



# Comparative Analysis of Driver Mutations and Transcriptomes in Papillary Thyroid Cancer by Region of Residence in South Korea

Jandee Lee<sup>1</sup>, Seonhyang Jeong<sup>2</sup>, Hwa Young Lee<sup>1</sup>, Sunmi Park<sup>2</sup>, Meesson Jeong<sup>3</sup>, Young Suk Jo<sup>2</sup>

<sup>1</sup>Department of Surgery, Open NBI Convergence Technology Research Laboratory, Severance Hospital, Yonsei Cancer Center, Yonsei University College of Medicine; <sup>2</sup>Division of Endocrinology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul; <sup>3</sup>Radiation Effect Research Section, Radiation Health Institute, Korea Hydro & Nuclear Power Co., Ltd., Gyeongju, Korea

**Background:** Radiation exposure is a well-known risk factor for papillary thyroid cancer (PTC). South Korea has 24 nuclear reactors in operation; however, no molecular biological analysis has been performed on patients with PTC living near nuclear power plants.

**Methods:** We retrospectively included patients with PTC ( $n=512$ ) divided into three groups according to their place of residence at the time of operation: inland areas ( $n=300$ ), coastal areas far from nuclear power plants ( $n=134$ ), and nuclear power plant areas ( $n=78$ ). After propensity score matching (1:1:1) by age, sex, and surgical procedure, the frequency of representative driver mutations and gene expression profiles were compared ( $n=50$  per group). Epithelial-mesenchymal transition (EMT), BRAF, thyroid differentiation, and radiation scores were calculated and compared.

**Results:** No significant difference was observed in clinicopathological characteristics, including radiation exposure history and the frequency of incidentally discovered thyroid cancer, among the three groups. *BRAF*<sup>V600E</sup> mutation was most frequently detected in the groups, with no difference among the three groups. Furthermore, gene expression profiles showed no statistically significant difference. EMT and BRAF scores were higher in our cohort than in cohorts from Chernobyl tissue bank and The Cancer Genome Atlas Thyroid Cancer; however, there was no difference according to the place of residence. Radiation scores were highest in the Chernobyl tissue bank but exhibited no difference according to the place of residence.

**Conclusion:** Differences in clinicopathological characteristics, frequency of representative driver mutations, and gene expression profiles were not observed according to patients' region of residence in South Korea.

**Keywords:** Thyroid cancer, papillary; Radiation genomics; Transcriptome; Geographic locations; Nuclear power plants

## INTRODUCTION

The incidence of thyroid cancer has increased in the last few decades, especially in South Korea [1-3]. This may be because of

the introduction of high-sensitivity neck ultrasound for detecting papillary thyroid cancer (PTC), mainly papillary thyroid microcarcinoma [4-6]. However, previous studies have reported that various risk factors for PTC, such as radiation exposure,

Received: 12 June 2023, Revised: 16 August 2023, Accepted: 25 September 2023

**Corresponding author:** Young Suk Jo  
Division of Endocrinology, Department of Internal Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea  
Tel: +82-2-2228-0752, Fax: +82-2-393-6884, E-mail: joys@yuhs.ac

Copyright © 2023 Korean Endocrine Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

high iodine intake, obesity, and family history of PTC, are responsible for the actual increase in the prevalence of thyroid cancer [7-11]. Among these risk factors, the best known is radiation exposure [12,13].

Furthermore, many studies targeting external exposure, such as radiation therapy, reported an increased incidence of PTC among survivors of the atomic bombings in Hiroshima and Nagasaki and residents near the Chernobyl nuclear power plant accident [12,14]. As the thyroid equivalent dose of children and adolescents increased because of the Chernobyl nuclear power plant accident in 1986, the equivalent thyroid dose from radioactive iodine at the time of the accident was estimated to be 100 to 2,000 mSv [15,16]. According to a report published by the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) in 2008, the prevalence of thyroid cancer is proportional to dose; however, no statistically significant increase is observed at doses <100 mSv [17]. In a study of thyroid cancer in adolescents and adults in Bryansk, Russia's most polluted region after the Chernobyl accident, the standardized incidence ratio of PTC was increased in both men and women; however, a relationship between radiation dose and PTC risk was not observed [18]. In addition, even if an increase in the incidence of thyroid cancer was observed in epidemiological studies, genetic susceptibility, iodine deficiency effect, and screening effect of thyroid cancer induced by the Chernobyl accident act as confounders, suggesting that careful interpretation is required [7].

In South Korea, there are 24 nuclear reactors in operation, located in Gori (Jangan-eup, Gijang-gun, and Busan), Shin-Gori (Seosaeng-myeon, Ulju-gun, and Ulsan), Wolsong/Sinwolseong (Yangnam-myeon, Gyeongju-si, and Gyeongsangbuk-do), Hanbit (Hongnong-eup, Yeonggwang-gun, and Jeollanam-do), and Hanul (Buk-myeon, Uljin-gun, and Gyeongsangbuk-do). This distribution indicates that four of the top 10 nuclear power plants density in the world are located in South Korea. Furthermore, 5.31 million people live within a 30 km radius of each of these nuclear power plants; considering that the population of Korea was 51.7 million as of 2019, more than one out of 10 people live within a 30 km radius of a nuclear power plant. Therefore, whether the occurrence of PTC is increased by radiation exposure in this region has been a very important social issue [19].

The driver mutations of radiation exposure-induced PTC are mainly RET/PTC rearrangements, as supported by Genomic Data from Chernobyl Studies [20,21]. However, RET/PTC rearrangements have been detected even in young patients with thy-

roid cancer who have not been exposed to radiation, and the cause remains unidentified [22]. *BRAF*<sup>V600E</sup> mutation has been frequently identified as the driver mutation of sporadic PTC in South Korea [23]. Furthermore, some studies reported that high iodine intake or obesity may be associated with *BRAF*<sup>V600E</sup> mutation; however, substantial evidence supporting this finding is lacking [24,25]. In addition, a study of atomic bomb survivors reported that the *BRAF*<sup>V600E</sup> mutation is highly prevalent at low doses, as in sporadic PTC [26,27].

In the present study, we compared the frequency of driver mutations according to the patients' region of residence to investigate the effect of nuclear power plants on the occurrence of PTC in South Korea. In addition, to overcome the limitation of driver mutation analysis, regions were compared in terms of transcriptome analysis. Considering that most nuclear power plant areas are distributed along the coast, a comparison was made between three groups: inland areas, coastal areas far from nuclear power plants, and nuclear power plant areas.

## METHODS

### Patients and tissue samples

Patients who underwent thyroidectomy at Yonsei Cancer Center (Seoul, Korea) between March 2003 and December 2019 were included in this study, and classical type PTC tissue samples were collected from the same patients. Patients with confirmed Hashimoto's thyroiditis pathologically were excluded. Patients with a first-degree relative (parent, brother, sister, or child) with thyroid cancer were defined as having a family history. Patients ( $n=512$ ) were divided into three groups according to their place of residence at the time of operation: inland areas ( $n=300$ ), coastal areas far from nuclear power plants ( $n=134$ ), and nuclear power plant areas ( $n=78$ ). All tissues were frozen in liquid nitrogen immediately after surgery and stored at  $-80^{\circ}\text{C}$  before use. To prepare tissue samples, three groups were identified by extracting the group of patients with thyroid cancer and matching age, sex, and surgical procedure in a ratio of 1:1:1 ( $n=50$  per group). The requirement for written informed consent was waived because of the study's retrospective nature. For tissue samples, consent was obtained from the institutional tissue banks at the time of surgery. The study protocol was approved by the Institutional Review Board of Yonsei Cancer Center, Severance Hospital (IRB-No. 4-2020-0815), Seoul, Korea.

### Identification of driver mutations

Genomic DNA from PTC tissue specimens was extracted using

the QIAamp DNA Mini kit (Qiagen Inc., Valencia, CA, USA) and used to amplify target genes with polymerase chain reaction (PCR) using specific primers, followed by Sanger sequencing (Applied Biosystems, Foster, CA, USA). The following primers were used: 5'-ATG CTT GCT CTG ATA GGA AA-3' (forward) and 5'-ATT TTT GTG AAT ACT GGG AA-3' (reverse) for BRAF; 5'-AGC ATT GCA TTC CCT GTG GT-3' (forward) and 5'-AGT GTG GTA ACC TCA TTT CCC C-3' (reverse) for NRAS<sup>O61R</sup>; 5'-TCT CAG CAC CCC AGG AGA G-3' (forward) and 5'-GAC ATG CGC AGA GAG GAC AG-3' (reverse) for HRAS<sup>O61R</sup>; 5'-CAC GTC TGC AGT CAA CTG GA-3' (forward) and 5'-ACC CTG ACA TAC TCC CAA GGA-3' (reverse) for KRAS<sup>G12V</sup>; 5'-CCG TCA TCT TTG GAG CAG GA-3' (sense) and 5'-TCC ACT GCT CTA ATC CCC CA-3' (reverse) for KRAS<sup>O61K</sup>; 5'-ATC GGG TTC CTT GCA TAG GG-3' (forward) and 5'-ACC CTG ACC ATC CTC TTT GA-3' (reverse) for eukaryotic translation initiation factor 1A X-linked (EIF1AX); 5'-GTG TTG CTT GAT GGC CCT TG-3' (sense) and 5'-CAG TGT TGA TCA GGC CCC AT-3' (reverse) for CAP-Gly domain containing linker protein 2 (CLIP2).

### RNA sequencing

Total RNA was extracted according to the manufacturer's protocol (TRIzol reagent, Qiagen). RNA quality was analyzed using a 2100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA, USA). MacroGen Inc. (Seoul, Korea) prepared a cDNA library using TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, USA) according to the sample preparation guide (Part #15031048 Rev E). NovaSeq 6000 (Illumina) was used for 100 nt paired-end sequencing. After trimming low-quality RNA or artifacts, such as adaptor sequences, contaminant DNA, and PCR duplicates from the raw data, data were mapped to the reference genome using HISAT2 (version 2.1.0), and aligned reads were generated using StringTie (version 2.1.3b). The transcripts were normalized to fragments per kilobase of transcript per million mapped reads and transcripts per million kilobases (TPM). Fold changes were calculated using the TPM log<sub>2</sub> value to analyze differentially expressed genes (DEGs). To minimize batch effects, we used same reagent lot and instrument used to conduct the experiment, and one researcher (S.P., PhD) performed the experiments, such as DNA and RNA extraction. RNA sequencing was randomly assigned and conducted simultaneously at the same institution (MacroGen Inc., Seoul, Korea). To validate the results from RNA sequencing, real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed for genes of interest

according to standard protocols (refer to Supplemental Fig. S1 for representative results) [28].

### Public data acquisition

Normalized raw data for RNA expression from The Cancer Genome Atlas (TCGA, <https://tcga-data.nci.nih.gov/tcga/>) and GSE33630 from the Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) were selected.

### Gene set enrichment analysis and generation of the scoring system

Genes identified as DEGs (fold changes >2 and two-tailed  $P < 0.05$ ) were subjected to pathway enrichment analysis using gene set enrichment analysis (GSEA) [29,30]. Hallmark gene sets were used for GSEA. Statistical significance was set at  $P < 0.05$  and false discovery rate at  $< 0.25$ . The results were presented using the dot plot to represent feature expression changes across different identity classes (groups). A single-sample GSEA (ssGSEA) was performed to calculate signature gene scores using signature gene sets [31]. The calculated signature scores were hallmark epithelial-mesenchymal transition (EMT), BRAF, thyroid differentiation (TD), and radiation scores [32]. The gene lists used for radiation scores are summarized in Supplemental Table S1.

### Fusion gene analysis

To identify fusion oncogenes, deFuse v0.8.1, FusionCatcher v1.00, and Arriba v1.2.0 were applied to the expression data. The fusion genes found simultaneously in all three algorithms were selected as valid results.

### Statistical analyses

Mean comparisons were performed using analysis of variance or Mann-Whitney  $U$  test, whereas group comparisons were performed using a two-tailed chi-square test, Fisher's exact test, or linear-by-linear associations accordingly. Statistical analysis was performed using Prism version 9 (GraphPad Software, San Diego, CA, USA) and SPSS version 25.0 for Windows (IBM Corporation, Armonk, NY, USA). All  $P$  values are two-sided.

## RESULTS

### Participants' baseline characteristics

First, we retrospectively included study participants according to their place of residence at the time of operation. Residents of the nuclear power plant area were selected from those who lived

**Table 1.** Participants' Clinicopathological Characteristics according to Residential Area (Inland, Coastal, and Near NPP) after Propensity Score Matching

Characteristic	NPP (n=50)	Inland (n=50)	Coastal (n=50)
Female sex	34 (68)	35 (70)	34 (68)
Age, yr	41.7±12.1	41.8±12.3	41.6±11.4
Extent of surgery (total thyroidectomy)	33 (66)	33 (66)	34 (68)
Extent of LN dissection (MRND)	16 (32)	17 (34)	17 (34)
Tumor size, cm	1.69±0.91	1.71±0.87	1.78±1.2
Extrathyroidal extension, minimal	40 (80)	35 (70)	37 (74)
Multiplicity, single/multiple/bilateral	32 (64)/5 (10)/13 (26)	34 (68)/3 (6)/13 (26)	27 (54)/5 (10)/18 (36)
T stage, T1/T2/T3/T4	13 (26)/6 (12)/22 (44)/9 (18)	20 (40)/4 (8)/23 (46)/3 (6)	15 (30)/2 (4)/28 (56)/5 (10)
N stage, N0/N1a/N1b	17 (34)/17 (34)/16 (32)	17 (34)/16 (32)/17 (34)	14 (28)/19 (38)/17 (34)
M stage (M1)	0	0	0
TNM stage, 1/2/3/4	37 (74)/0/9 (18)/4 (8)	39 (78)/1 (2)/4 (8)/6 (12)	36 (72)/2 (4)/7 (14)/5 (10)
Radiation exposure history	0	0	1 (2) <sup>a</sup>
Family history	9 (18)	5 (10)	8 (16)
Incidentaloma	45 (90)	41 (82)	38 (76)

Values are expressed as number (%) or mean±standard deviation. *P* values were calculated by analysis of variance (ANOVA), Mann-Whitney *U* test, or one-way ANOVA, whereas group comparisons were performed using a two-tailed chi-square test, Fisher's exact test, or linear-by-linear associations. However, *P* values are not reported because they were not significant.

NPP, nuclear power plant; LN, lymph node; MRND, modified radical neck dissection; TNM, tumor, node, and metastasis.

<sup>a</sup>The patient had been treated with radioactive iodine for hyperthyroidism. The driver mutation discovered in this patient was *BRAF*<sup>V600E</sup>.

within 30 km of the Kori and Wolsong nuclear power plants (*n*=78). Residents of the inland area were selected from those who lived in Eunpyeong-gu, Seodaemun-gu, Jongno-gu, and Mapo-gu, Seoul, Korea (*n*=300). Residents of the coastal area far from the nuclear power plant were from Incheon city (*n*=134). Participants' baseline characteristics according to their place of residence are presented in Supplemental Table S2.

To reduce the effect of confounding variables, we performed 1:1:1 propensity score matching using logistic regression for age, sex, and extent of surgical procedure (*n*=50 per group). As shown in Table 1, no significant difference was observed in clinicopathological characteristics, including radiation exposure history, family history of thyroid cancer, and the frequency of incidentally discovered thyroid cancer (incidentaloma) among the three groups.

### Comparison of driver mutation according to the place of residence

Representative driver genes, including point mutations and fusion genes, were analyzed to evaluate the difference in driver mutations. As shown in Table 2, *BRAF*<sup>V600E</sup> mutation was most frequently detected in all the groups, with no difference among

**Table 2.** Comparative Analysis of the Frequency of Driver Mutations and Fusion Genes according to Residential Area (Inland, Coastal, and Near NPP)

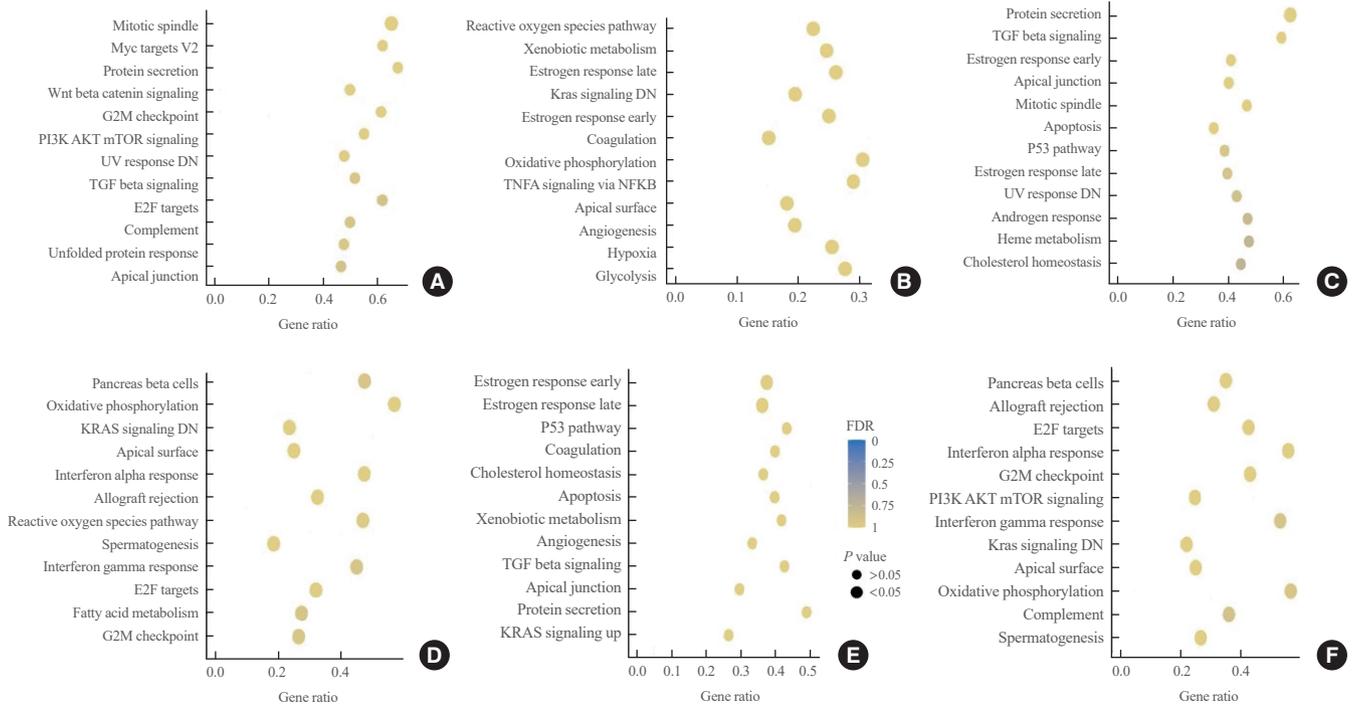
Variable	NPP (n=50)	Inland (n=50)	Coastal (n=50)	Total (n=150)
<i>BRAF</i> <sup>V600E</sup>	39/50 (78)	40/50 (80)	39/50 (78)	118/150 (78.6)
RET/PTC1	1 (2)	2 (4)	2 (4)	5/150 (3.3)
RET/PTC3	1 (2)	1 (2)	1 (2)	3/150 (2)
<i>NRAS</i> <sup>Q61R</sup> / <i>NRAS</i> <sup>Q61K</sup>	1 (2)/ <i>NRAS</i> <sup>Q61R</sup>	0	0	1/150 (0.67)
Other fusion	0	VIM/ NTRK3	0 <sup>a</sup>	1/150 (0.67)
All fusion	2 (4)	4 (8)	3 (6)	9/150 (6)
Unknown	8 (16)	6 (12)	8 (16)	23 (15.3)

Values are expressed as number (%).

NPP, nuclear power plant; PTC, papillary thyroid cancer; VIM, vimentin; NTRK3, neurotrophic receptor tyrosine kinase 3.

<sup>a</sup>Actinin alpha 1 (ACTN1)/*BRAF* was discovered, but is not an experimentally proven oncogene.

the three groups. According to an analysis of the other point mutations, one *NRAS*<sup>Q61R</sup> was found in the nuclear power plant group; however, it was statistically non-significant. According



**Fig. 1.** Dot plot for the gene set enrichment analysis (GSEA) results according to patients' residential area ( $n=50$  per group). GSEA presents gene sets enriched in the patient cohort in inland areas compared with that in coastal areas (A), patient cohort in coastal areas compared with that in inland areas (B), patient cohort in inland areas compared with that in areas near nuclear power plants (NPPs) (C), patient cohort in areas near NPPs compared with that in inland areas (D), patient cohort in coastal areas compared with that in areas near NPPs (E), and patient cohort in areas near NPPs compared with that in coastal areas (F). G2M, gap 2 mitosis; PI3K, phosphoinositide 3-kinase; mTOR, mammalian target of rapamycin; UV, ultraviolet; DN, down; TGF, transforming growth factor; E2F, E2F transcription factor; TNFA, tumor necrosis factor alpha; NFkB, nuclear factor kappa-light-chain-enhancer of activated B cells; FDR, false discovery rate.

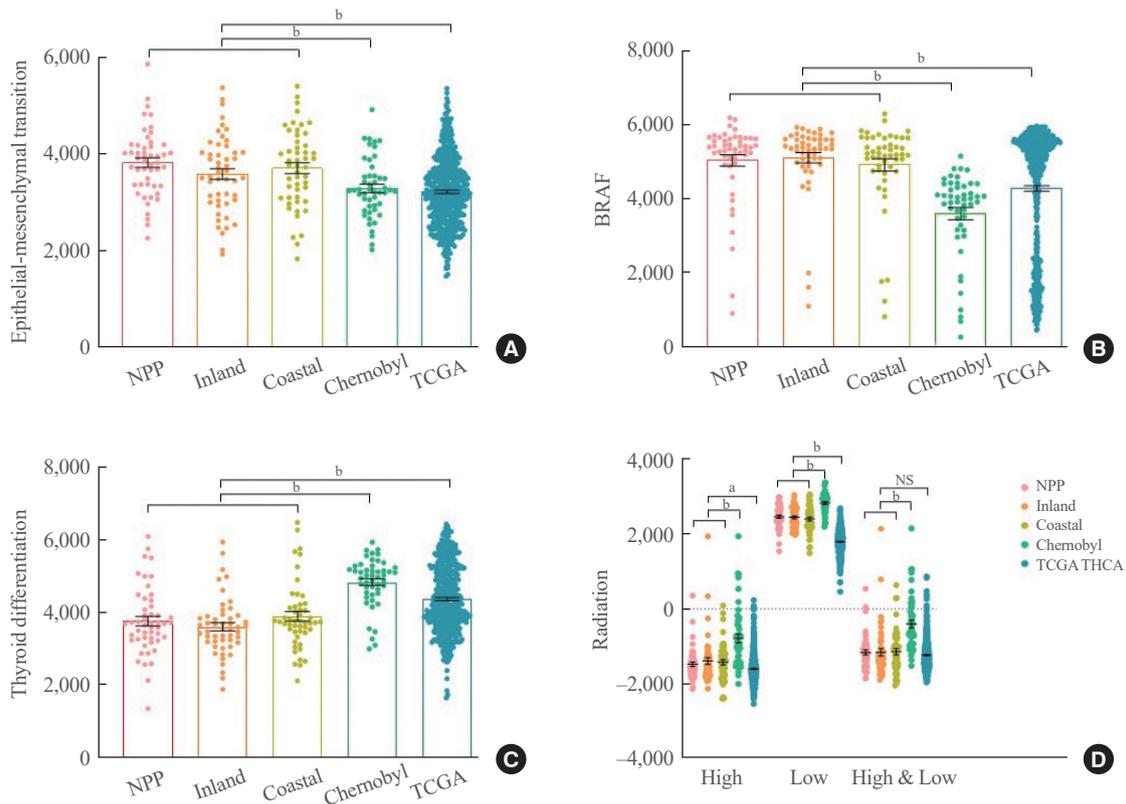
to the fusion gene analysis, eight cases harboring RET/PTC1 or RET/PTC3 were observed, with no difference among the three groups.

### Comparison of RNA expression according to the place of residence

To understand the biological characteristics of PTC according to the place of residence, we performed GSEA using bulk RNA sequencing. First, we compared gene expression between the inland area and the coastal area far from nuclear power plant groups. The inland area group was enriched in genes associated with the mitotic spindle and myc targets, whereas the coastal area group was enriched in genes associated with reactive oxygen species and xenobiotic metabolism. However, no statistically significant difference was observed (Fig. 1A, B). Furthermore, no significant difference was observed between the inland area and nuclear power plant groups and between the coastal area and nuclear power plant groups (Fig. 1C-F).

### Comparison of signature gene scores according to the place of residence

Finally, we generated and compared signature gene scores using genes associated with EMT, BRAF, TD, and radiation-induced damage. To achieve diverse data analysis, public repository data were used as control groups, including Chernobyl tissue bank and The Cancer Genome Atlas Thyroid Cancer (TCGA THCA). EMT scores were higher in the nuclear power plant, inland area, and coastal area groups than in the Chernobyl tissue bank and TCGA THCA groups. However, no difference was observed according to the place of residence (Fig. 2A). The results of the comparisons of BRAF scores were similar to those of the comparisons of EMT scores (Fig. 2B). TD scores were highest in the Chernobyl tissue bank group, followed by the TCGA THCA group (Fig. 2C). Radiation scores were also highest in the Chernobyl tissue bank group (Fig. 2D). Notably, no difference was observed in these signature gene scores according to the place of residence.



**Fig. 2.** Comparisons of signature gene scores among our cohort, the Chernobyl tissue bank, and The Cancer Genome Atlas Thyroid Cancer (TCGA THCA). Scores representing the expression status of genes associated with epithelial-mesenchymal transition (A), BRAF (B), thyroid differentiation (C), and radiation exposure (D). Group comparisons were performed using Mann-Whitney *U* test and Student *t* test accordingly. Error bars are presented with standard errors of the mean. NPP, nuclear power plant. <sup>a</sup>*P*<0.05; <sup>b</sup>*P*<0.01.

## DISCUSSION

In this study, we aimed to investigate the genomic and transcriptomic differences in patients with PTC according to the region of residence in South Korea. Therefore, we first performed direct sequencing to find representative driver gene mutations, such as *BRAF*<sup>V600E</sup>, *NRAS*<sup>Q61R</sup>, *HRAS*<sup>Q61R</sup>, *KRAS*<sup>G12V</sup>, *KRAS*<sup>Q61K</sup>, *EIF1AX*<sup>G9D</sup>, *EIF1AX*<sup>G9R</sup>, *EIF1AX*<sup>G8R</sup>, and *CLIP2*<sup>S17C</sup> [33,34]. As previously reported, the *BRAF*<sup>V600E</sup> mutation was most frequently detected, whereas the other point mutations were not observed except for *NRAS*<sup>Q61R</sup> harbored in one PTC case. *BRAF*<sup>V600E</sup> mutation has been reported to result from low dose ionizing radiation and high iodide intake; however, the present study's data did differ according to the region of residence. Next, we performed fusion gene analysis using bulk RNA sequencing data and found eight *RET*/*PTC1* or three positive PTC cases. However, there was no difference according to geographic locations. *RET*/*PTC1* and *RET*/*PTC3* were representative genetic alterations induced by radiation exposure [20]. However, be-

cause these rearrangements have been reported sporadically independent from radiation [22], we postulated that these eight cases are associated with sporadic occurrence. Furthermore, we found two fusion genes: vimentin (*VIM*)/neurotrophic receptor tyrosine kinase 3 (*NTRK3*) and actinin alpha 1 (*ACTN1*)/*BRAF*. *VIM*/*NTRK3* has been reported previously; however, *ACTN1*/*BRAF* is the first discovery [35]. The frequency of these rearrangements was too small, and the relationship with radiation exposure needs further investigation in future studies.

Before performing transcriptome analysis, we performed propensity score matching (1:1:1) to reduce the effect of some confounding variables, including age, sex, and surgical procedures. We postulated that concerns for radiation-induced cancer among patients in the nuclear power plant area could affect the early diagnosis rate or the incidence rate in men. The surgical method was matched as the possibility of finding indolent metastasis increases as the surgical range widens [36]. Thus, after correcting for differences in demographic features and intervention methods, we investigated the genomic and transcriptomic differences

between groups. Notably, there were no differences in important clinicopathological factors, such as the prevalence of incidentally discovered PTC, radiation exposure, family history, tumor size, and tumor, node, and metastasis (TNM) stage before or after matching, among the groups [37]. Because the age of patients in inland areas was higher than that of those in the other two groups, we needed to perform propensity score matching. In addition, because bulk RNA sequencing is a relatively expensive research technique [38], it was challenging to completely enumerate all collected PTC cases. After matching, we performed bulk RNA sequencing using 50 cases per group and GSEA to determine whether priori-defined gene sets showed statistically significant differences according to the region of residence [39,40]. GSEA could compare only two groups; therefore, we compared two groups each. The results suggested that the expression of several gene sets differed between the groups; however, the difference was not statistically significant. Therefore, we concluded that the clinicopathological characteristics, driver mutation, and gene expression profiles were not different according to the region of residence.

To verify our finding, we further calculated signature gene scores using our bulk RNA sequencing data and public repository, including the Chernobyl tissue bank, a representative radiation-induced PTC model [41], and TCGA THCA, a representative sporadically occurred PTC model [33]. Our transcriptome data was generated using the Illumina NovaSeq platform, whereas transcriptome data from the Chernobyl tissue bank and TCGA THCA were generated using Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA) and IlluminaGA\_mRNA\_DGE platform. Direct gene expression analysis, such as GSEA, was impossible because of these different data-generating platform technologies. In addition, to validate that the effect of radiation-induced carcinogenesis is not associated with tumor behavior in our cohort, we generated three radiation scores from previous studies [42]. The high, low, and high/low dose radiation scores included 34, 46, and 29 genes, respectively.

EMT and BRAF scores were higher in our cohort than that in the other two public repository cohorts, indicating that our cohort may have more biologically aggressive PTC. Furthermore, our cohort had the lowest TD scores, which represents differentiation status. However, there was no difference in these three scores according to the region of residence. The cohort from the Chernobyl tissue bank had the highest EMT, BRAF, and TD scores, whereas the cohort from TCGA THCA had the lowest scores. Radiation-related genes include many genes associated

with DNA damage response. This represents radiation exposure-induced changes in gene expression profiles, as indicated in the present study's results. However, it is also known that the expression of some of these genes increases with genetic damage induced by other factors generated by tumor progression [29,30]. These findings suggest that our cohort, which had lower scores than did the cohort from the Chernobyl tissue bank but higher scores than did the cohort from TCGA THCA, has a molecular biologically aggressive PTC among sporadic thyroid cancers, which provides an insight into understanding the molecular biology of PTC in Koreans.

Our study had some limitations. First, the retrospective nature of this study may pose various limitations that could introduce bias, and there was a lack of randomization. Although we conducted propensity score matching, it cannot account for unobserved confounders. Second, we could not prove the direct relationship between the region of residence and radiation exposure. Notwithstanding, like the Korea Radiation Effect and Epidemiology Cohort (KREEC), we also classified an around area as within 5 km, an intermediate proximate area as 5 to 30 km, and a far-distance area as more than 30 km [31,32] and drew comparisons according to these distances. Third, residence information provided by the patients may be inaccurate, even though we included patients with no change of residence on their medical records. Fourth, our study could not investigate patient exposure levels by residence area. However, according to the World Nuclear Association, the annual exposure for people living within a few kilometers of a nuclear power plant is 0.001 mSV. This appears to be a low dose, compared to the amount exposed by flying from New York to Tokyo (0.1 mSV) and typical chest X-ray exposure (0.1 mSV) (<https://www.world-nuclear.org/>). Finally, owing to funding constraints, bulk RNA sequencing has not been performed in large numbers. Regarding the safety of nuclear power plants, a study that establishes and prospectively tracks a cancer registry is required in the future.

In conclusion, we compared clinicopathological characteristics, frequency of representative driver mutations, and gene expression profiles according to the patients' region of residence in South Korea. Notably, some data supporting the understanding of the molecular biological characteristics of PTC in Koreans were identified; however, regional differences were not observed.

## CONFLICTS OF INTEREST

This study was funded by the Radiation Health Institute of Korea Hydro & Nuclear Power (A19LP20). No potential conflict

of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

Young Suk Jo was supported by the Korean Thyroid Association Basic Medical Research Award 2022.

## AUTHOR CONTRIBUTIONS

Conception or design: J.L., M.J., Y.S.J. Acquisition, analysis, or interpretation of data: J.L., S.J., H.Y.L., S.P. Drafting the work or revising: J.L., M.J., Y.S.J. Final approval of the manuscript: J.L., S.J., H.Y.L., S.P., M.J., Y.S.J.

## ORCID

Jandee Lee <https://orcid.org/0000-0003-4090-0049>

Young Suk Jo <https://orcid.org/0000-0001-9926-8389>

## REFERENCES

- Chen AY, Jemal A, Ward EM. Increasing incidence of differentiated thyroid cancer in the United States, 1988-2005. *Cancer* 2009;115:3801-7.
- Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 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* 2016;26:1-133.
- Oh CM, Won YJ, Jung KW, Kong HJ, Cho H, Lee JK, et al. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2013. *Cancer Res Treat* 2016;48:436-50.
- American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer; Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009;19:1167-214.
- Hughes DT, Haymart MR, Miller BS, Gauger PG, Doherty GM. The most commonly occurring papillary thyroid cancer in the United States is now a microcarcinoma in a patient older than 45 years. *Thyroid* 2011;21:231-6.
- Ito Y, Miyauchi A, Oda H. Low-risk papillary microcarcinoma of the thyroid: a review of active surveillance trials. *Eur J Surg Oncol* 2018;44:307-15.
- Shakhtarin VV, Tsyb AF, Stepanenko VF, Orlov MY, Kopecky KJ, Davis S. Iodine deficiency, radiation dose, and the risk of thyroid cancer among children and adolescents in the Bryansk region of Russia following the Chernobyl power station accident. *Int J Epidemiol* 2003;32:584-91.
- Kitahara CM, Platz EA, Freeman LE, Hsing AW, Linet MS, Park Y, et al. Obesity and thyroid cancer risk among U.S. men and women: a pooled analysis of five prospective studies. *Cancer Epidemiol Biomarkers Prev* 2011;20:464-72.
- Lee J, Lee CR, Ku CR, Kang SW, Jeong JJ, Shin DY, et al. Association between obesity and BRAFV600E mutation status in patients with papillary thyroid cancer. *Ann Surg Oncol* 2015;22 Suppl 3:S683-90.
- Son H, Lee H, Kang K, Lee I. The risk of thyroid cancer and obesity: a nationwide population-based study using the Korea National Health Insurance Corporation cohort database. *Surg Oncol* 2018;27:166-71.
- Stone TW, McPherson M, Gail Darlington L. Obesity and cancer: existing and new hypotheses for a causal connection. *EBioMedicine* 2018;30:14-28.
- Yamamoto H, Hayashi K, Scherb H. Association between the detection rate of thyroid cancer and the external radiation dose-rate after the nuclear power plant accidents in Fukushima, Japan. *Medicine (Baltimore)* 2019;98:e17165.
- Zidane M, Cazier JB, Chevillard S, Ory C, Schlumberger M, Dupuy C, et al. Genetic susceptibility to radiation-related differentiated thyroid cancers: a systematic review of literature. *Endocr Relat Cancer* 2019;26:R583-96.
- Albi E, Cataldi S, Lazzarini A, Codini M, Beccari T, Ambe-ssi-Impiomato FS, et al. Radiation and thyroid cancer. *Int J Mol Sci* 2017;18:911.
- Furukawa K, Preston D, Funamoto S, Yonehara S, Ito M, Tokuoka S, et al. Long-term trend of thyroid cancer risk among Japanese atomic-bomb survivors: 60 years after exposure. *Int J Cancer* 2013;132:1222-6.
- Nagataki S, Takamura N. Radioactive doses - predicted and actual - and likely health effects. *Clin Oncol (R Coll Radiol)* 2016;28:245-54.
- Rodgers BE, Holmes KM. Radio-adaptive response to environmental exposures at Chernobyl. *Dose Response* 2008;6:209-21.
- Ivanov VK, Gorski AI, Maksioutov MA, Vlasov OK, Godko AM, Tsyb AF, et al. Thyroid cancer incidence among adolescents and adults in the Bryansk region of Russia following the Chernobyl accident. *Health Phys* 2003;84:46-60.
- Kim JM, Kim MH, Ju YS, Hwang SS, Ha M, Kim BK, et

- al. Reanalysis of epidemiological investigation of cancer risk among people residing near nuclear power plants in South Korea. *Int J Environ Res Public Health* 2018;15:481.
20. Thomas GA, Bunnell H, Cook HA, Williams ED, Nerovnya A, Cherstvoy ED, et al. High prevalence of RET/PTC rearrangements in Ukrainian and Belarussian post-Chernobyl thyroid papillary carcinomas: a strong correlation between RET/PTC3 and the solid-follicular variant. *J Clin Endocrinol Metab* 1999;84:4232-8.
  21. Morton LM, Karyadi DM, Stewart C, Bogdanova TI, Dawson ET, Steinberg MK, et al. Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident. *Science* 2021;372:eabg2538.
  22. Nikiforov YE. RET/PTC rearrangement in thyroid tumors. *Endocr Pathol* 2002;13:3-16.
  23. Jo YS, Li S, Song JH, Kwon KH, Lee JC, Rha SY, et al. Influence of the BRAF V600E mutation on expression of vascular endothelial growth factor in papillary thyroid cancer. *J Clin Endocrinol Metab* 2006;91:3667-70.
  24. Guan H, Ji M, Bao R, Yu H, Wang Y, Hou P, et al. Association of high iodine intake with the T1799A BRAF mutation in papillary thyroid cancer. *J Clin Endocrinol Metab* 2009;94:1612-7.
  25. Kim HJ, Park HK, Byun DW, Suh K, Yoo MH, Min YK, et al. Iodine intake as a risk factor for BRAF mutations in papillary thyroid cancer patients from an iodine-replete area. *Eur J Nutr* 2018;57:809-15.
  26. Mitsutake N, Fukushima T, Matsuse M, Rogounovitch T, Saenko V, Uchino S, et al. BRAF(V600E) mutation is highly prevalent in thyroid carcinomas in the young population in Fukushima: a different oncogenic profile from Chernobyl. *Sci Rep* 2015;5:16976.
  27. Iwadata M, Mitsutake N, Matsuse M, Fukushima T, Suzuki S, Matsumoto Y, et al. The clinicopathological results of thyroid cancer with BRAFV600E mutation in the young population of Fukushima. *J Clin Endocrinol Metab* 2020;105:dga573.
  28. Seo DH, Lee SG, Lee HY, Jeong S, Park S, Lee J, et al. Lymph node metastasis-dependent molecular classification in papillary thyroid carcinoma defines aggressive metastatic outgrowth. *Clin Transl Med* 2023;13:e1211.
  29. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
  30. Groelly FJ, Fawkes M, Dagg RA, Blackford AN, Tarsounas M. Targeting DNA damage response pathways in cancer. *Nat Rev Cancer* 2023;23:78-94.
  31. Cho IS, Song M, Choi Y, Li ZM, Ahn YO. Power estimation and follow-up period evaluation in Korea radiation effect and epidemiology cohort study. *J Prev Med Public Health* 2010;43:543-8.
  32. Ahn YO, Li ZM; KREEC Study Group. Cancer risk in adult residents near nuclear power plants in Korea: a cohort study of 1992-2010. *J Korean Med Sci* 2012;27:999-1008.
  33. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 2014;159:676-90.
  34. Song YS, Park YJ. Genomic characterization of differentiated thyroid carcinoma. *Endocrinol Metab (Seoul)* 2019;34:1-10.
  35. Wu S, Liu Y, Li K, Liang Z, Zeng X. Molecular and cytogenetic features of NTRK fusions enriched in BRAF and RET double-negative papillary thyroid cancer. *J Mol Diagn* 2023;25:569-82.
  36. Wada N, Duh QY, Sugino K, Iwasaki H, Kameyama K, Mimura T, et al. Lymph node metastasis from 259 papillary thyroid microcarcinomas: frequency, pattern of occurrence and recurrence, and optimal strategy for neck dissection. *Ann Surg* 2003;237:399-407.
  37. Tuttle RM, Haugen B, Perrier ND. Updated American Joint Committee on Cancer/tumor-node-metastasis staging system for differentiated and anaplastic thyroid cancer (eighth edition): what changed and why? *Thyroid* 2017;27:751-6.
  38. Martin SA, Dehler CE, Krol E. Transcriptomic responses in the fish intestine. *Dev Comp Immunol* 2016;64:103-17.
  39. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately down-regulated in human diabetes. *Nat Genet* 2003;34:267-73.
  40. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545-50.
  41. Thomas GA, Bethel JA, Galpine A, Mathieson W, Krznaric M, Unger K. Integrating research on thyroid cancer after Chernobyl: the Chernobyl Tissue Bank. *Clin Oncol (R Coll Radiol)* 2011;23:276-81.
  42. Lu TP, Hsu YY, Lai LC, Tsai MH, Chuang EY. Identification of gene expression biomarkers for predicting radiation exposure. *Sci Rep* 2014;4:6293.