CANCER

Characterizing dedifferentiation of thyroid cancer by integrated analysis

Han Luo[1†], Xuyang Xia[2†], Gyeong Dae Kim[3†], Yang Liu[1†], Zhinan Xue[2], Lingyun Zhang[1], Yang Shu[2,4], Tian Yang[5], Yan Chen[2], Shouyue Zhang[2], Haining Chen[4], Weihan Zhang[4], Ruicen Li[6], Huairong Tang[6], Birong Dong[7], Xianghui Fu[2], Wei Cheng[2], Wei Zhang[8], Li Yang[2], Yong Peng[2], Lunzhi Dai[2], Hongbo Hu[2], Yong Jiang[5], Changyang Gong[2], Yiguo Hu[1,2], Jingqiang Zhu[1], Zhihui Li[1], Carlos Caulin[9], Tao Wei[1*], Jihwan Park[3*], Heng Xu[2,10*]

Understanding of dedifferentiation, an indicator of poo prognosis for patients with thyroid cancer, has been hampered by imprecise and incomplete characterization of its heterogeneity and its attributes. Using single-cell RNA sequencing, we explored the landscape of thyroid cancer at single-cell resolution with 46,205 cells and delineated its dedifferentiation process and suppressive immune microenvironment. The developmental trajectory indicated that anaplastic thyroid cancer (ATC) cells were derived from a small subset of papillary thyroid cancer (PTC) cells. Moreover, a potential functional role of *CREB3L1* on ATC development was revealed by integrated analyses of copy number alteration and transcriptional regulatory network. Multiple genes in differentiation-related pathways (e.g., EMT) were involved as the downstream targets of CREB3L1, increased expression of which can thus predict higher relapse risk of PTC. Collectively, our study provided insights into the heterogeneity and molecular evolution of thyroid cancer and highlighted the potential driver role of *CREB3L1* in its dedifferentiation process.

INTRODUCTION

Thyroid cancer is the most common endocrine malignancy (*1*). Multiple histological types of thyroid cancer have been described, including papillary thyroid cancer (PTC) and anaplastic thyroid cancer (ATC). PTC, the well-differentiated type, accounts for more than 80% of all thyroid cancers and has a clinically favorable prognosis with more than 90% 10-year disease-specific survival (*2*). In contrast, ATC is an undifferentiated type that exhibits stem cell–like properties, highly proliferative potential, and resistance to current therapies (*3*). As a result, ATC is one of the most lethal tumors, with a median survival of only 6 to 8 months, and it is responsible for nearly half of the thyroid cancer–related deaths (*4*). Histologically, PTC progresses to ATC by dedifferentiation, which is a biological process in pervasive cancers that induces the transition of cancer from highly differentiated to poorly differentiated status (*5*).

Cellular plasticity endows cancer cells with the ability to shift between differentiated and undifferentiated state, and thus contributes to dedifferentiation in cancer (*6*, *7*). This fluctuation may involve varied genetic alterations in different cancer types. For thyroid cancer, although *BRAF*[V600E] mutation frequently occurred in both ATC and PTC, common mutations in *TP53* and the *TERT* promoter that present in ATC are rare in PTC (*8*–*11*). It suggests that the accumulation of genetic alterations may facilitate PTC to dedifferentiate into ATC by influencing cellular plasticity. In addition, the derivation of ATC from PTC is also suggested by other indirect evidence, including frequent histopathologic coexistence of PTC components in ATC lesions and history of PTC in most ATC patients (*4*). In contrast, with whole-exome sequencing (WES) of microdissected PTC and ATC lesions of the concomitant cases, somatic mutation–based genomic evolution analysis revealed that ATC lesions diverged from PTC lesions in the early phase of tumor development (*12*). However, the large genetic distance between ATC and PTC components from the same patient cannot rule out the possibility that ATC cells may derive from a subset of PTC cells through genetic and/or epigenetic mechanisms. Collectively, no direct evidence has indicated whether or not ATC evolves from PTC at the cellular level.

The accurate identification and characterization of individual cell states through single-cell RNA sequencing (scRNA-seq) provides opportunities to detect fluctuating cell fate in different cell types continuously and uncover the gene regulatory network in cancer development (*13*, *14*). Therefore, scRNA-seq makes it possible to investigate intratumor heterogeneity and cellular plasticity at the single-cell level, which is beneficial for understanding and unveiling the dedifferentiation of thyroid cancer cells and development of ATC (*15*). In this study, we generated scRNA-seq transcriptomes to characterize the cellular composition of PTC and ATC and performed integrated analysis to investigate the shared and distinct features at single-cell resolution. Moreover, we aim to use single-cell profiling to delineate the molecular evolutionary trajectory from PTC to ATC and to identify the key genomic events involved in this transition.

[1]Department of Thyroid and Parathyroid Surgery, Laboratory of Thyroid and Parathyroid Disease, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, Sichuan, China. [2]State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University and Collaborative Innovation Center, Chengdu, Sichuan, China. [3]School of Life Sciences, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea. [4]Department of Gastrointestinal Surgery, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, Sichuan, China. [5]Department of Pathology, West China Hospital, Sichuan University, Chengdu, Sichuan, China. [6]Health Promotion Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China. [7]National Clinical Research Center for Geriatrics, Chengdu, Sichuan, China. [8]Department of Clinical Pharmacology, Hunan Key Laboratory of Pharmacogenetics, Xiangya Hospital, Central South University, Changsha, Hunan, China. [9]Department of Otolaryngology—Head & Neck Surgery and University of Arizona Cancer Center, University of Arizona, Tucson, AZ, USA. [10]Department of Laboratory Medicine, Precision Medicine Center, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, Sichuan, China.
*Corresponding author. Email: xuheng81916@scu.edu.cn (H.X.); jihwan.park@gist.ac.kr (J.P.); surgeonwei@163.com (T.W.)
†These authors contributed equally to this work.

RESULTS

scRNA-seq analysis and cellular profile of thyroid cancer

To explore the cellular composition of thyroid cancers, we profiled clinically annotated PTC/ATC tumor and normal thyroid tissue from one of the PTC patients by scRNA-seq and WES (Fig. 1A and tables S1 and S2). After quality control (fig. S1A and table S3), we acquired the transcriptomes of 46,205 single cells and conducted differential gene expression analysis to identify cluster-specific markers (table S4). In total, 16 main clusters were identified, including 3 clusters for thyroid cancer cells (e.g., PTC and ATC), 2 for fibroblasts, and 8 for immune cells (e.g., B cell, T cell, and macrophage) (Fig. 1B). In most cases, well-established canonical markers can generally distinguish different clusters, such as *TPO* for normal follicular cells, *CALCA* for parafollicular cells, *CD3D* for T cells, *RAMP2* for endothelial cells, *ACTA2* for fibroblasts, and *CD68* for macrophages (Fig. 1C and table S4). Certain cell types can be further separated into subgroups by specific markers, such as *FAP* for fibroblasts and *S100B/CD163* for macrophages (Fig. 1C). For thyroid tissues, normal follicular cells can be defined by highly expressed *TG*, expression of which decreased markedly during thyroid cancer progression, as previously reported (Fig. 1D and fig. S1B). Intriguingly, PTC cells were further divided into two main clusters (referred to as PTC1 and PTC2) (Fig. 1B), identified by a subset of differentially expressed genes (DEGs), including *NEAT1, SLC34A2, S100A13*, and *MDK* (Fig. 1, D and E), suggesting heterogeneity of PTC cells. Although both PTC and ATC cells expressed epithelium-related genes (e.g., *KRT19*) (Fig. 1C), they can be separated by several markers. Consistent with previous reports (*16*), *PBK* was specifically expressed in ATC (Fig. 1D), which was validated by immunohistochemistry (IHC) (Fig. 1F). In addition, ATC and PTC cells can also be discriminated by the high expression of markers for proliferation (e.g., *TOP2A*) in ATC cells (Fig. 1, C and D). Substantial cell proportion changes in thyroid cancer samples were observed. As expected, normal follicular cells were depleted in tumor samples, while PTC and ATC cells were highly enriched in PTC and ATC samples, respectively. For nontumor cells, fibroblasts and immune cells were increased in our PTC and ATC samples (Fig. 1G). This variation in cell proportion clustering was verified by deconvolution analysis using bulk transcriptome profiles of thyroid cancer tissues (45 normal, 49 PTC, and 11 ATC samples) (Fig. 1G) (*17*). Intriguingly, endothelium cells were nearly absent in ATC samples, which was validated by IHC staining with CD31 (fig. S1C).

To describe the tumor microenvironment, ATC-derived natural killer (NK) and T cells highly expressed lymphoid dysfunctional markers (e.g., *PDCD1, CTLA4, LAG3*, and *TIGIT*) (fig. S2A), while regulatory T cells (*FOXP3*) and tumor-associated myeloid cells (TAMCs) (*SIGLEC15*) were nearly exclusively enriched in ATC (fig. S2B). In addition, according to ssGSEA (single-sample gene set enrichment analysis) enrichment score of well-established signatures for M1- and M2-polarizated tumor-associated macrophage (TAM) (*18*), macrophages 1 and 2 were defined as M1 and M2 TAM, respectively (fig. S3A). A significantly decreased M1/M2 ratio was observed in ATC (fig. S3B), suggesting tumor progression and immunosuppression. Because *FAP*[+] fibroblasts promote an immunosuppressive tumor microenvironment (*19, 20*), the lower Fibroblast1 (*FAP*[−])/Fibroblast2 (*FAP*[+]) ratio in ATC suggested suppression of antitumor immunity in ATC (Fig. 1C and fig. S3B). We next evaluated the interactions of thyroid cells with other cell types. By CellphoneDB analysis, fibroblast2 had a specific strong interaction

with ATC cells among cancer cells and intended to interact with M2 and TAMC (fig. S3C). All of the evidence suggested an immunosuppressive microenvironment in ATC.

Dynamic transcriptional changes in thyroid cancer dedifferentiation

To gain insight into the cellular progression of PTC to ATC, we first divided PTC cells into PTC-derived PTC cells (pPTC) and ATC-derived PTC cells (aPTC) (Fig. 2A). To explore divergent characteristics of different PTC cells, we performed subclustering analysis for all PTC cells and identified 10 subclusters (Fig. 2B). Intriguingly, most aPTC cells aggregated in subcluster 10 (Fig. 2B), and PTC2 cells matched to subcluster 6 (fig. S4A). The evidence indicated that aPTC cells were a distinct subset of PTC1.

To determine the development of ATC cells, we performed cell trajectory analysis, revealing that ATC and PTC had different progression routes (Fig. 2C). aPTC cells localized on the ATC progression trajectory and a subset of them close to ATC cells after the bifurcation, indicating the similarity of aPTC cells with ATC cells and its potential association with ATC development (Fig. 2C). Next, we investigated canonical gene expression and the well-established thyroid cancer–related score along the trajectory. By ATC progression route, thyroid differentiation score (TDS) (*9*) was markedly reduced, while genes in extracellular signal–regulated kinase (ERK) and phosphatidylinositol 3-kinase (PI3K)–mammalian target of rapamycin (mTOR) signaling pathways were primarily activated in ATC cells (Fig. 2D and fig. S4B); the previous two of which can be affected by mutation status of *BRAF*[V600E] (fig. S4C). As expected, follicular canonical genes (e.g., *TPO, TG*, and *TFF3*) were expressed at the initial stage (fig. S4D), while ATC-specific markers (e.g., *PBK*) localized at the end of the ATC progression route (Fig. 2E). Intriguingly, the PTC-related genes (e.g., *LGALS3, NPC2*, and *S100A13*) were expressed along not only PTC progression but also ATC progression (Fig. 2E and fig. S4D), which was consistent with the presence of aPTC cells on the ATC progression trajectory (Fig. 2C). Experimentally, a subset of ATC cells (tagged as *PBK*[+]) can also be stained with PTC-related pathologic marker of galectin-3 (Fig. 2F) through immunofluorescence (IF), further supporting the possibility that ATC may originate from a subset of PTC cells (i.e., aPTC). Along with the PTC to ATC cell trajectory, we identified expression changes of many well-known genes that are involved in cancer development (e.g., *VIM, TWIST1*, and *COL1A1*) and different activation of several cancer-associated pathways. Despite the enrichment of epithelial-mesenchymal transition (EMT) and mTORC1 (Fig. 2, G and H), negative regulation of p53 was also identified along ATC development, which was tagged by several regulators (e.g., *MIF, CD44, SNAI2*, and *TWIST1*) rather than *TP53* expression itself (fig. S4D) (*21, 22*).

Together, these findings indicated that follicular cells progressed to PTC or ATC following different evolution routes, involving specific cell subcluster and activation of multiple pathways.

CNAs in thyroid cancer dedifferentiation

To identify the differences of genomic profile between ATC and PTC cells, we first inferred copy number alterations (CNAs) through the single cell–based approach (*23*). A higher burden of CNAs was identified in ATC cells than in PTC cells, indicating increased chromosome instability in ATC cells (Fig. 3A and fig. S5A). For PTC cells, PTC1 and PTC2 cells from each PTC patient shared similar CNA patterns, suggesting a common genomic evolution (fig. S5B).
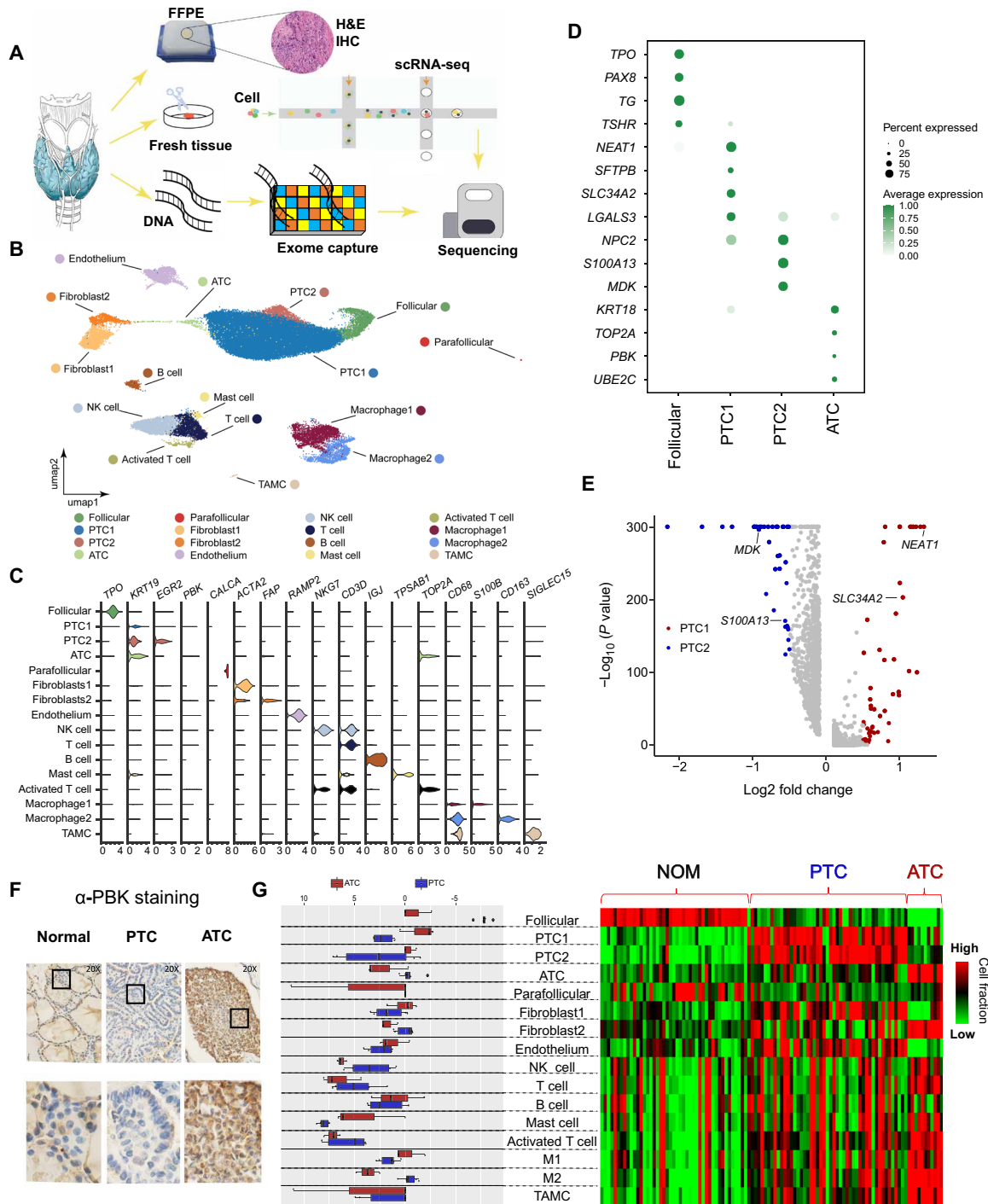
Fig. 1. Single-cell profile of thyroid cancers. (**A**) Overview of the experiment procedure. Ten scRNA-seq data were generated from normal, PTC, and ATC tissues using 10× Genomics protocol. We analyzed the transcriptome of 46,205 individual cells. In parallel, WES was performed. FFPE, formalin-fixed, paraffin-embedded; H&E, hematoxylin and eosin. (**B**) UMAP demonstrates the 16 clusters. Assigned cell types were labeled on the figure. TAMC, tumor-associated myeloid cell. (**C**) Violin plots showing the normalized expression level of cell type–specific markers. (**D**) Bubble plot showing the genes that were uniquely expressed in normal follicular and each cancer cell type. The degree of color represents the average expression value, and the size of dot represents the expression percentage in each cluster. (**E**) Volcano plot showing the DEGs between two types of PTC clusters (PTC1 and PTC2). The x axis represents $\log_2$ fold change, and the y axis represents $-\log_{10}$ P value. (**F**) Representative IHC staining for PBK in normal, PTC, and ATC tissues. (**G**) Left: Cell type proportion of cancer samples compared to a normal sample. The x axis represents each cell type, and the y axis represents the relative cell type proportion fold change compared to the normal sample. Blue and red boxes indicate PTC and ATC, respectively. Right: Deconvolution of bulk thyroid RNA-seq data (GSE33630) consisting of 45 normal samples, 49 PTC samples, and 11 ATC samples by using scRNA-seq data. The color indicates relative cell fraction between the samples.
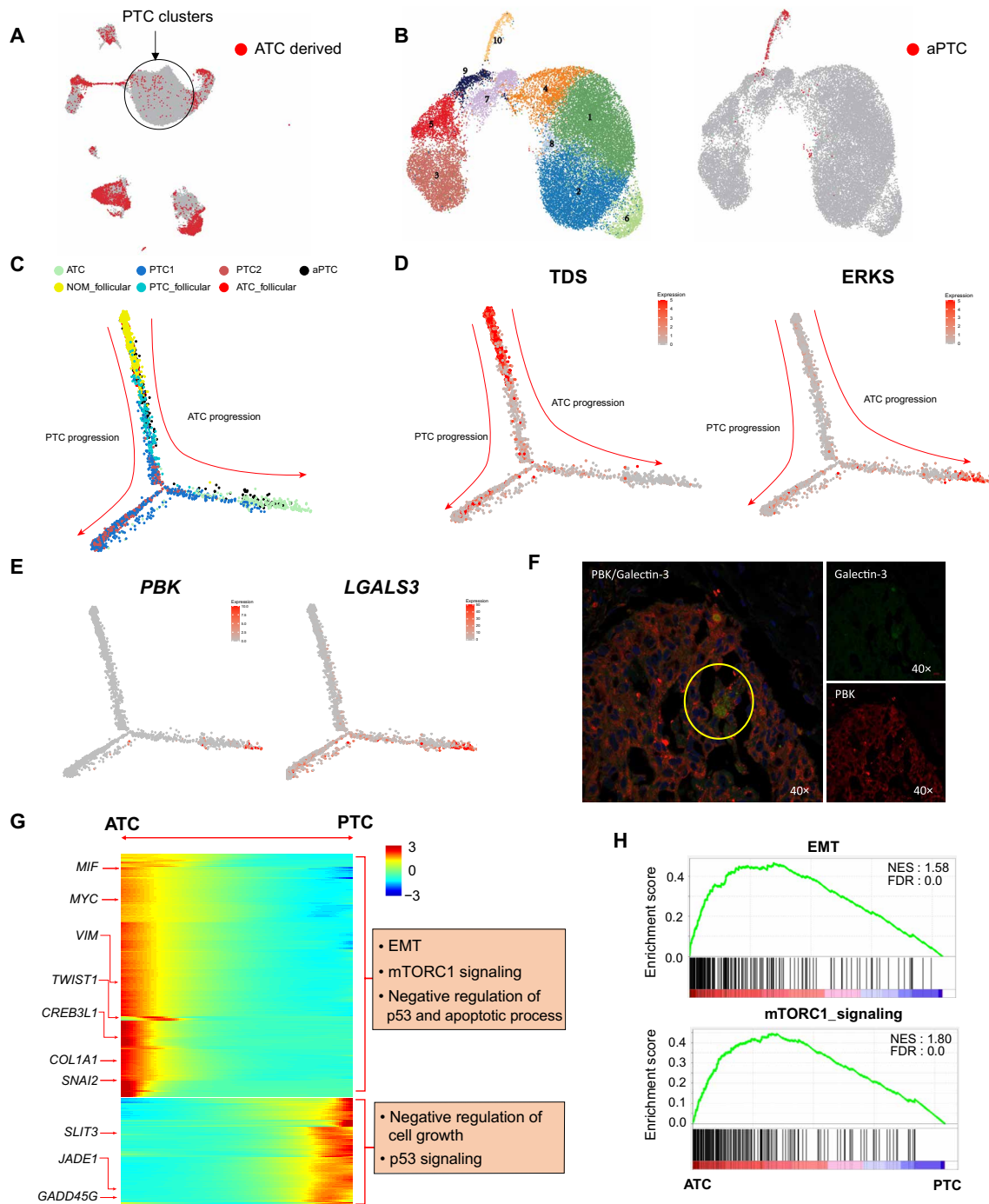
Fig. 2. Evolution trajectory and transcriptional fluctuation during dedifferentiation of thyroid cancer. (**A**) Distribution pattern of the cells derived from ATC samples. Red color indicates the cells derived from the ATC, and the circle indicates PTC clusters. (**B**) The t-Stochastic Neighbor Embedding (T-SNE) plot showing the 10 subclusters of PTC cell clusters (PTC1 and PTC2). On the left side, clusters of PTC cells were divided into 10 subclusters and each color represents each subcluster. On the right side, the cells derived from ATC samples are marked. Red color indicates the cells derived from the ATC samples. (**C**) Cell trajectory analysis of tumor progression from normal to cancer cells. Cells are labeled with each cell cluster: Follicular cells derived from normal (NOM_follicular), PTC (PTC_ Follicular), and ATC (ATC_ Follicular) samples and PTC1, PTC2, aPTC, and ATC cells. (**D**) Levels of TDS and ERKS (ERK score) on the trajectory. TDS consists of 16 genes and ERKS consists of 51 genes. Each score is represented by the average expression of the genes corresponding to each pathway. Color scale represents expression level. (**E**) *PBK* and *LGALS3* expression levels on the trajectory. Color scale represents expression level. (**F**) IF staining showed a few ATC cells double-stained by PBK (red fluorescence) and galectin-3 (green fluorescence). Yellow circle indicates double-stained ATC cells. (**G**) Heatmap showing the expression changes of the 800 highly variable genes along the ATC-PTC axis of the trajectory. Significantly enriched functional annotations are shown on the right side of the heatmap. Color scale indicates *z* score. (**H**) Example GSEA plots showing the gene enrichment patterns of EMT and mTORC1 signaling genes on the ATC-PTC axis of the trajectory. Normalized enrichment score (NES) and false discovery rate (FDR) values are shown on the plots.
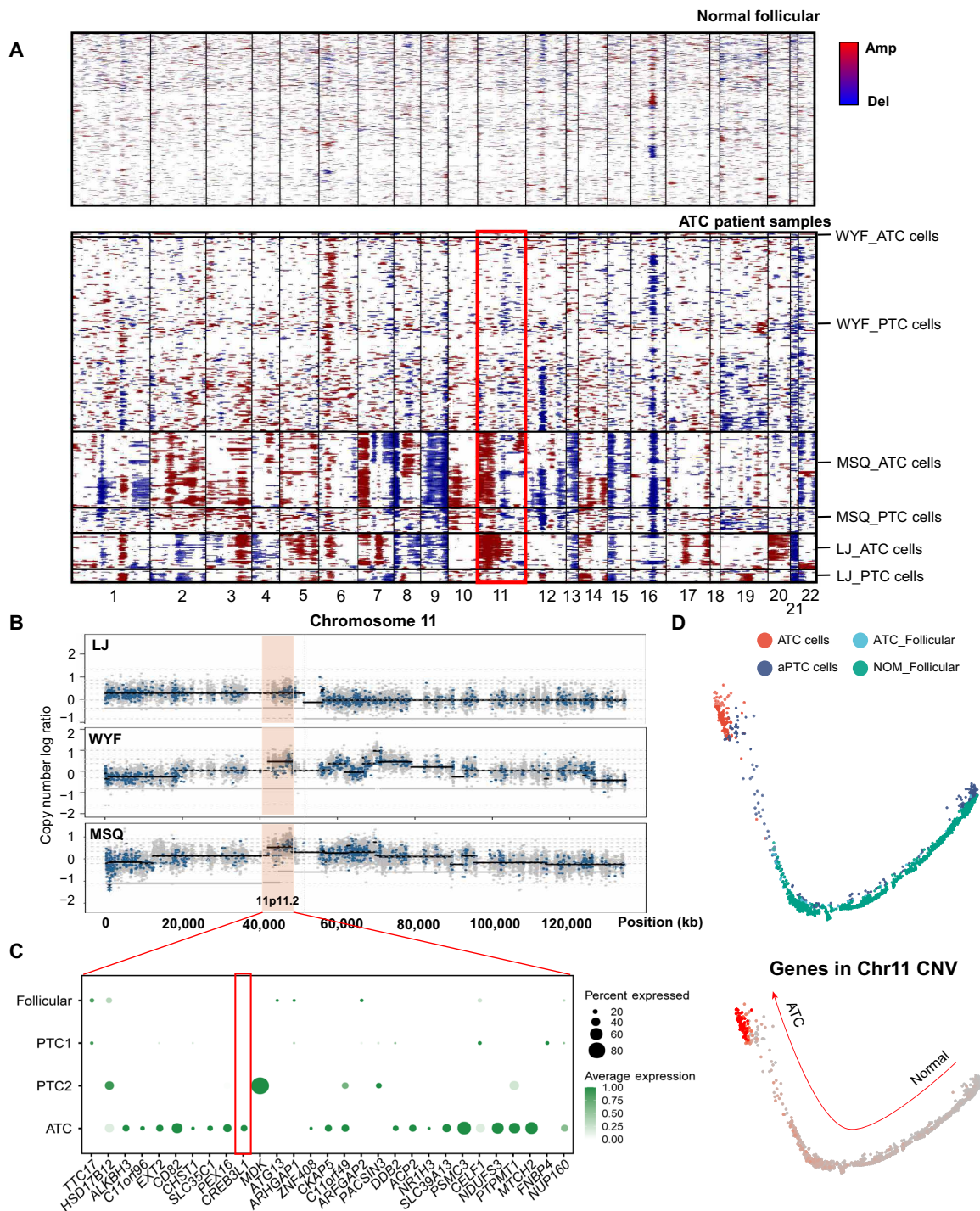
Fig. 3. CNAs during dedifferentiation of thyroid cancer. (**A**) Large-scale CNAs of ATC and normal follicular cells identified using scRNA-seq data. Normal follicular cells were used as a reference. The CNA patterns for the normal follicular cells, ATC cells, and PTC cells from three ATC patients are shown. *X* axis, chromosome position; *y* axis, individual cells. Red color represents amplification, while blue color represents deletion. (**B**) CNA plots showing the relative CNA patterns in chromosome 11 that were identified by both scRNA-seq and WES data. *X* axis, position; *y* axis, copy number. (**C**) Bubble plot showing expression of the amplified genes in the Chr11-CNA. The location of the gene set is marked by box in (B). Dot sizes and color indicate the expression percentage and average expression, respectively. (**D**) Top: Cell trajectory analysis of tumor progression from normal to ATC cells. Cells are labeled with each cell cluster. Bottom: Average expression level of the genes in Chr11 CNA on the trajectory. Color scale represents expression level.

ATC cells carried not only most CNAs found in aPTC from the same patient but also a large number of de novo CNAs, which may be acquired during the progression of aPTC to ATC (Fig. 3A). To validate and narrow down the relevant regions of interest, we inferred CNAs through the WES-based approach. Of the five large overlapping CNAs identified in ATC patients, CNA (41.1 to 47.9 M) in Chr11 (Chr11-CNA) was observed in all three ATC tumors (Fig. 3B and fig. S6), which was also observed in an independent large cohort (chromosome 11 amplification in 11 of 27 samples) (*24*). Bubble plots illustrated that most genes located in this CNA region were highly expressed in ATC cells compared with normal and PTC cells (Fig. 3C and fig. S6). Moreover, this Chr11-CNA began to increase in the midway of cell trajectory from normal to ATC cells and was highly enriched in ATC cells, but not in PTC cells (Fig. 3D), suggesting that a particular gene in this region may be involved in the ATC progression from aPTC.

CREB3L1 as a master transcriptional regulator in ATC

Because multiple pathways described above were associated with ATC progression, we speculated that an upstream transcription factor (TF) might function as a key regulator of ATC progression. Therefore, we performed SCENIC analysis with scRNA-seq data (*25*) to reconstruct coexpression modules and identify the key TF regulators during ATC progression. As a positive control, a well-established TF (i.e., *NKX2-1*, also known as thyroid transcription factor 1) was identified as a key regulator in both pPTC and aPTC cells, but not in ATC cells (Fig. 4A). Other TFs (e.g., *RXRG*) were enriched in pPTC cells, but not in aPTC cells, suggesting common and divergent regulatory networks in pPTC and aPTC cells. We also observed that ATC-derived follicular cells showed enhanced activities of several TFs (e.g., *FOS* and *JUN*) compared to normal follicular cells. These TFs were also enriched in aPTC cells, but largely absent in pPTC and ATC cells (Fig. 4A), suggesting their potential role in early ATC development.

CREB3L1 was enriched in ATC-derived thyroid cancer cells, particularly high in ATC cells, suggesting the association of *CREB3L1* expression with ATC progression. Most *CREB3L1*-positive cells were observed in ATC cells (Fig. 4B) and gradually enriched along with the follicular-aPTC-ATC trajectory (Fig. 4C). *CREB3L1* was located exactly in the ATC-associated Chr11-CNA described above (Fig. 3C), providing additional support for its potential role in ATC progression. Through IHC, CREB3L1 was absent or weakly stained in normal follicular and PTC but strongly positive in ATC (Fig. 4D and fig. S7A). Furthermore, positive staining of CREB3L1 was detected in both PTC and ATC components in cases of concomitant ATC and PTC, with stronger staining in the ATC component than the adjacent PTC (Fig. 4D). In addition, 17 of 28 available formalin-fixed, paraffin-embedded samples from patients with ATC were positively stained with CREB3L1, independently of common somatic mutations in driver genes (i.e., *TP53, BRAF*, and *TERT*) (fig. S7B) (*10*).

To investigate the functional relevance of *CREB3L1* in ATC development, we next characterized the target genes of CREB3L1 and their related pathways in ATC. CREB3L1-binding sites, identified by chromatin immunoprecipitation sequencing (ChIP-seq), were obtained from ENCODE (*26*). The binding regions of CREB3L1 were highly enriched in gene promoters (Fig. 4E). Considering both gene expression and ChIP-seq signals, 2222 genes were identified as the potential targets of CREB3L1, expression of which tends to increase along with ATC progression (Fig. 4F). Particularly, the CREB3L1

target genes were enriched in EMT and mTOR signaling in ATC cells compared to normal thyroid cells (Fig. 4G), which was consistent with pathway enrichment analysis of ATC transcriptomes generated by scRNA-seq (Fig. 2, G and H). As an example, CREB3L1-binding signals were localized near or within the gene loci of EMT markers (e.g., *COL1A1* and *MMP11*), expression of which increased along with the dedifferentiation process (Fig. 4H). Overall, these findings provided strong supports for a master transcriptional regulator role of *CREB3L1* in ATC development via the EMT and mTOR pathway.

Correlation of *CREB3L1*-driven pathways with thyroid cancer progression

The impact of *CREB3L1* on ATC development was further evaluated by GSEA on bulk RNA expression data from four independent publicly accessible thyroid cancer studies [accession GSE33630 (*17*), GSE27155 (*27*), GSE65144 (*28*), and GSE126698 (*29*)]. Consistently, expression of *CREB3L1* and genes involved in EMT/mTORC1 pathways was significantly enriched in ATC compared with normal samples (Fig. 5A and fig. S7C). Furthermore, to validate the fate determination role of *CREB3L1* in PTC patients, we divided thyroid cancer samples from The Cancer Genome Atlas (TCGA) into two groups according to their *CREB3L1* expression level. A total of 1115 DEGs were identified between the two groups, including up-regulation of EMT-related genes (e.g., *COL1A1* and *MMP11*) in *CREB3L1* high group (Fig. 5B). GSEA and hierarchical clustering analyses also confirmed the enrichment of EMT/mTOR signaling genes in *CREB3L1*-high tumors (Fig. 5, C and D). Notably, extremely high *CREB3L1* expression can predict significantly worse overall survival and disease-specific survival of PTC patients in TCGA (Fig. 5E), which was not biased by subtype of PTC (table S5). Next, we evaluated the predictive value of CREB3L1 IHC staining with PTC samples from the same center (recurrent group: 2-year clinically recurrent PTC patients; nonrecurrent group: 10-year disease-free PTC patients). The ratio of positive CREB3L1 staining in the recurrent group was significantly higher than that in the nonrecurrent group independent of pathological subtypes ($P = 0.006$) (Fig. 5F), which was consistent with the observation in TCGA cohort (Fig. 5E).

To further estimate the function of *CREB3L1*, we performed in vitro experiments by overexpressing *CREB3L1* in a PTC cell line (i.e., TPC-1^CREB3L1-OE), which significantly increased the migration ability ($P = 0.004$) (Fig. 5G), induced the expression of several EMT-related markers (e.g., Twist1, Vimentin, and Snail), and repressed the expression of E-cadherin that is negatively correlated with EMT (Fig. 5H). Besides, CREB3L1 staining was absent from thyroid metastasis of other cancers (e.g., lung adenocarcinoma, non-Hodgkin's lymphoma, and squamous cell carcinoma) (fig. S7D), which was hardly distinguished from ATC because of the similar clinical sign. Therefore, positive IHC staining of CREB3L1 can facilitate discrimination diagnosis of this lethal disease and guide the subsequent clinical treatment. Together, these findings suggested that *CREB3L1* expression may determine the cell fate of thyroid cancer and promote ATC progression and poor clinical outcomes for thyroid cancer patients.

DISCUSSION

Stepwise progression of cancer evolution has been reported in most human malignancies, including the presence of precancerous lesions that progress to malignant tumors (*30*). Dedifferentiation and
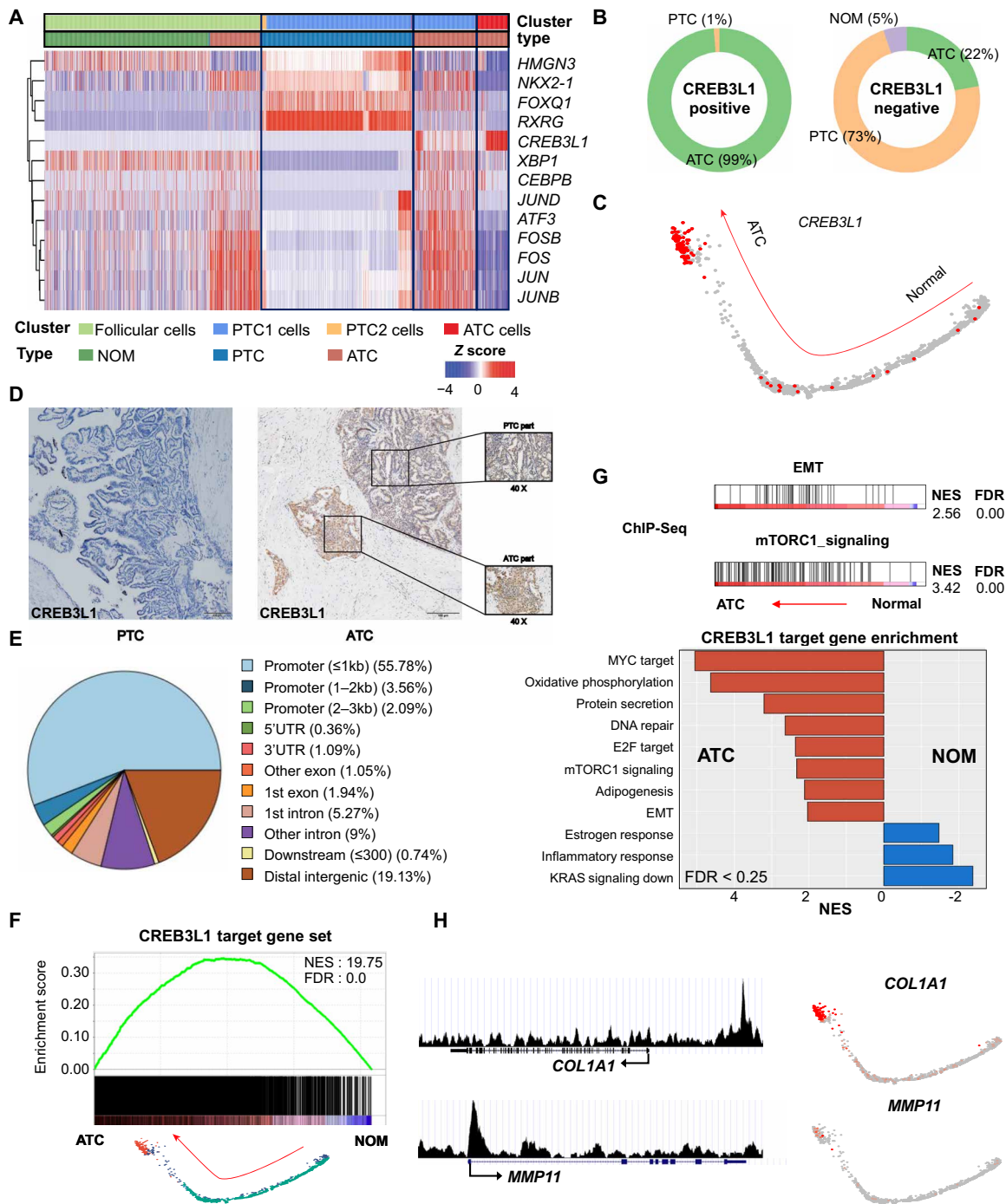
Fig. 4. Key transcriptional regulator of *CREB3L1* in thyroid cancer evolution. (**A**) Coexpression modules in each cell cluster and their master regulators were identified using SCENIC. Cell clusters and their origin samples are shown at the top of the figure. Color scale indicates regulon activity levels. (**B**) Ratio of samples (NOM, PTC, and ATC) between CREB3L1-positive and CREB3L1-negative in the ATC cell cluster. Color indicates each sample. (**C**) *CREB3L1* expression level on the trajectory. Color scale represents expression level. (**D**) Staining of CREB3L1 in PTC sample and ATC and PTC parts from one concurrent ATC sample. Lowercase represents the ATC part, which was cytoplasm-positive in CREB3L1 with high density; uppercase represents the PTC part, which was also positive in CREB3L1. (**E**) Distribution pattern of the CREB3L1-binding peaks in different genomic elements. Genomic elements and percent of peaks are shown on the right side. (**F**) GSEA plot showing the gene enrichment patterns of CREB3L1 target genes on the trajectory from normal to ATC cells. Normalized enrichment score and FDR values are shown on the plots. (**G**) Bottom: CREB3L1 target gene enrichment scores for top significant hallmarks of GSEA analysis (FDR < 0.25). The *x* axis represents normalized enrichment scores. Top: Enrichment of overlapped genes between EMT- or mTORC1-related genes and CREB3L1 targets along normal thyroid tissue to ATC. (**H**) CREB3L1-binding and expression patterns of CREB3L1 target genes. Left: Genome browser screenshots showing the CREB3L1-binding patterns near *COL1A1* and *MMP11* gene loci. Right: Cell trajectory plots showing the expression levels of *COL1A1* and *MMP11*.
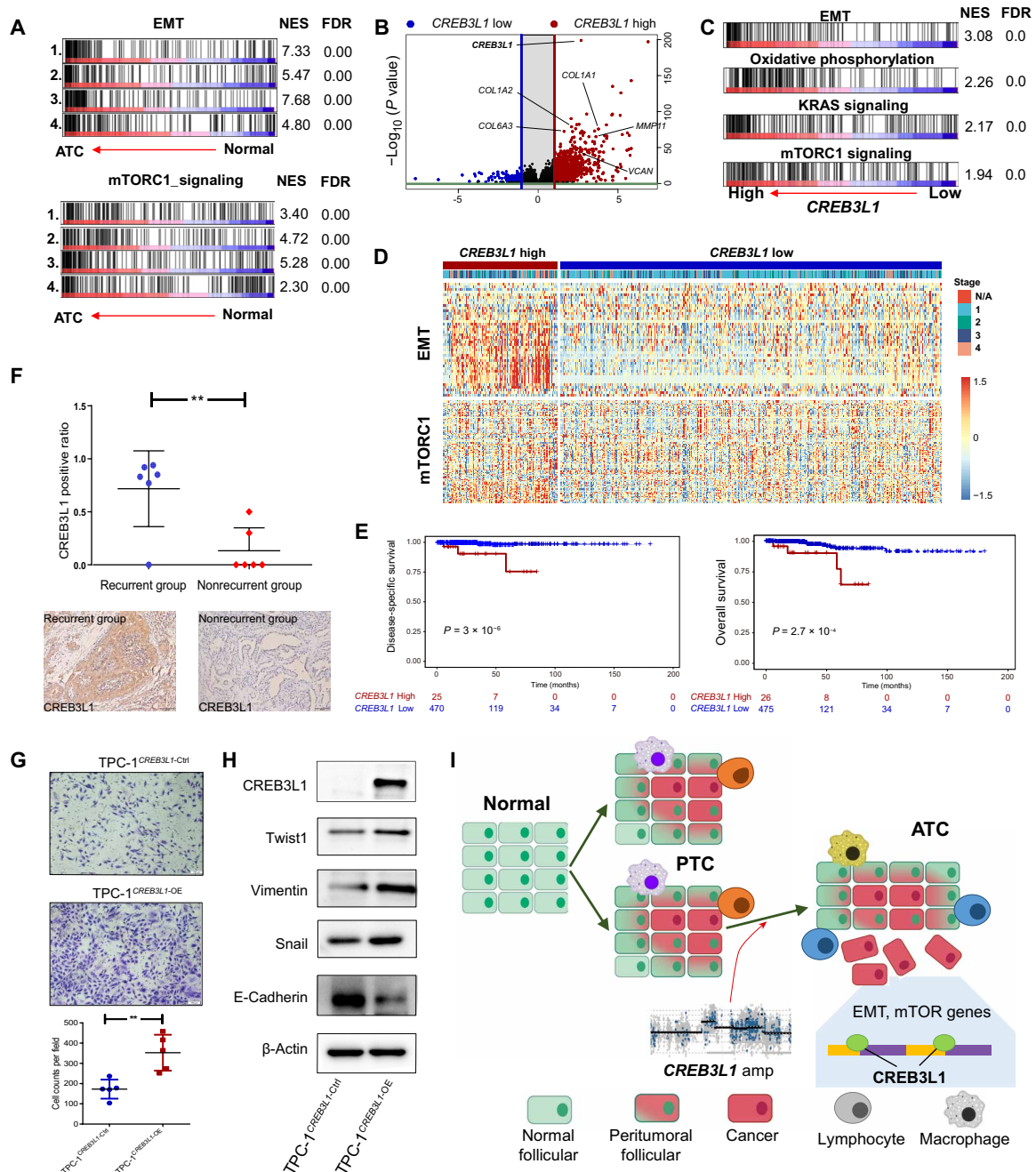
Fig. 5. Involvement and functional status of *CREB3L1* in thyroid cancer progression. (**A**) GSEA plots showing the gene enrichment patterns of EMT and mTORC1 signaling genes between normal and ATC samples from four different publicly available bulk RNA-seq datasets. Normalized enrichment score and FDR values are shown on the right side of each figure. 1: GSE27155, 2: GSE33630, 3: GSE65144, and 4: GSE126698. (**B**) Volcano plot showing the DEGs between *CREB3L1* high and low thyroid cancer groups. Red and blue indicate up-regulated genes in *CREB3L1* high and low, respectively. The *x* axis represents $\log_2$ fold change, and the *y* axis represents $-\log_{10}$ *P* value. (**C**) GSEA plots showing the gene enrichment patterns of EMT, oxidative phosphorylation, KRAS, and mTORC1 signaling genes between *CREB3L1* high and low thyroid cancer samples from TCGA. (**D**) Heatmap showing the expression patterns of EMT and mTORC1 signaling genes in *CREB3L1* high and low groups. (**E**) Kaplan-Meier survival curves showing the overall survival and disease-specific survival in thyroid cancer patients from TCGA separated by extremely high *CREB3L1* expression. (**F**) Comparison of positive ratio of CREB3L1 staining between recurrent group (2-year clinical recurrent PTC samples) and nonrecurrent group (10-year disease-free PTC samples). Representative images of CREB3L1 staining in both groups are shown below the *x* axis. (**G**) *CREB3L1* overexpression significantly increased invasion ability of the TPC-1 cell line compared with control ($P = 0.004$). (**H**) Expression changes of EMT-related markers in TPC-1 cells with *CREB3L1* overexpression versus control. (**I**) Schematic summary of the thyroid cancer progression revealed by this study.

EMT are often involved in cancer progression, associated with poor prognosis in thyroid cancer patients (*31*). Although the importance of EMT has been indicated in ATC tumorigenesis (*32*), no direct evidence of precancerous step for ATC has been revealed, particularly the possibility of its dedifferentiation from PTC. In this study, the systematic investigation of human thyroid cancer at single-cell resolution through scRNA-seq not only profiled intratumoral heterogeneity and specific tumor microenvironment content in ATC and PTC but also provided evidence for ATC progression from highly differentiated PTCs by inferring evolutionary trajectories.

Before our study, indirect evidence for the ATC development exhibited a paradoxical fact. Genomic evidence indicated that ATC and its adjacent PTC lesions had independent evolutionary trajectories, which shared very limited somatic mutations and diverged at the initial stages of tumor development (*12*). However, the extreme contrast of growth rate in PTC and ATC cells excludes the possibility of simultaneous initiation of PTC and ATC lesions from their progenitors, because concomitant cases of PTC and ATC are very common in ATC patients, and most ATC patients experience a history of PTC (*4*). Therefore, our findings postulated a model that reconciles previous views of ATC development, that ATC is dedifferentiated from a subset of aPTC cells rather than the PTC cells from the adjacent PTC lesion, which perfectly explains the 75 to 100% positive IHC staining of specific marker for PTC cells (e.g., galectin-3) in ATC lesions (*33*). As the hypothesis of ATC evolution, the progenitor of aPTC cells diverges from the PTC cells at an early stage and evolve independently. Key molecular events (e.g., overexpression of *CREB3L1* and/or gain of *TP53* mutations) are expected to happen during the transition from a particular aPTC cell to ATC, which would normally take a long time and confer an extreme proliferative capacity to the ATC cells. Collectively, the single-cell analysis is instrumental in delineating thyroid cancer evolution and possible other types of cancer with dedifferentiation event.

As described above, ATC transition may be attributed to the key event of copy number gain of the specific region in Chr11 and subsequent increased *CREB3L1* expression. *CREB3L1* encodes a TF that is normally located in the membrane of the endoplasmic reticulum (ER) and translocates into the nucleus to induce downstream cascade under ER stress (*34*). Recently, a landscape of human cells at the single-cell level grouped 140 orthologous TF regulons into 15 major modules, and *CREB3L1* was one of the representative TFs in stromal cell–associated module, which embodied its well-known mesenchyme-like characteristic (*35*). The direct regulation of CREB3L1 on *COL1A1* (*36*), which is a well-known EMT effector in the dedifferentiation process of thyroid cancer (*37*), has been revealed. Clinically, *CREB3L1* can predict distant metastasis in the mesenchymal subtype of triple-negative breast cancer (*38*), which was consistent with the prognostic value of *CREB3L1* expression for PTC-related death. Accordingly, it implied a stratified management strategy of PTC patients and is beneficial to define the risk level and avoid the overtreatment dilemma (*39*). Moreover, being different from other TFs, CREB3L1 could be a druggable target by the protease inhibitor because of its specific activation mechanism (*38*). However, due to the nonspecific binding of the chemical inhibitor, optimization is expected in the future to treat mesenchyme-like aggressive cancers, such as ATC and mesenchymal subtype of breast cancer.

Another contributor to cancer progression is cancer cell plasticity, which is a collective change of molecular and phenotypic fluctuation (*15*). These changes may involve not only genetic change (*40*) but

also epigenetic and microenvironment alterations. Notably, a switch in the macrophage population from proinflammatory M1-like in PTC to pro-tumorigenic an M2-like phenotype in ATC and the increase in *FAP*⁺ fibroblasts observed in ATCs may contribute to thyroid cancer progression. Moreover, both M2-like macrophages (*41*, *42*) and *FAP*⁺ fibroblasts (*43*, *44*) generate an immunosuppressive tumor microenvironment. Together with the increasing levels of dysfunctional lymphoid cells, regulatory T cells, and *Siglec-15*⁺ tumor-infiltrating myeloid cells, the immunosuppressive landscape of ATC may facilitate ATC cancerous cells to escape immune surveillance. Collectively, these tumor microenvironment changes may explain the negative result of the pilot study (NCT03122496)—nonresponsive to the combination of durvalumab and tremelimumab (*45*).

Several limitations in our studies should be noticed. First, our study was limited by the sample size, because it is hard to acquire fresh tumor tissue to conduct scRNA-seq due to the low prevalence of ATC. Consequently, the common mutation of *RAS* and *TP53* genes in ATC was not observed in our patients probably because of randomness. As the potential key molecular event for ATC transition from aPTC, activation of p53-negative regulatory genes instead of direct *TP53* mutation was observed in all three ATC patients, all of whom carried copy number gain of *CREB3L1*. It would be interesting to investigate the role of interaction between *CREB3L1* overexpression and p53 signaling pathway or mitogen-activated protein kinase (MAPK) signaling (*46*, *47*) in dedifferentiation of thyroid cancer in future study. Considering the limited sample size in the study, CNA of *CREB3L1* should be interpreted cautiously in other cohort and needs further confirmation in larger study. Second, the transition of ATC from aPTC is better to be validated with a specific genomic marker, such as *TP53* mutation carried by a subset of aPTC cells and all ATC cells. However, the mutation information is not available because only a small fraction of 3′ end of each transcript can be sequenced with current scRNA-seq approach, which would be solved by single-cell DNA sequencing at a much higher expense.

Collectively, we characterized the shared and distinct features of normal follicular, PTC, and ATC at single-cell resolution and delineated a dedifferentiation evolutionary trajectory of ATC development. CNA gain and overexpression of *CREB3L1* may play a key role in determining cell fate of thyroid cancer through regulation of EMT and mTOR signaling in thyroid cancer progression and stratifying management of thyroid cancer patients (Fig. 5I).

MATERIALS AND METHODS
Clinical information
The single-cell suspensions were prepared from six patients who underwent thyroid surgery from March to July 2019 at West China Hospital (Chengdu, China). Clinical and histology information were collected from medical record system. Two experienced pathologists independently reconfirmed sample histology in a blinded manner and classified them according to the eighth edition of the AJCC TNM (American Joint Committee on Cancer Tumor Node Metastasis classification) system. The procedure was carried out in accordance with the Declaration of Helsinki and the guidelines of the Ethical Committee of the West China Hospital (2019-821).

Tissue dissociation and purification
We processed thyroid tissues immediately after resection and generated single-cell suspensions. To prepare single-cell suspension,

the dissected tissues were washed with Hanks' balanced salt solution (HBSS) and sheared on ice. Then, the tissues were digested by collagenase I (2 mg/ml) (Gibco 1710-0017), collagenase IV (1 mg/ml) (Gibco 1710-4019), and 0.25% pancreatic enzymes (Gibco 25200-056) for 1 hour at 37°C. The digested tissues were filtered through 40-mm strainer to remove cell debris and large clumps. The suspensions were pelleted at 500g for 5 min at 4°C. After lysis of the red blood cells with 10× RBC Lysis Buffer (Thermo Fisher Scientific, 00-4300-54), the pellets were resuspended in HBSS with 0.04% bovine serum albumin (BSA) to count cell viability and concentration (Counting Star, Aber Instruments Ltd.). The remaining cells were pelleted at 500g for 5 min at 4°C and preserved at −80°C. Single-cell suspensions were finally diluted to 5×10^6 cells per milliliter. All tissues were acquired from West China Biobanks, Department of Clinical Research Management, West China Hospital, Sichuan University.

Library preparation and sequencing
The Chromium Single Cell 3′ Library & Gel Bead Kit v2 (PN-120237), Chromium Single Cell 3' Chip Kit v2 (PN-120236), and Chromium i7 Multiplex Kit (PN-120262) were used according to the manufacturer's instructions in the Chromium Single Cell 3′ Reagents Kits v2 User Guide. The single-cell suspension was washed twice with 1× phosphate-buffered saline (PBS) + 0.04% BSA. Cell number and concentration were confirmed with TC20 Automated Cell Counter. Cells were subjected immediately onto a 10× Genomics Chromium Controller machine for gel beads-in-emulsion (GEM) generation. Barcoded complementary DNAs (cDNAs) were prepared using a 10× Genomics Chromium Single Cell 3′ reagent kit (V2 chemistry), which was subsequently recovered, purified, and amplified to generate sufficient quantities for library preparation. Library quality and concentration were assessed using Agilent Bioanalyzer 2100. Libraries were run on the NovaSeq platform of Illumina for PE150 sequencing.

Single-cell bioinformatics analysis
Raw base call files were demultiplexed using mkfastq application (Cell Ranger v3.1.0) to make FASTQ files. FASTQ files were mapped to the human reference genome (GRCh38 v3.0.0) using count application (Cell Ranger v3.1.0) with default settings. Read10× function from the Seurat package (3.1.2) in R (3.6.0) was used to merge all sample data into an aggregate object. RenameCells function was used to ensure that the cell name is unique. We set several criteria to filter low-quality cells and genes: minimal expression of 200 genes per cell, mitochondrial content less than 15%, and genes that are expressed in more than 3 cells. After filtering, we obtained 46,205 cells. Data were normalized using the "LogNormalize" method and setting a scale factor of 10,000. ScaleData function of Seurat was used to regress out number of unique molecular identifiers (UMI), number of genes, and percent mitochondrial content to remove unwanted sources of variation. Top 2000 variably expressed genes were identified by the FindVariableFeatures function with "vst" option. Batch effects between the samples were removed by using RunFastMNN function in SeuratWrappers package (v0.1.0). After principal components analysis (PCA), cells were clustered using the FindClusters (resolution = 0.7) on the basis of shared nearest neighbor (SNN) using the identified 20 PCs. The cells were visualized by using Uniform Manifold Approximation and Projection (UMAP) embedding. DEGs between every pair of clusters were identified by using FindMarkers (p_val_adj < 0.01 and avg_logFC > 1). Clusters were merged if the number of DEGs

was less than 10 between two clusters. The biological processes that were expressed according to tumor progression were identified by WebGestalt DAVID (48) and GSEA (49).

Deconvolution analysis
Deconvolution analysis for the bulk RNA-seq data (GSE33630) was performed to validate the cell proportion change by using the CellCODE (v0.99.0). For this, information of bulk RNA-seq expression matrix and marker genes of scRNA-seq were used to compute the surrogate proportion variables (SPVs). Marker genes were selected as well known or based on a P value high-ranked. The results of SPVs were visualized by Multiple Experiment Viewer (MEV) (v4.9.0).

Cell-cell communication analysis
CellPhoneDB (https://github.com/Teichlab/cellphonedb), a public repository of ligands, receptors, and their interactions, was used to perform the cell-cell communication analysis. Normalized UMI count and cluster identities were used as the input file for statistical_analysis function with a P value of 0.05. Visualization was done using dot_plot and heatmap_plot function.

Cell trajectory analysis
Monocle2 (v2.12.0) was used to reconstruct the single-cell trajectory from normal to cancer cells. Raw UMI count was used as input for Monocle2, and pseudo time was calculated using reduceDimension and orderCells function. Through differentialGeneTest function, trajectory patterns were determined. Heatmaps were generated with plot_pseudotime_heatmap for the highly variable genes along the trajectory.

InferCNV analysis
A raw count matrix of scRNA-seq expression extracted from the Seurat object was used as input to inferCNV packages (https://github.com/broadinstitute/inferCNV). For both analysis, follicular cells derived from normal tissue were selected as normal reference. To investigate whether the copy number variation (CNV) of PTC cells is heterogeneous, we put PTC1 and PTC2 cells derived from PTC patients as tumor. In addition, we also put PTC cells and ATC cells derived from ATC patients as tumor for the sake of exploring their heterogeneity of CNV. For the inferCNV analysis, we set the parameters "denoise," "default hidden markov model (HMM)," and "cluster_by_groups" to TRUE, a value of 0.1 for "cutoff," and "random_trees" for "tumor_subcluster_partition_method."

SCENIC analysis and ssGSEA
Follicular cells, PTC1, PTC2, and ATC cells were included in the Single-Cell Regulatory Network Inference and Clustering analysis (SCENIC, R package, https://github.com/aertslab/SCENIC) and then sorted according to source and cell types. In the same cell type, every 3 and 50 cells derived from normal (NOM) and PTC patients, respectively, were merged into a "new cell" on the basis of their average expression levels. Genes that expressed either at very low levels or in too few cells were removed by geneFiltering function with default settings. Only the genes that were available in Rcis Target databases (https://resources.aertslab.org/cistarget) were kept. GENIE3 was used for gene regulatory network reconstruction and to detect the association between the TF and the potential target. The cisTarget Human motif database v9 (https://resources.aertslab.org/cistarget/motif2tf/motifs-v9-nr.hgnc-m0.001-o0.0.tbl) of 24,453 motifs was

used for enrichment of gene signatures and pruned for targets from this signature based on cis-regulatory cues with default settings. The "aucell" positional argument was used to find enrichment of regulons across single cells. The results were visualized with the pheatmap (1.0.12) R package.

To investigate the M1/M2 polarization of each single cell, ssGSEA was performed to assess the levels of M1 macrophage and M2 macrophage (recorded as ssGSEA score) in each cell according to the expression levels of macrophage-specific marker genes. The list of marker genes was obtained from the article published by Cassetta *et al.* (*18*).

WES and somatic alteration calling

DNA was extracted from seven samples by using the DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. DNA libraries were constructed using an Agilent SureSelect Human All Exon V6 kit (Agilent Technologies, CA, USA) following the manufacturer's instructions. Tagged libraries were sequenced on Illumina NovaSeq platform to generate PE150 sequencing reads. The pipeline to call somatic mutation and CNAs was proceeded, as we previously described (*40*, *50*). Briefly, after quality control and trimming barcode sequences, clean reads were mapped to the human reference genome (GRCh38/hg38) by Burrows-Wheeler Aligner. Best practices of Genome Analysis ToolKit were used for BAM file sorting, duplicate read marking, local realignment, and base quality recalibration. Mutect2, Strelka2, and Varscan2 were used for calling somatic variants and significantly mutated genes from tumor samples with matched normal samples. Further variant filtration was performed with homemade scripts. WES data were used as input to PureCN packages (https://github.com/lima1/PureCN). For preparation of environment and reference, the agilent_v6 target gene information and mappability and reptiming of wgEncode from the UCSC genome browser were used. VCF files made by Mutect2(GATK v4.1.6.0) were used as input to PureCN packages and matched normal samples for coverage normalization. For the PureCN analysis, we set parameters "VCF Filter" and "VCF Filter" to 0.001 and 0.15, respectively.

ChIP sequencing data analysis

ChIP sequencing data that were performed using a CREB3L1 antibody in K562 cell line was downloaded from ENCODE project (library ENCLB300HOB). Reads were aligned to the hg19 reference genome. The ChIP signal generation and peak calling were performed by MACS (v2.1.0) and SPP (v1.14), respectively. Binding signals of CREB3L1 were displayed by the UCSC genome browser for hg19 ($P < 0.01$). Distribution of CREB3L1-binding region was identified by ChIPseeker (v1.20.0).

IHC and IF staining

IHC: For preparation of tumor and normal sections, tissues were fixed in 10% formalin overnight at 4°C and then embedded in paraffin. Blocks were serially sectioned (thickness of 5 μm). Antigen retrieval was performed with tris-EDTA in microwave for 8 min. The sections were incubated with primary antibody overnight at 4°C, and a biotin-conjugated secondary anti-mouse antibody (DAKO) was used to incubate with primary antibody for 45 min at 37°C. Last, the DAB Chromogenic Reagent Kit (Noble-Ryder ZLI-9018) was used for visualization. IF: After deparaffinization in xylene and alcohol, sections were heated with tris-EDTA in microwave for 8 min

for antigen retrieval. The sections were blocked with 4% normal goat serum (Solaribio, SL038) for 40 min at 37°C. Then slides were washed with PBS (Thermo Fisher Scientific, pH 6), and the primary antibodies were combined with appropriate secondary antibody (Alexa 488 546, Thermo Fisher Scientific) for 1 hour at room temperature. Last, sections were combined with 4′,6-diamidino-2-phenylindole (DAPI) (Vector Labs, H1200) and stored at −20°C. Immunohistochemical staining was evaluated by two independent pathologists according to German semiquantitative scoring system, and the microscopic cell staining (yellowish to brown) was defined as positive cells. The whole slides were scored in terms of both percentage of positive cells and intensity of cell staining. The antibodies against PBK (ProMab, 30329, 1:100), CREB3L1 (Santa Cruz Biotechnology, sc-514635, 1:200), and galectin-3 (Abcam, ab76245, 1:50) were used.

Cell culture and overexpression of *CERB3L1*

The human PTC cell line (i.e., TPC-1) was provided by Shanghai Branch of Chinese Academy of Science and cultured in RPMI 1640 (HyClone) medium containing 10% fetal bovine serum (FBS) (PAN Seratech) and 1% penicillin/streptomycin (Gibco). The cells were grown at 37°C in 5% CO$_2$ and 95% humidified air. The TPC-1 cell line was authenticated by shore tandem repeat profiling. The mycoplasma contamination was tested as negative.

The plasmids (control and *CREB3L1* insertion) were obtained from GeneChem (Shanghai) and confirmed by Sanger sequencing, followed by packaging with helper vectors (PCMV-VSV-G and PCL-Eco) to prepare retrovirus with 293T cells. TPC-1 cells were infected by 24 hours and treated with puromycin for positive selection.

Transwell assay

Transwell assay was performed following standard methods as we previously described (*40*). Briefly, 24-well cell culture chambers (BD Biosciences) were used according to the manufacturer's instructions. TPC-1 cells were seeded at a density of 2×10^4 cells per well in the upper chamber with culture medium (200 μl) alone, while the bottom of the plate was filled with culture medium (800 μl) supplemented with 15% or 10% FBS as a chemoattractant. After 24 hours, cells that invaded the underside of the membrane were in 4% methanol and stained by 1% crystal violet. The same number of cells was also plated onto plates separately to determine the total number of attached cells to normalize the relative cell migration. For each experiment, the number of cells were counted using an inverted microscope in three random fields (magnification, ×100), and three independent filters were analyzed.

Western blotting

Western blotting was performed following standard methods, as we previously described (*40*). Briefly, cells were seeded and allowed to reach 70 to 80% of confluence. Proteins were extracted using radioimmunoprecipitation assay buffer, supplemented with protease inhibitors and phosphatase inhibitors (Beyotime), and quantified using the BCA assay (Beyotime). The relative molecular mass of the immunoreactive bands was determined using PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific). The semiquantitative analysis was performed using β-actin as reference proteins for loading control. Antibodies against CREB3L1 (Santa Cruz Biotechnology, sc-514635, 1:1000), Twitst1 (huaan, RT1635, 1:1000), vimentin (BOSTER,

PB9359, 1:1000), Snail (CST, 3895S, 1:1000), E-cadherin (BOSTER, PB9561, 1:1000), and β-actin (Santa Cruz Biotechnology, sc-69879, 1:1000) were used.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/31/eabf3657/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. R. L. Siegel, K. D. Miller, A. Jemal, Cancer statistics, 2019. *CA Cancer J. Clin.* **69**, 7–34 (2019).
2. J. A. Fagin, S. A. Wells Jr., Biologic and clinical perspectives on thyroid cancer. *N. Engl. J. Med.* **375**, 1054–1067 (2016).
3. J. A. Knauf, X. Ma, E. P. Smith, L. Zhang, N. Mitsutake, X. H. Liao, S. Refetoff, Y. E. Nikiforov, J. A. Fagin, Targeted expression of BRAFV600E in thyroid cells of transgenic mice results in papillary thyroid cancers that undergo dedifferentiation. *Cancer Res.* **65**, 4238–4245 (2005).
4. E. Molinaro, C. Romei, A. Biagini, E. Sabini, L. Agate, S. Mazzeo, G. Materazzi, S. Sellari-Franceschini, A. Ribechini, L. Torregrossa, F. Basolo, P. Vitti, R. Elisei, Anaplastic thyroid carcinoma: From clinicopathology to genetics and advanced therapies. *Nat. Rev. Endocrinol.* **13**, 644–660 (2017).
5. S. Boumahdi, F. J. de Sauvage, The great escape: Tumour cell plasticity in resistance to targeted therapy. *Nat. Rev. Drug Discov.* **19**, 39–56 (2020).
6. V. da Silva-Diz, L. Lorenzo-Sanz, A. Bernat-Peguera, M. Lopez-Cerda, P. Muñoz, Cancer cell plasticity: Impact on tumor progression and therapy response. *Semin. Cancer Biol.* **53**, 48–58 (2018).
7. C. E. Meacham, S. J. Morrison, Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328–337 (2013).
8. I. Landa, T. Ibrahimpasic, L. Boucai, R. Sinha, J. A. Knauf, R. H. Shah, S. Dogan, J. C. Ricarte-Filho, G. P. Krishnamoorthy, N. Xu, N. Schultz, M. F. Berger, C. Sander, B. S. Taylor, R. Ghossein, I. Ganly, J. A. Fagin, Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. *J. Clin. Invest.* **126**, 1052–1066 (2016).
9. S. K. Yoo, Y. S. Song, E. K. Lee, J. Hwang, H. H. Kim, G. Jung, Y. A. Kim, S. J. Kim, S. W. Cho, J. K. Won, E. J. Chung, J. Y. Shin, K. E. Lee, J. I. Kim, Y. J. Park, J. S. Seo, Integrative analysis of genomic and transcriptomic characteristics associated with progression of aggressive thyroid cancer. *Nat. Commun.* **10**, 2764 (2019).
10. L. Zhang, Z. Ren, Z. Su, Y. Liu, T. Yang, M. Cao, Y. Jiang, Y. Tang, H. Chen, W. Zhang, R. Gong, T. Wei, Y. Peng, B. Liu, W. Zhang, L. Yang, Y. Hu, Z. Li, J. Zhu, H. Xu, Y. Shu, H. Luo, Novel recurrent altered genes in Chinese patients with anaplastic thyroid cancer. *J. Clin. Endocrinol. Metab.* **106**, 988–998 (2021).
11. D. G. McFadden, A. Vernon, P. M. Santiago, R. Martinez-McFaline, A. Bhutkar, D. M. Crowley, M. M. Mahon, P. M. Sadow, T. Jacks, p53 constrains progression to anaplastic thyroid carcinoma in a *Braf*-mutant mouse model of papillary thyroid cancer. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1600–E1609 (2014).
12. J. Capdevila, R. Mayor, F. M. Mancuso, C. Iglesias, G. Caratù, I. Matos, C. Zafón, J. Hernando, A. Petit, P. Nuciforo, J. M. Cameselle-Teijeiro, C. Álvarez, J. A. Recio, J. Tabernero, X. Matias-Guiu, A. Vivancos, J. Seoane, Early evolutionary divergence between papillary and anaplastic thyroid cancers. *Ann. Oncol.* **29**, 1454–1460 (2018).
13. J. Park, R. Shrestha, C. Qiu, A. Kondo, S. Huang, M. Werth, M. Li, J. Barasch, K. Suszták, Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).
14. Y. Liu, X. Fan, R. Wang, X. Lu, Y. L. Dang, H. Wang, H. Y. Lin, C. Zhu, H. Ge, J. C. Cross, H. Wang, Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res.* **28**, 819–832 (2018).
15. S. Yuan, R. J. Norgard, B. Z. Stanger, Cellular plasticity in cancer. *Cancer Discov.* **9**, 837–851 (2019).
16. R. F. Rodrigues, L. Roque, T. Krug, V. Leite, Poorly differentiated and anaplastic thyroid carcinomas: Chromosomal and oligo-array profile of five new cell lines. *Br. J. Cancer* **96**, 1237–1245 (2007).
17. G. Tomás, M. Tarabichi, D. Gacquer, A. Hébrant, G. Dom, J. E. Dumont, X. Keutgen, T. J. Fahey III, C. Maenhaut, V. Detours, A general method to derive robust organ-specific gene expression-based differentiation indices: Application to thyroid cancer diagnostic. *Oncogene* **31**, 4490–4498 (2012).
18. L. Cassetta, S. Fragkogianni, A. H. Sims, A. Swierczak, L. M. Forrester, H. Zhang, D. Y. H. Soong, T. Cotechini, P. Anur, E. Y. Lin, A. Fidanza, M. Lopez-Yrigoyen, M. R. Millar, A. Urman, Z. Ai, P. T. Spellman, E. S. Hwang, J. M. Dixon, L. Wiechmann, L. M. Coussens, H. O. Smith, J. W. Pollard, Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. *Cancer Cell* **35**, 588–602.e10 (2019).
19. C. Feig, J. O. Jones, M. Kraman, R. J. B. Wells, A. Deonarine, D. S. Chan, C. M. Connell, E. W. Roberts, Q. Zhao, O. L. Caballero, S. A. Teichmann, T. Janowitz, D. I. Jodrell, D. A. Tuveson, D. T. Fearon, Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20212–20217 (2013).
20. V. Cremasco, J. L. Astarita, A. L. Grauel, S. Keerthivasan, K. MacIsaac, M. C. Woodruff, M. Wu, L. Spel, S. Santoro, Z. Amoozgar, T. Laszewski, S. C. Migoni, K. Knoblich, A. L. Fletcher, M. LaFleur, K. W. Wucherpfennig, E. Pure, G. Dranoff, M. C. Carroll, S. J. Turley, FAP delineates heterogeneous and functionally divergent stromal cells in immune-excluded breast tumors. *Cancer Immunol. Res.* **6**, 1472–1485 (2018).
21. D. Dhar, L. Antonucci, H. Nakagawa, J. Y. Kim, E. Glitzner, S. Caruso, S. Shalapour, L. Yang, M. A. Valasek, S. Lee, K. Minnich, E. Seki, J. Tuckermann, M. Sibilia, J. Zucman-Rossi, M. Karin, Liver cancer initiation requires p53 inhibition by CD44-enhanced growth factor signaling. *Cancer Cell* **33**, 1061–1077.e6 (2018).
22. S. Piccinin, E. Tonin, S. Sessa, S. Demontis, S. Rossi, L. Pecciarini, L. Zanatta, F. Pivetta, A. Grizzo, M. Sonego, C. Rosano, A. P. D. Tos, C. Doglioni, R. Maestro, A "twist box" code of p53 inactivation: Twist box:p53 interaction promotes p53 degradation. *Cancer Cell* **22**, 404–415 (2012).
23. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Register, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. K. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jane-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
24. J. J. Lee, A. Y. M. Au, T. Foukakis, M. Barbaro, N. Kiss, R. Clifton-Bligh, J. Staaf, A. Borg, L. Delbridge, B. G. Robinson, G. Wallin, A. Hoog, C. Larsson, Array-CGH identifies cyclin D1 and UBCH10 amplicons in anaplastic thyroid carcinoma. *Endocr. Relat. Cancer* **15**, 801–815 (2008).
25. S. Aibar, C. B. González-Blas, T. Moerman, V. A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J. C. Marine, P. Geurts, J. Aerts, J. van den Oord, Z. K. Atak, J. Wouters, S. Aerts, SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
26. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
27. T. J. Giordano, R. Kuick, D. G. Thomas, D. E. Misek, M. Vinco, D. Sanders, Z. Zhu, R. Ciampi, M. Roh, K. Shedden, P. Gauger, G. Doherty, N. W. Thompson, S. Hanash, R. J. Koenig, Y. E. Nikiforov, Molecular classification of papillary thyroid carcinoma: Distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* **24**, 6646–6656 (2005).
28. C. A. von Roemeling, L. A. Marlow, A. B. Pinkerton, A. Crist, J. Miller, H. W. Tun, R. C. Smallridge, J. A. Copland, Aberrant lipid metabolism in anaplastic thyroid carcinoma reveals stearoyl CoA desaturase 1 as a novel therapeutic target. *J. Clin. Endocrinol. Metab.* **100**, E697–E709 (2015).
29. J. Haase, D. Misiak, M. Bauer, N. Pazaitis, J. Braun, R. Pötschke, A. Mensch, J. L. Bell, H. Dralle, U. Siebolts, C. Wickenhauser, K. Lorenz, S. Hüttelmaier, IGF2BP1 is the first positive marker for anaplastic thyroid carcinoma diagnosis. *Mod. Pathol.* **34**, 32–41 (2021).
30. V. Bernard, A. Semaan, J. Huang, F. A. San Lucas, F. C. Mulu, B. M. Stephens, P. A. Guerrero, Y. Huang, J. Zhao, N. Kamyabi, S. Sen, P. A. Scheet, C. M. Taniguchi, M. P. Kim, C. W. Tzeng, M. H. Katz, A. D. Singhi, A. Maitra, H. A. Alvarez, Single-cell transcriptomics of pancreatic cancer precursors demonstrates epithelial and microenvironmental heterogeneity as an early event in neoplastic progression. *Clin. Cancer Res.* **25**, 2194–2205 (2019).
31. A. V. Nikitski, S. L. Rominski, V. Condello, C. Kaya, M. Wankhede, F. Panebianco, H. Yang, D. L. Altschuler, Y. E. Nikiforov, Mouse model of thyroid cancer progression and dedifferentiation driven by STRN-ALK expression and loss of p53: Evidence for the existence of two types of poorly differentiated carcinoma. *Thyroid* **29**, 1425–1437 (2019).
32. J. A. Knauf, M. A. Sartor, M. Medvedovic, E. Lundsmith, M. Ryder, M. Salzano, Y. E. Nikiforov, T. J. Giordano, R. A. Ghossein, J. A. Fagin, Progression of BRAF-induced thyroid cancer is associated with epithelial-mesenchymal transition requiring concomitant MAP kinase and TGFβ signaling. *Oncogene* **30**, 3153–3162 (2011).
33. C. G. Chiu, S. S. Strugnell, O. L. Griffith, S. J. M. Jones, A. M. Gown, B. Walker, I. R. Nabi, S. M. Wiseman, Diagnostic utility of galectin-3 in thyroid cancer. *Am. J. Pathol.* **176**, 2067–2081 (2010).
34. R. N. Vellanki, L. Zhang, M. A. Guney, J. V. Rocheleau, M. Gannon, A. Volchuk, OASIS/CREB3L1 induces expression of genes involved in extracellular matrix production but not classical endoplasmic reticulum stress response genes in pancreatic beta-cells. *Endocrinology* **151**, 4146–4157 (2010).
35. X. Han, Z. Zhou, L. Fei, H. Sun, R. Wang, Y. Chen, H. Chen, J. Wang, H. Tang, W. Ge, Y. Zhou, F. Ye, M. Jiang, J. Wu, Y. Xiao, X. Jia, T. Zhang, X. Ma, Q. Zhang, X. Bai, S. Lai, C. Yu, L. Zhu,

35. R. Lin, Y. Gao, M. Wang, Y. Wu, J. Zhang, R. Zhan, S. Zhu, H. Hu, C. Wang, M. Chen, H. Huang, T. Liang, J. Chen, W. Wang, D. Zhang, G. Guo, Construction of a human cell landscape at single-cell level. *Nature* **581**, 303–309 (2020).

36. T. Murakami, A. Saito, S. I. Hino, S. Kondo, S. Kanemoto, K. Chihara, H. Sekiya, K. Tsumagari, K. Ochiai, K. Yoshinaga, M. Saitoh, R. Nishimura, T. Yoneda, I. Kou, T. Furuichi, S. Ikegawa, M. Ikawa, M. Okabe, A. Wanaka, K. Imaizumi, Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. *Nat. Cell Biol.* **11**, 1205–1211 (2009).

37. L. A. Jolly, S. Novitskiy, P. Owens, N. Massoll, N. Cheng, W. Fang, H. L. Moses, A. T. Franco, Fibroblast-mediated collagen remodeling within the tumor microenvironment facilitates progression of thyroid cancers driven by Braf^{V600E} and Pten loss. *Cancer Res.* **76**, 1804–1813 (2016).

38. Y. X. Feng, D. X. Jin, E. S. Sokol, F. Reinhardt, D. H. Miller, P. B. Gupta, Cancer-specific PERK signaling drives invasion and metastasis through CREB3L1. *Nat. Commun.* **8**, 1079 (2017).

39. M. Li, L. Dal Maso, S. Vaccarella, Global trends in thyroid cancer incidence and the impact of overdiagnosis. *Lancet Diabetes Endocrinol.* **8**, 468–470 (2020).

40. H.-N. Chen, Y. Shu, F. Liao, X. Liao, H. Zhang, Y. Qin, Z. Wang, M. Luo, Q. Liu, Z. Xue, M. Cao, S. Zhang, W.-H. Zhang, Q. Hou, X. Xia, H. Luo, Y. Zhang, L. Yang, J.-K. Hu, X. Fu, B. Liu, H. Hu, C. Huang, Y. Peng, W. Cheng, L. Dai, L. Yang, W. Zhang, B. Dong, Y. Li, Y. Wei, H. Xu, Z.-G. Zhou, Genomic evolution and diverse models of systemic metastases in colorectal cancer. *Gut* , (2021).

41. M. H. M. Barros, F. Hauck, J. H. Dreyer, B. Kempkes, G. Niedobitek, Macrophage polarisation: An immunohistochemical approach for identifying M1 and M2 macrophages. *PLOS ONE* **8**, e80908 (2013).

42. R. Noy, J. W. Pollard, Tumor-associated macrophages: From mechanisms to therapy. *Immunity* **41**, 49–61 (2014).

43. M. Kraman, P. J. Bambrough, J. N. Arnold, E. W. Roberts, L. Magiera, J. O. Jones, A. Gopinathan, D. A. Tuveson, D. T. Fearon, Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. *Science* **330**, 827–830 (2010).

44. E. Sahai, I. Astsaturov, E. Cukierman, D. G. DeNardo, M. Egeblad, R. M. Evans, D. Fearon, F. R. Greten, S. R. Hingorani, T. Hunter, R. O. Hynes, R. K. Jain, T. Janowitz, C. Jorgensen, A. C. Kimmelman, M. G. Kolonin, R. G. Maki, R. S. Powers, E. Puré, D. C. Ramirez, R. Scherz-Shouval, M. H. Sherman, S. Stewart, T. D. Tlsty, D. A. Tuveson, F. M. Watt, V. Weaver, A. T. Weeraratna, Z. Werb, A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186 (2020).

45. E. J. Sherman, C. J. Tsai, W. I. Zhi, J. V. Fetten, V. Wu, A. L. Ho, N. Riaz, D. G. Pfister, N. Y. Lee, Pilot study combining PD-L1 antibody durvalumab (D) with CTLA-4 antibody tremelimumab (T) and stereotactic body radiotherapy (SBRT) to treat metastatic anaplastic thyroid cancer (ATC). *J. Clin. Oncol.* **37**, –6088 (2019).

46. M. Drosten, E. Y. M. Sum, C. G. Lechuga, L. Simón-Carrasco, H. K. C. Jacob, R. García-Medina, S. Huang, R. L. Beijersbergen, R. Bernards, M. Barbacid, Loss of p53 induces cell proliferation via Ras-independent activation of the Raf/Mek/Erk signaling pathway. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15155–15160 (2014).

47. J. A. Knauf, K. A. Luckett, K. Y. Chen, F. Voza, N. D. Socci, R. Ghossein, J. A. Fagin, Hgf/Met activation mediates resistance to BRAF inhibition in murine anaplastic thyroid cancers. *J. Clin. Invest.* **128**, 4086–4097 (2018).

48. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

49. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).

50. Y. Shu, W. Zhang, Q. Hou, L. Zhao, S. Zhang, J. Zhou, X. Song, Y. Zhang, D. Jiang, X. Chen, P. Wang, X. Xia, F. Liao, D. Yin, X. Chen, X. Zhou, D. Zhang, S. Yin, K. Yang, J. Liu, L. Fu, L. Zhang, Y. Wang, J. Zhang, Y. An, H. Cheng, B. Zheng, H. Sun, Y. Zhao, Y. Wang, D. Xie, L. Ouyang, P. Wang, W. Zhang, M. Qiu, X. Fu, L. Dai, G. He, H. Yang, W. Cheng, L. Yang, B. Liu, W. Li, B. Dong, Z. Zhou, Y. Wei, Y. Peng, H. Xu, J. Hu, Prognostic significance of frequent CLDN18-ARHGAP26/6 fusion in gastric signet-ring cell cancer. *Nat. Commun.* **9**, 2447 (2018).

Citation: H. Luo, X. Xia, G. D. Kim, Y. Liu, Z. Xue, L. Zhang, Y. Shu, T. Yang, Y. Chen, S. Zhang, H. Chen, W. Zhang, R. Li, H. Tang, B. Dong, X. Fu, W. Cheng, W. Zhang, L. Yang, Y. Peng, L. Dai, H. Hu, Y. Jiang, C. Gong, Y. Hu, J. Zhu, Z. Li, C. Caulin, T. Wei, J. Park, H. Xu, Characterizing dedifferentiation of thyroid cancer by integrated analysis. *Sci. Adv.* **7**, eabf3657 (2021).