

## Article

# Profiling of the Prognostic Role of Extracellular Matrix-Related Genes in Neuroblastoma Using Databases and Integrated Bioinformatics

Leila Jahangiri <sup>1,2,3</sup> 

- <sup>1</sup> Department of Pathology, Division of Cellular and Molecular Pathology, University of Cambridge, Addenbrookes Hospital, Cambridge CB2 0QQ, UK; leila.jahangiri@ntu.ac.uk; Tel.: +44-1153-0235-8447
- <sup>2</sup> Department of Life Sciences, Birmingham City University, Birmingham B15 3TN, UK
- <sup>3</sup> School of Science & Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK

**Simple Summary:** The microenvironment in which the tumours reside highly influences them and in this study, a list of 964 genes linked to the tumour microenvironment was studied from the viewpoint of their contribution to favourable or unfavourable patient disease course prediction (prognosis), in neuroblastoma, a cancer of children, using various databases including cBioPortal and PCAT. Of this list, 12 genes, *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* showed links to patient prognosis and these genes correlated with important neuroblastoma patient clinical attributes. Further, the association between these genes with other genes and factors including long non-coding RNAs (lncRNAs) and miRNAs involved in cancer-related processes was established using tools including Cytoscape, STRING, MSigDB/BioGRID, GeneMANIA and Omicsnet. This study revealed the importance of these 12 genes as potential patient prognosis predictors in neuroblastoma and other cancers.



**Citation:** Jahangiri, L. Profiling of the Prognostic Role of Extracellular Matrix-Related Genes in Neuroblastoma Using Databases and Integrated Bioinformatics. *Onco* **2022**, *2*, 85–112. <https://doi.org/10.3390/onco2020007>

Academic Editors: Youngchul Kim and Fred Saad

Received: 14 April 2022

Accepted: 18 May 2022

Published: 19 May 2022

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

**Abstract:** A complex interaction occurs between cancer cells and the extracellular matrix (ECM) in the tumour microenvironment (TME). In this study, the expressions and mutational profiles of 964 ECM-related genes and their correlations with patient overall survival (OS) in neuroblastoma, an aggressive paediatric malignancy, were investigated using cBioPortal and PCAT databases. Furthermore, extended networks comprising protein-protein, protein-long non-coding RNA (lncRNA), and protein-miRNA of 12 selected ECM-related genes were established. The higher expressions of 12 ECM-related genes, *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* in neuroblastoma patients displayed a significant correlation with patient OS, while similar associations with neuroblastoma patient risk groups, histology and MYCN amplification were obtained. Furthermore, extended gene networks formed by these 12 ECM-related genes were established using Cytoscape, STRING, MSigDB/BioGRID, GeneMANIA and Omicsnet. Finally, the implications of the 12 ECM-related genes in other cancers were revealed using GEPIA2 and the Human Pathology Atlas databases. This meta-analysis showed the significance of these 12 ECM-related genes as putative prognostic predictors in neuroblastoma and other cancers.

**Keywords:** neuroblastoma; solid tumours; tumour microenvironment; gene networks; extracellular matrix; lncRNAs and miRNAs



**Copyright:** © 2022 by the author. 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

Neuroblastoma (NB), an aggressive paediatric malignancy of the peripheral sympathetic nervous system, accounts for 7.5% of all cancer diagnoses in children, with 1200 new cases per annum in the United States and Europe [1,2]. Of these cases, approximately half can be categorised as international neuroblastoma risk group (INRG) high-risk [3]. At the time of diagnosis, risk groups correlate with patient prognoses and reflect various clinical attributes including MYCN amplification status (increase in copy numbers of the

MYCN gene in relation to chromosome 2), stage, grade, age, histological characteristics and the presence of recurrent segmental chromosomal alterations such as 11q loss [3–5]. For instance, NB cases featuring MYCN amplification represent 50% of high-risk cases and have a 5-year survival rate of 40–50%. NB is also categorised into five stages, including 1–4 and 4S, with stages 3–4 demonstrating metastatic characteristics [2–5]. NB tumours display significant heterogeneity, and this property is widely viewed as a hallmark of this cancer that may reflect morphological, genetic and clinical attributes [6,7]. Accordingly, some NB tumours undergo spontaneous regression, while others are extremely aggressive and demonstrate relentless progression [6]. Heterogeneity, hence, can influence the clinical course of the disease and limit the efficacy of current treatment modalities in NB patients [6–8]. For instance, the detection of heterogeneous MYCN amplification reflecting subclones with differential MYCN amplification status is a diagnostic and therapeutic dilemma in the field of NB patient treatment. Accordingly, patients presenting with heterogeneous MYCN subclones may receive the same stratification and treatment as their counterparts displaying homogenous MYCN amplification. As a result, the patient may experience overtreatment and undergo unnecessary exposure to high-risk treatments, and hence suffer long-term adverse effects [8].

Multiple processes and mechanisms may contribute to tumour heterogeneity; for instance, the role of the extracellular matrix (ECM) in instigating and propagating signals that may influence tumour cells and therapy response has been reported, while NB tumours were heterogeneous with respect to the repertoire of ECM proteins they produced [9–11]. Accordingly, the ECM is one of the components of the tumour microenvironment (TME), along with immune and stromal cells, cytokines and various signalling molecules [12]. The ECM constantly undergoes deposition and degradation, which influences all cancer hallmarks including cellular bioenergetic alterations and invasion [13]. Recent studies have shown that the interactions between NB tumour cells and the ECM were based on both biochemical and biophysical interactions. For instance, ligands in the ECM may be recognised by tumour cells in the TME, while the ECM stiffness, a biophysical property, can reduce NB cell proliferation and enhance neuritogenesis [10]. Furthermore, in high-risk NB, vitronectin, an ECM anchorage glycoprotein related to rigidity and stiffness, was highly increased [9]. Consistent with the role of the physicochemical properties of the ECM, tumour cells may be impacted by gene regulatory networks, and cellular and molecular signalling cascades. Notably, a recent study using *in vitro* and *in vivo* models showed that when *MYCN* and *LMO1* were highly expressed in NB, this in turn led to increased levels in ECM-related genes such as *ITGA3*, *LOXL3* and *ITGA5* [14]. This upregulation enhanced invasion and metastasis by remodelling the ECM, restructuring the cytoskeleton and the assembly of adhesion proteins [14].

The presence of large-scale genomic and transcriptomic data deposited to publicly available databases has provided opportunities for the cancer scientific community to analyse these data using integrated statistical methods, in order to predict the prognostic roles of various ECM-related genes in cancers, paving the way for future studies aimed at the biochemical validation of *in silico* predictions [15]. Given this background, this study specifically aimed to use genomic and transcriptomic data deposited in databases to investigate ECM-related genes in NB in order to identify genes that may predict altered patient prognoses. In addition, despite the paucity of somatic genetic alterations in NB [16], the presence of genetic alterations with potential oncogenic or unknown roles may also be significant; hence, the mutational profiles of ECM-related genes and their putative links with patient survival were investigated. Additionally, the wider gene networks formed by the ECM-related genes with other proteins, long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) may be significant. Accordingly, lncRNAs were often upregulated in various cancers, and their potential as therapeutic targets and biomarkers has been established [17]. Further, accumulating evidence implicates numerous lncRNAs in tumour progression and various cancer hallmarks [18], while miRNAs have long been implicated in NB through the gene-lncRNA-miRNA regulatory networks they form [19]. Therefore,

the extended networks formed between ECM-related genes, lncRNAs and miRNAs in NB may facilitate an understanding of their influence on tumour behaviour, and were thus investigated in this study.

This meta-analysis found that 12 ECM-related genes correlated with NB patient prognosis based on TARGET expression data in cBioPortal, and patient-derived xenografts (PDX) for childhood therapeutics (PCAT) databases. These 12 ECM-related genes displayed a correlation with MYCN amplification, histology status and risk groups, while they were co-expressed with multiple potentially oncogenic lncRNAs. Using Cytoscape, STRING, MsigDB/BioGRID, GeneMANIA and Omicsnet, the regulatory networks formed by these ECM proteins with other proteins and miRNAs were also revealed. This meta-analysis displayed the significance of these ECM-related genes as indicators of patient prognoses in NB and other cancers, which warrants further preclinical investigation.

## 2. Materials and Methods Used to Address the Significance of ECM Genes in NB

### 2.1. Gene Ontology, Expression and Mutational Profile of ECM-Related Genes, Their Correlation with Overall Survival, Risk Groups and Histology Status

Using the keyword “extracellular matrix”, relevant genes were obtained from the Gene Ontology database (<http://geneontology.org>) (accessed on 15 January 2022) [20] and were compared to lists from previously published studies [15], resulting in a list of 964 ECM-related genes. The Gene Ontology database aimed to develop computational models of biological systems by curating biological functions and molecular processes. In addition, this consortium endeavoured to cluster genes based on their functional similarities. The entries in this database varied from results obtained from experimental evidence, or those inferred from high-throughput genetic interaction screens [20].

Despite the rarity of somatic mutations in NB [16], this list was subjected to the cBioPortal genetic alteration discovery tool [21] using 1472 NB patient data, including AMC Amsterdam, 2012 (87 samples); Broad 2013 (240 samples); Broad 2015 (56 samples), and TARGET 2018 (1089 samples) (Supplementary File S1). The rationale was to conduct a thorough characterisation of all relevant genetic alterations in the ECM-related genes that may have oncogenic roles, and are hence relevant to patient prognoses. Briefly, all 4 data sets were selected in cBioPortal (1472 NB patient data in total), and the molecular profiles including mutations, structural variants and copy number alterations were selected.

Furthermore, the TARGET group comprising 1089 samples, including 143 samples with RNA-sequencing expression data under the  $Z = 2$  setting, was used to establish the percentage of NB patients that displayed high levels of expression of these genes (Supplementary File S1). Briefly, of the TARGET data set (TARGET 2018, 1089 samples), the 143 NB RNA-seq sample subsection was selected for both genomic profiles and patient sets in order to query the genes of interest.

Notably, the definition of “higher expression” must be understood in the context of the  $Z = 2$  setting used to query the indicated genes in cBioPortal. The Z score stands for distance from the average reported in multiples of standard deviation (SD), which signifies that the cohort of NB cases expressing higher levels of a particular gene was located 2 SD intervals away from the mean of the population of patients. This patient-patient sample comparison may indeed be viewed as a limitation of the cBioPortal database syntax. Further, the expression-based heatmap generation option in PCAT was utilised to test the hierarchical clustering of most of the 964 ECM-related genes and the selected 12 ECM-related genes based on NB patient disease stages. Briefly, under the heatmap visualisation module, TARGET, TARGET-NBL and gene clustering were selected, and the genes of interest were queried.

The list of 964 ECM-related genes was also subjected to the “KEGG\_2019\_Human” gene ontology module in PCAT [http://pedtranscriptome.org/?multiple\\_genes\\_enrich\\_r](http://pedtranscriptome.org/?multiple_genes_enrich_r) (accessed on 10 February 2022) [22] (Supplementary File S2). Briefly, the functional enrichment tool using the “KEGG\_2019\_Human” option was utilised to query the gene list.

The correlation of ECM-related genes with the overall survival (OS) of NB patients was established using a 2- or 3-method selection process. In the former, OS was determined using 143 NB patient samples with RNA-sequencing expression data in cBioPortal (TARGET) by firstly selecting the  $Z = 2$  setting, followed by generating and inspecting the comparison/survival tab on the query result page. This tab also offered subsections, including survival, that could produce survival plots for individual queried genes. The second part of the 2-method selection was OS estimation for NB patient samples deposited to TARGET-NBL in PCAT ([http://pedtranscriptome.org/?survival\\_analysis](http://pedtranscriptome.org/?survival_analysis), accessed on 12 February 2022) using TARGET, TARGET-NBL and OS options and auto-calculate cut-offs, as shown in the workflow depicted in Supplementary Figure S1A. The “auto-calculate” option tested all cut-offs between the top and bottom 20% of expression levels of the patient data and selected the value that best separated high and low expressing patient groups. As a result of the 2-method selection process, the higher expression of 12 ECM-related genes (i.e., *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD*) significantly correlated with altered patient OS. The 3-method selection process featured the 2-method selection steps, in addition to using PCAT OS estimation with TARGET, TARGET-NBL and OS options, and “mean” cut-offs; this revealed that the higher expressions of only 6 of the 12 ECM-related genes (i.e., *AMBN*, *HSPE1*, *MUC6*, *RAMP2*, *RUVBL2* and *SSBP1*) were significantly correlated with OS. In order to encapsulate a greater view of these genes, the 12 ECM-related genes were taken forward for further analyses (Supplementary File S3). It is noteworthy that through communication with the PCAT database group, the overlap between NB TARGET at cBioPortal and TARGET-NBL in the PCAT database was revealed to be substantial, although some differences in cohort size and the statistical methods used could be envisaged. For all OS predictions in this section, the built-in Kaplan–Meier calculation tools estimated the survival curves and log-rank tests based on comparing the two groups involved. In these analyses, greater or lower OS accounted for the period, long or short, respectively, for which the patient was alive. Additionally, 249 NB patient data sets with microarray data deposited to cBioPortal (TARGET) were also studied using  $Z = 2$  and compared to the previous OS data obtained. Briefly, after choosing the TARGET data set (TARGET, 2018 1089 samples), the last option, “mRNA expression Z scores relative to all samples (log microarray)”, was chosen for genomic profiles, and the “somatic mutations” and “putative copy-number alterations” were deselected; meanwhile, the 249 NB patient samples were selected for patient sets, allowing for the individual gene of interest to be queried.

The Molbiotools software (<https://molbiotools.com/randomgenesetgenerator.php>, accessed on 9 April 2022) was used to generate several gene lists equal in size to the original ECM-related gene list (as negative controls). In evidence, the organism was selected as “homo sapiens”, the gene list was set at “964 genes” and 3 random gene sets were generated. OS was determined using 143 NB patient samples with RNA-sequencing expression data in cBioPortal (TARGET) by first selecting the  $Z = 2$  setting, followed by generating and inspecting the comparison/survival tab on the query result page. This result was compared to the output of establishing OS using the original 964 ECM-related genes.

Further to OS estimation, the PCAT database enabled covariate analyses using Cox proportional hazard model and for instance, the HR for clinical attributes, such as histology in NB patient groups, was determined. Briefly, by using TARGET, TARGET-NBL, and OS options and auto-calculate cut-offs, and selecting the covariate as “histology”, this analysis was possible. By definition, HR compared the relative risk of death between groups. Furthermore, the correlation of the selected ECM-related genes with MYCN amplification was also established using 143 NB patient samples bearing RNA-sequencing expression information in cBioPortal (TARGET) by applying Chi-square ( $Z = 2$ ). Briefly, of the TARGET data set (TARGET 2018, 1089 samples), the 143 NB RNA-seq sample subsection was selected for both select genomic profiles and patient sets to query the genes of interest ( $Z = 2$ ). Under the comparison/survival tab of the results section, the “clinical” subsection

was selected; this option provided numerous statistical tests relating to clinical attributes. MYCN amplification, for instance, could be queried in the search bar.

For risk stratification of the TARGET data set (TARGET 2018, 1089 samples), the 143 NB RNA-seq sample subsection was selected for both genomic profiles and patient sets in order to query the genes of interest. Further, the plots tab of the cBioPortal results page was selected, and on the horizontal axis, mRNA profile was selected, while on the vertical axis clinical attributes including risk groups were chosen. The raw data for the risk groups were downloaded and analysed with GraphPad using Welch's *t*-test since the raw data did not necessarily follow a normal distribution.

Risk stratification in NB was based on a 5-year event-free survival (EFS) reported by the INRG, whereby very low, low, intermediate and high-risk NB patients displayed a 5-year EFS of >85%, 75–85%, 50–75% and <50%, respectively [3]. EFS was defined as the timeframe after obtaining the primary treatment in which the patient did not experience events associated with cancer that the treatment aimed to prevent or delay. Consistently, low-risk NB were patients displaying MYCN non-amplification, presenting localised or metastatic disease, and who were younger than the age of 18 months. Approximately 50% of high-risk cases displayed MYCN amplification, while the other 50% demonstrated other high-risk stratification criteria [3–5]. In order to gain more information about the biological roles of the 12 ECM-related genes, the compartment subcellular database was utilised. The compartment subcellular database integrated data from large screens, text mining and in silico predictions about the subcellular locations of proteins (<https://www.genecards.org/Guide/GeneCard#compartments>, accessed on 13 February 2022) (<https://compartments.jensenlab.org/Search>, accessed on 13 February 2022).

The list of 12 ECM-related genes was also investigated using the gene set enrichment analysis (GSEA) database under the molecular signatures database tab; the “investigate gene sets” and hallmarks module were selected, reporting *p*- and *q*-values assessing statistical significance (<0.05) and false discovery rate (FDR) (<0.05), respectively [23,24].

## 2.2. cBioPortal, PCAT, MSigDB/BioGRID, STRING, GeneMANIA, Omicsnet and Cytoscape for Network Analysis for the Selected ECM-Related Genes

To ascertain the network connectivity of the ECM-related genes, multiple integrated bioinformatics tools were used. Briefly, of the TARGET data set (TARGET 2018, 1089 samples), the 143 NB RNA-seq sample subsection was selected for both genomic profiles and patient sets in order to query the genes of interest. Further, the co-expression tab of the results section was selected, and the “show only positively correlated genes” option was selected, while the plot could be modified by adding log scales and regression lines. The assumption of “find genes in mRNA expression (RNA-seq-143 samples), that are correlated with the gene of interest in (RNA-seq-143 samples)” was also selected. This tool featured built-in Pearson and Spearman tests. Notably, Pearson correlations evaluated linear relationships among two variables, while Spearman's correlations evaluated monotonic relationships. Accordingly, a Spearman correlation of >0.7 was considered a strong positive correlation, while each relevant ECM-related gene correlated positively or negatively with numerous other lncRNAs; only the top positively correlated lncRNA was chosen. In general, 2745 lncRNAs have been recognised by cBioPortal [25].

Further, in the PCAT portal, the pairwise correlation module was utilised to assess the Pearson correlation between the 12 ECM-related genes ([http://pedtranscriptome.org/?multiple\\_genes\\_pairwise\\_cor](http://pedtranscriptome.org/?multiple_genes_pairwise_cor), accessed on 14 February 2022) by selecting TARGET, TARGET-NBL and Pearson. The underlying R and *p*-values produced by this tool are reported in Supplementary File S4.

Furthermore, the protein-protein interaction module of STRING software was used to predict the protein-protein interactions between the 12 ECM-related genes [26]. This database allowed for the analysis of protein-protein network biology. Briefly, the “multiple proteins” option was selected and the gene list was provided.

Furthermore, MSigDB/BioGRID database prediction of the interaction of the 12 ECM-related genes was conducted using the integrated BioGRID module in GSEA/MSigDB database [23,24,27]. This module is comprised of annotated gene sets embedded in GSEA. Accordingly, using the GSEA database under the molecular signatures database tab, the “investigate gene sets” was selected and the “NDEx Biological Network Repository” was launched. On the result page, the protein-protein interaction tab was selected, resulting in the BioGrid predicted network [27,28].

In addition, GeneMANIA was utilised to expand the external networks of the 12 ECM-related genes (<https://genemania.org>, accessed on 20 February 2022) [29]. GeneMANIA used a gene ontology weighting-based system to generate scores for all biological processes attributed to interacting genes. Briefly, the 12 ECM-related genes were entered into the search bar located in the upper right section of the main tool page.

Cytoscape version 3.9.0 and NDEx v2.4.5 were also used to investigate and visualise the network connectivity of the 12 selected ECM-related genes with other genes and miRNAs [28]. Notably, Cytoscape is an open-access tool that facilitates the analysis and visualisation of relevant networks. Briefly, in the downloaded Cytoscape tool, a new session was launched and the ECM-related genes were investigated for NDEx, with the resulting network being exported to NDEx for better visual representation. Subsequently, the 1-step neighbourhood interactions were selected to zoom into individual interactions.

Finally, Omicsnet, a tool designed to visualise biological networks, was used for the establishment of network interactions (<https://www.omicsnet.ca>, accessed on 15 February 2022). Briefly, the 12 ECM-related genes were submitted to Omicsnet under the gene module by selecting official gene symbols; this allowed for proceeding to the selection of the protein-protein/IntAc interaction options. Due to the complexity of the network, the initial search was refined to the most highly connected nodes, including *LMAN1*, *HSPE1*, *RUVBL2* and *SSBP1*, and the analysis was repeated with these genes. Further, the “function explorer” and “module explorer” were submitted, and 3D visualisation and some stylistic changes were implemented. The resulting subnetwork was then exported. Omicsnet featured experimentally validated protein-protein interaction data using the official gene symbol option [30].

### 2.3. Validation of the Significance of the 12 ECM-Related Genes in Other Cancers

As a means of validating the significance of the 12 ECM-related genes in the wider context of cancer, the Human Pathology Atlas and TIMER2 databases were utilised. Briefly, the prognostic prediction of the 12 ECM-related genes was investigated across cancer types using the Human Pathology Atlas (<https://www.proteinatlas.org/humanproteome/pathology>, accessed on 18 February 2022). Further, by choosing the pathology tab in the results section, the Kaplan–Meier survival curve in addition to other information was obtained. The tool featured built-in Kaplan–Meier survival analyses to establish favourable or unfavourable categorisations.

Using the GEPIA2 DIY module (<http://gepia2.cancer-pku.cn/#index>, accessed on 19 February 2022) and the boxplot and “match TCGA normal tissue” options, the expressions of the relevant ECM-related genes were investigated in other cancers [31]. This programme reported log<sub>2</sub> fold change (log<sub>2</sub>FC) of the tumour, and matched normal tissue by applying a cut-off of 1 and a *p*-value of less than 0.01, while the statistical significance of the tests was determined using the built-in one-way ANOVA tool. All figures were generated in Adobe Photoshop software 2021.

### 2.4. Correction for Multiple Testing

In this section, FDRs representing correction for multiple testing were undertaken for all the results obtained in this study using the <https://tools.carbocation.com/FDR> (accessed on 9 April 2022) tool, and have been reported by figure number in Supplementary File S5. In evidence, when conducting parallel hypothesis tests, FDR (*q*-values) may be used to correct for false positives, while retaining potential biologically meaningful findings.

### 3. Results Supporting the Role of the 12 ECM Genes in NB

#### 3.1. ECM-Related Genes Were Overexpressed in NB Patients, Showed Mutational Profiles, and Correlated with Risk Groups, MYCN Amplification, Histology and Overall Survival

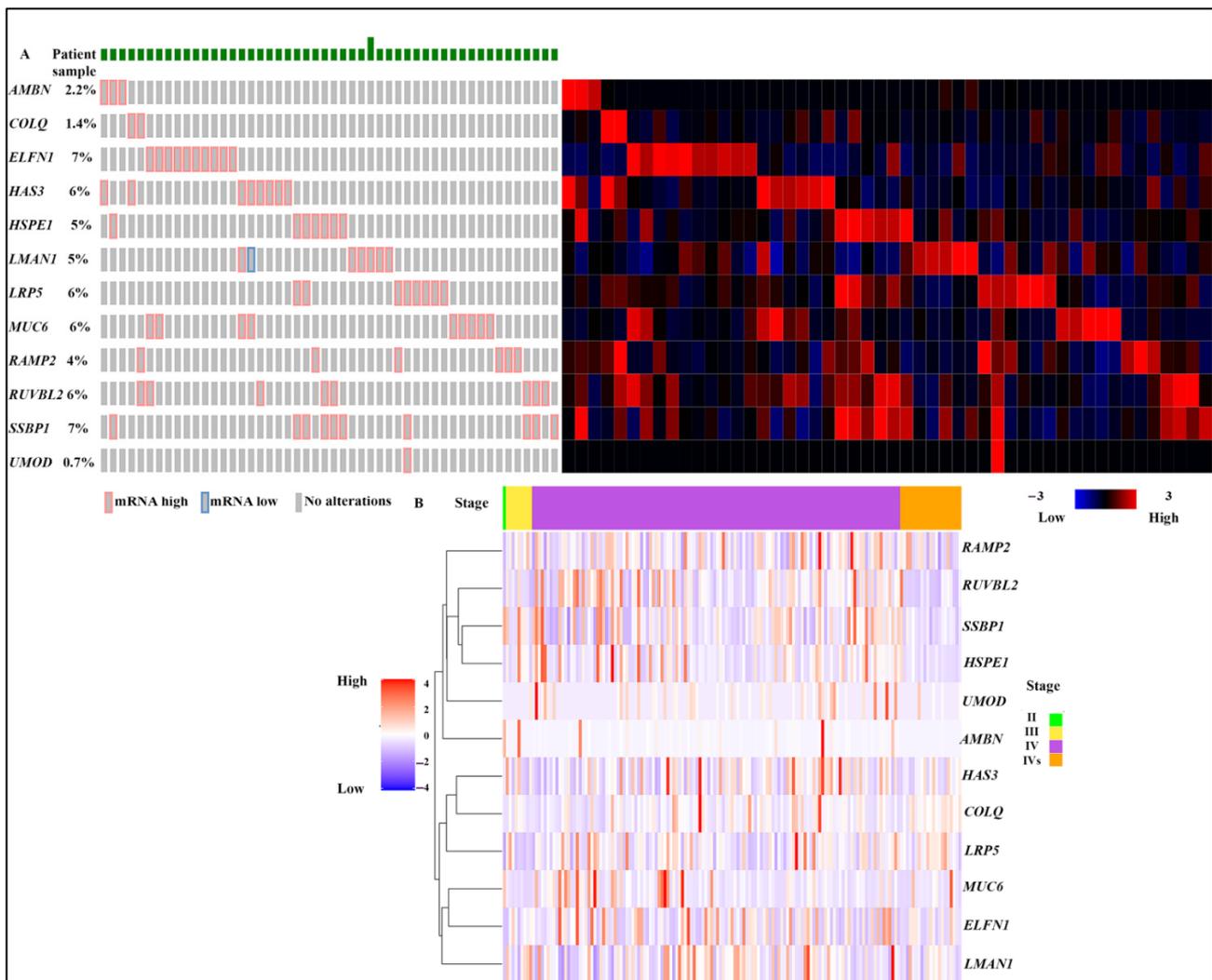
cBioPortal mutational profile analysis of 964 ECM-related genes was conducted on 1472 NB patient samples (Amsterdam, Broad 2013 and 2015, and TARGET), as mentioned in the Materials and Methods section (Supplementary File S1). Although multiple somatic genetic alterations have been identified and extensively investigated in NB, including ALK actionable mutations [32], this does not preclude characterising unappreciated alterations in other genes [33], including those linked to the ECM which provide a framework for pivotal interactions between the TME and tumour cells. These alterations included amplifications, fusions, missense and truncating mutations, deletions, splicing mutations and structural variants. Moreover, the corresponding patient sample identifier was recorded in Supplementary File S1, which facilitated obtaining clinical attributes from these anonymised samples. Accordingly, the vast majority of the genetic alterations were of unknown significance, and typically present in less than 1% of NB patient populations. For instance, *RAMP2* amplification was detected in 0.1% of NB patient samples deposited to cBioPortal including patient TARGET-30-PALJVX. A closer inspection of this sample revealed that the corresponding patient was a 5-year-old male whose OS was 95 months. Although NB molecular subtypes inclusive of mesenchymal and adrenergic have been reported [34], these subtypes have not been widely characterised across patient samples currently, nor have they been applied to patient sample databases; hence, these were not implemented in this analysis.

Notable genetic alterations obtained for the list of 964 ECM-related genes with putative driver predictions included *EPHA3* A629 (NBL44), *ITGA11* X931\_splice (TARGET-30-PASXIE), *NF1* X2397\_splice (TARGET-30-PATHYK) and *NF1* Q347\* (TARGET-30-PATVTL), among others, which may be of significance. Also reported in this datasheet was the percentage of patients that displayed higher expression levels of the ECM-related gene of interest, based on RNA-sequencing information of 143 NB patients in cBioPortal (TARGET). For instance, 6% of NB patients displayed higher expression levels of *A2M* compared to their non- or low-expressing NB patient counterparts, as defined earlier. In addition, using the TARGET-NBL expression data set in PCAT, the hierarchical clustering of most of the 964 ECM-related genes based on NB patient disease stage (II, III, IV, IVs) was determined, although not revealing an evident stage-based pattern (Supplementary File S1). Collectively, the mutational and expression profiles convey the significance of ECM-related genes in NB, and may support the establishment of a blueprint for future mutational studies aiming at linking these genetic alterations with phenotypes and patient prognosis. Furthermore, the list of 964 ECM-related genes was subjected to KEGG gene ontology analysis (Supplementary File S2). This analysis revealed the enrichment of terms including ‘extracellular matrix organisation’, ‘degradation of extracellular matrix’, ‘collagen formation’ and ‘integrin cell surface interactions’ (Supplementary File S2). This step was essential to validate the relevance of the gene list.

Furthermore, all 964 ECM-related genes were studied for patient OS based on RNA-sequencing information of 143 NB patients in cBioPortal (TARGET) under  $Z = 2$  setting and the TARGET-NBL data set in PCAT OS estimation using “auto-calculate” cut-offs (2-method selection) (Supplementary Figure S1A) [22]. The vast majority of the ECM-related genes did not correlate with patient prognoses ( $p$ -value > 0.05), hence were not taken forward (Supplementary File S3); approximately 5% were linked to altered patient prognoses.

Only 12 genes (i.e., *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD*) displayed significantly altered OS using both databases, and these were selected for further investigation (2-step selection process) (Supplementary File S3; Supplementary Figure S1A). This list of 12 ECM-related genes also displayed upregulation in 0.7–7% of NB patients compared to other NB patients who did not display elevated levels of these genes (Figure 1A). In this view, the unaltered samples have not been shown. Consistently, the hierarchical clustering of these 12 ECM-related

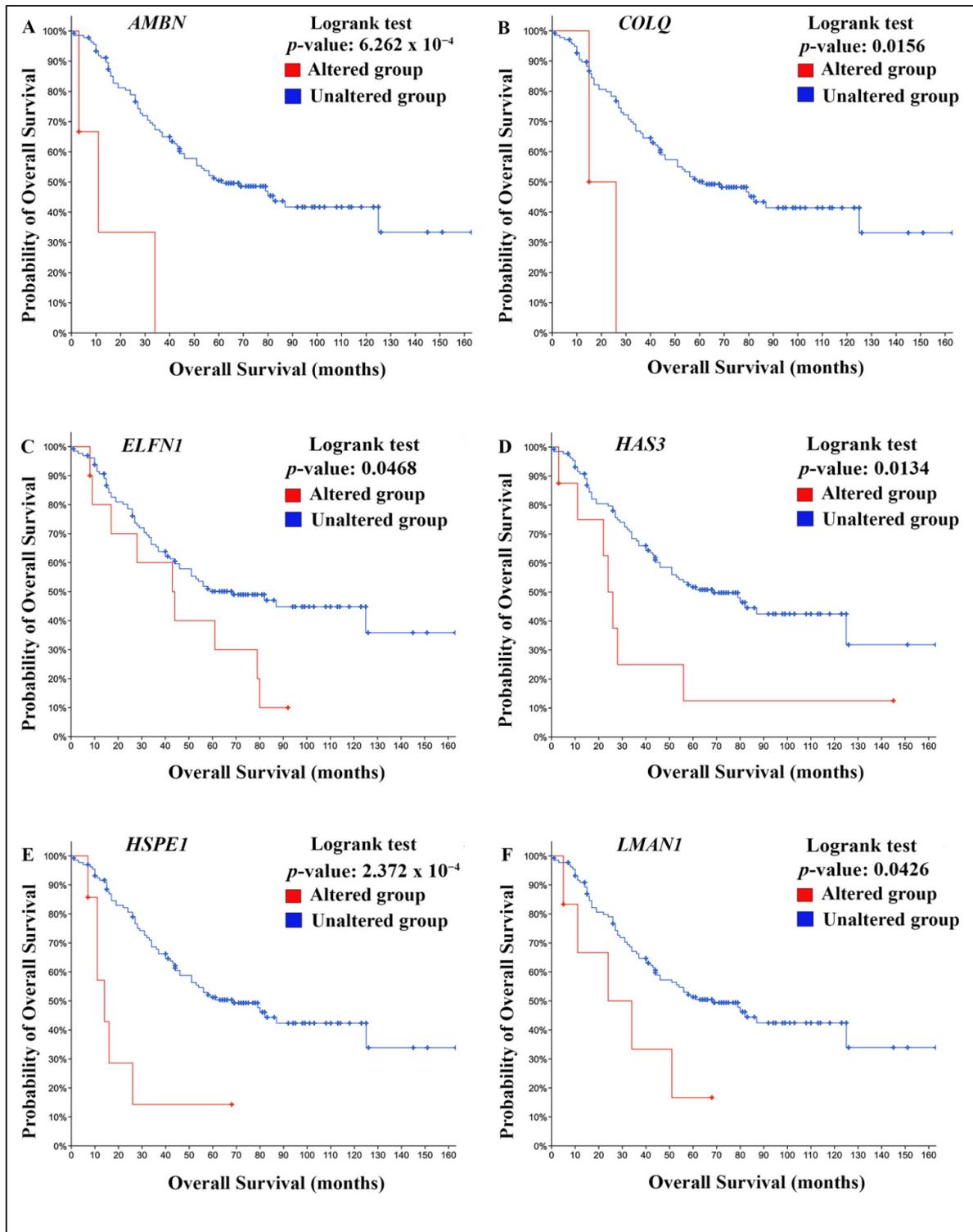
genes based on the TARGET-NBL data set in PCAT and patient disease stages, revealed the close clustering of *RAMP2*, *RUVBL2*, *SSBP1*, *HSPE1*, *UMOD* and *AMBN* (Figure 1B).



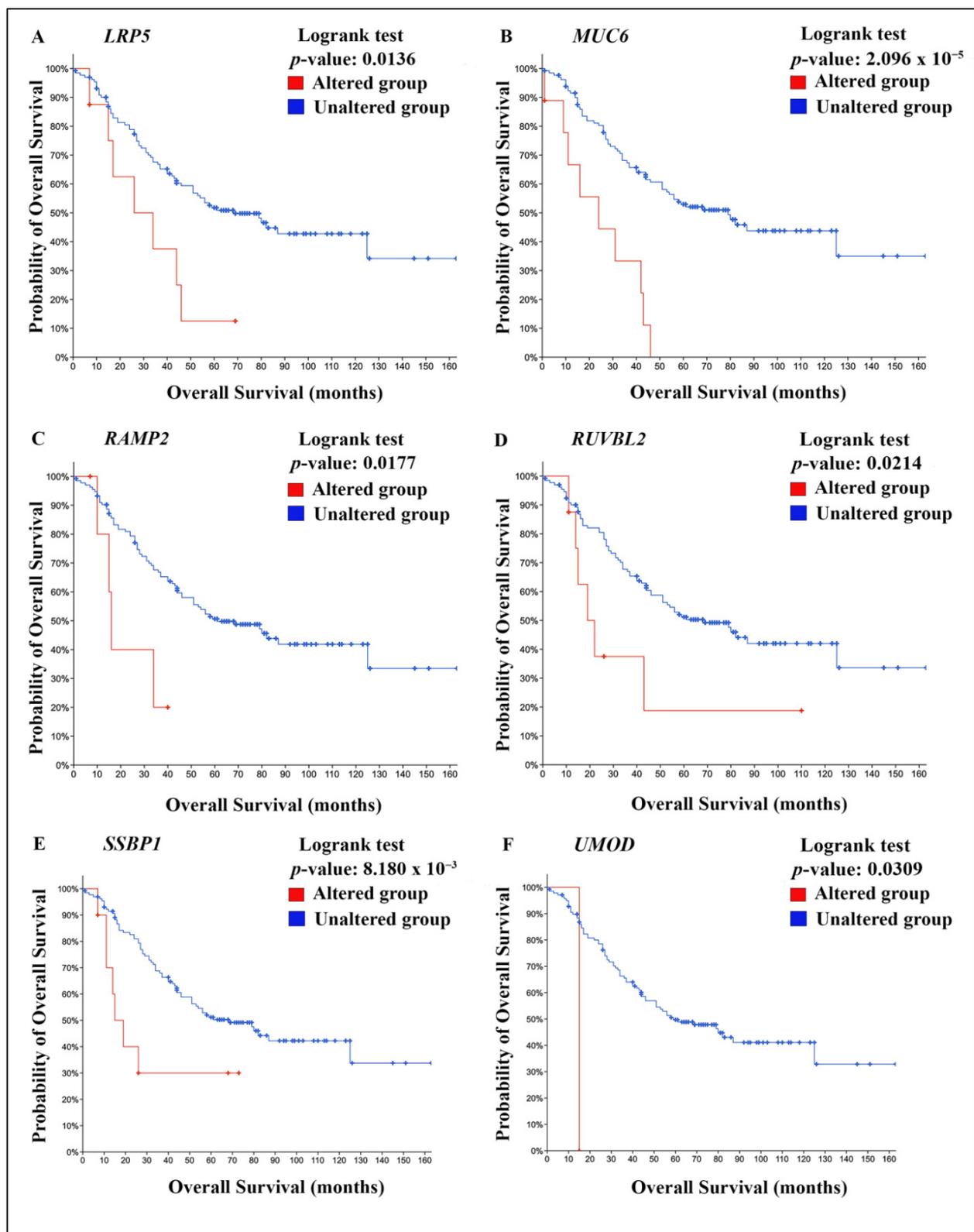
**Figure 1.** The expression profiles of selected ECM-related genes in NB patients. (A) The relevant percentages of NB patients displayed increased levels of each ECM-related gene. For instance, 2.2% of NB patients show upregulation of *AMBN* compared to their non- or low-expression NB patient counterparts. The pink and blue colour coding represents a high and low expression of the relevant genes in NB patients. The resulting heatmap is based on expression values in part A, and the red and blue colour coding of the heatmap represents higher and lower gene expressions, respectively. The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) ( $Z = 2$ ). The unaltered samples were excluded from this view in order to improve figure resolution. (B) TARGET-NBL expression data set in PCAT hierarchical clustering of the 12 ECM-related genes based on NB patient disease stage (II, III, IV and IVs), revealing the closer clustering of *RAMP2*, *RUVBL2*, *SSBP1*, *HSPE1*, *UMOD* and *AMBN*, although strict stage-based patterns were not evident.

Moreover, the expressions of 9 of 12 ECM-related genes (i.e., *AMBN*, *ELFN1*, *HSPE1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD*) correlated with reduced OS, based on RNA-sequencing information of 143 NB patients in cBioPortal (TARGET) (log-rank test  $p$ -values:  $6.262 \times 10^{-4}$ , 0.0468,  $2.372 \times 10^{-4}$ , 0.0136,  $2.096 \times 10^{-5}$ , 0.0177, 0.0214,  $8.18 \times 10^{-3}$  and 0.0309, respectively) (Supplementary File S3, Figure 2A,C,E and Figure 3A–F); this also applied to OS obtained from PCAT using auto-calculate cut-offs (log-rank  $p$ -values:

0.00874, 0.0234, 0.00005, 0.0065, 0.00001, 0.00464, <0.000001, 0.00002 and 0.00034, respectively) (Supplementary File S3). Figures 2 and 3 were not combined in order to prevent reduced resolution.



**Figure 2.** Kaplan–Meier curves for OS in NB patients. (A–F) The increased expressions of *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1* and *LMAN1* correlated with reduced OS in NB patients (log-rank test  $p$ -values:  $2.262 \times 10^{-4}$ , 0.0156, 0.0468, 0.0134,  $2.372 \times 10^{-4}$  and 0.0426, respectively). The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) ( $Z = 2$ ).



**Figure 3.** Kaplan–Meier curves for OS in NB patients. (A–F) The increased expressions of *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* correlated with reduced OS in NB patients (log-rank test  $p$ -values: 0.0136,  $2.096 \times 10^{-5}$ , 0.0177, 0.0214,  $8.180 \times 10^{-3}$  and 0.0309, respectively). The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) ( $Z = 2$ ).

Inversely, the higher expression of three genes (i.e., *COLQ*, *HAS3*, and *LMAN1*) showed conflicting correlations, based on RNA-sequencing information of 143 NB patients in cBioPortal (TARGET) and PCAT OS estimation using the auto-calculate setting. Namely, using the former method, *COLQ*, *HAS3* and *LMAN1* correlated with reduced patient OS (log-rank *p*-values: 0.0156, 0.0134 and 0.0426, respectively) (Supplementary File S3, Figure 2B,D,F); meanwhile, using the latter method, these three genes correlated with increased patient OS (log-rank *p*-values of 0.0468, 0.0049 and 0.0448, respectively) (Supplementary Figure S1A; Supplementary File S3). As mentioned in the Materials and Methods section, using a stricter 3-method selection workflow, 6 of the 12 ECM-related genes showed significantly reduced OS using all three methods (i.e., *AMBN*, *HSPE1*, *MUC6*, *RAMP2*, *RUVBL2* and *SSBP1*) (log-rank *p*-values for TARGET-NBL in the PCAT “mean” setting: 0.01113, 0.00159, 0.00051, 0.02307, 0.00031 and 0.00043, respectively) (Supplementary Figure S1A; Supplementary File S3); however, the list of 12 ECM-related genes was taken forward. Collectively, these results suggested the significant role of these 12 ECM-related genes in NB, and that their expression in NB patients may be of prognostic relevance.

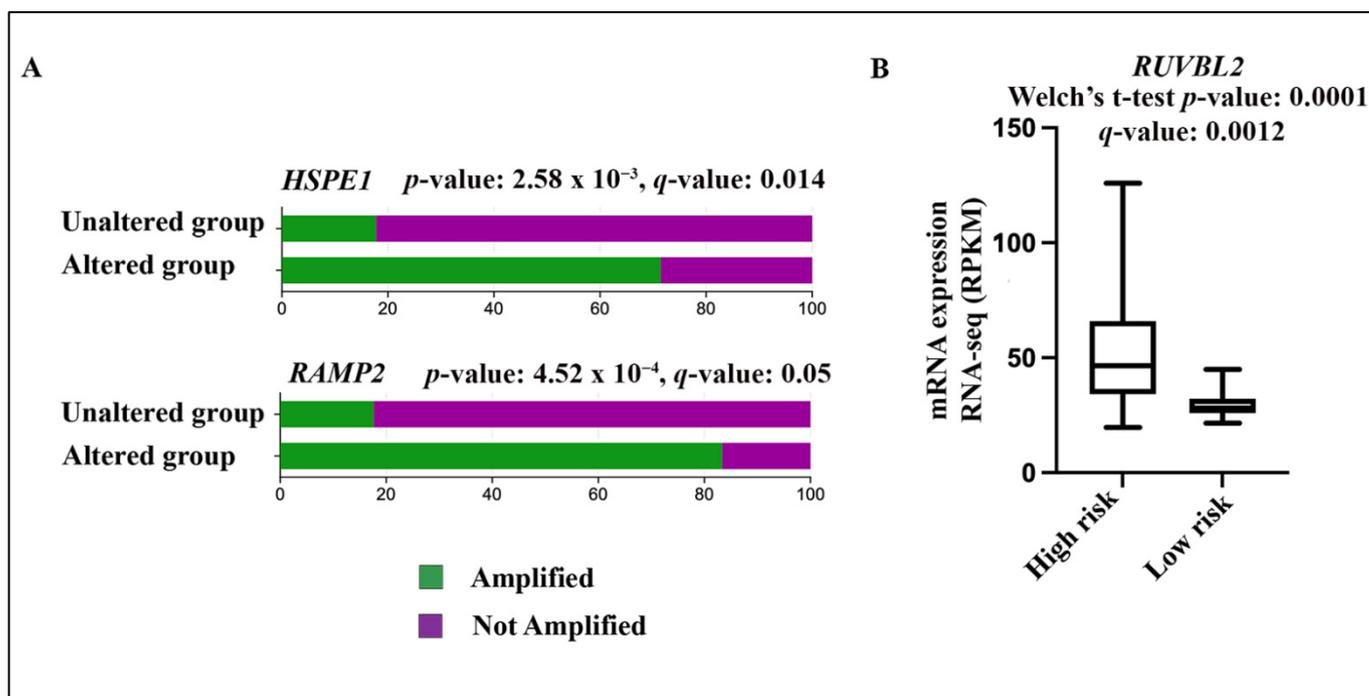
Finally, using 249 NB patient data sets with microarray expression information in cBioPortal (TARGET), the OS of the NB patients expressing any of the 12 ECM-related genes was determined. For example, overexpressing *LRP5* correlated with reduced OS (log-rank *p*-values:  $7.98 \times 10^{-6}$ ) (Supplementary Figure S1B).

The outcome of running OS estimation using 143 NB patient samples with RNA-sequencing expression data in cBioPortal (TARGET) under the  $Z = 2$  for three randomly generated gene lists of equal size to the original 964 ECM-related genes used in this study was 5–7% statistically significant hits. These hits were comparable to the percentage of statistically significant hits obtained from the original ECM-related gene list (around 5%= 52 genes). This was before implementing the 2- and 3-step selection processes described earlier that further restricted the 52 genes to 12 and 6 genes, respectively (Supplementary File S3). This result drove the premise that the current study has taken a “classical” approach to the ECM-related genes which showed statistically significant OS, meaning that each individual gene may indeed play a role in the tumourigenic processes, and are hence worth taking forward.

Of note, the OS establishment using the PCAT database (with, for instance, the auto-calculate setting) also allowed for the application of Cox proportional hazard ratio (HR) analyses for covariates, including histology status. NB patient samples upregulating *AMBN*, *HSPE1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* genes displayed statistically significant increased HR (and decreased OS) with the following HR, confidence interval (CI) and *p*-values: 0.5 (CI: 0.29–0.87; *p*-value: 0.014), 0.43 (CI: 0.26–0.71; *p*-value: 0.001), 0.48 (CI: 0.29–0.78; *p*-value: 0.003), 0.44 (CI: 0.27–0.74; *p*-value: 0.002), 0.45 (CI: 0.28–0.73; *p*-value: 0.001), 0.4 (CI: 0.24–0.66; *p*-value: <0.001), 0.46 (CI: 0.28–0.74; *p*-value: 0.002) and 0.48 (CI: 0.28–0.83; *p*-value: 0.008), respectively (Supplementary File S3). These specific patient groups also showed higher HR for unfavourable histology status (also linked to decreased OS) with the following HR, confidence interval (CI) and *p*-values: 4.3 (CI: 1.56–12.04; *p*-value: 0.005), 4.37 (CI: 1.58–12.07; *p*-value: 0.004), 5.42 (CI: 1.97–14.92; *p*-value: 0.001), 4.35 (CI: 1.57–12.04; *p*-value: 0.005), 5.3 (CI: 1.92–14.61; *p*-value: 0.001), 4.2 (CI: 1.50–11.55; *p*-value: 0.006), 4.11 (CI: 1.48–11.43; *p*-value: 0.007) and 4.61 (CI: 1.67–12.71; *p*-value: 0.003), respectively (Supplementary File S3). This result linked these patient groups with unfavourable histology profiles, suggesting that the expression of these genes and unfavourable histology status significantly reduced NB patient survival.

Given these results, the correlations between the 12 ECM-related genes with MYCN amplification status and risk groups were calculated based on RNA-sequencing information of 143 NB patients in cBioPortal (TARGET). *HSPE1* and *RAMP2* expressions showed a correlation with NB patient groups demonstrating MYCN amplification using cBioPortal built-in Chi-square tests (*p*-values:  $2.58 \times 10^{-3}$ , *q*-value: 0.014, and  $4.52 \times 10^{-4}$ , *q*-value: 0.05, respectively) (Figure 4A). *HAS3* was excluded since the *q*-value was not significant

based on the built-in cBioPortal Chi-square test, despite being significant using FDR tests (Supplementary File S5).



**Figure 4.** The correlation of selected ECM-related genes with patient clinical attributes. (A) The correlation of *HSPE1* and *RAMP2* genes with NB patient groups displaying MYCN amplification ( $p$ -values:  $2.58 \times 10^{-3}$ ,  $q$ -value: 0.014, and  $4.52 \times 10^{-4}$ ,  $q$ -value: 0.05, respectively) (Chi-square test). (B) *RUVBL2* was upregulated in high-risk NB patients, while it was downregulated in their low-risk counterparts based on Welch's  $t$ -test;  $p$ -value: 0.0001,  $q$ -value: 0.0012, showing a statistically significant difference between low and high-risk NB patient cases. The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) ( $Z = 2$ ).

Furthermore, *RUVBL2* was upregulated in high-risk NB patients compared to their low-risk counterparts using Welch's  $t$ -test for unequal variances in GraphPad (Figure 4B). Based on the resulting  $p$ -value: 0.0001,  $q$ -value: 0.0012, the null hypothesis of equal means was rejected, hence *RUVBL2* was statistically differentially expressed between high and low-risk groups. In addition, *HSPE1*, *SSBP1* and *HAS3* were also upregulated and significantly differentially expressed in high-risk NB patients compared to their low-risk counterparts ( $p$ -value: 0.001,  $q$ -value: 0.006;  $p$ -value: 0.0046,  $q$ -value: 0.0184;  $p$ -value: 0.006,  $q$ -value: 0.018, respectively) (Supplementary File S5).

The high and low-risk stratification was based on a 5-year EFS descriptor reported by the INRG. Accordingly, very low, low, intermediate and high-risk NB patients were defined as having a 5-year EFS of >85%, 75–85%, 50–75% and <50%, respectively [3]. These results collectively suggest that 12 ECM-related genes were significantly correlated with altered OS prognosis of NB patients, while a subset of three genes correlated with crucial clinical aspects of this cancer, including MYCN amplification and risk stratification; this lends support to their significance, warranting further preclinical investigation. Reviewing the NB literature using the combination of “neuroblastoma”, “extracellular matrix” and the 12 selected ECM-related genes did not yield any results, suggesting that these genes may have not been investigated within the framework of NB and the ECM. Widening the search by using “neuroblastoma” and the 12 selected ECM-related genes, however, yielded some supporting literature (Table 1).

**Table 1.** The relevance of the selected ECM-related genes to tumourigenesis in NB.

Gene Name	Cancer Type	Examples of Roles in Tumourigenesis	Reference
<i>HAS3</i>	Neuroblastoma	<i>HAS3</i> potentiated melatonin-based differentiation in neuroblastoma	[35]
<i>LRP5</i>	Neuroblastoma	<i>LRP5/6</i> protected SH-SY5Y cells against neurotoxicity induced by H <sub>2</sub> O <sub>2</sub>	[36]
<i>RAMP2</i>	Neuroblastoma	<i>RAMP2</i> expression was decreased in IMR-32 cells in hypoxic conditions	[37]
<i>RUVBL2</i>	Neuroblastoma	<i>RUVBL2</i> potentiated cell death mediated by PCI-24781 in SK-N-DZ cells	[38]
<i>SSBP1</i>	Neuroblastoma	A gene involved in mitochondrial DNA replication which may have promising therapeutic roles	[39]

In addition, the cellular and molecular functions that these genes played in cells were also investigated. Accordingly, the Genecard link to the compartment database revealed the roles of the ECM-related genes in subcellular organisation. For instance, *LMAN1*, an ER-Golgi intermediate protein, was involved in the recognition of molecules with sugar residues, and the recycling of molecules in the cell (Table 2). Thus, some of the ECM proteins on this list may indeed display functions within the cell or relate to the plasma membrane, which could then in turn impact ECM proteins, functions and processes, hence justifying the widening of the terms used to include “extracellular-matrix related”.

**Table 2.** The biological roles of the 12 ECM-related genes, based on GeneCards/compartment subcellular localisation database and GSEA/hallmark.

Gene Name	Subcellular Localisation Database Attributed Role
<i>AMBN</i>	Ameloblastin, involved in the structural organisation and mineralisation of the enamel
<i>COLQ</i>	Single-stranded homotrimer (collagen-like tail structure) of asymmetric acetylcholinesterase (AChE), attaches the catalytic subunits of asymmetric AChE to the basal lamina of the synapse
<i>ELFN1</i>	Fibronectin type III domain-containing protein 1 (extracellular leucine-rich repeat), a postsynaptic protein that regulated the circuit dynamics of the nervous system
<i>HAS3</i>	Hyaluronic acid synthase 3, essential for hyaluronan synthesis, that is a component of the ECM with structural roles
<i>HSPE1</i>	Heat shock protein family E, a co-chaperonin involved in mitochondrial protein transport and the assembly of macromolecules
<i>LMAN1</i>	Mannose-specific lectin, ER-Golgi intermediate protein, could recognise sugar residues in molecules and may be involved in recycling
<i>LRP5</i>	Low-density lipoprotein receptor-related protein 5 implicated in the Wnt-frizzled-LRP5-LRP6 complex that can instigate $\beta$ -catenin signalling
<i>MUC6</i>	Mucin-6, oligomeric mucus/gel-forming; played a role in the cryoprotection of epithelial layers and may be implicated as cancer markers
<i>RAMP2</i>	Receptor activity-modifying protein 2, involved in transport to the plasma membrane
<i>RUVBL2</i>	A member of the RuvB family, RuvB like AAA ATPase 2, and is involved in the endoplasmic reticulum (ER)-associated degradation pathways (ERAD) which impact ER stress-triggered responses
<i>SSBP1</i>	Single-stranded DNA-binding protein, this protein binds to single-stranded DNA, may be involved in mitochondrial DNA replication
<i>UMOD</i>	Involved in the synthesis of the apical membrane of epithelial cells and may be implicated in water barrier permeability, may also assist neutrophil migration

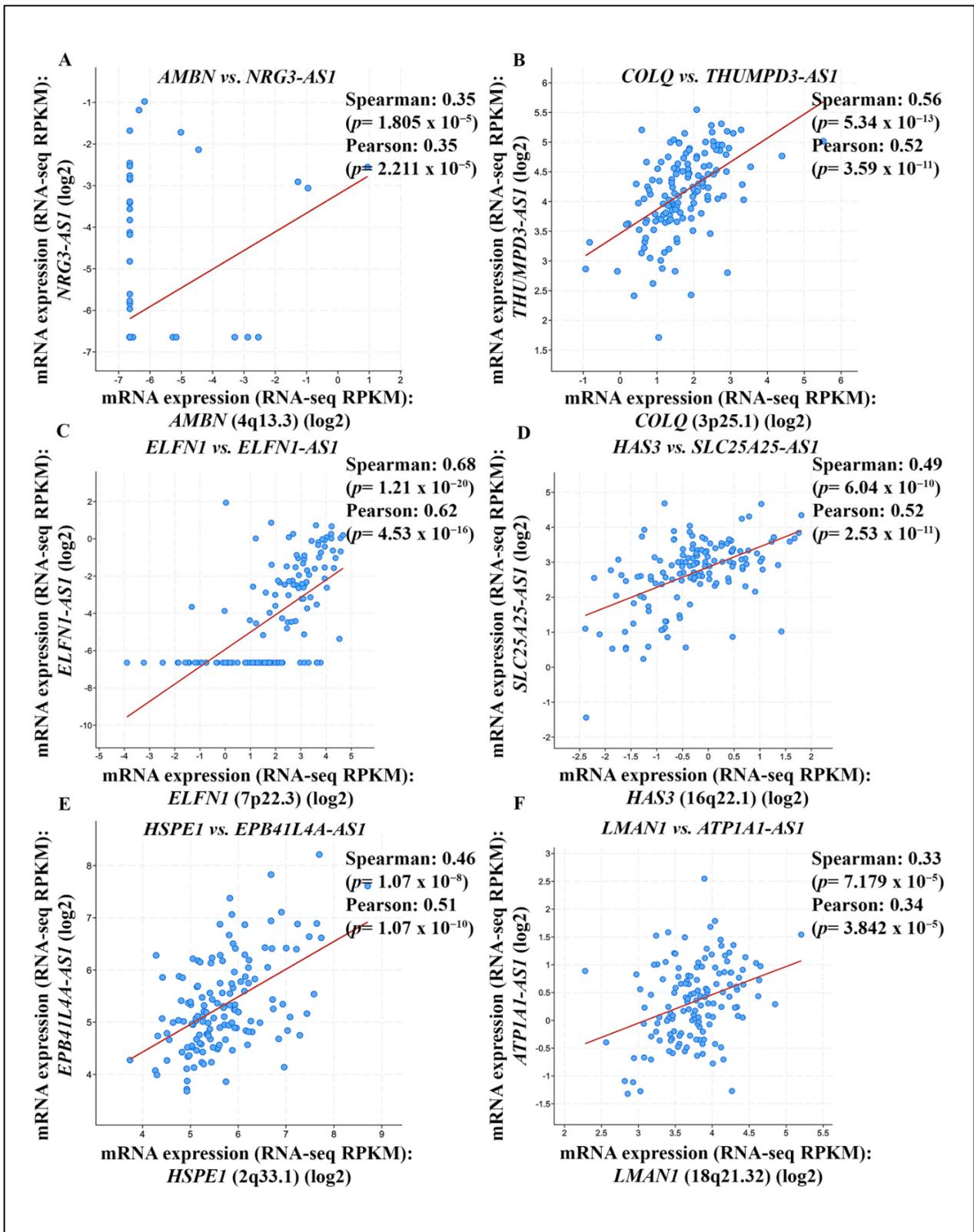
This list was also validated against the GSEA database under the gene sets/ hallmarks module, and revealed the enrichment of ‘hallmark MYC target genes’ ( $p$ -value:  $2.47 \times 10^{-5}$ ,  $q$ -value:  $1.24 \times 10^{-3}$ ) including *HSPE1*, *RUVBL2* and *SSBP1* [23,24]. This result was interesting, since links between MYC-regulated proteins, lncRNAs and miRNAs have been extensively reported in the literature [40]. Given these results, the gene networks formed by these 12 selected ECM-related genes with other genes, lncRNAs and miRNAs were investigated.

### 3.2. cBioPortal, PCAT, MSigDB/ BioGRID, STRING, GeneMANIA, Omicsnet and Cytoscape Revealed Network Connectivity of the Selected ECM-Related Genes

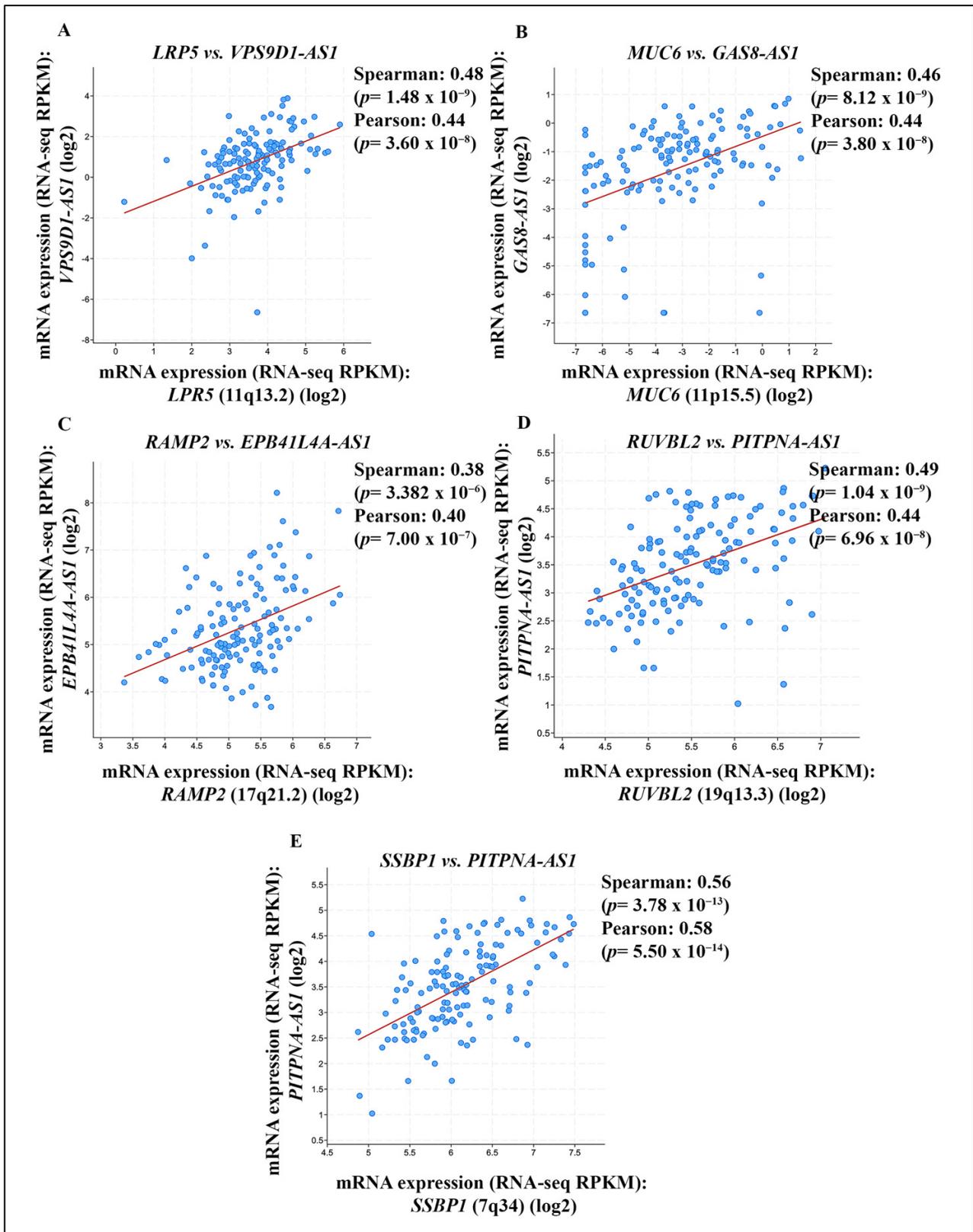
The Spearman and Pearson correlations of lncRNAs with each of the 12 ECM-related genes obtained from cBioPortal revealed evidence of positive correlation, reported in RPKM values. Accordingly, for each ECM-related gene, only the top positively correlated lncRNA has been reported, although each ECM-related gene correlated positively or negatively with numerous other lncRNAs. For instance, *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2* and *SSBP1* positively correlated with *NRG3-AS1*, *THUMPD3-AS1*, *ELFN1-AS1*, *SLC25A25-AS1*, *EPB41L4A-AS1*, *ATP1A1-AS1*, *VPS9D1-AS1*, *GAS8-AS1*, *EPB41L4A-AS1*, *PITPNA-AS1* and *PITPNA-AS1*, respectively (Spearman correlations: 0.35, 0.56, 0.68, 0.49, 0.46, 0.33, 0.48, 0.46, 0.38, 0.49 and 0.56, respectively) (Figures 5A–F and 6A–E). Figures 5 and 6 were not combined in order to prevent reduced resolution. These links were based on NB patient data in cBioPortal, hence would be directly applicable to this cancer.

No significant lncRNA correlations were obtained for *UMOD*, although none of the links presented displayed a strong Spearman correlation ( $>0.7$ ); however, these correlations were nonetheless significant, and may bear links to oncogenic pathways in NB and other cancers. In evidence, *THUMPD3-AS1* was associated with non-small cell lung cancer (NSCLC) self-renewal through the axis of *ONECUT2* and *miR543* [41], while *ELFN1-AS1* enhanced the progression of oesophageal cancer via the gene-miRNA network of *GFPT1-miR-183-3p* [42]. Collectively, these data suggest that the selected ECM-related genes in NB may form networks with lncRNAs that in turn have been implicated in oncogenic processes in other cancers [43].

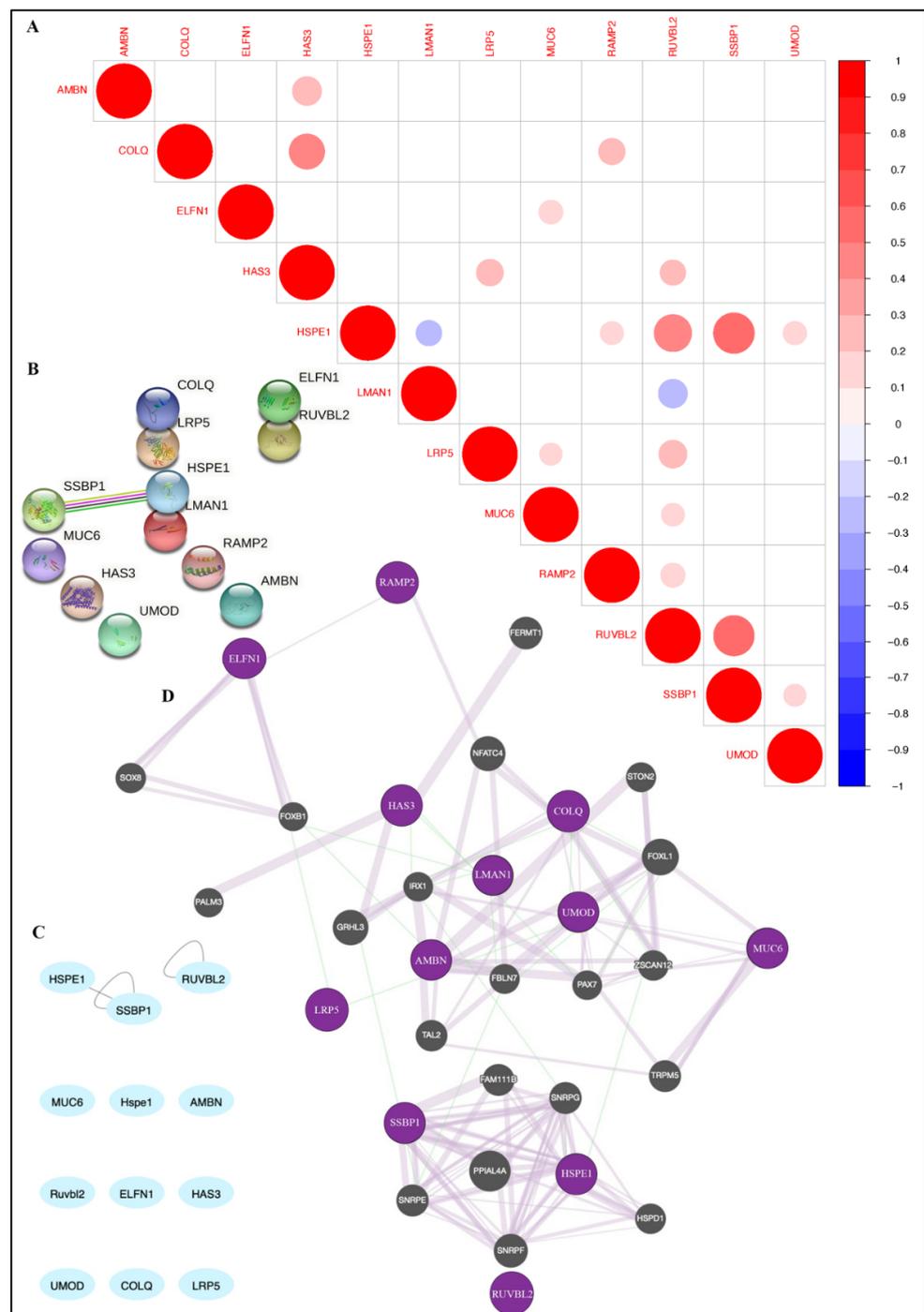
The links obtained between the 12 selected ECM-related genes and lncRNAs led to profiling the other networks formed by these genes in order to understand the extended gene networks they may form in NB and other cancers. Notably, the networks formed between the 12 selected ECM proteins were studied using the pairwise correlation module in PCAT, MSigDB/ BioGRID, and STRING [26]. The former predicted that, for instance, *HSPE1* positively correlated with *RAMP2*, *RUVBL2* and *SSBP1* (Figure 7A), with the following correlation ( $R$ ) and  $p$ -values: 0.17 (0.024), 0.46 ( $1.03 \times 10^{-9}$ ) and 0.54 ( $8.99 \times 10^{-14}$ ), respectively (Supplementary File S4). The red and blue colour coding suggest a positive and negative correlation, respectively. In agreement with this, the protein-protein interaction prediction module of STRING revealed interactions between *HSPE1* and *SSBP1* (co-expression score: 0.171) (Figure 7B). This score was estimated based on similar patterns of mRNA expression. Furthermore, MSigDB/ BioGRID database prediction of the interaction of the 12 ECM-related genes revealed interactions between *HSPE1* and *SSBP1* (score= 0.97) (Figure 7C) [27].



**Figure 5.** LncRNAs were co-expressed with ECM-related genes. (A–F) *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1* and *LMAN1* positively correlated with *NRG3-AS1*, *THUMPD3-AS1*, *ELFN1-AS1*, *SLC25A25-AS1*, *EPB41L4A-AS1* and *ATP1A1-AS1* (Spearman correlations: 0.35, 0.56, 0.68, 0.49, 0.46 and 0.33, respectively). The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) (Z = 2).



**Figure 6.** LncRNAs were co-expressed with ECM-related genes. (A–E) (*LRP5*, *MUC6*, *RAMP2*, *RUVBL2* and *SSBPI* positively correlated with *VPS9D1-AS1*, *GAS8-AS1*, *EPB41L4-AS1*, *PITPNA-AS1* and *PITPNA-AS1*, respectively) (Spearman correlations: 0.48, 0.46, 0.38, 0.49 and 0.56, respectively). The data were obtained from studying 143 NB tissue samples with RNA-sequencing information in cBioPortal (TARGET) (Z = 2).



**Figure 7.** Pairwise correlations and network connectivity of ECM-related genes. **(A)** The ECM-related genes were inter-correlated; for instance, *HSPE1* was positively correlated with *RAMP2*, *RUVBL2* and *SSBP1* with the following correlation (R) and *p*-values: 0.17 (0.024), 0.46 ( $1.03 \times 10^{-9}$ ) and 0.54 ( $8.99 \times 10^{-14}$ ), respectively. Red and blue colour coding represent positive and negative correlations, respectively, based on TARGET-NBL expression data in the PCAT database. **(B)** STRING protein-protein interaction module also predicted a correlation between *HSPE1* and *SSBP1* proteins (co-expression score: 0.171). **(C)** GSEA MSigDB/BioGRID database prediction of the interaction of the 12 ECM-related genes revealed the protein-protein interaction between *HSPE1* and *SSBP1* proteins (score = 0.97). **(D)** GeneMANIA network analysis revealed the co-expression of the 12 ECM proteins with other proteins such as *FOXL1*, *NFATC4*, *SNRPG*, *FAM111B* and *SNRPF* (co-expression 98.48%). The networks displayed functions in ribonucleoprotein assembly and spliceosome function. The 12 ECM-related genes are depicted in purple.

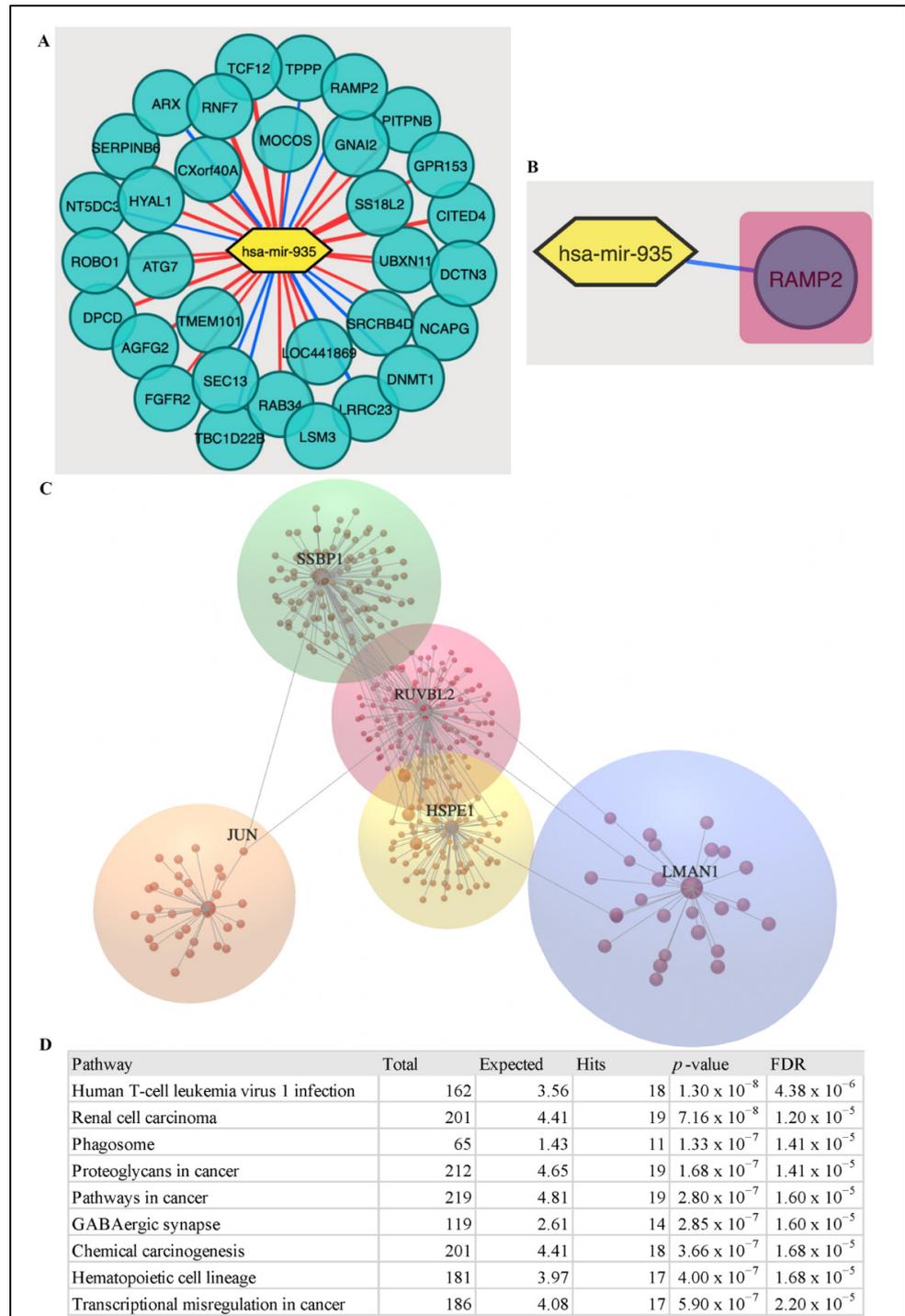
Furthermore, GeneMANIA network analysis revealed the co-expression of the 12 ECM proteins with other proteins such as FOXL1, NFATC4 and SNRPF (genetic interaction score: 1.52% and network co-expression score: 98.48%) (Figure 7D) [29]. These network proteins functioned in ribonucleoprotein assembly, transcription and spliceosome complexes, terms related to structural and molecular processes. The weights of each interaction are reported in Supplementary File S5.

In addition, Cytoscape was used to interrogate miRNA-related networks established by the ECM-related genes, which could further substantiate their extended networks. For instance, a network formed by *RAMP2* in the development of the nervous system described in TCGA-veal melanoma (UVM) was established by exporting these data from Cytoscape to NDEx and selecting 1-step neighbourhood interactions ( $q$ -value: 6.02) (Figure 8A) [28]. *RAMP2* negatively correlated with *hsa-mir-935* ( $p$ -value:  $1.087 \times 10^{-6}$ , correlation:  $-0.5139$ ) (depicted using blue edges) (Figure 8B). In evidence, the *mir-935*/HIF-1 $\alpha$  axis inhibited cell proliferation and invasive behaviour in gliomas [44], hence the roles of *RAMP2* and *mir-935* and their potential interplay in NB may be further investigated. Using the Omicsnet portal, the network connectivity of the 12 ECM-related genes was established. *RUVBL2*, *SSBP1*, *HSPE1* and *LMAN1* displayed the greatest degree of connectivity with 153, 119, 72 and 26 genes, respectively (Figure 8C) [30]. For instance, *SSBP1* and *RUVBL2* were linked to *JUN*, a proto-oncogene. This software allowed for the interrogation of gene ontology term enrichment of the networks formed. Accordingly, the described networks showed enrichment for 'cancers', 'signalling molecules and pathways' and 'transcriptional regulation', bringing into focus their links with oncogenic pathways (Figure 8D). The relevant  $p$ - and  $q$ -values (FDRs) were reported in the figure.

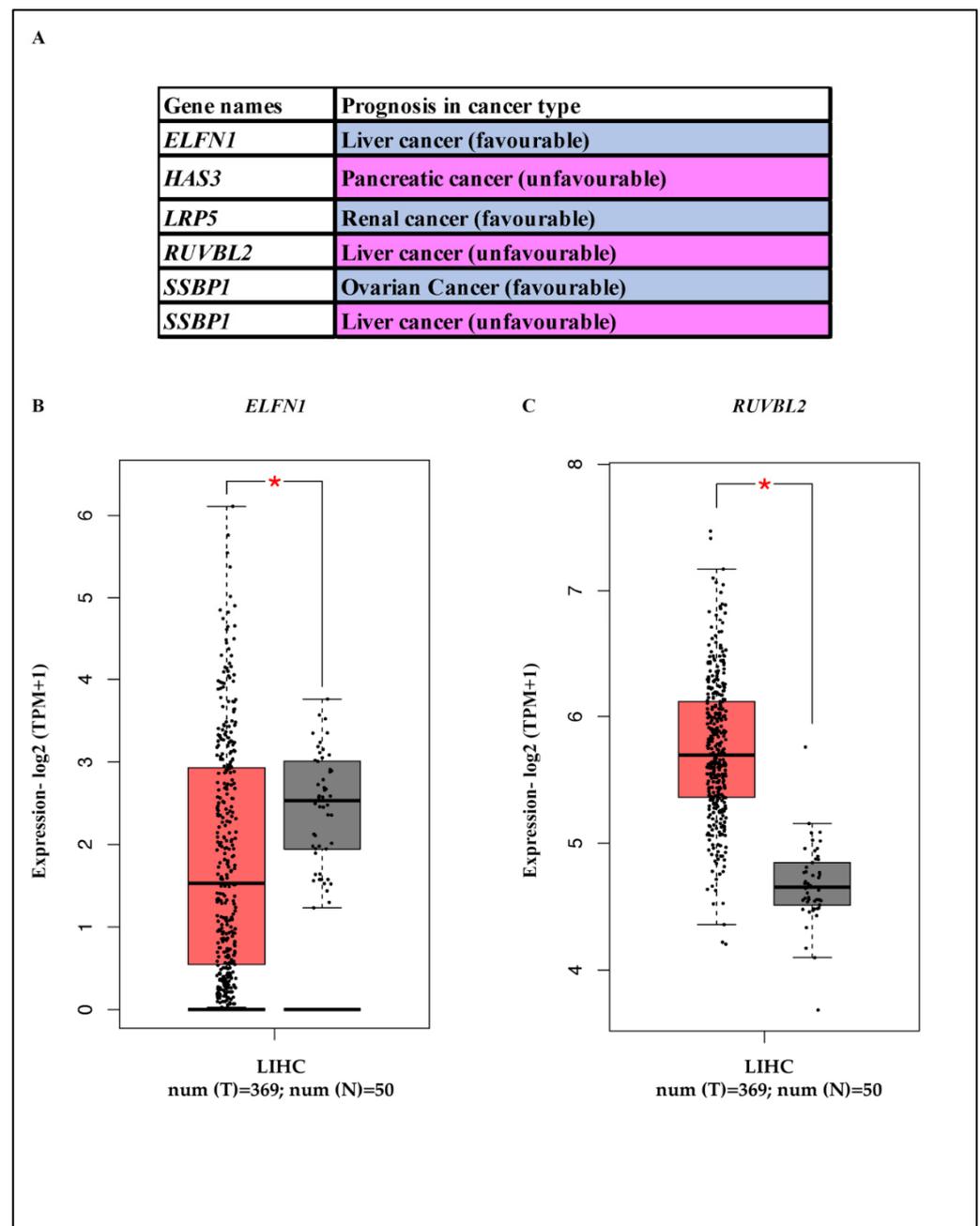
### 3.3. Selected ECM-Related Genes Were Correlated with Patient Prognoses in Other Cancers

Using the Human Pathology Atlas, the association of the 12 ECM-related genes with favourable or unfavourable patient prognoses in other cancers was reported. Namely, the expressions of *ELFN1*, *LRP5* and *SSBP1* were associated with favourable outcomes in renal, liver and ovarian cancers, respectively, while *HAS3*, *RUVBL2* and *SSBP1* were associated with unfavourable outcomes in pancreatic, liver and liver cancers, respectively (Figure 9A). Based on the GEPIA2 DIY module, for example, the expressions of *ELFN1* and *RUVBL2* were revealed to be significantly down- or upregulated in liver hepatocellular carcinoma (LIHC) compared to matched normal, which is in agreement with the favourable and unfavourable prognostic values obtained in Figure 9A–C. Notably, a log<sub>2</sub>-fold change ( $\log_2$  FC) = 1 and  $p$ -value <0.01 were applied, while a one-way ANOVA was used to test the statistical difference between the tumour and normal tissue groups. The expression levels were reported in  $\log_2$  TPM + 1 (transcript per million).

Reviewing the literature revealed that the role of these 12 ECM-related genes has been investigated in multiple cancers (Table 3). For instance, strong nuclear and cytoplasmic staining of *RUVBL2* in hepatocellular carcinoma (HCC) was correlated with reduced OS and recurrence-free survival (RFS), while higher levels of *RUVBL2* mRNA only correlated with reduced RFS in this cancer [45].



**Figure 8.** Cytoscape analysis of *RAMP2*-miRNA network and network connectivity of the ECM-related genes using Omicsnet. (A) Cytoscape analyses of *RAMP2* in TCGA-UVM (Uveal melanoma) in the development of the nervous system, established by exporting data to NDEX and selecting 1-step neighbourhood interactions of *RAMP2* with other genes and miRNA, yielding a gene ontology enrichment *q* value of 6.02. This approach generated correlated nodes for *RAMP2* and the edges between the nodes of this network (positive and negative correlations are displayed using red and blue colour coding, respectively). (B) The negative correlation of *RAMP2* and *hsa-miR-935* (*p*-value =  $1.087 \times 10^{-6}$ , correlation:  $-0.5139$ ). (C) *RUVBL2*, *SSBP1*, *HSPE1* and *LMAN1* displayed the highest degree of connectivity with 153, 119, 72 and 26 genes, respectively. (D) KEGG gene analysis revealed the enrichment of various ‘cancer’ types, ‘signalling’ and ‘transcriptional regulation’ terms. Relevant *p*- and *q*-values have been reported.



**Figure 9.** The validation of the prognostic value and expression of selected ECM-related genes in other cancers. **(A)** Data retrieved from the Human Pathology Atlas signified the association between *ELFN1*, *HAS3*, *LRP5*, *RUVBL2*, *SSBP1* and *SSBP1* with prognostic predictions in liver, pancreatic, renal, liver, ovarian and liver cancers, respectively. The pink and blue colour-coding represent favourable and unfavourable prognoses, respectively. **(B,C)** *ELFN1* and *RUVBL2* were down- or upregulated in liver hepatocellular carcinoma (LIHC), respectively. Notably, a log<sub>2</sub>-fold change (log<sub>2</sub>FC) = 1 was reported in TPM + 1, a *p*-value <0.01 was applied, and one-way ANOVA was used to test the statistical difference between the tumour and normal tissue groups in GEPIA2. Tumour (T) and matched normal (N) sample counts were 369 and 50, respectively.

**Table 3.** The relevance of the selected ECM-related genes to tumourigenesis pan-cancer.

Gene Name	Cancer Type	Examples of Roles in Tumourigenesis	Reference
<i>AMBN</i>	Osteosarcoma	<i>AMBN</i> promoted osteogenic differentiation by inactivating Src, <i>AMBN</i> induced tumour suppression and sensitivity to doxorubicin through the <i>AMBN</i> -Src-Stat3 axis	[46]
<i>HAS3</i>	Urothelial carcinoma	<i>HAS3</i> downregulation was correlated with tumour grade, invasion and metastasis in urothelial carcinoma	[47]
<i>HSPE1</i>	Clear cell renal cell carcinoma (ccRCC)	<i>HSPE1</i> was significantly downregulated in patient samples of ccRCC	[48]
<i>LMAN1</i>	Microsatellite instability (MSI)-high colorectal cell lines	<i>LMAN1</i> displayed a high mutation frequency in microsatellite instability-high colorectal cancer cell lines	[49]
<i>LRP5</i>	Gastric Cancer	<i>LRP5</i> promoted gastric cancer via the Wnt/ $\beta$ -catenin pathway	[50]
<i>MUC6</i>	Gastric Cancer	<i>MUC6</i> promoter methylation led to downregulating of this gene that promoted gastric cancer progression	[51]
<i>RAMP2</i>	Lung cancer	<i>RAMP2</i> downregulation in lung cancer was correlated with high tumour grade	[52]
<i>RUVBL2</i>	Hepatocellular carcinoma	<i>RUVBL2</i> mRNA overexpression was correlated with reduced recurrence-free survival (RFS) in hepatocellular carcinoma	[45]
<i>SSBP1</i>	Triple-negative breast cancer (TNBC)	<i>SSBP1</i> downregulation in TNBC promoted metastasis in both in vitro and in vivo models	[53]

#### 4. Discussion

Scientific views on the ECM have shifted from describing it as a structure with inert scaffolding properties, to a hugely dynamic network composed of proteins, proteoglycans and glycoproteins that collectively allow for the amplification of signals and the maintenance of its structural integrity [54,55]. ECM is present in all solid tissues, and is regarded as a regulator of cell phenotype and behaviour [56]. Accordingly, tumours respond to both biochemical signals and physical forces, including traction and compression [10,55], and changes in ECM stiffness and its dysregulation have been regarded as a hallmark of cancer [13]. Consistently, tumour cells with the greatest genotypic and phenotypic adaptability will respond to ECM stiffness alterations and dysregulation, thereby influencing ECM remodelling into more favourable environments to support their survival and migration [57–59]. In turn, aberrant ECMs can influence DNA repair mechanisms and induce genetic instability [56,59,60]; hence, the interplay between tumour cells and the ECM is of immense significance in tumour biology. On these grounds, the relationship between ECM-related genes in an aggressive paediatric malignancy of the sympathetic nervous system, NB, was investigated in this study. Previous studies have reported that the ECM composition in 3D cell line cultures and organoids allowed for the realistic recapitulation of the TME, and relevant ECM- and tumour cell-TME interactions. For instance, in one study, collagen-based scaffolds were supplemented with nanohydroxyapatite and glycosaminoglycans, usually found in the bone marrow, a common site for NB metastasis, which permitted NB cell migration and cluster formation [61]. Another study showed that NB patient-derived xenograft (PDX)-derived organoid systems responded to the addition of foetal bovine serum (FBS) and basic fibroblast growth factor (bFGF), and displayed more aggressive behaviour in culture [62]. Further to the significance of the composition of the ECM, the prognostic role of ECM-related genes was also of significance, and this was investigated in the current study using a myriad of databases and integrated bioinformatics

tools. Accordingly, a list of 964 ECM-related genes was obtained from Gene Ontology databases and published studies [15,20], and was subjected to cBioPortal mutational profile analyses based on 1472 NB patient samples from various studies deposited to this database. This analysis, for the most part, yielded genetic alterations of unknown significance with low prevalence (<0.5%) in the NB patient study cohort, while a small fraction of the genetic alterations displayed putative oncogenic activity. The latter included alterations in *EPHA3*, *ITGA11* and *NF1*, while the role of *NF1* as a tumour suppressor gene has been previously reported in NB [63]. 143 TARGET data sets for NB patients contained RNA-sequencing gene expression information, which was also reported for the 946 ECM-related genes. The identification of potential driver alterations, in addition to alterations of unknown significance along with their corresponding gene expression information, may inform future diagnostic and therapeutic investigations. Further, gene ontology enrichment studies of the 964 ECM-related genes using KEGG gene ontology, revealed high enrichment for collagen and integrin-related terms, and remodelling of ECM and platelet degranulation, suggesting a greater enrichment of ECM-related molecular functions. Subsequently, the list of 964 ECM-related genes was subjected to Kaplan–Meier OS analyses that utilised 143 NB patient RNA-sequencing data sets in the cBioPortal (TARGET), and TARGET-NBL in the PCAT databases. The data obtained from both databases showed that NB patients expressing *AMBN*, *ELFN1*, *HSPE1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* displayed lower OS. The predictions for three genes (i.e., *COLQ*, *HAS3* and *LMAN1*) using these two methods reduced and increased NB patient OS, respectively, which suggested that testing larger cohorts of NB patient data may help clarify this result; however, currently no other publicly available databases for NB patient tissue samples are available.

Reviewing the literature revealed that a fraction of the selected ECM-related genes has been studied previously in NB tumourigenesis, but not necessarily in the context of ECM, thus limiting the supporting evidence for the predictions made in this study. However, widening the scope of the search revealed the link between NB and the 12 ECM-related genes in various aspects of tumourigenesis. For example, a Wnt signalling cascade protein, *LRP5*, linked to low-density lipoprotein receptor, displayed a protective role against neurotoxicity [36], while *RAMP2* expression was decreased in NB IMR-32 cells under hypoxic conditions [37]. Additionally, the role of *RUVBL2*, an ER-related protein, in enhancing cell death mediated by a histone deacetylase inhibitor, PCI-24781, in NB has been reported [38]. These three links may be particularly significant about the TME, hypoxic conditions, and stress-survival mechanisms in NB [64], while being inherently associated with an important hallmark of cancer, the deregulation of cellular bioenergetics [65]. Furthermore, the role of *LRP5* and the *G171V* mutation in this gene in mice revealed that the *G171V* mutants developed stiffer bones than their wildtype counterparts, suggesting an exciting link. Despite this, the stiffness of the ECM (a hallmark of cancer) and *LRP5* in the framework of cancer, requires further investigation [66]. Moreover, *HAS3*, essential for hyaluronan synthesis, potentiated NB differentiation through melatonin [35], which conveys important ECM-related links to differentiation-based treatment options in this cancer [67]. Finally, the potential therapeutic roles of *SSBP1* in NB may be of diagnostic and therapeutic significance, perhaps from the angle of tumour cell bioenergetics and mitochondrial function [39].

In the current study, NB samples expressing *AMBN*, *HSPE1*, *LRP5*, *MUC6*, *RAMP2*, *RUVBL2*, *SSBP1* and *UMOD* genes showed a higher HR and reduced OS. An unfavourable histology status also reduced patient OS based on the Cox proportional HR model. The association between these two variables, however, requires further validation. Further, *HSPE1* and *RAMP2* correlated with NB patients with MYCN amplification which is the strongest predictor of poor prognosis in this cancer [68]. Further, *RUVBL2*, *HSPE1*, *SSBP1* and *HAS3* were upregulated in high-risk NB and correlated with reduced OS, suggesting the significance of this ECM-related gene. The cellular and molecular functions of the 12 ECM-related genes were also reported in the study (Table 2). For instance, *HAS3* was involved in the biosynthesis of hyaluronan, a component of the ECM that bears structural significance, which may be linked to oncogenic processes and mechanisms. In melanoma,

*HAS3* overexpression correlated with reduced migration [69]. Collectively, these data, while significant in forming the framework for the discovery of biomarkers, will require validation in preclinical study models and prospective clinical trials. In addition, another limitation of the study was that the 143 NB RNA-seq samples represented 139 NB patients, suggesting that a very small fraction of patients were overrepresented, which may be corrected for in future efforts that validate this study.

Furthermore, the positive correlation of the 12 ECM-related genes *AMBN*, *COLQ*, *ELFN1*, *HAS3*, *HSPE1*, *LMAN1*, *LPR5*, *MUC6*, *RAMP2*, *RUVBL2* and *SSBP1*, with *NRG3-AS1*, *THUMPD3-AS1*, *ELFN1-AS1*, *SLC25A25-AS1*, *EPB41L4A-AS1*, *ATP1A1-AS1*, *VPS9D1-AS1*, *GAS8-AS1*, *EPB41L4A-AS1*, *PITPNA-AS1* and *PITPNA-AS1*, respectively, was also significant since these lncRNAs have been linked to oncogenic processes in previous studies. In evidence, *SLC25A25-AS1* by regulating the *miR-195-5p/ITGA2* axis promoted oncogenic processes in NSCLC [70]. Furthermore, *VPS9D1-AS1* promoted prostate cancer progression through the *miR-4739-MEF2D* axis [71], while *EPB41L4A-AS1* displayed oncogenic functions in colorectal cancer through the Rho/ROCK pathway [72]. Finally, lncRNA *PITPNA-AS1* through the *SOX4-miR-92a-3p* axes promoted gastric cancer [73], collectively providing useful links for future molecular biochemical validations of ECM-related gene-lncRNA-miRNA-pathway axes in NB, although a mechanistic view of these links is currently absent from the literature.

Furthermore, in silico efforts using PCAT demonstrated that *HSPE1* was positively correlated with *RAMP2*, *RUVBL2* and *SSBP1* in NB. This result was confirmed by hierarchical clustering of gene expression, and STRING and MSigDB/BioGRID protein-protein interaction predictions, highlighting their relevance. The interactions of these genes have not been previously reported in the NB literature, which perhaps is the most significant finding of this study. It is plausible that the interaction of *HSPE1*, a heat shock-related protein, and *SSBP1*, a mitochondrial protein, may be linked through stress response-mediated processes in NB tumourigenesis, and through extended protein networks to ECM proteins and processes. Furthermore, GeneMANIA depicted protein-protein interactions between all 12 ECM proteins with other proteins, including those involved in key cellular processes such as ribonucleoprotein assembly and spliceosome complexes, which may be linked to the structural roles of these 12 ECM-related genes.

Cytoscape and Omicsnet revealed that *HSPE1*, *RAMP2*, *RUVBL2*, *SSBP1*, *LMAN1* and *LRP5* formed various protein-protein and protein-miRNA networks, while they displayed oncogenic roles in multiple cancers [45,48,52,53]. Accordingly, the networks formed by *RUVBL2*, *SSBP1*, *HSPE1* and *LPR5* were enriched for gene ontology terms such as 'renal cell carcinoma', 'cancer pathways', 'proteoglycans in cancers' and 'transcriptional misregulation in cancer', which further substantiated links to tumourigenesis, cellular, molecular and biochemical pathways, and transcriptional regulation input in cancers. Noteworthy is that in silico predictions of functional interactions between proteins, genes, miRNAs and lncRNAs are based on high-throughput genetic screens from large consortia.

The associations of selected ECM-related genes including *ELFN1*, *HAS3*, *LPR5*, *RUVBL2* and *SSBP1* with patient prognoses in other cancers including liver, pancreatic, renal and ovarian cancers was also insightful, and put into context the predictive value of these genes across cancer types. Previous studies have also implicated the 12 ECM-related genes in various cancer pathways and processes (Table 3). For instance, *LMAN1/ERGIC53*, whose functions pertain to the transit of glycosylated proteins, was highly mutated in microsatellite instability-high (MSI-H) colorectal cell lines [49]. Specifically, in *LMAN1*-deficient colorectal cell lines, a protein involved in angiogenesis inhibition, alpha 1 antitrypsin, was reduced. This downregulation was correlated with larger tumour size [49,74]. Also, previous studies have shown that *RAMP2* expression, when reduced in lung tumours correlated with higher tumour grades, signified the role of *RAMP2* as a suppressor of tumour growth. Interestingly, the transport of calcitonin-like receptors to the cell membrane fine-tuned the ligand affinity (i.e., adrenomedullin) to this receptor, and this was regulated by *RAMP2*, which could then impact tumour growth, since adrenomedullin suppressed

tumour growth [52]. Consistently, RNA interference (RNAi)-mediated downregulation of *RAMP2* led to increased cell proliferation in this cancer [52]. Furthermore, the role of single-stranded DNA-binding protein 1 (SSBP1) in triple-negative breast cancer (TNBC) was investigated. This study showed that reduced expression of *SSBP1* promoted metastasis by various mechanisms, including TGF $\beta$ -driven epithelial-to-mesenchymal transition (EMT). Reduced *SSBP1* levels also correlated with unfavourable patient prognoses [53]. The presence of conflicting prognostic values in different cancers attributed to a particular gene reflects biological diversity and context-dependent cellular and molecular landscapes [75]. Collectively, the evidence of the implication of the 12 ECM-related genes in various oncogenic processes in other cancers is significant, and lays the groundwork for future studies aimed at validating these links.

## 5. Conclusions

Although requiring prospective clinical trial validation, this study has provided significant links between these 12 selected ECM-related genes with NB patient prognostic predictions and pivotal clinicopathological attributes of NB, including patient risk groups and unfavourable histology, and *MYCN* amplification status. This study also provided strong computational evidence to support the interaction of HSPE1 and SSBP1 proteins in NB and potentially significant networks formed between the 12 ECM-related genes with various lncRNAs and miRNAs previously linked to tumour progression. Future preclinical studies may further validate ECM dynamics and the prognostic predictions reported in this study and pave the way for discovering NB-specific biomarkers linked to ECM and the TME, and assisting the design of ECM-linked therapeutics. Such biomarkers and therapeutics ultimately will improve patient survival and quality of life.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/onco2020007/s1>. Figure S1: Workflow of processes and examples of NB patient OS using microarray data in cBioPortal (TARGET); File S1: The interrogation of the genetic alteration profile, the percentage of patients upregulating the 964 ECM-related genes in NB and the hierarchical clustering of most of the 964 ECM-related genes based on TARGET-NBL expression data in PCAT; File S2: Gene Ontology term enrichment of the 964 ECM-related genes in NB using KEGG tools; File S3: Overall survival of NB patients upregulating 964 ECM-related genes based on TARGET data deposited to cBioPortal using Z = 2 setting, and OS and the corresponding patient number at risk obtained from the OS analysis module in PCAT with auto-calculate or mean settings, and finally the Cox histology status correlations with OS; File S4: Correlation (R) and *p*-values for the pairwise correlation studies presented in Figure 7A; File S5: Correction for multiple testing using FDR per figure number.

**Funding:** This research received no external funding.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data of this article has been obtained from public repositories including but not limited to PCAT and cBioPortal. <http://pedtranscriptome.org>, <http://www.cbioportal.org> (accessed on 12 February 2022), This study's data generated or analysed are included in this article and its Supplementary Files.

**Conflicts of Interest:** The author has no conflict of interest.

## Abbreviations

bFGF	Basic fibroblast growth factor
CI	Confidence interval
ECM	Extra-cellular matrix
EFS	Event-free survival
EMT	Epithelial-to-mesenchymal transition
ER	Endoplasmic reticulum
ERAD	Endoplasmic-associated degradation pathways
ENA	European Nucleotide Archive
FBS	Foetal bovine serum
FDR	False discovery rate
FISH	Fluorescent in situ hybridisation
GSEA	Gene set enrichment analysis
HCC	Hepatocellular carcinoma
HR	Hazard ratio
INRG	International Neuroblastoma Risk Group
LIHC	Liver hepatocellular carcinoma
LncRNAs	Long non-coding RNAs
miRNAs	MicroRNAs
NB	Neuroblastoma
NSCLC	Non-small cell lung cancer
OS	Overall survival
RNAi	RNA interference
RPKM	Reads per kilobase of transcripts per million
PCAT	PDX for childhood therapeutics
PDX	Patient-derived xenograft
RFS	Recurrence-free survival
SSBP1	Single-stranded DNA-binding protein 1
TME	Tumour microenvironment
TNBC	Triple-negative breast cancer
TPM	Transcript per million
UVM	Uveal Melanoma

## References

1. Park, J.R.; Eggert, A.; Caron, H. Neuroblastoma: Biology, prognosis, and treatment. *Pediatr. Clin. N. Am.* **2008**, *55*, 97–120. [[CrossRef](#)] [[PubMed](#)]
2. Newman, E.A.; Abdessalam, S.; Aldrink, J.H.; Austin, M.; Heaton, T.E.; Bruny, J.; Ehrlich, P.; Dasgupta, R.; Baertschiger, R.M.; Lautz, T.B.; et al. Update on neuroblastoma. *J. Pediatr. Surg.* **2019**, *54*, 383–389. [[CrossRef](#)] [[PubMed](#)]
3. Cohn, S.L.; Pearson, A.D.J.; London, W.B.; Monclair, T.; Ambros, P.F.; Brodeur, G.M.; Faldum, A.; Hero, B.; Iehara, T.; Machin, D.; et al. The International Neuroblastoma Risk Group (INRG) classification system: An INRG Task Force report. *J. Clin. Oncol.* **2009**, *27*, 289–297. [[CrossRef](#)] [[PubMed](#)]
4. Brodeur, G.M. Neuroblastoma: Biological insights into a clinical enigma. *Nat. Rev. Cancer* **2003**, *3*, 203–216. [[CrossRef](#)]
5. Maris, J.M.; Hogarty, M.D.; Bagatell, R.; Cohn, S.L. Neuroblastoma. *Lancet* **2007**, *369*, 2106–2120. [[CrossRef](#)]
6. Matthay, K.K.; Maris, J.M.; Schleiermacher, G.; Nakagawara, A.; Mackall, C.L.; Diller, L.; Weiss, W.A. Neuroblastoma. *Nat. Rev. Dis. Prim.* **2016**, *2*, 16078. [[CrossRef](#)]
7. Zafar, A.; Wang, W.; Liu, G.; Wang, X.; Xian, W.; McKeon, F.; Foster, J.; Zhou, J.; Zhang, R. Molecular targeting therapies for neuroblastoma: Progress and challenges. *Med. Res. Rev.* **2021**, *41*, 961–1021. [[CrossRef](#)]
8. Bogen, D.; Brunner, C.; Walder, D.; Ziegler, A.; Abbasi, R.; Ladenstein, R.L.; Noguera, R.; Martinsson, T.; Amann, G.; Schilling, F.H.; et al. The genetic tumor background is an important determinant for heterogeneous MYCN-amplified neuroblastoma. *Int. J. Cancer* **2016**, *139*, 153–163. [[CrossRef](#)]
9. López-Carrasco, A.; Martín-Vañó, S.; Burgos-Panadero, R.; Monferrer, E.; Berbegall, A.P.; Fernández-Blanco, B.; Navarro, S.; Noguera, R. Impact of extracellular matrix stiffness on genomic heterogeneity in MYCN-amplified neuroblastoma cell line. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 226. [[CrossRef](#)]
10. Lam, W.A.; Cao, L.; Umesh, V.; Keung, A.J.; Sen, S.; Kumar, S. Extracellular matrix rigidity modulates neuroblastoma cell differentiation and N-myc expression. *Mol. Cancer* **2010**, *9*, 35. [[CrossRef](#)]
11. Alitalo, K.; Keski-Oja, J.; Vaheri, A. Extracellular matrix proteins characterize human tumor cell lines. *Int. J. Cancer* **1981**, *27*, 755–761. [[CrossRef](#)] [[PubMed](#)]

12. Spill, F.; Reynolds, D.S.S.; Kamm, R.D.D.; Zaman, M.H.H. Impact of the physical microenvironment on tumor progression and metastasis. *Curr. Opin. Biotechnol.* **2016**, *40*, 41–48. [[CrossRef](#)] [[PubMed](#)]
13. Pickup, M.W.; Mouw, J.K.; Weaver, V.M. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep.* **2014**, *15*, 1243–1253. [[CrossRef](#)] [[PubMed](#)]
14. Zhu, S.; Zhang, X.; Weichert-Leahey, N.; Dong, Z.; Zhang, C.; Lopez, G.; Tao, T.; He, S.; Wood, A.C.; Oldridge, D.; et al. LMO1 Synergizes with MYCN to Promote Neuroblastoma Initiation and Metastasis. *Cancer Cell* **2017**, *32*, 310–323.e5. [[CrossRef](#)]
15. Ahluwalia, P.; Ahluwalia, M.; Mondal, A.K.; Sahajpal, N.; Kota, V.; Rojiani, M.V.; Rojiani, A.M.; Kolhe, R. Prognostic and therapeutic implications of extracellular matrix associated gene signature in renal clear cell carcinoma. *Sci. Rep.* **2021**, *11*, 7561. [[CrossRef](#)]
16. Durinck, K.; Speleman, F. Epigenetic regulation of neuroblastoma development. *Cell Tissue Res.* **2018**, *372*, 309–324. [[CrossRef](#)]
17. Mondal, T.; Juvvuna, P.K.; Kirkeby, A.; Mitra, S.; Kosalai, S.T.; Traxler, L.; Hertwig, F.; Wernig-Zorc, S.; Miranda, C.; Deland, L.; et al. Sense-Antisense lncRNA Pair Encoded by Locus 6p22.3 Determines Neuroblastoma Susceptibility via the USP36-CHD7-SOX9 Regulatory Axis. *Cancer Cell* **2018**, *33*, 417–434.e7. [[CrossRef](#)]
18. Lecerf, C.; Le Bourhis, X.; Adriaenssens, E. The long non-coding RNA H19: An active player with multiple facets to sustain the hallmarks of cancer. *Cell. Mol. Life Sci.* **2019**, *76*, 4673–4687. [[CrossRef](#)]
19. Zhu, K.; Wang, L.; Zhang, X.; Sun, H.; Chen, T.; Sun, C.; Zhang, F.; Zhu, Y.; Yu, X.; He, X.; et al. lncRNA HCP5 promotes neuroblastoma proliferation by regulating miR-186-5p/MAP3K2 signal axis. *J. Pediatr. Surg.* **2021**, *56*, 778–787. [[CrossRef](#)]
20. Carbon, S.; Ireland, A.; Mungall, C.J.; Shu, S.; Marshall, B.; Lewis, S.; AmiGO Hub; Web Presence Working Group. AmiGO: Online access to ontology and annotation data. *Bioinformatics* **2009**, *25*, 288–289. [[CrossRef](#)]
21. 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](#)]
22. Yang, J.; Li, Q.; Noureen, N.; Fang, Y.; Kurmasheva, R.; Houghton, P.J.; Wang, X.; Zheng, S. PCAT: An integrated portal for genomic and preclinical testing data of pediatric cancer patient-derived xenograft models. *Nucleic Acids Res.* **2021**, *49*, D1321–D1327. [[CrossRef](#)] [[PubMed](#)]
23. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [[CrossRef](#)] [[PubMed](#)]
24. Mootha, V.K.; Lindgren, C.M.; Eriksson, K.-F.; Subramanian, A.; Sihag, S.; Lehar, J.; Puigserver, P.; Carlsson, E.; Ridderstråle, M.; Laurila, E.; et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **2003**, *34*, 267–273. [[CrossRef](#)]
25. Liu, B.; Chen, Y.; Yang, J. lncRNAs are altered in lung squamous cell carcinoma and lung adenocarcinoma. *Oncotarget* **2017**, *8*, 24275–24291. [[CrossRef](#)]
26. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
27. Chatr-aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N.K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **2017**, *45*, D369–D379. [[CrossRef](#)]
28. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
29. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. [[CrossRef](#)]
30. Zhou, G.; Jianguo, X. OmicsNet: A web-based tool for creation and visual analysis of biological networks in 3D space. *Nucleic Acids Res.* **2018**, *46*, W514–W522. [[CrossRef](#)]
31. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [[CrossRef](#)] [[PubMed](#)]
32. Combaret, V.; Iacono, I.; Bellini, A.; Bréjon, S.; Bernard, V.; Marabelle, A.; Coze, C.; Pierron, G.; Lapouble, E.; Schleiermacher, G.; et al. Detection of tumor ALK status in neuroblastoma patients using peripheral blood. *Cancer Med.* **2015**, *4*, 540–550. [[CrossRef](#)] [[PubMed](#)]
33. Duan, C.; Wang, H.; Chen, Y.; Chu, P.; Xing, T.; Gao, C.; Yue, Z.; Zheng, J.; Jin, M.; Gu, W.; et al. Whole exome sequencing reveals novel somatic alterations in neuroblastoma patients with chemotherapy. *Cancer Cell Int.* **2018**, *18*, 21. [[CrossRef](#)]
34. van Groningen, T.; Koster, J.; Valentijn, L.J.; Zwijnenburg, D.A.; Akogul, N.; Hasselt, N.E.; Broekmans, M.; Haneveld, F.; Nowakowska, N.E.; Bras, J.; et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat. Genet.* **2017**, *49*, 1261–1266. [[CrossRef](#)]
35. Lee, W.-J.; Chen, L.-C.; Lin, J.-H.; Cheng, T.-C.; Kuo, C.-C.; Wu, C.-H.; Chang, H.-W.; Tu, S.-H.; Ho, Y.-S. Melatonin promotes neuroblastoma cell differentiation by activating hyaluronan synthase 3-induced mitophagy. *Cancer Med.* **2019**, *8*, 4821–4835. [[CrossRef](#)] [[PubMed](#)]

36. Zhang, L.; Bahety, P.; Ee, P.L.R. Wnt co-receptor LRP5/6 overexpression confers protection against hydrogen peroxide-induced neurotoxicity and reduces tau phosphorylation in SH-SY5Y cells. *Neurochem. Int.* **2015**, *87*, 13–21. [[CrossRef](#)] [[PubMed](#)]
37. Kitamuro, T.; Takahashi, K.; Totsune, K.; Nakayama, M.; Murakami, O.; Hida, W.; Shirato, K.; Shibahara, S. Differential expression of adrenomedullin and its receptor component, receptor activity modifying protein (RAMP) 2 during hypoxia in cultured human neuroblastoma cells. *Peptides* **2001**, *22*, 1795–1801. [[CrossRef](#)]
38. Zhan, Q.; Tsai, S.; Lu, Y.; Wang, C.; Kwan, Y.; Ngai, S. RuvBL2 is involved in histone deacetylase inhibitor PCI-24781-induced cell death in SK-N-DZ neuroblastoma cells. *PLoS ONE* **2013**, *8*, e71663. [[CrossRef](#)]
39. Batra, R.; Harder, N.; Gogolin, S.; Diessl, N.; Soons, Z.; Jäger-Schmidt, C.; Lawerenz, C.; Eils, R.; Rohr, K.; Westermann, F.; et al. Time-lapse imaging of neuroblastoma cells to determine cell fate upon gene knockdown. *PLoS ONE* **2012**, *7*, e50988. [[CrossRef](#)]
40. Jahangiri, L.; Pucci, P.; Ishola, T.; Trigg, R.M.; Williams, J.A.; Pereira, J.; Cavanagh, M.L.; Turner, S.D.; Gkoutos, G.V.; Tsaprouni, L. The Contribution of Autophagy and LncRNAs to MYC-Driven Gene Regulatory Networks in Cancers. *Int. J. Mol. Sci.* **2021**, *22*, 8527. [[CrossRef](#)]
41. Hu, J.; Chen, Y.; Li, X.; Miao, H.; Li, R.; Chen, D.; Wen, Z. THUMP3-AS1 Is Correlated With Non-Small Cell Lung Cancer And Regulates Self-Renewal Through miR-543 And ONECUT2. *Onco. Targets. Ther.* **2019**, *12*, 9849–9860. [[CrossRef](#)] [[PubMed](#)]
42. Zhang, C.; Lian, H.; Xie, L.; Yin, N.; Cui, Y. LncRNA ELFN1-AS1 promotes esophageal cancer progression by up-regulating GFPT1 via sponging miR-183-3p. *Biol. Chem.* **2020**, *401*, 1053–1061. [[CrossRef](#)] [[PubMed](#)]
43. Zhan, Y.; Chen, Z.; He, S.; Gong, Y.; He, A.; Li, Y.; Zhang, L.; Zhang, X.; Fang, D.; Li, X.; et al. Long non-coding RNA SOX2OT promotes the stemness phenotype of bladder cancer cells by modulating SOX2. *Mol. Cancer* **2020**, *19*, 25. [[CrossRef](#)] [[PubMed](#)]
44. Huang, G.; Chen, J.; Liu, J.; Zhang, X.; Duan, H.; Fang, Q. MiR-935/HIF1 $\alpha$  Feedback Loop Inhibits the Proliferation and Invasiveness of Glioma. *Onco. Targets. Ther.* **2020**, *13*, 10817–10828. [[CrossRef](#)] [[PubMed](#)]
45. Yan, T.; Liu, F.; Gao, J.; Lu, H.; Cai, J.; Zhao, X.; Sun, Y. Multilevel regulation of RUVBL2 expression predicts poor prognosis in hepatocellular carcinoma. *Cancer Cell Int.* **2019**, *19*, 249. [[CrossRef](#)]
46. Ando, T.; Kudo, Y.; Iizuka, S.; Tsunematsu, T.; Umehara, H.; Shrestha, M.; Matsuo, T.; Kubo, T.; Shimose, S.; Arihiro, K.; et al. Ameloblastin induces tumor suppressive phenotype and enhances chemosensitivity to doxorubicin via Src-Stat3 inactivation in osteosarcoma. *Sci. Rep.* **2017**, *7*, 40187. [[CrossRef](#)]
47. Chang, I.-W.; Liang, P.-I.; Li, C.-C.; Wu, W.-J.; Huang, C.-N.; Lin, V.C.-H.; Hsu, C.-T.; He, H.-L.; Wu, T.-F.; Hung, C.-H.; et al. HAS3 underexpression as an indicator of poor prognosis in patients with urothelial carcinoma of the upper urinary tract and urinary bladder. *Tumour Biol.* **2015**, *36*, 5441–5450. [[CrossRef](#)]
48. White, N.M.A.; Masui, O.; DeSouza, L.V.; Krakovska, O.; Metias, S.; Romaschin, A.D.; Honey, R.J.; Stewart, R.; Pace, K.; Lee, J.; et al. Quantitative proteomic analysis reveals potential diagnostic markers and pathways involved in pathogenesis of renal cell carcinoma. *Oncotarget* **2014**, *5*, 506–518. [[CrossRef](#)]
49. Roeckel, N.; Woerner, S.M.; Kloor, M.; Yuan, Y.-P.; Patsos, G.; Gromes, R.; Kopitz, J.; Gebert, J. High frequency of LMAN1 abnormalities in colorectal tumors with microsatellite instability. *Cancer Res.* **2009**, *69*, 292–299. [[CrossRef](#)]
50. Nie, X.; Wang, H.; Wei, X.; Li, L.; Xue, T.; Fan, L.; Ma, H.; Xia, Y.; Wang, Y.-D.; Chen, W.-D. LRP5 Promotes Gastric Cancer via Activating Canonical Wnt/ $\beta$ -Catenin and Glycolysis Pathways. *Am. J. Pathol.* **2022**, *192*, 503–517. [[CrossRef](#)]
51. Shi, D.; Xi, X.-X. Regulation of MUC6 Methylation Correlates with Progression of Gastric Cancer. *Yonsei Med. J.* **2021**, *62*, 1005–1015. [[CrossRef](#)] [[PubMed](#)]
52. Yue, W.; Dacic, S.; Sun, Q.; Landreneau, R.; Guo, M.; Zhou, W.; Siegfried, J.M.; Yu, J.; Zhang, L. Frequent inactivation of RAMP2, EFEMP1 and Dutt1 in lung cancer by promoter hypermethylation. *Clin. Cancer Res.* **2007**, *13*, 4336–4344. [[CrossRef](#)] [[PubMed](#)]
53. Jiang, H.-L.; Sun, H.-F.; Gao, S.-P.; Li, L.-D.; Huang, S.; Hu, X.; Liu, S.; Wu, J.; Shao, Z.-M.; Jin, W. SSBP1 Suppresses TGF $\beta$ -Driven Epithelial-to-Mesenchymal Transition and Metastasis in Triple-Negative Breast Cancer by Regulating Mitochondrial Retrograde Signaling. *Cancer Res.* **2016**, *76*, 952–964. [[CrossRef](#)] [[PubMed](#)]
54. Burgos-Panadero, R.; Lucantoni, F.; Gamero-Sandemetrio, E.; de la Cruz-Merino, L.; Álvaro, T.; Noguera, R. The tumour microenvironment as an integrated framework to understand cancer biology. *Cancer Lett.* **2019**, *461*, 112–122. [[CrossRef](#)] [[PubMed](#)]
55. Noguera, R.; Nieto, O.A.; Tadeo, I.; Fariñas, F.; Alvaro, T. Extracellular matrix, biotensegrity and tumor microenvironment. An update and overview. *Histol. Histopathol.* **2012**, *27*, 693–705. [[CrossRef](#)] [[PubMed](#)]
56. Filipe, E.C.; Chitty, J.L.; Cox, T.R. Charting the unexplored extracellular matrix in cancer. *Int. J. Exp. Pathol.* **2018**, *99*, 58–76. [[CrossRef](#)]
57. Schulte, M.; Köster, J.; Rahmann, S.; Schramm, A. Cancer evolution, mutations, and clonal selection in relapse neuroblastoma. *Cell Tissue Res.* **2018**, *372*, 263–268. [[CrossRef](#)]
58. Greaves, M. Nothing in cancer makes sense except . . . . *BMC Biol.* **2018**, *16*, 22. [[CrossRef](#)]
59. Berger, A.J.; Renner, C.M.; Hale, I.; Yang, X.; Ponik, S.M.; Weisman, P.S.; Masters, K.S.; Kreeger, P.K. Scaffold stiffness influences breast cancer cell invasion via EGFR-linked Mena upregulation and matrix remodeling. *Matrix Biol.* **2020**, *85–86*, 80–93. [[CrossRef](#)]
60. Discher, D.E.; Janmey, P.; Wang, Y.-L. Tissue cells feel and respond to the stiffness of their substrate. *Science* **2005**, *310*, 1139–1143. [[CrossRef](#)]
61. Gallagher, C.; Murphy, C.; O'Brien, F.J.; Piskareva, O. Three-dimensional In Vitro Biomimetic Model of Neuroblastoma using Collagen-based Scaffolds. *J. Vis. Exp.* **2021**, *9*, 173. [[CrossRef](#)] [[PubMed](#)]

62. Gavin, C.; Geerts, N.; Cavanagh, B.; Haynes, M.; Reynolds, C.P.; Loessner, D.; Ewald, A.J.; Piskareva, O. Neuroblastoma Invasion Strategies Are Regulated by the Extracellular Matrix. *Cancers* **2021**, *13*, 736. [[CrossRef](#)] [[PubMed](#)]
63. Hölzel, M.; Huang, S.; Koster, J.; Ora, I.; Lakeman, A.; Caron, H.; Nijkamp, W.; Xie, J.; Callens, T.; Asgharzadeh, S.; et al. NF1 is a tumor suppressor in neuroblastoma that determines retinoic acid response and disease outcome. *Cell* **2010**, *142*, 218–229. [[CrossRef](#)] [[PubMed](#)]
64. Jahangiri, L.; Ishola, T.; Pucci, P.; Trigg, R.M.; Pereira, J.; Williams, J.A.; Cavanagh, M.L.; Gkoutos, G.V.; Tsaprouni, L.; Turner, S.D. The Role of Autophagy and lncRNAs in the Maintenance of Cancer Stem Cells. *Cancers* **2021**, *13*, 1239. [[CrossRef](#)] [[PubMed](#)]
65. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [[CrossRef](#)]
66. Babij, P.; Zhao, W.; Small, C.; Kharode, Y.; Yaworsky, P.J.; Bouxsein, M.L.; Reddy, P.S.; Bodine, P.V.N.; Robinson, J.A.; Bhat, B.; et al. High bone mass in mice expressing a mutant LRP5 gene. *J. Bone Miner. Res.* **2003**, *18*, 960–974. [[CrossRef](#)]
67. Pählman, S.; Ruusala, A.I.; Abrahamsson, L.; Mattsson, M.E.; Esscher, T. Retinoic acid-induced differentiation of cultured human neuroblastoma cells: A comparison with phorbol ester-induced differentiation. *Cell Differ.* **1984**, *14*, 135–144. [[CrossRef](#)]
68. Pession, A.; De Bernardi, B.; Perri, P.; Mazzocco, K.; Rondelli, R.; Nigro, M.; Iolascon, A.; Forni, M.; Basso, G.; Conte, M.; et al. The prognostic effect of amplification of MYCN oncogene in neuroblastoma. The preliminary results of the Italian Cooperative Group for Neuroblastoma (GCINB). *La Pediatr. Medica Chir.* **1994**, *16*, 211–218.
69. Takabe, P.; Bart, G.; Ropponen, A.; Rilla, K.; Tammi, M.; Tammi, R.; Pasonen-Seppänen, S. Hyaluronan synthase 3 (HAS3) overexpression downregulates MV3 melanoma cell proliferation, migration and adhesion. *Exp. Cell Res.* **2015**, *337*, 1–15. [[CrossRef](#)]
70. Chen, J.; Gao, C.; Zhu, W. Long non-coding RNA SLC25A25-AS1 exhibits oncogenic roles in non-small cell lung cancer by regulating the microRNA-195-5p/ITGA2 axis. *Oncol. Lett.* **2021**, *22*, 529. [[CrossRef](#)]
71. Wang, X.; Chen, Q.; Wang, X.; Li, W.; Yu, G.; Zhu, Z.; Zhang, W. ZEB1 activated-VPS9D1-AS1 promotes the tumorigenesis and progression of prostate cancer by sponging miR-4739 to upregulate MEF2D. *Biomed. Pharmacother.* **2020**, *122*, 109557. [[CrossRef](#)] [[PubMed](#)]
72. Bin, J.; Nie, S.; Tang, Z.; Kang, A.; Fu, Z.; Hu, Y.; Liao, Q.; Xiong, W.; Zhou, Y.; Tang, Y.; et al. Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. *J. Cell. Physiol.* **2021**, *236*, 523–535. [[CrossRef](#)] [[PubMed](#)]
73. Liu, L.; Dai, A.; Zhang, Z.; Ning, M.; Han, D.; Li, L.; Li, Z. LncRNA PIPNA-AS1 promotes gastric cancer by increasing SOX4 expression via inhibition of miR-92a-3p. *Aging* **2021**, *13*, 21191–21201. [[CrossRef](#)] [[PubMed](#)]
74. Huang, H.; Campbell, S.C.; Nelius, T.; Bedford, D.F.; Veliceasa, D.; Bouck, N.P.; Volpert, O. V Alpha1-antitrypsin inhibits angiogenesis and tumor growth. *Int. J. Cancer* **2004**, *112*, 1042–1048. [[CrossRef](#)] [[PubMed](#)]
75. Jahangiri, L.; Pucci, P.; Ishola, T.; Pereira, J.; Cavanagh, M.L.; Turner, S.D. Deep analysis of neuroblastoma core regulatory circuitries using online databases and integrated bioinformatics shows their pan-cancer roles as prognostic predictors. *Discov. Oncol.* **2021**, *12*, 56. [[CrossRef](#)] [[PubMed](#)]