

## Article

# SETD7 Expression Is Associated with Breast Cancer Survival Outcomes for Specific Molecular Subtypes: A Systematic Analysis of Publicly Available Datasets

Fátima Liliana Monteiro <sup>1</sup>, Lina Stepanauskaite <sup>2,3</sup>, Cecilia Williams <sup>2,3,t</sup> and Luisa A. Helguero <sup>2,\*,t</sup>

<sup>1</sup> Department of Medical Sciences, Institute of Biomedicine—iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal

<sup>2</sup> SciLifeLab, Department of Protein Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden

<sup>3</sup> Department of Biosciences and Nutrition, Karolinska Institute, 141 83 Stockholm, Sweden

\* Correspondence: luisa.helguero@ua.pt

† These authors contributed equally to this study and shared the senior authorship.

**Simple Summary:** Breast cancer is the most common cancer among women, and it can be classified into subtypes with distinct biology and prognosis. The aim of our bioinformatic study was to assess the potential role of the protein methyltransferase SETD7 in breast cancer by using freely available resources. We saw that SETD7 is differentially expressed across subtypes, which may determine how SETD7 modulates cancer cell biological processes in each subtype. This translates into different prognosis and therapeutic response in patients stratified according to SETD7 levels. SETD7 might provide valuable additional information for discriminating patients based on subtypes and improve therapeutic decisions.



**Citation:** Monteiro, F.L.;

Stepanauskaite, L.; Williams, C.; Helguero, L.A. SETD7 Expression Is Associated with Breast Cancer Survival Outcomes for Specific Molecular Subtypes: A Systematic Analysis of Publicly Available Datasets. *Cancers* **2022**, *14*, 6029. <https://doi.org/10.3390/cancers14246029>

Academic Editors: Constantin N. Baxevasian and Cynthia Ma

Received: 17 November 2022

Accepted: 2 December 2022

Published: 7 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Abstract:** SETD7 is a lysine N-methyltransferase that targets many proteins important in breast cancer (BC). However, its role and clinical significance remain unclear. Here, we used online tools and multiple public datasets to explore the predictive potential of SETD7 expression (high or low quartile) considering BC subtype, grade, stage, and therapy. We also investigated overrepresented biological processes associated with its expression using TCGA-BRCA data. SETD7 expression was highest in the Her2 (*ERBB2*)-enriched molecular subtype and lowest in the basal-like subtype. For the basal-like subtype specifically, higher SETD7 was consistently correlated with worse recurrence-free survival ( $p < 0.009$ ). High SETD7-expressing tumours further exhibited a higher rate of *ERBB2* mutation (20% vs. 5%) along with a poorer response to anti-Her2 therapy. Overall, high SETD7-expressing tumours showed higher stromal and lower immune scores. This was specifically related to higher counts of cancer-associated fibroblasts and endothelial cells, but lower B and T cell signatures, especially in the luminal A subtype. Genes significantly associated with SETD7 expression were accordingly overrepresented in immune response processes, with distinct subtype characteristics. We conclude that the prognostic value of SETD7 depends on the BC subtype and that SETD7 may be further explored as a potential treatment-predictive marker for immune checkpoint inhibitors.

**Keywords:** SETD7; breast cancer; molecular subtypes; survival; gene expression; biological processes



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

SETD7 is a lysine N-methyltransferase that monomethylates the histone H3 lysine 4 (K4) and several other nonhistone proteins, including numerous transcription factors and epigenetic regulators (reviewed in [1]). Methylation by SETD7 can modulate a protein's stability, subcellular localization, and/or interactions with other proteins. For example, methylation by SETD7 improves the stability of ER $\alpha$  (*ESR1*) in breast cancer (BC), which may be of relevance to endocrine resistance [2]. SETD7 may also be important to prevent oxidative stress in BC cells by reducing KEAP1 and enhancing the expression of *GSTT2* and

*NFE2L2* (Nrf2), and promote metastasis by enhancing *VEGFA* or *RUNX2* expression [3]. On the other hand, *SETD7* can methylate oncogenic proteins (DNMT1, E2F1, and HIF1A), leading to their degradation (reviewed in [4]). Data obtained mainly from preclinical models point toward a context-dependent effect mediated by *SETD7*, and its role in BC remains controversial [4].

Several studies have compared *SETD7* mRNA or protein expression between BC and nontumorous tissue using public datasets or in-house cohorts, but with inconclusive results. Some studies showed that *SETD7* mRNA levels are lower in BC [5,6], others that BC has higher *SETD7* protein levels [7,8] or that no differences in *SETD7* expression between BC and normal tissue were observed [9]. The discrepancy between the studies also translates to the correlation analysis of *SETD7* expression with prognosis. While several studies found that high *SETD7* mRNA levels correlated with better overall survival (OS) [10] or disease-free survival (DFS) [11], others reported that higher *SETD7* levels were correlated with shorter OS and DFS [3,7,8]. Breast tumours are classified into distinct molecular subtypes based on gene expression (PAM50) profiles. The subtypes encompass luminal A, luminal B, Her2 (*ERBB2*)-enriched, normal-like, and basal-like subtypes, and exhibit fundamental differences in response to therapy and survival. Notably, in all studies that have analysed public datasets, no clarification as to whether the analysis was done by pooling all BC subtypes was available. Upon considering the number of cases in the studies analysing TCGA data, it appears that all subtypes were pooled, but this remains to be clarified in the other studies.

In a recent systematic review, we were able to associate high *SETD7* activity with inhibition of epithelial–mesenchymal transition in all the cancer types where this process had been studied, including BC [4]. Moreover, inhibition of *SETD7* function was associated with improved response to DNA-damaging agents in most of the analysed studies [4]. Thus, while effects mediated by *SETD7* are cell type- and signalling context-specific, the lack of clarity regarding the role and clinical significance of *SETD7* in BC may lead to stagnation in this field of research before clear conclusions can be drawn. Herein, an unbiased systematic analysis of public datasets was carried out. The main goal was to establish the predictive potential of *SETD7* expression in BC considering the impact of clinical factors such as subtype, grade, stage, and therapy on the association between *SETD7* expression and survival outcomes. The mutation frequency of *SETD7* and its target proteins was also investigated. Additionally, we identified the significantly overrepresented biological processes and pathways among the differentially expressed genes that emerge when *SETD7* expression is used to stratify the samples in each breast cancer subtype.

## 2. Materials and Methods

### 2.1. Description of Datasets

This study used previously published and publicly available data. No new sequencing or protein expression data were generated. A description of all the datasets that report *SETD7* expression, available for analysis within the different online tools used, is provided in Supplementary Table S1. Since cBioPortal (<https://www.cbioportal.org/>, v5.1.10, accessed on 5 August 2022) includes TCGA breast cancer data with different release dates, we used the PanCancer Atlas study for all analyses in cBioPortal (v5.1.10).

### 2.2. Analysis of *SETD7* Mutation and Copy Number

*SETD7* mutation and copy number were analysed using cBioPortal [12,13]. All datasets including mutation and copy number profiles (Supplementary Table S1) were pooled and analysis was carried out pooling samples from all BC subtypes [14–26], including data from The Metastatic Breast Cancer Project (<https://www.mbcproject.org/>, accessed on 5 August 2022) Count Me In (<https://joincountmein.org/>, accessed on 5 August 2022) (MBCproject cBioPortal data version February 2020).

### 2.3. Analysis of SETD7 Expression Using Online Tools

The analysis of *SETD7* expression in tumour and adjacent normal tissue was performed using RNA-seq data available in the TNMplot [27] online tool (<https://www.tnmplot.com>, accessed on 26 April 2022). The relationship between *SETD7* mRNA expression and clinicopathological characteristics, genomic alterations, DNA methylation, phosphoproteome, acetylproteome, and total proteome was explored using cBioPortal [12,13] online tool accessed between 20 January and 27 February 2022. RNA-seq and gene-chip data in cBioPortal are based on z-scores relative to all samples precomputed from the expression values in each dataset (fragments per kilobase of exon per million mapped fragments (FPKM), transcripts per million (TPM), or RNA-seq by expectation maximization (RSEM) for RNA-seq and log(microarray) for gene-chip). *SETD7* differential expression was set by comparing upper vs. lower quartiles (high and low expression, respectively). This analysis was done both by pooling all BC subtypes and for each subtype individually.

### 2.4. Correlation of SETD7 with Breast Cancer Outcomes

KM plotter, cBioPortal, and the Human Protein Atlas (HPA, [28]) were used to study the prognosis value of *SETD7* mRNA or corresponding protein. ROC plotter (<https://www.rocplot.org/>, accessed on 23 March 2022) [29] was used to study the potential predictive value of *SETD7*, using the recommended JetSet method [30] and without the ‘no outliers’ filter. ROC plotter uses 36 publicly available BC datasets that include chemotherapy (n = 2108), endocrine therapy (n = 971), and anti-Her2 (n = 267) treatment data. The patients are grouped into responders or non-responders by taking into consideration either the pathological complete response (n = 1775, incl. 639 responders and 1136 non-responders) or the relapse-free survival (n = 1329, incl. 978 responders and 351 non-responders) data provided by the studies. Differential expression of *SETD7* was set by comparing upper (high expression) vs. lower (low expression) quartiles, with exception of protein data in KM plotter and mRNA data in HPA where the differential expression was automatically set (median). Outcomes (OS, RFS—recurrence/relapse-free survival, PCR—pathological complete response, PFS—progression-free survival, DFS, DSFS—disease-specific free survival, DMFS—distant metastasis-free survival, PPS—palliative performance scale) could be evaluated in specific datasets depending on the patient data available for each dataset. Samples grouped by clinical factors or pooled BC subtypes were analysed.

### 2.5. Genes Associated with Differential SETD7 mRNA Expression in BC Subtypes

The TCGA-BRCA raw counts and FPKM data were downloaded on 20 March 2022 from NCI Genomic Data Commons (GDC) using the TCGAAbiolinks package (version 2.22.4) [31–33] in R (version 4.1.2). *SETD7* was defined as highly or lowly expressed based on upper and lower quartiles, respectively. The samples corresponding to the middle quartiles were considered unchanged and therefore removed. Genes with less than 1 FPKM in both high- and low-*SETD7* patients were considered not expressed and removed. Genes that were not present in at least a quarter of the samples were also filtered out. This was done based on counts per million using edgeR package (version 3.36.0) [34–36]. The raw counts for the remaining samples and genes were then processed using the default processing pipeline of DESeq2 (package version 1.34.0) [37]. Genes were considered significantly expressed if the Benjamini–Hochberg adjusted *p*-value (or *q*-value) for false-discovery rate (FDR) <0.05 and the absolute value of the log<sub>2</sub> fold change >0.4. Principal component analysis (PCA) on gene expression (after variance stabilizing transformation to the count data) was used for data visualization. The infiltrating immune and stromal scores for samples expressing high- and low-*SETD7* groups were calculated using immunodeconv (version 2.0.4) package [38]. Significantly differentially expressed genes between the high and low *SETD7* groups from each BC subtype were extracted using Venny 2.1 [39]. Gene ontology enrichment analysis was performed using DAVID [40,41]. Genes associated with high- or low-*SETD7* groups were analysed regardless of direction (up or down) and also separately in an attempt to distinguish which functional results are a subject of *SETD7* expression. The

default parameters with medium stringency were used. Biological processes containing at least two annotations and with adjusted  $p$ -value  $\leq 0.05$  are reported. The ggplot2 (version 3.3.5) [42] and GOpot (version 1.0.2) [43] packages were used for visualization.

### 3. Results

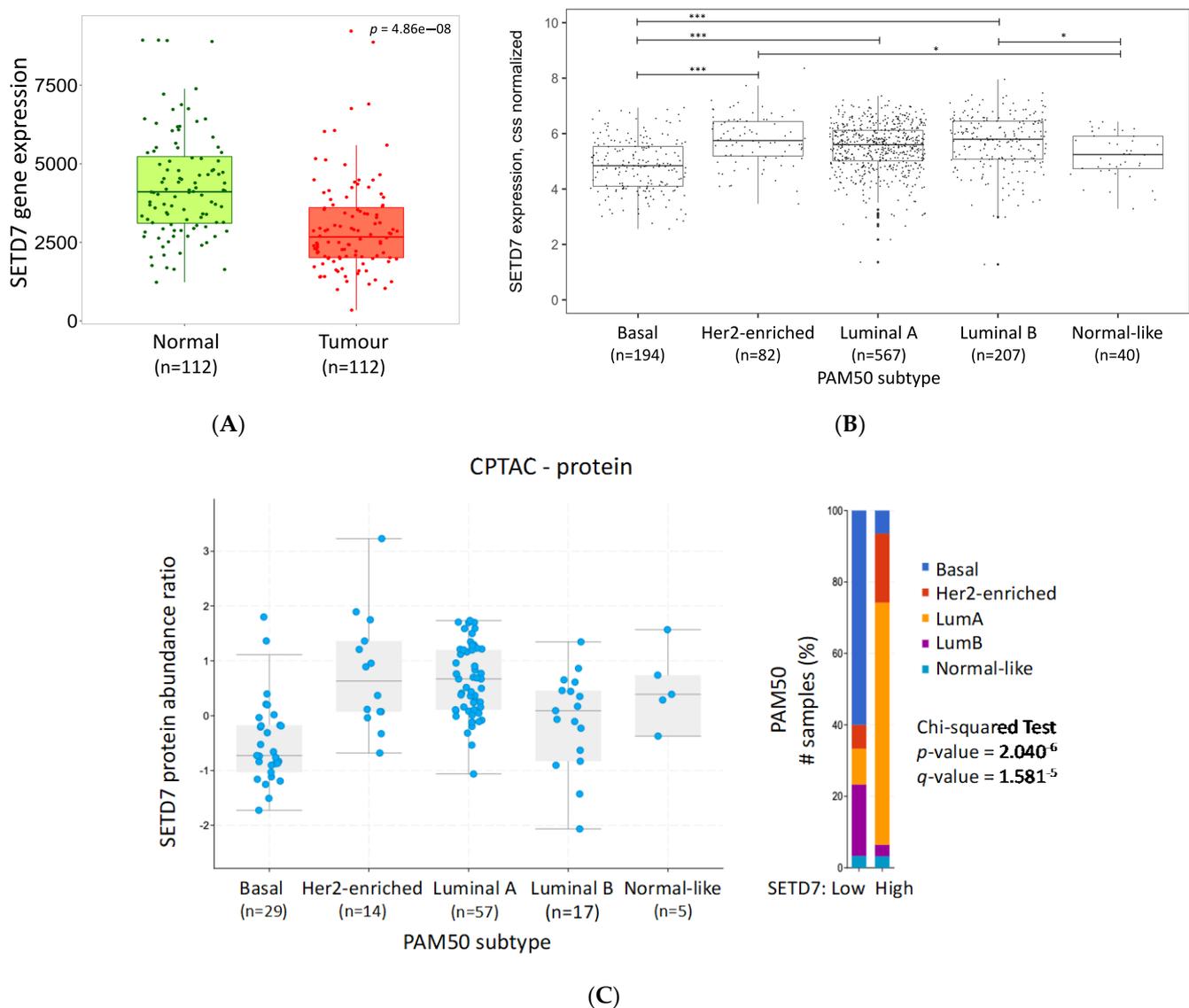
#### 3.1. Characterization of SETD7 Mutations, Copy Number, and Expression in BC

##### 3.1.1. SETD7 Mutation and Copy Number Profile

The frequency of *SETD7* mutations in BC was explored in publicly available data consisting of 8177 samples from 14 independent studies (whole exome sequencing, targeted sequencing, gene chip) [14–26]. *SETD7* was mutated in only 0.2% of BC cases (7/4378 profiled samples, Supplementary Figure S1A). These rare events corresponded to missense mutations of unknown significance, found randomly across the *SETD7* gene and across subtypes (Supplementary Figure S1B and File S1). *SETD7* copy number was altered with a slightly higher frequency of 12% (972/8177 patients). Shallow deletion (heterogeneous loss) of *SETD7* was observed in 17% of cases (761/4378 profiled samples, Supplementary Figure S1C), whereas deep deletion (deep loss, possibly a homozygous deletion) was only observed in 0.1% (3/4378), low-level gains (a few additional copies, often broad) in 4% (188/4378) and high-level amplification (more copies, often local) in 0.5% (23/4378) of cases (Supplementary Figure S1C). The shallow deletion was more often associated with the basal-like subtype (around 47% for basal-like vs. 35% for Her2-enriched, 21% for luminal A, 33% for luminal B, and 4% for normal-like subtypes). The low-level gain was more often associated with the Her2-enriched subtype (around 13% vs. 5% for basal, 6% for luminal A, 9% for luminal B, and 3% for normal-like subtypes). Overall, the genetic alteration of *SETD7* was not a common occurrence in BC, but a heterogeneous copy number loss was frequent (47%) in basal-like tumours specifically.

##### 3.1.2. Association of SETD7 Expression with Clinical Attributes

To compare the expression of *SETD7* in breast tumours and adjacent normal tissue, we used the publicly available online tool TNMplot comprising RNA-seq data of paired tissue samples from 112 patients. This analysis clearly showed that *SETD7* mRNA is significantly lower in breast tumours compared with the adjacent normal tissue (Figure 1A). Analysis by subtype was not supported by this tool. Next, the expression of *SETD7* was explored in the different BC datasets available from cBioPortal (RNA-seq, gene chip, and mass spectrometry). A significant correlation could be observed between *SETD7* expression and PAM50 subtype (in all datasets except METABRIC [ $q = 0.07$ ], Table 1). *SETD7* mRNA and protein expression were both consistently higher in the Her2-enriched and luminal A subtypes, and lower in the basal subtype (Figures 1B,C and S2, Table 1). The mean differences of each group (Cohen's  $d$ ) and the confidence interval for TCGA-BRCA data grouped by subtype were further analysed (Supplementary Figure S3). Luminal A vs. Her2-enriched ( $d = -0.25 [-0.54, 0.04]$ ), luminal B vs. Her2-enriched ( $d = -0.07 [-0.39, 0.25]$ ), and luminal B vs. luminal A ( $d = 0.18 [-0.02, 0.38]$ ) differences had a small effect size indicating little or no clinical relevance. However, the differences between normal-like and luminal B ( $d = -0.49 [-0.92, -0.06]$ ), Her2-enriched ( $d = -0.56 [-1.03, -0.08]$ ) or basal subtype ( $d = 0.36 [-0.07, 0.79]$ ) had a medium effect size, and most importantly, luminal A vs. basal ( $d = 0.67 [0.46, 0.88]$ ) or luminal B vs. basal ( $d = 0.85 [0.60, 1.10]$ ), and Her2-enriched vs. basal ( $d = 0.92 [0.60, 1.25]$ ) had strong effect sizes, supportive of relevant clinical differences (Supplementary Figure S3). Luminal B tumours exhibited varying mRNA levels dependent on the dataset (Supplementary Figure S2, Table 1) and low protein levels (Figure 1C). No correlation between *SETD7* differential expression and therapy, tumour grade, or stage was observed in pooled BC samples or when divided by subtype (Supplementary Table S2). In conclusion, our analysis across different large-scale datasets clearly shows that *SETD7* expression is significantly reduced in basal-like BC, which may be related to the copy number loss noted above for basal-like tumours.



**Figure 1.** SETD7 expression in breast cancer. (A) SETD7 mRNA expression in tumour and adjacent normal tissue using RNA-seq data available from TNMplot; (B) SETD7 mRNA expression across PAM50 subtypes using TCGA-BRCA data in R. ANOVA followed by Tukey's test: \* < 0.05; \*\*\* < 0.0001 (C) SETD7 protein expression across PAM50 subtypes using CPTAC data from cBioPortal.

### 3.1.3. Association of SETD7 Expression with Clinically Relevant Signatures

To investigate whether SETD7 differential expression was correlated with clinically relevant mRNA and protein signatures, we first analysed datasets in cBioPortal where this information was available (hypoxia scores were available for the TCGA PanCancer cohort, and stromal, immune, and stemness scores for CPTAC cohort), and as a second approach, we used a deconvolution method in R to further explore the tumour microenvironment infiltration mRNA signatures in the TCGA-BRCA cohort. High SETD7 mRNA correlated with lower hypoxia scores [44,45] in pooled samples from all subtypes in the TCGA PanCancer Atlas dataset (Figure 2A, left panel). High SETD7 protein correlated with higher stromal scores [46,47] in CPTAC dataset and high SETD7 mRNA with cancer-associated fibroblasts (CAFs), endothelial cells, and neutrophil signatures in the TCGA-BRCA dataset (Figure 2B). On the other hand, low SETD7 mRNA and protein levels were correlated with high xCell immune score and stemness score (CPTAC, pooled samples from all subtypes; Table 2 and Figure 2A, middle and right panels). Further analysis of the xCell immune score

showed enrichment of B and T cells (CD8+ T cells) in the low-SETD7 group, while, as mentioned above, enrichment of neutrophils was noted in the high-SETD7 group (TCGA-BRCA; Figure 2B).

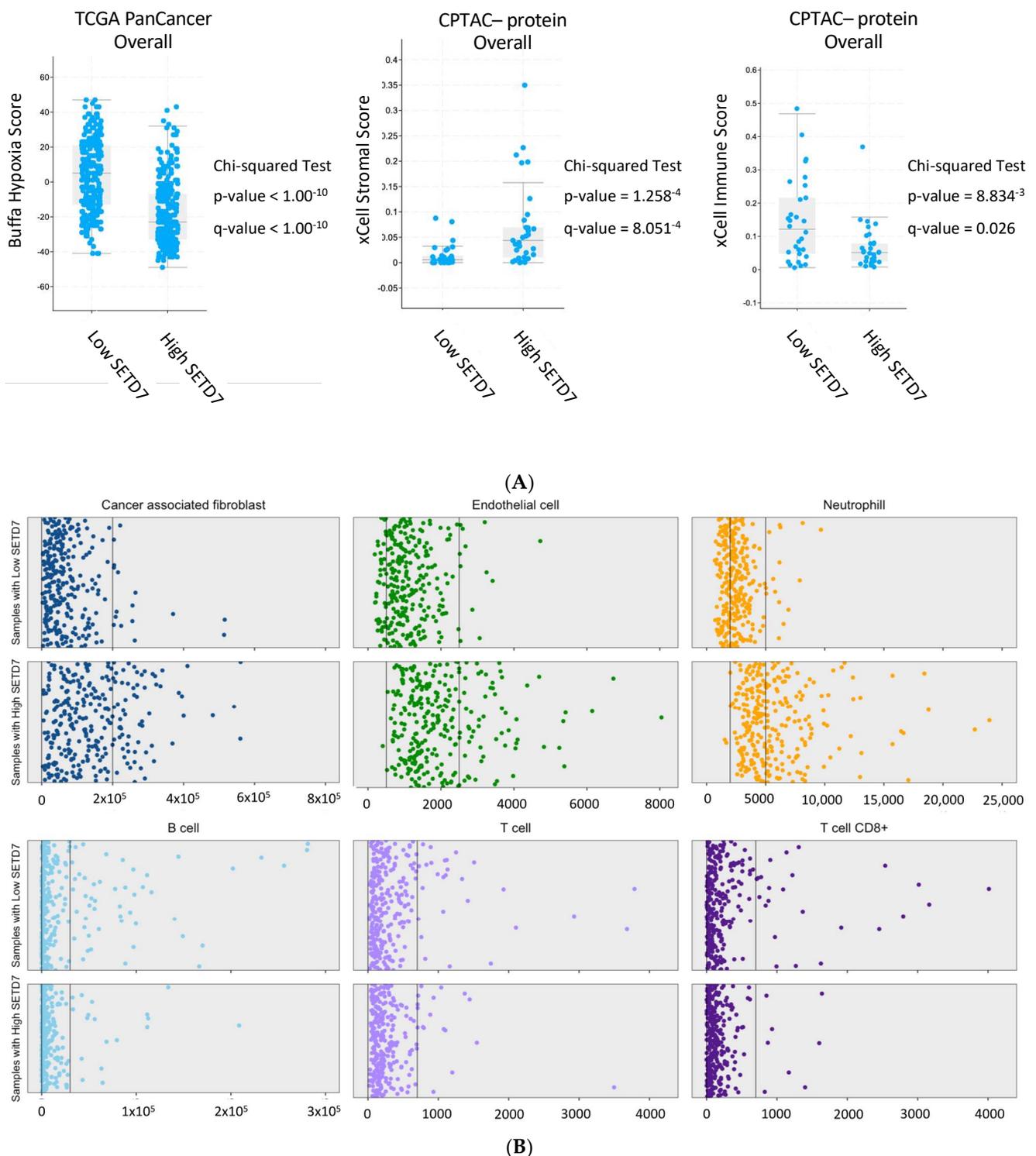
**Table 1.** *SETD7* expression per subtype. Significant values are highlighted in bold. Chi-squared test *p*-value and Benjamini–Hochberg FDR correction *q*-value. NA—not available; nSETD7 DE—number of samples with differentially expressed *SETD7*; nTotal—total number of samples.

cBioPortal (nSETD7 DE/n Total Samples)	PAM50	Luminal A	Luminal B	Her2-Enriched	Basal	Normal-Like
CPTAC-RNA (61/122)	<b><math>p = 1.59^{-5}</math></b> <b><math>q = 9.85^{-5}</math></b>	High	Low	High	Low	High
CPTAC-protein (61/122)	<b><math>p = 2.04^{-6}</math></b> <b><math>q = 1.58^{-5}</math></b>	High	Low	High	Low	Unchanged
METABRIC (952/2976)	<b><math>p = 4.23^{-3}</math></b> <b><math>q = 0.07</math></b>	High	Unchanged	High	Unchanged	Unchanged
SMC (84/187)	<b><math>p = 1.45^{-7}</math></b> <b><math>q = 2.33^{-6}</math></b>	High	High	High	Low	Unchanged
TCGA PanCancer Atlas (541/1084)	<b><math>p &lt; 10^{-10}</math></b> <b><math>q &lt; 10^{-10}</math></b>	High	High	High	Low	Unchanged

**Table 2.** *SETD7* association with stromal, immune, stemness and hypoxia scores. Significant values are highlighted in bold: Wilcoxon test *p*-value and Benjamini–Hochberg FDR correction *q*-value. NA—not available.

Scores	CPTAC			TCGA PanCancer Atlas			High <i>SETD7</i> Correlates with
	RNA	Protein		Overall	Luminal A	Luminal B	
	Overall	Luminal A					
xCell Strommal	<b><math>p = 4.06^{-4}</math></b> <b><math>q = 2.17^{-3}</math></b>	<b><math>p = 1.26^{-4}</math></b> <b><math>q = 8.05^{-4}</math></b>	<b><math>p = 2.09^{-3}</math></b> <b><math>q = 0.01</math></b>	NA	NA	NA	High
ESTIMATE Strommal	<b><math>p = 1.83^{-3}</math></b> <b><math>q = 7.33^{-3}</math></b>	<b><math>p = 5.96^{-4}</math></b> <b><math>q = 2.38^{-3}</math></b>	<b><math>p = 2.25^{-3}</math></b> <b><math>q = 0.01</math></b>	NA	NA	NA	High
xCell Immune	<b><math>p = 0.01</math></b> <b><math>q = 0.03</math></b>	<b><math>p = 8.83^{-3}</math></b> <b><math>q = 0.03</math></b>	<b><math>p = 0.68</math></b> <b><math>q = 0.81</math></b>	NA	NA	NA	Low
Stemness	<b><math>p = 0.01</math></b> <b><math>q = 0.03</math></b>	<b><math>p = 2.33^{-4}</math></b> <b><math>q = 5.96^{-4}</math></b>	<b><math>p = 0.06</math></b> <b><math>q = 0.20</math></b>	NA	NA	NA	Low
Buffa Hypoxia	NA	NA	NA	<b><math>p &lt; 1.00^{-10}</math></b> <b><math>q &lt; 1.00^{-10}</math></b>	<b><math>p &lt; 1.00^{-10}</math></b> <b><math>q &lt; 1.00^{-10}</math></b>	<b><math>p = 5.82^{-4}</math></b> <b><math>q = 0.01</math></b>	Low
Winter Hypoxia	NA	NA	NA	<b><math>p &lt; 1.00^{-10}</math></b> <b><math>q &lt; 1.00^{-10}</math></b>	<b><math>p = 1.11^{-8}</math></b> <b><math>q = 2.51^{-7}</math></b>	<b><math>p = 6.36^{-4}</math></b> <b><math>q = 0.01</math></b>	Low

Analysis by molecular subtype revealed that the lower hypoxia scores in high *SETD7* mRNA group was specific for luminal A and B subtypes (Table 2). Likewise, when each subtype was investigated separately, high stromal scores were significantly correlated with high *SETD7* expression in the luminal A subtype (Table 2) whereas enrichment of CAF, endothelial cell, and neutrophil signatures was noted in high-*SETD7* samples of all subtypes (Supplementary Figures S4 and S5).



**Figure 2.** Correlation of *SETD7* differential expression with clinical factors. **(A)** Association of *SETD7* differential expression (high vs. low) with Buffa Hypoxia Score (mRNA), xCell Stromal Score and xCell Immune Score (protein) when pooling all breast cancer types together. Wilcoxon test *p*-value and Benjamini–Hochberg FDR correction *q*-value; **(B)** Association of *SETD7* differential expression (high vs. low) with tissue-infiltrating immune and stromal cell populations using *mcp\_count* method [48] from *immunodeconv* package and TCGA-BRCA data overall.

In conclusion, we saw that reduced levels of *SETD7* are associated with high stemness and immune scores in general, while high expression is associated with increased stromal

score, including for CAFs, endothelial cells, and neutrophils, and reduced hypoxia scores in the luminal A subtype specifically.

### 3.2. Association of *SETD7* Expression with Genomic Alterations and DNA Methylation

*SETD7* histone methyltransferase can influence chromatin remodelling. The association of differential *SETD7* expression with genomic alterations and DNA methylation was explored in cBioPortal, using the TCGA PanCancer Atlas cohort to investigate associations with genomic alterations and METABRIC to analyse the impact on DNA methylation. No gene was significantly deleted or mutated in either high- or low-*SETD7* mRNA groups, even when specific mutation types were queried individually (missense, in-frame, truncating, structural variants, or CNA deletion), although there was a tendency for *TP53* (p53) gene alterations. Significant correlations between *SETD7* differential expression and other genomic alterations (such as amplifications) were observed (Supplementary File S2 and Figure S6). The high-*SETD7* group showed higher genomic alterations in *ERBB2* (Her2; 21% event frequency in the high-*SETD7* group vs. 6% in the low *SETD7* group; Supplementary Figure S6), especially higher copy number amplification (16.42% or 44/268 profiled samples in high *SETD7* compared with 4.85% or 13/268 profiled samples in low *SETD7* group). On the other hand, the low-*SETD7* group showed higher event frequency in the *TP53* gene (c.a. 28% in high *SETD7* vs. 48% in low *SETD7*, respectively; Supplementary Figure S6).

Analysis by molecular subtype did not disclose any significant correlation between *SETD7* differential expression and genomic alterations, although a clear tendency for a higher number of genomic alterations in the *ERBB2* gene in the high-*SETD7* group was observed for the Her2-enriched subtype (90% event frequency in high *SETD7* vs. 58% in low *SETD7*. Most of these alterations were copy number amplification (27/29 profiled samples).

*SETD7* impacts cancer-related processes, including in BC [4,49], but it was not mutated in BC (as shown above). Therefore, mutations in *SETD7* target genes were queried. Forty-two specific genes with known *SETD7* target methylation sites were analysed (reviewed in [1,4] and detailed in Supplementary Table S3). Only one mutation in a *SETD7* lysine methylation site was found, consisting of a K873E missense mutation in the tumour suppressor *RB1*, in only one sample. No mutations on sites previously reported to compete with *SETD7* methylation [4] were identified.

The correlation with general DNA methylation of *SETD7* target genes was investigated in the METABRIC dataset, which is the only set with information about DNA methylation. No correlation between *SETD7* differential expression and DNA methylation throughout the genome was observed, either when pooling all BC samples or when stratifying by molecular subtype.

Thus, high *SETD7* expression was related to increased *ERBB2* copy number in the Her2-enriched subtype, but not related with other genetic alterations.

### 3.3. Gene Expression and Biological Processes Associated with Differential *SETD7* mRNA Expression

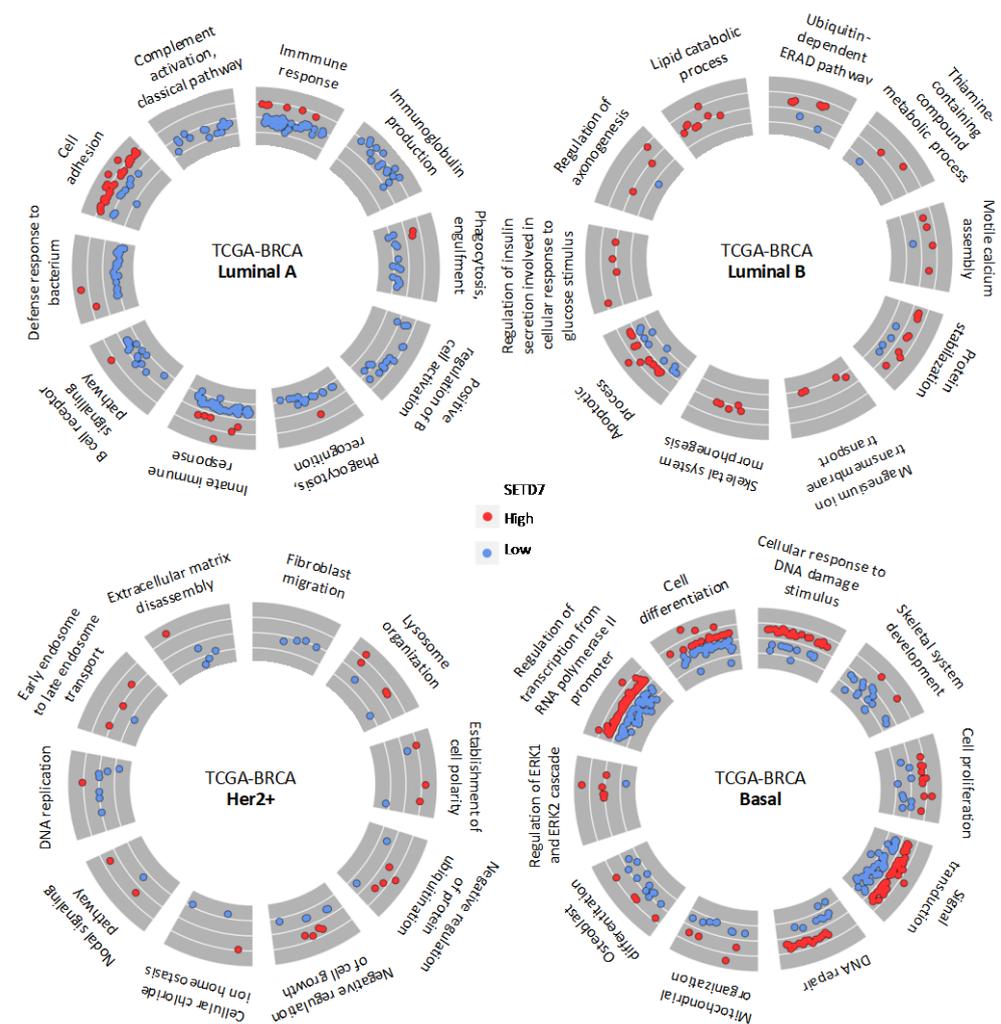
To avoid heterogeneity, this analysis was carried out on the TCGA data, which is the most powerful gene expression dataset available to date. We retrieved the TCGA-BRCA RNA-seq data, and analysis of differential gene expression between the high-*SETD7* and low-*SETD7* groups was carried out for each molecular subtype (Supplementary Figure S7A). The normal-like subtype was not included in the analysis due to the low number of samples available. First, the overall gene expression data were validated by the PCA plot clearly separating the basal-like and the Her2-enriched, luminal A and B subtypes, and further showing that the luminal A and B were more similar than the other subtypes, as expected (Supplementary Figure S7B).

Next, the comparison between high- and low-*SETD7* groups for each molecular subtype (Venn diagram in Supplementary Figure S8A) disclosed 2834 genes that were commonly associated with *SETD7* expression in all subtypes (Supplementary File S3). Of these, 1699 were highly expressed in the high-*SETD7* group and 1133 in the low-*SETD7* group.

Only two genes (*GPB1* and *CYP4F22*) were oppositely correlated with *SETD7* expression in different subtypes, being enriched in low-*SETD7* tumours for all subtypes except luminal B where they were upregulated in the high-*SETD7* group. The commonly upregulated genes in high-*SETD7* groups of all subtypes were overrepresented for biological functions related to protein phosphorylation and ubiquitination. Processes previously associated with *SETD7* (reviewed in [4]) also appeared overrepresented in the high-*SETD7* groups. These include cellular response to DNA damage stimulus, DNA repair, cell division, cell cycle, and cell migration (Supplementary File S4). The genes upregulated in low-*SETD7* tumours, on the other hand, were related to translation and mitochondrial respiration (Supplementary Figure S8B).

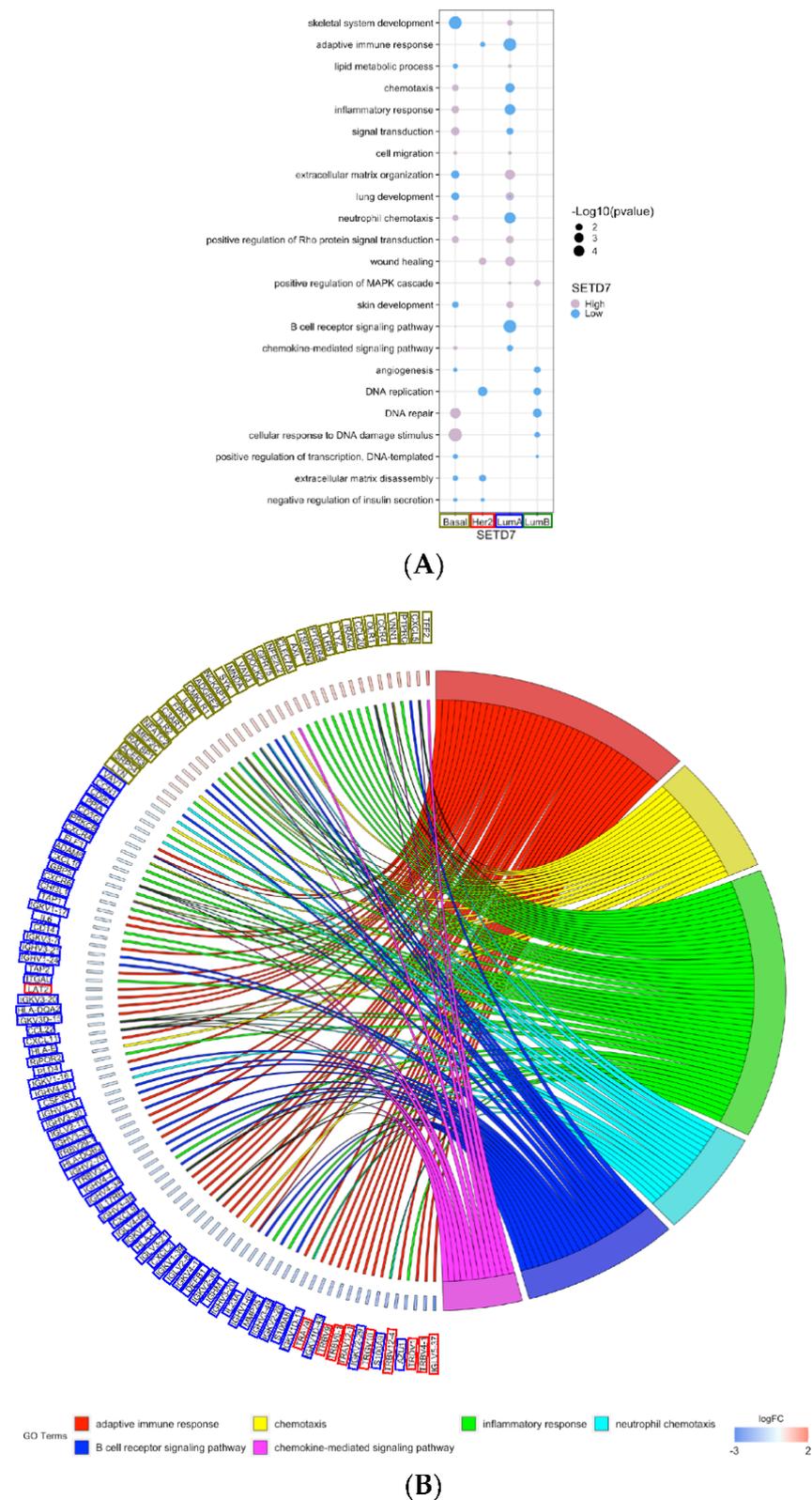
Further, the unique genes being differentially expressed between high- or low-*SETD7* groups within each subtype were analysed for enrichment of biological pathways (Figures 3 and S9). In luminal A subtype, the pathways related with immune response were overrepresented in low-*SETD7* group. The highly expressed genes that related more strongly ( $|\log_2\text{FC}| > 1$ ) with low-*SETD7* were mainly immunoglobulins, such as *IGKV2-29* ( $q\text{-value} = 7.39^{-05}$ ,  $\log_2\text{FC} = -1.57$ ) and other genes which trigger the immune response such as *AZU1* ( $q\text{-value} = 2.24^{-16}$ ,  $\log_2\text{FC} = -1.81$ ) and *S100A9* ( $q\text{-value} = 2.72^{-11}$ ,  $\log_2\text{FC} = -1.63$ ) (Supplementary File S3). On the other hand, pathways overrepresented in high-*SETD7* tumours were linked to cell adhesion-related pathways (Supplementary Figure S9). These included the genes *FGB* ( $q\text{-value} = 5.99^{-03}$ ,  $\log_2\text{FC} = 1.37$ ) and *ROBO2* ( $q\text{-value} = 2.54^{-06}$ ,  $\log_2\text{FC} = 1.32$ ). In the luminal B subtype, DNA repair and response to DNA damage-related pathways were the main biological processes overrepresented in low-*SETD7*, while the lipid catabolic process was strongly overrepresented in high-*SETD7* tumours. Interestingly, magnesium ion transmembrane transport and regulation of insulin secretion involved in cellular response to glucose stimulus were solely overrepresented in the luminal B subtype, where the genes *PNPLA3* ( $q\text{-value} = 5.46^{-03}$ ,  $\log_2\text{FC} = 1.27$ ) and *ADCY5* ( $q\text{-value} = 4.23^{-03}$ ,  $\log_2\text{FC} = 1.18$ ) stand out as strongly correlated with *SETD7* mRNA expression. In the Her2-enriched subtype, fibroblast migration and extracellular matrix disassembly were overrepresented in low-*SETD7* tumours, from which the genes *MMP7* ( $q\text{-value} = 3.25^{-02}$ ,  $\log_2\text{FC} = -1.32$ ), *KLK5* ( $q\text{-value} = 6.67^{-03}$ ,  $\log_2\text{FC} = -1.85$ ), and *KLK7* ( $q\text{-value} = 8.63^{-03}$ ,  $\log_2\text{FC} = -2.31$ ) were strongly regulated ( $|\log_2\text{FC}| > 1$ ). On the other hand, early endosome to late endosome transport was overrepresented in the high-*SETD7* group. Finally, in the basal-like subtype, cell differentiation-related pathways were strongly overrepresented in low-*SETD7* (Supplementary Figure S9) where keratins stand out (e.g., *KRT13*,  $q\text{-value} = 5.42^{-13}$ ,  $\log_2\text{FC} = -4.36$ ; *KRT6A*,  $q\text{-value} = 3.21^{-10}$ ,  $\log_2\text{FC} = -3.56$ ; and *KRT1*,  $q\text{-value} = 8.19^{-08}$ ,  $\log_2\text{FC} = -2.93$ ), along with other genes, such as *SNAI2* ( $q\text{-value} = 4.82^{-04}$ ,  $\log_2\text{FC} = -1.13$ ) and *IGF2* ( $q\text{-value} = 1.25^{-02}$ ,  $\log_2\text{FC} = -1.01$ ). In the high-*SETD7* basal-like tumours, the cellular response to DNA damage stimulus and DNA repair-related pathways were overrepresented. Also, *PPARGC1A* ( $q\text{-value} = 1.49^{-03}$ ,  $\log_2\text{FC} = 1.48$ ), a protein involved in cancer metabolic adaptation to stress, was upregulated.

Some biological processes were shared between subtypes, even though the genes for each subtype were unique (Figure 4A). These included many processes related to immune responses, such as chemotaxis, neutrophil chemotaxis and chemokine-mediated signalling, B cell receptor signalling pathway, and inflammatory response. The genes associated with immune response processes were often associated with low-*SETD7* in luminal A and Her2-enriched subtypes and with high-*SETD7* in the basal-like subtype (Figure 4B and Supplementary File S4). This aligns with the xCell immune score, B and T cell signatures shown above (Figure 2A,B).

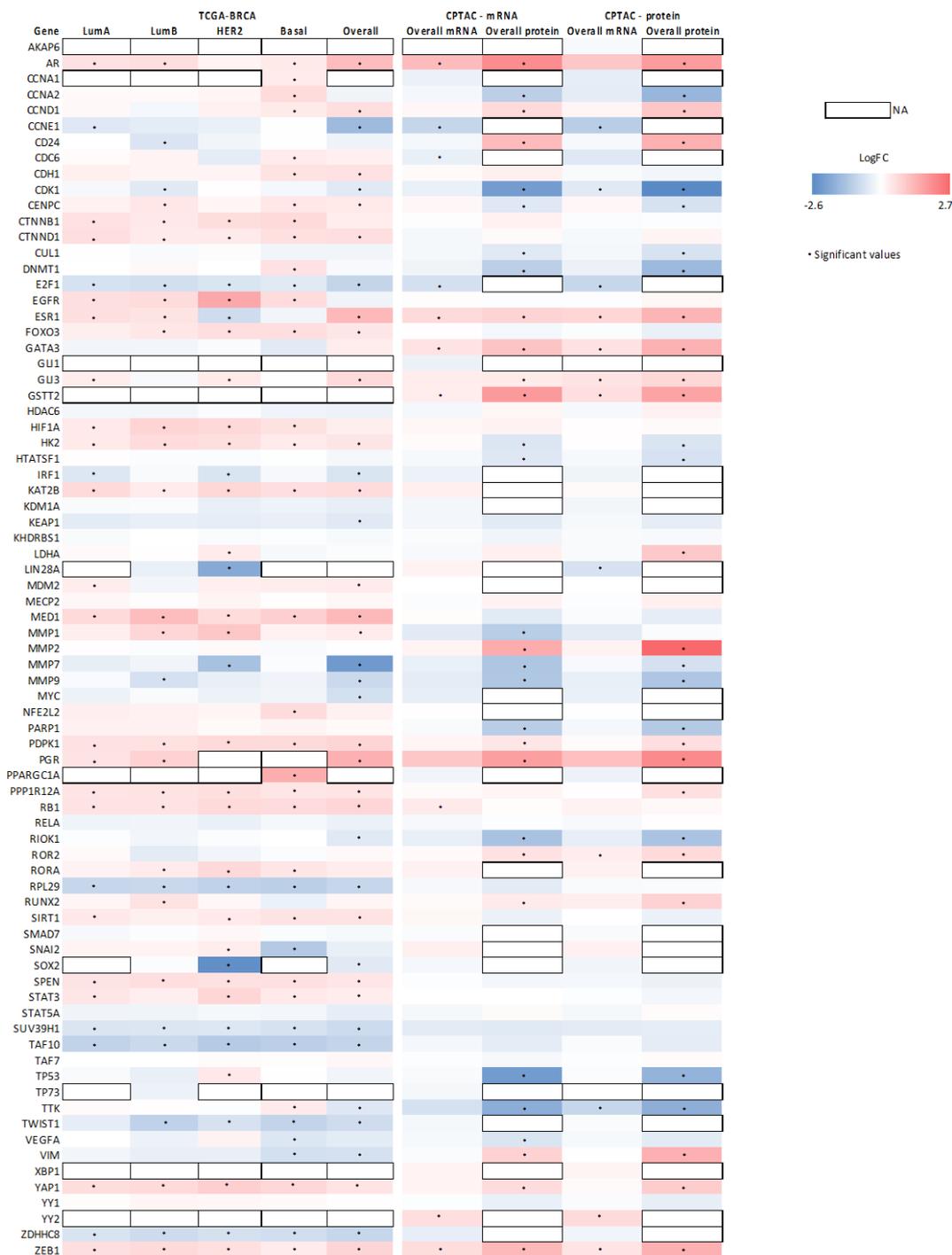


**Figure 3.** Top 10 biological processes overrepresented among the genes associated with *SETD7* differential expression (high and low) and unique for each subtype.

Next, the expression of genes with reported functional connection to *SETD7* was investigated. A list of 83 genes, including the 42 known *SETD7* targets plus other genes reported to be associated with *SETD7* function, was used to query the genes differentially expressed in high- versus low-*SETD7* tumours for each subtype (Figure 5). Interestingly, some of these genes were consistently associated with high-*SETD7* (*AR*, *CTNNB1*, *CTNND1*, *EGFR*, *FOXO3*, *HIF1A*, *HK2*, *KAT2B*, *MED1*, *NFE2L2*, *PDPK1*, *PPP1R12A*, *RB1*, *RORA*, *SIRT1*, *SPEN*, *STAT3*, *YAP1*, and *ZEB1*) or low-*SETD7* (*E2F1*, *IRF1*, *MMP7*, *MMP9*, *RPL29*, *SUV39H1*, *TAF10*, *TWIST1*, and *ZFHHC8*) independently of the subtype. Others were dependent on the subtype: *CCNA1*, *DNMT1*, *PPARGC1A*, and *TTK* were associated with high-*SETD7* and *SNAI2* with low-*SETD7* for the basal-like subtype; *LDHA* and *TP53* with high-*SETD7* and *ESR1* ( $\text{ER}\alpha$ ) and *SOX2* with low-*SETD7* for the Her2-enriched subtype. Moreover, in both luminal A and B subtypes, *ESR1* ( $\text{ER}\alpha$ ) and *PGR* ( $\text{PR}$ ) were associated with high-*SETD7*, highlighting an association of *SETD7* with endocrine treatment-predictive biomarkers.



**Figure 4.** Analysis of the shared biological processes overrepresented among the differentially expressed genes associated with SETD7 (high and low) and unique for each subtype. **(A)** Bubble plot representing all shared biological processes from the unique genes for each subtype; **(B)** GoChord showing all unique genes representing shared immune-related biological processes.



**Figure 5.** Heatmap showing the association of several genes of interest, including known SETD7 targets, with SETD7 differential expression using the TCGA-BRCA dataset (analysis performed both by subtype and by pooling all BC samples—overall) and CPTAC dataset (SETD7 stratified by mRNA or protein; analysis performed by pooling all BC—overall). Genes enriched in low-SETD7 group have negative log2 fold change (blue) and the ones enriched in the high-SETD7 group have positive log2 fold change (red). Genes with a significant adjusted *p*-value (Bonferroni post hoc, <0.05) have a star. Genes that were not present for a particular condition are represented as not available (NA, white boxes). Some genes were not detected in any dataset and were excluded from the figure: *ATOH1*, *ESR2* (ERβ), *GATA1*, *NANOG*, *NR1H4* (FXR), and *PDX1*.

In summary, our analysis showed that *SETD7* differential expression is correlated with the expression of different genes depending on the subtype, which may correspond to completely different biological processes (like cell adhesion for luminal A and lysosome organization and early endosome to late endosome transport for Her2-enriched subtype) or shared processes (like immune-related pathways). While no analysis per subtype showed significant results using CPTAC proteome data, it is to be noted that this is a relatively small cohort (total 122 samples), which does not reach a high power when dividing the samples by subtype. Thus, expanding this cohort would be beneficial for further studies.

### 3.4. Association between *SETD7* Expression and Its Target Proteins

*SETD7* is a methyltransferase with multiple known target proteins. Thus, the protein levels, phosphorylation, and acetylation patterns of *SETD7* targets were investigated in *SETD7*-high and -low groups, respectively. The proteins enriched in high- or low-*SETD7* groups (all BC subtypes pooled, mRNA data from TCGA PanCancer Atlas and protein data from CPTAC) were extracted and compared with a list of 42 known *SETD7* targets (Supplementary File S5 and Figure S10A). Nineteen targets were not found or were not significantly associated with *SETD7* expression in any dataset (light grey shade in Supplementary Table S3). For low-*SETD7* tumours, PARP-1 was present in all datasets from CPTAC; Cullin 1 in the total proteome; centromere protein C, HIV Tat (*HTATSF1*), RIO1, DNMT1, and TTK in total and phosphoproteome; Msx2-interacting protein (*SPEN*), catenin beta-1 (*CTNNB1*), TAF7, PPP1R12A, and SUV39H1 in phosphoproteome; MED1 and YY1 in phosphoproteome and acetylproteome; and Sam68 (*KHDRBS1*) and STAT3 in acetylproteome. For high-*SETD7* tumours, PPP1R12A, GLI3, YAP1, and AR were found in total and phosphoproteome from CPTAC; and pRb (*RB1*), RELA, ER $\alpha$ , MECP2, and *SPEN* in phosphoproteome. AR and ER $\alpha$  were also found in the TCGA PanCancer Atlas data.

When analysing the samples per subtype, only the total proteome showed significant correlations with *SETD7* mRNA or protein levels, mainly in the luminal A subtype (Supplementary Figure S10B), where ER $\alpha$  was associated with high *SETD7* mRNA expression (Supplementary Table S3). No correlations between *SETD7* differential protein levels and phosphoproteome or acetylproteome were observed by subtype.

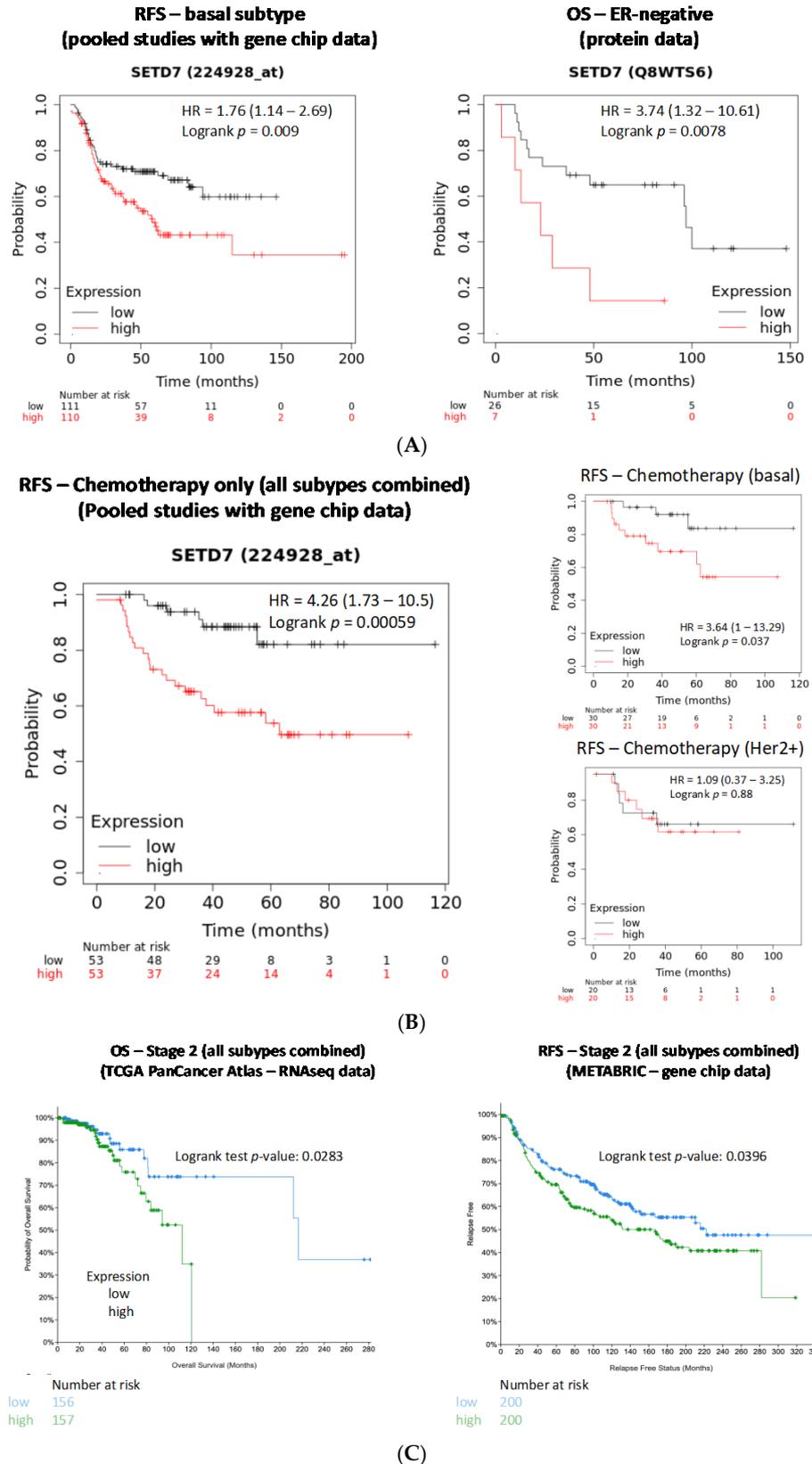
### 3.5. Association of *SETD7* Expression Levels with Breast Cancer Survival Outcomes

The prognostic value of *SETD7* in pooled samples from all BC subtypes was explored using the KMplotter online tool, HPA, and the datasets containing survival data available from cBioPortal. The association of high- or low-*SETD7* groups with RFS, DMFS, and OS was variable, varied between datasets, and did not show a clear association with either good or bad prognosis (Supplementary Table S4). This is in line with our recent findings reported in a systematic review [4]. Notably, a cohort analysing only ER (*ESR1*)-negative tumours showed that high-*SETD7* was significantly correlated with a poor prognosis. This led us to analyse the influence of clinical factors, including the molecular subtype, on the outcome of patients divided according to high or low *SETD7* expression.

#### 3.5.1. Influence of Histological and Molecular Subtype on Outcomes Associated with *SETD7* Expression

Survival outcomes were available in METABRIC and TCGA PanCancer Atlas datasets available at cBioPortal. Luminal A patients from the TCGA PanCancer Atlas cohort (244/499 total samples; RNA-seq data) exhibited a correlation between high *SETD7* mRNA and worse OS ( $p = 0.044$ ) and PFS ( $p = 0.032$ ) (Supplementary Table S5). However, when all microarray studies were combined (gene-chip data in KM plotter) high-*SETD7* correlated with good DMFS. In the luminal B subtype, high-*SETD7* correlated with bad DMFS only in one independent study (Supplementary Table S5). In the basal-like subtype, significant associations between high *SETD7* expression and worse RFS (gene-chip data; all the studies pooled, Figure 6A, left panel; Supplementary Table S5), DMFS and OS (one individual study). Analysis based on high *SETD7* protein data [50] in KM plotter also showed that

expression was associated with worse OS for ER-negative samples (33/65 total samples,  $p = 0.008$ ; Figure 6A, right panel).



**Figure 6.** Analysis of the effect of differential SETD7 expression on survival outcomes of BC patients. (A) Influence of SETD7 mRNA or protein expression levels on survival outcomes for patients with

the basal-like molecular subtype (left panel) or the ER-negative histological subtype (right panel), respectively; (B) Influence of *SETD7* expression levels (mRNA) for RFS outcomes following chemotherapy, combining all subtypes (left panel) or by subtype (right panel); (C) Influence of *SETD7* expression levels (mRNA) on survival outcomes for stage 2 tumours, using TCGA PanCancer Atlas (RNA-seq, left panel) or METABRIC cohort (gene chip, right panel). OS—overall survival; RFS—relapse / recurrence-free survival.

In conclusion, strong evidence suggests that expression of *SETD7* is predictive of a poor outcome for patients carrying basal-like tumours, even though this subtype has lower *SETD7* expression in comparison to the luminal A or Her2-enriched subtypes (Figure 1B,C). For all other subtypes, an association between *SETD7* expression and survival outcomes remains inconclusive.

### 3.5.2. Influence of *SETD7* Expression on Therapy Outcomes

Pooled gene-chip studies with information about therapy in KM plotter showed that for patients that only received chemotherapy, high *SETD7* mRNA was significantly correlated with bad RFS (106/211 total samples;  $p = 0.0006$ ; Figure 6B, left panel) and DMFS (84/168 total samples;  $p = 0.0012$ ). The same was observed when analysing two of the studies independently (Supplementary Table S6). Most of the patients who received chemotherapy only had basal-like (~60%) or Her2+ (~40%) tumours and, interestingly, high *SETD7* was correlated with worse RFS (Figure 6B, right panels) or DMFS (not shown) only for patients with the basal-like subtype. In patients receiving solely endocrine therapy, high *SETD7* was correlated with worse RFS but in only one study (METABRIC: 495/1025 total samples,  $p = 0.0325$ ; Supplementary Table S6). For patients that had not received any therapy or that had received both endocrine and chemotherapy, the results were inconclusive (Supplementary Table S6). Thus, *SETD7* could be a marker of chemoresistance for patients with basal-like tumours.

### 3.5.3. Influence of Tumour Stage on Outcomes Associated with *SETD7* Expression

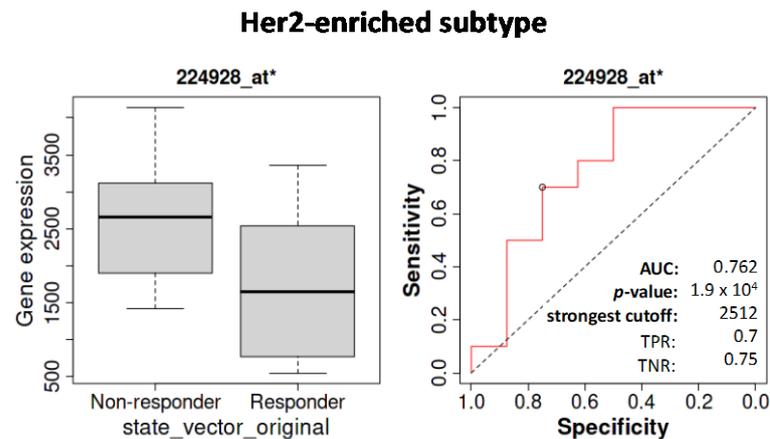
TCGA data analysed through the HPA showed that high *SETD7* expression (mean expression) was correlated with low survival for stage II patients (609 samples;  $p = 0.0003$ ). The same was observed using cBioPortal, where higher *SETD7* expression was correlated with lower OS (313/628 total samples;  $p = 0.0283$ ; Figure 6C, left panel) and DFS (312/628 total samples;  $p = 0.0569$ ; Supplementary Table S7) for stage 2 in TCGA PanCancer Atlas (confirming the results obtained in the same dataset using HPA; Supplementary Table S7) and lower RFS (400/979 total samples;  $p = 0.0396$ ) for stage 2 in METABRIC (Figure 6C, right panel). High *SETD7* protein expression [50] in KM plotter was also correlated with low OS ( $p = 0.036$ ) for stage 2 patients (46/65 total samples; Supplementary Table S7). It is important to note that stage 2 represents ~60% of all BC samples analysed, followed by stage 1 (~30%) and 3 (~10%). Stages 0 and 4 comprise the lowest percentages of tumours in the cohorts studied and analyses on these were limited. Also, most of the stage 2 tumours in these studies were of the luminal subtype (~60%). Even though one might assume these results suggest that *SETD7* could serve as a prognostic marker for luminal tumours in stage 2, this was not confirmed when we pooled stage 2 samples of luminal subtypes from TCGA PanCancer or METABRIC cohorts (not shown).

### 3.5.4. Influence of Tumour Grade, Lymph Node Status, and Metastasis on Survival Outcomes Associated with *SETD7* Expression

The influence of BC grade or lymph node status on the correlation of *SETD7* expression with survival outcomes was not clear, since few independent studies allowed this analysis, and the results did not agree (Supplementary Tables S8 and S9). No association between *SETD7* expression (lower vs. upper quartile) and metastasizing tumours was observed using MBC project in cBioPortal (not shown).

### 3.5.5. Predictive Power of SETD7

Using ROC plotter, no strong association between *SETD7* differential expression and hormone or chemotherapies was observed. However, patients that did not respond to anti-Her2 therapy expressed higher levels of *SETD7* (Figure 7). This may correlate with the higher *ERBB2* mutation rate in patients of the high-*SETD7* group.



**Figure 7.** ROCplotter analysis to study the correlation between *SETD7* mRNA expression (microarray data) and 5-year RFS for patients receiving any anti-Her2 therapy (29 responders and 21 non-responders). AUC—area under the curve; TNR—true-negative rate; TPR—true-positive rate.

## 4. Discussion

Current knowledge on *SETD7*'s impact on BC biology and its prognostic and predictive potential is scarce, with numerous contradictory findings [4,51]. In this work, we systematically analysed public datasets of BC samples to establish if *SETD7* expression is correlated with, or indicative of, diverse clinical conditions. The relevant biological processes associated with expression of *SETD7*, the genes involved, and their clinical significance were also evaluated. Stratification by molecular subtype, which has not previously been performed, showed that *SETD7* expression was dependent on subtype and that distinct processes were related to *SETD7* expression and could be clinically relevant.

Previous studies comparing *SETD7* expression between normal breast tissue and BC have not reported consistent results. We found that *SETD7* is significantly lower in BC than in adjacent normal tissue (TNMplot). This agrees with previous studies analysing mRNA [5,6] but not with studies analysing the protein level [7,8]. We observed a divergent relationship between mRNA and protein levels specifically for the luminal B subtype. This suggests that *SETD7* may be regulated post-transcriptionally, possibly by miR-372/373 [52], or post-translationally, possibly through TRIM21 [8]. For the remaining subtypes, *SETD7* mRNA and protein followed the same pattern. We observed a significantly higher expression in the Her2-enriched and luminal A compared with the basal-like subtype, which may be clinically relevant. This differential *SETD7* expression may be related to the higher frequency of *SETD7* gene loss that we noted for the basal-like subtype and, to some extent, the low-level gain of *SETD7* that we observed among the Her2-enriched subtype tumours.

A relationship between *SETD7* expression and prognosis was consistent only for patients with basal-like tumours, where high-*SETD7* was significantly associated with worse RFS, DMFS, and OS. This aligned with worse OS for ER-negative patients expressing high *SETD7* protein and with worse OS and DMFS in basal-like patients treated with chemotherapy. Higher *SETD7* mRNA and protein in Her2-enriched tumours was correlated with increased *ERBB2* amplification and corresponding *ERBB2* mRNA upregulation. This was associated with a significantly lower response to anti-Her-2 therapy in this subgroup. Still, no significant association with disease prognosis was found, but a trend of poorer OS and RFS could be observed in TCGA Pan Cancer Atlas dataset. Regarding luminal tumours, high *SETD7* was also correlated with worse RFS, but only in patients receiving

endocrine therapy and this association was not sufficiently strong, as the ROC plot did not support a prognostic value.

As the differential expression of *SETD7* between molecular subtypes may be clinically meaningful, we compared the transcriptomes between high- and low-*SETD7* groups. This showed how *SETD7* differential expression could impact the biology of the different molecular subtypes and reinforced the experimental data showing that *SETD7* function is context-dependent [4]. Thus, the hypothesis raised by this study should be validated in clinical specimens and stratified by molecular subtype. Although many genes associated with *SETD7* expression were different depending on subtype, in some cases, the same biological processes were overrepresented. This includes many immune-related processes. While immunotherapy has been increasingly used to treat cancer patients, this line of treatment has not been effective in BC, although some success has been noted for the triple-negative breast cancer subtype (mostly comprising the basal-like subtype) [53,54]. Herein, we show that the immune infiltration and response were highly correlated with *SETD7* expression, especially in luminal A and basal-like subtypes. The corresponding genes were primarily upregulated in low-*SETD7* luminal A tumours and also correlated with higher xCell Immune score (represented by signatures of B and T cells). Additionally, the upregulation of genes with a functional role in immune evasion (*PD1*, *FOXP3*, *CTLA4*, *IL17B* and the *IL17* receptors *IL17RE* and *IL17RC*) in the low-*SETD7* group of luminal A subtype supports the knowledge that lymphocyte infiltration is associated with worse prognosis in luminal subtypes [55,56]. Immunotherapy is not currently viewed as relevant in luminal A tumours; however, stratification by *SETD7* might improve the response rate of immune checkpoint inhibitors. On the contrary, in the basal-like subtype, immune-related genes were upregulated in the high-*SETD7* group, and this was correlated with higher T cell infiltration, including of CD8+, as inferred from their gene expression signatures. This usually corresponds to a better prognosis in the basal-like subtype [55,56], and thus suggests that the tumours expressing high *SETD7* might benefit from immunotherapy. However, future studies are needed to verify if stratification by *SETD7* alone or together with additional markers can improve selection of patients for immunotherapy in subgroups of luminal A and basal-like tumours.

In luminal subtypes, the two gold-standard biomarkers *ESR1* (ER $\alpha$ ) and *PGR* (PR) were associated with high *SETD7*. ER $\alpha$  is the target of endocrine treatments and a primary treatment-predictive marker in breast cancer [57]. However, many ER $\alpha$ -positive tumours develop endocrine resistance, where ER $\alpha$  is active in the absence of ligand. ER $\alpha$  is a known target of *SETD7* [2], which stabilizes ER $\alpha$  through methylation in lysine 302. It is not known if this stabilization contributes to endocrine resistance. However, the lower survival of luminal A high-*SETD7* patients from the TCGA dataset, along with the upregulation of *RUNX2* and *GPER1* (in luminal B, with reported roles in breast carcinogenesis [58] and endocrine resistance [59]), suggest a role for *SETD7* in endocrine resistance. The idea of targeting *SETD7* to overcome endocrine resistance thus deserves further testing. This may be specifically relevant for the luminal B subtype, where the most significant biological processes overrepresented in the high-*SETD7* (mRNA) group were the ubiquitin-dependent ERAD pathway and the positive regulation of autophagy, which is also linked to the ubiquitin-proteasome system (UPS, also overrepresented). The association with high *SETD7* was not strong, but given that these two pathways underly endocrine resistance [60,61], the additive contribution of all these genes to these processes should not be discarded. Autophagy is associated with the suppression of tumour initiation [62] and the survival of dormant BC stem cells and metastatic tumour recurrence [62,63]. Many preclinical studies have shown that autophagy inhibition improves endocrine therapy response [64]. Although the ROC plotter did not find a correlation between high *SETD7* mRNA stratification and 5-year RFS (suggestive of resistance to endocrine therapy), we need to consider that we could not show a correlation between *SETD7* mRNA and protein levels in this subtype. Further, these results deserve further validation, as the majority of patients were treated with tamoxifen.

In the Her2-enriched subtype, lysosome organization and early endosome to late endosome transport were overrepresented in the high-*SETD7* group. Activation of these two biological processes has been linked to anti-Her2 therapy resistance [65,66]. Remarkably, *EGFR* and *ERBB3* were strongly associated with high *SETD7* and also connected to anti-Her2 therapy resistance [66]. These results together with the correlation of high *SETD7* with higher amplification of *ERBB2* corroborates the ROC plotter results, where high *SETD7* was correlated with patients that did not respond well (shorter 5-year RFS) to anti-Her2 therapy. A recent study suggests *SORL1* to be a candidate therapeutic target to complement and potentiate anti-Her2 therapy [67]. Indeed, *SORL1* expression was significantly associated with high *SETD7* expression, pointing to a potential benefit of targeting *SETD7* alone or together with *SORL1* in patients with high *SETD7* in order to overcome resistance.

A limitation of our study is that multivariable adjustment was unavailable in the tools used to analyse the survival and prognosis value of *SETD7*. Moreover, information to correct for confounding effects was lacking, which restricts the conclusions that can be drawn about *SETD7* expression as an independent factor of diagnosis or resistance to therapies. Further studies will be needed to validate the clinical impacts.

In previous preclinical studies, *SETD7* function has consistently been associated with altered cellular response to DNA damage stimulus, including hypoxia and oxidative stress and independently of *TP53* status [4]. Genes related to the cellular response to DNA damage stimulus and DNA repair were common to all subtypes (Supplementary Files S3 and S4). Two of the major players in the DNA damage response pathway, *ATR* and *ATM* were highly expressed in the high-*SETD7* group in all subtypes. This indicates that chemotherapy might be less efficient in tumours expressing high *SETD7*. This was also supported by the poor prognosis associated with *SETD7* expression when analysing all BC subtypes from TCGA. When analysis was carried out by BC subtype, the basal-like subtype also showed unique genes associated with *SETD7* expression and strongly overrepresented in cellular response to DNA damage and DNA repair-related pathways were. This was in line with poor outcome after chemotherapy for patients with basal-like tumours expressing high *SETD7*. In the future and based on previous findings showing that inhibition of *SETD7* in other types of cancer improves response to chemotherapy [10,68–72], it would be interesting to explore if this subgroup of patients could benefit from targeting *SETD7* with inhibitors to improve chemotherapy response.

## 5. Conclusions

*SETD7* expression appears strongly associated with tumour stromal and immune signatures and related to therapy resistance. In the basal-like BC subtype, high *SETD7* expression was consistently predictive of bad prognosis, and this group was enriched in immune signatures. The unique genes associated with *SETD7* expression were strongly overrepresented in cellular response to DNA damage and DNA repair-related pathways, and this was aligned with poor outcome after chemotherapy. Future studies should focus on the identification of the differentially expressed genes that could constitute markers to aid decisions on prescribing immune therapy and test if inhibiting *SETD7* improves basal-like response to chemotherapy. In the Her2-enriched subtype, high *SETD7* may also have a predictive value, since *SETD7* expression was associated with *ERBB2* copy number amplification and worse response to anti-Her2 therapy, as well as upregulation of *EGFR*, *HER-3*, and overrepresentation of such biological processes as lysosome organization and early endosome to late endosome transport known to be underlying mechanisms of anti-Her-2 therapy resistance. In the luminal subtype, high *SETD7* expression was associated with higher *ESR1* ( $ER\alpha$ ), *PGR* (PR), *RUNX2* and *GPER1*, which together with previous findings on the role of *SETD7* maintaining  $ER\alpha$  protein stability and activity, highlight the need for further studies on the role of *SETD7* in endocrine resistance. Still, no consistent relationship with prognosis was found, except for worse OS in tumours with high *SETD7* and treated with endocrine therapy.

In summary, this study emphasizes that there is clinical potential in the study of SETD7, which must be evaluated in the context of the BC molecular subtype.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/cancers14246029/s1>. Figure S1: SETD7 genomic alterations in breast cancer using cBioPortal; Figure S2: SETD7 mRNA expression in breast cancer PAM50 subtypes using cBioPortal datasets; Figure S3: Mean differences of each comparison done in Figure 1B and the confidence intervals associated with it; Figure S4: Predicted infiltration of cancer-associated fibroblasts (CAFs), endothelial cells and cytotoxicity score in samples with differential SETD7 mRNA expression (high vs. low) from different breast cancer PAM50 subtypes; Figure S5: Predicted infiltration of immune cells in samples with differential SETD7 mRNA expression (high vs. low) from different breast cancer PAM50 subtypes; Figure S6: Genomic alterations associated with SETD7 mRNA differential expression from TCGA PanCancer atlas in cBioPortal; Figure S7: TCGA-BRCA data processing; Figure S8: Analysis of genes associated with SETD7 mRNA differential expression per subtype from TCGA-BRCA; Figure S9: Top 10 biological processes overrepresented by the genes associated with SETD7 differential expression (high and low distinctly) and unique for each subtype; Figure S10: Analysis of proteins associated with SETD7 protein differential expression using CPTAC; Table S1: Breast cancer datasets with SETD7 expression (mRNA or protein), mutation or copy number information.; Table S2: Association of SETD7 expression with therapy, grade and stage using datasets in cBioPortal; Table S3: Association of SETD7 differential expression and known SETD7 targets and their methylation sites and other sites known to compete with SETD7 methylation; Table S4: SETD7 association with survival outcomes; Table S5: Prognosis associated with high SETD7 in BC subtypes; Table S6: Prognosis associated with high SETD7 in BC considering therapy; Table S7: Prognosis associated with high SETD7 in BC considering stage; Table S8: Prognosis associated with high SETD7 in BC considering grade; Table S9: Prognosis associated with high SETD7 in BC considering lymph node status; Supplementary File S1: Mutations in SETD7 gene; Supplementary File S2: Genomic alterations correlating with SETD7 differential expression; Supplementary File S3: List of significant genes per subtype and common to all subtypes; Supplementary File S4: Biological processes list; Supplementary File S5: Proteome enriched in high or low SETD7 protein from CPTAC [2,68,73–110].

**Author Contributions:** Conceptualization, L.A.H., C.W. and F.L.M.; methodology, F.L.M. and L.S.; formal analysis, F.L.M.; investigation, F.L.M.; data curation, F.L.M.; writing—original draft preparation, F.L.M.; writing—review and editing, F.L.M., L.A.H., C.W. and L.S.; supervision, L.A.H. and C.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by iBiMED research unit UIDB/04501/2020 and UIDP/04501/2020, MEDISIS (CENTRO-01-0246-FEDER-000018) supported by Comissão de Coordenação e Desenvolvimento Regional do Centro. FLM thanks the Portuguese Science and Technology Foundation—FCT for her PhD scholarship SFRH/BD/117818/2016; C.W acknowledges the Swedish Cancer Society (21 1632 Pj), Karolinska Institute PhD for support for L.S. (KID 2021-00501) and Region Stockholm (HMT RS2021-0316).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The results shown here are in part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>, accessed on 5 August 2022. We would like to thank Madeleine Birgersson for the helpful discussions that contributed to the development of this project.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Batista, I.D.A.A.; Helguero, L.A. Biological Processes and Signal Transduction Pathways Regulated by the Protein Methyltransferase SETD7 and Their Significance in Cancer. *Signal Transduct. Target. Ther.* **2018**, *3*, 19. [[CrossRef](#)] [[PubMed](#)]
2. Subramanian, K.; Jia, D.; Kapoor-Vazirani, P.; Powell, D.R.; Collins, R.E.; Sharma, D.; Peng, J.; Cheng, X.; Vertino, P.M. Regulation of Estrogen Receptor Alpha by the SET7 Lysine Methyltransferase. *Mol. Cell* **2008**, *30*, 336–347. [[CrossRef](#)] [[PubMed](#)]

3. Huang, R.; Li, X.; Yu, Y.; Ma, L.; Liu, S.; Zong, X.; Zheng, Q. SETD7 Is a Prognosis Predicting Factor of Breast Cancer and Regulates Redox Homeostasis. *Oncotarget* **2017**, *8*, 94080–94090. [[CrossRef](#)]
4. Monteiro, F.L.; Williams, C.; Helguero, L.A. A Systematic Review to Define the Multi-Faceted Role of Lysine Methyltransferase SETD7 in Cancer. *Cancers* **2022**, *14*, 1414. [[CrossRef](#)] [[PubMed](#)]
5. Gu, Y.; Wang, X.; Liu, H.; Li, G.; Yu, W.; Ma, Q. SET7/9 Promotes Hepatocellular Carcinoma Progression through Regulation of E2F1. *Oncol. Rep.* **2018**, *40*, 1863–1874. [[CrossRef](#)] [[PubMed](#)]
6. Song, Y.; Zhang, J.; Tian, T.; Fu, X.; Wang, W.; Li, S.; Shi, T.; Suo, A.; Ruan, Z.; Guo, H.; et al. SET7/9 Inhibits Oncogenic Activities through Regulation of Gli-1 Expression in Breast Cancer. *Tumor Biol.* **2016**, *37*, 9311–9322. [[CrossRef](#)]
7. Zhang, Y.; Liu, J.; Lin, J.; Zhou, L.; Song, Y.; Wei, B.; Luo, X.; Chen, Z.; Chen, Y.; Xiong, J.; et al. The Transcription Factor GATA1 and the Histone Methyltransferase SET7 Interact to Promote VEGF-Mediated Angiogenesis and Tumor Growth and Predict Clinical Outcome of Breast Cancer. *Oncotarget* **2016**, *7*, 9859–9875. [[CrossRef](#)] [[PubMed](#)]
8. Si, W.; Zhou, J.; Zhao, Y.; Zheng, J.; Cui, L. SET7/9 Promotes Multiple Malignant Processes in Breast Cancer Development via RUNX2 Activation and Is Negatively Regulated by TRIM21. *Cell Death Dis.* **2020**, *11*, 151. [[CrossRef](#)]
9. Duan, B.; Bai, J.; Qiu, J.; Wang, J.; Tong, C.; Wang, X.; Miao, J.; Li, Z.; Li, W.; Yang, J.; et al. Histone-Lysine N-Methyltransferase SETD7 Is a Potential Serum Biomarker for Colorectal Cancer Patients. *EBioMedicine* **2018**, *37*, 134–143. [[CrossRef](#)]
10. Lezina, L.; Aksenova, V.; Fedorova, O.; Malikova, D.; Shuvalov, O.; Antonov, A.V.; Tentler, D.; Garabadgiu, A.V.; Melino, G.; Barlev, N.A. KMT Set7/9 Affects Genotoxic Stress Response via the Mdm2 Axis. *Oncotarget* **2015**, *6*, 25843–25855. [[CrossRef](#)]
11. Montenegro, M.F.; Sánchez-Del-Campo, L.; González-Guerrero, R.; Martínez-Barba, E.; Piñero-Madrona, A.; Cabezas-Herrera, J.; Rodríguez-López, J.N. Tumor Suppressor SET9 Guides the Epigenetic Plasticity of Breast Cancer Cells and Serves as an Early-Stage Biomarker for Predicting Metastasis. *Oncogene* **2016**, *35*, 6143–6152. [[CrossRef](#)] [[PubMed](#)]
12. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the CBioPortal. *Sci. Signal.* **2013**, *6*, p11. [[CrossRef](#)] [[PubMed](#)]
13. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The CBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* **2012**, *2*, 401–404. [[CrossRef](#)] [[PubMed](#)]
14. Krug, K.; Jaehnig, E.J.; Satpathy, S.; Blumenberg, L.; Karpova, A.; Anurag, M.; Miles, G.; Mertins, P.; Geffen, Y.; Tang, L.C.; et al. Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell* **2020**, *183*, 1436–1456.e31. [[CrossRef](#)] [[PubMed](#)]
15. Curtis, C.; Shah, S.P.; Chin, S.F.; Turashvili, G.; Rueda, O.M.; Dunning, M.J.; Speed, D.; Lynch, A.G.; Samarajiwa, S.; Yuan, Y.; et al. The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups. *Nature* **2012**, *486*, 346–352. [[CrossRef](#)] [[PubMed](#)]
16. Liu, J.; Lichtenberg, T.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatich, A.J.; Benz, C.C.; Levine, D.A.; Lee, A.V.; et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **2018**, *173*, 400–416.e11. [[CrossRef](#)]
17. Lefebvre, C.; Bachelot, T.; Filleron, T.; Pedrero, M.; Campone, M.; Soria, J.C.; Massard, C.; Lévy, C.; Arnedos, M.; Lacroix-Triki, M.; et al. Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med.* **2016**, *13*, e1002201. [[CrossRef](#)]
18. Razavi, P.; Chang, M.T.; Xu, G.; Bandlamudi, C.; Ross, D.S.; Vasan, N.; Cai, Y.; Bielski, C.M.; Donoghue, M.T.A.; Jonsson, P.; et al. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **2018**, *34*, 427–438.e6. [[CrossRef](#)]
19. Razavi, P.; Dickler, M.N.; Shah, P.D.; Toy, W.; Brown, D.N.; Won, H.H.; Li, B.T.; Shen, R.; Vasan, N.; Modi, S.; et al. Alterations in PTEN and ESR1 Promote Clinical Resistance to Alpelisib plus Aromatase Inhibitors. *Nat. Cancer* **2020**, *1*, 382. [[CrossRef](#)]
20. Nixon, M.J.; Formisano, L.; Mayer, I.A.; Estrada, M.V.; González-Ericsson, P.I.; Isakoff, S.J.; Forero-Torres, A.; Won, H.; Sanders, M.E.; Solit, D.B.; et al. PIK3CA and MAP3K1 Alterations Imply Luminal A Status and Are Associated with Clinical Benefit from Pan-PI3K Inhibitor Buparlisib and Letrozole in ER+ Metastatic Breast Cancer. *NPJ Breast Cancer* **2019**, *5*, 31. [[CrossRef](#)]
21. Pareja, F.; Brown, D.N.; Lee, J.Y.; Paula, A.D.C.; Selenica, P.; Bi, R.; Geyer, F.C.; Gazzo, A.; da Silva, E.M.; Vahdatinia, M.; et al. Whole-Exome Sequencing Analysis of the Progression from Non-Low-Grade Ductal Carcinoma in Situ to Invasive Ductal Carcinoma. *Clin. Cancer Res.* **2020**, *26*, 3682–3693. [[CrossRef](#)] [[PubMed](#)]
22. Kan, Z.; Ding, Y.; Kim, J.; Jung, H.H.; Chung, W.; Lal, S.; Cho, S.; Fernandez-Banet, J.; Lee, S.K.; Kim, S.W.; et al. Multi-Omics Profiling of Younger Asian Breast Cancers Reveals Distinctive Molecular Signatures. *Nat. Commun.* **2018**, *9*, 1725. [[CrossRef](#)]
23. Shah, S.P.; Roth, A.; Goya, R.; Oloumi, A.; Ha, G.; Zhao, Y.; Turashvili, G.; Ding, J.; Tse, K.; Haffari, G.; et al. The Clonal and Mutational Evolution Spectrum of Primary Triple Negative Breast Cancers. *Nature* **2012**, *486*, 395–399. [[CrossRef](#)] [[PubMed](#)]
24. Banerji, S.; Cibulskis, K.; Rangel-Escareno, C.; Brown, K.K.; Carter, S.L.; Frederick, A.M.; Lawrence, M.S.; Sivachenko, A.Y.; Sougnez, C.; Zou, L.; et al. Sequence Analysis of Mutations and Translocations across Breast Cancer Subtypes. *Nature* **2012**, *486*, 405–409. [[CrossRef](#)] [[PubMed](#)]
25. Stephens, P.J.; Tarpey, P.S.; Davies, H.; Van Loo, P.; Greenman, C.; Wedge, D.C.; Nik-Zainal, S.; Martin, S.; Varela, I.; Bignell, G.R.; et al. The Landscape of Cancer Genes and Mutational Processes in Breast Cancer. *Nature* **2012**, *486*, 400. [[CrossRef](#)] [[PubMed](#)]
26. Li, Q.; Jiang, B.; Guo, J.; Shao, H.; Del Priore, I.S.; Chang, Q.; Kudo, R.; Li, Z.; Razavi, P.; Liu, B.; et al. INK4 Tumor Suppressor Proteins Mediate Resistance to CDK4/6 Kinase Inhibitors. *Cancer Discov.* **2022**, *12*, 356–371. [[CrossRef](#)]

27. Bartha, Á.; Gyórfy, B. TNMplot.Com: A Web Tool for the Comparison of Gene Expression in Normal, Tumor and Metastatic Tissues. *Int. J. Mol. Sci.* **2021**, *22*, 2622. [[CrossRef](#)]
28. Thul, P.J.; Lindskog, C. The Human Protein Atlas: A Spatial Map of the Human Proteome. *Protein Sci.* **2018**, *27*, 233–244. [[CrossRef](#)]
29. Fekete, J.T.; Gyórfy, B. ROCplot.Org: Validating Predictive Biomarkers of Chemotherapy/Hormonal Therapy/Anti-HER2 Therapy Using Transcriptomic Data of 3,104 Breast Cancer Patients. *Int. J. Cancer* **2019**, *145*, 3140–3151. [[CrossRef](#)]
30. Li, Q.; Birkbak, N.J.; Gyórfy, B.; Szallasi, Z.; Eklund, A.C. Jetset: Selecting the Optimal Microarray Probe Set to Represent a Gene. *BMC Bioinform.* **2011**, *12*, 474. [[CrossRef](#)]
31. Mounir, M.; Lucchetta, M.; Silva, T.C.; Olsen, C.; Bontempi, G.; Chen, X.; Noushmehr, H.; Colaprico, A.; Papaleo, E. New Functionalities in the TCGAbiolinks Package for the Study and Integration of Cancer Data from GDC and GTEx. *PLoS Comput. Biol.* **2019**, *15*, e1006701. [[CrossRef](#)] [[PubMed](#)]
32. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. TCGAbiolinks: An R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* **2016**, *44*, e71. [[CrossRef](#)] [[PubMed](#)]
33. Silva, T.C.; Colaprico, A.; Olsen, C.; D'Angelo, F.; Bontempi, G.; Ceccarelli, M.; Noushmehr, H. TCGA Workflow: Analyze Cancer Genomics and Epigenomics Data Using Bioconductor Packages. *F1000Research* **2016**, *5*, 1542. [[CrossRef](#)] [[PubMed](#)]
34. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)] [[PubMed](#)]
35. McCarthy, D.J.; Chen, Y.; Smyth, G.K. Differential Expression Analysis of Multifactor RNA-Seq Experiments with Respect to Biological Variation. *Nucleic Acids Res.* **2012**, *40*, 4288–4297. [[CrossRef](#)]
36. Chen, Y.; Lun, A.T.L.; Smyth, G.K.; Burden, C.J.; Ryan, D.P.; Khang, T.F.; Lianoglou, S. From Reads to Genes to Pathways: Differential Expression Analysis of RNA-Seq Experiments Using Rsubread and the EdgeR Quasi-Likelihood Pipeline. *F1000Research* **2016**, *5*, 1438.
37. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)] [[PubMed](#)]
38. Sturm, G.; Finotello, F.; Petitprez, F.; Zhang, J.D.; Baumbach, J.; Fridman, W.H.; List, M.; Aneichyk, T. Comprehensive Evaluation of Transcriptome-Based Cell-Type Quantification Methods for Immuno-Oncology. *Bioinformatics* **2019**, *35*, i436–i445. [[CrossRef](#)]
39. Oliveros, J.C. VENNY. An Interactive Tool for Comparing Lists with Venn's Diagrams. Available online: <https://bioinfogp.cnb.csic.es/tools/venny/> (accessed on 24 March 2022).
40. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* **2008**, *4*, 44–57. [[CrossRef](#)]
41. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res.* **2009**, *37*, 1–13. [[CrossRef](#)]
42. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
43. Walter, W.; Sánchez-Cabo, F.; Ricote, M. GOpot: An R Package for Visually Combining Expression Data with Functional Analysis. *Bioinformatics* **2015**, *31*, 2912–2914. [[CrossRef](#)]
44. Winter, S.C.; Buffa, F.M.; Silva, P.; Miller, C.; Valentine, H.R.; Turley, H.; Shah, K.A.; Cox, G.J.; Corbridge, R.J.; Homer, J.J.; et al. Relation of a Hypoxia Metagene Derived from Head and Neck Cancer to Prognosis of Multiple Cancers. *Cancer Res.* **2007**, *67*, 3441–3449. [[CrossRef](#)]
45. Buffa, F.M.; Harris, A.L.; West, C.M.; Miller, C.J. Large Meta-Analysis of Multiple Cancers Reveals a Common, Compact and Highly Prognostic Hypoxia Metagene. *Br. J. Cancer* **2010**, *102*, 428. [[CrossRef](#)]
46. Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Treviño, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* **2013**, *4*, 2612. [[CrossRef](#)] [[PubMed](#)]
47. Aran, D.; Hu, Z.; Butte, A.J. XCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol.* **2017**, *18*, 220. [[CrossRef](#)] [[PubMed](#)]
48. Becht, E.; Giraldo, N.A.; Lacroix, L.; Buttard, B.; Elarouci, N.; Petitprez, F.; Selves, J.; Laurent-Puig, P.; Sautès-Fridman, C.; Fridman, W.H.; et al. Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol.* **2016**, *17*, 218. [[CrossRef](#)]
49. Daks, A.; Vasileva, E.; Fedorova, O.; Shuvalov, O.; Barlev, N.A. The Role of Lysine Methyltransferase SET7/9 in Proliferation and Cell Stress Response. *Life* **2022**, *12*, 362. [[CrossRef](#)]
50. Tang, W.; Zhou, M.; Dorsey, T.H.; Prieto, D.A.; Wang, X.W.; Ruppin, E.; Veenstra, T.D.; Ambs, S. Integrated Proteotranscriptomics of Breast Cancer Reveals Globally Increased Protein-MRNA Concordance Associated with Subtypes and Survival. *Genome Med.* **2018**, *10*, 94. [[CrossRef](#)] [[PubMed](#)]
51. Gu, Y.; Zhang, X.; Yu, W.; Dong, W. Oncogene or Tumor Suppressor: The Coordinative Role of Lysine Methyltransferase SET7/9 in Cancer Development and the Related Mechanisms. *J. Cancer* **2022**, *13*, 623–640. [[CrossRef](#)] [[PubMed](#)]
52. Wang, L.Q.; Yu, P.; Li, B.; Guo, Y.H.; Liang, Z.R.; Zheng, L.L.; Yang, J.H.; Xu, H.; Liu, S.; Zheng, L.S.; et al. MiR-372 and MiR-373 Enhance the Stemness of Colorectal Cancer Cells by Repressing Differentiation Signaling Pathways. *Mol. Oncol.* **2018**, *12*, 1949–1964. [[CrossRef](#)] [[PubMed](#)]

53. Schmid, P.; Adams, S.; Rugo, H.S.; Schneeweiss, A.; Barrios, C.H.; Iwata, H.; Diéras, V.; Hegg, R.; Im, S.-A.; Shaw Wright, G.; et al. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. *N. Engl. J. Med.* **2018**, *379*, 2108–2121. [[CrossRef](#)] [[PubMed](#)]
54. Schmid, P.; Cortes, J.; Pusztai, L.; McArthur, H.; Kümmel, S.; Bergh, J.; Denkert, C.; Park, Y.H.; Hui, R.; Harbeck, N.; et al. Pembrolizumab for Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **2020**, *382*, 810–821. [[CrossRef](#)] [[PubMed](#)]
55. Pellegrino, B.; Hlavata, Z.; Migali, C.; De Silva, P.; Aiello, M.; Willard-Gallo, K.; Musolino, A.; Solinas, C. Luminal Breast Cancer: Risk of Recurrence and Tumor-Associated Immune Suppression. *Mol. Diagnosis Ther.* **2021**, *25*, 409–424. [[CrossRef](#)] [[PubMed](#)]
56. Goldberg, J.; Pastorello, R.G.; Vallius, T.; Davis, J.; Cui, Y.X.; Agudo, J.; Waks, A.G.; Keenan, T.; McAllister, S.S.; Tolaney, S.M.; et al. The Immunology of Hormone Receptor Positive Breast Cancer. *Front. Immunol.* **2021**, *12*, 1515. [[CrossRef](#)] [[PubMed](#)]
57. Frigo, D.E.; Bondesson, M.; Williams, C. Nuclear Receptors: From Molecular Mechanisms to Therapeutics. *Essays Biochem.* **2021**, *65*, 847. [[CrossRef](#)] [[PubMed](#)]
58. Wysokinski, D.; Blasiak, J.; Pawlowska, E. Role of RUNX2 in Breast Carcinogenesis. *Int. J. Mol. Sci.* **2015**, *16*, 20969–20993. [[CrossRef](#)] [[PubMed](#)]
59. Pepermans, R.A.; Prossnitz, E.R. ER $\alpha$ -Targeted Endocrine Therapy, Resistance and the Role of GPER. *Steroids* **2019**, *152*, 108493. [[CrossRef](#)]
60. Direito, I.; Fardilha, M.; Helguero, L.A. Contribution of the Unfolded Protein Response to Breast and Prostate Tissue Homeostasis and Its Significance to Cancer Endocrine Response. *Carcinogenesis* **2018**, *40*, 203–215. [[CrossRef](#)]
61. Direito, I.; Monteiro, L.; Melo, T.; Figueira, D.; Lobo, J.; Enes, V.; Moura, G.; Henrique, R.; Santos, M.A.S.; Jerónimo, C.; et al. Protein Aggregation Patterns Inform about Breast Cancer Response to Antiestrogens and Reveal the RNA Ligase RTCB as Mediator of Acquired Tamoxifen Resistance. *Cancers* **2021**, *13*, 3195. [[CrossRef](#)]
62. La Belle Flynn, A.; Schiemann, W.P. Autophagy in Breast Cancer Metastatic Dormancy: Tumor Suppressing or Tumor Promoting Functions? *J. Cancer Metastasis Treat.* **2019**, *2019*, 43. [[CrossRef](#)] [[PubMed](#)]
63. Vera-Ramirez, L.; Vodnala, S.K.; Nini, R.; Hunter, K.W.; Green, J.E. Autophagy Promotes the Survival of Dormant Breast Cancer Cells and Metastatic Tumour Recurrence. *Nat. Commun.* **2018**, *9*, 1944. [[CrossRef](#)]
64. Cook, K.L.; Shajahan, A.N.; Clarke, R. Autophagy and Endocrine Resistance in Breast Cancer. *Expert Rev. Anticancer Ther.* **2011**, *11*, 1283. [[CrossRef](#)] [[PubMed](#)]
65. Mishra, A.; Hourigan, D.; Lindsay, A.J. Inhibition of the Endosomal Recycling Pathway Downregulates HER2 Activation and Overcomes Resistance to Tyrosine Kinase Inhibitors in HER2-Positive Breast Cancer. *Cancer Lett.* **2022**, *529*, 153–167. [[CrossRef](#)]
66. Hunter, F.W.; Barker, H.R.; Lipert, B.; Rothé, F.; Gebhart, G.; Piccart-Gebhart, M.J.; Sotiriou, C.; Jamieson, S.M.F. Mechanisms of Resistance to Trastuzumab Emtansine (T-DM1) in HER2-Positive Breast Cancer. *Br. J. Cancer* **2019**, *122*, 603–612. [[CrossRef](#)] [[PubMed](#)]
67. Pietilä, M.; Sahgal, P.; Peuhu, E.; Jäntti, N.Z.; Paatero, I.; Närvä, E.; Al-Akhrass, H.; Lilja, J.; Georgiadou, M.; Andersen, O.M.; et al. SORLA Regulates Endosomal Trafficking and Oncogenic Fitness of HER2. *Nat. Commun.* **2019**, *10*, 2340. [[CrossRef](#)] [[PubMed](#)]
68. Kontaki, H.; Talianidis, I. Lysine Methylation Regulates E2F1-Induced Cell Death. *Mol. Cell* **2010**, *39*, 152–160. [[CrossRef](#)] [[PubMed](#)]
69. Lezina, L.; Aksenova, V.; Ivanova, T.; Purmessur, N.; Antonov, A.V.; Tentler, D.; Fedorova, O.; Garabadgiu, A.V.; Talianidis, I.; Melino, G.; et al. KMTase Set7/9 Is a Critical Regulator of E2F1 Activity upon Genotoxic Stress. *Cell Death Differ.* **2014**, *21*, 1889–1899. [[CrossRef](#)]
70. Daks, A.; Mamontova, V.; Fedorova, O.; Petukhov, A.; Shuvalov, O.; Parfenyev, S.; Netsvetay, S.; Venina, A.; Kizenko, A.; Imyaninov, E.; et al. Set7/9 Controls Proliferation and Genotoxic Drug Resistance of NSCLC Cells. *Biochem. Biophys. Res. Commun.* **2021**, *572*, 41–48. [[CrossRef](#)] [[PubMed](#)]
71. Chuikov, S.; Kurash, J.K.; Wilson, J.R.; Xiao, B.; Justin, N.; Ivanov, G.S.; McKinney, K.; Tempst, P.; Prives, C.; Gamblin, S.J.; et al. Regulation of P53 Activity through Lysine Methylation. *Nature* **2004**, *432*, 353–360. [[CrossRef](#)]
72. Wang, C.; Shu, L.; Zhang, C.; Li, W.; Wu, R.; Guo, Y.; Yang, Y.; Kong, A.N. Histone Methyltransferase Setd7 Regulates Nrf2 Signaling Pathway by Phenethyl Isothiocyanate and Ursolic Acid in Human Prostate Cancer Cells. *Mol. Nutr. Food Res.* **2018**, *62*, e1700840. [[CrossRef](#)]
73. Dhayalan, A.; Kudithipudi, S.; Rathert, P.; Jeltsch, A. Specificity Analysis-Based Identification of New Methylation Targets of the SET7/9 Protein Lysine Methyltransferase. *Chem. Biol.* **2011**, *18*, 111–120. [[CrossRef](#)]
74. Ko, S.; Ahn, J.; Song, C.S.; Kim, S.; Knapczyk-Stwora, K.; Chatterjee, B. Lysine Methylation and Functional Modulation of Androgen Receptor by Set9 Methyltransferase. *Mol. Endocrinol.* **2011**, *25*, 433–444. [[CrossRef](#)] [[PubMed](#)]
75. Gaughan, L.; Stockley, J.; Wang, N.; McCracken, S.R.; Treumann, A.; Armstrong, K.; Shaheen, F.; Watt, K.; McEwan, I.J.; Wang, C.; et al. Regulation of the androgen receptor by SET9-mediated methylation. *Nucleic Acids Res.* **2010**, *39*, 1266–1279. [[CrossRef](#)] [[PubMed](#)]
76. Shen, C.; Wang, D.; Liu, X.; Gu, B.; Du, Y.; Wei, F.; Cao, L.; Song, B.; Lu, X.; Yang, Q.; et al. SET7/9 regulates cancer cell proliferation by influencing  $\beta$ -catenin stability. *FASEB J.* **2015**, *29*, 4313–4323. [[CrossRef](#)] [[PubMed](#)]
77. Estève, P.-O.; Chin, H.G.; Benner, J.; Feehery, G.R.; Samaranayake, M.; Horwitz, G.A.; Jacobsen, S.E.; Pradhan, S. Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5076–5081. [[CrossRef](#)] [[PubMed](#)]

78. Xie, Q.; Bai, Y.; Wu, J.; Sun, Y.; Wang, Y.; Zhang, Y.; Mei, P.; Yuan, Z. Methylation-mediated regulation of E2F1 in DNA damage-induced cell death. *J. Recept. Signal Transduct.* **2011**, *31*, 139–146. [[CrossRef](#)]
79. Calnan, D.R.; Webb, A.E.; White, J.L.; Stowe, T.R.; Goswami, T.; Shi, X.; Espejo, A.; Bedford, M.T.; Gozani, O.; Gygi, S.P.; et al. Methylation by Set9 modulates FoxO3 stability and transcriptional activity. *Aging* **2012**, *4*, 462–479. [[CrossRef](#)]
80. Xie, Q.; Hao, Y.; Tao, L.; Peng, S.; Rao, C.; Chen, H.; You, H.; Dong, M.; Yuan, Z. Lysine methylation of FOXO3 regulates oxidative stress-induced neuronal cell death. *EMBO Rep.* **2012**, *13*, 371–377. [[CrossRef](#)] [[PubMed](#)]
81. Fu, L.; Wu, H.; Cheng, S.Y.; Gao, D.; Zhang, L.; Zhao, Y. Set7 mediated Gli3 methylation plays a positive role in the activation of Sonic Hedgehog pathway in mammals. *eLife* **2016**, *5*, e15690. [[CrossRef](#)]
82. Kim, Y.; Nam, H.J.; Lee, J.; Park, D.Y.; Kim, C.; Yu, Y.S.; Kim, D.; Park, S.W.; Bhin, J.; Hwang, D.; et al. Methylation-dependent regulation of HIF-1 $\alpha$  stability restricts retinal and tumour angiogenesis. *Nat. Commun.* **2016**, *7*, 10347. [[CrossRef](#)]
83. Pagans, S.; Kauder, S.E.; Kaehlcke, K.; Sakane, N.; Schroeder, S.; Dormeyer, W.; Trievel, R.C.; Verdin, E.; Schnolzer, M.; Ott, M. The Cellular Lysine Methyltransferase Set7/9-KMT7 Binds HIV-1 TAR RNA, Monomethylates the Viral Transactivator Tat, and Enhances HIV Transcription. *Cell Host Microbe* **2010**, *7*, 234–244. [[CrossRef](#)] [[PubMed](#)]
84. Ali, I.; Ramage, H.; Boehm, D.; Dirk, L.M.; Sakane, N.; Hanada, K.; Pagans, S.; Kaehlcke, K.; Aull, K.; Weinberger, L.; et al. The HIV-1 Tat Protein Is Monomethylated at Lysine 71 by the Lysine Methyltransferase KMT7. *J. Biol. Chem.* **2016**, *291*, 16240–16248. [[CrossRef](#)] [[PubMed](#)]
85. Masatsugu, T.; Yamamoto, K. Multiple lysine methylation of PCAF by Set9 methyltransferase. *Biochem. Biophys. Res. Commun.* **2009**, *381*, 22–26. [[CrossRef](#)] [[PubMed](#)]
86. Vasileva, E.; Shuvalov, O.; Petukhov, A.; Fedorova, O.; Daks, A.; Nader, R.; Barlev, N. KMT Set7/9 is a new regulator of Sam68 STAR-protein. *Biochem. Biophys. Res. Commun.* **2020**, *525*, 1018–1024. [[CrossRef](#)] [[PubMed](#)]
87. Kim, S.-K.; Lee, H.; Han, K.; Kim, S.C.; Choi, Y.; Park, S.-W.; Bak, G.; Lee, Y.; Choi, J.K.; Kim, T.-K.; et al. SET7/9 Methylation of the Pluripotency Factor LIN28A Is a Nucleolar Localization Mechanism that Blocks let-7 Biogenesis in Human ESCs. *Cell Stem Cell* **2014**, *15*, 735–749. [[CrossRef](#)] [[PubMed](#)]
88. Balasubramanian, N.; Ananthanarayanan, M.; Suchy, F.J. Direct methylation of FXR by Set7/9, a lysine methyltransferase, regulates the expression of FXR target genes. *Am. J. Physiol. Liver Physiol.* **2012**, *302*, G937–G947. [[CrossRef](#)] [[PubMed](#)]
89. Kassner, I.; Andersson, A.; Fey, M.; Tomas, M.; Ferrando-May, E.; Hottiger, M.O. SET7/9-dependent methylation of ARTD1 at K508 stimulates poly-ADP-ribose formation after oxidative stress. *Open Biol.* **2013**, *3*, 120173. [[CrossRef](#)]
90. Maganti, A.V.; Maier, B.; Tersey, S.A.; Sampley, M.L.; Mosley, A.L.; Özcan, S.; Pachaiyappan, B.; Woster, P.M.; Hunter, C.S.; Stein, R.; et al. Transcriptional Activity of the Islet  $\beta$  Cell Factor Pdx1 Is Augmented by Lysine Methylation Catalyzed by the Methyltransferase Set7/9. *J. Biol. Chem.* **2015**, *290*, 9812–9822. [[CrossRef](#)]
91. Aguilo, F.; Li, S.; Balasubramanian, N.; Sancho, A.; Benko, S.; Zhang, F.; Vashisht, A.; Rengasamy, M.; Andino, B.; Chen, C.-H.; et al. Deposition of 5-Methylcytosine on Enhancer RNAs Enables the Coactivator Function of PGC-1 $\alpha$ . *Cell Rep.* **2016**, *14*, 479–492. [[CrossRef](#)] [[PubMed](#)]
92. Cho, H.-S.; Suzuki, T.; Dohmae, N.; Hayami, S.; Unoki, M.; Yoshimatsu, M.; Toyokawa, G.; Takawa, M.; Chen, T.; Kurash, J.K.; et al. Demethylation of RB Regulator MYPT1 by Histone Demethylase LSD1 Promotes Cell Cycle Progression in Cancer Cells. *Cancer Res* **2011**, *71*, 655–660. [[CrossRef](#)]
93. Carr, S.M.; Munro, S.; Kessler, B.; Oppermann, U.; La Thangue, N.B. Interplay between lysine methylation and Cdk phosphorylation in growth control by the retinoblastoma protein. *EMBO J.* **2010**, *30*, 317–327. [[CrossRef](#)] [[PubMed](#)]
94. Munro, S.; Khaire, N.; Inche, A.; Carr, S.; La Thangue, N.B. Lysine methylation regulates the pRb tumour suppressor protein. *Oncogene* **2010**, *29*, 2357–2367. [[CrossRef](#)] [[PubMed](#)]
95. Ea, C.-K.; Baltimore, D. Regulation of NF- $\kappa$ B activity through lysine monomethylation of p65. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18972–18977. [[CrossRef](#)] [[PubMed](#)]
96. Yang, X.-D.; Huang, B.; Li, M.; Lamb, A.; Kelleher, N.L.; Chen, L.-F. Negative regulation of NF- $\kappa$ B action by Set9-mediated lysine methylation of the RelA subunit. *EMBO J.* **2009**, *28*, 1055–1066. [[CrossRef](#)] [[PubMed](#)]
97. Hong, X.; Huang, H.; Qiu, X.; Ding, Z.; Feng, X.; Zhu, Y.; Zhuo, H.; Hou, J.; Zhao, J.; Cai, W.; et al. Targeting posttranslational modifications of RIOK1 inhibits the progression of colorectal and gastric cancers. *eLife* **2018**, *7*, e29511. [[CrossRef](#)]
98. Song, H.; Chu, J.W.; Park, S.C.; Im, H.; Park, I.-G.; Kim, H.; Lee, J.M. Isoform-Specific Lysine Methylation of ROR $\alpha$ 2 by SETD7 Is Required for Association of the TIP60 Coactivator Complex in Prostate Cancer Progression. *Int. J. Mol. Sci.* **2020**, *21*, 1622. [[CrossRef](#)] [[PubMed](#)]
99. Hamidi, T.; Singh, A.K.; Veland, N.; Vemulapalli, V.; Chen, J.; Hardikar, S.; Bao, J.; Fry, C.J.; Yang, V.; Lee, K.A.; et al. Identification of Rpl29 as a major substrate of the lysine methyltransferase Set7/9. *J. Biol. Chem.* **2018**, *293*, 12770–12780. [[CrossRef](#)]
100. Liu, X.; Wang, D.; Zhao, Y.; Tu, B.; Zheng, Z.; Wang, L.; Wang, H.; Gu, W.; Roeder, R.G.; Zhu, W.-G. Methyltransferase Set7/9 regulates p53 activity by interacting with Sirtuin 1 (SIRT1). *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1925–1930. [[CrossRef](#)] [[PubMed](#)]
101. Elkouris, M.; Kontaki, H.; Stavropoulos, A.; Antonoglou, A.; Nikolaou, K.C.; Samiotaki, M.; Szantai, E.; Saviolaki, D.; Brown, P.J.; Sideras, P.; et al. SET9-Mediated Regulation of TGF- $\beta$  Signaling Links Protein Methylation to Pulmonary Fibrosis. *Cell Rep.* **2016**, *15*, 2733–2744. [[CrossRef](#)]
102. Fang, L.; Zhang, L.; Wei, W.; Jin, X.; Wang, P.; Tong, Y.; Li, J.; Du, J.X.; Wong, J. A Methylation-Phosphorylation Switch Determines Sox2 Stability and Function in ESC Maintenance or Differentiation. *Mol. Cell* **2014**, *55*, 537–551. [[CrossRef](#)]

103. Stark, G.R.; Wang, Y.; Lu, T. Lysine methylation of promoter-bound transcription factors and relevance to cancer. *Cell Res.* **2010**, *21*, 375–380. [[CrossRef](#)] [[PubMed](#)]
104. Wang, D.; Zhou, J.; Liu, X.; Lu, D.; Shen, C.; Du, Y.; Wei, F.-Z.; Song, B.; Lu, X.; Yu, Y.; et al. Methylation of SUV39H1 by SET7/9 results in heterochromatin relaxation and genome instability. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5516–5521. [[CrossRef](#)] [[PubMed](#)]
105. Couture, J.-F.; Collazo, E.; Hauk, G.; Trievel, R.C. Structural basis for the methylation site specificity of SET7/9. *Nat. Struct. Mol. Biol.* **2006**, *13*, 140–146. [[CrossRef](#)]
106. Kouskouti, A.; Scheer, E.; Staub, A.; Tora, L.; Talianidis, I. Gene-Specific Modulation of TAF10 Function by SET9-Mediated Methylation. *Mol. Cell* **2004**, *14*, 175–182. [[CrossRef](#)] [[PubMed](#)]
107. Ivanov, G.S.; Ivanova, T.; Kurash, J.; Ivanov, A.; Chuikov, S.; Gizatullin, F.; Herrera-Medina, E.M.; Rauscher, F.; Reinberg, D.; Barlev, N.A. Methylation-Acetylation Interplay Activates p53 in Response to DNA Damage. *Mol. Cell. Biol.* **2007**, *27*, 6756–6769. [[CrossRef](#)] [[PubMed](#)]
108. Oudhoff, M.; Freeman, S.A.; Couzens, A.L.; Antignano, F.; Kuznetsova, E.; Min, P.H.; Northrop, J.P.; Lehnertz, B.; Barsyte-Lovejoy, D.; Vedadi, M.; et al. Control of the Hippo Pathway by Set7-Dependent Methylation of Yap. *Dev. Cell* **2013**, *26*, 188–194. [[CrossRef](#)]
109. Zhang, W.J.; Wu, X.N.; Shi, T.T.; Xu, H.T.; Yi, J.; Shen, H.F.; Huang, M.F.; Shu, X.Y.; Wang, F.F.; Peng, B.L.; et al. Regulation of Transcription Factor Yin Yang 1 by SET7/9-Mediated Lysine Methylation. *Sci. Rep.* **2016**, *6*, 21718. [[CrossRef](#)]
110. Wu, X.-N.; Shi, T.-T.; He, Y.; Wang, F.-F.; Sang, R.; Ding, J.-C.; Zhang, W.; Shu, X.-Y.; Shen, H.-F.; Yi, J.; et al. Methylation of transcription factor YY2 regulates its transcriptional activity and cell proliferation. *Cell Discov.* **2017**, *3*, 17035. [[CrossRef](#)] [[PubMed](#)]