HGG Advances

REPORT

Single-cell analysis identified POSTN^{+} cells associated with the aggressive phenotype and risk of esophageal squamous cell carcinoma

Yuqian Tan,[1] Lina Song,[1] Jialing Ma,[1] Miaoxin Pan,[1] Siyuan Niu,[1] Xinying Yue,[1] Yueping Li,[1] Linglong Gu,[1] Shasha Liu,[1] and Jiang Chang[1,2,*]

Summary

Tumors are intricate and heterogeneous systems characterized by mosaic cancer cell populations with diverse expression profiles. Leveraging single-cell technologies, we employed the Scissor algorithm to delineate an epithelial subpopulation associated with the aggressive phenotype in esophageal squamous cell carcinoma (ESCC). This identified subpopulation exhibited elevated expression of genes involved in critical pathways, such as epithelial-mesenchymal transition and PI3K-Akt. Key signature genes within this subpopulation, namely *CAV1*, *COL3A1*, *COL6A1*, *POSTN*, and *TAGLN*, demonstrated significant upregulation concomitant with both tumorigenesis and tumor progression across independent single-cell datasets. Furthermore, we selected 1,450 expression quantitative trait loci of the top 62 signature genes of this cell subpopulation to investigate their potential in predicting ESCC risk. The results showed that the *POSTN* loci were predominantly associated with ESCC susceptibility. Through functional annotation and replication analyses, we identified that the rs1028728 in the *POSTN* promoter was significantly associated with increased ESCC risk in 7,049 ESCC cases and 8,063 controls (odds ratio = 1.29, 95% confidence interval: 1.18–1.42, $p = 4.03 \times 10^{-8}$). Subsequent biochemical experiments showed that the rs1028728[T] allele enhanced *POSTN* expression by affecting the binding of *PRRX1* in the *POSTN* promoter. In summary, our meticulous single-cell analysis delineates an invasive epithelial subpopulation in ESCC, with *POSTN* emerging as an important marker for the aggressive phenotype. These findings offer more insights into potential strategies for the prevention and intervention of ESCC, enriching our understanding of this complex cancer landscape.

Esophageal squamous cell carcinoma (ESCC [MIM: 133239]) poses a significant health challenge, characterized as a highly lethal form of cancer with a mere 20–30% 5-year survival rate. It primarily manifests in the middle or upper section of the esophagus, constituting over 90% of all esophageal cancer cases in China.[1,2] Typically originating from the lining of esophageal squamous epithelium, ESCC development is triggered by exposure to carcinogens or mechanical damage, instigating abnormal proliferation of epithelial cells and ultimately progression to invasive cancer.[3] Understanding the molecular mechanisms underpinning epithelial cell progression is crucial for identifying molecular markers and intervention targets in ESCC.

Traditionally, the search for cancer biomarkers has predominantly relied on bulk RNA sequencing (RNA-seq).[4] However, the limitation of this approach in capturing signals from rare cell populations or specific cell types driving tumorigenesis and progression has prompted the adoption of single-cell RNA-seq (scRNA-seq),[4] which is widely applied in investigating tumor pathogenesis, including ESCC.[2,5] It offers exceptional sensitivity and specificity[6] and the ability to discern diverse cell types, states, and cell subpopulations within heterogeneous tissue.[7–9] Multiple single-cell algorithms have been developed to identify precancerous cells[10] and those associated with specific cancer phenotypes, aiming to explore yet-undetermined genetic determinants and identify cell types and processes influenced by genetic and gene variations. Especially, an algorithm known as Scissor leverages bulk data and phenotype information to steer single-cell analysis and identifies subsets of cells that significantly contribute to phenotypic disparities and disease relevance.[11] Additionally, a previous study defined a confused cell identity (CCI) in ESCC at single-cell resolution, establishing its potential as a malignant cellular characteristic and an independent prognostic marker.[12]

The genetic etiology of ESCC remains largely unknown. Previous studies have extensively investigated the genomic characteristics and alterations in ESCC and normal esophageal tissues and across various stages of normal esophagus, early and late precancerous lesions, and esophageal cancer.[13,14] The major genomic changes observed in cancer genomes include more frequent chromosomal instability and copy-number alterations (CNAs) in ESCC compared to normal esophageal tissues, with the burden of somatic mutations and CNAs increasing with the progression of the disease. Several genome-wide association studies (GWASs) of germline variants in ESCC have been published. GWASs have discovered over two dozen risk

[1]Department of Health Toxicology, Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[2]Lead contact
*Correspondence: changjiang815@hust.edu.cn
https://doi.org/10.1016/j.xhgg.2024.100278.

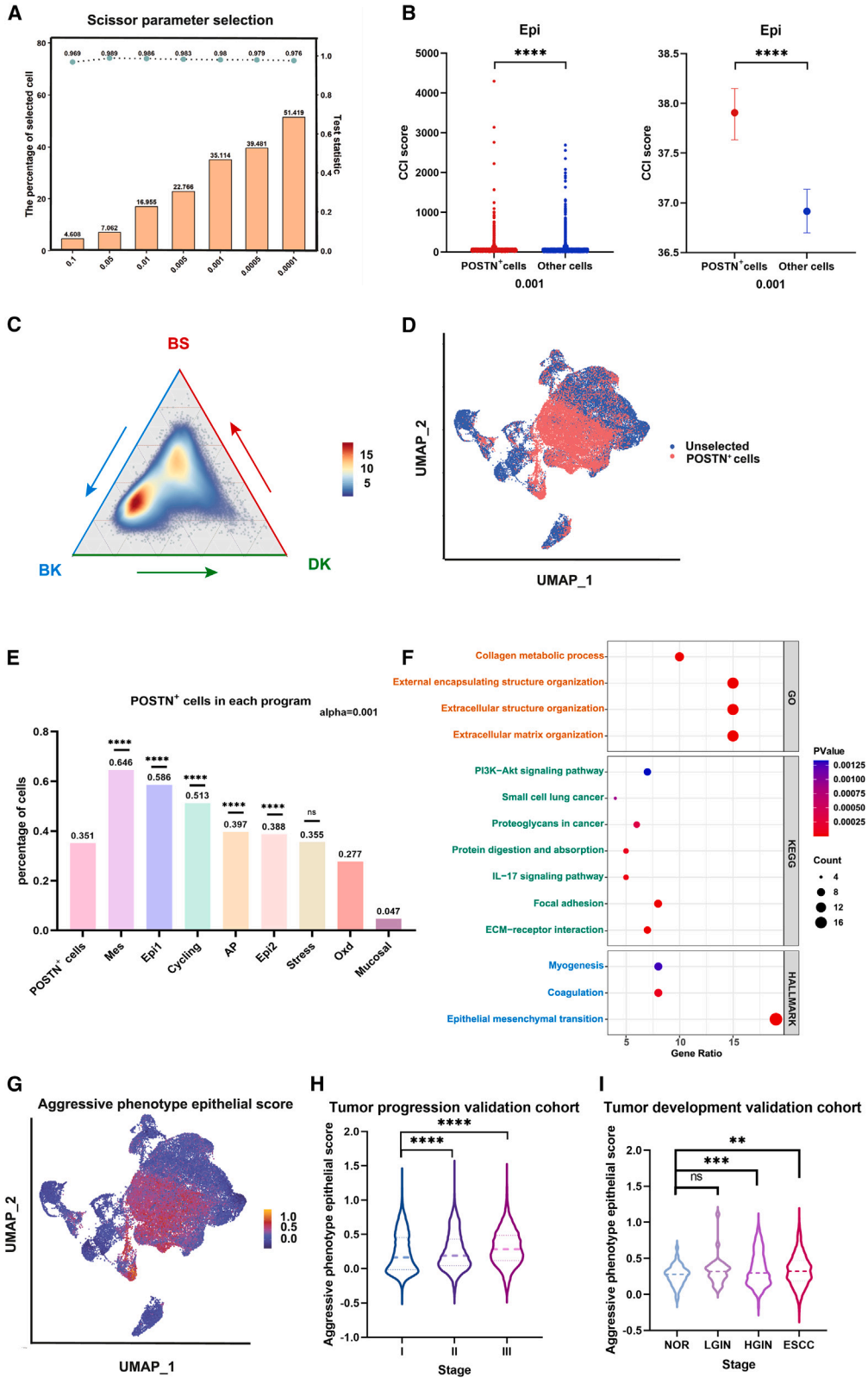

Figure 1. Identification of cell subpopulation associated with aggressive phenotype in ESCC

(A) The bar chart and line graph, respectively, present the proportions of POSTN^{+} cells selected by Scissor, as well as the values of significance test statistics, under different alpha parameters from 0.0001 to 0.1.

(B) CCI scores of all epithelial cells at 0.001 alpha parameters. The figure on the bottom compares the CCI scores of the POSTN^{+} cells with the rest of the cells by the Mann-Whitney test. ****$p < 0.0001$.

(legend continued on next page)

loci for ESCC.[15] Integrating scRNA-seq data with GWAS data can unveil pivotal genes vital for disease onset and progression.

Understanding the gene expression pattern and expression programs in invasive epithelial cells is crucial for unraveling the mechanisms of invasive carcinogenesis. In this study, we identified an aggressive tumor phenotype-related cell subpopulation and validated its heightened malignancy across multiple datasets and methods. Through case-control studies and multi-stage ESCC scRNA-seq data, we identified the signature gene *POSTN* (MIM: 608777) as a crucial regulator of ESCC susceptibility and its invasive phenotype. Notably, the rs1028728 variant within the *POSTN* promoter region emerged as a potential mechanism driving heightened ESCC invasion and susceptibility. These findings constitute a valuable resource for a deeper understanding of the intricate progression of epithelial cell invasion in ESCC.

To investigate the cellular subpopulations and biomarkers associated with the aggressive phenotype of ESCC, we integrated a bulk expression matrix and single-cell expression matrix using Scissor. The flowchart of this study can be found in Figure S1. The RNA-seq data for 174 ESCC tumors and 105 corresponding nontumor tissues were sourced from our prior study[16] and The Cancer Genome Atlas (TCGA) database. To integrate these two datasets into a unified expression matrix, we utilized the R package "sva" to mitigate batch effects in the combined samples (Figure S2). scRNA-seq data of 44,122 epithelial cells from 60 ESCC tumors were obtained from the GEO database[2] (Figure S3A). To stratify cell subpopulations associated with aggressive tumor phenotypes, we incorporated single-cell expression matrices into the Scissor algorithm. Concurrently, bulk RNA expression data from ESCC and matched normal samples were utilized to establish benchmarks for the aggressive tumor and normal phenotypes.

Using Scissor with various alpha parameter choices, we identified varying proportions of aggressive tumor phenotype-related epithelial cells (defined as $POSTN^+$ cells) , all exhibiting consistent reliability and significance (Figure 1A). We then introduced a novel feature, the CCI score, as reported in a previous study,[12] to validate the malignancy of the identified cell subpopulation of Scissor. We observed that the CCI scores of $POSTN^+$ cells were significantly higher than those of other cells under alpha = 0.001 (Figures 1B and S3B), confirming that the selected invasive subpopulation exhibited a higher degree of malignancy. A previous scRNA-seq analysis identified three principal cell populations constituting the normal esophageal squamous epithelium: postmitotic differentiated keratinocytes (DKs), transient proliferating basal keratinocytes (BKs), and basal stem cells (BSs).[17–19] We showed a ternary plot positioning BSs, BKs, and DKs at the vertices, allowing for the representation of $POSTN^+$ cell distribution in relation to the established populations of normal esophageal squamous epithelial cells. Consistent with previous results, this identified aggressive subpopulation almost fell in the center of the triangle, suggesting they may represent tumor cells acquiring a new identity with partial similarities to these three normal cell types[12] (Figure 1C). As a result, we identified a cell subpopulation associated with aggressive tumor phenotype, denoted as the $POSTN^+$ subpopulation, comprising 15,493 epithelial cells. This subpopulation passed a highly reliable significance test with a test statistic of 0.980 and a p value of less than 0.001, accounting for 35.11% of the total epithelial cells (Figures 1A, 1D, and 1E). Prior scRNA-seq analysis of 60 patients with ESCC identified eight expression programs, each associated with distinct cellular functions and states. Cells that express at least 70% of a program's gene set are deemed "activated" within that program.[2] To investigate the primary roles of $POSTN^+$ cells, we consequently computed the percentage of $POSTN^+$ cells within these program cells. As seen in Figure 1E, the $POSTN^+$ subpopulation predominantly exhibited features of the epithelial differentiation 1 (Epi1), the cell-cycle (Cycling) program features, and the mesenchymal stromal cell-like qualities (Mes). These findings deepen our understanding of the intricate landscape of tumor cells in ESCC.

We then sought to determine whether signature genes within the $POSTN^+$ subpopulation could serve as molecular markers for the occurrence and progression of ESCC. Through differential gene expression analysis between the $POSTN^+$ subpopulation and other cells across eight expression programs, we identified 62 signature gene candidates and aggregated them into a gene set for computing

(C) The ternary diagram shows the similarity and variation in scRNA-seq data of $POSTN^+$ cells, with the basal stem cells (BSs), transient proliferating basal keratinocytes (BKs), and postmitotic differentiated keratinocytes (DK) cells. The color represents the density levels of distribution.
(D) The uniform manifold approximation and projection (UMAP) visualization of the Scissor-selected cells. The red dots are cells associated with the aggressive phenotypes (alpha = 0.001).
(E) The percentage of $POSTN^+$ cells in the eight different programs. The percentages of invasive phenotype cells in each program and within epithelial cells were compared using a chi-squared test, $^{****}p < 0.0001$.
(F) GO, KEGG, and HALLMARK enrichment analyses were used to elucidate the potential functions of 62 signature genes in $POSTN^+$ cells.
(G) The UMAP plot maps the aggressive phenotype epithelial score for each cell.
(H) Violin plots of the aggressive phenotype epithelial scores in different stages are plotted. $^{****}p < 0.0001$ was calculated by the Mann-Whitney test.
(I) Violin plots of aggressive phenotypic epithelial scores for single-cell validation data of ESCC progression. The aggressive phenotype epithelial scores of cells in the HGIN and ESCC stages are significantly higher than those in the normal (NOR) stage. $^{**}p < 0.01$ and $^{***}p < 0.001$ were calculated by independent t test.

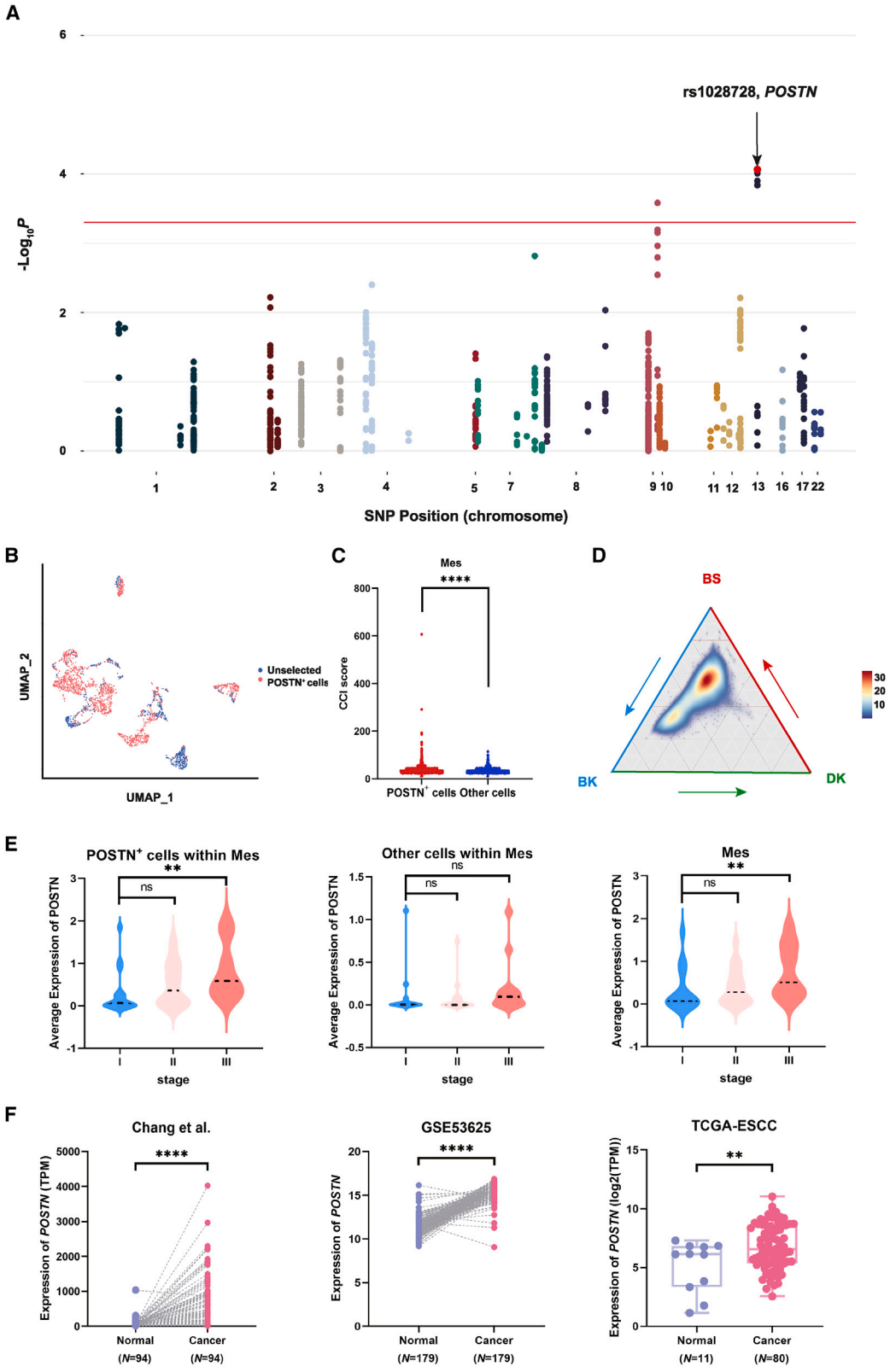

Figure 2. *POSTN* plays a vital role as a regulatory factor in ESCC susceptibility and its aggressive phenotype

(A) Genome-wide results for the relationship between 62 signature gene candidate expression quantitative trait locus (eQTL) SNPs and ESCC. Manhattan plot shows that *POSTN* was the susceptibility gene with the highest association with ESCC. $p < 5 \times 10^{-4}$ (false discovery rate [FDR] < 0.05) was considered statistically significant, and rs1028728 was annotated.

(B) The UMAP visualization of the Mes cells.

(C) CCI scores of all cells of Mes at 0.001 alpha parameters tested by the Mann-Whitney test. ****$p < 0.0001$.

(legend continued on next page)

the epithelial score associated with an aggressive phenotype for validation (Table S1). Notably, these genes were enriched in crucial cancer-associated pathways including epithelial-mesenchymal transition (19 genes), PI3K-Akt (7 genes), and proteoglycans in cancer (6 genes) (Figure 1F; Table S2). The aggressive phenotype epithelial score can characterize the features of POSTN$^+$ cells, with its distribution consistent with the clustering distribution of POSTN$^+$ cells, exhibiting significant differences between POSTN$^+$ cells and other cells (Figures 1G and S3C). The aggressive phenotype epithelial scores also exhibited a gradual increase along with the tumor stage (Figures 1H and S4). To validate transcriptional changes in POSTN$^+$ cells at various disease stages, we obtained a single-cell expression matrix from a validation cohort.[5] Our analysis revealed that the expression levels of the part of these 62 signature genes exhibited significant elevation in tumor cells or at late precancerous stage (high-grade intraepithelial neoplasia [HGIN]), with no corresponding increase in normal or early precancerous stage (low-grade intraepithelial neoplasia) (Figure S5). Specifically, the aggressive phenotype epithelial cell score exhibited a notable increase with tumor advancement in both ESCC and HGIN stages (Figure 1I). Moreover, the expression of these genes demonstrated a discernible upregulation concomitant with tumor progression, especially for *CAV1* (MIM: 601047), *POSTN*, *COL3A1* (MIM: 120180), *COL6A3* (MIM: 120250), and *TAGLN* (MIM: 600818) (Figures S4 and S5). Several identified signature gene candidates, such as *TNC*,[20] *COL3A1*,[21] and *COL6A3*,[22] have been widely implicated in cancer invasion. Caveolin-1 (*CAV1*), recognized as an oncogenic protein, is linked to tumor progression, invasive behavior, metastasis, and treatment resistance, including in ESCC.[23] We found that most of the signature genes associated with the aggressive phenotype were those traditionally thought to characterize fibroblasts. These genes exhibited high expression levels in invasive subpopulations of epithelial cells and either exclusively expressed or had elevated expression in HGIN and ESCC during the invasive progression of ESCC.

Genes identified by scRNA-seq data as pivotal in disease onset and progression may also contribute to disease susceptibility.[24] Consequently, we assessed whether these identified 62 signature genes in the POSTN$^+$ subpopulation could confer risk of ESCC. We utilized two GWAS datasets in Chinese populations, consisting of 3,929 individuals with ESCC and 4,144 controls,[25,26] with basic demographic information presented in Table S3. A total of 1,450 expression quantitative trait loci of the 62 signature genes were genotyped, revealing that the *POSTN* loci stood out as the most significant signal with a false discovery rate <0.05 (Figure 2A). *POSTN* was predominantly expressed in the Mes program, constituting 64.6% of the POSTN$^+$ subpopulation, significantly higher than the proportion of the malignant invasive subpopulation in all epithelial cells ($p < 0.0001$) (Figure 1E). The Mes program comprised a total of 2,743 epithelial cells featuring genes like *VIM* (MIM: 193060) and *SPARC* (MIM: 182120) and displayed activation of the epithelial-mesenchymal transition and angiogenic pathways.[2] Clustering of the POSTN$^+$ subpopulation with cells in the Mes program is depicted in Figure 2B, and the difference in cell identity between the CCI scores of the POSTN$^+$ subpopulation and those of the other cells confounded within the Mes program is highlighted in Figure 2C. Moreover, the POSTN$^+$ subpopulations in Mes program cells are located in the middle of the ternary phase diagram (Figure 2D), indicating a potential malignant profile for these cells.

POSTN also emerged as a potential key gene in the progression of ESCC, with significantly increased expression levels across disease stages, especially in the POSTN$^+$ subpopulation (Figure 2E). The expression of *POSTN* was also significantly upregulated in ESCC tumors in three independent bulk RNA-seq datasets compared to normal samples (Figure 2F). The dynamic interplay between the epithelium and mesenchymal stroma is crucial in triggering an invasive tumor cell phenotype, activating genes linked to cell proliferation, dedifferentiation, migration, and invasion.[27,28] *POSTN*, functioning as a nonstructural extracellular matrix protein, facilitates cell-matrix interactions.[29] During tumor progression, ECM proteins within the tumor stroma can establish a microenvironment conducive to growth and dissemination.[30] Previous research has identified *POSTN* as a crucial mediator of ESCC tumor invasion in invading EPC2-hTERT-EGFR-p53R175H cells.[31] Functional validation in TE11 and HCE4 ESCC cell lines confirmed *POSTN*'s role in promoting tumor growth and invasion in ESCC.[32]

Due to the commonality in the development process of SCC, there are also similarities in gene expression and mutation profiles. We further evaluated the expression of *POSTN* in head and neck squamous carcinoma (HNSC [MIM: 275355]), lung squamous carcinoma (LUSC [MIM: 211980]), and cervical SCC (MIM: 603956) in TCGA database. The results showed that the *POSTN* exhibited elevated expression levels in HNSC and LUSC, with expression levels increasing as the cancer progressed in HNSC (Figures S6A–S6C), suggesting that *POSTN* may be an invasive characteristic for SCCs. SCCs' single-cell data in HNSC, LUSC, and oropharyngeal SCC further reveal that, aside

(D) The ternary diagram shows the similarity and variation in scRNA-seq data of POSTN$^+$ cells within Mes.
(E) Violin plots of mean *POSTN* expression in POSTN$^+$ cells within Mes, other cells within Mes, and Mes are plotted for different stages. **$p < 0.01$ compared to controls by Mann-Whitney test.
(F) *POSTN* is significantly overexpressed in tumor tissues compared to normal tissues from multiple independent databases, including Chang et al.[16], GEO: GEO53625, and TCGA ESCC samples. ****$p < 0.0001$ was calculated by a Wilcoxon test in Chang et al. ESCC tissues and GEO: GEO53625. **$p < 0.01$ was calculated by Student's t test in TCGA.

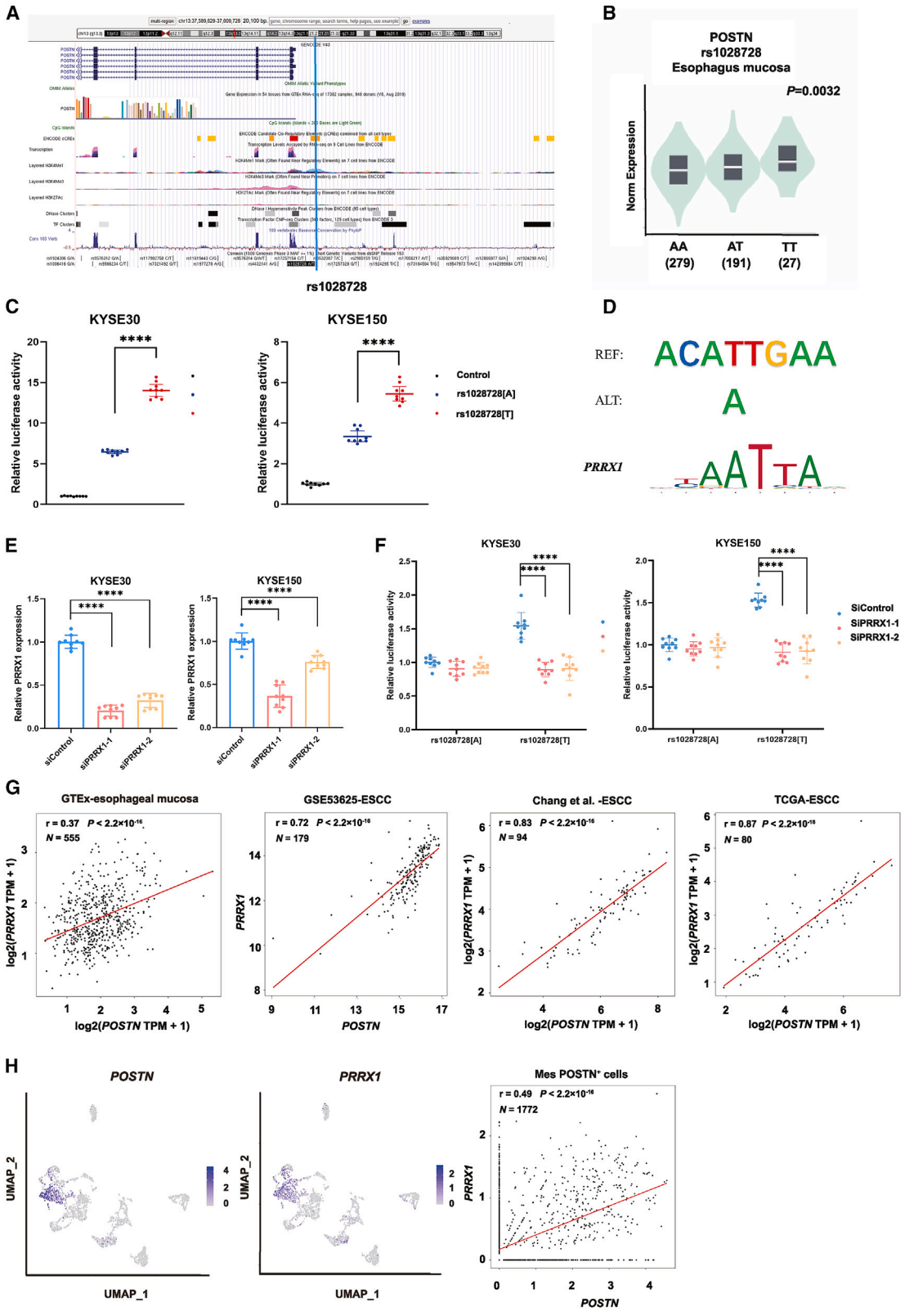

Figure 3. The rs1028728 variant exhibited an allele-specific effect on *POSTN* expression through binding with PRRX1

(A) Integrative analysis of the potential function of rs1028728 by querying UCSC. The SNP rs1028728 was in the *POSTN* promoter region.

(B) The violin plots of expression correlation between rs1028728 and *POSTN* through the GTEx v.8 eQTL Calculator.

(C) Relative reporter gene activity of plasmids containing the rs1028728[A] or -[T] allele in KYSE30 and KYSE150 cells.

(D) The rs1028728[T] allele resides within a *PRRX1* binding motif.

(legend continued on next page)

Table 1. The association between rs1028728 and the susceptibility to esophageal squamous carcinoma in Chinese populations

Stage	rs1028728	Cases (%)	Controls (%)	OR (95% CI)[a]	p
Discovery	AA	3,376 (85.9)	3,678 (89.0)	1.00 (Ref)	N/A
Discovery	AT	537 (13.7)	443 (10.7)	1.32 (1.15–1.51)	6.41×10^{-5}
Discovery	TT	16 (0.4)	14 (0.3)	1.23 (0.60–2.55)	0.5795
Discovery	dominant	N/A	N/A	1.31 (1.15–1.50)	5.79×10^{-5}
Discovery	recessive	N/A	N/A	1.19 (0.58–2.47)	0.6407
Discovery	additive	N/A	N/A	1.29 (1.14–1.47)	8.64×10^{-5}
Replication	AA	2,667 (85.5)	3,461 (88.3)	1.00 (Ref)	N/A
Replication	AT	432 (13.8)	446 (11.4)	1.24 (1.07–1.44)	4.25×10^{-3}
Replication	TT	21 (0.7)	12 (0.3)	1.95 (0.95–4.22)	0.0756
Replication	dominant	N/A	N/A	1.26 (1.09–1.46)	1.78×10^{-3}
Replication	recessive	N/A	N/A	1.91 (0.93–4.11)	0.0865
Replication	additive	N/A	N/A	1.26 (1.10–1.45)	9.88×10^{-4}
Combined samples	AA	6,043 (85.7)	7,148 (88.7)	1.00 (Refref)	N/A
Combined samples	AT	969 (13.8)	889 (11.0)	1.30 (1.16–1.43)	2.24×10^{-7}
Combined samples	TT	37 (0.5)	26 (0.3)	1.65 (1.00–2.77)	0.0513
Combined samples	dominant	N/A	N/A	1.31 (1.19–1.44)	6.14×10^{-8}
Combined samples	recessive	N/A	N/A	1.60 (0.97–2.68)	0.0681
Combined samples	additive	N/A	N/A	1.29 (1.18–1.42)	4.03×10^{-8}

OR, odds ratio; CI, confidence interval; Ref, reference; N/A, not applicable.
[a]Calculated by unconditional logistic regression model after adjusting for gender and age.

from fibroblasts, *POSTN* is expressed in tumor cells as well (Figures S6D–S6F). This aligns with Chao Liu et al.'s investigation into the onset and progression of human cervical SCC through single-cell sequencing.[33] Their study revealed heightened expression of *POSTN* within the Epi1 epithelial cell subset.[33] In particular, Luhai Wang's team in 2022 found that *POSTN* plays a vital role in ovarian tumor metastasis by enriching M2 macrophages and cancer associated fibroblasts (CAFs) through an autocrine effect in cancer cells.[34] These findings suggest that the *POSTN* expressed in epithelial cells may be a crucial molecular marker for ESCC tumorigenesis and progression.

We further conducted a fine-mapping analysis to identify functional variants influencing *POSTN* expression within this locus. Through bioinformatics annotation, we identified the rs1028728 in the *POSTN* promoter, exhibiting significant enrichment of H3K4me1 modifications and displaying the most promising potential to affect *POSTN* expression (Figure 3A). Subsequently, a significant correlation was observed between this variant and *POSTN* expression ($p < 0.05$) in 497 normal esophageal mucosa tissues from the GTEx cohort (Figure 3B). In the discovery stage, individuals carrying the risk allele of rs1028728 exhibited a 29% increased risk of ESCC (95% confidence interval [CI] 1.14–1.47, $p = 8.64 \times 10^{-5}$), and the association was successfully replicated in 3,120 ESCC cases and 3,919 controls[35] (odds ratio [OR] = 1.26, 95% CI 1.10–1.45, $p = 9.88 \times 10^{-4}$) (Table 1). In the combined samples, the association between rs1028728 and ESCC risk reached the genome-wide association significance level (OR = 1.29, 95% CI = 1.18–1.42, $p = 4.03 \times 10^{-8}$) (Table 1).

To test whether the rs1028728 variant influences the promoter activity of *POSTN*, we conducted luciferase reporter gene assays on KYSE30 and KYSE150 cells. The results revealed that the construct harboring the *POSTN* promoter

(E) Detection of *PRRX1* mRNA by RT-qPCR in KYSE30 and KYSE150 cells, respectively. After cell seeding in a 6-well plate, transfections were done with small interfering RNA (siRNA) targeting *PRRX1* or siControl. ****$p < 0.0001$ was calculated using a Mann-Whitney test in KYSE30, whereas it was calculated using a two-tailed Student's t test in KYSE150.
(F) The effect of *PRRX1* knockdown on the relative luciferase activity of constructs containing the rs1028728[A] or rs1028728[T] allele in KYSE30 and KYSE150 cells. ****$p < 0.0001$ values were calculated using a two-tailed Student's t test.
(G) The correlation between *POSTN* expression and *PRRX1* expression was calculated in normal esophageal mucosa from GTEx, GEO: GSE53625 dataset, Chang et al.[16], ESCC cohorts, and TCGA cohorts. All p and r values were using Pearson's correlation analysis.
(H) Correlation analysis of *POSTN* and *PRRX1* expression in Mes POSTN$^+$ cells was conducted, and the scatterplot was generated with corresponding r and p values, following the aforementioned methodology. UMAP plots were employed to visually represent the clustering of Mes cells and the expression patterns of *POSTN* and *PRRX1*.

with the rs1028728[A] allele displayed markedly lower luciferase expression than the construct containing the rs10287228[T] allele (Figure 3C). We further explored whether the rs1028728 variant affected *POSTN* expression by altering the binding of a transcription factor. Through motif enrichment analysis, we identified that the rs1028728[T] allele may preferentially bind to parental homology box1 (*PRRX1* [MIM: 167420]), a key inducer of epithelial-mesenchymal transition[36,37] (Figure 3D). We then knocked down the *PRRX1* expression by using two different small interfering RNAs (Figure 3E) and found a significant decrease in luciferase expression of the construct harboring the *POSTN* promotor containing the rs10287228[T] allele compared to the control vector (Figure 3F). Conversely, the luciferase expression of the construct harboring the *POSTN* promotor containing the rs10287228[A] allele remained largely unaffected upon *PRRX1* knockdown (Figure 3F). Moreover, significant positive correlations were observed between the expression of *PRRX*1 and *POSTN* in both bulk RNA-seq and scRNA-seq datasets (Figures 3G and 3H). Collectively, these findings indicate that the risk allele T of rs1028728 enhances the binding capability of *PRRX1*, thereby regulating the high expression of *POSTN*. The transcription factor *PRRX1* may be involved in promoting epithelial-mesenchymal transition through regulating *POSTN*.

This study employed single-cell integrated analysis to identify epithelial cells and characteristic genes associated with the invasive phenotype of ESCC, overcoming the limitations of bulk RNA-seq. The findings were robust and reliable, validated through multiple databases and diverse methodologies. Additionally, we conducted population and mechanistic studies, leading to the discovery of novel invasive biomarkers. Our study possesses certain limitations. The designation of rs1028728 as an ESCC risk SNP in populations beyond the Chinese demographic necessitates additional confirmation. Furthermore, the mechanism of action of *POSTN* and *PRRX1* in epithelial cells and its role in the interaction between epithelial cells and CAFs warrant further investigation. In conclusion, through single-cell sequencing analysis, bulk RNA-seq analysis, epidemiological studies, and a series of biochemical experiments, our findings advance our comprehension of the molecular mechanisms underpinning *POSTN* and offer valuable insights for future research endeavors.

Data and code availability

The published article includes all datasets generated or analyzed during this study.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2024.100278.

Acknowledgments

This study is supported by the National Key R&D Program of China (2021YFC2502000), the National Natural Science Foundation of China (82073654), and the Natural Science Fund for Distinguished Young Scholars of Hubei Province (2020CFA067) to J.C.

Author contributions

J.C., Y.T., and L.S. designed the study. Y.T., M.P., and Y.L. conducted computational analysis. Y.T., J.M., and S.N. performed the immunohistochemistry experiments. Y.T. and S.N. wrote the manuscript with comments from all authors. J.C. supervised the project. All authors discussed the results and reviewed and revised the manuscript. The work reported in the paper has been performed by the authors unless clearly specified in the text.

Declaration of interests

The authors declare no competing interests.

Received: December 16, 2023
Accepted: February 14, 2024

Web resources

Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/geo/.
Genotype-Tissue Expression, http://gtexportal.org.
GSA-Human, https://bigd.big.ac.cn/gsa-human.
HaploReg, v.4.1, http://compbio.mit.edu/HaploReg.
Jasper website prediction, https://jaspar.genereg.net/.
RegulomeDB, v.2.2, https://regulomedb.org/regulome-search/.
TCGA database, http://cancergenome.nih.gov.
Tumor Immune Single-cell Hub 2 (TISCH2), http://tisch.comp-genomics.org/home/.
UCSC Genome Browser, http://genome.ucsc.edu/.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. *68*, 394–424.
2. Zhang, X., Peng, L., Luo, Y., Zhang, S., Pu, Y., Chen, Y., Guo, W., Yao, J., Shao, M., Fan, W., et al. (2021). Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. Nat. Commun. *12*, 5291.
3. Abnet, C.C., Arnold, M., and Wei, W.Q. (2018). Epidemiology of Esophageal Squamous Cell Carcinoma. Gastroenterology *154*, 360–373.
4. Li, X., and Wang, C.Y. (2021). From bulk, single-cell to spatial RNA sequencing. Int. J. Oral Sci. *13*, 36.
5. Chen, Y., Zhu, S., Liu, T., Zhang, S., Lu, J., Fan, W., Lin, L., Xiang, T., Yang, J., Zhao, X., et al. (2023). Epithelial cells activate fibroblasts to promote esophageal cancer development. Cancer Cell *41*, 903–918.e8.
6. Sklavenitis-Pistofidis, R., Getz, G., and Ghobrial, I. (2021). Single-cell RNA sequencing: one step closer to the clinic. Nat. Med. *27*, 375–376.

7. Villani, A.C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science *356*, eaah4573.
8. Wagner, J., Rapsomaniki, M.A., Chevrier, S., Anzeneder, T., Langwieder, C., Dykgers, A., Rees, M., Ramaswamy, A., Muenst, S., Soysal, S.D., et al. (2019). A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. Cell *177*, 1330–1345.e18.
9. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401.
10. Liu, T., Zhao, X., Lin, Y., Luo, Q., Zhang, S., Xi, Y., Chen, Y., Lin, L., Fan, W., Yang, J., et al. (2022). Computational Identification of Preneoplastic Cells Displaying High Stemness and Risk of Cancer Progression. Cancer Res. *82*, 2520–2537.
11. Sun, D., Guan, X., Moran, A.E., Wu, L.Y., Qian, D.Z., Schedin, P., Dai, M.S., Danilov, A.V., Alumkal, J.J., Adey, A.C., et al. (2022). Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. Nat. Biotechnol. *40*, 527–538.
12. Pan, X., Wang, J., Guo, L., Na, F., Du, J., Chen, X., Zhong, A., Zhao, L., Zhang, L., Zhang, M., et al. (2022). Identifying a confused cell identity for esophageal squamous cell carcinoma. Signal Transduct. Target. Ther. *7*, 122.
13. Li, R., Di, L., Li, J., Fan, W., Liu, Y., Guo, W., Liu, W., Liu, L., Li, Q., Chen, L., et al. (2021). A body map of somatic mutagenesis in morphologically normal human tissues. Nature *597*, 398–403.
14. Chang, J., Zhao, X., Wang, Y., Liu, T., Zhong, C., Lao, Y., Zhang, S., Liao, H., Bai, F., Lin, D., and Wu, C. (2023). Genomic alterations driving precancerous to cancerous lesions in esophageal cancer development. Cancer Cell *41*, 2038–2050.e5.
15. Chang, J., Zhong, R., Tian, J., Li, J., Zhai, K., Ke, J., Lou, J., Chen, W., Zhu, B., Shen, N., et al. (2018). Exome-wide analyses identify low-frequency variant in CYP26B1 and additional coding variants associated with esophageal squamous cell carcinoma. Nat. Genet. *50*, 338–343.
16. Chang, J., Tan, W., Ling, Z., Xi, R., Shao, M., Chen, M., Luo, Y., Zhao, Y., Liu, Y., Huang, X., et al. (2017). Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. Nat. Commun. *8*, 15290.
17. Kalabis, J., Oyama, K., Okawa, T., Nakagawa, H., Michaylira, C.Z., Stairs, D.B., Figueiredo, J.L., Mahmood, U., Diehl, J.A., Herlyn, M., and Rustgi, A.K. (2008). A subpopulation of mouse esophageal basal cells has properties of stem cells with the capacity for self-renewal and lineage specification. J. Clin. Invest. *118*, 3860–3869.
18. DeWard, A.D., Cramer, J., and Lagasse, E. (2014). Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population. Cell Rep. *9*, 701–711.
19. Morrisey, E.E., and Rustgi, A.K. (2018). The Lung and Esophagus: Developmental and Regenerative Overlap. Trends Cell Biol. *28*, 738–748.
20. Kang, X., Xu, E., Wang, X., Qian, L., Yang, Z., Yu, H., Wang, C., Ren, C., Wang, Y., Lu, X., et al. (2021). Tenascin-c knockdown suppresses vasculogenic mimicry of gastric cancer by inhibiting ERK- triggered EMT. Cell Death Dis. *12*, 890.
21. Zhou, J., Yang, Y., Zhang, H., Luan, S., Xiao, X., Li, X., Fang, P., Shang, Q., Chen, L., Zeng, X., and Yuan, Y. (2022). Overexpressed COL3A1 has prognostic value in human esophageal squamous cell carcinoma and promotes the aggressiveness of esophageal squamous cell carcinoma by activating the NF-kappaB pathway. Biochem. Biophys. Res. Commun. *613*, 193–200.
22. Agarwal, S., Behring, M., Kim, H.G., Chandrashekar, D.S., Chakravarthi, B.V.S.K., Gupta, N., Bajpai, P., Elkholy, A., Al Diffalha, S., Datta, P.K., et al. (2020). TRIP13 promotes metastasis of colorectal cancer regardless of p53 and microsatellite instability status. Mol. Oncol. *14*, 3007–3029.
23. Jia, Y., Wang, N., Wang, J., Tian, H., Ma, W., Wang, K., Tan, B., Zhang, G., Yang, S., Bai, B., and Cheng, Y. (2014). Down-regulation of stromal caveolin-1 expression in esophageal squamous cell carcinoma: a potent predictor of lymph node metastases, early tumor recurrence, and poor prognosis. Ann. Surg Oncol. *21*, 329–336.
24. Li, Y., Ren, P., Dawson, A., Vasquez, H.G., Ageedi, W., Zhang, C., Luo, W., Chen, R., Li, Y., Kim, S., et al. (2020). Single-Cell Transcriptome Analysis Reveals Dynamic Cell Populations and Differential Gene Expression Patterns in Control and Aneurysmal Human Aortic Tissue. Circulation *142*, 1374–1388.
25. Abnet, C.C., Freedman, N.D., Hu, N., Wang, Z., Yu, K., Shu, X.O., Yuan, J.M., Zheng, W., Dawsey, S.M., Dong, L.M., et al. (2010). A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. Nat. Genet. *42*, 764–767.
26. Wu, C., Kraft, P., Zhai, K., Chang, J., Wang, Z., Li, Y., Hu, Z., He, Z., Jia, W., Abnet, C.C., et al. (2012). Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. Nat. Genet. *44*, 1090–1097.
27. Mueller, M.M., and Fusenig, N.E. (2004). Friends or foes - bipolar effects of the tumour stroma in cancer. Nat. Rev. Cancer *4*, 839–849.
28. Joyce, J.A., and Pollard, J.W. (2009). Microenvironmental regulation of metastasis. Nat. Rev. Cancer *9*, 239–252.
29. Chiodoni, C., Colombo, M.P., and Sangaletti, S. (2010). Matricellular proteins: from homeostasis to inflammation, cancer, and metastasis. Cancer Metastasis Rev. *29*, 295–307.
30. Christofori, G. (2006). New signals from the invasive front. Nature *441*, 444–450.
31. Michaylira, C.Z., Wong, G.S., Miller, C.G., Gutierrez, C.M., Nakagawa, H., Hammond, R., Klein-Szanto, A.J., Lee, J.S., Kim, S.B., Herlyn, M., et al. (2010). Periostin, a cell adhesion molecule, facilitates invasion in the tumor microenvironment and annotates a novel tumor-invasive signature in esophageal cancer. Cancer Res. *70*, 5281–5292.
32. Wong, G.S., Lee, J.S., Park, Y.Y., Klein-Szanto, A.J., Waldron, T.J., Cukierman, E., Herlyn, M., Gimotty, P., Nakagawa, H., and Rustgi, A.K. (2013). Periostin cooperates with mutant p53 to mediate invasion through the induction of STAT1 signaling in the esophageal tumor microenvironment. Oncogenesis *2*, e59.
33. Liu, C., Zhang, M., Yan, X., Ni, Y., Gong, Y., Wang, C., Zhang, X., Wan, L., Yang, H., Ge, C., et al. (2023). Single-cell dissection of cellular and molecular features underlying human cervical squamous cell carcinoma initiation and progression. Sci. Adv. *9*, d8977.

34. Lin, S.C., Liao, Y.C., Chen, P.M., Yang, Y.Y., Wang, Y.H., Tung, S.L., Chuang, C.M., Sung, Y.W., Jang, T.H., Chuang, S.E., and Wang, L.H. (2022). Periostin promotes ovarian cancer metastasis by enhancing M2 macrophages and cancer-associated fibroblasts via integrin-mediated NF-kappaB and TGF-beta2 signaling. J. Biomed. Sci. *29*, 109.
35. Chang, J., Zhong, R., Tian, J., Li, J., Zhai, K., Ke, J., Lou, J., Chen, W., Zhu, B., Shen, N., et al. (2018). Exome-wide analyses identify low-frequency variant in CYP26B1 and additional coding variants associated with esophageal squamous cell carcinoma. Nat. Genet. *50*, 338–343.
36. Jiang, Y.P., Tang, Y.L., Wang, S.S., Wu, J.S., Zhang, M., Pang, X., Wu, J.B., Chen, Y., Tang, Y.J., and Liang, X.H. (2020). PRRX1-induced epithelial-to-mesenchymal transition in salivary adenoid cystic carcinoma activates the metabolic reprogramming of free fatty acids to promote invasion and metastasis. Cell Prolif. *53*, e12705.
37. Feldmann, K., Maurer, C., Peschke, K., Teller, S., Schuck, K., Steiger, K., Engleitner, T., Öllinger, R., Nomura, A., Wirges, N., et al. (2021). Mesenchymal Plasticity Regulated by Prrx1 Drives Aggressive Pancreatic Cancer Biology. Gastroenterology *160*, 346–361.e24.