Long non-coding RNA landscapes in benign and malignant thyroid neoplasms of distinct histological subtypes

Valentina Yakushina (vdyakushina@gmail.com)
Research Centre of Medical Genetics https://orcid.org/0000-0001-5236-9297
Alexander Lavrov
Research Centre for Medical Genetics

Research article

Keywords: long non-coding RNA, SLC26A4-AS1, TNRC6C-AS1, CRNDE, RMST, follicular thyroid carcinomas, follicular adenoma, papillary thyroid carcinoma, anaplastic thyroid carcinoma

Posted Date: December 16th, 2019

DOI: https://doi.org/10.21203/rs.2.18943/v1

Abstract

Background: The main types of thyroid neoplasms, follicular adenoma (FA), follicular thyroid carcinoma (FTC), classical and follicular variants of papillary carcinoma (clPTC and fvPTC), anaplastic thyroid carcinoma (ATC), are differ in the prognosis, rate of progression and metastatic behavior. It can be supposed that there are specific patterns of lncRNAs involved in the development of clinical and morphological features. The lncRNA landscapes within distinct benign and malignant histological variants of thyroid neoplasm are unknown.

Methods: Comprehensive set of Microarray and RNA-Seq datasets was analyzed for the expression of lncRNAs in FA, FTC, fvPTC, clPTC and ATC. The potential biological functions were evaluated via coexpression and enrichment analysis.

Results and conclusion: Abberant expression of lncRNA in FA, FTC, fvPTC, clPTC and ATC was established. The lncRNAs common for benign and malignant neoplasms, specific for papillary carcinomas, specific for clPTC, fvPTC and ATC are determined. The determined common and specific lncRNAs are found to be putatively involved into L1CAM interactions; processing of capped intron-containing pre-mRNA; Tryptophan metabolism; PCP/CE pathway and Beta-catenin independent WNT signaling; extracellular matrix organization and cell cycle and mitotic. The patterns of lncRNA expression in FA and FTC are appeared to be similar with no genes significantly differentially expressed within these subtypes. Previously known oncogenic and supressor lncRNAs (NR2F1-AS1, LINC00511, SLC26A4-AS1, CRNDE, LINC01116, RMST) are found aberrantly expressed in thyroid carcinomas. The findings enhance the understanding of lncRNA landscape in thyroid neoplasms and its role in thyroid cancer progression.

Background

The most common types of thyroid cancer are papillary carcinomas (PTC) and follicular carcinomas (FTC). PTC and FTC account for about 70–80% and 10–15% of all thyroid cancers, respectively. Both PTC and FTC are differentiated carcinomas, but the mutational landscape and biological behavior, such as typical localization of metastasis (lymph nodes in the neck for PTC, distant organs, particularly, lungs and bones for FTC) and prognostic clinical markers are distinct [1]. Within PTC, several variants can be distinguished, where classical (clPTC) and follicular variants (fvPTC) are the most common. The fvPTC is composed of neoplastic follicles rather than papillae, but with follicular cells showing nuclear features characteristic of PTC, overall it has an indolent behavior. The benign counterpart of follicular carcinoma is follicular adenoma, and it is often challenging to differentiate them by cytology. Follicular adenomas (FA), FTC and fvPTC compose follicular-pattern thyroid tumors, sharing common mutational prevalence and clinical features [1, 2]. Anaplastic thyroid cancer (ATC) is the most advanced and aggressive thyroid cancer and the least likely to respond to treatment [3, 4].

The evidence of important role of long non-coding RNA (lncRNA) in tumor suppression, cancer progression, invasion and metastatic potential, and its prognostic and therapeutic value are increasing

[5]. LncRNAs are RNA molecules of more than 200 nucleotides, which typically do not have functional open reading frame (however, bifunctional RNAs were discovered which function as both, protein-coding and non-coding). Many lncRNA genes have two or more exons and display 5’-capping, poly-adenylation and alternative splicing. Functions of lncRNA are implemented through different ways: recruiting transcription factors, chromatin organizers, or chromatin modifiers, forming a DNA–RNA triplex anchoring effector proteins to the gene promoter, decoying miRNAs and proteins, or interfering with protein post-translational modification [5, 6, 7, 8]. Relative to the coding genes, lncRNA can be classified into intergenic (lincRNA); antisense (on the opposite strand of protein-coding locus); sense intronic or overlapping (on the same strand, transcript in introns of a coding gene, or contains a coding gene in its intron); retained intron (an alternatively spliced transcript containing intronic sequence); bidirectional (originates from the promoter region of a protein-coding gene, with transcription proceeding in the opposite direction on the other strand); 3-prime overlapping (overlap the 3'UTR of a protein-coding locus on the same strand). Today, the number of annotated lncRNA genes reached 14 720, according to Ensembl version 93 [9].

In thyroid cancer several lncRNAs were shown to have pathogenic and predictive role, including BANCR, FALEC, CNALPTC1, PVT1, NAMA, PTCSC1, PTCSC2, PTCSC3, TNRC6C-AS1 and others [10–21]. However, all of the studies considered only PTC and mostly none of the previous work took into account the difference between clPTC and fvPTC. There are no published studies describing landscapes of lncRNA in ATC, FTC and FA. Nevertheless, lncRNAs differently expressed in ATC could reflect anaplastic features and be strong prognostic factors. As morphology and behavior of FTC differ from PTC it can be proposed that the landscape of lncRNAs in FTC would be different from that of PTC. Investigation of lncRNAs common and specific for FA and FTC is important in understanding their relations and revealing differential diagnostic markers.

This study aimed to find out lncRNAs specific and common for main types of thyroid neoplasms (FA, FTC, fvPTC, clPTC and ATC). The expression data from microarray technology (8 datasets) and RNA-Seq technology (PRJEB11591 dataset and TCGA transcriptome data) were analyzed.

Materials & Methods

Microarray datasets

Microarray datasets from Affymetrix Human Genome U133 Plus 2.0 Array (Platform GPL570) were selected from GEO. The following datasets were included: GSE3467, GSE60542, GSE35570, GSE76039, GSE53157, GSE33630, GSE65144, GSE29265. A total of 107 samples of normal tissue and 32 fvPTC, 48 clPTC, 49 ATC were analyzed.

CEL files were downloaded and normalization was performed using gcrma R package. Microarray probes were annotated with Ensembl version 93 using biomaRt package [22].

RNA-Seq datasets

RNA-Seq dataset PRJEB11591 of Yoo SK et al. [23] was selected from EBI European Nucleotide Archive database (https://www.ebi.ac.uk/ena/data/view/PRJEB11591). PRJEB11591 is the most comprehensive available RNA-Seq dataset containing benign and malignant thyroid neoplasms (FA, FTC, fvPTC, clPTC). PRJEB11591 samples included 81 normal thyroid tissues (NT), 26 FA, 30 FTC, 48 fvPTC and 77 clPTC.

FASTQ files were downloaded, alignment was performed by hisat2 [24]. Counts were calculated using featureCounts (Rsubread package) with annotation by Ensembl version 93 and Ensemble gene ID as grouping attribute [25]. Genes with low counts (less than 2 count in number of samples exceeding the size of lowest sample group) were filtered out, TMM normalization (edgeR package) and voom method of limma R package were applied.

In TCGA transcriptomic data 58 NT, 356 clPTC and 101 fvPTC were selecte. Samples of metastases, and other minor histological subtypes were excluded. Raw counts (HTSeq – Counts Workflow Type, briefly, STAR 2-pass alignment followed by gene expression count assessment with HTSeq) were downloaded from Genomic Data Commons Data Portal (GDC, https://portal.gdc.cancer.gov/). Genes with low counts (les then 1 count in number of samples exceeding the size of lowest sample group) were filtered out, followed by TMM normalization (edgeR package) and voom analyses of limma [26].

Selection of lncRNA genes

Protein coding genes and genes attributed to Havana biotypes not related to lncRNA were filtered out of count matrices. Genes of the following Havana biotypes were included in the analyzes: lincRNA, antisense, 3-prime overlapping ncRNA, bidirectional promoter lncRNA, misc RNA, processed transcript, sense intronic, sense overlapping.

Statistical analysis

To identify differentially expressed lncRNAs, linear modelling using Limma package was performed [27]. Genes with FDR adjusted P-value ≤ 0.01 and fold change (FC) ≥ 2.0 were considered being differentially expressed. Hierarchical clustering heatmap analysis of differentially expressed genes was performed using coolmap of limma.

Validation

For clPTC and fvPTC, sets of genes found significantly differentially expressed at the previous step on Microarray, RNA-Seq PRJEB11591, and RNA-Seq TCGA datasets were processed with intersection. Genes found in all three datasets, and genes found in both RNA-Seq dataset but absent in microarray probes were considered as validated.

Evaluation of potential biological functions

To identify genes positively and negatively coexpressed with the differently expressed lncRNA pairwise Pearson correlation between the lncRNA and all the genes was calculated. Genes with an absolute $r \geq 0.7$ and a significant correlation (P-value < 0.05) were considered to be coexpressed. Enrichment of Gene Ontology (GO) Biological Process (2018), GO Molecular Function (2018), Kyoto Encyclopedia of Genes

and Genomes (KEGG, 2016) and Reactome (2016) terms was analyzed using Enrichr [28, 29]. Terms with adjusted P-value from Fisher's exact test ≤ 0.05 were considered significantly enriched.

Results

LncRNAs differentially expressed in thyroid neoplasms

Differential expression was evaluated in main histological subtypes of thyroid nodules: FA, FTC, fvPTC, clPTC, ATC. Differential expression was estimated in each of histological subtypes of thyroid nodule compared to normal tissue. Further, the sets of differentially expressed genes were compared against the other histological types to find specific patterns of expression.

Expression of 3910 lncRNA genes in microarray dataset, 2587 – in RNA-Seq PRJEB11591 dataset and 3009 – in RNA-Seq TCGA datasets was analyzed. Each dataset was examined separately with subsequent meta-analysis of robustly expressed genes via intersection of the results for clPTC and fvPTC (for the rest subtypes the only one dataset was available). For PTC differential expression of lncRNAs was considered validated in silico if it was registered in all 3 datasets, or in 2 RNA-seq datasets if a particular lncRNA is not covered by the microarray probes.

The numbers of differentially expressed lncRNA are represented in Table 1. The full lists of discovered differentially expressed lncRNA are in the Additional files 1–3. Volcano plots representing distribution of fold change and adjusted p-values in studied histological subtypes are in the Additional file 4.

By in silico validation, 116 genes were confirmed as validated in clPTC (45 genes found in all analyzed datasets, plus 71 genes without probes in microarray found in both RNA-Seq datasets; Fig. 1A), and 62 genes - in fvPTC (Fig. 1B) – robustly differentially expressed lncRNA.

Table 1. Number of lncRNAs differentially expressed in thyroid nodules compared to normal thyroid tissue

Histological type of the nodule	Microarray	RNA-Seq PRJEB11591	RNA-Seq THCA
FA		143	
FTC		213	
fvPTC	84	213	174
clPTC	137	401	308
ATC	330		

Clustering analyses was performed for microarray and PRJEB11591 datasets. It showed strong clustering of ATC, clustering of clPTC and weak clustering of fvPTC. Surprisingly there was no clustering within the groups of FTC and FA (Fig. 2).

LncRNAs common for benign and malignant thyroid neoplasms

There are LINC02555 and LINC02471 genes that are in Top 5 differently expressed lncRNAs in all studied thyroid neoplasms, including FA (Fig. 3, 4, Table 2). These lncRNA are validated in clPTC and fvPTC, and are differentially expressed in papillary carcinomas compared to FA.

LncRNAs common for differentiated thyroid carcinomas

There are 32 lncRNAs differentially expressed in all studied histological subtypes of differentiated carcinomas (FTC, clPTC, fvPTC) but not in FA (Fig. 3). Of them, 6 lncRNAs were validated and significantly differentially expressed in clPTC and fvPTC compared to FA (Fig. 4, Table 2). None of the 32 lncRNA was differentially expressed in FTC compared to FA.

Table 2. lncRNA common for thyroid neoplasms

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT			
			FA	FTC	fvPTC	clPTC
ENSG00000256268	LINC02454	lincRNA	-	2.0	4.2	4.5
ENSG00000225342	-	antisense	-	1.7	4.4	3.1
ENSG00000250343	STK32A-AS1	antisense	-	1.6	3.1	3.2
ENSG00000272384	-	lincRNA	-	1.1	1.9	2.5
ENSG00000233251	-	antisense	-	-1.4	-2.3	-1.6
ENSG00000254489	-	antisense	-	-1.8	-3.2	-3.9
ENSG00000260943	LINC02555	lincRNA	2.7	4.8	7.8	5.0
ENSG00000223914	LINC02471	lincRNA	2.0	4.2	7.2	6.5
Differential expression of all genes is validated in clPTC and fvPTC. Expression in fvPTC and clPTC differs significantly compared to FA.						

LncRNA specific for papillary carcinomas

There are 22 genes differentially expressed in both clPTC and fvPTC, but not in follicular neoplasms (Fig. 3), validated and significantly differentially expressed compared to FA and FTC (Fig. 4, Table 3) – lnRNAs associated with papillary features in thyroid carcinomas.

Table 3

LncRNA specific for papillary carcinomas

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT	
			fvPTC	clPTC
ENSG00000237463		antisense	4.2	6.4
ENSG00000203585	LINC02408	lincRNA	2.0	4.8
ENSG00000251002		antisense	4.2	4.7
ENSG00000272482		lincRNA	1.7	3.6
ENSG00000204282	TNRC6C-AS1	antisense	2.4	3.4
ENSG00000197301		antisense	2.4	3.4
ENSG00000267199		antisense	1.7	3.1
ENSG00000235978		antisense	2.2	2.5
ENSG00000230910		antisense	2.0	2.2
ENSG00000257989		lincRNA	2.5	2.2
ENSG00000224020	MIR181A2HG	antisense	1.7	1.8
ENSG00000272079		lincRNA	2.1	1.7
ENSG00000272512		lincRNA	1.0	1.6
ENSG00000237742		antisense	1.9	1.4
ENSG00000255366		lincRNA	1.1	1.3
ENSG00000265666	RARA-AS1	antisense	1.4	1.0
ENSG00000204934	ATP6V0E2-AS1	antisense	-1.1	-1.5
ENSG00000228559		lincRNA	-1.2	-1.6
ENSG00000234899	SOX9-AS1	lincRNA	-1.2	-1.7
ENSG00000228613		antisense	-2.2	-2.6
ENSG00000267034		lincRNA	-2.6	-2.9
ENSG00000261399		antisense	-2.0	-3.3

All these lncRNAs are differentially expressed compared to follicular thyroid carcinoma and follicular adenoma.

LncRNA specific for histological subtypes of differentiated carcinomas

There are 20 lncRNAs aberrantly expressed in FTC, but not in other studied neoplasms, and significantly differentially expressed compared to PTC (Table 4). However, none of these lncRNAs was differentially expressed compared to FA.

Table 4. lncRNA specific for FTC

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT
ENSG00000281383	-	lincRNA	1.4
ENSG00000272732	-	lincRNA	-1.0
ENSG00000224660	SH3BP5-AS1	antisense	-1.0
ENSG00000225855	RUSC1-AS1	antisense	-1.0
ENSG00000197989	SNHG12	antisense	-1.0
ENSG00000198221	AFDN-DT	lincRNA	-1.1
ENSG00000248019	FAM13A-AS1	antisense	-1.1
ENSG00000273576	-	lincRNA	-1.1
ENSG00000261087	-	lincRNA	-1.1
ENSG00000271895	-	antisense	-1.2
ENSG00000242282	-	lincRNA	-1.2
ENSG00000272374	-	lincRNA	-1.2
ENSG00000204584	-	antisense	-1.3
ENSG00000262370	-	lincRNA	-1.3
ENSG00000205959	-	lincRNA	-1.3
ENSG00000285103	-	bidirectional_promoter_lncRNA	-1.4
ENSG00000276007	-	sense_intronic	-1.4
ENSG00000226419	SLC16A1-AS1	antisense	-1.5
ENSG00000257671	KRT7-AS	antisense	-1.6

The 32 genes were found being differentially expressed in clPTC but not in other differentiated carcinomas and FA, validated, and significantly differentially expressed compared to fvPTC, FTC and FA - lncRNA specific for clPTC (Fig. 3, 4, Table 5).

Table 5

LncRNA specific for classical variant of papillary carcinoma

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT
ENSG00000227036	LINC00511	lincRNA	2.5
ENSG00000237187	NR2F1-AS1	antisense	2.5
ENSG00000260604	-	lincRNA	2.2
ENSG00000262903	-	antisense	2.1
ENSG00000261101	-	sense_overlapping	1.9
ENSG00000274021	-	antisense	1.8
ENSG00000281406	BLACAT1	lincRNA	1.5
ENSG00000245571	FAM111A-DT	lincRNA	1.3
ENSG00000253930	TNFRSF10A-AS1	antisense	1.3
ENSG00000235609	-	lincRNA	1.2
ENSG00000237943	PRKCQ-AS1	lincRNA	-1.0
ENSG00000260572	-	antisense	-1.0
ENSG00000204860	FAM201A	antisense	-1.0
ENSG00000177640	CASC2	antisense	-1.1
ENSG00000259704	-	sense_overlapping	-1.1
ENSG00000231769	-	antisense	-1.2
ENSG00000231231	LINC01423	lincRNA	-1.2
ENSG00000272622	-	lincRNA	-1.2
ENSG00000251602	-	antisense	-1.3
ENSG00000231856	-	antisense	-1.3
ENSG00000249249	-	antisense	-1.3
ENSG00000205791	LOH12CR2	lincRNA	-1.3
ENSG00000232415	ELN-AS1	antisense	-1.5
ENSG00000262185	-	sense_overlapping	-1.8
ENSG00000224885	EIPR1-IT1	sense_intronic	-1.9
ENSG00000256151	ADGRD1-AS1	lincRNA	-1.9

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT
ENSG00000231107	LINC01508	lincRNA	-2.0
ENSG00000267128	RNF157-AS1	antisense	-2.0
ENSG00000229457	LINC01789	lincRNA	-2.0
ENSG00000249487	LINC01586	lincRNA	-2.7
ENSG00000224568	LINC01886	lincRNA	-2.7
ENSG00000233705	SLC26A4-AS1	antisense	-3.0

Of 29 genes differently expressed in fvPTC but not in other differentiated carcinomas or FA (Fig. 3), only ENSG00000257647 gene is specific for fvPTC - validated and significantly differentially expressed in fvPTC compared to FA, FTC and clPTC.

LncRNA specific for ATC

ATC samples were available only in microarray dataset, which also included two variants of PTC. Out of 376 lncRNAs differentially expressed in ATC, 252 were not differentially expressed in other investigated histological subtypes, and 185 were significantly differentially expressed compared to clPTC and fvPTC – lncRNAs specific for ATC. Top 30 genes are represented in Table 6, the full list is in the Additional file 5.

Table 6
Top 30 lncRNA specific for anaplastic carcinoma

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT
ENSG00000272872	-	sense_intronic	3.5
ENSG00000244158	-	antisense	3.2
ENSG00000245694	CRNDE	lincRNA	3.2
ENSG00000240476	LINC00973	lincRNA	3.1
ENSG00000282638	-	lincRNA	3.0
ENSG00000247134	-	lincRNA	3.0
ENSG00000254615	-	lincRNA	2.9
ENSG00000280018	-	lincRNA	2.8
ENSG00000233682	-	antisense	-2.5
ENSG00000266904	LINC00663	lincRNA	-2.6
ENSG00000228506	-	antisense	-2.6
ENSG00000275234	-	antisense	-2.7
ENSG00000232229	LINC00865	lincRNA	-2.8
ENSG00000269609	RPARP-AS1	lincRNA	-2.9
ENSG00000270820	-	antisense	-2.9
ENSG00000271474	-	antisense	-3.0
ENSG00000284644	-	antisense	-3.0
ENSG00000273015	-	lincRNA	-3.1
ENSG00000260686	-	sense_overlapping	-3.2
ENSG00000180769	WDFY3-AS2	antisense	-3.2
ENSG00000247400	DNAJC3-DT	lincRNA	-3.2
ENSG00000271858	-	antisense	-3.3
ENSG00000236155	-	processed_transcript	-3.3
ENSG00000224078	SNHG14	antisense	-3.5
ENSG00000250073	-	antisense	-3.8
ENSG00000203709	MIR29B2CHG	lincRNA	-4.1

ENSG ID	HGNC symbol	Gene biotype	Log(FC) compared to NT
ENSG00000261183	SPINT1-AS1	antisense	-4.2
ENSG00000229891	LINC01315	lincRNA	-4.3
ENSG00000257151	PWAR6	lincRNA	-4.8
ENSG00000255794	RMST	lincRNA	-5.7

Potential biological functions of aberrantly expressed lncRNAs

For lncRNAs common for studied histological subtypes and top 5 specific lncRNA in each subtype, enrichment of Gene Ontology (GO) biological processes, GO molecular functions, KEGG, and Reactome terms with coexpressed coding genes was established. The main related functions of the common lncRNAs is L1CAM interactions; lncRNA specific for FTC - processing of capped intron-containing pre-mRNA; specific for papillary carcinomas - Tryptophan metabolism; specific for fvPTC - PCP/CE pathway and Beta-catenin independent WNT signaling; specific for clPTC - extracellular matrix organization; specific for ATC - cell cycle and mitotic (Fig. 5).

Discussion

Histological subtypes of follicular cell-derived thyroid carcinomas (FTC, PTC, ATC) significantly differ in their mutational landscapes and clinical characteristics. Although FTC and clPTC both are differentiated carcinomas, FTC are characterized by follicular growth pattern and tend more often to spread metastases to distant organs, while clPTC typically have papillary architecture and spread more often to lymph nodes in the neck. In FTC K/H/NRAS and PAX8/PPARG mutations are prevalent, whereas BRAF mutations and tyrosine kinase fusions prevail in clPTC [1]. fvPTC are composed of neoplastic follicles rather than papillae, but with follicular cells showing nuclear features of PTC. The clinical characteristics of fvPTC are intermediate [30]. The mutational profile of fvPTC is more similar to FTC: prevalence of K/H/NRAS and PAX8/PPARG mutations. In the TCGA study, fvPTC were characterized as Ras-like tumors and their classification as papillary carcinomas was questioned [31]. Recently, reclassification of encapsulated fvPTC as “noninvasive follicular thyroid neoplasm with papillary-like nuclear features” (NIFTP) was proposed [2]. FA are thought to be benign counterpart of FTC and understanding the common and different molecular features of these neoplasms is important for the development of diagnostic and therapeutic strategy. ATC are thought to be an advanced stage of thyroid neoplasms, it is the most aggressive thyroid cancer and it is expected that there are specific molecular features, including lncRNA, associated with its aggressive behavior.

In this study, expression of lncRNA was evaluated in the main histological subtypes of thyroid neoplasm: FA, FTC, fvPTC, clPTC and ATC. Datasets analyzed in the study (Microarray dataset of 8 independent experiments; RNA-Seq PRJEB11591 and RNA-Seq TCGA) allowed to perform robust in silico validation of

the results for clPTC and fvPTC, and to include representative set of FA, FTC and ATC samples. LncRNA landscapes in FA, FTC and ATC were analyzed for the first time.

LncRNAs general for thyroid neoplasms, specific for clPTC, fvPTC, ATC and general for papillary carcinomas (associated with papillary features) were discovered. The highest number of differently expressed genes was in ATC (330 lncRNAs) followed by clPTC (137), FTC and fvPTC that reflects the more advanced stage of ATC. The lncRNAs specific for ATC and probably associated with anaplastic features were discovered for the first time. None of lncRNAs specific for FTC is differentially expressed compared to FA. The only lncRNA ENSG00000257647 is specific for fvPTC that might be explained by its intermediate morphology combining features of both papillary and follicular carcinomas and debatable classification.

Aberrant expression of some lncRNAs previously found by Liyanarachchi S. et all. (2016) was confirmed in our study. Most of these lncRNAs occurred to be general for thyroid carcinomas [14]. Previously known promoters of cancer progression were found upregulated in thyroid carcinomas: NR2F1-AS1 and LINC00511 – in clPTC; TNRC6C-AS1 – in clPTC and fvPTC; CRNDE – in ATC. Known putative tumor suppressors were identified within downregulated lncRNAs: SLC26A4-AS1 – in clPTC and ATC; RMST – in ATC [32–38].

Putative biological process involving common and specific lncRNAs were established. LncRNAs common for all studied thyroid neoplasms might be involved in L1CAM interactions; common for follicular and classical variants of papillary carcinoma – in Tryptophan metabolism. Tryptophan degradation to kynurenine by the Indoleamine 2,3-Dioxygenase 1 (IDO1) is a well characterized immunosuppressive mechanism in cancer progression, including thyroid cancer [39].

Biological processes involving lncRNAs specific for FTC include processes that are associated with splicing (Processing of Capped Intron-Containing Pre-mRNA, mRNA Splicing, RNA processing). Accumulating evidence suggests that aberrant RNA splicing is a common and driving event in cancer development and progression. For instance, oncogenic Ras signaling via ERK and PI3-K/Akt pathways is described to regulate phosphorylation of splicing factors such as SRSF1, SRSF7, SPF45 and drive switching active and inactive states of tumor promoters and suppressors (MST1R, FAS, CD44, LBR, Casp-9, KLF6, and others) via alternative splicing [40, 41].

LncRNA ENSG00000257647 specific for fvPTC appeared to be involved in WNT signaling, predominantly Beta-catenin independent WNT pathway (especially, planar cell polarity that modulates cytoskeleton rearrangements through the activation of the small GTPases RhoA and Rac and their downstream effectors Rock and JNK). WNT signaling is described to play a crucial role in thyroid carcinogenesis, several mechanisms of its deregulation were described, including inhibition of β-catenin degradation complex via its phosphorylation by RET/PTC and decrease of E-cadherin expression by MAPK/ERK pathway activated by BRAF mutations. RAS mutations are described to activate both, canonical and non-canonical Wnt pathways in thyroid carcinomas [42, 43].

LncRNAs specific for clPTC are involved into extracellular matrix organization and collagen formation. Extracellular matrix (ECM) disorganization is known to play a pivotal role in cancer initiation and progression. There is emerging evidence of ECM remodeling induced by BRAF p.V600E in PTCs [44]. Notably, it was shown that extracellular matrix of PTCs driven by BRAF p.V600E (but not mutant HRAS) is enriched with stromal-derived fibrillar collagen and it facilitates cancer progression [45].

For lncRNAs specific for ATC there is a strong enrichment of cell cycle and mitotic pathways which possibly reflects involvement of these lncRNAs in the loss of differentiation and high proliferation rate characteristic for ATC.

Conclusion

LncRNAs common for thyroid neoplasms, common for carcinomas with papillary features, specific for clPTC, fvPTC, FTC and ATC were discovered in the performed analyses of the most comprehensive dataset (combined of Microarray dataset and two RNA-Seq datasets). Similarity of lncRNA landscapes in FTC and FA was revealed. LncRNAs found to be specific for ATC are probably associated with anaplastic features and cancer progression.

Declarations

Ethics approval and consent to participate. Not applicable

Consent for publication. Not applicable

Availability of data and materials. The datasets used and analysed during the current study are available in the following repositories:

GEO (https://www.ncbi.nlm.nih.gov/geo/, GSE3467, GSE60542, GSE35570, GSE76039, GSE53157, GSE33630, GSE65144, GSE29265);

Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/);

EBI European Nucleotide Archive database (https://www.ebi.ac.uk/ena/data/view/PRJEB11591).

Competing interests. The authors declare that they have no competing interests

Funding. The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation

Authors' contributions. VY performed bioinformatic and statistical analysis and wrote the manuscript. AL supervised the work. All authors read and approved the final manuscript.

Acknowledgements. Not applicable

References

1. Acquaviva G, Visani M, Repaci A, Rhoden KJ, de Biase D, Pession A, et al. Molecular pathology of thyroid tumours of follicular cells: a review of genetic alterations and their clinicopathological relevance. Histopathology. 2018;72(1):6–31.
2. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma. JAMA Oncol. 2016;2(8):1023-1029.
3. Cabanillas ME, Zafereo M, Gunn GB, Ferrarotto R. Anaplastic Thyroid Carcinoma: Treatment in the Age of Molecular Targeted Therapy. J Oncol Pract. 2016;12(6): 511–518.
4. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J Clin Invest. 2016;126(3): 1052-1066.
5. Bure IV, Kuznetsova EB, Zaletaev DV. Long Noncoding RNAs and Their Role in Oncogenesis. Mol Biol. 2018;52(6):787-798.
6. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012;22(9):1775–1789.
7. Fang Y, Fullwood MJ. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. Genomics Proteomics Bioinformatics. 2016;14(1):42–54.
8. Li R, Zhu H, Luo Y. Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures. Int J Mol Sci. 2016;17(5):E702.
9. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Research. 2018;46(D1):D754-D761.
10. Yoon H, He H, Nagy R, Davuluri R, Suster S, Schoenberg D, 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. Int J Cancer. 2007;121(4):767–775.
11. He H, Nagy R, Liyanarachchi S, Jiao H, Li W, Suster S, et al. A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. Cancer Res. 2009;69:625–631.
12. Jeong S, Lee J, Kim D, Seol MY, Lee WK, Jeong JJ, et al. Relationship of Focally Amplified Long Noncoding on Chromosome 1 (FAL1) lncRNA with E2F Transcription Factors in Thyroid Cancer. Med. 2016;95:e2592.
13. Liao T, Qu N, Shi RL, Guo K, Ma B, Cao YM, et al. BRAF-activated LncRNA functions as a tumor suppressor in papillary thyroid cancer. Oncotarget. 2016;8(1):238-247.
14. Liyanarachchi S, Li W, Yan P, Bundschuh R, Brock P, Senter L, et al. Genome-wide expression screening discloses long noncoding RNAs involved in thyroid carcinogenesis. J Clin Endocrinol Metab. 2016;101(11):4005-4013.

15. Wang Q, Yang H, Wu L, Yao J, Meng X, Jiang H, et al. Identification of Specific Long Non-Coding RNA Expression: Profile and Analysis of Association with Clinicopathologic Characteristics and BRAF Mutation in Papillary Thyroid Cancer. Thyroid. 2016;26 (12):1719–1732.
16. Zheng H, Wang M, Jiang L, Chu H, Hu J, Ning J, et al. BRAF-Activated Long Noncoding RNA Modulates Papillary Thyroid Carcinoma Cell Proliferation through Regulating Thyroid Stimulating Hormone Receptor. Cancer Res Treat. 2016;48(2):698–707.
17. Zhou Q, Chen J, Feng J, Wang J. Long noncoding RNA PVT1 modulates thyroid cancer cell proliferation by recruiting EZH2 and regulating thyroid-stimulating hormone receptor (TSHR). Tumor Biol. 2016;37(3):3105–3113.
18. Muhanhali D, Zhai T, Jiang J, Ai Z, Zhu W, Ling Y. Long Non-coding Antisense RNA TNRC6C-AS1 Is Activated in Papillary Thyroid Cancer and Promotes Cancer Progression by Suppressing TNRC6C Expression. Front Endocrinol (Lausanne). 2018;9:360.
19. Hou S, Lin Q, Guan F, Lin C. LncRNA TNRC6C-AS1 regulates UNC5B in thyroid cancer to influence cell proliferation, migration, and invasion as a competing endogenous RNA of miR-129-5p. J Cell Biochem. 2018;119:8304–8316.
20. Chen C, Zhou L, Wang H, Chen J, Li W, Liu W, et al. Long noncoding RNA CNALPTC1 promotes cell proliferation and migration of papillary thyroid cancer via sponging miR-30 family. Am J Cancer Res. 2018;8(1):192–206.
21. Huang H, Chen J, Ding C-M, Jin X, Jia Z-M, Peng J. LncRNA NR2F1-AS1 regulates hepatocellular carcinoma oxaliplatin resistance by targeting ABCC1 via miR-363. J Cell Mol Med. 2018;22:3238–3245.
22. Yoo SK, Lee S, Kim SJ, Jee HG, Kim BA, Cho H, et al. Comprehensive Analysis of the Transcriptional and Mutational Landscape of Follicular and Papillary Thyroid Cancers. PLoS Genet. 2016; 12(8):e1006239.
23. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4:1184–1191.
24. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–360.
25. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 2013;41(10):e108.
26. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40(10):4288–4297.
27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43(7):e47.
28. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.
29. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1): W90-7.

30. Daniels GH. Follicular Variant of Papillary Thyroid Carcinoma: Hybrid or Mixture?. Thyroid. 2016;26(7):872-874.
31. Agrawal N, Akbani R, Aksoy BA, Ally A, Arachchi H, Asa SL, et al. Integrated genomic characterization of papillary thyroid carcinoma. Cell. 2014;159(3):676–690.
32. Sun C-C, Li S-J, Li G, Hua R-X, Zhou X-H, Li D-J. Long Intergenic Noncoding RNA 00511 Acts as an Oncogene in Non–small-cell Lung Cancer by Binding to EZH2 and Suppressing p57. Mol Ther - Nucleic Acids. 2016;5;e385.
33. Xu S, Kong D, Chen Q, Ping Y, Pang D. Oncogenic long noncoding RNA landscape in breast cancer. Mol Cancer. 2017;16(1):129.
34. Han P, Li JW, Zhang BM, Lv JC, Li YM, Gu XY, et al. The lncRNA CRNDE promotes colorectal cancer cell proliferation and chemoresistance via miR-181a-5p-mediated regulation of Wnt/β-catenin signaling. Mol Cancer. 2017;16(1):9.
35. Xie H, Ma B, Gao Q, Zhan H, Liu Y, Chen Z, et al. Long non-coding RNA CRNDE in cancer prognosis: Review and meta-analysis. Clin Chim Acta. 2018;485:262–71.
36. Brodie S, Lee HK, Jiang W, Cazacu S, Xiang C, Poisson LM, et al. The novel long non-coding RNA TALNEC2, regulates tumor cell growth and the stemness and radiation response of glioma stem cells. Oncotarget. 2017;8(19):31785–1801.
37. He H-T, Xu M, Kuang Y, Han X-Y, Wang M-Q, Yang Q. Biomarker and competing endogenous RNA potential of tumor-specific long noncoding RNA in chromophobe renal cell carcinoma. Onco Targets Ther. 2016;9:6399–6406.
38. Wang L, Liu D, Wu X, Zeng Y, Li L, Hou Y, et al. Long non-coding RNA (LncRNA) RMST in triple-negative breast cancer (TNBC): Expression analysis and biological roles research. J Cell Physiol. 2018;233(10):6603–6612.
39. Moretti S, Menicali E, Voce P, Morelli S, Cantarelli S, Sponziello M, et al. Indoleamine 2,3-Dioxygenase 1 (IDO1) Is Up-Regulated in Thyroid Carcinoma and Drives the Development of an Immunosuppressant Tumor Microenvironment. J Clin Endocrinol Metab. 2014;99:E832–40.
40. Gonçalves V, Pereira JFS, Jordan P. Signaling Pathways Driving Aberrant Splicing in Cancer Cells. Genes (Basel). 2017;9(1):9.
41. Yea S, Narla G, Zhao X, Garg R, Tal-Kremer S, Hod E, et al. Ras promotes growth by alternative splicing-mediated inactivation of the KLF6 tumor suppressor in hepatocellular carcinoma. Gastroenterology. 2008;134(5):1521-1531.
42. Sastre-Perona A, Riesco-Eizaguirre G, Zaballos MA, Santisteban P. β-catenin signaling is required for RAS-driven thyroid cancer through PI3K activation. Oncotarget. 2016;7:49435–49449.
43. Ely KA, Bischoff LA, Weiss VL. Wnt Signaling in Thyroid Homeostasis and Carcinogenesis. Genes (Basel). 2018;9(4):E204.
44. Nucera C, Lawler J, Parangi S. BRAF(V600E) and microenvironment in thyroid cancer: a functional link to drive cancer progression. Cancer Res. 2011;71(7): 2417–2422.

45. Jolly LA, Novitskiy S, Owens P, Massoll N, Cheng N, Fang W, et al. Fibroblast-Mediated Collagen Remodeling Within the Tumor Microenvironment Facilitates Progression of Thyroid Cancers Driven by BrafV600E and Pten Loss. Cancer Res. 2016;76(7):1804-1813.

Additional Material Information

Additional file 1.xlsx – lncRNA differently expressed in thyroid neoplasms in Microarray dataset;

Additional file 2.xlsx – lncRNA differently expressed in thyroid neoplasms in RNA-Seq PRJEB11591 dataset;

Additional file 3.xlsx – lncRNA differently expressed in thyroid neoplasms in RNA-Seq TCGA dataset;

Additional file 4.pdf – Volcano plots of lncRNA differently expressed in thyroid neoplasms in Microarray, RNA-Seq PRJEB11591 and TCGA dataset;

Additional file 5.txt –lncRNA specific for anaplastic thyroid cancer.

Figures

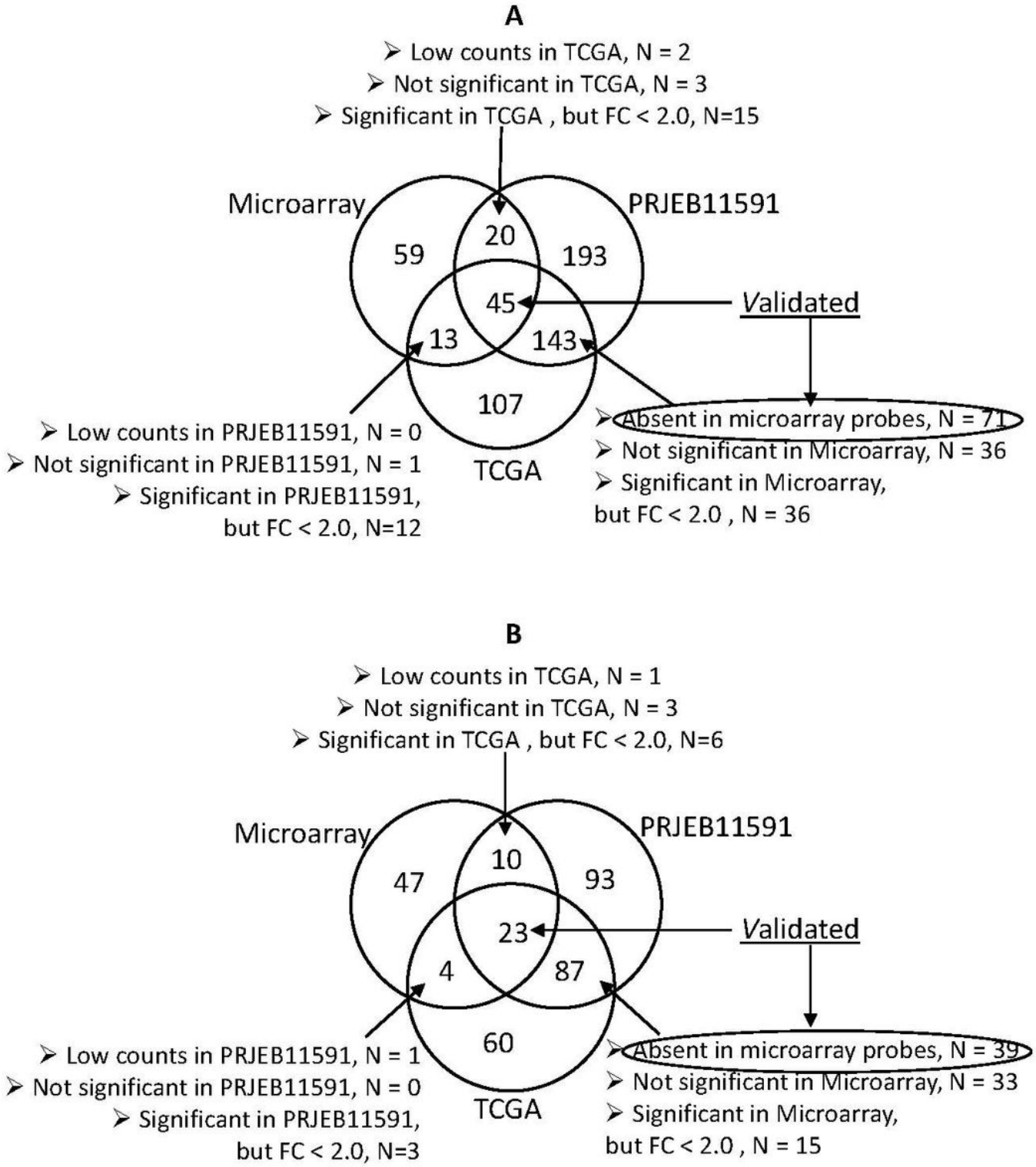

Figure 2

In silico validation of differently expressed lncRNA in clPTC (A) and fvPTC (B). In clPTC, 116 genes were considered to be validated (differently expressed in all datasets, or differently expressed in both RNA-Seq datasets, but absent in microarray probes). In fvPTC 62 genes can be considered to be validated.

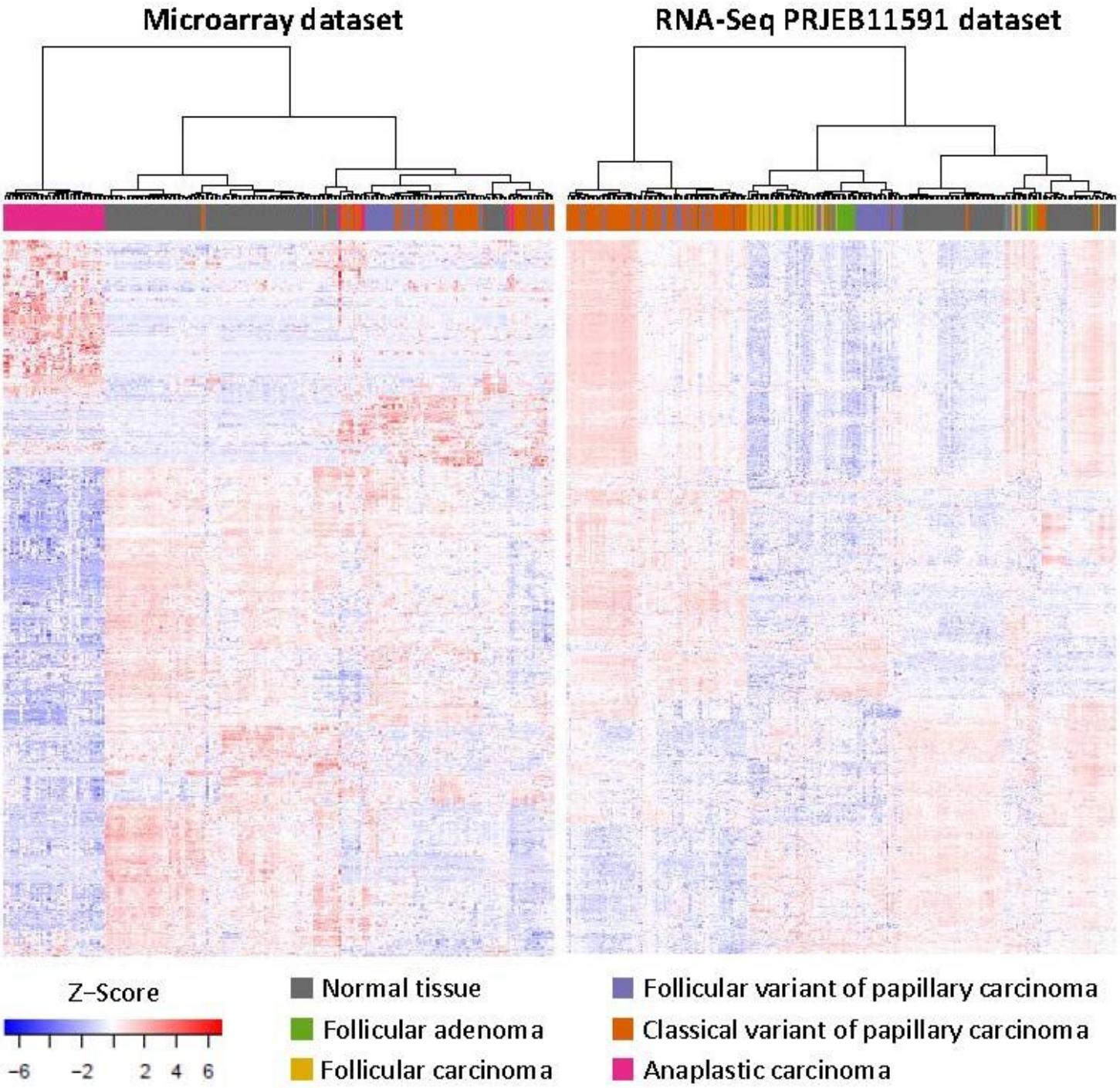

Figure 4

Clustering of FA, FTC, fvPTC, clPTC and ATC by the expression of lncRNA. A - Microarray dataset; B - RNA-Seq PRJEB11591 dataset. Genes differently expressed in each histological subtype are included.

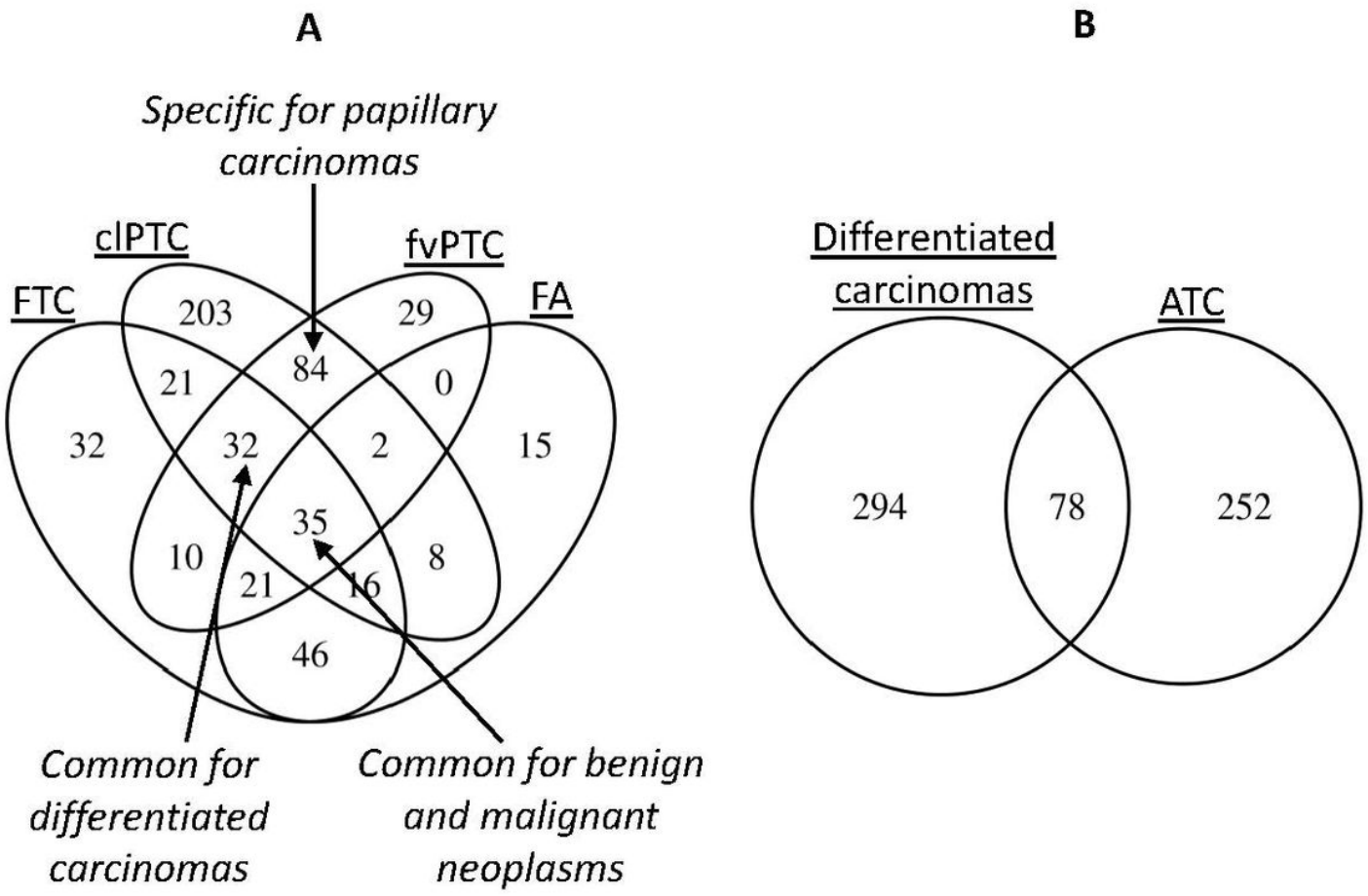

Figure 6

Overlapping of lncRNA landscapes in thyroid neoplasms. A – Overlapping of lncRNA landscapes in differentiated neoplasms – FA, FTC, fvPTC, clPTC (RNA-Seq PRJEB11591 dataset is used). B – Overlapping of lncRNA landscapes in differentiated neoplasms (total list of lncRNA differentially expressed in any of FA, FTC, fvPTC, clPTC) and ATC.

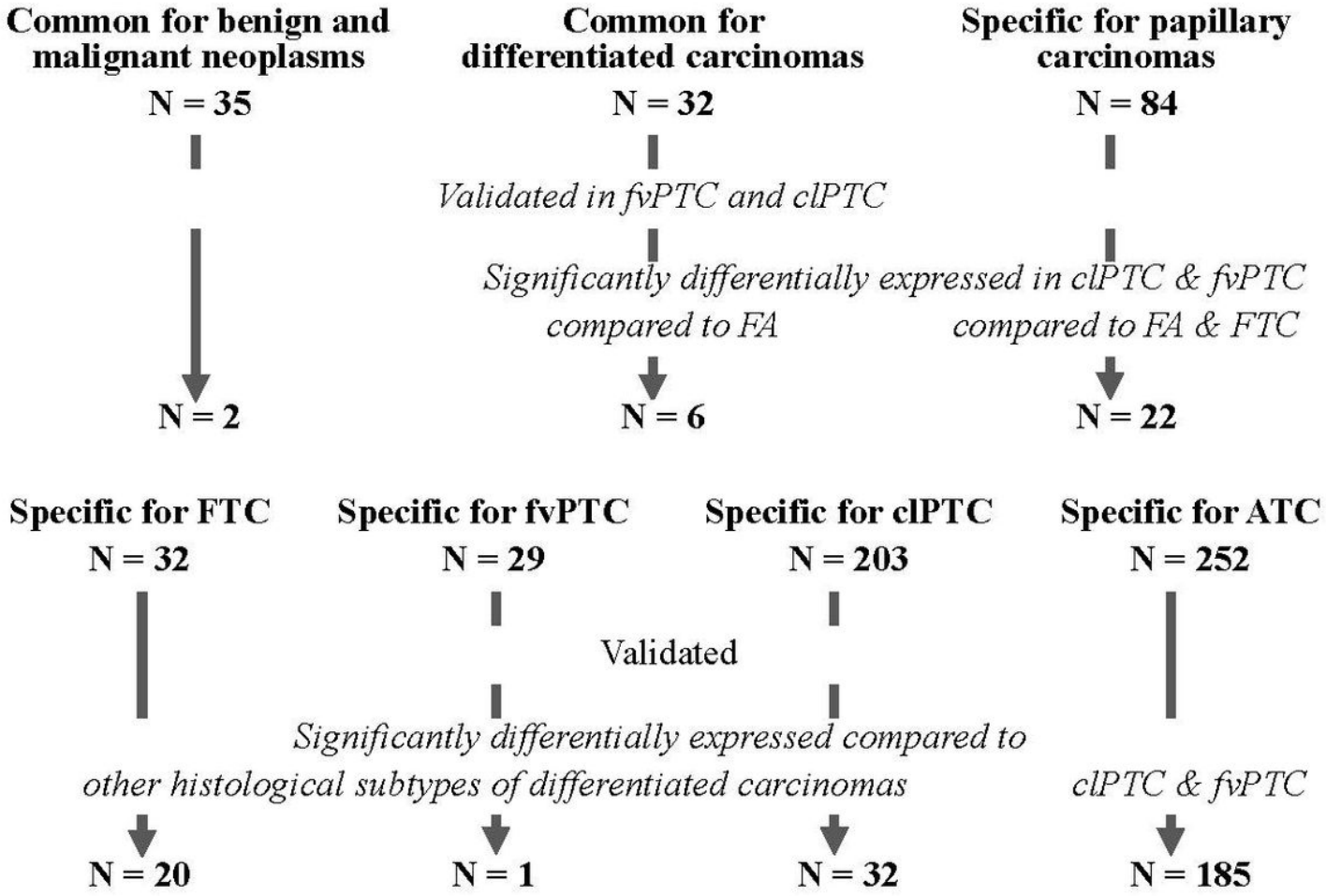

Figure 7

Results of the selection of common and specific lncRNA.

Figure 10

Putative biological process employing aberrantly expressed lncRNAs in thyroid neoplasms. Enrichment analysis of GO Biological Process, KEGG, and Reactome terms was performed for lncRNA common for thyroid neoplasms, and top 5 lncRNA general for papillary carcinomas, specific for fvPTC, clPTC, ATC, or FTC. Terms with adjusted P value <= 0.01 were considered significantly enriched. For fvPTC and ATC top 10 significantly enriched terms are represented.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile4.pdf
- Additionalfile5.txt
- Additionalfile3.xlsx
- Additionalfile2.xlsx
- Additionalfile1.xlsx
- Additionalfile4.pdf
- Additionalfile2.xlsx
- Additionalfile1.xlsx
- Additionalfile5.txt
- Additionalfile3.xlsx