Endocrine (2019) 66:607–613
https://doi.org/10.1007/s12020-019-02063-z

ORIGINAL ARTICLE

Check for updates

Heterogenous nature of gene expression patterns in BRAF-like papillary thyroid carcinomas with BRAF^{V600E}

Yuji Nagayama[1] · Hiroyuki Mishima[2]

Received: 25 June 2019 / Accepted: 20 August 2019 / Published online: 2 September 2019

Abstract

Purpose Papillary thyroid cancers (PTCs) are the most common type of thyroid cancers, in which BRAF^{V600E} is the most prevalent driver mutation. It is known that BRAF^{V600E}-positive PTCs are clinically and molecularly heterogenous in terms of aggressiveness and prognosis. The molecular mechanisms of this heterogeneity were evaluated.

Methods The publicly available RNA-seq data for 26 classical (c) and 5 follicular variant (fv) PTCs with BRAF^{V600E} mutation and the BRAF-like expression signature in the *BRAF^{V600E}-RAS* score (BRS) and their respective normal adjacent tissues were downloaded, and analyzed for differentially expressed genes (DEGs). The DEGs were then analyzed with the Gene Ontology annotation and the KEGG pathway dataset.

Results We found four lines of evidence for heterogeneity of cPTCs. First, the cluster dendrogram and principle component analyses could not completely distinguish the cancer tissues from normal tissues. Second, the DEGs identified in each sample were highly diverse from one another. Third, although the DEGs were enriched in many terms containing the word "extracellular" ("extracellular region", "extracellular space", and so on) when analyzed as groups, the degree of this enrichment was variable when analyzed individually. Fourth, there are only a few intersections in the over-/underexpressed genes annotated with the terms containing the word "extracellular" among the samples examined. Essentially same results were obtained with BRAF-like, fvPTCs with BRAF^{V600E}. Nevertheless, some frequently over-/underexpressed genes were detected, of which *LIPH* (*lipase H*) expression was found to be prognostic and its high expression was favorable for PTCs.

Conclusion Groups of BRAF-like, BRAF^{V600E}-positive cPTCs and fvPTCs that are homogenous in regard to histopathology, driver mutation and BRS were found to be highly heterogenous in terms of gene expression patterns. Yet, among the genes that were annotated with the terms containing the word "extracellular" and frequently over-/underexpressed, *LIPH* is a favorable prognostic marker for PTCs.

Keywords Thyroid cancer · RNA-seq · Differentially expressed genes · Heterogeneity

Introduction

Thyroid cancer is the most common endocrine disease, and its incidence has been increasing in recent decades [1]. Thyroid cancers include differentiated types of papillary (PTC) and follicular (FTC) thyroid cancers and poorly differentiated (PDTC) and anaplastic (ATC) thyroid cancers. PTCs account for >80% of all thyroid cancers. Although PTCs in general possess good prognosis, a small portion of patients are associated with poor prognosis [2] and sometimes experience metastasis and/or progression to PDTC/ATC, which have extremely poor prognosis. Differences in histopathology of PTCs help prognostic prediction to some extent; for example, tall cells, columnar, hobnail, and solid variants have been reported to be associated with more aggressive phenotypes and less favorable prognosis [3]. Variations in the genetic/genomic abnormalities may also explain the prognostic heterogeneity of PTCs. It has been widely demonstrated that the mutant BRAF (BRAF^{V600E}) is the most common driver mutation,

Supplementary information The online version of this article (https://doi.org/10.1007/s12020-019-02063-z) contains supplementary material, which is available to authorized users.

✉ Yuji Nagayama
nagayama@nagasaki-u.ac.jp

[1] Department of Molecular Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan

[2] Department of Human Genetics, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan

followed by RET/PTC chromosomal rearrangement and the mutant RAS, in PTCs [4]. The recent prevalence of techniques for high-throughput analysis of genomic/genetic landscapes of PTCs such as DNA microarray and next generation sequencing has revealed that PTCs can be divided into two subtypes: *BRAF*-like and *RAS*-like based on the 71-gene expression signature (the $BRAF^{V600E}$-*RAS* score (BRS)) [4]. The *BRAF*-like PTCs have classical or tall cell morphology, high levels of mitogen-activated protein kinase (MAPK) pathway signaling, less differentiated status and high risk for recurrence, while the *RAS*-like PTCs have follicular growth pattern with opposite characteristics [4, 5]. However, it is also known that even $BRAF^{V600E}$-positive PTCs are clinically and molecularly heterogenous when compared with mutant RAS-positive PTCs [5, 6], and the mechanisms for the heterogenous nature of $BRAF^{V600E}$-positive PTCs remain to be fully elucidated.

To attempt to solve this issue, we used publicly available RNA-seq data from classical (c) and follicular variant (fv) PTCs with $BRAF^{V600E}$ mutation and the *BRAF*-like expression signature and matched cancer-adjacent normal tissues [7], and analyzed differentially expressed genes (DEGs) as groups and in individual cases. The DEGs identified were then subjected to The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. We here identified the highly heterogenous nature of the *BRAF*-like cPTCs and fvPTCs with $BRAF^{V600E}$, but also some frequently over-/under-expressed genes, of which *LIPH* expression was found to be favorable prognostic marker for PTCs.

Materials and methods

The RNA sequencing data (fastq) for 15 cPTCs (SNU-GMI-PTC26-33, 35-37, 39, 42, 43 and 45), 5 fvPTCs (SNU-GMI-PTC30-34 and 37) and the respective matched cancer-adjacent normal tissues were downloaded from the EBI European Nucleotide Archive database with accession number PRJEB11591 (http://www.ebi.ac.uk/ena/data/view/PRJEB11591). Following assessment for their quality using the FastQC v0.11.3 program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) [8], the reads were aligned to GRCh38 reference genome using hisat2 version 2.1.0 (https://ccb.jhu.edu/software/hisat2/index.shtml) [9]. The transcripts were quantified using featureCounts (http://bioinf.wehi.edu.au/featureCounts/) [10]. Then, by using the read counts for quantity of gene expression, DEGs were determined using the edgeR version 3.6 program (https://bioconductor.org/packages/release/bioc/html/edgeR.html) [11], which is based on the negative binomial model. The edgeR filters genes with low transcript counts (average mapped reads/sample <1), followed by normalization of all the samples. The DEGs were defined as a false discovery rate (FDR) < 0.05 and |fold-change (FC)| > 2.

GO and KEGG pathway enrichments were performed through the online Database for Annotation, Visualization and Integrated Discovery (DAVID) 6.8 (https://david.ncifcrf.gov/). P values < 0.05 by Benjamini−Hochberg correction were judged as significant. The GO database is currently the most widely used gene annotation system for gene functions and products. It can perform functional enrichment analyses of target genes, and provide a better understanding of the relationship between genes and diseases. Broad classification of GO categories includes three groups—biological process (BP), cellular components (CC) and molecular function (MF). The KEGG database combines genetic information with functional information, and can be used to systematically analyze the relationship between gene functions and enriched pathways [12].

Prognosis was evaluated with The Human Protein Atlas (https://www.proteinatlas.org/), which is connected to The Cancer Genome Atlas (TCGA) dataset (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga).

Results

Samples used and cluster dendrogram, and principle component analyses (PCA)

The dataset we used in this study [7] includes 180 thyroid tumors with 25 thyroid follicular adenomas, 30 follicular thyroid cancers, 51 fvPTCs and 74 cPTCs. In these samples, 52 cPTCs and 10 fvPTCs had $BRAF^{V600E}$ and *BRAF*-like gene expression signature in the BRS, of which 26 cPTCs and 5 fvPTCs had matched cancer-adjacent normal tissues. We used these 26 pairs of cPTCs and 5 pairs of fvPTCs.

The cluster dendrogram and PCA were performed with the R package edgeR. The cancer tissues and their matched normal counterparts were distinctly separated in fvPTCs, but not in cPTCs, based on their global gene expression (Supplementary Fig. 1), suggesting that the global gene expression patterns cannot completely distinguish between cancerous and normal tissues in cPTCs.

Determination of DEGs in cPTCs

DEGs were analyzed in 26 cPTCs and the matched normal tissues (Table 1). Comparison was first made as groups in a nonpaired condition (biological replicates = 26 for cancers and normal tissues) (hereafter "group analysis"). A total of 3381 genes were found to be differentially expressed, of which 1725 genes (51%) were overexpressed and 1656 (49%) were underexpressed. Next, DEGs were determined

Table 1 Differentially expressed genes in 26 cPTCs as compared to matched normal tissues, analyzed as groups or in individual cases

	Group analyses		Individual analyses		
cPTC samples	# of DEGs (FDR < 0.05 and |FC| ≥ 2)	# of over-/underexpressed genes	# of DEGs (FDR < 0.05 and |FC| ≥ 2)	# of over-/underexpressed genes	# of commonly over-/underexpressed genes
cPTC-26			439	268/171	
cPTC-27			862	405/457	
cPTC-28			472	156/316	
cPTC-29			431	200/231	
cPTC-30			301	186/115	
cPTC-31			254	212/42	
cPTC-32			307	147/160	
cPTC-33			291	235/56	
cPTC-35			659	396/263	
cPTC-36			849	694/155	
cPTC-37			1046	385/661	
cPTC-39			818	358/460	
cPTC-42			942	469/473	
cPTC-43	3381	1725/1656	788	499/289	0/0
cPTC-45			730	400/330	
cPTC-50			113	72/41	
cPTC-51			530	119/409	
cPTC-52			192	12/180	
cPTC-53			630	88/142	
cPTC-54			511	194/317	
cPTC-55			1164	773/391	
cPTC-57			316	210/106	
cPTC-58			525	191/334	
cPTC-67			477	284/193	
cPTC-72			294	130/164	
cPTC-74			517	476/41	

individually between each cPTC and the matched normal tissue (hereafter "individual analysis"). Approximately 100 to 1000 genes were found to be differentially expressed, of which ~10 to 800 were overexpressed and ~40 to 660 were underexpressed. However, among these, there were no commonly over-/underexpressed genes in 26 samples. Instead, there were some genes that were frequently over-/underexpressed. For example, 1 gene (*ADCY8*) was overexpressed in 22 out of 26 samples, 14 genes (*SYT12, DNAPTP3, FN1, SLC22A31, TMPRSS6, LIPH, GABRB2, TM7SF4, HMGA2, RXRG, B3GNT3, PRR15, LOC400794, KLK10*) in 21 samples, and 10 genes (*ODZ1, ARHGAP36, KLK7, TMEM92, SYTL5, TMPRSS4, NGEF, PVRL4, CHI3L1, GAP43*) in 20 samples; and 1 gene (*CCL21*) was underexpressed in 17 out of 26 samples, 3 genes (*RELN, MYOC, TFF3*) in 14 samples and 4 genes (*GDF10, CNTFR, TPO, TFF2*) in 12 samples.

GO and KEGG functional analyses in cPTCs

The 1725 overexpressed and 1656 underexpressed genes that were identified by the group analysis (Table 1) were subjected to functional annotation using the GO enrichment analysis for clarifying the biological roles of the DEGs. The overexpressed 1725 DEGs were enriched in 109 functions and pathways (87 BP categories, 11 CC categories and 11 MF categories); the top ten of the most significantly enriched processes in each category are shown in Supplementary Table 1. Of interest, many terms that contained the word "extracellular" such as "extracellular structure organization" and "extracellular matrix organization" in the BP category, "extracellular region", and "extracellular space" in the CC category and "extracellular matrix structural constituent" in the MF category were observed. The KEGG pathway analysis was also performed with the 1725 DEGs

to identify the significant pathways that were altered in cPTC samples (Supplementary Table 1). Again a pathway whose name included the word "extracellular" (i.e., "extracellular matrix-receptor interaction") was listed. The underexpressed 1656 DEGs were enriched in 110 BP, 30 CC and 26 MF categories in the GO analyses, and 7 pathways in the KEGG analyses (only top ten of each GO category are shown in Supplementary Table 2), where we found again two terms containing the word "extracellular" (i.e., "proteinaceous extracellular matrix" and "extracellular space") in the CC category. Due to these apparently discrepant data, that is, both over- and underexpressed DEGs were enriched in the terms containing the word "extracellular", we focused on the terms containing the word "extracellular" in the subsequent study.

As the next step, the over- and underexpressed DEGs identified in the individual analysis (~10 to 800 and ~40 to 660 genes; see Table 1) were subjected to the GO and KEGG analyses, and the data on the terms containing the word "extracellular" are summarized in Fig. 1. Many over- and underexpressed DEGs were enriched in the terms containing the word "extracellular" in many, but not all, samples. Furthermore, the degree of enrichment for the terms containing the word "extracellular" was highly variable in both over- and underexpressed DEGs.

The lists of the over-and underexpressed genes that were annotated with the terms containing the word "extracellular" in 20 and 19, respectively, cPTC samples (excluding cPTC-28, 32, 50~52 and 58 in the overexpressed genes and cPTC-35, 43, 45, 52, 57, 67 and 74 in the underexpressed genes, because of no enrichment in the terms containing the word "extracellular" in these samples; see Fig. 1) were obtained from the GO database (644 and 577 genes, respectively). As shown in Supplementary Tables 3 and 4, there were no genes that were commonly over-/underexpressed in all the samples. *COL10A1, TMPRSS6, LIPH* and *FN1* were overexpressed in 18 out of 20 samples, *COL11A1, KLK7, SFTPB, MMP13* and *ACTBL2* in 17 samples, *TACSTD2, LCN2, CHI3L1* and *GDF15* in 16 samples. CCL21 was underexpressed in 17 out of 19 samples, *RELN, MYOC* and *TFF3* in 14 samples, *DPT* in 13 samples, and *GDF10, CNTFR, TPO* and *TFF2* in 12 samples. On the other hand, the majority (388/644 (60%) and 342/557 (61%) genes) appeared only once or twice in 20 and 19, respectively, samples. In addition, 147 genes were found in both lists of the over- and underexpressed genes (red in Supplementary Tables 3 and 4). Thus, although the over-/underexpressed DEGs were enriched in the terms containing the word "extracellular" in many, but not all, samples, the genes annotated with these terms were quite diverse among each sample.

Finally, prognostic analyses using the Human Protein Atlas and the TCGA dataset revealed that, among the frequently over- or underexpressed genes mentioned above,

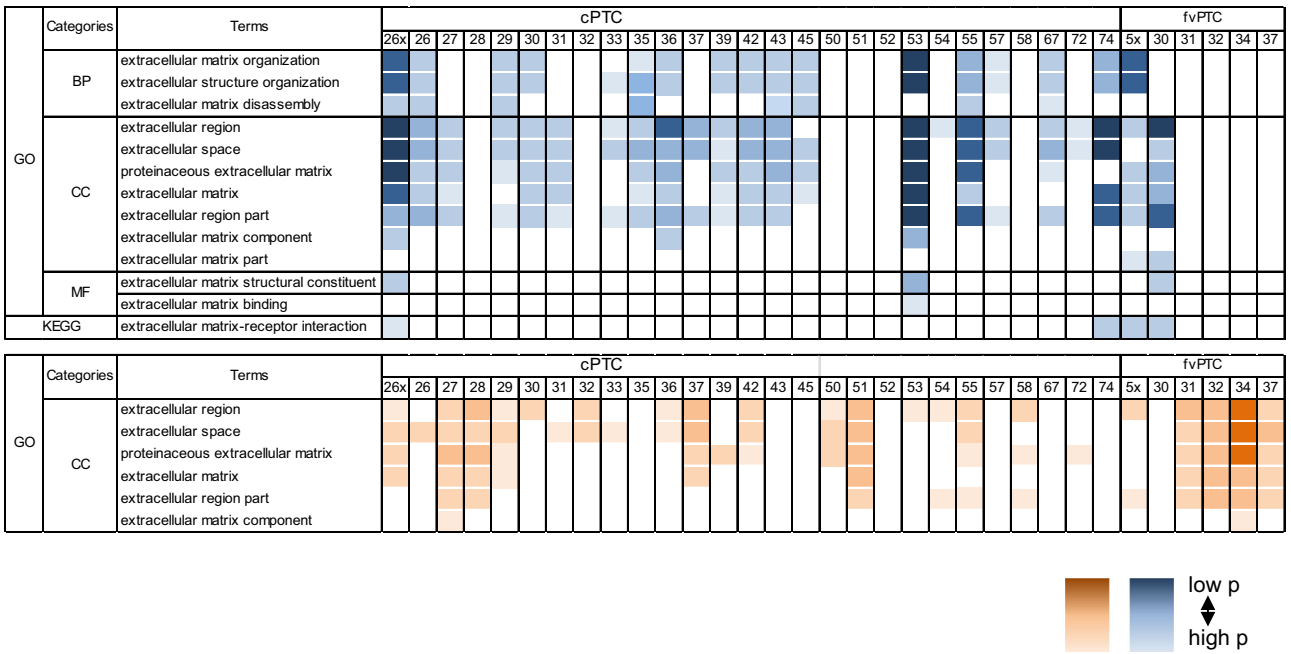

Fig. 1 Heatmap for GO and KEGG analyses of 26 cPTCs and 5 fvPTCs in group analyses (26× and 5×) and individual analyses (cPTC-26~74 and fvPTC-30~37). Only the terms containing the word "extracellular" are shown. The terms enriched with the over- and underexpressed genes are shown in blue and red, respectively. Darker colors indicate lower adjusted *P* values. GO, BP, CC, MF, and KEGG refer to Gene Ontology, biological process, cellular component, molecular function, and Kyoto Encyclopedia of Genes and Genomes, respectively

Table 2 Differentially expressed genes in five fvPTCs as compared to matched normal tissues, analyzed as groups or in individual cases

fvPTC samples	Group analyses		Individual analyses		
	# of DEGs (FDR < 0.05 and |FC| ≥ 2)	# of over-/underexpressed genes	# of DEGs (FDR < 0.05 and |FC| ≥ 2)	# of over-/underexpressed genes	# of commonly over-/underexpressed genes
fvPTC-30			462	333/129	
fvPTC-31			770	401/368	
fvPTC-32	1,711	1060/651	894	273/621	21/17
fvPTC-34			1039	381/658	
fvPTC-37			179	370/350	

LIPH (*lipase H*) expression was found to be prognostic and its high expression was favorable for PTCs (Supplementary Fig. 2; $p = 0.00042$).

Determination of DEGs in fvPTCs

Similar analyses were also performed with five fvPTCs and matched normal tissues. In the group analyses, 1711 genes were differentially expressed, of which 1060 genes were overexpressed and 651 were underexpressed. In the individual analyses, ~180 to 1000 genes were differentially expressed, of which ~270 to 400 genes were overexpressed and ~130 to 660 were underexpressed (Table 2). The number of commonly over- and underexpressed genes in the five samples were 21 (*ZCCHC16, LIPH, GJB3, ZCCHC12, GABRB2, HPCAL4, QPCT, ULBP1, SLIT1, HMGA2, CDH3, KLHDC8A, TRPC5, PLEKHN1, SPTBN2, HCN4, KCNQ3, EPHA10, C6orf138, TMEM215* and *PVRL4*) and 17 (*KCNA1, SLC6A15, FOXA2, LIPI, ODAM, LRP1B, CWH43, TPTE2P1, SLC5A5, CDH16, SEMA3D, UGT2B,ZMAT4, LOC401134, RAG2, MAPK4* and *TFF3*), respectively.

GO and KEGG functional analyses in fvPTCs

In the GO and KEGG analyses, many over- and underexpressed DEGs were again enriched in the terms containing the word "extracellular" (Supplementary Tables 5 and 6, and Fig. 1). Of interest, however, in the individual analyses, significant enrichment for the terms containing the word "extracellular" was detected only in overexpressed genes in fvPTC-30 and in underexpressed genes in the others (Fig. 1).

Comparison of the data on cPTCs and fvPTCs

Four genes (*LIPH, GABRB2*, *HMGA2* and *PVRL4*) were commonly observed between the frequently overexpressed genes in 26 PTCs mentioned above and the commonly overexpressed genes in 5 fvPTCs, but no gene were commonly observed between the frequently underexpressed genes in 26 PTCs and the commonly underexpressed genes in 5 fvPTCs.

Discussion

The present study examined the heterogenous nature of PTCs that were homogenous in terms of histopathology, the driver mutation and the BRS (i.e., classical or follicular variant papillary histology, $BRAF^{V600E}$-mutation and the *BRAF*-like signature, respectively) using publicly available RNA-seq data [7]. Many gene expression data on thyroid tumors, either microarray or RNA-seq, are available through NCBI-GEO, Array Express and The Cancer Genome Atlas. RNA-seq has in general more benefits when compared with microarray platforms, including broader dynamic range and increased specificity and sensitivity. Furthermore, RNA-seq data for both cancer and matched normal tissues were available in the report by Yoo et al. [7].

The inability to completely distinguish cPTCs from the normal tissues by analysis of global gene expression patterns with the cluster dendrogram and PCA is the *first* evidence of heterogeneity for cPTCs.

The DEGs determined by the group analysis identified thousands of genes, while those by the individual analyses tens to hundreds of genes (Table 1). Among the latter, there were no commonly over-/underexpressed genes in 26 PTCs, demonstrating the *second* evidence of heterogeneity in gene expression patterns in this homogenous PTC group. Huang et al. also compared gene expression levels in 8 PTC tumor/normal pairs, and found 24 commonly overexpressed genes in all the tumors [13]. The difference in number of samples examined might explain the different results between their data and our data (in general, more samples, less commonly dysregulated genes).

Functional annotation of over- and underexpressed DEGs identified by group analyses demonstrated that genes annotated with the terms containing the word "extracellular" were highly enriched in cPTCs (Supplementary

Tables 1 and 2). These results were not surprising because similar results have previously been reported by other groups including Yoo et al. (see Supplementary Fig. 5S in ref. [7]) and others [12, 14, 15]. Particularly, as in our study, Liu et al. reported enrichment of both over- and under-expressed genes in the terms containing the word "extracellular" in the group analyses between PTC and normal tissues [16]. However, a novel finding in this study is the highly variable degree of enrichment in the terms containing the word "extracellular" in the individual analyses (Fig. 1), showing the *third* evidence of heterogeneity in gene expression patterns in homogenous cPTC samples. Furthermore, when focused on the genes annotated with the terms containing the word "extracellular", no genes were commonly over-/underexpressed in all samples (Supplementary Tables 3 and 4), providing the *fourth* evidence of heterogeneity. Altogether, our analyses clearly demonstrate the highly heterogenous nature of cPTCs, which are homogenous in terms of histopathology, type of driver mutation and BRS.

Essentially the same results were obtained with five fvPTC samples with the same driver mutation and BRS (Table 2, Supplementary Tables 5 and 6 and Fig. 1), suggesting that, regardless of distinct histology, classical and follicular variant PTCs share similar heterogeneity in gene expression patterns, which is therefore likely attributed to the presence of $BRAF^{V600E}$. As PTCs usually have a single or a few driver mutations [17], microRNA profiles and epigenetic changes may have caused the heterogeneity found in this study.

This study was performed to evaluate the heterogenous nature of apparently homogenous PTC samples. Thus, evaluation of the functional significance of genes over- or underexpressed and annotated with the terms containing the word "extracellular" is not the primary purpose of this study, although it is also true that comprehensive expression analysis is often useful to identify diagnostic, prognostic and/or therapeutic biomarker(s). In this regard, the frequently or commonly overexpressed genes in cPTCs and fvPTCs may be candidates for biological markers. Many of these have previously been reported as potential diagnostic biomarkers for PTCs (*FN1, GABRB2, HMGA2, KLK7, KLK10, TM7SF4, LIPH, ODZ1, RXRG, SYT12, CHI3L1, NGEF, TMPRSS4* and *ADCY8*) [15, 18–25]. In addition, our study identified *LIPH* to be a favorable prognostic marker for cPTCs. *LIPH* encodes lipase H, hydrolyzing specifically phosphatidic acid to produce 2-acyl lysophosphatidic acid and fatty acid. Similar data were reported in esophageal cancers [26]. However, a recent study demonstrated promotion by LIPH of migration of thyroid cancer cell lines [27]. Furthermore, other studies identified *FN1* [28], *GABRB2* [29] and *CHI3L1* [23] as biomarkers for aggressive behavior of PTCs. Functional studies will be required to clarify the significance of these gene products in thyroid cancers.

The limitation in this study includes the small number of samples examined, particularly the fvPTC cases, with limited clinical data. It should be noted here that we could not find any correlation between the clinical data provided in ref. [7] and gene expression patterns found in this study, because the latter was so diverse. Further study with larger cPTC and fvPTC sample sizes and detailed clinical data will be of interest in the future.

In conclusion, our analyses of RNA-seq data for DEGs and functional annotation revealed the heterogeneity of gene expression patterns in a homogenous group of classical PTC samples with $BRAF^{V600E}$ and *BRAF*-like expression signature. These heterogeneities might at least in part explain the variable prognosis of cPTCs. We also found that expression levels of *LIPH* would be a favorable prognostic biomarker for PTCs.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

References

1. T. Carling, R. Udelsman, Thyroid cancer. Annu Rev. Med **65**, 125–137 (2014). https://doi.org/10.1146/annurev-med-061512-105739
2. D. Tumino, F. Frasca, K. Newbold, Updates on the management of advanced, metastatic, and radioiodine refractory differentiated thyroid cancer. Front. Endocrinol. **8**, 312 (2017). https://doi.org/10.3389/fendo.2017.00312
3. B.R. Haugen, E.K. Alexander, K.C. Bible, G.M. Doherty, S.J. Mandel, Y.E. Nikiforov, F. Pacini, G.W. Randolph, A.M. Sawka, M. Schlumberger, K.G. Schuff, S.I. Sherman, J.A. Sosa, D.L. Steward, R.M. Tuttle, L. Wartofsky, 2015 American Thyroid Association Management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid **26**(1), 1–133 (2016). https://doi.org/10.1089/thy.2015.0020
4. The Cancer Genome Atlas Research, N., Integrated genomic characterization of papillary thyroid carcinoma. Cell **159**(3), 676–690 (2014). https://doi.org/10.1016/j.cell.2014.09.050
5. G. Riesco-Eizaguirre, P. Santisteban, Endocrine tumours: advances in the molecular pathogenesis of thyroid cancer: lessons from the cancer genome. Eur. J. Endocrinol. **175**(5), R203–R217 (2016). https://doi.org/10.1530/eje-16-0202
6. K. Kim, S. Jeon, T. M. Kim, C. K. Jung, Immune gene signature delineates a subclass of papillary thyroid cancer with unfavorable clinical outcomes. Cancers (Basel) **10**(12) (2018). https://doi.org/10.3390/cancers10120494
7. S.K. Yoo, S. Lee, S.J. Kim, H.G. Jee, B.A. Kim, H. Cho, Y.S. Song, S.W. Cho, J.K. Won, J.Y. Shin, J. Park do, J.I. Kim, K.E. Lee, Y.J. Park, J.S. Seo, Comprehensive analysis of the transcriptional and mutational landscape of follicular and papillary

thyroid cancers. PLoS Genet **12**(8), e1006239 (2016). https://doi.org/10.1371/journal.pgen.1006239
8. S.W. Wingett, S. Andrews, FastQ Screen: a tool for multi-genome mapping and quality control. F1000Res **7**, 1338 (2018). https://doi.org/10.12688/f1000research.15931.2
9. D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. Nat. Methods **12**(4), 357–360 (2015). https://doi.org/10.1038/nmeth.3317
10. Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics **30**(7), 923–930 (2014). https://doi.org/10.1093/bioinformatics/btt656
11. M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**(1), 139–140 (2010). https://doi.org/10.1093/bioinformatics/btp616
12. J. Han, M. Chen, Y. Wang, B. Gong, T. Zhuang, L. Liang, H. Qiao, Identification of biomarkers based on differentially expressed genes in papillary thyroid carcinoma. Sci. Rep. **8**(1), 9912 (2018). https://doi.org/10.1038/s41598-018-28299-9
13. Y. Huang, M. Prasad, W.J. Lemon, H. Hampel, F.A. Wright, K. Kornacker, V. LiVolsi, W. Frankel, R.T. Kloos, C. Eng, N.S. Pellegata, A. de la Chapelle, Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. Proc. Natl. Acad. Sci. USA **98**(26), 15044–15049 (2001). https://doi.org/10.1073/pnas.251547398
14. Y. Xu, J. Chen, Z. Yang, L. Xu, Identification of RNA expression profiles in thyroid cancer to construct a competing endogenous RNA (ceRNA) network of mRNAs, long Noncoding RNAs (lncRNAs), and microRNAs (miRNAs). Med. Sci. Monit.: Int. Med. J. Exp. Clin. Res. **25**, 1140–1154 (2019). https://doi.org/10.12659/msm.912450
15. I. Reyes, N. Reyes, R. Suriano, C. Iacob, N. Suslina, A. Policastro, A. Moscatello, S. Schantz, R.K. Tiwari, J. Geliebter, Gene expression profiling identifies potential molecular markers of papillary thyroid carcinoma. Cancer Biomark. **24**(1), 71–83 (2019). https://doi.org/10.3233/cbm-181758
16. L. Liu, C. He, Q. Zhou, G. Wang, Z. Lv, J. Liu, Identification of key genes and pathways of thyroid cancer by integrated bioinformatics analysis. J. Cell. Physiol. (2019). https://doi.org/10.1002/jcp.28932
17. I. Landa, T. Ibrahimpasic, L. Boucai, R. Sinha, J.A. Knauf, R.H. Shah, S. Dogan, J.C. Ricarte-Filho, G.P. Krishnamoorthy, B. Xu, N. Schultz, M.F. Berger, C. Sander, B.S. Taylor, R. Ghossein, I. Ganly, J.A. Fagin, Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J. Clin. Investig. **126**(3), 1052–1066 (2016). https://doi.org/10.1172/jci85271
18. O.L. Griffith, A. Melck, S.J. Jones, S.M. Wiseman, Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. J. Clin. Oncol. **24** (31), 5043–5051 (2006). https://doi.org/10.1200/jco.2006.06.7330
19. M.C. Barros-Filho, F.A. Marchi, C.A. Pinto, S.R. Rogatto, L.P. Kowalski, High diagnostic accuracy based on CLDN10, HMGA2, and LAMB3 transcripts in papillary thyroid carcinoma. J. Clin. Endocrinol. Metab. **100**(6), E890–E899 (2015). https://doi.org/10.1210/jc.2014-4053
20. H.S. Kim, D.H. Kim, J.Y. Kim, N.H. Jeoung, I.K. Lee, J.G. Bong, E.D. Jung, Microarray analysis of papillary thyroid cancers in Korean. Korean J. Intern. Med. **25**(4), 399–407 (2010). https://doi.org/10.3904/kjim.2010.25.4.399
21. Q.X. Wang, E.D. Chen, Y.F. Cai, Q. Li, Y.X. Jin, W.X. Jin, Y.H. Wang, Z.C. Zheng, L. Xue, O.C. Wang, X.H. Zhang, A panel of four genes accurately differentiates benign from malignant thyroid nodules. J. Exp. Clin. Cancer Res. **35**(1), 169 (2016). https://doi.org/10.1186/s13046-016-0447-3
22. S.P. Cheng, M.J. Chen, M.N. Chien, C.H. Lin, J.J. Lee, C.L. Liu, Overexpression of teneurin transmembrane protein 1 is a potential marker of disease progression in papillary thyroid carcinoma. Clin. Exp. Med. **17**(4), 555–564 (2017). https://doi.org/10.1007/s10238-016-0445-y
23. D. Luo, H. Chen, P. Lu, X. Li, M. Long, X. Peng, M. Huang, K. Huang, S. Lin, L. Tan, Y. Zhu, Z. Chen, N. Ouyang, H. Li, CHI3L1 overexpression is associated with metastasis and is an indicator of poor prognosis in papillary thyroid carcinoma. Cancer Biomark. **18**(3), 273–284 (2017). https://doi.org/10.3233/cbm-160255
24. Y. Zhang, Z. Zhang, J. Ma, J. Pu, P. Hou, Q. Yang, High-accuracy detection of preoperative thyroid nodules using combination of BRAF(V600E) mutation and TMPRSS4 mRNA level. Arch. Med. Res. **49**(6), 365–372 (2018). https://doi.org/10.1016/j.arcmed.2018.11.003
25. J. Tang, D. Kong, Q. Cui, K. Wang, D. Zhang, Q. Yuan, X. Liao, Y. Gong, G. Wu, Bioinformatic analysis and identification of potential prognostic microRNAs and mRNAs in thyroid cancer. PeerJ **6**, e4674 (2018). https://doi.org/10.7717/peerj.4674
26. H. Ishimine, R. Zhou, K. Sumitomo, Y. Ito, Y. Seki, Y. Yoshida, A. Kurisaki, Lipase member H frequently overexpressed in human esophageal adenocarcinomas. Tumour Biol. **37**(2), 2075–2081 (2016). https://doi.org/10.1007/s13277-015-3985-y
27. Y. Li, X. Zhou, Q. Zhang, E. Chen, Y. Sun, D. Ye, O. Wang, X. Zhang, J. Lyu, Lipase member H is a downstream molecular target of hypoxia inducible factor-1alpha and promotes papillary thyroid carcinoma cell migration in BCPAP and KTC-1 cell lines. Cancer Manag. Res. **11**, 931–941 (2019). https://doi.org/10.2147/cmar.S183355
28. M. Sponziello, F. Rosignolo, M. Celano, V. Maggisano, V. Pecce, R.F. De Rose, G.E. Lombardo, C. Durante, S. Filetti, G. Damante, D. Russo, S. Bulotta, Fibronectin-1 expression is increased in aggressive thyroid cancer and favors the migration and invasion of cancer cells. Mol. Cell. Endocrinol. **431**, 123–132 (2016). https://doi.org/10.1016/j.mce.2016.05.007
29. Y. Jin, W. Jin, Z. Zheng, E. Chen, Q. Wang, Y. Wang, O. Wang, X. Zhang, GABRB2 plays an important role in the lymph node metastasis of papillary thyroid cancer. Biochem. Biophys. Res. Commun. **492**(3), 323–330 (2017). https://doi.org/10.1016/j.bbrc.2017.08.114