Translational Oncology 43 (2024) 101889

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

# Multi-omics analysis-based macrophage differentiation-associated papillary thyroid cancer patient classifier

Hanlin Sun [a,1], Zhengyan Chang [b,1], Hongqiang Li [a], Yifeng Tang [a], Yihao Liu [a], Lixue Qiao [a], Guicheng Feng [a], Runzhi Huang [c,***], Dongyan Han [b,**], De-tao Yin [a,d,e,*]

[a] Department of Thyroid Surgery, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, PR China
[b] Department of Pathology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai, PR China
[c] Department of Burn Surgery, the First Affiliated Hospital of Naval Medical University, Shanghai 200433, PR China
[d] Engineering Research Center of Multidisciplinary Diagnosis and Treatment of Thyroid Cancer of Henan Province, Zhengzhou 450052, Henan, PR China
[e] Key Medicine Laboratory of Thyroid Cancer of Henan Province, Zhengzhou 450052, Henan, PR China

ARTICLE INFO

*Keywords:*
Multi-omics
Single-cell RNA sequencing
Papillary thyroid carcinoma
Tumor immune microenvironment
Macrophage differentiation

ABSTRACT

*Background:* The reclassification of Papillary Thyroid Carcinoma (PTC) is an area of research that warrants attention. The connection between thyroid cancer, inflammation, and immune responses necessitates considering the mechanisms of differential prognosis of thyroid tumors from an immunological perspective. Given the high adaptability of macrophages to environmental stimuli, focusing on the differentiation characteristics of macrophages might offer a novel approach to address the issues related to PTC subtyping.
*Methods:* Single-cell RNA sequencing data of medullary cells infiltrated by papillary thyroid carcinoma obtained from public databases was subjected to dimensionality reduction clustering analysis. The RunUMAP and FindAllMarkers functions were utilized to identify the gene expression matrix of different clusters. Cell differentiation trajectory analysis was conducted using the Monocle R package. A complex regulatory network for the classification of Immune status and Macrophage differentiation-associated Papillary Thyroid Cancer Classification (IMPTCC) was constructed through quantitative multi-omics analysis. Immunohistochemistry (IHC) staining was utilized for pathological histology validation.
*Results:* Through the integration of single-cell RNA and bulk sequencing data combined with multi-omics analysis, we identified crucial transcription factors, immune cells/immune functions, and signaling pathways. Based on this, regulatory networks for three IMPTCC clusters were established.
*Conclusion:* Based on the co-expression network analysis results, we identified three subtypes of IMPTCC: Immune-Suppressive Macrophage differentiation-associated Papillary Thyroid Carcinoma Classification (ISMPTCC), Immune-Neutral Macrophage differentiation-associated Papillary Thyroid Carcinoma Classification (INMPTCC), and Immune-Activated Macrophage differentiation-associated Papillary Thyroid Carcinoma Classification (IAMPTCC). Each subtype exhibits distinct metabolic, immune, and regulatory characteristics corresponding to different states of macrophage differentiation.

## Introduction

Recent worldwide epidemiological data indicate a notable rise in the prevalence of thyroid cancer over the past few decades [1]. Thyroid cancer (TC) is the most widespread endocrine tumor and ranks as the eighth most common cancer globally [2,3]. Although thyroid cancer has long been recognized as a less malignant malignancy with a better prognosis, its high incidence still leads to a significant number of

* Corresponding author at: Department of Thyroid Surgery, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, PR China.
** Corresponding author at: Department of Pathology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, 301 Yanchang Road, Shanghai 200072, PR China.
*** Corresponding author at: Department of Burn Surgery, the First Affiliated Hospital of Naval Medical University, 168 Changhai Road, Shanghai 200433, PR China.
*E-mail addresses:* runzhihuang2022@163.com (R. Huang), 1500057@tongji.edu.cn (D. Han), detaoyin@zzu.edu.cn (D.-t. Yin).
[1] Co-first authorship: Hanlin Sun and Zhengyan Chang have contributed equally to this work.

https://doi.org/10.1016/j.tranon.2024.101889
Received 11 October 2023; Received in revised form 2 January 2024; Accepted 21 January 2024

patients experiencing poorer outcomes. Different TC types originating from the same differentiated cells show varying prognoses. Except for medullary thyroid carcinoma (MTC), which originates from parafollicular cells, most TCs develop from follicular epithelial cells. The majority of cases of thyroid carcinoma consist of papillary thyroid cancer (PTC) and follicular thyroid cancer (FTC). Despite being less common, anaplastic thyroid cancer (ATC) plays a significant role in the mortality associated with thyroid cancer [4,5]. However, despite the generally favorable prognosis associated with PTC, a subset of patients still experience adverse outcomes, including recurrence and metastasis [6].

PTC, FTC, and ATC exhibit distinct molecular mechanisms despite their shared origin. These differences are also believed to potentially demonstrate a progressive nature [7]. The variations in prognosis observed within the same pathological type of TC can be attributed, on the one hand, to clinical characteristics and, on the other hand, to molecular mechanisms that influence tumor prognosis. However, taking PTC as an example, there is still no consensus on whether molecular pathogenic mechanisms contribute to prognostic differences, as evidenced by the ongoing debate regarding the impact of *BRAF* gene mutations on PTC prognosis [8]. Therefore, we speculate that progressive molecular mechanisms may also be present within the same pathological types of tumors.

At this juncture, considering the prognosis differences of tumors from an immunological perspective offers a new approach to unraveling this complexity. Ever since Rudolf Virchow proposed the connection between inflammation, immune response, and cancer in 1863, understanding how the immune system intervenes in tumor progression has been a crucial direction in exploring the pathogenic mechanisms of cancer [9]. This direction was further explored when Hanahan and Weinberg defined the characteristics of tumors [10]. The notion of the tumor microenvironment was presented, emphasizing its crucial function in the onset and progression of tumors. The normal tissue microenvironment can suppress malignant tumors, while certain pathological tissue features are crucial for tumor progression [11,12]. It is of paramount importance to simultaneously acquire a comprehensive understanding of the tumor microenvironment and delineate the immunological aspects of the tumorigenic process to assist in halting tumor progression via immune surveillance. Given the numerous connections between thyroid cancer, inflammation, and immune response [13], it becomes increasingly crucial to focus on changes in the immune environment within the realm of TC. Tumor-infiltrating myeloid cells (TIMs) [14] are crucial regulators in tumor progression, operating within the intricate cellular networks of the tumor microenvironment. They have significant involvement in modulating tumor inflammation and angiogenesis, as highlighted by studies [15,16]. Macrophages, as important constituents of TIMs, are involved in various aspects of tumor immunity. Furthermore, macrophages have become a central focus in tumor immunology research due to their remarkable adaptability to environmental cues and their complex roles influenced by both temporal factors and the surrounding microenvironment [17,18].

The objective of this research was to uncover the complete picture of tumor pathogenicity by utilizing single-cell RNA sequencing (scRNA-Seq) technology. Our focus was on examining the immune microenvironment to elucidate the prognostic differences in PTC. Based on the differentiation of tumor-infiltrating macrophages and their immunological status within the tumor microenvironment, we established a clinical classification of PTC, termed Immune status and macrophage differentiation-associated papillary thyroid cancer classification (IMPTCC). Based on the immune progression status, we classified IMPTCC into three subtypes: Immune-suppressive macrophage differentiation-associated papillary thyroid carcinoma classification (ISMPTCC), Immune-neutral macrophage differentiation-associated papillary thyroid carcinoma classification (INMPTCC), and Immune-activated macrophage differentiation-associated papillary thyroid carcinoma classification (IAMPTCC). Each subtype exhibits distinct pathogenic mechanisms based on differences in the immune environment and molecular regulatory networks. Moreover, these subtypes provide novel insights into the potential pathogenic mechanisms of thyroid cancer at the levels of the immune microenvironment and metabolism.

## Methods

### *Data acquisition*

The scRNA-seq information of thyroid macrophages was acquired from the GSE154763 database, comprising gene expressions of 5312 cells from 10 patients who were diagnosed with papillary thyroid carcinoma. The Cancer Genome Atlas (TCGA) database (https://cancerdata.nci.nih.gov/) provided data for 512 tumor tissues and 5 metastases, along with genomic profiles, RNA sequencing profiles, Reverse Phase Protein Array (RPPA) profiles, Transposase-Accessible Chromatin sequencing (ATAC-seq) profiles, population statistics, and clinical information including overall survival rates. From 2016 to 2018, the Tenth People's Hospital affiliated with Tongji University recruited a total of 72 patients who were diagnosed with PTC for histopathological validation. Furthermore, all primary data were obtained from openly accessible repositories, and no further ethical clearance was necessary. Table S1 lists all the datasets, code packages, and software information used in this study. Supplementary materials containing clinical data confirmed by pathological histology and the original code of this study have been uploaded.

### *Batch correction and data quality control*

Our scRNA-seq data were obtained from batch-corrected PTC sample datasets publicly released by Professor Zemin Zhang's research [19]. These datasets were processed using Scanorama [20], an algorithm designed to identify and merge shared cell types across multiple datasets, thus eliminating batch effects within the datasets. The sequencing data from the GEO database, consisting of scRNA-seq information, were imported into the R software package. After applying the CreateSeuratObject function, the data were converted into a Seurat object, followed by quality control analysis using the Seurat R package. For further analysis, only cells that expressed more than 100,000 transcripts and had less than 10 % mitochondrial genes were considered. Further analysis also included genes that were expressed in a minimum of 3 cells. After applying the NormalizeData function, the data were normalized and the resulting values saved in the pbmc dataset. The function FindVariableFeatures was utilized for the identification of genes that exhibit high variability. The vst technique was utilized to identify the top 2000 genes with significant variation among all the genes, which were then employed for subsequent analysis to reduce dimensionality. Subsequently, the ScaleData function was used to linearly transform the data, normalizing the expression values of each gene to a mean of 0 and a variance of 1, thereby eliminating scale differences between different genes.

### *Data dimension reduction*

The data from individual cells were subjected to principal component analysis (PCA) using the RunPCA function. To find an optimal number of principal components for further analysis, the ElbowPlot function was used and detected a notable change in direction at the 10th PC. Subsequently, the Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) nonlinear dimensionality reduction technique was employed using the RunUMAP function to visually represent the complex structure of the data and the outcomes of clustering. UMAP was performed using the initial 20 principal components as input. The clustering process used the FindNeighbors function to perform clustering, while the FindClusters function was employed to

group cells at a clustering resolution of 0.5. The differential gene expression analysis involved the utilization of the FindAllMarkers function to identify markers specific to each cell type. To select genes showing differential expression, certain criteria were considered: a positive correlation, expression in a minimum of 25 % of cells, and a log2 fold change threshold of at least 0.25.

*Cell cycle analysis*

Cell cycle scoring of the scRNA-seq data was conducted using the Seurat package, employing well-established gene sets associated with the G2/M and S phases. The CellCycleScoring function within Seurat was utilized to convert gene expression values into G2M and S phase scores. Additionally, the VlnPlot function was used to generate violin plots of the G2M and S phase scores, facilitating visualization of score distribution across distinct cell populations.

*Cell subpopulation annotation*

The annotation of cells within each UMAP cluster was achieved using established cell surface markers, derived from the differential expression genes in each cluster and the CellMarker database [21]. The identification of marker genes for each cell type was conducted using the FindAllMarkers function of the Seurat package. The visualization of the expression trends of these markers across various cell subpopulations was achieved through the generation of heatmaps. The normalized data were represented through feature plots or violin plots. To evaluate the spatiotemporal distribution of distinct subpopulations within each cell type, individual cell types were isolated and organized as Seurat objects for subgroup analysis. The RunUMAP function, in conjunction with the FindAllMarkers function and Wilcoxon rank-sum test, was employed to detect differentially expressed genes (DEGs) among subgroups of the 13 cell types. Ultimately, the variations in spatial distribution and expression traits of the 13 cell subpopulations were depicted through cell feature plots, spatial feature plots, and heatmaps.

*Macrophage subtyping and differentiation trajectory analysis*

To explore the distribution of macrophage subtypes at different differentiation stages in thyroid cancer, we first employed the RunUMAP and FindAllMarkers functions from the Seurat package, along with the Wilcoxon rank-sum test, to construct an independent Seurat object for all macrophages across the samples. Macrophage subtyping was achieved using the Seurat package's FindClusters function, following PCA dimensionality reduction. We employed the first 20 principal components to discern co-expression patterns among cells. Subsequently, cells were clustered at a resolution of 0.1, leading to the identification of distinct macrophage subtypes. These subtypes were visualized and differentiated using UMAP plots, reflecting their unique expression profiles. The identification of differentially expressed genes (DEGs) within macrophage subtypes 1–5 was performed using the FindAllMarkers function. This function identifies genes significantly upregulated within each subgroup compared to all other cells, based on predefined thresholds of minimum percentage expression (min. pct = 0.25) and log fold change (logfc. threshold = 0.25). This analysis entails comparing each macrophage subgroup against the entirety of other cells in the dataset. We then applied additional filtering criteria, selecting genes with an average log2 fold change (avg_log2FC) greater than 0.5 to pinpoint the most significant DEGs. For each macrophage subgroup, we extracted the top five genes exhibiting the highest avg_log2FC changes as characteristic DEGs, providing a detailed insight into the unique expression profiles of each subgroup. Subsequently, we performed a subpopulation analysis of the data. Next, we utilized the Monocle package in R (version: 2.18.0) to analyze the differentiation trajectory of macrophages [22]. We employed the newCellDataSet function to categorize the UMI expression matrix, sample metadata, and gene expression information of all macrophages based on their differentiation trajectory. Unsupervised ordering was applied to all cells. The estimateSizeFactors function was used for data normalization. The detectGenes function was employed to filter out genes with low expression levels. Cell numbers at each time point were calculated to determine the developmental or differentiation starting point. After performing gene filtering and identifying differentially expressed genes, a total of 1000 genes were obtained as the foundation for cell ordering. Subsequently, the DDRTree method was used to reduce the dimensionality of cell clusters. The main chart was integrated into the high-dimensional scRNA-seq dataset, revealing the correlation between the gene expression space with many dimensions and a space with fewer dimensions. The differentiation trajectory between cells was built by projecting the expression data onto the low-dimensional space through the construction of a Minimum Spanning Tree (MST) using DDRTree. MST was recursively calculated to assign pseudotime to cells for ordering. The plot_cell_trajectory method was employed to display the developmental path of the cells, illustrating their various stages of differentiation.

Furthermore, we applied the Branch expression analysis modeling (BEAM) approach to infer the cell differentiation trajectory, comparing the differences between branching points and terminal branches to hypothesize cell fate in differentiation. Finally, we computed the differentially expressed genes (DEGs) among macrophage subtypes in different differentiation stages (States) and defined them as macrophage differentiation-related genes (MDGs). We selected the MDGs based on a q-value $< 0.05$ derived from the DDRTree algorithm and used them to sort the differentiation trajectory.

*Consistency cluster analysis*

To develop the immune status in Macrophage-infiltrated Thyroid Cancer Classification (IMPTCC), additional screening of MDGs is necessary. This screening procedure involved the integration of bulk-RNA sequencing data sourced from the TCGA database. Through the correlation of gene expression data with prognostic information derived from clinical samples, we utilized the survminer and limma packages to calculate Cox proportional hazards models for each gene. Genes with KM p-value $< 0.05$ and Cox regression p-value $< 0.05$ were selected for inclusion in the IMPTCC. Integration of significantly expressed genes from Bulk-RNA sequencing data analysis and scRNA-seq data pseudotime analysis retained genes that satisfied the following four criteria: [1] MDGs with q-values $< 0.05$ obtained through DDRTree algorithm and genes used for differentiation trajectory order; [2] MDGs exhibiting significant (non-parametric test p-value $< 0.05$) correlation between gene expression and pseudotime of PTC samples; [3] Genes showing significant (non-parametric test p-value $< 0.05$) results in Kaplan-Meier survival analysis; [23] Genes demonstrating significant (non-parametric test p-value $< 0.05$) results in univariate Cox survival analysis. Consequently, the integration of cellular differentiation features and clinical prognostic features yielded Macrophage differentiation and prognosis-related genes (MD&PRGs).

Subsequently, we employed the packages igraph, psych, reshape2, and RColorBrewer to construct and visualize gene co-expression networks for 19 MD&PRGs. Initially, the correlation between gene expressions was computed, and significant gene pairs were selected based on a p-value threshold; the correlation coefficients and p-values stored in a matrix. Next, the igraph package was utilized to generate the co-expression network graph. Nodes were assigned colors based on the grouping of genes according to their cellular differentiation states and risk factors, while node sizes were allocated based on the results of Cox proportional hazards regression analysis. The relationships between genes were represented using line attributes, with line width indicating the magnitude of p-values, and line color indicating the direction of correlation (pink for positive correlation and blue for negative correlation).

By using the ConsensusClusterPlus package in the R programming

language, the consensus clustering analysis (CCA) algorithm was applied to cluster the observed samples based on the expression profiles of key genes and to integrate their intrinsic characteristics. Observational samples with similar features, as determined by their close Euclidean distances, were defined as belonging to the same cluster subtype. To enhance clustering stability and representative grouping, the optimal clustering model was selected by constraining the cumulative distribution function (CDF) growth rate, the proportions of samples within each group, and both intra-group and inter-group correlations [24].

*Principal component analysis and clinical correlation analysis*

To convert the clustering results into quantitative variables, we performed PCA analysis on the sample data using the prcomp and predict functions. The gene expression data of the samples were dimensionally reduced, and the scores of each sample on the first two principal components were computed, yielding PCA scores for each sample. Based on these scores, the samples were divided into high and low groups. The surv_cutpoint, survdiff, and survfit functions were applied to integrate the survival data of the samples for survival analysis, and the ggsurvplot and ggplot functions were used for visualization. The results were visualized using box plots, bar graphs, violin plots, and heatmaps ($P<0.05$ indicating statistical significance). To obtain more reliable results, we chose progression-free survival (PFS) as the clinical survival outcome endpoint for analysis. Furthermore, with the aid of PCA scores, we conducted multi-omics analysis at the genomic, transcriptomic, proteomic, and epigenomic levels for the 19 MD&PRGs.

*Survival analysis and prediction model construction and diagnosis*

Differential expression analysis was performed using primary tumor and metastasis RNA-seq data from the TCGA database. Survival-associated genes were identified through univariate Cox regression analysis, and the top 20 prognostic genes were determined using Lasso regression based on the survival-associated genes. A multifactorial prognostic model was established by identifying the differentiation cell fate genomic signature in THCA. Subsequently, we employed random number table allocation to divide all samples into a training set and a validation set in a 6:4 ratio, and the prior model was trained based on the training set. Finally, the hypothesis was validated using the cross-validation set to determine the optimal Receiver Operating Characteristic (ROC) area. Samples were scored based on the median risk score, and risk curves, scatter plots, and expression heatmaps were generated. The area under the ROC curve was used to evaluate the accuracy of the model. Univariate and multivariate Cox regression analyses were applied to evaluate whether the risk score, age, gender, and TNM staging were independent prognostic factors.

*Genomics correlation analysis*

To calculate the percentage of copy number gains and losses for important genes in the samples, we utilized Perl scripts and R code for copy number variation (CNV) analysis. Circular chromosome plots displaying the chromosomal location information of the key genes were generated using the RCircos package. To analyze the correlation between tumor mutation burden (TMB) and PCA scores, as well as microsatellite instability (MSI), and to evaluate the prognostic differences of the samples by integrating TMB and PCA scores, the ggpubr and reshape2 packages were utilized. The oncoplot package was utilized to visualize and analyze gene mutation data.

*Gene-set enrichment analysis*

In this study, we used gene set variation analysis (GSVA) [25], gene set enrichment analysis (GSEA) [26], and Overall Represent analysis (ORA) [27] methods to better explore gene expression differences and their underlying functional implications.

The gene sets utilized in this research were acquired from the Molecular Signatures Database (MSigDB) (version 7.1) [28]. To assess the variation in the activity of each sample within gene sets, we initially utilized GSVA to produce enrichment scores for hallmark gene sets for each sample using the ordered gene expression profiles. GSVA assesses the variation of biological processes without predefined phenotype information, thus utilizing all available gene expression information. The GSVA package was utilized to conduct GSVA analysis on the gene expression matrix, whereas the limma package was employed for differential analysis, filtering significantly differentially expressed pathways using a logFC cutoff of 0.0001 and adjPvalue cutoff of 0.05. The ggplot2 package was used to create stacked bar plots and volcano plots, while the pheatmap package was employed to visualize the GSVA analysis results using heatmap plots.

Furthermore, we utilized the ORA technique to investigate genes exhibiting notable variations in expression. To assess the degree of enrichment in a particular set of genes, the comparison was made based on the proportion of genes that showed significant differential expression within the said gene set. By utilizing the ORA analysis, we conducted a comparison between the 19 essential genes and the gene sets in the reference. We evaluated the enrichment of the important genes in nine sets of genes by conducting hypergeometric distribution testing. The gene sets were classified into nine functional categories according to various functional characteristics, which include groups of gene sets labeled as C1-C8 and the HALLMARK gene set group.

To further elucidate underlying biological phenomena and mitigate biases arising from threshold selection, we employed the GSEA method to identify gene sets with similar expression patterns and examine their enrichment between the high-risk and low-risk groups. Specifically, we conducted GSEA analyses on the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) gene set databases. Significantly enriched gene sets with p-values less than 0.05 were identified for each database. GSEA plots depicting the top five significantly enriched gene sets were generated separately for the high-risk and low-risk groups [29,30].

*Immune infiltration and immune checkpoint analysis*

We analyzed the correlation between the expression of the PD-L1 gene and PCA scores by utilizing data acquired from The Cancer Immunome Atlas (TCIA) [31]. Furthermore, we investigated the variances in the quantities of immune cells infiltrating tumors among PCA score groups categorized as high and low. This analysis involved assessing the levels of expression for two immune checkpoint proteins, namely CTLA-4 and PD-L1. The limma and ggpubr packages were utilized for the analysis of immune checkpoint data processing and visualization.

We employed single-sample gene set enrichment analysis (ssGSEA) [32] to assess the immune infiltration status and accurately evaluate the differences in immune expression among different subtypes and between primary and metastatic samples. By utilizing known immune-related gene sets, we calculated immune infiltration scores for each sample to evaluate the degree of immune infiltration. The immune infiltration status was visualized using the corrplot and ggpubr packages, allowing for intuitive observation and comparison of the differences in immune infiltration levels among different subtypes.

To investigate the composition of infiltrating immune cells in complex tissues, we employed a technique known as Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) [33]. This method relies on the characteristics of gene expression profiles to estimate the relative subsets of RNA transcripts. To determine the correlation between MDGs and immune cell infiltration in PTC tissues, we uploaded the microarray expression data and RNA-seq gene expression matrix data to the Cell-type Identification By Estimating Relative

Subsets Of RNA Transcripts (CIBERSORT) database to assess the immune infiltration status of the samples.

*Proteomic analysis*

Protein expression values were calculated using gene expression profile data and Reverse Phase Protein Array (RPPA) data. Furthermore, the correlation between the gene expression profile and protein expression values was computed.

*Regulatory network construction*

To further explore the potential regulatory mechanisms of various subtypes of IMPTCC, a standardized approach was used to analyze quantitative results obtained from MD&PRGs, transcription factors, immune cells/immune function, RPPA protein arrays, and signaling pathways. We performed co-expression analysis by utilizing Pearson/Spearman correlation analysis. The analysis was based on data normality and rank, with a threshold set at $R > 0.300$ and $P < 0.001$. With minimum screening criteria of $R > 0.300$ and $P < 0.001$, we constructed an intricate regulatory network focused on the MD&PRGs of every IMPTCC subtype. The regulatory network was visualized using the igraph package in the R programming language (http://igraph.sourceforge.net) [34].

*Targeted drug prediction analysis for each subtype*

The Connectivity Map (Cmap) is an algorithm that explores the statistical matching relationship between small molecule active drugs and target genes based on whole-genome transcriptional expression data from human cells treated with small molecule compounds. The Cmap algorithm incorporates over 7000 cell culture systems and 1309 small molecule active drugs, all of which are FDA-approved drugs, and establishes a drug-function gene-disease network. This network provides a statistical basis for identifying small molecule inhibitors targeting genes associated with therapeutically relevant diseases (https://portals.broadinstitute.org/cmap/) [35]. In this study, the Cmap algorithm was employed to explore small molecule drugs targeting the regulatory networks of different IMPTCC subtypes. Subsequently, a dot plot heatmap analysis was used to identify the top 10 drugs with the lowest p-values as targeted drugs for each IMPTCC subtype. Furthermore, by integrating the gene expression profiles with the subtype classification, the most effective active drugs were selected. Additionally, we linked pharmacological data with genomic information using information from eight common thyroid cancer cell lines and visualized the data as heatmaps and box plots. Moreover, GDSC genomic data were utilized for predictive modeling of pharmacological sensitivity [36].

*Chromatin accessibility assessment*

ATAC-seq [37] is a remarkable, versatile, uncomplicated, and potent approach for the comprehensive characterization of chromatin regions across the genome, in contrast to traditional methodologies such as functional assays or sequence conservation analyses. To establish the accessibility of crucial genes on chromatin, we procured ATAC-seq data pertaining to these genes.

*Histopathological validation*

Approval for histopathological validation using clinical samples was obtained from the Clinical Research Institution Review Committee of the Tenth People's Hospital, which is affiliated with Tongji University in Shanghai. Before participation, all patients provided informed written consent. Samples for additional validation were gathered from the Tenth People's Hospital, affiliated with Tongji University, from 2016 to 2018.

Samples of initial tumors collected from a group of 72 patients diagnosed with papillary thyroid carcinoma underwent fixation and were then embedded in formalin and paraffin. Subsequently, these specimens underwent standard procedures of sectioning, dehydration, antigen retrieval through rehydration, and blocking, and were sliced into sections measuring 4 μm in thickness. The sections underwent incubation with an *APOE* antibody (1:3000 proteintech) at 4 °C overnight. Afterward, polymer HRP was used to label all sections for 40 min and then counterstained with hematoxylin for 5 min at room temperature. The evaluation of all slides was conducted by two pathologists, whereby positive cytoplasmic staining in the cancer cells was defined. It is worth noting that both cytoplasmic and membranous *APOE* were stained utilizing the same antibody. To create a negative control, the primary antibody was replaced with a buffer solution. We utilized non-parametric tests and Spearman's correlation tests to assess the association between the *APOE* staining intensity scores and the clinical features of the individuals.

*Statistical analysis*

The R programming language was utilized for processing and analyzing all the data. The normality of quantitative variables was assessed using the Shapiro-Wilk test, while Levene's test was employed to examine the homogeneity of variances. Comparisons between the two groups were conducted using either the Mann-Whitney U test (also known as the Wilcoxon rank sum test) or the *t*-test, depending on the distribution of the data. The chi-squared test or Fisher's exact test was used for categorical variables, depending on the anticipated frequencies. The level of statistical significance was determined to be $p<0.05$.

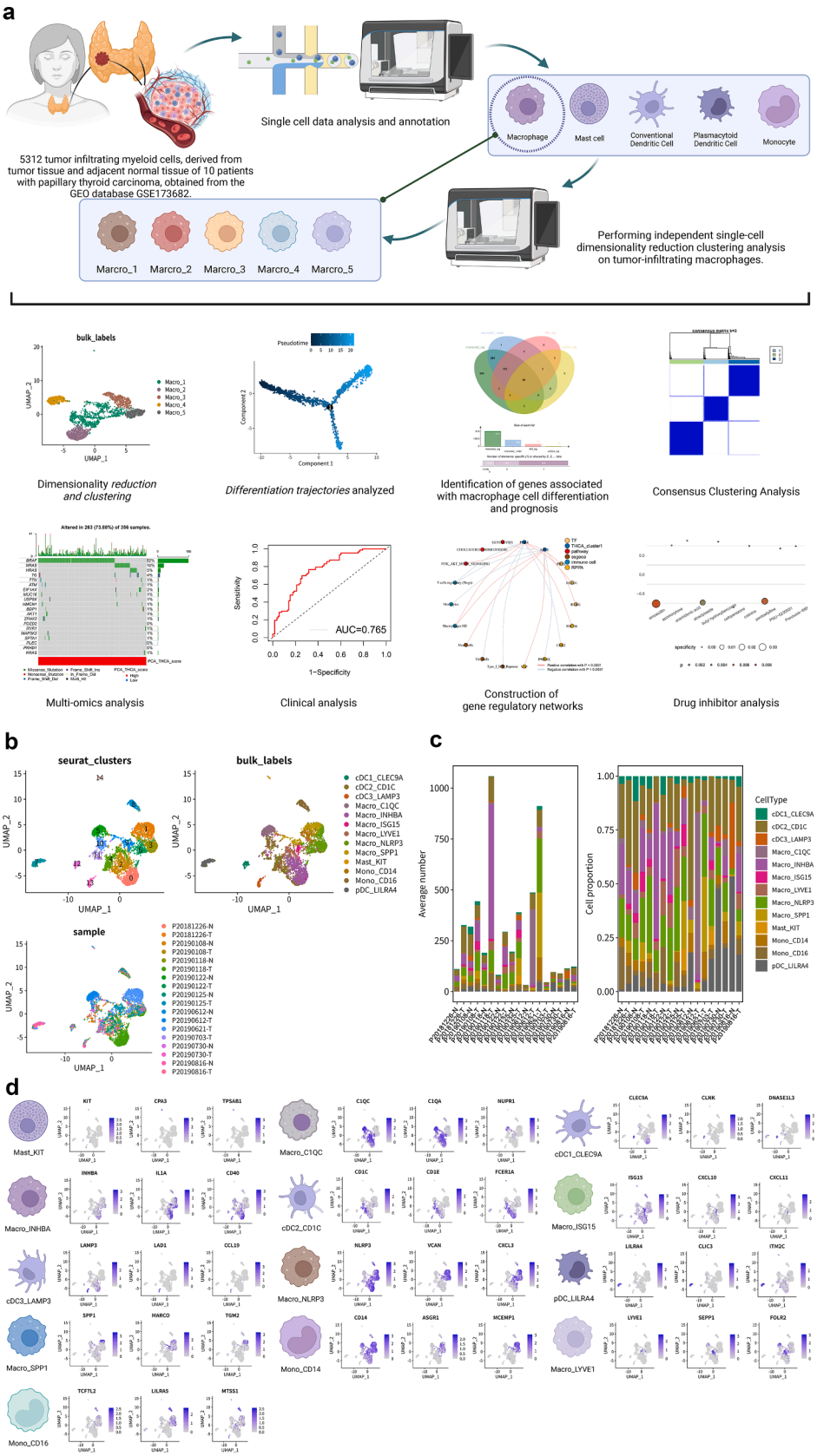
## Result

*ScRNA-seq cell subtype annotation*

Fig. 1a and Figure S1 display the workflow and schematic of the article. To explore the tumor microenvironment and immune landscape of thyroid cancer, we analyzed 5312 TIMs from 10 thyroid cancer patients obtained from the GEO database. Dimensionality reduction and clustering of scRNA-seq data using the Uniform Manifold Approximation and Projection method identified 15 unsupervised clusters, comprising 13 unique cell populations, and 5 cell types (monocytes, adipocytes, macrophages, cDC, and pDC) (Fig. 1b). The average number and percentage distribution of the 13 cell subgroups in 10 cancer and adjacent samples are displayed in a bar chart (Fig. 1c). It can be observed that, compared to normal tissues, tumor tissues exhibit significant macrophage infiltration. Similarly, the bar chart in Figure S2a also demonstrates a comparable distribution, and the Cleveland dot plot analysis of three marker genes further corroborates this observation. The macrophage-associated marker, *CD68* [38], exhibits high expression within the macrophage-related subgroups, while the mast cell-associated marker, *KIT* [39], shows elevated expression within the mast cell subgroups. As observed from the bar graph, tumor tissues distinctly possess a higher expression proportion within the macrophage subgroups. The classical macrophage marker gene *CD68* is significantly overexpressed in the macrophage subgroup, while *ITGAX* is expressed across all subtypes. The expression patterns of marker genes for each cell subtype are visualized in Fig. 1d. Considering the significant contribution of macrophages that infiltrate tumors to the advancement of tumors, we isolated 3326 macrophages individually and conducted separate analyses to reduce their dimensions and cluster them. This process sets the stage for additional investigation into the characteristics of macrophages that infiltrate tumors.

*Differential expression analysis and cell cycle analysis for each cell*

To provide a deeper understanding of the immune composition in the 5312 individual cell samples, we conducted an analysis of differentially




*(caption on next page)*

**Fig. 1.** scRNA-seq analysis of tumor-infiltrating myeloid cells
(a) The experimental design flow. The figure was created using BioRender.com.
(b) The UMAP scatter plot illustrates the results of dimensionality reduction clustering, annotations of cellular subtypes, and the distribution of samples. Notably, 5312 cells from 10 papillary thyroid cancer patients were clustered into 15 unsupervised clusters, which were subsequently annotated as 13 cell populations (Macro_NLRP3, Macro_INHBA, Macro_ISG15, Macro_LYVE1, Macro_C1QC, Macro_SPP1, cDC1_CLEC9A, cDC2_CD1C, cDC3_LAMP3, Mono_CD14, Mono_CD16, Mast_KIT, pDC_LILRA4). The distribution of tumor and normal tissue samples from the 10 patients exhibited differential patterns across these cellular subgroups.
(c) The bar graph depicting the distribution of 13 cell populations across 10 sample sources.
(d) The feature map illustrating the expression patterns of signature genes within 13 cell populations.

expressed genes (DEGs) on the top 2000 genes that exhibit high variability across all samples. A heatmap was created to exhibit the levels of expression of the top 5 differentially expressed genes (DEGs) for every subtype in every cell type (Figure S2b). To visualize the communication among the 13 subtypes of cells, cell communication networks and circular plots were employed (Figure S2c-d). Analysis of the cell cycle (Figure S2e-f) revealed that macrophages were primarily active during the G1 and G2M stages.

*Macrophage subtyping analysis*

The independent dimensionality reduction and clustering analysis of macrophages was performed by constructing separate Seurat objects for all samples. This analysis identified five subtypes of macrophages (Macro_1–5) (Fig. 2a). The distribution of 10 tumors and adjacent non-tumor samples within the five macrophage clusters was visualized using a bar plot (Fig. 2b). Among them, macrophages 2–4 predominantly originated from tumor tissue, while macrophages 1 and 5 exhibited a more complex origin, including both tumor tissue and adjacent normal tissue. The Cleveland dot plot depicted the expression levels of four classical marker genes associated with immune regulation and immune response (*MRC1, CD40, SPP1, C1QC*), revealing distinct immunological expression profiles among the five macrophage subtypes. Fig. 2c presents the top 5 differentially expressed genes (DEGs) within macrophages 1–5. Based on the gene expression characteristics in Fig. 2c, we functionally categorized five macrophage subgroups. Macro_1 expresses M2-related marker MRC1 [40] and stress-response genes HSPA1A and HSPA1B, indicating potential dual roles in pro- and anti-inflammatory processes [41]. Macro_2 is characterized by the expression of TNF and CD40, associated with pro-inflammatory responses [42]. Macro_3 expresses SPP1 and CXCL1, underscoring its role in tissue repair and pro-inflammatory processes [43,44]. Macro_4 exhibits lipid metabolism genes APOE and APOC1, involved in lipid transport dysregulation and inflammation [45]. Finally, Macro_5 expresses CFP and FCN1, consistent with its function in complement activation and inflammation [46,47]. These subgroup characteristics emphasize the versatility of macrophages in the PTC microenvironment and their potential roles in disease progression. Finally, to further elucidate the malignancy-associated expression features of the five macrophage subtypes, Gene Set Variation Analysis (GSVA) was applied to generate hallmark gene set expression scores based on ranked gene changes in the sample expression profiles. In the results of gene set enrichment analysis shown in Fig. 2d, Macro_2 specifically exhibits a high expression of the classical inflammation-related TNFα pathway, highlighting its key role in pro-inflammatory responses [42]. Macro_4 shows elevated expression of the oxidative phosphorylation pathway, emphasizing its specialized function in energy metabolism. The IL6-JAK-STAT3 signaling pathway, associated with immune response and inflammation regulation [48], aligns with the functions of Macro_1. The commonly high expression of downregulated KRAS signaling and WNT/β-catenin pathways across all subtypes indicates a shared mechanism in regulating the PTC tumor microenvironment, consistent with our following co-expression network analysis results. These differences will be comprehensively explained in subsequent subtype analyses.

*Macrophage differentiation trajectory analysis*

To delineate the temporal changes in the immune microenvironment across the five macrophage subtypes, the Monocle package in the R language was employed to infer the differentiation order and formulate cell fate hypotheses. Pseudotime analysis, visualized through cell developmental trajectories, was used to depict the differentiation states of the macrophage subtypes. Figs. 3a-d depict the distribution and differentiation trajectory of macrophage subtypes along three distinct cell fates. Each macrophage subgroup corresponds to a different state of cell differentiation. Macro_1 is involved throughout the differentiation process, while Macro_2 primarily exists in THCA 1, Macro_4 in THCA 2, and Macro_3 and Macro_5 in THCA 3. The time distribution suggests that THCA 1–3 represent early, mid, and late stages of pseudotime differentiation. With a polar division based on the key differentiation node 1 in Fig. 3a, two cell fates (cell fate 1, cell fate 2) are identified, as shown in the heatmap of Fig. 3e. THCA 1 and THCA 3 exhibit different cell fates, while THCA 2 appears to be in a transitional state. This comprehensive analysis led to a hypothesis about the sequence of macrophage subtype differentiation: THCA1 (early differentiation) corresponds to Macrophages 1 and 2; THCA2 (early/mid-differentiation) to Macrophage 4; and THCA3 (late differentiation) to Macrophages 1, 3, and 5.

*MD&PRGs identification*

During the analysis of pseudotime, we incorporated differentially expressed genes across various differentiation stages of macrophage subtypes, as well as relevant genes utilized for trajectory ordering through the DDRTree algorithm. This approach allowed us to identify crucial genes that are associated with the process of macrophage differentiation. Additionally, we obtained bulk RNA sequencing data from the TCGA database, matching gene expression data with clinical sample prognosis information, to identify prognostically relevant key genes. By intersecting these two gene sets, we effectively screened genes based on four criteria: [1] Macrophage Differentiation Genes (MDGs) with q-values $< 0.05$ obtained through the DDRTree algorithm and genes used for differentiation trajectory ordering from scRNA-seq samples; [2] MDGs showing significant correlation (non-parametric test p-value $< 0.05$) with the pseudotime of Papillary Thyroid Cancer (PTC) samples in scRNA-seq; [3] Genes exhibiting significant results in Kaplan-Meier survival analysis (non-parametric test p-value $< 0.05$) from TCGA thyroid cancer samples; [23] Genes demonstrating significant results in univariate Cox survival analysis (non-parametric test p-value $< 0.05$) from TCGA thyroid cancer samples. This process led to the identification of a key gene set, termed Macrophage Differentiation and Prognosis-Related Genes (MD&PRGs), including 19 critical genes (Figure S3a). The prognostic information of these 19 genes is displayed in Figure S3b-c.

*Construct imptcc using consistent cluster analysis*

By conducting a correlation analysis on MD&PRGs, we performed a co-expression network analysis to elucidate the expression patterns of these genes across three cellular differentiation states (THCA1–3) (Fig. 4a). The ConsensusClusterPlus package in R was used, applying the consensus clustering algorithm, to construct the IMPTCC based on the expression profiles of the MD&PRGs. Subsequently, an analysis of

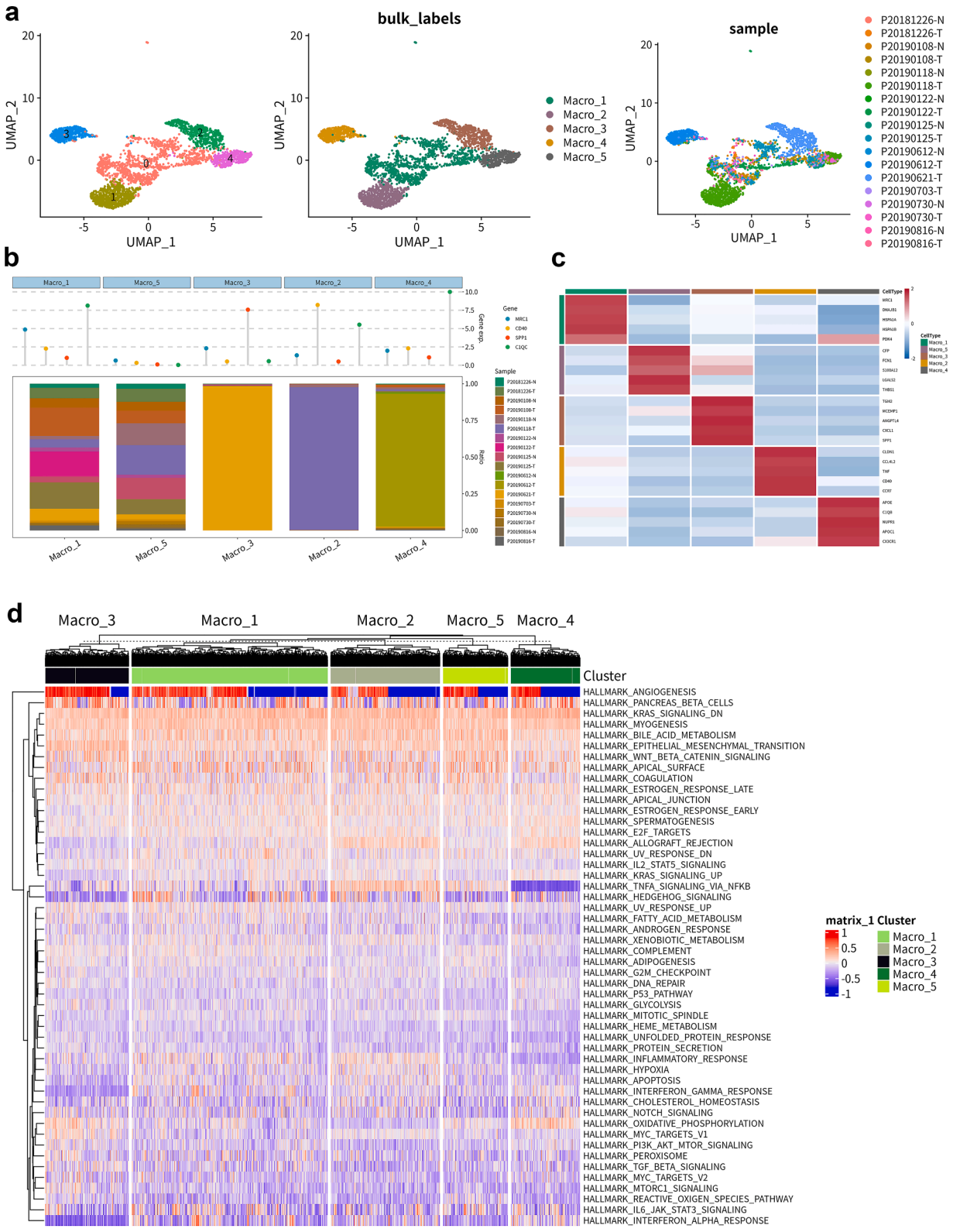



**Fig. 2.** scRNA-seq analysis of macrophage subpopulations.
(a) The UMAP scatter plot revealing the dimensionality-reduced clustering of isolated macrophages, with annotations for cell subtypes and sample distributions. A total of 3326 cells were clustered into 5 cell subgroups (Macro_1, Macro_2, Macro_3, Macro_4, Macro_5). Tumoral and adjacent non-tumoral samples from the 10 sources exhibit differential distribution across these cell subgroups.
(b) The bar graph presenting the distribution of 10 tumor and adjacent non-tumor samples across the 5 macrophage clusters. Cleveland dot plot delineates the expression levels of four canonical markers (*MRC1, CD40, SPP1, C1QC*) associated with immunoregulation and immune response.
(c) The heat map demonstrating the top 5 marker DEG expression levels for each cell type.
(d) The heatmap displays the GSVA gene set enrichment analysis results for the 5 macrophage subpopulations.

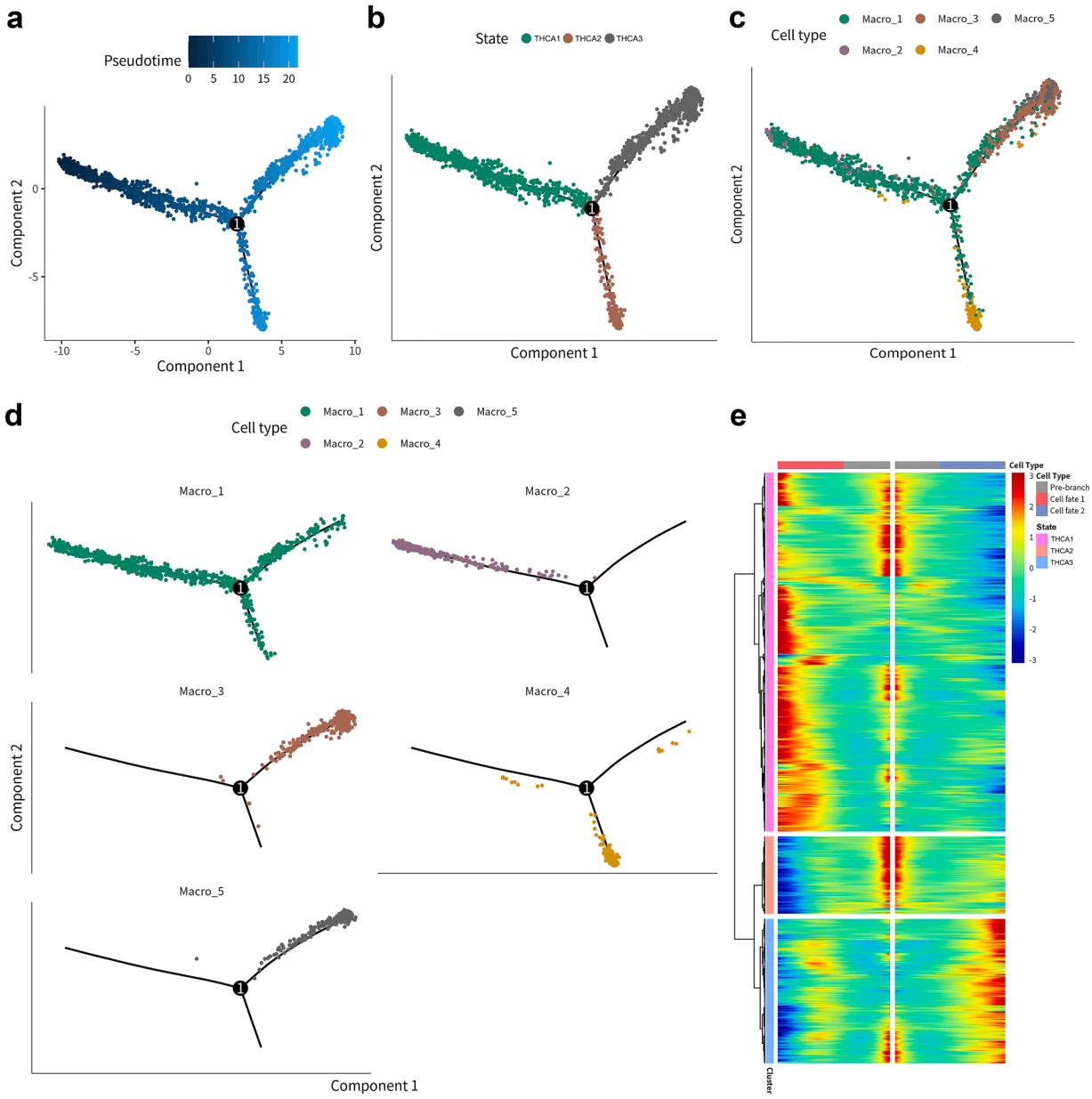



**Fig. 3.** Pseudo-time analysis traces the differentiation trajectory of macrophage subpopulations
(a) Pseudo-time analysis reveals the order in which cell fate differentiation takes place.
(b) Pseudo-time analysis illustrates the distribution of three cell fates.
(c-d) Pseudo-time analysis demonstrates the distribution of 5 macrophage subpopulations.
(e) The heatmap illustrates two distinct cell differentiation fate directions in the overall presentation of three cellular differentiation states (THCA1–3).

feature proportions within the MD&PRGs gene set was conducted, with features subsampled for all specimens. The results suggest that $k = 3$ is the most optimal value (Figs. 4b-c). When there are three subtypes in IMPTCC, their respective expression statuses of the 19 MD&PRGs in each sample are visualized via a heatmap (Fig. 4d).

Integrating Macro_1–5, derived from dimensionality reduction clustering, with the cell fate classifications THCA 1–3 from further analysis, we developed the IMPTCC. This classification, grounded in macrophage differentiation characteristics within the tumor microenvironment and correlated with clinical prognosis, reveals significant clinical and environmental implications. Fig. 4d demonstrates potential correlations between the three IMPTCC subtypes (Cluster 1–3) and the THCA1–3 cell fates, with Subtype 3 aligning with THCA1–2, Cluster 1 with THCA3, and Cluster 2 appearing in a quiescent phase of differentiation. This correlation will be further elucidated in upcoming co-expression network analyses. Further, Kaplan-Meier survival analysis confirmed that patients in Cluster 1 had the poorest prognosis, whereas those in Cluster 3 exhibited the best prognosis (Fig. 4e).

Considering that the categorical results of IMPTCC obtained from consensus clustering might introduce bias in further analyses, we employed the Principal Component Analysis (PCA) algorithm to compute the PCA scores of the samples. All samples were categorized into high/low PCA score groups, and as depicted in Fig. 4f, there was a notable correlation between the PCA scores and prognosis. Furthermore, Fig. 4g and Figure S4a showed the PCA scores of the three IMPTCC subtypes. Figure S4b visually displays the prognostic outcomes of the three subtypes through a Sankey diagram.

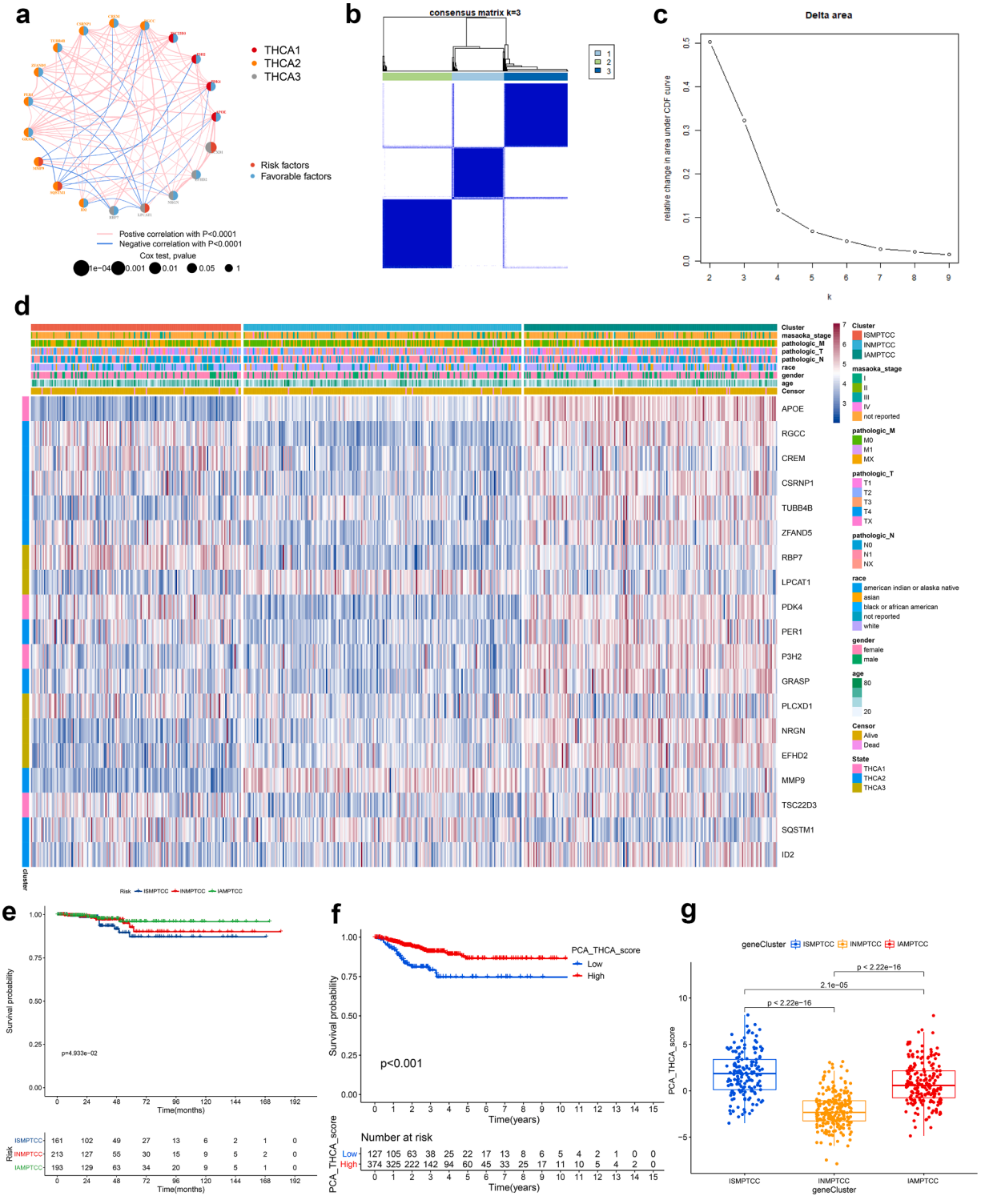



**Fig. 4.** Identification of MD&PRGs and IMPTCC construction

(a) Co-expression network analysis elucidating the co-expression relationships of 19 MD&PRGs across three cellular differentiation states.

(b-c) Bivariate clustering results from consistency matrices suggest optimal clustering into three groups at $k = 3$, evident from pronounced inter-group differences, minimal intra-group heterogeneity, and an expansive area under the CDF curve. Considering clustering distinction and maximal curve area principles, $k = 3$ is deemed optimal.

(d) The heatmap displays the differential expression of 19 MD&PRGs across three IMPTCC clusters and three cellular differentiation states, correlating with their respective clinical characteristics.

(e) Kaplan-Meier survival curves contrasting survival outcomes of the three clusters.

(f) Kaplan-Meier survival curves comparing survival outcomes between high and low PCA groups.

(g) Box plot delineating PCA score variations among the three clusters.

*Construction, validation, and diagnosis of the predictive model*

Fig. 5a presents a box plot illustrating the differential expression of MD&PRGs among the three clinical subtypes. To confirm the accuracy of MD&PRGs in predicting the clinical results of thyroid cancer, we combined the prognostic data of the expression pattern of 19 MD&PRGs from 503 TCGA database samples, excluding those without available metastasis information. This was subjected to univariate Cox regression, and 14 genes with statistical significance were incorporated into the subsequent analysis (Fig. 5b). Using Lasso regression, we identified that the ideal number of variables was nine (Fig. 5c, S4c). Following the analysis of multiple variables, a regression model was performed to create the predictive framework, affirming its significance as a distinct prognostic determinant (Figs. 5d-e). By employing a random number technique, the entire set of samples was partitioned into training and validation sets with a proportion of 6 to 4. The nine key genes exhibited differential expression patterns across the full sample set, training set, and validation set (Fig. 5f). To test the hypothesis, cross-validation was used, resulting in the optimal area under the ROC curve. Afterward, samples were classified into high-risk and low-risk groups based on the median risk score. The KM survival curve (Figure S4d), expression heatmap (Figure S4e), scatter plot (Figure S4f), and risk curve (Figure S4g) indicate that the risk model effectively differentiates the high-risk and low-risk groups. Figure S4h illustrates the quantitative evaluation of the model's accuracy through the area beneath the ROC curve. The heatmap elucidates the baseline distribution of samples within the risk model (Fig. 5g).

*Multi-omics analysis*

Using PCA scores, we performed a co-expression analysis of MD&PRGs at the multi-omics level. The analysis of mutation probability visualized the patterns of gene mutation in the groups with high and low PCA scores, suggesting notable disparities in gene expression patterns between the two groups. Significantly, the group with the highest score displayed an increased likelihood of mutation in the RAS gene, aligning with the unfavorable prognosis observed in the high PCA group (Figure S5a-b). Based on the co-expression analysis presented in Figures S5c-d, it is inferred that the group with a low PCA score exhibits a higher TMB score. This finding suggests that this particular group may display increased sensitivity to immunotherapeutic agents like immune checkpoint inhibitors (for example, PD-1/PD-L1 inhibitors). Figures S5e and S5f respectively display the copy number variations of key genes and their chromosomal location information.

From the results in Figure S6, the application of ORA and GSEA for enrichment analysis of various gene sets indicates that key genes consistently show enrichment for inflammation and immune-related gene sets across multiple databases.

We employed the CIBERSORT algorithm and the ssGSEA algorithm to quantify 23 types of immune cells and immune functions in THCA and compared the correlation of these immune cells/features with MD&PRGs. The results reveal that most immune cells/functions exhibit significant differences across different IMPTCC subtypes (Figure S7a). In the analysis related to immune infiltration, subtype 2 demonstrated relatively active immune cell expression levels, consistent with our definition of subtype 2, suggesting that it might have relatively high immune cell expression levels. However, the intermediate prognosis outcome of subtype 2 suggests that this active immune state does not directly lead to tumor progression. In the combined PCA analysis, consistent with the TMB analysis results, the PD-L1 expression level in the low PCA score group was significantly elevated (Figure S7b). Concurrently, the combined analysis of PD-L1 and CTLA-4 indicates that the low PCA group has a better immune therapy response (Figures S7c-f). Consistently, PCA scores are negatively correlated with immune cell expression levels (Figure S7g).

*Construction of the co-expression network*

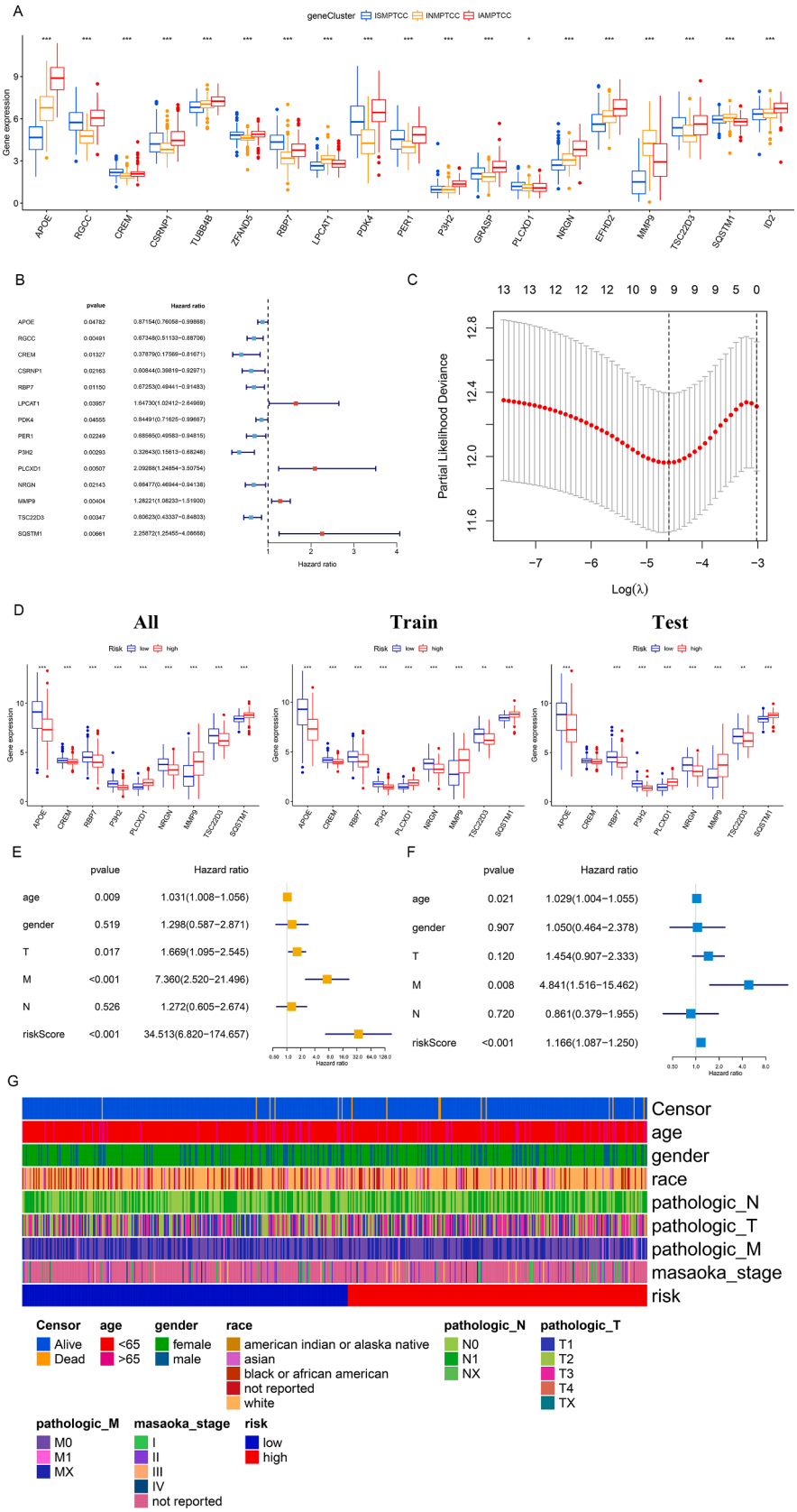
Upon determining the immune infiltration related to key genes and the enrichment of gene sets, we utilized the transcription factor data from the Cistrome database and the RPPA protein chip data obtained from the TCGA database to identify the related transcription factors upstream of the key genes and the downstream RPPA protein expression. In Figure S8, we visualized the key differential genes and the key components of the multi-omics analysis, including transcription factors, immune cells, and hallmark gene sets, using volcano plots and heatmaps. As a result, we discerned the relationships among the key components required to construct the regulatory network, including DETFs, MD&PRGs, immune cells/immune functions, and signaling pathways.

*Construction of the regulatory network for each IMPTCC subtype*

To further investigate the potential key regulatory mechanisms of each IMPTCC subtype, we conducted a co-expression analysis based on the standardized DETFs, MD&PRGs, immune cells/immune functions, and signaling pathways quantitative results for each IMPTCC subtype. Using a correlation coefficient $R>0.300$ and $P<0.001$ as the minimum screening criteria, we constructed three complex regulatory networks centered on MD&PRGs, which encompassed upstream TFs, downstream signaling pathways, and potential regulatory immune cells/functions (Fig. 6a). Based on gene expression patterns, inhibitor analysis, and integrated regulatory network analysis, we named the three types from cluster1–3 as Immune-Suppressive Macrophage-Infiltrated Thyroid Carcinoma Classification (ISMPTCC), Immune-Neutral Macrophage-Infiltrated Thyroid Carcinoma Classification (INMPTCC), and Immune-Activated Macrophage-Infiltrated Thyroid Carcinoma Classification (IAMPTCC). We also used a co-expression heatmap to display the specific correlation coefficients in each regulatory network (Fig. 6b). The detailed regulatory relationships within the co-expression network will be further elaborated in the discussion section. At this stage, we have understood the expression characteristics of the three IMPTCC subtypes. This clarifies the unique correspondence among macrophage subtypes (Macro_1–5), cell fate differentiation states (THCA1–3), and IMPTCC subtypes: IAMPTCC (related to M1 macrophage differentiation direction) correlates with THCA1 and corresponds to macrophages 1, 2; ISMPTCC (related to M2 macrophage differentiation direction) with THCA3, corresponding to macrophages 1, 3, 5. INMPTCC, without a clear macrophage subgroup correspondence, suggests an intermediary differentiation state. Integrating these relationships, we find that macrophages infiltrating PTC exhibit characteristics of transitioning from M1 polarized macrophages (anti-cancer, pro-inflammatory) to M2 polarized macrophages (anti-inflammatory, pro-cancer). This aligns with our definitions of macrophage subtypes, cell fate differentiation subtypes, and IMPTCC subtypes.

*Identification of inhibitors through C-map and CCLE algorithms*

To further integrate our proposed clinical staging with clinical treatment decisions, we initially employed the Cmap algorithm to identify small molecule drugs targeting MD&PRGs and TFs within the regulatory network of each IMPTCC subtype. We then selected the top 10 drugs with the smallest P-values and visualized them using a dot heatmap as target drugs for each IMPTCC subtype (Fig. 6c). Subsequently, with the assistance of the GDSC database, an analysis of drug sensitivity related to cell fate genes identified PIK-93 as a specific inhibitor for ISMPTCC (Figs. 6d-f). By integrating the results from both Cmap and GDSC databases, PI3K pathway inhibitor PIK93 [49], WNT pathway inhibitor LGK-974 [50,51], and M2 polarization inhibitor MS-275 [52] were found to be specifically effective against ISMPTCC, INMPTCC, and IAMPTCC, respectively.




(caption on next page)

**Fig. 5.** Construction, Validation, and Diagnosis of the Predictive Model
(a) Box plot illustrating the expression level disparities of 19 MD&PRGs across three subtypes. (Statistical significance denoted as: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p<0.05$).
(b) Forest plot detailing the univariate Cox regression analysis of the 19 MD&PRGs, with 14 genes displaying statistical significance earmarked for subsequent selection.
(c) Lasso path diagram demonstrating that post-lasso regression, the number of key genes incorporated into the predictive model was finalized at nine.
(d) Box plot showcasing the differential expression of the nine pivotal genes across all sample groups, training sets, and validation sets, stratified by risk.
(e) Univariate Cox regression analysis validates the risk score, computed from the predictive model, as a potential independent prognostic factor for PTC.
(f) Multivariate Cox regression analysis confirms the risk score, derived from the predictive model, as an independent prognostic determinant for PTC.
(g) Heatmap elucidates the baseline distribution of samples within the risk model.

*Histopathological validation*

To enhance the clinical applicability of IMPTCC subtyping, the present study adopted the differential expression of the *APOE* gene as a marker to delineate the three subtypes of IMPTCC. A total of 72 patients with a confirmed diagnosis of PTC were included, and based on the staining results of tumor pathological sections, they were categorized into groups with high and low staining intensities for *APOE* (Fig. 7a-b). Detailed clinical and immunohistochemical data were carefully recorded for all samples. To discern disparities in clinical information and prognosis between the two groups, we conducted an association analysis concerning quantitative variables such as age, maximal dimension, and Ki67, along with categorical variables like metastasis, *BRAF*, and *TERT*, contingent on the high and low *APOE* expression categorizations. Given the inherent characteristics of the data distribution, the Mann-Whitney U test was deployed for quantitative variable comparisons. Categorical variables were evaluated using chi-square tests or Fisher's exact tests. Table 1 showcases the baseline clinical and statistical analysis outcomes of the 72 PTC samples. We focused on the intensity of Ki67 expression, an immunohistochemical index associated with the prognosis of PTC [53], and the average intensity of Ki67 expression was higher in the *APOE* low expression group than in the *APOE* high expression group (Fig. 7c). Regarding clinical features, the percentage of lymph node metastasis and the mean tumor diameter were significantly higher in the *APOE* low expression group than in the *APOE* high expression group (Fig. 7d-e). Additionally, there was a significant difference in the positivity rate of HBME-1, an immunohistochemical marker used for the diagnosis of PTC [54], between the *APOE* high/low groups (Fig. 7f). These observations resonate with our antecedent subtype analysis, suggesting an association between diminished APOE expression and a less favorable clinical prognosis.

*Sequencing to detect transposase accessible chromatin*

Moreover, the accessibility of chromatin fragments for important genes (SQSTM1, RGCC, NRGN, PDK4, PLCXD1, APOE) was demonstrated by high-throughput sequencing findings of transposase-accessible chromatin (ATAC-seq) (Figure S9).

## Discussion

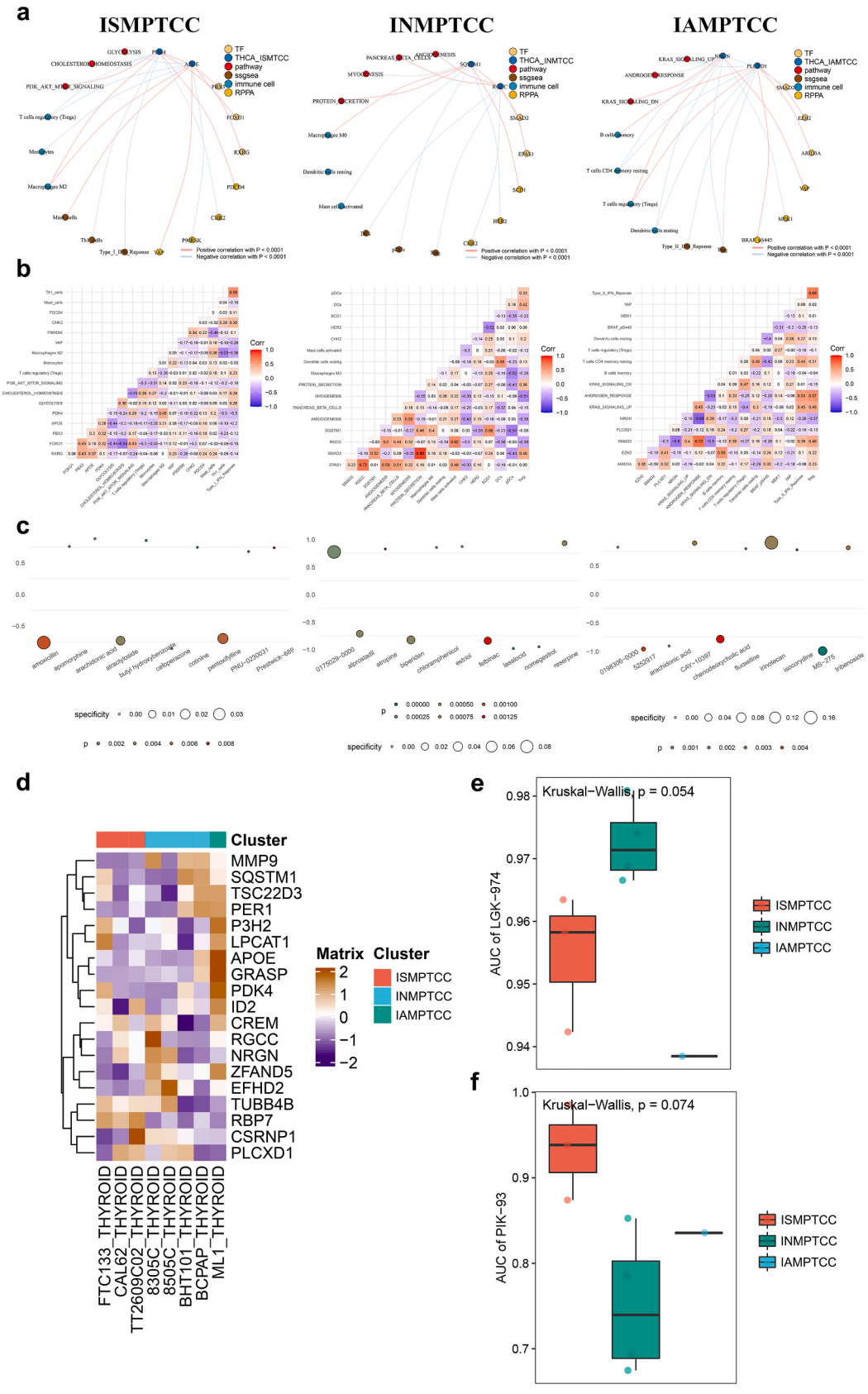
Tumors exhibit a complicated ecological environment in which heterogeneous malignant cells engage in complex interactions with both immune and non-immune cells, establishing the intricate cell network of the tumor microenvironment [10,55]. In the tumor microenvironment, TIMs [14] are crucial components of the immune ecology of the TME [15]. TIMs consist of several distinct major lineages, including adipocytes, plasmacytoid dendritic cells (pDC), conventional dendritic cells (cDC), monocytes, and macrophages [12]. Macrophages are considered a double-edged sword with dual potential in cancer; their plasticity in response to environmental cues offers rich possibilities for macrophage-related research [56–58]. Owing to their function in mediating anti-tumor activity and inducing adaptive immune responses in immune regulation, macrophages have become one of the primary targets of current checkpoint inhibitor immunotherapy [59,60]. Additionally, studies on colorectal cancer in humans and mice have revealed that different subgroups of myeloid cells, while presenting functional differences, have variable sensitivities to therapeutic strategies targeting immune cells [61]. This result suggests that different differentiation states of macrophages may lead to varied clinical prognostic outcomes and responses to targeted treatments in tumors, which is also one of the significant reasons for our study on macrophage differentiation in the thyroid cancer environment.

During the research on macrophage differentiation, the polarization system of "classically activated" M1 and "alternatively activated" M2 macrophages has been used to describe the in vitro activation status of macrophages [62]. However many studies believe that the binary M1-M2 model is insufficient to describe the activation of macrophages [63–65]. Owing to the plasticity of macrophages and their dependency on the surrounding environment, macrophages in the actual in vivo environment may exhibit more complex differentiation forms [66].

Through the integration of scRNA-seq and bulk RNA sequencing data, we analyzed the differentiation status of macrophages in tumor-related myeloid cells. Based on the differentiation characteristics of macrophages, we further reclassified the immunological environment levels of thyroid cancer. Using multi-omics analysis based on both scRNA-seq and bulk-sequencing data, we categorized the differentiation states of macrophages into three subtypes (ISMPTCC, INMPTCC, and IAMPTCC). These subtypes are classified based on their immune characteristics into immune homeostasis, immune activation, and immune repair states, respectively reflecting the traits of M1-M0-M2 macrophages. However, unlike the classical binary description, this immune progression state, based on co-expression analysis results, does not strictly follow the classical characteristics of macrophage M1/M2 differentiation and shows different molecular mechanisms of action.

ISMPTCC is associated with Type I IFN response, M2 macrophages, and immune cells involved in immune repair. Concurrently, within the co-expression network, genes, transcription factors, and signaling pathways expressed by ISMPTCC manifest cellular metabolic traits. Notably, the expression patterns of glycolytic and lipid metabolic markers align with the anti-inflammatory metabolic characteristics typically seen in classically activated M2 macrophages. Macrophage polarization involves intricate metabolic shifts. Lipid and glucose metabolism are at the forefront of these changes. In M2 polarized macrophages, research indicates certain trends. Compared to M1 polarized macrophages, which have pronounced inflammatory traits, M2 macrophages show decreased glycolytic metabolism levels [67]. They also demonstrate increased lipid metabolism, represented by processes like fatty acid uptake and oxidation [68–70]. These observations match the gene expression traits we identified in our study. Within the co-expression network of Fig. 6, cluster1, key genes *PDK4, APOE*, transcription factor FOXO1, and the downstream glycolytic pathway, PI3K/AKT/mTOR pathway hint at the potential regulatory mechanisms of this subtype.

The transcription factor FOXO1 is a key member of the Forkhead transcription factors (FOX) family [71]. Depending on the regulatory targets it modulates, FOXO1 exhibits intricate regulatory activities, encompassing processes such as cell apoptosis, cellular metabolism, and immunological processes [72]. Its role in metabolic processes extends to both glucose metabolism and lipid metabolism. In terms of glucose




*(caption on next page)*

**Fig. 6.** Construction of multi-omics regulatory network and target drug prediction for three subtypes in IMPTCC typing
(a) Multi-omics analysis results, amalgamating both scRNA-seq and bulk-sequencing data, utilizing a minimum selection criterion of correlation coefficient $R>0.400$ and $P<0.001$, structured around the MD&PRGs to form intricate regulatory networks—consisting of upstream TFs, downstream signaling pathways, and potential immune cell/function modulators—for the ISMPTCC, INMPTCC, and IAMPTCC subtypes.
(b) Co-expression heatmap manifesting specific correlation coefficients within each regulatory network.
(c) Dot map spotlighting target small molecule drugs identified for each IMPTCC subtype's regulatory network via the Cmap algorithm, with the ten drugs boasting the lowest P-values proposed as prospective inhibitors for each IMPTCC subtype.
(d-f) Utilizing data from the CCLE tumor cell line database for drug susceptibility analysis on genes affiliated with cellular destiny, it was discerned that the inhibitor PIK-93 exhibits specificity within ISMPTCC, while LGK-974 is specific to INMPTCC.

metabolism, FOXO1 has been reported to be activated in diabetic cardiomyocytes and attenuates glycolysis by stimulating *PDK4* [73]. *PDK4* acts by phosphorylating and inhibiting the activity of another protein kinase, PDH, thereby suppressing glycolysis by limiting pyruvate production [74]. Consistent with this, *PDK4* is expressed at high levels in ISMPTCC (Fig. 5a), suggesting that ISMPTCC may exhibit a suppressed state in glycolytic metabolism.

On the other hand, in terms of lipid metabolism regulation, FOXO1 has been reported to modulate lipid synthesis through a principal lipogenesis regulator, PPARγ [75]. Concurrently, in ISMPTCC, *RBP7* is distinctly overexpressed (Fig. 5a). *RBP7* is a target gene of PPARγ [76]. A dedicated study on this gene revealed that overexpression of *RBP7* in adipocytes enhances lipid droplet formation and TG accumulation. This is accompanied by an increased expression of crucial lipogenic proteins including Pparγ, Fabp4, C/ebpα, and AdipoQ [77], suggesting an active state of lipid synthesis in ISMPTCC. Beyond its evident lipid synthesis characteristics, regarding lipid breakdown regulation, *APOE* is specifically underexpressed in ISMPTCC (Fig. 5a). Apolipoprotein E (apoE) is a multifunctional protein involved in clearing apoB-containing chylomicrons and very low-density lipoprotein (VLDL) remnants from plasma, thereby reducing plasma lipid levels [78]. The diminished expression of *APOE* indicates reduced lipid transport in ISMPTCC, and lipid retention is one of the crucial metabolic features of tumors [79]. This aligns with the traits of M2 macrophages promoting tumor progression.

The metabolic characteristics of ISMPTCC in terms of carbohydrate and lipid metabolism align with the metabolic features of M2 polarized macrophages, both exhibiting inhibition of glycolysis and enhancement of lipid metabolism. We also observed that the PI3K/AKT/mTOR pathway is an integral part of the ISMPTCC expression network. Intriguingly, both FOXO1 and MTOR are activated by the PI3K/AKT signal [80], and MTOR also plays a role in metabolic regulation [81]. This suggests a consistent role in metabolic regulation by MTOR and FOXO1, both downstream molecules of PI3K/AKT, and hints at a new direction for targeting the pathogenic mechanisms of ISMPTCC. In line with this, based on our drug sensitivity analysis conducted on tumor cell lines, the PI3K inhibitor PIK-93 demonstrated favorable sensitivity [82]. Being an upstream target of both the PI3K/AKT/MTOR and PI3K/AKT/FOXO1 pathways, inhibiting PI3K might play a role in targeting the metabolic regulation of FOXO1 and MTOR, assisting TAMs in adjusting their polarization state.

INMPTCC expresses M0 macrophages and immune cells related to immune activation. Within its co-expression network, the regulatory relationships among molecules suggest that the autophagy mechanism might exert its influence by affecting tumor angiogenesis. Fig. 6, showcasing the INMPTCC co-expression network, highlights key genes *SQSTM1*, protein SCD1, and transcription factors FOXO1, SMAD2, and EPAS1. The downstream angiogenesis pathway hints at the potential regulatory mechanisms inherent to this subtype.

Autophagy exhibits bidirectional and contradictory roles in both tumor suppression and tumor progression [83]. SMAD2 is a member of the SMAD protein family, mediating the signaling of transforming growth factor-beta (TGF-β) to regulate various cellular processes [84]. Research has identified that the TGF-β signaling pathway can activate the autophagy process in human cancer cells [85], thereby linking SMAD2 to the activation of autophagy. Concurrently, *P62/SQSTM1*, as a selective autophagy receptor and a crucial molecule in autophagosome formation, extensively participates in the autophagy process as a principal component and an initiator of WNT-regulated autophagy functions [86]. On the other hand, SCD1 has been established as an essential molecule during the early stages of autophagosome formation. The SCD1 inhibitor 28c inhibits the formation of p62/SQSTM1 puncta [87]. The presence of these three molecules in the INMPTCC co-expression network suggests the significance of the autophagy mechanism in the pathogenic mechanism of INMPTCC tumors. Autophagy in macrophages suppresses M1-type polarization of macrophages, alleviating chronic inflammation and organ fibrosis. The autophagy signaling in INMPTCC, at the molecular mechanism level, indicates that INMPTCC may serve as an intermediate state transitioning from M1 polarization to M2 polarization [88].

Further, the presence of the EPAS1 molecule suggests a connection between autophagy and the angiogenic phenotype. Angiogenesis is closely linked with both macrophages and tumors. Tumor-associated macrophages (TAMs) can promote tumor progression by producing angiogenic growth factors [18]. Angiogenesis-associated phenotypes are considered to be prominently enriched pathways in a macrophage-infiltrated immune microenvironment [61]. This perspective is validated in a single-cell study encompassing 15 cancer types, where angiogenesis-promoting tumor-associated macrophages were associated with poor prognoses in multiple tumors [19]. As a member of the hypoxia-inducible factors (HIFs), the family of transcription factors activated by hypoxia [89], EPAS1 can stabilize and activate cancer cells. Here, it induces angiogenesis-associated genes like VEGF [90], promoting angiogenesis and hence tumor progression. Concurrently, EPAS1 has been targeted therapeutically to inhibit vascular and tumor growth [91]. Importantly, autophagy activation involving *SQSTM1* has been proven to suppress the angiogenic phenotype by degrading EPAS1, thereby inhibiting renal cell carcinoma. This indicates the adverse effect of autophagy activation on the angiogenic phenotype. In INMPTCC, aside from the high expression of genes associated with angiogenesis inhibition, *MMP9* is specifically overexpressed (Fig. 5a). *MMP9*, a member of the pericellular proteases, plays a crucial role in initiating angiogenesis. It stimulates the recruitment and function of auxiliary cells from blood or bone marrow, enhancing angiogenesis [92]. Concurrently, the secretion of *MMP9* is regulated by EPAS1, and overexpression of EPAS1 can boost the secretion of both VEGF and *MMP9* [93].

In a comprehensive view, INMPTCC demonstrates intricate characteristics of autophagy and angiogenesis. The suppressive effect on macrophage M1 polarization brought about by autophagy and the negative regulatory influence on the angiogenic phenotype might counterbalance the positive regulatory role of angiogenesis directly mediated by the activated *MMP9*. Together, these factors might collectively constitute the regulatory milieu of INMPTCC, potentially explaining the intermediate prognostic state of INMPTCC. Concurrently, in our drug analysis, we observed that WNT pathway inhibitors, LGK-974 and WNT-C59, exhibited specificity in this subtype. Given the pivotal role of the WNT pathway in autophagy mechanisms, we hypothesize that the autophagic mechanism in INMPTCC might predominantly manifest tumor-promoting effects. Nonetheless, this conjecture warrants further experimental validation.

IAMPTCC manifests a type II IFN response as well as immune cells associated with other immune homeostatic states. Correspondingly, the

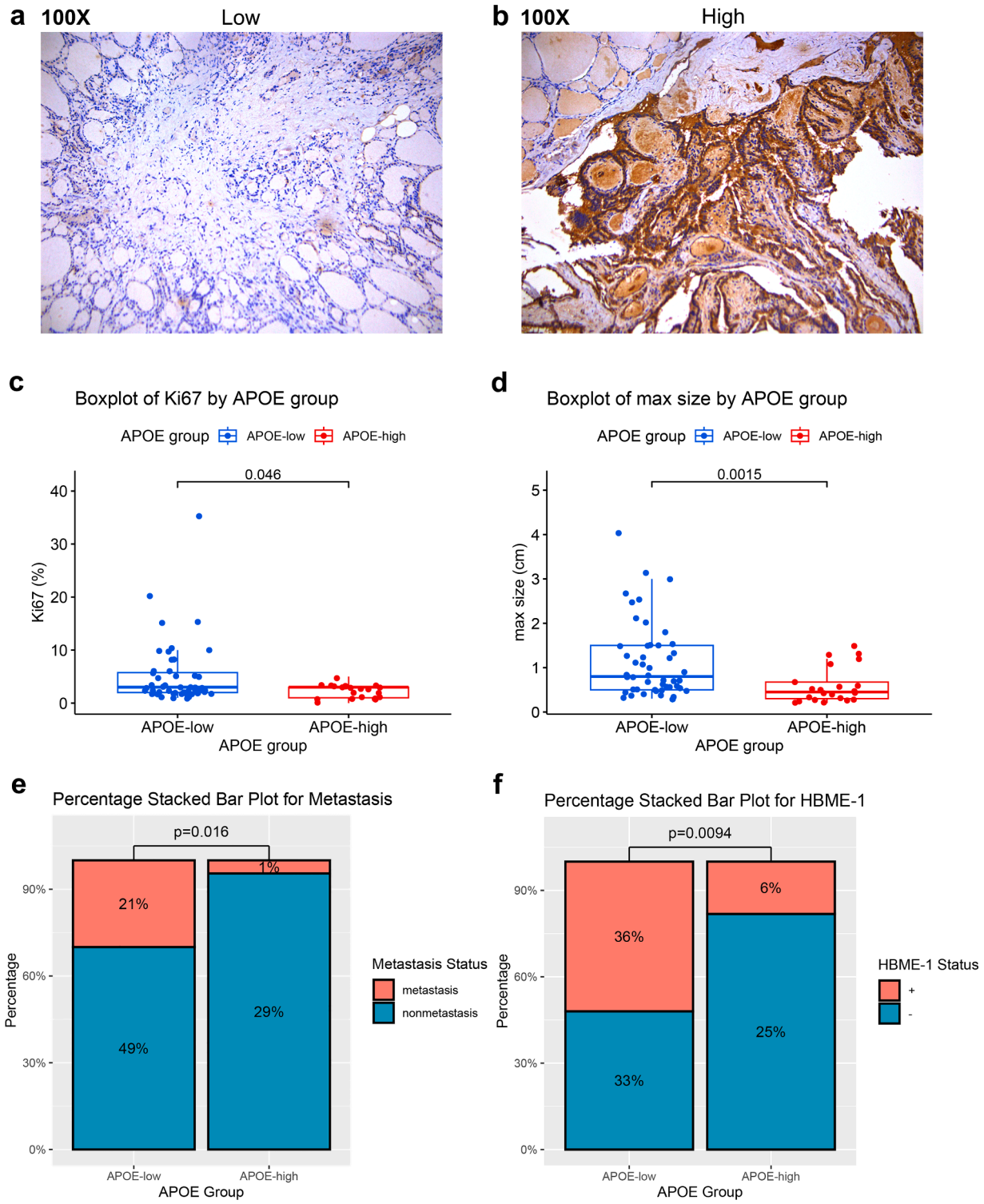



**Fig. 7.** Histopathological Validation
(a) Representative H&E images for the *APOE*-low staining group.
(b) Representative H&E images for *APOE*-high staining group.
(c) Stacked bar chart showing the differences in metastasis rates between the *APOE* high and low staining groups.
(d) Stacked bar chart showing the differences in HBME-1 positivity rates between the *APOE* high and low staining groups.
(e) Boxplot showing the differences in maximum tumor diameter between the *APOE* high and low staining groups.

molecular regulation evidence in IAMPTCC directly suggests its association with the M1 polarization state of macrophages. IAMPTCC's immune milieu displays characteristics related to immune homeostasis, accompanied by a type II IFN response. Interferons (IFNs) were initially identified as substances that "interfere" with viral replication in vitro, comprising type I interferons (IFN-α/β and related molecules), type II interferon (IFN-γ), and type III interferons (IFN-lambda) [94]. Type II IFN response is considered a classical hallmark of activated M1

**Table 1**
Clinical baseline information and differential analysis results of PTC patients with high/low APOE expression.
Significant statistical differences were observed between patients with high and low APOE expression in terms of metastasis rate, maximum tumor diameter, Ki67 mutation proportion, and HBME-1 positivity rate.

| | APOE-low ($N = 50$) | APOE-high ($N = 22$) | P-value |
|---|---|---|---|
| **Metastasis** | | | |
| Metastasis | 15 (30.0 %) | 1 (4.5 %) | 0.0161 |
| Nonmetastasis | 35 (70.0 %) | 21 (95.5 %) | |
| **Max size (cm)** | | | |
| Mean (SD) | 1.12 (0.844) | 0.595 (0.408) | 0.00145 |
| Median [Min, Max] | 0.800 [0.300, 4.00] | 0.450 [0.200, 1.50] | |
| **Ki67 (%)** | | | |
| Mean (SD) | 5.00 (5.94) | 2.23 (1.19) | 0.0457 |
| Median [Min, Max] | 3.00 [1.00, 35.0] | 3.00 [0, 5.00] | |
| **Age (years)** | | | |
| Mean (SD) | 48.2 (12.6) | 47.3 (12.5) | 0.651 |
| Median [Min, Max] | 47.5 [22.0, 70.0] | 41.0 [30.0, 69.0] | |
| **HBME-1** | | | |
| - | 24 (48.0 %) | 18 (81.8 %) | 0.00938 |
| + | 26 (52.0 %) | 4 (18.2 %) | |
| **BRAF** | | | |
| - | 17 (34.0 %) | 8 (36.4 %) | 1 |
| + | 33 (66.0 %) | 14 (63.6 %) | |
| **TERT** | | | |
| - | 50 (100 %) | 22 (100 %) | 1 |
| + | 0 (0 %) | 0 (0 %) | |
| **CK19** | | | |
| - | 2 (4.0 %) | 1 (4.5 %) | 1 |
| + | 48 (96.0 %) | 21 (95.5 %) | |
| **TPO** | | | |
| - | 44 (88.0 %) | 22 (100 %) | 0.168 |
| + | 6 (12.0 %) | 0 (0 %) | |
| **Galectin-3** | | | |
| - | 2 (4.0 %) | 0 (0 %) | 1 |
| + | 48 (96.0 %) | 22 (100 %) | |
| **CD56** | | | |
| - | 46 (92.0 %) | 21 (95.5 %) | 1 |
| + | 4 (8.0 %) | 1 (4.5 %) | |
| **34βE12** | | | |
| - | 5 (10.0 %) | 2 (9.1 %) | 1 |
| + | 45 (90.0 %) | 20 (90.9 %) | |
| **TG** | | | |
| - | 2 (4.0 %) | 1 (4.5 %) | 1 |
| + | 48 (96.0 %) | 21 (95.5 %) | |

macrophage polarization [95,96]. *EFHD2*, which exhibits specific overexpression in IAMPTCC (Fig. 5a), is believed to be a crucial molecule for the binding of IFN-γ to functional IFN-γR [97]. We also observed the appearance in IAMPTCC's molecular regulatory network of the protein YAP1, known to promote M1 polarization [98], molecules related to the MAPK pathway involved in M1 polarization regulation, namely BRAF, MEK1 [99], and EZH2, which enhances the chemotaxis of inflammatory macrophages in the tumor environment [100].

In terms of pathway expression features, the KRAS pathway emerges in the co-expression network [101], aligning with a breast cancer-related study that found the enrichment of KRAS signaling-associated genes to be associated with an anti-tumor immune microenvironment. Likewise, the androgen receptor pathway, which also appears in the co-expression network, has similar evidence [102]. Evidently, in terms of immune characteristics and gene and pathway expression patterns, IAMPTCC displays differentiation characteristics of M1 polarized macrophages, distinguishing it significantly from subtypes 1 and 2 in molecular expression features.

We acknowledge the limitations of our study. As a project reliant on publicly available data analysis, we utilized thyroid cancer scRNA-seq data from Peking University Cancer Hospital & Institute and bulk RNA sequencing data obtained from the TCGA database. This cross-verification approach may introduce biases stemming from ethnic and regional differences. Additionally, we made limited inferences based on existing research conclusions regarding the regulatory network relationships indicated by our bioinformatics analysis results. Preliminary experimental validations were conducted based on these interpretations. Building upon our preliminary validations, our future research will delve into the transdifferentiation of macrophages from anti-inflammatory to pro-inflammatory states. We plan to conduct comprehensive experiments, encompassing both upstream and downstream validation of the potential molecular regulatory relationships identified in our initial study. This will involve detailed functional assays to understand the mechanistic underpinnings of macrophage behavior in the context of papillary thyroid carcinoma. These studies aim to validate and expand upon our predictive model and candidate molecules, enriching our understanding of their roles in tumor biology.

## Conclusion

In summary, through the integration of scRNA-seq data from the PTC environment and bulk RNA-seq data from THCA for multi-omics analysis, our study established a thyroid cancer typing system based on macrophage differentiation characteristics. Guided by intricate upstream and downstream co-expression network regulatory clues, we identified potential thyroid cancer metabolism and regulation features under different macrophage differentiation states. The differentiation variance in macrophage states may indeed offer some discriminative power in the reclassification of PTC. Research into the specific mechanisms underlying the prognostic variations of PTC should consider the tumor immune microenvironment, especially the differential characteristics of tumor-infiltrating macrophages. We hope our analysis can offer new insights and research directions for risk stratification and precision treatment of PTC.

## Funding

This study was supported by the General Project of Natural Science Foundation of Henan Province (222300420568); Key Medical Science and Technology Project of Henan Province (SBGJ202101014); Major Scientific Research Projects of Traditional Chinese Medicine in Henan Province (20–21ZYZD14); Cultivation of Young and Middle-aged Health Science and Technology Innovation Leading Talents in Henan Province (YXKC2020015); National Natural Science Foundation of China (Grant No. 82002923) and Shanghai Rising-Star Program (Sailing Special Program) (No. 23YF1458400).

## Availability of data and materials

Table S1 displays all the datasets, code packages, and software information employed in this study. The clinical patient information and all the codes generated during the data analysis process in this study are stored in Supplementary material.

## Ethics approval and consent to participate

Approval for histopathological validation using clinical samples was obtained from the Clinical Research Institution Review Committee of the Tenth People's Hospital, which is affiliated with Tongji University in Shanghai. Before participation, all patients provided informed written consent. Samples for the additional validation were gathered from the Tenth People's Hospital, which is affiliated with Tongji University, from 2016 to 2018.

## Consent for publication

Not applicable.

## CRediT authorship contribution statement

**Hanlin Sun:** Formal analysis, Software, Validation, Writing – original draft. **Zhengyan Chang:** Funding acquisition, Methodology, Resources, Validation. **Hongqiang Li:** Resources, Supervision, Writing – review & editing. **Yifeng Tang:** Resources, Supervision, Writing – review & editing. **Yihao Liu:** Investigation, Supervision, Writing – review & editing. **Lixue Qiao:** Investigation, Writing – review & editing. **Guicheng Feng:** Writing – review & editing. **Runzhi Huang:** Conceptualization, Data curation, Formal analysis, Methodology, Software. **Dongyan Han:** Funding acquisition, Supervision. **De-tao Yin:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We would like to express our gratitude to the team at TCGA and GEO for allowing us to use their data.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2024.101889.

## References

[1] J.J. Wiltshire, T.M. Drake, L. Uttley, S.P. Balasubramanian, Systematic review of trends in the incidence rates of thyroid cancer, Thyroid. 26 (11) (2016) 1541–1552.
[2] A. Antonelli, C. La Motta, Novel therapeutic clues in thyroid carcinomas: the role of targeting cancer stem cells, Med. Res. Rev. 37 (6) (2017) 1299–1317.
[3] H. Qiu, S. Cao, R. Xu, Cancer incidence, mortality, and burden in China: a time-trend analysis and comparison with the United States and United Kingdom based on the global epidemiological data released in 2020, Cancer Commun. (Lond) 41 (10) (2021) 1037–1048.
[4] S.M. Ferrari, G. Bocci, T. Di Desidero, I. Ruffilli, G. Elia, F. Ragusa, et al., Vandetanib has antineoplastic activity in anaplastic thyroid cancer, in vitro and in vivo, Oncol. Rep. 39 (5) (2018) 2306–2314.
[5] A. Antonelli, S.M. Ferrari, P. Fallahi, P. Berti, G. Materazzi, I. Marchetti, et al., Evaluation of the sensitivity to chemotherapeutics or thiazolidinediones of primary anaplastic thyroid cancer cells obtained by fine-needle aspiration, Eur. J. Endocrinol. 159 (3) (2008) 283–291.
[6] L. Davies, J.K. Hoang, Thyroid cancer in the USA: current trends and outstanding questions, Lancet. Diabetes Endocrinol. 9 (1) (2021) 11–12.
[7] M. Xing, Molecular pathogenesis and mechanisms of thyroid cancer, Nat. Rev. Cancer 13 (3) (2013) 184–199.
[8] R.S. Scheffel, J.M. Dora, A.L. Maia, BRAF mutations in thyroid cancer, Curr. Opin. Oncol. 34 (1) (2022) 9–18.
[9] F. Balkwill, A. Mantovani, Inflammation and cancer: back to Virchow? Lancet 357 (9255) (2001) 539–545.
[10] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (5) (2011) 646–674.
[11] M.J. Bissell, W.C. Hines, Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression, Nat. Med. 17 (3) (2011) 320–329.
[12] L.M. Coussens, L. Zitvogel, A.K. Palucka, Neutralizing tumor-promoting chronic inflammation: a magic bullet? Science 339 (6117) (2013) 286–291.
[13] S.M. Ferrari, P. Fallahi, G. Elia, F. Ragusa, I. Ruffilli, S.R. Paparo, et al., Thyroid autoimmune disorders and cancer, Semin. Cancer Biol. 64 (2020) 135–146.
[14] E. Le Chatelier, T. Nielsen, J. Qin, E. Prifti, F. Hildebrand, G. Falony, et al., Richness of human gut microbiome correlates with metabolic markers, Nature 500 (7464) (2013) 541–546.
[15] C. Engblom, C. Pfirschke, M.J. Pittet, The role of myeloid cells in cancer therapies, Nat. Rev. Cancer 16 (7) (2016) 447–462.
[16] M.C. Schmid, J.A. Varner, Myeloid cells in the tumor microenvironment: modulation of tumor angiogenesis and tumor inflammation, J. Oncol. 2010 (2010) 201026.
[17] F. Ginhoux, J.L. Schultze, P.J. Murray, J. Ochando, S.K. Biswas, New insights into the multidimensional concept of macrophage ontogeny, activation and function, Nat. Immunol. 17 (1) (2016) 34–40.
[18] D.G. DeNardo, B. Ruffell, Macrophages as regulators of tumour immunity and immunotherapy, Nat. Rev. Immunol. 19 (6) (2019) 369–382.
[19] S. Cheng, Z. Li, R. Gao, B. Xing, Y. Gao, Y. Yang, et al., A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells, Cell 184 (3) (2021) 792–809.e23.
[20] B. Hie, B. Bryson, B. Berger, Efficient integration of heterogeneous single-cell transcriptomes using Scanorama, Nat. Biotechnol. 37 (6) (2019) 685–691.
[21] X. Zhang, Y. Lan, J. Xu, F. Quan, E. Zhao, C. Deng, et al., CellMarker: a manually curated resource of cell markers in human and mouse, Nucleic. Acids. Res. 47 (D1) (2019) D721–D7d8.
[22] X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H.A. Pliner, et al., Reversed graph embedding resolves complex single-cell trajectories, Nat. Methods 14 (10) (2017) 979–982.
[23] V. Tremaroli, F. Karlsson, M. Werling, M. Ståhlman, P. Kovatcheva-Datchary, T. Olbers, et al., Roux-en-Y gastric bypass and vertical banded gastroplasty induce long-term changes on the human gut microbiome contributing to fat mass regulation, Cell Metab 22 (2) (2015) 228–238.
[24] M. Reich, T. Liefeld, J. Gould, J. Lerner, P. Tamayo, J.P. Mesirov, GenePattern 2.0, Nat. Genet 38 (5) (2006) 500–501.
[25] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.
[26] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. USA 102 (43) (2005) 15545–15550.
[27] D.A. Hosack, G. Dennis Jr., B.T. Sherman, H.C. Lane, R.A Lempicki, Identifying biological themes within lists of genes with EASE, Genome Biol. 4 (10) (2003) R70.
[28] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J.P. Mesirov, P. Tamayo, The molecular signatures database (MSigDB) hallmark gene set collection, Cell Syst. 1 (6) (2015) 417–425.
[29] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic. Acids. Res. 28 (1) (2000) 27–30.
[30] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, et al., Gene ontology: tool for the unification of biology. The gene ontology consortium, Nat. Genet 25 (1) (2000) 25–29.
[31] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep 18 (1) (2017) 248–262.
[32] B. Xiao, L. Liu, A. Li, C. Xiang, P. Wang, H. Li, et al., Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma, Front. Oncol. 10 (2020) 607622.
[33] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457.
[34] A. Mora, I.M. Donaldson, iRefR: an R package to manipulate the iRefIndex consolidated protein interaction database, BMC Bioinf. 12 (2011) 455.
[35] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, et al., The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease, Science 313 (5795) (2006) 1929–1935.
[36] W. Yang, J. Soares, P. Greninger, E.J. Edelman, H. Lightfoot, S. Forbes, et al., Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells, Nucleic. Acids. Res. 41 (Database issue) (2013) D955–D961.
[37] J.D. Buenrostro, P.G. Giresi, L.C. Zaba, H.Y. Chang, W.J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position, Nat. Methods 10 (12) (2013) 1213–1218.
[38] Y.Y. Wang, H. Jiang, J. Pan, X.R. Huang, Y.C. Wang, H.F. Huang, et al., Macrophage-to-Myofibroblast Transition Contributes to Interstitial Fibrosis in Chronic Renal Allograft Injury, J. Am. Soc. Nephrol. 28 (7) (2017) 2053–2067.
[39] A.M. Gilfillan, J. Rivera, The tyrosine kinase network regulating mast cell activation, Immunol. Rev. 228 (1) (2009) 149–169.
[40] J.K. Niehaus, B. Taylor-Blake, L. Loo, J.M. Simon, M.J. Zylka, Spinal macrophages resolve nociceptive hypersensitivity after peripheral injury, Neuron 109 (8) (2021) 1274. -82.e6.
[41] D. Xu, Z. Liu, M.X. Liang, W.Q. Chen, Y.J. Fei, S.J. Yang, et al., Hyperthermia promotes M1 polarization of macrophages via exosome-mediated HSPB8 transfer in triple negative breast cancer, Discov Oncol 14 (1) (2023) 81.
[42] P.S. Liu, Y.T. Chen, X. Li, P.C. Hsueh, S.F. Tzeng, H. Chen, et al., CD40 signal rewires fatty acid and glutamine metabolism for stimulating macrophage anti-tumorigenic functions, Nat. Immunol. 24 (3) (2023) 452–462.
[43] A. Yim, C. Smith, A.M. Brown, Osteopontin/secreted phosphoprotein-1 harnesses glial-, immune-, and neuronal cell ligand-receptor interactions to sense and regulate acute and chronic neuroinflammation, Immunol. Rev. 311 (1) (2022) 224–233.
[44] K. Schuch, B. Wanko, K. Ambroz, A. Castelo-Rosa, V. Moreno-Viedma, N.G. Grün, et al., Osteopontin affects macrophage polarization promoting endocytic but not inflammatory properties, Obesity (Silver Spring) 24 (7) (2016) 1489–1498.
[45] E.V. Fuior, A.V. Gafencu, Apolipoprotein C1: its Pleiotropic Effects in Lipid Metabolism and Beyond, Int. J. Mol. Sci. 20 (23) (2019).
[46] M. Abe, [Complement activation and inflammation], Rinsho Byori 54 (7) (2006) 744–756.

[47] M. Katayama, K. Ota, N. Nagi-Miura, N. Ohno, N. Yabuta, H. Nojima, et al., Ficolin-1 is a promising therapeutic target for autoimmune diseases, Int. Immunol. 31 (1) (2019) 23–32.
[48] M.M. Zegeye, M. Lindkvist, K. Fälker, A.K. Kumawat, G. Paramel, M. Grenegård, et al., Activation of the JAK/STAT3 and PI3K/AKT pathways are crucial for IL-6 trans-signaling-mediated pro-inflammatory response in human vascular endothelial cells, Cell Commun. Signal. 16 (1) (2018) 55.
[49] C.M. Hanes, A.E. D'Amico, T. Ueyama, A.C. Wong, X. Zhang, W.F. Hynes, et al., Golgi-Associated Protein Kinase C-ε Is Delivered to Phagocytic Cups: role of Phosphatidylinositol 4-Phosphate, J. Immunol. 199 (1) (2017) 271–277.
[50] Y. Tang, M. Jiang, A. Chen, W. Qu, X. Han, J. Zuo, et al., Porcupine inhibitor LGK-974 inhibits Wnt/β-catenin signaling and modifies tumor-associated macrophages resulting in inhibition of the malignant behaviors of non-small cell lung cancer cells, Mol Med Rep 24 (2) (2021).
[51] Y. Yang, Y.C. Ye, Y. Chen, J.L. Zhao, C.C. Gao, H. Han, et al., Crosstalk between hepatic tumor cells and macrophages via Wnt/β-catenin signaling promotes M2-like macrophage polarization and reinforces tumor malignant behaviors, Cell Death. Dis. 9 (8) (2018) 793.
[52] L. He, J.H. Jhong, Q. Chen, K.Y. Huang, K. Strittmatter, J. Kreuzer, et al., Global characterization of macrophage polarization mechanisms and identification of M2-type polarization inhibitors, Cell Rep. 37 (5) (2021) 109955.
[53] J. Tang, C. Gui, S. Qiu, M. Wang, The clinicopathological significance of Ki67 in papillary thyroid carcinoma: a suitable indicator? World J Surg Oncol 16 (1) (2018) 100.
[54] Y.J. Chen, R.M. Zhao, Q. Zhao, B.Y. Li, Q.Y. Ma, X. Li, et al., Diagnostic significance of elevated expression of HBME-1 in papillary thyroid carcinoma, Tumour Biol. 37 (7) (2016) 8715–8720.
[55] R.A. Whitmer, E.P. Gunderson, E. Barrett-Connor, C.P. Quesenberry Jr., K Yaffe, Obesity in middle age and future risk of dementia: a 27 year longitudinal population based study, BMJ 330 (7504) (2005) 1360.
[56] A. Mantovani, F. Marchesi, A. Malesci, L. Laghi, P. Allavena, Tumour-associated macrophages as treatment targets in oncology, Nat. Rev. Clin. Oncol. 14 (7) (2017) 399–416.
[57] L. Cassetta, J.W. Pollard, Targeting macrophages: therapeutic approaches in cancer, Nat. Rev. Drug Discov. 17 (12) (2018) 887–904.
[58] M. Locati, G. Curtale, A. Mantovani, Diversity, Mechanisms, and Significance of Macrophage Plasticity, Annu Rev Pathol 15 (2020) 123–147.
[59] I. Kryczek, L. Zou, P. Rodriguez, G. Zhu, S. Wei, P. Mottram, et al., B7-H4 expression identifies a novel suppressive macrophage population in human ovarian carcinoma, J. Exp. Med. 203 (4) (2006) 871–881.
[60] I.S. Kim, Y. Gao, T. Welte, H. Wang, J. Liu, M. Janghorban, et al., Immuno-subtyping of breast cancer reveals distinct myeloid cell profiles and immunotherapy resistance mechanisms, Nat. Cell Biol. 21 (9) (2019) 1113–1126.
[61] L. Zhang, Z. Li, K.M. Skrzypczynska, Q. Fang, W. Zhang, S.A. O'Brien, et al., Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer, Cell 181 (2) (2020) 442. -59.e29.
[62] D.Y. Vogel, J.E. Glim, A.W. Stavenuiter, M. Breur, P. Heijnen, S. Amor, et al., Human macrophage polarization in vitro: maturation and activation methods compared, Immunobiology 219 (9) (2014) 695–703.
[63] O. Helm, J. Held-Feindt, E. Grage-Griebenow, N. Reiling, H. Ungefroren, I. Vogel, et al., Tumor-associated macrophages exhibit pro- and anti-inflammatory properties by which they impact on pancreatic tumorigenesis, Int. J. Cancer 135 (4) (2014) 843–861.
[64] F. Kratochvill, G. Neale, J.M. Haverkamp, L.A. Van de Velde, A.M. Smith, D. Kawauchi, et al., TNF Counterbalances the Emergence of M2 Tumor Macrophages, Cell Rep. 12 (11) (2015) 1902–1914.
[65] M.J. Stables, S. Shah, E.B. Camon, R.C. Lovering, J. Newson, J. Bystrom, et al., Transcriptomic analyses of murine resolution-phase macrophages, Blood 118 (26) (2011) e192–e208.
[66] A.K. Shalek, R. Satija, J. Shuga, J.J. Trombetta, D. Gennert, D. Lu, et al., Single-cell RNA-seq reveals dynamic paracrine control of cellular variation, Nature 510 (7505) (2014) 363–369.
[67] S. Saha, I.N. Shalova, S.K. Biswas, Metabolic regulation of macrophage phenotype and function, Immunol. Rev. 280 (1) (2017) 102–111.
[68] F.O. Martinez, S. Gordon, M. Locati, A. Mantovani, Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression, J. Immunol. 177 (10) (2006) 7303–7311.
[69] J.I. Odegaard, A. Chawla, Alternative macrophage activation and metabolism, Annu Rev Pathol 6 (2011) 275–297.
[70] L. Zhu, Q. Zhao, T. Yang, W. Ding, Y. Zhao, Cellular metabolism and macrophage functional polarization, Int. Rev. Immunol. 34 (1) (2015) 82–100.
[71] V. Kandula, R. Kosuru, H. Li, D. Yan, Q. Zhu, Q. Lian, et al., Forkhead box transcription factor 1: role in the pathogenesis of diabetic cardiomyopathy, Cardiovasc Diabetol 15 (2016) 44.
[72] Y.Q. Xing, A. Li, Y. Yang, X.X. Li, L.N. Zhang, H.C. Guo, The regulation of FOXO1 and its role in disease progression, Life Sci. 193 (2018) 124–131.
[73] D. Yan, Y. Cai, J. Luo, J. Liu, X. Li, F. Ying, et al., FOXO1 contributes to diabetic cardiomyopathy via inducing imbalanced oxidative metabolism in type 1 diabetes, J. Cell. Mol. Med. 24 (14) (2020) 7850–7861.
[74] Z. Li, Y. Peng, J. Li, Z. Chen, F. Chen, J. Tu, et al., N(6)-methyladenosine regulates glycolysis of cancer cells through PDK4, Nat. Commun. 11 (1) (2020) 2578.
[75] J. Chen, Y. Lu, M. Tian, Q. Huang, Molecular mechanisms of FOXO1 in adipocyte differentiation, J. Mol. Endocrinol. 62 (3) (2019) R239–Rr53.
[76] A.W. Woll, F.W. Quelle, C.D. Sigmund, PPARγ and retinol binding protein 7 form a regulatory hub promoting antioxidant properties of the endothelium, Physiol. Genomics 49 (11) (2017) 653–658.
[77] D.H. Kim, J. Ahn, Y. Suh, O. Ziouzenkova, J.W. Lee, K. Lee, Retinol Binding Protein 7 Promotes Adipogenesis in vitro and Regulates Expression of, Genes Involved in Retinol Metabolism. Front Cell Dev Biol. 10 (2022) 876031.
[78] G.S. Getz, C.A. Reardon, Apoprotein E and Reverse Cholesterol Transport, Int. J. Mol. Sci. 19 (11) (2018).
[79] X. Bian, R. Liu, Y. Meng, D. Xing, D. Xu, Z. Lu, Lipid metabolism and cancer, J. Exp. Med. 218 (1) (2021).
[80] M. Song, A.M. Bode, Z. Dong, M.H Lee, AKT as a Therapeutic Target for Cancer, Cancer Res. 79 (6) (2019) 1019–1031.
[81] J. Feng, S. Qiu, S. Zhou, Y. Tan, Y. Bai, H. Cao, et al., mTOR: a Potential New Target in Nonalcoholic Fatty Liver Disease, Int. J. Mol. Sci. 23 (16) (2022).
[82] Z.A. Knight, B. Gonzalez, M.E. Feldman, E.R. Zunder, D.D. Goldenberg, O. Williams, et al., A pharmacological map of the PI3-K family defines a role for p110alpha in insulin signaling, Cell 125 (4) (2006) 733–747.
[83] A. Eisenberg-Lerner, A. Kimchi, The paradox of autophagy and its implication in cancer etiology and therapy, Apoptosis 14 (4) (2009) 376–391.
[84] S. Abdollah, M. Macías-Silva, T. Tsukazaki, H. Hayashi, L. Attisano, J.L. Wrana, TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling, J. Biol. Chem. 272 (44) (1997) 27678–27685.
[85] K. Kiyono, H.I. Suzuki, H. Matsuyama, Y. Morishita, A. Komuro, M.R. Kano, et al., Autophagy is activated by TGF-beta and potentiates TGF-beta-mediated growth inhibition in human hepatocellular carcinoma cells, Cancer Res. 69 (23) (2009) 8844–8852.
[86] S. Lorzadeh, L. Kohan, S. Ghavami, N. Azarpira, Autophagy and the Wnt signaling pathway: a focus on Wnt/β-catenin signaling, Biochim. Biophys. Acta. Mol. Cell Res. 1868 (3) (2021) 118926.
[87] F. Ascenzi, C. De Vitis, M. Maugeri-Saccà, C. Napoli, G. Ciliberto, R. Mancini, SCD1, autophagy and cancer: implications for therapy, J. Exp. Clin. Cancer Res. 40 (1) (2021) 265.
[88] J.H. Wen, D.Y. Li, S. Liang, C. Yang, J.X. Tang, H.F. Liu, Macrophage autophagy in macrophage polarization, chronic inflammation and organ fibrosis, Front. Immunol. 13 (2022) 946832.
[89] F.J. Gonzalez, C. Xie, C. Jiang, The role of hypoxia-inducible factors in metabolic diseases, Nat. Rev. Endocrinol. 15 (1) (2018) 21–32.
[90] Y. Liu, S.R. Cox, T. Morita, S. Kourembanas, Hypoxia regulates vascular endothelial growth factor gene expression in endothelial cells. Identification of a 5′ enhancer, Circ. Res. 77 (3) (1995) 638–643.
[91] M. Arjaans, C.P. Schröder, S.F. Oosting, U. Dafni, J.E. Kleibeuker, E.G. de Vries, VEGF pathway targeting agents, vessel normalization and tumor drug uptake: from bench to bedside, Oncotarget 7 (16) (2016) 21247–21258.
[92] V.W. van Hinsbergh, M.A. Engelse, P.H. Quax, Pericellular proteases in angiogenesis and vasculogenesis, Arterioscler. Thromb. Vasc. Biol. 26 (4) (2006) 716–728.
[93] Q. Zhang, Y. Lou, J. Zhang, Q. Fu, T. Wei, X. Sun, et al., Hypoxia-inducible factor-2α promotes tumor progression and has crosstalk with Wnt/β-catenin signaling in pancreatic cancer, Mol. Cancer 16 (1) (2017) 119.
[94] S.Y. Zhang, S. Boisson-Dupuis, A. Chapgier, K. Yang, J. Bustamante, A. Puel, et al., Inborn errors of interferon (IFN)-mediated immunity in humans: insights into the respective roles of IFN-alpha/beta, IFN-gamma, and IFN-lambda in host defense, Immunol. Rev. 226 (2008) 29–40.
[95] F. Wang, S. Zhang, R. Jeon, I. Vuckovic, X. Jiang, A. Lerman, et al., Interferon Gamma Induces Reversible Metabolic Reprogramming of M1 Macrophages to Sustain Cell Viability and Pro-Inflammatory Activity, EBioMedicine 30 (2018) 303–316.
[96] S. Pagie, N. Gérard, B. Charreau, Notch signaling triggered via the ligand DLL4 impedes M2 macrophage differentiation and promotes their apoptosis, Cell Commun. Signal. 16 (1) (2018) 4.
[97] X. Xu, J. Xu, J. Wu, Y. Hu, Y. Han, Y. Gu, et al., Phosphorylation-Mediated IFN-γR2 Membrane Translocation Is Required to Activate Macrophage Innate Response, Cell 175 (5) (2018) 1336. -51.e17.
[98] X. Zhou, W. Li, S. Wang, P. Zhang, Q. Wang, J. Xiao, et al., YAP Aggravates Inflammatory Bowel Disease by Regulating M1/M2 Macrophage Polarization and Gut Microbial Homeostasis, Cell Rep. 27 (4) (2019) 1176. -89.e5.
[99] C. Liu, F. Hu, G. Jiao, Y. Guo, P. Zhou, Y. Zhang, et al., Dental pulp stem cell-derived exosomes suppress M1 macrophage polarization through the ROS-MAPK-NFκB P65 signaling pathway after spinal cord injury, J Nanobiotechnology 20 (1) (2022) 65.
[100] X. Zhang, Y. Wang, J. Yuan, N. Li, S. Pei, J. Xu, et al., Macrophage/microglial Ezh2 facilitates autoimmune inflammation through inhibition of Socs3, J. Exp. Med. 215 (5) (2018) 1365–1382.
[101] Y. Tokumaru, M. Oshi, E. Katsuta, L. Yan, V. Satyananda, N. Matsuhashi, et al., KRAS signaling enriched triple negative breast cancer is associated with favorable tumor immune microenvironment and better survival, Am J Cancer Res 10 (3) (2020) 897–907.
[102] R. Salehi, M. Asare-Werehene, B.A. Wyse, A. Abedini, B. Pan, A. Gutsol, et al., Granulosa cell-derived miR-379-5p regulates macrophage polarization in polycystic ovarian syndrome, Front. Immunol. 14 (2023) 1104550.