Am J Cancer Res 2022;12(2):585-600
www.ajcr.us /ISSN:2156-6976/ajcr0140533

Original Article

PBX4 functions as a potential novel oncopromoter in colorectal cancer: a comprehensive analysis of the *PBX* gene family

Eirini G Martinou[1,2], Carla S Moller-Levet[3], Angeliki M Angelidi[4]

[1]*Hepatobiliary and Pancreatic Surgery Department, Royal Surrey County Hospital, Guildford GU2 7XX, UK;* [2]*Faculty of Health and Medical Sciences, University of Surrey, Guildford GU2 7XH, UK;* [3]*Bioinformatics Department, Faculty of Health and Medical Sciences, University of Surrey Guildford, GU2 7XH, UK;* [4]*Department of Medicine, Beth Israel Deaconess Medical Centre, Harvard Medical School, Boston, MA 02215, USA*

Received November 13, 2021; Accepted January 10, 2022; Epub February 15, 2022; Published February 28, 2022

Abstract: Pre-B-cell leukaemia (PBX) is a transcription factor family (*PBX1, PBX2, PBX3* and *PBX4*) that regulates important cellular functions and has been identified to be involved in human cancers. This study aimed to explore the expression of *PBX* genes and their clinical significance in colorectal cancer (CRC). We analysed the differential expression of *PBX* genes in CRC vs. normal tissue, using the Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) and ONCOMINE platform (https://www.oncomine.org/). The UALCAN (http://ualcan.path.uab.edu/) interactive OMICS web-server was used to evaluate the epigenetic regulation of *PBX* genes via their promoter methylation status. We found that only *PBX4* was upregulated whereas *PBX1* and *PBX3* were downregulated (644 tumour vs. 51 normal samples) ($P<0.001$). The methylation status of *PBX4* promoter appeared to be decreased ($P=1.4e-07$) whereas the methylation status of *PBX1* and *PBX3* promoters was increased ($P=3.8e-04$ and $P=3.2e-07$, respectively) in cancer vs. normal samples. To determine the prognostic value of *PBX*s, we conducted a Kaplan-Meier survival analysis and multivariable COX regression. We observed that high *PBX4* expression was associated with increased risk for a worse overall survival (OS) in the TCGA CRC patient cohort (n=639), (HR 1.46, 95% CI 1.14-1.88, $P=0.003$) adjusted for age, gender, tumour location and metastases. We conducted in vitro gene expression modulation experiments to investigate the impact of *PBX4* overexpression in CRC cell (HCT116) growth. Additionally, we evaluated the RNA expression of epithelial-mesenchymal transition (EMT) and angiogenesis markers. *In vitro* studies showed that *PBX4* overexpression increased CRC cell proliferation ($P<0.001$) and upregulated the expression of EMT markers *VIM, CDH1, CDH2, ZEB1, SNAI1* ($P<0.05$) and angiomarker *VEGFA* ($P<0.0001$). Lastly, through the Cistrome data browser (http://dbtoolkit.cistrome.org/) we investigated putative transcriptional regulators and we performed gene set enrichment analysis in Enrichr server (https://maayanlab.cloud/Enrichr/) to identify related biological processes. Nineteen factors were identified to be putative regulators of *PBX4* and gene set enrichment analysis showed that biological processes related to cell cycle and cell proliferation were enriched (GO:0051726: *CDK8*, *JUN*, *JUND*, and *IRF1*, $P=0.001$). In conclusion, our study identified *PBX4* as a potential novel oncopromoter in CRC and its overexpression was found to be associated with increased risk for worse survival rate.

Keywords: Pre-B-cell leukaemia, PBX, colorectal, cancer, cell proliferation, cancer progression, survival

Introduction

Colorectal cancer (CRC) is an alarming public health threat as its incidence is rising, especially in the younger population [1]. Epidemiological data reveal that about 20-25% of patients with CRC, harbour metastatic disease at the time of their diagnosis and a further 40-50% will develop metastases, usually within the first 3 years of follow up after successful resection of the primary tumour [2]. Despite the improvement in surgical and systemic therapies, the prognosis for CRC patients remains poor, indicating that cancerous cells are not entirely eradicated by the current treatments available, leading to metastatic disease [2]. In the era of precision medicine, genome-based studies highlight the importance of identifying transcription factors that could serve as markers of tumour biology, prognosis or those that may be used as therapeutic targets in CRC [3].

Pre-B-Cell leukaemia homeobox (PBX), is a family of transcription factors that belongs to a larger highly conserved group called the TALE (three amino acid loop extension) family and has been found to interact with other proteins to control the transcription of target genes [4]. Humans have four *PBX* homologues in their genome *PBX1, PBX2, PBX3* and *PBX4* which are mostly known as regulatory proteins that control haematopoiesis [5]. PBX proteins mainly interact with Homeobox proteins (HOX) and therefore are widely known as HOX cofactors; however, emerging evidence shows that they may have a broader role by demonstrating both HOX-dependent and HOX-independent roles [6].

In recent years, research on *PBX* genes has attracted interest due to the increasing amount of evidence that exists regarding the important role of *HOX* genes in cancer pathogenesis and progression [7]. Direct targeting of HOX/PBX dimers has been reported to impair tumour growth and sensitise cells to standard chemotherapy in malignancy [4]. *PBX* genes have been found to be dysregulated, affect tumour progression and play an important prognostic role in various types of cancers such as breast, gastric cancer and haematological malignancies [8, 9]. Regarding CRC, limited evidence exists in terms of *PBX* dysregulation and its role in prognosis and progression. This study for the first time aimed to investigate the differential expression of the *PBX* gene family in CRC and its impact on overall survival (OS). Additionally, we aimed to uncover the impact of *PBX4* gene overexpression in CRC progression in vitro.

Materials and methods

Bioinformatics TCGA cancer to normal analysis and data mining

Tissue Cancer Genome Atlas (TCGA) expression data regarding colon (TCGA-COAD) and rectal adenocarcinoma (TCGA-READ) were downloaded from the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/) consisting of High-Throughput Sequencing count files [10]. The edgeR Bioconductor package was used for data pre-processing and differential expression analysis [11]. Lowly expressed genes were defined as having less than 8 Counts Per Million (CPM) in more than 51 samples and were filtered out. Data were normalised using the weighted trimmed mean of the log expression ratios (trimmed mean of M values: TMM method) using the R package edgeR (edgeR_3.24.3) [11]. A negative binomial generalised log-linear model was fitted to the read counts for each gene and likelihood ratio tests regarding differences in tumour vs. normal tissue were conducted. *P*-values were adjusted for multiple comparisons using the Benjamini-Hochberg (BH) approach [12]. TMM normalised log2 CPM was used to cluster the gene expression matrix. Clustered heatmaps were produced based on z-scored standardisation of gene values across samples. The ONCOMINE platform (https://www.oncomine.org/) was used for data mining and the fold change in expression in CRC vs. normal tissue from each reporter was recorded. The UALCAN (http://ualcan.path.uab.edu/) interactive OMICS web-server was utilised to evaluate the epigenetic regulation of *PBX* genes via promoter methylation status [13].

Bioinformatics TCGA survival analysis

For the survival TCGA analysis, the corresponding metadata and clinical data files of the TCGA COAD and READ datasets were downloaded as described above. After filtering out lowly expressed genes, the median CPM value across samples was used as a cut-off for minimum gene expression. Genes with a CPM value above the cut-off in at least 25% of all samples were retained. Data normalisation was performed as described above. Univariable Cox regression analysis was conducted, and the variables included were age, gender, tumour location, presence of metastases, the *PBX* gene family and all *HOX* genes. Gene expression and age were analysed as continuous variables, whereas tumour location, gender and metastases were analysed as binary categorical variables. Multivariable analysis was performed to assess whether *PBX* genes are independently associated with OS rate. Multivariable model 1 was adjusted for age, metastases and *HOX* genes and model 2 was additionally adjusted for gender and tumour location. Kaplan-Meier analysis was performed between patients with high vs. those with low gene expression. As a cut-off value, the value that produced the lowest *P*-value was selected. All survival analyses were conducted using the R packages (R version 3.5.1) survival_2.43-3 and survminer_0.4.3. The hazard ratio (HR), as well as the 95% confidence interval (CI) of HR, was calculated.

In vitro gene expression modulation

Study design and protocol: For the *in vitro* studies, the human HCT116 colon adenocarcinoma cell line was acquired from the American Type Culture Collection (ATCC, USA) which has low endogenous *PBX4* expression. Cells were maintained in McCoys 5A modified medium (Fisher Scientific, UK), supplemented with 10% foetal bovine serum, 100 μg/ml penicillin and 100 μg/ml streptomycin (Fisher Scientific, Gibco, UK). Cells were cultured at 37°C in a humidified 5% CO_2 atmosphere incubator. Cell culture was carried out in accordance with the guidelines for the use of cell lines in biomedical research [14].

Plasmid vectors tagged with Green Fluorescence Protein (GFP) (OriGene, Germany) were used in the overexpression mechanistic studies. The pCMV6-AC-PBX4-GFP (RC217742) and pCMV6-AC-GFP (PS100010) vectors were used as *PBX4* overexpressing and empty control vectors, respectively. HCT116 cells were transiently transfected with either a *PBX4*-overexpressing or empty vector in a 24-well plate by following the reverse transfection technique. A transfection mixture was prepared by mixing 500 ng of plasmid DNA with 2 μl of ViaFect (Promega, UK) transfection reagent in 100 μl of OptiMem medium (Life Technologies, UK) and was incubated for 20 mins at room temperature. HCT116 cells were seeded in 24 well plates at 7×10^4 density containing 100 μl of transfection mixture to initiate *PBX4* overexpression. After either 24- or 48-hours, cells were harvested and counted using a Countess II Cell Counter (Thermofisher, UK). HCT116 cells from each group were subcultured in 96 well plates at 2.5×10^3 density for use in the downstream functional assay. PBX4 overexpression was evaluated at mRNA and protein level using a Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR) and Western blot, respectively, as described below. The main outcome was cell proliferation over time which was assessed using AlamarBlue fluorescence proliferation assay every 24 hours up to 5 days post-transfection [15]. The study consisted of four biological experiments and contained eight to twelve technical replicates in each experiment. Additionally, we calculated the fold change in expression of epithelial-mesenchymal transition (EMT) markers (*CDH1, CDH2, VIM, SNAI1, SNAI2, ZEB1, ZEB2* and *TWIST*) and the angiogenesis marker (*VEGFA*). The flow chart of the *in vitro* study methodology is shown in **Figure 1**. The normality of the data was evaluated using the Shapiro-Wilk test. An independent-samples T-test (for normally distributed values) or Mann-Whitney test (for non-normally distributed values) was used to compare the differences between the control and the treated groups using GraphPad Prism 8 and SPSS v27. A two-tailed *P*-value ≤ of 0.05 was considered statistically significant.

RNA isolation and real-time quantitative polymerase chain reaction (RT-qPCR): RNA isolation was performed using the RNeasy® Plus Micro Kit (Qiagen, UK) following the manufacturer's instructions. The concentration and purity of RNA were determined spectrophotometrically by measuring its optical density (260/280; 260/230 ratios) using a Nanodrop ND-1000 (Labtech International, UK) [16].

Gene expression was assessed with RT-qPCR using the AgPath-ID™ One-Step RT-PCR Reagents mix (Life Technologies, UK) following the manufacturer's instructions. The reaction was conducted in the Stratagene Mx3005P qPCR machine (Agilent Technologies, USA). TaqMan assays (Life Technologies, UK) containing forward and reverse primers as well as hydrolysis probes were selected based on the criteria recommended by the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines for *PBX4* (Hs00257935_m1), *CDH1* (Hs01023894_m1), *CDH2* (Hs00983056_m1), *VIM* (Hs00185584_m1), *SNAI1* (Hs00195591_m1), *SNAI2* (Hs00161904_m1), *ZEB1* (Hs00232783_m1), *ZEB2* (Hs00207691_m1), *TWIST* (Hs00361186_m1), *VEGFA* (Hs00900055_m1) and *ACTB* (Hs01060665_g1). The relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method and *ACTB* was used as an endogenous control gene [16].

Western blotting: Cells were lysed in cell lysis buffer (Sigma, UK) supplemented with a protease inhibitor cocktail (Roche, UK) and were quantified using the Pierce™ Rapid Gold BCA Protein Assay Kit (Life Technologies, UK). Equal amounts of whole-cell lysates were separated using SDS-PAGE gels and mini-PROTEAN® electrophoresis apparatus (Bio-Rad, UK). Proteins were transferred onto polyvinylidene fluoride (PVDF) membranes and were incubated with

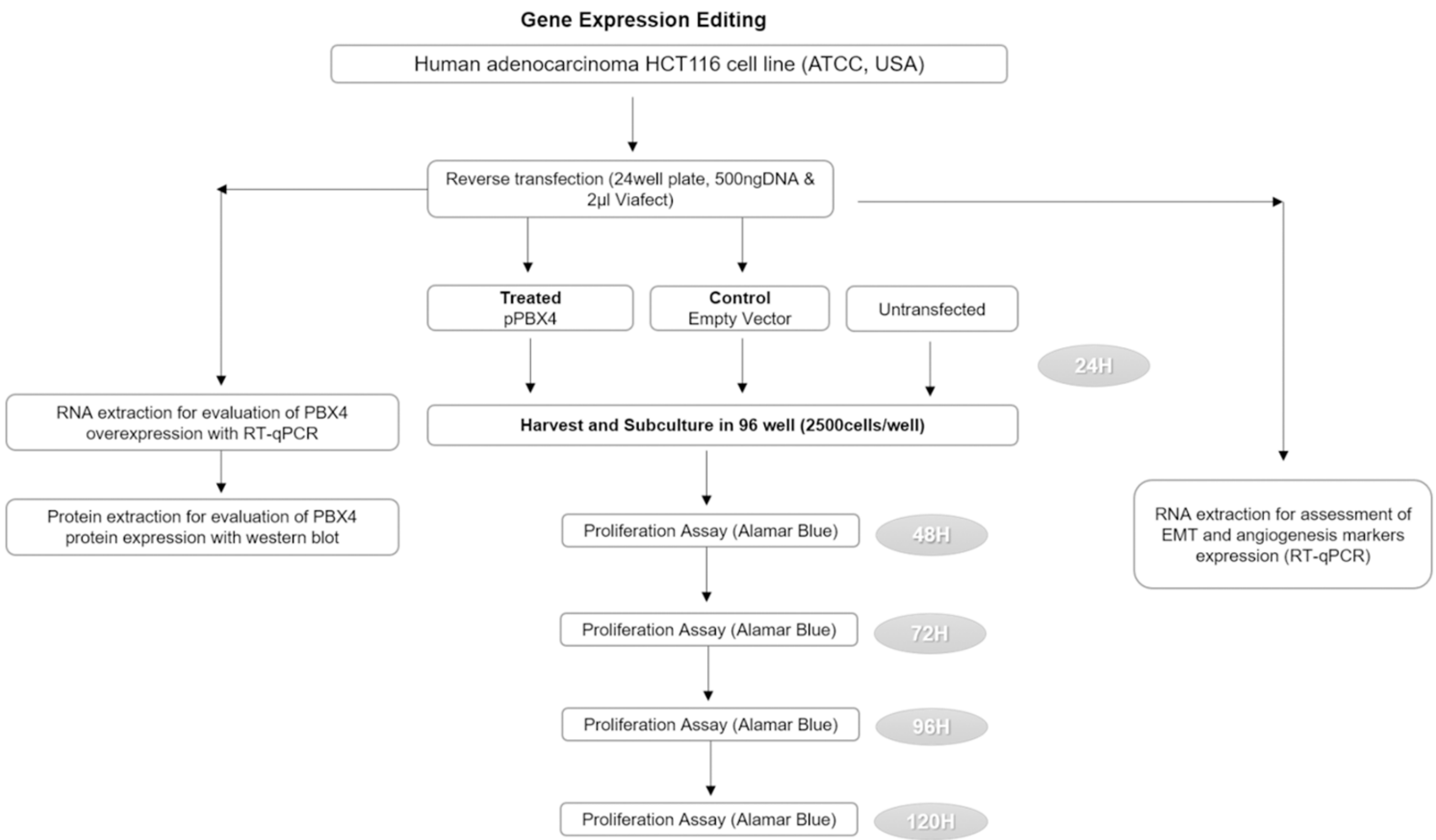


Figure 1. Study methodology flow chart of *in vitro* PBX4 overexpression in HCT116 cell line.

primary antibodies: anti-rabbit PBX4 1:500 (PA5-104032) (Life Technologies, UK) and anti-mouse β-actin 1:2000 (8H10D10) (Cell Signalling Technology), at 4°C overnight [17]. Detection was performed with IRDye 800CW goat anti-rabbit IgG 1:12000 (926-32211) and IRDye 680RD goat anti-mouse IgG 1:12000 (926-68070), (LI-COR Biosciences, UK). Membranes were imaged using the Odyssey CLx infrared system (LI-COR Biosciences, UK). The expected size of PBX4 protein according to the UniProt database was predicted to be 40,854 Da.

AlamarBlue proliferation assay: Cells were subcultured 24-48 h post-transfection and were seeded at 2500 cell density per well in 96 well plates containing 100 µL of medium per well and were placed for overnight incubation at 37°C at 5% CO_2. AlamarBlue (Life Technologies, UK) was added to each well at a 1:9 ratio, followed by incubation at 37°C for one hour. Fluorescent measurements (Relative Fluorescent Units, RFU) were performed on the SpectraMax i3 plate reader (Molecular Devices, UK) at 550 excitations and 590 emission wavelengths. Data were analysed by subtracting the background containing no cells from the overage fluorescence of each well-containing cells.

Cistrome data browser and gene set enrichment analysis (GSEA)

The Cistrome Data Browser (http://dbtoolkit.cistrome.org/) was used to identify factors that are predicted to bind the *PBX4* promoter and potentially regulate its transcription [18, 19]. Gene set enrichment analysis was performed in the Enrichr server (https://maayanlab.cloud/Enrichr/) to identify the biological processes and collective pathways of the included genes [20].

Results

Differential expression of PBX gene family in CRC

The differential transcriptional expression of *PBX1, PBX2, PBX3,* and *PBX4* genes between normal colon and CRC was analysed using the TCGA database. There were 644 primary solid tumours and 51 normal samples available in the combined TCGA COAD and READ datasets. Bioinformatics analysis showed that the RNA levels of *PBX1* (P<0.0001) and *PBX3* (P<0.0001) were significantly downregulated, whereas *PBX4* was the only gene that was significantly upregulated in CRC in comparison with NC (P<0.0001), as shown in **Figure 2A-E**. ONCOMINE platform data mining findings were in alignment with the bioinformatics analysis results. Specifically, most reporters showed that *PBX1* and *PBX3* genes are downregulated in CRC vs. NC whereas mixed findings exist for *PBX2*, as shown in **Figure 2F-H**. Nevertheless, all ONCOMINE reporters showed that *PBX4* seems to be upregulated in CRC tissue vs. NC, as shown in **Figure 2I**. The UALCAN web-server revealed that the methylation status of *PBX1* and *PBX3* promoters appear to be significantly increased in CRC vs. normal tissues (P=3.8e-04 and P=3.2e-07, respectively), whereas no difference was observed in the *PBX2* promoter methylation (P=0.070). On the contrary, the methylation status of the *PBX4* promoter was shown to be significantly decreased in CRC tissues (P=1.4e-07).

Association of PBX genes with OS in CRC

We investigated the clinical significance of the *PBX* gene family by exploring their association with OS in the TCGA COADREAD patient cohort (n=639). Since *PBX* genes are known as HOX co-factors and usually interact with *HOX* genes we added *HOX* genes in the survival analysis. All genes passed the filtering criteria except *HOXB1, HOXC12* and *HOXC13*. Univariable Cox regression model identified that age, presence of metastatic disease*, HOXA1, HOXC4, HOXC8, HOXC9, HOXC10, HOXC11, HOXD4, HOXD9, HOXD10, HOXD11*, *HOXD13* and *PBX4* gene were significantly related to OS rates in patients with CRC, as shown in **Table 1**. The multivariable analysis showed that high *PBX4* expression (HR 1.46, 95% CI: 1.14-1.88, P=0.003), as well as the presence of metastases (HR 4.9, 95% CI: 3.02-7.83, P=7.7E-11), were associated with the highest increased risk for worse OS rates in both models, as shown in **Table 1**. Kaplan-Meier analysis also showed that patients with high *PBX4* expression had worse OS rates in comparison with those with low expression (P=0.029).

Impact of PBX4 overexpression in CRC progression in vitro

To explore the potential oncopromoting role of *PBX4* we conducted mechanistic *in vitro* stud-

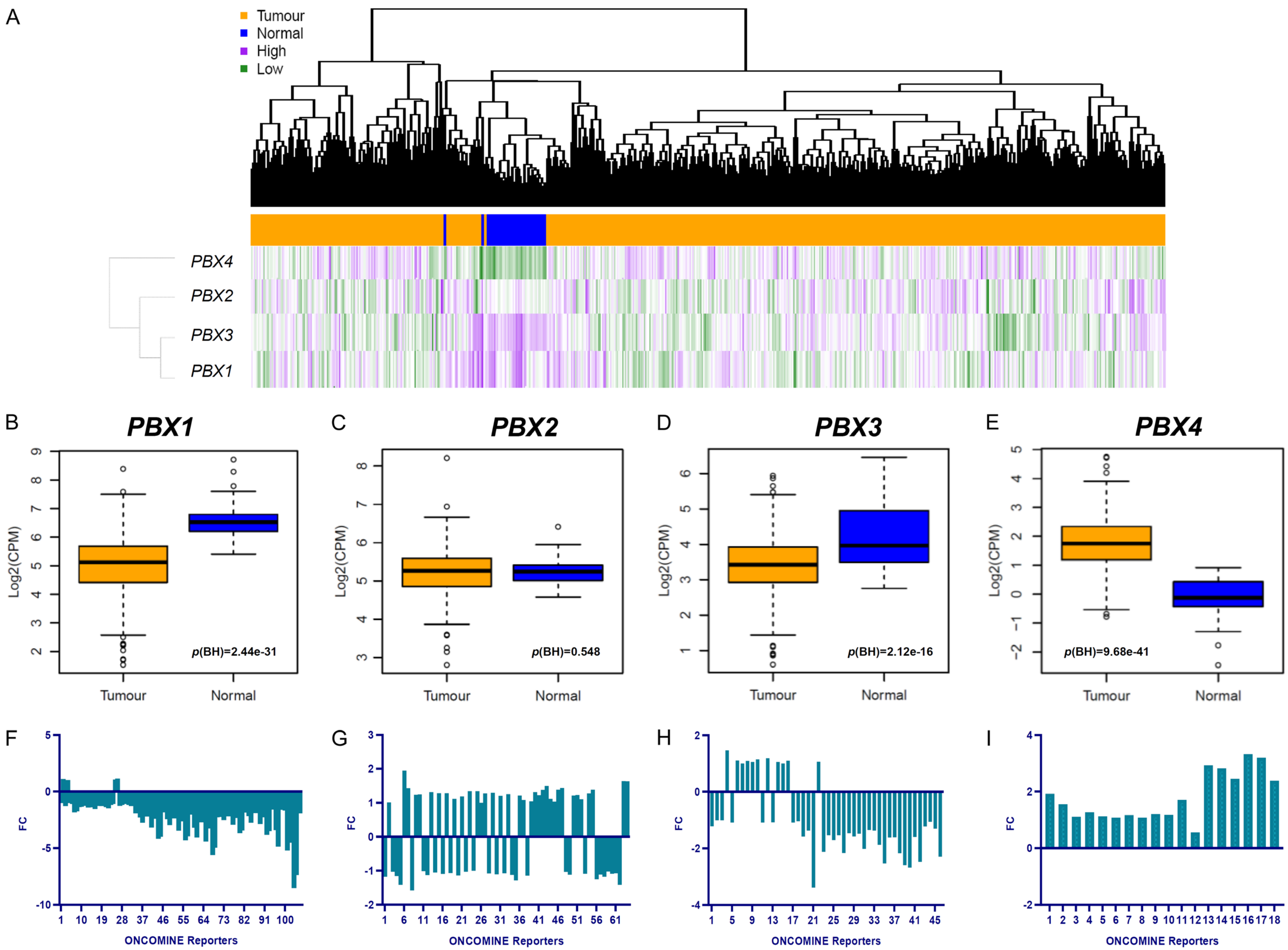


Figure 2. Differential expression of *PBX* genes in CRC vs. normal tissue. A. Heatmap of *PBX1, PBX2, PBX3 and PBX4* gene expression across TCGA COADREAD samples. B-E. Box plots of differential expression of *PBX1, PBX2, PBX3* and *PBX4* in CRC (n=644) vs. normal tissue (n=51) from TCGA COADREAD datasets. *Y*-axis represents the expression level as Log2counts per million (Log2CPM). F-I. Fold change (FC) of differential *PBX* genes expression in CRC vs. normal tissue using the ONCOMINE platform. Bars represent the findings from different reporters.

Table 1. Univariable and multivariable Cox survival analysis of PBX genes with HOX genes in the TCGA COADREAD patient cohort

Variable	Univariable Analysis HR (95% CI), *P*-value	Multivariable Model 1 HR (95% CI), *P*-value	Multivariable Model 2 HR (95% CI), *P*-value
Age	1.03 (1.01-1.04), 3.1E-4	1.03 (1.02-1.05), 1.5E-4	1.04 (1.02-1.05), 1.8E-4
Gender (M vs. F)	1.0 (0.73-1.5), 0.84		0.98 (0.65-1.49), 0.933
Location (right vs. left)	1.4 (0.93-2.0), 0.110		1.41 (0.85-2.35), 0.184
Metastases (yes vs. no)	3.8 (2.6-5.6), 9.5E-12	4.35 (2.85-6.64), 1.06E-11	4.9 (3.02-7.83), 7.7E-11
PBX1	1.1 (0.89-1.3), 0.530		
PBX2	1.3 (0.96-1.7), 0.094		
PBX3	1.1 (0.84-1.3), 0.660		
PBX4	1.3 (1-1.6), 0.020	1.34 (1.07-1.69), 0.012	1.46 (1.14-1.88), 0.003
HOXA1	1.1 (0.92-1.2), 0.40		
HOXA2	1 (0.93-1.2), 0.480		
HOXA3	1.1 (0.92-1.2), 0.420		
HOXA4	1.2 (1-1.4), 0.047	1.03 (0.87-1.22), 0.730	1.02 (0.85-1.23), 0.813
HOXA5	1 (0.93-1.2), 0.510		
HOXA6	1 (0.94-1.2), 0.410		
HOXA7	1 (0.93-1.1), 0.590		
HOXA9	1 (0.91-1.1), 1.0		
HOXA10	1 (0.89-1.1), 0.910		
HOXA11	1 (0.87-1.1), 0.970		
HOXA13	0.97 (0.9-1.1), 0.490		
HOXB2	1.1 (0.96-1.3), 0.180		
HOXB3	1.1 (0.92-1.2), 0.430		
HOXB4	1.1 (1-1.3), 0.0580		
HOXB5	1.1 (0.96-1.3), 0.160		
HOXB6	1.1 (0.95-1.2), 0.240		
HOXB7	1.1 (0.89-1.3), 0.540		
HOXB8	1.1 (0.96-1.2), 0.280		
HOXB9	1 (0.92-1.1), 0.620		
HOXB13	0.97 (0.9-1.1), 0.510		
HOXC4	1.2 (1.1-1.3), 0.00035	1.03 (0.83-1.27), 0.791	1.1 (0.86-1.39), 0.453
HOXC6	1.1 (1.1-1.2), 0.00015	1.15 (0.95-1.39), 0.165	1.1 (0.91-1.42), 0.279
HOXC8	1.2 (1.1-1.3), 0.0012	1.0 (0.76-1.31), 0.998	1.07 (0.79-1.44), 0.679
HOXC9	1.1 (1-1.3), 0.006	0.86 (0.67-1.12), 0.285	0.74 (0.56-0.99), 0.047
HOXC10	1.1 (1-1.2), 0.024	1.03 (0.89-1.19), 0.638	0.96 (0.82-1.13), 0.649
HOXC11	1.1 (1-1.1), 0.29	1.03 (0.94-1.12), 0.554	1.06 (0.96-1.17), 0.242
HOXD1	1.1 (0.98-1.2), 0.130		
HOXD3	1 (0.88-1.2), 0.880		
HOXD4	1.2 (1-1.4), 0.0084	1.06 (0.88-1.27), 0.554	1.06 (0.87-1.29), 0.557
HOXD8	1.1 (0.98-1.3), 0.098		
HOXD9	1.1 (1-1.3), 0.013	1.09 (0.89-1.32), 0.396	1.09 (0.88-1.36), 0.411
HOXD10	1.1 (1-1.2), 0.034	1.01 (0.79-1.29), 0.928	0.99 (0.76-1.28), 0.934
HOXD11	1.1 (1-1.2), 0.027	0.96 (0.74-1.25), 0.774	1.01 (0.76-1.34), 0.952
HOXD13	1.1 (1-1.2), 0.0072	1.07 (0.98-1.18), 0.149	1.05 (0.94-1.17), 0.397

ies in the HCT116 cell line which has low endogenous *PBX4* gene expression. Gain-of-function experiments showed that *PBX4* overexpression significantly increased cell prolifera-

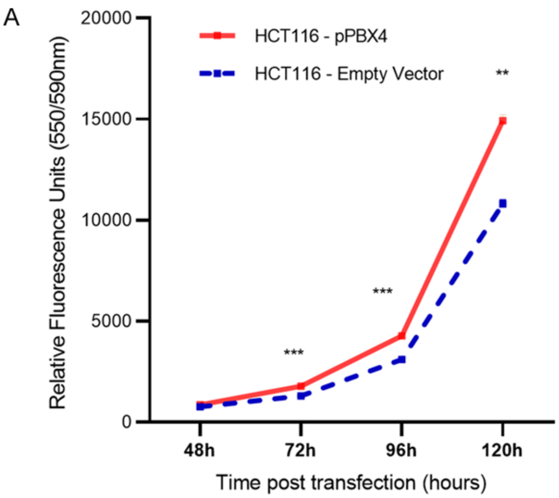
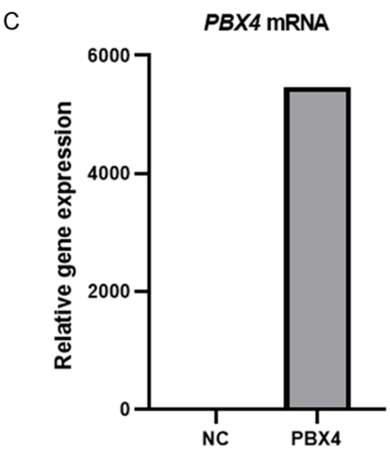
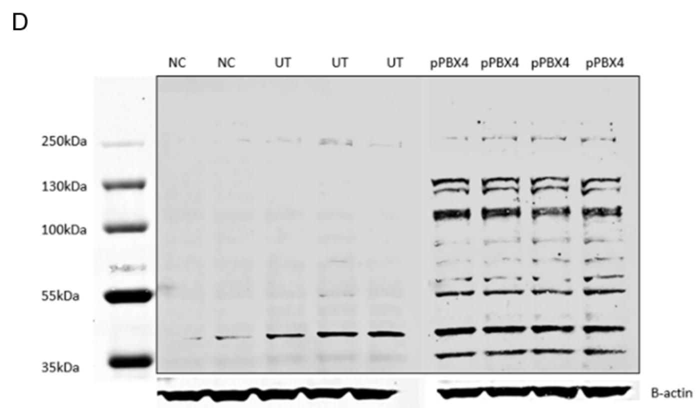


Figure 3. Impact of PBX4 gene overexpression on HCT116 cell proliferation in vitro. (A) HCT116 cell proliferation was measured as relative fluorescence (RFU) after PBX4 overexpression. Comparison groups were pPBX4 (overexpressing) and the control group which was transfected with an empty vector. Y-axis represents time points post-transfection. Data are presented as mean ± standard error of the mean (SEM) from 4 independent experiments containing 8-12 technical replicates and the values are shown in (B) alongside the *P*-values. (C) Relative PBX4 expression in HCT116 negative control group and in the overexpressing group from three independent experiments analysed in triplicates. (D) Western blot evaluation of PBX4 overexpression in HCT116 cells. NC represents the group transfected with an empty vector, UT represents the untransfected group and pPBX4 is the overexpressing group, The bottom band shows the expression of β-actin which was used as a loading control. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

B

	Time post transfection	HCT116-empty vector	HCT116-pPBX4	*P*-value
RFU (mean ± SEM), replicates (n)	48h	769.0±56.5, (n=56)	865.5±75.8, (n=52)	0.105
	72h	1303.4±90.7, (n=40)	1787.0±141.5, (n=40)	0.0005
	96h	3103.5±233.3, (n=54)	4275.0±337.4, (n=54)	0.0009
	120h	10808.1±620.5, (n=46)	14290.6±1078.8, (n=56)	0.0017





tion in the overexpressing group in comparison with the control group over 5 days post-transfection, as shown in **Figure 3**.

Relative gene expression of important EMT molecules and the angiogenesis marker VEGFA was compared between the *PBX4*-overexpressing group and the control group. The EMT molecular markers *VIM* (P<0.0001), *CDH1* (P≤0.0001), *CDH2* (P=0.0002), *ZEB1* (P=0.007) and *SNAI1* (P=0.001) were significantly upregulated when the *PBX4* gene was overexpressed, as shown in **Figure 4**. Regarding ZEB2 (P=0.105) and TWIST (P=0.129), no difference was observed between the two groups. On the contrary *SNAI2* (P=0.022) was found to be downregulated. *PBX4* overexpression was additionally found to markedly upregulate the expression of the *VEGFA* angiogenic marker (P<0.0001) *in vitro.*

In silico transcriptional regulation prediction of PBX4 and related biological processes

Overall, the Cistrome DB Toolkit identified 111 factors that potentially bind the promoter of

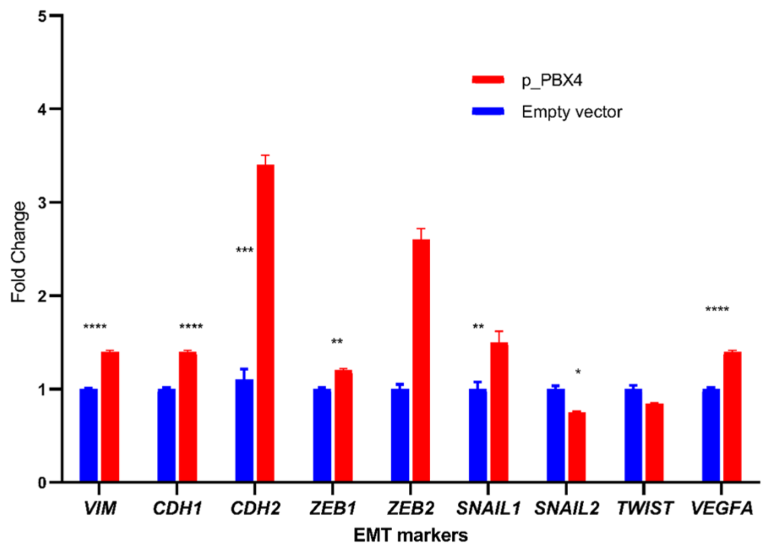


Figure 4. Impact of *PBX4* overexpression on EMT and angiogenesis markers *in vitro*. RNA fold change expression of EMT-related and angiogenesis markers in HCT116 overexpressing PBX4 vs. empty vector (control) group. Data are presented as mean ± SEM from three independent experiments analysed in triplicates. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. VIM: Vimentin; CDH1: Cadherin-1; CDH2: Cadherin-2; ZEB1: Zinc Finger E-Box Binding Homeobox 1; ZEB2: Zinc Finger E-Box Binding Homeobox 2; SNAI1: Snail Family Transcriptional Repressor 1; SNAI2: Snail Family Transcriptional Repressor 2; TWIST: Twist Family BHLH Transcription Factor 1; VEGFA: Vascular Endothelial Growth Factor A.

PBX4 and are predicted to regulate its expression. Nineteen genes (*IRF1, JUN, JUND, CEBPB, E2F3, SP1, RAD21, MED12, NR1H3, HEXIM1, MAX, POLR2A, PAF1, CDK8, SMC3, H2AZ, FOSL1, HOXA4* and *TCF7L2*) were identified in the Cistrome Chromatin Immunoprecipitation sequencing (ChIP-seq) experimental datasets to bind the *PBX4* promoter in CRC, as shown in **Table 2**. GSEA of these 19 factors showed that biological processes related to the regulation of cell cycle and cell proliferation were enriched with GO:0051726 (*CDK8, JUN, JUND,* and *IRF1*) being highly significant (P=0.001). Additionally, amongst the related predicted pathways, the Activator Protein-1 (AP-1) transcription network (*FOSL1, JUN, JUND, TCFL2,* and *SP1*) was found to be the most highly enriched (P<0.0001). A full interactive GSEA analysis in the Enrichr server is available through the following link: https://maayanlab.cloud/Enrichr/enrich?dataset=3bb2aca99def224443639fbba1b34389.

Enrichment analysis was expanded to include all the identified predicted factors (n=111) that bind the *PBX4* promoter. Pathways related to cell cycle regulation (G2/M, G1/S pathway, P<0.01) and proliferation such as the AP-1 network (P<0.0001) were significantly enriched. Additionally, pathways related to CRC such as the MYC pathway (P<0.0001), β-catenin (P<0.0001), and WNT pathway (P<0.0001) were also markedly enriched. A full interactive GSEA analysis in the Enrichr server is available through the following link: https://maayanlab.cloud/Enrichr/enrich?dataset=32108d2a9ff52c53b3cd948328a7bac8.

Discussion

Since the identification of the chimeric fusion oncoprotein E2A-PBX1 that promotes the progression of pre-B cell acute lymphoblastic leukaemia, emerging evidence suggests that the *PBX* gene family become dysregulated and plays an important role in the development and progression of many human cancers [21]. However, the expression profile of *PBX* genes in CRC remains unknown. Herein, in our study using TCGA and ONCOMINE analysis, we found that *PBX1* and *PBX3* appear to be significantly downregulated, whereas *PBX4* is the only gene that was found to be upregulated in CRC vs. normal tissue. The mechanism of how *PBX* genes become dysregulated in CRC is unknown. As CRC evolves through genetic and epigenetic changes, with aberrant DNA methylation being characterised as a cancer hallmark, we further explored the *PBX* gene methylation promoter status in cancer vs. normal tissue [22]. Through the UALCAN database, we observed, that the promoter methylation status of the *PBX* gene family appears to align with its dysregulation profile. Aberrant DNA methylation leading to gene promoter hypo or hypermethylation has been identified as an important regulatory transcriptional mechanism of *HOX* gene expression [22]. Since *PBX* genes are largely known as HOX cofactors, as they interact with *HOX* genes, it could be hypothesised that *PBX* genes may become dysregulated through DNA methylation. However, further studies are needed to explore this hypothesis.

Table 2. Putative PBX4 regulators from Cistrome DB

Factor	Biosource	RPS	Description (https://www.genecards.org/)
IRF1	LoVo	0.79	Interferon Regulatory Factor 1. Tumour suppressor involved in cell proliferation, apoptosis, immune response, and DNA damage response
JUN	LoVo	0.79	JUN proto-oncogene. It encodes a protein which interact with specific DNA sequences to regulate gene expression
CEBPB	LS180	0.7	CCAAT Enhancer Binding Protein Beta. Transcription factor important in the regulation of genes involved in immune and inflammatory responses. It is associated is Juvenile Polyposis
JUND	HCT116	0.62	JunD Proto-Oncogene. The protein encoded is a member of the JUN family and has been proposed to protect cells from p53-dependent senescence and apoptosis
E2F3	LoVo	0.61	E2F transcription factor 3. The encoded protein is a member of a small family of transcription factors that recognises a specific sequence motif in DNA and interacts directly with the retinoblastoma protein (pRB) to regulate the expression o of genes involved in the cell cycle
SP1	HCT116	0.57	Sp1 Transcription Factor. The encoded protein is involved in many cellular processes, including cell differentiation, cell growth, apoptosis, immune responses, response to DNA damage, and chromatin remodelling
RAD21	HCT116	0.56	RAD21 Cohesin Complex Component. The encoded protein is a nuclear phospho-protein, which becomes hyperphosphorylated in cell cycle M phase. RAD21 is involved in the repair of DNA double-strand breaks, as well as in chromatid cohesion during mitosis
MED12	LoVo	0.56	Mediator Complex Subunit 12. The encoded protein binds with CDK8 and other proteins involved in cell cycle including cyclin C. This subunit may regulate transcription targets of the Wnt signalling pathway
POLR2A	HCT116, DLD1, HT29	0.56	RNA Polymerase II Subunit A. This gene encodes the largest subunit of RNA polymerase II, the polymerase responsible for synthesizing messenger RNA in eukaryotes
NR1H3	HT29	0.55	Nuclear Receptor Subfamily 1 Group H Member 3. The protein encoded by this gene belongs to the NR1 subfamily of the nuclear receptor superfamily. The NR1 family members are key regulators of macrophage function, controlling transcriptional programs involved in lipid homeostasis and inflammation
HEXIM1	HCT116	0.54	HEXIM P-TEFb Complex Subunit 1. HEXIM1 is transcriptional regulator which functions as a general RNA polymerase II transcription inhibitor
MAX	HCT116	0.54	MYC Associated Factor X. The encoded protein is able to form homodimers and heterodimers with other family members, which include MAD and MYC, implicated in cell proliferation, differentiation and apoptosis
PAF1	HCT116, DLD1	0.53	Polymerase Associated Factor 1. This gene encodes a subunit of the polymerase associated factor (PAF1) complex. The PAF1 complex interacts with RNA polymerase II and plays a role in transcription elongation as well as histone modifications. PAF1C is required for transcription of HOX and WNT target genes
CDK8	Colon	0.53	Cyclin Dependent Kinase 8. This gene encodes a member of the cyclin-dependent protein kinase (CDK) family. CDK family members are known to be important regulators of cell cycle progression
SMC3	HCT116	0.53	Structural Maintenance of Chromosome 3. The encoded protein is a component of the multimeric cohesin complex that holds together sister chromatids during mitosis, enabling proper chromosome segregation
H2AZ	HCT116	0.51	Variant histone H2A. The encoded protein plays a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability
FOSL-1	HCT116	0.51	FOS Like 1, AP-1 Transcription Factor Subunit. The Fos gene family consists of 4 members: FOS, FOSB, FOSL1, and FOSL2. These genes encode leucine zipper proteins that can dimerize with proteins of the JUN family, thereby forming the transcription factor complex AP-1. As such, the FOS proteins have been implicated as regulators of cell proliferation and differentiation
HOXA4	LoVo	0.51	Homeobox A4
TCF7L2	HCT116	0.51	Transcription Factor 7 Like 2, This gene encodes a high mobility group (HMG) box-containing transcription factor that plays a key role in the Wnt signalling pathway

The clinical significance of *PBX* dysregulation is still unknown and few studies have investigated the association of *PBX* dysregulation with OS rates in cancer. For instance, *PBX1* was found to be upregulated in breast cancer and it was shown that high expression levels were correlated with worse survival in patients with oestrogen-receptor (ER)-positive, luminal A and luminal B subtypes of breast cancer [9]. Similarly, high *PBX2* expression appears to be related to a worse prognosis in gastric cancer and oesophageal squamous cell carcinoma [23]. Studies regarding CRC have focused on *PBX3,* showing that protein overexpression is associated with advanced disease stage and worse survival rates in CRC patients [8]. To the best of our knowledge, no previous study has comprehensively analysed the impact of *PBX* genes dysregulation in OS. For the first time, our study explored the *PBX* gene association with OS in the TCGA COADREAD patient cohort, including all *PBX* and *HOX* genes in our univariable and multivariable analysis. Interestingly, the multivariable survival analysis showed that only high *PBX4* expression levels appeared to be associated with an increased risk for worse OS in patients with CRC in both models. This association remained significant after being adjusted for the clinicopathological variable and *HOX* genes. Advanced age, as well as the presence of metastatic disease, is well-established prognostic factors in patients with CRC [24, 25]. The fact that in our study, high *PBX4* levels appeared to potentially increase the likelihood of worse OS, highlights the importance of *PBX4* as a prognostic risk factor. Additionally, the survival analysis findings suggest that *PBX4* may play an oncopromoting role in CRC.

We aimed to further explore our hypothesis on the potential role of *PBX4* as an oncopromoter in CRC by conducting *in vitro* mechanistic studies. We selected as the main outcome, the cell proliferation *in vitro* and as a secondary aim, we evaluated the changes in gene expression of EMT and angiogenesis markers. Our rationale for choosing these outcomes is based on the revisited Hallmarks II theory of cancer by Fouad *et al*. The Hallmarks II theory proposes seven cancer hallmarks: proliferative advantage, altered stress response favouring overall survival of tumour cells, vascularisation, invasion and metastasis, metabolic rewiring, an abetting tumour microenvironment (TME) and immune modulation. Increased proliferation, activation of invasion/metastasis and vascularisation are characterised as the cornerstones of cancer progression and are present in every hallmark of cancer theory [26-28]. Therefore, cell proliferation was selected as the main outcome and gene expression of EMT and angiogenesis markers were also evaluated as additional outcomes to explore our hypothesis on *PBX4* as a potential promoter of CRC progression *in vitro*.

We have shown that *PBX4* overexpression significantly increased cell proliferation in HCT116 cells. This study firstly reports the potential role of *PBX4* in cancer cell proliferation ability which is an important hallmark in cancer. Among the *PBX* gene family, studies regarding CRC have focused on *PBX3* and have reported that it may increase cell proliferation by regulating the expression of p53 through the suppression of its promoter activity [29]. The mechanism through which *PBX4* may promote cell proliferation remains unknown. Given the lack of relevant information in the literature, to shed more light on the potential role of *PBX4* in CRC we aimed to identify factors that bind to the promoter of *PBX4* and are predicted to regulate its transcription from an established portal (Cistrome) containing ChiP-seq experimental information. We speculated that *PBX4* is likely to be involved in the same pathways as those of its regulators. Interestingly, among the most highly predicted *PBX4* regulators were proteins involved in the AP-1 transcription network such as *JUN* and *JUND,* which are known to be linked with cellular growth control and malignant transformation [27]. There is evidence that the JUN and JUND oncoproteins contribute to tumorigenesis and progression in CRC and that targeting the AP-1 components has therapeutic potential in CRC patients [30, 31]. Our GSEA included all putative *PBX4* regulators and revealed that biological processes and pathways related to cell cycle regulation and proliferation were enriched, suggesting that *PBX4* could affect cell growth by being an important molecular component or the target gene of these pathways.

Post-translational modification [PTM] could be theoretically another potential mechanism with which PBX4 may be associated with cell proliferation. In cancer cells, oncogenes produce proliferative signals by adjusting the state of PTMs of proteins that are involved in the regulation of cell cycle and growth [32]. It is recognised that HOX transcription factors undergo significant PTMs and since PBX proteins act as HOX co-factors, they may be also subjects to multiple forms of PTMs. A recent review by Reichlmeir *et al*. reported that PBX proteins are subjected to several PTMs such as phosphorylation, ubiquitination, and methylation with the former being the most common PTM [33, 34]. Specifically for PBX4, phosphorylation appeared to be the most common PTM. Several studies have identified phosphorylated tyrosine, serine and threonine residues in the PBX family raising the possibility that PBX proteins could be substrates of pro-oncogenic kinases which play an important role in the progression of various tumours [33, 35]. Our western blot showed multiple bands which could raise the possibility of the presence of PTMs in PBX4 overexpressing protein [36]. Additionally, ur *in silico* analysis identified cyclin-dependent kinase 8 (CDK8), which is recognised as an oncoprotein and cell cycle regulator in CRC, as a putative PBX4 regulator. CDKs have been found to be important for efficient PTM of human proteins and especially phosphorylation which may affect the cell cycle [37, 38]. Nevertheless, there is no current direct experimental evidence on the mechanistic explanation of how PBX4 protein may promote cell proliferation by undergoing PTMs through protein-dependent kinases which highlights the need for further research to explore if PBX4 protein is subjected to PTMs and the potential impact of these PTMs on its functionality and the subsequent effect on CRC proliferation.

In furtherance to the observation of several protein bands, it is worth highlighting the potential proteomic diversity which may occur as a result of alternative splicing. Despite that little is known regarding the *PBX4* gene, it has been found that it contains 5 exons and alternatively, spliced variants have been observed according to the Ensemble genome browser database [39]. A recent systematic review by Zhang *et al*. provided cumulative evidence on the potential role of alternative splicing in the development of cancer and its progression [40]. Alternative splicing may regulate cell proliferation through the formation of different protein isoforms [41]. For instance splicing of the epidermal growth factor 2 has been found to promote cell proliferation whereas its variant Δ16HER2 was found to inhibit cell growth in breast cancer [41]. Stadler *et al*. showed that *HOXA9* which is a well-known oncogene in haematopoietic malignancies relies on alternative splicing to mediate leukemogenesis [42]. Alternative splicing could only be speculated as a potential mechanism of the *PBX4* gene to promote cell proliferation as no relevant data are available from either the literature or our study. Nevertheless, since alternative splicing may play an important impact in the development and progression of cancer its role in *PBX4* is worth being investigated further.

EMT is a hallmark in cancer and a pivotal step in the natural history of metastases in CRC. EMT is a transcriptional programme that leads cancer cells to show less epithelial differentiation in favour of a mesenchymal morphology enabling them to gain motility, invade adjacent tissues and metastasise [43]. This programme is regulated by important transcription factors, such as the ZEB and SNAIL family and is characterised by a decrease in the epithelial marker E-cadherin with a subsequent increase in mesenchymal markers Vimentin and N-cadherin known as the “cadherin switch” [43]. Studies have shown that *PBX3* can promote cancer invasion and metastases by inducing EMT through its regulatory pathways such as WNT, MAPK and AKT signalling, causing the upregulation of EMT transcription factors [44, 45]. However, no studies have examined the role of *PBX4* in EMT to date. We noted that *PBX4* overexpression upregulated important EMT molecules and increased the gene expression encoding for the mesenchymal markers N-cadherin and Vimentin. Although the downregulation of *CDH2* was not observed as one would expect in the case of a “cadherin switch”, important EMT activators were upregulated, indicating that *PBX4* may be important in the EMT process. Furthermore, it is worth highlighting that EMT and its opposite process, mesenchymal-epithelial transition (MET) are not clear binary processes but instead, they may co-exist in cancer cells and potentially may enable them to exhibit more aggressive behaviour [43]. An exciting additional finding was that we noticed a marked upregulation of the angiogenic marker *VEGFA* in the *PBX4* overexpressing cell group, indicating a potentially novel role of *PBX4* in promoting angiogenesis in CRC. Our *in vitro* finding was in agreement with a complementary *in silico* correlation analysis based on the TCGA COADREAD datasets (data not shown). No relevant information exists on the potential role of *PBX4* in activating angiogenesis through *VEGFA*. However, considering the fact that *HOX* genes have been found to promote angiogenesis in cancer it would be reasonable to hypothesise that their PBX co-factors may also play an important role in angiogenesis promotion [34]. Indeed, further research is warranted to explore this research field.

The strength of this study is that it provides the first investigation of *PBX* genes in colorectal cancer. The prognostic role of the *PBX4* gene in CRC has not been previously reported and no data exist with regard to its role, not only in CRC but also in other types of cancers. Our study is novel; however, it has limitations that should be considered in the context of the interpretation of its findings. Firstly, to explore the differential expression of the *PBX* family we performed cancer to normal analysis at an mRNA level. As changes at the RNA level do not always correspond to a phenotypic change, a differential expression at a protein level is indeed important. However, to overcome this limitation and obtain more evidence on their dysregulation profile we expanded our initial TCGA analysis by including ONCOMINE analysis from all reporters as well as by looking at the *PBX* gene promoter methylation status. Secondly, it is known that variables including Body Mass Index, lymph node status and tumour depth are also important prognostic factors in CRC. However, these data were not available in the TCGA to be included in the analysis. Nevertheless, our findings were adjusted for age, tumour location, gender and metastases and we also considered the impact of *HOX* genes. Another limitation of our study is that our *in vitro* results derive from only one cell line [HCT116]. However, in our methodology-optimisation period, we have explored 2 other cell lines [HT29 and SW620] and we found that both CRC cell lines were of a limited transfection efficiency which made these cell lines unsuitable for *in vitro* gene expression modulation experiments. It should be noted that based on our literature search and *in silico* analysis, the HCT116 is the most widely used CRC cell line amongst studies [29, 44, 46]. An additional limitation of our study is that only the changes in RNA levels of EMT and angiomarkers were reported. To obtain more evidence into the role of PBX4 in the ability of CRC cells to metastasise, the phenotypic changes need to be explored in the context of protein analysis as well as with invasion and migration functional assays [15]. In the present study, we have selected cell proliferation as our main outcome while we evaluated the changes in gene expression of EMT and angiogenesis markers as a secondary analysis. We found our initial results from our supplementary analysis interesting, and we believe that they enhance the hypothesis that *PBX4* potentially may have oncopromoting properties. It is worth investigating this field further with additional research to explore further the potential impact of PBX4 in the CRC metastatic process. Nevertheless, our study provides early evidence and sets an important foundation for future investigation on the role of *PBX4* in CRC progression and highlights that further research is needed to obtain more evidence on the proposed oncopromoting role of *PBX4* in CRC. Lastly, we investigated putative regulators of *PBX4* to explore its potential involvement in signalling pathways. Investigation of the target genes through chromatin immunoprecipitation experiments could further reveal the underlying biological processes and pathways in which *PBX4* is involved.

In conclusion, this study investigates the *PBX* gene family in CRC for the first time and suggests that *PBX4* may be a potential novel promoter of CRC progression and is associated with worse survival. This study highlights the need to further explore the specific role of *PBX4* in the hallmarks of CRC.

Acknowledgements

We thank Dr. Aikaterini Chatzipli, Department of Medical Informatics, Harvard Medical School, USA, for providing scientific advice on mechanistic experimental studies. We thank the staff at the molecular laboratory, Royal Surrey County Hospital, UK for their scientific advice on molecular analysis. We thank the scientific personnel of the University of Surrey, UK and Imperial Healthcare Tissue Bank, UK for the technical and material help. We thank the BRIGHT Cancer Care Charity and its trustees for arranging to fund this research. We thank the Royal Society of Medicine Oncology department for awarding our research with the Sylvia Lawler Prize.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Eirini G Martinou, Hepatobiliary and Pancreatic Surgery Department, Royal Surrey County Hospital, Egerton Road, Guildford GU2 7XX, UK. E-mail: eirini.martinou@nhs.net; Dr. Angeliki M Angelidi, Department of Medicine, Beth Israel Deaconess Medical Centre, Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA. E-mail: aangelid@broadinstitute.org

References

[1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020; 70: 7-30.
[2] Avolio M and Trusolino L. Rational treatment of metastatic colorectal cancer: a reverse tale of men, mice, and culture dishes. Cancer Discov 2021; 11: 1644-1660.
[3] Singh MP, Rai S, Pandey A, Singh NK and Srivastava S. Molecular subtypes of colorectal cancer: an emerging therapeutic opportunity for personalized medicine. Genes Dis 2019; 8: 133-145.
[4] Morgan R, El-Tanani M, Hunter KD, Harrington KJ and Pandha HS. Targeting HOX/PBX dimers in cancer. Oncotarget 2017; 8: 32322-32331.
[5] Li Z, Zhang Z, Li Y, Arnovitz S, Chen P, Huang H, Jiang X, Hong GM, Kunjamma RB, Ren H, He C, Wang CZ, Elkahloun AG, Valk PJ, Döhner K, Neilly MB, Bullinger L, Delwel R, Löwenberg B, Liu PP, Morgan R, Rowley JD, Yuan CS and Chen J. PBX3 is an important cofactor of HOXA9 in leukemogenesis. Blood 2013; 121: 1422-1431.
[6] Selleri L, Zappavigna V and Ferretti E. 'Building a perfect body': control of vertebrate organogenesis by PBX-dependent regulatory networks. Genes Dev 2019; 33: 258-275.
[7] Paço A, Aparecida de Bessa Garcia S, Leitão Castro J, Costa-Pinto AR and Freitas R. Roles of the hox proteins in cancer invasion and metastasis. Cancers (Basel) 2021; 13: 10.
[8] Morgan R and Pandha HS. PBX3 in cancer. Cancers (Basel) 2020; 12: 431.
[9] Ao X, Ding W, Zhang Y, Ding D and Liu Y. PBX1 is a valuable prognostic biomarker for patients with breast cancer. Exp Ther Med 2020; 20: 385-394.
[10] Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013; 45: 1113-1120.
[11] Robinson MD, McCarthy DJ and Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2009; 26: 139-140.
[12] Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995; 57: 289-300.
[13] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[14] Geraghty RJ, Capes-Davis A, Davis JM, Downward J, Freshney RI, Knezevic I, Lovell-Badge R, Masters JR, Meredith J, Stacey GN, Thraves P and Vias M; Cancer Research UK. Guidelines for the use of cell lines in biomedical research. Br J Cancer 2014; 111: 1021-1046.
[15] Menyhárt O, Harami-Papp H, Sukumar S, Schäfer R, Magnani L, de Barrios O and Győrffy B. Guidelines for the selection of functional assays to evaluate the hallmarks of cancer. Biochim Biophys Acta 2016; 1866: 300-319.
[16] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J and Wittwer CT. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009; 55: 611-622.
[17] Mahmood T and Yang PC. Western blot: technique, theory, and trouble shooting. N Am J Med Sci 2012; 4: 429-434.
[18] Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, Chen CH, Brown M, Zhang X, Meyer CA and Liu XS. Cistrome data browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res 2019; 47: D729-D735.
[19] Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG and Ma'ayan A. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) 2016; 2016: baw100.
[20] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW and Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44: W90-97.
[21] Lu Q, Wright DD and Kamps MP. Fusion with E2A converts the Pbx1 homeodomain protein into a constitutive transcriptional activator in human leukemias carrying the t(1;19) translocation. Mol Cell Biol 1994; 14: 3938-3948.

[22] Paço A, de Bessa Garcia SA and Freitas R. Methylation in HOX clusters and its applications in cancer therapy. Cells 2020; 9: 1613.

[23] Qiu Y, Song B, Zhao G, Deng B, Makino T, Tomita Y, Wang J, Luo W, Doki Y, Aozasa K and Morii E. Expression level of Pre B cell leukemia homeobox 2 correlates with poor prognosis of gastric adenocarcinoma and esophageal squamous cell carcinoma. Int J Oncol 2010; 36: 651-663.

[24] Liu Z, Xu Y, Xu G, Baklaushev VP, Chekhonin VP, Peltzer K, Ma W, Wang X, Wang G and Zhang C. Nomogram for predicting overall survival in colorectal cancer with distant metastasis. BMC Gastroenterol 2021; 21: 103.

[25] Yusup A, Wang HJ, Rahmutula A, Sayim P, Zhao ZL and Zhang GQ. Clinical features and prognosis in colorectal cancer patients with different ethnicities in Northwest China. World J Gastroenterol 2013; 19: 7183-7188.

[26] Fouad YA and Aanei C. Revisiting the hallmarks of cancer. Am J Cancer Res 2017; 7: 1016-1036.

[27] Hanahan D and Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144: 646-674.

[28] Hanahan D and Weinberg RA. The hallmarks of cancer. Cell 2000; 100: 57-70.

[29] Li WF, Herkilini A, Tang Y, Huang P, Song GB, Miyagishi M, Kasim V and Wu SR. The transcription factor PBX3 promotes tumor cell growth through transcriptional suppression of the tumor suppressor p53. Acta Pharmacol Sin 2021; 42: 1888-1899.

[30] Ashida R, Tominaga K, Sasaki E, Watanabe T, Fujiwara Y, Oshitani N, Higuchi K, Mitsuyama S, Iwao H and Arakawa T. AP-1 and colorectal cancer. Inflammopharmacology 2005; 13: 113-125.

[31] Xu H, Liu L, Li W, Zou D, Yu J, Wang L and Wong CC. Transcription factors in colorectal cancer: molecular mechanism and therapeutic implications. Oncogene 2021; 40: 1555-1569.

[32] Hitosugi T and Chen J. Post-translational modifications and the Warburg effect. Oncogene 2014; 33: 4279-4285.

[33] Reichlmeir M, Elias L and Schulte D. Posttranslational modifications in conserved transcription factors: a survey of the tale-homeodomain superclass in human and mouse. Front Cell Dev Biol 2021; 9: 648765.

[34] Yu M, Zhan J and Zhang H. HOX family transcription factors: related signaling pathways and post-translational modifications in cancer. Cell Signal 2020; 66: 109469.

[35] Primon M, Hunter KD, Pandha HS and Morgan R. Kinase regulation of hox transcription factors. Cancers (Basel) 2019; 11: 508.

[36] Bass JJ, Wilkinson DJ, Rankin D, Phillips BE, Szewczyk NJ, Smith K and Atherton PJ. An overview of technical considerations for Western blotting applications to physiological research. Scand J Med Sci Sports 2017; 27: 4-25.

[37] Sanchez V and Spector DH. Cyclin-dependent kinase activity is required for efficient expression and posttranslational modification of human cytomegalovirus proteins and for production of extracellular particles. J Virol 2006; 80: 5886-5896.

[38] Cuijpers SAG and Vertegaal ACO. Guiding mitotic progression by crosstalk between post-translational modifications. Trends Biochem Sci 2018; 43: 251-268.

[39] Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, Garcia Giron C, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Gonzalez Martinez J, Marugán JC, Maurel T, McMahon AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Salam AIA, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, De Silva N, Flint B, Frankish A, Hunt SE, Ilsley GR, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR and Flicek P. Ensembl 2021. Nucleic Acids Res 2021; 49: D884-D891.

[40] Zhang Y, Qian J, Gu C and Yang Y. Alternative splicing and cancer: a systematic review. Signal Transduct Target Ther 2021; 6: 78.

[41] Ouyang J, Zhang Y, Xiong F, Zhang S, Gong Z, Yan Q, He Y, Zhang W, Zhou M, Xiang B, Wang F, Li X, Li Y, Li GY, Zeng Z, Guo C and Xiong W. The role of alternative splicing in human cancer progression. Am J Cancer Res 2021; 11: 4642-4667.

[42] Stadler CR, Vegi N, Mulaw MA, Edmaier KE, Rawat VP, Dolnik A, Bullinger L, Heilmeier B, Quintanilla-Fend L, Spiekermann K, Hiddemann W, Döhner K, Döhner H, Feuring-Buske M and Buske C. The leukemogenicity of Hoxa9 depends on alternative splicing. Leukemia 2014; 28: 1838-1843.

[43] Loh CY, Chai JY, Tang TF, Wong WF, Sethi G, Shanmugam MK, Chong PP and Looi CY. The E-cadherin and N-cadherin switch in epithelial-to-mesenchymal transition: signaling, therapeutic implications, and challenges. Cells 2019; 8: 1118.

[44] Lamprecht S, Kaller M, Schmidt EM, Blaj C, Schiergens TS, Engel J, Jung A, Hermeking H, Grünewald TGP, Kirchner T and Horst D. PBX3

is part of an EMT regulatory network and indicates poor outcome in colorectal cancer. Clin Cancer Res 2018; 24: 1974-1986.
[45] Wang S, Li C, Wang W and Xing C. PBX3 promotes gastric cancer invasion and metastasis by inducing epithelial-mesenchymal transition. Oncol Lett 2016; 12: 3485-3491.
[46] Martinou E, Falgari G, Bagwan I and Angelidi AM. A systematic review on HOX genes as potential biomarkers in colorectal cancer: an emerging role of HOXB9. Int J Mol Sci 2021; 22: 13429.