



Article

Small Non-Coding RNAs and Their Role in Locoregional Metastasis and Outcomes in Early-Stage Breast Cancer Patients

Daniel Escuin ^{1,*} , Olga Bell ¹, Bárbara García-Valdecasas ^{1,2}, Montserrat Clos ^{1,2}, Itziar Larrañaga ^{1,2}, Laura López-Vilaró ^{1,2}, Josefina Mora ², Marta Andrés ^{1,2} , Cristina Arqueros ^{1,2} and Agustí Barnadas ^{1,2,3,4}

¹ Institut de Recerca Sant Pau (IR Sant Pau), 08041 Barcelona, Spain; obell@santpau.cat (O.B.); bgvaldecasas@santpau.cat (B.G.-V.); mclos@santpau.cat (M.C.); ilarranaga@santpau.cat (I.L.); llopezv@santpau.cat (L.L.-V.); mandresg@santpau.cat (M.A.); carqueros@santpau.cat (C.A.); abarnadasm@santpau.cat (A.B.)

² Hospital de la Santa Creu i Sant Pau, 08041 Barcelona, Spain; jmora@santpau.cat

³ School of Medicine, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Spain

⁴ Centro de Investigación Biomédica en Red Cáncer (CIBERONC), 28029 Madrid, Spain

* Correspondence: descuin@santpau.cat

Abstract: Deregulation of small non-coding RNAs (sncRNAs) has been associated with the onset of metastasis. We evaluated the expression of sncRNAs in patients with early-stage breast cancer, performing RNA sequencing in 60 patients for whom tumor and sentinel lymph node (SLN) samples were available, and conducting differential expression, gene ontology, enrichment and survival analyses. Sequencing annotation classified most of the sncRNAs into small nucleolar RNA (snoRNAs, 70%) and small nuclear RNA (snRNA, 13%). Our results showed no significant differences in sncRNA expression between tumor or SLNs obtained from the same patient. Differential expression analysis showed down-regulation (n = 21) sncRNAs and up-regulation (n = 2) sncRNAs in patients with locoregional metastasis. The expression of *SNHG5*, *SNORD90*, *SCARNA2* and *SNORD78* differentiated luminal A from luminal B tumors, whereas *SNORD124* up-regulation was associated with luminal B HER2+ tumors. Discriminating analysis and receiver-operating curve analysis revealed a signature of six snoRNAs (*SNORD93*, *SNORA16A*, *SNORD113-6*, *SNORA7A*, *SNORA57* and *SNORA18A*) that distinguished patients with locoregional metastasis and predicted patient outcome. Gene ontology and Reactome pathway analysis showed an enrichment of biological processes associated with translation initiation, protein targeting to specific cell locations, and positive regulation of Wnt and NOTCH signaling pathways, commonly involved in the promotion of metastases. Our results point to the potential of several sncRNAs as surrogate markers of lymph node metastases and patient outcome in early-stage breast cancer patients. Further preclinical and clinical studies are required to understand the biological significance of the most significant sncRNAs and to validate our results in a larger cohort of patients.

Keywords: sncRNAs; snoRNAs; early breast cancer; sentinel lymph node; metastasis; biomarkers



Citation: Escuin, D.; Bell, O.; García-Valdecasas, B.; Clos, M.; Larrañaga, I.; López-Vilaró, L.; Mora, J.; Andrés, M.; Arqueros, C.; Barnadas, A. Small Non-Coding RNAs and Their Role in Locoregional Metastasis and Outcomes in Early-Stage Breast Cancer Patients. *Int. J. Mol. Sci.* **2024**, *25*, 3982. <https://doi.org/10.3390/ijms25073982>

Academic Editor: Bruna Scaggiante

Received: 1 March 2024

Revised: 18 March 2024

Accepted: 27 March 2024

Published: 3 April 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/>).

1. Introduction

Cancer metastases are responsible for most breast cancer deaths. Despite intensive research in this field, our comprehension of the molecular events that drive metastatic progression remains largely incomplete. The identification of predictive and prognostic biomarkers is needed for early diagnosis of cancer and response monitoring of available therapies [1]. Ideally, these biomarkers should be highly specific and sensitive and several strategies are currently being developed to detect low expressed cancer-related biomarkers in liquid biopsies and other tumor samples [2]. However, more information on their expression in tumors, and thus suitability as biomarkers, is still required.

In recent years, small non-coding RNAs (sncRNAs) have emerged as important regulators of many cellular processes, including various steps of the metastatic process. SncRNAs

encompass an abundant variety of RNA molecules, 15–300 nucleotides (nt) in length, including microRNAs (miRNAs), small interfering RNAs (siRNAs), piwi-interacting RNA (piRNAs), small nucleolar RNAs (snoRNAs) and small nuclear (snRNAs).

SnoRNAs, which have long been known, represent a class of abundantly expressed sncRNAs, primarily present in the nucleolus and play pivotal roles in post-transcriptional rRNA processing and modification, thereby contributing significantly to the maintenance of cellular functions related to protein synthesis. SnoRNAs are approximately 30–300 nt long and, based on conserved sequence elements, are classified into C/D box snoRNAs (SNORDS) or H/ACA box snoRNA (SNORA), which determine the binding and modification of the RNA target. A third class of snoRNAs contains both sequence motifs and localizes to the nuclear Cajal bodies (SCARNA) [3]. However, approximately half of human snoRNA have no predictable rRNA targets, and numerous snoRNAs have been discovered to possess the ability to influence cell fate and alter disease progression; they therefore hold immense potential in terms of controlling human diseases, including Prader–Willi syndrome, Duplication15q syndrome and cancer [4]. It has been suggested that snoRNA dysregulation exhibits differential expression across various cancer types, stages, metastasis, treatment response and/or prognosis in patients [5]. This new role of snoRNAs has been addressed by recent studies showing that snoRNA can act to regulate pre-mRNA alternative splicing and mRNA abundance, as well as activate enzymes and be processed into shorter sncRNAs resembling miRNAs and piRNAs [6–8]. Furthermore, recent biochemical studies have shown that a given snoRNA can form both methylating and non-methylating ribonucleoprotein complexes, providing clues to the likely physical basis for such diverse new functions [9]. SnoRNAs are evidently more structurally and functionally diverse than previously thought, and their role in gene expression is under-appreciated.

SnRNAs are a class of highly abundant sncRNA molecules with an average size of 150 nt present in the cell nucleus and are involved in intron removal from pre-mRNA. SnRNA form a large particulate complex (spliceosome) along with ribonucleoprotein particles (snRNPs) and additional proteins, which binds to the primary RNA transcripts to mediate the splicing. Additional evidence indicates that snRNPs function in nuclear maturation of primary transcripts in mRNAs, gene expression regulation, splice donor in non-canonical systems and in 3'-end processing of replication-dependent histone mRNAs [10]. Accumulating evidence demonstrate that snRNA dysregulation are closely related to the progression of cancer through different mechanisms, such as transcriptional inhibition and post-transcriptional regulation [11].

In this study, we performed RNA sequencing to profile the sncRNA expression in 60 patients with early-stage breast cancer for whom tumor tissue and SLNs samples were available. We identified most of the sncRNAs as snoRNAs or snRNAs, and we found that, overall, down-regulation of snoRNAs was associated with patient locoregional metastatic status. Furthermore, our classifier model yielded a 6-snoRNA signature that clearly differentiated between negative and positive metastatic SLN and correlates with patient outcome. Deregulated snoRNAs showed a significant enrichment of biological processes associated with translation initiation, protein targeting to various organelles and regulation of Wnt and NOTCH signaling pathways. Our data highlight the potential use of sncRNAs as surrogate markers of locoregional metastases and patient outcome in breast cancer.

2. Results

2.1. Patients

For each of the 60 female patients who were included in this study, we analyzed paired tumor tissues and SLNs. The main clinicopathological characteristics of the patients are described in Table 1. Of the 60 patients, 40 (67%) had SLN-positive tumors, 20 were diagnosed as micrometastasis and 20 were diagnosed as macrometastasis.

Table 1. Basic patient and tumor characteristics.

Variable		Total (%)	LN Negative (%)	LN Positive (%)
Patients (%)		60 (100)	20 (100)	40 (100)
Age (years)	Mean + SD	63.6 ± 14.2	68.8 ± 11.9	60.9 ± 14.7
	Median (range)	64.3 (26–88)	71 (47–87)	60 (26–88)
	<40 years	3 (5)	0 (0)	3 (8)
	40–50 years	6 (10)	2 (10)	4 (10)
	>50 years	51 (85)	18 (90)	33 (82)
Tumor status	T1c	32 (53)	12 (60)	20 (50)
	T2	28 (47)	8 (40)	19 (50)
Node status	N0	20 (33)	20 (100)	0 (0)
	N1	40 (67)	0 (0)	40 (100)
Tumor stage	IA	12 (20)	12 (60)	0 (0)
	IIA	28 (47)	8 (40)	20 (50)
	IIB	20 (33)	0 (0)	20 (50)
Tumor grade	1	3 (5)	1 (5)	2 (5)
	2	34 (57)	14 (70)	20 (50)
	3	23 (38)	5 (25)	18 (45)
Tumor focality	Unifocal	38 (63)	13 (65)	25 (62)
	Multifocal	19 (32)	7 (35)	12 (30)
	Multicentric	3 (5)	0 (0)	3 (8)
Ki67 status	≤14%	29 (48)	7 (35)	22 (55)
	>14%	31 (52)	13 (65)	58 (18)
ER status	Negative	6 (10)	2 (10)	4 (10)
	Positive	54 (90)	18 (90)	36 (90)
PR status	Negative	22 (37)	6 (30)	16 (40)
	Positive	38 (63)	14 (70)	24 (60)
Molecular subtype	Luminal A	25 (42)	6 (30)	19 (48)
	Luminal B	21 (35)	10 (50)	11 (28)
	Luminal B	8 (13)	3 (15)	6 (15)
	HER2+	0 (0)	0 (0)	0 (0)
	HER2+	0 (0)	0 (0)	0 (0)
LVI	TN	5 (8)	1 (5)	4 (9)
	Negative	52 (87)	19 (95)	33 (83)
Menopausal status	Positive	8 (13)	1 (5)	7 (17)
	Premenopausal	10 (17)	2 (10)	8 (20)
Breast affected	Postmenopausal	50 (83)	18 (90)	32 (80)
	Left	28 (47)	10 (50)	18 (45)
Breast surgery	Right	32 (53)	10 (50)	22 (55)
	Mastectomy	18 (32)	5 (25)	13 (33)
	Lumpectomy	41 (68)	15 (75)	27 (67)

LN: lymph node; ER: estrogen receptor; PR: progesterone receptor; TN: triple negative; LVI: lymphovascular invasion.

2.2. RNA Sequencing

A total of 117 samples (98%) from 59 tumors and 58 SLNs passed the pre- and post-sequencing quality check, which confirmed average read quality and base quality Q-scores > 30 (99.9% correct) [12]. Three samples (1 tumor and 2 SLNs) were excluded from further analyses. In total, we analyzed 57 patients with paired samples (n = 117). The mean read number for tissues and SLNs were 3.8 million and 4.4 million, respectively. Following sequencing and trimming, reads were collapsed into a single read and passed into the analysis pipeline. This allowed for true quantification of the sncRNAs by eliminating library amplification bias and a better representation of the RNA molecules in the sample. We obtained an average 0.96 million and 1 million of collapsed reads for tissues and SLNs, respectively, and an average genome mapping rate of 25% and 23% for tissues and SLNs, respectively. The raw counts yield a total of 4207 sncRNAs that was reduced to 536 sncRNAs after performing a filtering step of at least 1 CPM in half of the samples. Count data were normalized and log2 transformed using the regularized log (rlog) method from the DESeq2

package (Table S1). The resulting sncRNAs were classified as snoRNAs (69.5%), snRNA (12.5%), miscellaneous RNAs (7.1%) and rRNAs (9.3%). Within the snoRNAs category, the majority of them were SNORDS (67.8%), followed by SNORA (29.7%) and SCARNA (2.6%).

2.3. Correlations and Clustering Analyses

To investigate whether patients were assigned into biological groups based on their sncRNA expression, we performed supervised hierarchical clustering using 50 sncRNAs with the largest coefficient of variation based on rlog-normalized counts. Our data indicated no significant differences between tumor and SLNs (Figure 1A). Similar results were obtained using a principal component analysis (PCA). Despite our data showing some differences between the two samples, those differences were not sufficiently large to cluster samples into different groups (Figure 1B). We also performed a tumor-to-SLN Spearman's correlation analysis (r_s). Our results showed that sncRNA expression in tumor and SLN samples from the same patient were highly correlated, with an average value for all patients of $r_s = 0.955$ (0.904–0.975) (Figure 1C, Table S2).

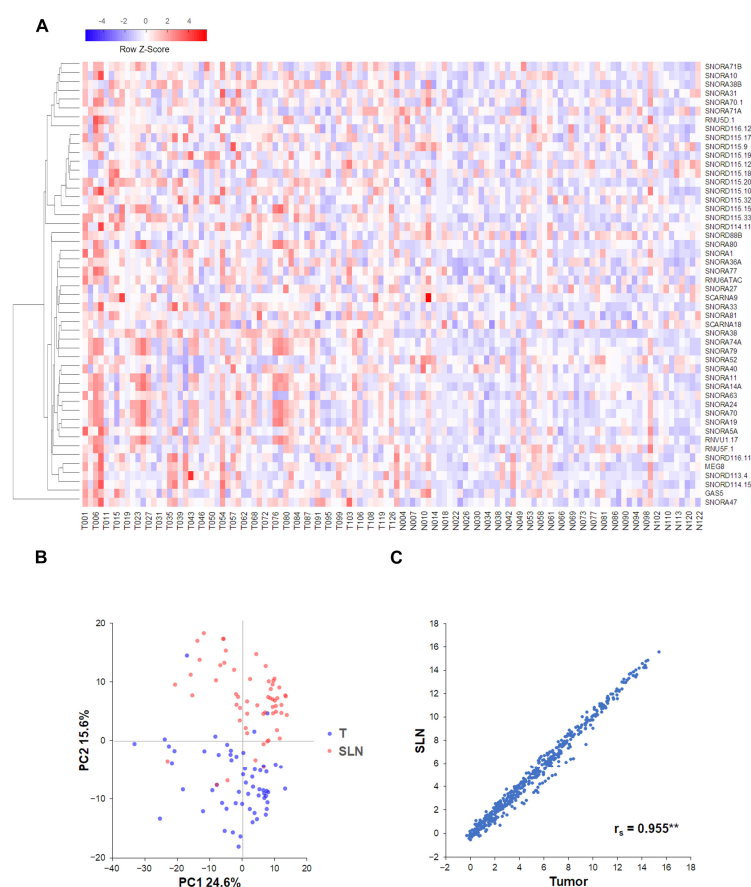


Figure 1. Class discovery associated with SLN metastatic status. The analysis was performed using 50 sncRNAs with the largest coefficient of variation based on rlog-normalized values. (A) Heatmap and unsupervised hierarchical clustering. Each row represents one sncRNA and each column represents one sample. The row Z-score scaling method was used to represent expression level above (red) and below (blue) the mean. (B) Principal component analysis shows sample clusters arising naturally based on the sncRNA expression profile. (C) Scatterplots depicting tumor-to-SLN comparison between samples from the same patient show the log expression of sncRNA expression for each sample type. The average Spearman's correlation coefficient (r_s) for all tumor–SLN comparisons is shown. ** Two-tailed $p < 0.01$.

Figure 2. Differentially expressed snRNAs. The volcano plots show differentially expressed snRNAs in tumor samples according to patient locoregional metastatic status (A–C) and molecular subtype (D–F). The data show the logarithmic relationship between false discovery rate-adjusted *p* values (q value) (*y*-axis) and the log2 fold change expression (*x*-axis). Red, blue and grey dots show q values < 0.05, non-adjusted *p* values < 0.05 and non-significant *p* values > 0.05, respectively. Only snoRNAs with an absolute log2 fold change ≥ 1.5 are labeled.

We next investigated DE sncRNAs based on breast cancer molecular subtypes. Our series included mainly patients with luminal A (n = 25) and luminal B (n = 21), followed by luminal B HER2+ (n = 8) and TN (n = 5) tumors (Table 1). Analyzing the first three subgroups (Figure 2D–F), our results show down-regulation of *SNHG5*, *SNORD90*, *SCARNA2* and up-regulation of *SNORD78* associated with luminal B compared to luminal A tumors. On the other hand, *SNORD124* up-regulation was associated with luminal B HER2+ compared to either luminal A (Figure 2E) or luminal B tumors (Figure 2F).

2.5. Biological Significance and Enriched Analysis of sncRNAs

We performed a biological significance analysis using DE sncRNAs based on patient locoregional metastatic status ($p < 0.05$ and absolute log fold change > 0.3). In contrast to other sncRNAs such as microRNAs, annotation of snoRNAs in functional databases such as gene ontology (GO), KEGG or Reactome is scarce. Therefore, biological significance analysis was assessed using three different gene list, including the snoRNAs host genes and gene targets retrieved from the snoDB database, and genes correlated selected sncRNA expression in the TCGA-BRCA and SNOric databases (Table S4). Overall, our data show that the host genes of DE sncRNAs in patients with positive locoregional metastasis included GO categories associated with translational initiation, various processes targeting specific proteins to particular regions of the cell during or after the translational process, regulation of the Wnt signaling pathway (Figure 3A). Likewise, the top GO categories of targets genes associated with DE sncRNAs were also involved in the same processes.

We carried out an enrichment analysis to determine those pathways associated with the DE sncRNAs. GO biological processes analysis using the SnoDB showed that the target genes were associated with Wnt signaling, protein translation, targeting proteins to particular cell locations, histone methylation, neutrophil activation and cell maturation and development (Figure 3B,C). To further understand the signaling pathways involved in the regulation of DE snoRNAs according to locoregional metastasis status, we performed a similar analysis using the Reactome, which, in contrast to the GO biological processes, makes extensive use of protein complex interactions in its representation, thus given a more detail picture of the pathways involved with a particular set of snoRNAs. Our results show an enrichment of DE sncRNAs associated with the NOTCH processing, resolution of sister chromatids and chemokine binding pathways (Figure 3D, Table 2).

Table 2. GO and Reactome enrichment analyses associated with Wnt and NOTCH signaling pathways. Data show sncRNAs associated with each biological term and their target genes.

ID	Description	Ratio	p	sncRNAs	Target Gene
GO:0090263	Positive regulation of canonical Wnt signaling pathway	3/37	0.003	<i>SNORA80E</i> , <i>SNORD10</i> , <i>SNORD15A</i>	<i>PSMD11</i> , <i>LGR5</i> , <i>DDX3X</i>
GO:0030177	Positive regulation of Wnt signaling pathway	3/37	0.005	<i>SNORA80E</i> , <i>SNORD10</i> , <i>SNORD15A</i>	<i>PSMD11</i> , <i>LGR5</i> , <i>DDX3X</i>
GO:0060070	Canonical Wnt signaling pathway	3/37	0.026	<i>SNORA80E</i> , <i>SNORD10</i> , <i>SNORD15A</i>	<i>PSMD11</i> , <i>LGR5</i> , <i>DDX3X</i>
GO:0016055	Wnt signaling pathway	3/37	0.083	<i>SNORA80E</i> , <i>SNORD10</i> , <i>SNORD15A</i>	<i>PSMD11</i> , <i>LGR5</i> , <i>DDX3X</i>
GO:0198738	Cell–cell signaling by wnt	3/37	0.084	<i>SNORA80E</i> , <i>SNORD10</i> , <i>SNORD15A</i>	<i>PSMD11</i> , <i>LGR5</i> , <i>DDX3X</i>
GO:0007219	Notch signaling pathway	12/474	0.003	<i>SNORD15B</i> , <i>SNORA68</i> , <i>SNORD10</i> , <i>SNORD93</i>	<i>DLL1</i> , <i>EGFL7</i> , <i>MFNG</i> , <i>MIB2</i> , <i>NOTCH4</i> , <i>NOTCH1</i> , <i>DNER</i> , <i>FOXA1</i> , <i>FOXC1</i> , <i>YBX1</i> , <i>TBX2</i> , <i>DLGAP5</i>
R-HSA-8951430	RUNX3 regulates WNT signaling	2/305	0.021	<i>SNORD93</i>	<i>RUNX3</i> , <i>TCF7L1</i>
R-HSA-157118	Signaling by NOTCH	14/305	0.008	<i>SNORD15B</i> , <i>SNORA68</i> , <i>SNORD10</i> , <i>SNORD93</i>	<i>DLL1</i> , <i>HDAC7</i> , <i>MFNG</i> , <i>MIB2</i> , <i>NOTCH4</i> , <i>PSMD12</i> , <i>YWHAZ</i> , <i>DNER</i> , <i>NOTCH1</i> , <i>TLE4</i> , <i>YBX1</i> , <i>FLT4</i> , <i>DLGAP5</i> , <i>ST3GAL3</i>

Table 2. Cont.

ID	Description	Ratio	p	sncRNAs	Target Gene
R-HSA-9012852	Signaling by NOTCH3	5/305	0.013	SNORD15B, SNORA68, SNORD10, SNORD93	DLL1, NOTCH1, YBX1, DLGAP5, MIB2
R-HSA-2691230	Signaling by NOTCH1 HD Domain Mutants in Cancer	3/305	0.008	SNORD15B, SNORA68, SNORD10, SNORD93	DLL1, NOTCH1, MIB2
R-HSA-2691232	Constitutive Signaling by NOTCH1 HD Domain Mutants	3/305	0.008	SNORD15B, SNORA68, SNORD10, SNORD93	DLL1, NOTCH1, MIB2
R-HSA-9013695	NOTCH4 Intracellular Domain Regulates Transcription	3/305	0.019	SNORD15B, SNORA68, SNORD10, SNORD93	NOTCH4, NOTCH1, FLT4
R-HSA-9013507	NOTCH3 Activation and Transmission of Signal to the Nucleus	3/305	0.034	SNORD15B, SNORA68, SNORD10, SNORD93	DLL1, YBX1, MIB2
R-HSA-350054	Notch-HLH transcription pathway	3/305	0.045	SNORD15B, SNORA68, SNORD10, SNORD93	HDAC7, NOTCH4, NOTCH1
R-HSA-1912399	Pre-NOTCH Processing in the Endoplasmic Reticulum	2/305	0.011	SNORD15B, SNORA68, SNORD10, SNORD93	NOTCH4, NOTCH1
R-HSA-2660825	Signaling by NOTCH1 t(7;9) (M1580_K2555) Translocation Mutant	2/305	0.016	SNORD15B, SNORA68, SNORD10/SNORD93	DLL1, NOTCH1
R-HSA-2660826	Constitutive Signaling by NOTCH1 t(7;9)(M1580_K2555) Translocation Mutant	2/305	0.016	SNORD15B, SNORA68, SNORD10/SNORD93	DLL1, NOTCH1
R-HSA-9013700	NOTCH4 Activation of Signal to the Nucleus	2/305	0.038	SNORD15B, SNORA68, SNORD10	NOTCH4, YWHAZ

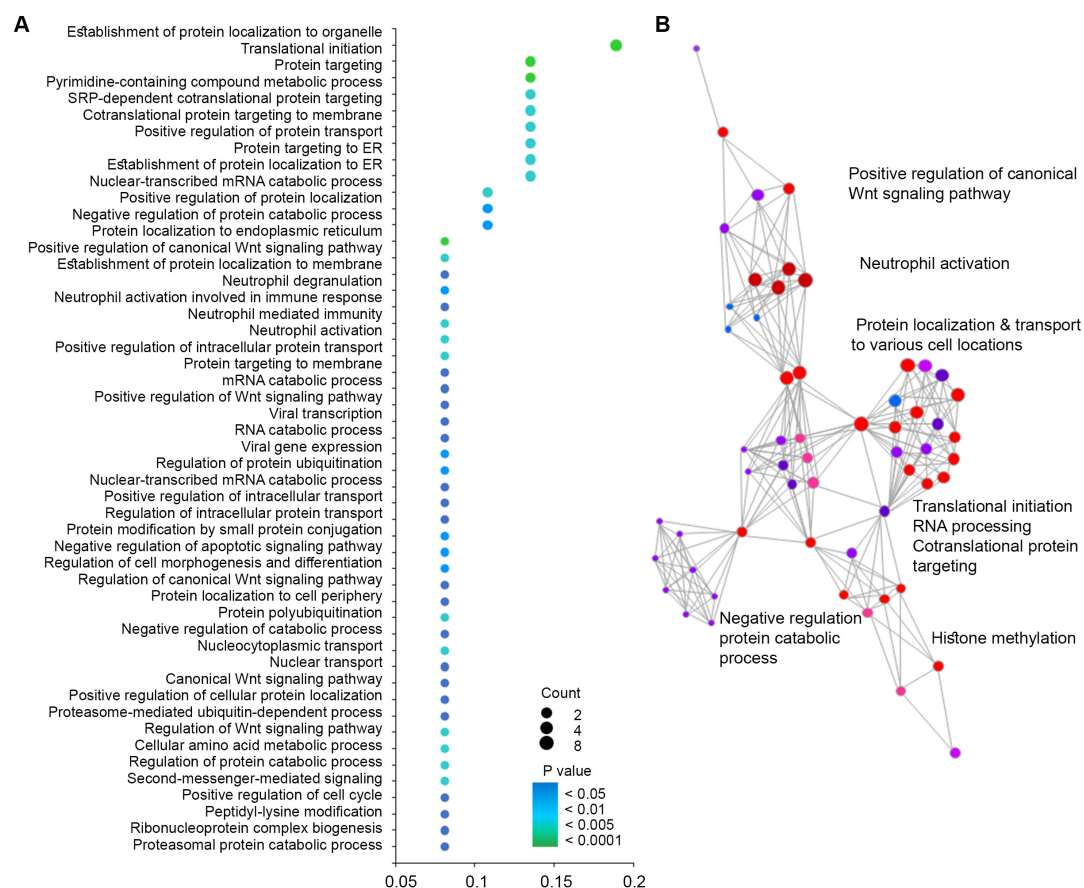


Figure 3. Cont.

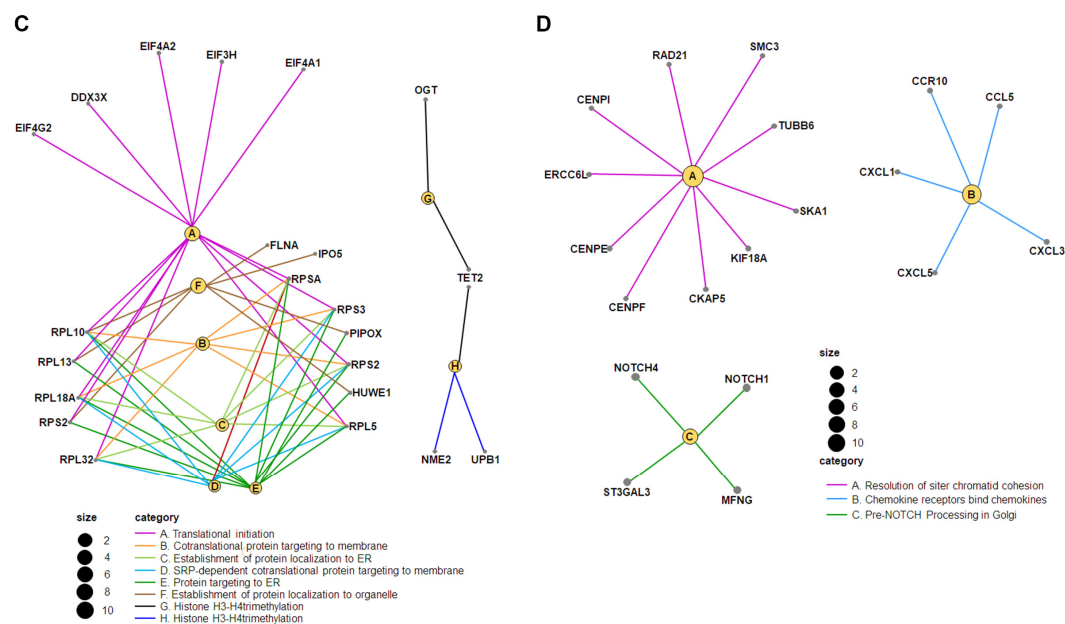


Figure 3. Gene ontology (GO) enrichment analysis for significant biological processes associated with positive SLNs. (A) The dot plot graph shows the 50 most significant biological process GO terms (*y*-axis) and the ratio between the number of expressed sncRNAs associated with the GO term and the number of significantly differentially expressed genes associated with the GO term (*x*-axis). The color of the nodes indicates the *p* value and the size of the nodes the number of sncRNAs associated with a specific GO term. (B) Enrichment map of the top 60 sncRNAs, with pathways grouped by similarity. Node size indicates the number of sncRNAs found in a pathway and node color reflects the significance of the *p* value. (C,D) The neural plots show the link between genes and terms associated with the most significant GO terms or Reactome pathways, respectively.

2.6. Clinicopathological Correlation with DE sncRNAs

In our series of 60 patients with early-stage breast cancer, we observed recurrence in 11 (18%) patients. Median follow-up time was 9.6 years (range 0.4–12.5 years). At the last follow-up, nine (15%) patients were dead. The univariate analysis showed several sncRNAs associated with tumor grade (*n* = 134), lymphovascular invasion (*n* = 20) and tumor focality (*n* = 12) and, to a lesser degree, with menopausal status (*n* = 2) and tumor stage (*n* = 1) (Table 3 and Table S5).

Table 3. Univariate analysis shows the number of significant (*q* < 0.01) sncRNAs associated with the patient clinicopathological characteristics of the patients. The names of the top 10 most significant sncRNAs are shown.

Variable	N	sncRNA Name
Tumor grade	134	SNORD105B, SNORD19C, SNORD101, SNORD102, SNORD107, SNORD114-3, SNORD126, SNORD18A, SNORD61, SNORD64
Lymphovascular invasion	20	RNU5D-1, RNU5E-1, RNU7-1, SNORA15B-1, SNORA2B, SNORA36B, SNORA36C, SNORA38B, SNORA41, SNORA53
Tumor focality	12	RNU5D-1, RNU5E-1, RNU7-1, SNORA15B-1, SNORA2B, SNORA36B, SNORA36C, SNORA38B, SNORA41
Menopausal status	2	RNU5E-1, RNU5D-1
Tumor stage	1	SNORD115-8

We further analyzed significant DE sncRNAs with an absolute log2 fold change ≥ 1.5 in terms of classifying patients according to the metastatic status. Various combinations of sncRNAs were examined using a classifier model producing a 6-snoRNA signature (*SNORD93*, *SNORA16A*, *SNORD113-6*, *SNORA7A*, *SNORA57* and *SNORA18A*) that clearly separated non-metastatic from metastatic tumors (Figure 4A). The receiver-operating characteristic (ROC) area under the curve (AUC) was 0.855, with a sensitivity of 89.7% and a specificity of 75% (Figure 4B). For results categorized into low or high expression based on the ROC analysis score, Kaplan–Meier analysis indicated that patients with low snoRNA score expression showed a trend toward improved disease-free survival (DFS), albeit not significant (HR = 3.012, 95% 0.92–9.92, $p = 0.057$) (Figure 4C) and better overall survival (HR = 4.067, 95% 1.06–9.92, $p = 0.038$) (Figure 4D).

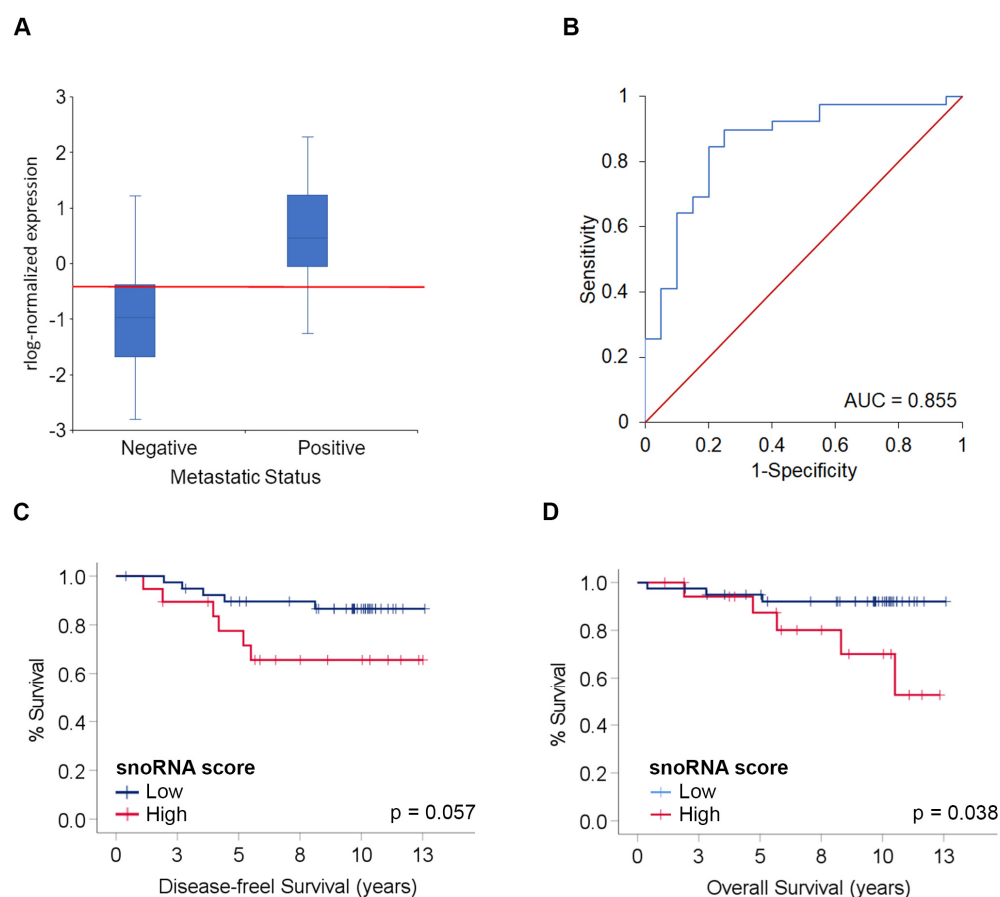


Figure 4. Classifier model and association with patient outcome. (A) Boxplot shows the normalized expression of the snoRNA score (y-axis) and the metastatic status of patients. The red line indicates the cutoff value. (B) ROC curve analysis of the snoRNAs score for discriminating patients with locoregional metastasis (blue). The reference line for random classification is shown in red. (C,D) Kaplan–Meier survival curves and log-rank tests for disease-free survival and overall survival based on snoRNAs categorized as low or high expression.

3. Discussion

Breast cancer is one of the most prevalent cancers among women and the leading cause of cancer mortality in women [13]. Currently, LN affection remains the most important prognosis factor in breast cancer [14,15] and the presence of metastasis in the SLNs is still currently the recommended procedure for axillary staging of early breast cancer [16]. Our recent research has focused on the involvement of miRNAs in the development of locoregional metastases in patients with early-stage breast cancer [17–19]. We did not, however, investigate other classes of sncRNAs that have emerged in recent years as

important regulators in cancer development and in the various steps of the metastatic process [20].

In this study, we investigated the expression of sncRNAs in paired primary tumor and SLNs from early-stage breast cancer patients and correlated the results with SLN metastatic status. Our RNA sequencing data show that 83% of the annotated sncRNAs were classified as snoRNAs (70%) or snRNAs (13%), whereas the rest belong to miscellaneous RNAs, including some long non-coding RNAs (lncRNAs). Nonetheless, the full landscape of sncRNAs may have not been revealed in our dataset as (1) we could only use tools as currently available for gene annotation are constantly improving (and those tools are constantly improving), (2) known RNA sequencing biases still exist [9] and (3) RNA library preparation limitations may prevent certain sncRNAs from being amplified [20].

Our data show that sncRNA expression in tumor and SLNs is similar, with minor non-significant changes that are likely due to histological differences between the SLN and the tumor, whereas SLNs are constituted mainly by lymphoid and monocytes cells the tumor tissue is formed mostly from epithelial and mesenchymal cells. Since our data showed high tumor–SLN correlation for patients, we performed further analyses on tumor samples.

We identified 23 DE sncRNAs associated with patient locoregional metastatic status. Overall, and similar to our previously reported data on miRNAs, we found that most DE sncRNAs were down-regulated (adjusted $q < 0.05$), suggesting that the expression of all sncRNAs follows a similar pattern in early-stage breast cancer patients [17,19]. Interestingly, down-regulated sncRNA expression was similar in patients affected by micro- and macrometastases, but up-regulated sncRNAs (*SNORD93*, *SNORD114-20* and *SNORD116-24*) was only significant in the micrometastatic group. However, the number of patients in each group was small ($n = 20$) and it remains to be elucidated whether loss of the expression of specific sncRNAs is part of the natural history course of breast cancer tumors.

Among the DE sncRNAs in our list, four snoRNAs (*SNORA47*, *SNORD94*, *SNORA70* and *SNORD10*) have been documented to be involved in various human cancers [5]. *SNORA47* has been reported to be up-regulated in human hepatocellular carcinoma and associated with intrahepatic metastasis and lymphatic invasion. In addition, a high expression of *SNORA47* predicted worse patient outcome [21]. *SNOR94*, *SNORA70* and *SNORD10* have been reported up-regulated in a p53 oncogenic gain-of-function mutant mouse osteosarcoma model. The authors showed by RNA-seq that a cluster of snoRNAs were highly up-regulated in p53 mutant tumors in association with the Ets2 transcription factor-binding site. Homozygous deletion of Ets2 resulted in down-regulation of these snoRNAs and reversed the pro-metastatic phenotype of p53 mutant tumors [22]. The results for those four snoRNAs suggest that they act as oncogenes, contrast with our results and suggest that, in breast cancer *SNORA47*, *SNOR94*, *SNORA70* and *SNORD10* may play a role as a tumor-suppressor gene (TSG) by yet unknown mechanisms. In support of this argument, down-regulation of *SNORD10* has been associated with epigenetic promoter silencing in stage IV melanoma cell lines [23]. Thirteen sncRNAs in our dataset (some of which with a non-adjusted $p < 0.05$) (*RMRP*, *RN7SK*, *SNORA47*, *SNORA50C*, *SNORA71A*, *SNORA73B*, *SNORA7B*, *SNORA80E*, *SNORD10*, *SNORD112*, *SNORD12B*, *SNORD15A* and *VTRNA2-1*) were found annotated in the DisGeNet database [24], a curated database integrating information on human–disease associations from various repositories and inferred associations from literature text mining. *RMRP*, *SNORA7B*, *SNORD15A*, *SNORA71A* and *VTRNA2-1* have been previously reported in breast carcinomas [25–28]. *RMRP*, part of the RNase mitochondrial RNA processing (MRP) complex, has been found to be regulated by the oncogenic Wnt/ β -catenin and Hippo/YAP pathways [25]. *SNORA7B* has been reported up-regulated in breast tumors compared to normal tissue [26], and *SNORD15A* and *SNORA71A* have been found up-regulated in brain metastases [27]. *VTRNA2-1* has been shown to be involved in the inhibition of protein kinase R (PKR) activity and act as a TSG in several cancer types. Increased breast cancer risk has been associated with down-regulation of *VTRNA2-1* linked to five methylation marks within the *VTRNA2-1* promoter region [28].

An interesting picture emerges from our data compared to the aforementioned reported data. First, our dataset included only tumor samples since normal tissue from the same patients was not available for study. We therefore could not address the relative expression of the DE sncRNAs found in our study compared to normal breast tissue, which would have helped to elucidate whether the studied sncRNAs act as oncogenes or TSG. However, three DE sncRNAs associated with locoregional metastases *SNORA80E*, *SNOR15B* and *SNORD114-20* have been previously reported to be deregulated in invasive local BC compared to benign breast tissue [29]. Second, our biological significance and enrichment analyses—based on various databases that included host genes, target genes and interactions between genes and sncRNAs from the TCGA-BRCA atlas—showed similar signaling pathways as those described in the literature [20,30]. For instance, our data identified various GO biological processes as part of the Wnt signaling pathway. WNTs and WNT pathway components are also frequently over- or under-expressed in various cancers, and these changes are correlated with epigenetic regulation of promoter activity. In some contexts, both the canonical and non-canonical WNT signaling, which governs processes such as cell polarity and morphogenesis, may also contribute to tumor formation by promoting cell migration, invasiveness and metastasis [31]. In addition to the GO biological processes that focus on the activities of individual genes, we used the Reactome pathway, which makes extensive use of protein complexes in their pathway representations and describes their formation, dissociation and activities [32]. The main Reactome pathways were associated with the Notch signaling pathway (NSP), a highly conserved pathway for cell–cell communication involved in the regulation of cellular differentiation, proliferation, and specification [33]; chemokine receptors and their interaction with various chemokines that activates integrins for leukocyte adherence on endothelial cells and induces chemotaxis of leukocytes in tissue microenvironments [34]; and various processes associated with mitosis such as resolution of sister chromatids during mitotic prometaphase that indicates the involvement of sncRNAs in cell proliferation.

In addition to finding a correlation of the DE of sncRNAs with patient metastasis status, we found that a large number of sncRNAs correlated with various clinicopathological features, especially tumor grade, lymphovascular invasion, tumor focality and breast cancer molecular subtypes, in agreement with a recent study [29]. Interestingly, we have also described *SNORD124* up-regulation in tumors expressing HER2, suggesting that *SNORD124* could serve as a diagnostic biomarker for HER2-positive tumors. More importantly, our six-sncRNAs signature (*SNORD93*, *SNORA16A*, *SNORD113-6*, *SNORA7A*, *SNORA57* and *SNORA18A*) accurately distinguished between patients with locoregional metastasis. We also found that a low expression of the sncRNA discriminant score was associated with better patient outcome.

The main limitation of our study is the small number of samples used to assess the potential of sncRNAs as surrogate biomarkers of the lymph node metastasis in breast cancer. Therefore, our results are preliminary and must be interpreted with caution. Nonetheless, in this study, we provide evidence that several sncRNAs are associated with the locoregional metastatic status and patient outcome in early-stage breast cancer. Further studies are required in a larger number of patients to clinically validate our results and to unveil the molecular mechanisms of the sncRNAs described in this study.

4. Materials and Methods

Patients. We studied 60 patients with early-stage breast cancer treated with surgery. Male patients were excluded from this study. Sample size was determined according to the model developed by Dobin, K et al. [35]. None of the patients had been previously treated with surgery, chemotherapy or radiation. All patients had confirmed diagnoses based on tumor biopsy histopathology and intraoperative SLN tissue evaluated using the OSNA assay [36]. All tumors were invasive ductal carcinomas (IDCs) with or without an in situ component. The following clinical and pathological parameters were recorded: age, menopausal status, personal and family disease precedents, clinical follow-up, tumor

stage determined according to the UICC system [37], histological grade determined using the Elston–Ellis grading system [38], tumor histology, presence of associated carcinoma in situ, presence of vascular and lymphatic invasion, tumor infiltrating lymphocytes, tumor focality, tumor necrosis; proliferation of non-tumoral tissue. For each patient, we collected paired tumor and SLNs ($n = 120$ samples). Samples were classified according to the SLN status as negative ($n = 20$) or positive ($n = 40$). Positive samples were sub-classified as macrometastatic ($n = 20$) or micrometastatic ($n = 20$) [36].

RNA isolation. Tumor and SLNs were processed as previously described [17]. RNA was isolated from tumor and SLNs samples using miRNeasy (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions and eluted in a volume of 30 μ L. The RNA integrity (RIN) level was measured for each RNA sample using Agilent TapeStation (Santa Clara, CA, USA). All samples used in this study had a RIN value > 7 . A range of spike-ins were added to all samples prior to RNA isolation. A Pre-sequencing quality check by q-PCR was performed on all samples to control for the quality of the RNA and inhibition in downstream enzymatic reactions as previously described [18].

RNA sequencing. All steps required to perform next-generation sequencing and genome annotation were performed as previously described [17,18]. Genome annotation was performed using the QIAGEN CLC Genomics Server v20.0.4 (Qiagen, Germantown, MD, USA). Following sequencing, Cutadapt (1.9.1) [39] was used to trim adaptor sequences. A quality check was performed to ensure Q-scores > 30 ($>99.9\%$ correct) for our data [12]. Reads with the correct length were collapsed into FASTQ files. Bowtie2 software (2.2.6) was used to map the reads. The mapping criterion for aligning reads to spike-ins, abundant sequences and databases was for reads to perfectly match the reference sequences. To map the genome, two mismatches were allowed in the sequences. Small insertions and deletions were not allowed. The resulting sequences were annotated using the human assembly GRCh38 (Ensembl) and the snoDB database v1.2.1 [40]. The raw data were filtered to keep only sncRNAs with at least 1 CPM in half of the samples. Count data were then normalized and log2 transformed using the regularized log (rlog) method from the DESeq2 package [41] to eliminate biases in the composition of the sequencing libraries and to stabilize variance–mean dependence in count data.

Correlation, hierarchical clustering and differential expression analyses. Correlation analyses were performed using the rlog-normalized counts matrix. Spearman's rho (r_s) statistic and a heatmap plotted with Euclidian distances were used to measure similarities between samples from the same patients. To visualize sample expression profiles, hierarchical clustering was performed using Euclidian distances and scaled and centered rlog-normalized counts. PCA was performed to reduce the rlog-normalized counts in two dimensions. Differential expression analyses were performed using the trimmed mean of M values (TMM) normalization method [42], converted to log2 scale, the R statistical software package v3.6.3 and libraries from the Bioconductor Project (www.bioconductor.org, accessed on 26 March 2024) [43].

Biological significance and enrichment analyses. Functional annotation of selected sncRNAs with a $p < 0.05$ was performed using Ensembl, NCBI resources and snoDB database v1.2.1 [40]. Validated gene targets were searched for using R software v3.6.3 to retrieve sncRNA-target interactions from DisGeNet v7.0 database [24]. The biological significance analysis was performed using gene and host set lists from the snoDB database and the TCGA_BRCA dataset from the SNORic [44] data portal. Enrichment analyses were conducted using GO [45] and the Reactome pathway database [46]. The analyses were performed using the R/Bioconductor's cluster Profiler package v3.12.0 [47].

Classifier model building. Briefly, several iterations were performed during the resampling step in a balanced, random manner. The data were split into training and test cohorts, which were used to build and validate the model, respectively. Within each iteration, biomarker candidates were selected using different methods (t -test, lasso, random forest), and choosing a fixed number of features (3, 5, 10 and 25). Once candidates were selected, classification profiles were created using penalized logistic regression, partial

least squares-linear discriminant analysis, and support vector machines to find the best candidate in terms of accuracy and robustness of the data. A ROC analysis was performed to assess diagnostic score sensitivity and specificity. Optimization of the linear combination of biomarkers, was performed using the total area under the ROC curve (AUC) [48,49].

Statistics. Differentially expressed snoRNAs from RNA sequencing data were detected by an exact test based on conditional maximum likelihood included in the R Bioconductor edgeR package [50]. *p* values from RNA sequencing were corrected (*q* values) for multiple testing using the Benjamini–Hochberg procedure [51]. A false discovery rate $q < 0.05$ was considered significant. In all group comparisons, missing expression values were treated as zero. Differences in total snoRNAs numbers between groups were analyzed by using two-sided parametric *t*-tests. Clinicopathological analysis was performed using Student’s *T*-test to compare quantitative variables and the X2 or Fisher exact tests to compare qualitative variables. DFS was defined as the time from diagnosis to either date of first relapse (local, regional, contralateral or metastatic) or second primary cancer. OS was defined as the time from sample collection to death from any cause. Patients lost to follow-up were censored at the last contact. Kaplan–Meier and log-rank analyses were used to compare DFS and OS and a Cox regression model was used to perform the multivariate analysis. A two-sided *p* value ≤ 0.05 was considered significant.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25073982/s1>.

Author Contributions: Formal analysis, D.E. and O.B.; funding acquisition, D.E. and A.B.; investigation, O.B., B.G.-V., M.C., I.L., L.L.-V., J.M., M.A., C.A. and A.B.; methodology, D.E., B.G.-V., M.C., I.L., L.L.-V., J.M., C.A. and A.B.; project administration, D.E.; resources, D.E., B.G.-V., M.C., I.L., L.L.-V., J.M., C.A. and A.B.; validation, D.E.; writing—original draft, D.E.; writing—review and editing, D.E., O.B., B.G.-V., M.C., I.L., L.L.-V., J.M., M.A., C.A. and A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Instituto de Salud Carlos III (ISCIII) and European Regional Development Fund PI13/00110 to D.E. and PI19/00362 to D.E. and A.B.; Thematic area for Cancer (CIBERONC) CB16/12/00471 to A.B. and Catalanian Support to Research Groups (2021_SGR-00827 to A.B., D.E. and C.A.).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki, and approved by the Clinical Research Ethics Committee of “Institut de Recerca Sant Pau” (PI19/00362; 14 March 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The original contributions presented in this study are publicly available at the Sequence Research Archive under ID PRJNA663033 and are available for download here <http://www.ncbi.nlm.nih.gov/bioproject/663033> (accessed on 26 March 2024).

Acknowledgments: We thank the patients and their caregivers for their participation and the Research Centers of Catalonia (CERCA)/Generalitat de Catalunya for its structural support.

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

References

- Huijbers, A.; Velstra, B.; Dekker, T.J.; Mesker, W.E.; van der Burgt, Y.E.; Mertens, B.J.; Deelder, A.M.; Tollenaar, R.A. Proteomic serum biomarkers and their potential application in cancer screening programs. *Int. J. Mol. Sci.* **2010**, *11*, 4175–4193. [CrossRef] [PubMed]
- Pilvenyte, G.; Ratautaite, V.; Boguzaitė, R.; Ramanavicius, A.; Viter, R.; Ramanavicius, S. Molecularly Imprinted Polymers for the Determination of Cancer Biomarkers. *Int. J. Mol. Sci.* **2023**, *24*, 4105. [CrossRef] [PubMed]
- Maxwell, E.S.; Fournier, M.J. The small nucleolar RNAs. *Annu. Rev. Biochem.* **1995**, *64*, 897–934. [CrossRef] [PubMed]
- Zhang, X.; Wang, C.; Xia, S.; Xiao, F.; Peng, J.; Gao, Y.; Yu, F.; Wang, C.; Chen, X. The emerging role of snoRNAs in human disease. *Genes. Dis.* **2023**, *10*, 2064–2081. [CrossRef]
- Liang, J.; Wen, J.; Huang, Z.; Chen, X.P.; Zhang, B.X.; Chu, L. Small Nucleolar RNAs: Insight into Their Function in Cancer. *Front. Oncol.* **2019**, *9*, 587. [CrossRef] [PubMed]

6. Brameier, M.; Herwig, A.; Reinhardt, R.; Walter, L.; Gruber, J. Human box C/D snoRNAs with miRNA like functions: Expanding the range of regulatory RNAs. *Nucleic Acids Res.* **2011**, *39*, 675–686. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Falaleeva, M.; Welden, J.R.; Duncan, M.J.; Stamm, S. C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: Old dogs show new tricks. *Bioessays* **2017**, *39*, 1600264. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Scott, M.S.; Ono, M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* **2011**, *93*, 1987–1992. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Shi, J.; Zhou, T.; Chen, Q. Exploring the expanding universe of small RNAs. *Nat. Cell Biol.* **2022**, *24*, 415–423. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Hernandez, N. Small nuclear RNA genes: A model system to study fundamental mechanisms of transcription. *J. Biol. Chem.* **2001**, *276*, 26733–26736. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Xiao, L.; Wang, J.; Ju, S.; Cui, M.; Jing, R. Disorders and roles of tsRNA, snoRNA, snRNA and piRNA in cancer. *J. Med. Genet.* **2022**, *59*, 623–631. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Cock, P.J.; Fields, C.J.; Goto, N.; Heuer, M.L.; Rice, P.M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **2010**, *38*, 1767–1771. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Humphrey, K.L.; Saksena, M.A.; Freer, P.E.; Smith, B.L.; Rafferty, E.A. To do or not to do: Axillary nodal evaluation after ACOSOG Z0011 Trial. *Radiographics* **2014**, *34*, 1807–1816. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Zurrida, S.; Veronesi, U. Milestones in breast cancer treatment. *Breast J.* **2015**, *21*, 3–12. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Veronesi, U.; Paganelli, G.; Galimberti, V.; Viale, G.; Zurrida, S.; Bedoni, M.; Costa, A.; de Cicco, C.; Geraghty, J.G.; Luini, A.; et al. Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes. *Lancet* **1997**, *349*, 1864–1867. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Escuin, D.; Lopez-Vilaro, L.; Bell, O.; Mora, J.; Moral, A.; Perez, J.I.; Arqueros, C.; Ramon, Y.C.T.; Lerma, E.; Barnadas, A. MicroRNA-1291 Is Associated with Locoregional Metastases in Patients with Early-Stage Breast Cancer. *Front. Genet.* **2020**, *11*, 562114. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Escuin, D.; Lopez-Vilaro, L.; Mora, J.; Bell, O.; Moral, A.; Perez, I.; Arqueros, C.; Garcia-Valdecasas, B.; Ramon, Y.C.T.; Lerma, E.; et al. Circulating microRNAs in Early Breast Cancer Patients and Its Association with Lymph Node Metastases. *Front. Oncol.* **2021**, *11*, 627811. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Escuin, D.; Lopez-Vilaro, L.; Bell, O.; Mora, J.; Garcia-Valdecasas, B.; Moral, A.; Clos, M.; Boronat, L.; Arqueros, C.; Barnadas, A. Circulating miRNA Expression Is Inversely Correlated with Tumor Tissue or Sentinel Lymph Nodes in Estrogen Receptor-Positive Early Breast Cancer Patients. *Int. J. Mol. Sci.* **2023**, *24*, 13293. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Zhang, Z.; Zhang, J.; Diao, L.; Han, L. Small non-coding RNAs in human cancer: Function, clinical utility, and characterization. *Oncogene* **2021**, *40*, 1570–1577. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Li, G.; He, Y.; Liu, X.; Zheng, Z.; Zhang, M.; Qin, F.; Lan, X. Small nucleolar RNA 47 promotes tumorigenesis by regulating EMT markers in hepatocellular carcinoma. *Minerva Med.* **2017**, *108*, 396–404. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Pourebrahim, R.; Zhang, Y.; Liu, B.; Gao, R.; Xiong, S.; Lin, P.P.; McArthur, M.J.; Ostrowski, M.C.; Lozano, G. Integrative genome analysis of somatic p53 mutant osteosarcomas identifies Ets2-dependent regulation of small nucleolar RNAs by mutant p53 protein. *Genes. Dev.* **2017**, *31*, 1847–1857. [\[CrossRef\]](#)
23. Li, J.L.; Mazar, J.; Zhong, C.; Faulkner, G.J.; Govindarajan, S.S.; Zhang, Z.; Dinger, M.E.; Meredith, G.; Adams, C.; Zhang, S.; et al. Genome-wide methylated CpG island profiles of melanoma cells reveal a melanoma coregulation network. *Sci. Rep.* **2013**, *3*, 2962. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Pinero, J.; Ramirez-Angueta, J.M.; Sauch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* **2020**, *48*, D845–D855. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Park, J.; Jeong, S. Wnt activated beta-catenin and YAP proteins enhance the expression of non-coding RNA component of RNase MRP in colon cancer cells. *Oncotarget* **2015**, *6*, 34658–34668. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Sun, Y.; Chen, E.; Li, Y.; Ye, D.; Cai, Y.; Wang, Q.; Li, Q.; Zhang, X. H/ACA box small nucleolar RNA 7B acts as an oncogene and a potential prognostic biomarker in breast cancer. *Cancer Cell Int.* **2019**, *19*, 125. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Schulten, H.J.; Bangash, M.; Karim, S.; Dallol, A.; Hussein, D.; Merdad, A.; Al-Thoubaity, F.K.; Al-Maghrabi, J.; Jamal, A.; Al-Ghamdi, F.; et al. Comprehensive molecular biomarker identification in breast cancer brain metastases. *J. Transl. Med.* **2017**, *15*, 269. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Joo, J.E.; Dowty, J.G.; Milne, R.L.; Wong, E.M.; Dugue, P.A.; English, D.; Hopper, J.L.; Goldgar, D.E.; Giles, G.G.; Southey, M.C. kConFab, Heritable DNA methylation marks associated with susceptibility to breast cancer. *Nat. Commun.* **2018**, *9*, 867. [\[CrossRef\]](#)
29. Karkkainen, E.; Heikkinen, S.; Tengstrom, M.; Kosma, V.M.; Mannermaa, A.; Hartikainen, J.M. Expression profiles of small non-coding RNAs in breast cancer tumors characterize clinicopathological features and show prognostic and predictive potential. *Sci. Rep.* **2022**, *12*, 22614. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Xu, S.; Wang, L.; Zhao, Y.; Mo, T.; Wang, B.; Lin, J.; Yang, H. Metabolism-regulating non-coding RNAs in breast cancer: Roles, mechanisms and clinical applications. *J. Biomed. Sci.* **2024**, *31*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Polakis, P. Wnt signaling in cancer. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a008052. [\[CrossRef\]](#) [\[PubMed\]](#)

32. Good, B.M.; Van Auken, K.; Hill, D.P.; Mi, H.; Carbon, S.; Balhoff, J.P.; Albou, L.P.; Thomas, P.D.; Mungall, C.J.; Blake, J.A.; et al. Reactome and the Gene Ontology: Digital convergence of data resources. *Bioinformatics* **2021**, *37*, 3343–3348. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Schweisguth, F. Regulation of notch signaling activity. *Curr. Biol.* **2004**, *14*, R129–R138. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Kim, C.H. Chemokine-chemokine receptor network in immune cell trafficking. *Curr. Drug Targets Immune Endocr. Metabol. Disord.* **2004**, *4*, 343–361. [\[CrossRef\]](#)
35. Dobbin, K.K.; Zhao, Y.; Simon, R.M. How large a training set is needed to develop a classifier for microarray data? *Clin. Cancer Res.* **2008**, *14*, 108–114. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Tsujimoto, M.; Nakabayashi, K.; Yoshidome, K.; Kaneko, T.; Iwase, T.; Akiyama, F.; Kato, Y.; Tsuda, H.; Ueda, S.; Sato, K.; et al. One-step nucleic acid amplification for intraoperative detection of lymph node metastasis in breast cancer patients. *Clin. Cancer Res.* **2007**, *13*, 4807–4816. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Webber, C.; Gospodarowicz, M.; Sobin, L.H.; Wittekind, C.; Greene, F.L.; Mason, M.D.; Compton, C.; Brierley, J.; Groome, P.A. Improving the TNM classification: Findings from a 10-year continuous literature review. *Int. J. Cancer* **2014**, *135*, 371–378. [\[CrossRef\]](#)
38. Elston, C.W.; Ellis, I.O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. *Histopathology* **1991**, *19*, 403–410. [\[CrossRef\]](#)
39. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **2011**, *17*, 3. [\[CrossRef\]](#)
40. Bouchard-Bourelle, P.; Desjardins-Henri, C.; Mathurin-St-Pierre, D.; Deschamps-Francoeur, G.; Fafard-Couture, E.; Garant, J.M.; Elela, S.A.; Scott, M.S. snoDB: An interactive database of human snoRNA sequences, abundance and interactions. *Nucleic Acids Res.* **2020**, *48*, D220–D225. [\[CrossRef\]](#) [\[PubMed\]](#)
41. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **2010**, *11*, R25. [\[CrossRef\]](#) [\[PubMed\]](#)
43. R_Development_Core_Team. *R: A Language and Environment for Statistical Computing*; The R Foundation for Statistical Computing: Vienna, Austria, 2011; ISBN 3-900051-07-0. Available online: <http://www.R-project.org/> (accessed on 26 March 2024).
44. Gong, J.; Li, Y.; Liu, C.J.; Xiang, Y.; Li, C.; Ye, Y.; Zhang, Z.; Hawke, D.H.; Park, P.K.; Diao, L.; et al. A Pan-cancer Analysis of the Expression and Clinical Relevance of Small Nucleolar RNAs in Human Cancer. *Cell Rep.* **2017**, *21*, 1968–1981. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Gene Ontology, C. Gene Ontology Consortium: Going forward. *Nucleic Acids Res.* **2015**, *43*, D1049–D1056. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Fabregat, A.; Jupe, S.; Matthews, L.; Sidiropoulos, K.; Gillespie, M.; Garapati, P.; Haw, R.; Jassal, B.; Korninger, F.; May, B.; et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **2018**, *46*, D649–D655. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Pepe, M.S.; Cai, T.; Longton, G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* **2006**, *62*, 221–229. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Yin, J.; Tian, L. Optimal linear combinations of multiple diagnostic biomarkers based on Youden index. *Stat. Med.* **2014**, *33*, 1426–1440. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [\[CrossRef\]](#)

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.