

# SCIENTIFIC REPORTS

OPEN

## Identification of Long Noncoding RNAs Deregulated in Papillary Thyroid Cancer and Correlated with BRAF<sup>V600E</sup> Mutation by Bioinformatics Integrative Analysis

Lucas Goedert<sup>1,2</sup>, Jessica Rodrigues Praça<sup>2,3</sup>, Cesar Seigi Fuziwara<sup>4</sup>, Maiaro Cabral Rosa Machado<sup>1</sup>, Desirée Rodrigues Praça<sup>5</sup>, Palloma Porto Almeida<sup>6</sup>, Talita Perez Sanches<sup>1</sup>, Jair Figueredo dos Santos<sup>7</sup>, Amanda Cristina Corveloni<sup>8</sup>, Illy Enne Gomes Pereira<sup>1</sup>, Marcela Motta de Castro<sup>1</sup>, Edna Teruko Kimura<sup>4</sup>, Wilson Araújo Silva Jr.<sup>2,9</sup> & Enilza Maria Espreafico<sup>1,2</sup>

Papillary Thyroid Cancer (PTC) is an endocrine malignancy in which BRAF<sup>V600E</sup> oncogenic mutation induces the most aggressive phenotype. In this way, considering that lncRNAs are arising as key players in oncogenesis, it is of high interest the identification of BRAF<sup>V600E</sup>-associated long noncoding RNAs, which can provide possible candidates for secondary mechanisms of BRAF-induced malignancy in PTC. In this study, we identified differentially expressed lncRNAs correlated with BRAF<sup>V600E</sup> in PTC and, also, extended the cohort of paired normal and PTC samples to more accurately identify differentially expressed lncRNAs between these conditions. Indirectly validated targets of the differentially expressed lncRNAs in PTC compared to matched normal samples demonstrated an involvement in surface receptors responsible for signal transduction and cell adhesion, as well as, regulation of cell death, proliferation and apoptosis. Targets of BRAF<sup>V600E</sup>-correlated lncRNAs are mainly involved in calcium signaling pathway, ECM-receptor interaction and MAPK pathway. In summary, our study provides candidate lncRNAs that can be either used for future studies related to diagnosis/prognosis or as targets for PTC management.

Thyroid cancer is the endocrine malignancy<sup>1</sup> that, although stable until the 1990s, has progressively and greatly increased thereafter<sup>2,3</sup>. The vast majority of the thyroid cancers originate from the follicular cell epithelium<sup>1</sup>, which includes papillary thyroid carcinoma (PTC) that accounts for approximately 80% of all thyroid malignancies<sup>4</sup>.

Thyroid oncogenesis is still under investigation, however a high frequency (70%) of activating mutations in components of the mitogen-activated protein kinase (MAPK) pathway was reported, such as BRAF<sup>V600E</sup><sup>5,6</sup> and HRAS/NRAS/KRAS point mutations<sup>7,8</sup>. Also, fusions involving the RET<sup>9</sup> and NTRK1 tyrosine kinases<sup>10</sup> were described to promote thyroid cancers. More recently, the set of known PTC driver alterations was extended to include EIF1AX, PPM1D, and CHEK2<sup>7</sup>.

<sup>1</sup>Department of Cell and Molecular Biology, Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, São Paulo, Brazil. <sup>2</sup>National Institute of Science and Technology in Stem Cell and Cell Therapy and Center for Cell-Based Therapy, Ribeirão Preto, São Paulo, Brazil. <sup>3</sup>Clinical Oncology, Stem Cell and Cell Therapy Program, Ribeirão Preto Medical School, Ribeirão Preto, São Paulo, Brazil. <sup>4</sup>Department of Cell and Developmental Biology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, São Paulo, Brazil. <sup>5</sup>Biotechnology Program, Federal University of São Paulo, São José dos Campos, São Paulo, Brazil. <sup>6</sup>Biotechnology Program, Federal University of Bahia, Vitória da Conquista, Bahia, Brazil. <sup>7</sup>Biomedicine Program, Federal University of Pernambuco, Recife, Pernambuco, Brazil. <sup>8</sup>Biomedicine Program, State University of Londrina, Londrina, Paraná, Brazil. <sup>9</sup>Department of Genetics, Ribeirão Preto Medical School, and Center for Integrative System Biology – CISBi-NAP/USP, University of São Paulo, Ribeirão Preto, São Paulo, Brazil. Lucas Goedert and Jessica Rodrigues Praça contributed equally to this work. Correspondence and requests for materials should be addressed to E.M.E. (email: [emesprea@fmrp.usp.br](mailto:emesprea@fmrp.usp.br))

Additionally of being the most frequent mutation in many types of cancers including PTC<sup>7,11</sup>, BRAF<sup>V600E</sup> confers poorer prognosis compared to other oncogenes. There is a growing number of evidence demonstrating that BRAF<sup>V600E</sup> correlates with metastasis, cancer recurrence<sup>12</sup> and higher mortality in PTC<sup>13</sup>. BRAF<sup>V600E</sup>-expressing cells have a diversity of malignant characteristics, including increased DNA synthesis, dedifferentiation, and chromosomal instability<sup>14</sup>. Also, BRAF<sup>V600E</sup> stimulates more actively MEK-dependent invasion than the expression of RET/PTC oncoprotein through the expression of matrix metalloproteinases (e.g. MMP-3, MMP-9 and MMP-13), which, in part, can explain the more aggressive BRAF<sup>V600E</sup>-induced phenotype<sup>15</sup>.

Similarly to melanoma<sup>16</sup>, BRAF mutation occurs at early stages of PTC development<sup>11,17</sup>. Besides all BRAF oncogenic activities, its single exacerbated stimulation of the MAPK pathway is not sufficient to sustain malignant transformation, resulting in induced senescence<sup>18</sup> that confers a barrier to tumor progression<sup>19</sup>. To bypass BRAF-induced senescence, cells may suffer a second event that allows malignant transformation, as possibly the epigenetic silencing of tumor suppressors DAPK, TIMP3<sup>20</sup>, SLC5A8<sup>20,21</sup> and hMLH1<sup>22</sup> and other BRAF-induced mechanisms that remain to be discovered<sup>11</sup>. In thyroid cancers, Thyroid-stimulating Hormone (TSH) is more involved in overcoming senescence; while BRAF overexpression suppresses thyroid hormone biosynthesis and leads to elevated TSH levels *in vivo*<sup>14</sup>; it was shown that TSH signaling inhibits BRAF<sup>V600E</sup>-induced senescence through DUSP6<sup>23</sup>.

Recently, BRAF<sup>V600E</sup>-associated mRNA signature was determined in a mouse model and human samples<sup>24</sup>, which identified new genes not previously reported as related to BRAF mutation in thyroid cancer (e.g. MMD, ITPR3, AACs, LAD1, PVRL3, ALDH3B1, and RASA1) that will provide further support for future research on BRAF-induced PTC<sup>24</sup>. However, this analysis did not evaluate the expression of long noncoding RNAs (lncRNAs), which are progressively shown to be of fundamental importance in other types of cancer<sup>25,26</sup>. Such analysis is necessary for the identification of BRAF<sup>V600E</sup>-correlated long noncoding RNAs.

lncRNAs are RNAs longer than 200 nucleotides that have no coding potential<sup>27</sup> and are involved in several processes, such as gene expression regulation through chromatin modulation<sup>28,29</sup>, epigenetic control<sup>30</sup>, association with translational apparatus<sup>31</sup>, improving other mRNA stability<sup>32</sup>, serving as a scaffold for protein<sup>33</sup>, acting as decoys for miRNAs<sup>34</sup>, altering protein turnover<sup>35</sup>, among others.

To date, as per the authors' knowledge, only a few published studies identified differentially expressed (DE) lncRNAs between normal (N) and tumoral thyroid (T) in a limited set of paired samples<sup>36,37</sup>. Although these findings laid the foundation for further investigation of lncRNAs related to PTC<sup>36</sup>, they need to be confirmed in a more numerous group of patients.

In this study we confirmed previously reported lncRNAs and determined new DE lncRNAs in PTC in a larger set of samples and also identified BRAF<sup>V600E</sup>-correlated lncRNAs, providing possible candidates that can constitute secondary mechanisms of BRAF-induced malignancy in PTC.

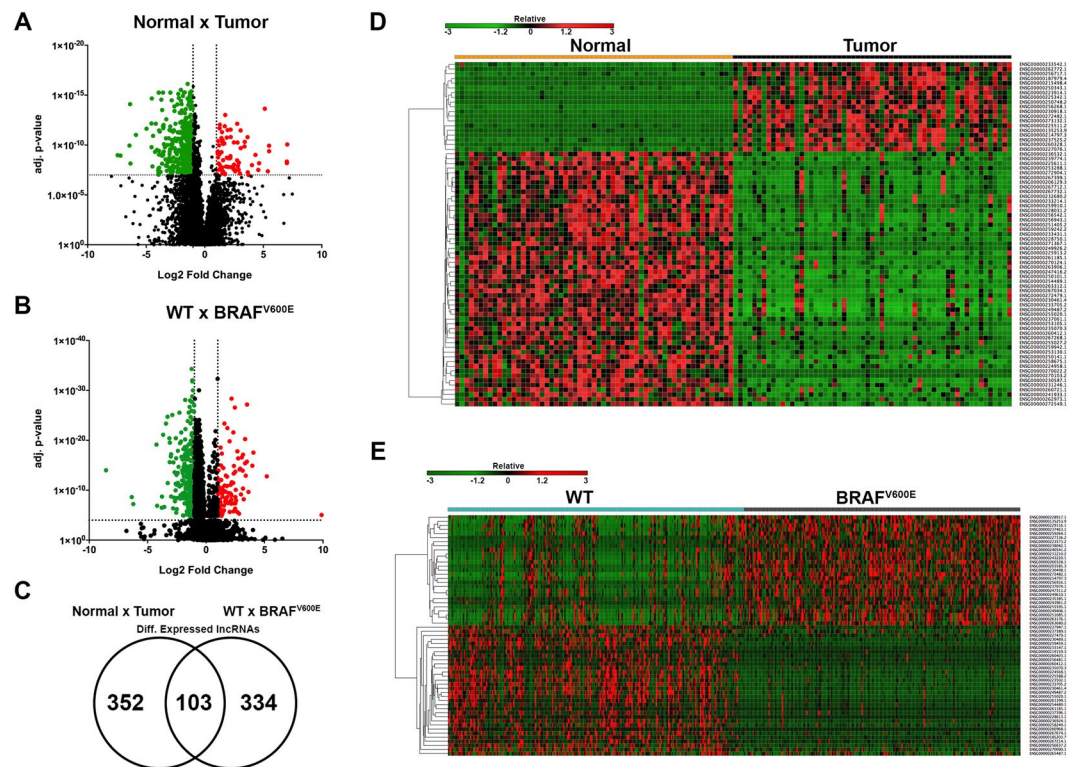
## Results

### Comparative analyses identified lncRNAs deregulated in PTC and correlated with BRAF<sup>V600E</sup> mutation.

Comparative analysis between 59 pairs of matched normal and papillary thyroid cancer samples identified 455 differentially expressed lncRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value < 1 × 10<sup>-7</sup>; Fig. 1A), being 71 lncRNAs upregulated and 384 lncRNAs downregulated in PTC (Supplemental Table S1). The same samples presented a total of 2016 mRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value < 0.05; Supplemental Table S2) and 186 microRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value < 0.05; Supplemental Table S3) differentially expressed.

Differential expression analyses were also performed to identify BRAF<sup>V600E</sup>-correlated lncRNAs. The comparison between BRAF wild type (WT) patients (n = 242) and BRAF<sup>V600E</sup> patients (n = 226), determined 437 differentially expressed lncRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value 1 × 10<sup>-4</sup>; Fig. 1B), being 117 upregulated and 320 downregulated (Supplemental Table S4). The same comparison found a total of 924 mRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value < 0.05; Supplemental Table S5) and 94 microRNAs (log<sub>2</sub> fold change > 1 or < -1; adj. p-value < 0.05; Supplemental Table S6) differentially expressed. A total of 103 lncRNAs was differentially expressed in both analyses [(Normal × Tumor and WT × BRAF<sup>V600E</sup>), (Fig. 1C, Table 1 and Supplemental Table S7)].

Experimental validation using qRT-PCR was performed to demonstrate the reliability of the bioinformatics analyses applied. From the top 25 positively DE lncRNAs and from the top 20 negatively DE lncRNAs, it were selected for validation those lncRNAs with low adjusted p-values to minimize expression variability, especially in the comparison BRAF<sup>WT</sup> × BRAF<sup>V600E</sup> tumor, among others characteristics (for detailed information see Methods). From a total of 5 DE lncRNAs selected for validation from the TCGA analysis (Fig. 2 and Supplemental Fig. S1, upper part), 4 DE lncRNAs were validated using thyroid cell lines (Fig. 2 and Supplemental Fig. S1, lower part), which strengthens the reliability of this bioinformatics analysis, although the experimentally tested set of lncRNAs constitutes a relatively small sampling. Considering all comparisons, we obtained a very expressive validation efficiency [from a total of 8 different comparisons (Normal × Tumor and BRAF<sup>WT</sup> × BRAF<sup>V600E</sup>), 6 were validated]. Downregulation of *ENSG00000235070.3* and *ENSG00000255020.1* in PTC was confirmed in the tumor cell lines TPC1 (BRAF<sup>WT</sup>) and BCPAP (BRAF<sup>V600E</sup>) compared to the normal immortalized cell line NTHY (Fig. 2). Also, downregulation of their expression was in accordance to the bioinformatics analysis, since lower expression for both of them was observed in BCPAP (BRAF<sup>V600E</sup>) compared to TPC1 (BRAF wild type) (Fig. 2). Noteworthy, is that due to the very low abundance of *ENSG00000255020.1* in the BCPAP cell line, qRT-PCR resulted in two unspecific melting peaks, which did not influenced the results. Upregulation of *ENSG00000273132.1* in PTC was confirmed, however its overexpression in BRAF<sup>V600E</sup> tumors was not observed in the cell line BCPAP compared to TPC1 (Fig. 2), maybe due to the small log<sub>2</sub> fold change value (1.69) of this comparison. Overexpression of *ENSG00000230498.1* in BRAF<sup>V600E</sup> PTC compared to BRAF wild type tumors was



**Figure 1.** Differentially expressed lncRNAs between  $N \times T$  and  $WT \times BRAF^{V600E}$  were identified. (A) Volcano plot of DE lncRNAs between  $N \times T$  ( $\log_2$  fold change  $> 1$  or  $< -1$ ; adj. p-value  $< 1 \times 10^{-7}$ ). (B) Volcano plot of DE lncRNAs between  $WT \times BRAF^{V600E}$  ( $\log_2$  fold change  $> 1$  or  $< -1$ ; adj. p-value  $< 1 \times 10^{-4}$ ). (C) Venn Diagram of common DE lncRNAs between  $N \times T$  and  $WT \times BRAF^{V600E}$ . (D) Heatmap\* of DE lncRNAs between  $N \times T$  ( $\log_2$  fold change  $> 3$  or  $< -3$ ; adj. p-value  $< 1 \times 10^{-7}$ ). (E) Heatmap\* of DE lncRNAs between  $WT \times BRAF^{V600E}$  ( $\log_2$  fold change  $> 2.5$  or  $< -2.5$ ; adj. p-value  $< 1 \times 10^{-4}$ ). \*For hierarchical clustering, one minus Spearman rank correlation was performed.

Ensembl	log <sub>2</sub> FC	Adj. p-value	log <sub>2</sub> FC	Adj. p-value
	$N \times T$	$N \times T$	$WT \times BRAF^{V600E}$	$WT \times BRAF^{V600E}$
ENSG00000214797.3	7.01	6.39E-09	5.19	4.92E-12
ENSG00000273132.1	5.49	4.17E-10	1.69	3.70E-12
ENSG00000230918.1	5.49	1.17E-10	1.36	3.89E-05
ENSG00000260328.1	5.39	4.27E-08	4.04	2.26E-16
ENSG00000256268.1	5.12	2.27E-14	1.00	4.52E-13
ENSG00000261185.1	-6.42	1.13E-10	-2.81	3.55E-06
ENSG00000254489.1	-6.37	7.98E-15	-3.27	8.33E-11
ENSG00000260412.1	-6.33	7.11E-09	-3.77	2.59E-08
ENSG00000235070.3	-4.97	2.28E-12	-3.01	1.10E-15
ENSG00000247416.2	-4.55	6.89E-13	-2.40	1.09E-16

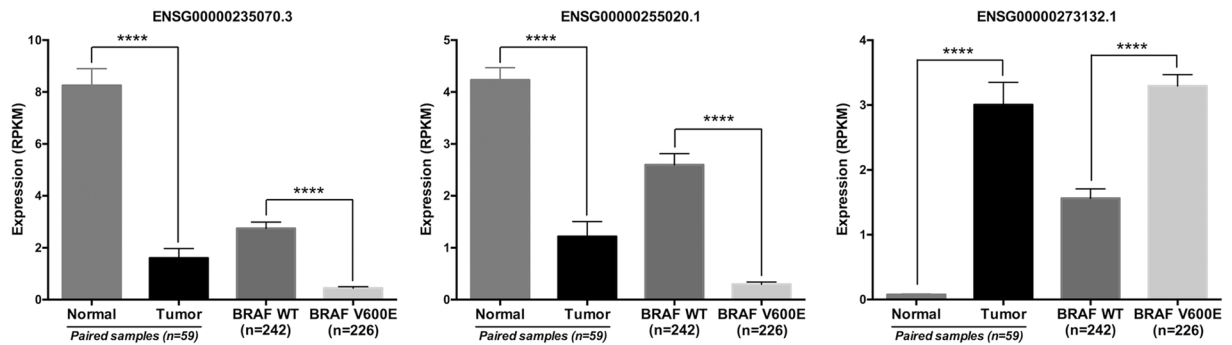
**Table 1.** Top 5 upregulated and 5 downregulated common DE lncRNAs between  $Normal \times Tumor$  and  $WT \times BRAF^{V600E}$ .

also confirmed (Supplemental Fig. S1); nevertheless ENSG00000247311.2 was undetectable in both TPC1 and BCPAP cells (Supplemental Fig. S1).

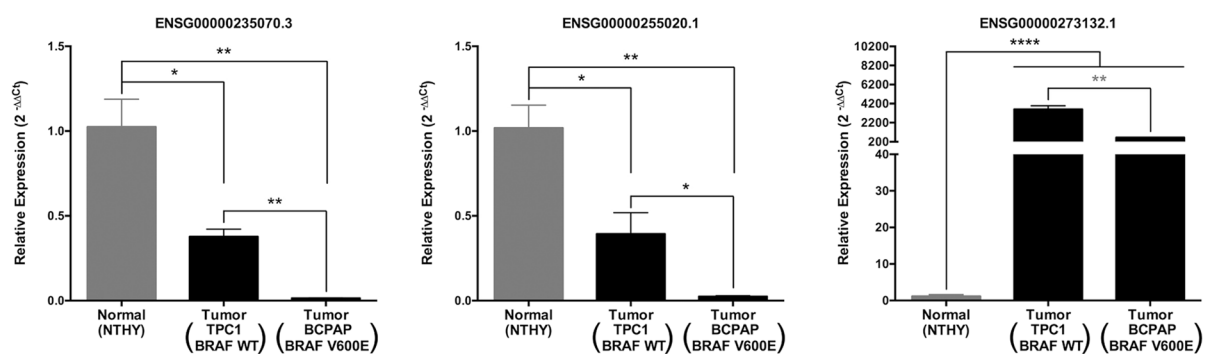
**Clustering lncRNAs identifies two groups with similar expression patterns.** For downstream analyses, we increased the stringency of differentially expressed lncRNAs between  $Normal \times Tumor$  ( $\log_2$  fold change  $> 3$  or  $< -3$ ; adj. p-value  $< 1 \times 10^{-7}$ ,  $n = 73$ ; Fig. 1D) and between  $WT \times BRAF^{V600E}$  ( $\log_2$  fold change  $> 2.5$  or  $< -2.5$ ; adj. p-value  $< 1 \times 10^{-4}$ ,  $n = 59$ ; Fig. 1E) to analyze the lncRNAs that were most DE. Hierarchical clustering was used to organize patients or lncRNAs into groups according to the expression levels of DE lncRNAs.

Results demonstrated that this set of lncRNAs is capable of clustering, majorly, normal and cancer patients in two distinct groups (Supplemental Fig. S2A). Clustering lncRNAs by Spearman correlation among all

## TCGA Analysis



## Experimental Validation



**Figure 2.** Experimental validation of DE lncRNAs. Upper part of panel displays the expression levels of the indicated lncRNAs in the TCGA analyses. The nonparametric Mann–Whitney test was applied due to the non-Gaussian expression distribution and p-value was assigned. Lower part of panel displays the experimental validation of these lncRNAs measured by qRT-PCR and calculated with  $2^{-\Delta\Delta C_t}$  method using RPL19 (Ribosomal Protein L19) as endogenous control. Experiments with three biological replicates were performed using two technical replicates for each sample. These results are representative of at least two independent experiments. Values are plotted as expression mean  $\pm$  Standard Error of Mean (SEM). Unpaired two-tailed t-Test assigned the p-value.

DE lncRNAs also identified two groups of lncRNAs highly positively correlated or negatively correlated (Supplemental Fig. S3A).

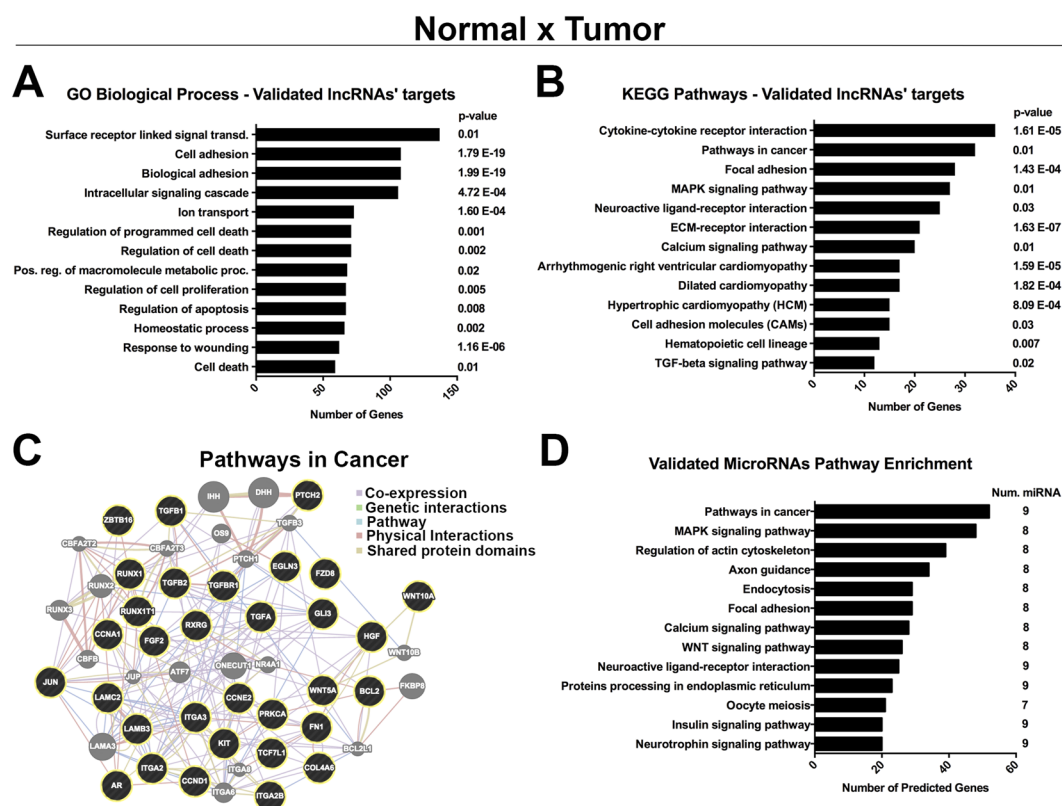
Hierarchical clustering was also performed with a more stringent set of DE lncRNAs between WT and BRAF<sup>V600E</sup>, which allowed the clustering of two groups enriched with WT and BRAF<sup>V600E</sup> patients, respectively (Supplemental Fig. S2B). Clustering lncRNAs by Spearman correlation among all DE lncRNAs also identified two groups highly positively correlated or negatively correlated lncRNAs (Supplemental Fig. S3B).

**Indirectly validated lncRNAs' targets are involved in several oncogenic processes.** As almost the totality of the identified DE lncRNAs in both conditions (Normal  $\times$  Tumor and WT  $\times$  BRAF<sup>V600E</sup>) is uncharacterized, we used prediction methods to identify a possible interaction between lncRNAs and mRNAs/microRNAs. Predicted mRNAs and microRNAs (targets of DE lncRNAs) were compared to differentially expressed mRNAs and microRNAs ( $\log_2$  fold change  $>1$  or  $<-1$ ; adj. p-value  $<0.05$ ) calculated from the same TCGA patients. Predicted mRNAs/microRNAs that were also identified as DE were considered as indirectly validated targets.

A total of 1109 DE mRNAs (Table 2 and Supplemental Table S8) and 26 DE microRNAs (Supplemental Table S9) were found to be predicted targets of the DE lncRNAs between Normal and Tumor samples and were considered as indirectly validated targets. Gene ontology and KEGG pathways enrichment of these validated mRNAs demonstrated that most of the genes are involved in surface receptors responsible for signal transduction and cell adhesion, as well as, regulation of cell death, proliferation and apoptosis (Fig. 3A). Enriched pathways (Fig. 3B) were composed of cytokine-cytokine receptor interaction, pathways in cancer (Fig. 3C), focal adhesion, MAPK pathway and calcium signaling pathway. Validated microRNAs were also used to determine enriched pathways based on their predicted targets calculated elsewhere (Fig. 3D). Genes involved in cancer and MAPK pathways were the most enriched pathways. Interestingly, some genes predicted to be targets of the validated microRNAs were also DE expressed in our analysis (Supplemental Table S2), such as upregulation of the MAPK constituents, CACNG4, CACNA1E, DUSP4, TGFBR1, FGF1, FGF2 and MAP3K1. Enriched pathways were extended to genes involved in cancer, focal adhesion and calcium signaling (Fig. 3D).

Ensembl	log2 FC	Adj. p-value	Upregulated indirectly validated targets	Downregulated indirectly validated targets
	N × T	N × T		
ENSG00000223914.1	7.04	8.69E-11	VGLL1, GDF6, FAM19A2, HRH1	RPS6KA5, RNF150, ANK2, FOSB
ENSG00000250748.2	7.01	4.56E-09	FUT3, GPR115, CAPN8, COL7A1	SVEP1, LMOD1, DPT, KCNA1
ENSG00000214797.3	7.01	6.39E-09	TMEM130, HES2, KCP, DTX4	CDHR3, PAK3, RASSF6, NWD1
ENSG00000273132.1	5.49	4.17E-10	KLK6, ELFN2, C19orf59, SHISA6	SRF, CPXM1, LAYN, FAM163A
ENSG00000230918.1	5.49	1.17E-10	GRM4, DPP4, LRP4, SHISA6	EGR1, TFPC2L1, FOXJ1, ABCA9
ENSG00000253288.1	-7.43	1.03E-09	SLC6A20, KLK10, FUT3, HRH1	HAP1, SH2D6, FOXP2, ADH1B
ENSG00000272479.1	-7.19	1.16E-09	DMBX1, TMPRSS6, TMPRSS4, PPP1R1B	CUX2, PAX1, CLCNKB, FOSB
ENSG00000261185.1	-6.42	1.13E-10	B3GNT3, ELFN2, LRP4, SHISA6	NR4A1, C1QTNF7, RNF150, FAM180B
ENSG00000254489.1	-6.37	7.98E-15	SYTL5, HPCAL4, KCNQ3, CPNE4	RBM24, PGA3, GFRA1, RNF150
ENSG00000260412.1	-6.33	7.11E-09	CLDN16, PDE4C, LRG1, SHROOM4	SLC26A4, CDHR3, PAK3, NWD1

**Table 2.** Top 5 upregulated and 5 downregulated DE lncRNAs between paired Normal × Tumor with examples of indirectly validated targets.



**Figure 3.** Indirectly validated targets of the DE lncRNAs between N × T are involved in cancer-related processes. (A) GO biological processes and (B) KEGG enriched pathways of the indirectly validated mRNA targets of the DE lncRNAs between N × T. (C) Proteins' network of genes linked to the pathways in cancer, where black circles are validated targets and grey circles are connective proteins. (D) KEGG Pathways enrichment of the indirectly validated microRNAs targets of the DE lncRNAs between N × T.

Ensembl	log2 FC	Adj. p-value	Upregulated indirectly validated targets	Downregulated indirectly validated targets
	WT × BRAF V600E	WT × BRAF V600E		
ENSG00000255595.1	9.87	5.68E-05	TCAP, ITGA2, LY6G6C, BEND6	SLC5A5, ASTN1, PART1, IRX6
ENSG00000214797.3	5.19	4.92E-12	HES2, DTX4, KCP, LDLR	RNF157, HAP1, NWD1, SLC14A2
ENSG00000260328.1	4.04	2.26E-16	TMPRSS4	PRND, TMPRSS3, PREX2, CNTNAP2
ENSG00000230498.1	3.95	5.74E-14	C1orf106, ADAMTS14, DMBX1, DUSP13	FCGBP, GCGR, SOX3, PPP2R2C
ENSG00000256916.1	3.62	3.51E-09	ELFN2, SPTBN2, MUC16, ZNF469	SSPO, ASXL3, CTNND2, CNTNAP2
ENSG00000267674.1	-8.55	3.89E-13	VSIG1, LDLR	ADM2, ST3GAL6, NWD1, SLC5A8
ENSG00000237396.1	-6.35	3.05E-08	SIGLEC6, SDK1, C1QL2, MUC21	TFCP2L1, SCUBE1, SBSN, PAX1
ENSG00000227947.1	-6.22	5.72E-07	ELFN2, EPHA10, SLC30A3, SYT1	SFTPC, GATA5, SFRP1, MATN1
ENSG00000224568.1	-4.23	8.95E-18	SLC6A14, C1orf106	HIF3A, SYT13, SLC29A4, ARSF
ENSG00000267214.1	-4.18	1.17E-06	ELFN2, COL7A1, B4GALNT3	FAM124A, SULT1A2, SLC29A4, CNTNAP2

**Table 3.** Top 5 upregulated and 5 downregulated DE lncRNAs between WT × BRAF<sup>V600E</sup> with examples of indirectly validated targets.

Between WT and BRAF<sup>V600E</sup>, 471 DE mRNAs (Table 3 and Supplemental Table S10) and 11 DE microRNAs (Supplemental Table S11) were indirectly validated. Gene ontology of these mRNAs demonstrated that most of the genes are also related to surface receptors involved in signal transduction and cell adhesion, but, additionally, with response to hormone stimulus and transmembrane transport (Fig. 4A). Enriched pathways (Fig. 4B) were constituted of calcium signaling pathway (Fig. 4C), cardiomyopathies and ECM-receptor interaction. KEGG enrichment pathway analysis of the validated DE microRNAs demonstrated participation of the MAPK and WNT pathways, as well as regulation of actin cytoskeleton and focal adhesion (Fig. 4D). Several pro-oncogenic genes were found to be upregulated in our analysis and were described as predicted targets of the validated DE microRNAs, as the example of MET and TGFBR1 genes (Supplemental Table S5).

## Discussion

Long noncoding RNAs are arising as key participants in cancer establishment and progression by several oncogenic mechanisms<sup>30,32</sup>. On the other hand, it is of urge interest the determination of how these lncRNAs are activated and how they can be associated with specific events or genotypes, such as point mutations. BRAF<sup>V600E</sup> is the driver oncogenic mechanism with the greatest incidence in PTC<sup>7</sup> and, therefore, any event correlated with this mutation will be necessary to understand BRAF<sup>V600E</sup>-induced aggressiveness.

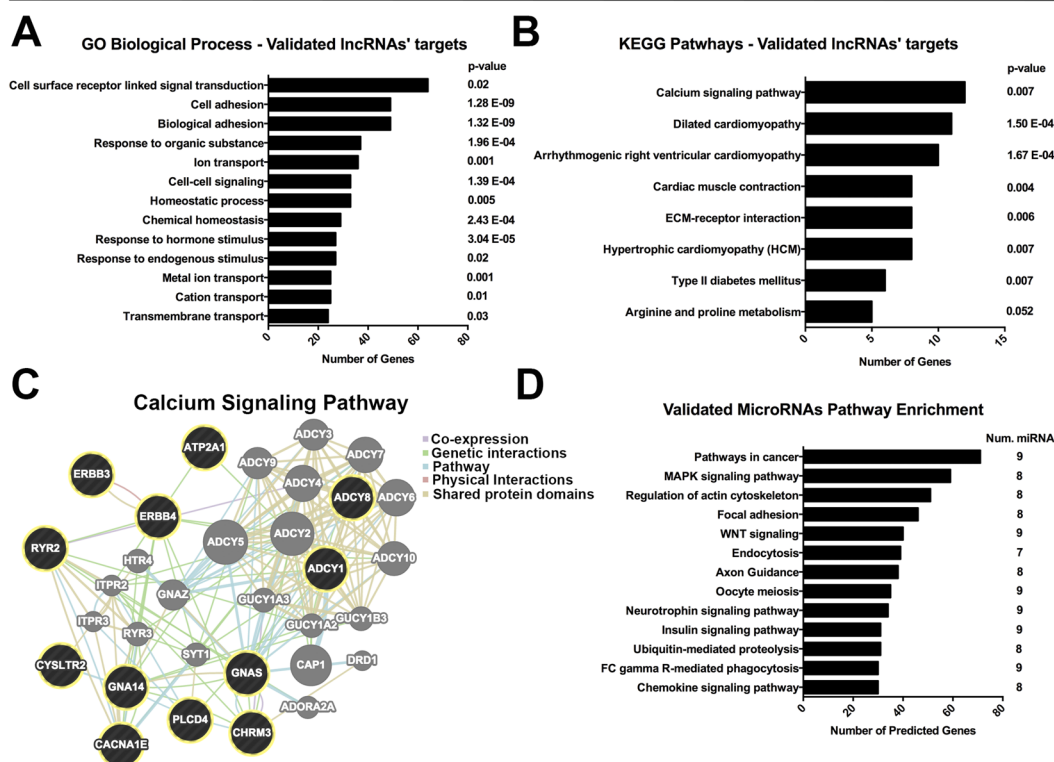
This is the first study to identify DE lncRNAs correlated with BRAF<sup>V600E</sup> in PTC and, besides that, we extended the cohort of paired normal and PTC samples to more accurately determine DE lncRNAs between these conditions.

We have identified 455 DE lncRNAs between paired normal and PTC samples. A total of 76 (log2 fold change > 1 or < -1; adj. p-value < 1 × 10<sup>-7</sup>) lncRNA were previously reported as DE in thyroid cancer compared to adjacent normal thyroid<sup>36</sup> (Fig. 5A and Supplemental Table S12). This validation set, together with the lncRNAs confirmed by experimental approaches (Fig. 2 and Supplemental Fig. S1), confers consistency to our analysis. Additionally, a diversity of DE lncRNAs identified in our analysis were reported in individual studies as altered in PTC samples, such as ENSG00000259104.2<sup>38,39</sup>, ENSG00000236130<sup>40</sup>, ENSG00000226363<sup>37</sup>, ENSG00000271086<sup>39,41</sup>, ENSG00000223914<sup>37</sup> and ENSG00000187979<sup>37</sup>.

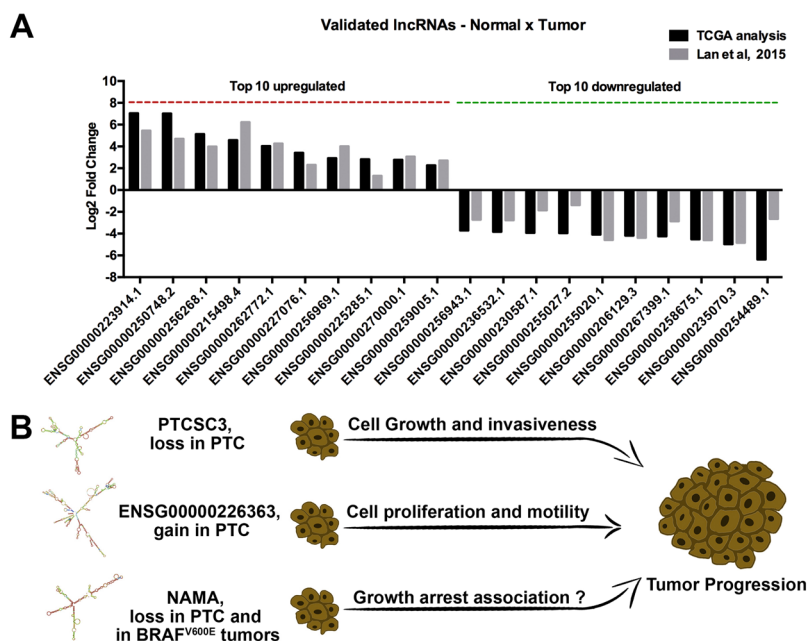
ENSG00000259104.2 (PTCSC3), which is downregulated in the tumor samples (log2 fold change -1.40; adj. p-value 1.11E-12) was previously reported as having thyroid-specific expression and decreased expression in PTC<sup>38,39</sup>. Interestingly, the risk allele [T] associated with SNP rs944289, located at PTCSC3's promoter, affects the binding site of C/EBPα and C/EBPβ (PTCSC3 activators), reducing its expression. Restoration of PTCSC3 expression in PTC cells inhibited cell growth and affected the expression of genes involved in DNA replication/repair, cellular movement and cell death<sup>38</sup>. Also PTCSC3 ectopic expression reduces cell proliferation and increases cell cycle arrest and apoptosis<sup>39</sup>, while reducing cell motility and invasiveness through S100A4 down-regulation<sup>42</sup> (Fig. 5B).

ENSG00000236130 (PTCSC2), was also reported as having decreased expression in PTC<sup>40</sup>, which was confirmed in our analysis (N × T log2 fold change -1.03; adj. p-value 3.12E-09). The risk allele [A] of rs965513 was significantly associated with low expression of unspliced PTCSC2 in unaffected thyroid tissue, however this correlation was not extended to PTC samples<sup>40</sup>.

### BRAF WT x BRAF<sup>V600E</sup>



**Figure 4.** Indirectly validated targets of the DE lncRNAs between WT x BRAF<sup>V600E</sup> are involved in oncogenic pathways. (A) GO biological processes and (B) KEGG enriched pathways of the indirectly validated mRNA targets of the DE lncRNAs between WT x BRAF<sup>V600E</sup>. (C) Proteins' network of genes linked to calcium signaling pathway, where black circles are validated targets and grey circles are connective proteins. (D) KEGG pathways enrichment of the indirectly validated microRNAs targets of the DE lncRNAs between WT x BRAF<sup>V600E</sup>.



**Figure 5.** Experimentally validated lncRNAs are important to tumor malignancy. (A) Examples of DE lncRNAs between Normal and PTC, which were confirmed in a validation set<sup>36</sup>. (B) Differentially expressed lncRNAs in PTC that alter tumor malignancy.

*ENST00000426615* (*ENSG00000226363*) is another lncRNA that we identified as upregulated in PTC (N×T log<sub>2</sub> fold change 3.87; adj. p-value 2.36E-05), which was experimentally demonstrated to be overexpressed in this cancer, inducing cell proliferation and motility<sup>37</sup> (Fig. 5B).

Our analysis also confirmed the differential expression (N × T: log<sub>2</sub> fold change -2.42; adj. p-value 3.96E-11) of the previously reported lncRNA *ENSG00000271086* (NAMA), which is downregulated in PTC compared to normal tissues<sup>39,41</sup> and in BRAF<sup>V600E</sup> tumors compared to wild type tumors<sup>39</sup> (Fig. 5B). NAMA is induced by inhibition of the MAPK pathway, growth arrest and DNA damage<sup>41</sup> and our analysis also demonstrated that NAMA is downregulated in BRAF<sup>V600E</sup> patients (WT × BRAF<sup>V600E</sup> log<sub>2</sub> fold change -1.66; adj. p-value 2.02E-15). All these independently validated lncRNAs demonstrate the reliability of our study (Fig. 5).

Similarity matrix based on Spearman correlation identified clusters of DE lncRNAs between Normal × Tumor (Supplemental Fig. S3A) and WT × BRAF<sup>V600E</sup> (Supplemental Fig. S3B) with similar expression patterns, which can provide evidence for further studies to determine common upstream regulators.

Indirectly validated targets of the DE lncRNAs between Normal × Tumor are involved in a diversity of biological processes (Fig. 3A). For instance, it was noticed an overrepresentation of adhesion molecules, such as downregulation of CDH16, which was already reported as a potential marker for PTC<sup>43</sup>. Along with CDH16, many other cadherins were identified as validated targets of DE lncRNAs, such as CDH2, CDH3, CDH4, CDH6, CDH11 and CDH24 (Supplemental Table S8). Another highly enriched biological process was the regulation of programmed cell death (Fig. 3A), represented by the upregulation of the antiapoptotic SOX4<sup>44</sup> and TP63<sup>45</sup> in PTC samples. Enriched pathways (Fig. 3B) as cytokine-cytokine receptor interaction, focal adhesion and MAPK pathways were already reported in the first study of DE lncRNA with paired Normal × PTC samples<sup>36</sup>, providing further support for future research. It was observed an enrichment of MAPK-related genes, represented in our results by upregulation, in the tumor samples, of TGB1, TGFB2 and TGFB1 that were shown to activate the MAPK pathway<sup>46</sup>. Interestingly, pathway enrichment analysis of indirectly validated microRNAs (Fig. 3D) demonstrated a convergent tendency to genes involved in cancer, MAPK pathway and focal adhesion, which were also observed with the validated mRNAs.

Indirectly validated targets of the DE lncRNAs between WT × BRAF<sup>V600E</sup> tumors were demonstrated to be involved with cell surface receptors responsible for signal transduction and with cell adhesion (Fig. 4A). Pathway enrichment analysis, identified genes involved in calcium signaling and ECM-receptor interaction, which were already reported as an early transcriptome change in BRAF<sup>V600E</sup>-associated mouse model<sup>24</sup>. Interestingly, we also observed several genes correlated with cardiomyopathies that are mostly related to calcium regulation in cardiac muscle cells (Fig. 4B). Calcium (Fig. 4C) and MAPK cascade (represented by BRAF<sup>V600E</sup> group) are tightly involved, where calcium modulates the protein interaction properties of ERKs, affecting the subcellular localization and influencing the distribution of their targets<sup>47</sup>. Calcium can also stimulates MEK through Ras activation<sup>48</sup>. Therefore, these results can support a future investigation to answer if BRAF<sup>V600E</sup>-stimulated MAPK activation can be reinforced by calcium modulation induced by the DE lncRNAs. Additionally, MAPK stimulation may be supported by the indirectly validated microRNAs, since pathway analysis demonstrated an enrichment of microRNAs' targets in cancer and MAPK pathways (Fig. 4D). Interestingly, predicted targets of the DE microRNAs were also differentially expressed in our analysis, such as the upregulation of TGFB1 and downregulation of PRKACB (Supplemental Table S5).

Concluding, our extended cohort of paired Normal and PTC patients identified new DE lncRNAs and confirmed many other lncRNAs already reported. Additionally, to our knowledge, this is the first study to identify BRAF<sup>V600E</sup>-correlated lncRNAs in PTC, which will provide support for future studies aiming to identify BRAF<sup>V600E</sup>-linked events in attempt to optimize therapeutic treatment and diagnosis/prognosis of this aggressive PTC genotype.

## Methods

**Data analysis.** Thyroid Carcinoma (THCA) clinical information, mRNA and microRNA data expression data were downloaded from The Cancer Genome Atlas (TCGA) online platform (<https://tcga-data.nci.nih.gov/tcga/>), as January 2016. Mutations data were retrieved through cBioPortal<sup>49</sup>. LncRNA RPKM expression levels corresponding to TCGA patients were downloaded through TANRIC<sup>50</sup>, which obtained the genomic coordinates of 13,870 human lncRNAs from the GENCODE Resource (version 19)<sup>51</sup> and further filtered out those lncRNA exons that overlapped with any known coding genes based on the gene annotations of GENCODE and RefGene, resulting in 12,727 lncRNAs<sup>50</sup>.

BRAF<sup>V600E</sup> patients were selected to form the BRAF<sup>V600E</sup> group (n = 226), which excluded any other type of BRAF mutations. Wild Type group (n = 242) was formed by patients without any somatic mutation in BRAF gene, but patients with mutations in HRAS, NRAS, KRAS, EIF1AX, PPM1D, RET and NTRK1 were considered. It was selected only the patients with papillary thyroid cancer diagnosis.

**Differential Expression Analysis.** For differential expression analysis of mRNA and microRNA was used edgeR package<sup>52</sup> through TCGAbiolinks<sup>53</sup>. To identify differentially expressed lncRNAs between groups, it was used the paired/unpaired Student t test to assess the statistical difference of mean expression values between the two groups<sup>50</sup>. LncRNA with median value equal zero were excluded and fold change was calculated using median expression values. In all differential analysis, p-values were adjusted for False Discovery Rate (FDR) < 0.05 as multiple hypothesis test correction method.

**RNA extraction, Reverse-transcription and qPCR.** Total RNA was phenol-chloroform extracted from cell lines Nthy-ori3-1 (NTHY-immortalized human thyroid follicular epithelial cell), TPC1 (papillary thyroid carcinoma- RET/PTC1 rearrangement) and BCPAP (papillary thyroid carcinoma-BRAF<sup>V600E</sup>) using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Four µg of total RNA was reverse transcribed using



M-MLV Reverse Transcriptase (Invitrogen) in the presence of 100 ng of random hexamers primers according to the manufacturer's protocol. qPCR reaction was performed using 100 ng of cDNA, 1X Power SYBR Green PCR Master Mix (Applied Biosystems) and specific primers. Amplification and detection were performed using ViiA7 TM Real-Time PCR System (Applied Biosystems). Relative gene expression was calculated using the QGENE program and calculated with  $2^{-\Delta\Delta Ct}$  method using RPL19 (Ribosomal Protein L19) as endogenous control.

Primers used (5'-3'): RPL19 (Fw-TCTCATGGAACACATCCACAA; Rv-TGGTCAGCCAGGAGCTTCTT), ENSG00000273132.1 (Fw-CTAGCTGCCAGCAGTGACAA; Rv-GCGAGAGCACAGATGACCAC), ENSG00000230498.1 (Fw-CCCTGGGTGATGAAGATGAG; Rv-TGGGATCCCTTTTGTGTCG), ENSG00000235070.3 (Fw-TGACTCCAAGTTCACGCAGC; Rv-GTGGATGAGTTGTGTGCTGG), ENSG00000255020.1 (Fw-AGTGACGTGGGGAAGAAACG; Rv-CGACATATTTCAAGGGCGCC) and ENSG00000247311.2 (Fw-GCTGTGAGTGACTCTTCAGC and ACAGACACACCCAGGAACAA).

To select the above DE lncRNAs for validation, it was taken in consideration one of the major characteristics of lncRNA, that is, high heterogeneous expression across the same tumor and even the same cell line. Long non-coding RNAs expression is tightly regulated by a wide range of cellular responses and, due to the markedly lower transcriptional levels of lncRNAs, the expression variability inside the same group of patients is expected. Taking this aspect in consideration, from the top 25 positively DE lncRNAs and from the top 20 negatively DE lncRNAs, it was selected for validation those lncRNAs with low adjusted p-values to avoid this variability, especially in the comparison BRAF<sup>WT</sup> × BRAF<sup>V600E</sup> tumor, which is the focus of this research. As another desirable characteristic, most lncRNAs selected, presented at least in one group (normal thyroid, PTC, BRAF<sup>WT</sup> or BRAF<sup>V600E</sup>) a median expression greater than 1 RPKM (reads per kilo base per million mapped reads). Added to that, it was given preference for those lncRNAs without isoforms (seen that many lncRNAs have annotation errors) and that present at least 2 exons, which are more stable and would allow the PCR primers to be located in different exons.

**Prediction of lncRNAs targets.** To identify the possible target genes of the selected (Fig. 1D and E) differentially expressed lncRNAs via cis- or trans-regulatory effects, two previously described approaches were used<sup>36,54</sup>. The genes transcribed within a 10 kb window upstream or downstream of lncRNAs were considered as cis-target genes<sup>36,54</sup>. The second method was used to identify trans-targets and is based on mRNA and microRNA sequence complementarity with the query lncRNA. For mRNA interactions we used a pre-computed database that catalogs the predicted lncRNA–RNA interactions<sup>55</sup>, where the accessible regions of the query lncRNA and possible targets (mRNA/lncRNA) are extracted, the binding energies of pairs of sequences (target and query) around the seed matches are evaluated and the minimum interaction energy of the joint secondary structures is calculated<sup>55</sup>. The 500 predicted targets (mostly constituting repeated targets with different interaction sites) with the lowest minimum free energies under  $-20$  kcal/mol were taken in consideration for downstream analysis. For lncRNA and microRNA interaction prediction it was used *rna22*, a method for identifying microRNA-binding sites and their corresponding heteroduplexes<sup>56</sup>. It was selected those microRNAs with a folding energy lower than  $-20$  kcal/mol.

**Indirect Validation.** As interaction prediction methods are susceptible to error and to minimize this, we compared the predicted targets of the differentially expressed lncRNAs with the differentially expressed mRNAs and microRNAs<sup>36</sup> calculated with the TCGA patients, because we consider that the targets of DE lncRNAs would possibly be DE in TCGA analysis. With this approach, we intended to enrich our analysis for targets with a greater propensity to be occurring biologically.

**Gene ontology, pathway enrichment and protein-protein interaction network.** Indirect validated targets of the DE lncRNAs were loaded into the Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>57</sup>, which returned the gene ontology of the query genes and identified enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways<sup>58</sup>. miRSystem<sup>59</sup> was used to calculate enriched pathways based on the predicted targets of the query microRNAs, which in this case, were the DE microRNAs in both conditions (Normal × Tumor and WT × BRAF<sup>V600E</sup>). For protein-protein interaction network it was used Genemania<sup>60</sup>.

## References

- Segev, D. L., Umbricht, C. & Zeiger, M. A. Molecular pathogenesis of thyroid cancer. *Surgical oncology* **12**, 69–90, doi:10.1016/S0960-7404(03)00037-9 (2003).
- Vigneri, R., Malandrino, P. & Vigneri, P. The changing epidemiology of thyroid cancer: why is incidence increasing? *Current opinion in oncology* **27**, 1–7, doi:10.1097/cco.000000000000148 (2015).
- Guay, B., Johnson-Obaseki, S., McDonald, J. T., Connell, C. & Corsten, M. Incidence of differentiated thyroid cancer by socioeconomic status and urban residence: Canada 1991–2006. *Thyroid: official journal of the American Thyroid Association* **24**, 552–555, doi:10.1089/thy.2013.0308 (2014).
- Fagin, J. A. & Mitsiades, N. Molecular pathology of thyroid cancer: diagnostic and clinical implications. *Best practice & research. Clinical endocrinology & metabolism* **22**, 955–969, doi:10.1016/j.beem.2008.09.017 (2008).
- Cohen, Y. *et al.* BRAF mutation in papillary thyroid carcinoma. *Journal of the National Cancer Institute* **95**, 625–627, doi:10.1093/jnci/95.8.625 (2003).
- Kimura, E. T. *et al.* High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. *Cancer research* **63**, 1454–1457 (2003).
- Integrated genomic characterization of papillary thyroid carcinoma. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **159**, 676–690, doi:10.1016/j.cell.2014.09.050 (2014).
- Lemoine, N. R. *et al.* Activated ras oncogenes in human thyroid cancers. *Cancer research* **48**, 4459–4463 (1988).
- Grieco, M. *et al.* PTC is a novel rearranged form of the ret proto-oncogene and is frequently detected *in vivo* in human thyroid papillary carcinomas. *Cell* **60**, 557–563, doi:10.1016/0092-8674(90)90659-3 (1990).
- Pierotti, M. A. *et al.* Rearrangements of TRK proto-oncogene in papillary thyroid carcinomas. *Journal of endocrinological investigation* **18**, 130–133, doi:10.1007/bf03349721 (1995).

11. Knauf, J. A. & Fagin, J. A. Role of MAPK pathway oncoproteins in thyroid cancer pathogenesis and as drug targets. *Current opinion in cell biology* **21**, 296–303, doi:10.1016/j.ceb.2009.01.013 (2009).
12. Kebebew, E. *et al.* The prevalence and prognostic value of BRAF mutation in thyroid cancer. *Annals of surgery* **246**, 466–470, discussion 470–461, doi:10.1097/SLA.0b013e318148563d (2007).
13. Xing, M. *et al.* Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *Jama* **309**, 1493–1501, doi:10.1001/jama.2013.3190 (2013).
14. Knauf, J. A. *et al.* Targeted expression of BRAFV600E in thyroid cells of transgenic mice results in papillary thyroid cancers that undergo dedifferentiation. *Cancer research* **65**, 4238–4245, doi:10.1158/0008-5472.can-05-0047 (2005).
15. Mesa, C. Jr. *et al.* Conditional activation of RET/PTC3 and BRAFV600E in thyroid cells is associated with gene expression profiles that predict a preferential role of BRAF in extracellular matrix remodeling. *Cancer research* **66**, 6521–6529, doi:10.1158/0008-5472.can-06-0739 (2006).
16. Pollock, P. M. *et al.* High frequency of BRAF mutations in nevi. *Nature genetics* **33**, 19–20, doi:10.1038/ng1054 (2003).
17. Adeniran, A. J. *et al.* Correlation between genetic alterations and microscopic features, clinical manifestations, and prognostic characteristics of thyroid papillary carcinomas. *The American journal of surgical pathology* **30**, 216–222, doi:10.1097/01.pas.0000176432.73455.1b (2006).
18. Michaloglou, C. *et al.* BRAF600-associated senescence-like cell cycle arrest of human naevi. *Nature* **436**, 720–724, doi:10.1038/nature03890 (2005).
19. Bartkova, J. *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444**, 633–637, doi:10.1038/nature05268 (2006).
20. Hu, S. *et al.* Association of aberrant methylation of tumor suppressor genes with tumor aggressiveness and BRAF mutation in papillary thyroid cancer. *International journal of cancer. Journal international du cancer* **119**, 2322–2329, doi:10.1002/ijc.22110 (2006).
21. Porra, V. *et al.* Silencing of the tumor suppressor gene SLC5A8 is associated with BRAF mutations in classical papillary thyroid carcinomas. *The Journal of clinical endocrinology and metabolism* **90**, 3028–3035, doi:10.1210/jc.2004-1394 (2005).
22. Guan, H. *et al.* Hypermethylation of the DNA mismatch repair gene hMLH1 and its association with lymph node metastasis and T1799A BRAF mutation in patients with papillary thyroid cancer. *Cancer* **113**, 247–255, doi:10.1002/cncr.23548 (2008).
23. Kim, Y. H. *et al.* TSH signaling overcomes B-RafV600E-induced senescence in papillary thyroid carcinogenesis through regulation of DUSP6. *Neoplasia (New York, N.Y.)* **16**, 1107–1120, doi:10.1016/j.neo.2014.10.005 (2014).
24. Rusinek, D. *et al.* BRAFV600E-Associated Gene Expression Profile: Early Changes in the Transcriptome, Based on a Transgenic Mouse Model of Papillary Thyroid Carcinoma. *PLoS one* **10**, e0143688, doi:10.1371/journal.pone.0143688 (2015).
25. Flockhart, R. J. *et al.* BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. *Genome research* **22**, 1006–1014, doi:10.1101/gr.140061.112 (2012).
26. Goedert, L. *et al.* RMEL3, a novel BRAFV600E-associated long noncoding RNA, is required for MAPK and PI3K signaling in melanoma. *Oncotarget*, **8**, doi:10.18632/oncotarget.9164 (2016).
27. Kapusta, A. & Feschotte, C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. *Trends in genetics: TIG* **30**, 439–452, doi:10.1016/j.tig.2014.08.004 (2014).
28. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137, doi:10.1038/379131a0 (1996).
29. Saxena, A. & Carninci, P. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. *BioEssays: news and reviews in molecular, cellular and developmental biology* **33**, 830–839, doi:10.1002/bies.201100084 (2011).
30. Huang, M. D. *et al.* Long non-coding RNA ANRIL is upregulated in hepatocellular carcinoma and regulates cell apoptosis by epigenetic silencing of KLF2. *Journal of hematology & oncology* **8**, 50, doi:10.1186/s13045-015-0146-0 (2015).
31. Yoon, J. H. *et al.* LincRNA-p21 suppresses target mRNA translation. *Molecular cell* **47**, 648–655, doi:10.1016/j.molcel.2012.06.027 (2012).
32. Yuan, S. X. *et al.* Antisense long non-coding RNA PCNA-AS1 promotes tumor growth by regulating proliferating cell nuclear antigen in hepatocellular carcinoma. *Cancer letters* **349**, 87–94, doi:10.1016/j.canlet.2014.03.029 (2014).
33. Yoon, J. H. *et al.* Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nature communications* **4**, 2939, doi:10.1038/ncomms3939 (2013).
34. Karreth, F. A. *et al.* The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma *in vivo*. *Cell* **161**, 319–332, doi:10.1016/j.cell.2015.02.043 (2015).
35. Greco, S., Gorospe, M. & Martelli, F. Noncoding RNA in age-related cardiovascular diseases. *Journal of molecular and cellular cardiology* **83**, 142–155, doi:10.1016/j.yjmcc.2015.01.011 (2015).
36. Lan, X. *et al.* Genome-wide analysis of long noncoding RNA expression profile in papillary thyroid carcinoma. *Gene* **569**, 109–117, doi:10.1016/j.gene.2015.05.046 (2015).
37. Xu, B. *et al.* The Long Non-Coding RNA ENST00000537266 and ENST00000426615 Influence Papillary Thyroid Cancer Cell Proliferation and Motility. *Cellular physiology and biochemistry: international journal of experimental cellular physiology, biochemistry, and pharmacology* **38**, 368–378, doi:10.1159/000438637 (2016).
38. Jendrzewski, J. *et al.* The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8646–8651, doi:10.1073/pnas.1205654109 (2012).
39. Zheng, H. *et al.* BRAF-activated Long Non-coding RNA Modulates Papillary Thyroid Carcinoma Cell Proliferation through Regulating Thyroid Stimulating Hormone Receptor. *Cancer research and treatment: official journal of Korean Cancer Association*. doi:10.4143/crt.2015.118 (2015).
40. He, H. *et al.* Genetic predisposition to papillary thyroid carcinoma: involvement of FOXE1, TSHR, and a novel lincRNA gene, PTCSC2. *The Journal of clinical endocrinology and metabolism* **100**, E164–172, doi:10.1210/jc.2014-2147 (2015).
41. Yoon, H. *et al.* Identification of a novel noncoding RNA gene, NAMA, that is downregulated in papillary thyroid carcinoma with BRAF mutation and associated with growth arrest. *International journal of cancer. Journal international du cancer* **121**, 767–775, doi:10.1002/ijc.22701 (2007).
42. Jendrzewski, J. *et al.* PTCSC3 Is Involved in Papillary Thyroid Carcinoma Development by Modulating S100A4 Gene Expression. *The Journal of clinical endocrinology and metabolism* **100**, 1370–1377, doi:10.1210/jc.2015-2247 (2015).
43. Li, W. B. *et al.* Identification of Genes Associated with Papillary Thyroid Carcinoma (PTC) for Diagnosis by Integrated Analysis. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme*, doi:10.1055/s-0035-1569289 (2016).
44. Pramoongjago, P., Baras, A. S. & Moskaluk, C. A. Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. *Oncogene* **25**, 5626–5639, doi:10.1038/sj.onc.1209566 (2006).
45. Matin, R. N. *et al.* p63 is an alternative p53 repressor in melanoma that confers chemoresistance and a poor prognosis. *The Journal of experimental medicine* **210**, 581–603, doi:10.1084/jem.20121439 (2013).
46. Mulder, K. M. Role of Ras and Mapks in TGFbeta signaling. *Cytokine & growth factor reviews* **11**, 23–35 (2000).
47. Chuderland, D. & Seger, R. Calcium regulates ERK signaling by modulating its protein-protein interactions. *Communicative & integrative biology* **1**, 4–5 (2008).

48. Rosen, L. B., Ginty, D. D., Weber, M. J. & Greenberg, M. E. Membrane depolarization and calcium influx stimulate MEK and MAP kinase via activation of Ras. *Neuron* **12**, 1207–1221, doi:10.1016/0896-6273(94)90438-3 (1994).
49. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* **2**, 401–404, doi:10.1158/2159-8290.cd-12-0095 (2012).
50. Li, J. *et al.* TANRIC: An Interactive Open Platform to Explore the Function of lncRNAs in Cancer. *Cancer research* **75**, 3728–3737, doi:10.1158/0008-5472.can-15-0273 (2015).
51. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108, doi:10.1038/nature11233 (2012).
52. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139–140, doi:10.1093/bioinformatics/btp616 (2010).
53. Colaprico, A. *et al.* TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Research*. <https://www.bioconductor.org/packages/release/bioc/html/TCGAAbiolinks.html> (2015).
54. Han, L. *et al.* lncRNA profile of glioblastoma reveals the potential role of lncRNAs in contributing to glioblastoma pathogenesis. *International journal of oncology* **40**, 2004–2012, doi:10.3892/ijo.2012.1413 (2012).
55. Terai, G., Iwakiri, J., Kameda, T., Hamada, M. & Asai, K. Comprehensive prediction of lncRNA-RNA interactions in human transcriptome. *BMC genomics* **17**(Suppl 1), 12, doi:10.1186/s12864-015-2307-5 (2016).
56. Miranda, K. C. *et al.* A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126**, 1203–1217, doi:10.1016/j.cell.2006.07.031 (2006).
57. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37**, 1–13, doi:10.1093/nar/gkn923 (2009).
58. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27–30, doi:10.1093/nar/28.1.27 (2000).
59. Lu, T. P. *et al.* miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. *PLoS one* **7**, e42390, doi:10.1371/journal.pone.0042390 (2012).
60. Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research* **38**, W214–220, doi:10.1093/nar/gkq537 (2010).

## Acknowledgements

This work was supported by grants to EME from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP- 2014/18189-5) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Research Fellowship- 309187/2015-0 and 506780/2013-9). FAPESP provided fellowships to LG (2014/07726-0), C.S.F. (2014/50521-0) and M.C.R.M. (2015/13396-5). Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) provided fellowship to L.G., J.R.P. and J.F.S. (CAPES-PROCAD-2961/2014). Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) provided fellowship to P.P.A. (6693/2015). CNPq provided fellowship to D.R.P. and A.C.C. and provided financial support to E.T.K. (458505/2014-5 and 308527/2013-5). L.G., J.R.P., W.A.S.J. and E.M.E. are members of Center For Cell-Therapy, CEPID/FAPESP (grant 2013/08135-2). Some of the results published here are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

## Author Contributions

L.G., J.R.P., C.S.F., M.C.R.M., D.R.P., P.P.A., T.P.S., J.F.S., A.C.C., I.E.G.P. and M.M.C. conducted the analyses and experiments. L.G., J.R.P., D.R.P., P.P.A., T.P.S., J.F.S., A.C.C., I.E.G.P., M.M.C., E.T.K., W.A.S.J. and E.M.E. contributed to the writing of the manuscript and jointly developed the structure and arguments for the paper. All authors reviewed and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-01957-0

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017