Title page

Molecular and Proteomic Profiles of Radioiodine Refractory Papillary Thyroid Cancer

Running title: Proteomics & RR-PTC

Author

Hanqing Liu [1, #], Jiaxi Wang [1, #], Yan Zhou [2, 3, 4, #], Deguang Kong [1, #], Pingping Hu [2, 3, 4], Lu Li [5, 2], Dan Yang [6], Zhiliang Xu [1], Yaoting Sun [2, 3, 4, *], Chuang Chen [1, *]

Addresses

[1] Department of Breast and Thyroid Surgery, Renmin Hospital of Wuhan University, Wuhan University at Jiefang Road 238, Wuhan 430060, China.

[2] School of Medicine, Westlake University, Hangzhou 310024, Zhejiang Province, China.

[3] Westlake Center for Intelligent Proteomics, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou 310024, Zhejiang Province, China.

[4] Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou 310024, Zhejiang Province, China.

[5] College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310024, China.

[6] Department of Cardiology, the Second Affiliated Hospital of Chongqing Medical University, Chongqing 400010, China.

[#] Equal contribution.

* Correspondence:

Prof. Chuang Chen, M.D.,
Department of Thyroid and Breast Surgery,
Renmin Hospital of Wuhan University,
Wuhan University at Jiefang Road 238, Wuhan 430060, RP China.
Tel: +86(0)27 88041911; Fax: +86(0)27 88041911, Email: chenc2469@whu.edu.cn.

Dr. Yaoting Sun, Ph.D.,
Westlake Laboratory of Life Sciences and Biomedicine,
School of Medicine,
Westlake University (Yunqi Campus) at No. 18 Shilongshan Rd., Hangzhou 310024, PR China,
Tel: +86(0)571 8688 6859; Fax: +86(0)571 8688 6859; Email: sunyaoting@westlake.edu.cn.

Abstract and keywords

Abstract:

Background: Despite the generally favorable prognosis of papillary thyroid cancers (PTCs) following thyroidectomy and potential radioactive iodine (RAI) therapy, approximately one-third of patients experiencing recurrence and metastasis eventually develop resistance to RAI, leading to a poor outcome. However, the mechanisms underlying RAI-refractoriness remain elusive. The present study aimed to assess the molecular and proteomic characteristics of RAI-refractory PTC (RR-PTC) to gain a deeper understanding of this condition.

Methods: The medical records at our institution were reviewed for the selection and grouping of RR-PTC patients and RAI-sensitive controls. RR-PTC patients were divided into three subgroups: continuous RAI uptake (ID), loss of uptake at the first I-131 treatment (iDF) and RAI uptake lost gradually (iDG). Proteomic profiling and targeted next-generation sequencing were performed on formalin-fixed paraffin embedded primary lesions. The incidence of gene mutations and fusions was compared across groups. Bioinformatic analysis was subsequently conducted to identify the differentially expressed proteins and enriched pathways. The correlation of protein expression and gene variances was assessed.

Results: A total of 48 PTC patients with recurrence and/or metastasis were included. The expression profiles of the RR-PTC and control groups were similar. In the subgroup comparison, enriched pathways related to MAPK and TNF signaling were associated with negative I-131 uptake and tumor tolerance with positive I-131 uptake. The *BRAF*^{V600E} mutation was less common in the ID group, whereas the *TERT* promoter mutation was more common in the iDF group. *NCOA4-RET* fusion was more common in the ID group. In addition, four proteins were dysregulated in *BRAF*-mutated PTCs.

Conclusion: The present study constructed the first proteomic profile of RR-PTC. The proteins and pathways identified in the analysis may be promising biomarkers and drug targets. Gene alterations and fusions can aid in the early diagnosis of RR-PTC.

Keywords: thyroid carcinoma, radioactive iodine, refractoriness, proteomics, NGS, *RET/PTC3*.

1 Introduction

Thyroid cancer ranks eighth among female malignancies [1]. The incidence of this disease has increased threefold or greater during the past four decades in many countries [2-4]. Papillary thyroid cancer (PTC) is the most common histological type and accounts for 70-96% of all thyroid cancers [4,5]. Most patients with PTC achieve complete remission after undergoing lobectomy or total thyroidectomy and neck dissection. However, approximately half of PTC patients present with aggressive tumor behaviors, such as tumor invasion into perithyroidal tissues, aggressive histology, vascular invasion, multiple lymph node metastases, and distant metastases [6,7]. These patients have an intermediate or high risk of recurrence and are thus considered or recommended for postsurgical radioactive iodine (RAI) therapy [8]. Unfortunately, one to two-thirds of patients in this subset ultimately develop RAI-refractoriness (RAIR), especially those with a high risk of recurrence [9,10]. RAI therapy barely improves patient prognosis under these conditions. The 10-year survival rate of patients with RAI-refractory PTC (RR-PTC) has decreased to 20-65%, and their life expectancy has substantially decreased [10,11].

Currently, the diagnosis of RR-PTC depends on an I-131 whole-body scan (WBS) and evidence of regional and/or distant metastatic lesions [8]. Owing to the high affinity of thyroid follicular cells for iodine, radioactive isotopes are concentrated in PTC cells and exert a tumor-killing effect by emitting β-rays. In most cases, two or more I-131 WBSs are required for the final confirmation of RAIR. Given that the routine time interval between two I-131 WBSs is six months [12], the diagnosis of RAIR is made nine to twelve months after the initial surgery. The time lag postpones the start of systemic therapy (e.g. multi-kinase inhibitors, MKIs) and/or local treatments. Early identification of the RAIR can avoid unnecessary RAI therapy and thus improve patient prognosis by shifting to alternative treatments.

Many studies have focused on the early diagnosis of RAIR cancer during the past decade. Positron emission tomography (PET) is a valuable tool. Previous studies have shown that the "flip-flop" phenomenon of I-131 WBSs and PET (negative I-131 and positive F-18-FDG uptake) is indicative of the RAIR [13]. In addition to radiological tools, *BRAF*^{V600E} and *TERT* promoter mutations can also contribute to the diagnosis of RAIR. The *BRAF*^{V600E} mutation can impair the expression of the sodium iodide symporter (NIS) and promote malignancy in tumor cells [14,15]. A previous retrospective study revealed that a combination of *BRAF*^{V600E} and *TERTp* mutations, which presented in one-fifth of the PTC cohort, strongly predicts RR-PTC at a positive predictive value of 97.4% [16]. However, for approximately 80% of patients without the genetic duet, the negative predictive value was only 47.6%. A subsequent study revealed that 52.9% of patients with the genetic duet presented with RAIR [17]. Generally, several factors, including high cost and unsatisfactory accuracy, have restricted the use of these diagnostic tools in clinical practice. Currently, the four criteria for RAIR differentiated thyroid cancer (DTC) are still based on the I-131 WBS and additional evidence for recurrence or metastasis [8]. In addition, although the RAIR is related to gene alterations in clinical studies [18,19], the mechanism underlying refractoriness is not fully understood. For example, the NIS protein was not necessarily downregulated, as the above article revealed. It can be upregulated and exert a non-pump pro-tumorigenic effect on thyroid cancer cells [20,21].

Proteins are important participants in tumorigenic processes and dedifferentiation. Mass-based proteomics is a rapidly developing technique that facilitates comprehensive investigations into molecular activities and biological processes involved in cancer initiation and progression. Lots of diagnostic biomarkers and therapeutic targets in clinical practice are based on proteins. Proteomics has been widely used to reveal the cellular activities of thyroid cancer cells since its first appearance [22,23]. Recently,

proteomics has been used in several areas of thyroid cancer research, including discrimination of benign and malignant nodules, subtype identification, and risk stratification [24,25]. During the past two years, Sun *et al.* constructed a comprehensive thyroid tissue proteomic spectral library and distinguished between follicular adenomas and follicular thyroid carcinomas [26]. In addition, Shi *et al.* plotted the multi-omic atlas of medullary thyroid carcinoma via the proteomic technique [27]. The proteomic profile of RR-PTC has not yet been revealed, which may help improve the early diagnosis of RR-PTC in clinical practice and understand the mechanism of RAIR.

This study aimed to construct 23-gene panel-based molecular and proteomic profiles of RR-PTC and to identify promising biomarkers related to RAIR, which could aid in early identification and drug target selection.

2 Results

2.1 Patient cohort

Based on the inclusion and exclusion criteria in the Patients and Methods section, forty-eight patients with locoregional or distant metastatic PTCs were recruited for our study (Table 1). The average age of all patients was 38.4 ± 12.69 years, and the patients in the RR-PTC group were slightly older than those in the control group ($P = 0.421$). Interestingly, all five patients ≥ 55 years old were in the RR-PTC group. Although females were more common in the control group (72.7% *vs.* 54.1%, $P = 0.319$), the difference was not significant. In addition, no differences were found in the tumor diagnosis, smoking status or alcohol consumption, body mass index or the incidence of concomitant thyroid diseases between the two groups.

The histopathological subtypes of PTC were identically distributed between the two groups. No significant differences were found in terms of tumor size, multifocality or bilateral lesions. The incidence of extrathyroidal extension (ETE) was greater in the RR-PTC group with marginal statistical significance ($P = 0.095$). Owing to differences in age and ETE, T2-4 categories were more common in patients with RR-PTC (91.9% *vs.* 63.6%, $P = 0.039$). The presence of lymph node metastatic lesion (LNM) was similar in the two groups, but distant metastasis was much more common in the patients with RR-PTC (56.8% *vs.* 9.1%, $P = 0.006$). Consequently, higher AJCC stages were observed in RR-PTC patients ($P = 0.001$). Compared with patients in the control group, patients in the RR-PTC group received larger doses of I-131 after the initial resection of primary tumors. Notably, all six patients who died were in the refractory group.

2.2 Molecular profile of PTCs correlated with I-131 resistance

A total of 23 genes were sequenced (Sup table S1). Four genes were found to harbor driver mutations (Table 2). The $BRAF^{V600E}$ mutation was identified in 25 samples (56%), with no significant difference between the RR-PTC group and the control group (54.3% *vs.* 60.0%, $P = 1.000$). All six samples with *TERTp* mutations (17.1% *vs.* 0%, $P = 0.312$) were in the RR-PTC group. In addition, *RET* and *TP53*

mutations were found in one and two patients, respectively. *RAS* mutations, including *HRAS*, *KRAS* and *NRAS*, were not detected in any samples. In addition, four kinds of gene fusions were found, namely, eight samples with *NCOA4-RET*, two with *CCDC6-RET*, three with *ETV6-NTRK3* and one with *STRN-ALK*. Fusions were more frequent in the control group (28.6% *vs.* 40.0%, $P = 0.700$), whereas *RET*-related fusions were almost equally distributed (22.9% *vs.* 20.0%, $P = 1.000$). However, no significant differences in the expression of these genes were found between the two groups.

The forty-five samples were then divided into four subgroups. The $BRAF^{V600E}$ mutation was less common in the continuous RAI uptake (ID) group (11.1%, $P < 0.05$), whereas the *TERTp* mutation was more frequent in the loss of uptake at the first I-131 treatment (iDF) group (33.3%, $P < 0.05$). Notably, *NCOA4-RET* fusion was more common in the ID group (44.4%, $P < 0.05$).

The present study further evaluated the associations between gene variants and I-131 uptake. The samples in the ID and RAI-sensitive PTC (Id) groups were categorized as positive for I-131 uptake. $BRAF^{V600E}$ mutation was more frequent in the negative I-131 uptake group with statistical significance (69.2% *vs.* 37.8%, $P = 0.039$). The incidence of *TERTp* mutations was also greater in this group (19.2% *vs.* 5.3%, $P = 0.222$). In contrast, fusions were more common in the positive uptake group (23.1% *vs.* 42.1%, $P = 0.206$).

The distribution of gene variants with unknown clinical significance was also assessed (Sup table S2). Four variants were subjected to nonparametric tests after those with a frequency < 3 were filtered out. The four mutations were located on *ATM* (c.1236-2A>T), *PTEN* (c.402G>A) and *RET* (c.2071G>A and c.2671T>G). Three variants were more common in the RR-PTC group, except for *PTEN* c.402G>A, although the difference was not statistically significant. In subgroup analysis, the incidence of *PTEN* c.402G>A was greater in the ID group (22.2%, $P < 0.05$). In addition, *PTEN* c.402G>A was associated with positive I-131 uptake with marginal significance (0% *vs.* 15.8%, $P = 0.068$).

2.3 Overview of the proteomic profiling

To profile the proteomic characteristics of RR-PTC, we collected 168 tissue slides from 73 patients. A total of 9769 proteins were identified (Sup figure 1A), which indicated the high quality of our data. Since the number of proteins identified in all the samples exceeded the threefold interquartile range (≤ 1614), no sample was excluded from subsequent analysis. The principal component analysis (PCA) plot annotated by batch groups indicates no observable batch effect during data acquisition (Sup figure 1B). Although samples were well resolved by their site, tumor tissues mixed with little difference but negative lymph nodes were clustered (Sup figure 1C). The *R* values of correlation analyses exceeded 0.97 in both intra-batch and inter-batch tests, indicating high technical stability across the data acquisition (Sup figure 1D&E). Overall, no observable batch effect was detected (Sup figure 1F).

Next, the raw files were converted to an expression matrix. We applied a filtering criterion to select proteins with missing values less than 50%. These 7406 filtered proteins were then subjected to data normalization and outlier substitution (Sup figure 2A-C). After ID conversion, an expression profile encompassing 7394 proteins across 168 samples was compiled (Sup figure 2D). As indicated by the *ESTIMATE* algorithm, the median tumor purity was 76.2% (Sup figure 2E), indicating the reliability of our dataset for further analysis.

The global proteomic profiles of the original PTC lesions were subjected to dimensionality reduction to evaluate the clustering of the samples. The area covered by the samples represented similarity within

a group. The RAI-sensitive samples were more tightly clustered than the RR-PTC samples (Figure 1A). The overlap of the two groups might be explained by the similarity between tumor tissues and the intrinsic heterogeneity of RR-PTC samples. To minimize potential heterogeneity, we furtherly divided RR-PTCs into three subgroups (iDF, iDG and ID) as described in the patients and tissue selection section. The diversity of the ID group was the lowest among the three subgroups, while the diversity of the RAI uptake lost gradually (iDG) group was the highest (Figure 1B). Interestingly, when labeled with the *BRAF*^{V600E}-mutation status, samples in the two groups were separately clustered (Figure 1C).

Subsequently, 107 differentially expressed proteins (DEPs) with P values < 0.05 and absolute fold change (FC) > 1.414 were identified via the multiple linear regression algorithm (Figure 1D). Eighty-nine proteins were upregulated in the RR-PTCs, whereas the remaining eighteen proteins were downregulated. The top upregulated protein in the RR-PTCs was KRT71 (keratin, type II cytoskeletal 71, $\log_2$FC = 1.422). Interestingly, the top downregulated protein also belonged to the keratin family (keratin, type I cytoskeletal 16, KRT16, $\log_2$FC = -0.955), which suggested the potential involvement of the cytoskeleton in the progression of RAIR. Samples were clustered unsupervisely based on the 107 DEPs and then labeled with clinicopathological features (Figure 1E). The borders of the clinicopathological labels were generally indistinct. However, the thyroid differentiation score (TDS) and extracellular signal-regulated kinase (ERK) score were correlated with the groups to some extent (Figure 1F). The TDS was lower in the RR-PTC group, which could indicate poorer differentiation (P = 0.288). Conversely, the ERK score was greater in the RR-PTC group (P = 0.215). Unfortunately, none of these scores reached statistical significance.

Gene set enrichment analysis (GSEA) based on the Reactome database revealed 64 pathways with statistical significance (Sup figure 3A). Twenty-nine pathways with positive normalized enrichment scores (ESs) were unregulated in RR-PTC samples. The pathway with the greatest increase was gamma carboxylation hypusine formation and arylsulfatase activation (nES = 1.687). In addition, Akt phosphorylates targets in the cytosol was also upregulated (nES = 1.552). The remaining 39 pathways were downregulated. Regulation of mRNA stability by proteins that bind Au-rich elements was mostly suppressed for RR-PTCs (nES = -1.960). Notably, *RAS* processing was also downregulated (nES = -1.728). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was also used for the annotation of pathways closely related to oncogenesis and dedifferentiation (Sup figure 3B). P53 signaling was significantly enriched (nES = 1.5440, P = 0.0136). Interestingly, the MAPK signaling pathway showed a slight positive correlation with RR-PTCs (nES = 1.2623, P = 0.1642). The gene set variation analysis (GSVA) results supported the above results (Sup figure 3C).

A previously published study assessed the transcript profiles of RR-PTC via a gene microarray [28]. The DEPs identified in the present study were assessed via this profile to determine whether they were robust in another cohort. Ninety-nine of the 107 DEPs were identified in this cohort, and only three of them were dysregulated (Sup figure 3D). BET1 and C11orf96 were downregulated in the RR-PTC group (BET1 median: 6.534 *vs.* 6.105, P = 0.032; C11orf96 median: 7.338 *vs.* 7.028, P = 0.032). The expression of SLC4A9 was greater in the RR-PTC group (median: 5.216 *vs.* 5.392, P = 0.045). Notably, the regulation direction of BET1 and SLC4A9 was consistent with the results of the present study, whereas C11orf96 was oppositely expressed in the two analyses. GSVA based on the Reactome and KEGG databases revealed two differentially downregulated pathways (N-glycan antennae elongation and GAB1 signalosome), none of which were enriched in our study (Sup figure 3E&F).

To assess the spectral change in RR-PTC, the Id, ID and iDF subgroups were defined as having low, intermediate and high severity, respectively. A total of 403 proteins with significance difference in

analysis of variance were divided into six clusters (Sup figure 4). In two clusters, 160 proteins were monotonically regulated (Sup figure 5A). Collagen biosynthesis and modifying enzymes were significantly enriched in Metascape (Sup figure 5B). In addition, regulation of protein autophosphorylation and regulation of extrinsic apoptotic signaling pathway were also associated with the severity of RR-PTC. The results were supported by the ingenuine pathway analysis (IPA) database (Sup figure 5C).

2.4 Proteomic biomarkers and pathways associated with I-131 uptake

For analysis of proteomic biomarkers and the underlying mechanisms for the loss of I-131 uptake in PTC, two subgroups, ID and iDF, were included in the subsequent analysis of this section. The ability of these two groups to take up I-131 was completely different at the first radiotherapy. The demographic and clinicopathologic characteristics of the two groups were roughly comparable (Sup table S3). No significant differences were observed in histopathological features, tumor staging or risk stratification. Interestingly, four patients in the iDF group drank alcohol and/or smoked, while no individual in the ID group reported these behaviors ($P = 0.033$). In addition, the patients in the iDF group had greater weights than their counterparts ($P = 0.010$). Since the diagnosis of RR-PTC in the ID group required multiple rounds of I-131 treatment, the cumulative dose of I-131 was inevitably greater than that in the iDF group ($P = 0.001$).

Although the ID and iDF groups overlapped with each other to some extent, a total of 145 DEPs were identified with P values < 0.05 and absolute FC > 1.414 (Figure 2A&B). Ninety-five proteins were upregulated in the ID group, whereas the remaining 50 proteins were highly expressed in the iDF group. The 145 DEPs could barely separate the two groups. The TDS was greater in the ID group, whereas the ERK score was lower. However, none of these scores reached statistical significance (Figure 2C&D).

Pathway enrichment analysis was then conducted. Only seven pathways in the KEGG database were significantly enriched. Among the seven pathways, glutathione metabolism and Notch signaling were enriched in the ID group whereas TGF-beta signaling and ECM receptor interaction were enriched in the iDF group (Figure 2E). Fifty-one pathways in the Reactome database were significantly enriched. The top enriched pathway was “activated TAK1 mediates P38 MAPK activation” with an nES of 1.9182 (Figure 2F). In addition, several MAPK-related pathways, including ERK/MAPK targets and prolonged ERK activation events, were correlated with positive I-131 uptake. In addition, TNF signaling and glutathione conjugation were associated with positive I-131 uptake.

The results of pathway enrichment were then assessed via GSVA. Sixty-three pathways were significantly enriched and 18 of these pathways overlapped with the GSEA results (Figure 3A). After the associations between pathways and tumor progression were evaluated, sixteen pathways were found to play roles in the loss of I-131 uptake in PTC, eleven of which overlapped with the GSEA results (Figure 3B). Six pathways were closely related to MAPK signaling. Three pathways are associated with transforming growth factor-β activated kinase 1 (TAK1) signaling. In addition, one pathway was correlated with both MAPK and TAK1 signaling. The other six pathways involved glutathione metabolism, TNF signaling and MET proto-oncogene (MET) signaling. The expression of proteins involved in these thirteen pathways was compared between the RAI-positive and RAI-negative groups. Proteins with an absolute FC > 1.2 were displayed (Figure 3C).

The results of our proteomic profile were validated with a transcriptomic profile of RR-PTCs

(GSE151179). The samples in GSE151179 were categorized as ID or iD based on their I-131 uptake. The iD group was not subdivided into iDF or iDG since the information was not provided. A linear regression algorithm identified three common differentially expressed genes (DEGs) (Sup figure 6A). However, the regulatory direction was completely different. In the proteomic profile, MYH7 was up-regulated, whereas in the transcriptomic profile, its expression was greater in the iD group. S100B and UXT also exhibited inverse changes in the two profiles. Only MYH7 significantly differed according to the nonparametric test ($P = 0.035$). GSVA revealed five common pathways (Sup figure 6B), and the direction of the change in expression was consistent between the two profiles. Notably, three TAK1-associated pathways were significantly enriched. All genes involved in the 16 selected pathways and these five common pathways were tested. Only three genes were differentially expressed between the two groups. PI4K2B and PIK3R4 participate in the synthesis of PIPs at the early endosome membrane, and CEACAM8 is involved in fibronectin matrix formation (Sup figure 6C).

Immune analysis based on the ssGSEA matrix revealed an abundancy of memory B cell and a deficiency of macrophage and immature T cell in the ID samples (Sup figure 7A). Notably, eosinophils were more abundant in RAI-sensitive samples compared to RR-PTCs. However, immunohistochemistry results did not show any statistically significant differences (Sup figure 7B).

2.5 Proteomic biomarkers and pathways for different responses to positive I-131 uptake

The data of twenty-one patients in the ID group and the Id group (or control group) were compared to determine the potential mechanisms underlying the different responses to positive I-131 uptake in PTC patients. The baseline characteristics of the two groups were compared (Sup table S4). No significant differences in demographic or tumor features were found, except for a greater incidence of ETE in the ID group ($P = 0.035$). More distant metastases were found in the ID group (80% *vs.* 9.1%, $P = 0.002$), and consequently, the disease stages were greater ($P < 0.001$). According to the criteria of the ID group, the cumulative I-131 dose was inevitably greater ($P < 0.001$).

Similar to the previous comparison, the two groups could be distinctly separated from each other (Figure 4A). Twenty-one samples were roughly divided into two groups with 125 DEPs identified via the multiple linear regression algorithm (Figure 4B&D). No significant differences were found in the TDS, ERK score, T cell infiltration score (TIS), immune infiltration score (IIS) or immune cytolytic activity score (CYT) (Figure 4C).

GSEA revealed twelve enriched pathways based on KEGG annotations. The p53 signaling pathway ($P = 0.0498$) and mTOR signaling pathway ($P = 0.0346$) were both promoted in the RAI-resistant samples (Figure 4E). Interestingly, inositol phosphate (IP) metabolism was also activated in the ID group ($P = 0.0190$). No significant difference was revealed in the MAPK signaling pathway. Forty-five pathways were significantly enriched according to the Reactome annotation file. TP53-related pathways, including regulation of TP53 through association with co-factors ($P = 0.0260$) and methylation ($P = 0.0463$), were increased in ID samples (Figure 4F). TNF signaling ($P = 0.0188$) and its participant TNFR1-induced NF-kappa-B signaling pathway ($P = 0.0021$) were both associated with RAI uptake and resistance. In addition, the pathway "activated TAK1 mediates p38 MAPK activation" was closely related to RAI-resistance ($P < 0.0001$). An increase in IP metabolism was also detected ($P = 0.0081$).

GSVA based on the Reactome annotation file was conducted to verify the previous results. A total of 49 pathways were enriched with significant differences, eleven of which overlapped with previous results

(Figure 3D). After the search and selection, four clusters comprising thirteen pathways were pooled for subsequent analysis (Figure 3E). Eight of the thirteen pathways were also dysregulated according to GSEA. A total of 39 proteins involved in TNF signaling, TAKI-related pathways, plasma lipoprotein remodeling and IP-related pathways were identified with an absolute FC > 1.2 (Figure 3F).

The expression levels in the transcriptomic profiles of the ID and Id groups were then compared for validation. Only PGS1 was downregulated in the ID group (median: 6.080 *vs.* 6.456, $P = 0.012$; Sup figure 6D), which was consistent with the proteomic data. According to the pathway analysis, only GAB1 signalosome was significantly decreased in the ID group ($P = 0.012$, Sup figure 6E). However, none of the genes in the GAB1 signalosome were differentially expressed between the two groups. Among those involved in the above thirteen pathways, three genes (MAPKAPK2, NFKB1 and PLKHA6) were downregulated in the ID group, in contrast to the expression pattern observed in our proteomic profile (Sup figure 6F).

2.6 Association of gene variance and protein expression

The 7394 proteins were divided into 21 modules via weighted correlation network analysis (WGCNA) (Figure 5A&B). *BRAF* mutation was significantly correlated with seven gene modules, among which the blue module had the highest absolute correlation coefficient (CC) (CC = 0.61, $P = 8.4e-6$). The most prominent module associated with *TERTp* mutations was pink (CC = -0.45, $P = 1.9e-3$). The green-yellow module was closely related to the gene fusions (CC = 0.42, $P = 3.7e-3$) and the *NCOA4-RET* fusion (CC = 0.32, $P = 0.03$). For other gene variants of unknown significance, *ATM* c.1236-2A>T was related to the green module (CC = 0.41, $P = 5.0e-3$). In addition, *RET* c.2671T>G was closely associated with the yellow module (CC = 0.50, $P = 4.5e-4$).

Three machine learning algorithms (LASSO, RF and SVM-RFE) were used to select biomarker proteins from the most correlated modules of each gene mutation and fusion. Since other gene mutations were generally rare (< 10%) in the present study, samples with *BRAF* mutation, *TERTp* mutation, gene fusion and *NCOA4-RET* fusion were retained for subsequent analysis. A total of 28, 32 and 39 proteins were used for model construction of *BRAF* mutations via the least absolute shrinkage and selection operator (LASSO) regression, random forest (RF), and support vector machine-recursive feature elimination (SVM-RFE) algorithms, respectively, and four of these proteins were common biomarkers (Figure 5C). In addition, only PTCD1 and AHCYL2 were revealed to be common biomarkers for gene fusions. According to the receiver operating characteristic (ROC) curves of The Cancer Genome Atlas Thyroid Cancer (TCGA-THCA) dataset, the areas under the curve (AUCs) of three genes exceeded 0.9 (Figure 5D). PTPRE had the highest accuracy (AUC = 0.93), followed by TMEM43 (AUC = 0.92) and FN1 (AUC = 0.91). According to the merged Gene Expression Omnibus (GEO) profile, the most accurate gene was TMEM43 (AUC = 0.80), followed by FN1 (AUC = 0.73) and SUSD2 (AUC = 0.69). Interestingly, PTPRE did not significantly differ. The AUCs of AHCYL2 for gene fusions in TCGA-THCA and merged GEO cohorts were 0.65 and 0.68, respectively. Disease free survival analysis revealed that dysregulation of FN1 (HR = 2, $P = 0.026$) and AHCYL2 (HR = 0.53, $P = 0.038$) was associated with poor prognosis in patients with thyroid cancer (Sup figure 8).

To explore the potential biomarkers and pathways in case of *BRAF* mutation and wild-type *TERTp*, only 19 *BRAF*-mutated and *TERTp*-unmutated samples were analyzed. A total of 184 proteins were significantly different between the RAIR-refractory and RAIR-sensitive groups (Sup figure 9A&B).

Metascape revealed that several pathways related to the immune response, including T cell activation, regulation of lymphocyte activation and phagocytosis, were enriched with statistical significance (Sup figure 9C). IPA results also supported these findings (Sup figure 9D). Granzyme A signaling was downregulated in *BRAF*-wild-type samples, whereas neutrophil degranulation, B cell development and phagosome formation were upregulated in *BRAF*-mutated *TERTp*-unmutated samples.

3 Discussion

To our knowledge, the present study is the first to construct a proteomic profile of RR-PTC on primary lesions. We evaluated the differences in the expression of genes in the RAI-refractory and RAI-sensitive samples. Three genes were potential biomarkers for RR-PTC. To minimize the intrinsic heterogeneity due to different criteria, we subdivided RR-PTC primary lesions based on the response to radiotherapy and the presence of I-131 uptake. Proteins and potential pathways associated with these two tumor behaviors were evaluated. MAPK signaling was the most commonly involved pathway. TGF-beta and TAK1 signaling was revealed to be associated with the loss of I-131 uptake. In contrast, tumor necrosis factor (TNF) signaling might affect cancer cell death in positive foci. Targeted deep sequencing revealed that wild-type *BRAF* and *TERTp* mutation were associated with the ID and iDF phenotypes, respectively. *NCOA4-RET* fusion was more frequent in the ID group. With respect to protein expression and gene variance, four proteins were associated with $BRAF^{V600E}$ mutation.

Many previous studies have investigated the dedifferentiation mechanism and redifferentiation targets of RR-PTC. NIS is the most important transporter of iodide and is highly expressed in thyroid epithelium [29]. Previously, researchers hypothesized that the downregulation of NIS in primary tumors contributed to the loss of RAI uptake in metastatic RR-DTC lesions [30]. This hypothesis was soon challenged by the intracellular overexpression of NIS in ~70% of thyroid cancer samples [21]. In 2018, Feng *et al.* reported that increased intracellular NIS could exert a non-pump pro-tumorigenic effect [20]. The relationship between NIS expression and the RAIR has yet to be determined [31,32]. Owing to the methodological limitation of proteomics, we failed to quantify NIS protein expression. However, seven molecules indicating thyroid differentiation, including PAX8, TG and TPO, were identified. Although the TDS tended to increase in the RR-PTC samples, the difference did not reach statistical significance. Consequently, our results do not fully support that thyroid-specific proteins in primary lesions could be used as diagnostic biomarkers for RR-PTC.

The associations between thyroid cancer and gene variants have been widely discussed. The pathway most likely related to dedifferentiation and RAI refractoriness was MAPK signaling [33]. MAPK signaling regulates cell proliferation, dedifferentiation and death. *BRAF* is an essential kinase in the cascade. The $BRAF^{V600E}$ mutation has been reported to be associated with aggressive tumor behavior and poor prognosis [34,35]. Recent studies have revealed a strong correlation between $BRAF^{V600E}$ mutation and negative iodine uptake [16,28], which was confirmed by a meta-analysis [36]. The overall prevalence of *BRAF* mutation was ~55% in our study, but ID samples were less likely to harbor *BRAF* mutation (~10%) than other subgroups. Given that *BRAF* mutation can impair the expression of the NIS protein [37], recurrent or metastatic lesions without *BRAF* mutation are prone to be positive in I-131 WBSs. Interestingly, a more recent study by Mu *et al.* revealed a lower rate of *BRAF* mutation in the "continuously RAI-avid but RAI-refractory" group (~14%) than in the partial and gradual RAI-refractory groups [38]. Although few

studies have subdivided RR-PTC based on different criteria, we could hypothesize that ID tumors have distinct molecular features and tumorigenic mechanisms. However, it is difficult to explain the high prevalence of *BRAF* mutation in the Id group in our study, which was not consistent with many previously published articles [31,38,39]. Some studies have reported no significant correlation between *BRAF* status and RAI uptake [40,41]. It was assumed that clinical pathways in different countries/regions and patient selection protocols contributed to the diverse observations to some extent [35].

In addition, *TERTp* mutation is a promising predictive factor for the RAIR. The prevalence of this mutation was lower than that of the *BRAF* mutation, but patients harboring *TERTp* mutations had poorer prognoses and greater mortality [42]. Liu *et al.* reported that the genetic duet of *BRAF* and *TERTp* mutations robustly predicted the loss of RAI avidity in PTCs with a high positive predictive value (97.4%) [16]. However, the negative predictive value was lower than 50%, which limits its utility in clinical practice. Subsequent studies have shown similar but somewhat different results [17]. The genetic duet could predict RAIR with a positive predictive value barely higher than 50%. A recent study from Shanghai revealed that *TERT* accelerated *BRAF*-mutated thyroid cancer dedifferentiation and progression by regulating ribosome biogenesis [43]. Our study also revealed a greater rate of *TERTp* mutation in the RR-PTC group. Interestingly, *TERTp* mutation was significantly more frequent in iDF samples, indicating its strong association with the loss of RAI uptake. This observation was in accordance with a previously published article [38]. Since all six *TERTp*-mutated samples harbored the $BRAF^{V600E}$ mutation, we did not test the combined efficacy of the two mutations.

Of interest, our study revealed a significantly greater incidence of *NCOA4-RET* (*PTC3-RET*) fusion in the ID group. This phenomenon was also observed in two studies. An Italian study revealed that fusion genes (especially *RET-PTC*) were more common in ID (39%) than in iD (16%) and Id (13%) patients ($P = 0.075$) [28]. Mu's study revealed that the incidence of *RET*-fusions in the ID subgroup (~40%) was greater than that in other RR-DTC subgroups and the Id group [38]. However, statistical significance was not reached in the above two studies. A subsequent study from Mu's institution revealed that fusion oncogenes in pediatric DTCs were associated with RAI-refractoriness ($P = 0.017$) [44]. But *RET*-fusions were not analyzed separately. The radioactive isotope is concentrated in thyroid epithelium-derived tumor cells and exerts tumor-killing effect by emitting β-rays. Considering that *RET* rearrangements can activate its kinase and inhibit apoptosis via the MAPK and PI3K signaling pathways [45,46], the apoptosis triggered by radioisotope-induced DNA damage could also be impaired by *RET* alterations, especially the *NCOA4-RET* fusion. The potential causal relationship between *RET* fusion and the ID phenotype has not been verified in cell or animal experiments. Recent clinical studies have shown that the *RET* fusion-directed therapy can restore RAI avidity in patients with RR-PTC [47,48].

The accuracy of gene mutations and fusions are helpful but not satisfactory. Proteomic analysis was conducted to identify promising biomarkers related to the RAIR. Although the expression differences between phenotypes were generally not obvious, some signaling pathways were significantly enriched. In the comparison of the RAI-sensitive and RAI-refractory groups, P53 signaling was the pathway most likely associated with the loss of iodine uptake. Although some scholars have suggested that *TP53*-mutated follicular adenomas are precursors for the dedifferentiation of anaplastic thyroid cancer [49], few studies have investigated the correlation between the RAIR and mutations in P53 signaling. At that time, we assumed that the heterogeneity in RR-PTC reduced the power of the test. In the subsequent comparisons of subgroups (ID *vs.* iDF and ID *vs.* Id), enriched pathways clustered into several major pathways, including the MAPK signaling and TNF signaling pathways. The MAPK pathway is strongly related to thyroid cancer behaviors, and *BRAF* mutation impairs the expression of the NIS protein by

deacetylating its gene promoter histones [37]. MAPK inhibitors and histone deacetylase inhibitors can restore iodine uptake and re-differentiate PTC cells [50,51]. Interestingly, our study revealed that TAK1-associated pathways were also correlated with that RAIR. A recent study suggested that silencing TAK1 inhibited the proliferation and migration of thyroid cancer cells via that suppression of p38 MAPK signaling [52]. We hypothesize that TAK1 is a novel potential molecular target of RAIR. The TNF signaling is another pathway of interest. TNF is an important cytokine that triggers inflammation. A retrospective study by Gheorghe *et al.* suggested that TNF-α might exert different antitumor effects in response to RAI therapy depending on the patient's immune profile [53]. The activation of TNF signaling might restore I-131 uptake and promote redifferentiation in thyroid cancer cells. Unfortunately, most proteins involved in these pathways did not significantly differ.

Most of the molecules identified in the article were not the main regulators of their pathways. However, these three proteins warrant further discussion. There are few reports of S100B in thyroid cancer, but its high expression promotes cancer metastasis via interaction with P53 signaling in other glandular epithelium-derived carcinomas [54]. NFKB1 belongs to the well-known NF-κB signaling, which is closely related to cancer initiation and progression. A functional polymorphism in its promoter increases the risk of PTC [55]. In addition, FN1 overexpression was found in aggressive thyroid cancer and promoted its migration and invasion [56]. The expression of this molecule was greater in *BRAF*-mutated PTCs and was indictive of poor prognosis [57]. These proteins are associated with aggressive cancer behaviors and could be promising biomarkers for RR-PTC.

The current study did not focus on the iDG group because of its high intrinsic heterogeneity, but one hypothesis could help explain the process of loss of I-131 uptake in this group. The original tumor may comprise heterogeneous cancer cells with different tolerances to I-131 treatment. I-131 exerted a selective effect on cells, and those with greater tolerance survived, which ultimately results in the formation of negative uptake foci on I-131 WBSs.

Several previous studies have plotted the molecular and omic atlas of RR-DTC (Sup table S6). Sabra and Shobab independently constructed the genomic landscape for RR-DTC [58,59]. Mutations in the MAPK signaling pathway accounted for ~50% of oncogenic drivers, which was similar with our results. In 2020, Colombo *et al.* constructed molecular and gene/miRNA profiles for RR-PTC [28]. *BRAF* mutation was more frequent in the RAI-negative RR-PTC group. Although PTCs were clearly distinguished from normal thyroid tissues, no distinct expression patterns were found between RAI-refractory and RAI-sensitive PTCs. In general, our results were consistent with those of Colombo's study. A distinguishable feature for RAIR was not found based on either the mRNA/miRNA or protein profile, regardless of subclassification. The insufficient sample size may partly explain this phenomenon. One hypothesis is that the RAIR occurs during recurrence and metastasis. The hypothesis, however, is not convincing in the presence of some RAI negative lesions at the first I-131 treatment. Another competing hypothesis is that the underlying mechanisms of the RAIR are trivial and are covered by the heterogeneity of other tumor behaviors. More homogenous samples and proper subclassification are needed to eliminate confounding effects and uncover the underlying mechanisms involved. Two precursor studies provided novel insights into the proteomic expression of RR-PTC but were limited by small sample sizes and the lack of primary tumors [60,61].

The current study has several shortcomings. First, although the medical records of more than ten thousand patients were reviewed thoroughly by independent researchers, the intrinsic nature of this single-center, retrospective study and the relatively small sample size limit the generalizability of our conclusions. In addition, the lack of Benjamini-Hochberg correction for differential expression analysis

undermined the reliability of the results. In addition, due to the lack of previous proteomic profiles, potential protein biomarkers were tested with several external transcriptomic profiles. The consistency of expression across mRNAs and corresponding proteins is debatable [27]. Moreover, FTC samples and metastatic lesions were excluded from the present study because of the small sample size and potential histopathological heterogeneity. The potential biomarkers and mechanisms were not verified with *in vivo* experiments. The above observations of our study could be confirmed with multicenter large-sample proteomic profiling. Cell and animal experiments could help verify our results.

In conclusion, a proteomic profile based on RR-PTC primary lesions was constructed for the first time in the present study. Our current work improves the molecular and biological understanding of RR-PTC, which could enlighten future preclinical and clinical studies toward molecule-guided treatment. The dataset created in this study could serve as an important resource for further investigations of RR-PTC biology and therapeutic targets.

4 Patients and Methods

4.1 Study design and ethics approval

This retrospective, case-control study was conducted at Renmin Hospital of Wuhan University. The Institutional Ethical Committee of the hospital reviewed and approved the study design (No. WDRY2021-K032). The requirement for obtaining informed consent from the involved patients was waived due to the retrospective nature of the study. The study was conducted in accordance with the Declaration of Helsinki [62]. The study was performed in accordance with the STROBE checklist for case-control study (version 4).

4.2 Patient and tissue selection

The medical records of patients who were diagnosed with thyroid cancer at our tertiary center from Jan. 1st, 2016 to Dec. 30th, 2022 were reviewed (Figure 6A). Patient demographic and clinicopathological characteristics, laboratory test results and radiological results were collected by two researchers and independently evaluated. During the initial screening, patients who a) had incomplete medical records and b) did not undergo I-131 WBS were excluded. In the second round of assessment, histopathological data and radiological images, including chest computed tomography (CT), cervical magnetic resonance imaging (MRI), whole-body bone scanning and positron emission tomography (PET), were used to evaluate local and distant metastasis at the first admission for surgery. Patients without cervical or distant metastasis at diagnosis were excluded from the subsequent evaluation. Patients with RAI uptake limited to thyroid bed and disease remission were also excluded. In the final assessment, RAI uptake was determined by I-131 WBSs or I-131 single photon emission computerized tomography (SPECT). A total of 73 patients were identified to have locoregional or distant metastasis, which was confirmed by radiological, cytopathological or histopathological solid evidence. Patients with primary lesions not resected in our center or other histopathological types other than PTC were also excluded from further

statistical analysis. Finally, forty-eight PTC patients with locoregional and/or distant metastasis were included in this study and their primary tumor samples (one sample from each patient) were used for subsequent analysis.

The included patients were then divided into two groups based on their response to RAI treatment: RAI-sensitive (RAI uptake positive and disease remission, Id) and RAI-refractory (disease persistence, Figure 6B). RAI refractoriness was identified in accordance with the 2015 American Thyroid Association (ATA) guideline [8]. Briefly, patients with metastatic lesions that a) did not ever concentrate RAI, b) lost the ability to concentrate RAI after previous evidence of RAI avidity, and c) concentrated in some lesions but not in others were considered to have negative RAI uptake. Individuals with d) metastatic disease despite a significant concentration of RAI were defined as having positive uptake. Owing to the potential heterogeneity of RR-PTC, patients in the RAI-refractory group were further categorized into three subgroups: a) negative RAI uptake at the first RAI treatment with disease persistence (iDF), b) RAI uptake lost gradually after previous RAI treatments with disease persistence (iDG), and c) positive RAI uptake but with disease persistence (ID). Tumor stage and the risk of recurrence were stratified using the American Joint Committee on Cancer (AJCC) staging system and the 2015 ATA guideline [8,63].

Resected primary and metastatic lesions were preserved in formalin-fixed paraffin-embedded (FFPE) tissue blocks at room temperature. A total of 168 FFPE samples were collected from the above 73 patients (Figure 6C). The tissue types included primary tumors, synchronous cervical LNM, LNM after I-131 treatment, negative lymph nodes, and regional and distant metastatic lesions. Hematoxylin and eosin-stained slides from all samples were reviewed by two independent pathologists with expertise in thyroid pathology. The tissue blocks with the highest tumor purity in each patient were selected for subsequent analysis. The borderlines of the tumor and adjacent tissues were carefully marked by expert pathologists.

4.3 Proteomic data acquisition and preprocessing

The FFPE samples were prepared for subsequent proteomic analysis as described previously [26,64]. The samples were allocated into ten batches. The tissues were subjected to a series of manipulations, including dewaxing, rehydration and lysis, for peptide extraction and digestion via pressure cycling technology (PCT). Peptides were then quantified via a liquid chromatography (LC) system coupled with a trapped ion mobility spectrometry mass spectrometer (MS). Data-independent acquisition (DIA) files were acquired and then analyzed against a thyroid tissue specific spectral library [26] using DIA-NN (v1.8.1) [65]. Correlation analysis was conducted to assess the inter-batch and intra-batch stability of data acquisition. The profile was then processed for protein filtration, imputation and ID conversion. A detailed description of proteomic data acquisition, quality control and preprocessing can be found in the supplementary methods.

4.4 Targeted next-generation sequencing (TNGS)

The primary lesions from 48 patients with PTC were subjected to molecular profiling via TNGS. The protocol was performed as described previously [66]. Briefly, DNA was extracted from FFPE tissue sections using the QIAamp DNA FFPE Tissue Kit (Qiagen, Germany) following the manufacturer's

instructions. After amplification of targeted DNA fragments and removal of primers, the products were purified using an Ion AmpliSeq Library Kit (Thermo Fisher Scientific, USA). The concentrations of samples were quantitated by a NanoDrop system (Thermo Fisher Scientific, USA). The quality of the purified DNA was evaluated by 1% agarose gel electrophoresis. Severe degradation was detected in three samples, which were excluded from further analysis.

The remaining 45 library products were sequenced via 150 bp paired-end runs on the NextSeq 500 platform (Illumina, Inc., USA). The medians of sequencing depth and coverage were 5136× and 98.7%, respectively. Sequencing data were aligned to a reference human genome dataset (hg19/GRCh37). Subsequently, read mapping, quality control, variant calling and genotyping were performed following the protocols of the OncoAim® Thyroid Cancer Multigene Assay Kit (Singlera Genomics, Inc., China). Mutations and fusions were evaluated for 23 genes (Sup table S1). The minimum confidence threshold for variant allele frequency was 5%. The ENSEMBL Variant Effect Predictor (v90) was used for variant functional annotation. In addition, the gene variants were searched against the ClinVar database (v2020006) [67]. Somatic gene variants were categorized for clinical significance based on the Catalogue Of Somatic Mutations In Cancer database (COSMIC, v97) [68].

4.5 Bioinformatic analysis

DEPs were identified using a multiple linear regression algorithm and analysis of variance (ANOVA). Differential expression analysis was conducted using the *limma* R package. Proteins with an absolute FC > 1.414 and a P value < 0.05 were defined as DEPs. The dimension of the profile was reduced and visualized using the PCA algorithm. ANOVA was used to identify proteins correlated with the severity of RR-PTC. Proteins with statistical significance were clustered using the *mfuzz* R package. Proteins in monotonically regulated clusters were regarded as DEPs.

Pathway enrichment analysis and gene ontology (GO) were conducted using four tools. GSEA was conducted using GSEA software to evaluate potential pathways and molecular mechanisms. For GSVA, ESs were calculated with the *GSEA* R package. Predefined gene sets were downloaded from the Reactome and KEGG databases. The top enriched GO processes were identified via the Metascape web-based platform. Another network tool, IPA, identifies most significantly relevant pathways with the overall activation or inhibition states based on DEPs.

Immune infiltration analysis was performed via the *CIBERSORTx*, *ESTIMATE* and *single sample GSEA* (*ssGSEA*) algorithms. Several scores, including the TDS, ERK score and CYT, were calculated based on the normalized protein profile. TIS and IIS were calculated using the Z score-standardized ssGSEA matrix. *BRAF-RAS* score (BRS) was not calculated since no *RAS*-mutated samples were identified.

WGCNA was performed to cluster genes with high correlation and assess the correlation between protein modules and gene alterations. Proteins in the modules that were highly associated with gene mutations or fusions were selected. Least absolute shrinkage and selection operator (LASSO) regression, random forest (RF), and support vector machine-recursive feature elimination (SVM-RFE) were used to identify biomarker proteins based on selected proteins.

Several external datasets were used for the validation of our results (Sup Table S7). GSE151179 from the GEO database included 52 samples derived from radioiodine-refractory and radioiodine-avid PTC patients. This dataset was used to verify the results related to RAI refractoriness. The associations between gene alterations and protein expression were assessed with the THCA (thyroid cancer) program

605 from TCGA and four additional datasets from the GEO database. The predictive performance of the
606 genes in the external datasets was evaluated with ROC curves.
607 A detailed description of the bioinformatic analysis and references to methodological articles can be
608 found in the supplementary methods.
609

610 4.6 Statistical analysis

611 Quantitative variables were displayed as the means ± standard deviations or medians ± quartiles,
612 whereas qualitative variables were presented as numbers and ratios. The significant differences in the
613 quantitative variables were determined via two-tailed independent *t* test or Mann-Whitney *U* test, as
614 appropriate. The *chi*-square test was used to evaluate the differences in the distributions of qualitative
615 variables. When multiple groups were present, the *z* test with a Bonferroni correction was used to assess
616 the intergroup difference in every group. Statistical analysis was conducted using SPSS software (IBM,
617 US, v26).
618

Abbreviations

ANOVA, analysis of variance; ATA, American Thyroid Association; AUC, area under the curve (of ROC); *BRAF*, B-Raf proto-oncogene, serine/threonine kinase; BRS, *BRAF-RAS* score; CC, correlation coefficient; COSMIC, the Catalogue Of Somatic Mutations In Cancer; CT, computed tomography; CYT, immune cytolytic activity score; DEG, differentially expressed gene; DEP, differentially expressed protein; DIA, data-independent acquisition; DTC, differentiated thyroid cancer; ECP, eosinophil cationic protein; ES, enrichment score; ERK, extracellular signal-regulated kinase; ETE, extrathyroidal extension; FC, fold change; FFPE, formalin-fixed paraffin-embedded; F-18-FDG, 2-[^{18}F]fluoro-2-deoxy-D-glucose; GEO, the Gene Expression Omnibus database; GO, gene ontology; GS, gene significance; GSEA, gene set enrichment analysis; GSVA, gene set variation analysis; HR, hazard ratio; ID, positive RAI uptake and disease persistence; Id, RAI uptake positive and disease remission; iDF, negative RAI uptake at the first time of RAI treatment with disease persistence; iDG, RAI uptake lost gradually after previous RAI treatments with disease persistence; IIS, immune infiltration score; IP, inositol phosphate; IPA, ingenuine pathway analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; KNN, K-nearest neighbor; LASSO, least absolute shrinkage and selection operator; LC, liquid chromatography; LNM, lymph node metastatic lesion; MAPK, mitogen-activated protein kinase; MET, MET proto-oncogene; MKI, multi-kinase inhibitor; MM, module membership; MS, mass spectrometry; MRI, magnetic resonance imaging; NIS, sodium iodide symporter; PCA, principal component analysis; PCT, pressure cycling technology; PET, positron emission tomography; RAI, radioactive iodine; RAIR, RAI-refractoriness; RF, random forest; RFS, recurrence-free survival; RMA, robust multiarray average; ROC, receiver operating characteristic; RR-PTC, radioactive iodine refractory papillary thyroid cancer; SPECT, single photon emission computerized tomography; ssGSEA, single sample GSEA; SVM-RFE, support vector machine-recursive feature elimination; TAK1, transforming growth factor-β activated kinase 1; TCGA, The Cancer Genome Atlas Program; *TERT*, telomerase reverse transcriptase; *TERTp*, *TERT* promoter; TDS, thyroid differentiation score; TIS, T cell infiltration score; TNF, tumor necrosis factor; TNGS, targeted next-generation sequencing; TOM, topological overlap matrix; WBS, whole-body scanning; WGCNA, weighted correlation network analysis.

Disclosure

Ethics approval statement

The Institutional Ethical Committee of the hospital reviewed and approved the study design (No. WDRY2021-K032). The requirement for obtaining informed consent from the involved patients was waived due to the retrospective nature of the study. The study was conducted in accordance with the Declaration of Helsinki.

Acknowledgement

We appreciate the great support by four expert pathologists, Dr. Xiaokang Ke, Dr. Jiacai Ren, Dr. Xiaoyan Wu and Dr. Feng Guan, and the secretary of the Pathology Department, Mrs. Lingli Xia, at Renmin Hospital of Wuhan university. We give special thanks to Dr. Jun Liang in the Department of Nuclear Medicine at Renmin Hospital of Wuhan University and Dr. Jie Tan in the Department of Breast the Thyroid Surgery at Wuhan Union Hospital for their contribution and guidance in patient selection and grouping. We appreciate the great support from Prof. Tiannan Guo in School of Medicine at Westlake University for providing the platform of proteomics. We thank Prof. Katherine Hoadley at the University of North Carolina at Chapel Hill for her kind help in building a profile-based scoring system. We thank Dr. Zhou Liu from the Breast Tumor Center of Sun Yat-Sen Memorial Hospital and Dr. Nancy Li from the Reactome HelpDesk for their timely replies on two independent technical problems. We pay special thanks to authors who provided the external expression profiles for validation in the present study.

Funding

This research was supported by grants from the Interdisciplinary Innovative Talents Foundation from Renmin Hospital of Wuhan University (JCRCFZ-2022-015), the Fundamental Research Funds for the Central Universities (2042019kf0229), the Natural Science Foundation of Hubei Province, China (2023AFB701), and the Thyroid Research Project for Young and Middle-aged Doctors from Bethune Charitable Foundation (JKM2022-B12). All these funds were given to Prof. Chuang Chen to cover the costs during sample collection, TNGS, proteomics and other costs related to this academic research.

Conflicts of interest

All authors declare no competing interests.

Data and code availability

The proteomic data have been deposited in the iProX database with the project ID IPX0009103001. Calculation files and additional data are available in the Mendeley database (DOI: 10.17632/yfpfvktrxn). No custom code was used in the current study.

Author contribution

Conceptualization: HQ Liu (supporting) and C Chen (lead).
Data curation: HQ Liu, JX Wang, Y Zhou, DG Kong, PP Hu, L Li and ZL Xu (all equal).
Formal analysis: HQ Liu (lead), JX Wang (supporting), Y Zhou (supporting) and YT Sun (supporting).
Funding acquisition: C Chen (lead).
Investigation: HQ Liu (supporting), JX Wang (lead), DG Kong (supporting), D Yang (supporting) and ZL Xu (supporting).
Methodology: HQ Liu (supporting), Y Zhou (lead), and YT Sun (supporting).
Project administration: HQ Liu, YT Sun and C Chen (all equal).
Resources: YT Sun and C Chen (all equal).
Software: Y Zhou (supporting), PP Hu (supporting), L Li (supporting) and YT Sun (lead).
Supervision: YT Sun and C Chen (all equal).
Validation: D Yang (lead).
Visualization: HQ Liu (supporting) and D Yang (lead).
Writing - original draft: HQ Liu (lead), JX Wang (supporting), Y Zhou (supporting) and DG Kong (supporting).
Writing – review & editing: D Yang (supporting), YT Sun (lead) and C Chen (lead).

References

1 Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA: a cancer journal for clinicians* **74**, 12-49, doi:10.3322/caac.21820 (2024).

2 Li, M. *et al.* Changing incidence and projections of thyroid cancer in mainland China, 1983-2032: evidence from Cancer Incidence in Five Continents. *Cancer Causes & Control* **32**, 1095-1105, doi:10.1007/s10552-021-01458-6 (2021).

3 Pereira, M., Williams, V. L., Johnson, J. & Valderrabano, P. Thyroid Cancer Incidence Trends in the United States: Association with Changes in Professional Guideline Recommendations. *Thyroid* **30**, 1132-1140, doi:10.1089/thy.2019.0415 (2020).

4 Miranda-Filho, A. *et al.* Thyroid cancer incidence trends by histology in 25 countries: a population-based study. *Lancet Diabetes & Endocrinology* **9**, 225-234, doi:10.1016/s2213-8587(21)00027-9 (2021).

5 Lim, H., Devesa, S. S., Sosa, J. A., Check, D. & Kitahara, C. M. Trends in Thyroid Cancer Incidence and Mortality in the United States, 1974-2013. *Jama-Journal of the American Medical Association* **317**, 1338-1348, doi:10.1001/jama.2017.2719 (2017).

6 van Velsen, E. F. S. *et al.* Evaluating Disease-specific Survival Prediction of Risk Stratification and TNM Systems in Differentiated Thyroid Cancer. *Journal of Clinical Endocrinology & Metabolism* **108**, E267-E274, doi:10.1210/clinem/dgac721 (2023).

7 Manzardo, O. A. *et al.* TNM 8th edition in thyroid cancer staging: is there an improvement in predicting recurrence? *Endocrine-Related Cancer* **27**, 325-336, doi:10.1530/erc-19-0412 (2020).

8 Haugen, B. R. *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* **26**, 1-133, doi:10.1089/thy.2015.0020 (2016).

9 van Velsen, E. F. S. *et al.* Evaluating the 2015 American Thyroid Association Risk Stratification System in High-Risk Papillary and Follicular Thyroid Cancer Patients. *Thyroid* **29**, 1073-1079, doi:10.1089/thy.2019.0053 (2019).

10 Nunes, K. S. *et al.* Risk factors associated with disease-specific mortality in papillary thyroid cancer patients with distant metastases. *Endocrine* **75**, 814-822, doi:10.1007/s12020-021-02901-z (2022).

11 Durante, C. *et al.* Long-term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: Benefits and limits of radioiodine therapy. *Journal of Clinical Endocrinology & Metabolism* **91**, 2892-2899, doi:10.1210/jc.2005-2838 (2006).

12 Ciarallo, A. & Rivera, J. Radioactive Iodine Therapy in Differentiated Thyroid Cancer: 2020 Update. *American Journal of Roentgenology* **215**, 285-291, doi:10.2214/ajr.19.22626 (2020).

13 Sakulpisuti, C., Charoenphun, P. & Chamroonrat, W. Positron Emission Tomography Radiopharmaceuticals in Differentiated Thyroid Cancer. *Molecules* **27**, 4936, doi:10.3390/molecules27154936 (2022).

14 Durante, C. *et al.* BRAF mutations in papillary thyroid carcinomas inhibit genes involved in iodine metabolism. *Journal of Clinical Endocrinology & Metabolism* **92**, 2840-2843, doi:10.1210/jc.2006-2707 (2007).

15 Riesco-Eizaguirre, G. *et al.* The BRAF V600E Oncogene Induces Transforming Growth Factor β Secretion Leading to Sodium Iodide Symporter Repression and Increased Malignancy in Thyroid Cancer. *Cancer Research* **69**, 8317-8325, doi:10.1158/0008-5472.Can-09-1248 (2009).

16 Liu, J. *et al.* The Genetic Duet of BRAF V600E and TERT Promoter Mutations Robustly Predicts Loss of Radioiodine Avidity in Recurrent Papillary Thyroid Cancer. *Journal of Nuclear Medicine* **61**, 177-182, doi:10.2967/jnumed.119.227652 (2020).

17 Cao, J. *et al.* The genetic duet of BRAF V600E and TERT promoter mutations predicts the poor curative effect of radioiodine therapy in papillary thyroid cancer. *European Journal of Nuclear Medicine and Molecular Imaging* **49**, 3470-3481, doi:10.1007/s00259-022-05820-x (2022).

18 Groener, J. B. *et al.* BRAF V600E and Retinoic Acid in Radioiodine-Refractory Papillary Thyroid Cancer. *Hormone and Metabolic Research* **51**, 69-75, doi:10.1055/a-0765-9078 (2019).

19 Yang, X. *et al.* TERT Promoter Mutation Predicts Radioiodine-Refractory Character in Distant Metastatic Differentiated Thyroid Cancer. *Journal of Nuclear Medicine* **58**, 258-265, doi:10.2967/jnumed.116.180240 (2017).

20 Feng, F. *et al.* A nonpump function of sodium iodide symporter in thyroid cancer via cross-talk with PTEN signaling. *Cancer Research* **78**, 6121-6133, doi:10.1158/0008-5472.Can-18-1954 (2018).

21 Dohán, O., Baloch, Z., Bánrévi, Z., Livolsi, V. & Carrasco, N. Predominant intracellular overexpression of the Na+/I- symporter (NIS) in a large sampling of thyroid cancer cases. *Journal of Clinical Endocrinology & Metabolism* **86**, 2697-2700, doi:10.1210/jc.86.6.2697 (2001).

22 Marsee, D. K. *et al.* Inhibition of heat shock protein 90, a novel RET/PTC1-associated protein, increases radioiodide accumulation in thyroid cells. *Journal of Biological Chemistry* **279**, 43990-43997, doi:10.1074/jbc.M407503200 (2004).

23 Paron, I. *et al.* Nuclear localization of Galectin-3 in transformed thyroid cells: a role in transcriptional regulation. *Biochemical and Biophysical Research Communications* **302**, 545-553, doi:10.1016/s0006-291x(03)00151-7 (2003).

24 Huang, D. *et al.* Proteotypic Differences of Follicular-Patterned Thyroid Neoplasms. *Frontiers in endocrinology* **13**, 854611, doi:10.3389/fendo.2022.854611 (2022).

25 Qu, N. *et al.* Integrated proteogenomic and metabolomic characterization of papillary thyroid cancer with different recurrence risks. *Nature Communications* **15**, 3175, doi:10.1038/s41467-024-47581-1 (2024).

26 Sun, Y. *et al.* Stratification of follicular thyroid tumours using data-independent acquisition proteomics and a comprehensive thyroid tissue spectral library. *Molecular oncology* **16**, 1611-1624, doi:10.1002/1878-0261.13198 (2022).

27 Shi, X. *et al.* Integrated proteogenomic characterization of medullary thyroid carcinoma. *Cell discovery* **8**, 120, doi:10.1038/s41421-022-00479-y (2022).

28 Colombo, C. *et al.* The molecular and gene/miRNA expression profiles of radioiodine resistant papillary thyroid cancer. *Journal of Experimental & Clinical Cancer Research* **39**, 245, doi:10.1186/s13046-020-01757-x (2020).

29 Xing, M. Molecular pathogenesis and mechanisms of thyroid cancer. *Nature Reviews Cancer* **13**, 184-199, doi:10.1038/nrc3431 (2013).

30 Arturi, F. *et al.* Iodide symporter gene expression in human thyroid tumors. *Journal of Clinical Endocrinology & Metabolism* **83**, 2493-2496, doi:10.1210/jc.83.7.2493 (1998).

31 Anekpuritanang, T. *et al.* The Association Between Radioiodine Refractory in Papillary Thyroid Carcinoma, Sodium/Iodide Symporter Expression, and BRAFV600E Mutation. *Oncotargets and Therapy* **14**, 3959-3969, doi:10.2147/ott.S308910 (2021).

32 Zhang, R. *et al.* Association between sodium iodide symporter and differentiated Thyroid cancer: a meta-analysis of 9 studies. *International Journal of Clinical and Experimental Medicine* **8**, 17986-17994 (2015).

33 Braga-Basaria, M. & Ringel, M. D. Beyond radioiodine: A review of potential new therapeutic approaches for thyroid cancer. *Journal of Clinical Endocrinology & Metabolism* **88**, 1947-1960, doi:10.1210/jc.2002-021863 (2003).

34 Agrawal, N. *et al.* Integrated Genomic Characterization of Papillary Thyroid Carcinoma. *Cell* **159**, 676-690, doi:10.1016/j.cell.2014.09.050 (2014).

35 Xing, M. *et al.* Association Between BRAF V600E Mutation and Recurrence of Papillary Thyroid Cancer. *Journal of Clinical Oncology* **33**, 42-U79, doi:10.1200/jco.2014.56.8253 (2015).

36 Luo, Y. *et al.* Clinical, Pathological, and Molecular Characteristics Correlating to the Occurrence of Radioiodine Refractory Differentiated Thyroid Carcinoma: A Systematic Review and Meta-Analysis. *Frontiers in Oncology* **10**, 549882, doi:10.3389/fonc.2020.549882 (2020).

37 Zhang, Z., Liu, D., Murugan, A. K., Liu, Z. & Xing, M. Histone deacetylation of NIS promoter underlies BRAF V600E-promoted NIS silencing in thyroid cancer. *Endocrine-Related Cancer* **21**, 161-173, doi:10.1530/erc-13-0399 (2014).

38 Mu, Z. *et al.* Characterizing Genetic Alterations Related to Radioiodine Avidity in Metastatic Thyroid Cancer. *Journal of Clinical Endocrinology & Metabolism* **109**, 1231-1240, doi:10.1210/clinem/dgad697 (2024).

39 Collina, F. *et al.* AXL Is a Novel Predictive Factor and Therapeutic Target for Radioactive Iodine Refractory Thyroid Cancer. *Cancers* **11**, 785, doi:10.3390/cancers11060785 (2019).

40 Meng, Z. *et al.* TERT promoter mutation in primary papillary thyroid carcinoma lesions predicts absent or lower (131)i uptake in metastases. *Iubmb Life* **71**, 1030-1040, doi:10.1002/iub.2056 (2019).

41 Jung, C. K. *et al.* Risk Stratification Using a Novel Genetic Classifier Including PLEKHS1 Promoter Mutations for Differentiated Thyroid Cancer with Distant Metastasis. *Thyroid* **30**, 1589-1600, doi:10.1089/thy.2019.0459 (2020).

42 Liu, R. *et al.* Mortality Risk Stratification by Combining BRAF V600E and TERT Promoter Mutations in Papillary Thyroid Cancer Genetic Duet of BRAF and TERT Promoter Mutations in Thyroid Cancer Mortality. *Jama Oncology* **3**, 202-208, doi:10.1001/jamaoncol.2016.3288 (2017).

43 Yu, P. *et al.* TERT accelerates BRAF mutant-induced thyroid cancer dedifferentiation and progression by regulating ribosome biogenesis. *Science Advances* **9**, eadg7125, doi:10.1126/sciadv.adg7125 (2023).

44 Ju, G. *et al.* Fusion Oncogenes in Patients With Locally Advanced or Distant Metastatic Differentiated Thyroid Cancer. *Journal of Clinical Endocrinology & Metabolism* **109**, 505-515, doi:10.1210/clinem/dgad500 (2024).

45 Arai, S. *et al.* In vitro and in vivo anti-tumor activity of alectinib in tumor cells with NCOA4-RET. *Oncotarget* **8**, 73766-73773, doi:10.18632/oncotarget.17900 (2017).

46 Paratala, B. S. *et al.* RET rearrangements are actionable alterations in breast cancer. *Nature*

Communications **9**, 4821, doi:10.1038/s41467-018-07341-4 (2018).

47 Groussin, L. *et al.* Selpercatinib-Enhanced Radioiodine Uptake in RET-Rearranged Thyroid Cancer. *Thyroid* **31**, 1603-1604, doi:10.1089/thy.2021.0144 (2021).

48 Lee, Y. A. *et al.* NTRK and RET fusion-directed therapy in pediatric thyroid cancer yields a tumor response and radioiodine uptake. *Journal of Clinical Investigation* **131**, e144847, doi:10.1172/jci144847 (2021).

49 Nikitski, A. V. *et al.* Can TP53-mutant follicular adenoma be a precursor of anaplastic thyroid carcinoma? *Endocrine-Related Cancer* **28**, 621-630, doi:10.1530/erc-21-0095 (2021).

50 Fu, H. *et al.* MAPK Inhibitors Enhance HDAC Inhibitor-Induced Redifferentiation in Papillary Thyroid Cancer Cells Harboring BRAFV600E : An In Vitro. *Molecular Therapy-Oncolytics* **12**, 235-245, doi:10.1016/j.omto.2019.01.007 (2019).

51 Rothenberg, S. M., McFadden, D. G., Palmer, E. L., Daniels, G. H. & Wirth, L. J. Redifferentiation of Iodine-Refractory BRAF V600E-Mutant Metastatic Papillary Thyroid Cancer with Dabrafenib. *Clinical Cancer Research* **21**, 1028-1035, doi:10.1158/1078-0432.Ccr-14-2915 (2015).

52 Zhang, C. *et al.* Inhibitory effects of siRNA targeting silencing TAK1 gene on proliferation and migration of thyroid cancer cells and p38 MAPK signaling pathway. *Journal of Jilin Univeristy. Medicine edition* **47**, 110-117 (2021).

53 Gheorghe, D. C., Stanciu, M. M., Zamfirescu, A. & Stanciu, A. E. TNF-α May Exert Different Antitumor Effects in Response to Radioactive Iodine Therapy in Papillary Thyroid Cancer with/without Autoimmune Thyroiditis. *Cancers* **13**, 3609, doi:10.3390/cancers13143609 (2021).

54 Yen, M.-C. *et al.* S100B expression in breast cancer as a predictive marker for cancer metastasis. *International Journal of Oncology* **52**, 433-440, doi:10.3892/ijo.2017.4226 (2018).

55 Wang, X. *et al.* A Functional Insertion/Deletion Polymorphism in the Promoter Region of the NFKB1 Gene Increases the Risk of Papillary Thyroid Carcinoma. *Genetic Testing and Molecular Biomarkers* **19**, 167-171, doi:10.1089/gtmb.2014.0271 (2015).

56 Sponziello, M. *et al.* Fibronectin-1 expression is increased in aggressive thyroid cancer and favors the migration and invasion of cancer cells. *Molecular and Cellular Endocrinology* **431**, 123-132, doi:10.1016/j.mce.2016.05.007 (2016).

57 Chen, X., Zheng, J., Zhang, A. & You, Z. Increased FN1 expression correlates to poor prognosis, BRAF mutation and immune infiltrates in papillary thyroid cancer. *Chinese Journal of Immunology* **37**, 1841-1847 (2021).

58 Sabra, M. M. *et al.* Clinical Outcomes and Molecular Profile of Differentiated Thyroid Cancers With Radioiodine-Avid Distant Metastases. *Journal of Clinical Endocrinology & Metabolism* **98**, E829-E836, doi:10.1210/jc.2012-3933 (2013).

59 Shobab, L. *et al.* Clinical, Pathological, and Molecular Profiling of Radioactive Iodine Refractory Differentiated Thyroid Cancer. *Thyroid* **29**, 1262-1268, doi:10.1089/thy.2019.0075 (2019).

60 Song, H.-J., Xue, Y.-L., Qiu, Z.-L. & Luo, Q.-Y. Comparative serum proteomic analysis identified afamin as a downregulated protein in papillary thyroid carcinoma patients with non-131I-avid lung metastases. *Nuclear Medicine Communications* **34**, 1196-1203, doi:10.1097/mnm.0000000000000001 (2013).

61 Li, Y. *et al.* Proteomic analysis of radioiodine-refractory differentiated thyroid cancer identifies CHI3L1 upregulation in association with dysfunction of the sodium-iodine symporter.

Oncology Letters **25**, 36, doi:10.3892/ol.2022.13622 (2023).

62 World Med, A. World Medical Association Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. *Jama-Journal of the American Medical Association* **310**, 2191-2194, doi:10.1001/jama.2013.281053 (2013).

63 Perrier, N. D., Brierley, J. D. & Tuttle, R. M. Differentiated and Anaplastic Thyroid Carcinoma: Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA: a cancer journal for clinicians* **68**, 56-63, doi:10.3322/caac.21439 (2018).

64 Cai, X. *et al.* High-throughput proteomic sample preparation using pressure cycling technology. *Nature Protocols* **17**, 2307-2325, doi:10.1038/s41596-022-00727-1 (2022).

65 Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nature Methods* **17**, 41-+, doi:10.1038/s41592-019-0638-x (2020).

66 Xiong, Y. *et al.* Application of biomarkers in the diagnosis of uncertain samples of core needle biopsy of thyroid nodules. *Virchows Archiv* **479**, 961-974, doi:10.1007/s00428-021-03161-y (2021).

67 Landrum, M. J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Research* **44**, D862-D868, doi:10.1093/nar/gkv1222 (2016).

68 Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* **47**, D941-D947, doi:10.1093/nar/gky1015 (2019).

897 **Tables, figures and supplementary materials**

898 Tables

899 Table 1. Demographic and clinicopathologic characteristic of included patients

Clinicopathological features	All patients (48)	RR-PTC (37)	Control (11)	*P* value
Age at diagnosis				
Years	38.4 ± 12.69	39.2 ± 13.29	35.6 ± 10.49	0.421
≥55 y	5 (10.4)	5 (13.5)	0 (0)	0.576
Female sex	28 (58.3)	20 (54.1)	8 (72.7)	0.319
Tumor diagnosis				0.302
Incidental	22 (45.8)	15 (40.5)	7 (63.6)	
Symptomatic	26 (54.2)	22 (59.5)	4 (36.4)	
Smoking or alcohol	9 (18.8)	8 (21.6)	1 (9.1)	0.662
BMI (kg/m^2)	24.1 ± 5.36	24.6 ± 5.75	22.5 ± 3.47	0.261
Hashimoto's thyroiditis	1 (2.1)	1 (2.7)	0 (0)	1.000
Nodular goiter	10 (20.8)	7 (18.9)	3 (27.3)	0.675
Histological subtypes [a]				
cPTC	39 (81.3)	29 (78.4)	10 (90.9)	0.662
fvPTC	13 (27.1)	12 (32.4)	1 (9.1)	0.246
tcv/svPTC	6 (12.5)	5 (13.5)	1 (9.1)	1.000
Tumor size (cm)	2.1 ± 1.40	2.1 ± 1.38	2.1 ± 1.40	0.996
Multifocality	36 (75.0)	28 (75.7)	8 (72.7)	1.000
Bilateral lesion	27 (56.3)	20 (54.1)	7 (63.6)	0.733
ETE	37 (77.1)	31 (83.8)	6 (54.5)	0.095
TNM				
T1	7 (14.6)	3 (8.1)	4 (36.4)	0.039 *
T2-4	41 (85.4)	34 (91.9)	7 (63.6)	
N0	3 (6.3)	3 (8.1)	0 (0)	1.000
N1	45 (93.8)	34 (91.9)	11 (100)	
M0	26 (54.2)	16 (43.2)	10 (90.9)	0.006 **
M1	22 (45.8)	21 (56.8)	1 (9.1)	
AJCC stage				0.001 **
I	22 (45.8)	12 (32.4)	10 (90.9)	
II/III/IV	26 (54.2)	25 (67.6)	1 (9.1)	
ATA risk stratification				1.000
Low	0 (0)	0 (0)	0 (0)	
Intermediate/high	48 (100)	37 (100)	11 (100)	
Cumulative dose of RAI (mCi)	370 (220, 570)	430 (220, 430)	220 (220, 250)	0.002 **
Follow up				0.313

Alive	42 (87.5)	31 (83.8)	11 (100)
Dead	6 (12.5)	6 (16.2)	0 (0)

Note: Quantitative variables were displayed as mean ± standard deviation, whereas qualitative variables were displayed as number (ratio). The cumulative dose of RAI was presented as medians (quartiles) due to inconformity to Gaussian distribution.

[a] Multiple tumors in one patient may exhibit different histopathological types. The diameters of the largest tumor were measured when multiple lesions existed.

* $P < 0.05$, ** $P < 0.01$.

Abbreviations: AJCC, American Joint Committee on Cancer; ATA, American Thyroid Association; BMI, body mass index; cPTC, classic papillary thyroid cancer; ETE, extrathyroidal extension; fvPTC, follicular variant of papillary thyroid cancer; RAI, radioactive iodine; svPTC, solid variant of papillary thyroid cancer; tcvPTC, tall cell variant of papillary thyroid cancer.

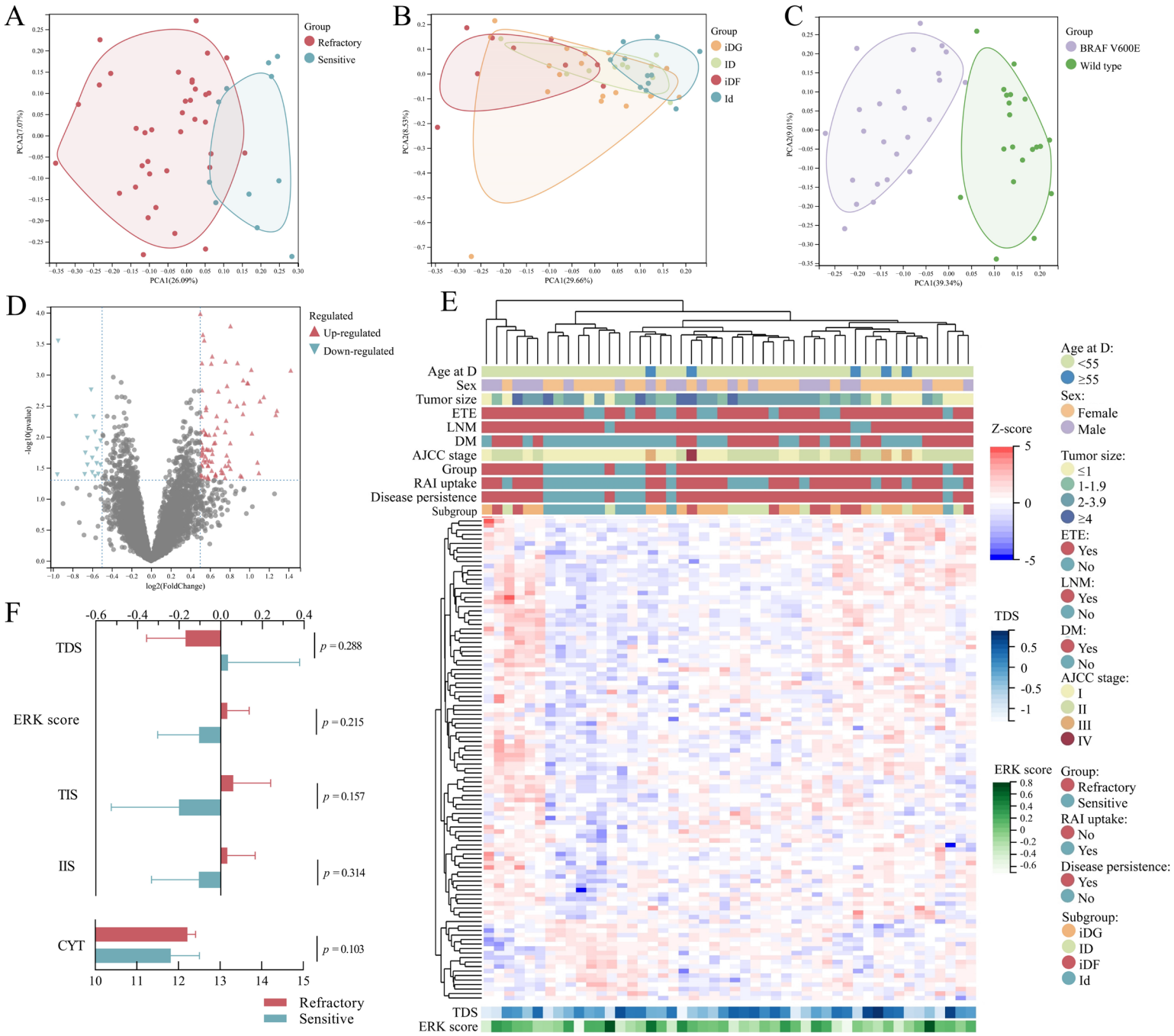
Table 2. The comparison of gene mutations and fusions in RAI-refractory and RAI-sensitive PTC

Comparison	BRAF	RET	TERTp	TP53	Fusion [a]	RET fusion	NCOA4-RET
RR-PTC *vs.* Control							
RR-PTC (35)	19 (54.3)	0 (0)	6 (17.1)	1 (2.9)	10 (28.6)	8 (22.9)	7 (20.0)
Control (10)	6 (60.0)	1 (10.0)	0 (0)	1 (10.0)	4 (40.0)	2 (20.0)	1 (10.0)
P value	1.000	0.222	0.312	0.399	0.700	1.000	0.661
Subgroups							
ID (9)	**1 (11.1) ***	0 (0)	1 (11.1)	1 (11.1)	4 (44.4)	4 (44.4)	**4 (44.4) ***
iDF (9)	6 (66.7)	0 (0)	**3 (33.3) ***	0 (0)	1 (11.1)	1 (11.1)	0 (0)
iDG (17)	12 (70.6)	0 (0)	2 (11.8)	0 (0)	5 (29.4)	3 (17.6)	3 (17.6)
Id (10)	6 (60.0)	1 (10.0)	0 (0)	1 (10.0)	4 (40.0)	2 (20.0)	1 (10.0)
P value	0.026 *	0.311	0.193	0.411	0.421	0.327	0.081
I-131 uptake							
Negative (26)	18 (69.2)	0 (0)	5 (19.2)	0 (0)	6 (23.1)	4 (15.4)	3 (11.5)
Positive (19)	7 (36.8)	1 (5.3)	1 (5.3)	2 (10.5)	8 (42.1)	6 (31.6)	5 (26.3)
P value	0.039 *	0.422	0.222	0.173	0.206	0.281	0.253

* $P < 0.05$. Bold font indicated statistical significance revealed by *z* test.

[a] Detected gene fusions included *ETV6-NTRK3* in three samples, one *STRN-ALK*, two *CCDC6-RET* and eight *NCOA4-RET*. The latter two fusions were collectively called *RET* fusion.

Figures


A
Group
Refractory
Sensitive
PCA1(26.09%)
PCA2(7.07%)
B
Group
iDG
ID
iDF
Id
PCA1(29.66%)
PCA2(8.53%)
C
Group
BRAF V600E
Wild type
PCA1(39.34%)
PCA2(9.01%)
D
Regulated
Up-regulated
Down-regulated
log2(FoldChange)
-log10(pvalue)
E
Age at D
Sex
Tumor size
ETE
LNM
DM
AJCC stage
Group
RAI uptake
Disease persistence
Subgroup
Z-score
TDS
ERK score
Age at D:
<55
≥55
Sex:
Female
Male
Tumor size:
≤1
1-1.9
2-3.9
≥4
ETE:
Yes
No
LNM:
Yes
No
DM:
Yes
No
AJCC stage:
I
II
III
IV
Group:
Refractory
Sensitive
RAI uptake:
No
Yes
Disease persistence:
Yes
No
Subgroup:
iDG
ID
iDF
Id
F
TDS
ERK score
TIS
IIS
CYT
p = 0.288
p = 0.215
p = 0.157
p = 0.314
p = 0.103
Refractory
Sensitive

917

Figure 1. Proteomic overview of RAI-refractory and RAI-sensitive PTC samples. (A-C) Dimension reduction visualization of the analyzed PTC samples. The samples were labeled according to their response to I-131 treatment (refractory or sensitive), subgroup (iDF, iDG, ID or Id) and *BRAF* mutation status. Three samples whose quality control for sequencing failed were not included in Figure C. (D) Differential expression analysis of samples in the RR-PTC and control groups. Unadjusted P values < 0.05 and absolute FC > 1.414 were defined as thresholds for DEPs. (E) Overview of the proteomic profiles. Z scores of the DEPs were calculated. (F) Correlation of scores and groups. Error bar denoted 95% CI. The TDS was lower in the RR-PTC group (-0.17 ± 0.562 *vs.* 0.03 ± 0.512, P = 0.288). Conversely, the ERK score was greater in the RR-PTC group (0.03 ± 0.317 *vs.* -0.10 ± 0.297, P = 0.215). The means and SDs are presented with boxes and error bars.

Abbreviations: age at D, age at diagnosis; DM, distant metastasis; ETE, extrathyroidal extension; LNM, lymph node metastasis.

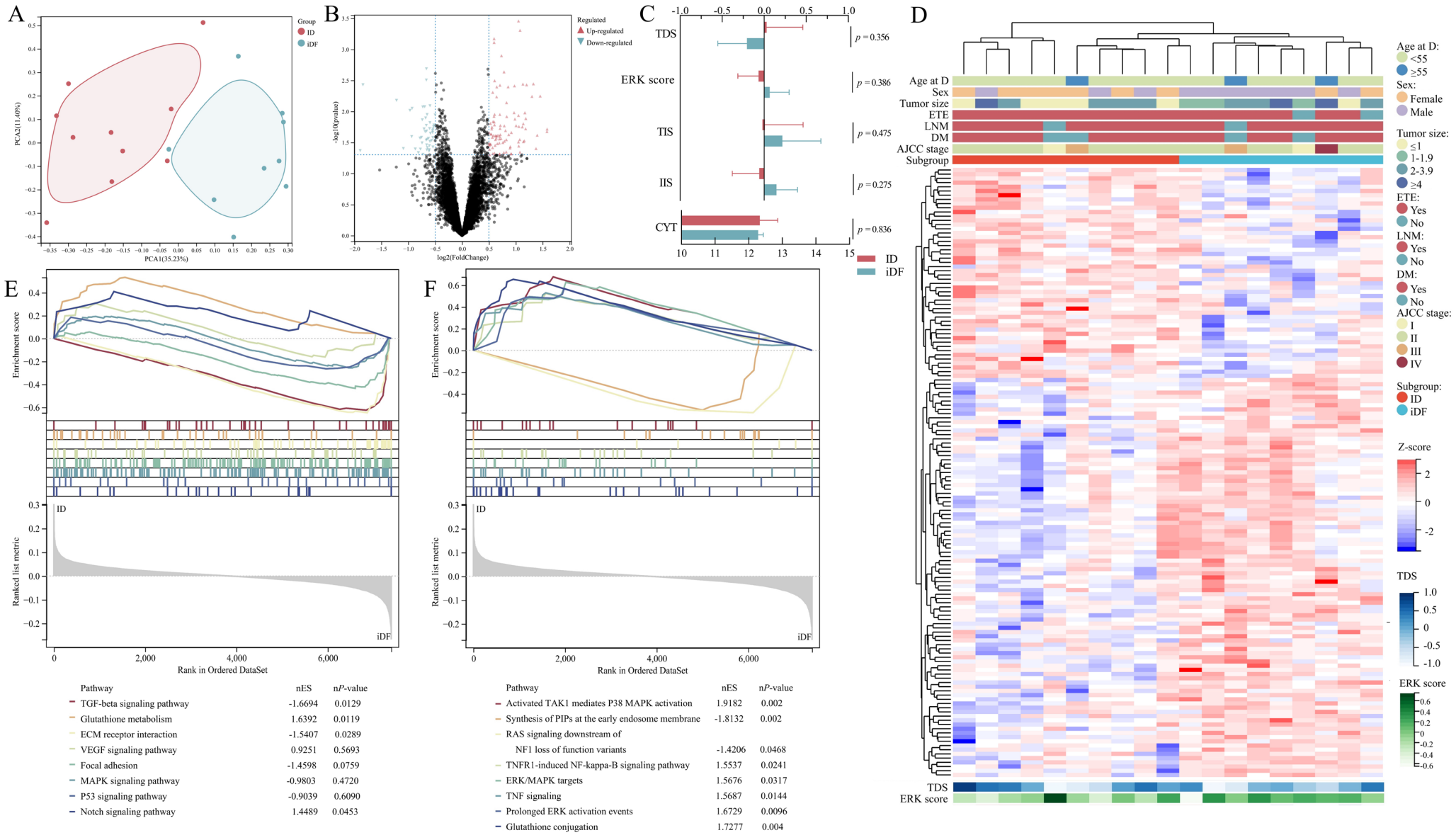


Figure 2. Differential expression analysis of the ID and iDF groups. (A) PCA plot of the two groups. (B) Differential expression analysis of samples in the two groups. Unadjusted *P* values < 0.05 and absolute FC > 1.414 were defined as thresholds for DEPs. (C) Correlations of scores and groups. Error bar denoted 95% CI. The TDS was greater in the ID group (0.03 ± 0.599 *vs.* -

928 0.20 ± 0.454, $P = 0.356$), while the ERK score was greater in the iDF group (-0.07 ± 0.345 *vs.* 0.07 ± 0.300, $P = 0.386$). (D) Expression profiles of 145 DEPs. (E&F) GSEA based on the KEGG
929 and Reactome databases. Sixteen pathways are presented in the plot.
930

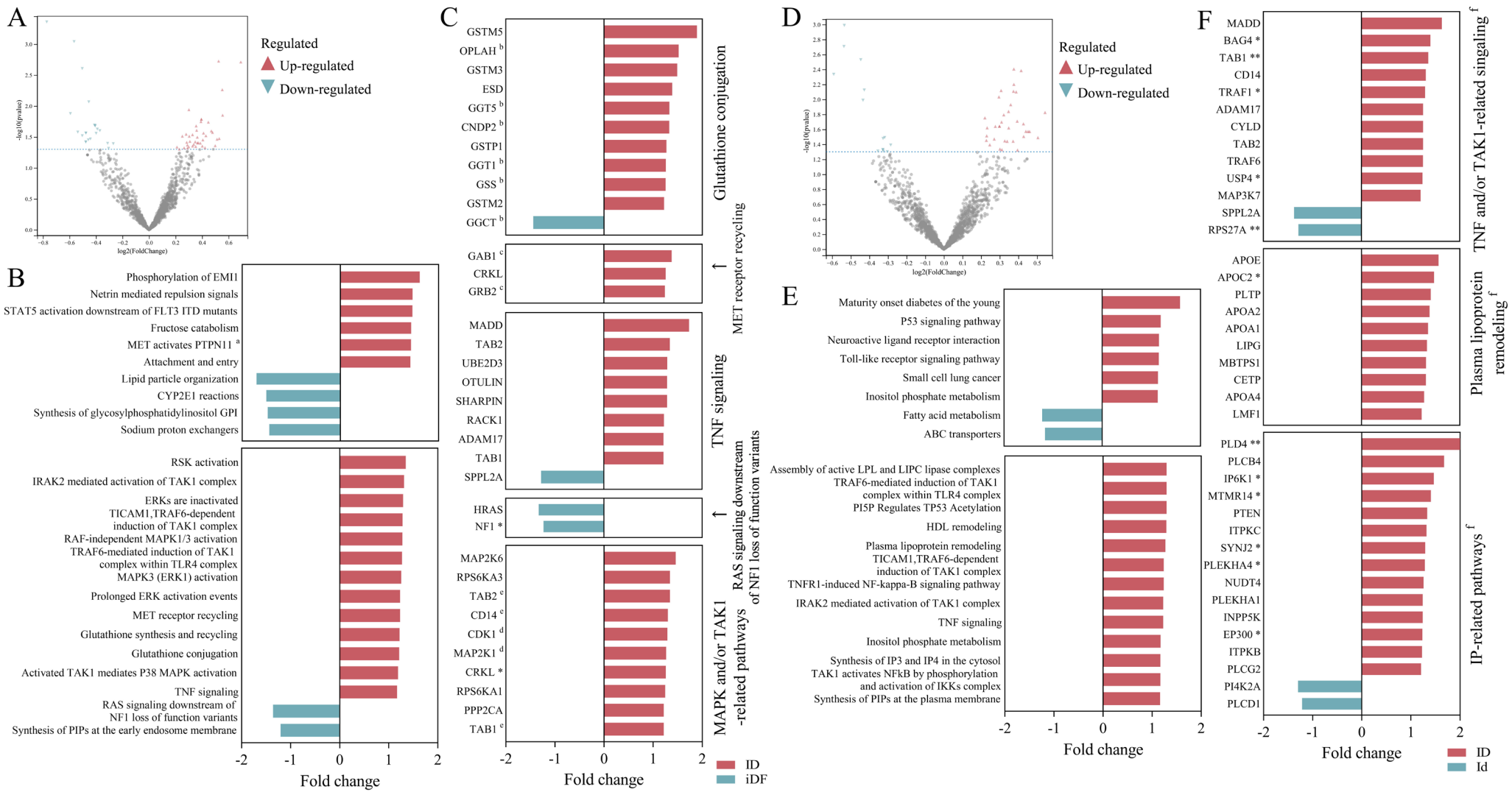


Figure 3. Expression of selected genes according to the enriched pathways identified in the ID *vs.* iDF and ID *vs.* Id comparisons. (A&D) Dysregulated pathways identified in the ID *vs.* iDF and ID *vs.* Id comparisons. The *P* value threshold was 0.05. No threshold was set for fold change. (B) The ten most significantly enriched pathways and the thirteen selected pathways. (C) The selected proteins from the thirteen selected pathways. The threshold of absolute fold change was > 1.2. (E) The eight significantly enriched pathways and thirteen selected pathways. (F) The selected proteins from the thirteen selected pathways. The TNF signaling pathway and the TAK1-related pathways were combined in a single plot since the proteins strongly overlapped. * $P < 0.05$, ** $P < 0.01$.

937 [a] The pathway "MET activates PTPN1" was also selected for subsequent identification of proteins.

938 [b, c] Involved in the selected pathways "glutathione synthesis and recycling" and "MET activates PTPN11", respectively.

939 [d] Both CDK1 and MAP2K1 participate in the pathways "MAPK3 (ERK1) activation" and "RAF-independent MAPK1/3 activation", and MAP2K1 is also a participant in "prolonged ERK
940 activation events".

941 [e] TAB2 and TAB1 are involved in both MAPK and TAK1-associated pathways, while CD14 participates in TAK1-associated pathways.

942 [f] Molecules in these figures might participate in several pathways. The associations of molecules and pathways are provided in Sup table S5.

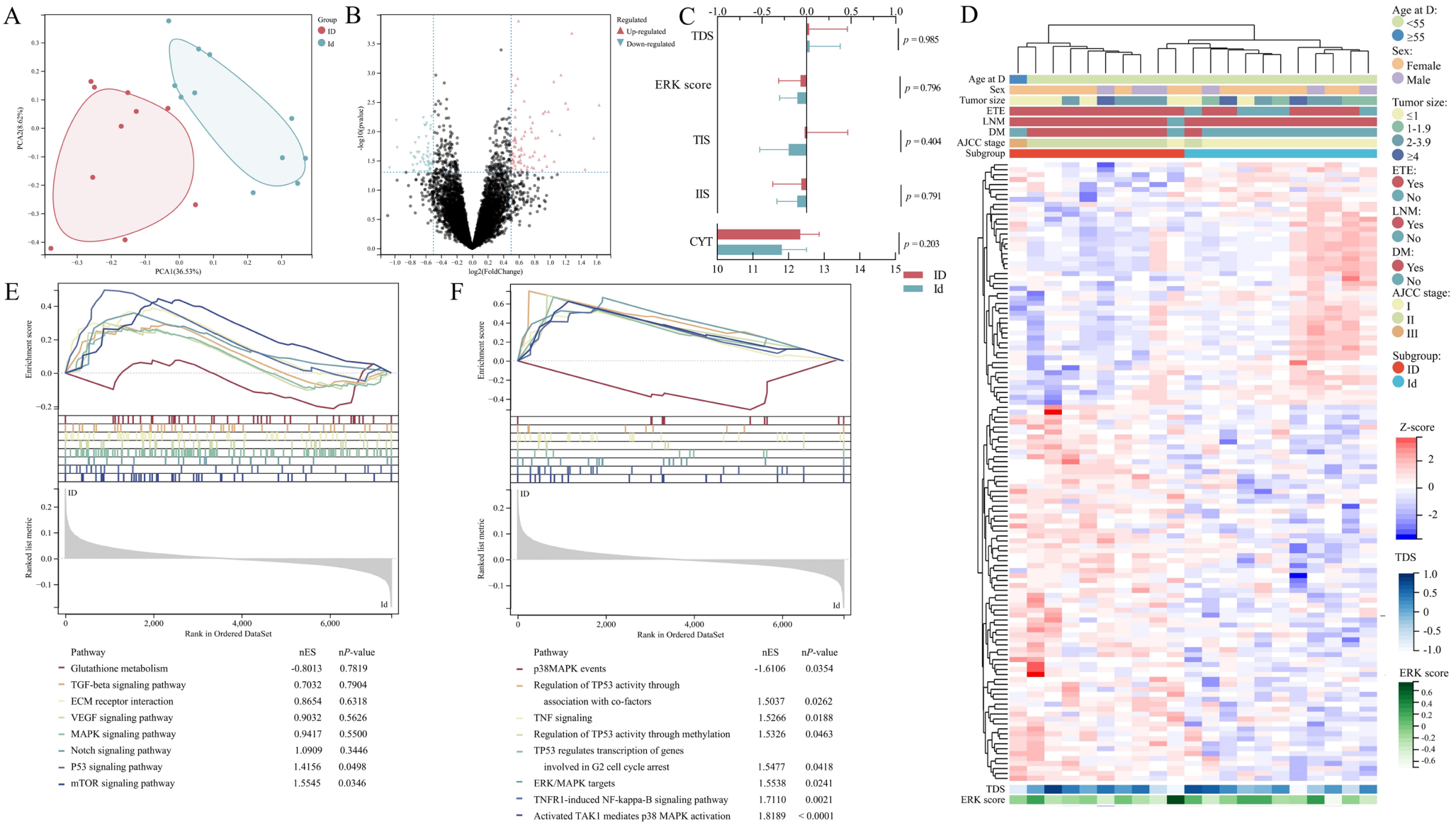
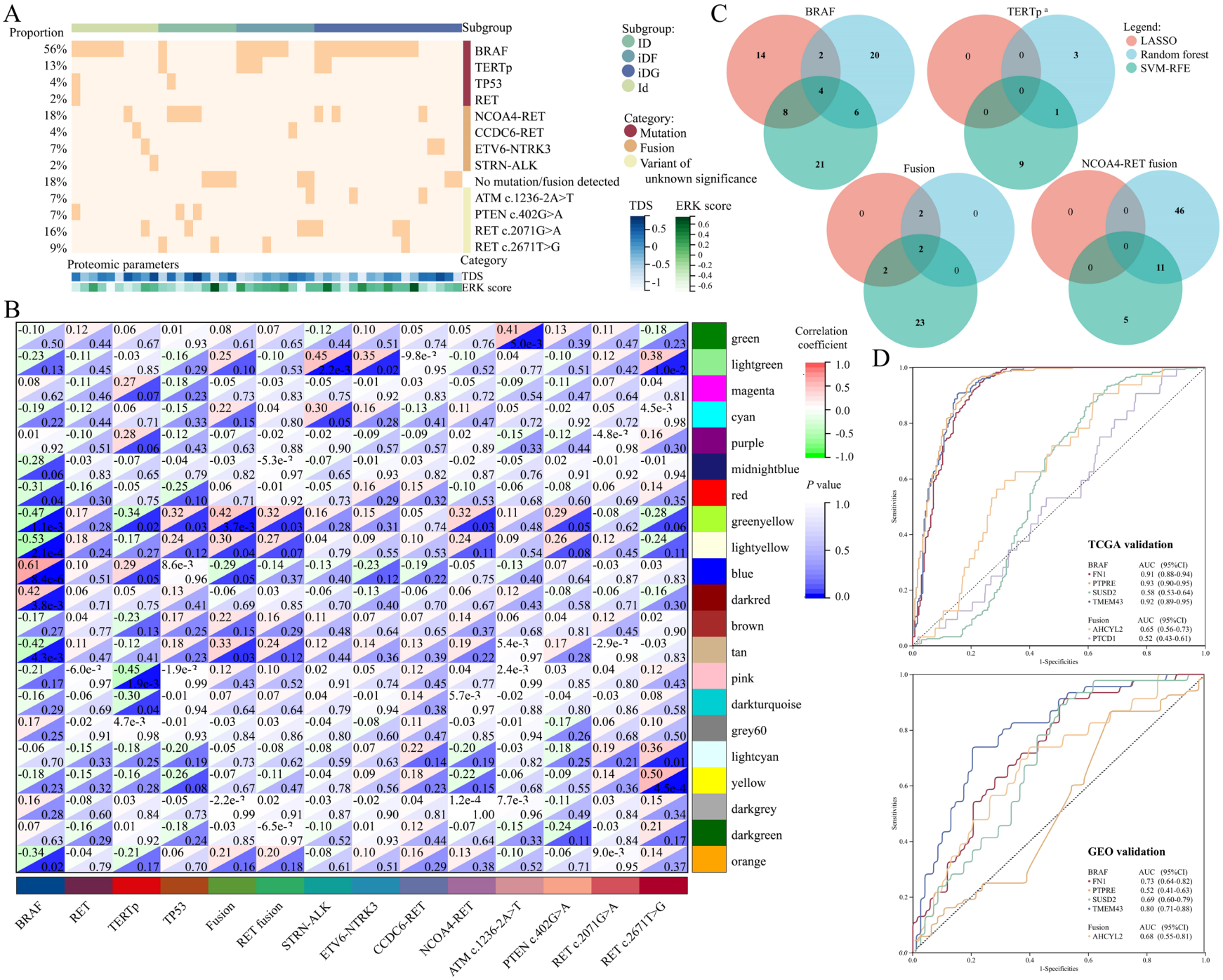


Figure 4. Differential expression analysis of the ID and Id groups. (A) PCA plot of the two groups. (B) Differential expression analysis of samples in the two groups. Unadjusted *P* values < 0.05 and absolute FC > 1.414 were defined as thresholds for DEPs. (C) Correlations of scores and groups. Error bar denoted 95% CI. No significant difference was found in the TDS (0.03 ± 0.599

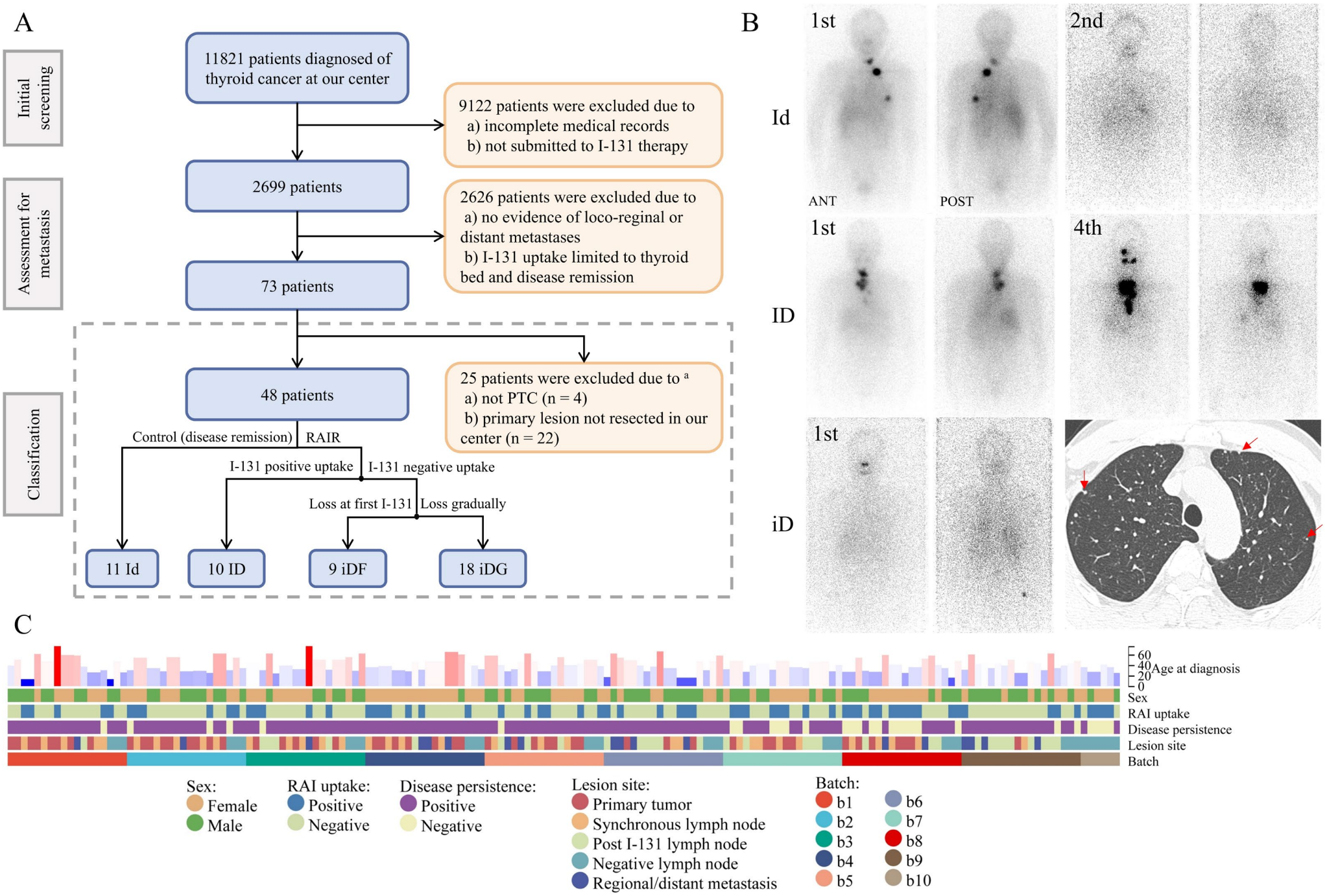
946 *vs.* 0.03 ± 0.512, $P = 0.985$) or ERK score (-0.07 ± 0.345 *vs.* -0.10 ± 0.297, $P = 0.796$). (D) Expression profiles of 125 DEPs. (E&F) GSEA based on the KEGG and Reactome databases. Sixteen
947 pathways are presented in the plot.


A
Proportion
Subgroup
56%
13%
4%
2%
18%
4%
7%
2%
18%
7%
7%
16%
9%
BRAF
TERTp
TP53
RET
NCOA4-RET
CCDC6-RET
ETV6-NTRK3
STRN-ALK
No mutation/fusion detected
ATM c.1236-2A>T
PTEN c.402G>A
RET c.2071G>A
RET c.2671T>G
Proteomic parameters
Category
TDS
ERK score
Subgroup:
ID
iDF
iDG
Id
Category:
Mutation
Fusion
Variant of unknown significance
TDS
ERK score
B
green
lightgreen
magenta
cyan
purple
midnightblue
red
greenyellow
lightyellow
blue
darkred
brown
tan
pink
darkturquoise
grey60
lightcyan
yellow
darkgrey
darkgreen
orange
Correlation coefficient
P value
BRAF
RET
TERTp
TP53
Fusion
RET fusion
STRN-ALK
ETV6-NTRK3
CCDC6-RET
NCOA4-RET
ATM c.1236-2A>T
PTEN c.402G>A
RET c.2071G>A
RET c.2671T>G
C
BRAF
TERTp a
Fusion
NCOA4-RET fusion
Legend:
LASSO
Random forest
SVM-RFE
D
TCGA validation
BRAF AUC (95%CI)
FN1 0.91 (0.88-0.94)
PTPRE 0.93 (0.90-0.95)
SUSD2 0.58 (0.53-0.64)
TMEM43 0.92 (0.89-0.95)
Fusion AUC (95%CI)
AHCYL2 0.65 (0.56-0.73)
PTCD1 0.52 (0.43-0.61)
GEO validation
BRAF AUC (95%CI)
FN1 0.73 (0.64-0.82)
PTPRE 0.52 (0.41-0.63)
SUSD2 0.69 (0.60-0.79)
TMEM43 0.80 (0.71-0.88)
Fusion AUC (95%CI)
AHCYL2 0.68 (0.55-0.81)
Sensitivities
1-Specificities

948

Figure 5. Gene variants and weighted gene co-expression network analysis. (A) Gene variance landscape. (B) Correlations of gene modules and gene phenotypes. The cells were divided into upper left (correlation coefficient) and lower right (*P* value) parts. (C) The intersection of LASSO-, RF- and SVM-RFE-identified biomarkers for *BRAF*, and *TERTp* mutations, gene fusions and *NCOA4-RET* fusion. (D) Validation of six genes in the TCGA-THCA cohort and merged GEO cohort. PTCD1 was not detected in the merged GEO dataset.

[a] The low number of *TERTp*-positive samples (n = 6) prevented model construction and error analysis for *TERTp* mutations via the SVM algorithm. The first ten proteins with the top average rank were processed for biomarker identification.


A
Initial screening
Assessment for metastasis
Classification
11821 patients diagnosed of thyroid cancer at our center
9122 patients were excluded due to
a) incomplete medical records
b) not submitted to I-131 therapy
2699 patients
2626 patients were excluded due to
a) no evidence of loco-reginal or distant metastases
b) I-131 uptake limited to thyroid bed and disease remission
73 patients
48 patients
25 patients were excluded due to a
a) not PTC (n = 4)
b) primary lesion not resected in our center (n = 22)
Control (disease remission)
RAIR
I-131 positive uptake
I-131 negative uptake
Loss at first I-131
Loss gradually
11 Id
10 ID
9 iDF
18 iDG
B
Id
ID
iD
1st
2nd
4th
ANT
POST
C
Age at diagnosis
60
40
20
0
Sex
RAI uptake
Disease persistence
Lesion site
Batch
Sex:
Female
Male
RAI uptake:
Positive
Negative
Disease persistence:
Positive
Negative
Lesion site:
Primary tumor
Synchronous lymph node
Post I-131 lymph node
Negative lymph node
Regional/distant metastasis
Batch:
b1
b2
b3
b4
b5
b6
b7
b8
b9
b10

954

Figure 6. (A) Flow chart of patient and sample selection. Strict criteria were followed to identify patients with RAI-refractory or RAI-sensitive thyroid cancers. Individuals with insufficient radiological evidence were also excluded from the second round of selection. A total of 73 patients (in the dashed box) were revealed to have regional or distant metastatic lesions with solid evidence. The FFPE tissues were collected for subsequent proteomic analysis and/or targeted deep sequencing. Notably, 25 patients were excluded from this study due to a lack of primary lesions or not having PTC. (B) Grouping of included patients. Three samples are shown here for the Id, ID and iD groups. The first patient belonged to the Id group. Her or his lung metastases, which were confirmed by I-131 WBS and chest CT, were eliminated by two I-131 treatments. The second patient was included in the ID group. She or he received total thyroidectomy but her or his cervical regional metastatic lesion progressed during the four rounds of I-131 treatment. The first and second I-131 WBSs of the third patient revealed no positive focus. However, PET/CT during the same period revealed several suspicious pulmonary nodules (red arrow), which were confirmed to be metastases by subsequent surgery and histopathology. (C) Information on the collected FFPE samples. The protein profiles of 168 samples were evaluated in 10 batches. The sample types included primary lesions, synchronous and post I-131 cervical LNM, negative cervical lymph node and regional or distant metastatic lesions. Notably, the regional metastatic lesions were tumors near the primary site, such as the esophagus, trachea and cervical skin. LNM was not included in regional metastatic lesions. In addition, forty-eight primary tumors were examined via targeted deep sequencing.

[a] The primary lesion of one patient was neither resected in our center nor examined to be PTC.

Supplementary materials

Supplementary methods
Supplementary tables 1-7
Supplementary figure 1-12
References