Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cancer-derived IGKC Suppresses Interferon Signaling and Promotes Tumor Metastasis

Chaoyang Sun

suncydoctor@gmail.com

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

xucui zhuang

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Zixuan Fang

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

Yiting Liu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Junpeng Fan

Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology https://orcid.org/0000-0002-9413-408X

Xin Hou

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Hao Li

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Xingyuan Hu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Jingrui Jiang

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Wanwan Luo

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Di Wu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Zimeng Wu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Chenzhao Feng

National Clinical Research Center for Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Zhe Hu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Chun Gao

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology,

Yi Ding

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Tianyu Qin

Tongji Hospital

Ruqi Chen

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Dongli Kong

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Yingjian Zhou

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Ding Ma

Cancer Biology Research Center (Key laboratory of the ministry of education)

Nan Yu

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology

Gang Chen

Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology
https://orcid.org/0000-0002-8320-6594

Article

Keywords:

Posted Date: June 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9600876/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: (Not answered)

Cancer-derived IGKC Suppresses Interferon Signaling and Promotes Tumor Metastasis

Xucui Zhuang[1,2,3], Zixuan Fang[1,2,3],Yiting Liu[1,2,3], Junpeng Fan[1,2,3], Xin Hou[1,2], Hao Li[1,2], Xingyuan Hu[1,2], Jingrui Jiang[1,2], Wanwan Luo[1,2], Di Wu[1,2], Zimeng Wu[1,2], Chenzhao Feng[1,2], Zhe Hu[1,2], Chun Gao[1,2], Yi Ding[1,2], Tianyu Qin[1,2], Ruqi Chen[1,2], Dongli Kong[1,2], Yingjian Zhou[1,2], Ding Ma[1,2], Nan Yu[1,2]*, Gang Chen[1,2]*, Chaoyang Sun[1,2,4]*

[1] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

[2] National Clinical Research Center for Obstetrics and Gynecology, Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

[3] These authors contributed equally

[4] Lead contact

*Correspondence: nny86@126.com (N. Y.) , tjchengang@hust.edu.cn (G. C.), suncydoctor@gmail.com (C. S.)

Declaration of interests

The authors declare no potential conflicts of interest.

Abstract

Immunoglobulins (Igs) have traditionally been recognized for their role in antibody production by B cells. However, their expression in cancer cells and their contributions to tumorigenesis and disease progression remain poorly understood. In this study, using single-cell resolution analysis, we identified a distinct subset of tumor cells with high immunoglobulin gene expression, particularly IGKC, which was significantly enriched in metastatic tumors. The overexpression of IGKC in metastatic tumors was further validated by in situ hybridization across multiple types of clinical tumor samples. Integrated analyses of whole-genome and transcriptome sequencing from pan-cancer cell lines in the CCLE revealed that, beyond B-cell lymphomas, a subset of tumor cells from most carcinoma types also overexpress IGKC. Unlike in the B-cell lineage, IGKC expression in tumor cells is driven by alternative splicing, generating fusion transcripts rather than the conventional DNA rearrangements seen in B cells. CUT&Tag and RNA sequencing demonstrated that IGKC transcription is directly regulated by TFAP2C, a transcription factor highly expressed in tumors and associated with poor prognosis. Mechanistically, IGKC interacts with hnRNPM, reducing the accumulation of double-stranded RNA, suppressing interferon responses, and ultimately promoting metastatic progression. In vivo, IGKC overexpression markedly accelerated tumor progression in multiple mouse models. Together, these findings identify tumor-derived IGKC as a key driver of metastasis and highlight the TFAP2C–IGKC–hnRNPM axis as a promising target for prognostic assessment and therapeutic intervention in metastatic cancer.

Introduction

Immunoglobulins (Igs) have traditionally been recognized for their role in antibody production in B cells[1]. Although previous studies have reported that Igs can also be expressed in non-B cells[2, 3], this phenomenon has received limited attention and remains poorly understood. Only in recent years, emerging research has revealed that tumor cells are capable of secreting Igs, and cancer-derived IgG contributes to tumor progression and survival[4-6]. However, the precise role and underlying mechanisms of Igs in tumor development and metastasis remain largely unclear. Due to the extensive infiltration of B cells within tumor tissues, studies on the transcriptional patterns and functions of IgG in tumor cells have been largely limited. The advent of single-cell RNA sequencing (scRNA-seq), coupled with advanced computational analyses, has enabled high-resolution profiling of cancer-derived Igs across individual cells, epithelial subpopulations, and diverse tumor types.

IgG molecules secreted by B cells are composed of two heavy chains (55 kDa) and two light chains (25 kDa)[7]. The first functional domain at the N-terminus of either the heavy or light chain varies among different IgG molecules and is referred to as the variable region, whereas the remaining domains are conserved and collectively termed the constant region. IgG variable regions are generated by rearrangement of V, D, and J gene segments-one V, D, and J for the heavy chain, and one V and J for the light chain. Multiple segment copies together ensure antibody diversity[8, 9]. However, the expression pattern of IgG in tumor cells and whether it exhibits similar characteristics across different tumor types remain unexplored.

In this study, we found that tumor cells aberrantly express immunoglobulin κ constant (IGKC) through a non-canonical mechanism that promotes metastatic progression. Single-cell transcriptomic analyses identified an IGKC-high tumor cell subset enriched in metastatic lesions, which was validated in clinical specimens. Pan-cancer analyses of CCLE datasets revealed widespread IGKC expression beyond B-cell malignancies. Mechanistically, tumor-derived IGKC is generated by alternative splicing rather than immunoglobulin DNA rearrangement and is directly transcriptionally activated by TFAP2C. Functional assays demonstrate that IGKC interacts with the RNA-binding protein hnRNPM, limiting double-stranded RNA accumulation and suppressing interferon signaling. Consequently, IGKC expression enhances tumor growth and metastasis in multiple in vivo models. Together, these findings uncover a previously unrecognized TFAP2C-IGKC-hnRNPM axis that enables immune evasion and drives metastatic competence in cancer.

Results

1. IG-featured tumor cells were enriched in tumors with metastasis

To investigate lymph node metastasis in cervical cancer, we collected 18 samples for single-cell RNA sequencing, including 3 normal cervix tissue, 9 primary tumors without metastasis, 6 primary tumors with metastasis, and 3 matched metastatic lymph nodes form the 6 metastasis patients **(Fig. 1A)**. 214,490 high-quality cells identified 14 distinct clusters **(Fig. 1B and Fig. S1A-S1C)** and malignant epithelial cells were classified by inferCNV analysis **(Fig. S1D)**. After stringent quality control, single-cell transcriptomic data of 27,422 tumor cells were divided into 8 subgroups and visualized using a UMAP plot (**Fig. 1C**). Among these subpopulations, the KRT6B-epithelial (KRT6B-epi) cluster resembled the previously described HNSCC epithelial differentiation state 2 (Epi diff.2)[10], characterized by enrichment of keratinization-related genes such as *KRT6A* and *SPRR1B*, hallmark of squamous cell carcinoma. The MUC20-epithelial (MUC20-epi) cluster corresponded to an epithelial state previously reported to exhibit mixed keratinization and glandular differentiation, with high expression of *MUC20, MUC1, MUC*4, and *SPRR3*[11]. Notably, this cluster displayed enrichment of IFNγ and MHCI associated gene signatures **(Fig. S1F)**. Additional epithelial states, the pEMT-epithelial (pEMT-epi) cluster closely resembled the partial epithelial-mesenchymal transition (pEMT) state described in head and neck squamous cell carcinoma (HNSCC)[10], and ISG15-epi, Prolif-epi, Stress-epi and Metabolism-epi were consistent with epithelial programs previously identified across multiple cancer types in pan-cancer analyses[12]. Notably, we identified a distinct tumor cell subset (C1)

characterized by high expression of immunoglobulin genes (*IGKC, IGHG1, IGHG4*, etc.) together with keratin markers (KRT6A), which we designated as the IG-epithelial (IG-epi) subgroup. Because Igs genes are highly expressed in B cells, to exclude potential confounding effects from contaminating B cells and ensure accurate tumor cell annotation, we applied the B-lineage and plasma-lineage gene signature (*PTPRC, CD79A, CD19, MS4A1, JCHAIN, CD38, CD27 and PRDM1*) to score all cells and determined a stringent cutoff, retaining only cells below this threshold for downstream analysis **(Fig. S1E)**. We next compared the distribution of the eight epithelial subpopulations across non-metastatic primary tumor cells (no-met primT), metastatic primary tumor cells (met primT), and metastatic lymph node tumor cells (Met LN). This analysis revealed that MUC20-epi was enriched in metastatic lymph nodes within an immune cell infiltrated environment, which was consistent with the conclusion of CESC immune epithelium (Epi-Imm) more relevant to immune neighborhoods in previous report[11, 13]. KRT6B-epi cells were particularly enriched in metastatic primary tumors, which likely represent a highly differentiated squamous epithelial state associated with advanced tumor progression[14]. MUC20-epi and KRT6B-epi cells showed distinct distributions between metastatic lymph nodes and primary tumor tissues, which may reflect changes in tumor cell differentiation and epithelial characteristics following metastasis to the lymph nodes. Importantly, the IG-epi subgroup was markedly enriched in metastatic primary tumor cells and metastatic lymph node tumor cells, with particularly high abundance observed in metastatic lymph nodes **(Fig. 1F)**.

To further define metastasis-associated transcriptional programs, we performed differential gene expression analysis between tumor cells from metastatic and non-metastatic tumors. This analysis revealed marked upregulation of keratinization-related genes (*KRT3, SPRR3*, etc.) and Igs genes (*IGKC, IGHG4, IGLC2*, etc.) in metastatic tumors, with IGKC showing the most pronounced increase **(Fig. 1G and Fig. S1G)**. To assess the generalizability of these findings, we collected publicly available single-cell datasets from solid tumors that included paired primary tumors and lymph node metastases. In total, we identified five suitable datasets, comprising head and neck squamous cell carcinoma (HNSCC)[15], breast cancer (BRCA) [12], lung adenocarcinoma (LUAD)[16], hepatocellular carcinoma (HCC)[17], and gastric cancer (GC)[18]. The results showed that Ig genes were upregulated in metastatic primary tumors and metastatic lymph node lesions of breast cancer, head and neck squamous cell carcinoma, and hepatocellular carcinoma, while in gastric cancer and lung adenocarcinoma, certain Ig genes such as IGLC, IGHA, and IGHG were more highly expressed in non-metastatic primary tumors, possibly reflecting pathological differences among cancer types. Notably, however, IGKC consistently showed higher expression in metastatic lymph node lesions across all tumor types, indicating a conserved association between IGKC expression and lymphatic metastasis **(Fig. 1H)**.

Previous studies have reported the presence of Igs expression in malignant solid carcinoma cells, with minimal or no expression in normal epithelial tissues. To further assess the similarity between our findings and these prior observations, we calculated an Igs gene-set score based on a panel of Ig genes (*IGKC, IGLC1, IGLC2, IGLC3,*

IGHA1, IGHG1, IGHG3, IGHG4, and *IGHM*). Using receiver operating characteristic (ROC) analysis, we defined a stringent threshold (AUC score > 0.09) to classify tumor cells as Ig-positive **(Fig. S1H)**. Applying this criterion, 95.7% of cells within the IG-epi cluster fell within the Ig-positive range, supporting the robustness and specificity of this classification strategy. Comparative analysis across tissue types revealed virtually no Ig-positive epithelial cells in healthy cervical tissues, whereas a stepwise increase was observed from non-metastatic primary tumors (0.14%) to metastatic primary tumors (3.0%), and further to metastatic lymph node lesions (14.6%) **(Fig. 1I)**. Collectively, these results demonstrate that Igs gene expression is a widespread and tumor-associated feature of malignant epithelial cells, rather than a rare or artifactual event. Importantly, our findings highlight cancer-derived IGKC as a key contributor to metastatic progression, supporting its potential role as a driver gene in tumor metastasis.

2. Immunoglobulin gene fusion transcripts in non-B-lineage malignant cell lines were produced by RNA alternative splicing but not DNA rearrangement

IgG variable regions are generated by rearrangement of V, D, and J gene segment in B cells[19-21]. However, the expression patterns of Igs genes in tumor cells remain largely unclear. To address this question, we analyzed whole-genome sequencing (WGS) and RNA-seq data from the Cancer Cell Line Encyclopedia (CCLE), which, after multiple passages, provide a B cell-free system for reliable genomic and transcriptomic evaluation of Igs gene expression in carcinomas. In total, we included 925 RNA-seq datasets from the database, of which 305 datasets had matched WGS data available for paired analysis. **(Fig. 2A)**. Given that not all hematologic malignancies originate from

B cells, we classified them into B-lineage lymphoid carcinomas and non-B-lineage lymphoid carcinomas. As expected, we observed that Igs genes, including *IGKC, IGLC1,* and *IGHG1*, were highly expressed in the majority of B-lineage lymphoid carcinoma cell lines. Interestingly, a substantial fraction of non-B-lineage lymphoid malignancies and solid tumor cell lines also exhibited Igs gene expression, although at levels lower than those in B-lineage lymphomas. Statistical analysis of *IGKC, IGLC1*, and *IGHG1* revealed that, across most tumor types, IGKC expression was consistently higher than that of IGLC1 and IGHG1. Notably, IGKC with highest mean expression observed in urinary tract tumors, ovarian carcinomas, and central nervous system malignancies **(Fig. 2B, and Fig. S2A)**.

To investigate the compositional patterns of Igκ genes, we analyzed the Burkitt's lymphoma cell line Raji and the endometrial carcinoma cell line RL95-2, and visualized their Igκ loci using the Integrative Genomics Viewer (IGV). The Raji cell line represents a classical B-cell lymphoma and has been well documented to harbor Igs gene rearrangements, serving as an appropriate reference for canonical Igκ recombination. The results showed that the Raji cell line exhibited a truncated genomic pattern at the chr2: igκ region, consistent with canonical Igs gene rearrangement. In contrast, RL95-2 displayed a coherent peak coverage across the same region, with no evidence of Igs gene recombination. Considering rearrangement occurs in the genomic level, we tend to believe the alternative splicing detected in RL952 transcriptome level is irrelevant to genetic rearrangement (**Fig. 2C**). Based on this discovery, we further visualized all paired WGS and RNA-seq in CCLE (n = 305) and confirmed Igs

rearrangement exclusively in B-lineage lymphoid carcinomas, accounting for 86%, while non-B-lineage malignant cell lines exhibited only transcriptional splicing events **(Fig. 2D)**.

To further investigate the arrangement patterns of IGK transcripts，we analyzed it in B-lineage lymphoid carcinomas versus non-B-lineage lymphoid and solid carcinomas. Using the Raji and RL952 cell lines as examples, we found that Raji cells generated full-length IGKV-IGKJ2-IGKC recombined transcripts, whereas RL952 cells produced only a simpler IGKJ2-IGKC fusion transcript. Building on this observation, we performed a systematic analysis across 925 transcriptome datasets to characterize IGK transcript patterns across tumor types. We found that, in B-lineage lymphoid carcinomas, the IGKC fusion transcripts are composed of IGKV-IGKJ1/2/3/4/5-IGKC, with IGKV-IGKJ1-IGKC and IGKV-IGKJ2-IGKC being the most prevalent subtypes, and generated through gene rearrangement, which referred to as the classical modules **(Fig. 2E and 2F)**. However, in non-B-lineage lymphoid and solid carcinomas, the IGKC fusion transcripts are composed of IGKJ1/2/3/4/5-IGKC or only IGKC, with the predominant forms being IGKJ1-IGKC, IGKJ2-IGKC and IGKJ4-IGKC, which referred to as the no-classical modules, and exhibited a monotypic specificity, which contained single AS module within an individual **(Fig. 2C).** Furthermore, the de novo assembly by StringTie to quantitate full-length tarnscripts from CCLE RNA-seq demonstrated that lymphoid carcinomas yielded full-length Igs structure, whereas recombined transcripts in non-B-lineage lymphoid and solid carcinomas were predominantly incomplete, lacking variable (V) region genes **(Fig.**

2F). These findings suggested that IGKC fusion transcripts in tumor cells are generated through RNA alternative splicing, rather than gene rearrangement.

3. IGKC expression in non-B-lineage malignant cell lines and clinical samples from diverse tumor types

To further characterize the expression of IGKC in tumor cells, we designed IGKC-specific primers and performed quantitative RT-PCR to systematically assess IGKC transcript levels across a panel of tumor cell lines. IGKC was abundantly expressed in multiple solid tumor cell lines, with particularly high levels observed in endometrial, ovarian, and lung cancer cells **(Fig. 3A)**. In addition, we collected metastatic and non-metastatic cervical cancer tissues to assess IGKC expression. To exclude the possibility that IGKC signals originated from contaminating B cells, we established primary tumor cell cultures from these tissues and maintained them in tumor cell-specific medium for at least four passages, effectively eliminating B-cell contamination. Potential B-cell contamination was effectively eliminated, as confirmed by the absence of detectable expression of the B-cell specific marker MZB1 by RT-PCR **(Fig. 3B)**. Quantitative RT-PCR analysis of these primary tumor cells revealed that IGKC expression was elevated in cells derived from metastatic lesions compared with those from non-metastatic tumors **(Fig. 3B and 3C)**, supporting a tumor-intrinsic and metastasis-associated pattern of IGKC expression.

Additionally, to determine the spatial expression pattern of IGKC in tumor tissues, we performed RNAscope in situ hybridization on primary cervical and ovarian cancer specimens with or without distant metastases. An IGKC-specific probe was used to

detect IGKC transcripts, while pan-cytokeratin (pan-CK) immunostaining was performed to identify tumor cells, and CD20 immunostaining was used to label B cells. IGKC+ tumor cells were stringently defined as cells co-expressing IGKC and pan-CK, but lacking CD20 expression. Consistent with our prior findings, IGKC exhibited diffuse and high-level expression in B cells, predominantly localized within the stromal compartment. In contrast, IGKC expression in tumor cells was detected at relatively low levels, with IGKC-positive tumor cells preferentially enriched at the tumor invasive front **(Fig. 3D and Fig. S3A)**, implicating IGKC in tumor cell invasion and metastatic progression. Subsequently, we quantified the proportion of IGKC+ tumor cells among total tumor cells in primary cervical and ovarian cancer tissues with or without metastasis. This analysis revealed that the proportion of IGKC+ tumor cells was significantly higher in metastatic tumors than in non-metastatic tumors in both ovarian **(Fig. 3E)** and cervical cancers **(Fig. S3B)**.

Through systematic analysis of the CCLE database, we found that IGKC fusion transcripts are widely present in tumor cells; however, their precise transcript structures have not been experimentally validated. To address this, we designed IGKC-specific primers and employed 5′ rapid amplification of cDNA ends (5′ RACE) to characterize full-length IGKC transcripts. We first performed IGKC 5′ RACE in multiple tumor cell lines, using the B-lineage lymphoid carcinoma cell lines Raji as a positive control to validate the reliability of the assay. As expected, multiple intact, canonical IGKV-IGKJ-IGKC fusion transcripts were readily detected in Raji cells, consistent with immunoglobulin gene rearrangement in B cells **(Fig. 3F and Fig. S3C).** In contrast, in

the ovarian cancer cell line OVCAR3 and the lung cancer cell line HCC827, only two transcript forms were detected: IGKJ-IGKC fusion transcripts and structurally simpler IGKC-only transcripts **(Fig. 3F)**. In other tumor cell lines, including the breast cancer cell line MDA-MB-468 and the endometrial cancer cell lines HEC-1-A and RL952, no fusion transcripts were detected; instead, only simple IGKC transcripts were observed **(Fig. S3D)**. We further amplified IGKC transcripts from primary cervical cancer cells using 5′ RACE, with lymphocytes serving as positive controls. Consistent with the results observed in tumor cell lines, intact IGKV-IGKJ-IGKC fusion transcripts were detected in lymphocytes, whereas primary tumor cells contained only IGKJ-IGKC fusion transcripts **(Fig. 3G and Fig S3E)**. Comparative analysis of IGKC transcript structures among lymphocytes, tumor cell lines, and primary tumor cells revealed that cancer-derived IGKC transcripts predominantly consist of a non-coding region-(IGKJ)-IGKC architecture. Importantly, regardless of whether IGKJ fusion was present, all IGKC transcripts detected in tumor cells lacked a complete open reading frame (ORF), indicating limited protein-coding potential.

Collectively, IGKC transcripts in tumor cells exhibit relatively conserved transcriptional features across different tumor types and are structurally and mechanistically distinct from canonical immunoglobulin transcripts in lymphocytes. Based on these characteristics, we designate these aberrant, tumor-specific IGKC transcripts as cancer-derived IGKC (cIGKC).

4. TFAP2C regulates cIGKC transcription in tumor cells.

Given that IGKC transcriptional patterns in tumor cells are markedly different

from those in B cells, we hypothesized that IGKC expression in tumor cells may be regulated by specific transcription factors (TFs). To investigate the transcriptional regulatory landscape of IG-epithelial cells, we compared TF activity between IG tumor cells and other tumor cells. The analysis revealed that, among the B-cell specific TFs reported in the literature[22], only SPI1 and MEF2C were significantly active in IG-epithelial cells, whereas other B-cell factors, such as PAX5, IKZF1, EBF1, etc. showed no detectable activation **(Fig. 4A)**. Next, we used pySCENIC to identify the TFs with the highest activity scores in IG-epithelial cells. The analysis revealed that BATF3, HES6, MEF2C, and TFAP2C were the most active TFs in IG-epithelial cells **(Fig. 4B)**. To further investigate the transcriptional regulation of IGKC in the majority of non-B-cell tumors, we compared TF activity between CCLE non-B-cell tumor cell lines with high versus low IGKC expression. Differentially active TFs were intersected with the top 15 TFs ranked by activity in IG-epithelial cells from our single-cell analysis, yielding 7 candidate TFs. Promoter analysis within the 2 kb upstream region of IGKC revealed that only TFAP2C and BACH2 were predicted as potential regulators of IGKC transcription **(Fig. 4C)**. However, when comparing the regulon activity Z-scores of TFAP2C and BACH2 between IG-epithelial cells and other tumor epithelial cells, we found that BACH2 activity was not particularly high in IG-epithelial cells (TFAP2C=2.249 vs BACH2=0.684) **(Fig. S4A)**. Furthermore, UMAP mapping of binarized regulon activity scores (RAS) for BACH2 revealed that, in addition to IG-epithelial cells, BACH2 exhibited activity in other epithelial subpopulations as well. These observations indicate that BACH2 lacks specificity in regulating IGKC

transcription in IG-epithelial cells **(Fig. S4B)**. In contrast, TFAP2C exhibited higher activity specifically in IG-epithelial cells **(Fig. 4D)**, and JASPAR analysis predicted the presence of the 5'-GCCnnnnGGC-3' motif in the cIGKC promoter **(Fig. 4D)**, suggesting that TFAP2C may serve as a specific transcriptional regulator of IGKC in tumor cells. In addition, previous studies reported that TFAP2C implicated invarious biological processes and tumor progression[23-26], we evaluated its transcriptional activity across multiple cancer types. Found that elevated TFAP2C activity was consistently associated with poor patient prognosis in cervical, ovarian, lung, and colorectal cancers **(Figure 4J)**, indicated that the regulatory mechanism of TFAP2C on cIGKC warrants further investigation.

To validate the role of TFAP2C in regulating cIGKC transcription, we established TFAP2C-overexpressing (OE-TFAP2C) cell lines in HCC827 and OVCAR3 cell lines and observed elevated TFAP2C expression led to a significant upregulation of cIGKC expression **(Fig. 4E)**. Next, we performed CUT&Tag assay with TFAP2C-specific antibody to enrich TFAP2C-bound DNA and subsequently sequenced the captured fragments in OE-IGKC and control HCC827 cell line. The peaks of TFAP2C detected on the upstream of the cIGKC locus were enriched in OE-IGKC cells campared with control cells **(Fig. 4F)**. Motif analysis using HOMER predicted the binding motif (5'-GCCnnnGGC-3', $p = 4.08\times10^{-7}$) in peaks of cIGKC promoter within OE-cIGKC program **(Fig. 4G).** Based on these predicted motifs, we constructed three mutant reporter plasmids by substituting purines with pyrimidines at the respective binding motifs，and the mutant (Mut1-3) and wild-type (WT) reporter plasmids were then

317 individually transfected into both OE-TFAP2C and control HCC827 and OVCAR3 cell
318 lines, dual-luciferase reporter constructs were generated to validate the specificity of
319 TFAP2C-cIGKC interactions **(Figure 4H and 4I)**. These findings suggest that in tumor
320 cells, the expressionc of cIGKC may be specifically regulated by TFAP2C, thereby
321 promoting tumor metastasis.

322 **5. cIGKC suppresses IFNγ responses in tumor cells**

323 To further investigate the role of Igs in tumor metastasis, we compared significantly
324 enriched gene signatures between IG-epithelial (IG-epi) cells derived from metastatic
325 lymph nodes (LN) and those from primary tumors (primT). Gene set enrichment
326 analysis (GSEA) was performed using the clusterProfiler package, and pathway
327 enrichment was quantified by normalized enrichment scores (NES) (solid bars, $p \leq 0.05$;
328 hollow bars, $p > 0.05$). The results showed that cell cycle related pathways, including
329 *E2F targets, G2M checkpoint, MYC and DNA repair signaling*, were significantly
330 suppressed in IG-epi cells from metastatic lymph nodes **(Fig. 5A)**. This suppression
331 likely reflects a dormancy-like state adopted by tumor cells after dissemination to
332 lymph nodes, where they encounter a complex and restrictive microenvironment[27, 28].
333 Notably, interferon γ (IFNγ) signaling pathways were also markedly suppressed in IG-
334 epi cells from metastatic lymph nodes **(Fig. 5A)**. This observation is particularly
335 intriguing, as previous studies have reported that attenuation of IFNγ signaling can
336 facilitate tumor immune evasion and promote metastatic progression[29, 30]. Motivated
337 by this finding, we systematically compared IFNγ pathway activity between IG-
338 epithelial cells and other epithelial tumor cells across our scRNA dataset as well as

multiple independent external datasets. Consistently, IG-epithelial cells exhibited significantly reduced IFNγ pathway activity relative to other epithelial tumor populations **(Fig. 5B)**, supporting the notion that suppression of IFN signaling is a conserved and defining feature of IG-epithelial cells associated with metastasis.

Based on these observations, we hypothesized that IG-epithelial cells may promote tumor metastasis by suppressing IFN signaling by IGKC. To test this hypothesis, we perturbed IGKC expression in tumor cell lines and examined the resulting transcriptional changes. We designed IGKC-specific siRNAs to selectively suppress IGKC expression. Following siRNA transfection in the OVCAR3 cell line, we performed RNA sequencing to assess global gene expression changes. Transcriptomic analysis revealed that siIGKC led to marked upregulation of key genes involved in the IFN signaling pathway, including *2'-5'-oligoadenylate synthase-like (OASL), interferon induced protein with tetratricopeptide repeats 2 (IFIT2), interferon induced protein with tetratricopeptide repeats 3 (IFIT3)* and *2'-5'-oligoadenylate synthetase 2 (OAS2) ,C-X-C Motif Chemokine Ligand 10 (CXCL10)* **(Fig. 5C)**. Consistently, GSEA demonstrated significant activation of both IFNγ and IFNα response pathways upon IGKC knockdown **(Fig. 5D)**. Notably, apoptosis-related pathways were also significantly enriched after siIGKC **(Fig. 5D)**, suggesting that IGKC may function as a critical protective factor that promotes tumor cell survival. RT-PCR and Western blot analyses were performed to assess key cytokines of the IFNγ response pathway, including ISG15, OAS1, OASL, JAK2, STAT1 and STAT2. The results demonstrated that cIGKC overexpression significantly suppressed IFNγ signaling activity, whereas

IGKC knockdown reactivated it **(Fig. 5E-5G and Fig. S5B and S5C)**. Functionally, cIGKC overexpression significantly promoted, while its knockdown inhibited, tumor cell invasion **(Fig. 5H)**, supporting the role of cIGKC in facilitating immune evasion and metastatic progression. Collectively, these results indicate that cIGKC promotes tumor progression at least in part by suppressing IFNγ signaling and protecting tumor cells from innate immune responses.

6. cIGKC-hnRNPM Interaction Attenuates Intracellular dsRNA and Suppress IFN response

To elucidate how cIGKC regulates downstream interferon (IFN) signaling, we first performed Gene Ontology (GO) biological process analysis based on RNA-seq data from OE-IGKC HCC827 and OVCAR3 cells. The results revealed significant enrichment of pathways related to RNA splicing and mRNA processing **(Fig. S6A)**, suggesting a potential role of cIGKC in post-transcriptional regulation. Given that RNA-binding proteins (RBPs) are central regulators of RNA metabolism-including RNA stability, splicing, transport and translation， we next sought to identify RBPs that interact with cIGKC. To this end, we designed probes specifically binding with cIGKC, using antisense probes (AS) as controls, and performed RNA pull-down assay coupled with quantitative mass spectrometry (MS) in HCC827 and OVCAR3 cell lines. Among the top 20 enriched proteins, 18 were RNA-binding proteins (RBPs), with heterogeneous nuclear ribonucleoprotein M (hnRNPM) showing the highest relative enrichment in both cell lines **(Fig. 6A)**. The cIGKC-hnRNPM interaction was independently validated by RNA pull-down and RNA immunoprecipitation (RIP)

assays in HCC827 and OVCAR3 cell lines **(Fig. 6B-6C, Fig S6B-6C)**. Given the established role of hnRNPM in regulating alternative RNA splicing, we hypothesized that its interaction with cIGKC may alter intracellular splicing programs, thereby modulating IFN signaling.

To further investigate the mechanism of cIGKC-hnRNPM interaction regulates the IFN pathway, we performed eCLIP-seq by hnRNPM specific antibody in control and OE-cIGKC HCC827 cell line, to pull down the RNA fragments bound to it **(Fig. 6D)**. Significant hnRNPM binding peaks were identified in the OE-cIGKC program (52,934 peaks) and the control program (26,333 peaks) ($P < 0.05$). The majority of peaks were localized within intronic regions of target genes, particularly deep intron, consistent with previous reports[31]. Further analysis of hnRNPM binding peaks within the cIGKC region revealed significant enrichment on cIGKC exonic regions in the OE-cIGKC program **(Fig. 6E)**. HOMER motif analysis predicted GU-rich binding motifs at the interaction sites **(Fig. 6F)**. In addition, we employed AlphaFold3 to predict interactions between multiple hnRNPM proteins and cIGKC mRNA fragments. The prediction indicated a single predominant binding site on the cIGKC mRNA (5'-nUGUGUGn-3') under conditions simulating multiple protein presence **(Fig. 6G)**. Adopting the approach described by Zheng's study, we normalized hnRNPM binding peaks by genomic length[31]. This analysis revealed that hnRNPM binding remained significantly enriched within 5'UTRs and deep introns. Among different repetitive element types, hnRNPM exhibited specific enrichment primarily on LINEs **(Fig. 6H-6I)**. Comparative analysis of binding site regions between the OE-IGKC and control programs revealed

a significantly increased proportion of hnRNPM binding occurring within LINE regions, particularly L1 elements (the most abundant LINE in humans), in the OE-IGKC program **(Fig. 6J-6K)**.

Previous study demonstrated that LINE elements are a major intracellular source of double-stranded RNA (dsRNA)[31, 32]. We hypothesized that elevated IGKC expression may reduce dsRNA production by altering hnRNPM binding to LINE elements, thereby limiting LINE exposure and subsequent dsRNA formation. We examined the intracellular levels of dsRNA by immunofluorescence using anti-dsRNA antibody (J2). Consistent with our expectation, cIGKC overexpression markedly reduced dsRNA accumulation, whereas cIGKC knockdown led to a significantly increase dsRNA levels **(Fig. 6M and Fig. S6H)**. Furthermore, hnRNPM silencing in cIGKC-overexpressing cells restored dsRNA formation **(Fig. 6N)** and reactivated the IFNγ response **(Fig. 6Q and Fig. S6I)**. Collectively, these results indicate that elevated IGKC expression in tumor cells promotes hnRNPM binding to LINE regions, resulting in decreased dsRNA generation and suppression of the IFNγ response.

7. cIGKC suppresses IFNγ responses and promotes tumor metastasis *in vivo*

To further investigate IGKC expression in tumors across species, we designed mouse-specific RT-PCR and 5′ RACE primers to assess IGKC expression and fusion transcript status in murine tumor cell lines. We collected a panel of mouse tumor cell lines derived from lung cancer, ovarian cancer, melanoma, colon cancer, breast cancer, and liver cancer. RT-PCR analysis revealed faint amplification bands in CT26, 4T1, and Hepa-16 cells, although all were markedly lower than the expression levels observed in

mouse lymphocytes **(Fig. 7A)**. Next, we performed 5′ RACE to amplify full-length IGKC transcripts in mouse tumor cell lines. Notably, in the B16 melanoma cell line, we detected IGKC transcripts that lacked IGKV–IGKJ fusion, resembling the simple IGKC transcripts observed in human tumor–derived samples. These transcripts consisted of a non-coding region-IGKC structure and lacked a complete ORF **(Fig. 7B)**. These results indicate that, although tumor-derived IGKC expression is low in mice, its transcriptional pattern is conserved with human tumor cells, suggesting that cancer-derived IGKC may have similar functional roles across species. To test this, we constructed a mouse IGKC-overexpressing lentiviral vector and established stable overexpression lines in the 4T1 breast cancer and TC-1 cervical cancer cell lines. RT-PCR analysis revealed that mIGKC overexpression significantly suppressed the expression of genes in the IFN response pathway in both cell lines, consistent with the immunosuppressive effects observed in human tumor cells **(Fig. 7C and 7D)** .

To further investigate the effect of IGKC expression on tumor metastasis *in vivo,* we established a mouse footpad lymph node metastasis model by injecting control and mcIGKC-overexpressing (OE-IGKC) TC-1 cell lines into the footpads, BALB/c mouse model constructed with 4T1 cell lines. Tumor cells were injected into the footpads of mice, and tumor growth and lymph node metastasis were assessed four weeks later. Compared with controls group, mcIGKC-overexpressing tumors of both models significantly promoted primary tumor growth **(Fig. 7E and 7F)**. We then dissected the popliteal, inguinal, and axillary lymph nodes for analysis. Compared with controls, lymph nodes from the IGKC-overexpressing group were noticeably enlarged

and firm, suggesting increased metastatic burden **(Fig. 7E)**. Histological examination using H&E staining confirmed a higher incidence of popliteal lymph node metastasis in the IGKC-overexpressing group **(Fig. 7G and 7H)**. Notably, in the 4T1 mouse model, the overexpression group also displayed a significantly increased rate of lung metastasis **(Fig. 7I)**. These results indicate that IGKC overexpression enhances both primary tumor growth and lymphatic metastasis, supporting a pro-metastatic role for tumor-derived IGKC in vivo.

Accumulating evidence indicates that suppression of the IFNγ response can create an immunologically permissive environment that facilitates tumor progression and metastasis[29, 30, 33]. To investigate this *in vivo,* we assessed IFNγ pathway activity in primary tumors by RT-PCR. Consistent with our hypothesis, expression of IFNγ responsive genes was significantly reduced in mcIGKC-overexpressing tumors compared with controls **(Fig. 7J)**. Together with our prior in vitro and single-cell analyses, these results support the model that aberrant upregulation of cIGKC in tumor cells suppresses IFNγ signaling, fostering an immunosuppressive microenvironment that facilitates immune evasion and promotes tumor metastasis.

Discussion

Tumor metastasis remains a major obstacle to improving clinical outcomes in oncology and is the primary cause of cancer-related mortality despite substantial therapeutic advances[34, 35]. Although previous studies have suggested that tumor cells may express immunoglobulins to enhance malignancy and contribute to poor prognosis[6, 36], the expression dynamics of immunoglobulin genes within tumor cells

remain poorly understood. Here, through single-cell RNA sequencing, we identify a distinct epithelial subset defined by immunoglobulin gene expression and demonstrate that tumor cell-derived immunoglobulin kappa constant region (IGKC) functions as a critical regulator of metastatic progression. Using RNA Scope and 5′RACE, we further delineated the expression patterns of IGKC in tumor cells and uncovered the underlying mechanisms by which IGKC drives tumor metastasis.

The extensive infiltration of B cells within tumor tissues makes it challenging to directly extract cIgG information through bulk genomic or transcriptomic sequencing. Current approaches largely depend on microdissection to eliminate B cells, which inherently limits both sampling efficiency and sequencing depth[37, 38]. The advent of scRNA-seq has provided a powerful tool to dissect gene expression and functional states across diverse cell populations in the tumor microenvironment[39]. In this study, we identified a distinct immunoglobulin-expressing epithelial, IG cluster across, multiple tumor types, including cervical cancer, ovarian cancer, lung cancer, and head and neck squamous cell carcinoma. Notably, IGKC expression was significantly upregulated in the metastatic primary tumors compared with non-metastatic counterparts. This observation was further validated using metastatic and non-metastatic tumor tissue sections, primary tumor cells, and mouse models. Collectively, our findings suggest that IGKC may serve as a potential biomarker and predictive target for tumor metastasis, although larger cohorts and additional tumor types are required for further validation.

It is well established that in lymphocytes, immunoglobulin variable region gene

fragments undergo recombination to generate complete variable region sequences. This process, known as immunoglobulin gene rearrangement, occurs during lymphocyte development and involves random recombination of V, D, and J gene segments[8, 9]. The heavy-chain variable region is formed by one V, one D, and one J fragment, while the light-chain variable region consists of one V and one J fragment. Multiple copies of these gene segments ensure the diversity of antibodies. To investigate whether a similar mechanism exists in tumor cells, we analyzed WGS and RNA-seq data from the CCLE pan-cancer cell line database. Interestingly, we found that immunoglobulin fusion transcripts are widely present in tumor cells; however, unlike lymphocytes, they are generated through alternative RNA splicing rather than gene rearrangement. Moreover, IGKC transcripts in both primary tumor cells and tumor cell lines exhibited a degree of conservation but encoded proteins with limited functionality. These findings suggest that the immunoglobulin structures in tumor cells differ fundamentally from those in lymphocytes, indicating a tumor-specific immunoglobulin configuration with potential for therapeutic intervention. Nevertheless, this observation challenges the conventional understanding of cancer cell-encoded cIgG, and further studies are required to elucidate the precise structure of immunoglobulin light chains in tumor cells.

We compared transcription factors between B cells and the IG-expressing tumor cell population and found that IGKC expression in tumor cells is not regulated by the conventional immunoglobulin-associated transcription factors observed in B cells. To further elucidate the transcriptional regulation of IGKC in tumor cells, we integrated CUT&Tag and RNA-seq analyses, which revealed that TFAP2C specifically regulates

cIGKC transcription. Moreover, analysis of lung cancer and other tumor datasets demonstrated that TFAP2C transcriptional activity is aberrantly elevated and correlates with poor patient prognosis[24, 40, 41]. Collectively, these findings suggest that targeting TFAP2C transcriptional activity may represent a promising strategy to suppress IGKC-driven tumor progression.

RBPs are key post-transcriptional regulators that modulate gene expression through mechanisms such as alternative splicing, polyadenylation, mRNA stability, and translation, thereby contributing to tumorigenesis, progression, and metastasis[42, 43]. To further elucidate the mechanism underlying IGKC-mediated tumor metastasis, we combined RNA-seq, RNA pull-down MS, and RIP assays and identified a specific interaction between IGKC and hnRNPM, with binding affinity increasing in response to elevated IGKC expression. hnRNPM, an RBP containing three RNA recognition motifs with splicing activity, regulates mRNA stability and translation largely through alternative splicing. Our results showed that enhanced IGKC-hnRNPM binding promoted hnRNPM association with repetitive genomic elements, particularly LINE regions, which represent a critical source of dsRNA. This interaction reduced intracellular dsRNA accumulation and consequently suppressed IFN signaling. Importantly, knockdown of hnRNPM restored IFN responses, further supporting its role in this pathway. Based on these findings, we propose that elevated IGKC expression in tumor cells enhances hnRNPM binding, leading to diminished dsRNA production and attenuated IFN signaling, thereby facilitating tumor metastasis. Nevertheless, further studies are required to validate these mechanistic insights and fully delineate how IGKC

drives metastatic dissemination.

In conclusion, our study provides compelling evidence that IGKC is a critical driver of tumor metastasis. By leveraging single-cell transcriptome sequencing, we mapped the gene expression landscape of Ig-expressing tumor cell subsets and demonstrated that IGKC expression is significantly upregulated in metastatic tumor cells. Mechanistically, IGKC is transcriptionally regulated by TFAP2C, and functionally, it promotes metastasis by binding to hnRNPM, suppressing dsRNA production, and dampening IFN responses. Together, these findings expand our understanding of immunoglobulin expression in tumor cells and highlight IGKC as a potential biomarker and therapeutic target for metastasis prediction and intervention.

Experimental model and subject details

Sample collection

This project was conducted with approval from the Institutional Review Board of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology (TJ-IRB202509038). The study adhered rigorously to the principles outlined in the Declaration of Helsinki. All human samples, encompassing sequencing samples, IHC staining specimens and organoid culturing samples, were collected only with written informed consent from all participants. After collection, fresh tissue samples from patients with CESC or healthy donors were surgically excised and meticulously preserved in MACS tissue storage solution (Miltenyi Biotec, Cat#130-100-008) until subsequent processing.

scRNA-seq library preparation and sequencing

Tumors were dissociated to single-cell suspensions using digestion buffer containing collagenase IV and deoxyribonuclease I. Red blood and dead cells were removed using EasySep Dead Cell Removal. The remaining suspensions were processed using the Chromium Next GEM Single Cell 3' Reagent Kits v.3.1 (10x Genomics) according to the manufacturer's instructions. The resulting cDNA libraries were sequenced on a DNBSEQ-T7 (RRID: SCR_017981) platform.

Cell lines and cell culture

All cell lines were authenticated by short tandem repeat profile analysis and verified as free for Mycoplasma contamination. All medium was supplemented with penicillin and streptomycin, and all cell lines were maintained in a 5% CO2 incubator at 37 °C. Detail information about cells are provided in Table S2. Whole Genome sequencing data and bulk RNA sequencing data was gathered public databases: CCLE (https://www.cbioportal.org/study/summary?id=ccle_broad_2019, RRID: SCR_013836).

Primary tumor cell culture

Fresh tumor specimens were collected immediately after surgical resection and kept in ice-cold PBS. Tissue dissociation was performed using the Tumor Dissociation Kit, human (Miltenyi Biotec, Cat#130-095-929) according to the manufacturer's protocol. Briefly, tumor tissues were minced into ~1–2 mm^3 pieces and enzymatically digested in the enzyme mix provided by the kit using the gentleMACS™ Octo Dissociator with Heaters (Miltenyi Biotec, Cat#130-096-427). Single-cell suspensions

were filtered through 70μm strainers, washed, and cultured in DMEM/F12 supplemented with 10% FBS, 1% penicillin–streptomycin, and maintained in a 5% CO2 incubator at 37°C. To enrich for epithelial tumor cells and minimize fibroblast contamination, cultures were subjected to brief trypsinization (1–2min), which preferentially detached fibroblasts. After 3-5 passages in culture, RT-PCR was performed to assess MZB1 expression, ensuring the absence of B-cell or plasma cell contamination in the primary tumor cell cultures.

RNA Scope assay

Formalin-fixed paraffin-embedded (FFPE) tissue sections (4μm) deparaffinized in xylene, rehydrated through graded ethanol series, and subjected to endogenous peroxidase blocking. RNA in situ hybridization was performed using the RNAscope Multiplex Fluorescent Reagent Kit (Advanced Cell Diagnostics, ACD, cat#323180, cat#322381, cat#323100, cat#322809, cat#310091) according to the manufacturer's instructions. Human-specific IGKC probes were used to detect cIGKC mRNA, while anti-pan-CK antibody (CST, cat#4545, RRID: AB_490860) was used to mark epithelial cells and anti-CD20 antibody (Abcam, cat#ab64088, RRID: AB_1139386) to identify B cells. Nuclei were counterstained with DAPI, and slides were mounted with an anti-fade medium. Fluorescent signals for RNA and protein were imaged using a fluorescence microscope Nikon ECLIPSE Ti inverted microscope system (RRID: SCR_021242). Image Pro Plus (RRID: SCR_007369) was used for image analysis. Paired t test was performed in GraphPad Prism (RRID: SCR_002798) software for statistical analysis.

5′ Rapid Amplification of cDNA Ends (5′ RACE)

Total RNA was extracted from cells and tissues using the RNeasy Mini Kit (QIAGEN, cat#74104) following the manufacturer's instructions. 5′ RACE was performed using the HiScript-TS 5'/3' RACE Kit (Vazyme, cat#RA101-01) according to the manufacturer's protocol. Briefly, 1μg of total RNA was reverse transcribed with a 5′ RACE CDS primer and 5' TS Oligo to generate first-strand cDNA. The 5′ ends of cIGKC transcripts were amplified using a gene-specific reverse primer and a universal primer mix provided by the kit. Nested PCR was subsequently performed with a nested universal primer and a nested gene-specific primer to increase specificity. PCR products were resolved on a 1.5% agarose gel, purified with a FastPure Gel DNA Extraction Mini Kit (Vazyme, cat#DC301-01), and cloned into the pMD^{TM}-T vector cloning kit (Takara, cat#6013) for Sanger sequencing. The obtained sequences were analyzed to determine the transcriptional start sites and fusion patterns of IGKC. All primers sequences are listed in Table S3.

Stable over expression cell generation

Concentrated lentiviral particles encoding human IGKC or corresponding empty-vector control were used to infect target cells. OVCAR3 and HCC827 Cells were seeded to achieve 40–60% confluence on the day of infection and exposed to viral supernatant, and supplemented with 10 ug/mL ploybrene (CAT#40804ES76, YEASEN). After 16–24 h, viral supernatant was removed and replaced with fresh complete medium. At 48–72 h post-infection, cells were subjected to selection with puromycin 2 μg/mL until non-transduced control cells were completely eliminated.

Stable polyclonal cultures were expanded and cIGKC overexpression was validated by RT-PCR. Empty-vector transduced cells served as negative controls.

Construction of knockout cell lines

hnRNPM knockout cell lines were generated using the CRISPR-Cas9 system. Single-guide RNAs (sgRNAs) were designed with the CRISPick platform (https://portals.broadinstitute.org/gppx/crispick/public, RRID: SCR_025148), and the top three candidate sequences were synthesized and cloned into expression vectors. The optimal sequence was selected by knockout efficiency verification for subsequent experiments. The sgRNAs targeting hnRNPM were synthesized (Tsingke, Beijing, China) and cloned into pLentiCRISPR v2. Lentivirus was produced in HEK293T cells co-transfected with pCMV-R8.91 (RRID: Addgene_202687), pCMV-VSVG (RRID: Addgene_240630), and pLentiCRISPR v2 constructs (RRID: Addgene_140206). OVCAR3 and HCC827 cells were infected with viral supernatant in the presence of 10 μg/mL polybrene (YEASEN, Cat#40804ES76) and selected with 2 μg/mL puromycin. hnRNPM knockout efficiency was confirmed by Western blot. All sg-RNA sequences are listed in Table S3.

RNA Interference and Transfection

Targeted cIGKC-specific siRNA (Ruibo Biotech) was transfected at a final concentration of 100 nM using Lipofectamine3000 (Thermo Fisher Scientific, cat#L3000015) in serum-free medium (Opti-MEM). Transfection complexes were incubated at room temperature for 15-20 min prior to application to cells. Knockdown efficiency was assessed 24-48 h post-transfection by RT-PCR, with β-actin serving as

internal controls. All siRNA sequences are listed in Table S3.

Transwell assay

Cell migration ability was examined using 24-well transwell chambers (pore size, 8μm; Costar, Cambridge, USA). 3×10^4 OVCAR3 or HCC827 cells, respectively, were seeded on the upper chamber while complete medium with were placed in the lower chamber. After 24h, the inserts were removed and migrated cells were stained with 0.05% crystal violet.

Bulk RNA isolation and sequencing

Snap frozen fresh harvested cell were processed with QIAGEN's AllPrep RNA Mini Kit for RNA extraction. RNA concentration was measured using a Qubit fluorometer, and quality was assessed with an Agilent 2100 Bioanalyzer. Samples passing quality control with a minimum RNA integrity number (RIN), only samples with an RNA integrity number (RIN) ≥7 were used for library preparation. 2 μg of total RNA per sample was employed to construct sequencing libraries using the VAHTS mRNA-seq v2 Library Prep Kit (Illumina). Libraries were quantified, pooled, and sequenced on an Illumina platform to generate paired-end reads of 150bp.

Western Blot Analysis

Western blot was performed as previously described[44]. Harvested cells were lysed on the ice for 15 min with RIPA buffer (Servicebio, cst#G2002) supplemented with 1% PMSF and 2% cocktail. Lysates were lysed on ice after sonication and supernatant was collected for further studies. Protein concentration was measured using Coomassie (Beyotime, cat#ST1119, cat#P0006C). Equivalent protein lysates proteins were

separated in SDS-PAGE gels (BioSci, cat#8012011), and electro-transferred onto polyvinylidene fluoride membranes (PVDF, Cytiva Life Sciences, cat#10600023). The PVDF membrances were blocked using 5% BSA and incubated with the following primary antibody at 4℃ overnight: anti-pJAK2 (Y1007, Y1008, Abcam, cat#ab32101, RRID: AB_775808), anti-pSTAT1 (Tyr701, CST, cat#9167, RRID: AB_561284), anit-IRF1 (Proteintech, cat #11335-1-AP, RRID: AB_2877759), anti-IRF3 (Abcam, cat#ab25950, RRID: AB_2126991), anti-OASL (CST, cat#36845, RRID: AB_3712914), anti-ISG15 (Abcam, cat#ab38639, RRID: AB_775793), anti-GAPDH (Abclonal, cat#A19056, RRID: AB_2862549), anti-β-actin (Abclonal, cat#AC026, RRID: AB_2768234), anti-Vinculin (Proteintech, cat#26520-1-AP, RRID: AB_2868558), anti-HNRNPM (Proteintech, cat #26897-1-AP, RRID: AB_2880673), anti-TFAP2C (Abcam, cat#AB218107, RRID: AB_2891087), followed by incubation with 1:5000 secondary antibodies (Abclonal, cat#AS014, RRID: AB_2769854, cat#AS003, RRID: AB_2769851) for 1 h at room temperature. Bands were visualized by WesternBright ECL using a Western blotting detection kit (Advansta, cat#K-12045-D50) in the ChemiDoc Imaging System (Bio-Rad).

RNA Extraction and RT-PCR

Total RNA was isolated from cultured cells using a total RNA extraction kit (TIANGEN, cat#DP419). cDNA was synthesized using RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, cat#K1622) on an RT-PCR System (Bio-Rad). RT-qPCR was performed on a CFX Connect quantitative real-time PCR System (Bio-Rad) using SYBR Green Master Mix (Vazyme; cat#R223-01). Relative mRNA

691 expression was determined by the ΔΔCt method and normalized by β-actin. All primers
692 sequences are listed in Table S3.

693 **CUT&Tag library preparation and sequencing**

694 CUT&Tag assays on OVCAR3 and HCC827 cells were performed using the
695 Hyperactive Universal CUT&Tag Kit (Vazyme, cat#TD904). Cells were bound to
696 concanavalin A–coated beads, permeabilized with digitonin, and incubated with
697 TFAP2C antibody (Abcam, cat#ab218107, RRID: AB_2891087), followed by pA–Tn5
698 transposase treatment. After tagmentation, DNA was extracted, amplified, and purified
699 to generate libraries, which were sequenced on an Illumina NovaSeq 6000 with 150 bp
700 paired-end reads (RRID: SCR_016387).

701 **Dual-Luciferase Reporter Assay System**

702 The dual-luciferase reporter assay was performed to evaluate the specific
703 interaction between TFAP2C and cIGKC promoter. Reporter constructs were generated
704 by cloning predicted TFAP2C-binding sites of cIGKC promoter, along with their
705 corresponding three site-directed mutants, into the pGL3-basic vector (TsingKe Co.,
706 Ltd, RRID: Addgene_48743). Cells were co-transfected with these reporter constructs
707 and a TFAP2C overexpression plasmid using Lipofectamine 3000 (Invitrogen). A
708 Renilla luciferase vector (pRL-TK, Promega, RRID: Addgene_11313) was co-
709 transfected as an internal control. Luciferase activity was measured 48h post-
710 transfection using the Dual-Luciferase Reporter Assay System (Beyotime, cat#RG029)
711 following the manufacturer's protocol and measured on a BioTek Synergy HT Plate
712 Reader (Biotek, RRID: SCR_020536). Relative firefly luciferase activity was

normalized to Renilla luciferase activity.

RNA pull down and Mass spectrometry

RNA pull-down experiments were performed using the RNA Pull-Down Kit (The Thermo Fisher, cat#20164) according to the manufacturer's instructions. Biotin-labeled CHIRP probes targeting IGKC were custom-synthesized and using the antisense (AS) prob as control probe (Ruibo Biotech). In brief, cell lysates were prepared in lysis buffer containing protease and RNase inhibitors and incubated with RNA-bound streptavidin-coated magnetic beads at 4 °C for 1-2 h. After extensive washing, bound proteins were either eluted for western blotting to validate specific interactions or subjected to mass spectrometry (MS) for unbiased identification of interacting proteins.

RNA immunoprecipitation (RIP)

RIP experiments were performed using the RIP Kit (Sigma, cat#17-700) according to the manufacturer's instructions. Briefly, cells were harvested and lysed in RIP lysis buffer containing RNase and protease inhibitors. Lysates were incubated with magnetic beads conjugated to anti-HNRNPM antibody (Abcam, cat#128143, RRID: AB_11141263) or control IgG (Abcam, cat#ab172730, RRID: AB_268793) at 4 °C overnight with gentle rotation. After washing to remove non-specifically bound material, RNA-protein complexes were eluted, and co-precipitated RNAs were purified and analyzed by RT-qPCR to assess the enrichment of target RNA.

eCLIP and sequencing

Samples were irradiated twice with 254 nm UV-C light on ice at 400 mJ and lysed on ice for 15 min in lysis buffer (20 mM Tris, pH 8.0; 1 mM EDTA; 400 mM NaCl;

0.1% SDS; 0.5% Triton X-100; 1 x protease inhibitor; 200 U/mL RNasin Plus; 20 U/mL DNase I). Lysates were treated with 40 U RNase I (The Thermo Fisher, cat#EN0601) and incubated at 37 °C for 5 min with rotation. The reaction was quenched by adding RNasin Plus, and samples were centrifuged at 15,000 g for 15 min at 4 °C. 2% of the supernatant was reserved as input, and the remainder was incubated with anti-HNRNPM antibody (Abcam, cat#128143, RRID: AB_11141263) at 4 °C for 4 h with slow rotation. Protein A/G magnetic beads (The Thermo Fisher, cat#88802) were added and incubated for 2 h, followed by successive washes. 3′ phosphate groups were removed using FastAP (The Thermo Fisher, cat#EF0654), and a 3′ App-DNA linker was ligated to RNA using T4 RNA ligase (NEB, cat#M0204). RNA-protein complexes were separated on NuPAGE 4-12% Bis-Tris gels (The Thermo Fisher, cat#WG1401BX10) and transferred to a nitrocellulose membrane. RNA bound to the protein of interest was excised, digested with proteinase K (NEB, cat#P8107) at 37°C for 20 min, and extracted with TRIzol (The Thermo Fisher, cat#15596018CN). Purified RNA was concentrated, desalted, and reverse transcribed. The resulting cDNA was ligated with a 3′ UMI adapter, amplified by PCR, and sequenced on an Illumina NovaSeq X Plus using PE150 (RRID:SCR_024568).

Detection of dsRNA by immunofluorescence

Cells were seeded on 20 mm glass coverslips in 12-well plates at ~70% confluence. Coverslips were washed with PBS, fixed in 4% paraformaldehyde, and permeabilized with 0.25% Triton X-100. When indicated, cells were incubated with 4 U RNase III (Biosystems, cat#A2290) in reaction buffer at 37 °C for 4 h after permeabilization.

Following blocking with 5% goat serum in PBS for 30min, samples were incubated overnight at 4 °C with SCICONS anti-dsRNA antibody (J2) (Jena Bioscience, cat#RNT-SCI-10010; 1:100, RRID: AB_2922431), followed by Alexa Fluor 488-conjugated Goat Anti-Mouse IgG (ABclonal, cat#AS037, RRID: AB_2768319) or for 1 h at room temperature. Nuclei were counterstained with DAPI (Servicebio, cat#G1012). Images were acquired on a Nikon ECLIPSE Ti inverted microscope system (RRID: SCR_021242) and analyzed using Image-Pro Plus 6.0 (RRID: SCR_007369).

IHC staining

Tissues were fixed overnight in 4% paraformaldehyde and embedded in paraffin. After deparaffinization, rehydration, antigen unmasking, and endogenous peroxidase blocking, sections were blocked in 5% BSA with 0.1% Triton X-100 in PBS, which were subsequently incubated overnight at 4°C with primary antibodies, anti-PanCK antibody (CST, cat#4545). Tissue sections were then stained using a Mouse/Rabbit Enhanced polymer detection system (ZSGB-BIO, PV-9000) and the target proteins were visualized by DAB chromogenic reaction (Servicebio, cat#G1212). Tissues were observed under a microscope and photographed after being counterstained with hematoxylin, dried, and mounted.

Lymph node metastasis model in mice

All animal experiments were conducted with the approval of the Ethics Committee of Tongji Medical College（Approval No.3665). Female C57BL/6 mice (RRID:RRID: MGI:2159769) (6-8 weeks old), each group (n=5), were anesthetized and injected subcutaneously in the right hind paw with 1 × 10^5 control TC-1 cells or OE-IGKC

TC-1 cells in 25 μL PBS. Female BALB/c mice (RRID: MGI:2161072) (6-8 weeks old) were anesthetized and injected subcutaneously in the right hind paw with 1× 10^4 control 4T1 cells or OE-IGKC 4T1 cells in 25 μL PBS. Tumor growth at the injection site was monitored every 2-3 days using calipers. At predetermined time points, mice were euthanized, and primary paw tumors, popliteal and inguinal lymph nodes, and lungs were harvested. Tissues were fixed in 4% paraformaldehyde for histological analysis or snap-frozen for molecular assays. Lymph node and lung metastatic lesions were evaluated by H&E staining and PanCK immunohistochemistry to confirm epithelial tumor cell presence.

Method details

scRNA-seq data processing

Quantification and quality control (QC)

The sequencing data were processed using Cell Ranger (v6.1.1, RRID: SCR_017344) against the GRCh38 human reference genome (RRID: SCR_027738). Further quality control and doublet removal were applied to the obtained cells based on unique molecular identifier counts, the number of genes detected and the proportion of mitochondrial gene counts per cell using Scanpy (v1.9.3, RRID: SCR_018139). Cell with more than 200 genes, unique molecular identifier n_genes_by_counts under 10000 and total_counts under 80000, and a proportion of mitochondrial gene counts less than 20% were retained for further analysis. After that, we applied library-size correction and logarithmized normalization using Scanpy, and the normalized count matrix was used for downstream analysis.

802 **Dimension reduction and unsupervised clustering**

803 Dimension reduction and unsupervised cluster were applied to the single-cell data
804 following the workflow in Scanpy (RRID: SCR_018139). The top 3000 highly variable
805 genes were selected and the effect of total count per cell and the percentage of
806 mitochondrial genes were regressed from the dataset. We performed principal
807 component analysis with 30 components and batch effects removal from patients using
808 the harmony (v0.0.5, RRID: SCR_022798) algorithm. Finally, we applied the UMAP
809 algorithm (RRID: SCR_018217) for dimension redution and visualization of the dataset
810 and an unsupervised graph-based clustering algorithm called Leiden (resolution=1) to
811 cluster single cells.

812 **Cell-type annotation**

813 Cell types were annotated by evaluating the expression of canonical markers and
814 cluster-specific markers identified using the scanpy.tl.rank_gene_groups function. Any
815 clusters expressing more than two canonical cell-type markers were excluded for
816 further analysis.

817 **Single-cell CNV estimation**

818 Using infercnvpy (v0.2.0)[45] (RRID: SCR_025804) , we were able to identify
819 malignant epithelial cells in cervical cancer patients. Vary large-scale chromosomal
820 CNVs are identified by infercnvpy as cells with unequal mean expression of genes
821 across chromosomes compared to control cells. Epithelial cells constituted the query
822 set, with additional non-malignant endothelial cells derived from cervical cancer
823 serving as reference controls.

Differential expression and gene set enrichment analysis

DEGs from scRNA-seq were identified using Wilcoxon rank-sum tests *P*-values from Wilcoxon rank-sum tests were adjusted using the BenjaminiHochberg procedure. For gene set enrichment analysis, the gene list was ranked by log2(fold change) in ascending order and used as input for gene set enrichment analysis using the clusterProfiler (v4.10.1, RRID: SCR_016884)[46] package with default options.

Calculation of functional enrichment scores and setting therasholds

The scores of significantly enriched gene signatures were calculated for each group of interest by using the AUCell_calcAUC in AUCell (RRID: SCR_021327). The likely therasholds were setting for tumor cells by using AUCell_exploreThresholds with aucMaxRank defined as the top 1% of genes (ceiling(0.01 * nrow(rankings)).

Public scRNA datasets

The following publicly available scRNA datasets were downloaded:

- HNSCC (Wass et al.[15] , PMID: 39738080, GSE181919)
- BRCA (Mao et al.[47], PMID: 36040519, HRA002051)
- LUAD (Kim et al.[16], PMID: 32385277, GSE131907)
- HCC (Lu et al[17], PMID: 35933472, GSE149614)
- GC (Jiang et al.[18], PMID: 35184420, GSE163558)

For each dataset, cells were filtered for >200 detected genes, and a mitochondrial read fraction <20 %.

SCENIC analysis

Activated regulons in different tumor cells were analyzed using pySCENIC[48]

(v0.11.2, RRID: SCR_025802), using the expression matrix as an input. The co-expression network was computed by GRNBoost2 (RRID: SCR_025614), and the regulons were identified through the RcisTarget (RRID: SCR_024860) method. Following this, the AUCell (RRID: SCR_021327) algorithm was employed to score the activity of each regulon in individual cells.

Kaplan-Meier (KM) survival curves

The 33 kinds of TCGA tumor count data (RRID: SCR_003193) and clinical parameters were obtained from UCSC Xena (http://xena.ucsc.edu/, RRID: SCR_018938). Subsequently, TPM values were computed and utilized for characterizing tumor subsets and assessing patient survival. TF activities for 33 tumor types were computed using DoRothEA[49] (v1.14.1, RRID: SCR_027126). Survival analyses were performed using the R survival package (v3.40, RRID: SCR_021137), with dichotomization at the median value or the best cut-off point determined by the "surv_cutpoint" function in the survminer package (v0.4.9, SCR_021094). Kaplan-Meier method was used for survival curves, and *P*-values were calculated using a log-rank test unless otherwise specified.

Bulk RNA-seq analysis

Data process

To quantify gene expression, raw data were trimmed by Trim Galore (v0.6.7, RRID: SCR_011847) with default settings and then aligned to the human genome (hg38) with STAR (v2.7.10a, SCR_004463). The read count was generated by subread (v2.0.1, RRID: SCR_009803)[50] with default parameters and the transcripts per million (TPMs)

of genes were quantified using RSEM (v.1.3.1, RRID: SCR_000262).

Differentially expressed analysis

DEGs were identified with DESeq2 (RRID: SCR_015687) for pairwise comparisons between samples of different group. An adjusted P-value threshold of 0.05 was used.

Enrichment Analysis

Rank-based gene set enrichment (GSEA, RRID: SCR_003199) analyses were conducted using the R package clusterProfiler (v4.10.1) [46] with default options. Gene sets of molecular pathways from the KEGG (RRID: SCR_012773) /Hallmark/Reactome (RRID: SCR_003485) databases were applied to compute pathways.

WGS analysis

Data process

The paired-end WGS reads to human reference genome (hg38) with BWA-mem[51] v0.7.15 (RRID: SCR_010910). The BAM files were reordering and sorting for visualization using by Intergrative Genomics Viewer (IGV, RRID: SCR_011793).

CUT&Tag analysis

Data process

The paired-end reads were trimmed by Trim Galore (v0.6.7) and aligned to hg38 using Bowtie2[52] (v2.4.5, RRID: SCR_016368) using parameters --local --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700. Mapped read pairs were filtered and retained using samtools (RRID: SCR_002105), and the resulting BAM file

was converted to BED format using bedtools[53] (v2.31.1, RRID: SCR_006646). Then read pairs on the same chromosome with a fragment length less than 1000bp were retained. Read coverage over hg38 genome was estimated by bamCoverage from deepTools[54] (v3.1.3, RRID: SCR_016366) with parameters --binSize 50 --normalizeUsing RPKM and visualized by UCSC browser. To Assess the reproducibility between replicates and across conditions, the genome is split into 200-bp bins, and a Pearson correlation of the log2-transformed values of read counts in each bin is calculated between replicate datasets.

Peak calling

Peaks were called using MACS2[55] (v2.2.9.1, RRID: SCR_013291). Peaks in OE-IGKC group were identified with peaks derived from the control group serving as controls. ChIPseeker[56] (v1.32.1, RRID: SCR_021322) was employed to retrieve the nearest genes around the peak and annotate the genomic distribution. Peak calls within replicate datasets were combined using bedtools.

Motif analysis

Motif analyses were carried out with HOMER[57] (v4.11, RRID: SCR_010881). De novo motifs of TFAP2C CUT&Tag peaks were identified by findMotifsGenome.pl. Motifs were screened within the binding region by Find Individual Motif Occurrences (FIMO) tool in the MEME suite[58] (v5.5.5, RRID: SCR_001783) and MotifScan[11] with the default threshold $P < 1e\text{-}4$, based on position weight matrices (PWMs) from JASPAR (ID:MA0524.1, RRID: SCR_003030). The transcription factor binding site (TFBS) was further predicted in JASPAR with the relative profile score threshold of

80%.

eCLIP-seq analysis

Data analysis

Raw data were performed to filter out low quality reads by using fastp (v0.23, RRID: SCR_016962)[59]. Clean Reads with UMI were further treated with in-house scripts to eliminate PCR duplication. In brief, clean reads were first clustered according to the UMI sequences, in which reads with the same UMI sequence were grouped into the same cluster. Reads in the same cluster were compared to each other by pairwise alignment, and then reads with sequence identity over 95% were extracted to a new sub-cluster. After all sub-clusters were generated, multiple sequence alignment was performed to get one consensus sequence for each sub-clusters. After these steps, the errors and biases introduced by PCR amplification and sequencing were eliminated. Deduplicated consensus sequences mapped to rRNAs were removed with SortMeRNA[60] (v4.3.3, RRID: SCR_014402). The filtered reads were mapped to GRCh38 (RRID: SCR_027738) by using STAR software (v2.5.3a, RRID: SCR_004463) with default parameters. The MACS2 (v2.2.9.1, RRID: SCR_013291) software was used for peak calling. Peaks were annotated using bedtools (v2.31.1, RRID: SCR_006646). The Guitar R (v 3.18) was used for peak density visualization. The differentially binding peaks were identified by QNB (v1.1.11). Sequence motifs enriched in peak regions were identified using HOMER (v4.11). Gene ontology (GO, RRID: SCR_002811) analysis and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis for annotated genes were both implemented by KOBAS software

(v2.1.1, RRID: SCR_006350) with a *P*-value cutoff of 0.05 to judge statistically significant enrichment.

Statistical analysis

Statistical analysis Data was presented as mean (±SEM) of three independent experiments. Comparisons between two groups of samples were evaluated using the Mann–Whitney U test. All *P*-values reported were two tailed and statistical significance was defined as *P* less than 0.05. All statistical analyses were conducted using R software (R Foundation for Statistical Computing, Vienna, Austria) and GraphPad prism 8.0.2 (RRID: SCR_002798) .

Acknowledgments

This work is supported by grants from the National Key Research and Development Program of China (2025YFC2708404), the Scientific Research Innovation Capability Support Project for Young Faculty (ZYGXQNJSKYCXNLZCXM-H9), the HE Research Fellowship Program (HERF2025005), the China Postdoctoral Science Foundation (2024M761037), the National Natural Science Foundation of China (824B2091 and 825B2092). The authors thank the Laboratory Animal Center of Huazhong University of Science and Techoology for the raising and care of animals.

Author contributions

Conceptualization, X.Z., Z.F., and Y.L.; methodology, X.Z., Z.F., and Y.L.; investigation, X. Z., Z. F., Y. ., J. F., X. H., H. L., X. H., J. J., W. ., D. W., Z. W., C. F., Z. H., C. G., Y. D., T. Q., R. C., D. K. and Y. Z.; writing-original draft, X.Z., Z.F., and Y.L., writing-review and editing, C.S., X.Z., and Z.F.; funding acquisition, N. Y, C.S., X.C., T. Q. and

X.Y. ; resources, C.S., G.C., and N.Y.; supervision, C.S., G.C., and N.Y.

Declaration of interests

The authors declare no competing interests.

Data and materials availability

All data associated with this study are present in the paper or the Supplementary Materials. The data generated in this study, including scRNA-seq, bulk RNA-seq, joint CUT&Tag and rna-seq datasets, RNA-pull down MS and eCLIP-seq, are publicly accessible in the Genome Sequence archive for Human (https://ngdc.cncb.ac.cn/gsa-human/) with accession number HRA010561 and HRA010561 under the project PRJCAXXX (The submission of bulk RNA-seq, joint CUT&Tag and rna-seq datasets, RNA-pull down MS and eCLIP-seq data is currently under review by the GSA curation team, and we anticipate receiving the accession number). Previously published datasets used in this study are listed in the Methods section with their original accession numbers.

References

1. O'Brien, R.L., R.L. Brinster, and U. Storb, *Somatic hypermutation of an immunoglobulin transgene in kappa transgenic mice.* Nature, 1987. **326**(6111): p. 405–9.
2. Shi, W., et al., *Dysregulation of circulating follicular helper T cells in nonsmall cell lung cancer.* DNA Cell Biol, 2014. **33**(6): p. 355–60.
3. Guo, J., et al., *Hepatocyte-derived Igκ promotes HCC progression by stabilizing electron transfer flavoprotein subunit α to facilitate fatty acid β-oxidation.* J Exp Clin Cancer Res, 2024. **43**(1): p. 280.
4. Wang, Q., et al., *A widely expressed free immunoglobulin κ chain with a unique Vκ4-1/Jκ3 pattern promotes colon cancer invasion and metastasis by activating the integrin β1/FAK pathway.* Cancer Lett, 2022. **540**: p. 215720.
5. Huang, X., et al., *A Self-Propagating c-Met-SOX2 Axis Drives Cancer-Derived IgG Signaling That Promotes Lung Cancer Cell Stemness.* Cancer Res, 2023. **83**(11): p. 1866–1882.
6. Cui, M., et al., *Cancer-cell-derived sialylated IgG as a novel biomarker for predicting poor pathological response to neoadjuvant therapy and prognosis in pancreatic cancer.* Int J Surg, 2023. **109**(2): p. 99–106.
7. Vidarsson, G., G. Dekkers, and T. Rispens, *IgG subclasses and allotypes: from structure to effector functions.* Front Immunol, 2014. **5**: p. 520.
8. Constantinescu, A. and M.S. Schlissel, *Changes in locus-specific V(D)J recombinase activity induced by immunoglobulin gene products during B cell development.* J Exp Med, 1997. **185**(4): p. 609–20.
9. Alt, F.W., *From gene amplification to V(D)J recombination and back: a personal account of my early years in B cell biology.* Eur J Immunol, 2007. **37 Suppl 1**(Suppl 1): p. S138–47.
10. Puram, S.V., et al., *Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer.* Cell, 2017. **171**(7): p. 1611–1624.e24.
11. Sun, H., et al., *Quantitative integration of epigenomic variation and transcription factor binding using MAmotif toolkit identifies an important role of IRF2 as transcription activator at gene promoters.* Cell Discov, 2018. **4**: p. 38.
12. Gavish, A., et al., *Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours.* Nature, 2023. **618**(7965): p. 598–606.
13. Fan, J., et al., *Multiomic analysis of cervical squamous cell carcinoma identifies cellular ecosystems with biological and clinical relevance.* Nat Genet, 2023. **55**(12): p. 2175–2188.
14. Liu, Z.L., et al., *Single cell deciphering of progression trajectories of the tumor ecosystem in head and neck cancer.* Nat Commun, 2024. **15**(1): p. 2595.
15. Waas, M., et al., *Molecular correlates for HPV-negative head and neck cancer engraftment prognosticate patient outcomes.* Nat Commun, 2024. **15**(1): p. 10869.
16. Kim, N., et al., *Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma.* Nat Commun, 2020. **11**(1): p. 2285.
17. Lu, Y., et al., *A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma.* Nat Commun, 2022. **13**(1): p. 4594.
18. Jiang, H., et al., *Revealing the transcriptional heterogeneity of organ-specific metastasis in human gastric cancer using single-cell RNA Sequencing.* Clin Transl Med, 2022. **12**(2): p. e730.
19. Rees, A.R., *Understanding the human antibody repertoire.* MAbs, 2020. **12**(1): p. 1729683.

20. Jeon, B.C., et al., *Dynamic O-GlcNAcylation governs long-range chromatin interactions in V(D)J recombination during early B-cell development.* Cell Mol Immunol, 2025. **22**(1): p. 68–82.
21. McConnell, S.A. and A. Casadevall, *New insights into antibody structure with implications for specificity, variable region restriction and isotype choice.* Nat Rev Immunol, 2025. **25**(8): p. 621–632.
22. Tangye, S.G., et al., *Inborn errors of human B cell development, differentiation, and function.* J Exp Med, 2023. **220**(7).
23. Kim, M., et al., *The transcriptional regulatory network modulating human trophoblast stem cells to extravillous trophoblast differentiation.* Nat Commun, 2024. **15**(1): p. 1285.
24. Kim, W., et al., *TFAP2C-mediated upregulation of TGFBR1 promotes lung tumorigenesis and epithelial-mesenchymal transition.* Exp Mol Med, 2016. **48**(11): p. e273.
25. Quirico, L., et al., *The chimeric aptamer axl-miR-214sponge inhibits breast cancer and melanoma dissemination.* Mol Ther, 2025.
26. Wang, Y., et al., *LncRNA MILIP links YBX1 to translational activation of Snai1 and promotes metastasis in clear cell renal cell carcinoma.* J Exp Clin Cancer Res, 2022. **41**(1): p. 260.
27. Flores-Guzmán, F., J. Utikal, and V. Umansky, *Dormant tumor cells interact with memory CD8(+) T cells in RET transgenic mouse melanoma model.* Cancer Lett, 2020. **474**: p. 74–81.
28. Zhang, J., et al., *Inflammation awakens dormant cancer cells by modulating the epithelial-mesenchymal phenotypic state.* Proc Natl Acad Sci U S A, 2025. **122**(36): p. e2515009122.
29. Gocher, A.M., C.J. Workman, and D.A.A. Vignali, *Interferon-γ: teammate or opponent in the tumour microenvironment?* Nat Rev Immunol, 2022. **22**(3): p. 158–172.
30. Garnier, L., et al., *IFN-γ-dependent tumor-antigen cross-presentation by lymphatic endothelial cells promotes their killing by T cells and inhibits metastasis.* Sci Adv, 2022. **8**(23): p. eabl5162.
31. Zheng, R., et al., *hnRNPM protects against the dsRNA-mediated interferon response by repressing LINE-associated cryptic splicing.* Mol Cell, 2024. **84**(11): p. 2087–2103.e8.
32. Zheng, R., et al., *LINE-associated cryptic splicing induces dsRNA-mediated interferon response and tumor immunity.* bioRxiv, 2023.
33. Ivashkiv, L.B., *IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy.* Nat Rev Immunol, 2018. **18**(9): p. 545–558.
34. Liu, Z., et al., *Multi-stage mechanisms of tumor metastasis and therapeutic strategies.* Signal Transduct Target Ther, 2024. **9**(1): p. 270.
35. Han, B., et al., *Cancer incidence and mortality in China, 2022.* J Natl Cancer Cent, 2024. **4**(1): p. 47–53.
36. Wang, Z., et al., *Cancer-derived sialylated IgG promotes tumor immune escape by binding to Siglecs on effector T cells.* Cell Mol Immunol, 2020. **17**(11): p. 1148–1162.
37. Zheng, J., et al., *Immunoglobulin gene transcripts have distinct VHDJH recombination characteristics in human epithelial cancer cells.* J Biol Chem, 2009. **284**(20): p. 13610–13619.
38. Wang, J.P., H.Y. Chen, and H. Peng, *[Functional analysis of cancer-derived immunoglobulin G whole molecule-interacting proteins identified by LC-MS/MS].* Nan Fang Yi Ke Da Xue Xue Bao, 2015. **35**(1): p. 93–7.
39. Deng, Z., et al., *Single-cell RNA sequencing confirms IgG transcription and limited diversity of V(H)DJ(H) rearrangements in proximal tubular epithelial cells.* Sci Rep, 2020. **10**(1): p. 19657.
40. Wang, X., et al., *TFAP2C promotes stemness and chemotherapeutic resistance in colorectal cancer via inactivating hippo signaling pathway.* J Exp Clin Cancer Res, 2018. **37**(1): p. 27.

41. Mughal, M.J., et al., *TFAP2C-DDR1 axis regulates resistance to CDK4/6 inhibitor in breast cancer.* Cancer Lett, 2025. **610**: p. 217356.
42. Tao, Y., et al., *Alternative splicing and related RNA binding proteins in human health and disease.* Signal Transduct Target Ther, 2024. **9**(1): p. 26.
43. Huang, S., et al., *RNA Epigenetics in Cancer: Current Knowledge and Therapeutic Implications.* MedComm (2020), 2025. **6**(8): p. e70322.
44. Zhuang, X., et al., *MRE11:p.K464R mutation mediates olaparib resistance by enhancing DNA damage repair in HGSOC.* Cell Biosci, 2023. **13**(1): p. 178.
45. Wolf, F.A., P. Angerer, and F.J. Theis, *SCANPY: large-scale single-cell gene expression data analysis.* Genome Biol, 2018. **19**(1): p. 15.
46. Xu, S., et al., *Using clusterProfiler to characterize multiomics data.* Nat Protoc, 2024. **19**(11): p. 3292–3320.
47. Mao, X., et al., *Single-cell and spatial transcriptome analyses revealed cell heterogeneity and immune environment alternations in metastatic axillary lymph nodes in breast cancer.* Cancer Immunol Immunother, 2023. **72**(3): p. 679–695.
48. Obradovic, A., et al., *Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages.* Cell, 2021. **184**(11): p. 2988–3005.e16.
49. Garcia-Alonso, L., et al., *Benchmark and integration of resources for the estimation of human transcription factor activities.* Genome Res, 2019. **29**(8): p. 1363–1375.
50. Liao, Y., G.K. Smyth, and W. Shi, *The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads.* Nucleic Acids Res, 2019. **47**(8): p. e47.
51. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform.* Bioinformatics, 2010. **26**(5): p. 589–95.
52. Langmead, B., et al., *Scaling read aligners to hundreds of threads on general-purpose processors.* Bioinformatics, 2019. **35**(3): p. 421–432.
53. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841–2.
54. Ramírez, F., et al., *deepTools2: a next generation web server for deep-sequencing data analysis.* Nucleic Acids Res, 2016. **44**(W1): p. W160–5.
55. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
56. Wang, Q., et al., *Exploring Epigenomic Datasets by ChIPseeker.* Curr Protoc, 2022. **2**(10): p. e585.
57. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576–89.
58. Bailey, T.L., et al., *The MEME Suite.* Nucleic Acids Res, 2015. **43**(W1): p. W39–49.
59. Chen, S., *fastp 1.0: An ultra-fast all-round tool for FASTQ data quality control and preprocessing.* Imeta, 2025. **4**(5): p. e70078.
60. Kopylova, E., L. Noé, and H. Touzet, *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.* Bioinformatics, 2012. **28**(24): p. 3211–7.

Figures and Figure Legends

Figure 1

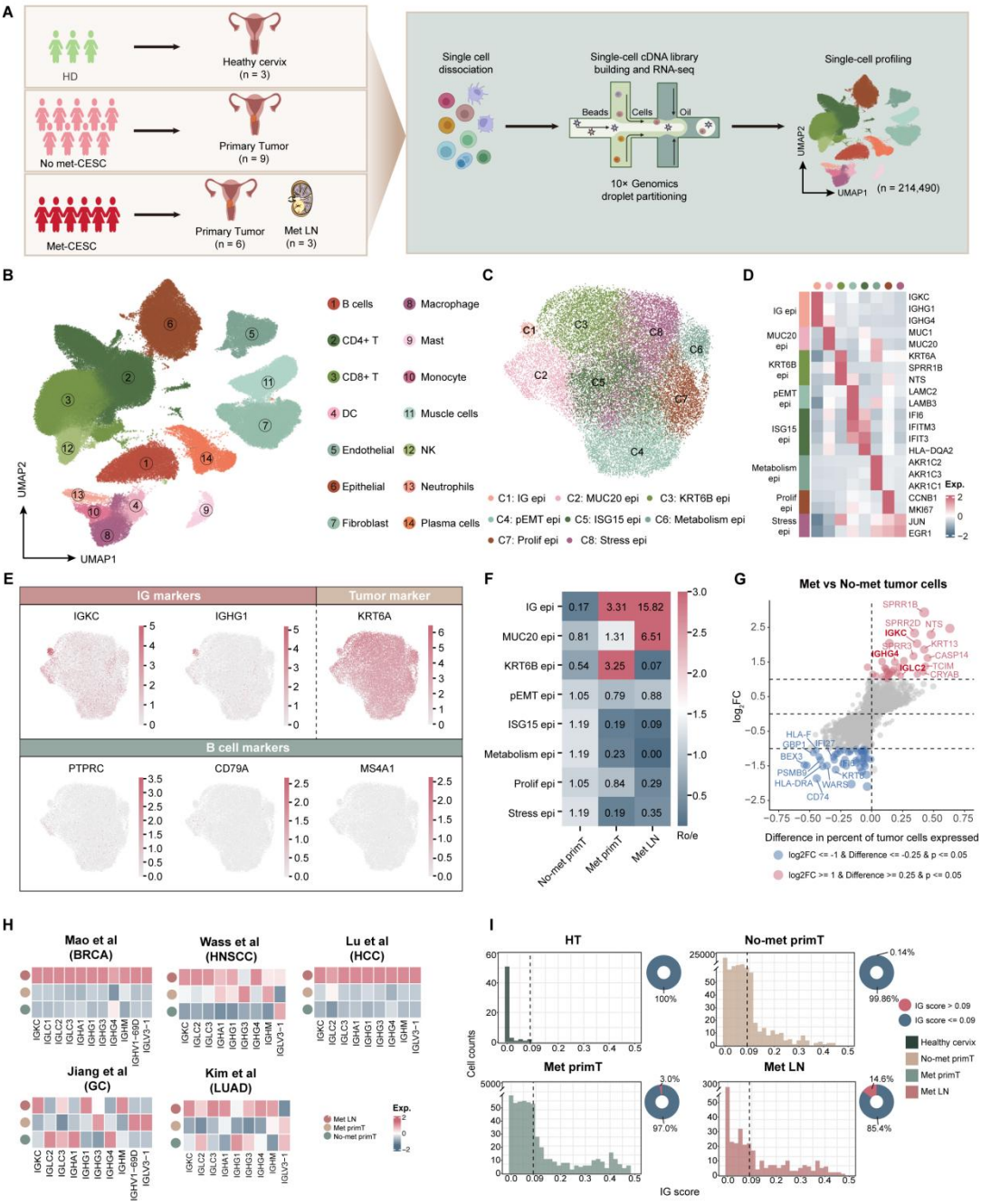


Figure 1. Immunoglobulin Genes Expression Elevates in Cervical Cancer Lymph Node Metastases.

(A) Overview of samples collection and analysis workflow. The study of CESC patients and HDs, including 9 without metastasis and 6 with lymph node metastasis, along with 3 metastatic lymph node samples and 3 healthy cervical tissue samples were collected. (**B and C**) Uniform manifold approximation and projection (UMAP) plot of

all aggregated patient samples (**B**) and tumor cells (**C**) colored by cell types. (**D**) Heatmap shows the expression of marker genes for each tumor subcluster. Key genes are indicated at the right. (**E**) Expression levels of immunoglobulin, tumor and B cells marker genes in tumor cells. (**F**) Heatmap shows the odds ratios (ORs) of tumor cells subset distribution in each group. (**G**) Vlocano plot of upregulated and downregulated genes between tumor cells from metastatic patients and non-metastases. Significance is determined as an adjusted $p \leq 0.05$ (two-sided Wilcoxon rank-sum test) & log2FC ≥ 1 & difference (subtraction between pct.1 and pct.2) ≥ 0.25 (pink dots), or adjusted $p \leq 0.05$ & log2FC ≤ 1 & difference ≤ -0.25 (blue plots). Key genes are labelled. (**H**) Heatmaps show the expression of immunoglobulin genes from different groups, incorporating lymph node metastases (Met LN), metastases (Met primT) and non-metastases (No-met primT) in external validation datasets. (**I**) The histograms of density distribution show IG score in healthy human tissue epithelium, cervix and lymph node tumors. Epithelial cells defined as IG-positive cells and IG-negative cells depend on IG score > 0.09 or ≤ 0.09, and the pie charts exhibit the percentage of IG-positive and IG-negative cells in tissue epithelium.

Figure 2

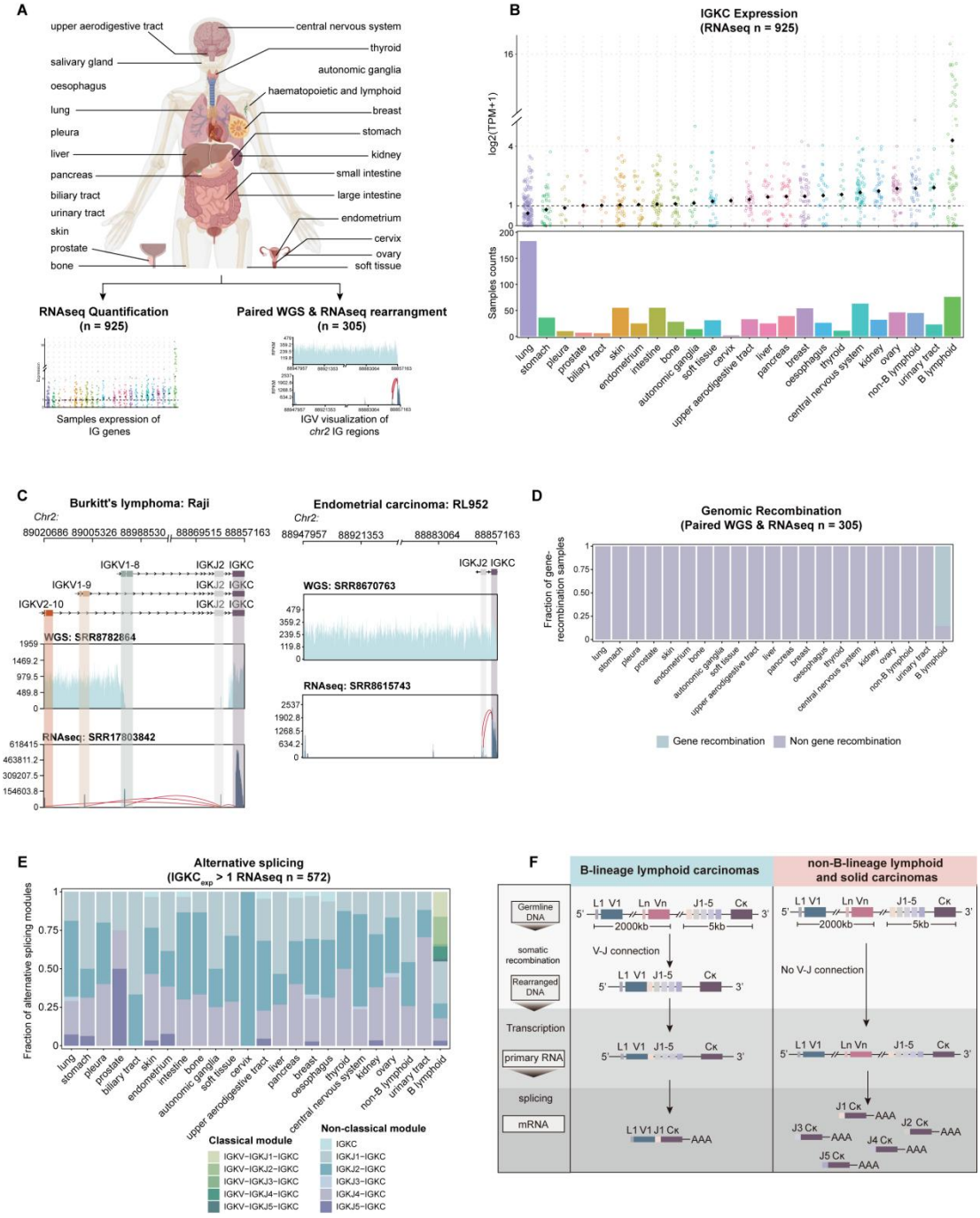


Figure 2. Recombinational Rearrangement of Immunoglobulin Genes in CCLE Occurs at Transcriptomic, not Genomic level.

(**A**) Workflow of whole genome sequencing (WGS) and RNA sequencing (RNA-seq) in Cancer Cell Line Encyclopedia (CCLE) involved in quantitative and rearrangement analysis. (**B**) Top dot plot shows immunoglobulin gene IGKC expression in CCLE. Median expressions in each cancer type are labelled. Line cutoff is set at log2 (TPM+1) =1. Bottom barplot shows the sample counts in each cancer type. (**C**) IGV

visualization of samples (endometrial carcinoma cell line RL952 and Burkitt's lymphoma cell line Raji) immunoglobulin κ light region(chr2: 88,857,163-89,330,429) in WGS and RNA-seq. Models are showed above. (**D**) Stacked bar plot shows the proportion of gene recombination samples in each cancer type in CCLE. Paired WGS and RNA-seq of CCLE are used in Integrative Genomics Viewer (IGV) visualization at immunoglobulin κ light region (chr2: 88,857,163 - 89,330,429). (**E**) Stacked bar plot shows the proportion of AS models in each cancer type in CCLE. Classical models include IGKV-IGKJ1/2/3/4/5-IGKC, non-classical models include IGKC, IGKJ1/2/3/4/5-IGKC. RNA-seq data of CCLE are used in IGV visualization at immunoglobulin κ light region and transcription reassembly by StringTie (*Nature Biotechnology*, 2015). (**F**) Overview of immunoglobulin gene recombination and AS process in B-lineage lymphoid carcinomas and non-B-lineage lymphoid and solid carcinomas cell lines.

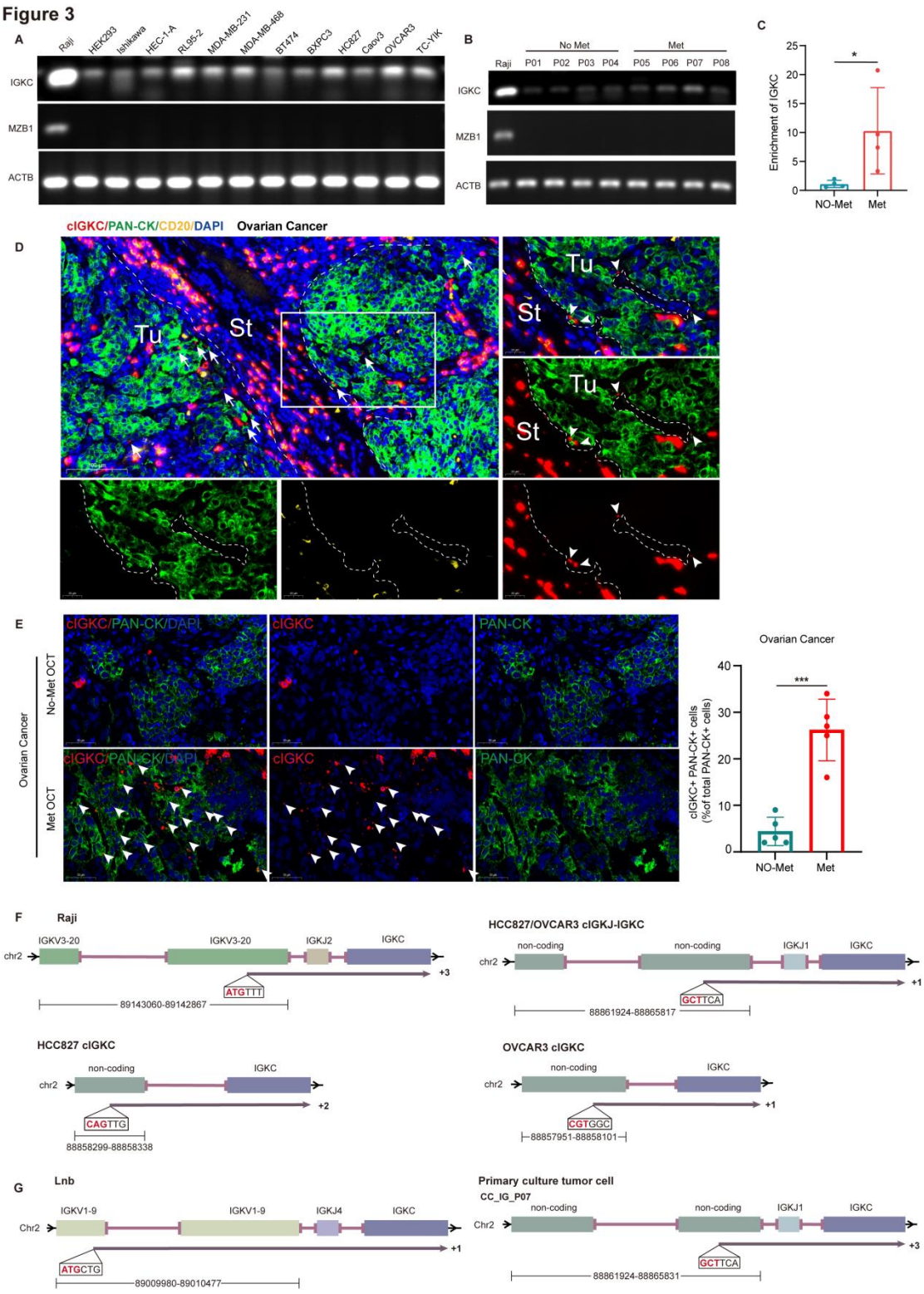


Figure 3. IGKC expression pattern in tumor samples and cell lines.

(A) Agarose gel electrophoresis shows IGKC mRNA amplification bands by RT-PCR across multiple human tumor cell lines, with Raji cell line serving as positive controls. **(B)** RT-PCR analysis of IGKC expression in primary tumor cells extracted

from metastatic (n=4) and non-metastatic (n=4) cervical cancer. Agarose gel electrophoresis shows IGKC amplification bands, with lymph node samples serving as positive controls. **(C)** Quantification of IGKC mRNA levels in primary tumor cell lines by RT-PCR. **(D)** RNAscope detection of IGKC mRNA expression (red), PAN-CK antibodies (green) and CD20 (yellow) were simultaneously used to identify tumor cells in metastatic ovarian cancer. **(E)** RNAscope detection of IGKC expression (red), and PAN-CK antibodies were simultaneously used to identify tumor cells (green) in metastatic (n=5) and non-metastatic (n=5) ovarian cancer tissues. Representative images (left) and quantification (right) are shown. **(F)** Alternative splicing patterns in immunoglobulin κ light region in HCC827and OVCAR3 cell lines, with Raji cells as positive controls. Start codon has been marked in red. 5' RACE and Sanger Sequencing are used for the sequences in cell lines, which blat with GRCh38 in UCSC Genome Browser. **(G)** Alternative splicing patterns in immunoglobulin κ light region in primary cervical cancer cells. Lymphocytes samples were used as positive controls. Start codon has been marked in red. 5' RACE and Sanger Sequencing are used for the sequences in cell lines, which blat with GRCh38 in UCSC Genome Browser. Data are shown as mean±SEM. Statistical significance was determined using a two-tailed Students't test; $p < 0.05$ was considered significant. *$p < 0.05$, **$p < 0.01$.

Figure 4

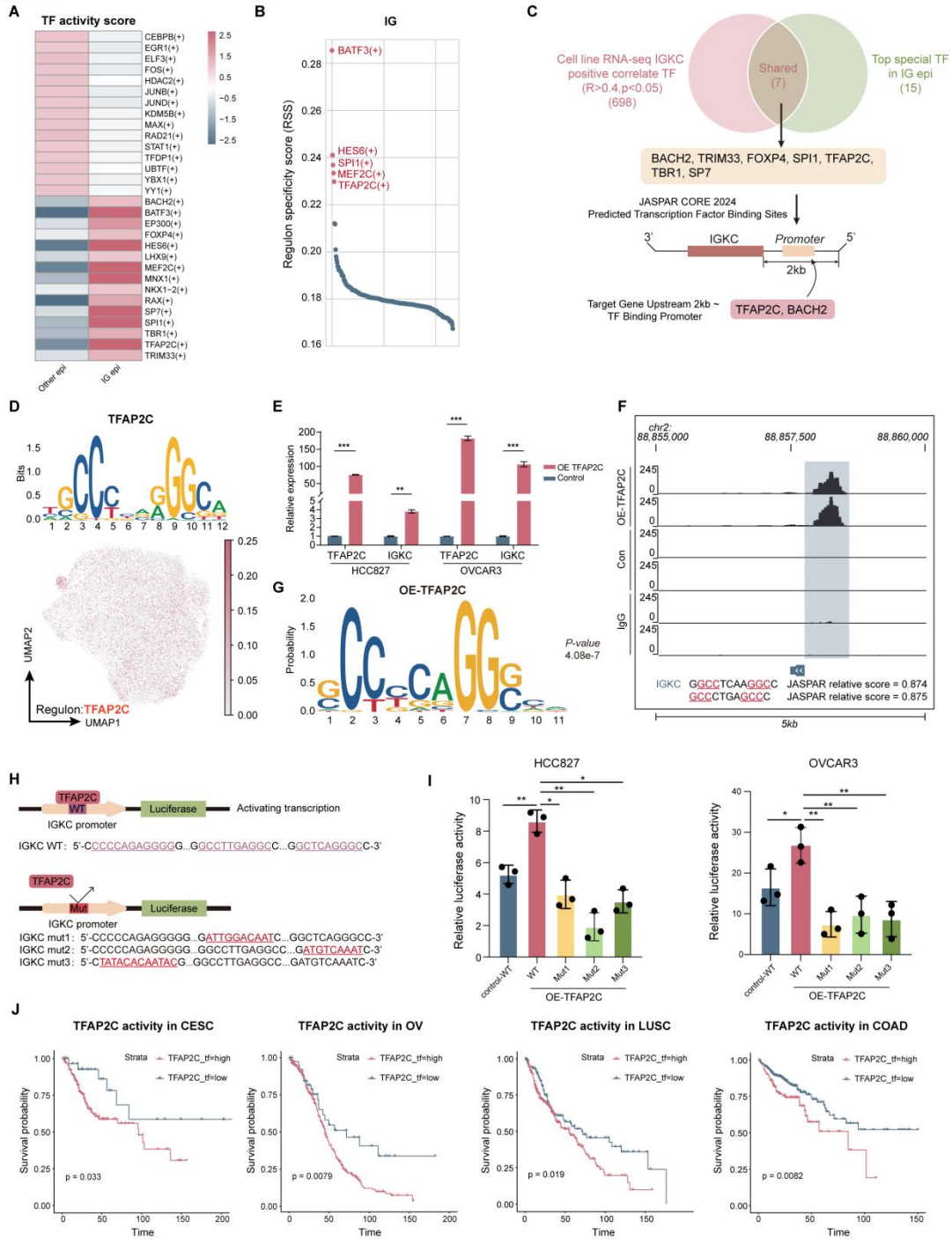


Figure 4. TFAP2C Regulates IGKC Specific Transactivation.

(**A**) Heatmap of transcription factors (TFs) activity score across tumor cells. Top 15 TFs are shown betweeen IG epi and other epi. (**B**) Rank for regulons in IG epi based on regulon specificity score (RSS). (**C**) Flow diagram for IG epi specific TFs selection. (**D**) Binding motif of TFAP2C (top) and UMAP map with binarized regulon activity scores (RAS) for regulon TFAP2C (red dots) (bottom) in tumor. (**E**) IGKC expression

was markedly increased following TFAP2C overexpression. RT-PCR detection of IGKC expression levels in OVCAR3 and HCC827 cells transfected with TFAP2C-overexpressing plasmids or control vectors. (**F**) Motif with 200-bp regions surrounding IGKC-binding peaks in OE-TFAP2C, inferred by JASPAR. (**G**) IGV browser representation of TFAP2C binding in IGKC region. (**H**) A schematic representation of the IGKC promoter sequence and its mutated variants, which were individually or jointly cloned and inserted into the pGL3 vector for the luciferase reporter assay. (**I**) Relative luciferase activity was measured in HCC827 cells (left) and OVCAR3 (right), transfected with wild-type (WT) or mutant IGKC promoter, co-transfected with control or TFAP2C vector. Data are shown as mean±SEM. Statistical significance was determined using a two-tailed Students't test; $p < 0.05$ was considered significant. *$p <$ 0.05, **$p < 0.01$. (**J**) Kaplan–Meier curves comparing the probability of survival according to TFAP2C activity score in cancer patients from TCGA-CESC, TCGA-OV, TCGA-LUSC and TCGA-COAD. Patients were dichotomized as high or low score using the TFAP2C activity score as the cutoff. *p* values are calculated with log-rank tests.

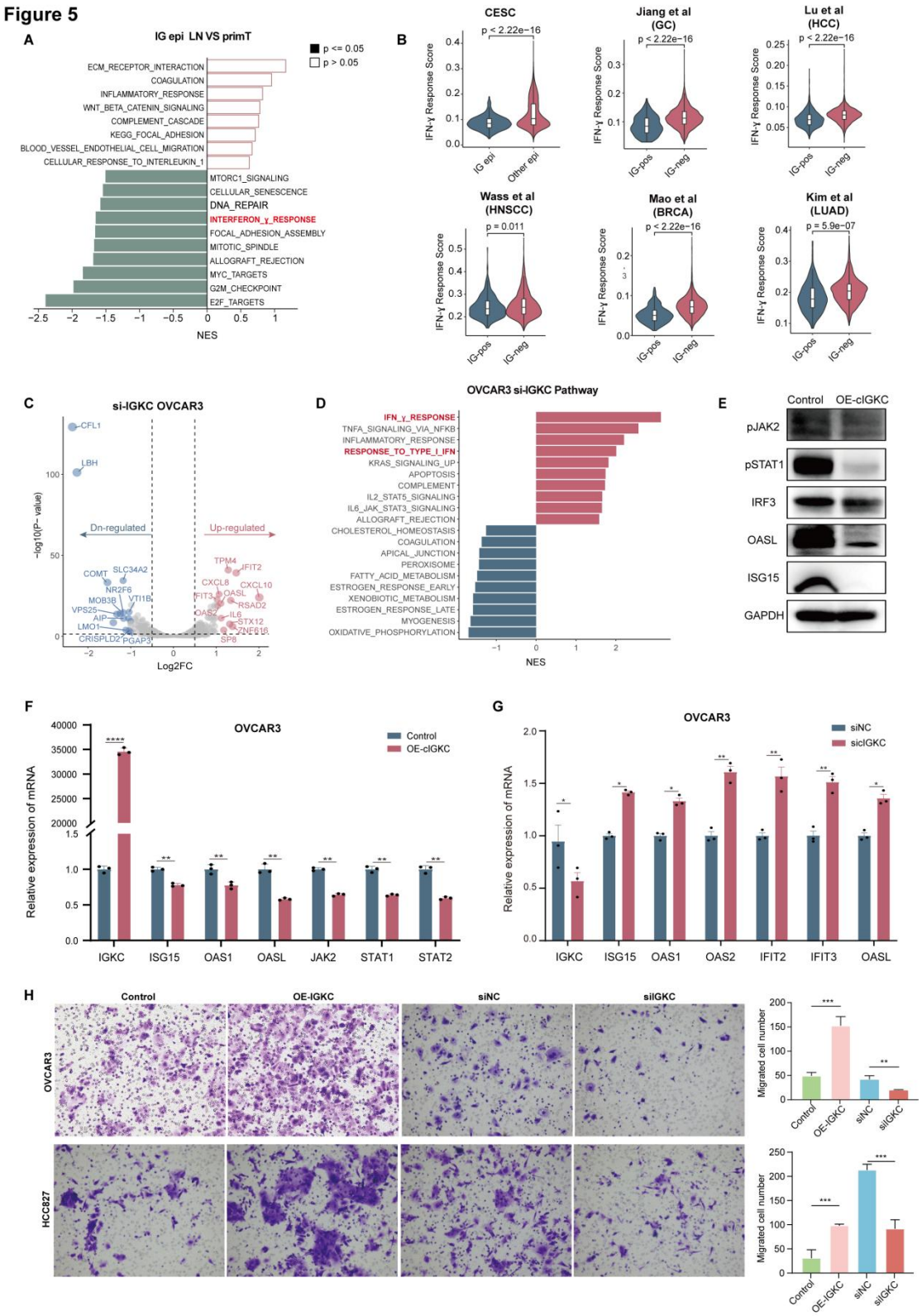


Figure 5. Functional analysis of cIGKC in regulating IFN signaling and tumor cell invasion.

(A) Significantly enriched gene signatures from HALLMARK and Gene Ontology in IG epi of tumor cells from lymph node vs. primary lesion (solid bar plot,

$p \leq 0.05$; hollow bar plot, $p > 0.05$). Normalized enrichment score (NES) of enriched terms are derived from the clusterProfiler package GSEA. (**B**) Violin plots of IFNγ score in-house tumor cells (**CESC**) and outer datasets are included. Likely therasholds of IG-pos and IG-neg are calculated separately in each outer datasets. The inner boxes display the interquartile range (IQR = Q3–Q1; the 25th (Q1) to the 75th percentiles (Q3)), with the centerline denoting the median. *p* value was calculated by two-sided Mann-Whitney U test. (**C**) Vlocano plot of upregulated and downregulated genes in si-IGKC OVCAR3 cell line. Significance is determined as an adjusted $p < 0.05$ (two-sided Wilcoxon rank-sum test), log2FC ≥ 1. Key genes are labelled. (**D**) Bar plots of top enrich terms from HALLMARK and Gene Ontology using DEGs of OVCAR3 si-IGKC cells RNA-seq. Normalized enrichment score (NES) of enriched terms are derived from the clusterProfiler package GSEA. **(E)** Western blot analysis showing the effects of IGKC overexpression on IFNγ response signaling pathway proteins. **(F)** RT–PCR analysis of IFNγ response pathway genes in OVCAR3 cells overexpressing IGKC. **(G)** RT–PCR analysis of IFNγ response pathway genes after siIGKC in HCC827 cells. **(H)** Transwell invasion assays showing the effects of IGKC silencing on the invasive capacity of OVCAR3 and HCC827 cells. Data represent mean ± SEM; $*p < 0.05$, $**p < 0.01$,$***p < 0.001$.

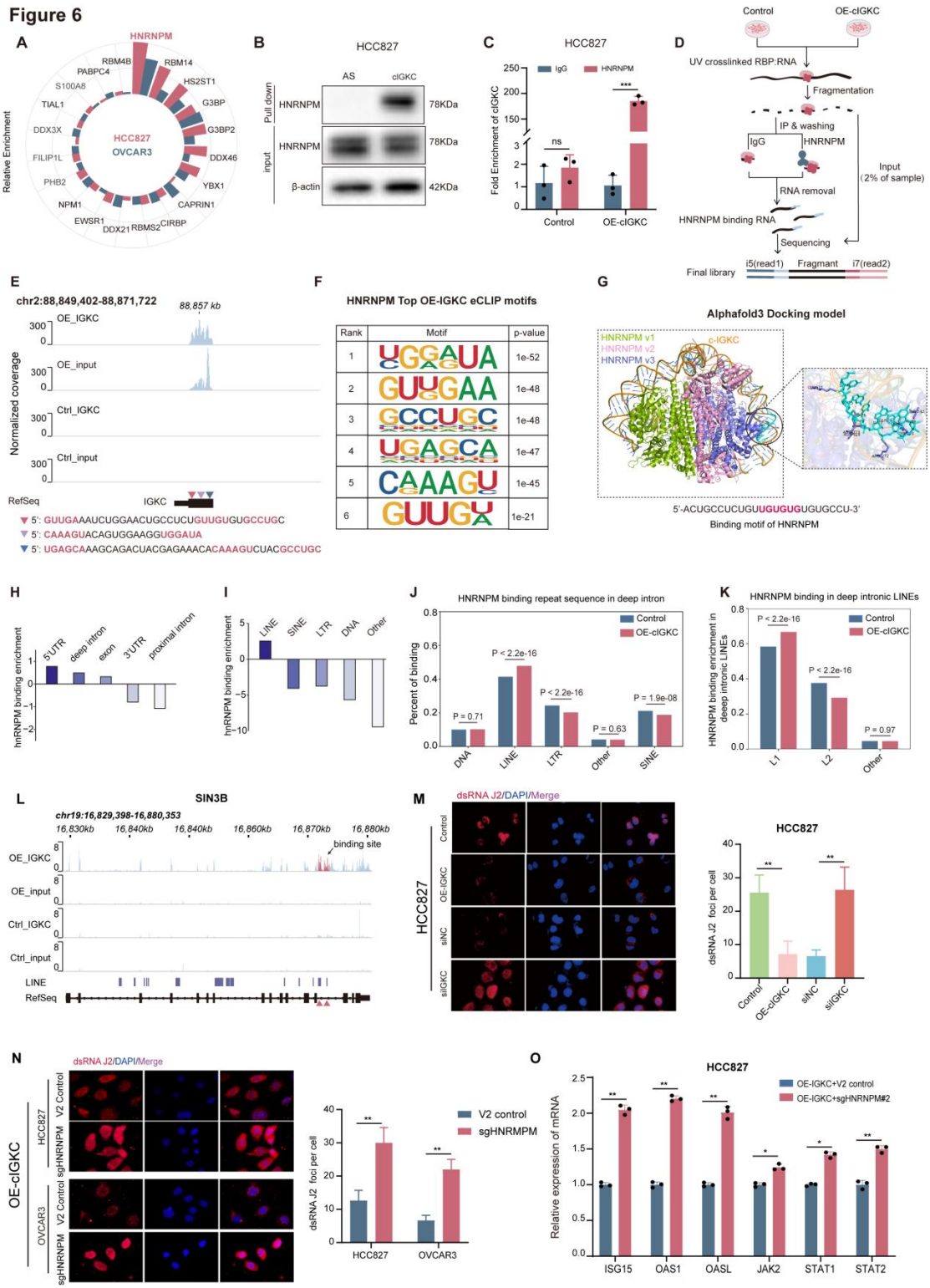


Figure 6. IGKC interacts with hnRNPM to suppress dsRNA formation and IFNγ response signaling.

(A) RNA Binding Proteins binding distribution in IGKC region. Circular bars indicate relative binding enrichment in HCC827 (Red) and OVCAR3 (Blue) cell. **(B)** RNA pull-down assay detecting the interaction between cIGKC and hnRNPM in

HCC827 cells. **(C)** RNA immunoprecipitation showing enhanced cIGKC and hnRNPM binding after IGKC overexpression in HCC827 cells. **(D)** Schematic workflow of eCLIP-seq analysis used to assess the effect of cIGKC overexpression on hnRNPM-bound RNAs.（**E)** hnRNPM eCLIP and input plots of cIGKC over-expression (OE-IGKC) and control. IGV browser representation of hnRNPM binding in IGKC region. Red indicates the UG-rich motifs. (**F**) HOMER motif analysis of with 200-bp regions surrounding IGKC-binding peaks in OE-IGKC reveals 6-nucleotide or 5-nucleotide UC motifs. (**G**) The structure and site of the hnRNPM-IGKC interaction complex were predicted by AlphaFold3. (**H**) Bar plot shows hnRNPM binding enrichment at different elements. (**I**) Bar plot shows hnRNPM binding enrichment at different repeats. (**J**) Bar plot shows percentage of repeats elements bound by hnRNPM in deep intron. *p* values are calculated with Chi-Square Test. (**K**) Bar plot shows percentage of LINE subtypes bound by hnRNPM in deep intronic LINEs. *p* values are calculated with Chi-Square Test. **(L)** IGV browser representation of hnRNPM binding in SIN3B region. Red indicates the LINE-rich motifs. **(M)** Immunofluorescence staining with dsRNA (J2) antibody detecting reduced intracellular dsRNA levels upon cIGKC knockdown in HCC827 cells. **(N)** dsRNA (J2) immunofluorescence showing restored dsRNA levels after hnRNPM interference in cIGKC-overexpressing HCC827 and OVCAR3 cells. **(O)** RT–PCR analysis showing recovery of IFNγ signaling pathway gene expression after hnRNPM knockdown in cIGKC-overexpressing HCC827 cells. Data represent mean ± SEM; $*p < 0.05$, $**p < 0.01$.

Figure 7

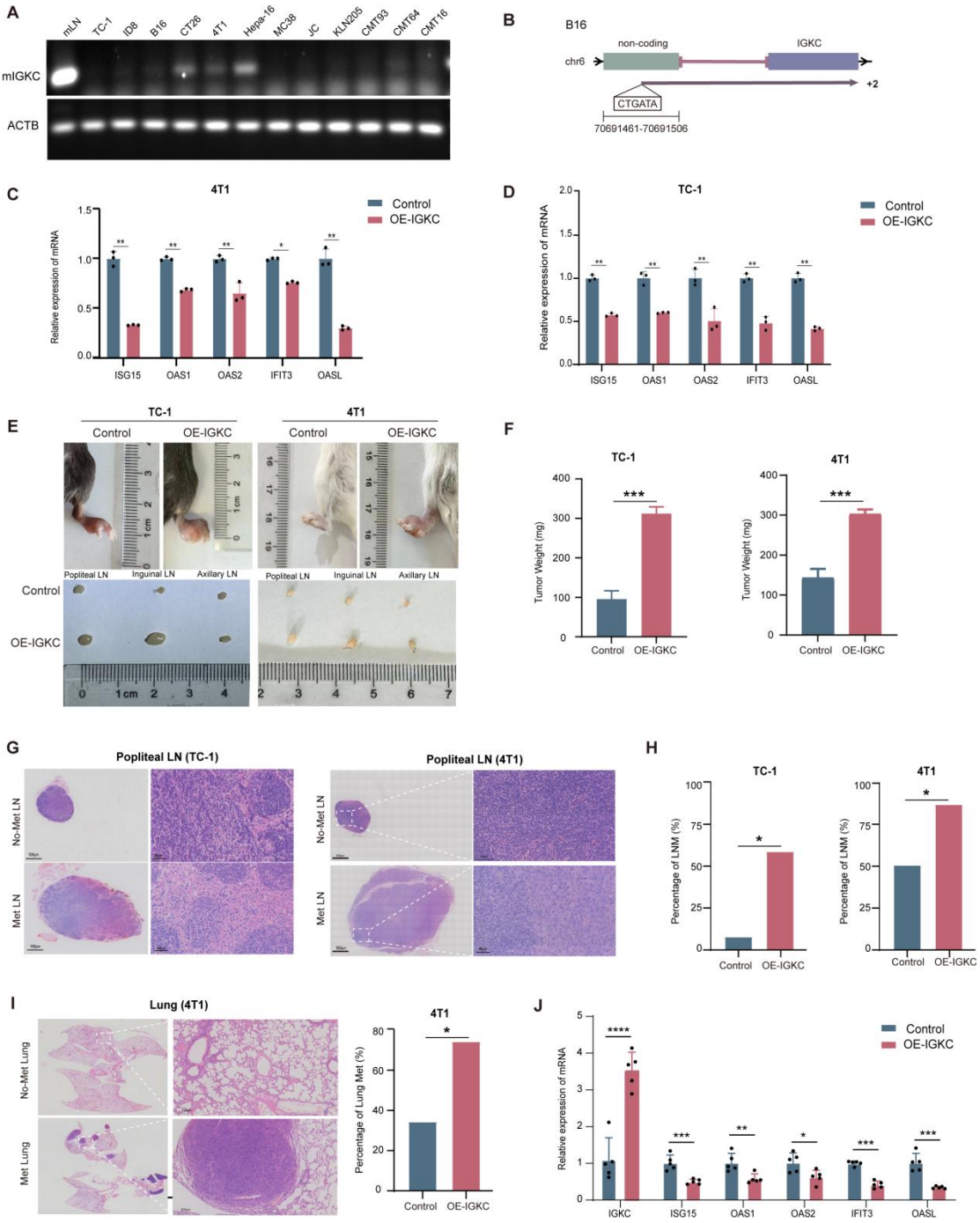


Figure 7. cIGKC suppresses IFNγ responses and promotes tumor metastasis in vivo.

(A) Agarose gel electrophoresis shows IGKC mRNA amplification bands by RT-PCR across multiple murine tumor cell lines, with mouse lymph node serving as positive controls. **(B)** Schematic illustration of IGKC fusion transcript identified by 5′RACE in B16 cell line. **(C and D)** RT-PCR analysis of IFN responds pathway gene

expression in 4T1 **(C)** and TC-1 **(D)** cells with or without mcIGKC overexpression. **(E)** Representative images of primary tumors and draining lymph nodes from footpad lymphatic metastasis models established with TC-1 cell line (right) or 4T1 cells line (left) with or without IGKC overexpression. **(F)** Quantification of primary tumor weights in TC-1 (right) and 4T1 (left) models comparing IGKC-overexpressing (OE-IGKC) (n=5) and control (n=5) groups. **(G)** Representative H&E staining of metastatic and non-metastatic popliteal lymph nodes from footpad lymphatic metastasis models established with TC-1 (right) and 4T1 (left) cells. **(H)** Statistical analysis of lymph node metastasis rates in TC-1 (right) and 4T1 (left) models comparing IGKC-overexpressing (OE-IGKC) and control groups. **(I)** Representative H&E (left) staining and quantification (right) of metastatic and non-metastatic lung tissues from footpad lymphatic metastasis models established with 4T1 cells (left). **(J)** RT-PCR analysis of IFN responds signaling gene expression in primary tumors from footpad lymphatic metastasis models established with TC-1 cells. Data are presented as mean ± SEM. Statistical significance was determined by unpaired two-tailed Student's t-test (for two groups) or one-way ANOVA with Tukey's post hoc test (for multiple groups). $*p < 0.05$, $**p < 0.01$,$*** p < 0.001$.

Figure 8

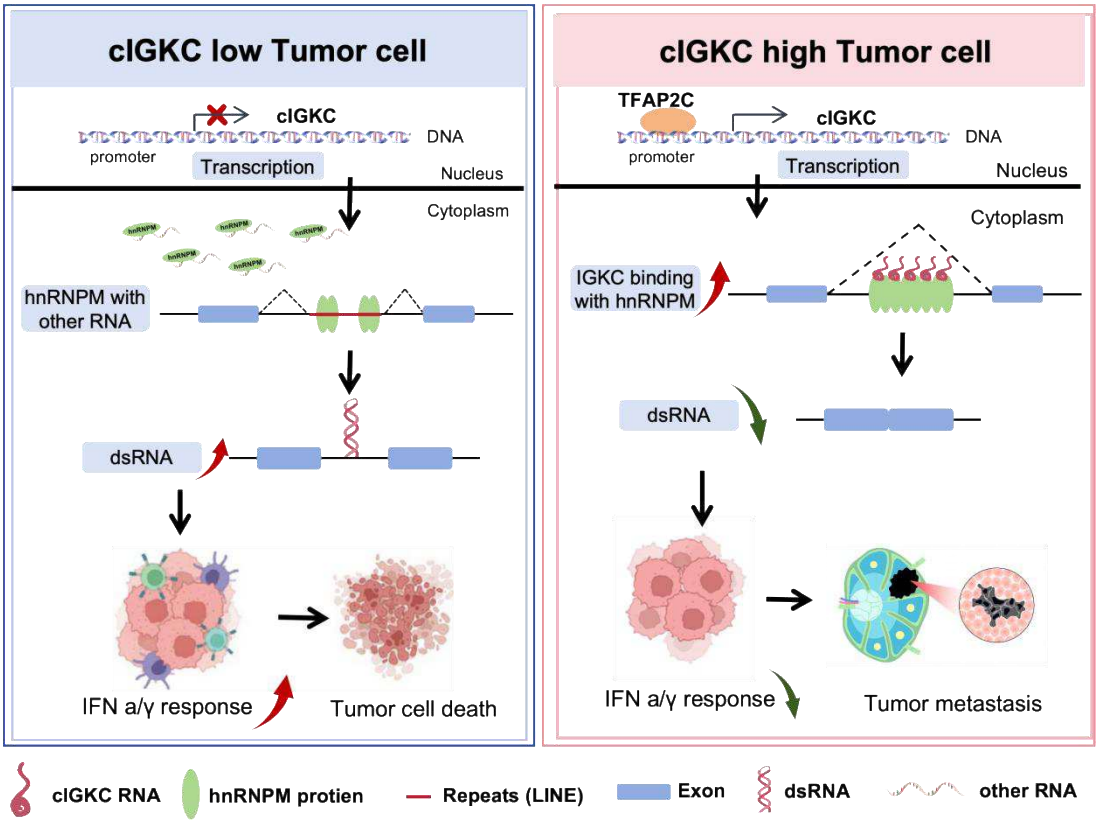


Figure 8. TFAP2C-IGKC-hnRNPM regulatory axis modulates dsRNA sensing and interferon activation to control tumor immune fate.

In tumor cells with low IGKC expression (the left panel) , low TFAP2C transcriptional activity results in minimal IGKC expression. HNRNPM preferentially binds other cellular RNAs, leading to exposure of intronic LINE elements within pre-mRNAs and increased dsRNA formation. Accumulated dsRNA activates interferon signaling, promoting immune-mediated tumor cell elimination. In tumor cells with high IGKC expression (the right panel) , elevated TFAP2C activity induces IGKC transcription. IGKC transcripts competitively bind hnRNPM and direct it to intronic LINE elements within pre-mRNAs, reducing dsRNA formation. Attenuated interferon signaling facilitates immune evasion and tumor progression.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalMaterial.docx
- Figs2.tif
- Figs6.tif
- FigS5.tif
- Figs3.tif
- Figs4.tif
- Figs1.tif
- FigS6G.pdf
- PCR.xlsx
- FigS6B.pdf
- Fig6B.pdf
- Fig5E.pdf