Wiltshire *et al. Human Genomics* (2024) 18:24
https://doi.org/10.1186/s40246-024-00592-x

Human Genomics

RESEARCH Open Access

Cellular and clinical impact of protein phosphatase enzyme epigenetic silencing in multiple cancer tissues

Edward Wiltshire[1], Manuel Castro de Moura[2], David Piñeyro[2] and Ricky S. Joshi[1*]

Abstract

Background Protein Phosphatase Enzymes (PPE) and protein kinases simultaneously control phosphorylation mechanisms that tightly regulate intracellular signalling pathways and stimulate cellular responses. In human malignancies, PPE and protein kinases are frequently mutated resulting in uncontrolled kinase activity and PPE suppression, leading to cell proliferation, migration and resistance to anti-cancer therapies. Cancer associated DNA hypermethylation at PPE promoters gives rise to transcriptional silencing (epimutations) and is a hallmark of cancer. Despite recent advances in sequencing technologies, data availability and computational capabilities, only a fraction of PPE have been reported as transcriptionally inactive as a consequence of epimutations.

Methods In this study, we examined promoter-associated DNA methylation profiles in Protein Phosphatase Enzymes and their Interacting Proteins (PPEIP) in a cohort of 705 cancer patients in five tissues (Large intestine, Oesophagus, Lung, Pancreas and Stomach) in three cell models (primary tumours, cancer cell lines and 3D embedded cancer cell cultures). As a subset of PPEIP are known tumour suppressor genes, we analysed the impact of PPEIP promoter hypermethylation marks on gene expression, cellular networks and in a clinical setting.

Results Here, we report epimutations in PPEIP are a frequent occurrence in the cancer genome and manifest independent of transcriptional activity. We observed that different tumours have varying susceptibility to epimutations and identify specific cellular signalling networks that are primarily affected by epimutations. Additionally, RNA-seq analysis showed the negative impact of epimutations on most (not all) Protein Tyrosine Phosphatase transcription. Finally, we detected novel clinical biomarkers that inform on patient mortality and anti-cancer treatment sensitivity.

Conclusions We propose that DNA hypermethylation marks at PPEIP frequently contribute to the pathogenesis of malignancies and within the precision medicine space, hold promise as biomarkers to inform on clinical features such as patient survival and therapeutic response.

Keywords Cancer, Epigenetics, DNA methylation, Protein prosphatase enzymes, Transcriptomics, RNA-seq, Hypermethylation, Gene-silencing, Biomarker

*Correspondence:
Ricky S. Joshi
rsj17@leicester.ac.uk

[1]Leicester Cancer Research Centre, Department of Genetics and Genome Biology, Robert Kilpatrick Clinical Sciences Building, Leicester Royal Infirmary, Leicester, UK
[2]Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Spain

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Protein phosphorylation is a post-translational modification that is vital for controlling signalling pathways and homeostatic maintenance of human cells such as metabolism, transcription/translation and cell division [1]. Protein phosphatases (PPE) are enzymes that catalyse the removal of phosphate groups from cell signalling proteins by hydrolysis, reversing the action of protein kinases (PK). The activities of both are highly coordinated and orchestrate responses from external stimuli and/or relay information for key transcriptional responses that stimulate or inhibit growth and protect the cell which are highly sensitive to small changes in PPE activity [2, 3]. PPE are categorized into two subtypes according to their specificity and action; Protein Tyrosine Phosphatases (PTP) and Protein Serine/Threonine Phosphatases (PSTP). PTP are further divided into (a) receptor-like and non-receptor PTP and (b) Dual-Specificity Phosphatases (DUSP). Typical PTP catalyse phosphotyrosine residues whereas DUSP dephosphorylate phosphoserine/threonine and phosphotyrosine amino acids. PSTP are also divided into two sub-groups; PhosphoProtein Phosphatases (PPP) and Metal-dependent Protein Phosphatases (PPM) [4, 5]. . PPE and PK represent arguably the most studied group of proteins in the literature to date. The uncontrolled suppression of PPE has often been described as a cause of dysregulated cell signalling programs that lead to human disease [3, 6]. In cancer, an imbalance in phosphorylation equilibrium has been reported as a cause of abnormal cell proliferation, dissemination and insensitivity to therapeutic treatment (extensively reviewed in Turdo et al., 2021 [7]). Many PK are well-known oncogenes, therefore PPE that counteract PK are assumed to be tumour suppressors [8]. To date, numerous PPE have been reported with tumour suppressor activity, PTEN as the most documented example. Identified as a tumour suppressor in 1997, *PTEN* was initially observed to be deleted in brain, breast and prostate cancer tissue [9]. Protein Phosphatase 2 A (PP2A) is the most expressed PSTP and described as a tumour suppressor in several cancers, e.g. breast, lung and melanomas [10] due to its role in inhibiting signal transduction pathways such RAF-MEK-ERK and Ras/PI3K/PTEN/Akt/mTOR. These pathways favour a number of cellular functions vital for tumour growth, such as cancer proliferation, reduced sensitivity to apoptotic signals and activation of pro-survival pathways [11, 12]. This dysregulation is essential to create an environment conducive to malignancies. PTP are one of the most recognized group of genes that make up the tumour suppressor family and are frequently inactivated and/or mutated in a variety of cancers [13, 14]. For example, PTPRT and PTPRD negatively regulate the JAK/STAT pathway [15]. PTPRH and PTPRB suppress downstream signalling of PI3K/Akt/mTOR and MEK/MAPK pathways by dephosphorylation of the epidermal growth factor receptor (EGFR) [16] and PTPN13 in non-small cell lung cancer by the phosphorylation control of EGFR and HER2 [17]. As well as the PTP, several DUSP are crucial regulators of MAPK proteins, such as ERK and JNK and therefore considered critical tumour suppressors [18]. Other types of phosphatases not discussed above also have important regulatory roles as tumour suppressors such as cellular prostatic acid phosphatase (PAcP) whose loss of expression leads to prostate carcinogenesis [19] or INPP5K, a phosphoinositide phosphatase gene associated with tumour suppressor activity in endometrial carcinoma [20].

Epigenetic dysregulation has been identified in cancer cells and mainly consists of global DNA hypomethylation with DNA hypermethylation at promoters of specific tumour-suppressor genes resulting in transcriptional silencing [21]. The clinical implications of DNA methylation have been an important characteristic to understand cellular transformation and is an important tool for cancer diagnosis, prognosis and therapy monitoring [22]. In spite of the overwhelming evidence for PPE as tumour suppressors and the role of DNA hypermethylation plays in transcription silencing, only a handful of PPE have been epigenetically characterized in primary tumours [23–30].

Advances in sequencing technologies, data availability and bioinformatic approaches have enabled simultaneous analyses of large sample sizes and -omic datasets. In this study, we performed an exhaustive, systematic analysis of promoter DNA methylation of genes that encode PPE and PPE-interacting proteins (PPEIP) in five different malignant tissues to identify aberrant DNA hypermethylation profiles that are absent in healthy controls (epimutations). We report the frequency of epimutations discerned in Colorectal, Oesophageal, Lung, Pancreatic and Stomach cancers and their effect on transcription, gene regulatory networks/pathways as well as providing examples of their clinical implications such as survival and response to treatment.

Materials and methods

Genome-wide DNA methylation array samples

A total of 729 genome-wide DNA methylation array datasets were initially downloaded from publicly available databases and processed in this study. Raw intensity (idats) files produced using the Infinium® HumanMethylation450 (450 K) BeadChip (Illumina) allowed for genome-wide interrogation of 482,000 CpG's. Raw idat files from 500 primary tumours from The Cancer Genome Atlas (TCGA) legacy database were downloaded. This consisted of 100 samples each from five cancer subtypes; Colorectal cancer (TCGA-COAD and READ), Oesophageal cancer (TCGA-ESCA), Lung

cancer (TCAG-LUAD), Pancreatic cancer (TCGA-PAAD), and Stomach cancer (TCGA-STAD). In order to identify cancer-associated changes in DNA methylation, 450 K data from 50 healthy samples (10/tissue type) from each tissue were used as controls in this study. 40 were downloaded from TCGA (10 each from Colorectal, Oesophageal, Lung and Pancreatic tissue) and 10 healthy stomach 450 K datasets were downloaded from Gene Expression Omnibus (GEO) under the accession number GSE127857. Additionally, 450 K data from a second independent test cohort of healthy tissue controls from 47 individuals were downloaded from TCGA (Large intestines ($n=16$), Oesophagus ($n=8$), Lung ($n=16$), Pancreas ($n=5$), and Stomach ($n=2$)) and analysed to identify hypermethylation events at PPEIP gene promoters in a population of healthy (non-cancer) tissues. The paucity of healthy tissue where 450 K data is available, limited the test cohort to 47 samples. 204 cancer cell line 450 K datasets were acquired from the Catalogue Of Somatic Mutations In Cancer (COSMIC) database from the Wellcome Trust Sanger Institute as detailed in [31]. Raw intensity files were downloaded corresponding to the same five cancer subtypes as described above; Colorectal ($n=49$), Oesophageal ($n=35$), Lung ($n=61$), Pancreas ($n=31$) and stomach ($n=28$) cancer cell lines. Finally, idat files for 25 embedded 3D cultures (organoid) representing the same 5 cancer subtypes were downloaded from GEO under the accession number GSE144213 (Colorectal $n=11$, Oesophageal $n=4$, Lung $n=1$, Pancreatic $n=7$ and Stomach ($n=2$) .

DNA methylation quality control, normalization and filtering

Raw signal intensity values were initially QC'd and pre-processed from subsequent idat files in R statistical environment (v3.6.1) (r-project.org) [32] using minfi Bioconductor package (v1.32.0) [33, 34] and processed in batches by tissue, cancer cell model and healthy samples independently. Quality control steps were applied to minimize errors and remove poor probe signals. Putative labelling errors were ascertained by examining methylation status at sex chromosomes of each individual. Duplicate samples were discerned by the SNP analysis feature in minfi. Vigorous quality control steps were performed on all samples and are detailed in [35]. Briefly problematic probes such as failed probes (detection p value>0.01), cross-reacting probes and probes that overlapped single nucleotide variants within +/-1 bp of CpG sites were removed. Background correction and dye-based normalization was performed using ssNoob algorithm (single-sample normal-exponential out-of-band). Probes hybridising to Chromosomes X and Y were removed from the final analysis. DNA methylation values for each CpG, represented as β-values with 1 representing fully methylated CpG and 0, fully unmethylated CpG were used for analysis. To further assess sample clusters to identify mislabelled samples based on DNA methylation similarity profiles (cancer vs. healthy controls) in the TCGA primary tumour dataset, robust correlations using multi-step bootstrap resampling and Euclidian distance measures for unsupervised hierarchical clustering were performed using the R package pvclust [36]. Ward.D2 minimum variance method for hierarchical clustering formation was produced using the R package hclust function [37]. Final clusters were plotted using the t-Distributed Stochastic Neighbour Embedding (t-SNE) technique using the R package Rtsne [38, 39]. For this, fifty thousand β values from across the genome in all tissue types were randomly selected in 5000 iterations to assemble and visualize inherent sample similarity. A total of 32 duplicate datasets and mislabelled samples (24 primary tumour and 8 healthy controls) were identified and removed from this analysis. Therefore, 705 cancer samples, 42 healthy controls and 47 healthy test controls were used in the final analysis (a breakdown of all sample ID's, tissues and cell models used in the analysis is provided in Table 1, Supplementary data S2 and S3). All QC, normalization and filtering steps mentioned above were performed separately for each cancer subtype and cell model. The control samples were ran separately to the tumour samples. Downstream analyses were performed under R statistical environment (v3.6.1).

Gene expression profiles

In order to assess the putative consequence of DNA methylation on the expression status of all TCGA individuals that harboured outlier DNA methylation at promoters of Protein Phosphatase Enzymes and interacting proteins (PPEIP) and controls, RNA-seq Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values were downloaded using the R Bioconductor package TCGAbiolinks [40, 41] for TCGA primary tumours and healthy tissues. Transcription profiles were only downloaded from matched patients where both DNA methylation and RNA-seq data was available. For cancer cell lines, gene expression profiles (Transcripts Per Kilobase Million or TPM) were downloaded directly from the Cancer Cell Line Encyclopedia (https://portals.broadinstitute.org/ccle) website. No direct analysis was performed between the two different types of RNA-seq expression counts (FPKM vs. TPM).

Clinical data

Extensive clinical information for all samples from the TCGA project (primary tumours) were retrieved from TCGA using TCGAbiolinks and integrated into R statistical environment (v3.6.1) for overall survival analyses.

Table 1 Breakdown of the Infinium Human Methylation 450 BeadChip (450 K) samples used and excluded in this study. Initial dataset refers to the number of cancer patients and control datasets downloaded and pre-processed for 450 K DNA methylation analysis. Datasets for analysis is the final number of 450 K patient and control samples analysed in this study. Datasets removed were the number of samples excluded after pre-processing and quality control steps, representing incorrect or mislabeled samples.

Tissue dataset	TCGA cancer	CCL	Organoids	Controls	TCGA healthy test
Dataset source:					
Colorectum	100	49	11	10	16
Oesophagus	100	35	4	10	8
Lung	100	61	1	10	16
Pancreas	100	31	7	10	5
Stomach	100	28	2	10*	2
Total:	**500**	**204**	**25**	**50**	**47**
Datasets used for analysis:					
Colorectum	95	49	11	10	16
Oesophagus	92	35	4	8	8
Lung	94	61	1	10	16
Pancreas	97	31	7	4	5
Stomach	98	28	2	10	2
Total:	**476**	**204**	**25**	**42**	**47**
Datasets excluded:	24	0	0	8	0
Colorectum	5	0	0	0	0
Oesophagus	8	0	0	2	0
Lung	6	0	0	0	0
Pancreas	3	0	0	6	0
Stomach	2	0	0	10	0
Total:	**24**	**0**	**0**	**8**	**0**

TCGA cancer = Primary tumour samples from The Cancer Genome Atlas data repository, CCL = Cancer Cell Line, Organoids = embedded 3D cultures. Controls = healthy tissue samples used as controls to identify cancer-associated hypermethylation at PPEIP gene promoters. TCGA healthy test = separate subset of healthy tissue samples used to identify non-cancer associated hypermethylation events in PPEIP gene promoters in the healthy population. * = Stomach samples were downloaded from GEO under the accession number: GSE127857 (due to the lack of available samples in TCGA data repository)

Data analysis pipeline

A curated list of 726 unique Protein Phosphatase Enzymes and interacting proteins (PPEIP) genes was downloaded from Ensembl (http://www.ensembl.org). PPEIP were defined as genes that encode a protein with previously described phosphatase activity and/or encode proteins that form PPEIP subunits, interact with PPEs, or potentially alter their function. Of the 726 PPEIP genes curated, 203 (28%) lacked a 450 K probe for CpG methylation profiling at its associated promoter region and therefore excluded from this study. The full list of 523 PPEIP genes can be found in Supplementary data S1. Promoter associated probes were defined as 450 K probes that measure CpG DNA methylation levels +/- 1500 bp and 200 bp from the transcription start site, present in the 5'UTR and 1st exon of all PPEIP genes. Probe annotation (such as those located in promoters, CpG island/shores, enhancers, etc.) was provided by Illumina's 450 K manifest file. Average promoter methylation for each PPEIP gene was calculated in all samples. Cancer-associated epimutations in PPEIP gene promoters were defined as absolute increases of 0.66 average beta value in cancer cell lines and absolute increases of 0.33 average beta value in primary tumour and organoid samples compared to average PPEIP promoter methylation in baseline healthy controls. These cut off values were consistent with previously published studies where hypermethylation in cancer cell lines applied a stringent > 0.66 beta value cut off for cellular homogeneity [31, 42], while for primary tumours and organoids that constitute diverse cell types, the cut off is reduced to 0.33 [43, 44]. Moreover, the epigenetic landscape of organoids is inherently closer to primary tumours than cancer cell lines [35]. A full list of average promoter methylation values for the 523 PPEIP genes in all 705 cancer cases and 42 baseline healthy controls can be found in Supplementary data S9. PPEIP epimutations present in the healthy population were defined as > 0.33 average beta value in test control individuals compared to the average baseline controls. This analysis and all downstream applications were performed using R statistical environment.

Epimutation tissue distribution and statistical analysis

In an even epimutation distribution ($n = 5007$), one would expect the number of epimutations to be proportional to the number of individuals per tissue set. To test this, we calculated the observed epimutation distribution ratio (OEDR) for each cancer tissue. The equation is shown below. The sample share is the fraction of individuals in each tissue tested for epimutations. Epimutation

share per tissue is the fraction of epimutations observed in each tissue and expected to be the same value as the tissue sample share in an even distribution model. The observed epimutation distribution ratio (OEDR) is the tissue sample share / epimutation share. For visualization purposes, the OEDR was converted to log2 values. This was calculated for each tissue.

$$\mathrm{OEDR} = \frac{\left(\frac{n\text{ of tissue samples}}{\text{Total }n\text{ of tissue samples}}\right)}{\left(\frac{\text{Observed }n\text{ of Epimutations/tissue}}{\text{Total }n\text{ of Epinutations}}\right)}\}\mathrm{Log2}$$

For Parametric tests, Pearson correlation and unpaired students t-tests were applied and for non-parametric data distributions, Spearmans' rank correlation and Wilcoxon rank test were calculated. For enrichment analyses, Fishers exact and Chi-squared tests were used. Data normality was assessed using a Shapiro-Wilk test. Linear regression models were used to estimate statistical relationships between DNA methylation and drug sensitivity. Kaplan-Meier plots and Log-rank (Mantel-Cox) tests were used to estimate overall survival (OS) in pancreatic cancer individuals with outlier *PTPRM* expression. Individuals in the top 25% quartiles of *PTPRM* expression were considered as "high expression" and "low expression" the bottom 25% quartile. The extreme 25% quartiles were used as standard cut off points to allow for robust statistical analysis performed through Cox proportional hazards regression models. All statistical analyses were carried out with the R statistical environment (v3.6.1) and p values <0.05 were considered as statistically significant.

Pathway and transcription factor binding analysis

Biocarta, KEGG and curated wiki pathway analyses were carried out to highlight cellular features and cancer networks that are putatively affected by hypermethylated PPEIP gene promoters. Transcription factor (TF) gene target examination using Encode data was also performed to inform on TF regulatory networks. All analyses were performed using the R package Enrichr [45].

Drug sensitivity

IC_{50} Z scores corresponding to drug sensitivity were downloaded from the Genomics of Drug Sensitivity in Cancer database (https://www.cancerrxgene.org/). This repository provides drug response data and genomic markers of sensitivity for 809 cancer cell lines and 198 compounds as part of their GDSC2 data release [31]. All cancer cell lines pertinent to this study were downloaded and processed according to tissue type.

A simplified overview of the analysis workflow is provided in Supplementary Fig. 1.

Results

Pan-cancer promoter hypermethylation analysis of protein phosphatase enzymes and interacting proteins

The Infinium HumanMethylation450 BeadChip (450K) platform allows for genome-wide CpG methylation analysis at 485,000 CpG's located at various genomic regions including 98% of all promoters in refseq-annotated genes. In order to dispel potential probe representation issues, we first analysed the distribution of probes at PPEIP promoters against all other 450K represented genes. 6296 450K probes informed CpG DNA methylation levels in PPEIP promoters (12 probes per gene) and 3260 (52%) of those, in CpG islands associated with PPEIP promoters (6 probes per CpG island) (Fig. 1A). In comparison, 168,664 probes were annotated to all gene promoters (9 probes per gene) and 79,008 (47%) to all promoter associated CpG islands (4 probes per gene) indicating a higher representation of probes for PPEIP gene promoters compared to all other gene promoters (Fig. 1A). The definition of "promoter probes" were those probes that provided semi-quantitative methylation values for CpG's located within 1500 bp and 200 bp of the transcription start site (TSS1500, TSS200), 5' untranslated region (5'UTR) or 1st exon of each gene. The majority of PPEIP gene promoter probes were found in the TSS1500 (30%) and the least in the 1st exon (13%) (Fig. 1B). Half of all PPEIP probes were distributed between the 5'UTR (29%) and the TSS200 (21%) (Fig. 1B). The median number of probes per gene was 11, and 7 when considering only CpG islands in PPEIP promoters (Fig. 1C). The Protein Tyrosine Phosphatase gene *PTPMT1* was the PPEIP with the most probes associated with its promoter ($n=60$).

PPEIP promoter hypermethylation profiles reveal distinct tissue susceptibility and frequency associated with cancer

Overall, 5007 hypermethylated Protein Phosphatase Enzymes and Interacting Proteins (PPEIP) gene promoters (epimutations) were identified in 593 cancer samples (84%), a median of 4 cancer-associated epimutations per individual (Supplementary Fig. 4B). The distribution of epimutations across samples was significantly imbalanced ($P=2.2\ 10^{-16}$). Stomach cancer patients demonstrated the highest percentage of PPEIP epimutation distribution of 27% however, stomach cancer cases only made up 18% of the total sample number (Supplementary Fig. 2A). This represented the highest ratio of epimutation vs. sample share, or observed epimutation distribution ratio (OEDR) of 0.59 (expected OEDR=0) (Supplementary Fig. 2B). Lung cancer patients showed with the lowest epimutation share with 9%, although 22% of all samples were from lung malignancies (-1.25 OEDR) (Supplementary Fig. 2A and B). Colorectal, oesophageal and stomach cancers showed positive OEDR (0.59–0.25) while lung and pancreatic cancer patients

showed negative OEDR (-1.25 and −0.68 respectively) (Supplementary Fig. 2B). 593 cancer patients (84%) presented with at least one hypermethylated PPEIP promoter in 160 (31%) PPEIP genes analysed as compared with healthy controls. Oesophagus represented the cancer tissue with the highest number of individuals with at least one epimutation (98%) (Supplementary Fig. 3A). This slight inconstancy with the OEDR stomach cancer data can be explained by the individual epimutation frequency in both tissues. 44% of all epimutations detected in the top 10% of individuals with the most epimutations accumulated were attributed to stomach cancer patients compared to only 28% in the oesophagus (Supplementary data S4). This suggests that a subset of stomach cancer patients accumulate higher quantities of epimutations in fewer individuals as compared with oesophageal cancer patients where epimutations are procured less inter-individually but more consistently across individuals. Lung cancer patients presented with the least number of epimutations (67%) consistent with OEDR data (Supplementary Fig. 3A). 96% of all organoid samples harboured at least one hypermethylated PPEIP promoter compared to 86% and 77% of all cancer cell lines and primary tumours respectively (Supplementary Fig. 3B). To assess the data further, we applied our epimutation detection pipeline to 450 K data from a separate test cohort of 47 control individuals presenting the five analysed tissues to identify epimutations in a population of healthy individuals. 11 individuals (23%) were identified carrying 79 hypermethylated PPEIP promoters as compared to 84% in cancer samples. A median epimutation count of 0 per individual was observed (Supplementary Fig. 4A) as compared to 4 in the same tissues in a cancer context (Supplementary Fig. 4B). Epimutations we detected were overwhelmingly enriched in cancer patients as compared to healthy individuals ($P=7.02\times10^{-18}$). Interestingly, epimutations detected in healthy samples were only observed in two tissues (Oesophagus $n=7$ and Pancreas $n=4$) (Supplementary data S5). This gave us confidence that the epimutations we identified through our bespoke bioinformatic pipeline were cancer-associated.

Next, we examined the frequency of PPEIP-associated epimutations in cancer patients. 5007 epimutations were detected in 160 PPEIP genes (Supplementary data S6). Of the 160 PPEIP genes, 88 (55%) were considered "rare" (observed in <1% of all cancer cases) and 41 (26%) as "recurrent" (identified in >5% of cancer samples) (Supplementary data S6). Many recurrent genes were known tumour suppressors with previously described cancer-related promoter hypermethylation anomalies (*PTPN13*, *DUSP5*, [2] *PPP1R14A* [26], *PPP1R3C* [27], *PTPRM* [28] and *IGFBP3* [46–48] validating the robustness of our approach. We also detected several PPEIP genes with previously undescribed recurrent methylation changes (Fig. 1D). INPP5B (Inositol Polyphosphate-5-Phosphatase) is an anti-apoptotic protein with a proliferative role in different cancer types [49, 50] and epimutations were observed in 130 individuals (Fig. 1D) in all 5 tissues (Supplementary Fig. 5). Proline-serine-threonine phosphatase interacting protein 2 or *PSTPIP2* promoter hypermethylation was observed in 37 cases (Fig. 1D) in all 5 tissues examined (Supplementary Fig. 6). PSTPIP2 is required for correct cell cycle function and dysregulation of PSTPIP2 contributes to abnormal proliferation and terminal differentiation in megakaryocytes [51]. In addition to recurrent differentially methylated promoters, several patients displayed rare, previously undescribed epimutations at PPEIP promoters. For example, an epimutation was detected at the promoter of *PPP3CC* (Protein Phosphatase 3 Catalytic Subunit Gamma) in 1 individual (colorectal cancer). PPP3CC repression has been reported to contribute to invasion and growth of glioma cells [52]. Loss-of-function genetic mutations in *PPP1R3B* gene have been associated with lung cancer [53], similarly, DNA methylation associated transcription silencing mimics loss-of-function properties. We observed one individual with oesophageal cancer phenotype harbouring an epimutation at the *PPP1R3B* promoter (Fig. 1D). Beta values for each probe and individuals (cancer and healthy controls) are provided in supplementary data S7.

We examined highly epimutated PPEIP genes and their described roles in human malignancies. Collectively, *PTPRT*, *CDH2*, *EYA4*, *SLITRK5*, *NTRK3*, *ADCY8*, *DNAJC6*, *PPM1E*, *FBP2* and *GRIN3A* were identified as those genes where most cancer individuals were observed to harbour hypermethylated PPEIP promoters. An even distribution of epimutation count was not observed across all cancer tissue types, consistent with the OEDR data. A breakdown of this is presented in Fig. 1E. *PTPRT* was the most epimutated PPEIP with 43% of all cancer cases showing epimutations in this gene. PTPRT is a tyrosine phosphatase with a previously described role as a tumour suppressor in colorectal cancer [13]. Interestingly, the authors also demonstrated that the most frequent genetically mutated tyrosine phosphatase gene was *PTPRT* in their colorectal cancer (CRC) cohort. In line with this finding, *PTPRT* was also the most epimutated gene in our CRC cohort; 76% of all CRC patients carried an epimutation. On the contrary, only 21% and 19% of all lung and pancreatic cancer patients respectively carried an epimutation in *PTPRT* (Fig. 1E). Several of the highest epimutable genes have disparate prevalence of hypermethylated promoters between cancer tissue types. Eyes absent 4 (EYA4) is a threonine-tyrosine phosphatase [54] previously described as a tumour suppressor in multiple cancers examined in this study (CRC [55], oesophagus [56], lung [57] and pancreas [58]). EYA4 promoter DNA

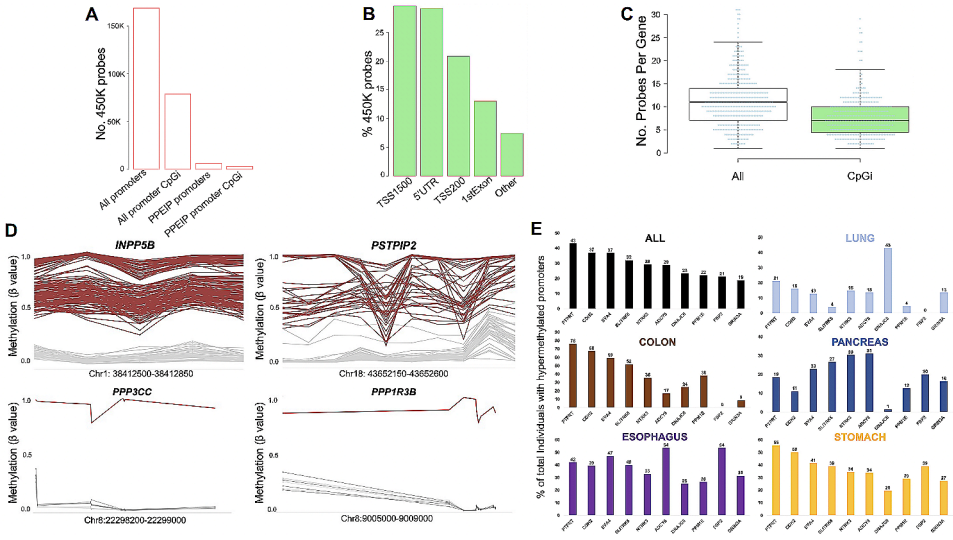

Fig. 1 Infinium Human Methylation 450K BeadChip (450K) probe distribution and promoter hyper-methylation analysis reveals cancer-associated epimutations occur with varying frequencies in multiple tissues. **A**. Distribution of 450K probes across all gene promoters (and CpG island-specific) and PPEIP promoters (also showing CpG island specific probes). **B**. % of promoter CpG genomic annotations. TSS1500 = CpG present within 1500 bp from transcription start site (TSS), TSS200 = within 200 bp of TSS, 5'UTR = CpG is found within the 5' untranslated region of the PPEIP and 1stExon = CpG found in the first exon of the PPEIP gene. The annotation "others" refers to promoter CpG's that annotate to other genomic regions (such as gene bodies or introns) of overlapping genes. **C**. Distribution of promoter probes per gene annotated to both CpG islands and non-CpG islands. **D**. Examples of hyper-methylated PPEIP promoters for both recurrent (observed in > 5% of the cancer population) and rare (< 1% % of the cancer population) epimutations. Red lines present individuals with outlier epimutations and grey lines, healthy controls. **E**. Bar graph showing the ten most epimutated PPEIP genes. Numbers on the y-axis (and above each bar) represent the % of total cancer individuals in all cancer tissues (ALL) as well as a breakdown of epimutations identified in individuals of specific cancer related tissue

methylation has been reported to be negatively correlated with gene expression and plays an important role in cell proliferation inhibition via Wnt and MAPK signalling pathways [59]. The frequency of epimutation is highly contrasting as 59% of CRC patients carry hypermethylated EYA4 promoters compared with only 13% of lung cancer patients (Fig. 1E). Another example is the NTRK3 gene. NTRK3 has been described in important cancer related pathways that promote both survival and cell proliferation and so its role as an oncogene [60] and tumour suppressor [61] is not unexpected. We observed at least a twofold difference in individuals with NTRK3 hypermethylated promoters between lung cancer (15%) and the other 4 cancers (CRC; 36%, stomach; 34%, Oesophagus; 33% and Pancreas 30%) (Fig. 1E).

Pan-cancer promoter hypermethylation of PPEIP affect cellular pathways and networks that favour tumour success

To further decipher the role of the 160 cancer-associated promoter hypermethylation susceptible PPEIP genes, we performed an enrichment analysis for gene networks, cellular pathways and transcription factor (TF) binding (Fig. 2). For gene network and cellular pathway analysis, three highly cited software were used; Biocarta [62], Kyoto Encyclopedia of Gene and Genomes (KEGG) [63] and curated WikiPathways [64]. All three software demonstrated high overlap of well-known pathways described in cancer cells such PI3K-AKT [65], MAPK [66] and cellular metabolism [67] (Fig. 2A,C,D). All three present actionable targets for anti-cancer drugs [68–70]. Other interesting pathways include angiogenesis related VEGFA-VEGFR2 signalling [71], regulatory circuits of STAT3 signalling pathways [72] as well as gene networks involved in aging [73]. We also interrogated transcription factor (TF) targets computed from ChIP-seq data from the ENCODE project [74]. The genes most affected by promoter hypermethylation are also targets for TF that are highly mutated in cancer such as chromatin remodelers (EP300, HDAC2, KDM4A among others) (extensively reviewed in [75]), cell cycle regulator (SIN3A [76]) and cell proliferation (YY1 [77]) (Fig. 2B).

Pan-cancer examination of epimutations in protein tyrosine and dual specific phosphatases exhibit aberrations in transcriptomic profiles affecting key cellular networks related to cancer

A number of Protein Tyrosine Phosphatases (PTPs) have been described in human cancers as tumour suppressor genes [2] and among the most studied include *PTPRM*; [78, 79], *PTPN13*; [17] and PTPRG; [80]. Therefore, in primary tumours, PTPs would represent a subset of key genes where the effects of promoter DNA methylation-induced transcriptional silencing are detrimental. In

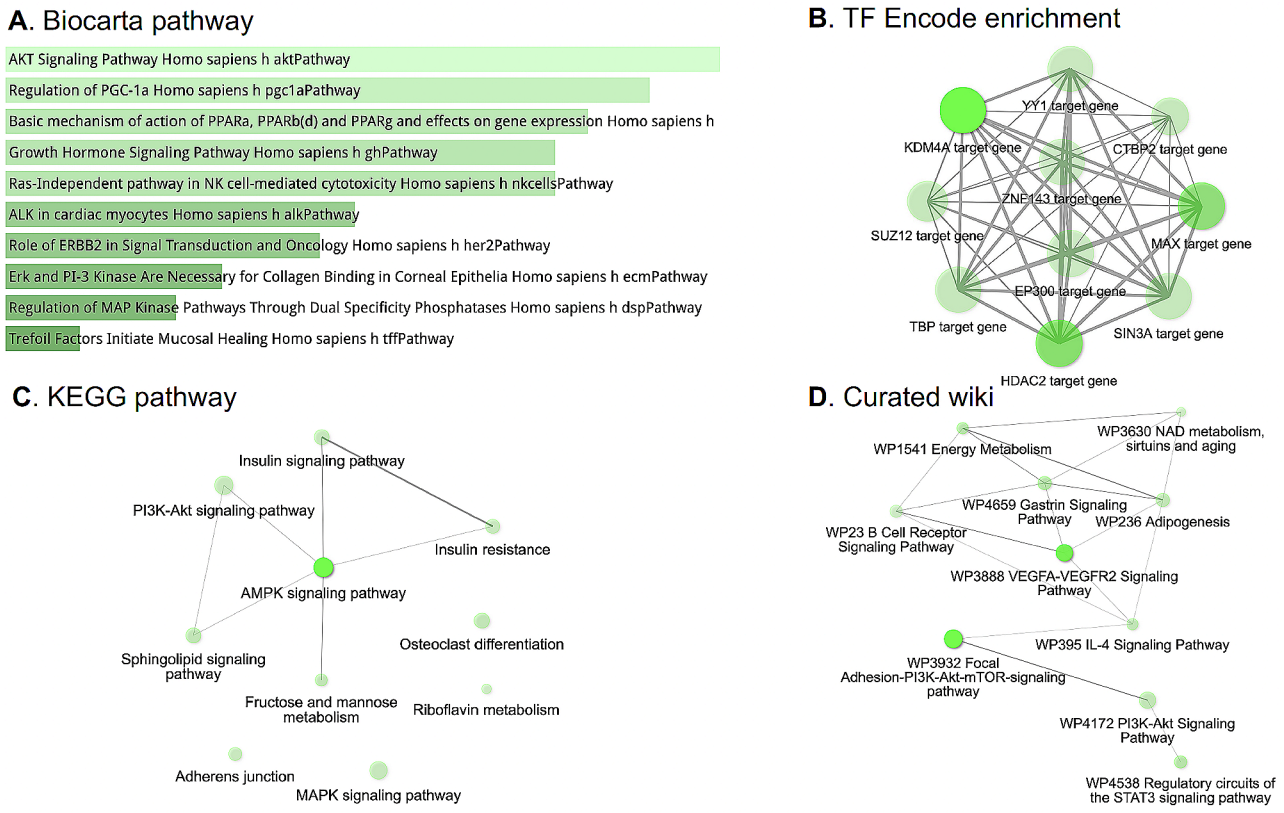

Fig. 2 PPEIP genes with hypermethylated promoters are involved in several cancer-associated cellular pathways and mechanisms. (**A**) Biocarta pathway analysis for pathway gene enrichment. The shade of green presents the significance of the specific gene-set or term. (**B**) Enrichment of transcription factor (TF) binding associated with hyper-methylated PPEIP genes from ChIP-seq experiments of > 300 TFs from the ENCODE project. Bright green presents higher number of genes. **C-D**. Orthogonal cellular pathway analyses from two independent resources; Kyoto Encyclopedia of Genes and Genomes (KEGG) and Curated wiki. For Figures A-D, the brighter the tone of green, the more significant the term is. In all network images (**B-D**), the grey lines represent gene content similarity

this regard, we closely examined PTPs for methylation sensitivity in primary tumours, cancer cell lines and 3D embedded cell cultures (organoids). We also performed the same analysis in dual-specificity phosphatases (DUSP) given their activity and role in cancer [81]. Only few individuals presented hypermethylated promoters for serine / threonine phosphatases (<2%) and therefore we focused our attention on PTP and DUSP genes. A list of PTPs and DUSPs was compiled from Ensembl and DNA methylation profiles generated for 43 PTP and 24 DUSP genes (Supplementary data S8). 17 PTP (40%) and 7 (29%) DUSP were observed with hypermethylated promoters in 410 (57%) and 102 (14%) cancer cases respectively. Colorectal cancer (CRC) patients demonstrated the highest number of PTP epimutations (81%) and pancreatic cancer patients the lowest (27%) (Fig. 3A). Stomach cancers presented the highest number of individuals with DUSP epimutations (23%) with lung and pancreatic malignancies the least (4%) (Fig. 3B). Further analysis into the role of PTPs and DUSP revealed PTPRT is the most ubiquitously epimutated PTP, 43% of all cancer individuals showed hypermethylated PTPRT promoters. DNAJC6 (23%) and PTPRM (16%) were the second and third most epimutated PTP. PTPRT was also found to be the most epimutated PTP in 4 of the 5 cancer tissues analysed (CRC = 78%, Stomach = 55%, Oesophagus = 42% and Pancreas = 19%) and second most epimutated in Lung (21%). DNAJC6 was the most epimutated PTP in Lung (43%) (Fig. 3C). CRC, stomach and oesophageal cancer showed overall high levels of individuals with PTP hypermethylated promoters. Of the top 10 most epimutated PTPs, Stomach (9/10), Oesophageal (7/10) and CRC (6/10) showed > 5% of individuals with epimutations in PTPs. Pancreas and Lung (2/10) presented with low epimutated PTPs (Fig. 3C). To a lesser extent, DUSP genes were also found to be highly epimutated (Fig. 3B). DUSP26 was the gene with the highest number of individuals with hypermethylated promoters (10%), with DUSP5 (6%), DUSP23 (3%) and DUSP15 (2%) also showing epimutations (Fig. 3D). CRC individuals showed the highest number of epimutated DUSP26 (19%) followed by Stomach (14%) and oesophagus (9%). Again, lung and Pancreas showed the least (4% and 1% respectively). DUSP5 was the most

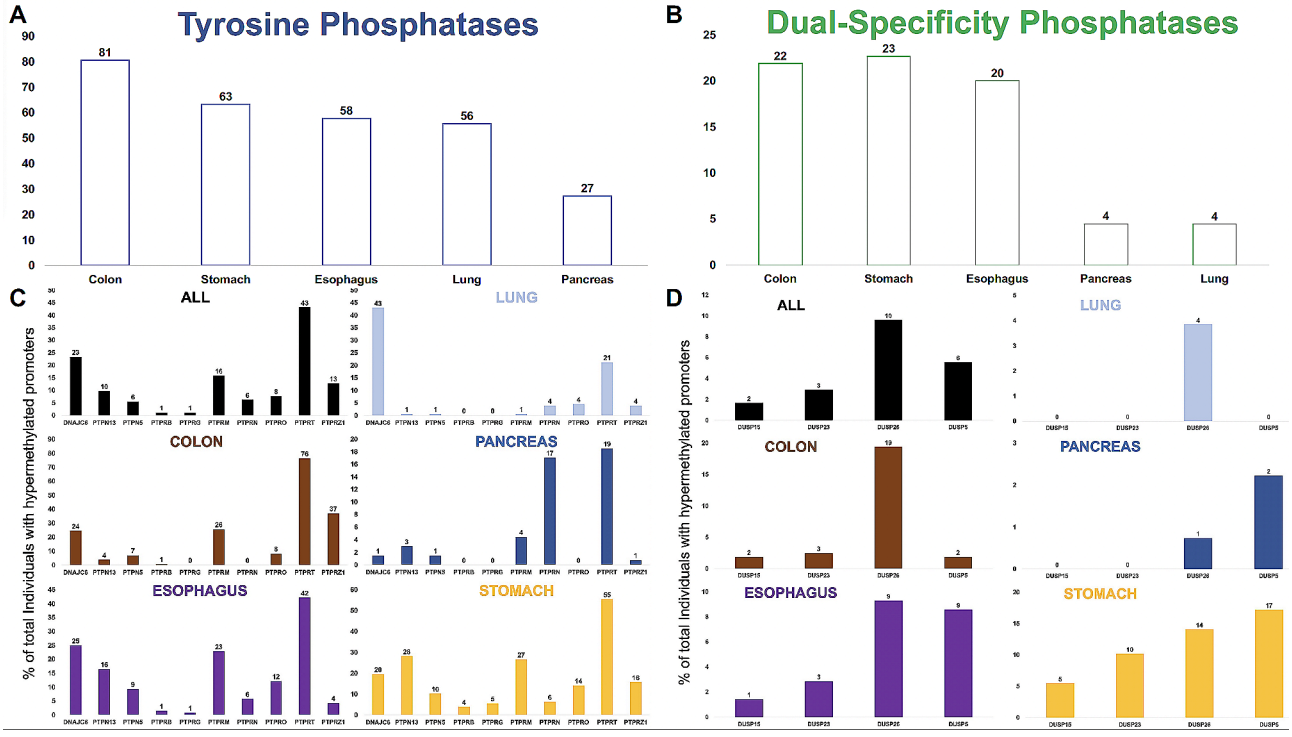

Fig. 3 Cancer-associated epimutations are commonly observed in Protein Tyrosine (PTP) and Dual-Specificity Phosphatase (DUSP) gene promoters. **A-B**. Bar graph demonstrating the % of individuals in each cancer subtype that carry at least one cancer-specific epimutation in tyrosine (**A**, blue) and dual specificity (**B**, green) phosphatase gene promoters. **C-D**. Bar graph representing the most ubiquitous PTP (**C**) and DUSP (**D**) gene promoters that harbour cancer-specific hyper-methylated promoters. Numbers above each bar represents % of cancer individuals with epimutations in the genes highlighted on the x-axis. The genes represented in (**C**) and (**D**) were detected in at least 5 individuals

epimutated DUSP in Stomach (17%), Pancreas (2%) and equal to DUSP26 in Oesophageal cancer (9%) (Fig. 3D).

Next, we analysed the effect of epimutations on PTP and DUSP gene expression given not all are expressed in the 5 tissues. 17 PTP and 7 DUSP were detected to contain at least one cancer-associated hypermethylated promoter. An initial analysis of gene expression in normal tissues was conducted using the GTEx portal (gtex.org, [82]) in the five tissues analysed in this study to determine the potential effect of cancer-associated epimutations on PTP and DUSP transcription. 9 of 17 PTPs and 5 of 7 DUSP were observed to be expressed at high levels in at least one healthy tissue type (>5 TPM) (Supplementary Fig. 7A and B). Further investigation revealed that 4 PTP (PTPRM, PTPN13, PTPRG and PTPRB) and 3 DUSP (DUSP23, DUSP5 and DUSP2) were ubiquitous epigenetic outliers in all cancer cell models, highly expressed in at least one normal tissue (>5 TPM) and maintained expression in their malignant counterpart prior to segregation based on promoter DNA methylation (Supplementary Fig. 8). Expression profiles from primary tumours (TCGA) and cancer cell lines (CCLE) for the 4 PTP and 3 DUSP are presented in Fig. 4. In each boxplot, gene expression is partitioned by individuals with hypermethylated promoters (>0.33 average promoter beta value in TCGA and >0.66 in CCLE vs. healthy controls) (see methods for details). A significant negative correlation between promoter methylation and gene expression was observed in all genes and cell models. We expect these data to also be representative of cancer organoid cell models [35]. Although hypermethylated promoters negatively correlate with gene expression in our cancer cohort (Fig. 4), one exception was observed in the gene *DNAJC6*, where an increase in promoter methylation was positively correlated with gene expression (Supplementary Fig. 9A and C). *DNAJC6* is the second most epimutated gene in all cancer samples and the most epimutated gene in the lung cancer cohort (43%) (Figs. 1 and 3). Although a high number of epimutations were identified in our analysis cohort, several showed very low expression in their pertinent tissues. As mentioned above, PTPRT is the most epimutated gene found in this study however its expression is extremely low (or non-existent) in the 5 tissues analysed (Supplementary Fig. 9B and D) with the highest expression observed in brain tissue (Supplementary Fig. 7A). This interesting finding demonstrates that epigenetic dysregulation in cancer cells occurs independent of an active transcriptional program and may provide important genomic information for other tissues.

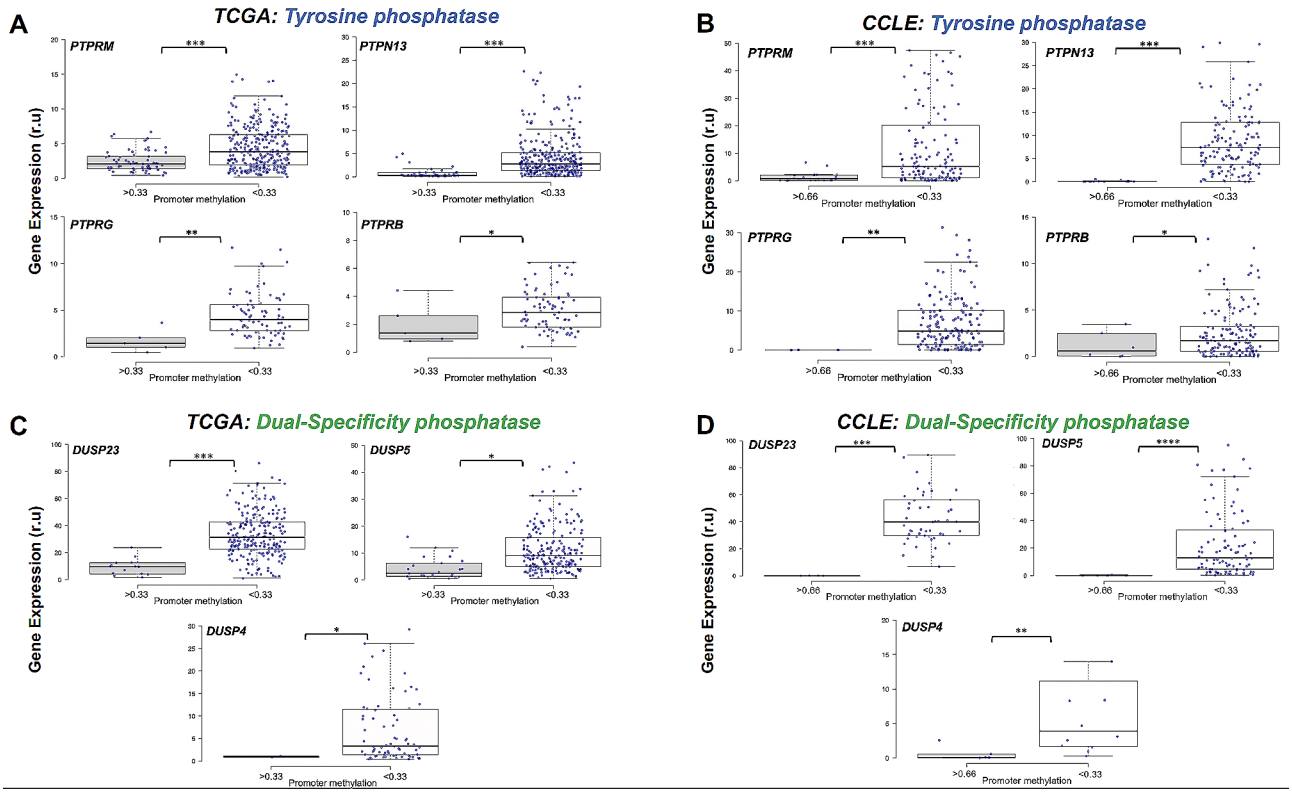

Fig. 4 Cancer-specific promoter DNA hypermethylation in Protein Tyrosine (PTP) and Dual-Specificity Phosphatase (DUSP) are associated with gene expression silencing. **A-B**. Boxplots representing the effect of hypermethylated promoters of PTP genes in primary tumours (**A**) and cancer cell lines (**B**). **C-D**. Boxplots representing the effect of epimutations discerned in promoters of DUSP genes from primary tumours (**C**) and cancer cell lines (**C**). Primary tumour and cancer cell line gene expression data was downloaded from The Cancer Genome Atlas (TCGA) and the Cancer Cell Line Encyclopedia (CCLE) projects. Gene expression values are presented as relative units (r.u) and are specific to each project (TCGA = FPKM and CCLE = TPM). Gene names shown at the top left-hand corner of each boxplot and blue dots presents individual (and cancer cell lines) expression values where DNA methylation profiles were also available. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

PTPRM epigenetic transcription silencing correlates with poor clinical outcome and reduced anti-cancer drug sensitivity

The role of PTP as tumour suppressors has been described in great detail in a number of tissues [2] and are targets for anti-cancer therapy [83]. In our cancer cohort, PTPRM represented the PTP with highest number of individuals with epimutations and high median expression values in healthy subjects (>5 TPM). Having demonstrated the presence of *PTPRM* promoter hypermethylation associated transcription silencing (Fig. 4A and B), we studied if PTPRM epigenetic loss in cancer patients had any impact on the clinical outcome in these patients. For this, we leveraged complete clinical and transcriptomic data from all individuals used in this study available in TCGA data repository. We identified that *PTPRM* gene silencing (top 25% quartile expression vs. bottom 25%) in pancreatic cancer patients showed a significant association with poor overall survival probability (Log Rank P = < 0.05, hazard ratio [HR] = 1.82; 95% confidence interval [CI] = 0.995–4.336; P = < 0.05).

PTPRM is an important component of STAT3 regulation with downstream effects on proliferation and metastasis in lung cancer [84] therefore we speculated whether PTPRM epigenetically deficient cancer cells could be exploited for therapeutic purposes, specifically drugs that target the JAK/STAT cell signalling pathway in other tissues. We downloaded IC_{50} concentration Z scores from GDSC2 (Genomics of Drug Sensitivity in Cancer, dataset2) database [31] for antitumour drugs that target key cancer-related cellular pathways (Fig. 5B). We identified 4 compounds, AZ960 [85], JAK8517, JAK18709 [86] and Ruxolitinib [87], that specifically target proteins in the JAK/STAT pathway and compared their drug sensitivity to *PTPRM* promoter methylation status in pancreatic cancer cell lines. We observed that drug sensitivity was significantly proportional to DNA methylation levels of the *PTPRM* gene promoter in pancreatic cancer cell lines (rho = 0.3, $P = 0.00462$) (Fig. 5C) suggesting that *PTPRM* promoter methylation profiles maybe used as a potential biomarker for clinical treatment response in pancreatic cancer patient therapy.

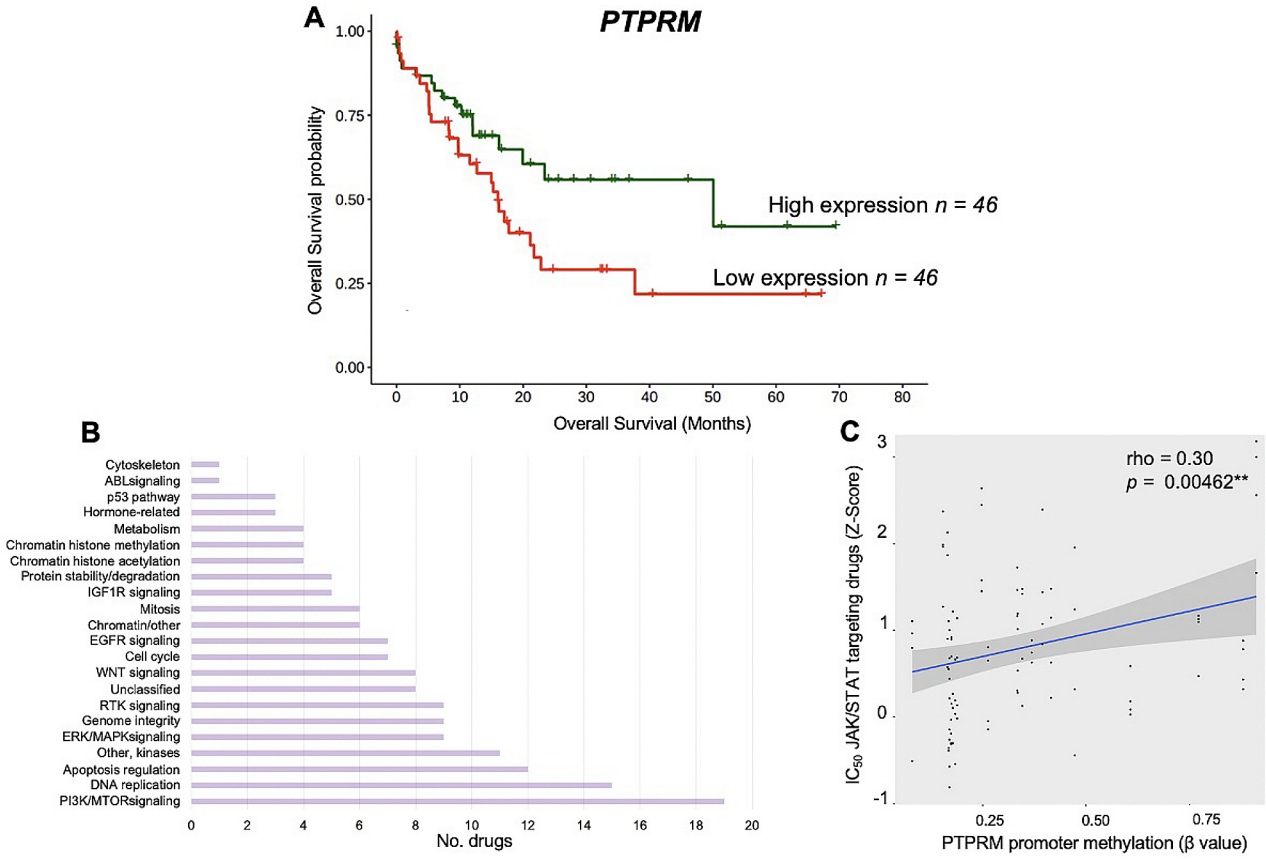

Fig. 5 *PTPRM* reduced expression via epigenetic silencing is associated with poor survival and reduced sensitivity to JAK/STAT targeted anti-cancer therapy in pancreatic cancer patients. (**A**) Kaplan–Meier curve showing reduced expression of PTPRM in patients with pancreatic cancer was significantly associated with shorter overall survival, Log Rank P = < 0.05 (HR = 2; 95% CI = 0.757–4.336; P = < 0.05). Green line represents patients in the top 25-percentile of PTPRM expression and the red line, bottom 25-percentile *PTPRM* expression in the pancreatic primary tumour cohort. (**B**) Horizontal bar graph demonstrates the cellular pathways targeted by anti-cancer drugs from GDSC2 (Genomics of Drug Sensitivity in Cancer) project. Drugs targeting "other" category were excluded. (**C**) Scatter plot illustrating IC_{50} concentration of JAK/STAT pathway targeting compounds is significantly proportional to DNA methylation of the PTPRM gene promoter in pancreatic cancer cell lines (rho = 0.3, P = < 0.005)

Discussion

In this study, we systematically surveyed hypermethylation profiles of Protein Phosphatase Enzymes and interacting protein (PPEIP) gene promoters to highlight the role of epigenetic marks that alter transcription of tumour suppressor genes and its effects within a clinical setting. In this regard, we detected 5007 hypermethylation events in 160 PPEIP gene promoters that were absent in healthy controls (epimutations). These epimutations were detected in 539 cancer patients (84%) across 5 tissues (colorectal, oesophageal, lung, pancreatic and stomach) and may disrupt the delicate balance of protein phosphorylation in signalling networks and promote malignancies. This dysregulation of key genes is not uniform across all tissue types. By assaying epimutations in multiple tissues, we observed that stomach cancer patients accumulate the highest number of epimutations (Supplementary figure S2) while lung cancer cases produce the least. This is an interesting finding as epimutations in PPEIP have been described in both tissues, (examples include; [88, 89]) although a cumulative comparison across tissues and individuals has not been previously reported. This disparity of epimutations may relate to their function defined by regulatory signatures underlying each tissue (such as active chromatin marks H3K4me3 and H3K4me1) that accompany differences in DNA methylation patterns by tissue-specific mechanisms in malignant cells [90]. This decrease in PPEIP epimutation susceptibility (*OEDR=log2 -1.25*) in lung tumours suggests that DNA methylation is more tightly controlled in lung tumorigenesis and likely reduces the role DNA methylation plays in highly mutable signalling pathways such as PI3K-AKT, JAK/STAT and MAPK (Fig. 2) in lung cancers as compared to other tissues. The oncogenic programs that establish a favourable molecular environment for these signalling pathways likely rely on other genomic mechanisms [91]. These findings may have important

clinical implications for designing treatment strategies that target PPEIP genes and their disrupted pathways.

The role of phosphatases, in particular Protein Tyrosine Phosphatases or PTP (and to a less extent, Dual Specific Phosphatases, DUSP), as tumour suppressors have been studied for decades and are emerging targets for novel technologies for oncogenic therapy [81, 92]. 57% of all cancer patients revealed epimutations in PTP genes and were overwhelmingly enriched in Colorectal Cancer probands (81%) as compared to pancreatic cancer (23%). Interestingly, epimutations in *PTPRT* were discerned in the most individuals despite very low expression in all tissues analysed (Supplementary Fig. 7A and 9B and D). This finding provides evidence that abnormal promoter hypermethylation mechanisms target certain tumour suppressors independent of their transcriptional activity. PTPRT is a negative regulator of signal transducer and activator of transcription 3 (STAT3) and therefore promoter hypermethylation events in addition to deleterious genetic mutations may be utilized as biomarkers to inform on potential neoplasm growth in multiple tissues, responsiveness to STAT3 inhibitors [93] and predictors of standard treatments against cancer [94]. In contrast, we discerned 4 PTP and 3 DUSP genes where epimutations consistently resulted in transcription silencing in all cell models (Fig. 4). All 7 genes have been implicated in several cancer-specific cellular programs as tumour suppressors and DNA methylation profiles associated with these genes may have compelling clinical implications [2]. For example, *PTPRM* epimutations were observed in the most cancer cases of the 7 genes (16%) in all 5 tissues. PTPRM is a receptor PTP and its intracellular compartment is responsible for phosphatase activity, whereas the extracellular section serves in cell–cell and cell–matrix contact [3]. STAT3 inactivation is catalysed by PTPRM dephosphorylation and leads to cancer cell death, therefore errors incurred in STAT3 dephosphorylation, such as promoter hypermethylation induced *PTPRM* transcriptional silencing, may lead to cancer initiation and progression with poor clinical prognosis [84, 95]. Furthermore, our data demonstrated that low PTPRM expression is associated with poor overall survival (Fig. 5A) and PTPRM methylation marks maybe used as biomarkers for JAK/STAT targeting anti-cancer drugs response. Additionally, the other epimutated genes showed equally intriguing roles in cancer. For example, in breast cancer, PTPN13 was reported to inhibit cancer aggressiveness by Src dephosphorylation [96] which is upregulated in tamoxifen-resistant ER-positive breast cancer patients [97], or DUSP4 repressed expression was identified as a mechanism of neoadjuvant drug chemoresistance and frequently depleted in chemotherapy refractory breast tumours [98]. Together, this highlights the role of epimutations in PTP and DUSP genes as putative cancer biomarkers for diagnosis, prognosis and treatment response.

Interestingly, we identified one example where hypermethylated promoters resulted in an increase of transcription (Supplementary Fig. 9A and C). *DNAJC6* was the second most epimutated PTP gene in cancer patients and the most in lung cancer patients (Fig. 3C). A literature search revealed DNAJC6 possesses oncogenic properties and promotes hepatocellular carcinoma (HCC) progression through induction of epithelial–mesenchymal transition. Overexpression of DNAJC6, as shown with increased promoter methylation in cancer subjects, was observed to enhance cell proliferation and invasion suggesting *DNAJC6* hypermethylation may be assayed as a putative biomarker for poor outcome in HCC [99].

Conclusions

Overall, our data illustrates the cellular and clinical relevance of DNA hypermethylation in a subset of PPEIP. We have broadened our knowledge of PPEIP epimutations across previously undescribed genes and tissues as well as providing an insight into epimutation susceptibility and distribution and its subsequent role in tumourigenesis and treatment. Considering the clinical implications, this type of phosphatase-wide analysis of promoter hypermethylation would benefit patients with other cancer types with high levels of phosphatase activity such as brain and neuroblastomas (Supplementary Fig. 7). With advances in single cell sequencing technologies, targeting epimutations in phosphatase enzymes will reveal precise signalling pathways and transcriptional programs in clonal sub-populations of tumour cells previously undetectable with bulk technologies and vastly improve cancer therapeutics. Together, it's seductive to posit that the emergence of epi-drugs that rehabilitate genes inactivated through epigenetic mechanisms [100] hold promise for re-activating tumour-suppressing cellular programs through genes such as *PTPRM*, that leads to improved treatment strategies and reduction in mortality of high-risk patients.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40246-024-00592-x.

Supplementary Material 1: Supplementary figures 1–9

Supplementary Material 2: Supplementary data S1–S9

Acknowledgements

We thank the University of Leicester for institutional support. We also thank Prof. Manel Esteller at the Josep Carreras Leukaemia Research Institute for providing both DNA methylation samples and guidance for this study.

Author contributions
R.S.J. conceived and designed the study; E.W. M.C.dM., D.P. and R.S.J. analyzed the multi-omics and clinical data. R.S.J. wrote the manuscript with contributions and approval from all authors.

Funding
Open Access funding enabled and organized by the University of Leicester. This work was funded in part by the Olga Torres Foundation (FOT) and Centro de Investigación Biomédica en Red (CIBER).

Data availability
Raw DNA methylation .idat files for primary tumours and healthy tissues were downloaded from the TCGA website. We incorporated 450 K DNA methylation data from healthy stomach tissue available in GEO under accession number GSE127857. Raw DNA methylation .idat files corresponding to the cancer cell lines for the 5 organs were downloaded from GEO under accession number GSE68379 and idat files for the embedded 3D cultures (organoid) representing the 5 cancer subtypes were downloaded from the GEO website under the accession number GSE144213. Finally, RNA-seq data corresponding to primary tumours and cancer cell lines was downloaded from TCGA and CCLE data repository respectively.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 3 November 2023 / Accepted: 27 February 2024
Published online: 12 March 2024

References

1. Cohen P. The structure and regulation of protein phosphatases. Annu Rev Biochem. 1989;58:453–508.
2. Stebbing J, Lit LC, Zhang H, Darrington RS, Melaiu O, Rudraraju B, et al. The regulatory roles of phosphatases in cancer. Oncogene. 2014;33:939–53.
3. Tonks NK. Protein tyrosine phosphatases: from genes, to function, to disease. Nat Rev Mol Cell Biol. 2006;7:833–46.
4. Barford D, Das AK, Egloff M-P, THE STRUCTURE AND, MECHANISM OF PROTEIN PHOSPHATASES. Insights into Catalysis and Regulation. Annu Rev Biophys Biomol Struct. 1998;27:133–64.
5. Li X, Wilmanns M, Thornton J, Köhn M. Elucidating human phosphatase-substrate networks. Sci Signal. 2013;6.
6. Bononi A, Agnoletto C, De Marchi E, Marchi S, Patergnani S, Bonora M, et al. Protein kinases and phosphatases in the control of cell fate. Enzyme Res. 2011;2011:1–26.
7. Turdo A, D'Accardo C, Glaviano A, Porcelli G, Colarossi C, Colarossi L et al. Targeting phosphatases and kinases: how to Checkmate Cancer. Front Cell Dev Biol. 2021;9.
8. Fontanillo M, Köhn M, Phosphatases. Their Roles in Cancer and Their Chemical Modulators. 2016. p. 209–40.
9. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science. 1997;275:1943–7.
10. Calin GA, Di Iasio MG, Caprini E, Vorechovsky I, Natali PG, Sozzi G, et al. Low frequency of alterations of the alpha (PPP2R1A) and beta (PPP2R1B) isoforms of the subunit a of the serine-threonine phosphatase 2A in human neoplasms. Oncogene. 2000;19:1191–5.
11. Zhang Q, Claret FX, Phosphatases. The New brakes for Cancer Development? Enzyme Res. 2012;2012:1–11.
12. Steelman LS, Chappell WH, Abrams SL, Kempf CR, Long J, Laidler P, et al. Roles of the Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR pathways in controlling growth and sensitivity to therapy-implications for cancer and aging. Aging. 2011;3:192.
13. Wang Z. Mutational analysis of the Tyrosine Phosphatome in Colorectal cancers. Sci (80-). 2004;304:1164–6.
14. FUNATO K, YAMAZUMI Y, ODA T. Tyrosine phosphatase PTPRD suppresses colon cancer cell migration in coordination with CD44. Exp Ther Med. 2011;2:457–63.
15. Veeriah S, Brennan C, Meng S, Singh B, Fagin JA, Solit DB, et al. The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. Proc Natl Acad Sci U S A. 2009;106:9435–40.
16. Yao Z, Darowski K, St-Denis N, Wong V, Offensperger F, Villedieu A, et al. A Global Analysis of the receptor tyrosine kinase-protein phosphatase interactome. Mol Cell. 2017;65:347–60.
17. Scrima M, De Marco C, De Vita F, Fabiani F, Franco R, Pirozzi G, et al. The nonreceptor-type tyrosine phosphatase PTPN13 is a tumor suppressor gene in non–small cell Lung Cancer. Am J Pathol. 2012;180:1202–14.
18. Caunt CJ, Keyse SM. Dual-specificity MAP kinase phosphatases (MKPs). FEBS J. 2013;280:489–504.
19. Muniyan S, Ingersoll MA, Batra SK, Lin M-F. Cellular prostatic acid phosphatase, a PTEN-functional homologue in prostate epithelia, functions as a prostate-specific tumor suppressor. Biochim Biophys Acta - Rev Cancer. 2014;1846:88–98.
20. Hedberg Oldfors C, Dios DG, Linder A, Visuttijai K, Samuelson E, Karlsson S, et al. Analysis of an independent tumor suppressor locus telomeric to Tp53 suggested Inpp5k and Myo1c as novel tumor suppressor gene candidates in this region. BMC Genet. 2015;16:80.
21. Esteller M. Epigenetics in Cancer. N Engl J Med. 2008;358:1148–59.
22. Ortiz-Barahona V, Joshi RS, Esteller M. Use of DNA methylation profiling in translational oncology. Semin Cancer Biol. 2020. https://doi.org/10.1016/j.semcancer.2020.12.011.
23. Luo S, Chen J, Mo X. The association of PTEN hypermethylation and breast cancer: a meta-analysis. Onco Targets Ther. 2016;9:5643–50.
24. Vanaja DK, Ballman KV, Morlan BW, Cheville JC, Neumann RM, Lieber MM, et al. PDLIM4 repression by hypermethylation as a potential biomarker for prostate cancer. Clin Cancer Res. 2006;12:1128–36.
25. Ying J, Li H, Cui Y, Wong AHY, Langford C, Tao Q. Epigenetic disruption of two proapoptotic genes MAPK10/JNK3 and PTPN13/FAP-1 in multiple lymphomas and carcinomas through hypermethylation of a common bidirectional promoter. Leukemia. 2006;20:1173–5.
26. Li D, Guo J, Wang S, Zhu L, Shen Z. Identification of novel methylated targets in colorectal cancer by microarray analysis and construction of co–expression network. Oncol Lett. 2017;14:2643–8.
27. Takane K, Midorikawa Y, Yagi K, Sakai A, Aburatani H, Takayama T, et al. Aberrant promoter methylation of PPP1R3C and EFHD1 in plasma of colorectal cancer patients. Cancer Med. 2014;3:1235–45.
28. Barazeghi E, Hellman P, Westin G, Stålberg P. PTPRM, a candidate tumor suppressor gene in small intestinal neuroendocrine tumors. Endocr Connect. 2019;8:1126–35.
29. Tögel L, Nightingale R, Wu R, Chüeh AC, Al-Obaidi S, Luk I, et al. DUSP5 is methylated in CIMP-high colorectal cancer but is not a major regulator of intestinal cell proliferation and tumorigenesis. Sci Rep. 2018;8:1767.
30. Schmid CA, Robinson MD, Scheifinger NA, Müller S, Cogliatti S, Tzankov A, et al. DUSP4 deficiency caused by promoter hypermethylation drives JNK signaling and tumor cell survival in diffuse large B cell lymphoma. J Exp Med. 2015;212:775–92.
31. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, et al. A Landscape of Pharmacogenomic interactions in Cancer. Cell. 2016;166:740–54.
32. Team RC, R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
33. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30:1363–9.
34. Fortin J-P, Triche TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2016. btw691.
35. Joshi R, Castro De Moura M, Piñeyro D, Alvarez-Errico D, Arribas C, Esteller M. The DNA methylation landscape of human cancer organoids available at the American type culture collection. Epigenetics. 2020;15:1167–77.
36. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006;22:1540–2.

37. Müllner D. Fastcluster: fast hierarchical, agglomerative clustering routines for R and Python. J Stat Softw. 2013;53.
38. van der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-SNE. J Mach Learn Res. 2008;9:2579–605.
39. LVD M. Accelerating t-SNE using tree-based algorithms. J Mach Learn Res. 2014;15:3221–45.
40. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44:e71–1.
41. Silva TC, Colaprico A, Olsen C, D'Angelo F, Bontempi G, Ceccarelli M, et al. TCGA Workflow: analyze cancer genomics and epigenomics data using Bioconductor packages. F1000Research. 2016;5:1542.
42. Simo-Riudalbas L, Diaz-Lagares A, Gatto S, Gagliardi M, Crujeiras AB, Matarazzo MR, et al. Genome-wide DNA methylation analysis identifies Novel hypomethylated non-pericentromeric genes with potential clinical implications in ICF Syndrome. PLoS ONE. 2015;10:e0132517.
43. Llinàs-Arias P, Rosselló-Tortella M, López-Serra P, Pérez-Salvia M, Setién F, Marin S et al. Epigenetic loss of the endoplasmic reticulum–associated degradation inhibitor SVIP induces cancer cell metabolic reprogramming. JCI Insight. 2019;4.
44. Rosselló-Tortella M, Llinàs-Arias P, Sakaguchi Y, Miyauchi K, Davalos V, Setien F, et al. Epigenetic loss of the transfer RNA-modifying enzyme TYW2 induces ribosome frameshifts in colon cancer. Proc Natl Acad Sci. 2020;117:20785–93.
45. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90–7.
46. Ye P, Qu C-F, Hu X-L. Impact of IGF-1, IGF-1R, and IGFBP-3 promoter methylation on the risk and prognosis of esophageal carcinoma. Tumor Biol. 2016;37:6893–904.
47. Kawasaki T, Nosho K, Ohnishi M, Suemoto Y, Kirkner GJ, Fuchs CS, et al. IGFBP3 promoter methylation in Colorectal Cancer: relationship with microsatellite instability, CpG Island Methylator phenotype, p53. Neoplasia. 2007;9:1091–8.
48. Torng P-L, Lin C-W, Chan MW, Yang H-W, Huang S-C, Lin C-T. Promoter methylation of IGFBP-3 and p53 expression in ovarian endometrioid carcinoma. Mol Cancer. 2009;8:120.
49. Akl H, Monaco G, La Rovere R, Welkenhuyzen K, Kiviluoto S, Vervliet T, et al. IP3R2 levels dictate the apoptotic sensitivity of diffuse large B-cell lymphoma cells to an IP3R-derived peptide targeting the BH4 domain of Bcl-2. Cell Death Dis. 2013;4:e632.
50. Kopacek J, Ondrias K, Sedlakova B, Tomaskova J, Zahradnikova L, Sedlak J, et al. Type 2 IP(3) receptors are involved in uranyl acetate induced apoptosis in HEK 293 cells. Toxicology. 2009;262:73–9.
51. Liu L, Wen Q, Gong R, Gilles L, Stankiewicz MJ, Li W, et al. PSTPIP2 dysregulation contributes to aberrant terminal differentiation in GATA-1-deficient megakaryocytes by activating LYN. Cell Death Dis. 2014;5:e988–8.
52. Wang H, Zhao S, Chen B, Fu C, Dang Y, Fang P, et al. Repression of the expression of PPP3CC by ZEB1 confers activation of NF-κB and contributes to invasion and growth in glioma cells. Jpn J Clin Oncol. 2018;48:175–83.
53. Kohno T, Takakura S, Yamada T, Okamoto A, Tanaka T, Yokota J. Alterations of the PPP1R3 gene in human cancer. Cancer Res. 1999;59:4170–4.
54. Okabe Y, Sano T, Nagata S. Regulation of the innate immune response by threonine-phosphatase of eyes absent. Nature. 2009;460:520–4.
55. Kim S-J, Tae CH, Hong SN, Min B-H, Chang DK, Rhee P-L, et al. EYA4 acts as a new tumor suppressor gene in Colorectal Cancer. Mol Carcinog. 2015;54:1748–57.
56. Luo M, Li Y, Shi X, Yang W, Zhou F, Sun N, et al. Aberrant methylation of EYA4 promotes epithelial-mesenchymal transition in esophageal squamous cell carcinoma. Cancer Sci. 2018;109:1811–24.
57. Wilson IM, Vucic EA, Enfield KSS, Thu KL, Zhang YA, Chari R, et al. EYA4 is inactivated biallelically at a high frequency in sporadic lung cancer and is associated with familial lung cancer risk. Oncogene. 2014;33:4464–73.
58. Mo S-J, Liu X, Hao X-Y, Chen W, Zhang K-S, Cai J-P, et al. EYA4 functions as tumor suppressor gene and prognostic marker in pancreatic ductal adenocarcinoma through β-catenin/ID2 pathway. Cancer Lett. 2016;380:403–12.
59. Miller SJ, Lan ZD, Hardiman A, Wu J, Kordich JJ, Patmore DM, et al. Inhibition of eyes absent homolog 4 expression induces malignant peripheral nerve sheath tumor necrosis. Oncogene. 2010;29:368–79.
60. Vaishnavi A, Le AT, Doebele RC. TRKing down an old oncogene in a new era of targeted therapy. Cancer Discov. 2015;5:25–34.
61. Genevois A-L, Ichim G, Coissieux M-M, Lambert M-P, Lavial F, Goldschneider D, et al. Dependence receptor TrkC is a putative colon cancer tumor suppressor. Proc Natl Acad Sci U S A. 2013;110:3017–22.
62. Nishimura D, BioCarta. Biotech Softw Internet Rep. 2001;2:117–20.
63. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44:D457–62.
64. Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C. WikiPathways: pathway editing for the people. PLoS Biol. 2008;6:e184.
65. Vara JÁF, Casado E, de Castro J, Cejas P, Belda-Iniesta C, González-Barón M. PI3K/Akt signalling pathway and cancer. Cancer Treat Rev. 2004;30:193–204.
66. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene. 2007;26:3279–90.
67. Seth Nanda C, Venkateswaran SV, Patani N, Yuneva M. Defining a metabolic landscape of tumours: genome meets metabolism. Br J Cancer. 2020;122:136–49.
68. Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. Exploiting the PI3K/AKT pathway for Cancer Drug Discovery. Nat Rev Drug Discov. 2005;4:988–1004.
69. Lee S, Rauch J, Kolch W. Targeting MAPK signaling in Cancer: mechanisms of Drug Resistance and Sensitivity. Int J Mol Sci. 2020;21.
70. Ngoi NYL, Eu JQ, Hirpara J, Wang L, Lim JSJ, Lee S-C, et al. Targeting cell metabolism as Cancer Therapy. Antioxid Redox Signal. 2020;32:285–308.
71. Zhao D, Pan C, Sun J, Gilbert C, Drews-Elger K, Azzam DJ, et al. VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate Myc and Sox2. Oncogene. 2015;34:3107–19.
72. Yu H, Lee H, Herrmann A, Buettner R, Jove R. Revisiting STAT3 signalling in cancer: new and unexpected biological functions. Nat Rev Cancer. 2014;14:736–46.
73. Hoeijmakers JHJ. DNA damage, aging, and Cancer. N Engl J Med. 2009;361:1475–85.
74. Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, Hitz BC, et al. ENCODE data at the ENCODE portal. Nucleic Acids Res. 2016;44:D726–32.
75. Nair SS, Kumar R. Chromatin remodeling in cancer: a gateway to regulate gene transcription. Mol Oncol. 2012;6:611–9.
76. Shi X, Garry DJ. Sin3 interacts with Foxk1 and regulates myogenic progenitors. Mol Cell Biochem. 2012;366:251–8.
77. Gordon S, Akopyan G, Garban H, Bonavida B. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. Oncogene. 2006;25:1125–42.
78. Sun P-H, Ye L, Mason MD, Jiang WG. Protein tyrosine phosphatase μ (PTP μ or PTPRM), a negative Regulator of Proliferation and Invasion of breast Cancer cells, is Associated with Disease Prognosis. PLoS ONE. 2012;7:e50183.
79. Sudhir P-R, Lin S-T, Chia-Wen C, Yang S-H, Li AF-Y, Lai R-H et al. Loss of PTPRM Associates with the Pathogenic Development of Colorectal Adenoma-Carcinoma Sequence. Sci Reports 2015 51. 2015;5:1–9.
80. Shu ST, Sugimoto Y, Liu S, Chang H-L, Ye W, Wang L-S, et al. Function and regulatory mechanisms of the candidate tumor suppressor receptor protein tyrosine phosphatase gamma (PTPRG) in breast cancer cells. Anticancer Res. 2010;30:1937–46.
81. Gao P-P, Qi X-W, Sun N, Sun Y-Y, Zhang Y, Tan X-N, et al. The emerging roles of dual-specificity phosphatases and their specific characteristics in human cancer. Biochim Biophys Acta - Rev Cancer. 2021;1876:188562.
82. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. Nat Genet. 2013;45:580–5. The Genotype-Tissue Expression (GTEx) project.
83. Köhn M. Turn and face the strange: a New View on Phosphatases. ACS Cent Sci. 2020;6:467–77.
84. Im J-Y, Kim B-K, Lee K-W, Chun S-Y, Kang M-J, Won M. DDIAS promotes STAT3 activation by preventing STAT3 recruitment to PTPRM in lung cancer cells. Oncogenesis. 2020;9:1.
85. Yang J, Ikezoe T, Nishioka C, Furihata M, Yokoyama A. AZ960, a novel Jak2 inhibitor, induces growth arrest and apoptosis in adult T-cell leukemia cells. Mol Cancer Ther. 2010;9:3386–95.
86. Mongre RK, Mishra CB, Shukla AK, Prakash A, Jung S, Ashraf-Uz-zaman M et al. Emerging importance of tyrosine kinase inhibitors against Cancer: Quo Vadis to cure? Int J Mol Sci. 2021;22.
87. Han ES, Wen W, Dellinger TH, Wu J, Lu SA, Jove R, et al. Ruxolitinib synergistically enhances the anti-tumor activity of paclitaxel in human ovarian cancer. Oncotarget. 2018;9:24304.
88. Omerovic J, Clague MJ, Prior IA. Phosphatome profiling reveals PTPN2, PTPRJ and PTEN as potent negative regulators of PKB/Akt activation in ras-mutated cancer cells. Biochem J. 2010;426:65–72.
89. Wu C-W, Chen J-H, Kao H-L, Li AFY, Lai C-H, Chi C-W et al. PTPN3 and PTPN4 tyrosine phosphatase expression in human gastric adenocarcinoma. Anticancer Res 26:1643–9.

90. Zhang B, Zhou Y, Lin N, Lowdon RF, Hong C, Nagarajan RP, et al. Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Res. 2013;23:1522–40.
91. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007;446:153–8.
92. Stanford SM, Bottini N. Targeting tyrosine phosphatases: time to end the Stigma. Trends Pharmacol Sci. 2017;38:524–40.
93. Peyser ND, Freilino M, Wang L, Zeng Y, Li H, Johnson DE, et al. Frequent promoter hypermethylation of PTPRT increases STAT3 activation and sensitivity to STAT3 inhibition in head and neck cancer. Oncogene. 2016;35:1163–9.
94. Hsu H-C, Lapke N, Chen S-J, Lu Y-J, Jhou R-S, Yeh C-Y, et al. PTPRT and PTPRD deleterious mutations and deletion predict Bevacizumab Resistance in Metastatic Colorectal Cancer patients. Cancers (Basel). 2018;10:314.
95. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. Nat Rev Cancer. 2009;9:798–809.
96. Glondu-Lassis M, Dromard M, Lacroix-Triki M, Nirdé P, Puech C, Knani D, et al. PTPL1/PTPN13 regulates breast Cancer Cell aggressiveness through direct inactivation of Src Kinase. Cancer Res. 2010;70:5116–26.
97. Planas-Silva MD, Bruggeman RD, Grenko RT, Stanley Smith J. Role of c-Src and focal adhesion kinase in progression and metastasis of estrogen receptor-positive breast cancer. Biochem Biophys Res Commun. 2006;341:73–81.
98. Balko JM, Cook RS, Vaught DB, Kuba MG, Miller TW, Bhola NE, et al. Profiling of residual breast cancers after neoadjuvant chemotherapy identifies DUSP4 deficiency as a mechanism of drug resistance. Nat Med. 2012;18:1052–9.
99. Yang T, Li X-N, Li X-G, Li M, Gao P-Z. DNAJC6 promotes hepatocellular carcinoma progression through induction of epithelial–mesenchymal transition. Biochem Biophys Res Commun. 2014;455:298–304.
100. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. Nat Rev Genet. 2019;20:109–27.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.