oncogene*, *26*(22), 3203–3213. https://doi.org/10.1038/sj.onc.1210412

Plotnik, J. P., Budka, J. A., Ferris, M. W., & Hollenhorst, P. C. (2014). ETS1 is a genome-wide effector of RAS/ERK signaling in epithelial cells. *Nucleic Acids Research*, *42*(19), 11928–11940. https://doi.org/10.1093/nar/gku929

Pulido, R., & Lang, R. (2019). Dual specificity phosphatases: From molecular mechanisms to biological function. *International Journal of Molecular Sciences*, *20*(18), 15–18. https://doi.org/10.3390/ijms20184372

Saigusa, S., Inoue, Y., Tanaka, K., Toiyama, Y., Okugawa, Y., Shimura, T., Hiro, J., Uchida, K., Mohri, Y., & Kusunoki, M. (2013). Decreased expression of DUSP4 is associated with liver and lung metastases in colorectal cancer. *Medical Oncology*, *30*(3), 620–627. https://doi.org/10.1007/s12032-013-0620-x

Shen, W. H., Wang, J., Wu, J., Zhurkin, V. B., & Yin, Y. (2006). Mitogen-activated protein kinase phosphatase 2: A novel transcription target of p53 in apoptosis. *Cancer Research*, *66*(12), 6033–6039. https://doi.org/10.1158/0008-5472.CAN-05-3878

Sieben, N. L. G., Oosting, J., Flanagan, A. M., Prat, J., Roemen, G. M. J. M., Kolkman-Uljee, S. M., Van Eijk, R., Cornelisse, C. J., Fleuren, G. J., & Van Engeland, M. (2005). Differential gene expression in ovarian tumors reveals Dusp 4 and Serpina 5 as key regulators for benign behavior of serous borderline tumors. *J. Clin. Oncol.*, *23*(29), 7257–7264. https://doi.org/10.1200/JCO.2005.02.2541

Sim, J., Yi, K., Kim, H., Ahn, H., Chung, Y., Rehman, A., Jang, S. M., Lee, K. H., Jang, K., & Paik, S. S. (2015). Immunohistochemical expression of dual-specificity protein phosphatase 4 in patients with colorectal adenocarcinoma. *Gastroenterology Research and Practice*, *2015*, 283764. https://doi.org/10.1155/2015/283764

Stelzer, Y., Bar, S., Bartok, O., Afik, S., Ronen, D., Kadener, S., & Benvenisty, N. (2015). Differentiation of human parthenogenetic pluripotent stem cells reveals multiple tissue- and isoform-specific imprinted transcripts. *Cell Reports*, *11*(2), 308–320. https://doi.org/10.1016/j.celrep.2015.03.023

Torrano, V., Chernukhin, I., Docquier, F., D'Arcy, V., León, J., Klenova, E., & Delgado, M. D. (2005). CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *Journal of Biological Chemistry*, *280*(30), 28152–28161. https://doi.org/10.1074/jbc.M501481200

Varela, T., Laizé, V., Conceição, N., Caldeira, P., Marreiros, A., Guerreiro, H., & Cancela, M. L. (2020). Expression of DUSP4 transcript variants as a potential biomarker for colorectal cancer. *Biomarkers in Medicine*, *14*(8), 639–650. https://doi.org/10.2217/bmm-2019-0369

Waha, A., Felsberg, J., Hartmann, W., Von Dem Knesebeck, A., Mikeska, T., Joos, S., Wolter, M., Koch, A., Yan, P. S., Endl, E., Wiestler, O. D., Reifenberger, G., Pietsch, T., & Waha, A. (2010). Epigenetic downregulation of mitogen-activated protein kinase phosphatase MKP-2 relieves its growth suppressive activity in glioma cells. *Cancer Res.*, *70*(4), 1689–1699. https://doi.org/10.1158/0008-5472.CAN-09-3218

Wang, H., Lu, Y., Huang, W., Papoutsakis, E. T., Fuhrken, P., & Eklund, E. A. (2007).

HoxA10 activates transcription of the gene encoding mitogen-activated protein kinase phosphatase 2 (Mkp2) in myeloid cells. *Journal of Biological Chemistry*, *282*(22), 16164–16176. https://doi.org/10.1074/jbc.M610556200

Wang, H. Y., Cheng, Z., & Malbon, C. C. (2003). Overexpression of mitogen-activated protein kinase phosphatases MKP1, MKP2 in human breast cancer. *Cancer Lett.*, *191*(2), 229–237. https://doi.org/10.1016/S0304-3835(02)00612-2

Wang, J., Shen, W. H., Jin, Y. J., Brandt-Rauf, P. W., & Yin, Y. (2007). A molecular link between E2F-1 and the MAPK cascade. *Journal of Biological Chemistry*, *282*(25), 18521–18531. https://doi.org/10.1074/jbc.M610538200

Xiong, A., Yang, Z., Shen, Y., Zhou, J., & Shen, Q. (2014). Transcription factor STAT3 as a novel molecular target for cancer prevention. *Cancers*, *6*(2), 926–957. https://doi.org/10.3390/cancers6020926

Yang, L., Wang, J., Lv, Y., Hao, D., Zuo, Y., Li, X., & Jiang, W. (2014). Characterization of TATA-containing genes and TATA-less genes in S. cerevisiae by network topologies and biological properties. *Genomics*, *104*(6), 562–571. https://doi.org/10.1016/j.ygeno.2014.10.005

Yant, S. R., Zhu, W., Millinoff, D., Slightom, J. L., Goodman, M., & Gumucio, D. L. (1995). High affinity YY1 binding motifs: Identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human β globin cluster. *Nucleic Acids Research*, *23*(21), 4353–4362. https://doi.org/10.1093/nar/23.21.4353

Yip-Schneider, M. T., Lin, A., & Marshall, M. S. (2001). Pancreatic tumor cells with mutant K-ras suppress ERK activity by MEK-dependent induction of MAP kinase phosphatase-2. *Biochemical and Biophysical Research Communications*, *280*(4), 992–997. https://doi.org/10.1006/bbrc.2001.4243

Zhang, Q., Stovall, D. B., Inoue, K., & Sui, G. (2011). The oncogenic role of Yin Yang 1. *Critical Reviews in Oncogenesis*, *16*(3–4), 163–197. https://doi.org/10.1615/CritRevOncog.v16.i3-4.30

Table 1. List of the primers used to clone and sub-clone human *DUSP4* gene promoters (Fw, forward primer, Rev, reverse primer). Underlined sequences indicate restriction site for endonucleases cited in primer name.

Name	Sequence (5′ to 3′)
Promoter cloning primers	
HsDUSP4_F1.1_Fw	AGAAATATCTTGATGCAATGCCTCTT
HsDUSP4_F1.2_Fw	CTAATGCATACGCTAACGTCTCAG
HsDUSP4_F1.3_Fw	CCCTGTTTACTCCGCTCTTTGTG
HsDUSP4_F1_Rev	GAGAGCCTCTTCTTCCCTGTCC
HsDUSP4_F2.1_Fw	GCTACATCCTAGGTTCGGTCAA
HsDUSP4_F2.2_Fw	CCGACATCTGCCTGCTCAAAGG
HsDUSP4_F2.3_Fw	CAAATGCAGAAAGGGAGATTGG
HsDUSP4_F2_Rev	CTTAGCAGTTCAACCAAAGGTCAA
Promoter sub-cloning primers	
HsDUSP4_Promoter_F2.1_BglII	GGAGATCTGCTACATCCTAGGTTCGGTCAA
HsDUSP4_Promoter_F2.2_BglII	GGAGATCTCCGACATCTGCCTGCTCAAAGG
HsDUSP4_F2_Rev_HindIII	CTAAAGCTTCTTAGCAGTTCAACCAAAGGTCAA
Site-directed mutagenesis primers	
HsDUSP4_F2.1_mut-1267/-1251_Fw	CCCGGATTTTCGACTATCTGCGTGCGGCGCGTGCCGG
HsDUSP4_F2.1_mut-1267/-1251_Rev	CCGGCACGCGCCGCACGCAGATAGTCGAAAATCCGGG
HsDUSP4_F2.3_mut-293/-288_Fw	GTCAACGTGCGCTGTAACAAAATCGTGCGGCGGC
HsDUSP4_F2.3_mut-293/-288_Rev	GCCGCCGCACGATTTTGTTACAGCGCACGTTGAC

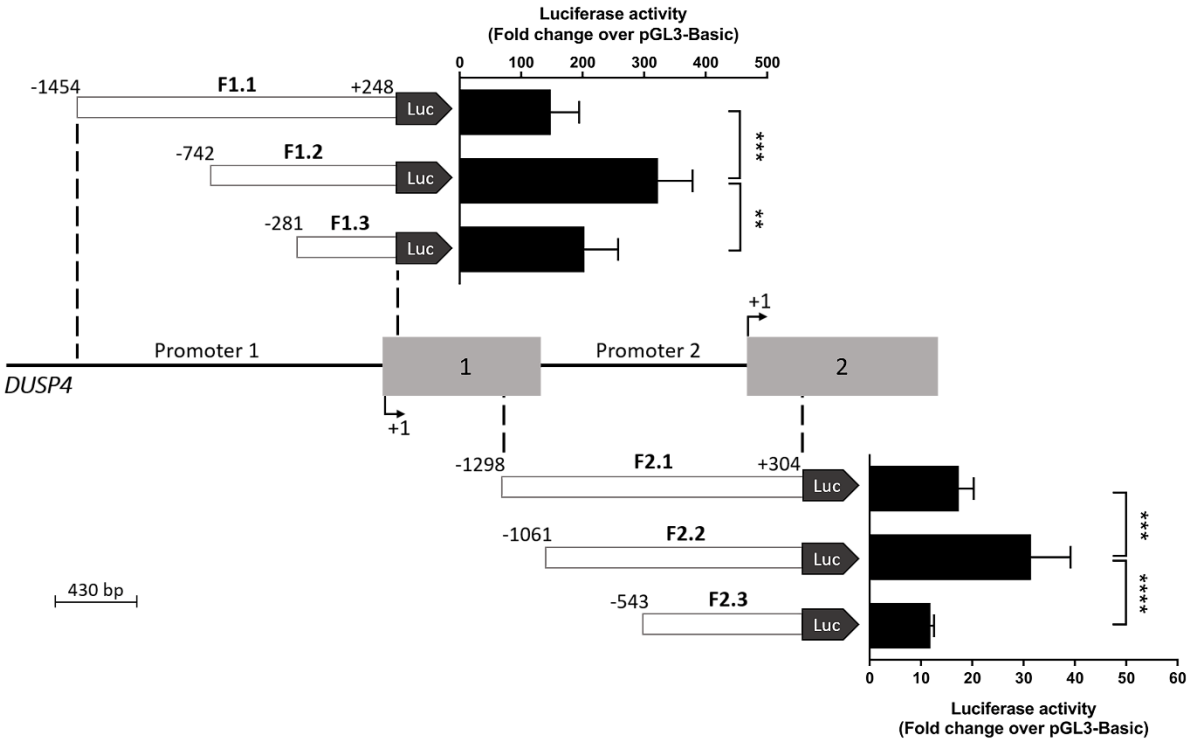

Figure 1. Relative luciferase activity of *DUSP4* promoter constructs. Exons are represented by grey boxes. Promoter fragments inserted upstream of the luciferase (Luc) gene are represented as white boxes and start/end positions are indicated according to respective transcription initiation (+1). Data are presented as the mean of luciferase activity ± SD. **, *** and **** indicate $p<0.01$, $p<0.001$ and $p<0.0001$, respectively (one-way ANOVA followed by Tukey's post-test).

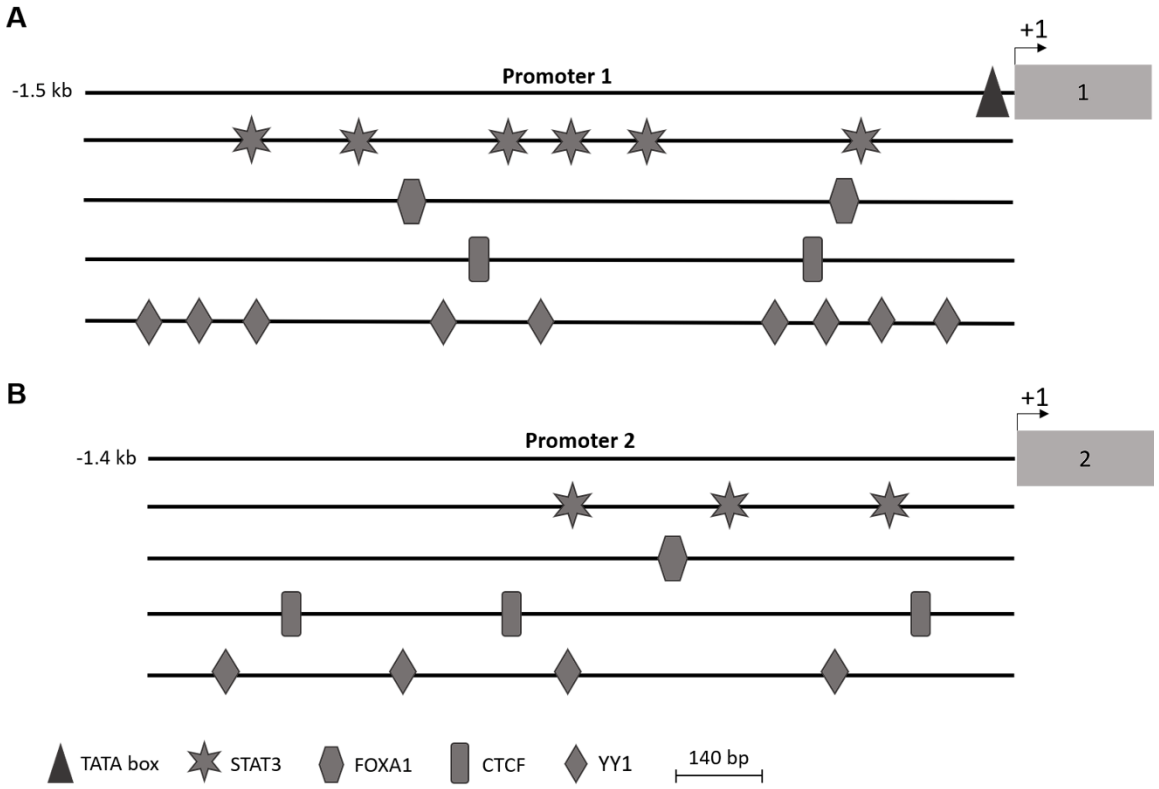

Figure 2. Transcription factor binding sites predicted in *DUSP4* promoter region 1 (A) and 2 (B). Positions -1.5 kb and -1.4 kb are relative to transcription initiation sites (+1) in exon 1 and exon 2, respectively. Black triangle indicates the TATA box in promoter 1. Grey geometric shapes represent transcription factor binding for STAT3 (stars), FOXA1 (hexagons), CTCF (rectangles) and YY1 (diamonds). Exons are not in scale.

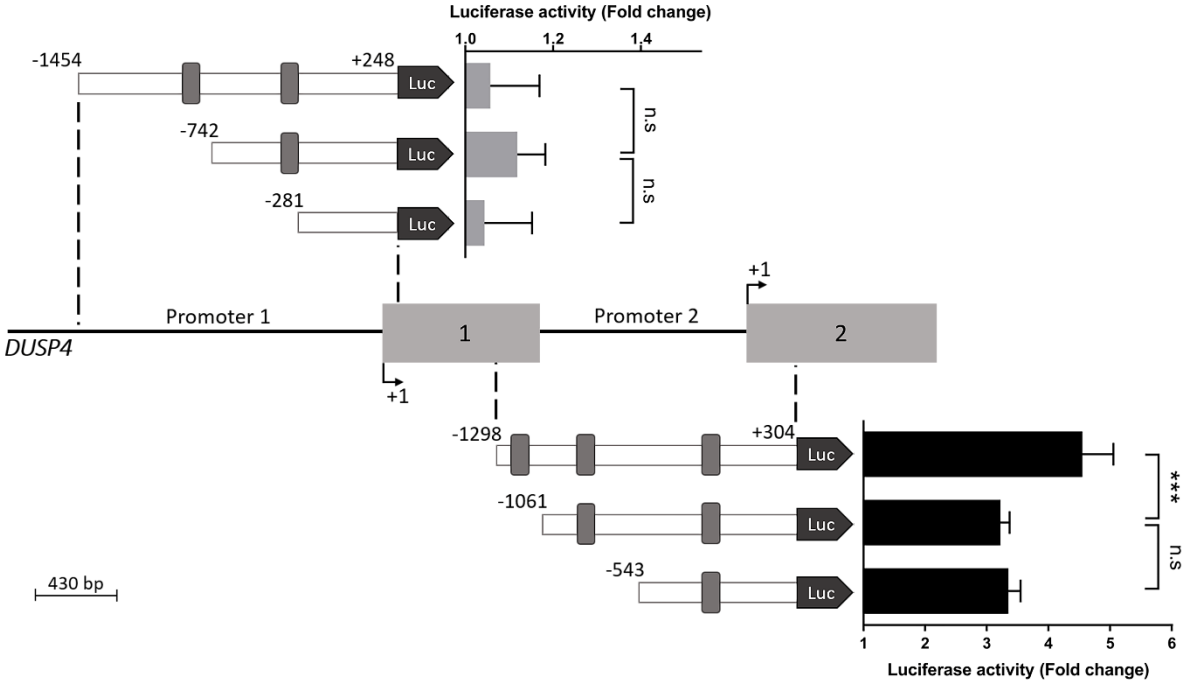

Figure 3. *DUSP4* transcriptional regulation by CTCF. Exons are represented by grey boxes. Promoter fragments inserted upstream of the luciferase (Luc) gene are represented as white boxes and start/end positions are indicated according to respective transcription initiation (+1). Solid rectangles represent putative CTCF binding sites. Black and grey bars represent luciferase values that are significantly different ($p<0.05$) and not significantly different ($p>0.05$) from the activity driven by the empty expression vector, respectively. Data are presented as the mean of luciferase activity ± SD. *** indicates $p<0.001$. n.s indicates no statistical difference (one-way ANOVA followed by Tukey's post-test).

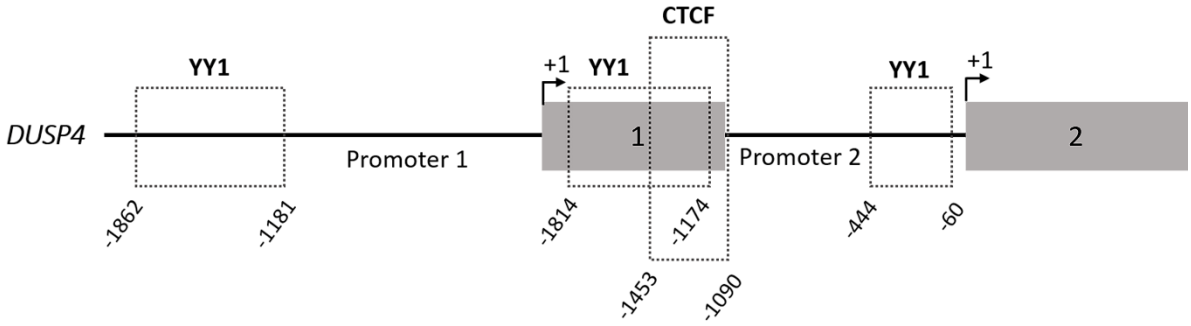

Figure 4. Schematic representation of human *DUSP4* CHIP-seq data retrieved from the UCSC Genome Browser. Exons 1 and 2 are represented by grey boxes. Dotted rectangles indicate the promoter regions bound by transcription factors CTCF and YY1, and the numbers below represent the start/end positions of these regions according to respective transcription initiation sites (+1).

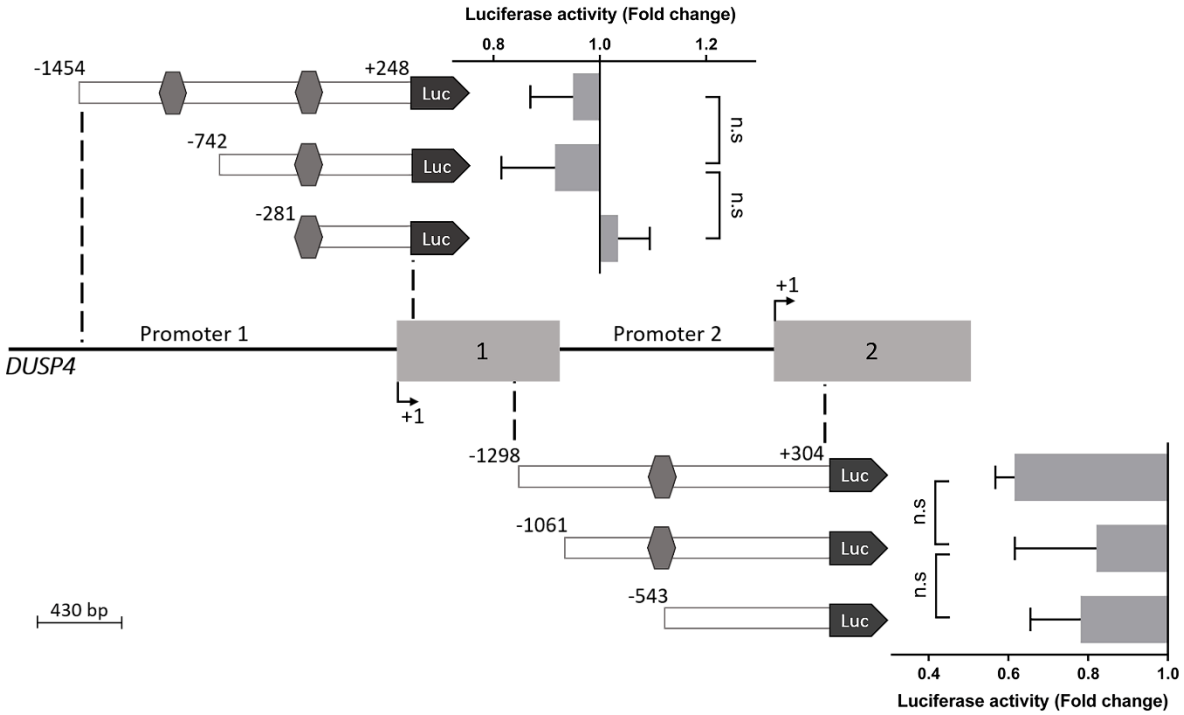

Figure 5. *DUSP4* transcriptional regulation by FOXA1. Exons are represented by grey boxes. Promoter fragments inserted upstream of the luciferase (Luc) gene are represented as white boxes and start/end positions are indicated according to respective transcription initiation (+1). Solid hexagons represent putative FOXA1 binding sites. Grey bars represent luciferase values that are not significantly different (p>0.05) from the activity driven by the empty expression vector. Data are presented as the mean of luciferase activity ± SD. n.s indicates no statistical difference (one-way ANOVA followed by Tukey's post-test).

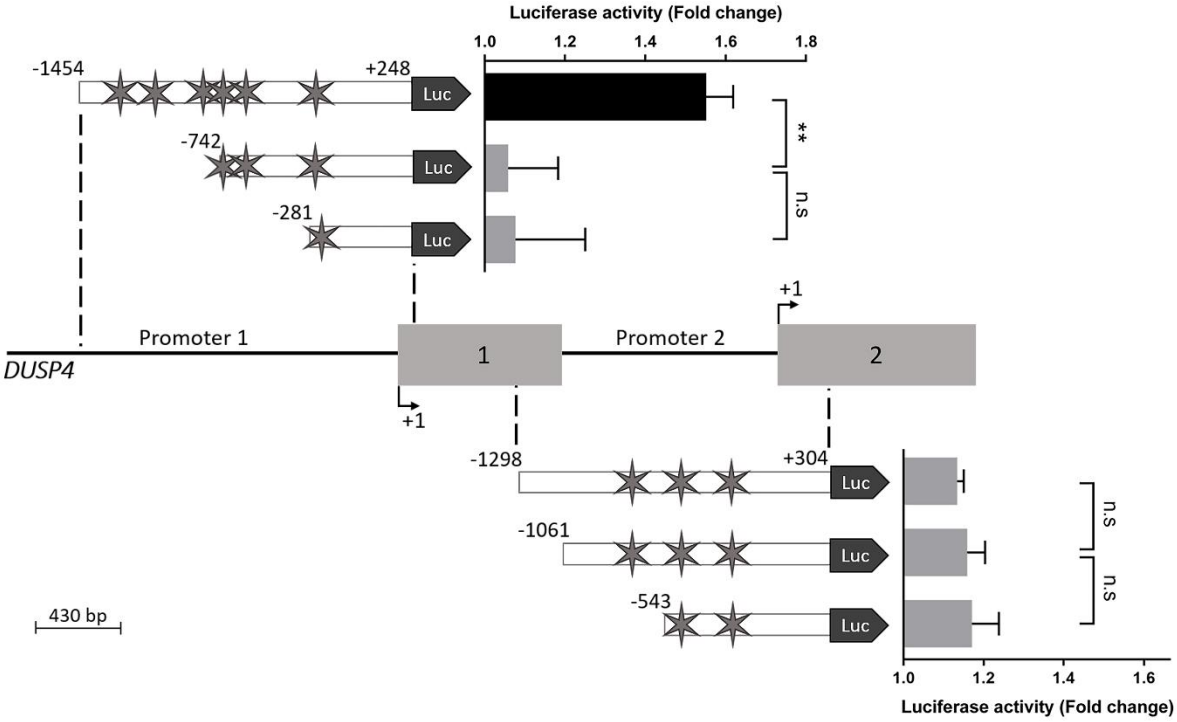

Figure 6. *DUSP4* transcriptional regulation by STAT3. Exons are represented by grey boxes. Promoter fragments inserted upstream of the luciferase (Luc) gene are represented as white boxes and start/end positions are indicated according to respective transcription initiation (+1). Solid stars represent putative STAT3 binding sites. Black and grey bars represent luciferase values that are significantly different ($p<0.05$) and not significantly different ($p>0.05$) from the activity driven by the empty expression vector, respectively. Data are presented as mean of luciferase activity ± SD. ** indicates $p<0.01$. n.s indicates no statistical difference (one-way ANOVA followed by Tukey's post-test).

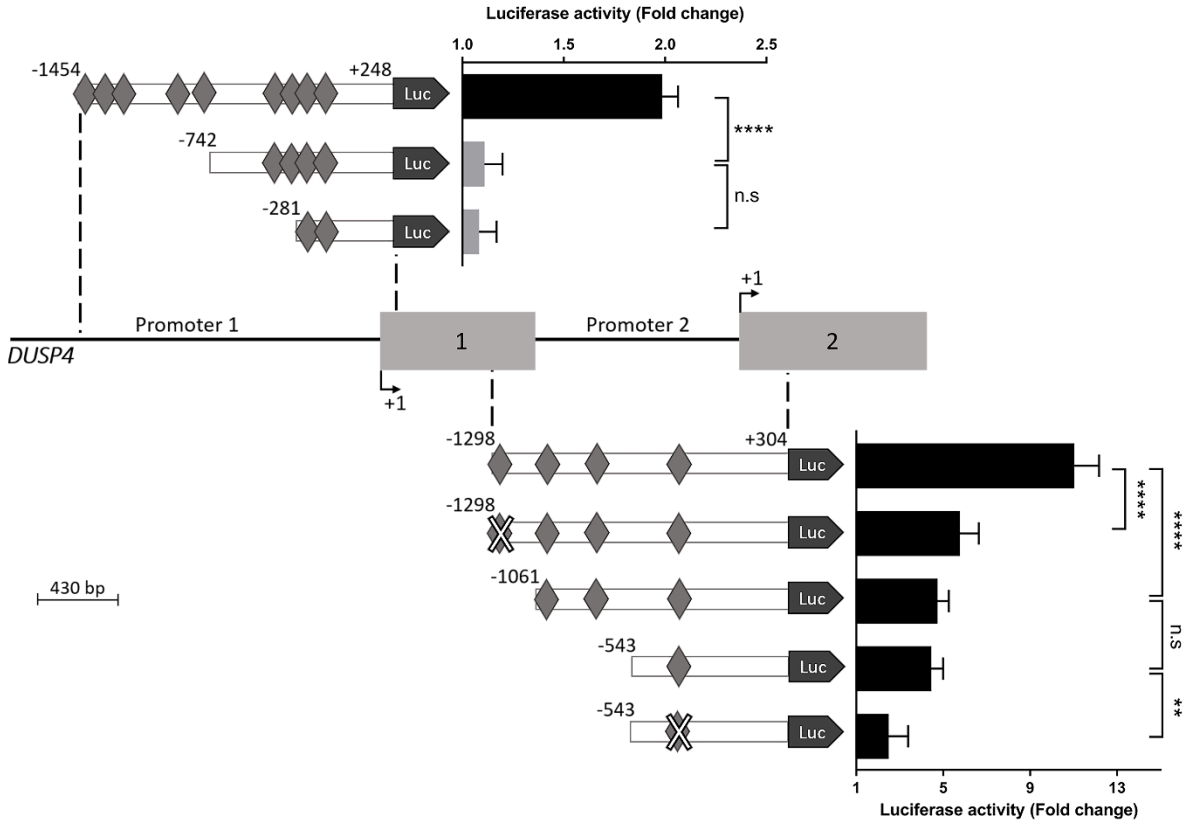

Figure 7. *DUSP4* transcriptional regulation by YY1. Exons are represented by grey boxes. Promoter fragments inserted upstream of the luciferase (Luc) gene are represented as white boxes and start/end positions are indicated according to respective transcription initiation (+1). Solid and crossed diamonds represent non-mutated and mutated putative YY1 binding sites, respectively. Black and grey bars represent luciferase values that are significantly different (p<0.05) and not significantly different (p>0.05) from the activity driven by the empty expression vector, respectively. Data are presented as the mean of luciferase activity ± SD. **** indicates p<0.0001. n.s indicates no statistical difference (one-way ANOVA followed by Tukey's post-test).

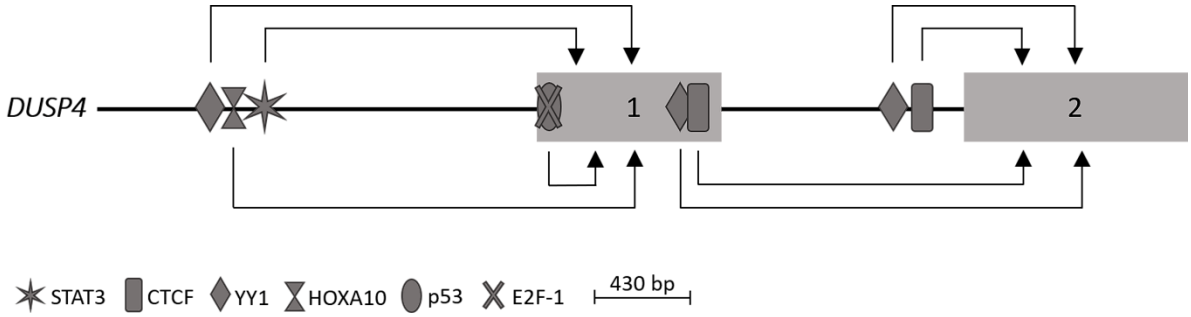

Figure 8. Schematic representation of human *DUSP4* transcriptional data. Transcription factor binding sites identified in this study (CTCF, STAT3 and YY1) and in previous studies (P53, E2F1 and HOXA10) are indicated with different shapes. Exons and promoter regions are represented by grey boxes and black lines, respectively. Arrows indicate promoter-specific activation of gene transcription. A single binding site is represented in promoter 1 for YY1 and STAT3.