Abstract

PTOV1 is a transcription and translation regulator and a promoter of cancer progression. Its overexpression in prostate cancer induces transcription of drug resistance and self-renewal genes, and docetaxel resistance. Here we studied PTOV1 ability to directly activate the transcription of *ALDH1A1* and *CCGN2* by binding to specific promoter sequences. Chromatin immunoprecipitation and electrophoretic mobility shift assays identified a DNA-binding motif inside the PTOV-A domain with similarities to known AT-hooks that specifically interacts with *ALDH1A1* and *CCGN2* promoters. Mutation of this AT-hook-*like* sequence significantly decreased the expression of *ALDH1A1* and *CCGN2* promoted by PTOV1. Immunohistochemistry revealed the association of PTOV1 with mitotic chromosomes in high grade prostate, colon, bladder, and breast carcinomas. Overexpression of *PTOV1*, *ALDH1A1*, and *CCNG2* significantly correlated with poor prognosis in prostate carcinomas and with shorter relapse-free survival in colon carcinoma. The previously described interaction with translation complexes and its direct binding to *ALDH1A1* and *CCNG2* promoters found here reveal the PTOV1 capacity to modulate the expression of critical genes at multiple levels in aggressive cancers. Remarkably, the AT-hook motifs in PTOV1 open possibilities for selective targeting its nuclear and/or cytoplasmic activities.

A novel DNA-binding motif in Prostate Tumor Overexpressed-1 (PTOV1) required for the expression of *ALDH1A1* and *CCNG2* in cancer cells

Valentina Maggio[a], Verónica Cánovas[a], Alex J. Félix[b], Valentí Gómez[a], Inés de Torres[c], María Eugenia Semidey[c], Juan Morote[a,d], Verónique Noé[b], Carlos J. Ciudad[b] and Rosanna Paciucci[a*].

[a] Biochemistry Service and Biomedical Research Group of Urology, Vall d'Hebron University Hospital and Universitat Autònoma de Barcelona, Pg. Vall d'Hebron 119-129, Barcelona 08035, Spain.

[b] Department of Biochemistry and Physiology, School of Pharmacy, and Institute of Nanoscience and Nanotechnology, University of Barcelona, Avinguda Joan XXIII 27, 08028 Barcelona, Spain.

[c] Department of Pathology, Vall d'Hebron University Hospital, Barcelona, Spain.

[d] Deparment of Urology, Vall d'Hebron University Hospital, and Universitat Autònoma de Barcelona.

*** Corresponding author.** Rosanna Paciucci, Ph.D. Biochemistry Service and Biomedical Research Group of Urology, Laboratory 211, Collserola building, Vall d'Hebron Institute of Research, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain. Phone: +34 93 4894063, E-mail: rosanna.paciucci@vhir.org

Highlights

- A novel DNA-binding motif in PTOV1 is required for *ALDH1A1* and *CCNG2* overexpression.
- PTOV1 is a direct transcriptional activator of gene expression in prostate cancer.
- Distinct DNA- RNA-binding motifs might direct PTOV1 nuclear and cytoplasmic actions.
- *PTOV1, ALDH1A1, CCNG2* expression associated with prostate and colon cancer bad prognosis.

Keywords:

PTOV1, DNA-binding motif, AT-hook, aggressive prostate cancer, chromatin immunoprecipitation, EMSA

Abbreviations

PTOV1: Prostate Tumor Overexpressed-1; PCa: Prostate Cancer; CRPC: Castration Resistant Prostate Cancer; ADT: Androgen Deprivation Therapy; ALDH1A1: Aldehyde Dehydrogenase 1 family member A1; CCNG2: Cyclin G2; DKK1: Dickkopf WNT signaling pathway inhibitor 1; HES1: Hes family bHLH transcription factor 1; HMGA1: High Mobility Group Protein A1; RACK1: Receptor of Activated protein C Kinase 1.

Funding

This work was supported by the Ministry of Economy MINECO [SAF2014-59958-R to R.P. and SAF2014-51825-R to V.N. and C.C.]; the AGAUR [2014 SGR733]; and Industex S.L.

Acknowledgements

Authors wish to thank Daniele Quintavalle for his help in chromatin immunoprecipitation experiments, the Vall d'Hebron Research Institute, and the University of Barcelona for support.

1. Introduction

Prostate cancer (PCa) is the most frequent neoplasia in men and the third leading cause of death for cancer in western Countries[1]. The majority of patients are cured by radical prostatectomy but late-diagnosed or aggressive cancers have decreased survival rates of 29%[2]. The mainstay therapy for PCa is androgens deprivation (ADT) that, although initially very effective, leads the majority of patients to develop a Castration Resistant Prostate Cancer (CRPC) with poor prognosis[3,4]. Recently, docetaxel, a taxane that arrests mitotic division, has been introduced in combination with ADT as first line therapy in metastatic hormone-naïve PCa with improved results[5]. However, the development of resistance remains the most critical unsolved problem in prostate cancer[2,6].

Our previous findings described the action of the oncogenic protein PTOV1 in the progression of PCa to a metastatic disease and docetaxel resistance[7,8]. Suggestive of a role for PTOV1 in the acquisition of a more aggressive phenotype, we and others showed that the increased expression of this protein significantly associated with a higher grade of malignancy in PCa and numerous other cancer types[7,9-14]. The structure of the protein consists mainly of two highly homologous sequences arranged in tandem identified as PTOV-A and -B domains[15]. PTOV1 is able to shuttle between nucleus and cytoplasm, its overexpression induces cell proliferation, tumor growth, and increases the motility of PCa and breast cancer cells *in vitro*, and metastasis *in vivo*[16,17]. In the cytoplasm, PTOV1 was shown to interact with the receptor of activated protein C kinase 1 (RACK1), and positively regulated protein synthesis, in particular c-Jun translation and activity[16]. In the nucleus, PTOV1 was shown to decrease Notch signaling in metastatic prostate tumors by repressing the transcription of Notch target genes *HES1* and *HEY1*, an action that was associated to active histone deacetylases[18]. In

breast cancer cells, PTOV1 was found to repress the expression of Dickkopf 1 (DKK1), a major secreted Wnt signaling antagonist that resulted in the activation of downstream Wnt/β catenin signaling and cancer progression[17].

Recently, PTOV1 was described able to directly bind nucleic acids through an amino acid sequence with similarities to known AT-hook motifs[19]. The first AT-hook motif, discovered in the high mobility group proteins A1 (HMGA1), confers these proteins a global role as master regulator of chromatin structure in addition to physically interact with a large variety of different transcription factors[20]. Remarkably, HMGA1 proteins are over-expressed in virtually every type of cancer, where their expression levels correlate with tumor malignancy and poor outcome[21]. An AT-hook consensus motif was defined as composed by a core of arginines (R) and prolines (P) that allows the amino acid stretch to bind the minor groove of A-T rich sequences of RNA and DNA[22]. The first 43 amino acids at the N-terminal of PTOV1 contain a new type of AT-hook motif defined as *extended AT-hook* (eAT-hook) that is 10-15 amino acids longer and differs from the canonical motif[19].

The progression of CRPC PC3 and Du145 cells to a docetaxel resistant phenotype promoted by the overexpression of PTOV1 is associated to increased expression of genes involved in resistance to docetaxel and self-renewal (*e.g. ABCB1, CCNG2* and *ALDH1A1*)[8], suggesting that through the expression of these genes PTOV1 conferred cells a higher resistance to chemotherapy and higher plasticity[8,23-26].

Here, we studied whether PTOV1 may directly bind and activate the promoter regions of these genes in LNCaP androgen-dependent prostate cancer cells. We report the identification and localization of a novel DNA-binding motif in PTOV1 that allows the protein to directly and specifically bind to *ALDH1A1* and *CCNG2* promoters and it is required for their full activation.

2. Materials and methods

2.1. Cell cultures

LNCaP androgen dependent prostate cancer cells were obtained from the American Type Culture Collection (Rockville, MD). Cells were periodically confirmed free of contaminated cell lines by in house authentication of cell cultures by STR-fingerprints comparing these with those published by ATCC. LNCaP cells were maintained in RPMI 1640 medium (Life Technologies, Inc., Grand Island, NY) supplemented with 10% heat-inactivated fetal bovine serum (Life Technologies) at 37°C in an atmosphere of 5% CO2. Antibodies to PTOV1 antibody were produced and purified as previously described[15,27]. Wnt3a conditioned medium was kindly provided by D. Arango (Vall d'Hebron Institute of Research, Barcelona).

2.2. Plasmid and peptides

The lentiviral HAPTOV1-ires-GFP and GFP-PTOV1 vectors were previously described[18]. Short-hairpin shRNA sequences 1397 and 1439 (Sigma-Aldrich, St. Louis, MO) targeting the human PTOV1 mRNA were as described[8] and are shown in Supplementary Table 3. PTOV1 mutant at amino acids 98-100 (the sequence KRRP was changed to EGGP) was obtained with the plasmid GFP-PTOV1 using the QuikChange™ Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. The primers used for mutagenesis were: 5'-GAGTGGCAGGAGGAGGGCGGACCCTAGTCTGAC-3' (Forward) and 5'-GTCAGAGTAGGGTCCGCCCTCCTCCTGCCACTC-3' (Reverse). The eAT-hook-wild type and eAT-hook-mutant peptides were purchased from PepMic (Suzhou, China). The A-domain peptide was described previously[9].

2.3. Real Time qPCR

Total RNA was extracted with the RNeasy mini kit (QIAGEN, Hilden, Germany), reverse transcribed with the Mooney Murine Leukemia Virus Reverse Transcription (M-MLV-RT) kit (Promega, Madison, USA) and real-time qPCR performed with the Universal Probe Library (Roche, Basilea, Switzerland) on a LightCycler 480 RealTime PCR instrument (Roche). Primers are shown in Supplementary Table 1. The ΔΔCt method was applied to estimate relative transcript levels. TBP, IPO8, or HMBS were used as endogenous control genes. Values are presented as mean + SD.

2.4. Chromatin immunoprecipitation (ChIP)

Chromatin was immunoprecipitated using EZ-chip Chromatin Immuno Precipitation kit (Millipore, Burlington, USA) according to the manufacturer, as previously described[16]. Briefly, after a mild formaldehyde crosslinking step, cells were sonicated, lysates incubated with primary antibodies and precipitated with protein A/G-Sepharose. Crosslinking of DNA-protein complexes was reversed, DNA purified and used as a template for PCR reactions. Primers used for PCR in ChIP experiments are described in Table 3.

2.5. Electrophoretic mobility shift assay (EMSA)

For binding assays, we used the following double-stranded DNAs corresponding to regions of: *ALDH1A1* promoter (from -395- to -363), *CCNG2* promoter (from -269 - to -239) and *HES1* gene (from +45 to +69). The dsDNA probes were formed by mixing 20 µg of each single-stranded oligodeoxynucleotide in a 150 mM NaCl solution. After incubation at 90ºC for 5 min, solutions were allowed to cool down slowly to room temperature. The duplexes were purified in a nondenaturing 20% polyacrylamide gel

electrophoresis and DNA concentration was determined by measuring its absorbance (260 nm) at 25ºC in a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Barcelona, Spain). Duplexes were 5'-end-labeled with [γ-^{32}P]-ATP (Perkin Elmer, Madrid, Spain) by T4 polynucleotide kinase (New England BioLabs) in a 10 µL reaction mixture, according to the manufacturer's protocol. After incubation at 37 ºC for 1h, 90 µL of Tris-EDTA buffer (1 mM EDTA and 10 mM Tris, pH 8.0; Sigma-Aldrich) were added to the reaction mixture, which was subsequently filtered through a Sephadex G-25 (Sigma-Aldrich) spin-column to eliminate the unincorporated [γ-^{32}P]-ATP. Radio-labelled probes (100.000 cpm, [γ-^{32}P]-ATP) were placed in ice for 30 minutes with 10 µg of PTOV1 domains or 0.3 µg of peptides using either 5 µg or 0.15 µg poly(dI:dC), respectively, as unspecific competitor, in the presence of the binding buffer (5% Glycerol, 0.5 mM DTT, 4 mM $MgCl_2$, 36 mM KCl, 0.5 mM EDTA, 25 mM Tris-HCl pH 8.0; all reagents were purchased from Sigma-Aldrich). The products of the binding reactions were electrophoretically resolved in 5% polyacrylamide and 5% glycerol native gels at a fixed voltage of 220V and 4ºC. Gels were dried at 80ºC and visualized on a Storm 840 PhosphorImager (Molecular Dynamics, GE Healthcare Life Sciences, Barcelona, Spain). ImageQuant software v5.2 was used to visualize the results.

2.6. In vivo growth of prostate cancer cells

LNCaP cells (1x10^6) in 100 µl mix of PBS and Matrigel (mix 1:1), were inoculated subcutaneously in the right flank of 6-week-old male mice (Rj:NMRI-*Foxn1*$^{nu/nu}$, n=6) (Charles River Laboratories, Barcelona, Spain). All animal experimental procedures were approved by the Vall d'Hebron Hospital Animal Experimentation Ethics Committee. After tumors reached 1.5 cm in diameter, mice were euthanized and tumors

excised. Tumors were fixed in formol, embedded in paraffin and processed for histopathology. Hematoxylin & eosin staining to verify the histopathological findings were performed.

2.7. Immunohistochemistry

Four-micron sections from human and mice tumors were used for antigen retrieval in citrate buffer, as described[18]. After blocking endogenous peroxidase activities and non-specific labeling, sections were labeled with PTOV1 antibody for 2 hours at room temperature. As a negative control, primary antibody was omitted or replaced with non-specific rabbit antibody, obtaining clean negative results in all cases (data not shown). Reaction was revealed by the avidin–biotin complex and staining with diaminobenzidine (ABC Elite kit, Burlingame, CA, USA). Nuclei we counterstained with Hematoxylin.

2.8. DNA analyses in human prostate cancer samples

The mRNA expression of *PTOV1*, *ALDH1A1* and *CCNG2* genes was analyzed in dataset from human prostate tumors (GSE97284 n=188) using the R2 bioinformatics platform (http://r2.amc.nl). The most informative probeset, according to its average present signal (APS) and average value (Avg) was selected. The presence of DNA amplification, point mutations and deletions in the genes of interest were analyzed from publicly available datasets of human prostate tumors using the cBioPortal platform[28,29]. Correlations between *PTOV1*, *ALDH1A1*, and *CCNG2* genes and survival of patients with colon carcinoma were obtained using the dataset GSE14333.

2.9. Western blotting

Western blotting analyses were performed as previously described[18]. Specific reactivity to antibodies was detected with a chemiluminescent substrate (GE Healthcare). Densitometric analysis of western blot signals with antibody to PTOV1 was performed using ImageJ software.

2.10. Statistics

Results are expressed as means $\pm$ standard deviation of the means. For statistical analysis, according to whether data were sampled from a Gaussian distribution or not, unpaired t test or Mann-Whitney U test was used to compare two groups. A p value $\leq$ 0.05 was taken as the level of significance. These analyses were performed using GraphPad Prism 5 software.

3. Results

3.1. PTOV1 induces the expression of self-renewal genes in prostate cancer cells

Our prior studies identified *PTOV1, ALDH1A1* and *CCNG2* genes as potential significant markers of metastasis and poor prognosis when overexpressed in primary androgen dependent prostate tumors[8]. ALDH1A1 is a hallmark of both normal and cancer stem cells (CSCs) and an enabler of drug resistance in different types of cancers[14,30], while CCNG2 is a promoter of G2/M cell cycle arrest that contributes to thiopurine and doxorubicin resistance[31,32]. In metastatic CRPC Du145 and PC3 cells, the overexpression of PTOV1 provoked an increase in the expression of these genes (Figure 1A)[8]. Additionally, in androgen dependent LNCaP PCa cells PTOV1 also significantly induced the expression of *ALDH1A1* and *CCNG2* (Figure 1B and C). On the other hand, the knockdown of *PTOV1* by short hairpin RNAs provoked a significant reduction of endogenous *PTOV1* levels and a parallel significant decrease in *ALDH1A1* and *CCNG2* mRNA expression with respect to a control shRNA (Figure 1B and C).

3.2. PTOV1 directly induces the expression of ALDH1A1 and CCNG2 in PCa cells

We next interrogated the mechanisms by which PTOV1 drives the transcription of *ALDH1A1* and *CCNG2*. Prior work had shown that the engagement of two signaling networks, Wnt/β-catenin and Jun kinase (JNK), activate PTOV1-mediated functions[16,17]. Upon inhibition of these pathways by iCRT14 or JNK inhibitor II, specific inhibitors of Wnt and JNK signaling, respectively, the expression of *ABCB1*, driven by PTOV1, was abrogated (Supplementary Figure S1). In contrast, the PTOV1-dependent expression of *ALDH1A1* and *CCNG2* was not affected, suggesting that PTOV1 induced their transcription independently from these pathways (Supplementary Figure S1C)[16,17]. Based on these observations, we hypothesized that PTOV1 mediates the transcription of *ALDH1A1* and *CCNG2* genes by direct association with their promoters. We thus performed chromatin immunoprecipitation (ChIP) in LNCaP cells, and observed a specific binding of PTOV1 to the chromatin of the *ALDH1A1* and *CCNG2* promoters (Figure 2). In contrast, the binding of PTOV1 to the *ABCB1* promoter resulted in unclear or non specific, suggesting that *ABCB1* transcriptional activation occurs through other circuits promoted by PTOV1, as suggested above (Supplementary Figure S1). As expected, no binding of PTOV1 was observed to internal sequences of the *HES1* gene [18] .

3.3. PTOV1 directly binds to ALDH1A1 and CCNG2 promoters through a new motif in its A domain.

AT-hook amino acid motifs confer proteins the ability to bind DNA or RNA. The N-terminal region of PTOV1 contained an extended AT-hook (eAT-hook) motif[19]. We tested different regions of PTOV1, including the eAT-hook motif, the A domain, and

the B domain for their ability to bind the promoter sequences of *ALDH1A1* and *CCNG2* genes using electrophoretic mobility shift assays (EMSA). We synthesized small DNA probes (32 nucleotides) from the *ALDH1A1* and *CCNG2* promoters containing putative AT-hook target sequences (Figure 3), and from the *HES1* intron-1 as negative control. These probes were labeled with [^{32}P] and used in binding reactions with recombinant glutathione-S-transferase (GST), full-lenght GST-PTOV1 protein, GST-A domain, GST-B domain, and two short peptides containing the wild type eAT-hook or the mutated eAT-hook described previously[19] (Figure 3). A specific shifted band is detected when recombinant GST-PTOV1 is incubated with either *ALDH1A1* or *CCNG2* probes (Figure 3B). Surprisingly, the GST-A domain of PTOV1, but not the GST-B domain, showed a strong binding activity with both *ALDH1A1* and *CCNG2* probes. However, either the wild-type or the mutated eAT peptides did not generate any shifted band with these probes, and no shifted band was visible with the *HES1* negative control probe. (Figure 3B).

The analysis of the amino acids sequence of the A domain revealed a motif with similarities to a 'classic AT-hook' (Figure 3A)[33]. This sequence is not conserved in the corresponding homologous B domain of PTOV1. Therefore, we synthesized a short peptide of 15 amino acids that corresponds to the motif present in the A domain (Figure 3A). Interestingly, both labeled probes from *ALDH1A1* and *CCNG2* produced light shifted bands in the presence of this short peptide from the A domain (Figure 3B), indicating that PTOV1 is able to bind to these promoters through a new amino acid sequence present in the A domain, but not with the eAT-hook previously described, nor with the B domain of the protein.

3.4. The new PTOV1 AT-hook-like motif is necessary to modulate ALDH1A1 and CCNG2 expression.

In order to study the functional relevance of the KRRP sequence resembling the core AT-hook (Figure 3A) present in the peptide that binds to *ALDH1A1* and *CCNG2* promoters, we changed the motif to EGGP, by means of site-directed mutagenesis. This mutant was created in the GFP-full-length PTOV1 plasmid, so as to study the function of the protein as a transcription activator. As it is shown in Figure 4, both wild-type or mutant GFP-PTOV1 were efficiently expressed in transfected cells. Importantly, the expression of the mutant caused a moderate but significant decrease in *ALDH1A1* and *CCNG2* transcript levels when compared to wild-type PTOV1 (Figure 4B). The fused GFP-PTOV1 plasmids express comparable levels of the exogenous protein, indicating similar levels of transfection in these experiments (Figure 4C).

Together with the above demonstration of direct binding to *ALDH1A1* and *CCNG2* promoter sequences, these results lend support to the hypothesis that the AT-hook-*like* motif in the A domain of PTOV1 is critical for the PTOV1 promoted activation of transcription of these genes.

3.4. PTOV1, ALDH1A1 and CCNG2 expression levels are associated with aggressiveness in prostate and colon carcinomas.

To understand the significance of the ability of the oncoprotein PTOV1 to directly bind and activate the expression of *ALDH1A1* and *CCNG2* genes in tumor cells, we interrogated several publicly available databases containing expression data, clinical and pathological information of untreated patients with prostate cancer for the association of expression of these genes. Data derived from micro-dissected untreated prostate tumors specimens show that *PTOV1*, *ALDH1A1* and *CCNG2* transcript levels

are significantly increased in patients with high Gleason Score ($\geq$ 8) in comparison to patients with low Gleason score ($\leq$ 7) (Figure 5A) (GSE97284)[34]. The transcriptlevels of these genes are also significantly higher in prostate adenocarcinomas of patients who developed regional or distal metastasis after radical prostatectomy, suggesting their relationship with metastatic progression[8]. In addition, the expression levels of PTOV1 significantly correlated with the expression of *ALDH1A1* (Spearman 0.46, $p < 0.0001$) and *CCNG2* (Spearman 0.68, $p < 0.0001$) (GSE46691). Moreover, in a cohort of patients with prostate carcinoma including data derived from 10 studies[35-44], *PTOV1*, *ALDH1A1* and *CCNG2* genes also showed a highly significant co-occurrence of alterations at their DNA (Table 1). These observations suggest that the expression of *PTOV1*, *ALDH1A* and *CCNG2* genes is associated for a coordinated activity in aggressive prostate tumors.

Because PTOV1 can strongly and specifically bind to DNA *in vitro* and to chromatin *in vivo*, we searched for PTOV1 association with DNA in mitotic tumors cells using immunohistochemical analysis in different types of tumors. In xenografted mice tumors derived from LNCaP cells, a strong signal for PTOV1 was detected in mitotic cells where the staining appears associated to condensed chromosomes (Figure 5B). Similarly, in colon carcinoma tissues immunohistochemical analysis reveals a clear accumulation of PTOV1 in the nuclei of mitotic cells, confirming its strong association with chromatin in aggressive tumor cells (Figure 6A). Interrogating publicly available datasets of patients with colon carcinomas (GSE24551, GSE14333), high levels of *PTOV1*, *ALDH1A1* and *CCNG2* expression correlated with poor relapse-free survival and event-free survival, although for *ALDH1A1* the association did not reach significance (Figure 6B). Furthermore, immunohistochemical analyses of high grade urothelial bladder carcinoma and ductal breast carcinoma, previously shown to express

high levels of PTOV1[7], also revealed a clear association of the protein with condensed mitotic DNA (Figure 7).

4. Discussion

Here, we provide new insights into the mechanisms used by the protein PTOV1 to regulate the expression of *ALDH1A1* and *CCNG2*, relevant factors involved in tumor progression. Specifically, we show that PTOV1 is able to directly bind to the promoter sequences of these genes through a newly unveiled protein motif localized within the A domain of the protein[45]. This finding impacts on several aspects of the action of PTOV1 in cancer progression.

First, our findings define PTOV1 as a new nucleic acid binding protein containing two distinct motifs: one extended (eAT-hook) at the N-terminal[19] and a second AT-hook-*like* motif, identified and characterized in this work, within the A domain of PTOV1. The previously described eAT-hook does not bind to any of the DNA sequences tested here, in line with previous results that this motif has a higher affinity for RNA sequences[19]. In contrast, the newly identified AT-hook-*like* strongly binds to DNA sequences from the *ALDH1A1* and *CCNG2* promoters but not to *ABCB1* promoter or internal *HES1* gene sequences, indicating its specificity. Of note, mutations of the 'core' sequence in the AT-hook-*like* motif, decreases the expression of *ALDH1A1* and *CCNG2* indicating that this motif is functionally relevant for the transcriptional activity of PTOV1. Since PTOV1 was shown to function both at specific promoter sites to regulate transcription[17,18] and at ribosomes[16] to regulate mRNA translation, our present findings reveal potential new ways to selectively mitigate the nuclear or cytoplasmic functions of PTOV1 in cancer cells: the identification of inhibitors of these 'micro-handles' in the

protein could prevent its specific binding to either DNA or RNA, causing functionally diverse consequences on cell fate.

Secondly, we have shown that the transcriptional activation of *ALDH1A1* and *CCNG2* by PTOV1 is independent of the canonical Wnt or JNK pathways, being instead mechanistically explained by the direct binding of PTOV1 to specific sequences of these promoters. ALDH1A1 is an established hallmark of CSCs and promoter of drug resistance in different types of cancers, including colon cancer[14,30] and CCNG2 is an unconventional cyclin whose function in cancer progression remains to be determined. CCNG2 was first described as a tumor suppressor in several cancer types[46-48], being significantly upregulated in response diverse growth inhibitory stimuli, contributing to induce G2/M checkpoint and cell cycle arrest in response to doxorubicin[32] and to thiopurine resistance in lymphoblastoid cells[31]. However, it was also reported to modulate invasion in glioblastoma cells[49]. In these tumors, CCNG2 is remarkably expressed at hypoxic sites, and was shown to cooperate with actin-binding proteins providing flexibility to actin filaments for glioblastoma cell invasion. Consistent with an oncogenic role, we found that *CCNG2* is significantly co-expressed with *PTOV1* and *ALDH1A1* in aggressive prostate tumors and colon carcinomas, where overexpression of these genes is linked to progression and decreased relapse-free survival (Figure 5A and Figure 6B). The association of ALDH1A1 with poor survival was very recently confirmed in colon carcinomas[50]. Lastly, the significant co-occurrence of DNA alterations in *PTOV1*, *ALDH1A1* and *CCNG2* in prostate carcinomas (Table 1) together with the above mentioned findings, indicate a concerted action of these genes in cancer progression[14,31,45,49]. These observations also suggest that, similarly to HMGA proteins, PTOV1 bound to *ALDH1A1* and *CCNG2* promoters may function as an epigenetic

factor that triggers the recruitment of chromatin regulators and transcription factors to activate gene expression.

The expression of CCNG2, increased at G2-M phase, blocks the cell cycle before mitotic entry and chromosome condensation and is associated to thiopurine and doxorubicin resistance[31,32]. In addition, *ALDH1A1* is one of the most overexpressed gene in Solitary Fibrous Tumors, a rare spindle cell tumor with high mitotic counts[51]. Thus, PTOV1 action on *CCNG2* and *ALDH1A1* promoters might occur at mitotic stages and, conceivably, its binding to chromatin may be detected in condensed chromosomes of tumors cells. In fact, we observed positive PTOV1 staining in mitotic DNA in cells of prostate, colon, bladder and breast carcinomas. In mice tumors from LNCaP cells, numerous mitotic cells had a strong chromatin associated PTOV1 staining, although strong reactivity was also observed in the cytoplasm, suggesting its actions in cytoplasm and nucleus in aggressive tumor stages[27].

Thirdly, the protein structure of PTOV1 shares similarities with other proteins that interact with nucleic acids, like the β-barrel present in Ku (Ku70/Ku80) heterodimers and the SMRT/HDAC1-associated repressor protein (SHARP)[52,53]. Ku70/Ku80 are the DNA binding subunits of the DNA-PK complex necessary for the assembling of the DNA repair machinery for double strand breaks and are also implicated in transcriptional regulation[54]. SHARP is an important transcriptional regulator of nuclear receptor-mediated responses, Notch-mediated transcription, and X-chromosome silencing during development[53,55]. Besides sharing characteristics like DNA binding abilities and transcription regulatory functions, the two tandem domains in PTOV1 are structurally related to the SPOC domain in Ku and SHARP[52,56]. Similarly to PTOV1, SHARP also contains RNA recognition motifs, but is not known to contain nucleic acid-binding AT-hook-*like* motifs[53]. Members of the HMGA family of proteins also

share structural (presence of several AT-hook motifs) and functional similarities with PTOV1, *e.g*. promote transcription[20-21], bind directly to gene promoters, are overexpressed in cancer, and are associated to metastasis and therapy resistance[57-59].

AT-hook variants containing modifications that still allow binding to the DNA minor groove have been discovered, suggesting that the consensus core motif based on the HMGA family might not be so strict[60]. Recently, different unconventional AT-hook motifs were identified in DNA/RNA binding proteins[61,62]. Similar to the unconventional AT-hook-*like* motif reported here for PTOV1, an unconventional AT-hook domain 2 found in MeCP2 with a KRGRK core it is still able to bind AT-rich DNA sequences[63]. Another example of unconventional, AT-hook sequence is the TAF1 protein of *Drosophila melanogaster* that significantly diverges from the HGMA consensus without losing DNA affinity[60]. Moreover, Tip5, a subunit of the nucleolar remodeling complex (NoRC) that provides the link between nucleolar matrix and rDNA, also contains different AT-hook motifs for interaction with nucleic acids[64], including an *extended*-AT-hook that preferentially binds RNA to modulate the association between NoRC and promoter-associated RNA (pRNA)[65]. Finally, RNA-binding proteins Rrp12 and Srsf10 contain AT-hook motifs, suggesting that similarly to PTOV1 these proteins may have roles in DNA and RNA processes[66]. We propose that the previously identified eAT-hook and the newly unveiled AT-hook-*like* motifs have key roles in the functions of PTOV1 in mRNA translation and transcription, respectively[16-18].

In summary, we report a novel DNA binding motif in PTOV1 that allows its specific and direct binding to sequences of the *ALDH1A1* and *CCNG2* promoters, genes implicated in self-renewal and drug resistance, and their expression is associated with progression of tumors to aggressive stages and worse patients outcome. The distinct DNA- RNA-binding AT-hook motifs might direct PTOV1 nuclear and cytoplasmic

actions. Noteworthy, these motifs reveal specificities potentially useful to screen for inhibitors of distinct oncogenic functions of PTOV1.

Author's contributions

VM and VC carried out the majority of the experimental work, the statistical analyses, analyses of data, and participated in the writing of the manuscript. AJ participates in the design of EMSA, realization and analysis. CC and VN designed and performed the experimental EMSA with VM. IdT and MES performed the immunohistochemistry analysis of carcinoma tissues. JM participated in data analyses and interpretation. RP conceived and designed the study, interpreted the data, and wrote the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

The authors declare no potential conflict of interest.

References

1 Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. **68**, 7-30, doi:10.3322/caac.21442 (2018).

2 Attard, G. *et al.* Prostate cancer. *Lancet (London, England)* **387**, 70-82, doi:10.1016/s0140-6736(14)61947-4 (2016).

3 Huggins, C., Stevens, R. E., Jr & Hodges, C. V. Studies on prostatic cancer: Ii. the effects of castration on advanced carcinoma of the prostate gland. *Archives of Surgery* **43**, 209-223, doi:10.1001/archsurg.1941.01210140043004 (1941).

4 Beltran, H. *et al.* New therapies for castration-resistant prostate cancer: efficacy and safety. *European urology* **60**, 279-290, doi:10.1016/j.eururo.2011.04.038 (2011).

5 James, N. D. *et al.* Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet (London, England)* **387**, 1163-1177, doi:10.1016/s0140-6736(15)01037-5 (2016).

6 Mottet, N. *et al.* Updated Guidelines for Metastatic Hormone-sensitive Prostate Cancer: Abiraterone Acetate Combined with Castration Is Another Standard. *European urology*, doi:10.1016/j.eururo.2017.09.029 (2017).

7 Fernandez, S. *et al.* PTOV1 is overexpressed in human high-grade malignant tumors. *Virchows Archiv : an international journal of pathology* **458**, 323-330, doi:10.1007/s00428-010-1018-1 (2011).

8 Canovas, V. *et al.* Prostate Tumor Overexpressed-1 (PTOV1) promotes docetaxel-resistance and survival of castration resistant prostate cancer cells. *Oncotarget* **8**, 59165-59180, doi:10.18632/oncotarget.19467 (2017).

9 Morote, J. *et al.* PTOV1 expression predicts prostate cancer in men with isolated high-grade prostatic intraepithelial neoplasia in needle biopsy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 2617-2622, doi:10.1158/1078-0432.ccr-07-4987 (2008).

10 Yang, Q. *et al.* Prostate Tumor Overexpressed 1 (PTOV1) Is a Novel Prognostic Marker for Nasopharyngeal Carcinoma Progression and Poor Survival Outcomes. *PloS one* **10**, e0136448, doi:10.1371/journal.pone.0136448 (2015).

11 Guo, F. *et al.* Increased PTOV1 expression is related to poor prognosis in epithelial ovarian cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **36**, 453-458, doi:10.1007/s13277-014-2662-x (2015).

12 Lei, F. *et al.* Overexpression of prostate tumor overexpressed 1 correlates with tumor progression and predicts poor prognosis in breast cancer. *BMC cancer* **14**, 457, doi:10.1186/1471-2407-14-457 (2014).

13 Chen, S. P. *et al.* Prostate tumor overexpressed 1 is a novel prognostic marker for hepatocellular carcinoma progression and overall patient survival. *Medicine* **94**, e423, doi:10.1097/md.0000000000000423 (2015).

14 Li, T. *et al.* ALDH1A1 is a marker for malignant prostate stem cells and predictor of prostate cancer patients' outcome. *Laboratory investigation; a journal of technical methods and pathology* **90**, 234-244, doi:10.1038/labinvest.2009.127 (2010).

15 Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* **20**, 1455-1464, doi:10.1038/sj.onc.1204233 (2001).

16 Marques, N. *et al.* Regulation of protein translation and c-Jun expression by prostate tumor overexpressed 1. *Oncogene* **33**, 1124-1134, doi:10.1038/onc.2013.51 (2014).

17 Cui, Y. *et al.* Prostate tumour overexpressed-1 promotes tumourigenicity in human breast cancer via activation of Wnt/beta-catenin signalling. *The Journal of pathology* **239**, 297-308, doi:10.1002/path.4725 (2016).

18 Alana, L. *et al.* Prostate tumor OVerexpressed-1 (PTOV1) down-regulates HES1 and HEY1 notch targets genes and promotes prostate cancer progression. *Molecular cancer* **13**, 74, doi:10.1186/1476-4598-13-74 (2014).

19 Filarsky, M. *et al.* The extended AT-hook is a novel RNA binding motif. *RNA biology* **12**, 864-876, doi:10.1080/15476286.2015.1060394 (2015).

20 Benecke, A. G. & Eilebrecht, S. RNA-Mediated Regulation of HMGA1 Function. *Biomolecules* **5**, 943-957 (2015).

21 Shah, S. N. & Resar, L. M. High mobility group A1 and cancer: potential biomarker and therapeutic target. *Histology and histopathology* **27**, 567-579, doi:10.14670/hh-27.567 (2012).

22 Huth, J. R. *et al.* The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nature structural biology* **4**, 657-665 (1997).

23 Demidenko, R. *et al.* Frequent down-regulation of ABC transporter genes in prostate cancer. *BMC cancer* **15**, 683, doi:10.1186/s12885-015-1689-8 (2015).

24 Burger, P. E. *et al.* High aldehyde dehydrogenase activity: a novel functional marker of murine prostate stem/progenitor cells. *Stem cells (Dayton, Ohio)* **27**, 2220-2228, doi:10.1002/stem.135 (2009).

25 Albino, D. *et al.* Activation of the Lin28/let-7 Axis by Loss of ESE3/EHF Promotes a Tumorigenic and Stem-like Phenotype in Prostate Cancer. *Cancer research* **76**, 3629-3643, doi:10.1158/0008-5472.can-15-2665 (2016).

26 Desarnaud, F., Geck, P., Parkin, C., Carpinito, G. & Makarovskiy, A. N. Gene expression profiling of the androgen independent prostate cancer cells demonstrates complex mechanisms mediating resistance to docetaxel. *Cancer biology & therapy* **11**, 204-212 (2011).

27 Santamaria, A. *et al.* PTOV-1, a novel protein overexpressed in prostate cancer, shuttles between the cytoplasm and the nucleus and promotes entry into the S phase of the cell division cycle. *The American journal of pathology* **162**, 897-905, doi:10.1016/s0002-9440(10)63885-0 (2003).

28 Cerami, E. *et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* **2**, 401-404, doi:10.1158/2159-8290.cd-12-0095 (2012).

29 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).

30 Ma, I. & Allan, A. L. The role of human aldehyde dehydrogenase in normal and cancer stem cells. *Stem cell reviews* **7**, 292-306, doi:10.1007/s12015-010-9208-4 (2011).

31 Chouchana, L. *et al.* Molecular insight into thiopurine resistance: transcriptomic signature in lymphoblastoid cell lines. *Genome medicine* **7**, 37, doi:10.1186/s13073-015-0150-6 (2015).

32 Zimmermann, M. *et al.* Elevated cyclin G2 expression intersects with DNA damage checkpoint signaling and is required for a potent G2/M checkpoint arrest response to doxorubicin. *The Journal of biological chemistry* **287**, 22838-22853, doi:10.1074/jbc.M112.376855 (2012).

33 Reeves, R. & Nissen, M. S. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *The Journal of biological chemistry* **265**, 8573-8582 (1990).

34 Tyekucheva, S. *et al.* Stromal and epithelial transcriptional map of initiation progression and metastatic potential of human prostate cancer. **8**, 420, doi:10.1038/s41467-017-00460-4 (2017).

35 Gao, D. *et al.* Organoid cultures derived from patients with advanced prostate cancer. *Cell* **159**, 176-187, doi:10.1016/j.cell.2014.08.016 (2014).

36 Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer cell* **18**, 11-22, doi:10.1016/j.ccr.2010.05.026 (2010).

37 Fraser, M. *et al.* Genomic hallmarks of localized, non-indolent prostate cancer. *Nature* **541**, 359-364, doi:10.1038/nature20788 (2017).

38 Grasso, C. S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239-243, doi:10.1038/nature11125 (2012).

39 Robinson, D. *et al.* Integrative clinical genomics of advanced prostate cancer. *Cell* **161**, 1215-1228, doi:10.1016/j.cell.2015.05.001 (2015).

40 Beltran, H. *et al.* Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. **22**, 298-305, doi:10.1038/nm.4045 (2016).

41 Baca, S. C. *et al.* Punctuated evolution of prostate cancer genomes. *Cell* **153**, 666-677, doi:10.1016/j.cell.2013.03.021 (2013).

42 Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nature genetics* **44**, 685-689, doi:10.1038/ng.2279 (2012).

43 Kumar, A. *et al.* Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. **22**, 369-378, doi:10.1038/nm.4053 (2016).

44 Hieronymus, H. *et al.* Copy number alteration burden predicts prostate cancer relapse. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 11139-11144, doi:10.1073/pnas.1411446111 (2014).

45 Canovas, V., Lleonart, M., Morote, J. & Paciucci, R. The role of prostate tumor overexpressed 1 in cancer progression. *Oncotarget* **8**, 12451-12471, doi:10.18632/oncotarget.14104 (2017).

46 Cui, D. W., Sun, G. G. & Cheng, Y. J. Change in expression of cyclin G2 in kidney cancer cell and its significance. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **35**, 3177-3183, doi:10.1007/s13277-013-1415-6 (2014).

47 Sun, G. G., Hu, W. N., Cui, D. W. & Zhang, J. Decreased expression of CCNG2 is significantly linked to the malignant transformation of gastric carcinoma. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **35**, 2631-2639, doi:10.1007/s13277-013-1346-2 (2014).

48 Sun, G. G., Zhang, J. & Hu, W. N. CCNG2 expression is downregulated in colorectal carcinoma and its clinical significance. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **35**, 3339-3346, doi:10.1007/s13277-013-1440-5 (2014).

49 Fujimura, A. *et al.* Cyclin G2 promotes hypoxia-driven local invasion of glioblastoma by orchestrating cytoskeletal dynamics. *Neoplasia (New York, N.Y.)* **15**, 1272-1281 (2013).

50 van der Waals, L. M., Borel Rinkes, I. H. M. & Kranenburg, O. ALDH1A1 expression is associated with poor differentiation, 'right-sidedness' and poor survival in human colorectal cancer. *PloS one* **13**, e0205536, doi:10.1371/journal.pone.0205536 (2018).

51 Bertucci, F. *et al.* Gene expression profiling of solitary fibrous tumors. *PloS one* **8**, e64497, doi:10.1371/journal.pone.0064497 (2013).

52 Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature structural & molecular biology* **18**, 404-409, doi:10.1038/nsmb.1997 (2011).

53 Oswald, F. *et al.* RBP-Jkappa/SHARP recruits CtIP/CtBP corepressors to silence Notch target genes. *Molecular and cellular biology* **25**, 10379-10390, doi:10.1128/mcb.25.23.10379-10390.2005 (2005).

54 Fell, V. L. & Schild-Poulter, C. The Ku heterodimer: function in DNA repair and beyond. *Mutation research. Reviews in mutation research* **763**, 15-29, doi:10.1016/j.mrrev.2014.06.002 (2015).

55 McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232-236, doi:10.1038/nature14443 (2015).

56 Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *Journal of structural biology* **174**, 245-251, doi:10.1016/j.jsb.2010.10.011 (2011).

57 Mendez, O. & Peg, V. Extracellular HMGA1 promotes tumor Invasion and Metastasis in Triple-Negative Breast Cancer. doi:10.1158/1078-0432.ccr-18-0517 (2018).

58 Liau, S. S., Rocha, F., Matros, E., Redston, M. & Whang, E. High mobility group AT-hook 1 (HMGA1) is an independent prognostic factor and novel therapeutic target in pancreatic adenocarcinoma. *Cancer* **113**, 302-314, doi:10.1002/cncr.23560 (2008).

59 Wood, L. J. *et al.* HMG-I/Y, a new c-Myc target gene and potential oncogene. *Molecular and cellular biology* **20**, 5490-5502 (2000).

60 Metcalf, C. E. & Wassarman, D. A. DNA binding properties of TAF1 isoforms with two AT-hooks. *The Journal of biological chemistry* **281**, 30015-30023, doi:10.1074/jbc.M606289200 (2006).

61 Blokken, J., De Rijck, J., Christ, F. & Debyser, Z. Protein-protein and protein-chromatin interactions of LEDGF/p75 as novel drug targets. *Drug discovery today. Technologies* **24**, 25-31, doi:10.1016/j.ddtec.2017.11.002 (2017).

62 Garabedian, A., Bolufer, A. & Leng, F. Peptide Sequence Influence on the Conformational Dynamics and DNA binding of the Intrinsically Disordered AT-Hook 3 Peptide. **8**, 10783, doi:10.1038/s41598-018-28956-z (2018).

63 Lyst, M. J., Connelly, J., Merusi, C. & Bird, A. Sequence-specific DNA binding by AT-hook motifs in MeCP2. *FEBS letters* **590**, 2927-2933, doi:10.1002/1873-3468.12328 (2016).

64 Zillner, K. *et al.* Large-scale organization of ribosomal DNA chromatin is regulated by Tip5. *Nucleic acids research* **41**, 5251-5262, doi:10.1093/nar/gkt218 (2013).

65 Zhou, Y. *et al.* Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. *Nature cell biology* **11**, 1010-1016, doi:10.1038/ncb1914 (2009).

66 Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nature reviews. Genetics* **15**, 829-845, doi:10.1038/nrg3813 (2014).

TABLE 1

Gene A	Gene B	Log Odds Ratio	Adjusted p-Value	Tendency
PTOV1	*ALDH1A1*	>3	**<0.001**	Co-occurrence
PTOV1	*CCNG2*	2.762	**<0.0001**	Co-occurrence

Table 1. Co-occurrence of DNA alterations in prostate tumors.

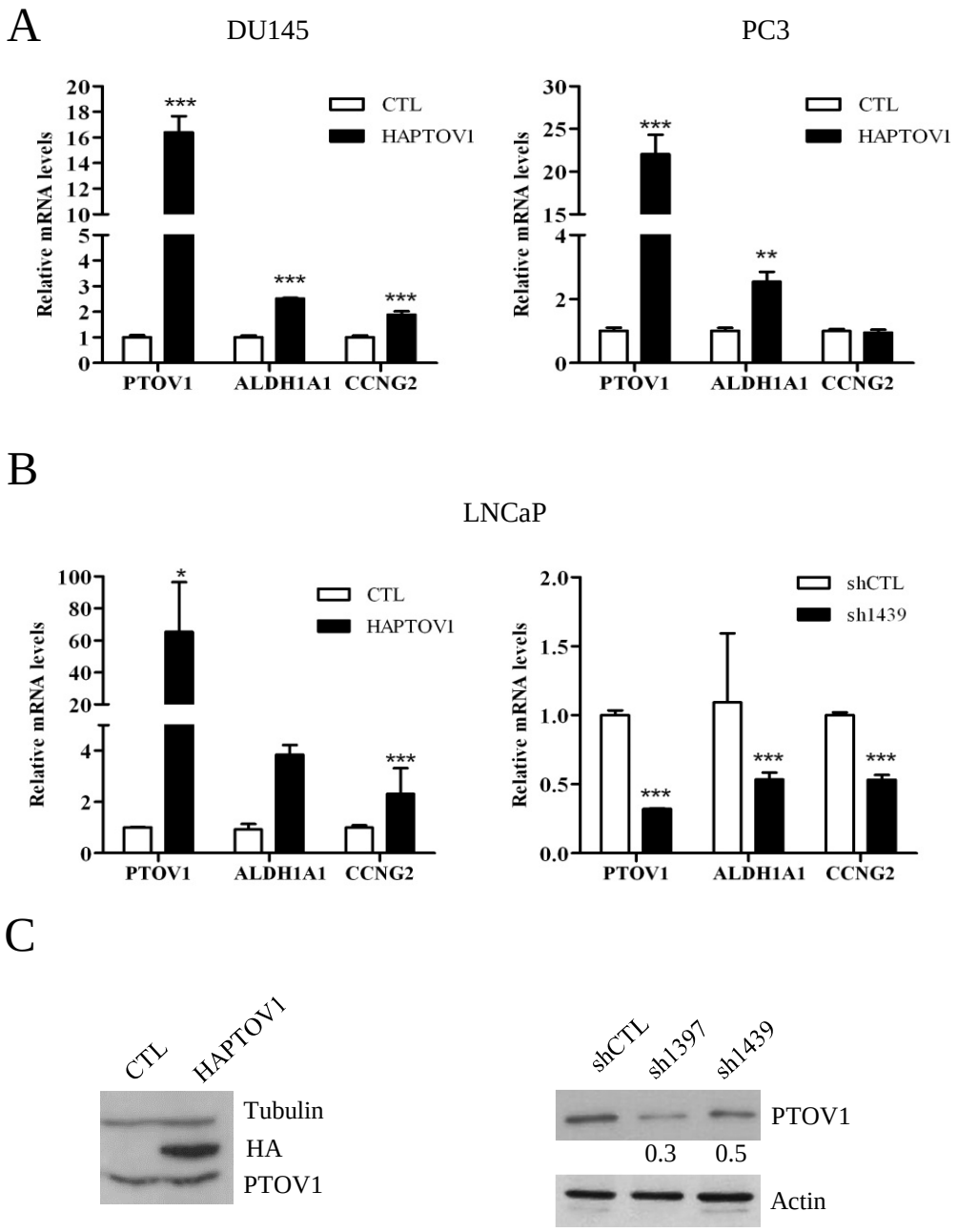

A
DU145
PC3
CTL
HAPTOV1
Relative mRNA levels
PTOV1
ALDH1A1
CCNG2
B
LNCaP
shCTL
sh1439
C
CTL
HAPTOV1
Tubulin
HA
PTOV1
shCTL
sh1397
sh1439
PTOV1
0.3
0.5
Actin

Figure 1

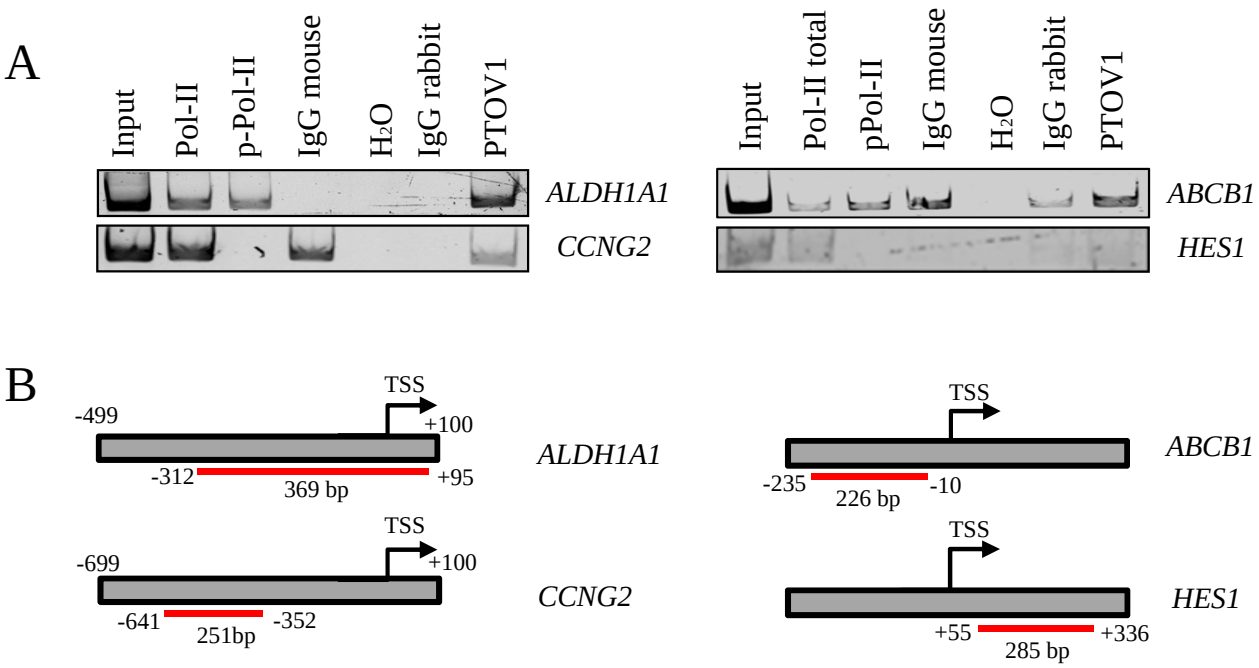
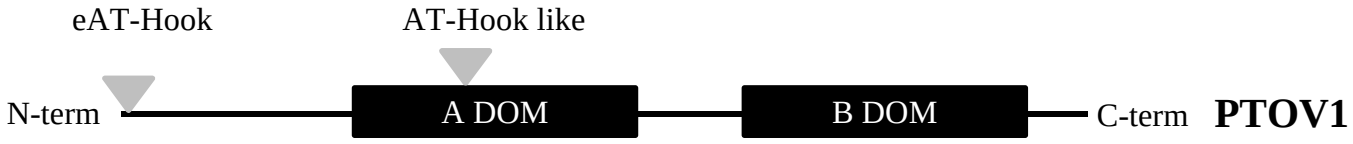

A
Input
Pol-II
p-Pol-II
IgG mouse
H2O
IgG rabbit
PTOV1
ALDH1A1
CCNG2
Input
Pol-II total
pPol-II
IgG mouse
H2O
IgG rabbit
PTOV1
ABCB1
HES1
B
TSS
-499
+100
-312
+95
369 bp
ALDH1A1
TSS
-699
+100
-641
-352
251bp
CCNG2
TSS
-235
-10
226 bp
ABCB1
TSS
+55
+336
285 bp
HES1

Figure 2

A



eAT-hook	M V R P R R A P Y R S G A G G P L G G R **G R P** P R P L V V R A V R S R S W P A S P R G
A Dom	S N K L L A W S G V L E W Q E K **R R P** Y S D S T A L K L K R T L P C Q A Y V N
B Dom	V N N K F L A W S G V M E W Q E P R P E P N S R S K R W L P S H V Y V N Q G E
A Dom peptide	K **R R P** Y S D S T A L K L K R
AT-hook consensus	T E K R G R **G R P** R K

B

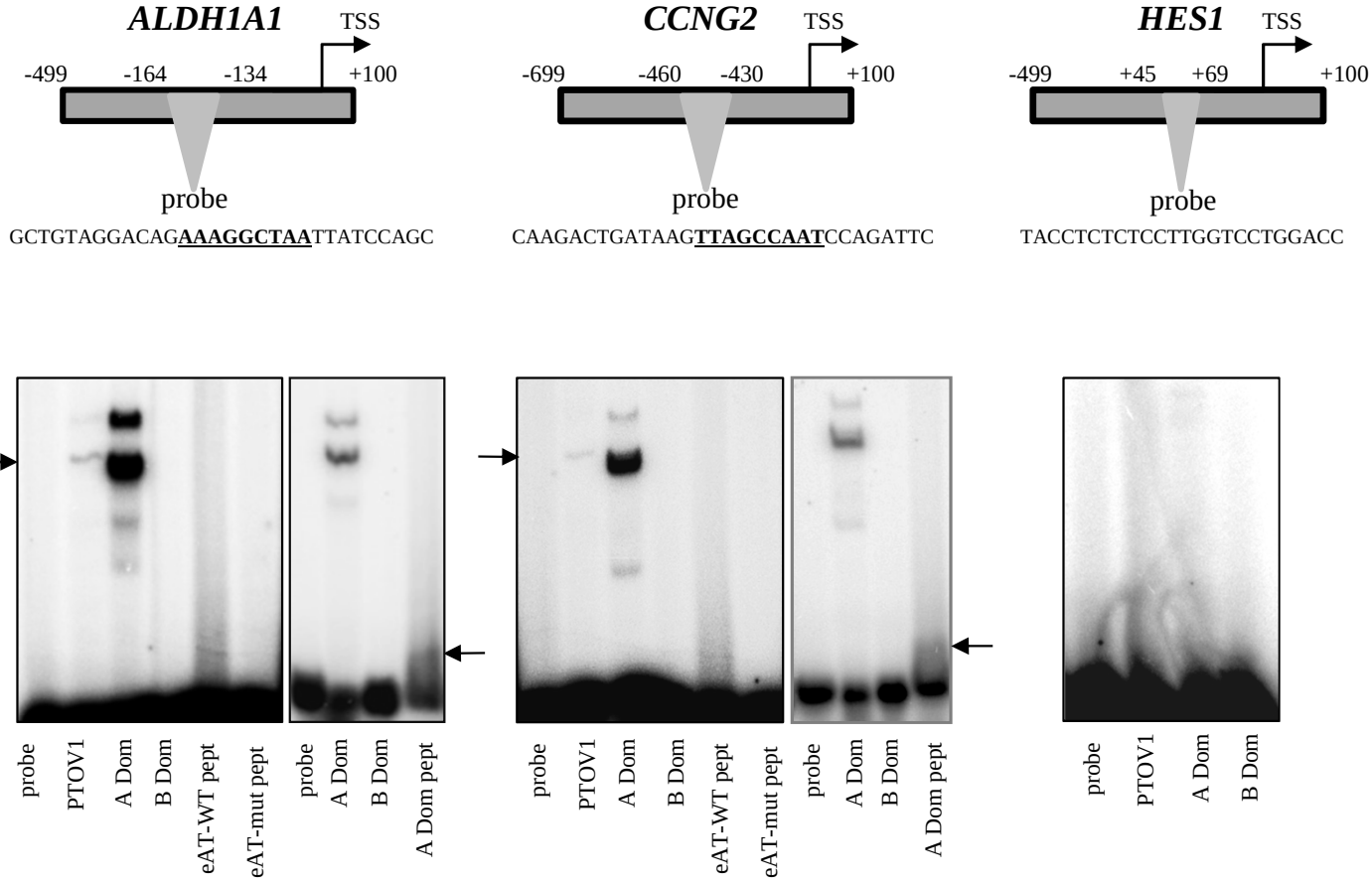
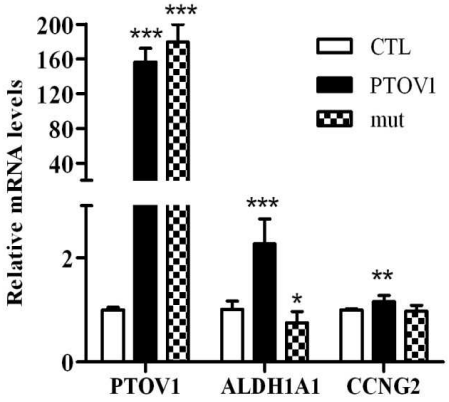
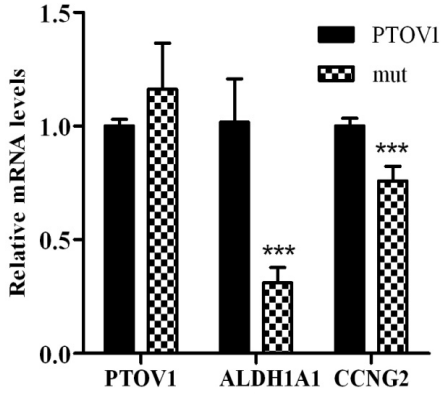


Figure 3

A


Relative mRNA levels
200
160
120
80
40
2
0
***
***
***
*
**
CTL
PTOV1
mut
PTOV1
ALDH1A1
CCNG2

B


Relative mRNA levels
1.5
1.0
0.5
0.0
***
***
PTOV1
mut
PTOV1
ALDH1A1
CCNG2

C

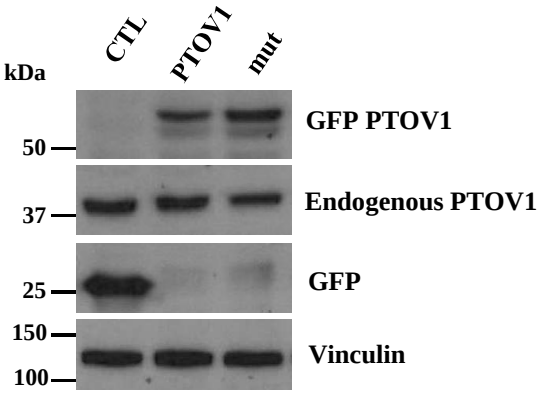
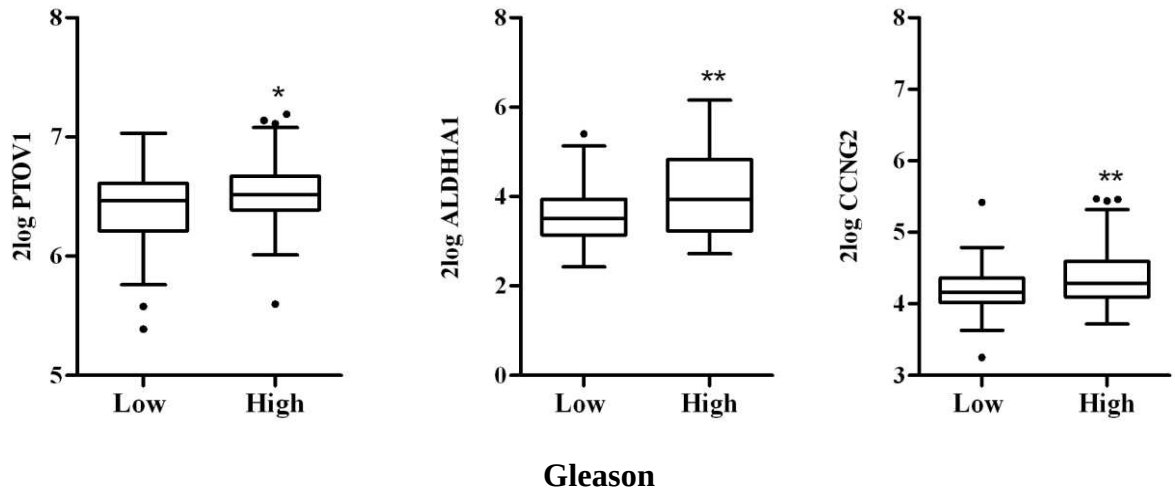

kDa
CTL
PTOV1
mut
50
37
25
150
100
GFP PTOV1
Endogenous PTOV1
GFP
Vinculin

Figure 4

A


2log PTOV1
8
7
6
5
*
Low
High
2log ALDH1A1
8
6
4
2
0
**
Low
High
2log CCNG2
8
7
6
5
4
3
**
Low
High
Gleason

B

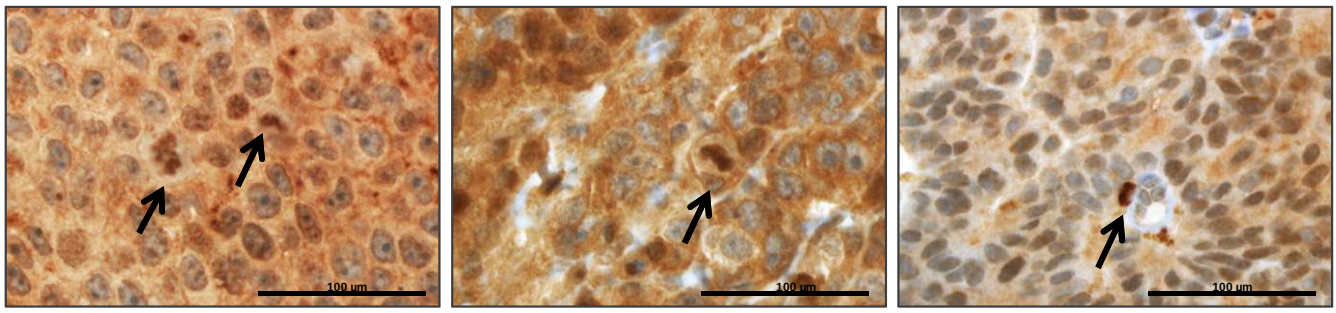
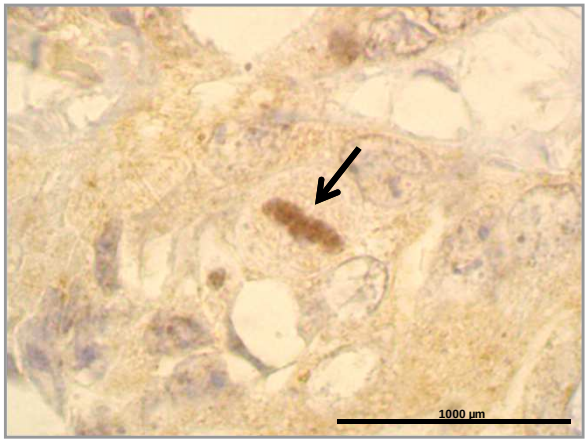
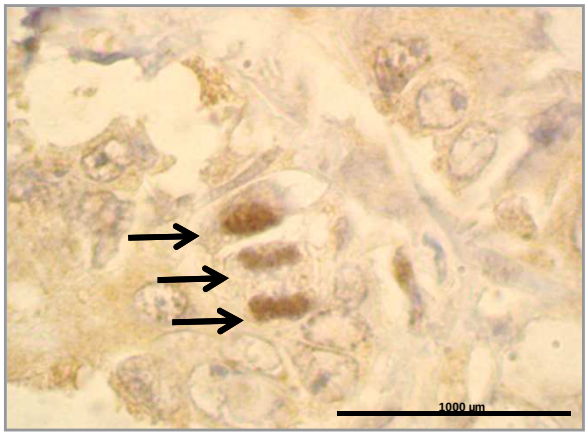
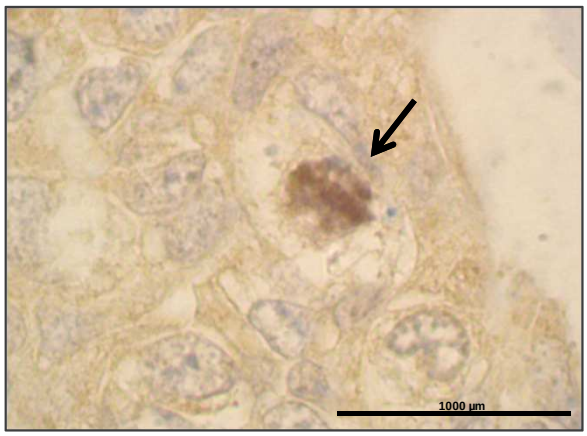

100 um
100 um
100 um

Figure 5

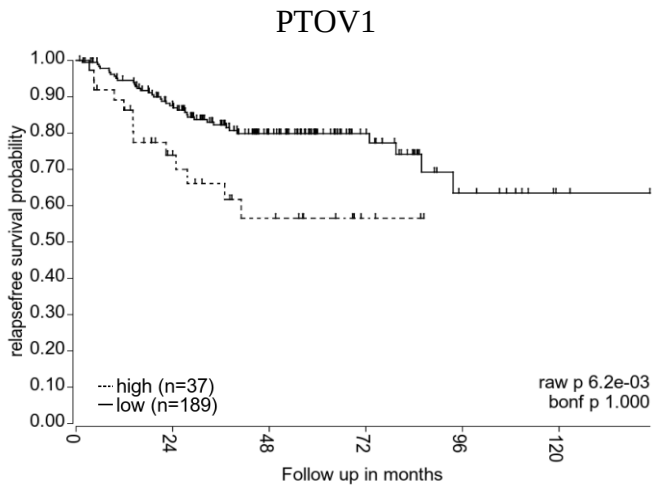
A

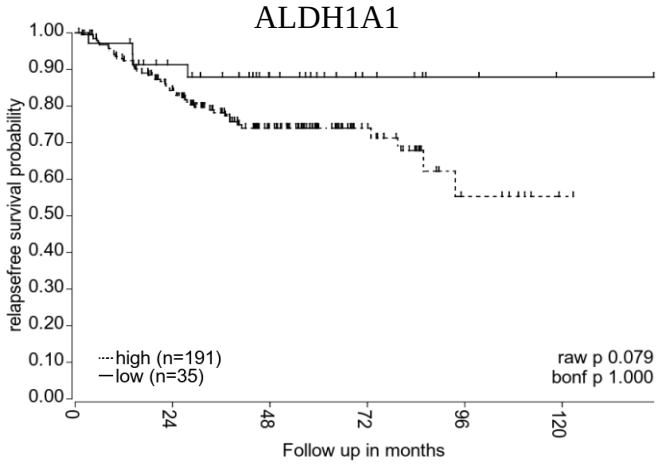

1000 µm


1000 µm


1000 µm

B


PTOV1
relapsefree survival probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
high (n=37)
low (n=189)
raw p 6.2e-03
bonf p 1.000
0
24
48
72
96
120
Follow up in months


ALDH1A1
relapsefree survival probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
high (n=191)
low (n=35)
raw p 0.079
bonf p 1.000
0
24
48
72
96
120
Follow up in months

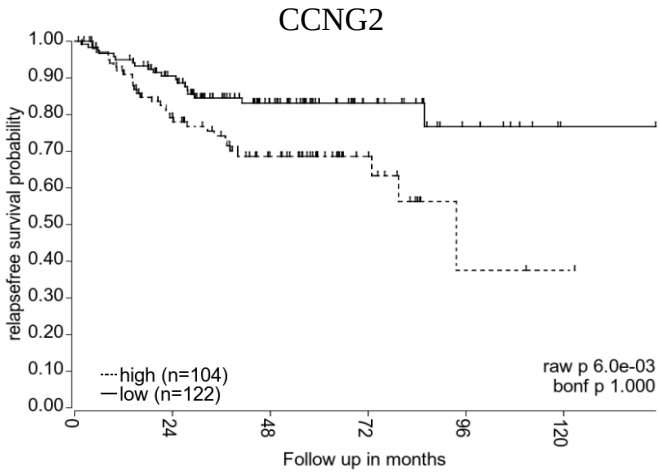

CCNG2
relapsefree survival probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
high (n=104)
low (n=122)
raw p 6.0e-03
bonf p 1.000
0
24
48
72
96
120
Follow up in months

Figure 6

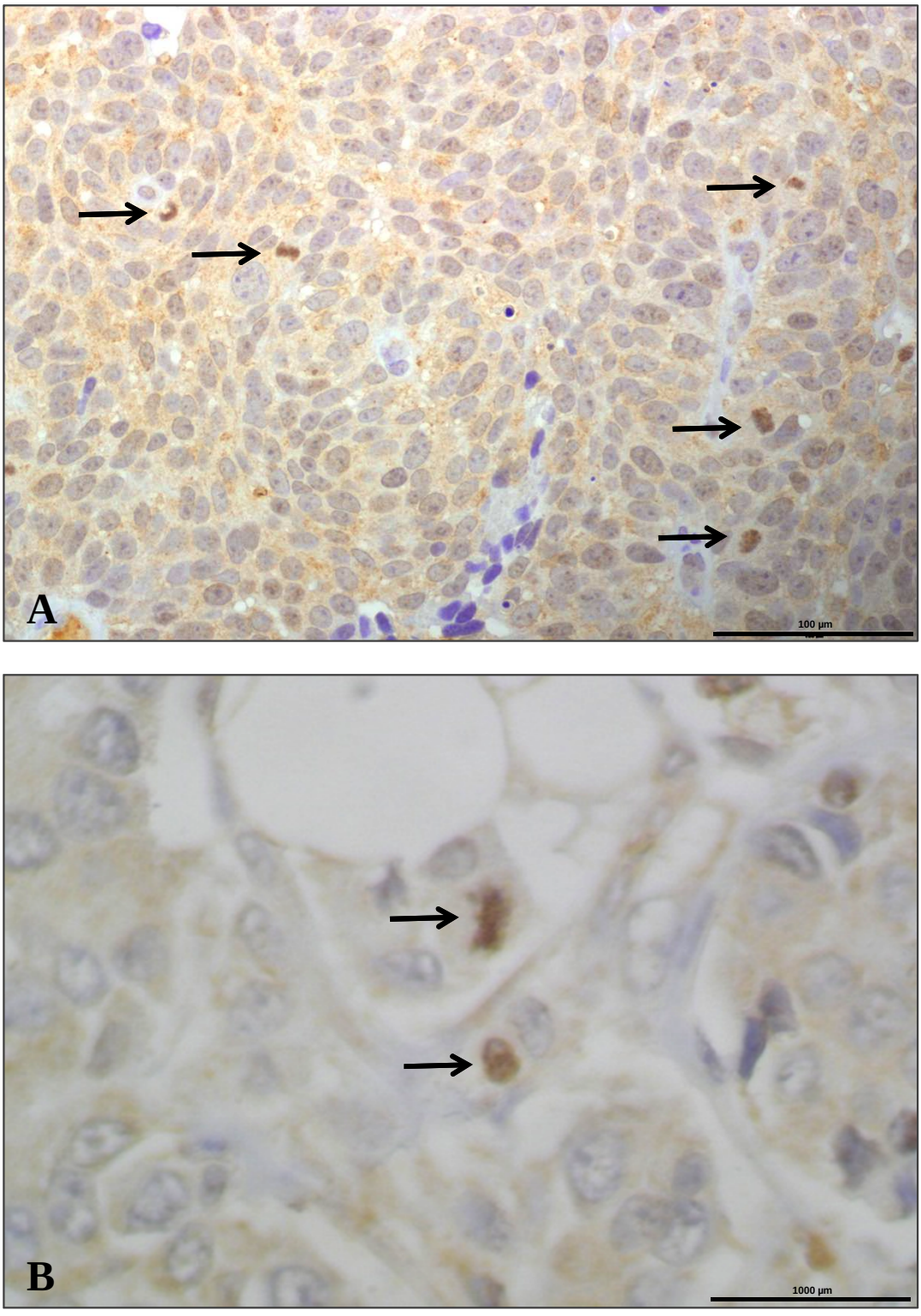


Figure 7

FIGURE LEGENDS

Figure 1. The ectopic expression of PTOV1 in prostate cancer cells promotes *ALDH1A1* and *CCNG2* expression. (**A**) CRPC Du145 and PC3 cells transduced with a lentiviral vector HAPTOV1-ires-GFP, or a control lentivirus (GFP), were analyzed by real time qPCR for *ALDH1A1* and *CCNG2* expression. **(B)** Left, LNCaP cells transduced with the lentiviral vector HAPTOV1, or the control lentivirus (GFP), were analyzed by real time qPCR for *gene* expression. Right, LNCaP cells transduced with lentiviral vectors bearing shRNA sequences (sh1397 or sh1439) and a control shRNA (shCTL) for PTOV1 knockdown, were analyzed by real time qPCR.(C) Immunoblots of LNCaP cells as in B, identify the endogenous or the HAPTOV1 protein. . Numbers in the right panel, express the reduction of PTOV1 protein levels as quantified by densitometric analysis of the signal with respect to actin using ImageJ software. p-value: * < 0.05; ** < 0.01; *** < 0.001.

Figure 2. PTOV1 is associated to the chromatin of *ALDH1A1* and *CCNG2* promoters. (**A**) Sheered chromatin from LNCaP cells transduced with a lentivirus encoding the fusion protein HAPTOV1 was immunoprecipitated with rabbit antibodies to PTOV1, total and phosphorylated polymerase II, and rabbit or mouse IgGs as controls. Co-immunoprecipitated DNA fragments were analyzed by PCR with specific primers for *ALDH1A1*,*CCNG2*, *ABCB1* and *HES1* promoter regions. (**B**) Graphical representation of primers localization and length of the amplified sequences on the promoters of the indicated genes. TSS: transcription start site.

Figure 3. EMSA identify a new AT-hook-*like* motif in the A domain of PTOV1. **(A)** Schematic of the protein structure of PTOV1 identifying the AT-hook domains present at the N-terminal (eAT-hook) and within the A domain (AT-hook-*like*) (not in scale). Bottom, eAT-hook, the first 43 amino acids from the N-terminal of PTOV1; A domain, the partial amino acid sequence (85 to 125) of the A domain; B domain, the partial amino acid sequence of the B domain (amino acid 249 to 292); A domain peptide, the amino acid sequence of the AT-hook-*like* motif peptide used for EMSA assays (amino acid 100 to 114); AT-hook consensus, the described AT-hook of HGMA1 (amino acid 21 to 31). Bold characters identify amino acids in the 'core' of AT-hook motifs. (**B**) Drawings show the localization of the DNA sequence probes with respect to the transcription start sites (TSS). Bottom panels: two gel shift assays are shown for each *ALDH1A1* and *CCNG2* probe: in the left gels, labeled sequences were incubated with recombinant GST-PTOV1 (PTOV1), GST-A domain (A dom), GST-B domain (B dom), eAT-hook wild type peptide (eAT-WT pept) and mutated eAT-hook peptide (eAT-mut pept). In the right gels, labeled sequences were incubated with the GST-A domain (A dom), GST-B domain (B dom), and the A domain AT-hook-*like* peptide (A dom pept). Arrows indicate the shift provoked by the binding of protein domains or peptides to labeled DNA. A gel shift assay performed with the *HES1* probe is shown as control.

Figure 4. Mutation of the newly discovered AT-hook-*like* 'core' motif of PTOV1 reduces the protein transcription capacity**.** The plasmid containing GFP/full-lenght PTOV1 mutated at the AT-hook (EGGP), or wild-type plasmid, and the control GFP

plasmid were transfected in HEK293T cells. (**A**) Cells were analyzed for gene expression by real-time PCR 72 h after transfection. The graph on the left shows the relative expression of *PTOV1, ALDH1A1* and *CCNG2* as compared to their endogenous levels (cells transfected with the GFP plasmid, CTL). (**B**) The levels of expression of GFP-PTOV1 (wild-type) taken as control compared to mutated PTOV1 from the same experiment as in A, show the effect that mutant exogenous PTOV1 has on the expression of downstream *ALDH1A1* and *CCNG2* genes in comparison to wild-type PTOV1. . (**C**) Western blot analysis of cells transfected with wild-type or mutant GFP-PTOV1 plasmids, shows similar levels of expression of the exogenous PTOV1, as indicated by the GFP antibody signals. Vinculin is shown as protein loading control.

Figure 5. The expression of *PTOV1, ALDH1A1* and *CCNG2* is significantly associated with the Gleason score in prostate carcinomas. (**A**) Box and whisker plots represent *PTOV1, ALDH1A1* and *CCNG2* expression levels in prostate tumors with different Gleason as obtained from published database (GSE 97284). Low grade represents the grade group of ≤7 n= 83; High grade indicates the grade groups 4 and 5 (Gleason score: 4+4, 3+5, 5+3, 4+5, 5+4, 5+5) n= 91. (**B**) Immunohistochemical stainings of antibodies to PTOV1 in LNCaP tumors formed in immunosuppressed mice. PTOV1 associates to condensed chromatin in aggressive mitotic cells (arrows). Scale bar: 100 μm. p-value: * <0.05; ** < 0.01; *** < 0.001.

Figure 6. PTOV1 accumulates in the nuclei of mitotic colon carcinoma cells. (**A**) Immunohistochemistry analyses of aggressive colon carcinoma cells show PTOV1 localized in condensed mitotic nuclei (arrows). Scale bar: 1000 μm. (**B**) The expression of PTOV1 significantly correlates with poor overall survival in colon tumors. Public

datasets (GSE62452, GSE3141, GSE45547, GSE24551, GSE21653, GSE14333) containing information of survival rates in patients with colon cancer were analyzed using R2 platform.

Figure 7. PTOV1 accumulates in the nuclei of mitotic bladder carcinoma and breast carcinoma cells. Immunohistochemical staining of antibodies to PTOV1 in high grade bladder carcinoma **(A)** and breast carcinoma **(B)** reveal strong staining associated to chromatin of mitotic cells (arrows). Scale bar: 1000 µm.

Highlights

- A novel DNA-binding motif in PTOV1 is required for *ALDH1A1* and *CCNG2* overexpression.
- PTOV1 is a direct transcriptional activator of gene expression in prostate cancer.
- Distinct DNA- RNA-binding motifs might direct PTOV1 nuclear and cytoplasmic actions.
- *PTOV1, ALDH1A1, CCNG2* expression associated with prostate and colon cancer bad prognosis.