

Multidimensional data analysis revealed thyroiditis-associated TCF19 SNP rs2073724 as a highly ranked protective variant in thyroid cancer

Xianhui Ruan^{1,*}, Yu Liu^{1,*}, Shuping Wu^{1,2}, Guiming Fu^{1,3}, Mei Tao¹, Yue Huang¹, Dapeng Li¹, Songfeng Wei¹, Ming Gao^{1,4,5}, Shicheng Guo⁶, Junya Ning^{1,4}, Xiangqian Zheng¹

¹Department of Thyroid and Neck Tumor, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin 300060, China

²Department of Head and Neck Surgery, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou 350014, Fujian, China

³Thyroid-Otolaryngology Department, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital and Institute, Sichuan Cancer Center, Affiliated Cancer Hospital of University of Electronic Science and Technology of China, Chengdu 610000, Sichuan, China

⁴Department of Thyroid and Breast Surgery, Tianjin Union Medical Center, Tianjin 300121, China

⁵Tianjin Key Laboratory of General Surgery in Construction, Tianjin Union Medical Center, Tianjin 300121, China

⁶Department of Medical Genetics, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53706, USA

*Equal contribution

Correspondence to: Shicheng Guo, Junya Ning, Xiangqian Zheng; email: shihcheng.guo@gmail.com, <https://orcid.org/0000-0002-7047-9972>; ningjunya123@tmu.edu.cn; xzheng05@tmu.edu.cn

Keywords: thyroid cancer, GWAS, biomarker, therapeutic target, SNP

Received: November 24, 2023

Accepted: March 14, 2024

Published: April 4, 2024

Copyright: © 2024 Ruan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Thyroid cancer represents the most prevalent malignant endocrine tumour, with rising incidence worldwide and high mortality rates among patients exhibiting dedifferentiation and metastasis. Effective biomarkers and therapeutic interventions are warranted in aggressive thyroid malignancies. The transcription factor 19 (TCF19) gene has been implicated in conferring a malignant phenotype in cancers. However, its contribution to thyroid neoplasms remains unclear.

Results: In this study, we performed genome-wide and phenome-wide association studies to identify a potential causal relationship between TCF19 and thyroid cancer. Our analyses revealed significant associations between TCF19 and various autoimmune diseases and human cancers, including cervical cancer and autoimmune thyroiditis, with a particularly robust signal for the deleterious missense variation rs2073724 that is associated with thyroid function, hypothyroidism, and autoimmunity. Furthermore, functional assays and transcriptional profiling in thyroid cancer cells demonstrated that TCF19 regulates important biological processes, especially inflammatory and immune responses. We demonstrated that TCF19 could promote the progression of thyroid cancer *in vitro* and *in vivo* and the C>T variant of rs2073724 disrupted TCF19 protein binding to target gene promoters and their expression, thus reversing the effect of TCF19 protein.

Conclusions: Taken together, these findings implicate TCF19 as a promising therapeutic target in aggressive thyroid malignancies and designate rs2073724 as a causal biomarker warranting further investigation in thyroid cancer.

INTRODUCTION

Thyroid cancer is the most common malignant endocrine tumour, accounting for 1-2% of all malignant tumours [1]. Thyroid cancer is an important public health concern because its incidence has been increasing globally [2, 3]. A study conducted by the International Agency for Research on Cancer for GLOBOCAN 2020 and WHO cancer statistics estimated 224,023 and 53,815 new thyroid cancer cases in China and the United States in 2022, respectively, with 9,915 deaths (4.4%) and 2,262 deaths (4.2%) [4, 5]. Most thyroid cancers are indolent, but some tumours undergo dedifferentiation and metastasis, resulting in a high mortality rate [6–8]. These patients do not respond well to conventional treatments, including surgery and radioactive iodine (I^{131}) [9, 10]. As a result, identifying effective biomarkers and therapeutic targets is crucial for aggressive thyroid cancer.

As a complex disease, thyroid cancer is related to multiple genetic factors. Recently, unbiased large-scale association studies have been conducted to identify genetic variations associated with an increased risk of developing thyroid cancer. For instance, a genome-wide association study (GWAS) identified 24 new genetic variants associated with thyroid stimulating hormone in 2020 [11]. Other studies have identified specific genes that may play a role in the development of thyroid cancer, including genes involved in DNA repair, cell cycle regulation, and immune function [12, 13]. In addition, a phenome-wide association study (PheWAS) was conducted to use electronic health record (EHR) data and identify genetic associations with multiple phenotypes, such as thyroid function, thyroid hormone levels, and thyroid nodules, simultaneously [14, 15]. Overall, these data-driven GWAS and PheWAS have provided valuable insights into the genetic basis of thyroid cancer and have the potential to lead to the development of new diagnostic and treatment strategies. However, further research is needed to fully understand the complex genetic and environmental factors that contribute to thyroid cancer development and progression. In addition, like all large-scale association studies, findings from both GWAS and PheWAS must be validated before they can be considered definitive.

In the current study, we comprehensively collected previous GWAS and PheWAS studies for thyroid-related diseases. By further leveraging large-scale genetic expression data, we prioritized a key gene, TCF19, which was demonstrated to have a potential causal relationship with thyroid cancer. Moreover, we performed *in vitro* functional studies and found that TCF19 could significantly promote the proliferation and

metastasis of thyroid cancer and increase the severity of disease progression. An important finding was that the C>T variant of rs2073724 partially blocked TCF19's function and played a protective role in thyroid cancer progression. RNA-seq analysis of TCF19 wild-type and SNP-overexpressing cells illustrated that TCF19 regulates many important biological processes in thyroid cancer, especially the inflammation pathway and the immune response [16].

RESULTS

Genome-wide and phenome-wide associations between TCF19 and immune diseases and human cancers

We collected multiple GWAS studies, including UKB, FinnGen, and GWAS catalogue, and identified that TCF19 is significantly associated with multiple autoimmune diseases and human cancers, for example, type 1 diabetes and autoimmune thyroid diseases ($P=1.0 \times 10^{-23}$), cervical cancer (4.0×10^{-25}) and lung cancer (5.0×10^{-19}) (Table 1). We also observed that rs2073724, a potential deleterious missense variation predicted by Sorting Intolerant From Tolerant (SIFT) [16], is associated with thyroid function ($P=2.9 \times 10^{-19}$), thyroiditis ($P=2.3 \times 10^{-7}$), hypothyroidism ($P=1.0 \times 10^{-8}$) and multiple autoimmune diseases ($P=2.7 \times 10^{-10}$) (Table 2).

Moreover, the GTEx dataset showed that compared with the wild type allele (Pro), the rs2073724 alternative allele (Leu) increased the expression level of TCF19 in almost all human tissues (thyroid tissue, $NES=0.99$, $P=9.6 \times 10^{-66}$, Supplementary Figure 1). The causal relationship between the alternative allele and increased aetiology of thyroiditis and thyroid cancer. Through analysis of the TCF19-related SNP in UCSC, we found that rs2073724 was located in exon 3 of TCF19 (Supplementary Figure 2). Hence, exonic SNPs are important SNPs, and the function of exonic SNPs in thyroid cancer remains largely unknown. Therefore, the potential role of TCF19 and its exonic SNP (rs2073724) in thyroid cancer tumorigenesis and the specific mechanism drew our attention.

TCF19 is upregulated in thyroid cancer and correlated with cancer progression and a poor prognosis

To further investigate the expression pattern of TCF19 in human cancers, we first analysed The Cancer Genome Atlas (TCGA) database and found that TCF19 mRNA levels were upregulated in most human cancers (meta-analysis, $\log_2FC=0.68$, $P=2.84 \times 10^{-13}$) including thyroid cancer (Supplementary Figure 3 and Figure 1A).

Table 1. Previous PheWAS studies that are reported to be associated with TCF19 gene.

Study ID	Reported trait	Study N cases	Index variant ID	Index variant RSID	p-value	PMID	Gene
GCST90011816	Cervical cancer	416913	6_31409942_T_C	rs2523496	4E-25	PMID:32887889	TCF19
GCST002876	Type 1 diabetes and autoimmune thyroid diseases	1598	6_31600692_G_A	rs2857595	2E-23	PMID:25936594	TCF19
GCST008365	Thyrotoxic hypokalemic periodic paralysis and Graves disease	3782	6_31090401_T_C	rs4947296	3E-22	PMID:31050781	TCF19
GCST90018770	Medication use (thyroid preparations)	178726	6_31593738_C_T	rs9357135	1E-19	PMID:34594039	TCF19
GCST012213	Aerodigestive squamous cell cancer (pleiotropy)	75848	6_31459618_G_C	rs9267123	2E-19	PMID:33667223	TCF19
GCST004748	Lung cancer	85716	6_31466334_A_G	rs3094604	5E-19	PMID:28604730	TCF19
GCST001900	Cervical cancer	7180	6_31422633_T_C	rs2516448	4E-18	PMID:23482656	TCF19
GCST012475	Cervical intraepithelial neoplasia grade 3 and invasive cervical cancer	150314	6_31410016_T_A	rs6938453	2E-17	PMID:33794208	TCF19
GCST012213	Aerodigestive squamous cell cancer (pleiotropy)	75848	6_30731245_G_A	rs3095327	2E-16	PMID:33667223	TCF19
GCST90018921	Skin cancer	670929	6_31356770_G_C	rs1065386	4E-16	PMID:34594039	TCF19
GCST006899	Thyroid stimulating hormone levels	76671	6_31140352_C_T	rs1265091	5E-15	PMID:30367059	TCF19
GCST90011822	Cancer (pleiotropy)	475312	6_31346329_C_G	rs3016018	2E-14	PMID:32887889	TCF19
GCST002876	Type 1 diabetes and autoimmune thyroid diseases	1598	6_30814225_C_T	rs886424	3E-14	PMID:25936594	TCF19
GCST004833	Cervical cancer	9347	6_31512891_G_A	rs3132461	2E-13	PMID:28806749	TCF19
GCST011008	Autoimmune traits	64039	6_31380039_G_A	rs4713462	2E-13	PMID:32534018	TCF19
GCST008870	Keratinocyte cancer (MTAG)	358840	6_31441506_G_A	rs61447909	4E-13	PMID:31174203	TCF19
GCST90011816	Cervical cancer	416913	6_31021915_T_G	rs114115115	6E-13	PMID:32887889	TCF19
GCST008573	Composite immunoglobulin trait (IgA/IgM)	16329	6_31358474_C_T	rs9405084	8E-13	PMID:28628107	TCF19
GCST90027058	Non-melanoma skin cancer	274830	6_31137533_A_G	rs1265100	8E-13	PMID:34290314	TCF19
GCST90085699	Human papilloma virus 16 positive oropharyngeal cancer	6334	6_31363011_G_A	rs9266329	1E-12	PMID:34642315	TCF19
GCST90018849	Gastric cancer	643238	6_31281955_T_C	rs2853940	8E-12	PMID:34594039	TCF19
GCST008870	Keratinocyte cancer (MTAG)	358840	6_31394488_T_C	rs28366119	4E-11	PMID:31174203	TCF19
GCST90018905	Prostate cancer	301559	6_31099487_C_T	rs9263523	1E-10	PMID:34594039	TCF19
GCST008870	Keratinocyte cancer (MTAG)	358840	6_31346869_A_G	rs2507999	1E-10	PMID:31174203	TCF19
GCST010286	Oropharynx cancer	11197	6_30991224_G_A	rs13211972	1E-10	PMID:32276964	TCF19
GCST90018629	Gastric cancer	167122	6_31271808_C_T	rs1050437	2E-10	PMID:34594039	TCF19
GCST90011823	Cancer (pleiotropy) (bidirectional)	475312	6_31045574_G_A	rs17190106	2E-10	PMID:32887889	TCF19
GCST011932	Thyrotoxic periodic paralysis	3822	6_31466032_T_G	rs112723370	5E-10	PMID:33105104	TCF19
GCST008500	Tetanus toxoid IgG concentrations post childhood immunization	1852	6_31481245_T_C	rs2523650	1E-09	PMID:31189108	TCF19
GCST003527_4	Anti-thyroid drug induced agranulocytosis	5209	6_31356759_T_A	rs1071816	2E-09	PMID:27157822	TCF19
GCST009825	Cervical cancer (MTAG)	40094	6_30719695_C_T	rs117670375	2E-09	PMID:31488892	TCF19
GCST90085699	Human papilloma virus 16 positive oropharyngeal cancer	6334	6_31466557_C_T	rs2523679	3E-09	PMID:34642315	TCF19
GCST009823_8	Gynecologic disease	47331	6_31404941_G_A	rs2507968	3E-09	PMID:31488892	TCF19
GCST007856	Colorectal cancer or advanced adenoma	125478	6_31042408_T_C	rs116353863	9E-09	PMID:30510241	TCF19
GCST001474	Hypothyroidism	39282	6_31050630_A_G	rs2517532	1E-08	PMID:22493691	TCF19
GCST009823_5	Gynecologic disease (multivariate analysis) [uterine cervical cancer]	47331	6_30719695_C_T	rs117670375	1E-08	PMID:31488892	TCF19
GCST90018638	Hepatic cancer	161323	6_31355203_C_T	rs1131500	2E-08	PMID:34594039	TCF19
GCST007992	Colorectal cancer	92967	6_30790689_A_G	rs3131043	3E-08	PMID:31089142	TCF19
GCST001148	Prostate cancer	73609	6_31150734_T_G	rs130067	3E-08	PMID:21743467	TCF19
GCST009825	Cervical cancer (MTAG)	40094	6_31330644_T_C	rs28752857	5E-08	PMID:31488892	TCF19

Note: previously GWAS and pheWAS studies are collected from multiple sources including GWAS catalog, Opentargets and ExPheWas.

Table 2. Previous PheWAS studies that are reported to be associated with rs2073724 in TCF19.

Study.ID	Trait	p-value	N.Overall	SNP	Gene	PMID
GCST90038635	Thyroid problem (not cancer)	2.9E-19	484598	rs2073724	TCF19	PMID:33959723
NEALE2_20002_1111	Asthma non-cancer illness code, self-reported	6.84289E-19	361141	rs2073724	TCF19	UKB Neale v2
NEALE2_20002_1456	Malabsorption/coeliac disease non-cancer illness code, self-reported	1.32417E-18	361141	rs2073724	TCF19	UKB Neale v2
NEALE2_20002_1226	Hypothyroidism/myxoedema non-cancer illness code, self-reported	8.05449E-14	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_D3_immuNEMECHANISM	Certain disorders involving the immune mechanism	1.35E-10	260405	rs2073724	TCF19	FINNGEN_R6
FINNGEN_R6_D3_BLOOD	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	1.82E-10	260405	rs2073724	TCF19	FINNGEN_R6
FINNGEN_R6_AUTOimmuNE	Autoimmune diseases	2.7E-10	260405	rs2073724	TCF19	FINNGEN_R6
FINNGEN_R6_CD2_FOLLICULAR_LYMPHOMA_EXALLC	Follicular lymphoma (controls excluding all cancers)	2.23E-09	209748	rs2073724	TCF19	FINNGEN_R6
GCST90038603	immunological or systemic disorder	0.000000029	484598	rs2073724	TCF19	PMID:33959723
NEALE2_20002_1428	Thyroiditis non-cancer illness code, self-reported	2.32757E-07	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_C3_UNCERTAIN_SECONDARY_EXALLC	Secondary uncertain malignant neoplasm (controls excluding all cancers)	0.00000067	206307	rs2073724	TCF19	FINNGEN_R6
FINNGEN_R6_C3_LYMPHNODES_SECONDARY_NAS_EXALLC	Secondary and unspecified malignant neoplasm of lymph nodes (controls excluding all cancers)	0.0000101	205451	rs2073724	TCF19	FINNGEN_R6
NEALE2_20002_1371	Sarcoidosis non-cancer illness code, self-reported	2.48599E-05	361141	rs2073724	TCF19	UKB Neale v2
NEALE2_20002_1453	Psoriasis non-cancer illness code, self-reported	5.71729E-05	361141	rs2073724	TCF19	UKB Neale v2
NEALE2_20002_1313	Ankylosing spondylitis non-cancer illness code, self-reported	9.10283E-05	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_CD2_INSITU_CERVIX_UTERI_EXALLC	Carcinoma <i>in situ</i> of cervix uteri (controls excluding all cancers)	0.00012	119317	rs2073724	TCF19	FINNGEN_R6
NEALE2_20002_1387	Hayfever/allergic rhinitis non-cancer illness code, self-reported	0.000127793	361141	rs2073724	TCF19	UKB Neale v2
NEALE2_20002_1471	Atrial fibrillation non-cancer illness code, self-reported	0.000132399	361141	rs2073724	TCF19	UKB Neale v2
GCST006085	Prostate cancer	0.000243	140254	rs2073724	TCF19	PMID:29892016
NEALE2_20002_1225	Hyperthyroidism/thyrotoxicosis non-cancer illness code, self-reported	0.000245311	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_C3_VULVA_EXALLC	Malignant neoplasm of vulva (controls excluding all cancers)	0.000323	117169	rs2073724	TCF19	FINNGEN_R6
NEALE2_20002_1473	High cholesterol non-cancer illness code, self-reported	0.000333684	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_C3_CERVIX_UTERI_EXALLC	Malignant neoplasm of cervix uteri (controls excluding all cancers)	0.000981	119210	rs2073724	TCF19	FINNGEN_R6
NEALE2_20002_1138	Gastro-oesophageal reflux (gord) / gastric reflux non-cancer illness code, self-reported	0.00153374	361141	rs2073724	TCF19	UKB Neale v2
GCST006465	Endometrial cancer (endometrioid histology)	0.00253884	54884	rs2073724	TCF19	PMID:30093612
NEALE2_20002_1202	Urinary frequency / incontinence non-cancer illness code, self-reported	0.00268392	361141	rs2073724	TCF19	UKB Neale v2
FINNGEN_R6_C3_DLCL_EXALLC	Diffuse large B-cell lymphoma (controls excluding all cancers)	0.00283	204397	rs2073724	TCF19	FINNGEN_R6
NEALE2_20002_1597	Tinnitus / tinititis non-cancer illness code, self-reported	0.0031828	361141	rs2073724	TCF19	UKB Neale v2
GCST006464	Endometrial cancer	0.0040379	121885	rs2073724	TCF19	PMID:30093612
FINNGEN_R6_CD2_NONHODGKIN_NAS_EXALLC	Other and unspecified types of non-Hodgkin lymphoma (controls)	0.0041	209777	rs2073724	TCF19	FINNGEN_R6

	excluding all cancers)					
NEALE2_20002_1575	Herpes simplex non-cancer illness code, self-reported	0.00427488	361141	rs2073724	TCF19	UKB Neale v2
SAIGE_149	Cancer of larynx, pharynx, nasal cavities	0.00458	407449	rs2073724	TCF19	UKB SAIGE
NEALE2_20002_1477	Psoriatic arthropathy non-cancer illness code, self-reported	0.00466974	361141	rs2073724	TCF19	UKB Neale v2

Note: previously GWAS and pheWAS studies are collected from multiple sources including GWAS catalog, Opentargets and ExPheWas.

Furthermore, in two independent GEO datasets (GSE29265 and GSE33630), TCF19 mRNA levels were significantly higher in thyroid cancer than in matched normal controls (Figure 1B). To confirm the results of the datasets, we used a quantitative reverse transcription PCR (RT-qPCR) method to quantify TCF19 mRNA levels in 10 pairs of thyroid cancer tissues and their corresponding adjacent normal thyroid tissues. Upregulation of TCF19 was detected in 10 pairs, accounting for 90% of all tested samples (Figure 1C). We also analyzed the TCF19 expression levels in relation to TNM stages in thyroid cancer patients using the TCGA database. We observed that TCF19 expression was elevated in T3 stage primary lesions as compared to T2 (Figure 1D). Similarly, patients with lymph node metastasis exhibited higher TCF19 levels than those without (Figure 1E). Although no significant differences were found in metastasis (M), this could be due to the small number of cases with distant metastasis (M1) in our sample (Figure 1F). Moreover, TCF19 expression was significantly higher in stage III and IV tumors compared to stage II (Figure 1G). Furthermore, patients with high TCF19 expression had shorter progression-free intervals (PFIs) than those with low TCF19 expression (HR=2.04, p=0.012) (Figure 1H). These results indicate that TCF19 is highly expressed in thyroid cancer and that patients with high TCF19 expression have a poor prognosis.

TCF19 promotes thyroid cancer progression, and the C>T variant of rs2073724 plays a protective role in thyroid cancer tumorigenesis

To detect the function of TCF19 and the SNP in thyroid cancer progression, we established TCF19[C] (rs2073724[C]) expression and TCF19[T] (rs2073724[T]) overexpression in thyroid cancer. RT-qPCR and Western blot analysis confirmed that TCF19 was strongly overexpressed. There was no significant difference in the expression level between TCF19[C] and TCF19[T] (Figure 2A and Supplementary Figure 4A). CCK-8 and EdU assays revealed that TCF19[C] overexpression promotes thyroid cancer cell proliferation, and TCF19[T] partly blocked these effects (Figure 2B, 2C). Additionally,

we also performed Transwell and wound healing assays. As we observed previously TCF19[C] significantly increased the migration and invasion of thyroid cancer cells, while overexpression of TCF19[T] notably reduced these abilities affected by TCF19[C] overexpression (Figure 2D). Taken together, TCF19 plays an oncogenic role in thyroid cancer progression. Notably, the alternative allele of the SNP is required for the effect of TCF19 on thyroid cancer progression. In addition, we also established stable TCF19-knock-down TC cells (Supplementary Figure 4A, 4B). As expected, downregulation of TCF19 suppressed the proliferation, migration, and invasion of thyroid cancer cells compared with control cells, as demonstrated by CCK-8, EdU, Transwell, and wound healing assays (Supplementary Figure 4C–4F). These results indicate that TCF19 regulates the proliferation, migration and invasion of thyroid cancer cells *in vitro*. To investigate TCF19's function in tumorigenesis *in vivo*, we conducted subcutaneous tumor formation assays using nude mice. We injected nude mice subcutaneously with C643 cells overexpressing either TCF19[C] or TCF19[T], along with a control group. The mice were euthanized 24 days post-injection to assess tumor development. Our findings revealed that TCF19[C] overexpression facilitated the growth of subcutaneous tumors. In contrast, TCF19[T] overexpression was observed to impede tumor progression when compared to TCF19[C] (Figure 2F–2H). More importantly, the C>T variant of rs2073724 diminishes the carcinogenic effect of TCF19.

Identification of TCF19[C] and TCF19[T] targets by RNA-seq

To identify the potential targets regulated by TCF19, RNA sequencing (RNA-seq) was performed in control, TCF19[C] and TCF19[T] overexpression cells. As shown in Supplementary Figure 5A, PCA analysis indicates that the different RNA-seq samples could be completely distinguished. Overexpression of TCF19[C] resulted in abnormal gene expression, with 339 upregulated genes and 424 downregulated genes. However, when TCF19[T] was overexpressed, a total of 723 genes were upregulated and 1865 genes were

downregulated in these cells compared to TCF19[C]-overexpressing cells (Figure 3A, 3B). The results indicate that TCF19[C] and TCF19[T] affect the expression of many genes in thyroid cancer cells. GO and KEGG analyses showed that the differentially expressed genes (DEGs) were mainly enriched in cancer-related programs or signalling pathways, such as the TNF signalling pathway, IL-17 signalling pathway and cytokine–cytokine receptor interaction (Figure 3C, 3D). And then we found that there were 31 genes associated with SNP (Figure 3E and Supplementary Figure 5B). Additionally, GO and KEGG analyses revealed enrichment of signatures associated with signalling receptor activator activity,

the IL-17 signalling pathway, and the TNF signalling pathway (Figure 3F, 3G). Based on these results, TCF19 appears to be a potent regulator of thyroid cancer progression, especially in regard to the inflammatory response. In addition, the SNP rs2073724 plays a crucial role in maintaining protein function.

The tumour microenvironment (TME) is a complex and evolving system. Tumour occurrence, growth and metastasis strongly depend on the internal environment of tumour cells. It is related not only to the structure, function and metabolism of tumour tissues but also to the internal environment of tumour cells [17, 18]. To further stress the role of TCF19 in regulating the TME

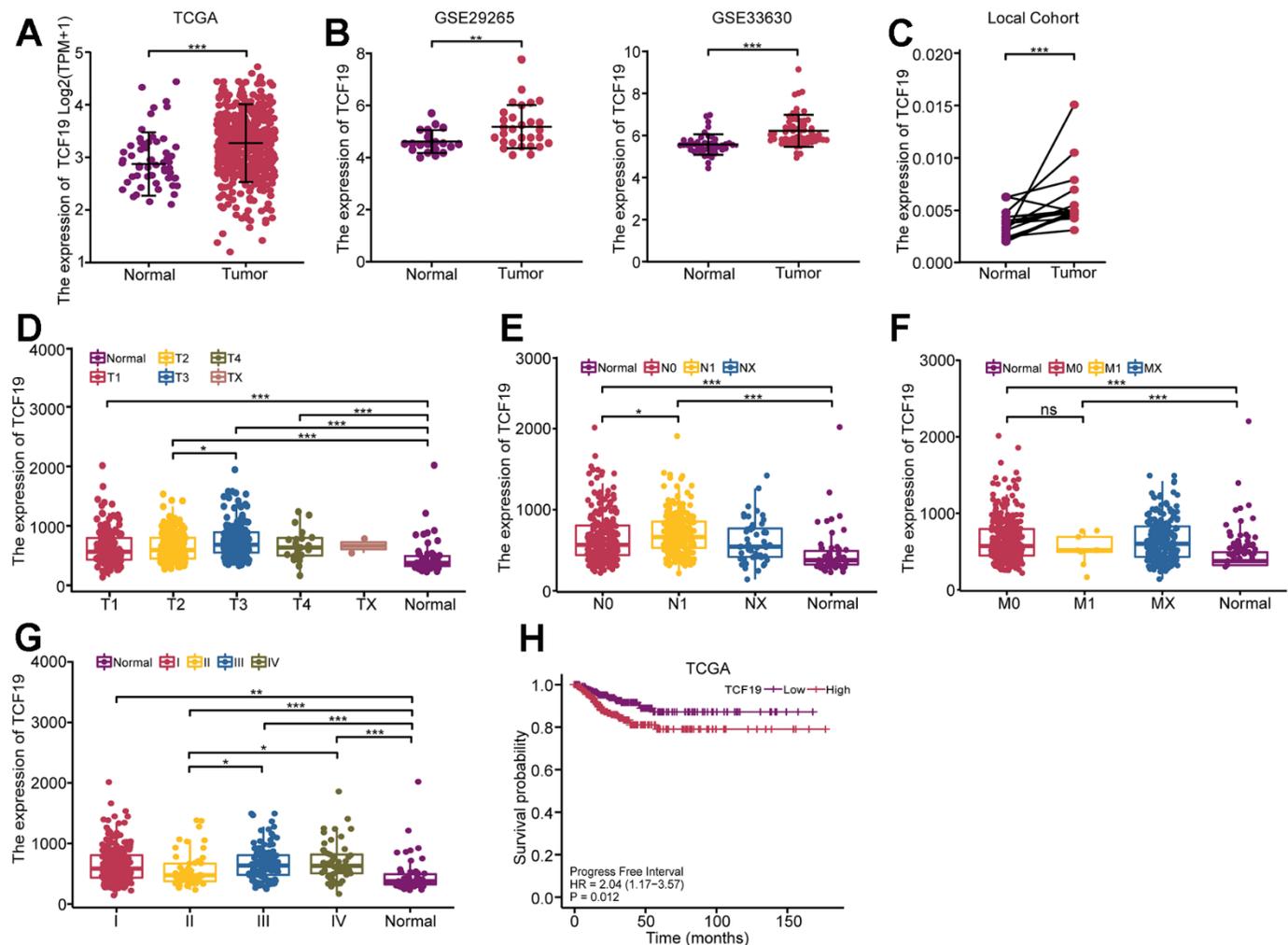


Figure 1. High expression of TCF19 is associated with poor prognosis in thyroid cancer. (A) The expression pattern of TCF19 in total was analyzed in 512 THCA tissues and 59 normal controls (TCGA database). (B) Kaplan-Meier analysis evaluating the association between Progress Free Interval and TCF19 expression in thyroid cancer patients (N=512) in the TCGA-THCA cohort. Progress Free Interval analysis was further stratified by TCF19-High and TCF19-Low characteristics for Kaplan–Meier analysis. (C) The expression pattern of TCF19 in GSE29265 and GSE33630, comparing normal thyroid tissues with thyroid cancer. (D–G) The relationship between TCF19 expression and T stage (D), N stage (E), M stage (F) and TNM stage (G) in TCGA cohort. (H) The expression pattern of TCF19 in 20 THCA tumor and their adjacent tissues. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS, not significant.

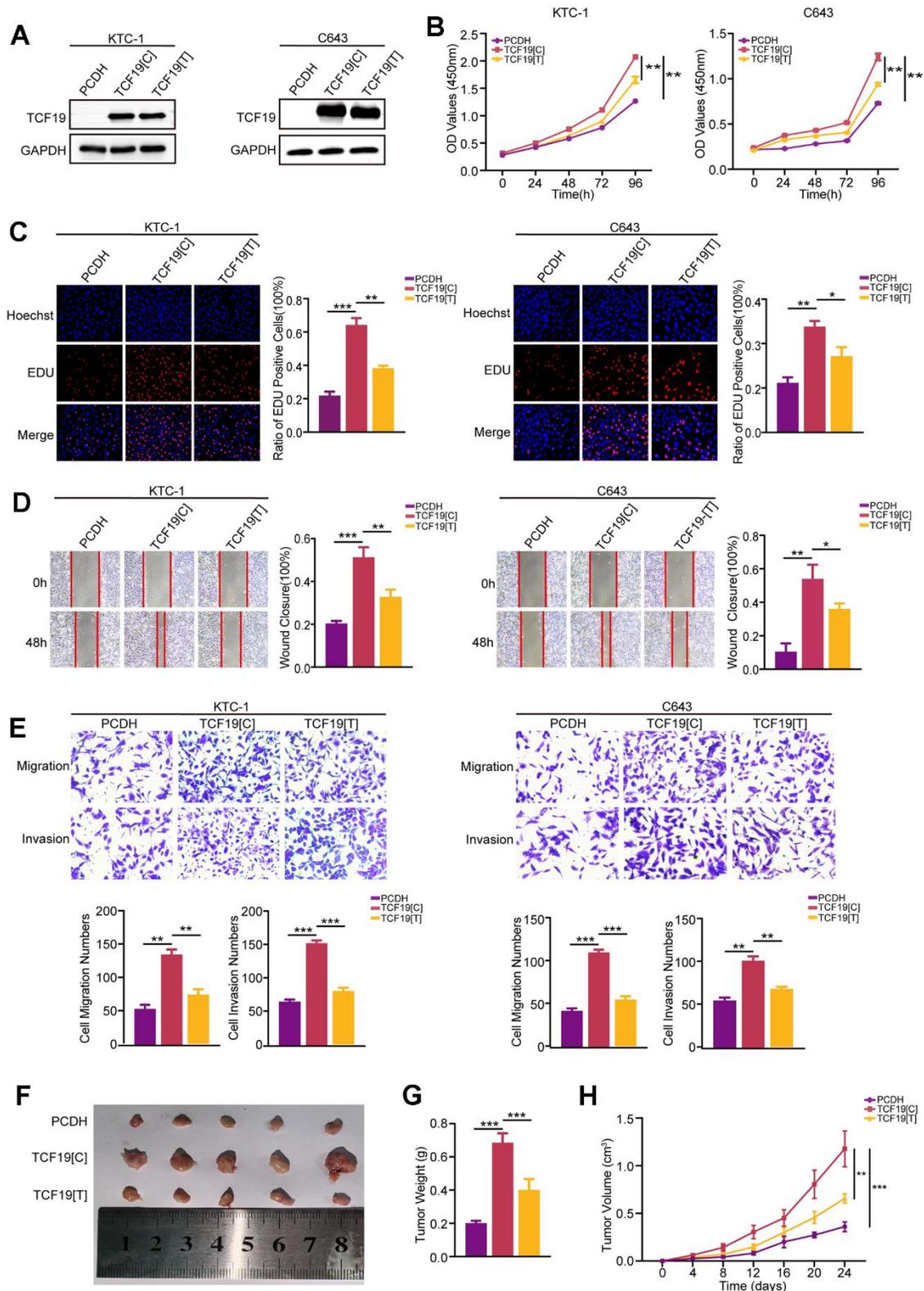


Figure 2. TCF19[C] and TCF19[T] expression impacts thyroid cancer progression. (A) TCF19 expression was detected in KTC-1 and C643 cells transfected with pcDNA, TCF19[C] and TCF19[T] plasmids by Western blotting. (B) CCK8 assay was used to detect cell viability. (C) The EdU method was used to detect DNA synthesis. (D) Wound healing assay was used to detect cell migration. (E) Transwell assay was used to detect cell invasion and migration in the cells indicated above. (F) Tumor images in tumor-bearing mouse groups. (G) Mouse tumor weights in various groups. (H) Mouse tumor volumes were measured every four days. Data are shown as means \pm S.D. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS, not significant.

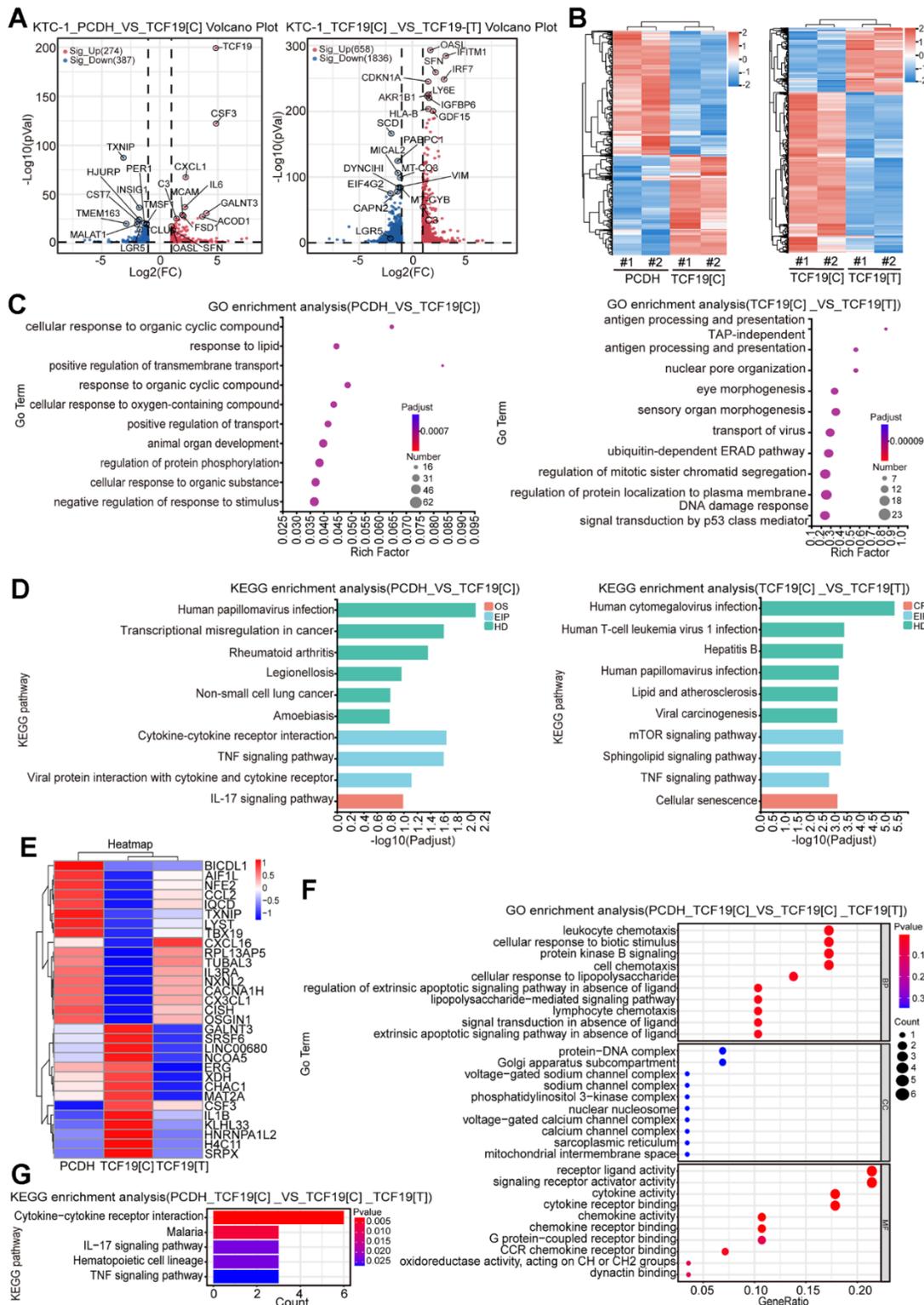


Figure 3. Identification of TCF19[C] and TCF19[T] targets by RNA-seq. (A) Volcano plots displaying differentially expressed genes (DEGs) in microarray data comparing PCDH with TCF19[C] and TCF19[C] with TCF19[T] KTC-1 cells. The numbers of significantly variant genes ($FC > 2.0$, $P < 0.05$) were shown. Vertical dashed lines indicate cut-off of FC (2.0), whereas the horizontal dashed lines indicate cut-off of p -value (0.05). (B) Heatmap of DEGs identified by RNA-seq. (C) Gene Ontology (GO) enrichment analysis of DEGs, p -value <0.05 . (D) Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs, p -value <0.05 . (E) Heatmap analyzed from 31 total genes obtained by the intersection of cluster 1 (PCDH vs TCF19[C]) and cluster 2 (TCF19[C] VS TCF19[T]). (F) GO analysis of genes described in (E). (G) KEGG analysis of genes described in (E).

of thyroid cancer. We analysed the TCGA databases and found a positive correlation between TCF19 and immune scores (Supplementary Figure 6A) and stromal scores (Supplementary Figure 6B) and a negative correlation between TCF19 and tumour purity (Supplementary Figure 6C) in thyroid cancer. We further performed xCell and EPIC analyses to explore whether TCF19 expression correlates with immune cell infiltration. The results showed that THCA samples with high TCF19 expression had higher levels of monocyte and macrophage infiltration than samples with low TCF19 expression (Supplementary Figure 6D, 6E). Next, we quantified 8 immune cell populations and 2 nonimmune stromal populations (endothelial cells and fibroblasts) by MCP-counter scoring. The results showed that the proportions of B lineage, monocytic lineage, CD8+ T cells, cytotoxic lymphocytes, fibroblasts and myeloid dendritic cells were higher, while the proportion of endothelial cells was lower in the TCF19 high group than in the TCF19 low group (Supplementary Figure 6F). We also used the Tumour Immune Dysfunction and Exclusion (TIDE) algorithm [19] to predict the response to immune checkpoint blockade therapy (ICB). In general, high TIDE score was associated with a lower ICB response rate. Our results suggest that TCF19 overexpression may confer resistance to ICB therapy (Supplementary Figure 6G).

TCF19[T] reduces its DNA binding ability, thus affecting protein function

To further validate that the SNP mutation could abolish TCF19 protein function, we performed RT-PCR to detect the downstream genes of TCF19 from the RNA-seq. The results indicated that TCF19 could significantly regulate its target genes, while TCF19[T] partly abolished this function (Figure 4A). To further stress this interesting question, we analysed the SNP mutation site of the protein. We found that the SNP mutation site was located in the proline-rich region of TCF19 (Supplementary Figure 7A, 7B). A previous study indicated that TCF19 contains a proline-rich region, which is a common characteristic of transactivating factors [20]. We noticed that the SNP was located in the proline-rich region of the TCF19 protein. Therefore, we hypothesized that the SNP mutation affects the transactivating ability of TCF19. Then, we analysed our unpublished CUT&Tag data for TCF19 and found that TCF19 could directly bind to the promoter of the SRSF6 gene (Figure 4B). To further verify whether TCF19 activates the SRSF6 promoter through direct binding, we performed chromatin immunoprecipitation (ChIP) assays. As expected, TCF19 was recruited to the promoter sites. However, TCF19[T] strongly reduced the binding ability at the promoter of SRSF6 (Figure 4C). Taken

together, TCF19 may be an exciting new therapeutic target for aggressive thyroid cancer. The alternative allele of the rs2073724 variant disrupts the oncogenic functions of TCF19, implicating the rs2073724 C>T variant as a causal genetic variant influencing thyroid cancer risk.

DISCUSSION

In this study, we first conducted a large-scale data integration of genome-wide association studies (GWAS) and phenome-wide association studies (pheWAS) to identify potential genetic risk factors for overall survival in cancer patients. The association results further overlapped with gene expression data from GTEx, which contains DNA and RNA data from samples acquired from donors postmortem across various tissue sites, and data from the TCGA project, which contains cancer samples. The gene expression level was log₂ transformed, and both fixed effect and random effect models were used for the aggregation. TCF19 and a potential deleterious missense variation (rs2073724) were identified to be associated with multiple autoimmune diseases and human cancers. Furthermore, we constructed TCF19[C] and TCF19[T] overexpression constructs and performed shRNA knockdown experiments to investigate the role of TCF19 in thyroid carcinoma cells. Additionally, we conducted RNA-seq to investigate the mechanisms of TCF19 in thyroid cancer progression.

Our meta-analysis of GWAS and pheWAS data identified an association between TCF19 expression levels and overall survival time in thyroid cancer patients. Specifically, high TCF19 expression levels were associated with a decreased overall survival time. These results suggest that TCF19 may be a potential prognostic biomarker for thyroid cancer. Moreover, *in vitro* and *vivo* functional studies highlighted that TCF19 plays an oncogenic role in thyroid cancer development, and RNA-seq indicated that TCF19 regulates many important biological processes, especially inflammation pathways. Interestingly, the SNP (rs2073724) located at TCF19 partly abolished its role in cancer progression, and GWAS data showed that the C>T variant of rs2073724 is a deleterious missense variation.

TCF19 contains a proline-rich region, a forkhead association (FHA) domain, and a PHD finger region [20], indicating that TCF19 is important transcription factor of transcriptional regulation. The FHA domain serves as a nuclear signalling domain or as a phosphoprotein binding domain [21]. Many proteins containing an FHA domain are found in the nucleus and are involved in DNA repair, cell cycle arrest, or pre-mRNA processing [22]. The PHD finger is located

at the carboxyl terminus, which may allow it to interact with chromatin via methylated histone H3 [23, 24]. A recent study showed that the PHD finger of TCF19 is necessary for regulating hepatocellular carcinoma cell proliferation through its H3K4me3 binding ability [25]. These characteristics support that TCF19 plays an important role in promoting thyroid cancer cell

proliferation as well as other human cancers. Additionally, TCF19 also contains a proline-rich region, a common characteristic of transactivating factors [20]. Importantly, we found that the missense variation (rs2073724) in our study is located in the proline-rich region of the TCF19 protein. Therefore, the TCF19 SNP may affect protein function by damaging its

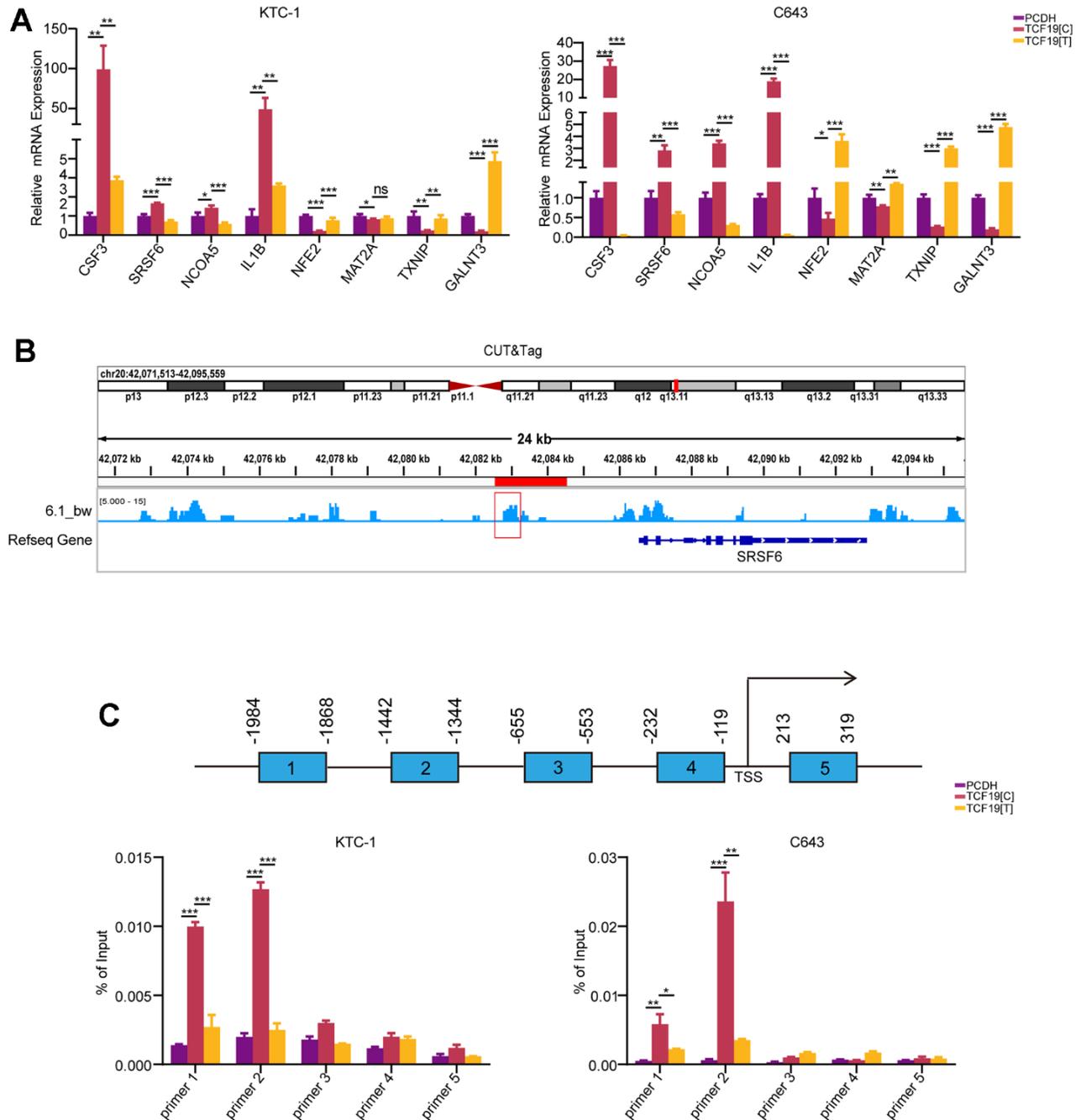


Figure 4. TCF19[T] reduces its DNA binding ability, thus affecting protein function. (A) CSF3, SRSF6, NCOA5, IL1B, NFE2, MAT2A, TXNIP, GALNT3 mRNA expression levels in cells stably transfected with PCDH, TCF19[C] and TCF19[T] plasmids were measured by qRT-PCR, respectively. (B) CUT&Tag data showed TCF19 could directly bind to the promoter of SRSF6 gene. (C) ChIP-qPCR analysis revealed potential TCF19-binding sites in the SRSF6 promoter region. After SNP mutation, the binding ability was weakened.

transactivating ability. To confirm this hypothesis, we conducted ChIP-PCR analysis, and the results demonstrated that TCF19[C] but not TCF19[T] could strongly bind to the promoter of the SRSF6 gene. Accumulated evidence has shown that SRSF6 is a potential oncogenic gene that promotes oncogenic splicing in many kinds of cancers [26, 27]. Together, our data suggest that the TCF19 SNP cannot transcriptionally activate oncogenes efficiently, thus playing a protective role in thyroid cancer progression (Figure 5).

A growing body of evidence suggests that TCF19 plays an important role in many types of cancer, including colorectal cancer, non-small cell lung cancer, renal clear cell carcinoma, hepatocellular carcinoma, and head and neck cancer [25, 28–31]. Studies report that TCF19 could promote cell proliferation and contribute to the malignant progression of cancer through mechanisms

such as regulating WWC1, inhibiting FOXO1, or activating the ATK/FOXO1 signalling pathway [28, 29, 31]. Moreover, genome-wide association studies (GWAS) have indicated that TCF19 increases the risk of certain types of cancer and multiple autoimmune diseases [32–35]. However, the underlying role of TCF19 in the pathology of thyroiditis and thyroid cancer has not been fully investigated. Our study shows that TCF19 regulates many inflammatory pathways in thyroid cancer, such as the TNF signalling pathway and IL-17 signalling pathway. TCGA analysis further demonstrated that TCF19 overexpression affects immune cell infiltration to promote thyroid cancer immune escape and ICB therapy resistance. TCF19 is a transcription factor involved in the regulation of multiple cellular signalling pathways. There is no direct evidence that TCF19 can promote immune escape. Therefore, more studies are needed to determine the specific role and mechanism of TCF19 in tumour immune escape.

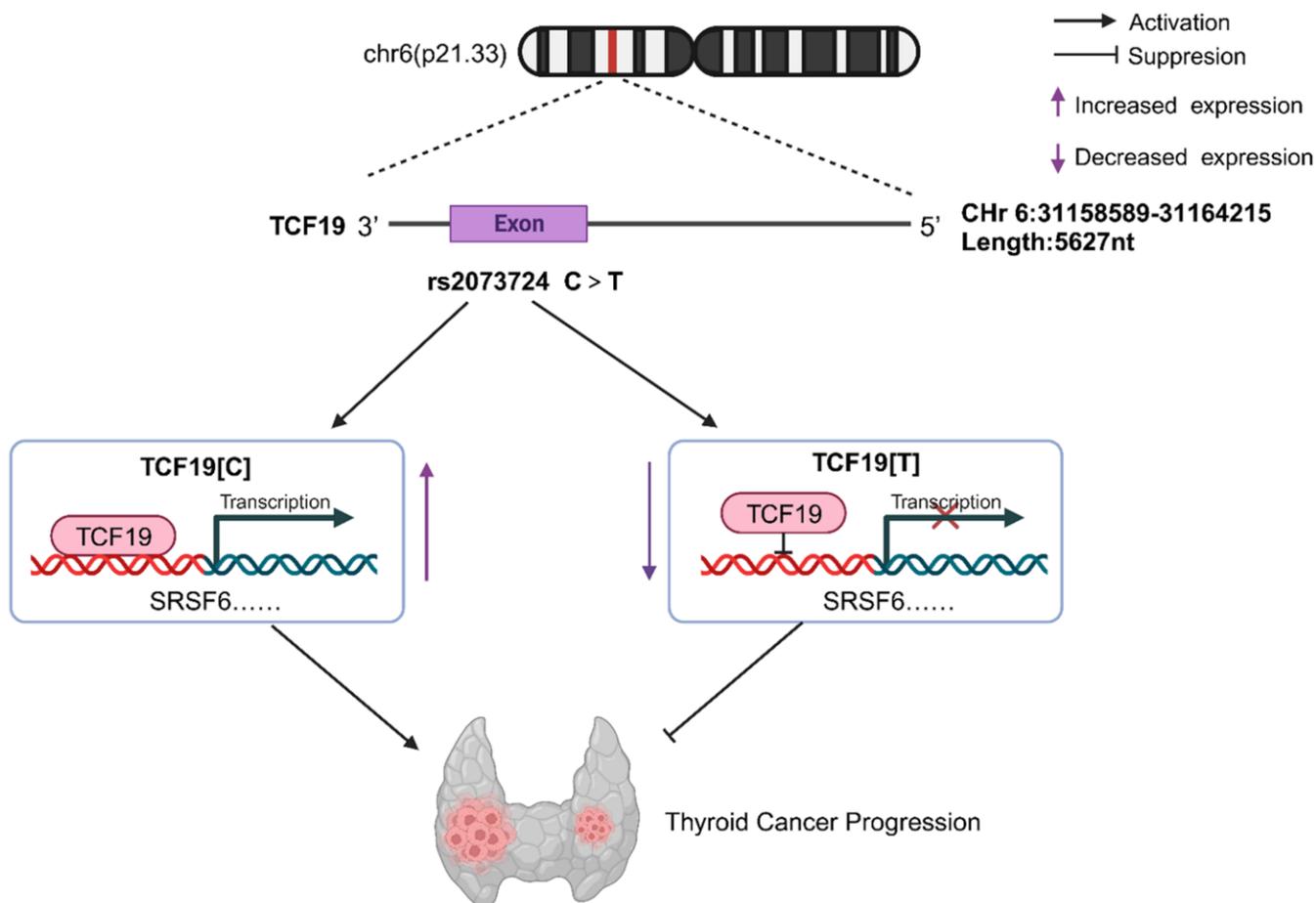


Figure 5. TCF19 on chromosome 6p21.33 can promote tumor cell proliferation, invasive and migratory ability by enhancing the transcription of SRSF6, etc., thus acting as a tumor-promoting factor in thyroid cancer, while Rs2073724 located in exon 3 of TCF19 negatively regulates the transcription and acts as a tumor suppressor. The C>T variant of rs2073724 affects the TCF19 function by damaging its transactivating ability.

Hashimoto's thyroiditis (HT) is the most common autoimmune inflammatory pathology of the thyroid and is the main cause of autoimmune hypothyroidism. The cooccurrence of Hashimoto's thyroiditis and papillary thyroid cancer is prevalent (8–36.4%) and increasing [36]. Thus, the association between thyroiditis and thyroid cancer remains an active focus of research and controversy. Both our study and other studies showed that coexistent HT has been reported to be significantly associated with less aggressive clinicopathologic characteristics of the thyroid [37–39]. In addition, one recent study highlights that thyroiditis plays a protective role in thyroid cancer progression and significantly reduces cancer-related mortality [40]. However, the important molecules that drive the development of inflammation are not particularly well understood. Our study finds that a potential deleterious missense variation (rs2073724) of TCF19 is associated with thyroiditis, and further functional studies demonstrate that this SNP plays a protective role in thyroid cancer development. Moreover, our RNA-seq data also suggest that compared with TCF19[C], TCF19[T] significantly increases the expression of proinflammatory cytokines, such as CCL2, CCL16 and CXCL1. Taken together, TCF19 SNPs could induce a strong inflammatory response in thyroid cancer patients with coexistent HT, thus playing a protective role in thyroid cancer progression.

Our study has several limitations. First, our meta-analysis is based on previously published data, and the quality of the data and potential confounding factors cannot be controlled. Second, we did not identify the TCF19 SNP as a protective factor in thyroid cancer patients with coexistent HT in the meta-analysis of GWAS and pheWAS data. There could be a few reasons: (1) The sample size was too small to detect a significant association. GWAS and pheWAS studies require large sample sizes to have enough power to detect genetic associations, especially for complex diseases such as cancer. (2) The effect size is small. It is possible that the TCF19 variant does confer some protective effect against thyroid cancer in the context of HT, but the magnitude of the effect is small. Even larger sample sizes are required to detect small effect sizes. (3) The TCF19 association is restricted to a specific subgroup. The effect of the variant could be limited to a certain subgroup of thyroid cancer/HT patients based on ethnicity, age, sex, environmental factors, etc. Therefore, the small sample size, small effect size, and subgroup-specific effects could explain the lack of significant TCF19 findings in the current study. In addition, we used an exogenous overexpression system to study the function of TCF19 SNPs in thyroid cancer progression. A CRISPR–Cas9 system to knock-in TCF19[T] in TCF19[C] thyroid cancer cells will lead to more solid

results. Finally, we also need transgenic animal models to demonstrate that TCF19[T] inhibits thyroid cancer progression and promotes inflammatory responses in thyroid cancer.

Our study first highlights that TCF19 and missense variation (rs2073724) of TCF19 are potential prognostic biomarkers for thyroid cancer patients. These findings may have important implications for the development of new therapeutic strategies and biomarkers for thyroid cancer and the identification of patients at higher risk for poor outcomes.

MATERIALS AND METHODS

Bioinformatics analysis

Previous GWAS and pheWAS studies were collected from multiple sources, including the GWAS catalogue, Opentargets and ExPheWas. The GTEx's final dataset (V8) contains DNA data from 838 postmortem donors and 17,382 RNA-seq across 54 tissue sites and two cell lines where eQTLs have been well identified and used in previous research. Cancer samples were collected from the TCGA project (N = 10,490). The gene expression level was log₂ transformed before the meta-analysis. A fixed effect model and a random effect model were both applied for the aggregation. The 95% CI was applied to show the risk and protective effect on overall survival time. To show more details for different studies, any standardized mean difference (SMD) higher than 3 and lower than -3 is shown with an arrow. Blue filled parallelograms represent the SMD for the fixed effect model and random effect model.

Cell lines and cell culture

KTC-1 and C643 cells were identified by STR analysis. All cell lines were cultured in RPMI-1640 or DMEM (Gibco, USA) supplemented with 10% foetal bovine serum, penicillin (Gibco, USA) and streptomycin (Gibco) and kept in a humid atmosphere of 5% CO₂ at 37° C.

Plasmid construction and lentivirus infection

The cDNA of the TCF19 CDS region was amplified and cloned and inserted into the lentiviral vector pCD-Puro-3×Flag. Catalytic mutant TCF19 was synthesized by GENEWIZ using the same vector as wild-type TCF19. The shRNA targeting human TCF19 was cloned and inserted into pLV-shRNA-puro-mCherry. All constructed vectors were verified by DNA sequencing. The target sequences were as follows: shTCF19-1, 5'-CCAAGGUACUUUGGUCAAUAA-3'; shTCF19-2, 5'-CCCAGGAAGAAACUCCGUGUA-3'; and shTCF19-3, 5'-ACACUGAUCCUAAACUCCAUA-3'. Plasmids

were transfected into 293T cells using PEI, Opti-MEM I reduced serum medium (Gibco, USA) and packaging vectors (PAX8 and PVSIG). Forty-eight hours and 72 hours after transfection, viral supernatants were collected, passed through a 0.45 µm filter, and infected at 50% confluency to target cells.

RNA extraction and RT-qPCR

Total RNA was isolated using RNA Extraction Solution (Servicebio, Wuhan, China), followed by cDNA synthesis using HiScript III RT SuperMix for qPCR (+g DNA wiper) (Vazyme Biotech, Nanjing, China). RNA expression levels were determined by SYBR qPCR Master Mix (Vazyme Biotech, Nanjing, China) on a Bio-Rad QX100 Droplet Digital PCR System (USA), and the RNA expression of each target was calculated by the $2^{-\Delta\Delta CT}$ method and normalized to that of β-actin. qPCR was performed as shown in Supplementary Table 1.

Western blotting

Total protein was extracted with RIPA lysis buffer (Solarbio, Beijing, China) containing protease inhibitors. Proteins were detected using BCA protein detection kit (Biosharp, Beijing, China), then separated by SDS-PAGE and migrated to PVDF membrane (Millipore, USA). The membrane was then incubated with primary antibody against GAPDH (1:5000) (Boster, USA, A00227-1), TCF19 (1:1000) (Abcam, UK, ab230005), and FLAG (1:1000) (Sigma, USA, F1804). Western blot bands were detected using a Sparkjade ECL super (Sparkjade, Shandong, China) and an imaging system (Bio-Rad, USA).

CCK-8 assay

THCA cells were seeded in 96-well plates at a density of 1×10^3 cells per well and then treated with CCK-8 reagent (APEX-BIO, USA). The absorbance was measured at 450 nm.

EdU assay

Cells were cultured in 24-well plates and treated with 1 ml medium containing 1 µl EdU (10 mM). After incubation for 2 hours at 37° C and 5% CO₂, the cells were fixed with 4% paraformaldehyde for 30 minutes and incubated with 0.5% Triton-X-100 in PBS for 20 minutes. The nuclei were stained with Hoechst 33342 dye. The proliferation rate was calculated according to the manufacturer's instructions (Bioscience, Shanghai, China). In each group, three regions were randomly selected and imaged with a fluorescence microscope (Leica, Wetzlar, Germany).

Wound healing assay

Cells were seeded and cultured until a 90% confluent monolayer formed and were then scraped with sterile pipette tips and processed in FBS-free medium as indicated in the text. Three areas were randomly selected under the microscope at 0 h and 48 h, and the distance of the cells that migrated to the scratched area was measured.

Migration and invasion assays

Migration or invasion assays were performed using 24-well plates immersed in 8.0 µm pore size Transwell filters (Corning, USA) with or without precoated diluted Matrigel (Corning, USA). THCA cells (2×10^4) in serum-free medium were added to the upper chamber, and medium containing 10% foetal bovine serum was added to the lower chamber. After incubation at 37° C for 24 h (migration) or 48 h (invasion), cells were fixed on slides and stained with crystal violet (Solarbio, Beijing, China). Infiltrating cells in 3 random fields were then counted under the microscope.

RNA-seq data analysis

Total RNA was harvested and isolated from stable TCF19 overexpression, TCF19 mutation and scramble control cells. The library was sequenced on an Illumina NovaSeq 6000. Differentially expressed genes were defined as those with fold change >2 or <0.5, $p < 0.05$ and were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis.

ChIP-PCR assay

ChIP detection was performed using the SimpleChIP® Plus Enzymatic Chromatin IP Kit (magnetic beads) (CST, USA). For ChIP-qPCR detection, the primers designed for the promoter region of SRSF6 are listed in Supplementary Table 1.

Statistical analysis

All the data presented are from at least three independent biological experiments and are shown as the mean ± SD. An unpaired two-tailed Student's t test was used for statistical analysis of the experiments. $P < 0.05$ was considered to indicate statistical significance. All data were statistically analysed with GraphPad Prism 7.0.

Abbreviations

TCF19: the transcription factor 19; GWAS: genome-wide association study; PheWAS: phenome-wide association study; HER: electronic health record;

SIFT: Sorting Intolerant from Tolerant; TCGA: The Cancer Genome Atlas; RT-Qpcr: quantitative reverse transcription PCR; PFIs: progression-free intervals; RNA-seq: RNA sequencing; DEGs: differentially expressed genes; TME: tumour microenvironment; TIDE: Tumour Immune Dysfunction and Exclusion; ChIP: chromatin immunoprecipitation; FHA: forkhead association.

AUTHOR CONTRIBUTIONS

ZXQ, NJY and GSC conceived and designed the experiments, revised the manuscript. RXH, LY, WSP and FGM performed the experiments and prepared the figures. HY and TM analysed the data. GM, ZXQ, LDP, RXH, and NJY provided the financial support. WSF and RXH wrote the paper. The manuscript includes contributions of all authors.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

ETHICAL STATEMENT

All mouse experimental procedures were approved by the Tianjin Medical University Cancer Institute and the Hospital Animal Ethics Committee (Protocol No. 2023032).

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (82372753, 82172821, 82103386, 82272721), Tianjin Municipal Science and Technology Project (19JCYBJC27400, 21JCZDJC00360), The Science and Technology Development Fund of Tianjin Education Commission for Higher Education (2021ZD033), Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-058B, TJYXZDXK-009A) and Tianjin Health Research Project (TJWJ2022XK024), The China Postdoctoral Science Foundation (2023T160481).

REFERENCES

1. Kitahara CM, Sosa JA. The changing incidence of thyroid cancer. *Nat Rev Endocrinol*. 2016; 12:646–53. <https://doi.org/10.1038/nrendo.2016.110> PMID:27418023
2. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet*. 2016; 388:2783–95. [https://doi.org/10.1016/S0140-6736\(16\)30172-6](https://doi.org/10.1016/S0140-6736(16)30172-6) PMID:27240885
3. Ruan X, Tian M, Kang N, Ma W, Zeng Y, Zhuang G, Zhang W, Xu G, Hu L, Hou X, Xie W, Gao M, Piao Y, et al. Genome-wide identification of m6A-associated functional SNPs as potential functional variants for thyroid cancer. *Am J Cancer Res*. 2021; 11:5402–14. PMID:34873468
4. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021; 71:209–49. <https://doi.org/10.3322/caac.21660> PMID:33538338
5. Xia C, Dong X, Li H, Cao M, Sun D, He S, Yang F, Yan X, Zhang S, Li N, Chen W. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J (Engl)*. 2022; 135:584–90. <https://doi.org/10.1097/CM9.0000000000002108> PMID:35143424
6. Ito Y, Miyauchi A, Kihara M, Fukushima M, Higashiyama T, Miya A. Overall Survival of Papillary Thyroid Carcinoma Patients: A Single-Institution Long-Term Follow-Up of 5897 Patients. *World J Surg*. 2018; 42:615–22. <https://doi.org/10.1007/s00268-018-4479-z> PMID:29349484
7. Luo H, Xia X, Kim GD, Liu Y, Xue Z, Zhang L, Shu Y, Yang T, Chen Y, Zhang S, Chen H, Zhang W, Li R, et al. Characterizing dedifferentiation of thyroid cancer by integrated analysis. *Sci Adv*. 2021; 7:eabf3657. <https://doi.org/10.1126/sciadv.abf3657> PMID:34321197
8. Jiao C, Li L, Zhang P, Zhang L, Li K, Fang R, Yuan L, Shi K, Pan L, Guo Q, Gao X, Chen G, Xu S, et al. REGy ablation impedes dedifferentiation of anaplastic thyroid carcinoma and accentuates radio-therapeutic response by regulating the Smad7-TGF- β pathway. *Cell Death Differ*. 2020; 27:497–508. <https://doi.org/10.1038/s41418-019-0367-9> PMID:31243343
9. Zhang W, Ruan X, Li Y, Zhi J, Hu L, Hou X, Shi X, Wang X, Wang J, Ma W, Gu P, Zheng X, Gao M. KDM1A promotes thyroid cancer progression and maintains stemness through the Wnt/ β -catenin signaling pathway. *Theranostics*. 2022; 12:1500–17. <https://doi.org/10.7150/thno.66142> PMID:35198054
10. Ruan X, Shi X, Dong Q, Yu Y, Hou X, Song X, Wei X, Chen L, Gao M. Antitumor effects of anlotinib in thyroid cancer. *Endocr Relat Cancer*. 2019; 26:153–64. <https://doi.org/10.1530/ERC-17-0558>

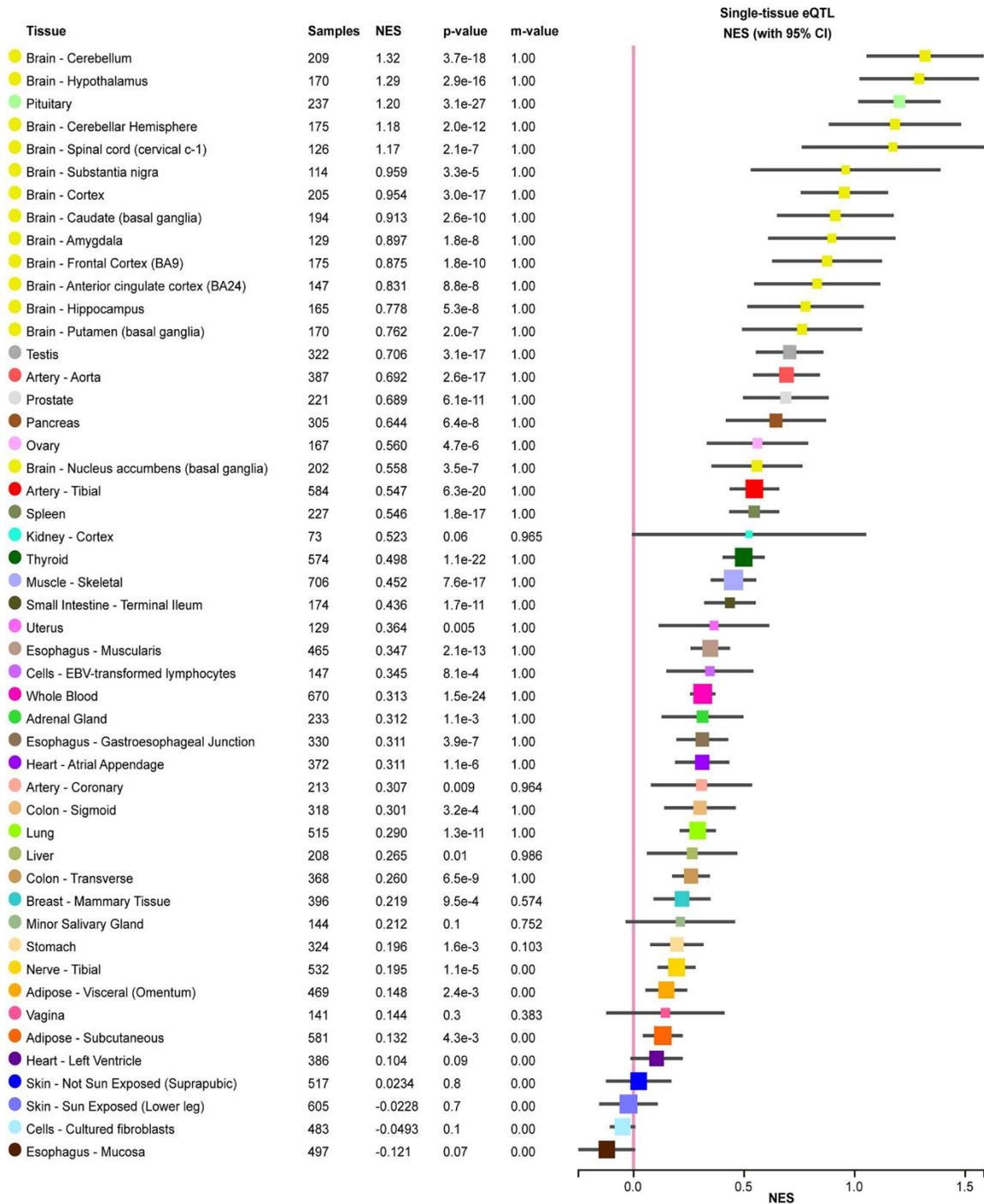
PMID:[30139768](#)

11. Zhou W, Brumpton B, Kabil O, Gudmundsson J, Thorleifsson G, Weinstock J, Zawistowski M, Nielsen JB, Chaker L, Medici M, Teumer A, Naitza S, Sanna S, et al. GWAS of thyroid stimulating hormone highlights pleiotropic effects and inverse association with thyroid cancer. *Nat Commun*. 2020; 11:3981.
<https://doi.org/10.1038/s41467-020-17718-z>
PMID:[32769997](#)
12. Gudmundsson J, Thorleifsson G, Sigurdsson JK, Stefansdottir L, Jonasson JG, Gudjonsson SA, Gudbjartsson DF, Masson G, Johannsdottir H, Halldorsson GH, Stacey SN, Helgason H, Sulem P, et al. A genome-wide association study yields five novel thyroid cancer risk loci. *Nat Commun*. 2017; 8:14517.
<https://doi.org/10.1038/ncomms14517>
PMID:[28195142](#)
13. Hwangbo Y, Park YJ. Genome-Wide Association Studies of Autoimmune Thyroid Diseases, Thyroid Function, and Thyroid Cancer. *Endocrinol Metab (Seoul)*. 2018; 33:175–84.
<https://doi.org/10.3803/EnM.2018.33.2.175>
PMID:[29947174](#)
14. Salem JE, Shoemaker MB, Bastarache L, Shaffer CM, Glazer AM, Kroncke B, Wells QS, Shi M, Straub P, Jarvik GP, Larson EB, Velez Edwards DR, Edwards TL, et al. Association of Thyroid Function Genetic Predictors With Atrial Fibrillation: A Phenome-Wide Association Study and Inverse-Variance Weighted Average Meta-analysis. *JAMA Cardiol*. 2019; 4:136–43.
<https://doi.org/10.1001/jamacardio.2018.4615>
PMID:[30673079](#)
15. Denny JC, Crawford DC, Ritchie MD, Bielinski SJ, Basford MA, Bradford Y, Chai HS, Bastarache L, Zuvich R, Peissig P, Carrell D, Ramirez AH, Pathak J, et al. Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical records for genome- and phenome-wide studies. *Am J Hum Genet*. 2011; 89:529–42.
<https://doi.org/10.1016/j.ajhg.2011.09.008>
PMID:[21981779](#)
16. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003; 31:3812–4.
<https://doi.org/10.1093/nar/gkg509> PMID:[12824425](#)
17. Fan T, Zhu M, Wang L, Liu Y, Tian H, Zheng Y, Tan F, Sun N, Li C, He J. Immune profile of the tumor microenvironment and the identification of a four-gene signature for lung adenocarcinoma. *Aging (Albany NY)*. 2020; 13:2397–417.
<https://doi.org/10.18632/aging.202269>
PMID:[33318300](#)
18. Privat-Maldonado A, Bengtson C, Razzokov J, Smits E, Bogaerts A. Modifying the Tumour Microenvironment: Challenges and Future Perspectives for Anticancer Plasma Treatments. *Cancers (Basel)*. 2019; 11:1920.
<https://doi.org/10.3390/cancers11121920>
PMID:[31810265](#)
19. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. 2018; 24:1550–8.
<https://doi.org/10.1038/s41591-018-0136-1>
PMID:[30127393](#)
20. Krautkramer KA, Linnemann AK, Fontaine DA, Whillock AL, Harris TW, Schleis GJ, Truchan NA, Marty-Santos L, Lavine JA, Cleaver O, Kimple ME, Davis DB. Tcf19 is a novel islet factor necessary for proliferation and survival in the INS-1 β -cell line. *Am J Physiol Endocrinol Metab*. 2013; 305:E600–10.
<https://doi.org/10.1152/ajpendo.00147.2013>
PMID:[23860123](#)
21. Durocher D, Henckel J, Fersht AR, Jackson SP. The FHA domain is a modular phosphopeptide recognition motif. *Mol Cell*. 1999; 4:387–94.
[https://doi.org/10.1016/s1097-2765\(00\)80340-8](https://doi.org/10.1016/s1097-2765(00)80340-8)
PMID:[10518219](#)
22. Li J, Lee GI, Van Doren SR, Walker JC. The FHA domain mediates phosphoprotein interactions. *J Cell Sci*. 2000; 113:4143–9.
<https://doi.org/10.1242/jcs.113.23.4143>
PMID:[11069759](#)
23. Shi X, Kachirskaia I, Walter KL, Kuo JHA, Lake A, Davrazou F, Chan SM, Martin DGE, Fingerma IM, Briggs SD, Howe L, Utz PJ, Kutateladze TG, et al. Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem*. 2007; 282:2450–5.
<https://doi.org/10.1074/jbc.C600286200>
PMID:[17142463](#)
24. Ramseier TM, Figge RM, Saier MH Jr. DNA sequence of a gene in *Escherichia coli* encoding a putative tripartite transcription factor with receiver, ATPase and DNA binding domains. *DNA Seq*. 1994; 5:17–24.
<https://doi.org/10.3109/10425179409039700>
PMID:[7894055](#)
25. Sen S, Sanyal S, Srivastava DK, Dasgupta D, Roy S, Das C. Transcription factor 19 interacts with histone 3 lysine 4 trimethylation and controls gluconeogenesis via the nucleosome-remodeling-deacetylase complex. *J Biol Chem*. 2017; 292:20362–78.

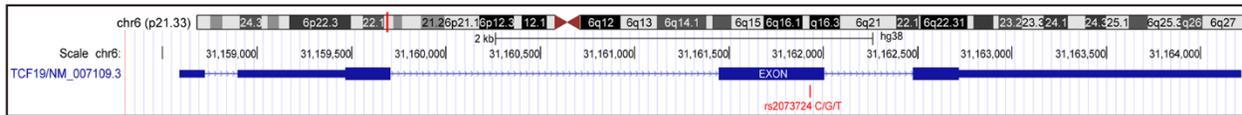
- <https://doi.org/10.1074/jbc.M117.786863>
PMID:[29042441](https://pubmed.ncbi.nlm.nih.gov/29042441/)
26. She W, Shao J, Jia R. Targeting Splicing Factor SRSF6 for Cancer Therapy. *Front Cell Dev Biol.* 2021; 9:780023.
<https://doi.org/10.3389/fcell.2021.780023>
PMID:[34917618](https://pubmed.ncbi.nlm.nih.gov/34917618/)
27. Wan L, Yu W, Shen E, Sun W, Liu Y, Kong J, Wu Y, Han F, Zhang L, Yu T, Zhou Y, Xie S, Xu E, et al. SRSF6-regulated alternative splicing that promotes tumour progression offers a therapy target for colorectal cancer. *Gut.* 2019; 68:118–29.
<https://doi.org/10.1136/gutjnl-2017-314983>
PMID:[29114070](https://pubmed.ncbi.nlm.nih.gov/29114070/)
28. Du WB, Huang Z, Luo L, Tong SP, Li HQ, Li X, Tong JH, Yao YL, Zhang WB, Meng Y. TCF19 aggravates the malignant progression of colorectal cancer by negatively regulating WWC1. *Eur Rev Med Pharmacol Sci.* 2020; 24:655–63.
https://doi.org/10.26355/eurrev_202001_20042
PMID:[32016966](https://pubmed.ncbi.nlm.nih.gov/32016966/)
29. Cheng X, Hou J, Wen X, Dong R, Lu Z, Jiang Y, Wu G, Yuan Y. Immunotherapeutic Value of Transcription Factor 19 (TCF19) Associated with Renal Clear Cell Carcinoma: A Comprehensive Analysis of 33 Human Cancer Cases. *J Oncol.* 2022; 2022:1488165.
<https://doi.org/10.1155/2022/1488165>
PMID:[36111242](https://pubmed.ncbi.nlm.nih.gov/36111242/)
30. Zhou ZH, Chen G, Deng C, Tang JM, Xie L, Zhou HY, Ye X, Zhang DK, Shi RQ, Tian D, Qiao GB, Ben XS. TCF19 contributes to cell proliferation of non-small cell lung cancer by inhibiting FOXO1. *Cell Biol Int.* 2019; 43:1416–24.
<https://doi.org/10.1002/cbin.11189>
PMID:[31141247](https://pubmed.ncbi.nlm.nih.gov/31141247/)
31. Ji P, Chang J, Wei X, Song X, Yuan H, Gong L, Li Y, Ding D, Zhang E, Yan C, Zhu M, Miao X, Wu C, et al. Genetic variants associated with expression of TCF19 contribute to the risk of head and neck cancer in Chinese population. *J Med Genet.* 2022; 59:335–45.
<https://doi.org/10.1136/jmedgenet-2020-107410>
PMID:[34085947](https://pubmed.ncbi.nlm.nih.gov/34085947/)
32. Rashkin SR, Graff RE, Kachuri L, Thai KK, Alexeeff SE, Blatchins MA, Cavazos TB, Corley DA, Emami NC, Hoffman JD, Jorgenson E, Kushi LH, Meyers TJ, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun.* 2020; 11:4423.
<https://doi.org/10.1038/s41467-020-18246-6>
PMID:[32887889](https://pubmed.ncbi.nlm.nih.gov/32887889/)
33. Tomer Y, Dolan LM, Kahaly G, Divers J, D’Agostino RB Jr, Imperatore G, Dabelea D, Marcovina S, Black MH, Pihoker C, Hasham A, Hammerstad SS, Greenberg DA, et al, and SEARCH for Diabetes in Youth Study. Genome wide identification of new genes and pathways in patients with both autoimmune thyroiditis and type 1 diabetes. *J Autoimmun.* 2015; 60:32–9.
<https://doi.org/10.1016/j.jaut.2015.03.006>
PMID:[25936594](https://pubmed.ncbi.nlm.nih.gov/25936594/)
34. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, Caporaso NE, Johansson M, Xiao X, Li Y, Byun J, Dunning A, Pooley KA, et al, and SpiroMeta Consortium. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet.* 2017; 49:1126–32.
<https://doi.org/10.1038/ng.3892> PMID:[28604730](https://pubmed.ncbi.nlm.nih.gov/28604730/)
35. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, Ishigaki K, Suzuki A, Suzuki K, et al, and FinnGen. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet.* 2021; 53:1415–24.
<https://doi.org/10.1038/s41588-021-00931-x>
PMID:[34594039](https://pubmed.ncbi.nlm.nih.gov/34594039/)
36. Feldt-Rasmussen U. Hashimoto’s thyroiditis as a risk factor for thyroid cancer. *Curr Opin Endocrinol Diabetes Obes.* 2020; 27:364–71.
<https://doi.org/10.1097/MED.0000000000000570>
PMID:[32773575](https://pubmed.ncbi.nlm.nih.gov/32773575/)
37. Huang D, Zhi J, Zhang J, Qin X, Zhao J, Zheng X, Gao M. Relationship Between Thyroid Autoantibodies and Recurrence of Papillary Thyroid Carcinoma in Children and Adolescents. *Front Oncol.* 2022; 12:883591.
<https://doi.org/10.3389/fonc.2022.883591>
PMID:[35756669](https://pubmed.ncbi.nlm.nih.gov/35756669/)
38. Xu J, Ding K, Mu L, Huang J, Ye F, Peng Y, Guo C, Ren C. Hashimoto’s Thyroiditis: A “Double-Edged Sword” in Thyroid Carcinoma. *Front Endocrinol (Lausanne).* 2022; 13:801925.
<https://doi.org/10.3389/fendo.2022.801925>
PMID:[35282434](https://pubmed.ncbi.nlm.nih.gov/35282434/)
39. Rotondi M, Gropelli G, Croce L, Latrofa F, Ancona G, Coperchini F, Pasquali D, Cappelli C, Fugazza A, Guazzoni V, Radetti G, Chiovato L. Patients with chronic autoimmune thyroiditis are not at higher risk for developing clinically overt thyroid cancer: a 10-year follow-up study. *Eur J Endocrinol.* 2020; 183:317–23.
<https://doi.org/10.1530/EJE-20-0350> PMID:[32717718](https://pubmed.ncbi.nlm.nih.gov/32717718/)
40. Xu S, Huang H, Qian J, Liu Y, Huang Y, Wang X, Liu S, Xu Z, Liu J. Prevalence of Hashimoto Thyroiditis in Adults With Papillary Thyroid Cancer and Its Association With Cancer Recurrence and Outcomes. *JAMA Netw Open.* 2021; 4:e2118526.
<https://doi.org/10.1001/jamanetworkopen.2021.18526>
PMID:[34313737](https://pubmed.ncbi.nlm.nih.gov/34313737/)

SUPPLEMENTARY MATERIALS

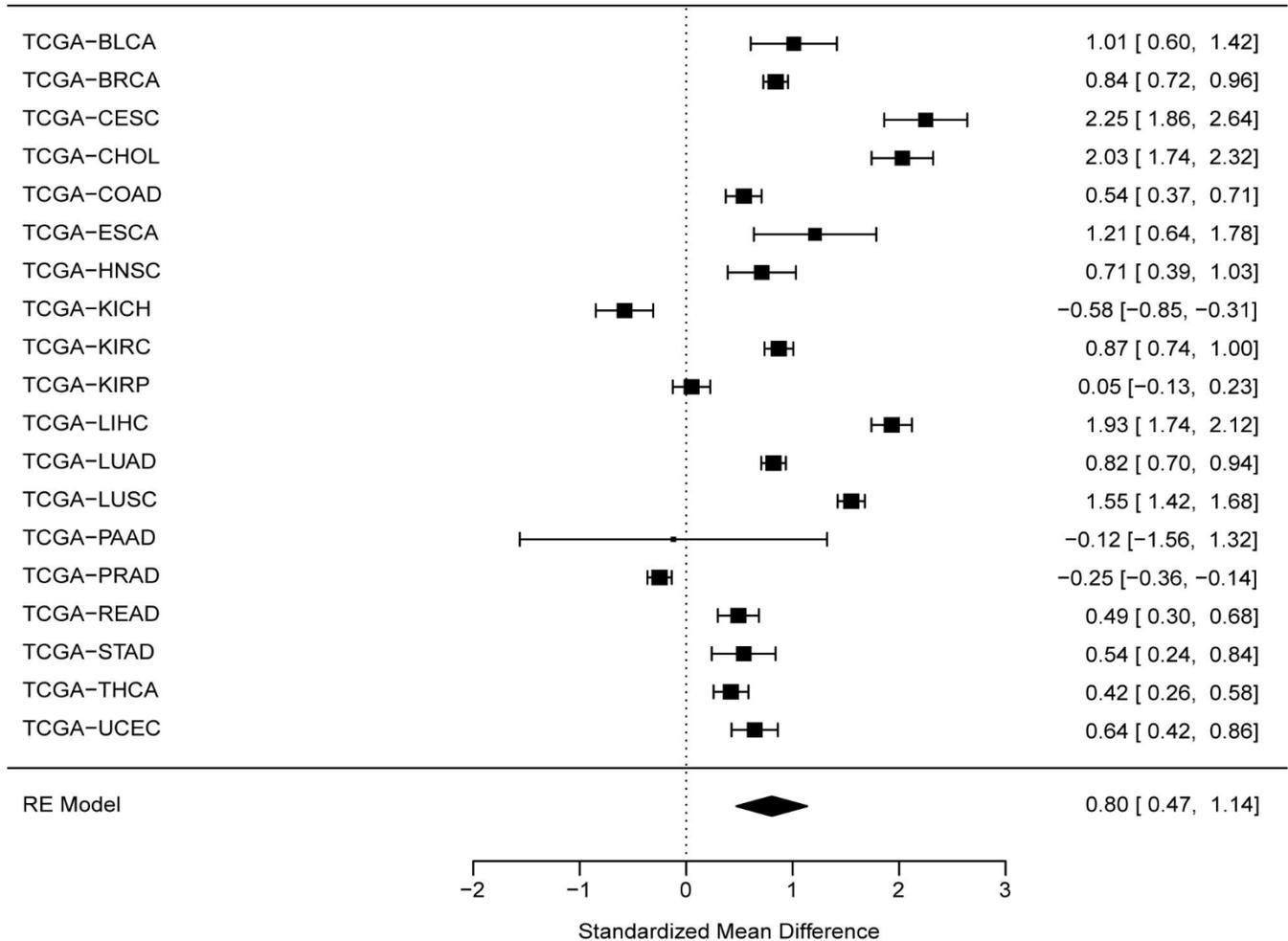
Supplementary Figures



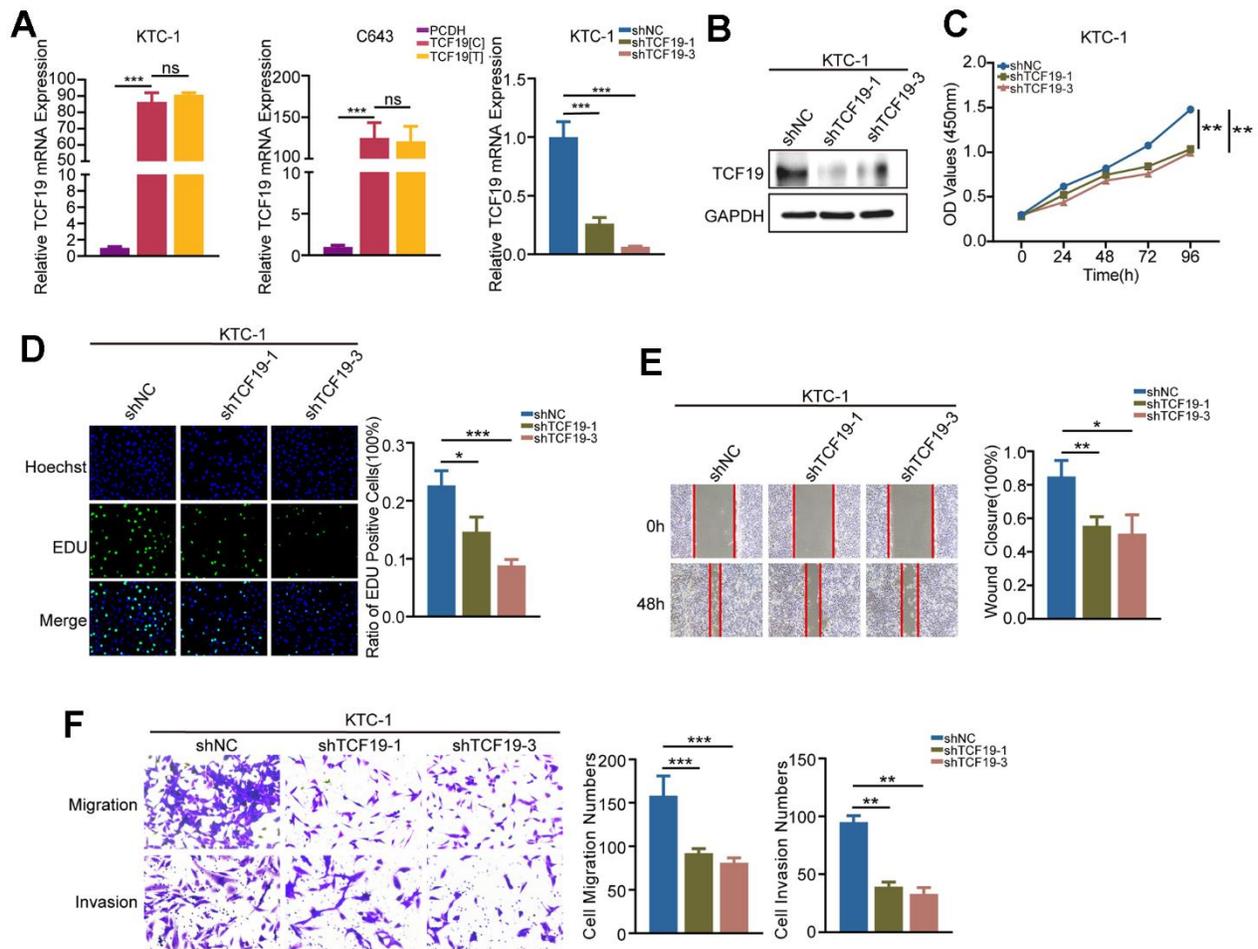
Supplementary Figure 1. Positive expression regulation to TCF19 caused by rs2073724 with GTEx dataset. The GTEx’s final dataset (V8) contains DNA data from 838 postmortem donors and 17,382 RNA-seq across 54 tissue sites and two cell lines. In this project, eQTL have been identified with genomics and expression data. GTEx dataset shows rs2073724 alternative allele (Leu) will increase the expression level of TCF19 compared with wild type (Pro) in almost all human tissues (thyroid tissue, NES=0.99, P=9.6x10⁻⁶⁶) the causal relationship between alternative allele with increased etiology of thyroiditis and thyroid cancer.



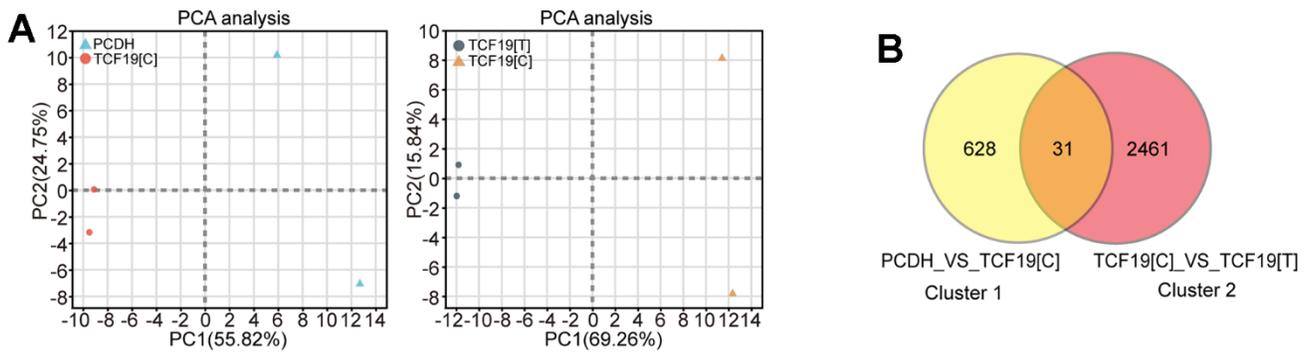
Supplementary Figure 2. RS1317082 is located in the third exon of TCF19 in UCSC.



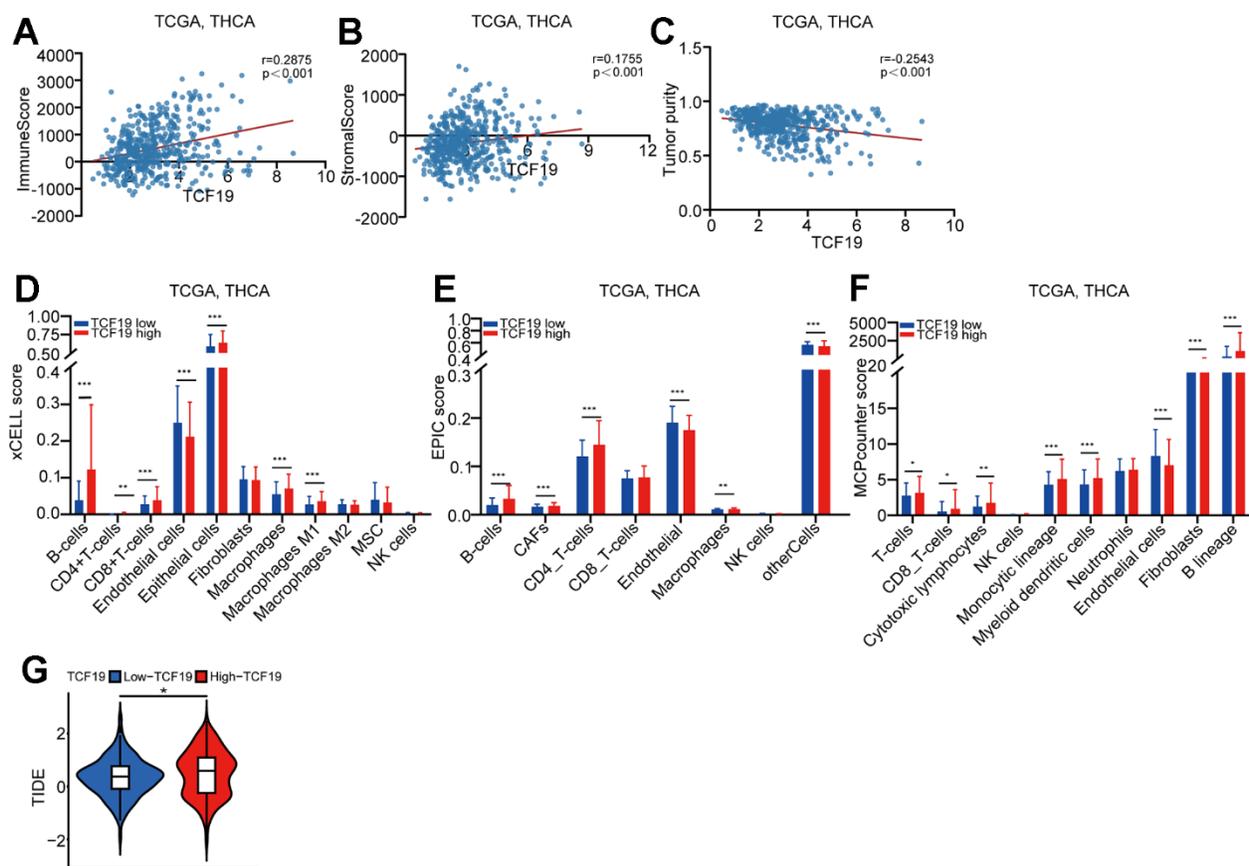
Supplementary Figure 3. Over-expression of TCF19 in human cancers showed with TCGA dataset. Pan-cancer down-regulation of TCF19 crosses 23 types of human cancers. Cancer samples were collected from the TCGA project (N = 10,490). Gene expression level was log2 transformed before the meta-analysis. Both fixed effect model and random effect model were applied for the aggregation. 95% CI was applied to show the risk and protective effect to overall survival time. In order to show more details for different studies, any standardized mean difference (SMD) higher than 3 and lower than -3 was showed with arrow. Blue filled parallelograms represent the SMD for fixed effect model and random effect model.



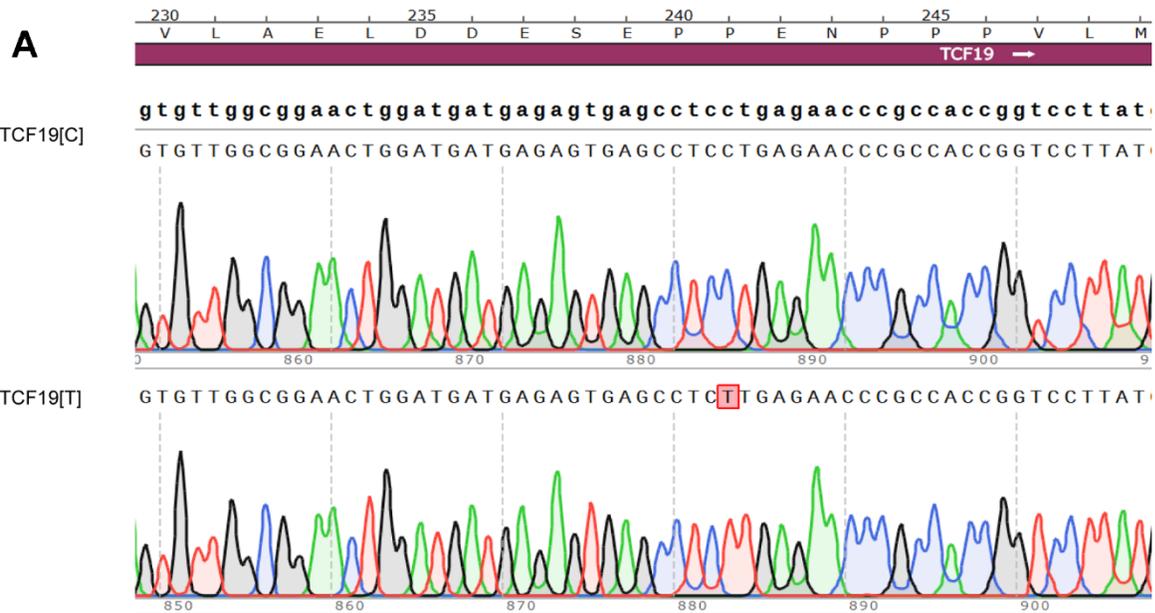
Supplementary Figure 4. Silencing of TCF9 inhibits the proliferation, migration and invasion of thyroid cancer cells. (A) TCF9 mRNA expression was detected in KTC-1 and C643 cells transfected with pcDNA, TCF9[C] and TCF9[T] plasmids by qRT-PCR. (B) Western blot was used to detect the protein expression of TCF9 in KTC-1 cells stably transfected with PCDH and shTCF9 plasmids. (C) CCK8 assay was used to detect cell viability, (D) The EdU method was used to detect DNA synthesis. (E) Wound healing assay was used to detect cell migration. (F) Transwell assay was used to detect cell invasion and migration in KTC-1 cells stably transfected with PCDH and shTCF9 plasmids. Data are shown as means \pm S.D. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS, not significant.



Supplementary Figure 5. Multivariate statistical analysis of RNA-seq samples from different origins. (A) PCA score plot based on RNA-seq data. (B) Microarray data revealed overlapping cluster 1 (PCDH vs TCF19[C]) and cluster 2 (TCF19[C] VS TCF19[T]) ($FC > 2.0$, $P < 0.05$) in KTC-1 cells, all 31 differential genes listed.



Supplementary Figure 6. TCF19 regulates the TME of thyroid cancer. (A–C) Correlation analysis of TCF19 expression and three items, including immune cell score, stromal score and tumor purity in THCA tissues from TCGA. (TCGA THCA RNA-seq, $n = 505$, Pearson correlation analysis). (D–G) XCell, EPIC and MCPcounter analysis of the microenvironmental cells populations and TIDE prediction score for ICB therapy in TCF19 high tissues and TCF19 low tissues. p -values were calculated using the Mann-Whitney test.



B transcription factor 19 [Homo sapiens]

1 20 40 70
M L P C F Q L L R I G G G R G G D L Y T F H P P A G A G C T Y R L G H R A D L C D V A L R P Q Q E P G L I S G I H A E L H A E P R G D D W R

71 91 111 140
V S L E D H S S Q G T L V N N V R L P R G H R L E L S D G D L L T F G P E G P P G T S P S E F Y F M F Q Q V R V K P Q D F A A I T I P R S R

141 161 181 210
G E A R V G A G F R P M L P S Q G A P Q R P L S T F S P A P K A T L I L N S I G S L S K L R P Q P L T F S P S W G G P K S L P V P A P P G E

211 231 241 251 280
M G T T P S A P P Q R N R R K S V H R V L A E L D D E S E P E N P P P V L M E P R K K L R V D K A P L T P T G N R R G R P R K Y P V S A P

281 301 345
M A P P A V G G G E P C A A P C C L P Q E E T V A W V Q C D G C D V F H V A C V G C S I Q A A R E A D F R C P G C R A G I Q T

Supplementary Figure 7. TCF19[T] reduces its DNA binding ability, thus affecting protein function. (A) Plasmid sequencing peak maps verified TCF19 mutation (P241A) from SnapGene. (B) The amino acid sequence of TCF19 protein was from NCBI.

Supplementary Table

Supplementary Table 1. List of primers.

Gene	Forward primer	Reverse primer
Human β -actin	CACCATTGGCAATGAGCGGTTC	AGGTCTTTGCGGATGTCCACGT
Human TCF19	TCAGCCTGGAAGACCACAGCAG	CCAAAGGTCAGGAGGTCTCCAT
Human CSF3	TCCAGGAGAAGCTGGTGAGTGA	CGCTATGGAGTTGGCTCAAGCA
Human SRSF6	GTGCTTTGGACAAACTGGATGGC	CTCCTACTTCGTGACCGTCTTC
Human NCOA5	AGATTCACCGCTCCTGCACAGT	CTGTCTGGCAATCTCCTCACGT
Human IL1B	CCACAGACCTTCCAGGAGAATG	GTGCAGTTCAGTGATCGTACAGG
Human NFE2	GGAGAGATGGAAGTACTGACTTGGC	GAATCTGGGTGGATTGAGCAGG
Human MAT2A	CTGGCAGAACTACGCCGTAATG	GTGTGGACTCTGATGGGAAGCA
Human TXNIP	CAGCAGTGCAAACAGACTTCGG	CTGAGGAAGCTCAAAGCCGAAC
Human GALNT3	CACCTGCAATACTGCTGAAGG	ACAGAGGTTCTAGCCAACCAT
Human SRSF6 Primer 1	GGTTCCTTCCTCCATCTTCAAAGC	AGAGCCACAGTCAGAGAAGCAG
Human SRSF6 Primer 2	GCGGTCCAGCCAGCAGTG	GGGTGACAAAGCAAGACTCCAAC
Human SRSF6 Primer 3	GATAGCCCAATCCACAAGTGAGTC	CCGCCGACATCAGCCTCTC
Human SRSF6 Primer 4	CTGCGGCTGGATTAGAGACATTG	CCCTGGGCTAGATGGTGACTG
Human SRSF6 Primer 5	GCCGCCTCCTCGAAGTAGAC	CCGCTTCTTTCCTTGGAGAGTTC