Transcriptional Subtypes on Immune Microenvironment and Predicting Postoperative Recurrence and Metastasis in Pheochromocytoma and Paraganglioma

Yang Liu[1*], Xu Yan[2*], Yibo Zhang[3*], Zhenfu Gao[1], Fengrui Nan[1], Siyu Shi[1], Jingyun Chen[1], Lingyu Li[1,4#]

[1]Cancer Institute, The First Hospital of Jilin University; Changchun, Jilin 130021, China.

[2]Pathology department, The First Hospital of Jilin University; Changchun, Jilin 130021, China.

[3]Clinical laboratory, The First Hospital of Jilin University; Changchun, Jilin 130021, China.

[4]Cancer Center, The First Hospital of Jilin University, Changchun, China

[#]Corresponding author Email: lilingyu@jlu.edu.cn

[*]These authors contributed equally to this work

Abstract

Pheochromocytomas and paragangliomas (PPGLs) exhibit substantial molecular and immune heterogeneity, complicating risk assessment and treatment. Here, we define three distinct transcriptional subtypes (C1, C2, C3) through integrative transcriptomic and immunogenomic profiling. C1 is characterized by hypoxia-driven pathways and an immunosuppressive microenvironment, correlating with poor prognosis. C2 exhibits a highly inflamed immune landscape with robust CD8+ T cell infiltration, suggesting potential sensitivity to immunotherapy. C3 is enriched in metabolic reprogramming pathways and displays intermediate clinical outcomes. Genetic analysis reveals subtype-specific mutational patterns, with pseudohypoxic driver mutations (SDHB, VHL, SDHA, SDHD) predominant in C1 and C3, while kinase pathway alterations (NF1, RET) define C2. Single-nucleus RNA sequencing further delineates immune ecosystem diversity. Notably, we identify ANGPT2, PCSK1N, and GPX3 as key subtype-specific biomarkers, with ANGPT2 driving tumor progression in C1 and emerging as a potential therapeutic target. Our findings provide a refined molecular classification integrating immune and genomic features, offering a framework for improved prognostication and precision therapies in PPGLs.

Keywords: Pheochromocytomas and paragangliomas, transcriptional subtypes, immune infiltration, prognosis assessment

Introduction

Pheochromocytomas and paragangliomas (PPGLs) are rare, highly heterogeneous neuroendocrine tumors that pose significant challenges in prognosis assessment due to their complex molecular landscape and variable clinical behavior [1, 2]. Postoperative recurrence and metastasis rates for PPGLs vary widely, ranging from 6.0 to 17.4% [3], complicating treatment and management strategies. An accurate assessment of the risk for recurrence and metastasis following surgery is essential for optimal management, as it directly informs decisions regarding postoperative surveillance and therapy.

Current genomic subtypes, such as pseudohypoxic, kinase signaling, and Wnt-altered subtypes, have shown promise in identifying the molecular characteristics of PPGLs [4]. However, these approaches have limited applicability, as the overall germline mutation rate in PPGL patients is only 30-40% [5, 6] leaving a substantial portion of PPGLs without actionable genomic markers. This creates a gap in our ability to predict recurrence and metastasis based solely on genetic profiles. While some genotype-phenotype correlations have been established [7-9], molecular subclassifications have advanced our understanding of PPGLs biology, the role of secondary molecular events, particularly in the tumor microenvironment (TME), remains underexplored. The TME has been shown to influence prognosis and response to treatment in other tumor types, and emerging evidence suggests that it may play a critical role in the development and progression of PPGLs. However, its precise contribution to disease progression in PPGLs has yet to be fully elucidated.

This study aims to bridge these gaps by analyzing multiple PPGLs cohorts with a focus on the transcriptional landscape and immune microenvironment. Using weighted gene co-expression network analysis (WGCNA) [10], we seek to define the transcriptional subtypes of PPGLs, explore how these subtypes are linked to immune infiltration patterns, and identify key transcriptional alterations associated with recurrence and metastasis. We also aim to propose potential prognostic markers and assess the impact of immune microenvironment profiles on disease progression. This comprehensive analysis will provide new insights into the biology of PPGLs and improve the predictive accuracy of postoperative recurrence and metastasis risk.

Results

Characterization of Three Transcriptional Subtypes in PPGLs via WGCNA

In this study, we employed WGCNA to explore the transcriptional landscape of PPGLs, successfully identifying three distinct transcriptional subtypes that are associated with varying clinical outcomes and may affect disease progression through mechanisms such as immune infiltration. Our analysis encompassed a comprehensive cohort of 87 PPGLs patients from the

First hospital of Jilin University (PPGLs cohort), supplemented by an extensive dataset from The Cancer Genome Atlas (TCGA) that included 179 samples (TCGA cohort), which allowed us to validate our findings on a broader scale. We utilized Immunohistochemistry (IHC), RNA sequencing, and whole-genome sequencing (WGS) to analyze the samples, integrating molecular profiles with clinical data to predict recurrence and metastasis based on the identified transcriptional subtypes (Fig. 1A and Table S1). we harmonized anatomic site annotations from our PPGL cohort and the TCGA cohort. Stacked bars depict the proportion of adrenal pheochromocytomas (PC, red) and extra-adrenal paragangliomas (PG, blue) within each subtype (Supplementary Fig. 1A). The site distribution is essentially the same across subtypes — approximately three-quarters PC and one-quarter PG—with only minimal, non-systematic variation. Using WGCNA on the PPGL cohort, we resolved 20 co-expression gene modules. Module-level functional enrichment (Fig. 1B) grouped these modules into two overarching biological programs—metabolism-associated pathways and HIF-1 signaling/proliferation—both central to tumor biology and immune modulation. Clustering tumors by these program signatures delineated three major transcriptional subtypes, designated Cluster 1 (C1), Cluster 2 (C2), and Cluster 3 (C3). Subtype C1 is enriched in pathways associated with hypoxia and HIF-1 signaling, which are vital for tumor survival and aggressiveness [11]. Subtype C3's enrichment in metabolic pathways, such as fatty acid metabolism and parathyroid metabolism, reflects its adaptation to metabolic stress, potentially revealing targets for metabolic intervention.

A detailed analysis of gene expression within each subtype revealed several critical marker genes that characterize each group (Fig. 1C). Subtype C1 is marked by genes involved in HIF-1 signaling and angiogenesis, such as *ANGPT2* and *CAV1* [12, 13]. In contrast, subtype C2 is defined by genes, including *PCSK1N* and *FTL*, while subtype C3 shows a significant upregulation of genes like *DYNLL2* and *GPX3*. The neuroendocrine (NE) features, assessed through a NE score, reveal distinct patterns among the three PPGL subtypes (Fig. 1D). Notably, subtype C2 demonstrates significantly elevated NE score in comparison to subtypes C1 and C3, suggesting a strong neuroendocrine phenotype. Subtype C1 exhibited higher cell cycle scores (Fig. 1E), expression of CDK1 (Fig. 1F)[2], epithelial-mesenchymal transition (EMT) scores (Fig. 1G), and hypoxia-inducible factor (HIF) signaling scores (Fig. 1H) compared to subtypes C2 and C3. This suggests that C1 may promote rapid tumor growth, metastasis, and potentially lead to a poorer prognosis. Subtype C2, on the other hand, shows an increased inflamed score compared to subtypes C1 and C3, suggesting a more inflamed tumor microenvironment that may be more responsive to immunotherapy (Fig. 1I).

We aggregated the snRNA-seq data to the sample level by summing gene counts across nuclei to generate pseudo-bulk profiles. These profiles were then normalized for library size, log-transformed (log1p), and z-scaled across samples. Using genesets scores derived from our

WGCNA analysis of PPGLs, we subsequently defined subtypes in the TCGA cohort (Supplementary Fig. 1B) and in Magnus's cohort (Supplementary Fig. 1C), revealing distinct transcriptional subtypes with unique molecular features that underscore disease heterogeneity and suggest mechanisms that may drive divergent clinical behaviors.

Genetic Mutation Patterns and Pathway Alterations Across Transcriptional Subtypes of PPGLs

To elucidate the relationship between transcriptional subtypes and previous genetic mutation classifications, we utilized 84 variation dates from 179 PPGLs patients in TCGA cohort and 30 PPGLs patients in Magnus's cohort. The mutation profile for each subtype highlights distinct genetic alterations in PPGLs (Fig. 2A and Table S2). Subtype C1 exhibits a high frequency of mutations in genes associated with the pseudohypoxia pathway, such as *VHL* and *SDHB* [14], suggesting a unique adaptation to hypoxic conditions that may drive tumor progression. In contrast, subtype C2 predominantly features mutations in genes involved in kinase signaling pathways, including *NF1* and *RET* [15], which imply active signaling networks that could influence tumor growth and response to therapies. Subtype C3 does not exhibit a clearly defined pathway but presents a diverse mutational landscape, indicating the complexity of tumor biology in this group.

When analyzing the proportion of genetic alterations in key pathways (Fig. 2B), we found that more than half of the cases in each subtype lacked detectable driver mutations. Among the patients with genetic alterations, subtypes C1 and C3 exhibited a significantly higher proportion of pseudohypoxic signaling alterations compared to C2. In contrast, subtype C2 was more strongly enriched for kinase pathway mutations and WNT alterations. A closer examination of pseudohypoxic mutations revealed a significant concentration of *SDHB* and *VHL* mutations in subtype C1, while subtype C3 exhibited a notable concentration of *SDHA* and *SDHD* mutations (Fig. 2C). Notably, subtype C2 patients exhibited the highest proportion of kinase pathway involvement, along with the greatest variety of kinase pathway mutations (Fig. 2D).

Mutation features, such as microsatellite instability (MSI), tumor mutational burden (TMB), and mutations in pathways like MYC-associated factor X (MAX), which are commonly used to assess the efficacy of clinical immunotherapy, do not show significant differences among subtypes (Fig. 2E). Additionally, principal component analysis (PCA) effectively distinguished transcriptional subtypes within the PPGLs cohort, further supporting the validity of the subtype classification [16]. In both the Magnus [8] and TCGA cohorts, PCA similarly confirmed that transcriptional subtypes could significantly differentiate PPGLs subgroups, and that transcriptional subtyping outperformed traditional genetic mutation-based classification in distinguishing PPGLs (Fig. 2F). In summary, our study offers a thorough molecular

characterization of PPGLs, identifying three distinct subtypes with unique transcriptional features.

Immune Infiltration Profiles Across Transcriptional Subtypes of PPGLs

To elucidate the immune-infiltration characteristics of different transcriptomic subtypes, we applied the xCell [17] algorithm to analyze the RNA-seq data from the PPGLs and TCGA cohorts, revealing tumor microenvironment infiltration features at the transcriptional level. The heatmap illustrates the distribution and abundance of different immune cells across the three subtypes (Fig. 3A). Subtype C2, which showed significantly higher inflamed scores compared to subtypes C1 and C3 (Fig. 1I), exhibited increased levels of $CD4^+$ Th1 cells and cytotoxic lymphocytes, such as T cells and natural killer (NK) cells. In contrast, hematopoietic stem cells (HSCs) and endothelial cells (ECs) were found to be reduced. Immunohistochemical (IHC) staining analysis of infiltrated $CD8^+$ T cells further supports the quantitative findings from the inflamed score, demonstrating that subtype C2 has more pronounced $CD8^+$ T cell infiltration at the tumor site compared to subtypes C1 and C3 (Fig. 3B). These findings reinforce the heatmap data and suggest an active cytotoxic immune response in C2. In further statistical analysis, both the PPGLs cohort and the TCGA cohort showed that the C2 subtype had significantly higher infiltration of cytotoxic lymphocytes and $CD4^+$ Th1 cells, while exhibiting low infiltration of immunosuppressive cells (such as HSC), reflecting an activated immune microenvironment. In contrast, the C1 subtype was enriched in HSC but showed the lowest levels of cytotoxic lymphocyte and $CD4^+$ Th1 cell infiltration. The immune microenvironment infiltration profile of the C3 subtype was intermediate between the two (Fig. 3C-D). This validation in the TCGA cohort, a larger, independent dataset, underscores the robustness of our findings and their potential relevance across different PPGL populations.

To further investigate the relationship between transcriptomic subtypes and the presence of tumor-infiltrating cytotoxic lymphocytes (TILs), we applied multiple linear regression models. After adjusting for various potential factors that could influence TIL infiltration, we aimed to determine whether transcriptomic subtypes still have an impact on the degree of TIL infiltration. In the PPGLs cohort, regardless of whether the crude model (Model 1), the age, gender, and primary site-adjusted Model 2, or Model 3 (which further adjusted for Ki-67, SSTR2, and SDHB expression based on Model 2) was used, the results showed that compared to the subtype C1, the subtype C2 exhibited the most significant TILs infiltration, followed by the subtype C3 (Fig. 3E). This conclusion was also validated in the TCGA cohort (Fig. 3F). In summary, our extensive immunological profiling across the transcriptional subtypes of PPGLs reveals distinct immune landscapes, with subtype C2 exhibiting an active immunological profile.

178 **Transcriptional Diversity and Immune Microenvironment Characteristics in PPGL**
179 **Subtypes**

180 To further clarify the tumor microenvironment characteristics of each transcriptional subtype,
181 we performed single-nucleus RNA sequencing (snRNA-seq) analysis on 32 PPGL patients in
182 Magnus's cohort [8]. The UMAP analysis successfully categorized the genetic profiles into
183 three distinct subtypes: C1, C2, and C3, while also identifying various cell types, including
184 immune cells, chromaffin, myeloid, and fibroblast cells across the samples (Fig. 4A). This
185 diverse cellular representation highlights the complex biological composition of PPGLs and
186 correlates with specific mutational profiles such as *SDHB*, *VHL*, and *RET*. Consistent with
187 transcriptomic sequencing data from the TCGA cohort (Fig. 2A-D), subtype C1 predominantly
188 harbors *VHL* and *SDHB* mutations, while subtype C2 is primarily enriched in mutations related
189 to kinase pathways such as *RET* and *HRAS*. Subtype C3 is mainly characterized by mutations
190 in *SDHA* and *SDHD*. After extracting the immune cells, we performed further cluster analysis.
191 UMAP visualization revealed that the infiltrating immune cells were categorized into 12 distinct
192 subsets, including $CD8^+$ T, $CD4^+$ T, NK, B cells, dendritic cells (DC), and others (Fig. 4B).
193 Further examination of immune cell distribution using UMAP in subtypes C1, C2, and C3,
194 compared to normal tissue, reveals distinct clustering patterns of immune cells (Fig. 4C-D).
195 This visualization indicates a higher proportion of cytotoxic lymphocyte immune infiltration,
196 specifically $CD8^+$ T cells, in subtype C2 compared to subtypes C1 and C3, with no significant
197 involvement of NK cells. Although no significant differences were observed in the overall
198 macrophage content analysis across the different transcriptional subtypes (Fig. 2C-D), further
199 clustering of macrophages using single-cell sequencing revealed an increased proportion of
200 tumor-suppressing M1 macrophages and a significantly reduced proportion of tumor-promoting
201 tumor-associated macrophages (TAMs) in the C2 subtype compared to the other two subtypes.

202 To determine whether the increased $CD8^+$ T cell infiltration in subtype C2 is associated with
203 enhanced immunoregulatory function, we analyzed ligand-receptor interactions using CellChat.
204 The results revealed that the interactions between $CD8^+$ T cells and themselves, as well as with
205 $CD4^+$ T cells and NK cells, were significantly enhanced through chemokine pathways such as
206 CCL-CCR. In contrast, the TGFB-TGFBR ligand-receptor interaction between $CD8^+$ T cells
207 and immunosuppressive M2 macrophages and TAM cells was significantly weakened (Fig. 4E).
208 In conclusion, the increased infiltration of $CD8^+$ T cells in subtype C2 may contribute to
209 immune activation by interacting with other immune cells in the microenvironment.

210 Pseudotime analysis was used to elucidate the developmental pathways of tumor cells in PPGLs,
211 highlighting dynamic shifts between transcriptional states, particularly the transition from a
212 naive to an activated state within the tumor microenvironment [18]. Pseudotime analysis reveals
213 that compared with normal adrenal cells, subtypes C2 and C1 follow distinct developmental
214 trajectories, while subtype C3 occupies an intermediate position between the C1 and C2

pathways. This suggests that C3 may represent a transitional state between the two, which helps explain the significant differences observed between C1 and C2 in terms of neuroendocrine characteristics, tumor proliferation, HIF-1 pathway activity, and the tumor immune microenvironment, with C3 consistently lying between them (Fig. 4F).

In summary, our multi-omics analysis illuminates the transcriptional diversity and intricate immune interactions present in PPGLs, providing a comprehensive overview of the genetic and immune profiles that characterize each subtype. These results not only deepen our understanding of PPGL biology but also open avenues for more targeted and effective therapeutic strategies tailored to the distinct features of each subtype.

Subtype-Specific Gene Expression Profiles and Clinical Outcomes in PPGLs

To convert this transcriptional subtype into a clinically applicable classification tool, we initially employed a linear regression model to compare the effect values (β value) of various candidate marker genes (Supplementary Fig. 1 D-F) on the transcriptional subtypes. We then identified the genes with the most significant effect values and statistical differences for each subtype as the marker genes for that subtype. Ultimately, we identified the genes *ANGPT2*, *PCSK1N*, and *GPX3*, which are significantly overexpressed and have the most significant β values in subtypes C1, C2, and C3, respectively, as the marker genes for these three subtypes (Fig. 5A and Supplementary Fig. 1 D-F). *ANGPT2*, involvement in angiogenesis and vascular remodeling [12], had significantly higher expression in subtype C1 compared to C2 and C3, which may contribute to the aggressive characteristics of this subtype. *PCSK1N*, encoding the neuroendocrine protein ProSAAS [19], exhibited notably high expression levels in subtype C2. This is consistent with our finding of significantly elevated NE scores in subtype C2 (Fig. 1D), further suggesting that this subtype possesses the most prominent neuroendocrine characteristics. Additionally, *GPX3*, a significantly upregulated molecule in the C3 subtype, is involved in the regulation of redox homeostasis [20], which may be closely associated with the active metabolic state in this subtype. Immunohistochemical staining confirmed the differential expression patterns of ANGPT2, PCSK1N, and GPX3 in C1, C2, and C3 subtypes, respectively (Fig. 5B). These findings not only validate the transcriptomic data but also emphasize the potential biological roles these genes may play in the pathophysiology of PPGLs.

A Kaplan-Meier survival curve analysis revealed significant differences in patient outcomes based on subtype classification (Fig. 5C). Patients in subtype C1 exhibited markedly worse progression-free survival (PFS) compared to those in subtypes C2 and C3, underscoring the critical role of subtype stratification in predicting clinical outcomes. Additionally, Cox regression analysis quantified the impact of these subtypes on PPGL patient prognosis (Fig. 5D). The models, adjusted for various factors, confirmed that subtype C1 was associated with

the highest risk of recurrence and metastasis. In contrast, subtype C2 demonstrated the most favorable PFS, significantly outperforming both subtype C1 and subtype C3.

In conclusion, our analysis highlights the potential of *ANGPT2*, *PCSK1N*, and *GPX3* as subtype-specific biomarkers, offering valuable insights into the molecular heterogeneity and clinical prognosis of PPGLs.

ANGPT2 as a Key Regulator and Diagnostic Marker in PPGL Subtype C1

ANGPT2 had significantly higher expression in subtype C1 compared to C2 and C3, indicating its potential involvement in angiogenesis and vascular remodeling, which may contribute to the aggressive characteristics of this subtype. The Receiver Operating Characteristic (ROC) curve highlights the diagnostic potential of *ANGPT2* in subtype C1, with an area under the curve (AUC) of 0.880 (Fig. 6A), demonstrating its high sensitivity and specificity in distinguishing between the subtypes based on their molecular profiles. The samples were categorized into *ANGPT2*^{high} and *ANGPT2*^{low} groups based on the cutoff of *ANGPT2*'s ROC in the PPGLs cohort. It was revealed that the *ANGPT2*^{high} group was associated with higher scores in the cell cycle, EMT, and HIF-1 signaling pathways (Fig. 6B-D). This suggests that *ANGPT2* may not only serve as a specific biomarker for subtype C1 but also act as a regulator of tumor progression and the tumor's ability to adapt to hypoxic conditions.

To further validate our findings, we employed the rat PPGL cell line PC12 and used CRISPR-Cas9 technology to knock out (KO) the *ANGPT2* gene, which was confirmed by Western blot analysis (Fig. 6E). We then compared the tumor growth curves of *ANGPT2* KO (*ANGPT2*$^{-/-}$) and *ANGPT2*^{WT} PC12 cells in subcutaneous xenograft models in BALB/c nude mice. The results showed that the *in vivo* growth of *ANGPT2*$^{-/-}$ cells was significantly slower (Fig. 6F-G). RNA-seq analysis confirmed that, after *ANGPT2* knockout, the expression scores of key genes involved in EMT, tumor metastasis, proliferation, and hypoxia-related pathways were markedly reduced in the PC12 cell line (Fig. 6H-L). This suggests that *ANGPT2*, a marker gene for subtype C1, functions as a regulatory molecule in the malignant biological behavior of PPGLs. Targeting this molecule may provide a new therapeutic approach for subtype C1 of PPGLs, which is associated with the highest malignancy and poorest prognosis.

Discussion

The significant biological heterogeneity and wide distribution of PPGLs contribute to substantial differences in patient outcomes, making their classification and risk assessment a critical focus in clinical practice. Traditional classification methods based on anatomical origin and pan-genomic approaches centered on pathogenic mutations have offered insights but fail to

fully capture the complexity of these tumors, given the relatively low prevalence of detectable mutations.

Our study addresses these gaps by introducing a novel transcriptional subtype classification framework that effectively distinguishes PPGL patients based on transcriptomic expression. This classification demonstrates significantly improved differentiation capability compared to pan-genomic classification and has been validated across independent cohorts. Immunohistochemical analysis further confirmed the robustness of this classification, revealing subtype-specific clinicopathological features that closely align with their molecular characteristics, underscoring its clinical applicability.

Analysis of genomic characteristics revealed that transcriptional subtypes C1 and C3 were enriched in pseudohypoxic genomic subtypes but exhibited distinct mutation profiles: C1 was primarily associated with *SDHB* and *VHL* mutations, while C3 was predominantly linked to *SDHA* and *SDHD* mutations. Notably, more than half of the patients across all subtypes lacked detectable pathogenic mutations, yet their transcriptome expression profiles mirrored those of patients with such mutations. This finding highlights a previously overlooked patient population in genomic studies.

Further investigation into the tumor microenvironment revealed substantial differences among subtypes. Subtype C1 exhibited the most immunosuppressive microenvironment, characterized by low infiltration of $CD8^+$ T cells and $CD4^+$ Th1 cells, high infiltration of HSCs, and an elevated proportion of TAMs. Immune activation pathways were weaker in C1, with enhanced inhibitory interactions involving M2 macrophages and TAMs. In contrast, the C2 subtype demonstrated the most activated and inflammatory tumor microenvironment, with the highest inflammatory scores and robust immune cell interactions. These findings provide a foundation for stratified immunotherapy strategies.

The assessment of postoperative recurrence and metastatic risk in PPGLs remains one of the most critical challenges in current clinical practice [21]. Identifying high-risk patients with precision has become a major focus of research. While several approaches, such as pathology-based scoring systems (e.g., PASS and PAGG scores) and genomic mutation-based evaluation systems [22], have been explored, none have achieved satisfactory outcomes.

Compared to existing clinical and molecular predictors, risk assessment in PPGL has long relied on the following indicators: clinicopathological features (e.g., tumor size, non-adrenal origin, specific secretory phenotype, Ki-67 index), histopathological scoring systems (such as PASS/GAPP), and certain genetic alterations (including high-risk markers like SDHB inactivation mutations, as well as susceptibility gene mutations in ATRX, TERT promoter, MAML3, VHL, NF1, among others). Although these metrics are highly actionable in clinical practice, they exhibit several limitations: first, current molecular markers only cover a subset

of patients, and technical constraints hinder the detection of many potentially significant variants (e.g., non-coding mutations), thereby compromising the comprehensiveness of prognostic evaluation; second, histopathological scoring is susceptible to interobserver variability; furthermore, the lack of standardized detection and evaluation protocols across institutions limits the comparability and generalizability of results.

Our transcriptomic classification system—comprising C1 (pseudohypoxic/angiogenic signature), C2 (kinase-signaling signature), and C3 (SDHx-related signature)—provides a complementary approach to PPGL risk assessment. These subtypes reflect distinct biological backgrounds tied to specific genetic alterations and can be approximated by measuring the expression of individual genes (e.g., *ANGPT2*, *PCSK1N*, or *GPX3*). This study demonstrates that the classifier offers three major advantages: first, it accurately distinguishes subtypes with coherent biological features; second, it retains significant predictive value even after adjusting for clinical covariates; third, it can be implemented using readily available assays such as immunohistochemistry. These findings suggest that integrating transcriptomic subtyping with conventional clinical markers may offer a more comprehensive and generalizable risk stratification framework. However, this strategy would require validation through multi-center prospective studies and standardization of detection protocols.

In this study, subtype C1 was also identified as the poorest prognostic group, exhibiting heightened tumor proliferation, enhanced EMT, and the highest recurrence and metastatic risk. Importantly, we identified *ANGPT2* was identified as a key marker gene and regulator driving the aggressive behavior of C1, providing both a novel tool for identifying high-risk patients and a potential therapeutic target.

In summary, this transcriptional classification framework advances our understanding of PPGL heterogeneity and provides a theoretical foundation for designing targeted therapeutic strategies. By integrating molecular and immune features, it enables precision medicine approaches that have the potential to improve outcomes for patients with this challenging cancer type.

Methods

Patient cohorts

We conducted a retrospective analysis of patients diagnosed with PPGLs who underwent surgery at the First Hospital of Jilin University between October 2019 and March 2022. Formalin-fixed paraffin-embedded (FFPE) tumor samples from these patients were retrieved and reevaluated by two pathologists using hematoxylin and eosin (H&E) staining and the original pathological reports. Patients with combined PPGLs or evidence of other tumors were excluded from the study. Clinical, pathological, and demographic data were obtained by

reviewing electronic medical records. Follow-up information was gathered through regular visits or telephone interviews. RNA bulk sequencing was performed on FFPE samples from 87 patients with complete clinical data and confirmed PPGL diagnoses, along with 5 FFPE samples of normal adrenal tissue as controls. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki and approved by the Ethics Committee of the First Hospital of Jilin University (23K101-001). Given its retrospective design, the need for informed consent was waived. Patient data were processed and analyzed anonymously to ensure confidentiality.

Cell lines

Rat PPGL cell line PC-12 was obtained from the American Type Culture Collection (ATCC). The identity of the cell line has been verified. Authentication was performed using short tandem repeat (STR) profiling. All cell lines were routinely tested and confirmed to be free of mycoplasma contamination. None of the cell lines used are listed among the commonly misidentified cell lines maintained by the International Cell Line Authentication Committee. It was cultured in RPMI-1640 medium. Penicillin/streptomycin (#15140-122; Gibco) and 10% fetal bovine serum (#35-081-CV; Corning) were added to all media. All cell lines were cultured at 37℃ with 5% CO2.

Mouse models

For CDX models, 1.0×10^6 PC-12 cell lines were subcutaneously injected into the flank of 6-week-old female BALB/c nude mice (Hafukang Co., Ltd., Beijing, China). Tumor size was measured every 2 days, and animal survival rate was recorded every day. Tumor size was calculated as $V = (L \times W^2)/2$, where V is tumor volume, L is the length of the tumor (longer diameter), and W is the width. Mice with tumor size larger than 2.0 cm^3 at the longest axis were euthanized for ethical consideration. The Animal Care and Use Committee of First Hospital of Jilin University approved all procedures concerning mouse studies.

RNA extraction and gene expression profiling

Total RNA was successfully extracted from 87 FFPE samples using TRIzol and RNeasy MinElute Cleanup Kit (Invitrogen). RNA purity was assessed using the NanoDrop Spectrophotometer (Thermo Fisher Scientific™, Waltham, USA). RNA integrity and concentration were measured with the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, Palo Alto, CA, USA). Subsequently, mRNA libraries were

created by using the NEB Next Ultra RNA Library Prep Kit (NEB, Beverly, MA, USA), following the manufacturer's protocol. Geneplus-2000 sequencing platform (Geneplus, Beijing, China) was utilized to sequence the constructed RNA-seq libraries. The sequencing reads containing adapter sequences and low-quality reads were removed to obtain high-quality reads. Reads passing quality control were aligned to the human genome hs37d5 using STAR software[29]. Transcript assembly was conducted by using StringTie2.

Immunohistochemical staining

Immunohistochemical (IHC) staining was performed to evaluate the expression of ANGPT2, PCSK1N and GPX3 in formalin-fixed, paraffin-embedded (FFPE) tissue sections. Tissue samples were first baked at 65 ℃ for 15 minutes. Then deparaffinize the sections in xylene (twice,5 minutes each) and rehydrate through graded ethanol solutions (100%,95%,70%, 5 minutes each). Antigen retrieval was conducted by heating the slides in Tris-EDTA buffer pH 9.0(MVS-0098, MXB) at 95 °C for 2 minutes using a pressure cooker. The slides were then allowed to cool to room temperature and washed with phosphate-buffered saline (PBS).

To block endogenous peroxidase activity, the sections were incubated with 3% hydrogen peroxide solution (BCCK5591, Sigma-Aldrich) for 20 minutes, and then washed with PBS. Non-specific antibody binding was minimized by pre-incubating the slides with 10% goat serum (C0265, Beyotime) for 1 hour at room temperature. The tissue sections were then incubated overnight at 4 °C with primary antibodies targeting ANGPT2, PCSK1N, and GPX3 (proteintech-ANGPT2-24613-1-AP, abnova-PCSK1N-H00027344-M02 and Proteintech-GPX3-13947-1-AP).

After washing with PBS, the sections were incubated with HRP signal enhancement solution (PV-9001, ORIGENE) for 20 minutes and washed, then incubated with enhanced enzyme-labeled goat anti-rabbit IgG polymer (PV-9001, ORIGENE) for 20 minutes at room temperature. For chromogenic detection, the signal was developed using DAB (3,3'-diaminobenzidine) (DAB-4032, MXB) and counterstained with hematoxylin (G1004, Servicebio), differentiation solution (G1039, Servicebio) and bluing reagent (G1040, Servicebio), washed with distilled water. Slides were dehydrated through graded ethanol solutions, cleared in xylene, and mounted using neutral baisam (10004160, SCR).

Stained sections were examined under a light microscope (eclipse e100, Nikon), and images were captured. Evaluate the staining pattern and intensity. Use Fiji image J to measure the percentage of the stained area. Use the statistical software GraphPad Prism to perform T-test.

Weighted gene co-expression network analysis (WGCNA)

Overview and rationale. We applied Weighted Gene Co-expression Network Analysis (WGCNA) to identify modules of genes that show coordinated expression across samples and to relate these modules to study traits. Conceptually, WGCNA transforms pairwise gene–gene correlations into a weighted network, detects modules (clusters) of highly co-expressed genes, summarizes each module by its module eigengene (first principal component of the standardized expression matrix for that module), and correlates module eigengenes with external traits to prioritize biology-linked modules and hub genes.

Preprocessing. Raw counts were normalized and variance-stabilized (log2(TPM+1) with library-size correction). Lowly expressed genes were filtered. Outlier samples were assessed by hierarchical clustering and removed if necessary.

Network construction and soft-threshold selection. We computed pairwise gene correlations and raised the absolute correlations to a soft-thresholding power β to emphasize strong connections while preserving continuous weights. We selected β = [value] using the pickSoftThreshold procedure, targeting a scale-free topology fit $R^2 \geq 0.8$ and mean connectivity within a reasonable range. A signed network (networkType="signed"), which preserves correlation direction, was used to focus modules on positively co-regulated genes.

Module detection and merging. Adjacency matrices were converted to a topological overlap matrix (TOM) to capture shared-neighbor structure. Genes were hierarchically clustered on TOM-based dissimilarity, and initial modules were identified using dynamic tree cutting. Highly similar modules were merged based on eigengene correlation. For large gene sets, we used blockwiseModules to ensure computational scalability.

The bulk-seq data of 87 patients were analyzed by WGCNA using the R package WGCNA, which was divided into three subtypes: C1, C2, and C3. Then, the clustering dendrogram and gene module heat map of the genes were constructed to observe the significant gene modules of each phenotype. At the same time, the correlation analysis heat map between the modules was constructed. The parameters were set as follows: minModuleSize = 15, softPower=7.

Analysis of Immune Infiltration Using xCell

The immune infiltration in tumor samples was assessed using the R package xCell to predict the relative proportions of 19 immune and stromal cell types within the tumor microenvironment. The final results were visualized using a heatmap generated by the R package pheatmap, displaying significant differences in cell enrichment across various samples. Violin plots were utilized to illustrate the differences in immune and stromal cell distributions among the three tumor subtypes in the PPGLs Cohort and TCGA datasets.

Gene Mutation Analysis

For Gene Mutation Landscape, mutation profiles of each sample of TCGA data for genes related to PCPG, were analyzed using the R package maftools. A waterfall plot of the gene mutation landscape was generated using the oncoplot function. For tumor mutation burden (TMB), microsatellite instability (MSI) and mutant-allele tumor heterogeneity (MATH) analysis, TCGA data was utilized, and the R package maftools was employed to read and calculate the MSI, MATH, and TMB scores. The MSI, MATH, and TMB scores across the three subtypes were compared to observe the mutation patterns in the three subtypes.

Single-cell nuclear transcriptome sequencing(snRNA-seq)

For analysis of snRNA-seq data, single cell nucleus RNA sequencing datasets were downloaded from the European Genome-Phenome Archive (EGA) under accession code EGAS00001005861. The snRNA-seq data was analyzed using the R package Seurat v4.4.1. The scRNA-seq data underwent filtration to incorporate genes expressed in no less than 3 cells, as well as cells that exhibited gene expression between 200 and 5000 genes, with mitochondrial transcripts comprising less than 5%. Given patient-driven clustering observed in the analysis of immune populations, we applied the SCTransform v2 method to correct the batch effects between different patients. To ascertain that SCTransform v2 solely addressed technical discrepancies and not biological variations, we manually compared the resulting clusters and their associated markers to those obtained from a single library cluster analysis. After normalizing the data, dimensionality reduction and clustering were performed, with unsupervised clustering analysis used to identify different cell subpopulations. The predominant cellular subpopulations within the and immune compartments were delineated based on cluster-averaged gene expression levels of the subsequent gene markers: 1) Chromaffin cells, PCPG and ATPG: *TH and CHGA*, 2) Adrenocortical cells: *STAR*, 3) *ECs*: *FLT1*, 4) *SCLCs*: *SOX10*, 5) Fibroblasts: *PDGFRB*, 6) Myeloid cells: *MSR1*, 7) T/NK cells: *CD2*, 8) Mast cells: *HDC*, 9) B cells: *CD79A*. For in-depth investigation of immune cell populations, the dataset was narrowed down to exclusively include these specific cell types. Subsequently, the data underwent comprehensive processing from its raw form, following the aforementioned methodology. The DimPlot function was used to visualize UMAP plots for all cell types, infiltrated immune cells, and different subtypes. The ggplot function was utilized to create stacked bar plots of cell type proportions.

Cellchat analysis

CellChat is an R package designed for inference, analysis, and visualization of cell-cell

communication from single-cell and spatially resolved transcriptomics. CellChat aims to enable users to identify and interpret cell-cell communication within an easily interpretable framework, with the emphasis of clear, attractive, and interpretable visualizations.

Pseudotime and enrichment analysis (monocle2)

The R package of monocle2 was used to perform the diffusion pseudotime analysis. Normalized data computed in Seurat was directly transferred to monocle2. The differential gene analysis function in monocle2 was used to identify genes with significant changes between C1, C2, C3, and Normal subtypes (q-value < 0.05), which were then utilized as input for temporal ordering of cells along the differentiation trajectory. The final results were subsequently transferred back to Seurat. Then, we performed KEGG and GO enrichment analysis by using clusterProfiler package (v3.18.1).

Survival Analysis

The data for survival analysis was derived from PPGLs Cohort. The R packages survival and survminer were utilized to plot Kaplan-Meier survival curves. The survfit function was applied to fit the survival model. Ultimately, the ggsurvplot was employed to plot the survival curves, depicting the survival status of patients in different subtypes.

Software and versions

Seurat (v4.4.0), sctransform (v0.4.2), CellChat (v2.2.0), monocle (v2.36.0; monocle2), pheatmap (v1.0.13), clusterProfiler (v4.16.0), survival (v3.8.3), and ggplot2 (v3.5.2).

Data availability

The bulk-tissue RNA-seq datasets generated in this study are available from the corresponding author upon reasonable request. We gratefully acknowledge Prof. Richard W. Tothill for providing the single-nucleus RNA-seq dataset, which has been deposited in the European Genome-Phenome Archive (EGA) under accession EGAS00001005861.

Reference

1. Dahia, P.L., *Pheochromocytomas and Paragangliomas, Genetically Diverse and Minimalist, All at Once!* Cancer Cell, 2017. **31**(2): p. 159-161.
2. Calsina, B., et al., *Genomic and immune landscape Of metastatic pheochromocytoma and paraganglioma.* Nat Commun, 2023. **14**(1): p. 1122.
3. Parisien-La Salle, S., et al., *Postoperative Recurrences in Patients Operated for Pheochromocytomas and Paragangliomas: New Data Supporting Lifelong Surveillance.* Cancers (Basel), 2022. **14**(12).
4. Burnichon, N., et al., *Integrative genomic analysis reveals somatic mutations in pheochromocytoma and paraganglioma.* Hum Mol Genet, 2011. **20**(20): p. 3974-85.
5. Horton, C., et al., *Universal Germline Panel Testing for Individuals With Pheochromocytoma and Paraganglioma Produces High Diagnostic Yield.* J Clin Endocrinol Metab, 2022. **107**(5): p. e1917-e1923.
6. Amar, L., et al., *Genetic testing in pheochromocytoma or functional paraganglioma.* J Clin Oncol, 2005. **23**(34): p. 8812-8.
7. Cui, Y., et al., *Local-Regional Recurrence of Pheochromocytoma/Paraganglioma: Characteristics, Risk Factors and Outcomes.* Front Endocrinol (Lausanne), 2021. **12**: p. 762548.
8. Zethoven, M., et al., *Single-nuclei and bulk-tissue gene-expression analysis of pheochromocytoma and paraganglioma links disease subtypes with tumor microenvironment.* Nat Commun, 2022. **13**(1): p. 6262.
9. Li, M., et al., *Genetic background and intraoperative haemodynamic instability in patients with pheochromocytoma and paraganglioma: a multicenter retrospective study.* Int J Surg, 2024.
10. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**: p. 559.
11. Domenech, M., et al., *Hypoxia: The Cornerstone of Glioblastoma.* Int J Mol Sci, 2021. **22**(22).
12. Kapiainen, E., et al., *The Amino-Terminal Oligomerization Domain of Angiopoietin-2 Affects Vascular Remodeling, Mammary Gland Tumor Growth, and Lung Metastasis in Mice.* Cancer Res, 2021. **81**(1): p. 129-143.
13. Nwosu, Z.C., et al., *Caveolin-1 in the regulation of cell metabolism: a cancer perspective.* Mol Cancer, 2016. **15**(1): p. 71.
14. Crona, J., et al., *Genotype-phenotype correlations in pheochromocytoma and paraganglioma: a systematic review and individual patient meta-analysis.* Endocr Relat Cancer, 2019. **26**(5): p. 539-550.
15. Fishbein, L., et al., *Comprehensive Molecular Characterization of Pheochromocytoma and Paraganglioma.* Cancer Cell, 2017. **31**(2): p. 181-193.
16. Chu, X., et al., *Integrative single-cell analysis of human colorectal cancer reveals patient stratification with distinct immune evasion mechanisms.* Nat Cancer, 2024. **5**(9): p. 1409-1426.
17. Aran, D., Z. Hu, and A.J. Butte, *xCell: digitally portraying the tissue cellular heterogeneity landscape.* Genome Biol, 2017. **18**(1): p. 220.

18. Qiu, X., et al., *Single-cell mRNA quantification and differential analysis with Census.* Nat Methods, 2017. **14**(3): p. 309-315.
19. Feng, Y., S.E. Reznik, and L.D. Fricker, *ProSAAS and prohormone convertase 1 are broadly expressed during mouse development.* Brain Res Gene Expr Patterns, 2002. **1**(2): p. 135-40.
20. Kritsiligkou, P., et al., *Unconventional Targeting of a Thiol Peroxidase to the Mitochondrial Intermembrane Space Facilitates Oxidative Protein Folding.* Cell Rep, 2017. **18**(11): p. 2729-2741.
21. Kiriakopoulos, A., P. Giannakis, and E. Menenakos, *Pheochromocytoma: a changing perspective and current concepts.* Ther Adv Endocrinol Metab, 2023. **14**: p. 20420188231207544.
22. Kim, H.S., J. Kweon, and Y. Kim, *Recent advances in CRISPR-based functional genomics for the study of disease-associated genetic variants.* Exp Mol Med, 2024. **56**(4): p. 861-869.

Author contributions

Yang Liu performed experiments, analyzed data, performed statistical analysis and wrote the manuscript; Xu Yan, Yibo Zhang, Zhenfu Gao, Siyu Shi and Jingyun Chen acquired data; Fengrui Nan performed statistical analysis; Lingyu Li conceived the experiments, performed statistical analysis and wrote the manuscript.

Funding

National Natural Science Foundation of China [grant number 82172690 and 82472799] to Lingyu Li

Figure Legends

Figure 1. Divergence in transcriptional subtypes defined three PPGL subtypes across different datasets.

(A) Overview of PPGLs Cohort (n = 87) and TCGA Cohort (n=179) detailing the number of cases by anatomical localization: HN-PG (head and neck paraganglioma), ME-PG (mediastinal paraganglioma), RP-PG (retroperitoneal paraganglioma), BC-PG (bladder paraganglioma), and PC (pheochromocytoma). Clinical data, whole genome sequencing (WGS), and RNA-seq were used for subtype identification and correlation analysis related to recurrence/metastasis. (B) Heatmap depicting key metabolic and biological pathways associated with each subtype and displaying the pathway enrichment analysis for three subtypes of pheochromocytoma: C1 (HIF-1), C2 (Inflamed), and C3 (Metabolism). Each row represents a specific biological pathway, and the color intensity represents the degree of enrichment across the subtypes. (C) Heatmap showing gene expression changes across subtypes. Specific genes like *ANGPT2* and *CAV1* are upregulated in C1 compared to other subtypes. (D-E) Violin plots displaying the enrichment scores for Neuroendocrine and cell cycle among the three subtypes (C1, C2, C3). (F) Violin plots displaying *CDK1* among the three subtypes (C1, C2, C3). (G-I) Violin plots displaying the enrichment scores for EMT, HIF and Inflamed among the three subtypes (C1, C2, C3). Comparisons were calculated by one-way ANOVA (D-I). Data are presented as mean ± SD. ns, $p > 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Figure 2. Mutational landscape and molecular characterization of PPGL cohorts.

(A) Overview of the genomic alterations across the three subtypes (C1, C2, C3) in TCGA cohorts. Genes associated with pseudohypoxic, kinase, and other pathways are shown, with individual rows representing distinct gene mutations and columns corresponding to samples. Bar Plot Showing the Proportional Distribution of Variant Classifications. (B) Proportional analysis of the transcriptional subtypes across C1, C2, and C3. The Pseudohypoxic, Kinase, WNT-Altered, and Wild Type (WT) subtypes (lacking pathogenic mutations) are depicted, illustrating the distribution of each subtype. (C) The number of patients with mutations in the pseudohypoxic-related genes (*SDHB*, *VHL*, *EGLN1*, *EPAS1*, *SDHA*, *SDHD*) across the three subtypes. (D) The number of patients with mutations in kinase-related genes (*NF1*, *RET*, *FGFR1*, *HRAS*, *MAX*, *TMEM127*) across subtypes. (E) Comparisons of Microsatellite Instability (MSI), Tumor Mutation Burden (TMB), and MATH (Mutant Allele Tumor Heterogeneity) scores across the three subtypes. (F) Linear regression analysis of beta coefficients (PC1 and PC2) for different cohorts including PPGLs Cohort, TCGA, and Magnus's Cohort. Comparisons were calculated by one-way ANOVA (E). Data are presented as mean ± SD. ns, $p > 0.05$.

Figure 3. Immune infiltration landscape across PPGLs and TCGA cohorts.

(A) Heatmap displaying the immune cell composition and clinical features, including nerve and vascular infiltration, gender, and age across C1, C2, and C3 subtype. The heatmap highlights variations of immune cell infiltration abundance across different samples in immune cell types such as dendritic cells (DC), macrophages, B cells, and NK cell. (B) Representative immunohistochemistry images (top) and the absolute number (bottom) of infiltrated $CD8^+$ T cells in C1, C2, and C3 subtypes. The infiltrated number of $CD8^+$ tumor-infiltrating lymphocytes (TILs) was determined by randomly selecting 10 fields at 40× magnification and calculating the total count of $CD8^+$ cells. (C-D) The xCell analysis in the PPGLs cohort (C) and TCGA cohort (D) depicting absolute abundance of immune cells across C1, C2, and C3 subtype. Th1, T helper 1 cells; Th2, T helper 2 cells; Mφ, macrophages; HSC, hematopoietic stem cell; EC, endothelial cell. Cytotoxic lymphocytes include both NK cells and T cells. (E-F) The forest plot illustrates tumor-infiltrating cytotoxic lymphocytes (TILs) across different transcriptional subtypes in the PPGL cohort (E) and the TCGA cohort (F), using linear regression models. Model 1 was crude model; Model 2 was adjusted for age, gender, and primary tumor location based on Model 1; Model 3 was adjusted for *MKI67*, *SDHB*, and *SSTR2* expression based on Model 2. Each model is compared to subtype C1 as the reference group. Orange represents subtype C2, and green represents subtype C3. Comparisons were calculated by one-way ANOVA (B-D). Data are presented as mean ± SD. Ns, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Figure 4. UMAP analysis of cell types and pseudotime trajectories in different subtypes

(A) UMAP plot showing the distribution of various cell types across three subtypes (C1, C2, and C3) and normal tissues. Each subtype is uniquely colored, with consistent coloring for the same subtype. Corresponding mutated genes are indicated. PC/AT, Paraganglioma/Adrenal Tumor; NAMs, Neuroblastic cells; ECs, Endothelial Cells; SCLCs, Schwann cell-like cells. (B) UMAP plot highlighting specific immune cells across three subtypes and normal tissues, with corresponding marker genes indicated. (C) UMAP plot highlighting specific immune cells in each subtype: C1, C2, C3, and normal tissues. (D) Bar plot represents the proportion of each immune cell type in the C1, C2, C3 and normal subtypes. (E) The dotplot represents a ligand-receptor interaction plot generated using CellChat, showcasing the communication network between various immune cell types across three different subtypes (C1, C2, C3). The rows represent different immune cell populations, while the columns depict various ligands and receptors involved in cell-cell communication. The color gradient from red to green indicates interaction strength, with red representing stronger interactions and green representing weaker ones. The size of the dots reflects the significance level of each interaction. Distinct regions of enhanced interactions are highlighted with colored boxes, emphasizing subtype-specific signaling patterns. (F) Pseudotime analysis revealing the developmental trajectories of cell

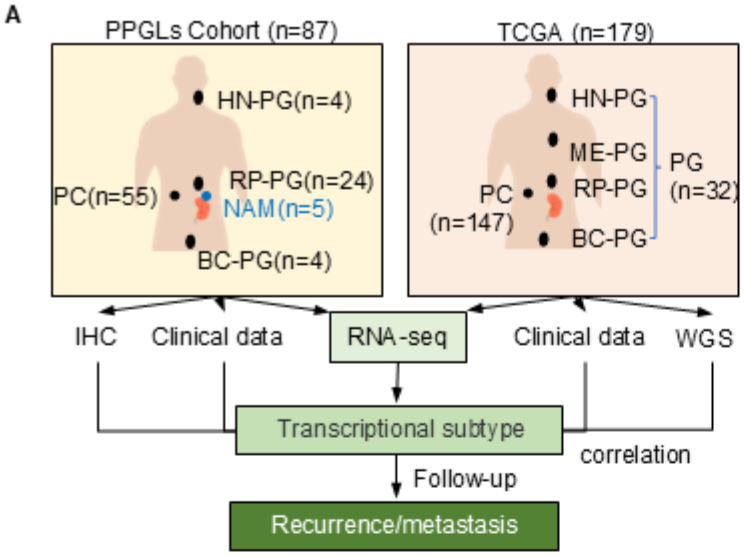
677 populations in different subtypes. Arrows indicate the progression paths of cell differentiation,
678 suggesting different developmental stages and lineage decisions across C1, C2, C3 and normal
679 subtype.

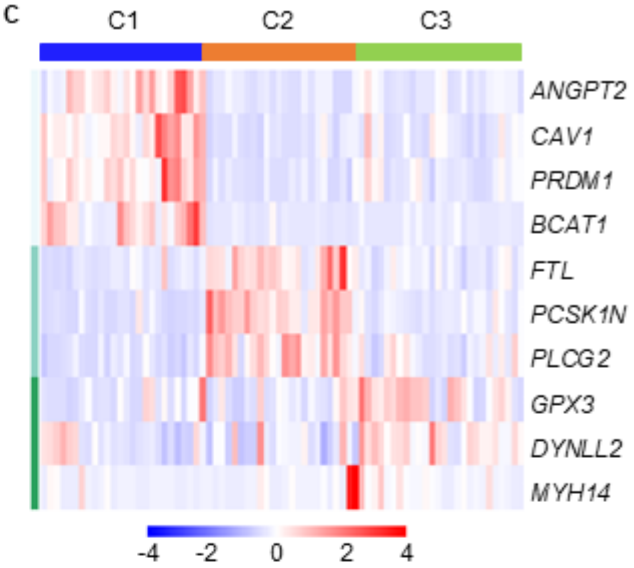
680 **Figure 5. The specific genetic features of different PPGL subtypes**

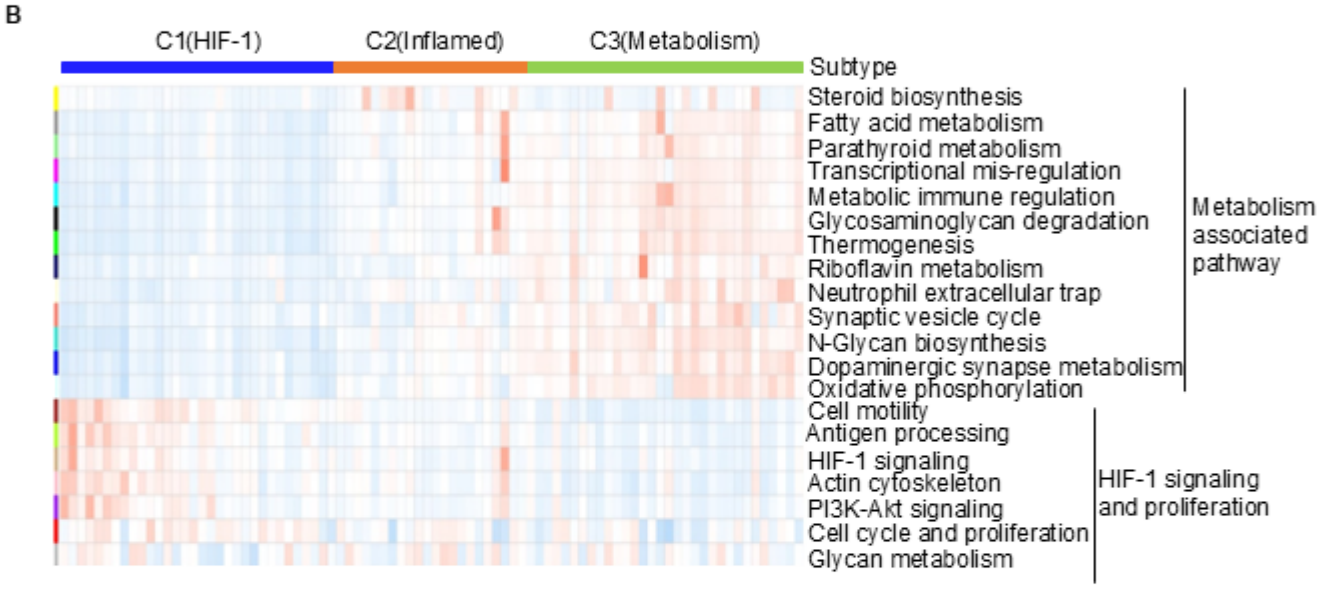
681 (A) Violin plots depicting the expression levels of *ANGPT2*, *PCSK1N*, and *GPX3* gene across
682 three identified subtypes in PPGLs cohort. (B) Immunohistochemistry images showing the
683 protein expression of ANGPT2, PCSK1N, and GPX3 across subtypes C1, C2, and C3. Scale
684 bars represent 50 μm. (C) Survival curves comparing the progression-free survival (PFS) of
685 patients across three subtypes of PPGLs, integrating data from both the PPGL and TCGA
686 cohorts. (D) The forest plot displaying the hazard ratios (HR) with 95% confidence intervals
687 (CI) for PFS across PPGL subtypes using Cox regression analysis. Model 1 was crude model;
688 Model 2 was adjusted for age, gender, and race based on Model 1; Model 3 was further adjusted
689 for tumor location and pathogenic mutation type based on Model 2. Each model is compared
690 to subtype C1 as the reference group. Orange represents subtype C2, and green represents
691 subtype C3. Comparisons were calculated by one-way ANOVA (A). Data are presented as mean
692 ± SD. ns, $p > 0.05$; ****, $p < 0.0001$.

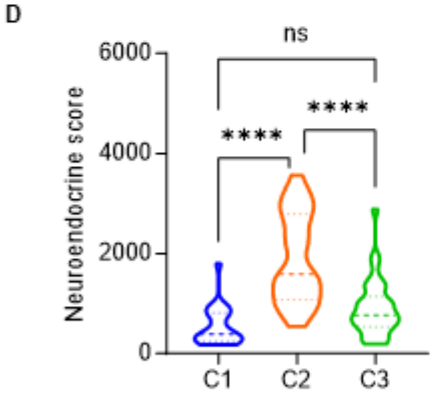
693 **Figure 6. The characteristics of *ANGPT2* knockout**

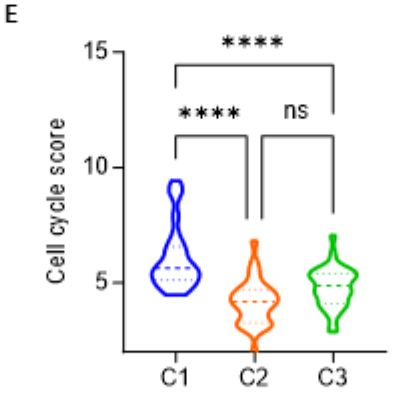
694 (A) The ROC curve illustrates the diagnostic ability to distinguish ANGPT2 expression in
695 PPGLs, specifically differentiating subtype C1 from non-C1 subtypes. The red dot indicates the
696 point with the highest sensitivity (88.9%) and specificity (83.3%). AUC, the area under the
697 curve. (B-D) Violin plots display the distribution of cell cycle scores, EMT scores, and HIF
698 scores in PPGLs between ANGPT2high and ANGPT2low Groups, using a cutoff value of 1.445
699 for ANGPT2 expression as shown in panel A. (E) Immunoblots showing ANGPT2 and Actin
700 expression in ANGPT2 wild-type (ANGPT2WT) and knockout (ANGPT2–/–) rat PPGL cell
701 line PC12. (E) The graph shows a series of tumors harvested from mice xenografted with either
702 ANGPT2 wild-type (WT) or knockout (KO) PPGL cells. (F) The graph shows a series of tumors
703 harvested from BALB/c nude mice xenografted with either ANGPT2WT or ANGPT2–/– PC12
704 tumors over 16 days. (G) The tumor growth volume of ANGPT2WT or ANGPT2–/– PC12
705 tumors in Balb/c nude mice. (H) Heatmap illustrates changes in gene expression related to cell
706 proliferation, tumor metastasis, and hypoxia between ANGPT2WT or ANGPT2–/– PC12 cells.
707 (I-L) The dotplots illustrate EMT scores, tumor metastasis scores, cell proliferation scores, and
708 hypoxia scores in ANGPT2 KO tumors. Comparisons were calculated by t tests (B-D, I-L).
709 Data are presented as mean ± SD. ***, $p < 0.001$; ****, $p < 0.0001$.

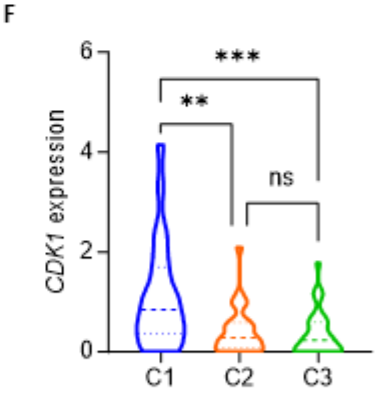

A
PPGLs Cohort (n=87)
HN-PG(n=4)
PC(n=55)
RP-PG(n=24)
NAM(n=5)
BC-PG(n=4)
TCGA (n=179)
HN-PG
ME-PG
PG (n=32)
PC (n=147)
RP-PG
BC-PG
IHC
Clinical data
RNA-seq
Clinical data
WGS
Transcriptional subtype
correlation
Follow-up
Recurrence/metastasis

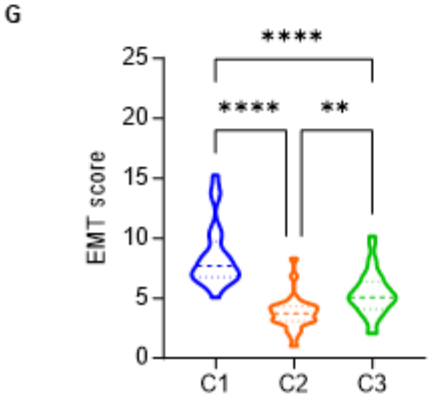

C
C1
C2
C3
ANGPT2
CAV1
PRDM1
BCAT1
FTL
PCSK1N
PLCG2
GPX3
DYNLL2
MYH14
-4
-2
0
2
4

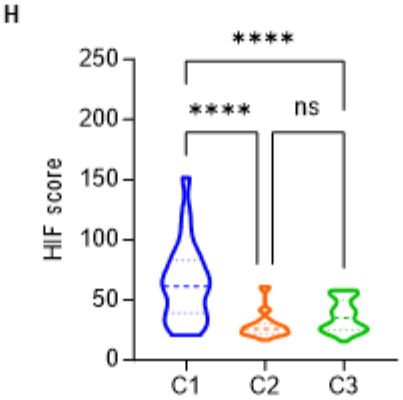

B
C1(HIF-1)
C2(Inflamed)
C3(Metabolism)
Subtype
Steroid biosynthesis
Fatty acid metabolism
Parathyroid metabolism
Transcriptional mis-regulation
Metabolic immune regulation
Glycosaminoglycan degradation
Thermogenesis
Riboflavin metabolism
Neutrophil extracellular trap
Synaptic vesicle cycle
N-Glycan biosynthesis
Dopaminergic synapse metabolism
Oxidative phosphorylation
Metabolism associated pathway
Cell motility
Antigen processing
HIF-1 signaling
Actin cytoskeleton
PI3K-Akt signaling
Cell cycle and proliferation
Glycan metabolism
HIF-1 signaling and proliferation

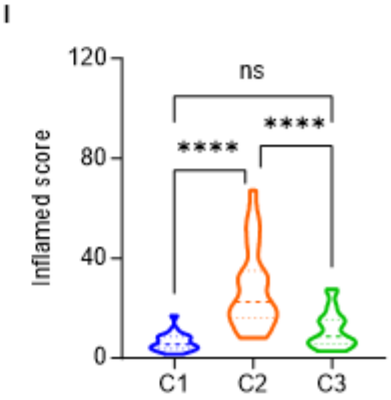

D
Neuroendocrine score
6000
4000
2000
0
ns
****
****
C1
C2
C3

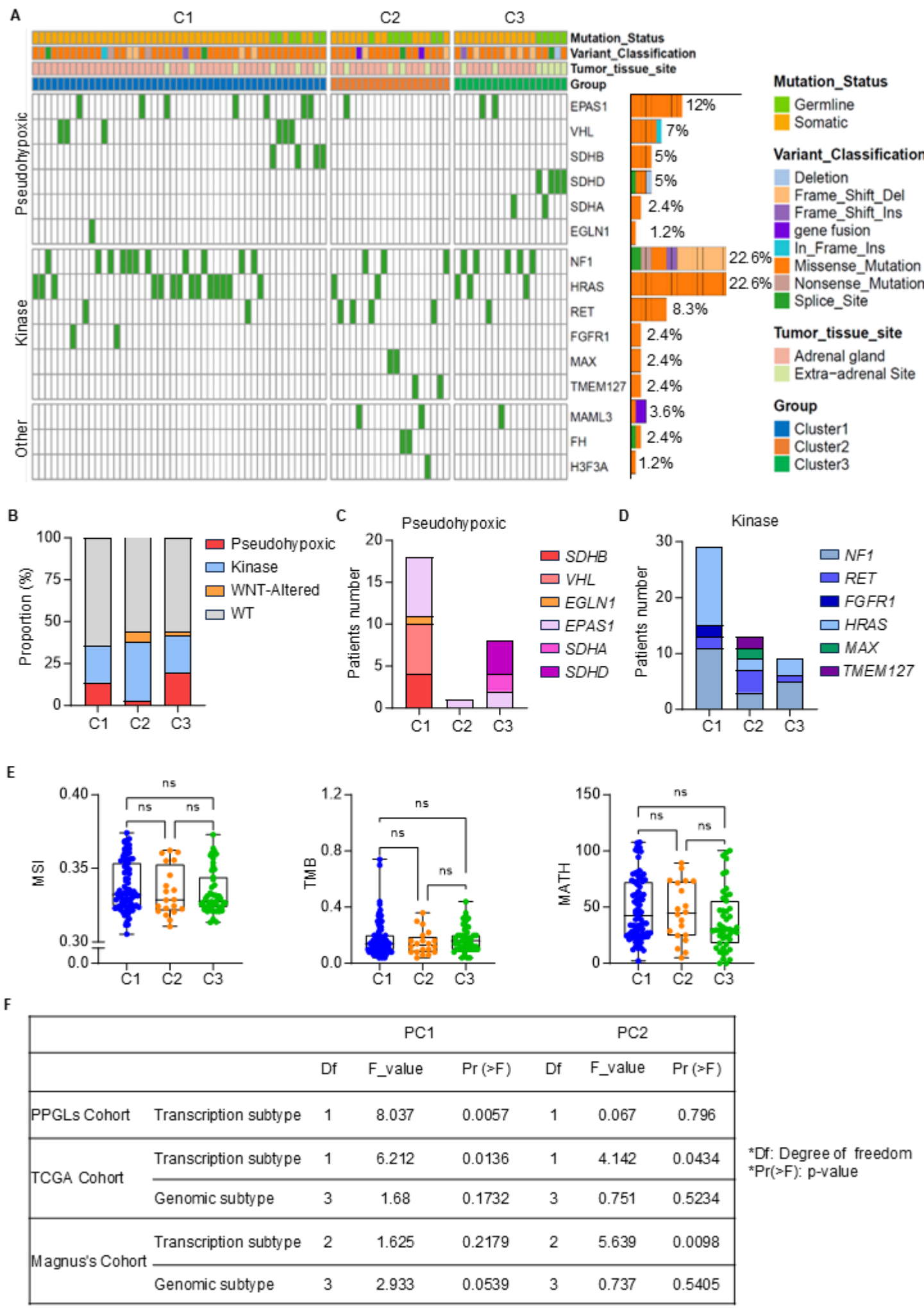

E
Cell cycle score
15
10
5
****
****
ns
C1
C2
C3


F
CDK1 expression
6
4
2
0
***
**
ns
C1
C2
C3


G
EMT score
25
20
15
10
5
0
****
****
**
C1
C2
C3


H
HIF score
250
200
150
100
50
0
****
****
ns
C1
C2
C3

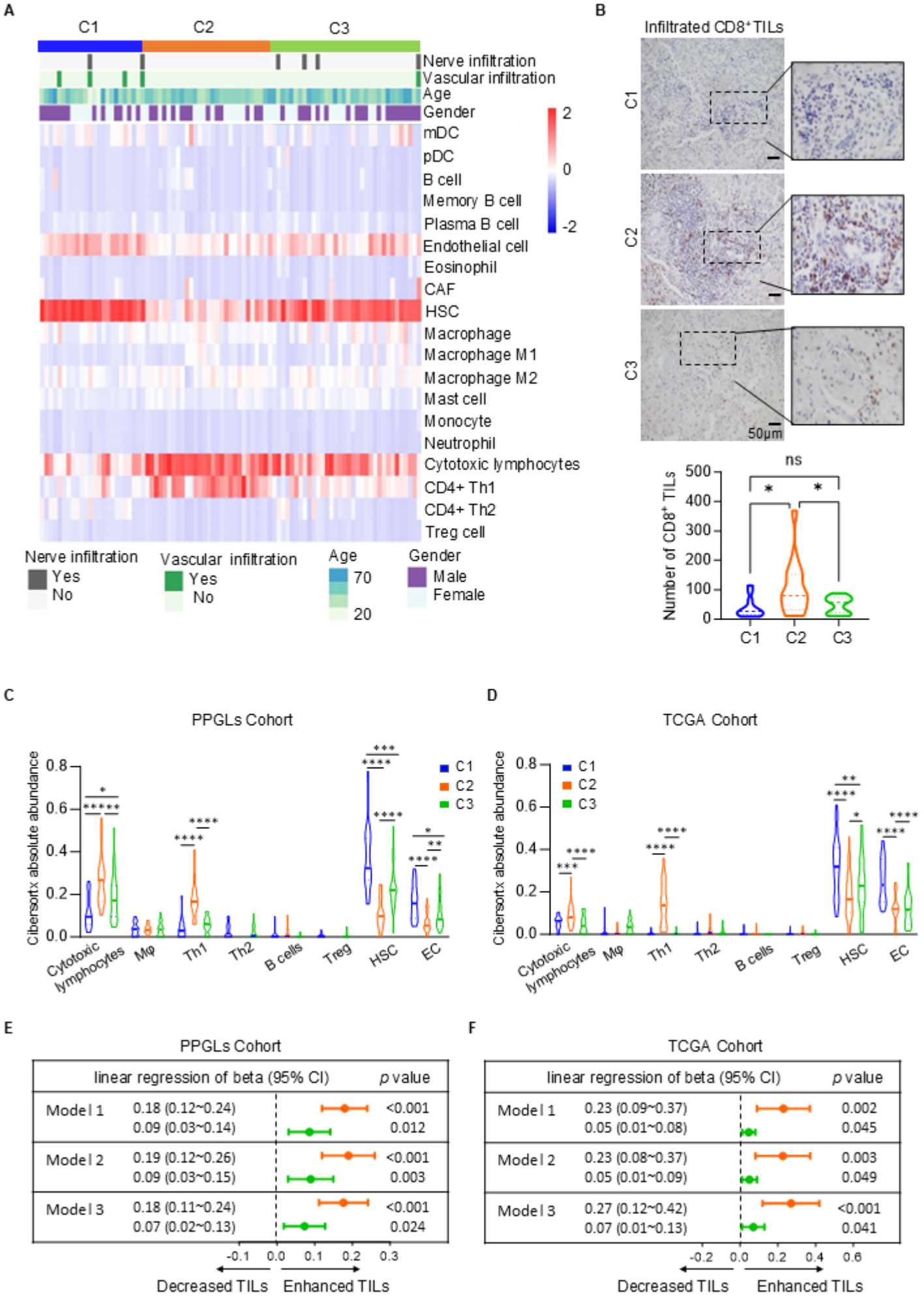

I
Inflamed score
120
80
40
0
ns
****
****
C1
C2
C3

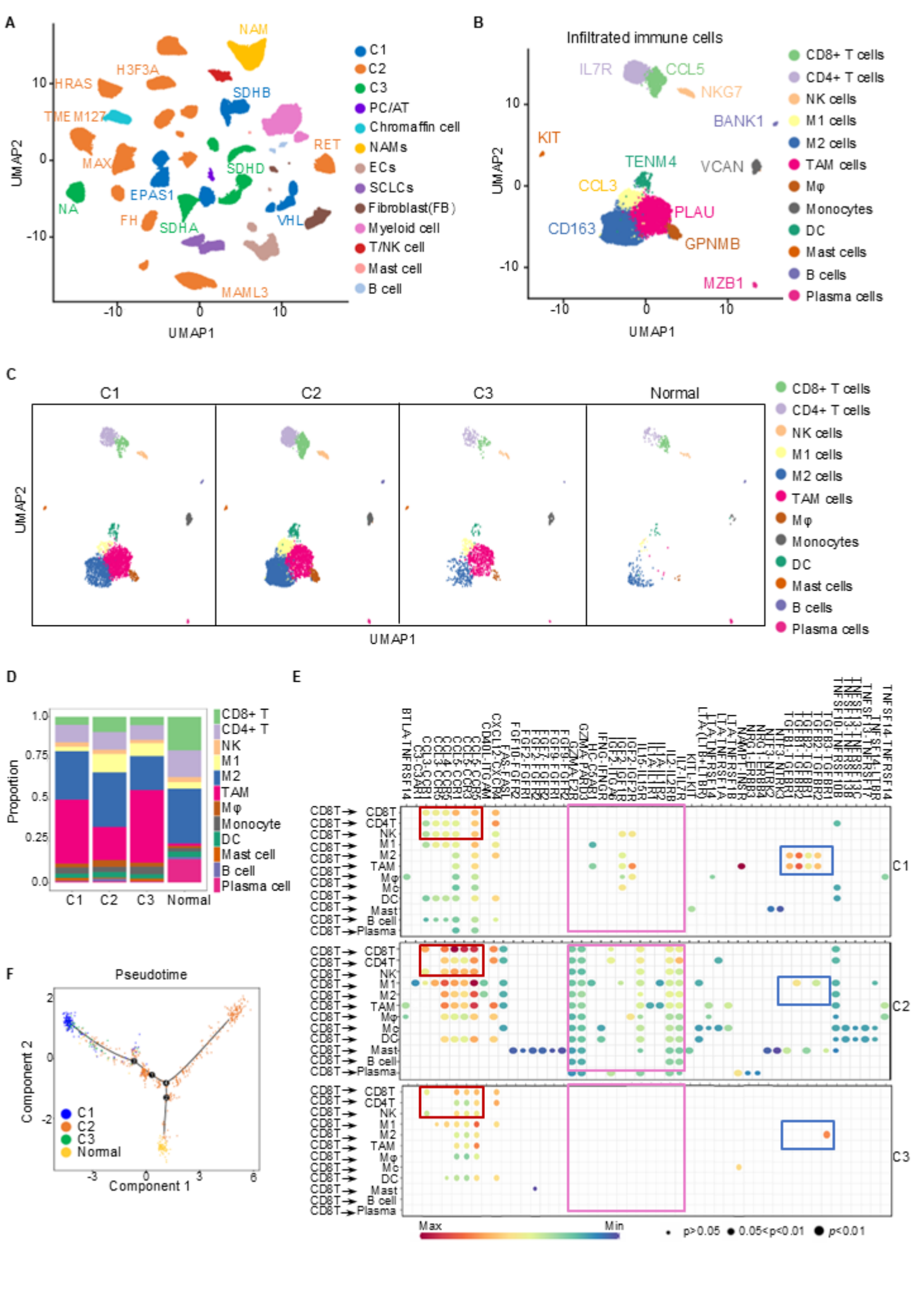


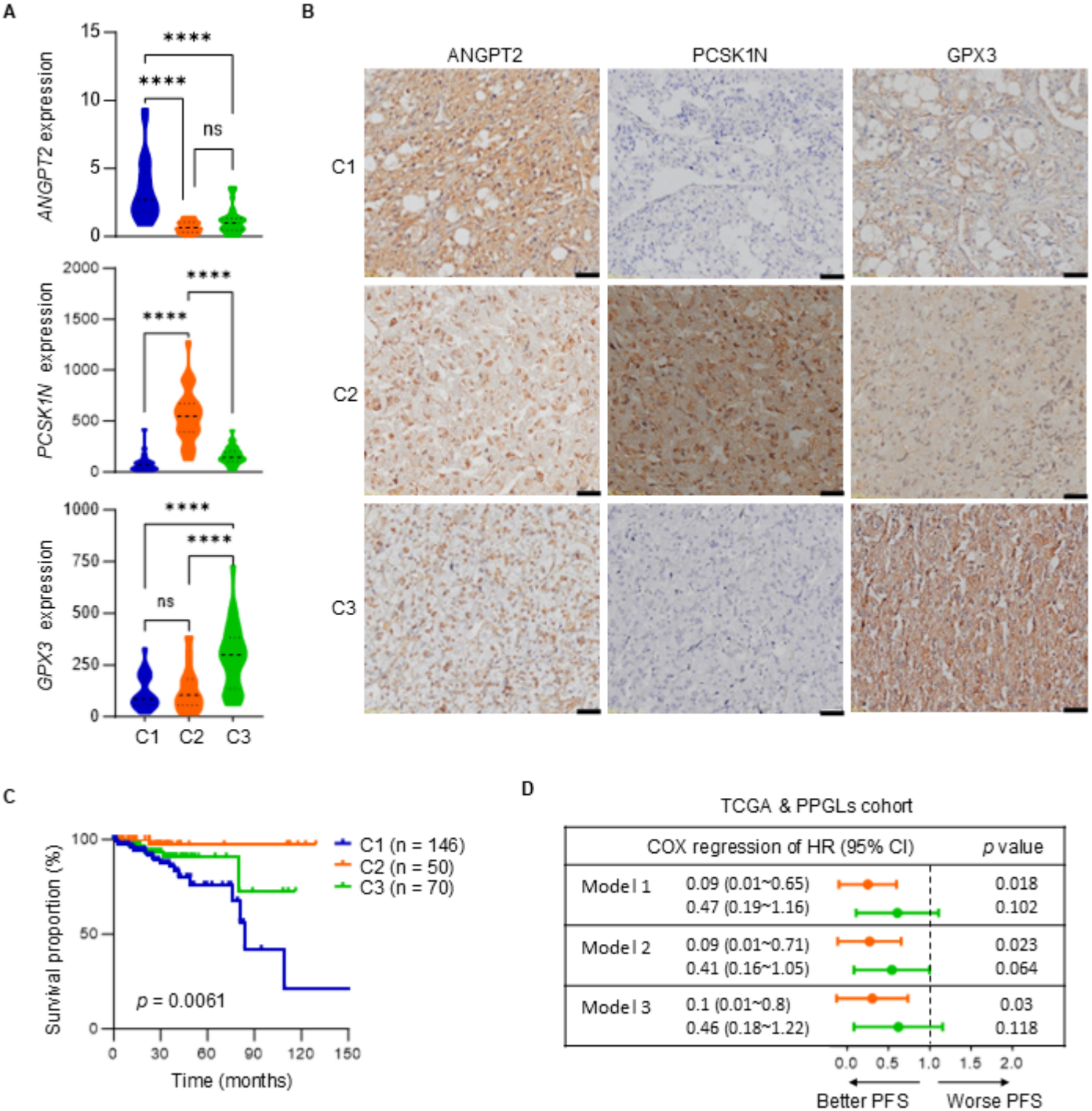
F

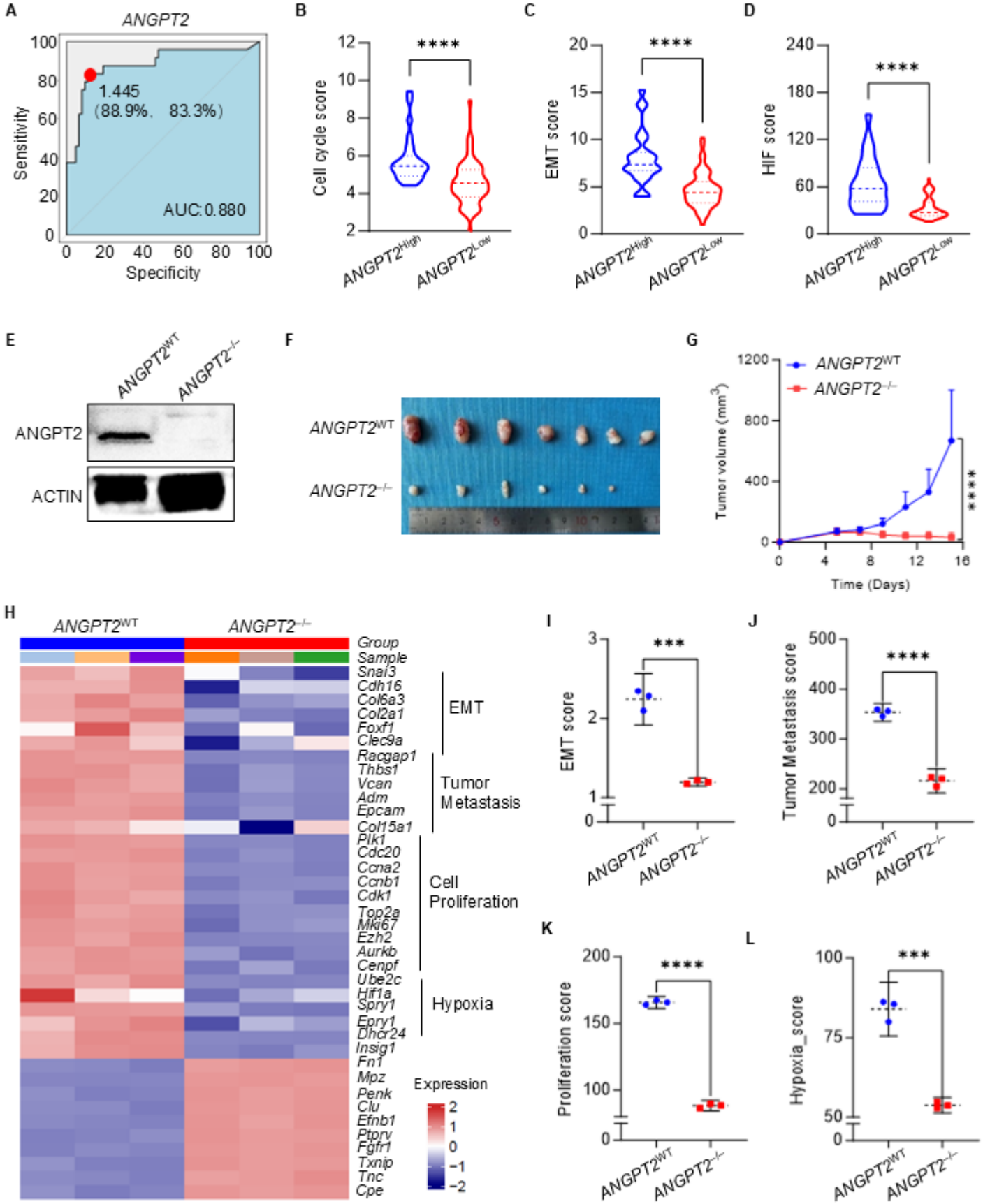
		PC1			PC2		
		Df	F_value	Pr (>F)	Df	F_value	Pr (>F)
PPGLs Cohort	Transcription subtype	1	8.037	0.0057	1	0.067	0.796
TCGA Cohort	Transcription subtype	1	6.212	0.0136	1	4.142	0.0434
	Genomic subtype	3	1.68	0.1732	3	0.751	0.5234
Magnus's Cohort	Transcription subtype	2	1.625	0.2179	2	5.639	0.0098
	Genomic subtype	3	2.933	0.0539	3	0.737	0.5405

*Df: Degree of freedom
*Pr(>F): p-value

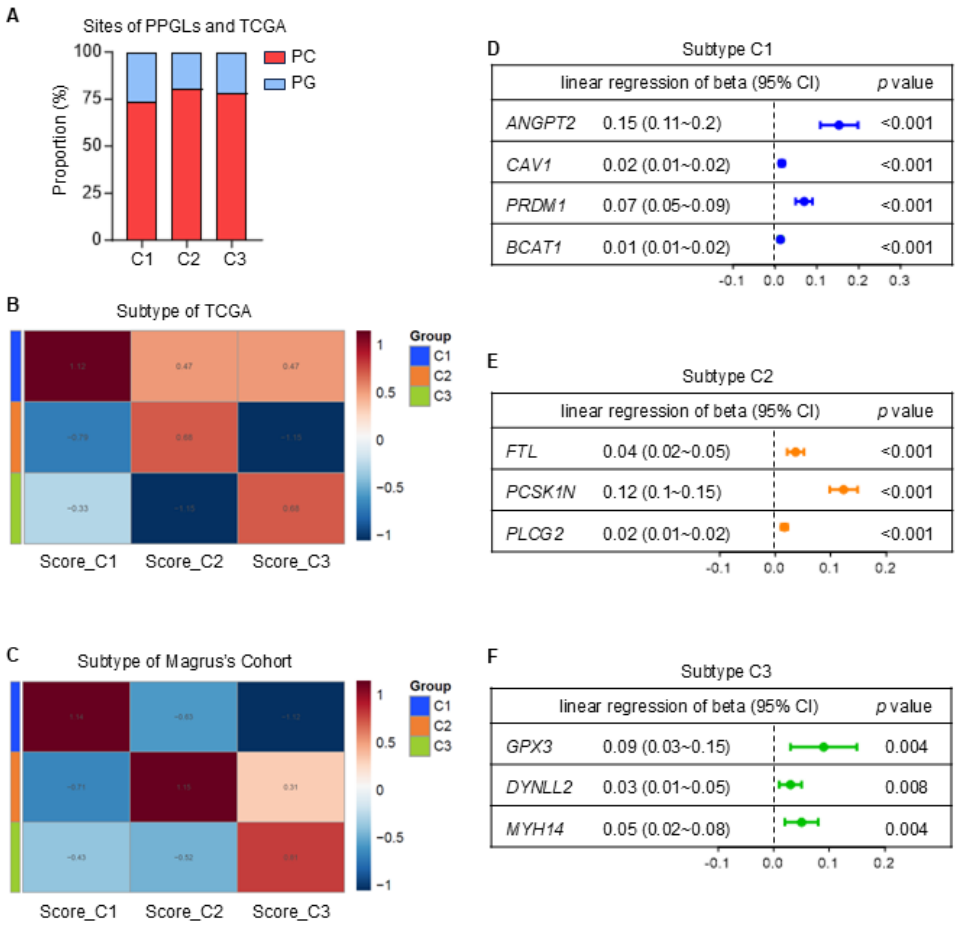

A
C1
C2
C3
Nerve infiltration
Vascular infiltration
Age
Gender
mDC
pDC
B cell
Memory B cell
Plasma B cell
Endothelial cell
Eosinophil
CAF
HSC
Macrophage
Macrophage M1
Macrophage M2
Mast cell
Monocyte
Neutrophil
Cytotoxic lymphocytes
CD4+ Th1
CD4+ Th2
Treg cell
2
0
-2
Nerve infiltration
Yes
No
Vascular infiltration
Yes
No
Age
70
20
Gender
Male
Female
B
Infiltrated CD8+ TILs
C1
C2
C3
50μm
ns
Number of CD8+ TILs
500
400
300
200
100
0
C1
C2
C3
C
PPGLs Cohort
Cibersortx absolute abundance
0.8
0.6
0.4
0.2
0.0
C1
C2
C3
Cytotoxic lymphocytes
Mφ
Th1
Th2
B cells
Treg
HSC
EC
D
TCGA Cohort
Cibersortx absolute abundance
0.8
0.6
0.4
0.2
0.0
C1
C2
C3
Cytotoxic lymphocytes
Mφ
Th1
Th2
B cells
Treg
HSC
EC
E
PPGLs Cohort
linear regression of beta (95% CI)
p value
Model 1
0.18 (0.12~0.24)
<0.001
0.09 (0.03~0.14)
0.012
Model 2
0.19 (0.12~0.26)
<0.001
0.09 (0.03~0.15)
0.003
Model 3
0.18 (0.11~0.24)
<0.001
0.07 (0.02~0.13)
0.024
-0.1
0.0
0.1
0.2
0.3
Decreased TILs
Enhanced TILs
F
TCGA Cohort
linear regression of beta (95% CI)
p value
Model 1
0.23 (0.09~0.37)
0.002
0.05 (0.01~0.08)
0.045
Model 2
0.23 (0.08~0.37)
0.003
0.05 (0.01~0.09)
0.049
Model 3
0.27 (0.12~0.42)
<0.001
0.07 (0.01~0.13)
0.041
-0.2
0.0
0.2
0.4
0.6
Decreased TILs
Enhanced TILs


A
NAM
H3F3A
HRAS
SDHB
TMEM127
RET
MAX
SDHD
EPAS1
NA
FH
SDHA
VHL
MAML3
UMAP2
UMAP1
C1
C2
C3
PC/AT
Chromaffin cell
NAMs
ECs
SCLCs
Fibroblast(FB)
Myeloid cell
T/NK cell
Mast cell
B cell
B
Infiltrated immune cells
IL7R
CCL5
NKG7
BANK1
KIT
TENM4
VCAN
CCL3
PLAU
CD163
GPNMB
MZB1
CD8+ T cells
CD4+ T cells
NK cells
M1 cells
M2 cells
TAM cells
Mφ
Monocytes
DC
Mast cells
B cells
Plasma cells
C
C1
C2
C3
Normal
D
Proportion
CD8+ T
CD4+ T
NK
M1
M2
TAM
Mφ
Monocyte
DC
Mast cell
B cell
Plasma cell
E
Max
Min
p>0.05
0.05<p<0.01
p<0.01
F
Pseudotime
Component 1
Component 2


A
ANGPT2 expression
PCSK1N expression
GPX3 expression
****
ns
C1
C2
C3
B
ANGPT2
PCSK1N
GPX3
C1
C2
C3
C
Survival proportion (%)
Time (months)
C1 (n = 146)
C2 (n = 50)
C3 (n = 70)
p = 0.0061
D
TCGA & PPGLs cohort
COX regression of HR (95% CI)
p value
Model 1
0.09 (0.01~0.65)
0.018
0.47 (0.19~1.16)
0.102
Model 2
0.09 (0.01~0.71)
0.023
0.41 (0.16~1.05)
0.064
Model 3
0.1 (0.01~0.8)
0.03
0.46 (0.18~1.22)
0.118
Better PFS
Worse PFS


A
ANGPT2
1.445
(88.9%, 83.3%)
AUC:0.880
Sensitivity
Specificity
B
Cell cycle score
****
ANGPT2High
ANGPT2Low
C
EMT score
****
D
HIF score
****
E
ANGPT2WT
ANGPT2−/−
ANGPT2
ACTIN
F
ANGPT2WT
ANGPT2−/−
G
ANGPT2WT
ANGPT2−/−
Tumor volume (mm3)
Time (Days)
****
H
ANGPT2WT
ANGPT2−/−
Group
Sample
Snai3
Cdh16
Col6a3
Col2a1
Foxf1
Clec9a
EMT
Racgap1
Thbs1
Vcan
Adm
Epcam
Col15a1
Tumor Metastasis
Plk1
Cdc20
Ccna2
Ccnb1
Cdk1
Top2a
Mki67
Ezh2
Aurkb
Cenpf
Cell Proliferation
Ube2c
Hif1a
Spry1
Epry1
Dhcr24
Insig1
Hypoxia
Fn1
Mpz
Penk
Clu
Efnb1
Ptprv
Fgfr1
Txnip
Tnc
Cpe
Expression
I
EMT score
***
J
Tumor Metastasis score
****
K
Proliferation score
****
L
Hypoxia_score
***

Supplementary Fig. 1



Supplementary Fig. 1. Subtype assignment and marker performance across cohorts.
(A) Distribution of tumor site categories—pheochromocytoma (PC) vs paraganglioma (PG)—by transcriptional subtype (C1–C3) in the aggregated dataset (TCGA + Magnus). (B) TCGA and (C) Magnus cohorts: heatmaps of WGCNA-derived gene-set scores (Score_C1, Score_C2, Score_C3) per sample; values are z-scores, with warmer colors indicating higher concordance with the corresponding subtype. (D–F) The forest plots (D) show the effect sizes (β values) of marker genes as continuous variables for distinguishing subtype C1 from non-C1, (E) for distinguishing subtype C2 from non-C2, and (F) for distinguishing subtype C3 from non-C3, all based on linear regression.