

## Article

# Pan-Cancer Analysis Identifies *MNX1* and Associated Antisense Transcripts as Biomarkers for Cancer

Denise Ragusa <sup>1,2,3</sup> , Sabrina Tosi <sup>1,2,3</sup>  and Cristina Sisu <sup>2,3,\*</sup> 

- <sup>1</sup> Leukaemia and Chromosome Research Laboratory, College of Health, Medicine and Life Sciences, Brunel University London, Kingston Lane, Uxbridge UB8 3PH, UK
- <sup>2</sup> Centre for Genome Engineering and Maintenance (CenGEM), College of Health, Medicine and Life Sciences, Brunel University London, Kingston Lane, Uxbridge UB8 3PH, UK
- <sup>3</sup> Department of Life Sciences, College of Health, Medicine and Life Sciences, Brunel University London, Kingston Lane, Uxbridge UB8 3PH, UK
- \* Correspondence: [cristina.sisu@brunel.ac.uk](mailto:cristina.sisu@brunel.ac.uk)

**Abstract:** The identification of diagnostic and prognostic biomarkers is a major objective in improving clinical outcomes in cancer, which has been facilitated by the availability of high-throughput gene expression data. A growing interest in non-coding genomic regions has identified dysregulation of long non-coding RNAs (lncRNAs) in several malignancies, suggesting a potential use as biomarkers. In this study, we leveraged data from large-scale sequencing projects to uncover the expression patterns of the *MNX1* gene and its associated lncRNAs *MNX1-AS1* and *MNX1-AS2* in solid tumours. Despite many reports describing *MNX1* overexpression in several cancers, limited studies exist on *MNX1-AS1* and *MNX1-AS2* and their potential as biomarkers. By employing clustering methods to visualise multi-gene relationships, we identified a discriminative power of the three genes in distinguishing tumour vs. normal samples in several cancers of the gastrointestinal tract and reproductive systems, as well as in discerning oesophageal and testicular cancer histological subtypes. Notably, the expressions of *MNX1* and its antisenses also correlated with clinical features and endpoints, uncovering previously unreported associations. This work highlights the advantages of using combinatory expression patterns of non-coding transcripts of differentially expressed genes as clinical evaluators and identifies *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions as robust candidate biomarkers for clinical applications.

**Keywords:** transcription; pan-cancer; *MNX1*; biomarker; diagnosis; prognosis; cancer cell development



**Citation:** Ragusa, D.; Tosi, S.; Sisu, C. Pan-Cancer Analysis Identifies *MNX1* and Associated Antisense Transcripts as Biomarkers for Cancer. *Cells* **2022**, *11*, 3577. <https://doi.org/10.3390/cells11223577>

Academic Editor: Amancio Carnero

Received: 6 September 2022

Accepted: 3 November 2022

Published: 11 November 2022

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



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

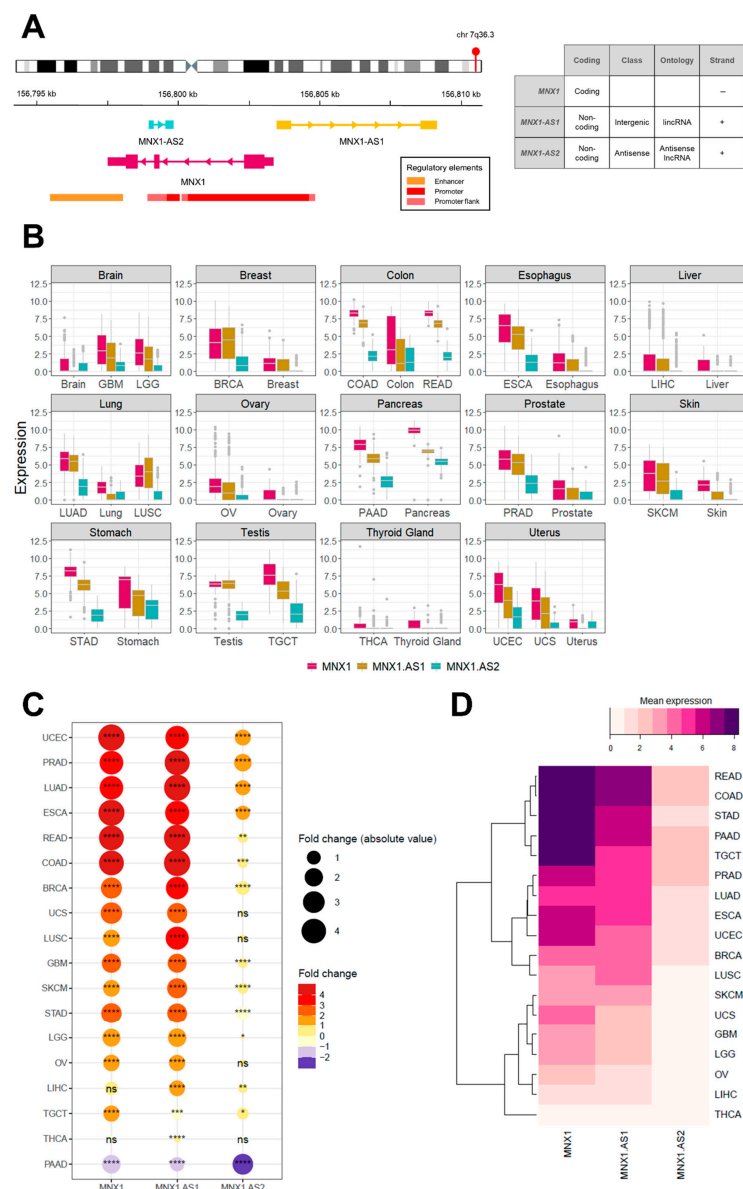
## 1. Introduction

Cancer represents a major contributor of morbidity and mortality worldwide, with current estimates suggesting that one in two people in the UK will develop cancer in their lifetime [1,2]. Despite significant improvements in therapeutic interventions and clinical outcomes in recent years, accurate patient diagnoses are imperative for maximising clinical success. Thus, the identification of genes as diagnostic and prognostic markers is a major objective in cancer research. The availability of large genome-wide studies is now facilitating the discovery of clinically relevant genes [3,4].

Motor neuron and pancreas homeobox protein 1 (*MNX1*), also known as *HLXB9*, is a key developmental homeobox gene coding for the transcription factor HB9. *MNX1* is involved in the differentiation and development of pancreatic and neuronal cells, and commonly expressed in these respective tissues [5–8]. The past decade saw a plethora of studies highlighting the dysregulation of *MNX1* in various cancers, including leukaemia [9], lymphoma [10], cancers of the breast [11,12], prostate [13], bladder [14], colorectal [15], brain [16,17], hepatocellular carcinomas [18], and pancreatic tumours [19]. However, the function and activity of *MNX1* in cancer is yet to be elucidated.

At the chromosomal locus 7q36.3, the *MNX1* gene is associated with two long non-coding RNAs (lncRNA)—the antisense transcripts *MNX1-AS1* (long intergenic non-coding RNA; lincRNA) and *MNX1-AS2* (antisense lincRNA). *MNX1* and *MNX1-AS1* share a promoter region, while *MNX1-AS2* is under the control of a second intergenic promoter (Figure 1A), however their transcriptional regulation remains elusive. Earlier studies have associated *MNX1-AS1* with the pathogenesis of a wide range of malignancies (e.g., bladder, breast, lung, prostate, oesophageal, gastrointestinal (GI), and gynaecological) in virtue of its regulatory activity on modulating malignant processes of proliferation, apoptosis, invasion/metastases, and migration (reviewed in [20]). The overexpression of *MNX1-AS1* has also been correlated with dismal clinical outcomes, such as shorter overall survival (OS), advanced clinical stage, and TNM classifiers [21]. By contrast, to date there are no reports on the function and activity of *MNX1-AS2* in cancer.

In this study, we leveraged the large-scale data from the Cancer Genome Atlas (TCGA) and Genotype Tissue Expression (GTEx) consortia projects to explore the transcriptional activity and function of the *MNX1* gene and its associated lncRNAs *MNX1-AS1* and *MNX1-AS2* in cancer and to uncover their clinical potential as biomarkers.



**Figure 1.** Overview of expression of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in cancers of the TCGA

cohort. (A) Schematic representation of the chromosome 7q36.3 locus and location of the *MNX1* gene and its associated non-coding transcripts *MNX1-AS1* and *MNX1-AS2*. Arrows within the gene structure indicate the direction of transcription and vertical boxes represent gene exons. Underneath, known regulatory elements are shown. The table on the right summarises the classification and genomic features of the three genes. (B) Expression values of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* by anatomical site in cancer (indicated by TCGA acronyms) and normal tissues (indicated in lower case). (C) Fold change in expression between tumour (TCGA) and normal (GTEx) by tumour site and statistical significance of the differential expression. The size of the dots represents the absolute fold change value and the colour gradients indicate increased (positive) or decreased (negative) fold change. (D) Hierarchical clustering of TCGA cancer by mean expression values of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* revealed similarities in patterns of expression of the three genes. +: positive strand; -: negative strand; Statistical significance is symbolised as:  $p$  value < 0.05 (\*), 0.001 (\*\*), 0.0001 (\*\*\*), 0.00001 (\*\*\*\*), ns: not significant.

## 2. Methods

### 2.1. Data Availability, Sample Selection and Expression Analysis

Clinical phenotype and expression data for unmatched healthy and tumour samples were extracted from the TCGA and GTEx data, available in the TCGA-TARGET-GTEX cohort from the University of California Santa Cruz public repository, Xena [22]. The data were processed by UCSC from raw RNAseq reads using the TOIL pipeline ensuring the quantification and cross-dataset normalisation of gene expression without any computational batch effects [23]. All gene expression values are presented in units of  $\log_2(\text{norm\_count} + 1)$ . Differential expression between normal and tumour sites was calculated as fold change, defined as the ratio between tumour and healthy tissue samples. In this study we used only primary tumour data, and further filtered the samples to remove those lacking expression information for the genes of interest. The summary of analysed samples is shown in Table 1. While we have used the largest uniformly processed and normalized dataset available for tumour and healthy tissue samples, we are aware that using a single data set can influence the outcome of our results and that this is a limitation of our study. However, we addressed this issue by evaluating the potential for analysis artifacts by splitting the data into two randomly selected groups and measuring the change in the average expression levels for our genes of interests (Supplementary Table S3). Our results show no change between the two groups suggesting that there is little chance for a data bias to impact the biological insights we have drawn from the data.

**Table 1.** Clinical details of TCGA and GTEx cohorts.

Organ	Breast	Colon		Oesophagus	Brain		Liver	Lung		Ovary	Pancreas	Prostate	Skin	Stomach	Testis	Thyroid	Uterus	
TCGA Cancer and Abbreviation	Breast Cancer BRCA	Colon Adenocarcinoma COAD	Rectum Adenocarcinoma READ	Oesophageal Carcinoma ESCA	Glioblastoma Multiforme GBM	Brain Lower Grade Glioma LGG	Liver Hepatocellular Carcinoma LIHC	Lung Adenocarcinoma LUAD	Lung Squamous Cell Carcinoma LUSC	Ovarian Serous Cystadenocarcinoma OV	Pancreatic Adenocarcinoma PAAD	Prostate Adenocarcinoma PRAD	Skin Cutaneous Melanoma SKCM	Stomach Adenocarcinoma STAD	Testicular Germ Cell Tumours TGCT	Thyroid Carcinoma THCA	Uterine Corpus Endometrial Carcinoma UCEC	Uterine Carcinoma UCS
Samples Normal tissue (GTEx)	179	307	307	650	1141	1141	109	288	288	86	167	100	788	173	164	279	77	77
Primary tumour (TCGA)	1212	331	103	195	171	523	421	574	548	427	183	548	102	450	154	571	204	57
Clinical stage																		
Stage I																		
Stage II																		
Stage III																		
Stage IV																		
N/A	97 (8.3%)	31 (10.1%)	21 (20.4%)	37 (19.0%)	N/A	N/A	84 (20.0%)	9 (1.6%)	8 (1.4%)	10 (2.3%)	8 (4.4%)	N/A	22 (18.3%)	61 (13.6%)	24 (15.6%)	70 (12.3%)	N/A	0 (0.0%)
	20 (1.7%)	40 (13.0%)	13 (12.6%)	9 (4.6%)	N/A	N/A	5 (1.2%)	28 (4.9%)	8 (1.4%)	63 (14.7%)	4 (2.2%)	N/A	27 (22.5%)	41 (9.1%)	0 (0.0%)	55 (9.6%)	N/A	10 (17.5%)
	248 (21.3%)	82 (26.7%)	33 (32.0%)	55 (28.2%)	N/A	N/A	77 (18.3%)	96 (16.7%)	89 (16.1%)	332 (77.2%)	3 (1.6%)	N/A	2 (1.7%)	169 (37.6%)	14 (9.1%)	112 (19.6%)	N/A	20 (35.1%)
	182 (15.6%)	44 (14.3%)	12 (11.7%)	18 (9.2%)	N/A	N/A	169 (40.1%)	306 (53.3%)	269 (48.7%)	1 (0.2%)	147 (80.3%)	N/A	3 (2.5%)	58 (12.9%)	104 (67.5%)	283 (49.6%)	N/A	22 (38.6%)
	617 (53.0%)	110 (35.8%)	24 (23.3%)	76 (39.0%)	N/A	N/A	86 (20.4%)	135 (23.5%)	178 (32.2%)	24 (5.6%)	147 (80.3%)	N/A	66 (55.0%)	121 (26.9%)	12 (7.8%)	51 (8.9%)	N/A	5 (8.8%)

Table 1. Cont.

TCGA Cancer and Abbreviation	Organ	Age GTEX (years)	Age < 50	Age > 50	Age TCGA (years)	Age < 50	Age > 50
Breast Cancer BRCA	Breast		115 (64.3%) 64 (35.7%)		763 (69.9%) 328 (30.1%)		
Colon Adenocarcinoma COAD	Colon		199 (64.9%) 108 (35.1%)		242 (84.6%) 44 (15.4%)		
Rectum Adenocarcinoma READ			74 (81.3%) 17 (18.7%)				
Oesophageal Carcinoma ESCA	Oesophagus		398 (61.3%) 252 (38.7%)		154 (85.1%) 27 (14.9%)		
Glioblastoma Multiforme GBM	Brain		963 (84.4%) 178 (15.6%)		116 (76.3%) 36 (23.7%)		
Brain Lower Grade Glioma LGG			155 (31.1%) 343 (68.9%)		155 (31.1%) 343 (68.9%)		
Liver Hepatocellular Carcinoma LIHC	Liver		80 (73.3%) 29 (26.6%)		291 (80.4%) 71 (19.6%)		
Lung Adenocarcinoma LUAD	Lung		195 (67.9%) 92 (32.1%)		375 (91.7%) 34 (8.3%)		
Lung Squamous Cell Carcinoma LUSC			519 (96.3%) 20 (3.7%)		519 (96.3%) 20 (3.7%)		
Ovarian Serous Cystadenocarcinoma OV	Ovary		48 (55.8%) 38 (44.2%)		322 (76.1%) 101 (23.9%)		
Pancreatic Adenocarcinoma PAAD	Pancreas		100 (59.9%) 67 (40.1%)		154 (86.5%) 24 (13.5%)		
Prostate Adenocarcinoma PRAD	Prostate		56 (56.0%) 44 (44.0%)		461 (93.1%) 34 (6.9%)		
Skin Cutaneous Melanoma SKCM	Skin		522 (66.2%) 266 (33.8%)		87 (85.3%) 15 (14.7%)		
Stomach Adenocarcinoma STAD	Stomach		101 (58.4%) 72 (41.6%)		375 (91.7%) 34 (8.3%)		
Testicular Germ Cell Tumours TGCT	Testis		102 (62.1%) 62 (37.9%)		5 (3.9%) 122 (96.1%)		
Thyroid Carcinoma THCA	Thyroid		193 (69.1%) 86 (30.9%)		215 (44.1%) 272 (55.9%)		
Uterine Corpus Endometrial Carcinoma UCEC	Uterus		39 (50.6%) 38 (49.3%)		166 (93.8%) 11 (6.2%)		
Uterine Carcinosarcoma UCS			57 (100.0%) 0 (0.0%)		57 (100.0%) 0 (0.0%)		

## 2.2. Prediction of Targets

The target prediction for *MNX1-AS1* and *MNX1-AS2* was performed on RNAInter 4.0 [24] by selecting the top 200 gene from the RNA-Protein hits based on the score. The top 200 targets of *MNX1* were extracted from JASPAR database of transcription factor binding sites [25] based on the score. The Venn diagrams were generated on VennPainter [26].

## 2.3. Differential Expression Analysis

RNA-sequencing values of transcripts were downloaded in counts units from the University of California Santa Cruz public repository [22]. The differential expression analysis was performed using limma (*limma* package version 3.48.3 in R statistical software environment version 4.0.3-The R Foundation for statistical Computing, Vienna, Austria implemented in R studio version 1.4.1717, RStudio, Boston, MA, USA). The statistically significant genes were extracted by filtering by  $p$  adjusted value  $< 0.01$ .

## 2.4. Gene Ontology Analysis

The Gene Ontology (GO) enrichment analysis and network construction was performed on ExpressAnalyst (available at [www.expressanalyst.ca](http://www.expressanalyst.ca), last accessed on 20 October 2022) using the PANTHER Biological Process database. The significance was determined using  $p$  value  $< 0.05$ .

## 2.5. Statistical Analysis

All statistical analyses were carried out using the standard libraries (*stats* 4.0.2, *ggpubr* 0.4.0) in R (statistical software environment version 4.0.3-The R Foundation for statistical Computing, Vienna, Austria implemented in R studio version 1.4.1717, RStudio, Boston, MA, USA). The Wilcoxon test was used to evaluate the statistical significance, with a set threshold for significance set to a  $p$  value  $\leq 0.05$ . The correlation between clinicopathological features and expression levels was calculated using the Chi-squared test. The gene expression was defined as “high” or “low” based on the absolute expression value relative to the mean average of the respective phenotype. The receiver-operating characteristic (ROC) was used to calculate the area under the curve (AUC) value to evaluate the diagnostic power in differentiating the tumour vs. normal samples.

## 2.6. Clustering Analysis

We used T-distributed Stochastic Neighbour Embedding (t-SNE), as implemented using the library *Rtsne* 0.15, to visualise and identify the differentiation between various phenotypes based on expression data. The t-SNE parameters were adjusted for perplexity (5–50) and numbers of iterations (5000) according to sample sizes, abiding to the ranges suggested by van der Maaten et al. [27]. The principal component analysis (PCA) was implemented in the R (statistical software environment version 4.0.3-The R Foundation for statistical Computing, Vienna, Austria implemented in R studio version 1.4.1717, RStudio, Boston, MA, USA) environment using the *factoextra* package version 1.0.7 for the analysis and identification of gene clusters. The contribution for each principal component (PC) was considered significant above the threshold of 33%.

## 2.7. Survival Analysis

The clinical outcomes were extracted from survival data for the TCGA cohort available for 10,496 samples [22]. The following survival parameters were included: overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI), and progression-free interval (PFI). The survival probability was analysed using Kaplan–Meier plots generated using R (statistical software environment version 4.0.3-The R Foundation for statistical Computing, Vienna, Austria implemented in R studio version 1.4.1717, RStudio, Boston, MA, USA) libraries *survminer* (0.4.9) and *survival* (3.2-11). The expression data were segregated in

“high” and “low” by comparison to the average expression recorded by the mean in each analysed group.

### 3. Results

#### 3.1. *MNX1, MNX1-AS1, and MNX1-AS2 Are Dysregulated in Most Cancers*

We explored the expression profile of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in 18 cancer tissues from the TCGA dataset (Table 1) and unmatched healthy samples from GTEx (Figure 1B).

The majority of cancers analysed showed a statistically significant increase in expression for all three genes in tumour samples. The highest fold changes were recorded for *MNX1* and *MNX1-AS1* in UCEC, PRAD, LUAD, ESCA, and COAD/READ (Figure 1C). By contrast, *MNX1-AS2* showed a modest upregulation in expression for the majority of cancers. Moreover, cancers of the gastrointestinal (GI) tract showed a rather homogenous expression profile across the three genes. LIHC, STAD, TGCT, and THCA were the only tumours that showed different expression profiles across the three genes in terms of both statistical significance as well as the differential expression (up vs. down regulation). *MNX1* expression was unchanged between normal and tumour in LIHC and THCA, however *MNX1-AS1* and *MNX1-AS2* showed a statistically significant upregulation. Notably, all three genes were significantly downregulated in pancreatic adenocarcinoma.

The average expression values in the different cancers revealed a similarity in expression patterns by hierarchical clustering (Figure 1D). READ, COAD, STAD, PAAD, and TGCT showed the most similar expression pattern characterised by a high expression of all three genes, with the highest expression being *MNX1*. A second cluster of tumours showed similar expression levels between *MNX1* and *MNX1-AS1* and included PRAD, LUAD, ESCA, UCEC, BRCA, and LUSC. Low expression values of all three genes clustered SKCM, UCS, GBM, LGG, OV, LIHC, and THCA. Furthermore, we have explored the correlate of the expression profiles of any of the three genes in the 18 studied cancers (Supplementary Figure S1). The results show a cancer/tissue specific relationship in particular for the associations with *MNX1-AS2*.

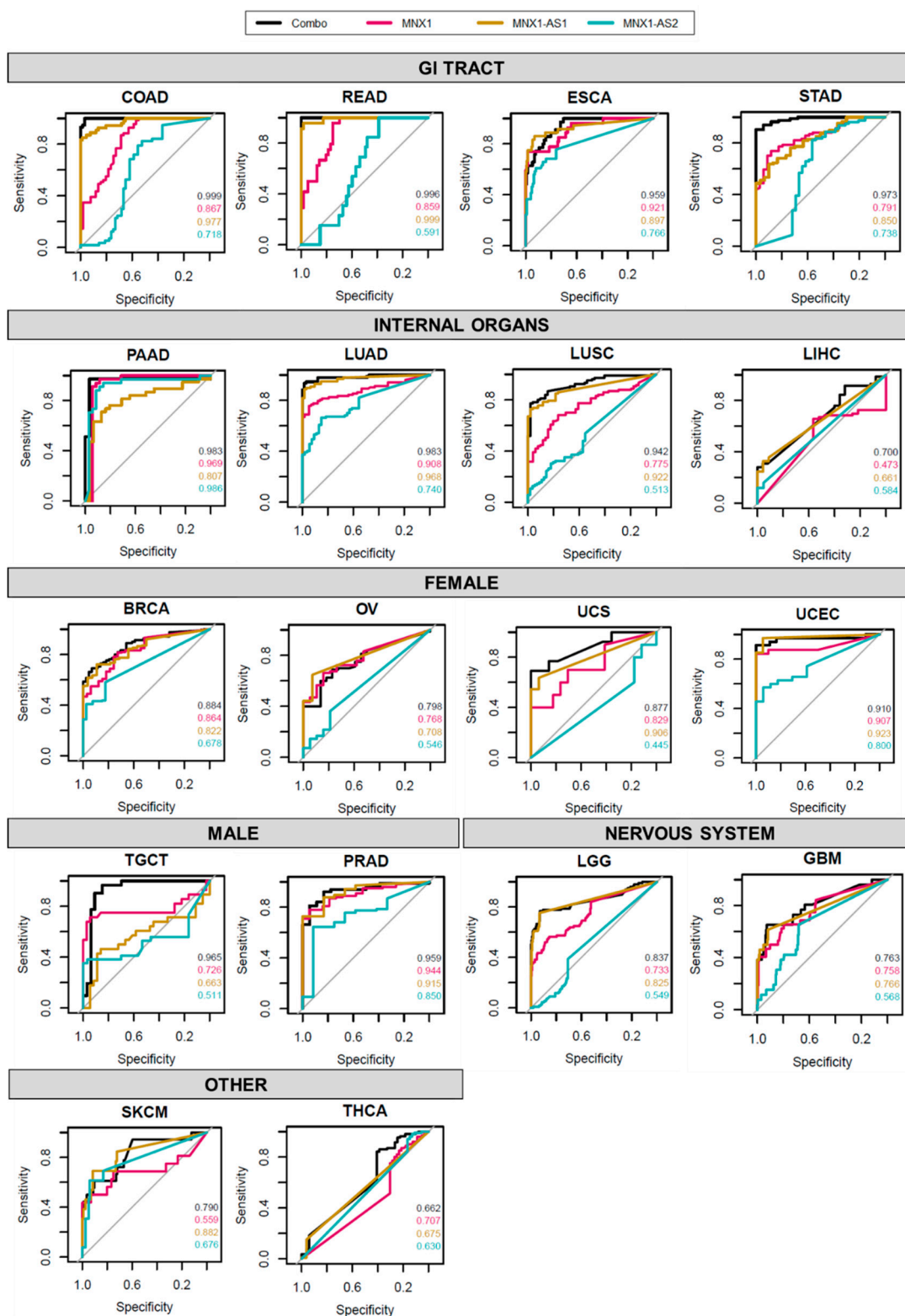
#### 3.2. *Diagnostic Biomarker Potential for MNX1, MNX1-AS1, and MNX1-AS2*

Given the differential expression profile shown by the three genes across a large variety of cancer types, we investigated their potential to be defined as individual or collective diagnostic biomarkers for cancer. We performed a ROC analysis to evaluate the sensitivity and specificity of the three genes as a collective diagnostic biomarker (Figure 2). The combinatorial expression profile of *MNX1* and its associated lncRNAs (“combo”) showed high AUC values for most cancers analysed, indicating a high sensitivity and specificity as a collective diagnostic marker. Cancers of the GI system achieved the best AUC scores, followed closely by LUAD, TGCT, PRAD, and UCEC. By contrast, none of the individual genes, nor their combined expression profiles produced an AUC higher than 0.7 in SKCM, THCA, GBM, and LIHC.

Overall, *MNX1-AS2* showed the lowest performance of AUC values across all cancer sites, with the exception of PAAD, where *MNX1-AS2* was comparable and even exceeded AUC values of *MNX1* and *MNX1-AS1*. High level AUC scores were observed for *MNX1-AS2* in PRAD and UCEC suggesting a moderate diagnostic power for these tumour types.

Interestingly, in several sites (e.g., COAD, READ, LUAD, and LUSC) the lncRNAs showed a superior performance to *MNX1*, suggesting a higher tissue specificity of the antisense transcripts in these tumours. Despite lower individual values, the combination expression patterns displayed the highest performance in the majority of cancers, indicating that the transcriptional patterns of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* are important determinants in distinguishing normal vs. tumour samples.





**Figure 2.** ROC analysis of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions as diagnostic evaluators. Each plot represents the ratio of sensitivity and specificity of the discriminatory power for each gene and their combination. The AUC values for each gene and combination are shown on the bottom right corner, with values closer to 1 indicating a higher performance in discerning normal vs. tumour samples.

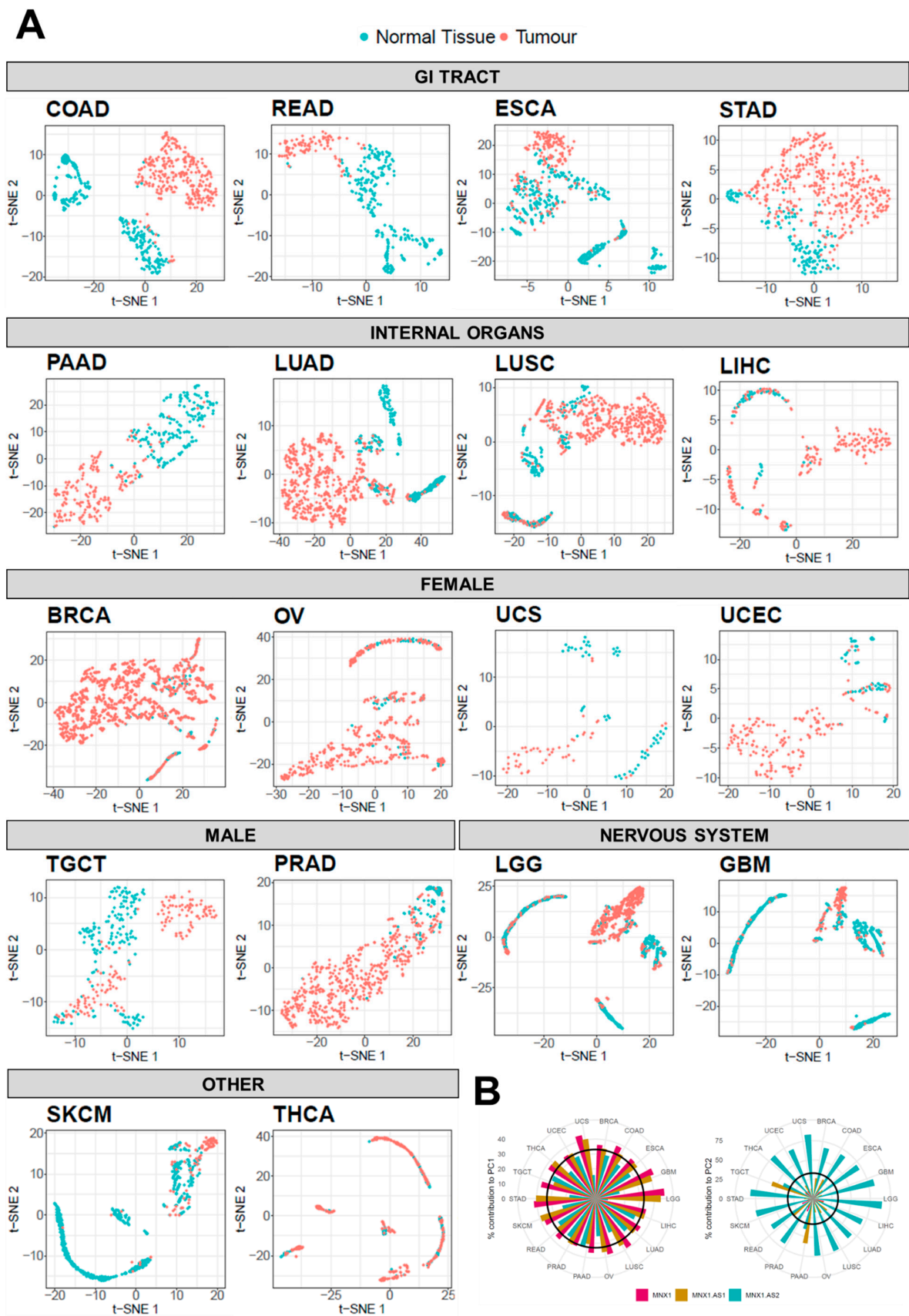


We used t-SNE analysis to assess whether the combinatory expression patterns of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* have discriminatory power in distinguishing tumour vs. normal tissue samples (Figure 3A). Overall, we found that the three genes are able to differentiate between the two states in a tissue specific manner (Figure 3A). Specifically, patients with GI tract tumours were clearly distinguished from health individuals using the combined expression profiles of the *MNX1* and its associated antisense transcripts. A poorer discriminatory power was observed for the other tumour types, namely TGCT, LUAD, LUSC, UCS, and UCEC. By contrast, BRCA, THCA, OV, LGG, GBM, SKCM, and LIHC cancers show no differentiation from the corresponding normal samples based on the transcription profiles of the three genes, reflecting the low expression and fold change at these sites.

We then quantified the observed t-SNE clustering and differentiation of tumour and normal samples using PCA. The contribution percentage of each gene for each principal component (PC) was extracted and is shown in Figure 3B. In the majority of tumours, *MNX1* and *MNX1-AS1* are the main contributors (i.e., accounting for more than 33% of the total contribution) to PC1, while *MNX1-AS2* is the highest contributor for PC2. PAAD and TGCT diverged from this trend, with *MNX1* being the sole significant contributor to PC1 and both *MNX1-AS1* and *MNX1-AS2* contributing to PC2. Consistent with the fold change patterns, PCA contribution reflected the changes in expression and the tissue specificity of the lncRNAs.

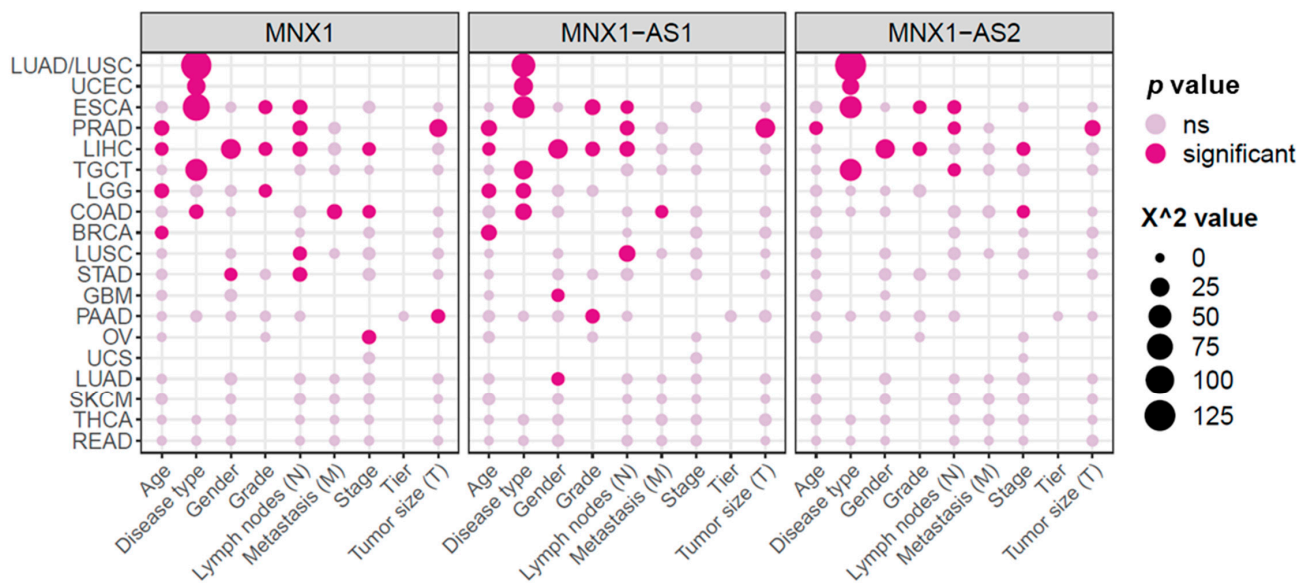
### 3.3. *MNX1*, *MNX1-AS1*, and *MNX1-AS2* Correlate with Oncologic Clinicopathological Features

Next, we looked at the association between clinicopathological features and the expression levels of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* (summarized in Figure 4). Statistically significant associations are reported in detail in Supplementary Table S1. The strongest associations in terms of  $X^2$  and  $p$  values were found for disease subtype in lung tumours (LUAD and LUSC), ESCA, UCEC, and TGCT. High expressions of all three genes were able to discern adenocarcinomas from squamous tumours in LUAD vs. LUSC, as well as in ESCA. A weaker association with adenocarcinomas was also identified in COAD, however for high *MNX1* and *MNX1-AS1* expressions only. In UCEC, high expressions of the three genes were also associated with adenocarcinomas, while lower expressions defined cystic, mucinous, and serous types. High expressions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* also correlated with the TGCT subtype of seminomas. The high expression of *MNX1-AS1* was only associated with oligodendrogliomas in LGG.



**Figure 3.** t-SNE clustering analysis of normal vs. tumour samples based on the combinatorial expressions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2*. (A) The visualization of multi-gene relationships

using t-SNE plots allowed the discrimination of tumour vs. normal samples in several cancer sites. A distinct clustering of normal (in green) and tumour (in red) samples, with no or minimal overlap between samples, defined a strong performance in diagnostic power. (B) PCA contribution of expressions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in clustering differentiation of cancer and normal samples. Percentage of contribution to PC1 (left) and PC2 (right). The black intercept indicates the significance threshold of 33%.



**Figure 4.** Summary of correlations between *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions and clinicopathological features. Each dot represents the value of Chi-square correlation expressed in  $\chi^2$  value, as well as the statistical significance in pink (at a threshold of  $p$  value < 0.05). Detailed sample information on statistically significant features is reported in Supplementary Table S1. ns: not significant.

We also observed a correlation between increased expression levels and clinical features associated with advanced disease progression (i.e., TNM features, staging, and grading), which was particularly evident for GI tract cancers. High expressions of all three genes correlated with a higher histological grade for ESCA and the presence of lymph node invasion in both ESCA, PRAD, and LIHC. In PRAD, high expression of all three genes also correlated with larger tumours. On the contrary, high expression of *MNX1*, but not of the antisenses, was associated with a smaller tumour size in PAAD.

Other clinical features mainly associated with *MNX1* and *MNX1-AS1* expressions. In COAD, *MNX1*, and *MNX1-AS1* high expression correlated with the presence of metastasis. Lymph node involvement was associated with increased expressions of these two genes in LUSC and LIHC. *MNX1* and *MNX1-AS1* also showed an association with age, with higher expression values in patients above the age of 50 years in BRCA, LGG, LIHC, and PRAD. Late disease staging was associated with increased expressions of *MNX1* and *MNX1-AS2* in COAD and a lower expression at early stages of LIHC.

Overall, *MNX1* expression had the highest number of statistically significant associations ( $n = 25$ ), followed by *MNX1-AS1* ( $n = 22$ ) and *MNX1-AS2* ( $n = 14$ ). No clinicopathological features were significantly linked to any of the *MNX1* transcripts in UCS, SKCM, THCA, and READ.

### 3.4. *MNX1*, *MNX1-AS1*, and *MNX1-AS2* Distinguish Tumour Subtypes

Building on the positive results showcased by the discriminatory power of the three genes to distinguish between tumour and healthy samples, we expanded our analysis to evaluate the biomarker potential in differentiating various histological subtypes. The

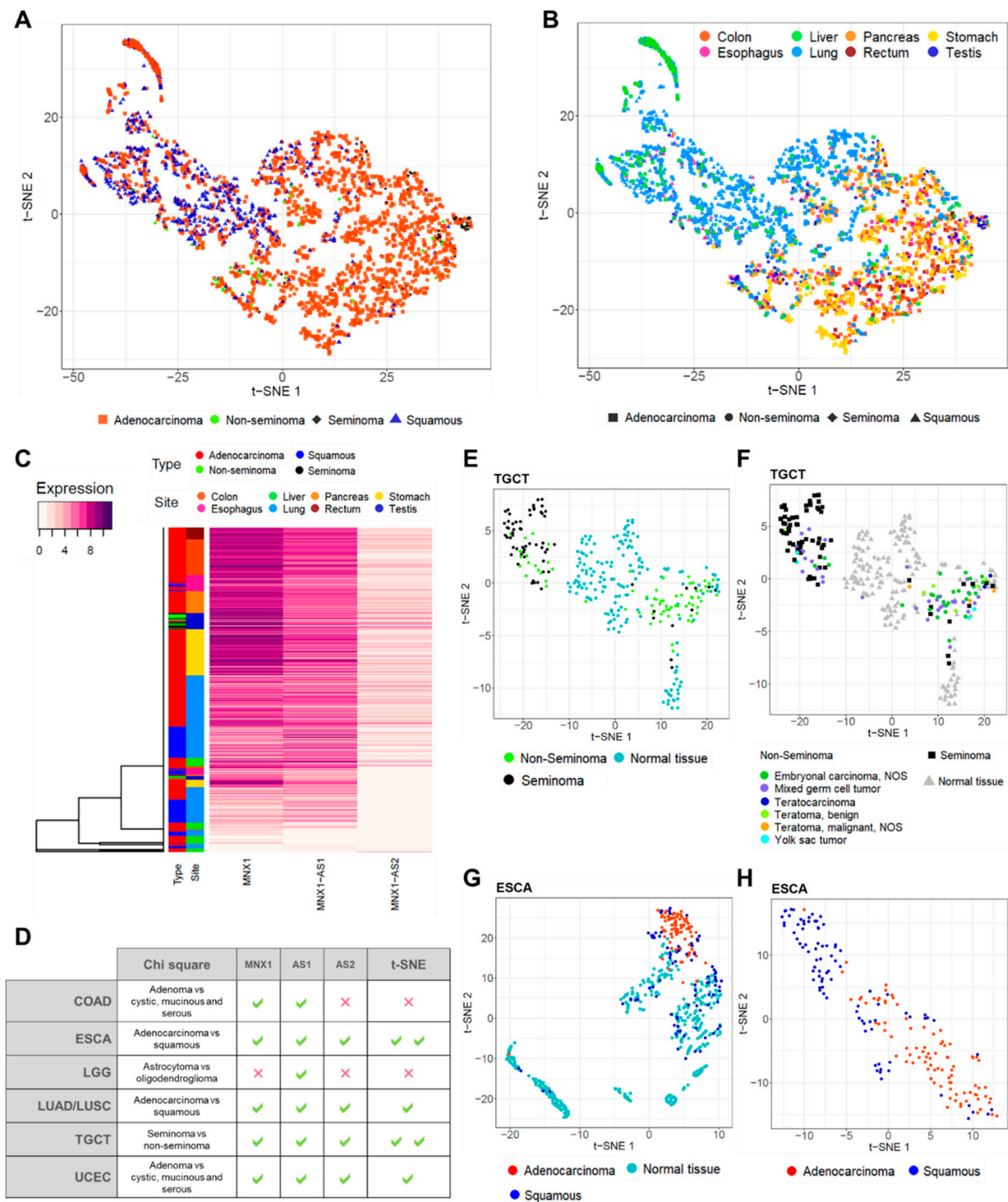
observation of defined tumour clusters in ESCA, STAD, TGCT, LUAD, and LUSC using t-SNE (Figure 3) suggested that specific tumour subtypes may be recognised by the combined expression profiles of the *MNX1*, *MNX1-AS1*, and *MNX1-AS2*. The Chi-square analysis of clinicopathological features (Figure 4) also revealed a defining power of these genes in distinguishing disease subtypes.

First, we combined multiple anatomical sites with adenocarcinomas, squamous tumours, seminomas, and non-seminomas. *MNX1*, *MNX1-AS1*, and *MNX1-AS2* were capable of discerning adenocarcinomas and squamous carcinoma samples by combinatorial analysis visualised using t-SNE. Testicular seminomas and non-seminomas also appeared distinctly clustered (Figure 5A). When colour-coded by anatomical site, adenocarcinomas of the GI tract (COAD, ESCA, PAAD, STAD, and READ) clustered closely together and were discernible from squamous ESCA tumours, LIHC, LUAD, and LUSC. TGCT seminomas formed a defined cluster within the majority of GI adenocarcinomas (Figure 5B).

This separation was also confirmed by hierarchical clustering on individual samples, classified by anatomical site and tumour type (Figure 5C and Supplementary Figure S2). The vast majority of adenocarcinomas of the GI organs were characterised by a high expression of the three genes. A second major cluster grouped squamous ESCA tumours, both LUAD and LUSC, and LIHC. The distinct seminoma TCGT clustering is also evident from the central cluster of the heatmap, which separates seminomas from non-seminomas. These results indicate that *MNX1*, *MNX1-AS1*, and *MNX1-AS2* are expressed at a higher level in adenocarcinomas and seminomas, compared to squamous and non-seminoma tumours, and that their discriminative power is visualisable using clustering methods.

Site-specific clustering using t-SNE revealed specific tumour clusters in TGCT and ESCA (Figure 5E–H), but weakly in COAD, LUAD/LUSC, and UCEC (Supplementary Figure S3). A summary of the results from Chi-square and t-SNE analyses is reported Figure 6D. *MNX1*, *MNX1-AS1*, and *MNX1-AS2* could distinguish seminomas against embryonal carcinomas in TGCT (Figure 5E) and adenocarcinomas against squamous subtypes in ESCA (Figure 5G–H). Detailed analysis of histological subtypes of TGCT revealed that the TGCT of the mixed subtype were equally distributed between the seminoma and non-seminoma clusters (Figure 5E), which is coherent with the biological properties of this subtype, yet clearly distinguishable from the corresponding healthy testicular tissue. ESCA tumours of the squamous type, on the other hand, showed overlap with healthy tissue of the oesophagus (Figure 5G), but a remarkable distinction when analysed in tumour tissues only (Figure 5H).





**Figure 5.** Site-specific t-SNE clustering for the identification of disease subtypes by combinatorial expressions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2*. **(A)** t-SNE visualisation of selected cancer sites by tumour type (adenocarcinoma, squamous, non-seminoma, and seminoma). **(B)** t-SNE clustering by tumour type and colour-coded by cancer site. **(C)** Hierarchical clustering of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expression levels, with cancer type and site information, revealing similar clustering patterns to t-SNE analyses. **(D)** Comparison between Chi-square results of distinction between disease subtypes and the correspondence in t-SNE visualisation of distinct clusters. One green tick indicates a statistically significant correlation in Chi-square tests and the ability to differentiate disease subtypes using t-SNE. Two green ticks indicate a strong t-SNE clustering performance. **(E)** Site-specific t-SNE clustering of TGCT tumours against healthy testicular tissue (GTEx) by TGCT

types of seminomas and non-seminomas. (F) Diagnostic subtype specification of samples within clusters identified using t-SNE in TGCT seminomas and non-seminomas. (G) t-SNE clustering of ESCA tumour subtypes (adenocarcinomas and squamous carcinomas) and healthy oesophageal tissue. (H) t-SNE clustering of ESCA adenocarcinomas and squamous carcinomas by analysis of tumour tissues only.

### 3.5. Prediction of Biological Functions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in Cancer

We attempted to explore biological functions of the two antisense transcripts by looking at the potential targets of *MNX1-AS1* and *MNX1-AS2*. We observed a large degree of overlap between the predicted targets of *MNX1-AS1* and *MNX1-AS2* and a smaller set of antisense specific gene targets (Figure 6A). As limited functional studies are available for these transcripts, we inferred their association with biological processes using Gene Ontology (GO) analysis (Figure 6B–D). Both *MNX1* and antisenses shared several biological processes related to transcription and developmental processes (e.g., anatomical structure morphogenesis, endoderm development, nervous system development), which supports a role in gene expression regulation during development. Notably, targets of *MNX1-AS1* were enriched in processes related to proliferation, cell cycle regulation, apoptosis, and angiogenesis (Figure 6C). *MNX1-AS2* also targets genes associated with the cell cycle and angiogenesis, as well as a unique association with epigenetic regulation of gene expression and chromatin remodelling (Figure 6D).

Based on the high scores observed in the chi-square analysis of clinicopathological features (Figure 4) and subtype prediction using t-SNE (Figure 5), we performed differential expression analysis of COAD, ESCA, PRAD, TGCT, LUSC, and UCEC to identify biologically relevant differences between samples expressing high and low *MNX1*, *MNX1-AS1*, and *MNX1-AS2*. In agreement with some redundancy of predicted functions (Figure 6E), we observed overlap between processes for *MNX1* and the antisenses, but also unique terms associated with specific antisenses in a cancer-specific fashion. The most reoccurring GO terms for all three genes encompassed biological processes relating to cellular transport, metabolism, transcription, and cellular differentiation. The patterns of biological process enrichment indicate varying roles for the two antisenses, which are likely modulators of transcription to accomplish changes in cellular metabolism and identity. In agreement with the main clinical feature associated with high expressions of *MNX1* and antisenses being disease subtype, we observed cell differentiation and lipid metabolism particularly enriched in UCEC and ESCA. We also observed the enrichment of nervous system processes (nervous system development and synaptic vesicle endocytosis), which involve the expression of cell adhesion genes that can modulate invasion behaviour. In TGCT, the highest enriched process was spermatogenesis, which suggests a modulation of testicular gene expression patterns that distinguishes seminomas from non-seminoma tumours.

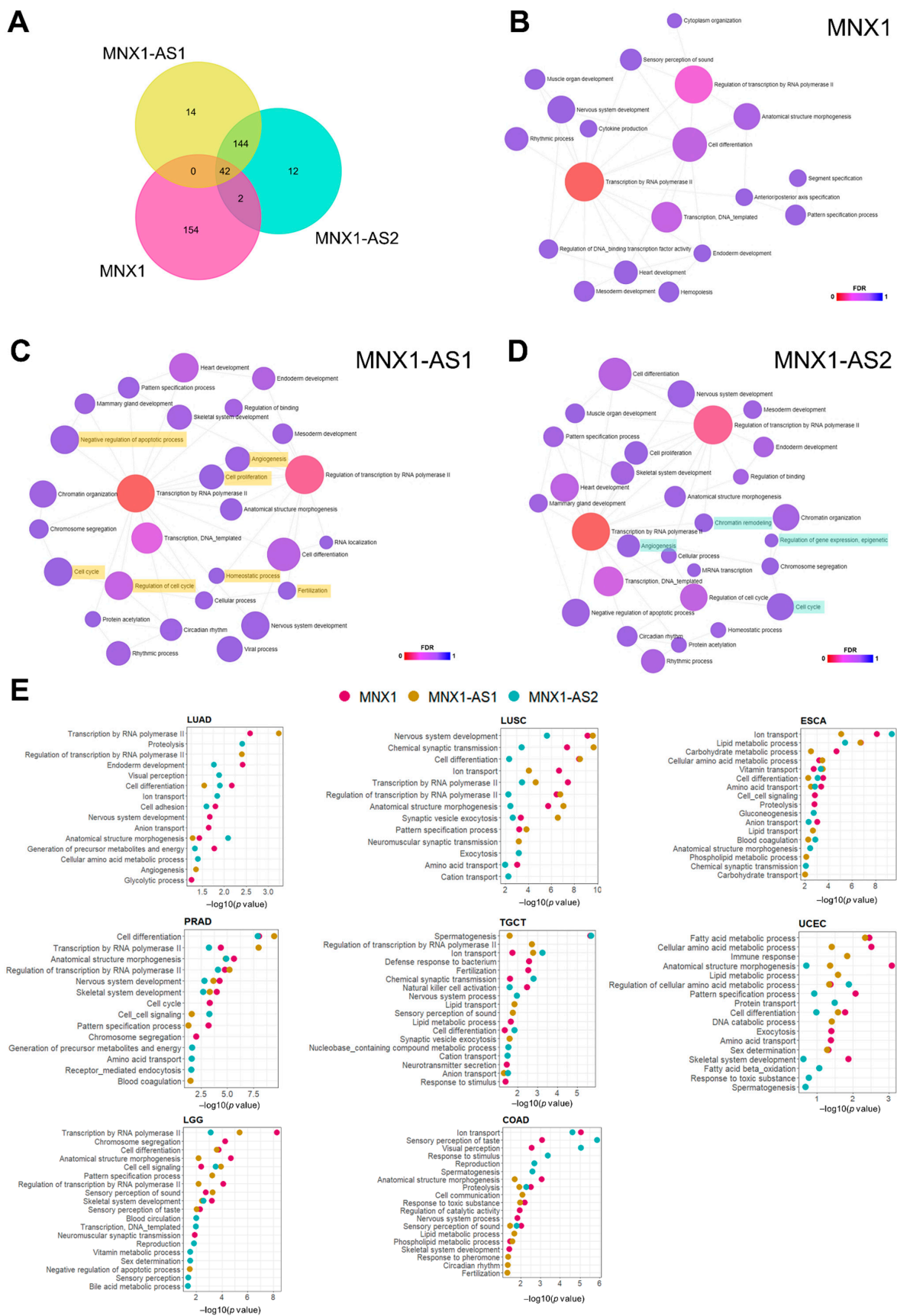


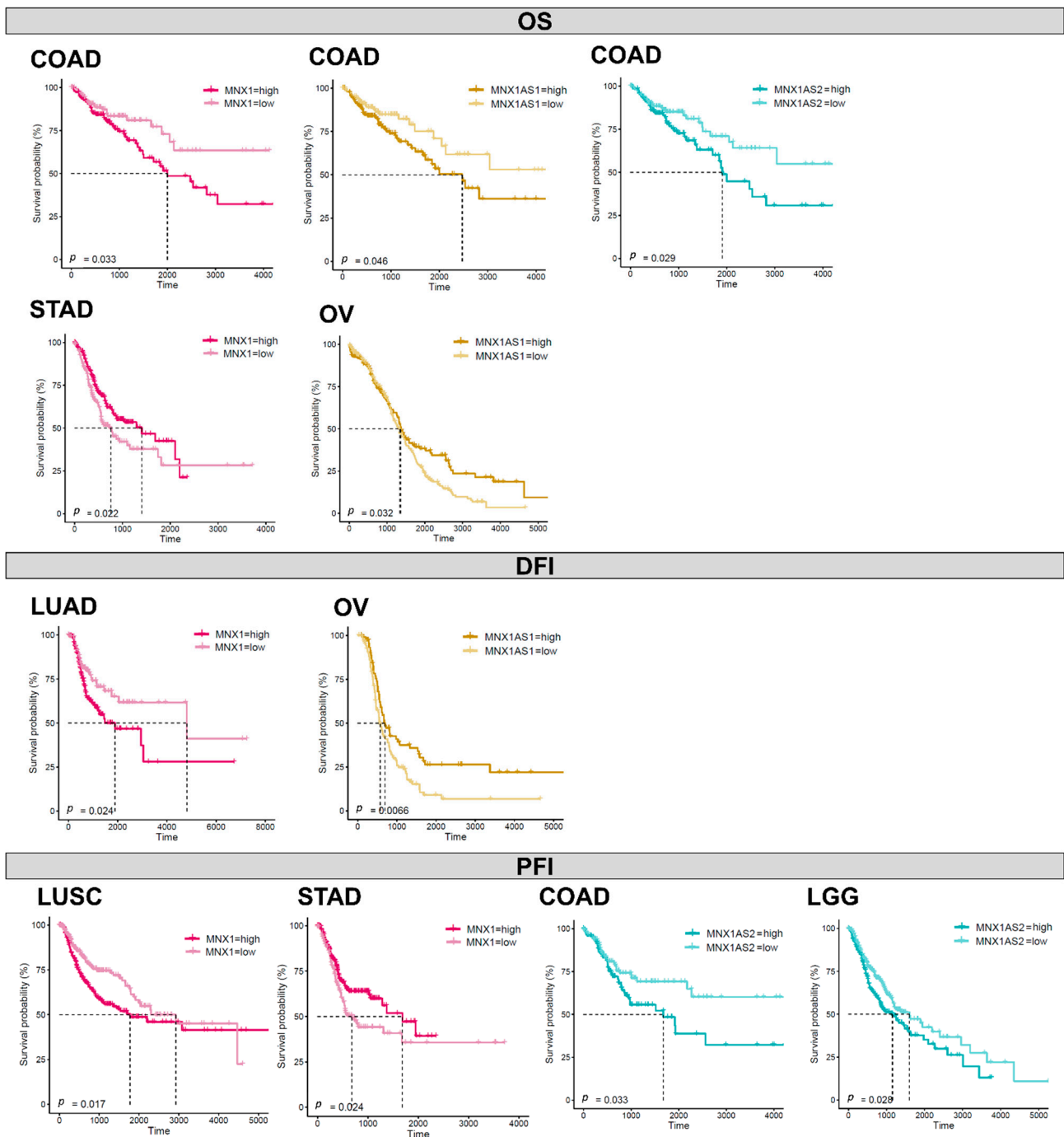
Figure 6. Functional prediction of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* functions in cancer using



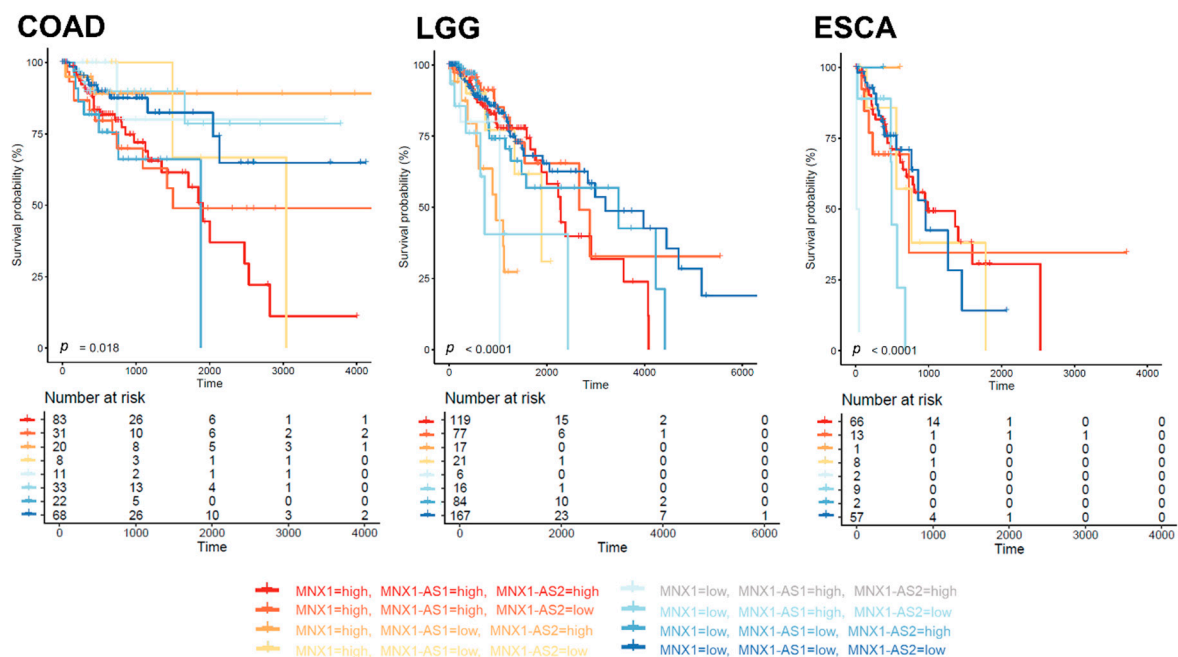
Gene Ontology (GO) analysis. (A) Venn diagram of shared and unique predicted targets of *MNX1*, *MNX1-AS1*, and *MNX1-AS2*. (B–D) Biological processes associated with targets of *MNX1* (B), *MNX1-AS1* (C), and *MNX1-AS2* (D), represented as network of biological processes terms from the PANTHER database and coloured by false discovery rate (FDR; shades of red representing lower values). Highlighted processes are unique to *MNX1-AS1* (yellow) or *MNX1-AS2* (turquoise). (E) Top 10 biological processes (PANTHER) enriched in differentially expressed genes of tumours expressing high levels of *MNX1* (in pink), *MNX1-AS1* (in yellow), and *MNX1-AS2* (turquoise).

### 3.6. Prognostic Biomarker Potential for *MNX1*, *MNX1-AS1*, and *MNX1-AS2*

In order to uncover potential prognostic value of *MNX1*, *MNX1-AS1* and *MNX1-AS2*, we performed Kaplan–Meier analysis based on their expression levels in all cancers. The individual expression levels of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* had a weak correlation with survival (overall—OS, disease free—DFI, and progression free—PFI) (Supplementary Table S2). However, statistical significance for segregating the patient prognostic profiles using gene expression was reached in colon, stomach, ovarian, lung (LUSC, LUAD) and brain (LGG) cancers (Figure 7). Higher expressions of all three genes were associated with inferior OS in COAD. A higher *MNX1-AS2* expression was associated with PFI in COAD ( $p = 0.033$ ) and LGG (0.028). Interestingly, in stomach cancer, a lower expression of *MNX1* associated with inferior clinical outcomes for both OS and PFI. A similar trend was observed for low expressions of *MNX1-AS1* in OV for poor outcomes in terms of OS and DFI. In lung cancers LUAD and LUSC, only a high *MNX1* expression was correlated with inferior DFI and PFI. Overall, these observations indicated that individual *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions harbour modest prognostic potential in selected tumours. Combinatorial analysis of expressions, however, showed stronger statistical differences in OS for COAD ( $p = 0.018$ ), ESCA ( $p < 0.0001$ ), and LGG ( $p < 0.001$ ), suggesting a value in multi-gene expression combinations (Figure 8). In particular, in all three cancers, the combinatorial high expression of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* indicated poor clinical outcomes.



**Figure 7.** Prognostic evaluation of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions. Kaplan–Meier plots of statistically significant associations between individual expressions of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in OS, DFI, and PFI.



**Figure 8.** Combinatorial survival analysis of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions. Kaplan-Meier plots of OS in COAD, LGG, and ESCA with patients stratified according to expression levels.

#### 4. Discussion

Several reports have highlighted *MNX1* and *MNX1-AS1* dysregulation in various malignancies [9–21], however the combined expression patterns of the locus comprising the *MNX1* gene and its associated lncRNAs *MNX1-AS1* and *MNX1-AS2* have not been explored. Here, we leveraged the TCGA dataset to show the clinical potential of the three genes in cancer, in order to address the clinical need for more accurate diagnostic and prognostic biomarkers.

We showed that *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions are dysregulated in most cancers analysed, with particularly high fold changes observed in uterine, prostate, oesophageal, colorectal, and lung cancers. The significant differential expression between normal and tumour tissues suggested that these genes can serve as hallmark diagnostic markers for various cancers. These observations are in line with previous reports describing overexpression of *MNX1* in cancers of the colon [15], breast [11,12], prostate [13], and glioma [17], as well as *MNX1-AS1* in breast [28], colon [29], oesophageal [30], stomach [31], glioblastoma [32], liver [33], lung [34], ovarian [35,36], and prostate [37] tumours. *MNX1* was not found to be significantly overexpressed in liver cancer samples from TCGA compared to healthy liver tissue from GTEx, despite reports of upregulation in hepatocellular carcinoma [18]. Nevertheless, *MNX1* and its antisenses showed statistically significant associations in grading, lymph node invasion, and staging in liver cancer. One of the most interesting results was the overexpression of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* in testicular tumours and their downregulation in pancreatic adenocarcinomas, which have not been previously reported.

In particular, we describe here for the first time the dysregulation of *MNX1-AS2* in cancer. While its fold changes between normal and tumour samples were lower than *MNX1* and *MNX1-AS1*, its expression levels were particularly high in gastro-intestinal tract and testicular tumours, with the highest negative fold change in pancreatic adenocarcinoma. The current literature on *MNX1-AS2* is scarce, with only one report on *MNX1-AS2* dysregulation associated with pre-eclampsia [38]. From our study, we identified individual and combinatorial diagnostic and prognostic potential of *MNX1-AS2* in several cancers. From the PCA analysis, *MNX1-AS2* showed a significant contribution in sample segregation in both pancreatic and testicular tumours. In terms of association with clinical features,

*MNX1-AS2* expression shows a strong discrimination of disease subtype (adenocarcinomas vs. squamous, or seminoma vs. non-seminoma).

Overall, *MNX1*, *MNX1-AS1*, and *MNX1-AS2* individual expressions exhibited limited prognostic potential, compared to their more widespread diagnostic power. Poor clinical outcomes have been previously associated with the overexpression of *MNX1-AS1* in various cancers [21]. In this study, while some association was found with OS, DFI, and PFI in colon, stomach, ovarian, lung, and brain cancers, stronger statistical differences were observed in combinatorial analyses of their expressions in colon and oesophageal cancers and lower grade glioma. The disparity in the distribution of clinical stages in the analysed cohorts may have impacted the evaluation of prognostic significance and especially in tumours where the range of overexpression of the genes is homogeneous (e.g., COAD).

Thanks to the plethora of transcriptomics data becoming available, lncRNAs are increasingly gaining clinical interest for their expression dysregulation and involvement in biological processes related to cancer. Nonetheless, the functions and mechanisms of the vast majority of identified lncRNAs remain to be elucidated [39,40]. The high tissue-specificity in expression renders lncRNAs attractive markers for specific cellular states, with potential correlations with specific disease characteristics [41,42]. The tissue tropism of lncRNA expressions is in fact an advantage in providing biomarker specificity. As revealed by the ROC analysis, the expression of the associated lncRNA *MNX1-AS1* and *MNX1-AS2* performed better than the corresponding coding gene in distinguishing cancerous vs. normal samples in some tissues. In virtue of lncRNA's regulatory function in gene expression modulation, the transcriptional relationship between coding genes and non-coding transcripts may also reveal biologically relevant clues on the transcriptional activity of specific loci [43,44]. In fact, in this study, combinatorial expression patterns displayed a high degree of accuracy in distinguishing tumour samples against normal tissues and tumour subtypes. Nevertheless, the combinatorial predictive power may only be clinically applicable in specific tumours rather than a collective pan-cancer biomarker.

While the functions of *MNX1-AS1* and *MNX1-AS2* remain elusive, we identified a high degree of overlap in predicted targeted genes and biological functions associated with the two antisense. By comparing the DEGs in high vs. low expressing patient samples, we observed different target genes but similar biological processes across tumours, mainly associated with cellular differentiation and metabolism. We can speculate that these genes function as selective transcriptional regulators with overlapping targets, with tissue/cancer-specific action. From the observations of the transcriptional elements regulating the expression of these genes (Figure 1A), distinct promoters and enhancer regions allow the selective expression of *MNX1* or its transcripts to modulate the regulation of their targets. In cancer, the dysregulation of this physiological transcriptional balance results in the activation or repression of key cellular processes.

The implementation of dimensionality reduction techniques (e.g. clustering methods such as t-SNE [27]) allowed us to uncover gene relationships with ease of visualisation and interpretation. Specifically, we were able to capture tissue-specific transcriptional patterns of the *MNX1* locus that are altered in the cancer state, with a higher specificity than measuring the expression levels of the individual genes. Interestingly, the *MNX1*, *MNX1-AS1*, and *MNX1-AS2* combined expression patterns defined specific tumour subtypes, displaying both tissue- and cell-type specificity of adenocarcinomas and seminomas. Transcriptional and molecular differences between adenocarcinomas and squamous cancers have been identified, with a high degree of similarity even across different anatomical sites [45]. While the role of *MNX1*-associated lncRNAs has not been investigated during embryonic development, *MNX1*'s role in instructing pancreatic beta cell identity (endocrine portion of the pancreas) may explain their reminiscent expression in the healthy tissues of the gastro-intestinal tract and the specificity in expression in glandular tumours. In addition to providing diagnostic specificity for these cancers, the dysregulation of the *MNX1* transcripts may also be interesting in the light of tumour genesis and development.

While the association with GI organs may derive from the embryonic functions of *MNX1*, the expression of *MNX1* and its antisenses in normal testis and testicular cancer does not seem to link with developmental activity. *MNX1* has not been described in genital development but has been reported to be expressed in the developing urinary tract [46]. It must be noted that testes are known to be a transcriptionally active tissue, with a large proportion of genes being expressed [47,48]. Nevertheless, *MNX1*, *MNX1-AS1*, and *MNX1-AS2* transcription profiles exhibited high specificity in differentiating seminomas from non-seminomas samples. The expression patterns of *MNX1* and its lncRNA's in these subtypes may be linked to the fact that the cell of origin of seminomas and non-seminomas appears to be distinct, with unique transcriptional identities [49]. Moreover, the distinction between the two subtypes also has therapeutic implications due to the radiosensitivity of seminomas compared to non-seminomas [50]. Thus, the use of the three genes as collective biomarkers with subtype discriminatory power can strongly impact the patient's course of treatment. Thus, it would be an interesting avenue to pursue in the future examination of the transcriptional activity based on the cellular subtype in the tumour environment.

The use of a next-generation sequencing (NGS) platform in clinical practice will enable these observations to be translated into diagnostic procedures [51]. Overall, *MNX1*, *MNX1-AS1*, and *MNX1-AS2* show a robust performance in distinguishing normal samples from tumours, with a good site-specific diagnostic power. The use of a simple three gene signature has the potential to improve histological classification workflows and showcase lncRNAs as diagnostic markers [52–54]. Currently, the use of NGS data has already shown promising results in supporting histopathological information [51,55,56] and the clinical potential of *MNX1* and its antisenses.

## 5. Conclusions

The availability of transcriptomics data and the use of an NGS platform in clinical applications will enable the use of both coding and non-coding transcripts as clinical individual and collective biomarkers. In this study, we identified the expression patterns of the *MNX1* gene and its antisense transcripts *MNX1-AS1* and *MNX1-AS2* as potential diagnostic and prognostic biomarkers in cancer. Based on the analysis of the TCGA cohort, these genes are dysregulated in the majority of cancers. By the use of clinicopathological features, clustering, and ROC analysis, we identified a strong diagnostic power for both individual and combinatorial expressions of these three genes in distinguishing tumour vs. healthy samples, as well as differentiating histological tumour subtypes. In particular, *MNX1*, *MNX1-AS1*, and *MNX1-AS2* could discern adenocarcinomas from squamous tumours at various sites, as well as testicular seminomas from non-seminomas. As the use of next generation sequencing platforms widens in the clinic, we foresee that robustly differentially expressed genes with tissue specificity, such as *MNX1* and its antisense transcripts, will be ideal diagnostic markers to be implemented in current clinical workflows.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells11223577/s1>, Figure S1: Hierarchical clustering on Pearson correlation values of expression between *MNX1-MNX1-AS1*, *MNX1-MNX1-AS2*, and *MNX1-AS1-MNX1-AS2* in each cancer type analysed; Figure S2: Complete hierarchical clustering of *MNX1*, *MNX1-AS1* and *MNX1-AS2* expression levels, with cancer type and site information; Figure S3: Site-specific t-SNE plots for discriminating disease subtypes based on *MNX1*, *MNX1-AS1*, and *MNX1-AS2* expressions; Table S1: Statistically significant correlations between expression levels of *MNX1*, *MNX1-AS1*, and *MNX1-AS2* and clinicopathological features; Table S2: P values of survival analysis by Kaplan-Meier analysis based on high and low individual expressions of *MNX1*, *MNX1-AS1* and *MNX1-AS2*, as well as combinatorial (“COMB”); Table S3: Evaluation of the sample data bias impact on the biological insights.

**Author Contributions:** Conceptualization, C.S., D.R. and S.T.; methodology, C.S. and D.R.; validation, C.S. and D.R.; formal analysis, D.R.; investigation, D.R.; writing—original draft preparation, D.R. and C.S.; writing—review and editing, D.R., C.S. and S.T.; visualization, D.R.; supervision, C.S. and S.T.; funding acquisition, S.T. All authors have read and agreed to the published version of the manuscript.



**Funding:** D.R. is the recipient of a Kidscan funded PhD studentship and partly supported by Brunel University London.

**Data Availability Statement:** Clinical phenotype and expression data is publicly available and freely accessible in the TCGA-TARGET-GTEX cohort from the University of California Santa Cruz public repository Xena accessible online at [www.xenabrowser.net](http://www.xenabrowser.net) (last accessed on 20 October 2022) [22].

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

## References

1. Ahmad, A.S.; Ormiston-Smith, N.; Sasieni, P.D. Trends in the Lifetime Risk of Developing Cancer in Great Britain: Comparison of Risk for those Born from 1930 to 1960. *Br. J. Cancer* **2015**, *112*, 943–947. [[CrossRef](#)] [[PubMed](#)]
2. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer Statistics, 2021. *CA Cancer J. Clin.* **2021**, *71*, 7–33. [[CrossRef](#)] [[PubMed](#)]
3. Chang, K.; Creighton, C.J.; Davis, C.; Donehower, L.; Drummond, J.; Wheeler, D.; Ally, A.; Balasundaram, M.; Birol, I.; Butterfield, Y.S. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat. Genet.* **2013**, *45*, 1113–1120.
4. Lonsdale, J.; Thomas, J.; Salvatore, M.; Phillips, R.; Lo, E.; Shad, S.; Hasz, R.; Walters, G.; Garcia, F.; Young, N. The Genotype-Tissue Expression (GTEx) Project. *Nat. Genet.* **2013**, *45*, 580–585. [[CrossRef](#)] [[PubMed](#)]
5. Arber, S.; Han, B.; Mendelsohn, M.; Smith, M.; Jessell, T.M.; Sockanathan, S. Requirement for the Homeobox Gene Hb9 in the Consolidation of Motor Neuron Identity. *Neuron* **1999**, *23*, 659–674. [[CrossRef](#)]
6. Li, H.; Arber, S.; Jessell, T.M.; Edlund, H. Selective Agenesis of the Dorsal Pancreas in Mice Lacking Homeobox Gene Hlxb9. *Nat. Genet.* **1999**, *23*, 67–70. [[CrossRef](#)]
7. Thaler, J.; Harrison, K.; Sharma, K.; Lettieri, K.; Kehrl, J.; Pfaff, S.L. Active Suppression of Interneuron Programs within Developing Motor Neurons Revealed by Analysis of Homeodomain Factor HB9. *Neuron* **1999**, *23*, 675–687. [[CrossRef](#)]
8. Harrison, K.A.; Thaler, J.; Pfaff, S.L.; Gu, H.; Kehrl, J.H. Pancreas Dorsal Lobe Agenesis and Abnormal Islets of Langerhans in Hlxb9-Deficient Mice. *Nat. Genet.* **1999**, *23*, 71–75. [[CrossRef](#)]
9. Tosi, S.; Kamel, Y.M.; Owoka, T.; Federico, C.; Truong, T.H.; Saccone, S. Paediatric Acute Myeloid Leukaemia with the T (7; 12)(q36; p13) Rearrangement: A Review of the Biological and Clinical Management Aspects. *Biomark. Res.* **2015**, *3*, 21. [[CrossRef](#)]
10. Nagel, S.; Scherr, M.; Quentmeier, H.; Kaufmann, M.; Zaborski, M.; Drexler, H.G.; MacLeod, R. HLXB9 Activates IL6 in Hodgkin Lymphoma Cell Lines and is Regulated by PI3K Signalling Involving E2F3. *Leukemia* **2005**, *19*, 841–846. [[CrossRef](#)]
11. Tian, T.; Wang, M.; Zhu, Y.; Zhu, W.; Yang, T.; Li, H.; Lin, S.; Dai, C.; Deng, Y.; Song, D. Expression, Clinical Significance, and Functional Prediction of MNX1 in Breast Cancer. *Mol. Ther.–Nucleic Acids* **2018**, *13*, 399–406. [[CrossRef](#)] [[PubMed](#)]
12. Neufing, P.J.; Kalionis, B.; Horsfall, D.J.; Ricciardelli, C.; Stahl, J.; Vivekanandan, S.; Raymond, W.; Tilley, W.D. Expression and Localization of Homeodomain Proteins DLX4/HB9 in Normal and Malignant Human Breast Tissues. *Anticancer. Res.* **2003**, *23*, 1479–1488. [[PubMed](#)]
13. Zhang, L.; Wang, J.; Wang, Y.; Zhang, Y.; Castro, P.; Shao, L.; Sreekumar, A.; Putluri, N.; Guha, N.; Deepak, S. MNX1 is Oncogenically Upregulated in African-American Prostate Cancer. *Cancer Res.* **2016**, *76*, 6290–6298. [[CrossRef](#)]
14. Chen, M.; Wu, R.; Li, G.; Liu, C.; Tan, L.; Xiao, K.; Ye, Y.; Qin, Z. Motor Neuron and Pancreas Homeobox 1/HLXB9 Promotes Sustained Proliferation in Bladder Cancer by Upregulating CCNE1/2. *J. Exp. Clin. Cancer Res.* **2018**, *37*, 154. [[CrossRef](#)]
15. Yang, X.; Pan, Q.; Lu, Y.; Jiang, X.; Zhang, S.; Wu, J. MNX1 Promotes Cell Proliferation and Activates Wnt/B-catenin Signaling in Colorectal Cancer. *Cell Biol. Int.* **2019**, *43*, 402–408. [[CrossRef](#)] [[PubMed](#)]
16. Leotta, C.G.; Federico, C.; Brundo, M.V.; Tosi, S.; Saccone, S. HLXB9 Gene Expression, and Nuclear Location during in Vitro Neuronal Differentiation in the SK-N-BE Neuroblastoma Cell Line. *PLoS ONE* **2014**, *9*, e105481. [[CrossRef](#)] [[PubMed](#)]
17. Jiang, L.; Chen, S.; Zhao, D.; Yan, J.; Chen, J.; Yang, C.; Zheng, G. MNX1 Reduces Sensitivity to Anoikis by Activating TrkB in Human Glioma Cells. *Mol. Med. Rep.* **2018**, *18*, 3271–3279. [[CrossRef](#)]
18. Wilkens, L.; Jaggi, R.; Hammer, C.; Inderbitzin, D.; Giger, O.; von Neuhoff, N. The Homeobox Gene HLXB9 is Upregulated in a Morphological Subset of Poorly Differentiated Hepatocellular Carcinoma. *Virchows Arch.* **2011**, *458*, 697–708. [[CrossRef](#)]
19. Desai, S.S.; Modali, S.D.; Parekh, V.I.; Kebebew, E.; Agarwal, S.K. GSK-3 $\beta$  Protein Phosphorylates and Stabilizes HLXB9 Protein in Insulinoma Cells to Form a Targetable Mechanism of Controlling Insulinoma Cell Proliferation. *J. Biol. Chem.* **2014**, *289*, 5386–5398. [[CrossRef](#)]
20. Shen, S.; Zhou, H. Clinical Effects and Molecular Mechanisms of lncRNA MNX1-AS1 in Malignant Tumors. *Am. J. Transl. Res.* **2020**, *12*, 7593.
21. Li, J.; Jin, W.; Zhang, Z.; Chu, J.; Yang, H.; Li, C.; Dong, R.; Zhao, C. The prognostic utility and clinical outcomes of MNX1-AS1 expression in cancers: A systematic review and meta-analysis. *Res. Sq.* **2020**. [[CrossRef](#)]
22. Goldman, M.J.; Craft, B.; Hastie, M.; Repečka, K.; McDade, F.; Kamath, A.; Banerjee, A.; Luo, Y.; Rogers, D.; Brooks, A.N. Visualizing and Interpreting Cancer Genomics Data Via the Xena Platform. *Nat. Biotechnol.* **2020**, *38*, 675–678. [[CrossRef](#)] [[PubMed](#)]
23. Vivian, J.; Rao, A.A.; Nothhaft, F.A.; Ketchum, C.; Armstrong, J.; Novak, A.; Pfeil, J.; Narkizian, J.; Deran, A.D.; Musselman-Brown, A. Toil Enables Reproducible, Open Source, Big Biomedical Data Analyses. *Nat. Biotechnol.* **2017**, *35*, 314–316. [[CrossRef](#)] [[PubMed](#)]

24. Kang, J.; Tang, Q.; He, J.; Li, L.; Yang, N.; Yu, S.; Wang, M.; Zhang, Y.; Lin, J.; Cui, T. RNAInter v4.0: RNA Interactome Repository with Redefined Confidence Scoring System and Improved Accessibility. *Nucleic Acids Res.* **2022**, *50*, D326–D332.
25. Fornes, O.; Castro-Mondragon, J.A.; Khan, A.; Van der Lee, R.; Zhang, X.; Richmond, P.A.; Modi, B.P.; Correard, S.; Gheorghe, M.; Baranašić, D. JASPAR 2020: Update of the Open-Access Database of Transcription Factor Binding Profiles. *Nucleic Acids Res.* **2020**, *48*, D87–D92. [[CrossRef](#)]
26. Lin, G.; Chai, J.; Yuan, S.; Mai, C.; Cai, L.; Murphy, R.W.; Zhou, W.; Luo, J. VennPainter: A Tool for the Comparison and Identification of Candidate Genes Based on Venn Diagrams. *PLoS ONE* **2016**, *11*, e0154315. [[CrossRef](#)] [[PubMed](#)]
27. Van der Maaten, L.; Hinton, G. Visualizing Data using T-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
28. Li, J.; Li, Q.; Li, D.; Shen, Z.; Zhang, K.; Bi, Z.; Li, Y. Long Non-Coding RNA MNX1-AS1 Promotes Progression of Triple Negative Breast Cancer by Enhancing Phosphorylation of Stat3. *Front. Oncol.* **2020**, *10*, 1108. [[CrossRef](#)]
29. Wu, Q.; Luo, X.; Liu, J.; Lu, Y.; Wang, Y.; Qi, J.; Liu, Z.; Huang, Q.; Liu, Z.; Lu, J. MYC-Activated lncRNA MNX1-AS1 Promotes the Progression of Colorectal Cancer by Stabilizing YB1. *Cancer Res.* **2021**, *81*, 2636–2650. [[CrossRef](#)]
30. Chu, J.; Li, H.; Xing, Y.; Jia, J.; Sheng, J.; Yang, L.; Sun, K.; Qu, Y.; Zhang, Y.; Yin, H. LncRNA MNX1-AS1 Promotes Progression of Esophageal Squamous Cell Carcinoma by Regulating miR-34a/SIRT1 Axis. *Biomed. Pharmacother.* **2019**, *116*, 109029. [[CrossRef](#)]
31. Zhang, W.; Huang, L.; Lu, X.; Wang, K.; Ning, X.; Liu, Z. Upregulated Expression of MNX1-AS1 Long Noncoding RNA Predicts Poor Prognosis in Gastric Cancer. *Bosn. J. Basic Med. Sci.* **2019**, *19*, 164. [[CrossRef](#)] [[PubMed](#)]
32. Gao, Y.; Xu, Y.; Wang, J.; Yang, X.; Wen, L.; Feng, J. lncRNA MNX1-AS1 Promotes Glioblastoma Progression through Inhibition of miR-4443. *Oncol. Res.* **2019**, *27*, 341. [[CrossRef](#)] [[PubMed](#)]
33. Ji, D.; Wang, Y.; Sun, B.; Yang, J.; Luo, X. Long Non-Coding RNA MNX1-AS1 Promotes Hepatocellular Carcinoma Proliferation and Invasion through Targeting miR-218-5p/COMMD8 Axis. *Biochem. Biophys. Res. Commun.* **2019**, *513*, 669–674. [[CrossRef](#)] [[PubMed](#)]
34. Liu, H.; Han, L.; Liu, Z.; Gao, N. Long Noncoding RNA MNX1-AS1 Contributes to Lung Cancer Progression through the miR-527/BRF2 Pathway. *J. Cell. Physiol.* **2019**, *234*, 13843–13850. [[CrossRef](#)]
35. Shen, Y.; Lv, M.; Fang, Y.; Lu, J.; Wu, Y. LncRNA MNX1-AS1 Promotes Ovarian Cancer Process Via Targeting the miR-744-5p/SOX12 Axis. *J. Ovarian Res.* **2021**, *14*, 1–14. [[CrossRef](#)] [[PubMed](#)]
36. Lv, Y.; Li, H.; Li, F.; Liu, P.; Zhao, X. Long Noncoding RNA MNX1-AS1 Knockdown Inhibits Cell Proliferation and Migration in Ovarian Cancer. *Cancer Biother. Radiopharm.* **2017**, *32*, 91–99. [[CrossRef](#)]
37. Li, Z.; Wang, F.; Zhang, S. Knockdown of Lnc RNA MNX1-AS1 Suppresses Cell Proliferation, Migration, and Invasion in Prostate Cancer. *FEBS Open Bio.* **2019**, *9*, 851–858. [[CrossRef](#)]
38. Long, W.; Rui, C.; Song, X.; Dai, X.; Xue, X.; Lu, Y.; Shen, R.; Li, J.; Li, J.; Ding, H. Distinct Expression Profiles of lncRNAs between Early-Onset Preeclampsia and Preterm Controls. *Clin. Chim. Acta.* **2016**, *463*, 193–199. [[CrossRef](#)]
39. Cao, H.; Wahlestedt, C.; Kapranov, P. Strategies to Annotate and Characterize Long Noncoding RNAs: Advantages and Pitfalls. *Trends Genet.* **2018**, *34*, 704–721. [[CrossRef](#)]
40. Carlevaro-Fita, J.; Lanzós, A.; Feuerbach, L.; Hong, C.; Mas-Ponte, D.; Pedersen, J.S.; Johnson, R. Cancer LncRNA Census Reveals Evidence for Deep Functional Conservation of Long Noncoding RNAs in Tumorigenesis. *Commun. Biol.* **2020**, *3*, 1–16. [[CrossRef](#)]
41. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G. The GENCODE V7 Catalog of Human Long Noncoding RNAs: Analysis of their Gene Structure, Evolution, and Expression. *Genome Res.* **2012**, *22*, 1775–1789. [[CrossRef](#)] [[PubMed](#)]
42. Mattioli, K.; Volders, P.; Gerhardinger, C.; Lee, J.C.; Maass, P.G.; Melé, M.; Rinn, J.L. High-Throughput Functional Analysis of lncRNA Core Promoters Elucidates Rules Governing Tissue Specificity. *Genome Res.* **2019**, *29*, 344–355. [[CrossRef](#)] [[PubMed](#)]
43. Ning, Q.; Li, Y.; Wang, Z.; Zhou, S.; Sun, H.; Yu, G. The Evolution and Expression Pattern of Human Overlapping lncRNA and Protein-Coding Gene Pairs. *Sci. Rep.* **2017**, *7*, 1–12. [[CrossRef](#)]
44. Song, X.; Cao, G.; Jing, L.; Lin, S.; Wang, X.; Zhang, J.; Wang, M.; Liu, W.; Lv, C. Analysing the Relationship between Lnc RNA and Protein-coding Gene and the Role of Lnc RNA as Ce RNA in Pulmonary Fibrosis. *J. Cell. Mol. Med.* **2014**, *18*, 991–1003. [[CrossRef](#)] [[PubMed](#)]
45. Lin, E.W.; Karakasheva, T.A.; Lee, D.; Lee, J.; Long, Q.; Bass, A.J.; Wong, K.K.; Rustgi, A.K. Comparative Transcriptomes of Adenocarcinomas and Squamous Cell Carcinomas Reveal Molecular Similarities that Span Classical Anatomic Boundaries. *PLoS Genet.* **2017**, *13*, e1006938. [[CrossRef](#)] [[PubMed](#)]
46. Ott, E.; Wendik, B.; Srivastava, M.; Pacho, F.; Töchterle, S.; Salvenmoser, W.; Meyer, D. Pronephric Tubule Morphogenesis in Zebrafish Depends on Mnx Mediated Repression of Irx1b within the Intermediate Mesoderm. *Dev. Biol.* **2016**, *411*, 101–114. [[CrossRef](#)]
47. Schmidt, E.E. Transcriptional Promiscuity in Testes. *Curr. Biol.* **1996**, *6*, 768–769. [[CrossRef](#)]
48. Xia, B.; Yan, Y.; Baron, M.; Wagner, F.; Barkley, D.; Chiodin, M.; Kim, S.Y.; Keefe, D.L.; Alukal, J.P.; Boeke, J.D. Widespread Transcriptional Scanning in the Testis Modulates Gene Evolution Rates. *Cell* **2020**, *180*, 248–262. [[CrossRef](#)]
49. Som, A.; Wen, S.; Tu, S. Stem Cell Origin of Testicular Seminoma. *Clin. Genitourin. Cancer* **2013**, *11*, 489–494. [[CrossRef](#)]
50. Liu, R.; Lei, Z.; Li, A.; Jiang, Y.; Ji, J. Differentiation of Testicular Seminoma and Nonseminomatous Germ Cell Tumor on Magnetic Resonance Imaging. *Medicine* **2019**, *98*, e17937. [[CrossRef](#)]
51. Conway, J.R.; Warner, J.L.; Rubinstein, W.S.; Miller, R.S. Next-Generation Sequencing and the Clinical Oncology Workflow: Data Challenges, Proposed Solutions, and a Call to Action. *JCO Precis. Oncol.* **2019**, *3*, 1–10. [[CrossRef](#)] [[PubMed](#)]



52. Hu, J.; Xu, L.; Shou, T.; Chen, Q. Systematic Analysis Identifies Three-lncRNA Signature as a Potentially Prognostic Biomarker for Lung Squamous Cell Carcinoma using Bioinformatics Strategy. *Transl. Lung Cancer Res.* **2019**, *8*, 614. [[CrossRef](#)] [[PubMed](#)]
53. He, J.; Wu, F.; Han, Z.; Hu, M.; Lin, W.; Li, Y.; Cao, M. Biomarkers (mRNAs and Non-Coding RNAs) for the Diagnosis and Prognosis of Colorectal Cancer—From the Body Fluid to Tissue Level. *Front. Oncol.* **2021**, *11*, 1478. [[CrossRef](#)]
54. Wang, L.; Yu, Z.; Sun, S.; Peng, J.; Xiao, R.; Chen, S.; Zuo, X.; Cheng, Q.; Xia, Y. Long Non-Coding RNAs: Potential Molecular Biomarkers for Gliomas Diagnosis and Prognosis. *Rev. Neurosci.* **2017**, *28*, 375–380. [[CrossRef](#)]
55. Yu, K.; Berry, G.J.; Rubin, D.L.; Re, C.; Altman, R.B.; Snyder, M. Association of Omics Features with Histopathology Patterns in Lung Adenocarcinoma. *Cell Syst.* **2017**, *5*, 620–627. [[CrossRef](#)] [[PubMed](#)]
56. Carrara, S.; Soldà, G.; Di Leo, M.; Rahal, D.; Peano, C.; Giunta, M.; Lamonaca, L.; Auriemma, F.; Anderloni, A.; Fugazza, A. Side-by-Side Comparison of Next-Generation Sequencing, Cytology, and Histology in Diagnosing Locally Advanced Pancreatic Adenocarcinoma. *Gastrointest Endosc.* **2021**, *93*, 597–604. [[CrossRef](#)]