



Article

# The Long Non-Coding RNA MALAT1 Modulates NR4A1 Expression through a Downstream Regulatory Element in Specific Cancer Cell Types

Sara Wernig-Zorc <sup>1,2</sup> , Uwe Schwartz <sup>3</sup>, Paulina Martínez-Rodríguez <sup>4</sup> , Josefa Inalef <sup>5</sup>, Francisca Pavicic <sup>5</sup>, Pamela Ehrenfeld <sup>5,6</sup> , Gernot Längst <sup>1</sup> and Rodrigo Maldonado <sup>7,\*</sup>

<sup>1</sup> Regensburg Center for Biochemistry [RCB], Universität Regensburg, 93053 Regensburg, Germany; sara.wernig-zorc@ccri.at (S.W.-Z.)

<sup>2</sup> St. Anna Children's Cancer Research Institute, 1090 Vienna, Austria

<sup>3</sup> NGS Analysis Center, Biology and Pre-Clinical Medicine, Universität Regensburg, 93053 Regensburg, Germany

<sup>4</sup> Programa de Doctorado en Ciencias, mención Biología Celular y Molecular Aplicada, Universidad de La Frontera, Temuco 4811230, Chile; paulinaconstanza.martinez@ufrontera.cl

<sup>5</sup> Institute of Anatomy, Histology, and Pathology, Faculty of Medicine, Universidad Austral de Chile, 5090000 Valdivia, Chile; ingridehrenfeld@uach.cl (P.E.)

<sup>6</sup> Center for Interdisciplinary Studies of the Nervous System [CISNe], Universidad Austral de Chile, 5090000 Valdivia, Chile

<sup>7</sup> Facultad de Medicina y Ciencias, Universidad San Sebastián, 5110246 Valdivia, Chile

\* Correspondence: rodrigo.maldonado@uss.cl



**Citation:** Wernig-Zorc, S.; Schwartz, U.; Martínez-Rodríguez, P.; Inalef, J.; Pavicic, F.; Ehrenfeld, P.; Längst, G.; Maldonado, R. The Long Non-Coding RNA MALAT1 Modulates NR4A1 Expression through a Downstream Regulatory Element in Specific Cancer Cell Types. *Int. J. Mol. Sci.* **2024**, *25*, 5515. <https://doi.org/10.3390/ijms25105515>

Academic Editors: Guilherme Tralhão and Rui Caetano Oliveira

Received: 16 April 2024

Revised: 7 May 2024

Accepted: 11 May 2024

Published: 18 May 2024



**Copyright:** © 2024 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/>).

**Abstract:** Long non-coding RNAs (lncRNAs) have been shown to modulate gene expression and are involved in the initiation and progression of various cancer types. Despite the wealth of studies describing transcriptome changes upon lncRNA knockdown, there is limited information describing lncRNA-mediated effects on regulatory elements (REs) modulating gene expression. In this study, we investigated how the metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) lncRNA regulates primary target genes using time-resolved MALAT1 knockdown followed by parallel RNA-seq and ATAC-seq assays. The results revealed that MALAT1 primarily regulates specific protein-coding genes and a substantial decrease in the accessibility downstream of the *NR4A1* gene that was associated with a decreased *NR4A1* expression. Moreover, the presence of an NR4A1-downstream RE was demonstrated by CRISPR-i assays to define a functional MALAT1/NR4A1 axis. By analyzing TCGA data, we identified a positive correlation between NR4A1 expression and NR4A1-downstream RE accessibility in breast cancer but not in pancreatic cancer. Accordingly, this regulatory mechanism was experimentally validated in breast cancer cells (MCF7) but not in pancreatic duct epithelial carcinoma (PANC1) cells. Therefore, our results demonstrated that MALAT1 is involved in a molecular mechanism that fine-tunes NR4A1 expression by modulating the accessibility of a downstream RE in a cell type-specific manner.

**Keywords:** chromatin-associated lncRNAs; MALAT1; NR4A1; chromatin accessibility; breast cancer

## 1. Introduction

Histones, chromatin-associated proteins, and RNA play essential roles in fine-tuning gene expression. Among them, non-coding RNAs have been identified as key factors modulating nuclear architecture, genome-wide chromatin interactions, gene expression, and genomic integrity [1]. Long non-coding RNAs (lncRNAs) are transcripts of at least 200 nucleotides (nt) in length transcribed by RNA polymerase II; in most cases, they are spliced, 3' polyadenylated, and 5' capped by 7-methyl guanosine [2]. lncRNAs function as cis- or trans-acting transcriptional regulators depending on their functional interactions with chromatin, chromatin-associated factors, and nuclear ribonucleoprotein particles [3].

According to information from different databases, almost 57,000 lncRNA genes have been found to generate more than 127,000 different transcripts from the human genome [4]. The expression of lncRNAs is highly tissue- and cell-type specific, and their regulation of stem cell differentiation and renewal, in combination with specific protein partners, demonstrates their remarkable transcriptional regulatory potential [5]. Lineage engagement for adipose, skeletal, cartilage, muscle, neuronal, and skin cell differentiation, in addition to embryonic stem cell pluripotency, has been described to be guided by different lncRNAs, such as PU.1 AS, Msx1as, HOTTIP, H19, HOTAIRM1, TINCR, RoR, Xist, Airn, and Firre, among others [6–8]. Thus, lncRNAs are associated with the onset and progression of cancer by regulating gene expression at the transcriptional and posttranscriptional levels, affecting intermediary metabolism and cell signaling [2,9–11].

The metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is one of the most abundant lncRNAs in human cells and is localized mainly on nuclear speckles [12]. In humans, MALAT1 is an 8 kb single-exon transcript expressed at chromosomal region 11q13, and studies on its mouse homolog have shown that RNaseP–RNaseZ posttranscriptional processing results in the retention of a long transcript on nuclear speckles and a 61 nt small RNA with a tRNA-like structure [13,14]. Intriguingly, after RNase processing, the 3′ end of the long MALAT1 transcript folds to generate a blunt-ended triple helical structure that stabilizes the transcript through protection from 3′-to-5′ exonucleases [15,16].

The role of the lncRNA MALAT1 has been studied by RNA–RNA interactions using RNA–antisense purification (RAP) and sequencing experiments in mouse embryonic stem cells, which revealed that MALAT1 interacts mainly with pre-mRNAs, such as arginine/serine-rich phosphoproteins, through proteins involved in alternative splicing [17]. Additionally, the function of MALAT1 has been studied by capture hybridization analysis of RNA targets (CHART) in MCF7 breast cancer cells, which was found to interact with DNA at numerous genomic sites, preferentially associated with the coding region and termination or polyadenylation sites of transcriptionally active and spliced genes [18]. Then, by combining mass spectrometry with the CHART assay (CHART-MS), the preferential interaction of MALAT1 with RNA processing factors, such as nuclear speckle and paraspeckle proteins and transcriptional regulators, was revealed [18].

Alterations in MALAT1 expression and/or function have been associated with different types of cancer. MALAT1 has been shown to act as a tumor suppressor or a tumor promoter, depending on the cancer type examined, or even to have opposite roles in the same cancer type [9,19,20]. These discrepancies could be explained by the diverse effects of MALAT1 downregulation on different cancer cells, where genome-wide transcriptomic changes have been reported. However, studies on the time-dependent modulation of MALAT1 expression and how this alteration in gene expression and chromatin accessibility affect gene-specific mechanisms of transcriptional regulation in cancer are still scarce.

In this study, we performed a time-resolved modulation of MALAT1 expression with locked nucleic acid (LNA) Gapmers followed by RNA-seq to define primary transcriptional targets in HeLa cells as a model system. In parallel, chromatin accessibility changes were analyzed by ATAC-seq, which revealed general changes and specific and considerable loss of accessibility for an element downstream of the nuclear receptor subfamily 4 Group A Member 1 (*NR4A1*) gene associated with its downregulation. Moreover, using CRISPR-i technology, we determined a functional axis in which MALAT1 modulates the accessibility of the *NR4A1*-downstream RE regulating *NR4A1* expression (the *NR4A1*/MALAT1 axis). *NR4A1* is a transcription factor that is dysregulated in different cancer types, involved in T-cell dysfunction, and associated with the development of fibrotic diseases [21–26]. Therefore, to understand the role of this axis in cancer, we correlated the chromatin accessibility of an *NR4A1*-downstream RE and the expression of the *NR4A1* gene in TCGA data. Interestingly, we found cancer type-specific correlations, such as in breast cancer, but no correlation was observed in the pancreatic cancer data. Then, we experimentally evaluated the effect of the *NR4A1*/MALAT1 axis on specific cancer types through MALAT1 downregulation and CRISPR-i assays. The results showed that the *NR4A1* gene was specifically regulated by

MALAT1 in MCF7 breast cancer cells but not in pancreatic ductal adenocarcinoma PANC1 cells. Our findings delineate the molecular mechanism of MALAT1 that fine-tunes the expression of specific cancer modulators, such as NR4A1, by regulating the chromatin accessibility of an NR4A1-downstream RE.

## 2. Results

### 2.1. MALAT1 Primarily Regulates Protein-Coding Genes

To determine the primary transcriptional targets involved, we performed a time-dependent modulation of MALAT1 expression by transfecting HeLa cells with LNA Gapmers and harvesting them after 24 or 48 h for RNA-seq or ATAC-seq (Figure 1A). The time-resolved MALAT1 knockdown (KD) effect was validated by real-time PCR (Figure 1B) and RNA-seq (Figure 1C) at both timepoints, where a Gapmer against Beta Galactosidase (LacZ) was used as a negative control. The genome-wide transcriptional changes showed that 24 h after MALAT1 KD, there was a greater number of downregulated transcripts (275) than upregulated transcripts (124) (Figure 1D,E; Supplementary Table S1). The vast majority of the differentially expressed transcripts were protein-coding genes (Figure 1E, Supplementary Table S1). At 48 h after MALAT1 Gapmer transfection, the number of differentially expressed transcripts strongly increased, reaching similar levels for both up- and downregulated transcripts (Figure 1D,E; Supplementary Tables S2 and S3). At this timepoint, in addition to the prevalence of protein-coding genes, we also observed the differential expression of non-coding genes (Figure 1E and Supplementary Tables S2 and S3). Gene ontology analysis by overrepresentation test [27,28] of the differentially expressed transcripts at 24 h after MALAT1 KD revealed that the downregulated genes were enriched in transcripts associated with transcriptional regulation (ATP-dependent activity, acting on DNA, and transcription coregulator activity) (Figure 1F, left). While upregulated genes were enriched in transcripts associated with the structural components of ribosomes and chromatin (Figure 1F, right).

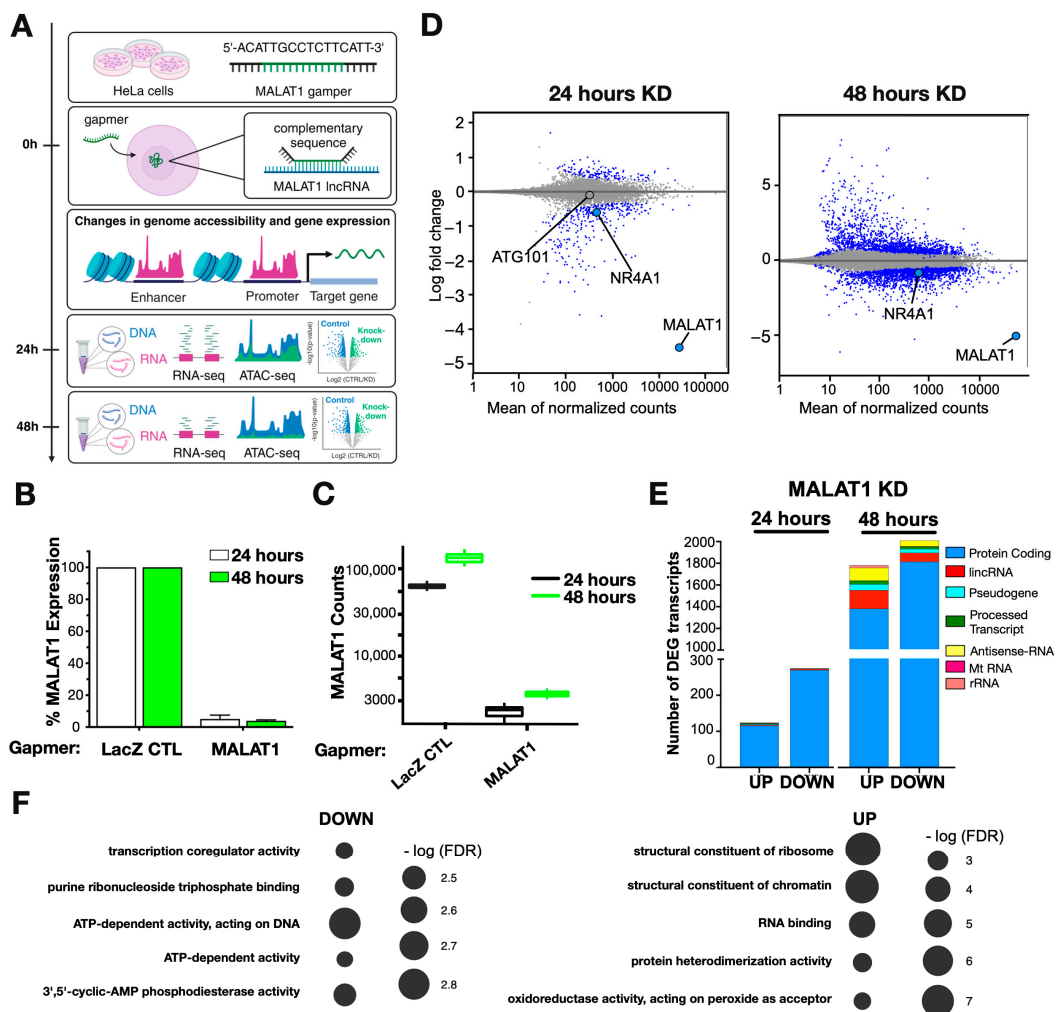
The differences observed between 24 and 48 h after MALAT1 KD suggest two important aspects of the function of this lncRNA. First, MALAT1 is predominantly associated with the regulation of protein-coding genes. Second, the time-dependent changes in the transcriptome suggested the presence of primary target genes after 24 h and the emergence of incidental targets 48 h after knockdown, when transcriptional regulators (proteins and nc-RNAs) become differentially expressed.

### 2.2. Chromatin Accessibility Changes after MALAT1 Downregulation

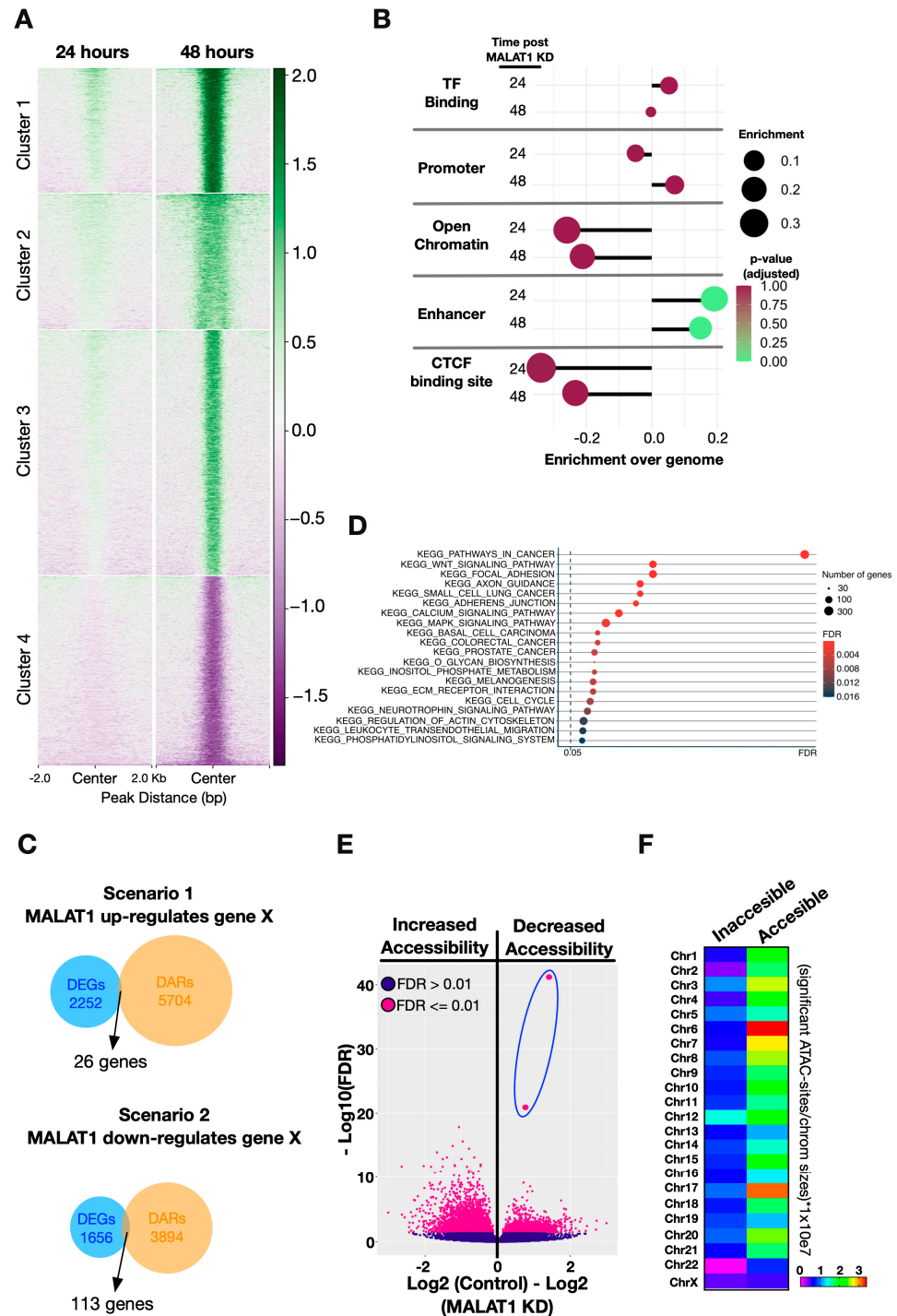
To delineate potential mechanisms by which MALAT1 regulates the expression of protein-coding genes through changes in chromatin accessibility, we performed ATAC-seq experiments 24 and 48 h after MALAT1 KD (Figure 1A). In general, after MALAT1 KD, sites became accessible (Figure 2A, Clusters 1–3) or inaccessible (Figure 2A, Cluster 4) after 24 h, and the magnitude of the change tended to increase at 48 h. Therefore, considering that MALAT1 primary target genes were found after 24 h of KD via RNA-seq, we focused on early structural changes at this time point. We identified numerous significantly altered ATAC sites with increased or decreased chromatin accessibility distributed across all chromosomes (Supplementary Figure S1A). Nevertheless, by comparing against genomic sites, we calculated the enrichment or reduction of the significantly altered ATAC sites per chromosome. The analysis revealed enrichment of significantly altered ATAC sites on chromosomes 1, 5, 6, 8, 11, 12, 15, 17, and 20 and reductions in chromosomes 4, 9, 13, 14, 16, 21, and X (Supplementary Figure S1B). Despite the broad changes in accessibility, these changes were concentrated on specific chromosomes.

Then, to determine whether the significant changes in accessibility were located within REs, we analyzed their enrichment over specific genomic elements and found that enhancer regions are the most affected after 24 and 48 h after MALAT1 KD (Figure 2B). Thus, to understand the accessibility changes in terms of gene expression, we defined two scenarios for MALAT1 gene regulation through accessibility changes. The first one

was with MALAT1 inducing the expression of a target gene maintaining the accessibility of the corresponding REs under normal conditions (scenario 1), and the second one was where MALAT1 represses the expression of a target gene by decreasing the accessibility of the corresponding REs under normal conditions (scenario 2) (Supplementary Figure S1C). Using both scenarios as parameters to find the overlap between differentially expressed genes (DEGs) and differentially accessible regions (DARs), we found a small number of genes up- or downregulated according to the accessibility changes after MALAT1 KD (Figure 2C). These results indicate that transcriptional changes following accessibility alterations due to the loss of MALAT1 are limited to specific genes. KEGG co-expression analysis showed that pathways related to cancer are the most enriched among DEGs and DARs together (Figure 2D), where nuclear glucocorticoid receptor binding is the only significant enriched GO term found for scenario 1 (with ETS2 and NR4A1 genes), while for scenario 2, we found more terms that are significantly enriched; nevertheless, they are more general compared to scenario 1 (Supplementary Figure S1D).



**Figure 1.** MALAT1 primarily modulates the expression of protein-coding genes. (A) Methodology for determining the time-resolved MALAT1 KD through the transfection of Gappers, followed by RNA-seq and ATAC-seq assays in HeLa cells. Validation of MALAT1 KD at 24 and 48 h post transfection by real-time PCR (B) and RNA sequencing (C). (D) MA plot of the RNA-seq data showing the MALAT1 KD values at 24 and 48 h post transfection. Significantly up- and downregulated transcripts are indicated in blue. (E) Quantification of significantly up- and downregulated MALAT1 transcripts from RNA-seq data at 24 and 48 h post transfection. (F) Gene ontology analysis by overrepresentation test of the differentially expressed transcripts 24 h post transfection.

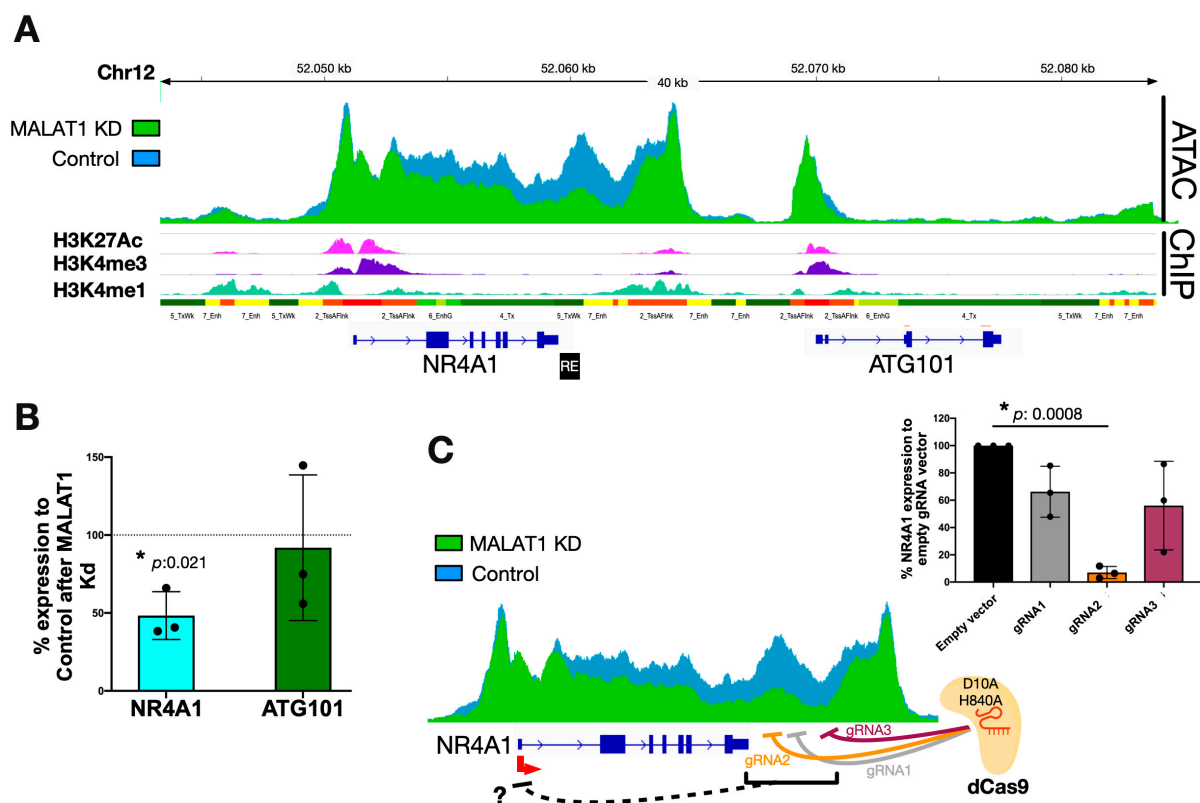


**Figure 2.** Chromatin accessibility changes after MALAT1 downregulation. (A) Heatmap showing the K-means clusters of differentially accessible ATAC-seq sites in MALAT1 KD versus control 24 and 48 h post transfection. Accessibility scores are shown on the right side. (B) Enrichment of significant ATAC sites over specific regulatory elements after 24 and 48 h post MALAT1 KD. (C) Overlap quantification of the differentially expressed genes (DEGs) and differentially accessible regions (DARs) after 24 and 48 h post MALAT1 KD, according to the scenarios described on the Supplementary Figure S1C. (D) KEGG co-expression analysis of the DEGs and genes associated with DARs after MALAT1 KD. (E) Extent and type of change in chromatin accessibility after MALAT1 KD for significant (purple dots) and non-significant (violet dots) ATAC-seq sites. (F) Chromosomal distribution of significant accessible/inaccessible ATAC-seq sites after MALAT1 KD.

Now, examining the sites that gained or lost accessibility, we focused on the two sites exhibiting the greatest extent of decrease in accessibility after 24 h of MALAT1 KD (Figure 2E, dots inside the blue oval). These sites are part of cluster 4 (Figure 2A) and are located on chromosome 12, a chromosome exhibiting a higher percentage of inaccessible sites than average (Figure 2F).

### 2.3. MALAT1 Regulates NR4A1 Expression by Modulating the Chromatin Accessibility of a Downstream RE

Specifically, the highly inaccessible sites found after MALAT1 KD were located downstream of the NR4A1 gene (Chr12: 52.022.832–52.059.507 on GRCh38) in a region flanked by additional strong ATAC peaks and coinciding with CTCF-ChIP-seq peaks (red and black squares, respectively; Supplementary Figure S2A). This finding suggested the presence of a topologically associated domain (TAD), in which the NR4A1 and ATG101 genes are expressed, as observed by our RNA-seq data (Supplementary Figure S2A). Additionally, the active histone marks H3K27Ac/H3K4me3/H3K4me1 indicate the NR4A1 and ATG101 promoters and enhancer regions upstream and downstream of the genes, respectively (Figure 3A). Cap analysis of gene expression (CAGE) data from HeLa cells revealed the location of the transcriptional start site, corresponding to the NR4A1 transcript ENST00000394825 (Supplementary Figure S2B). The promoter and coding region of this transcript cover a broad ATAC-accessible region, including the downstream element (Figure 3A).



**Figure 3.** MALAT1 modulates the accessibility of a regulatory element downstream of NR4A1. (A) Chromatin states throughout the NR4A1 gene. (B) Real-time PCR expression analysis of NR4A1 and ATG101 after MALAT1 KD. (D) Scheme depicting the CRISPR-i approach using 3 gRNAs against the NR4A1-downstream regulatory element. (C) Real-time PCR quantification of NR4A1 expression after CRISPR-i assays. \* Significant differences are shown with  $p$ -values obtained from  $t$  tests.

Twenty-four hours after MALAT1 KD, the NR4A1 gene was significantly downregulated without affecting the expression of ATG101 (Figure 1D, Supplementary Table S1).

This effect was verified by RT-qPCR (Figure 3B), which suggested a primary MALAT1-dependent loss of chromatin accessibility at a novel NR4A1-downstream RE (Figure 3A, black box). To test the functionality of the proposed RE, we used the CRISPR-i system, which consists of catalytically inactive Cas9 (dCas9) and three independent guide RNAs (gRNAs) targeting the NR4A1-downstream RE (Figure 3C). Guide RNAs were designed to bind directly to the proposed RE (gRNA2) or 1 kbp (gRNA1) and 2 kbp (gRNA3) downstream of the RE. HeLa cells were independently transfected with plasmids encoding guide RNAs (Supplementary Figure S3A) and dCas9. RT-qPCR analysis revealed the repression of NR4A1 expression by all three guide RNAs, with gRNA 2 directly targeting the NR4A1-downstream RE, being the most efficient at repressing NR4A1 expression (Figure 3C). These results suggest the presence of a potential downstream regulatory element for the NR4A1 gene.

Additionally, we tested whether the MALAT1 lncRNA is physically associated with the NR4A1-downstream RE to maintain accessibility at this locus. Using the RAP assay, we pulled-down MALAT1 from HeLa nuclear extracts (Supplementary Figure S3B). However, we did not find a direct physical interaction between MALAT1 and the NR4A1-downstream RE (Supplementary Figure S3C), suggesting that additional factors mediate the chromatin accessibility of this locus (see discussion).

Taken together, the results showed that MALAT1 regulates the accessibility of a novel downstream regulatory element downstream of the NR4A1 gene, fine-tuning its expression in HeLa cells.

#### *2.4. The Accessibility of the NR4A1-Downstream RE Correlates with NR4A1 Expression in Specific Cancer Types*

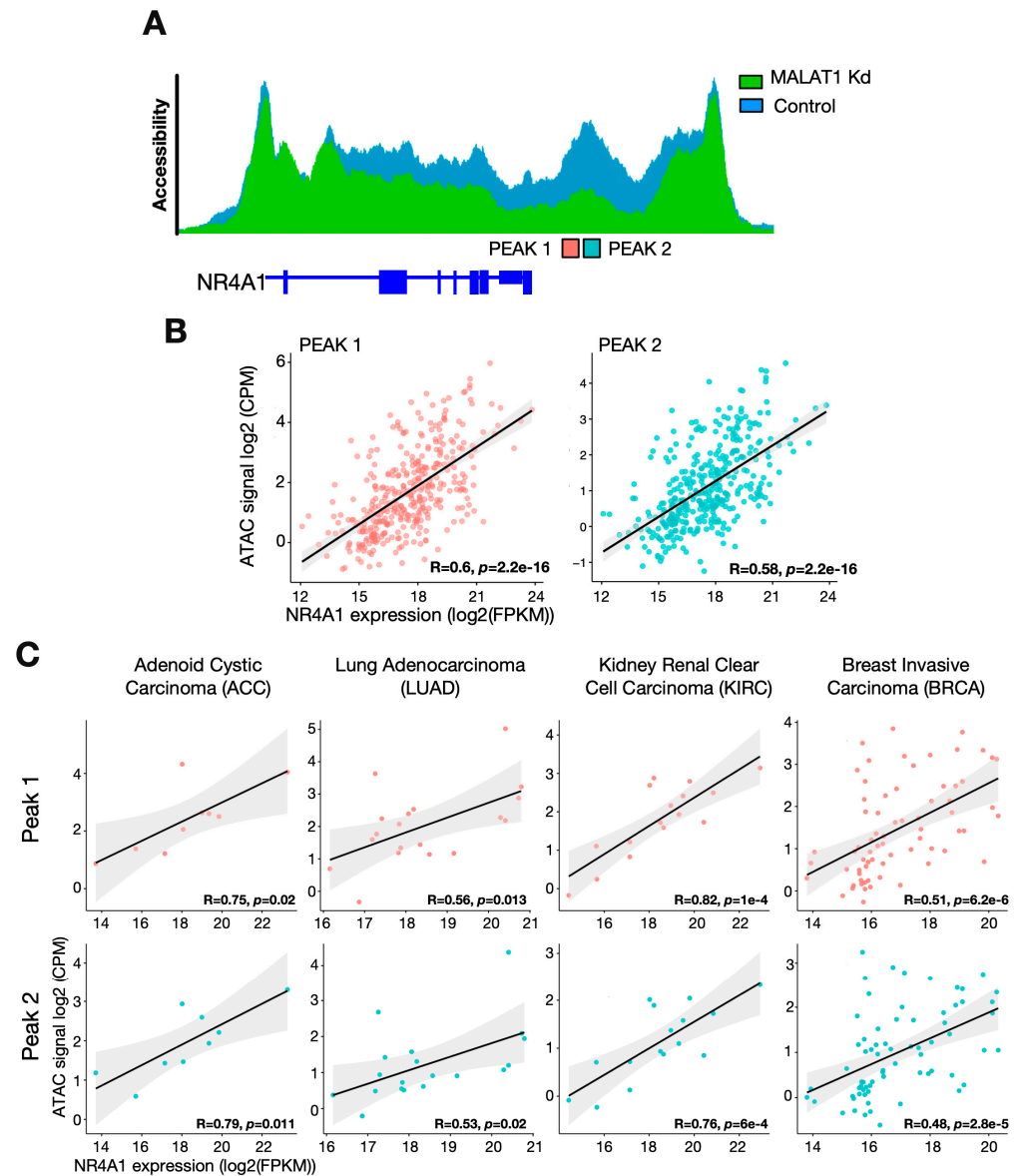
To address the functional role of the NR4A1-downstream RE, we screened the TCGA cancer database [29]. Indeed, we observed an overlap between the NR4A1-downstream RE and two ATAC-seq peaks derived from TCGA data (Figure 4A, PEAK1 and PEAK2). The intensity of both ATAC-seq peaks was positively correlated with NR4A1 expression in all cancer types (Figure 4B). Furthermore, analyzing the accessibility of the two ATAC peaks in individual cancer types showed that the accessibility of this region was cancer type-specific (Supplementary Figure S4). The highest level of accessibility was observed for pheochromocytoma and paraganglioma (PCPG), and the lowest ATAC-seq scores were found for kidney renal papillary cell carcinoma (KIRP) (Supplementary Figure S4). These cancer-specific effects led us to evaluate the correlation between accessibility for both peaks and NR4A1 expression in different cancer types. Here, we identified a strong correlation between breast invasive carcinoma (BRCA) and other cancer types (Figure 4C).

The direct correlation between NR4A1 expression and the accessibility of the NR4A1-downstream RE defined a MALAT1/NR4A1 axis, which, according to our initial results in HeLa cells and the bioinformatic analysis of TCGA data, suggested that MALAT1 regulates NR4A1 expression through accessibility changes in the NR4A1-downstream RE in a cancer-specific manner.

#### *2.5. The MALAT1/NR4A1 Axis Is Functional in Breast Cancer Cells but Not in Pancreatic Cancer Cells*

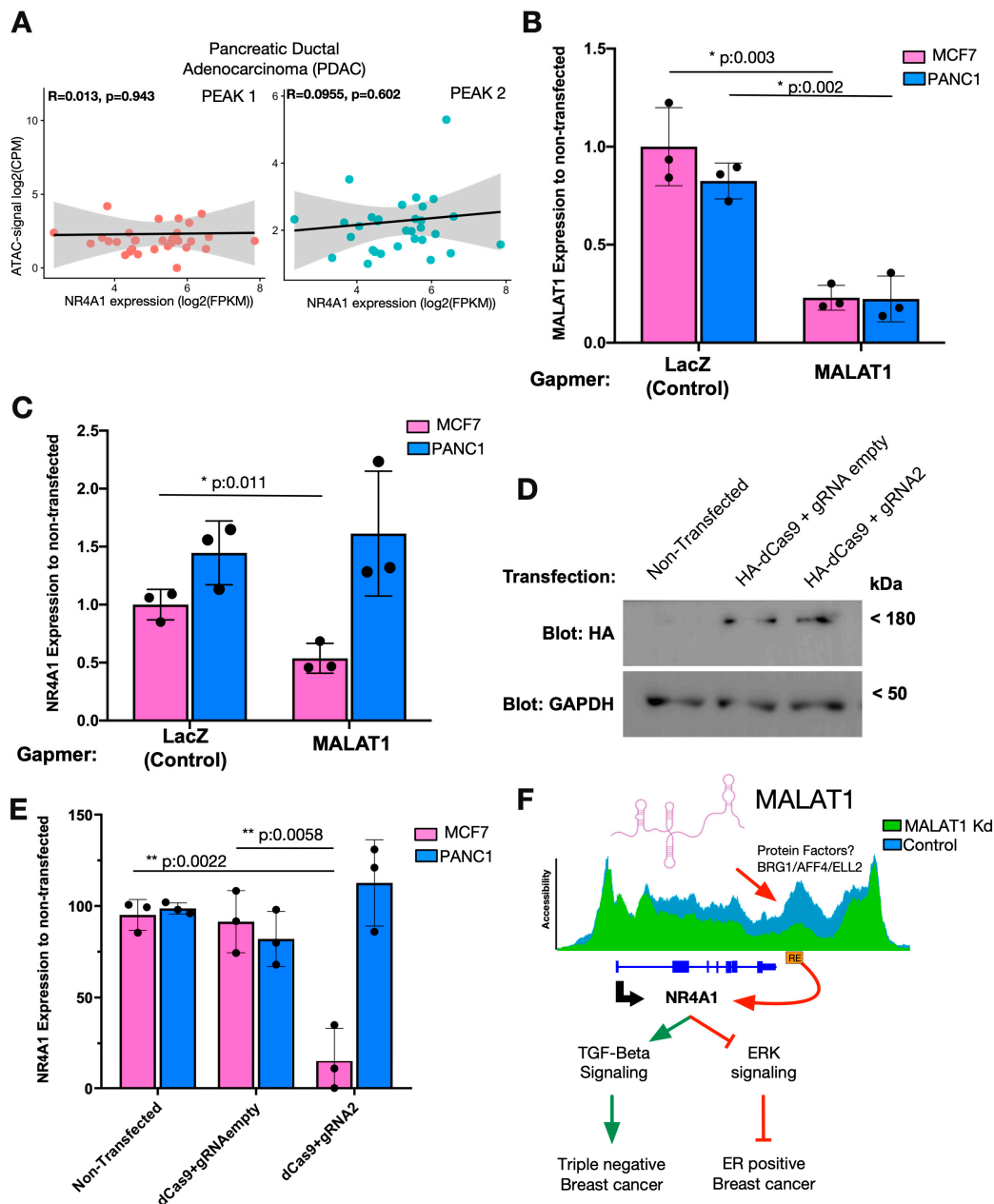
To test the hypothesis that a cancer-type-specific MALAT1/NR4A1 axis modulates the expression of NR4A1, we evaluated the effect of this axis in specific cancer cell types. We tested whether MALAT1 expression and NR4A1-downstream RE accessibility were strongly correlated in cancer types, such as BRCA (Figure 4C), versus low- or no-correlation cancers, such as pancreatic adenocarcinoma (PDAC) (Figure 5A). To assess the functionality of the MALAT1/NR4A1 axis, we transfected MCF7 breast cancer and PANC1 pancreatic adenocarcinoma cells with Gapmers to downregulate MALAT1 expression (Figure 5B). MALAT1 KD induced a significant decrease in NR4A1 expression in MCF7 cells, while no difference was observed in PANC1 cells compared to the controls (Figure 5C). Furthermore, we evaluated the regulatory function of the NR4A1-downstream RE by transfecting dCas9 and gRNA2 into MCF7 and PANC1 cells (Figure 5D) and evaluated the expression of

NR4A1. A significant reduction in the expression of NR4A1 was observed in MCF7 cells but not in PANC1 cells (Figure 5E). These results indicate that the NR4A1-downstream RE is sufficient to modulate the expression of the NR4A1 gene in MCF7 cells but is not functional in PANC1 cells.



**Figure 4.** Cancer-specific correlation between the accessibility of the NR4A1-downstream RE and the expression of NR4A1. (A) Schematic representation of the TCGA-derived peaks (PEAK1 and PEAK2) found on the NR4A1-downstream RE. (B) Correlation of TCGA-derived peak accessibility and NR4A1 expression in all TCGA cancer types shown in Supplementary Figure S4. (C) Correlations of the accessibility of TCGA-derived peaks with NR4A1-downstream RE and NR4A1 expression in individual TCGA cancer types.





**Figure 5.** MALAT1 regulates NR4A1 in breast cancer cells but not in pancreatic adenocarcinoma cells. (A) Correlation between TCGA-derived peak accessibility and NR4A1 expression in PDAC patients. (B) MALAT1 KD validation in MCF7 and PANC1 cells using Gapmers. (C) Real-time PCR analysis of NR4A1 expression after MALAT1 KD in MCF7 and PANC1 cells. (D) Western blot analysis of HA-tag expression to verify the effect of dCas9 transfection on MCF7 cells. (E) CRISPR-i assays with gRNA2 analyzing the expression of NR4A1 by real-time PCR in MCF7 and PANC1 cells. (F) Scheme representing the molecular mechanism underlying the MALAT1-mediated modulation of NR4A1 through the accessibility of a downstream RE. \* Significant differences are shown with  $p$ -values obtained from  $t$  tests.

### 3. Discussion

LncRNAs are key players in the regulation of cellular transcription. During cancer onset and progression, altered transcriptional programs trigger uncontrolled intracellular signaling that normally induces cell dedifferentiation and proliferation [30–32]. In this context, the lncRNA MALAT1 is involved in different cancer types and acts as a transcriptional regulator of oncogenes and tumor suppressors [9,19,20]. Our work focused on partially

deciphering the MALAT1-dependent molecular mechanism involved in gene expression regulation by altering the chromatin accessibility of primary gene targets.

### 3.1. MALAT1-Dependent Changes in Chromatin Accessibility

The time-resolved downregulation of MALAT1 demonstrated that its primary target genes corresponded mainly to protein-coding genes. MALAT1 KD induced alterations in the chromatin accessibility of a large number of loci. Such an increase or decrease in chromatin accessibility after the KD of a lncRNA has been described previously [33–36], confirming that lncRNAs are active chromatin regulators. Nevertheless, to our knowledge, no study has investigated the impact of MALAT1 on chromatin accessibility. Our ATAC-seq data revealed the largest changes in chromatin accessibility downstream of the *NR4A1* gene after MALAT1 KD. The local decrease within an open 36 kbp region was associated with the repression of *NR4A1* gene expression. We further showed that the accessibility of this downstream element was MALAT1-dependent, where CRISPR-i experiments targeting the *NR4A1*-downstream RE demonstrated the presence of a regulatory element for the *NR4A1* gene. Similar types of REs have been reported to regulate the *Cja1* (*Cx43*) and mouse globin genes [37,38]. This led us to think about a possible interaction between MALAT1 and the *NR4A1*-downstream RE locus. Nevertheless, MALAT1 RAP assays could not detect the suggested physical interaction, indicating that MALAT1 is not the sole factor involved in the accessibility maintenance of the *NR4A1*-downstream RE in HeLa cells. Consistent with these findings, MALAT1 CHART-seq data showing direct MALAT1-chromatin contacts did not reveal any interaction with the *NR4A1*-downstream RE in MCF7 breast cancer cells, confirming our MALAT1 RAP findings in HeLa cells (Supplementary Figure S5) [18]. Our ATAC-seq data showed that the *NR4A1* promoter and the downstream RE correspond to the region with the highest accessibility of the whole chromosome 12. Therefore, the decrease in the accessibility of the *NR4A1*-downstream RE after MALAT1 KD must involve additional factors that maintain the accessibility of this region under normal conditions. Data retrieved from CHIP-atlas (<https://chip-atlas.org/> accessed on 28 December 2021) in HeLa cells revealed that the transcription activator BRG1, the catalytic subunit of the SWI/SNIF chromatin remodeling complex, binds to the whole *NR4A1* gene body, including the *NR4A1*-downstream RE (Supplementary Figure S6A). A direct MALAT1-BRG1 interaction was previously shown to increase the inflammatory response in hepatocellular carcinoma [39]. Moreover, two other transcription factors involved in the super-elongation complex of RNA polymerase II, AFF4 and ELL2 [40], also bind to the *NR4A1* gene body and the *NR4A1*-downstream RE (Supplementary Figure S6B). MALAT1 KD resulted in a minor decrease in BRG1 and AFF4 expression but also in the significant downregulation of ELL2 (Supplementary Figure S6B and Table S1). Therefore, the putative MALAT1-BRG1 interaction, the significant decrease in ELL2 expression after MALAT1 KD, and the AFF4 occupancy over the RE could explain the MALAT1-dependent changes in chromatin accessibility at the *NR4A1*-downstream RE region (Figure 5F). Nevertheless, in addition to the additional protein factors, we cannot rule out the possibility that other ncRNAs are involved in maintaining the accessibility of this region.

Interestingly, the expression of an antisense lncRNA against *NR4A1*, *NR4A1AS*, was shown to modulate *NR4A1* expression in colorectal cancer cells [41]. The promoter region of *NR4A1AS1* coincides with the *NR4A1*-downstream RE found in this study, suggesting that the expression of this lncRNA is an interesting candidate for understanding the changes in *NR4A1*-downstream RE accessibility. However, our RNA-seq data did not reveal the presence of this lncRNA in HeLa cells. Moreover, these results were verified by RT-qPCR assays in HeLa and MCF7 cells, in which no expression of the lncRNA *NR4A1AS* was detected. These results indicate that *NR4A1AS* is not involved in the model systems used in this study but likely plays a tissue-specific role in other cellular models.

### 3.2. The Impact of the MALAT1/NR4A1 Axis on Breast Cancer Development

Analysis of TCGA data revealed that the chromatin accessibility of the NR4A1-downstream RE is positively correlated with NR4A1 expression in specific cancer types, as we observed in HeLa cells. Here, we focused on breast carcinoma cells due to their high correlation. We recapitulated the MALAT1/NR4A1 axis in ER-positive MCF7 breast cancer cells but did not find this axis in pancreatic adenocarcinoma PANC1 cells, as predicted by the lack of correlation in PDAC analyses. Additionally, this cell type-specific effect could be explained by mutations on the NR4A1-downstream RE or the TAD boundaries where this gene is located on the pancreatic cancer cells, but not on the breast cancer cells, following the enhancer retargeting process observed on the ZCCHC7 gene during B cell lymphoma progression [42].

The function of MALAT1 in breast cancer cells is still a matter of debate. Both, oncogenic and tumor suppressor roles have been described for this lncRNA [43]. Our results agree with the role of MALAT1 as a tumor suppressor maintaining NR4A1 expression in MCF7 cells. On one hand, NR4A1 was shown to upregulate ERK signaling at low levels and preserve the levels of proapoptotic proteins, regulating proliferation and controlled cell death [26]. Through these mechanisms, NR4A1 seems to maintain the tamoxifen sensitivity of MCF7 cells [26]. In contrast, in triple negative breast cancer (TNBC) MDA-MB-231 cells, NR4A1 promotes tumor invasion and metastasis by activating transforming growth factor beta (TGF-R) signaling [25]. These results suggest that, depending on the cellular environment, the MALAT1-NR4A1 axis can function as a tumor promoter or suppressor (Figure 5F).

The transcription factor NR4A1 has been described as an immediate early gene (IEG) that corresponds to a group of genes that rapidly respond to stress through their fast and transient transcriptional stimulation [44]. Replication stress in the absence of checkpoints and quality control machinery is one of the first steps leading to genomic instability during cancer development [45]. By comparing non-malignant MCF10 breast epithelial cells with TNBC MDA-MB-231 cells and circulating patient-derived breast tumor cells, NR4A1 was recently reported to modulate IEG expression through interaction with gene bodies and suppression of transcriptional elongation. Under oncogenic replication stress conditions, NR4A1 is released, triggering a burst of IEG expression [46].

Taken together, these data point to a mechanism by which MALAT1 modulates the accessibility of an NR4A1-downstream RE, probably through other protein factors, regulating NR4A1 expression levels. The presence of MALAT1 seems to fine-tune NR4A1 expression levels, which in turn are associated with low proliferative signaling and pro-apoptosis cues in ER-positive MCF7 cells. The discovery of the MALAT1-NR4A1 axis increases the potential of these genes as therapeutic targets for ER-positive breast cancer and drug resistance treatments.

## 4. Materials and Methods

### 4.1. Cell Culture

HeLa, MCF7, and PANC1 cells were cultured in HyClone Dubelco's modified Eagle medium (DMEM) supplemented with glucose (4.5 g/L), L-glutamine (584 mg/L), sodium pyruvate (110 mg/L), and 10% fetal bovine serum. The cells were maintained at 37 °C in 5% CO<sub>2</sub> and subcultured using 0.25% trypsin.

### 4.2. Western Blot

Protein extracts from HeLa and MCF7 cells were obtained after homogenization in RIPA buffer supplemented with protease inhibitors by pipetting and sonication at 4 °C. The proteins were resolved by SDS-PAGE and transferred to PVDF membranes. The membranes were blocked with 5% milk in TBST (TBS with 0.1% Tween-20<sup>®</sup>), probed with antibodies against GAPDH and the HA-tag, and then with horseradish peroxidase-coupled secondary antibodies. Luminescent blot signals were obtained by incubating the samples with peroxidase substrate and imaged on a Syngene G:box gel documentation system.

#### 4.3. Real-Time PCR

Total RNA (1 µg) from HeLa, MCF7, and PANC1 cells was used for cDNA preparation with a RevertAid RT Kit (Thermo Scientific, Waltham, MA, USA). Real-time PCRs were prepared in 0.1 mL tubes with a final volume of 20 µL containing specific primers (MALAT1, NR4A1, and GAPDH), variable amounts of cDNA (depending on the target), 10 µL of the 2X JumpStart™ ReadyMix™, 1 mM MgCl<sub>2</sub>, and SyberGreen (1:400,000 from the 10,000X stock, Thermo Scientific, Waltham, MA, USA). Reactions were analyzed on a Rotor-Gene RG3000 thermal cycler.

#### 4.4. Cloning of Guide RNAs

The backbone vector (pZDonor) was modified to include BbsI restriction sites, generating the pSPgRNA plasmid (#47108 Addgene), to allow the cloning of gRNAs of choice as described before [47]. In brief, guide RNAs were ordered as oligonucleotides from IDT Technologies (Figure 3E), hybridized, phosphorylated, ligated with the BbsI digested pSPgRNA plasmid, and cloned into DH5α cells. Positive clones were verified initially by colony PCR and then by sequencing (Supplementary Figure S3A).

#### 4.5. Transfections

HeLa cells were reverse-transfected with Lipofectamine 2000 for Gapmer-mediated MALAT1 downregulation and CRISPR-I assays. Gapmer transfections were carried out using 1.5 µL of 200 µM Gapmers against MALAT1 or LacZ and 5 µL of Lipofectamine for 1.2 million cells in a 6-well plate for 24 h post transfection. For 48 h post transfection, we used half the number of cells. For CRISPR-i, we used 5 µL of Lipofectamine, 750 µg of dCas9 plasmid (#61355 Addgene), and 750 µg of each gRNA plasmid (#47108 Addgene) for 0.5 million cells on a 6-well plate. The cells were harvested 48 h post transfection.

MCF7 cells were plated and transfected with Lipofectamine 3000 for Gapmer-mediated MALAT1 downregulation and CRISPR-I assays. Gapmer transfections were carried out using 2 µL of 200 µM Gapmers, against MALAT1 or LacZ, 2 µL of lipofectamine, and 1 µL of P3000 reagent; 0.5 million cells were seeded on a 6-well plate the day before transfection. Cells were harvested 24 h post transfection. For CRISPR-i, we used 2 µL of Lipofectamine, 650 µg of dCas9 plasmid (#47106 Addgene), 650 µg of each gRNA plasmid (#47108 Addgene), and 2.5 µL of P3000 reagent for 0.5 million cells seeded the day before they were seeded on a 6-well plate. The cells were harvested 48 h post transfection.

#### 4.6. Custom Genome

A custom version of the hg38 human genome was used for all sequencing analyses, in which all repeats of ribosomal genes were masked with N bases and a single repeat of the ribosomal gene was kept as an extra chromosome named “ChrR” [48].

#### 4.7. Annotation

A custom version of the gencode v27 *Homo sapiens* annotation was used, with ribosomal RNA present one time on an extra chromosome, ChrR, to fit our custom genome.

#### 4.8. RNA Purification

Total RNA purification for cDNA preparation was performed following the manufacturer’s instructions for TRIzol® Reagent (Ambion by Life Technologies, Carlsbad, CA, USA). For ChRIP assays, RNA purified from immunoprecipitated samples was extracted by resuspending the beads in proteinase K buffer (10% SDS, 50 mM EDTA, 10 mM Tris-Cl, pH 7.4) and incubating them with 50 µg of proteinase K for 1 h at 50 °C with gentle agitation. Then, the crosslinking was reversed at 65 °C for 2 h, followed by RNA purification with TRIzol and isopropanol precipitation.

#### 4.9. RNA Sequencing and Transcriptome Analysis

The quality and adapter contamination of the raw files were first checked with fastqc v0.11.9 and multiqc v1.13 [49,50]. Following the quality check, adapters were removed using Trimmomatic v0.38 [51], with the following flags: PE -phred33 ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 MINLEN:30 SLIDINGWINDOW:4:20 and cutadapt v4.0 [52] for library- and sequencer-specific trimming, with the following flags: -u 5-a "A{100}" -a "G{100}" -a "T{100}". The reads were then aligned to the reference genome hg38 using STAR v2.7.2a [53] and the following flags: --sjdbGTFfile gencode.v27.annotation.gtf --outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNoverReadLmax 0.04 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000 --outSAMtype BAM SortedByCoordinate --outFilterScoreMinOverLread 0 --outFilterMatchNminOverLread 0 --outFilterMatchNmin 0. Unmapped reads and reads with a MAPQ mapping quality lower than 30 were removed for downstream analysis using SAMtools v1.90 [54]. Next, read counts were obtained with FeatureCount from the Subread package v2.0.0 [55] using the following flags: -p -s 1 -t exon -g gene\_id -C -B. Differential expression analysis between the knockdown samples and the controls was performed in R v4.1.2 with the DESeq2 R package [56], with apeglm shrinkage estimation for ranking of genes and visualization [57]. Differentially expressed genes were filtered with an absolute log fold change > 1.0 and an FDR < 0.05.

#### 4.10. ATAC-seq Assays, Sequencing, and Analysis

The ATAC assay was performed following the instructions of the Omni-ATAC protocol [58] using the Nextera DNA Flex Library prep kit (Illumina, San Diego, CA, USA). Tagmented DNA was analyzed on a bioanalyzer and sequenced on a Next-Seq 550 device in the Genomics core unit of the Regensburger Centrum für Interventionelle Immunologie.

The quality and adapter contamination of the raw files were first checked with fastqc v0.11.9 and multiqc v1.13 [49,50]. Following the quality check, adapters were removed using Trimmomatic v0.38 [51]. Paired-end reads were then mapped to the reference genome hg38 using Bowtie2 [59] with the following flags: --local --very-sensitive-local -X 5000 --no-discordant -dovetail. Unmapped reads and reads with a MAPQ mapping quality lower than 30 were removed for downstream analysis using SAMtools v1.90 [54]. Reads mapping to chrM, blacklisted regions, and PCR duplicates were removed using SAMtools and Picard, respectively. Finally, the reads were shifted based on the Tn5 cuts using DeepTools v3.1.2 [60] alignmentSieve--ATACshift. Peak calling was performed using MACS2 v2.2.71 [61], with the following flags: callpeak -f BAMPE -g hs -broad -keep-dup all -buffer-size = 100. Differentially accessible regions were obtained with the DiffBind R package. Differentially accessible regions were filtered with an absolute Log fold change > 1.0 and an FDR < 0.01. Regulatory regions were assigned to genes using the GREAT tool [62] via the basal plus extension method (5 kb upstream, 1 kb downstream plus distal up to 50 kb).

#### 4.11. Correlations with TCGA and Pancreatic Cancer Data

The normalized ATAC-seq peak signals (ID: Peak1 = PRAD\_72732 and Peak2 = PRAD\_72733) of TCGA tumor samples were obtained from the ATAC-seq Hub (<https://ataseq.xenahubs.net> accessed on 6 February 2021) [29]. The corresponding gene expression data of the matching donors were quantified using the TCGAbiolinks Bioconductor package [63]. Here, we used the upper quartile-normalized FPKM values.

The accessibility signals for Peak1 and Peak2 were subsequently plotted against the NR4A1 expression. For the PANC1 cell line, ATAC-seq and RNA-seq data were obtained from the Gene Expression Omnibus (GEO) database (GSE124229 and GSE124231). First, the gene expression and accessibility signals were normalized to the library size for each patient. The accessibility signals for Peak1 and Peak2 were subsequently plotted against the NR4A1 expression. Linear regression was used to model the correlation between gene

accessibility and gene expression signals in R v.4.1.2. Heatmaps were generated using DeepTools v.3.4.3 [60] and the pheatmap R package.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ijms25105515/s1>.

**Author Contributions:** Conception and design: R.M., G.L. and S.W.-Z.; administrative support: R.M. and G.L.; provision of study materials: R.M., G.L. and P.E.; collection and assembly of data: R.M., S.W.-Z., U.S., P.M.-R., J.I. and F.P.; data analysis and interpretation: S.W.-Z., U.S., R.M. and G.L.; manuscript writing: S.W.-Z., G.L. and R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Agencia Nacional de Investigación y Desarrollo (ANID, FONDECYT de Iniciación 11190246 to R.M.) and by the Deutsche Forschungsgemeinschaft (SFB960 and LA1331 to G.L.).

**Informed Consent Statement:** No applicable.

**Data Availability Statement:** ChIP data for the histone marks H3K27Ac, H3K4me3, and CTCF were extracted from ENCODE under the accession GSE29611. The raw sequencing data were deposited in The Sequence Read Archive (SRA) under the accession ID PRJNA943483. The code used for the analysis is available on GitHub: <https://github.com/sarawernig/Cancer-specific-MALAT1-regulation>.

**Acknowledgments:** The results shown here are in part based upon the data generated by the TCGA Research Network <https://www.cancer.gov/tcga> accessed on 3 March 2020.

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

## References

1. Marchese, F.P.; Raimondi, I.; Huarte, M. The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* **2017**, *18*, 206. [[CrossRef](#)] [[PubMed](#)]
2. Statello, L.; Guo, J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. [[CrossRef](#)] [[PubMed](#)]
3. Li, X.; Fu, X.D. Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.* **2019**, *20*, 503–519. [[CrossRef](#)] [[PubMed](#)]
4. Volders, P.J.; Anckaert, J.; Verheggen, K.; Nuytens, J.; Martens, L.; Mestdagh, P.; Vandesompele, J. Lncipedia 5: Towards a reference set of human long non-coding rnas. *Nucleic Acids Res.* **2019**, *47*, D135–D139. [[CrossRef](#)]
5. Flynn, R.A.; Chang, H.Y. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell* **2014**, *14*, 752–761. [[CrossRef](#)]
6. Chen, L.; Zhang, S. Long noncoding RNAs in cell differentiation and pluripotency. *Cell Tissue Res.* **2016**, *366*, 509–521. [[CrossRef](#)] [[PubMed](#)]
7. Much, C.; Lasda, E.L.; Lewandowski, J.P.; Smallegan, M.J.; Rinn, J.L. The temporal dynamics of lncRNA *Firre*-mediated epigenetic and transcriptional regulation. *bioRxiv* **2022**. [[CrossRef](#)]
8. Arab, K.; Park, Y.J.; Lindroth, A.M.; Schäfer, A.; Oakes, C.; Weichenhan, D.; Lukanova, A.; Lundin, E.; Risch, A.; Meister, M.; et al. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Mol. Cell* **2014**, *55*, 604–614. [[CrossRef](#)]
9. Nandwani, A.; Rathore, S.; Datta, M. LncRNAs in cancer: Regulatory and therapeutic implications. *Cancer Lett.* **2021**, *501*, 162–171. [[CrossRef](#)]
10. Lin, W.; Zhou, Q.; Wang, C.Q.; Zhu, L.; Bi, C.; Zhang, S.; Wang, X.; Jin, H. LncRNAs regulate metabolism in cancer. *Int. J. Biol. Sci.* **2020**, *16*, 1194–1206. [[CrossRef](#)]
11. Peng, W.X.; Koirala, P.; Mo, Y.Y. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* **2017**, *36*, 5661–5667. [[CrossRef](#)] [[PubMed](#)]
12. Hutchinson, J.N.; Ensminger, A.W.; Clemson, C.M.; Lynch, C.R.; Lawrence, J.B.; Chess, A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genom.* **2007**, *8*, 39. [[CrossRef](#)] [[PubMed](#)]
13. Ji, P.; Diederichs, S.; Wang, W.; Böing, S.; Metzger, R.; Schneider, P.M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; et al. MALAT-1, a novel noncoding RNA, and thymosin  $\beta$ 4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **2003**, *22*, 8031–8041. [[CrossRef](#)] [[PubMed](#)]
14. Wilusz, J.E.; Freier, S.M.; Spector, D.L. 3' End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. *Cell* **2008**, *135*, 919–932. [[CrossRef](#)] [[PubMed](#)]

15. Wilusz, J.E.; JnBaptiste, C.K.; Lu, L.Y.; Kuhn, C.-D.; Joshua-Tor, L.; Sharp, P.A. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **2012**, *26*, 2392–2407. [[CrossRef](#)] [[PubMed](#)]
16. Brown, J.A.; Valenstein, M.L.; Yario, T.A.; Tycowski, K.T.; Steitz, J.A. Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MEN $\beta$  noncoding RNAs. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19202–19207. [[CrossRef](#)]
17. Engreitz, J.M.; Sirokman, K.; McDonel, P.; Shishkin, A.A.; Surka, C.; Russell, P.; Grossman, S.R.; Chow, A.Y.; Guttman, M.; Lander, E.S. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell* **2014**, *159*, 188–199. [[CrossRef](#)]
18. West, J.A.; Davis, C.P.; Sunwoo, H.; Simon, M.D.; Sadreyev, R.I.; Wang, P.I.; Tolstorukov, M.Y.; Kingston, R.E. The Long Noncoding RNAs NEAT1 and MALAT1 Bind Active Chromatin Sites. *Mol. Cell* **2014**, *55*, 791–802. [[CrossRef](#)]
19. Goyal, B.; Yadav, S.R.M.; Awasthee, N.; Gupta, S.; Kunnumakkara, A.B.; Gupta, S.C. Diagnostic, prognostic, and therapeutic significance of long non-coding RNA MALAT1 in cancer. *Biochim. Biophys. Acta-Rev. Cancer* **2021**, *1875*, 188502. [[CrossRef](#)]
20. Sun, Y.; Ma, L. New insights into long non-coding rna malat1 in cancer and metastasis. *Cancers* **2019**, *11*, 216. [[CrossRef](#)]
21. Hedrick, E.; Lee, S.O.; Safe, S. The nuclear orphan receptor NR4A1 regulates  $\beta$ 1-integrin expression in pancreatic and colon cancer cells and can be targeted by NR4A1 antagonists. *Mol. Carcinog.* **2017**, *56*, 2066–2075. [[CrossRef](#)]
22. Palumbo-Zerr, K.; Zerr, P.; Distler, A.; Fliehr, J.; Mancuso, R.; Huang, J.; Mielenz, D.; Tomcik, M.; Fürnrohr, B.G.; Scholtyssek, C.; et al. Orphan nuclear receptor NR4A1 regulates transforming growth factor- $\beta$  2 signaling and fibrosis. *Nat. Med.* **2015**, *21*, 150–158. [[CrossRef](#)] [[PubMed](#)]
23. Liu, X.; Wang, Y.; Lu, H.; Li, J.; Yan, X.; Xiao, M.; Hao, J.; Alekseev, A.; Khong, H.; Chen, T.; et al. Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. *Nature* **2019**, *567*, 525–529. [[CrossRef](#)] [[PubMed](#)]
24. Chen, J.; López-Moyado, I.F.; Seo, H.; Lio, C.-W.J.; Hempleman, L.J.; Sekiya, T.; Yoshimura, A.; Scott-Browne, J.P.; Rao, A. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **2019**, *567*, 530–534. [[CrossRef](#)] [[PubMed](#)]
25. Zhou, F.; Drabsch, Y.; Dekker, T.J.A.; de Vinuesa, A.G.; Li, Y.; Hawinkels, L.J.A.C.; Sheppard, K.-A.; Goumans, M.-J.; Luwor, R.B.; de Vries, C.J.; et al. Nuclear receptor NR4A1 promotes breast cancer invasion and metastasis by activating TGF- $\beta$  signalling. *Nat. Commun.* **2014**, *5*, 3388. [[CrossRef](#)]
26. Kim, Y.C.; Kim, C.Y.; Oh, J.H.; Kim, M.H. Nr4a1 regulates tamoxifen resistance by suppressing erk signaling in er-positive breast cancer. *Cells* **2021**, *10*, 1633. [[CrossRef](#)] [[PubMed](#)]
27. Thomas, P.D.; Campbell, M.J.; Kejariwal, A.; Mi, H.; Karlak, B.; Daverman, R.; Diemer, K.; Muruganujan, A.; Narechania, A. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res.* **2003**, *13*, 2129–2141. [[CrossRef](#)]
28. Thomas, P.D.; Kejariwal, A.; Guo, N.; Mi, H.; Campbell, M.J.; Muruganujan, A.; Lazareva-Ulitsky, B. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* **2006**, *34*, 645–650. [[CrossRef](#)]
29. Corces, M.R.; Granja, J.M.; Shams, S.; Louie, B.H.; Seoane, J.A.; Zhou, W.; Silva, T.C.; Groeneveld, C.; Wong, C.K.; Cho, S.W.; et al. The chromatin accessibility landscape of primary human cancers. *Science* **2018**, *362*, eaav1898. [[CrossRef](#)]
30. Lin, C.Y.; Loven, J.; Rahl, P.B.; Paranal, R.M.; Burge, C.B.; Bradner, J.E.; Lee, T.I.; Young, R.A. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **2012**, *151*, 56–67. [[CrossRef](#)]
31. Martini, M.; De Santis, M.C.; Braccini, L.; Gulluni, F.; Hirsch, E. PI3K/AKT signaling pathway and cancer: An updated review. *Ann. Med.* **2014**, *46*, 372–383. [[CrossRef](#)]
32. Altomare, D.A.; Testa, J.R. Perturbations of the AKT signaling pathway in human cancer. *Oncogene* **2005**, *24*, 7455–7464. [[CrossRef](#)] [[PubMed](#)]
33. Schubert, T.; Pusch, M.C.; Diermeier, S.; Benes, V.; Kremmer, E.; Imhof, A. Df31 Protein and snoRNAs Maintain Accessible Higher-Order Structures of Chromatin. *Mol. Cell* **2012**, *48*, 434–444. [[CrossRef](#)] [[PubMed](#)]
34. Fujita, R.; Yamamoto, T.; Arimura, Y.; Fujiwara, S.; Tachiwana, H.; Ichikawa, Y.; Sakata, Y.; Yang, L.; Maruyama, R.; Hamada, M.; et al. Nucleosome destabilization by nuclear non-coding RNAs. *Commun. Biol.* **2020**, *3*, 3–8. [[CrossRef](#)] [[PubMed](#)]
35. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R.; Chen, Y.; Lajoie, B.R.; Protacio, A.; Flynn, R.A.; Gupta, R.A.; et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **2011**, *472*, 120–126. [[CrossRef](#)]
36. Engreitz, J.M.; Pandya-Jones, A.; McDonel, P.; Shishkin, A.; Sirokman, K.; Surka, C.; Kadri, S.; Xing, J.; Goren, A.; Lander, E.S.; et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **2013**, *341*, 1237973. [[CrossRef](#)]
37. van Ouwwerkerk, A.F.; Bosada, F.M.; van Duijvenboden, K.; Hill, M.C.; Montefiori, L.E.; Scholman, K.T.; Liu, J.; de Vries, A.A.F.; Boukens, B.J.; Ellinor, P.T.; et al. Identification of atrial fibrillation associated genes and functional non-coding variants. *Nat. Commun.* **2019**, *10*, 4755. [[CrossRef](#)]
38. Drissen, R.; Palstra, R.-J.; Gillemans, N.; Splinter, E.; Grosveld, F.; Philipsen, S.; de Laat, W. The active spatial organization of the  $\beta$ -globin locus requires the transcription factor EKLF. *Genes Dev.* **2004**, *18*, 2485–2490. [[CrossRef](#)] [[PubMed](#)]
39. Huang, M.; Wang, H.; Hu, X.; Cao, X. lncRNA MALAT1 binds chromatin remodeling subunit BRG1 to epigenetically promote inflammation-related hepatocellular carcinoma progression. *Oncoimmunology* **2019**, *8*, e1518628. [[CrossRef](#)]
40. Lin, C.; Smith, E.R.; Takahashi, H.; Lai, K.C.; Martin-Brown, S.; Florens, L.; Washburn, M.P.; Conaway, J.W.; Conaway, R.C.; Shilatifard, A. AFF4, a Component of the ELL/P-TEFb Elongation Complex and a Shared Subunit of MLL Chimeras, Can Link Transcription Elongation to Leukemia. *Mol. Cell* **2010**, *37*, 429–437. [[CrossRef](#)]

41. Xie, X.; Lin, J.; Liu, J.; Huang, M.; Zhong, Y.; Liang, B.; Song, X.; Gu, S.; Chang, X.; Huang, D.; et al. A novel lncRNA NR4A1AS up-regulates orphan nuclear receptor NR4A1 expression by blocking UPF1-mediated mRNA destabilization in colorectal cancer. *Clin. Sci.* **2019**, *133*, 1457–1473. [[CrossRef](#)]
42. Leeman-Neill, R.J.; Song, D.; Bizarro, J.; Wacheul, L.; Rothschild, G.; Singh, S.; Yang, Y.; Sarode, A.Y.; Gollapalli, K.; Wu, L.; et al. Noncoding mutations cause super-enhancer retargeting resulting in protein synthesis dysregulation during B cell lymphoma progression. *Nat. Genet.* **2023**, *55*, 2160–2174. [[CrossRef](#)]
43. Chen, Q.; Zhu, C.; Jin, Y. The Oncogenic and Tumor Suppressive Functions of the Long Noncoding RNA MALAT1: An Emerging Controversy. *Front. Genet.* **2020**, *11*, 505991. [[CrossRef](#)]
44. Bahrami, S.; Drabløs, F. Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* **2016**, *62*, 37–49. [[CrossRef](#)] [[PubMed](#)]
45. Gorgoulis, V.G.; Vassiliou, L.V.F.; Karakaidos, P.; Zacharatos, P.; Kotsinas, A.; Liloglou, T.; Venere, M.; Ditullio, R.A., Jr.; Kastriakis, N.G.; Levy, B.; et al. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **2005**, *434*, 907–913. [[CrossRef](#)]
46. Guo, H.; Golczer, G.; Wittner, B.S.; Langenbucher, A.; Zachariah, M.; Dubash, T.D.; Hong, X.; Comaills, V.; Burr, R.; Ebright, R.Y.; et al. NR4A1 regulates expression of immediate early genes, suppressing replication stress in cancer. *Mol. Cell* **2021**, *81*, 4041–4058.e15. [[CrossRef](#)]
47. Perez-Pinera, P.; Kocak, D.D.; Vockley, C.M.; Adler, A.F.; Kabadi, A.M.; Polstein, L.R.; Thakore, P.I.; Glass, K.A.; Ousterout, D.G.; Leong, K.W.; et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **2013**, *10*, 973–976. [[CrossRef](#)] [[PubMed](#)]
48. Mars, J.; Sabourin-Felix, M.; Tremblay, M.; Moss, T. A deconvolution protocol for ChIP-seq reveals analogous enhancer structures on the mouse and human ribosomal RNA genes. *G3 Genes Genomes Genet.* **2018**, *8*, 303–314. [[CrossRef](#)] [[PubMed](#)]
49. Andrews, S. FASTQC: A Quality Control Tool for High Throughput Sequence Data. *Nucleic Acid Res.* **2010**, *41*, e108.
50. Ewels, P.; Magnusson, M.; Lundin, S.; Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047–3048. [[CrossRef](#)]
51. Bolger, A.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2015**, *30*, 2114. [[CrossRef](#)] [[PubMed](#)]
52. Martin, M. Cutadapt removes adapters from high-throughput sequencing reads. *EMBnet.J.* **2011**, *17*, 10–12. [[CrossRef](#)]
53. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)] [[PubMed](#)]
54. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
55. Liao, Y.; Smyth, G.; Shi, W. The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acid Res.* **2013**, *41*, e108. [[CrossRef](#)]
56. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)] [[PubMed](#)]
57. Zhu, A.; Ibrahim, J.G.; Love, M.I. Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **2019**, *35*, 2084–2092. [[CrossRef](#)] [[PubMed](#)]
58. Corces, M.R.; Trevino, A.E.; Hamilton, E.G.; Greenside, P.G.; Sinnott-Armstrong, N.A.; Vesuna, S.; Satpathy, A.T.; Rubin, A.J.; Montine, K.S.; Wu, B.; et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **2017**, *14*, 959–962. [[CrossRef](#)] [[PubMed](#)]
59. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
60. Ramírez, F.; Dündar, F.; Diehl, S.; Grüning, B.A.; Manke, T. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **2014**, *42*, 187–191. [[CrossRef](#)]
61. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [[CrossRef](#)] [[PubMed](#)]
62. McLean, C.Y.; Bristol, D.; Hiller, M.; Clarke, S.L.; Schaar, B.T.; Lowe, C.B.; Wenger, A.M.; Bejerano, G. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **2010**, *28*, 495–501. [[CrossRef](#)] [[PubMed](#)]
63. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **2016**, *44*, e71. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.