

Article

The Potential of NORAD–PUMILIO–RALGAPB Regulatory Axis as a Biomarker in Breast Cancer

Cristiane Sato Mara Muller ^{1,†}, Igor Samesima Giner ^{1,†}, Érika Pereira Zambalde ², Tamyres Mingorance Carvalho ¹, Enilze Maria de Souza Fonseca Ribeiro ¹, Jaqueline Carvalho de Oliveira ¹, Carolina Mathias ^{1,3} and Daniela Fiori Gradia ^{1,*} 

¹ Laboratory of Human Cytogenetics and Oncogenetics, Postgraduate Program in Genetics, Department of Genetics, Federal University of Paraná (UFPR), Curitiba 81531-980, Brazil

² Multidisciplinary Laboratory of Food and Health (LabMAS), School of Applied Sciences (FCA), University of Campinas (UNICAMP), Limeira 13484-350, Brazil

³ Laboratory for Applied Science and Technology in Health, Carlos Chagas Institute, Oswaldo Cruz Foundation (FIOCRUZ), Curitiba 81310-020, Brazil

* Correspondence: danielagrada@ufpr.br

† These authors contributed equally to this work.



Citation: Muller, C.S.M.; Giner, I.S.; Zambalde, É.P.; Carvalho, T.M.; Ribeiro, E.M.d.S.F.; Carvalho de Oliveira, J.; Mathias, C.; Gradia, D.F. The Potential of NORAD–PUMILIO–RALGAPB Regulatory Axis as a Biomarker in Breast Cancer. *Non-Coding RNA* **2022**, *8*, 76. <https://doi.org/10.3390/ncrna8060076>

Academic Editor: Luca Falzone

Received: 27 September 2022

Accepted: 4 November 2022

Published: 10 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/>).

Abstract: Introduction: Long non-coding RNAs (LncRNA) represent a heterogeneous family of RNAs that have emerged as regulators of various biological processes through their association with proteins in ribonucleoproteins complexes. The dynamic of these interactions can affect cell metabolism, including cancer development. Annually, breast cancer causes thousands of deaths worldwide, and searching for new biomarkers is pivotal for better diagnosis and treatment. **Methods:** Based on in silico prediction analysis, we focus on LncRNAs that have binding sites for PUMILIO, an RBP family involved in post-transcriptional regulation and associated with cancer progression. We compared the expression levels of these LncRNAs in breast cancer and non-tumor samples from the TCGA database. We analyzed the impact of overall and disease-free survival associated with the expression of the LncRNAs and co-expressed genes and targets of PUMILIO proteins. **Results:** Our results found NORAD as the most relevant LncRNA with a PUMILIO binding site in breast cancer, differently expressed between Luminal A and Basal subtypes. Additionally, NORAD was co-expressed in a Basal-like subtype (0.55) with the *RALGAPB* gene, a target gene of PUMILIO related to chromosome stability during cell division. **Conclusion:** These data suggest that this molecular axis may provide insights for developing novel therapeutic strategies for breast cancer.

Keywords: LncRNA; NORAD; PUMILIO; *RALGAPB*; breast cancer; biomarker; regulatory axis

1. Introduction

Except for skin cancers, breast cancer (BRCA) is the most common malignant disease among women. It is a heterogeneous disease that presents different pathological characteristics, and patients show distinct responses to treatments with varied survival along its course [1]. The molecular classification of BRCA is based on gene expression profiles for the stratification of patients, classifying them into four molecular subtypes: Luminal A and B, Her2-enriched, and Basal-like [1]. Luminal A (Lum A) patients show low-grade tumors and a good prognosis compared to Luminal B (Lum B), who have tumors with higher proliferation rates. Her2-enriched and Basal-like subtypes comprise patients with worse prognoses and higher-grade tumors [2]. The therapeutic management of breast cancer includes surgery, radiotherapy, and chemotherapy [2]. In addition, luminal subtypes respond to hormonal therapies, such as Tamoxifen, which benefits estrogen-positive patients [3]. Early-stage Her2-enriched patients may have a good outcome from the addition of trastuzumab along with chemotherapy, reducing the risk of relapse and death [4]. Lastly, patients with a mutated BRCA gene may benefit from the administration of PARP

inhibitors, associated with chemotherapy [5]. Due to this great diversity, the discovery of additional molecular markers is needed to understand the disease's outcome better, providing personalized therapies that improve breast cancer's survival rate [6].

Molecular complexes such as ribonucleoprotein particles (RNP) are key players in cell life, and these complexes' deregulation is often associated with genetic diseases, including cancer [7]. Proteins such as DNA methyltransferases (DNMTs), heterochromatin protein 1, MOF, MSL, DDP1, Polycomb group, and Trithorax group proteins are RBPs capable of binding to lncRNAs, and their levels fluctuate significantly in tumor samples. The lncRNA HOTAIR, one of the first known, is related to cancer, including breast cancer and hepatocellular carcinoma, and participates in chromatin modification complexes by interaction with the polycomb group protein PRC2, which is a histone methyltransferase, and with LSD1, which is a histone demethylase [8]. PUMILIO proteins are members of the PUF family of RNA-binding proteins (RBPs) that exhibit a sequence-specific and highly conserved recognition domain called the PUMILIO Homology Domain (Pum-HD) that interacts with the RNA [9,10]. The interaction occurs at the 3' end of target RNAs with an 8-nucleotide consensus sequence (UGUANAUA), known as the PUMILIO Response Element (PRE) [11]. PUMILIO proteins repress mRNA by promoting deadenylation of the poly-A tail through the CNOT complex. In addition, the PUMILIO, may also act antagonistically to the PABPC1 protein by repressing translation initiation, preventing the binding of the translation initiation factors eIF4E and eIF5B [11]. Expression of PUM in human tumorigenesis and characterization of PUM-mediated translation control directly in cancer cells has demonstrated its potential to regulate hundreds of transcripts, including many oncogenes and tumor suppressor genes [12].

Long non-coding RNAs (lncRNA) represent a heterogeneous family of RNAs that have emerged as regulators of various biological processes [13]. lncRNAs and RNA-binding proteins (RBP) can associate with ribonucleoprotein (RNP) complexes in both the cytoplasm and the nucleus [14]. They act through distinct mechanisms, controlling gene expression at multiple levels and are involved in physiological and pathological scenarios [15]. Increased expression levels of these lncRNAs can act as RBP decoys, sequestering these proteins and thus hindering their function in regulating their target genes [16]. In this study, we propose a new strategy for searching for potential new biomarkers in breast cancer focused on the regulatory axis formed by lncRNAs, PUMILIO proteins, and their targets in the breast cancer context.

2. Results

2.1. lncRNA with Sites for PUMILIO Are Differentially Expressed in BRCA

Of the RNAs containing PUMILIO binding sites identified in previous work [17,18], 47 were classified as lncRNAs (supplementary material). From these, by comparing linear models using the Limma package in R, we found 34 as differentially expressed ($p < 0.05$), between tumor and non-tumor samples, in BRCA (supplementary material). We then selected the lncRNAs that showed above the median— \log_2 FC cutoff expression in at least one BRCA subtype, resulting in 27 lncRNAs (Figure 1A,B).

Since lncRNAs have intrinsically low expression in human tissue, we selected those that showed detectable expression in at least 80% of the samples from the 1039 patients diagnosed with breast cancer in the Xena Browser platform (<https://tcga.xenahubs.net>, accessed on 16 August 2022). Of the twenty-seven differentially expressed lncRNAs with binding sites for PUMILIO, seven showed detectable expression: ASH1L-AS1 (ENSG00000235919), RHPN1-AS1 (ENSG00000254389), SERTAD4-AS1 (ENSG00000203706), NORAD (ENSG00000248360), LINC01011 (ENSG00000244041), LINC00504 (ENSG00000248360), and MALAT1 (ENSG00000251562).

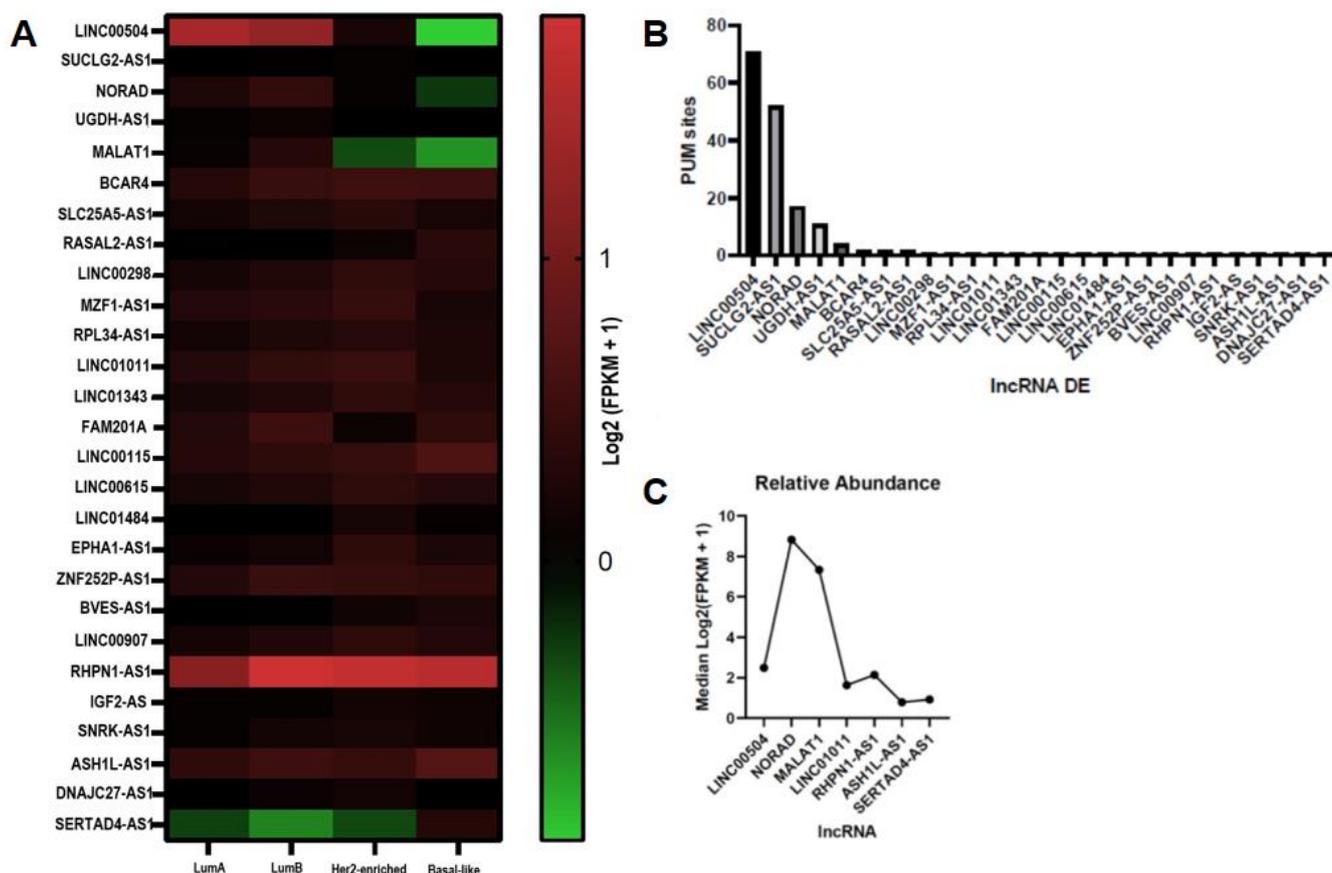


Figure 1. LncRNA with sites for PUMILIO differentially expressed in BRCA. Legend: (A) Heatmap representation of the activity of expression networks of 27 lncRNAs in breast cancer molecular subtypes. According to this metric, differentially expressed positive (above the median— $\log_2\text{FC}$ cutoff) ($p < 0.05$) are in red and differentially expressed negative (below the median— $\log_2\text{FC}$ cutoff) ($p < 0.05$) are in green. Samples are arranged in the columns of the heatmap and categorized according to the annotation bar on the side of the graph (-0.92 to 1.8). The TCGA BRCA cohort samples were classified according to molecular classification into Lum A ($n = 560$), Lum B ($n = 207$), Her2-enriched ($n = 82$), and Basal-like ($n = 190$). (B) Number of sites for PUMILIO in 27 differentially expressed lncRNAs in breast cancer (BRCA). (C). Relative abundance of seven lncRNA with detectable expression—median $\text{Log}_2(\text{FPKM} + 1)$ (Fragments Per Kilobase Million) in BRCA samples from the same TCGA BRCA cohort.

Next, we looked up the relative abundance of the lncRNAs using the same data available in the Xena Browser platform (<https://tcga.xenahubs.net>, accessed on 16 August 2022). LINC00504 has the highest number of sites for PUMILIO (71). Still, its relative abundance (FPKM = 2.491192) in the samples is much lower when compared to NORAD (FPKM = 8.832009) and MALAT1 (FPKM = 7.336683), which contain seventeen and four PRE (PUMILIO Response Element), respectively. The lncRNAs ASH1L-AS1, RHPN1-AS1, SERTAD4-AS1, and LINC01011 have only one PRE and low expression. Considering that the number of PRE and the relative abundance of each lncRNA can increase or decrease the ability to sequester PUMILIO, we assume that the lncRNAs NORAD, MALAT1, and LINC00504 have a higher potential to influence the regulatory activity of PUMILIO (Figure 1C).

2.2. LncRNA NORAD Shows Co-Expression Networks in All Four BRCA Subtypes

For the construction of the co-expression networks, we selected the lncRNAs NORAD, MALAT1, and LINC00504, which showed a higher abundance or more sites for PUMILIO than the other lncRNAs with detectable expression. The analysis did not point to any PUMILIO target genes significantly co-expressed with the MALAT1 or LINC00504 in

breast cancer samples. However, we found that PUMILIO targeted genes co-expressed for NORAD in all four subtypes. The *RALGAPB* gene appears co-expressed with NORAD in all four BRCA subtypes and shows the highest mutual information value in the Basal-like subtype (0.55) ($p < 0.00001$) (Figure 2).

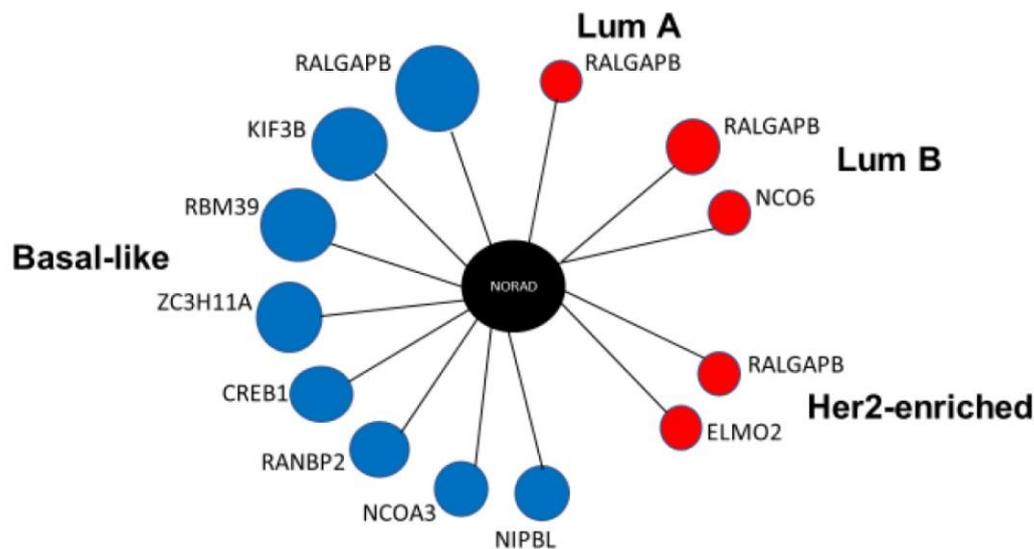


Figure 2. Co-expression network for NORAD lncRNA in BRCA subtypes. Legend: Network representation of NORAD extended in genes with higher values of mutual information in BRCA subtypes. Red circles refer to genes with high expression and blue to low expression compared to non-tumor samples in each BRCA subtype and all show a direct correlation with NORAD. The size of the dot refers to the mutual information value, whereas large dots indicate greater mutual information.

The lncRNA NORAD shows negative differential expression (are less expressed) in the Basal-like subtype, so we selected the target genes that were also under-expressed (below the median— \log_2FC cutoff) ($p < 0.05$).

2.3. NORAD and RALGAPB Are Differentially Expressed in Lum A and Basal-Like Subtypes Compared to Non-Tumoral Samples

NORAD is differentially expressed and inversely modulated between Lum A and Basal-like. While Lum A has a better prognosis with low-grade tumors, Basal-like is characterized by greater invasive potential [1]. In this context, our goal was to seek explanations that help to understand these differences. Our analyses showed higher expression of NORAD ($\text{Log}_2FC = 0.2541$, $p = 2.31 \times 10^{-5}$) and *RALGAPB* ($\text{Log}_2FC = 0.3526$, $p = 5.96 \times 10^{-11}$) in Lum A when compared to non-tumor samples and lower expression in Basal-like—NORAD ($\text{Log}_2FC = -0.2506$, $p = 4.664 \times 10^{-4}$) and *RALGAPB* ($\text{Log}_2FC = -0.1533$, $p = 1.54 \times 10^{-2}$) when compared to non-tumor samples. When comparing Lum A with Basal-like subtypes, both NORAD and *RALGAPB* showed higher expression in Lum A ($\text{Log}_2FC = 0.5048$, $p = 1.19 \times 10^{-16}$) ($\text{Log}_2FC = 0.5059$, $p = 6.56 \times 10^{-24}$), respectively. In the context of tumor phenotype, NORAD may act differently between the two subtypes and may present a tumor suppressor or oncogenic role [19,20]. Its higher expression in Lum A gives it a greater potential to sequester PUMILIO. Consequently, *RALGAPB*-increased levels may be a consequence of this regulatory axis. The opposite effect could be assumed between NORAD and *RALGAPB* in Basal-like samples (Figure 3A,B).

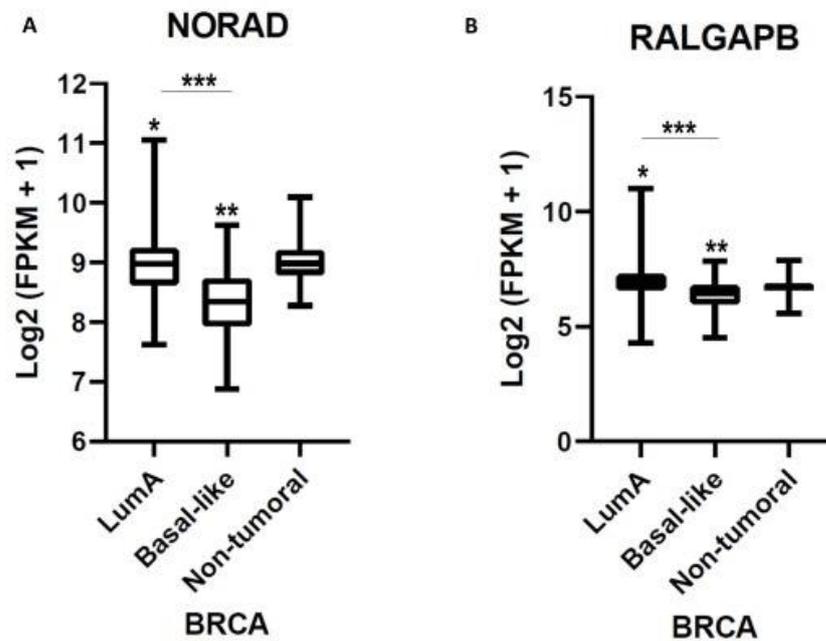


Figure 3. Differential expression of NORAD lncRNA and *RALGAPB* for Lum A and Basal-like subtypes. Legend: (A) Differential expression of NORAD lncRNA in comparing Lum A, Basal-like subtypes, and non-tumor samples. (NORAD) Lum A—Log₂FC = 0.254, * $p = 1.19 \times 10^{-16}$, Basal-like—Log₂FC = -0.2506, ** $p = 4.66 \times 10^{-4}$, Lum A versus Basal-like—Log₂FC = 0.5048, *** $p = 1.19 \times 10^{-16}$. (B) Differential expression of *RALGAPB* in comparing Lum A, Basal-like subtypes, and non-tumor samples. (*RALGAPB*) Lum A—Log₂FC = 0.3526, * $p = 5.96 \times 10^{-11}$, Basal-like—Log₂FC = -0.1533, ** $p = 1.542 \times 10^{-2}$, Lum A versus Basal-like—Log₂FC = 0.5059, *** $p = 6.56 \times 10^{-24}$.

2.4. NORAD and RALGAPB Are up Expressed in Breast Cancer and Related to OS in Patients

We performed Kaplan–Meier and Cox proportional hazards regression analyses to verify if the signatures of NORAD and *RALGAPB* could predict the overall survival (OS) in breast cancer patients.

In Lum A and Basal-like patients, NORAD expression was not found to be related to OS (Figure 4A,B). However, the up expression of *RALGAPB* was associated with a poor OS in Lum A patients ($p = 0.042$) (Figure 4C,D). In Lum A and Basal-like patients, the combined expression of NORAD and *RALGAPB* was also not found to be related to OS (Figure 4E,F).

A univariate Cox regression analysis shows, for both NORAD and *RALGAPB*, an increased expression was related to a worse prognosis and a poor OS in Lum A patients. A multivariate Cox regression analysis confirms these results for NORAD regarding patients’ age (Table 1).

Table 1. Univariate and Multivariate Cox regression analyses for NORAD and *RALGAPB* in Lum A patients.

Variables Lum A		Univariate Model			Variables Lum A		Multivariate Model		
	HR	CI	<i>p</i> -value		HR	CI	<i>p</i> -value		
NORAD	2	1.2–3.4	0.01	NORAD	2.73	1.07–6.9	0.035		
<i>RALGAPB</i>	1.6	0.95–2.8	0.079	<i>RALGAPB</i>	0.72	0.29–1.8	0.475		
				AGE	1.04	1.02–1.1	<0.001		

Legend: HR: Hazard Ratio; CI 95%: Confidence Interval of 95%.

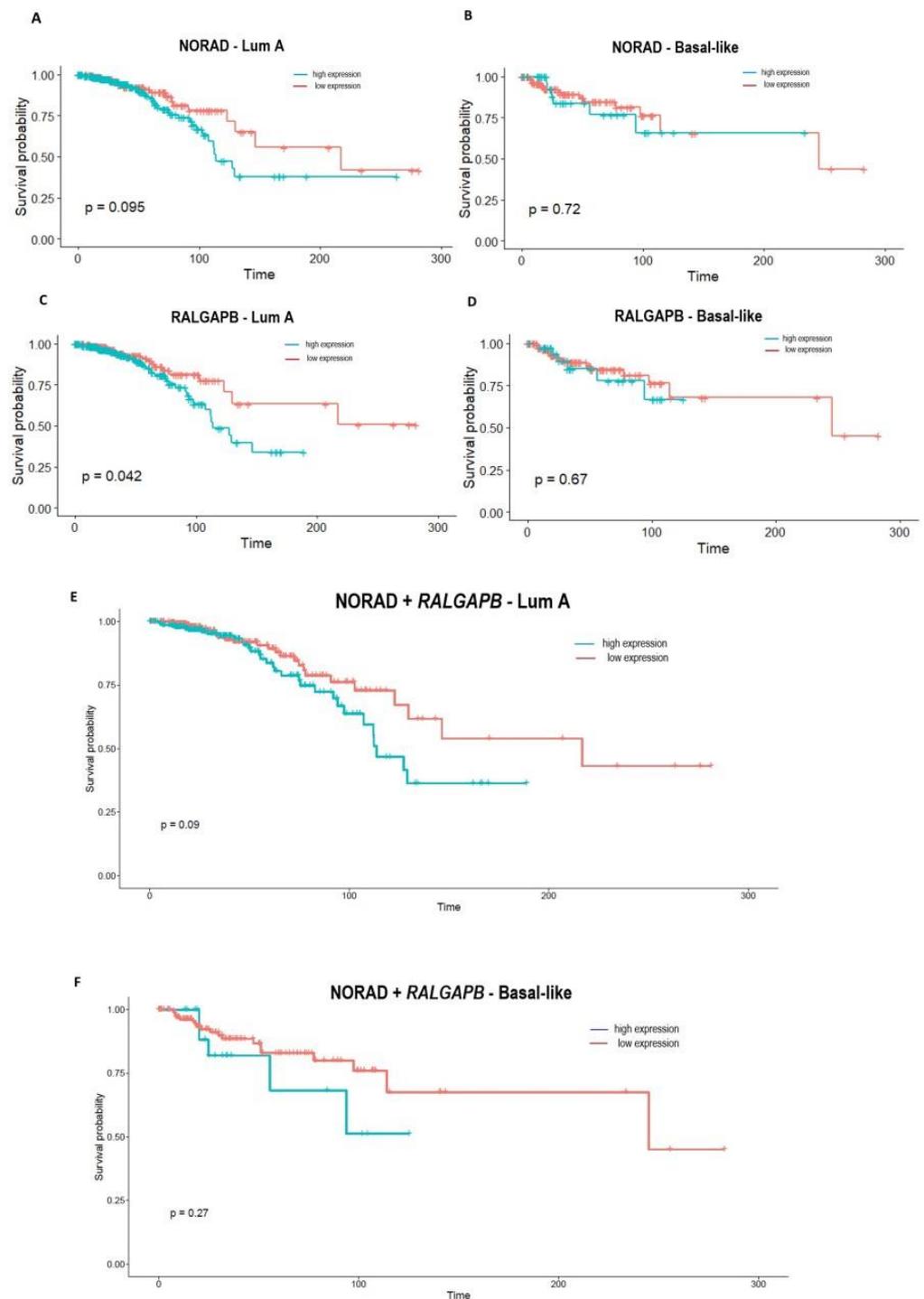


Figure 4. Overall survival analyses for Lum A and Basal-like patients related to lncRNA NORAD and RALGAPB expression. **Legend:** (A) Survival analysis considering the median value of NORAD in Lum A (High expression ($n = 304$), Low expression ($n = 185$)) and (B) Basal-like patients (High expression ($n = 37$), Low expression ($n = 131$)). (C) Survival analysis considering the median value of RALGAPB in Lum A (High expression ($n = 282$), Low expression ($n = 207$)), and (D) Basal-like patients (High expression ($n = 38$), Low expression ($n = 130$)). (E) Survival combined panel analysis with NORAD and RALGAPB in Lum A patients. (F) Survival combined panel analysis with NORAD and RALGAPB in Basal-like patients. Red and blue lines refer to patients with low and high expression, respectively. A p -value < 0.05 was considered significant.

2.5. NORAD and RALGAPB Showed a Biomarker Potential, Alone or Combined in Panel

We performed an ROC curve analysis to determine the biomarker potential of NORAD and RALGAPB. We found a potential for NORAD to discriminate Basal-like tumor samples ($n = 190$) from non-tumor samples (NT) ($n = 113$) with high sensibility (88.5%) and specificity (72.11%). Despite the limited specificity, the RALGAPB gene showed a high sensitivity to identify both Lum A and Basal-like subtypes compared to non-tumor samples, suggesting a modest potential as a biomarker. The combined panel analysis with two molecules increased the sensitivity and specificity of RALGAPB to discriminate both Lum A and Basal-like from NT samples. Particularly, NORAD and RALGAPB in the panel may distinguish Basal-like from NT samples with 91.15% specificity. Additionally, NORAD and RALGAPB combined showed a potential to discriminate Lum A samples from Basal-like (68.95% sensitivity and 76.61% specificity) (Table 2).

Table 2. Data about receiver operating characteristic (ROC) curves to investigate the diagnostic potential of RNAs in Lum A and Basal-like samples.

Comparison	RNA	AUC	Sensitivity	Specificity	<i>p</i> -Value
<i>LA</i> × <i>NT</i>	NORAD	0.5369	86.73	29.64	=0.2149
	<i>RALGAPB</i>	0.6574	88.50	46.96	<0.0001
	NORAD + <i>RALGAPB</i>	0.7242	78.79	56.79	<0.0001
<i>BL</i> × <i>NT</i>	NORAD	0.8496	88.50	72.11	<0.0001
	<i>RALGAPB</i>	0.6659	85.84	51.58	<0.0001
	NORAD + <i>RALGAPB</i>	0.8588	70.53	91.15	<0.0001
<i>LA</i> × <i>BL</i>	NORAD	0.7953	69.47	76.07	<0.0001
	<i>RALGAPB</i>	0.7555	81.58	57.86	<0.0001
	NORAD + <i>RALGAPB</i>	0.7980	68.95	76.61	<0.0001

Legend: Sensitivity and specificity are presented as a percentage (%). In bold and underlined are the highest AUC value for the group comparison. NT = non-tumor; LA = luminal A; BL = basal-like.

3. Materials and Methods

3.1. TCGA Analysis

The expression data for the lncRNAs were extracted using the GDCRNATools package [21] and the pipeline contained in the Limma package [22] from The Cancer Genome Atlas (TCGA) database through the Xena Browser platform (<https://tcga.xenahubs.net> accessed on 16 August 2022), following the program's ethical, legal, and policy criteria. An RNAseq-Illumina HiSeq table was used to extract the differential expression data as $\log_2(\text{FPKM} + 1)$ for lncRNA and mRNA. The clinical data of the patients were extracted using the cBioPortal platform. The BRCA patient data were analyzed into molecular subtypes: Luminal A (Lum A) ($n = 560$), Luminal B (Lum B) ($n = 207$), HER2-enriched ($n = 82$), and Basal-like ($n = 190$) [19].

3.2. LncRNAs Selection

The selection of the long non-coding RNAs with PUMILIO binding sites (UGUANAUA) was made using data from [17,18]. Smialek and collaborators [18] used RIP-Seq and RNA-Seq to identify mRNA regulated by PUM1 and PUM2 in the cell line Tcam-2, a human male germ cell model. Bohn and collaborators [17] validated a set of RNAs with PUMILIO binding sites in human HEK293 cells using RNA-Seq analysis, statistical testing, and experimental validation. We identified lncRNAs with sites for PUMILIO and with positive expression (above the median— $\log_2\text{FC}$ cutoff) ($p < 0.05$), when comparing tumor versus non-tumor tissues in at least one BRCA subtype. Considering the low intrinsic expression of lncRNAs in general, we used one more filter, selecting only lncRNAs expressed in at least 80% of tumor samples.

3.3. Co-Expression Regulatory Network

Co-expression networks were generated using the RTN package, which is designed for the reconstruction of transcriptional regulatory networks (TRNs) and analysis of regulons using mutual information (MI) metrics [23]. We tested the association between a given lncRNA and all potential targets using transcriptome data and applied 1000 permutations considering the significance level $p < 0.00001$. The same strategy was previously used to identify potential lncRNA co-expression networks in Lum A breast cancer tumors [21].

Finally, of all the genes selected for each lncRNA with a site for PUMILIO, only those that were also targets of PUMILIO according to [17,18] data and have a positive expression (above the median— \log_2 FC cutoff) ($p < 0.05$) when comparing tumor versus non-tumor tissue, were selected. The lncRNA-target gene pairs with co-expression values of >0.15 were chosen, interpreting a correlation above 10% between lncRNA and the target. Each target was analyzed for its expression in tumor tissue compared to normal tissue using only TCGA data.

3.4. Statistical Analysis

We analyzed univariate and combined analysis for disease-free (DFS) and overall survival (OS), based on disease-free days and the presence or absence of death events, respectively, for lncRNA and genes co-expressed using the R platform with the survival and survminer packages. We compared the low or high expression of each lncRNA and mRNA targeted by PUMILIO using long-rank (Mantel–Cox tests) and Kaplan–Meier curves. Patient survival was analyzed according to each gene expression and to the pairwise expression in each patient. A $p < 0.05$ was considered statistically significant.

To verify whether the signature of the two genes, NORAD and RALGAPB, can predict the survival of breast cancer patients independently of age, we applied multivariate Cox regression analyses.

ROC (receiver operating characteristic) curves were performed using 1039 patients (Lum A $n = 560$, Lum B $n = 207$, Her2-enriched $n = 82$, and Basal-like $n = 190$) and 113 non-tumor samples. Combined ROC curves were performed through binary logistic regression analysis using IBM SPSS Statistics 26.0 (IBM SPSS Statistics Inc., Armonk, NY, USA). Differential expression data of lncRNA and PUMILIO genes of the BRCA cohort were used.

The cutoff values of the ROC curves, sensitivity, and specificity were determined by the Youden index. Analyses were performed using GraphPad Prism v.9.0.0 (GraphPad Software Inc., San Diego, CA, USA).

4. Discussion

From the RNAs containing PUMILIO binding sites identified in previous work [17,18], we identify 27 lncRNA with sites for PUMILIO positively regulated (above the median— \log_2 FC cutoff) ($p < 0.05$) in the BRCA samples. Of these, the NORAD lncRNA showed the highest relative abundance in the cellular environment (FPKM = 8.832009) and appeared co-expressed with the RALGAPB gene in all four breast cancer subtypes: Lum A, Lum B, Her2-enriched, and Basal-like, with the highest mutual information value in the Basal-like subtype (0.55) ($p < 0.00001$).

Taking into account the heterogeneity among BRCA subtypes, in other words, between Lum A and Basal-like subtypes, Lum A has a better prognosis, characterized by low-grade tumors, while Basal-like has a worse prognosis and is characterized by tumors with greater invasive potential [1]. Thus, we verified the expression of both genes in these subtypes.

Our analyses showed higher expression of NORAD ($\text{Log}_2\text{FC} = 0.2541$, $p = 2.31 \times 10^{-5}$) and RALGAPB ($\text{Log}_2\text{FC} = 0.3526$, $p = 5.96 \times 10^{-11}$) in Lum A when compared to non-tumor samples and lower expression in Basal-like-NORAD ($\text{Log}_2\text{FC} = -0.2506$, $p = 4.664 \times 10^{-4}$) and RALGAPB ($\text{Log}_2\text{FC} = -0.1533$, $p = 1.54 \times 10^{-2}$) when compared to non-tumor samples. When comparing Lum A with Basal-like subtypes, both NORAD and RALGAPB showed higher expression in Lum A ($\text{Log}_2\text{FC} = 0.5048$, $p = 1.19 \times 10^{-16}$; $\text{Log}_2\text{FC} = 0.5059$,

$p = 6.56 \times 10^{-24}$, respectively). In the context of tumor phenotypes, NORAD may be acting in different ways between the two subtypes and may present a tumor suppressor or oncogenic role [19,24].

Our work suggests that the higher expression of NORAD in Lum A gives it a higher potential to sequester PUMILIO. Consequently, the increased levels of *RALGAPB* may be a consequence of this regulatory axis. The opposite effect could be assumed between NORAD and *RALGAPB* in Basal-like.

PUMILIO (PUM) proteins are members of an evolutionarily conserved family of RBPs known as PUF (Pumilio-Fem-3-binding factor) [25]. The characterizing feature of all PUF proteins is the presence of the C-terminal RNA-binding domain known as the PUMILIO homology domain (PUM-HD), formed by eight imperfect tandem repeats of 36 amino acids each [12]. PUMILIO acts in the repression of translation of messenger RNAs with which it interacts through domains in its structure capable of recognizing specific nucleotide sequences in the 3'UTR region of the target transcript. Translation repression can occur through the recruitment of factors that will prevent translation or through the degradation of the target mRNA by the recruitment of deadenylase complexes [25]. PUMILIO is also an important regulator of DICER1, indirectly affecting miRNA processing in DICER1 syndrome, a cancer predisposition disorder [26].

Loss of PUMILIO-mediated fine regulation could provide ideal conditions for tumor development, as cancer cells would benefit from the proteomic plasticity generated to meet their demands for growth and metastasis [27].

NORAD is highly conserved and an abundant lncRNA, with expression levels of approximately 500 to 1000 copies per cell. This lncRNA maintains genomic stability by sequestering PUMILIO proteins, preventing them from repressing target mRNA stability and translation. In the absence of NORAD, the hyperactivity of PUMILIO represses mitotic, DNA repair, and replication factors [13].

Similar to some other lncRNAs, such as XIST and FIRRE [28], the NORAD sequence contains repeat elements—12 NORAD repeat units (NRUs)—recognizable and sequence-like, which likely originated by tandem duplication in mammalian emergence and still share substantial sequence homology. Most NRUs contain one or two binding sites for the two homologs of PUMILIO in mammals: PUM1 and PUM2. NORAD contains at least 17 of these PUM recognition elements (PREs), which have the consensus sequence UGURNAUA (R = A/G). This structural domain organization facilitates its function on multiple levels [29]. NORAD interacts with the PUMILIO proteins (PUM1/PUM2), controlling their availability in the cytoplasm. Recently, it was shown that NORAD promotes the condensation and sequestration of PUM proteins into NP (NORAD–PUMILIO) bodies [29].

Deletion of NORAD in cell lines results in increased repression of genes carrying PREs in their 3'UTR regions, and overexpression has an inverse effect [13]. Since these proteins regulate a wide class of transcripts and many of them are involved with genome protection and cell cycle control, changes in NORAD levels could consequently unbalance the orchestration of essential mechanisms for cellular homeostasis [30].

We know that NORAD-regulated PUM targets are enriched in genes involved in cell division [30], and depletion of this lncRNA results in defects in mitosis and the accumulation of cells with instability in chromosome number [13].

Our analyses showed the *RALGAPB* gene co-expressed with NORAD in all four BRCA subtypes with the highest mutual information value in the Basal-like subtype (0.55) ($p < 0.00001$).

RalGAPs (Ral GTPase activating proteins) are multi-subunit complexes comprising a catalytic alpha (α) subunit and beta regulatory (β) subunits. They act as GTPase activators for the small GTPases RALA and RALB. The function and stability of the complex require the association of the α and β subunits.

During mitosis, *RALGAPB* shows a dramatic cell redistribution, shifting from the nucleus and Golgi complex during the prophase to the mitotic spindle and intercellular bridge at cytokinesis [31]. Since the *RALGAPB* is a target of PUMILIO, increased or

decreased expression of lncRNA NORAD interferes with the efficiency of interaction with PUMILIO, and should lead to increased or decreased *RALGAPB* gene expression in BRCA.

A study evaluating the role of *RALGAPB* during the process of cell mitosis showed that its depletion causes chromosomal misalignment and decreases the amount of mitotic cyclin B1, disturbing the transition from metaphase to anaphase. The over-expression of *RALGAPB* interferes with cell division, leading to binucleation and multinucleation, and cell death [31].

A study in oral squamous cell carcinoma [10] showed that *RALGAPB* silencing increased the level of RalA activation, promoting HSC-2 cell migration and invasion in vitro, a profile characteristic of metastatic cells. In HeLa cells, the depletion of the *RALGAPB* protein in the mitotic spindle and intercellular bridge caused a considerable increase in the number of metaphase cells with aberrant chromosomal misalignment in the equatorial plane [31].

Applying Kaplan–Meier analyses, we found that up expression of *RALGAPB* was related to the poor OS in patients with Lum A breast cancer ($p = 0.042$). The signature of this gene and the NORAD can predict a poor prognosis in Lum A. The results of the univariate and multivariate Cox analysis for NORAD and *RALGAPB* again point out that the higher expression of NORAD is related to a worse prognosis for Lum A patients compared to the *RALGAPB* gene. Furthermore, we also verified the independence of these results considering the patients' age. The multivariate Cox analysis showed that the poor OS found in Lum A patients regarding NORAD expression is independent of age. Moreover, the multivariate analyses also do not confirm the independence of the two genes in predicting prognosis, which may support our hypothesis that a NORAD–PUMILIO–*RALGAPB* regulatory axis is involved with the tumor phenotype and is impacting patient survival.

Finally, the results of the ROC curve analyses showed that the differential expression of NORAD has an excellent potential to discriminate Basal-like tumor samples ($n = 190$) from non-tumor samples ($n = 113$) (88.5% sensitivity, 72.11% specificity, AUC = 0.8496, CI: 0.80–0.89, $p < 0.001$), suggesting its potential as a biomarker in Basal-like patients. Additionally, the combined panel with NORAD and *RALGAPB* increased these molecules' biomarker potential, especially to distinguish Basal-like from NT samples. Nonetheless, the potential to discriminate Lum A and Basal-like in the panel is also achievable.

The signature of NORAD and *RALGAPB* in Lum A and Basal-like subtypes can predict a worse or better prognosis according to the survival of patients, and also has the potential to discriminate tumor and non-tumor groups. Functional analyses must be conducted to validate our findings.

5. Conclusions

In this study, we propose a new strategy for searching for potential new biomarkers in breast cancer focused on the regulatory axis formed by lncRNAs, PUMILIO proteins, and their targets in the breast cancer context. Our results pointed out NORAD as the most relevant lncRNA with PUMILIO binding site in breast cancer, differently expressed between Lum A and Basal-like subtypes and co-expressed significantly in Basal-like (0.55) with the *RALGAPB* gene, a target gene of PUMILIO, related to chromosome stability during cell division. This work makes *RALGAPB* a new player in genomic instability promoted by NORAD dysregulation in breast cancer.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ncrna8060076/s1>.

Author Contributions: C.S.M.M., D.F.G., J.C.d.O. and E.M.d.S.F.R. designed the study. C.S.M.M. and I.S.G. performed the in silico analyses. D.F.G., J.C.d.O. and C.M. supervised the analyses. C.S.M.M. drafted the manuscript, supervised by D.F.G., É.P.Z., T.M.C. and E.M.d.S.F.R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

BRCA	breast cancer
DFS	disease-free survival
FC	fold change
FPKM	Fragments Per Kilobase Million
HR	Hazard Ratio
Lum A	luminal A
Lum B	luminal B
LncRNA	Long non-coding RNAs
NRU	NORAD repeat unit
OS	overall survival
RBP	RNA-binding protein
RNP	ribonucleoprotein particle
ROC	receiver operating characteristic
PRE	PUM recognition element
TCGA	The Cancer Genome Atlas
TRN	transcriptional regulatory network

References

- Eroles, P.; Bosch, A.; Pérez-Fidalgo, J.A.; Lluch, A. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treat. Rev.* **2012**, *38*, 698–707. [[CrossRef](#)] [[PubMed](#)]
- Naito, Y.; Urasaki, T. Precision medicine in breast cancer. *Chin. Clin. Oncol.* **2018**, *7*, 29. [[CrossRef](#)] [[PubMed](#)]
- Fisusi, F.A.; Akala, E.O. Drug Combinations in Breast Cancer Therapy. *Pharm Nanotechnol* **2019**, *7*, 3–23. [[CrossRef](#)] [[PubMed](#)]
- Bradley, R.; Braybrooke, J.; Gray, R.; Hills, R.; Liu, Z.; Peto, R.; Davies, L.; Dodwell, D.; McGale, P.; Pan, H. Trastuzumab for early-stage, HER2-positive breast cancer: A meta-analysis of 13,864 women in seven randomised trials. *Lancet Oncol.* **2021**, *22*, 1139–1150. [[CrossRef](#)]
- Gonçalves, A.; Bertucci, A.; Bertucci, F. PARP inhibitors in the treatment of early breast cancer: The step beyond? *Cancers* **2020**, *12*, 1378. [[CrossRef](#)]
- Andre, F.; Ismaila, N.; Henry, N.L.; Somerfield, M.R.; Bast, R.C.; Barlow, W.; Collyar, D.E.; Hammond, M.E.; Kuderer, N.M.; Liu, M.C.; et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women with Early-Stage Invasive Breast Cancer: ASCO Clinical Practice Guideline Update—Integration of Results From TAILORx. *J. Clin. Oncol.* **2019**, *37*, 1956–1964. [[CrossRef](#)]
- Lavalée, M.; Curdy, N.; Laurent, C.; Fournié, J.-J.; Franchini, D.-M. Cancer cell adaptability: Turning ribonucleoprotein granules into targets. *Trends Cancer* **2021**, *7*, 902–915. [[CrossRef](#)]
- Ferrè, F.; Colantoni, A.; Helmer-Citterich, M. Revealing protein-lncRNA interaction. *Brief Bioinform* **2016**, *17*, 106–116. [[CrossRef](#)]
- Qian, X.; Zhao, J.; Yeung, P.Y.; Zhang, Q.C.; Kwok, C.K. Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. *Trends Biochem. Sci.* **2019**, *44*, 33–52. [[CrossRef](#)]
- Dai, H.; Shen, K.; Yang, Y.; Su, X.; Luo, Y.; Jiang, Y.; Shuai, L.; Zheng, P.; Chen, Z.; Bie, P. PUM1 knockdown prevents tumor progression by activating the PERK/eIF2/ATF4 signaling pathway in pancreatic adenocarcinoma cells. *Cell Death Dis.* **2019**, *10*, 595. [[CrossRef](#)]
- Goldstrohm, A.C.; Hall, T.M.T.; McKenney, K.M. Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. *Trends Genet.* **2018**, *34*, 972–990. [[CrossRef](#)] [[PubMed](#)]
- Galgano, A.; Forrer, M.; Jaskiewicz, L.; Kanitz, A.; Zavolan, M.; Gerber, A.P. Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. *PLoS ONE* **2008**, *3*, e3164. [[CrossRef](#)] [[PubMed](#)]
- Lee, S.; Kopp, F.; Chang, T.-C.; Sataluri, A.; Chen, B.; Sivakumar, S.; Yu, H.; Xie, Y.; Mendell, J.T. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **2015**, *164*, 69–80. [[CrossRef](#)] [[PubMed](#)]
- Statello, L.; Guo, C.-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](#)]

15. Mathias, C.; Groeneveld, C.; Trefflich, S.; Zambalde, E.; Lima, R.; Urban, C.; Prado, K.; Ribeiro, E.; Castro, M.; Gradia, D.; et al. Novel lncRNAs Co-Expression Networks Identifies LINC00504 with Oncogenic Role in Luminal A Breast Cancer Cells. *Int. J. Mol. Sci.* **2021**, *22*, 2420. [[CrossRef](#)]
16. Wang, X.; Zhang, X.; Dang, Y.; Li, D.; Lu, G.; Chan, W.Y.; Leung, P.C.K.; Zhao, S.; Qin, Y.; Chen, Z. Long noncoding RNA HCP5 participates in premature ovarian insufficiency by transcriptionally regulating MSH5 and DNA damage repair via YB1. *Nucleic Acids Res.* **2020**, *48*, 4480–4491. [[CrossRef](#)]
17. A Bohn, J.; Van Etten, J.L.; Schagat, T.L.; Bowman, B.M.; McEachin, R.C.; Freddolino, P.L.; Goldstrohm, A.C. Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins. *Nucleic Acids Res.* **2017**, *46*, 362–386. [[CrossRef](#)]
18. Smialek, M.J.; Ilaslan, E.; Sajek, M.P.; Swiercz, A.; Janecki, D.M.; Kusz-Zamelczyk, K.; Wozniak, T.; Kotecki, M.; Handschuh, L.; Figlerowicz, M.; et al. Characterization of RNP Networks of PUM1 and PUM2 Post-Transcriptional Regulators in TCam-2 Cells, a Human Male Germ Cell Model. *Cells* **2020**, *9*, 984. [[CrossRef](#)]
19. Zhao, W.; Wang, L.; Xu, F. lncRNA NORAD stimulates proliferation and migration of renal cancer via activating the miR-144-3p/MYC axis. *Eur. Rev. Med. Pharm. Sci.* **2020**, *24*, 10426–10432.
20. Shi, P.; Zhang, J.; Li, X.; Li, W.; Li, H.; Fu, P. Long non-coding RNA NORAD inhibition upregulates microRNA-323a-3p to suppress tumorigenesis and development of breast cancer through the PUM1/eIF2 axis. *Cell Cycle* **2021**, *20*, 1295–1307. [[CrossRef](#)]
21. Li, R.; Qu, H.; Wang, S.; Wei, J.; Zhang, L.; Ma, R.; Lu, J.; Zhu, J.; Zhong, W.-D.; Jia, Z. GDCRNATools: An R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics* **2018**, *34*, 2515–2517. [[CrossRef](#)] [[PubMed](#)]
22. Ritchie, M.E.; Belinda, P.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)] [[PubMed](#)]
23. Castro, M.A.; de Santiago, I.; Campbell, T.M.; Vaughn, C.; Hickey, T.E.; Ross, E.D.; Tilley, W.; Markowitz, F.; Ponder, B.A.; Meyer, K.B. Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nat. Genet.* **2015**, *48*, 12–21. [[CrossRef](#)] [[PubMed](#)]
24. Liu, W.; Zhou, X.; Li, Y.; Jiang, H.; Chen, A. Long non-coding rna norad inhibits breast cancer cell proliferation and metastasis by regulating mir-155-5p/ socs1 axis. *J. Breast Cancer* **2021**, *24*, 330–343. [[CrossRef](#)]
25. Spassov, D.S.; Jurecic, R. The PUF Family of RNA-binding Proteins: Does Evolutionarily Conserved Structure Equal Conserved Function? *IUBMB Life* **2003**, *55*, 359–366. [[CrossRef](#)]
26. Rajasekaran, S.; Khan, E.; Ching, S.R.; Khan, M.; Siddiqui, J.K.; Gradia, D.F.; Lin, C.; Bouley, S.J.; Mercadante, D.; Manning, A.L.; et al. PUMILIO competes with AUF1 to control DICER1 RNA levels and miRNA processing. *Nucleic Acids Res.* **2022**, *50*, 7048–7066. [[CrossRef](#)]
27. Wang, Y.; Chen, D.; Qian, H.; Tsai, Y.S.; Shao, S.; Liu, Q.; Dominguez, D.; Wang, Z. The Splicing Factor RBM4 Controls Apoptosis, Proliferation, and Migration to Suppress Tumor Progression. *Cancer Cell.* **2014**, *26*, 374–389. [[CrossRef](#)]
28. Hacisuleyman, E.; Shukla, C.J.; Weiner, C.L.; Rinn, J.L. Function and evolution of local repeats in the Firre locus. *Nat. Commun.* **2016**, *7*, 1–12. [[CrossRef](#)]
29. Elguindy, M.M.; Mendell, J.T. NORAD-induced Pumilio phase separation is required for genome stability. *Semin. Cell Dev. Biol.* **2021**, *595*, 303–308. [[CrossRef](#)]
30. Kedde, M.; Van Kouwenhove, M.; Zwart, W.; Oude Vrielink, J.A.F.; Elkon, R.; Agami, R. A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* **2010**, *12*, 1014–1020. [[CrossRef](#)]
31. Personnic, N.; Lakisic, G.; Gouin, E.; Rousseau, A.; Gautreau, A.; Cossart, P.; Bierne, H. A role for Ral GTPase-activating protein subunit β in mitotic regulation. *FEBS J.* **2014**, *281*, 2977–2989. [[CrossRef](#)] [[PubMed](#)]