JEM Journal of Experimental Medicine

ARTICLE

E-cadherin loss drives diffuse-type gastric tumorigenesis via EZH2-mediated reprogramming

Gengyi Zou[1*], Yuanjian Huang[1,2*], Shengzhe Zhang[1], Kyung-Pil Ko[1], Bongjun Kim[1], Jie Zhang[1], Vishwa Venkatesan[1], Melissa P. Pizzi[3], Yibo Fan[3], Sohee Jun[1], Na Niu[6], Huamin Wang[7], Shumei Song[3], Jaffer A. Ajani[3], and Jae-Il Park[1,4,5]

Diffuse-type gastric adenocarcinoma (DGAC) is a deadly cancer often diagnosed late and resistant to treatment. While hereditary DGAC is linked to *CDH1* mutations, the role of CDH1/E-cadherin inactivation in sporadic DGAC tumorigenesis remains elusive. We discovered CDH1 inactivation in a subset of DGAC patient tumors. Analyzing single-cell transcriptomes in malignant ascites, we identified two DGAC subtypes: DGAC1 (CDH1 loss) and DGAC2 (lacking immune response). DGAC1 displayed distinct molecular signatures, activated DGAC-related pathways, and an abundance of exhausted T cells in ascites. Genetically engineered murine gastric organoids showed that *Cdh1* knock-out (KO), *Kras*^{G12D}, *Trp53* KO (EKP) accelerates tumorigenesis with immune evasion compared with *Kras*^{G12D}, *Trp53* KO (KP). We also identified EZH2 as a key mediator promoting CDH1 loss-associated DGAC tumorigenesis. These findings highlight DGAC's molecular diversity and potential for personalized treatment in CDH1-inactivated patients.

Introduction

Gastric adenocarcinoma (GAC) is the fourth most common cause of cancer deaths worldwide (Sung et al., 2021). GAC is mainly divided into intestinal-type gastric adenocarcinoma (IGAC, 50%), diffuse-type gastric adenocarcinoma (DGAC, 30%), and mixed (Iyer et al., 2020). DGAC is histologically characterized by poor differentiation, loss of cell adhesion proteins, fibrosis, and infiltration. Unlike IGAC, DGAC is relatively more often observed in younger, female, and Hispanic populations than in older, male, and non-Hispanic ones (Chen et al., 2016; Wang et al., 2020c). While the incidence of IGAC has declined due to *Heliobacter pylori* (HP) therapy and lifestyle improvements over the past few decades, the number of DGAC cases has remained constant or has risen (Henson et al., 2004; Assumpção et al., 2020).

DGAC tends to metastasize to the peritoneal cavity, which makes it difficult to diagnose early by imaging. In addition, isolated tumor cells or small clusters of tumor cells infiltrate in unpredictable patterns. Thus, DGAC is often detected at a late stage, leading to a poor prognosis. For such patients, curative resection is not possible. Systemic therapy is the main option for potentially prolonging survival and improving symptoms (Muro et al., 2019; Ajani et al., 2022). Despite the distinct features of DGAC in both a molecular basis and therapy resistance, the first-line treatment options are not specific to DGAC (Garcia-Pelaez et al., 2021; Ajani et al., 2022). Systemic therapy with targeted therapy has shown limited benefits (Selim et al., 2019; Körfer et al., 2021). In parallel, immune checkpoint inhibitors (ICIs) have been used recently. The advent of first-generation ICIs that target cytotoxic T-lymphocyte antigen 4 (CTLA4) and programmed death-ligand 1 (PD-L1) has brought a paradigm shift in the treatment of various advanced cancers (Mazzarella et al., 2019). Nivolumab (PD-1 inhibitor) can be either combined with chemotherapy as first-line treatment or used as monotherapy as later-line treatment in Asia (Boku et al., 2021; Janjigian et al., 2021). Pembrolizumab (PD-1 inhibitor) showed a promising outcome in treating GAC with high microsatellite instability or high tumor mutational burden (Wainberg et al., 2021). However, DGAC imposes major difficulty in the clinic, and available therapies perform poorly. Generally, DGAC has immunosuppressed stroma and is genomically stable (Teng et al., 2015; Ge et al., 2018). Given the limited therapeutic options for DGAC, it is imperative to understand the biology of DGAC, which may establish a groundwork for developing new targeted therapies for DGAC. Furthermore, for maximizing

---

[1]Division of Radiation Oncology, Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [2]Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China; [3]Department of GI Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [4]Graduate School of Biomedical Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [5]Program in Genetics and Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [6]Department of Pathology, Yale School of Medicine, New Haven, CT, USA; [7]Division of Pathology/Lab Medicine, Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.

*G. Zou and Y. Huang contributed equally to this paper. Correspondence to Jae-Il Park: jaeil@mdanderson.org.

https://doi.org/10.1084/jem.20230561

therapeutic efficacy, it is crucial to identify patients who can most benefit from specific treatment options. Nevertheless, to date, DGAC patient stratification by molecular signatures has not been achieved.

Hereditary DGAC, as a minor proportion of DGAC (1–3%), is mainly characterized by germline mutations in the *CDH1* gene that encodes the E-cadherin protein (Blair et al., 2020). However, other than hereditary DGAC, the role of CDH1 loss in DGAC tumorigenesis is unclear. Cell-to-cell adhesion is a crucial phenomenon for maintaining tissue morphogenesis and homeostasis, as well as for regulating cell differentiation, survival, and migration. E-cadherin mediates cell-to-cell adhesion, which is essential for determining the proliferation specificity and differentiation of epithelial cells and preventing invasion (van Roy and Berx, 2008). To understand the impact of CDH1 loss on DGAC tumorigenesis, we analyzed single-cell transcriptomes of cryopreserved peritoneal carcinomatosis (PC) from 19 DGAC patients and identified two subtypes of DGACs exhibiting specific molecular signatures including E-cadherin loss and immune landscape remodeling. To further verify our in silico analysis, we generated and characterized a genetically engineered gastric organoid (GO) model that recapitulates E-cadherin inactivation-associated DGAC tumorigenesis. This study stratifies DGAC patients by single-cell transcriptomics and elucidates the unexpected role of E-cadherin loss in transcriptional reprogramming and immune evasion, providing novel insights into E-cadherin loss-associated DGAC tumorigenesis.

Results

CDH1 inactivation in DGAC

To explore the role of CDH1 in DGAC tumorigenesis, we examined the genetic alterations and protein levels of CDH1 in DGAC. According to cBioPortal, 25% of tumors from the DGAC patients showed *CDH1* gene alterations, including mutations and deep deletions (Fig. 1 A). We also assessed the CDH1/E-cadherin protein expression in the tissue microarray of 114 DGAC patients' tumor samples (patient information was listed in Table S4). Immunohistochemistry (IHC) showed that 37.72% of DGAC patients were CDH1 negative, 37.72% exhibited low CDH1 expression, and 24.56% displayed high CDH1 expression (Fig. 1 B), which was also quantified with histochemical scoring assessment (H-score) of each slide (Fig. 1 C). Next, we determined the transcriptional signature of DGAC at the single-cell transcriptomics level by analyzing single-cell RNA-seq (scRNA-seq) datasets of PC cells from 19 stage IV DGAC patients (Fig. 1 D and Table S5) (Wang et al., 2021). After data integration and normalization, a total of 30 cell clusters were generated according to distinctive gene expression patterns (Fig. 1 E; and Fig. S1, A and B; and Table S6). We reclustered the datasets as mega clusters according to Leiden-based UMAP (Fig. 1 F). To conduct the precise subtyping of DGAC, we reanalyzed the scRNA-seq datasets with only epithelial cells (Fig. 1 G, Fig. S1 C, and Table S7). An unsupervised pair-wise correlation analysis showed that the combined datasets of 19 DGAC patients were divided into two major subtypes (DGAC1 and DGAC2) (Fig. 1 H). To comparatively analyze DGAC 1 and 2 according to their clinical information (Table S5), we have thoroughly examined the available data and compared various clinical and pathological features between the two subtypes. Upon analysis, we did not observe significant differences in survival time, race, gender, or age between DGAC1 and DGAC2 subtypes (Fig. S1, D, E, G, and I). Regarding pathological features, DGAC1 had a higher proportion (DGAC1: 3/11, 27.3%; DGAC2: 1/8, 12.5%) of patients with non-signet ring cell carcinoma (Fig. S1 F). A notable distinction in metastatic sites was displayed. DGAC1 patients exhibited a higher prevalence of metastatic sites compared with DGAC2. This observation suggests potential differences in the metastatic behavior of the two subtypes (Fig. S1 H). The transcriptional signature of DGAC1 epithelial cell clusters was distinct from that of DGAC2 (Fig. 1, I and J; and Table S8). In line with the heterogeneity of *CDH1*'s genomic alterations and expression in DGAC patients (Fig. 1, A and B), the DGAC1 subtype exhibited a relatively lower expression of *CDH1* compared with DGAC2 (Fig. 1, K and L), indicating that the unsupervised pair-wise subtyping can also stratify DGAC patients by *CDH1* expression. While tissue microarray analysis showed that within the cohort of 114 DGAC patients, where 37.7% of patients exhibit low *CDH1* levels (Fig. 1, B and C), this subset of patients may have been classified into DGAC1 or DGAC2 based on the differential expression of other signature genes specific to each cluster, rather than solely relying on *CDH1* expression. We also identified the molecular signatures of DGAC1 and DGAC2 (Fig. 1 M). The gene list for calculating signature scores, including the DGAC1 and DGAC2 signatures, comprised the top 50 highly variable genes from each subgroup (Fig. 1 M and Table S8). These results identify two distinct subtypes of DGACs by distinct molecular signatures and CDH1/E-cadherin expression.

Molecular characterization of DGAC subtypes

Next, we characterized the molecular subtypes of DGAC. Given that E-cadherin downregulation is commonly observed in epithelial tumors and is a hallmark of the epithelial-to-mesenchymal transition (EMT), we checked the EMT scores based on the established gene set (Table S9). DGAC1 showed a higher EMT score compared with DGAC2 (Fig. S1 J). Extensive genomic analyses of GAC have found that DGACs display distinct activation of signaling pathways different from IGACs (Astudillo, 2020; Koushyar et al., 2020; Lei et al., 2022; Liu et al., 2022, 2023; Ooki and Yamaguchi, 2022; Seeneevassen et al., 2022; Messina et al., 2023). scRNA-seq-based gene set enrichment analysis (GSEA) showed that the WNT gene set scores were enriched in DGAC1, which corresponds to decreased *CDH1* expression compared with DGAC2 (Fig. S1 L). Conversely, the RHOA gene set score was relatively higher in DGAC2 compared with DGAC1 (Fig. S1 K). Additionally, we analyzed the copy number variation (CNV) of DGACs by using normal stomach samples as a reference. We combined 29 scRNA-seq datasets of normal stomach samples (Normal) with the previous 19 DGAC patients (Kim et al., 2022) (Fig. S2 A). Except for the endothelial cell markers, the same marker panel was utilized as the previous DGAC subcategory process to annotate the cells into epithelial cells, myeloid cells, B cells, plasma cells, T cells, effector T cells, naïve T cells, exhausted T cells, fibroblasts, and endothelial cells

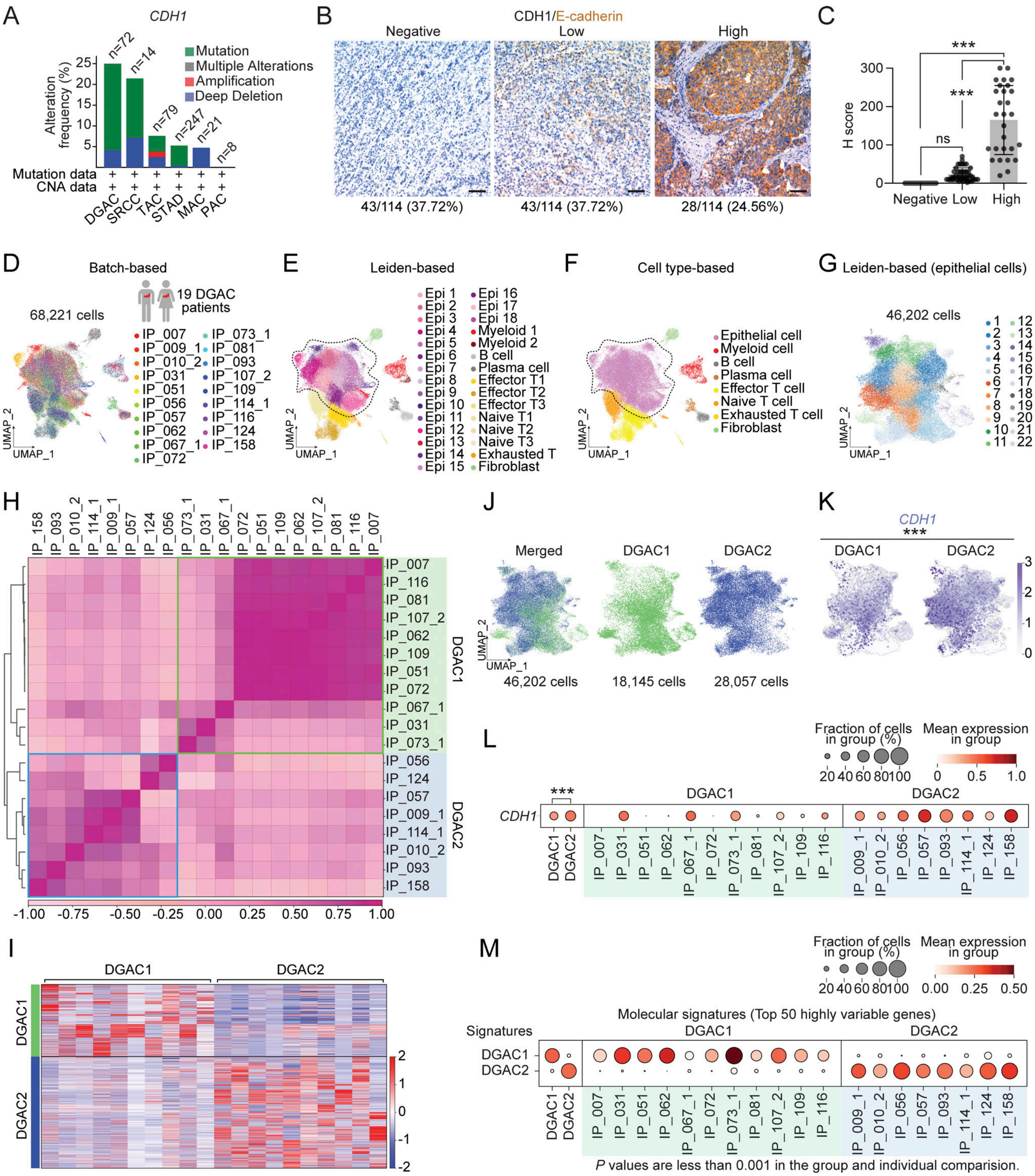


Figure 1. ***CDH1* inactivation in DGAC patient tumor cells.** **(A)** Genetic alterations of the *CDH1* based on the cBioPortal stomach cancer datasets (https://www.cbioportal.org). *n* represents the total patients number enrolled in each subtype. DGAC, diffuse-type gastric adenocarcinoma; SRCC, signet ring cell carcinoma; TAC, tubular adenocarcinoma; STAD, stomach adenocarcinoma; MAC, mucinous adenocarcinoma; PAC, papillary adenocarcinoma. **(B and C)** IHC staining of CDH1/E-cadherin in 114 DGAC patient tumor samples. The representative images are shown; scale bar: 50 µm (B). Quantification of H score of CDH1 expression (C). P values were calculated using the one-way ANOVA; error bars: standard deviation (SD). Clinical information of 114 DGAC patients is available in Table S4. **(D)** Integrated batch-based UMAPs of 19 DGAC patients; integration package: Harmony. Clinical information of 19 DGAC patients is available in Table S5. **(E)** Integrated Leiden-based UMAP of 19 DGAC patients. Dashed line circle: epithelial cells. Epi: epithelial cells; Myeloid: myeloid cells; Effector T: effector T cells; Naïve T: naïve T cells; Exhausted T: exhausted T cells. **(F)** Integrated cell type-based UMAP of 19 DGAC patients. All cells were reclustered according to the Leiden clusters and gathered as mega clusters. Dashed line circle: epithelial cells. **(G)** Epithelial cells were reclustered by Leiden. **(H)** Correlation matrix plot of epithelial cells showing pair-wise correlations among all samples above. The dendrogram shows the distance of each dataset based on principal component analysis, and the Pearson correlation is displayed with a color spectrum. Groups of patients were categorized by dendrogram and correlation. **(I)** Type-based

heatmap of epithelial cells of integrated datasets in 19 DGAC patients. The top 100 highly variable genes of each type are shown in Table S8. **(J)** Type-based integrated and separated UMAPs of DGAC1 and DGAC2. **(K)** Feature plots of epithelial cells displaying *CDH1* expression. **(L)** Dot plots of epithelial cells of *CDH1* expression in different DGAC groups and individual patients. P values were calculated by using a *t* test. **(M)** Molecular signatures of DGAC1 and DGAC2 patients. Top 50 highly variable genes were used to calculate the molecular signature of each group. Gene list is shown in Table S8. Dot plots of epithelial cells of each molecular signature in different subtypes and individual patient. P values were calculated by Mann–Whitney testing. ns: $P > 0.05$, ***: $P \leq 0.001$. All data are derived from two or more independent experiments with the indicated number of human donors.

(Fig. S1 A; and Fig. S2, B and C). Leiden-based UMAP exhibited the same cell types as the DGAC stratification analysis (Fig. S2, D and E; and Table S10), except that the endothelial cell cluster appeared due to the normal tissue (Fig. S2 C). According to the previously identified DGAC subgroups, we separated the UMAP as Normal, DGAC1, and DGAC2 (Fig. S2, F and G). Although the epithelial cells were defined as $EPCAM^{high}$ clusters among all groups, epithelial cells from the Normal group were clearly isolated from the major epithelial cell population of the integrated datasets (Fig. S2 G). CNV patterns were different between DGAC1 and DGAC2 (Fig. S2 H). We observed notable differences between DGAC1 and DGAC2 regarding copy number gains on specific chromosomes. In DGAC1, we observed more pronounced copy number gains on chromosomes 3, 9, 19, and X, while in DGAC2, there were increased copy number gains on chromosomes 1, 8, 11, 17, 20, and 21. These differences in copy number alterations were found to be statistically significant, as indicated by the adjusted P values (Fig. S2, I and J). These results indicate the heterogeneity of DGAC with differentially activated signaling pathways.

Immune landscape remodeling with T cell exhaustion in DGAC1

Having determined the molecular signatures of DGAC tumor cells, we next analyzed the immunological response associated with DGAC ascites. Intriguingly, scRNA-seq-based immune cell profiling showed that compared with DGAC2, where immune cells barely existed, DGAC1 was highly enriched with immune cells, including T cells, B cells, and myeloid cells (Fig. 2, A–C; and Fig. S2, K and L). Additionally, we examined cellular networks among all cell clusters (DGAC1 versus DGAC2) using a CellChat package that infers cell-to-cell functional interaction based on ligand–receptor expression (Jin et al., 2021). Compared with DGAC2, DGAC1 showed relatively more inferred interactions among different cell types (Fig. 2 D). According to the differential number of interactions, the interactions between fibroblast and epithelial and endothelial cells were decreased, while widespread increased interactions were found in DGAC1 compared with DGAC2 (Fig. 2 E). Notably, exhausted T cells, as a receiver, showed the most increased interactions compared with other T cells in DGAC1, which is the major population among all immune cells (Fig. 2 F). GSEA identified the pathways that are enriched in DGAC1 with six gene sets, including gene sets derived from the Gene Ontology Biological Process (GOBP), and five canonical pathways gene sets (REACTOME, WP, BIOCARTA, PID, and KEGG) (Fig. S3, A–F; and Tables S11, S12, S13, S14, S15, and S16). Except for REACTOME (Fig. S3 B), T cell-related immune response pathways were enriched in DGAC1 based on the other five gene sets (Fig. S3, A and C–F). Consistent with the CellChat prediction and GSEA results, DGAC1 showed significant upregulation of T cell exhaustion markers (*LAG3*, *TIGIT*, *CTLA4*, and *HAVCR2*) and increased T cell exhaustion score compared with DGAC2 (Fig. 2, G–I). Similarly, immune checkpoints-related genes (*CTLA4*, *PDCD1*, *PDCD1LG2*, and *CD274*) and their scores were markedly upregulated in DGAC1 over DGAC2 (Fig. 2, J–L). In addition to T cell analysis, we also examined myeloid-derived suppressor cells (MDSC) and macrophage polarization. MDSC score is relatively higher in DGAC1 than in DGAC2 (Fig. 2, M–O). Meanwhile, most of the M1 and M2 macrophage polarization maker expression is enriched in DGAC1 compared with DGAC2 (Fig. S3, G and H). These results suggest that compared with DGAC2, the DGAC1 subtype exhibits distinct immune remodeling featured by T cell exhaustion and increased expression of the genes associated with immune checkpoints.

E-cadherin loss induces neoplasia in conjunction with *Trp53* KO and $Kras^{G12D}$

To validate the in silico results, we utilized murine GOs that enable multiple genetic engineering with immediate phenotype analyses. *Cdh1* deficiency results in early-stage DGAC phenotype in a mouse model (Mimata et al., 2011; Hayakawa et al., 2015). Nevertheless, other genes need to be included to recapitulate DGAC tumorigenesis. The genes encoding the receptor tyrosine kinase (RTK)–RAS signaling pathway and the *TP53* gene were profoundly disrupted in DGAC (Cancer Genome Atlas Research Network, 2014; Cristescu et al., 2015). *KRAS* and *TP53* were genetically altered in 13.19% and 36.11% of DGAC cases, respectively, as per cBioPortal analysis (Fig. 3 A). Therefore, to create preneoplastic or neoplastic conditions to determine the impact of *CDH1* loss on DGAC tumorigenesis (Till et al., 2017), we genetically manipulated three genes (*Cdh1*, *Trp53*, and *Kras*) in GOs. Briefly, from the *Cdh1* wild type (WT) and $Kras^{LSL\text{-}G12D/+}$; $Trp53^{fl/fl}$ mice, gastric epithelial cells were isolated to culture them into GOs (Fig. 3 B). Subsequently, using the Cre-LoxP recombination and CRISPR-based genetic manipulation, we established two lines of GOs carrying $Kras^{G12D/+}$ and *Trp53* deletion in combination with *Cdh1* KO (KP: $Kras^{G12D/+}$; *Trp53* KO [KP], *Cdh1*/E-cadherin KO; $Kras^{G12D/+}$; *Trp53* KO [EKP]) (Fig. 3 B). Genetic modifications were validated by PCR-based genotyping, genomic DNA sequencing, and immunofluorescence (IF) staining (Fig. S4, A–C and Fig. 3 G). Meanwhile, we monitored their sizes and numbers by macroscopic analyses during passages to maintain the stable culture process during passages (Fig. 3, C and D). Unlike WT GOs growing as a single layer of epithelial cells, KP and EKP GOs displayed multilayered epithelium (Fig. 3 E). Notably, compared with WT and KP, EKP GOs exhibited abnormal morphology such as vacuolization and cell adhesion loss along with cell hyperplasia (Fig. 3 E). Additionally, EKP GOs were hyperproliferative compared with WT and KP GOs,

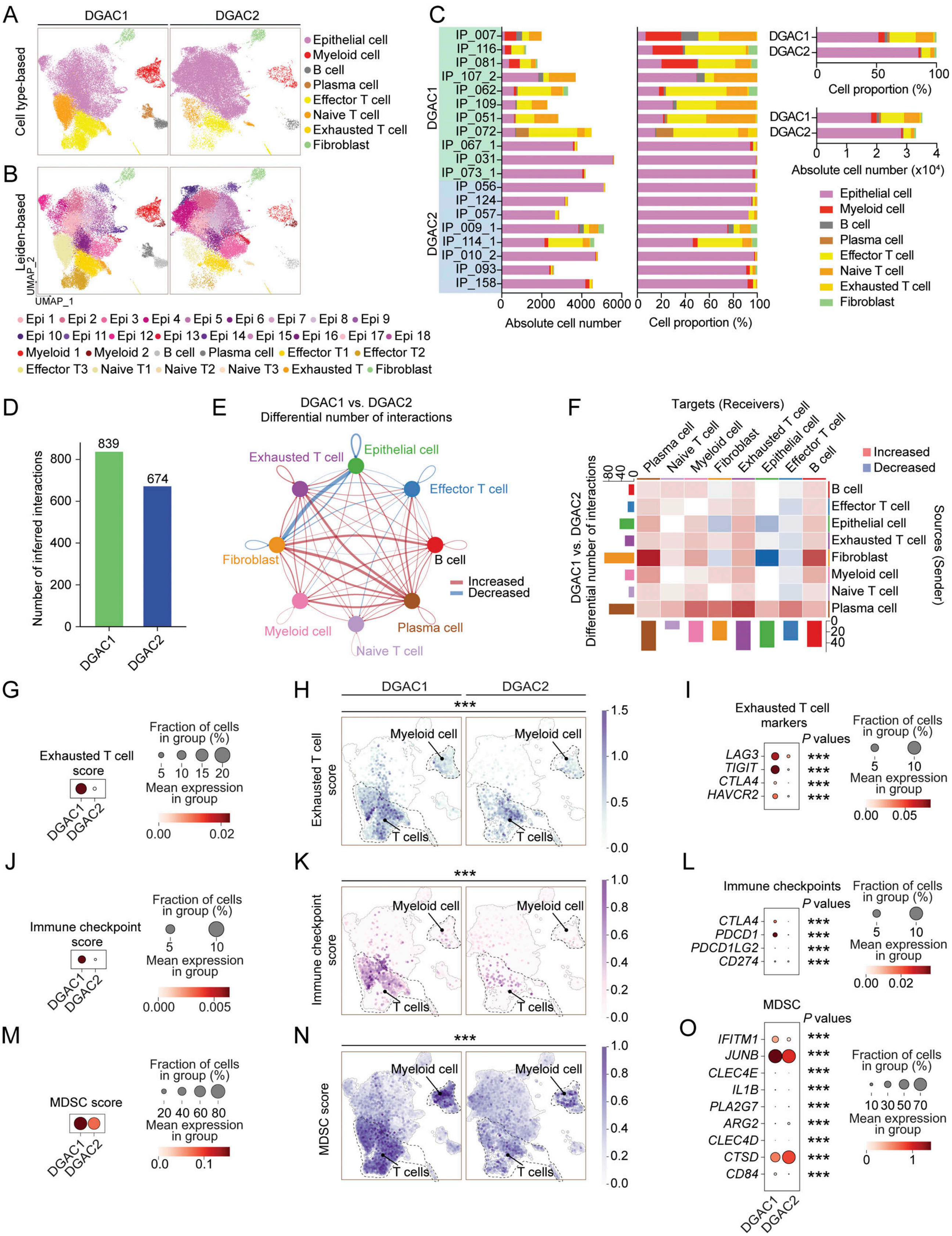


Figure 2. **Comparative analyses of immune landscapes of DGAC subtypes. (A and B)** Cell type-based and Leiden-based UMAPs of DGAC1 and DGAC2. **(C)** Absolute and relative cell proportions of individual patients and DGAC subtypes. Patients list was ranked by the DGAC group they belong. **(D)** Total cell–cell interactions from DGAC1 and DGAC2 were analyzed by using the CellChat package. More interactions were found in DGAC1. **(E and F)** Differential number of

interactions between DGAC1 and DGAC2 using circle plots (E) and heatmap (F). Red (or blue) colored edges (E) and squares (F) represent increased (or decreased) signaling in the DGAC1 compared to DGAC2. **(G–I)** Score-based dot plot (G), feature plots (H), and dot plot of individual marker gene (I) of exhausted T cell score (markers are included in that score: *LAG3*, *TIGIT*, *CTLA4*, and *HAVCR2*). Genes that were included in score analysis are shown in Table S9, P values were calculated by Mann–Whitney testing. For the dot plot of a single gene, P values were calculated by using a *t* test. **(J–L)** Score-based dot plot (J), feature plots (K), and dot plot of individual marker gene (L) of immune checkpoint score (markers are included in that score: *CTLA4*, *PDCD1*, *PDCD1LG2*, and *CD274*). Genes that were included in score analysis are shown in Table S9, P values were calculated by Mann–Whitney testing. For the dot plot of single gene, P values were calculated by using a *t* test. **(M–O)** Score-based dot plot (M), feature plots (N), and dot plot of individual marker gene (O) of exhausted T cell score (markers are included in that score: *IFITM1*, *JUNB*, *CLEC4E*, *IL1B*, *PLA2G7*, *ARG2*, *CLEC4D*, *CTSD*, and *CD84*). Genes that were included in score analysis are shown in Table S9, P values were calculated by Mann–Whitney testing. For the dot plot of single gene, P values were calculated by using a *t* test. ***: P ≤0.001. All data are derived from two or more independent experiments with the indicated number of human donors.

---

assessed by immunostaining of MKI67, a cell proliferation marker (Fig. 3, F and H).

We next interrogated the mechanism of E-cadherin loss-associated DGAC tumorigenesis by performing multiplex scRNA-seq of WT, KP, and EKP GOs (Fig. S4 D). Each group was tagged with two cell multiplexing oligo (CMO) tags and then pooled together with the same number of cells after being counted. All datasets were integrated with the Harmony algorithm (Korsunsky et al., 2019) to minimize the batch effect (Fig. S4 E). WT, KP, and EKP GOs were integrated well in a batch-based UMAP (Fig. 3 I). To identify the gene signature of each cell cluster, we generated a heatmap to calculate the top 5,000 highly variable genes (Fig. S4 F). Each UMAP and heatmap represented the different cell distribution among three types of GOs (Fig. 3, J and K; and Fig. S4, G–I; and Table S17). Notably, Aquaporin 5 (*Aqp5*), a gastric tissue stem cell marker (Tan et al., 2020), was decreased in EKP compared with WT and KP (Fig. 3 K).

To determine the pathological relevance of EKP GOs with human DGAC, we utilized a single-cell inferred site-specific omics resource (Scissor) analysis (Sun et al., 2022) and assessed the transcriptomic similarity between EKP GOs and the bulk RNA-seq data of patients diagnosed with DGAC from the TCGA database. While using WT organoids as a reference and comparing the transcriptional signature, we observed that EKP organoids displayed similarities in gene expression features of human DGAC (Fig. 3 L). To determine the subtype similarity, we compared the EKP scRNA-seq data with our own datasets (DGAC1 and DGAC2) rather than relying solely on the TCGA database. The analysis revealed that EKP organoids exhibited a greater resemblance to DGAC1 transcriptional signature compared with DGAC2 (Fig. 3 M). Next, by comparing the expression levels of DGAC1 and DGAC2 signatures in EKP (Fig. 1 M), we observed a higher presence of DGAC1 signature compared with DGAC2 (Fig. 3 N). These data suggest that CDH1 loss combined with TP53 inactivation and KRAS hyperactivation (EKP) is sufficient to induce transformation, and EKP organoids display similar transcriptional features to DGAC1, indicating pathological relevance of EKP GOs to human DGAC.

Cdh1 KO induces immune evasion of tumor cells

Having determined distinct immune remodeling with T cell exhaustion in the DGAC1 subtype where *CDH1* is downregulated (Fig. 2), we asked whether genetic ablation of *CDH1* contributes to immune evasion of DGAC. To test this, we established KP and EKP GO-derived cell lines in 2D culture with minimum growth factors (culture medium: DMEM Complete Medium with 10% fetal bovine serum) for allograft transplantation (Fig. 4 A). Unlike WT GOs that failed to grow in 2D culture, both KP and EKP cells grew in 2D culture and were maintained well at multiple passages. Then, KP and EKP cell lines derived from the C57BL/6 strain were used for transplantation into C57BL/6 mice. The morphological characteristics of KP and EKP cells exhibited notable differences. KP cells exhibited a compact and tightly packed phenotype, forming densely clustered colonies, while EKP cells displayed a more loosely arranged and dispersed morphology, lacking the cohesive structure of KP cells (Fig. 4 A). Of note, there was no significant difference in cell proliferation between KP and EKP cells (Fig. 4 B). However, transplantation results showed that tumor incidence and volume of EKP tumors were markedly higher than KP tumors (tumor incidence rates: EKP [91.7%] versus KP [16.7%]) (Fig. 4, C–E). Histologically, EKP tumors exhibited poorly differentiated tumor cells, the feature of DGAC (Fig. 4 F) with *CDH1* loss (Fig. 4 H) and increased cell proliferation (Fig. 4, G and I).

To further determine the impact of *CDH1* loss on immune evasion, we performed the immunostaining of KP and EKP tumors. We observed that CD3 (a marker for all T cells), CD4 (a marker for helper T cells), and TIM3 (a marker for exhausted T cells) were enriched in the EKP tumor cortex compared with KP cortex (Fig. 4, J, K, N, P, Q, and T), and the CD8 (a marker for killer T cells) expression is similar between KP and EKP tumors (Fig. 4, L and R). PDCD1, a marker for exhausted T cells, showed increased expression in the middle and cortex of EKP compared with the same part of KP tumors (Fig. 4, M and S). Furthermore, we performed LY6G (a marker for MDSCs) and CD11B (a marker for myeloid cells) co-staining on tumor slides of KP and EKP and observed relatively higher enrichment of MDSC markers in EKP (Fig. 4, O and U). These results suggest that CDH1 is a gatekeeper restricting the immune evasion of DGAC, confirming immune landscape remodeling associated with the DGAC1 subtype where CDH1 is inactivated.

E-cadherin depletion-activated *Ezh2* regulon promotes gastric tumorigenesis

Since E-cadherin loss is associated with distinct molecular signatures of DGAC1 (Fig. 1 M), we sought to identify key transcriptional regulatory modules (regulons) activated by *CDH1* depletion. We integrated the scRNA-seq datasets of WT, KP, and EKP into batch-based and regulon pattern-based UMAPs (Fig. 5 A). In the regulon activity-based UMAP, six major transcriptional clusters (0–5) were identified (Fig. 5 A). With the

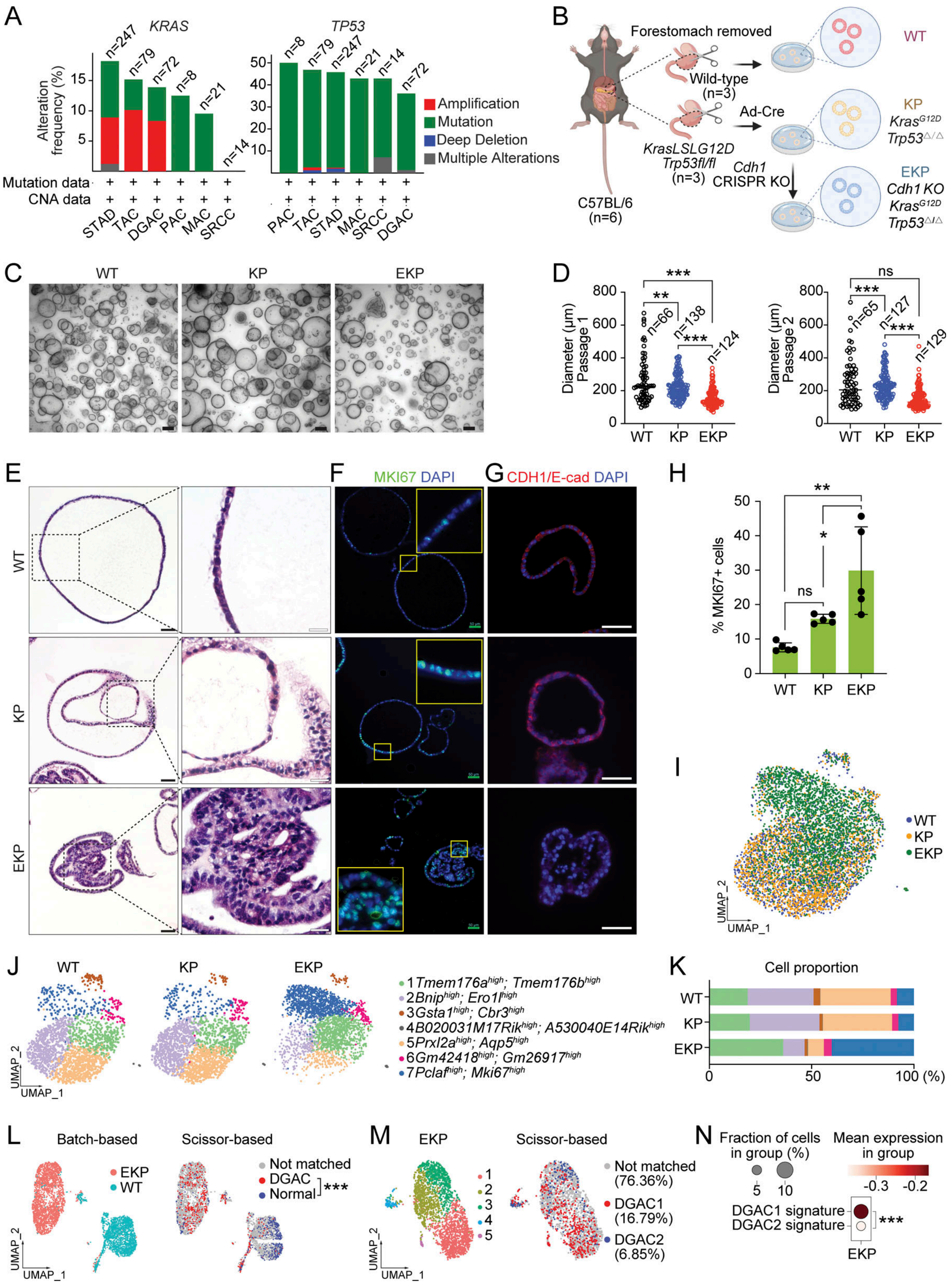


Figure 3. **Establishment of genetically engineered gastric organoids with *Cdh1*-inactivation. (A)** Genetic alteration of the *KRAS*, and *TP53* genes based on the cBioPortal. *n* represents the total number of patients enrolled in each subtype. DGAC, diffuse-type gastric adenocarcinoma; SRCC, signet ring cell carcinoma; TAC, tubular adenocarcinoma; STAD, stomach adenocarcinoma; MAC, mucinous adenocarcinoma; PAC, papillary adenocarcinoma. **(B)** Illustration of the workflow for stomach tissue collection and dissociation, gene manipulation of the gastric organoids (GOs), GOs culture, and representative image of GOs. Three GO lines were generated, including WT, KP, and EKP. WT mice and KP mice were sacrificed to collect stomach tissue. After removing the forestomach, stomach tissue was dissociated into a single cell and culture as organoids. Adeno-Cre virus was used to treat $Kras^{LSL\text{-}G12D}$; $Trp53^{fl/fl}$ organoids to generate KP organoids, followed by nutlin-3 selection. After selection, EKP organoids were generated using CRISPR-mediated *Cdh1* KO from KP GOs. **(C)** Representative images of WT, KP, and EKP GOs at passage day 8. Scale bars: 200 μm. **(D)** Growth analysis for WT, KP, and EKP GOs in two passages on day 8 of each passage.

P values were calculated using the one-way ANOVA; error bars: SD. Numbers below each label represent the number of organoids. **(E)** H&E staining of WT, KP, and EKP GOs. Scale bars: 50 µm (left panels); 200 µm (right panels). **(F)** MKI67 staining of WT, KP, and EKP GOs ($n = 5$). Scale bars: 50 µm. **(G)** CDH1 staining of WT, KP, and EKP GOs ($n = 5$). Scale bars: 50 µm. **(H)** Statistics analysis of MKI67 staining (Fig. 3 F). P values were calculated using the one-way ANOVA; error bars: SD. The representative images are shown. **(I)** Batch-based UMAPs of WT, KP, and EKP GOs. The Harmony integration package was used to remove the batch effect. **(J)** Leiden-based clustering UMAPs of WT, KP, and EKP GOs. Cell clusters were named by the most highly variable genes. **(K)** Cell proportion analysis of WT, KP, and EKP GOs. Each color represents a different cell type. The color code is based on the cell types shown in Fig. 3 J. **(L)** Batch-based and Scissor-based UMAP of WT and EKP GOs generated by the Scissor package. TCGA datasets of normal stomach and DGAC patients were utilized. **(M)** Cluster-based and Scissor-based UMAP of EKP GOs generated by Scissor package. DGAC1 and DGAC2 datasets were utilized to perform the comparison. **(N)** Dot plots of EKP GOs of DGAC1 and DGAC2 molecular signatures. The top 50 highly variable genes were used to calculate the molecular signature of each DGAC subtype. Gene list is shown in Table S8. P values were calculated by Mann–Whitney testing. ns: $P > 0.05$; *: $P \le 0.05$; **: $P \le 0.01$; ***: $P \le 0.001$. All data are derived from two or more independent experiments with the indicated number of mice.

separated UMAP, we observed that WT and KP shared somewhat similar transcriptional landscapes. However, EKP exhibited distinct features with an increased cluster 5 (Fig. 5 B). To pinpoint essential regulons, we created an unbiased workflow (Fig. 5 C). Based on the Z score of each regulon, we identified 32 regulons specific to the EKP transcriptional profile compared with those of WT and KO (Fig. 5 D). Additionally, the regulon specificity score (RSS) analysis showed the top 20 regulons specific to EKP (Fig. 5 E). RSS-based top 20 regulons belonged to Z score-based regulons (Fig. 5 F and Table S18). Both RSS and Z-score were used to quantify the activity of a gene or set of genes. Z-score was used to quantify the level of gene expression in a particular sample, while RSS was used to quantify the specificity of a gene set to a particular regulatory network or module (Kelley et al., 2016). According to TCGA-based upregulation in DGAC patients compared with normal stomach tissues, 13 regulons (Brca1, E2f1, E2f3, E2f7, E2f8, Ezh2, Gabpa, Gtf2b, Gtf2f1, Hmga2, Pole4, Sox4, and Tfdp1) were selected (Fig. S5 A). Next, we examined the regulons' expression in organoid datasets. Compared with WT and KP, the expression of Ezh2, Gtf2b, Pole4, and Sox4 was obviously increased in EKP GOs with over 40% fractions of clusters (Fig. 5 G). According to the regulon activity-based UMAP, Ezh2 displayed the highest score in EKP compared with WT and KP GOs (Fig. 5 H and Fig. S5 B). To assess the pathological relevance of EZH2 to DGAC, we analyzed the expression of downstream target genes of EZH2 in the DGAC datasets (Table S9) (Yu et al., 2023). One gene list included genes that were downregulated by EZH2 activation through histone modification (EZH2_histone_modification_down) (Fig. 5 I, Fig. S5 C, and Table S9), and the other gene list included genes that were downregulated by EZH2 activation that were reported in gastric cancer (EZH2_activation_down_in_GC) (Fig. 5 J, Fig. S5 D, and Table S9). Compared with DGAC2 (*CDH1* high), the EZH2_histone_modification_down and EZH2_activation_down_in_GC scores were relatively lower in DGAC1 (*CDH1* low) (Fig. 5, I and J). EZH2 is a histone methyltransferase catalyzing the methylation of histone H3 lysine 27 (H3K27) to generate H3K27me3, which is associated with gene repression (Lee et al., 2007). Consistent with *Ezh2* regulon activation by *Cdh1* KO, H3K27Me3 was also increased in EKP tumors compared with KP while there was no significant difference in H3K27Ac expression (Fig. 5 K). Next, we treated EKP cells with GSK343, a specific inhibitor of EZH2 methyltransferase (Verma et al., 2012). EKP cells were more sensitive to GSK343 than KP in cell proliferation in vitro (Fig. 5 L). Meanwhile, we conducted experiments to evaluate the effect of GSK343 on KP and EKP organoids (Fig. 5 M). We observed that the number of EKP GOs was significantly decreased (30.8%) after GSK343 treatment, while the number of KP GOs was marginally affected by GSK343 (92.6%) of the organoids initially seeded (Fig. 5 N). Additionally, allograft transplantation experiments showed the growth inhibitory effect of GSK343 on EKP tumorigenesis (Fig. 5, O–Q). These results identify Ezh2 as a key regulon contributing to tumorigenesis of E-cadherin inactivation-associated DGAC.

Discussion

The impact of CDH1/E-cadherin loss on sporadic DGAC tumorigenesis remains unknown. Single-cell transcriptomics-based unsupervised clustering identified two subtypes of DGAC: DGAC1 (CDH1-negative or downregulated) and DGAC2 (CDH1 high). Unlike DGAC2 lacking ascites tumor cells-associated immunologic response, the DGAC1 subtype is enriched with exhausted T cells. Single-cell transcriptomics and transplantation assays showed that *Cdh1* KO, in conjunction with *Trp53* KO and $Kras^{G12D}$, induces accelerated tumorigenesis and immune evasion. Moreover, Ezh2 regulon specifically activated by E-cadherin loss promotes DGAC tumorigenesis.

Patient stratification is crucial for improving therapeutic efficacy. Despite several studies classifying GAC patients (Ge et al., 2018; Fukamachi et al., 2019; Tong et al., 2019; Kim et al., 2020; Wang et al., 2020b), such subtyping did not consider single-cell level cellular convolution, which might be insufficient to represent the full spectrum of DGAC features. Our stratification approach was based on the high dimensional transcriptional signatures at the single-cell level, immune cell profiling, and cellular network, which may complement limitations from the bulk analyses and likely better stratify DGAC patients. Through scRNA-seq, we identified two distinct subtypes of DGAC. Molecular signatures for DGAC1 and DGAC2 were determined utilizing the top 50 highly variable genes (Fig. 1 M and Table S8). While single-cell transcriptomics emerges as a promising tool for stratifying cancer subtypes, running and analyzing scRNA-seq for every patient tumor sample is currently less practical. Therefore, further investigations and clinical validation of molecular signatures from scRNA-seq may reveal vital biomarkers for each DGAC subtype. For instance, discs large homolog 1 (*DLG1*) and fibroblast growth factor receptor substrate 2 (*FRS2*) are specifically expressed in DGAC1, whereas mucin 1 (*MUC1*) and keratin 7 (*KRT7*) are

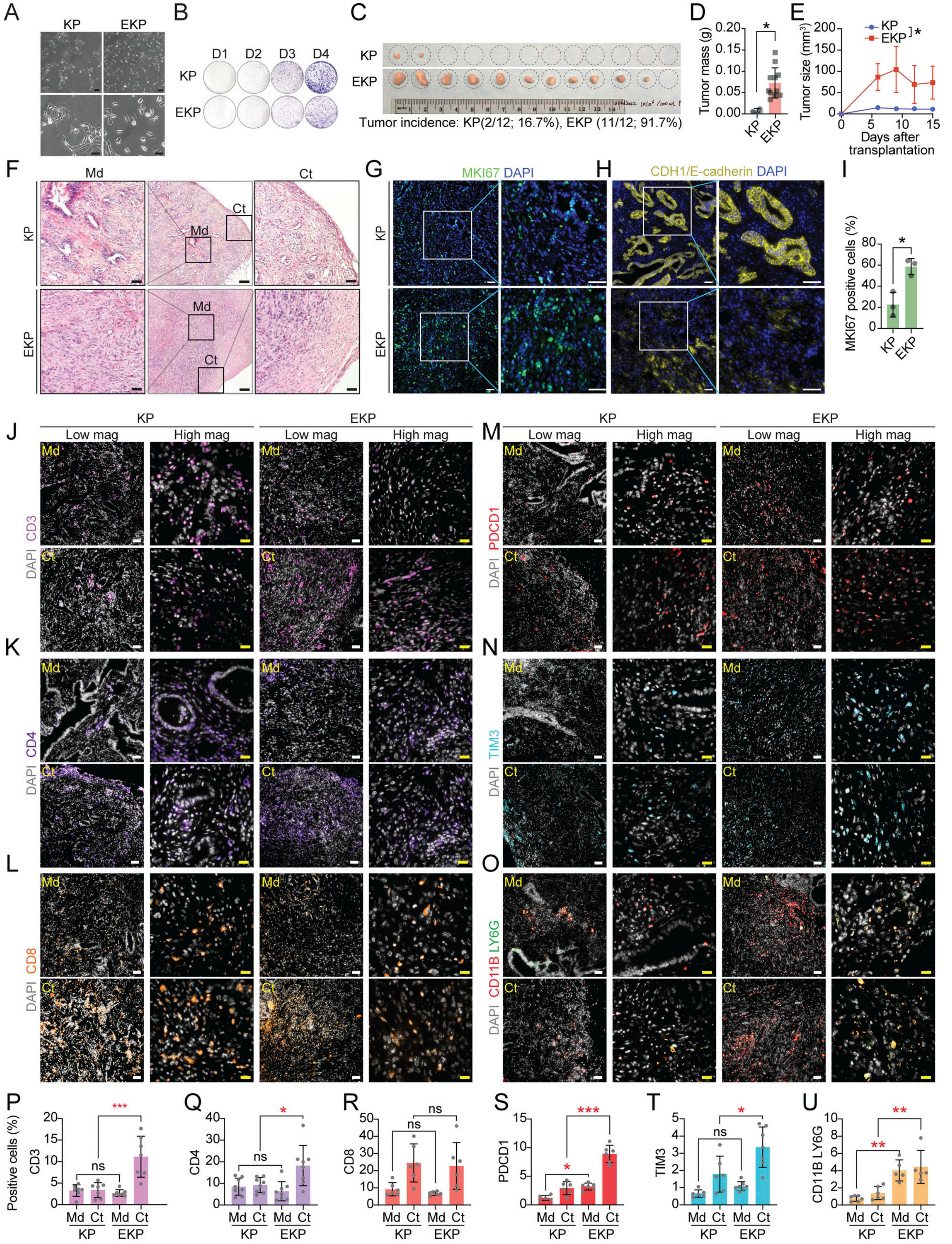


Figure 4. ***Cdh1* KO promotes KP-driven gastric tumorigenesis. (A)** Bright-field images of KP and EKP cells in low and high magnification. Scale bars: 100 µm (upper panels); 50 µm (lower panels). **(B)** Crystal violet staining of KP and EKP GOs-derived cells. **(C)** Bright-field images of KP and EKP allograft tumors; tumor incidence of allograft tumors. **(D and E)** Plot for tumor mass (D) and tumor size (E) assessment of KP and EKP allografts. **(F)** H&E staining of KP and EKP allograft tumors (*n* ≥ 3). Scale bars (from left to right): 50, 200, 50 µm; Md: middle; Ct: cortex. **(G and H)** Immunostaining of KP and EKP allograft tumors (*n* ≥ 3)

for MKI67 (G) and E-cadherin (H). Scale bars: 50 μm. **(I)** Statistics analysis of MKi67 staining in Fig. 4 G. P values were calculated using Student's *t* test; error bars: SD. **(J–O)** CD3 (J), CD4 (K), CD8 (L), PDCD1 (M), TIM3 (N) staining and CD11B/LY6G co-staining (O) of KP and EKP allograft tumors ($n \geq 3$). Left panels (low magnification [low mag]; right panels (high magnification [high mag]). Scale bars: 50 μm (low mag) and 20 μm (high mag). **(P–U)** Statistics analysis of CD3 (P), CD4 (Q), CD8 (R), PDCD1 (S), TIM3 (T) staining and CD11B/LY6G co-staining (U). The positive cell percentage indicates the area of cells expressing a specific marker divided by the total field-occupied cells stained by DAPI in the same area, which allows for normalization. Md: middle; Ct: cortex. P values were calculated using the one-way ANOVA; error bars: SD. ns: $P > 0.05$; *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$. All data are derived from two or more independent experiments with the indicated number of mice.

---

enriched in DGAC2 (Table S8). These potential markers exhibit positive correlations with cancer progression (Zhang et al., 2013; Huang et al., 2016, 2020; Chen et al., 2021; Kim et al., 2023). Such marker-based identification holds promise for guiding personalized treatment strategies in further clinical interventions. This study might also distinguish IGAC patients with *CDH1* deficiency not solely based on clinical history and histology but also by leveraging identified biomarkers, which still requires comprehensive analysis across a substantial cohort comprising DGAC and IGAC patient samples. This exploration of DGAC signatures may open a new avenue for pathological diagnosis, patient stratification, and therapeutic guidance. Except for analyzing tumor parenchymal cells, our unsupervised subtyping by tumor cell transcriptome matched well with distinct immune cell properties (Fig. 2, A–C). Furthermore, the application of CellChat and GSEA analysis led to the identification of T cell-related immune profiling as the dominant feature in DGAC1 (Fig. 2, D–F; and Fig. S3, A–F). Interestingly, T cell exhaustion and immune checkpoint-related genes were notably enriched in DGAC1 compared with DGAC2 (Fig. 2, G–L), confirmed by the transplantation experiments (Fig. 4). These results strongly suggest that DGAC1 patients might benefit from T cell-based ICIs, whereas DGAC2 patients might be ICI non-responders (Fig. 2).

Understanding the biology of cancer immune evasion is also imperative for improving cancer treatment. To date, how DGAC tumor cells evade immune surveillance remains elusive. Our transplantation assays showed that in conjunction with *Trp53* KO and $Kras^{G12D}$, *Cdh1* loss is sufficient for immune evasion of DGAC (Fig. 4). In line with this, EKP allografts displayed increased expression of CD3, CD4, PDCD1, TIM3, and LY6G (Fig. 4, J–U), also identified as molecular signatures of DGAC1 (Fig. 2, G–O). These tantalizing results suggest a new role of *CDH1* in restricting the immune evasion of tumor cells beyond its canonical role in cell–cell adhesion.

Tumors are immunogenically categorized into "hot," "altered-excluded," "altered-immunosuppressed," and "cold" (Galon and Bruni, 2019). The terms hot and cold describe T cell-infiltrated inflamed tumors and non-infiltrated tumors, respectively (Galon et al., 2006). Altered-immunosuppressed tumors have few $CD8^+$ T cells, mainly at the tumor's periphery, with immune-suppressing cells like MDSCs and regulatory T cells. In altered-excluded immune tumors, $CD8^+$ T cells are absent, and the tumor microenvironment is dense and hypoxic, hindering immune cell survival (Galon and Bruni, 2019). Cold tumors, altered-immunosuppressed, or immune-excluded tumors, respond less favorably to ICIs and generally have a poorer prognosis compared with hot tumors, which tend to respond well to ICIs (Galon and Bruni, 2019; Lee and Ruppin, 2019). According to the immune profiling of EKP tumors (Fig. 4), which mimic DGAC1, it is highly probable that DGAC1 may correspond to hot or altered-immunosuppressed mixed tumors, while DGAC2 is likely to be classified as either cold tumors or altered-excluded immune tumors. Emerging evidence suggests that E-cadherin loss may be associated with an inflamed phenotype (Stodden et al., 2015; Kaneta et al., 2020), which might contribute to T cell exhaustion induced by chronic inflammation (Fang et al., 2022). E-cadherin encoded by *CDH1* is an adhesion molecule responsible for maintaining cell–cell interactions and tissue integrity. Loss of E-cadherin disrupts adherens junctions of tumor cells and subsequently disorganizes tumor architecture (Bruner and Derksen, 2018), which likely promotes immune cell infiltration. Apart from E-cadherin, *CD44* and *CD52* emerged among the top 100 differentially expressed genes in DGAC1 (Table S8). These cluster of differentiation (CD) antigens, initially characterized on leukocytes, play a crucial role in cellular identification and have demonstrated significance in solid tumor prognosis (Chan et al., 1988; Yin et al., 2016; Wang et al., 2020a). CD44, notably upregulated in cancer cell subpopulations, serves as a molecular marker for cancer stem cells and is implicated in recruiting immune cells to infection sites (Ponta et al., 2003; Yin et al., 2016). In gastric cancer, elevated CD44 expression stands as an independent prognostic factor linked to immune invasion, potentially serving as a prognostic biomarker for evaluating outcomes associated with gastric immune invasion (Hou et al., 2022). Additionally, a breast cancer study indicated that a high-expression CD52 group was enriched in immune-related pathways (Wang et al., 2020a). These collective findings propose potential scenarios where E-cadherin loss and the collaborative upregulation of CD44 and CD52 could jointly enhance immune evasion, which remains to be determined.

Previously, two distinct molecular subtypes of GAC were introduced: mesenchymal phenotype (MP) and epithelial phenotype (EP) (Oh et al., 2018; Wang et al., 2020b). Since its association with E-cadherin downregulation and EMT (Fig. S1 J), the DGAC1 subtype might belong to the MP subtype, which displays poor survival and chemotherapy resistance (Oh et al., 2018). Unlike DGAC1, DGAC2 does not show E-cadherin loss and EMT. DGAC exhibits frequent mutations in the *TP53*, *CDH1*, *RHOA*, *APC*, *CTNNB1*, *ARID1A*, *KMT2C*, and *PIK3CA* genes (Cancer Genome Atlas Research Network, 2014; Kakiuchi et al., 2014; Oliveira et al., 2015; Cho et al., 2019). Among these genes, *CDH1* and *RHOA* mutations are mainly observed in DGAC and not found in IGAC (Cancer Genome Atlas Research Network, 2014; Kakiuchi et al., 2014; Cho et al., 2019). As CDH1/E-cadherin and RHOA both play a crucial role in modulating the cytoskeleton, cell morphology, and cell migration (Handschuh et al., 1999; McBeath et al., 2004; O'Connor and Chen, 2013; Al-Ahmadie et al., 2016), the general histological features of DGAC are

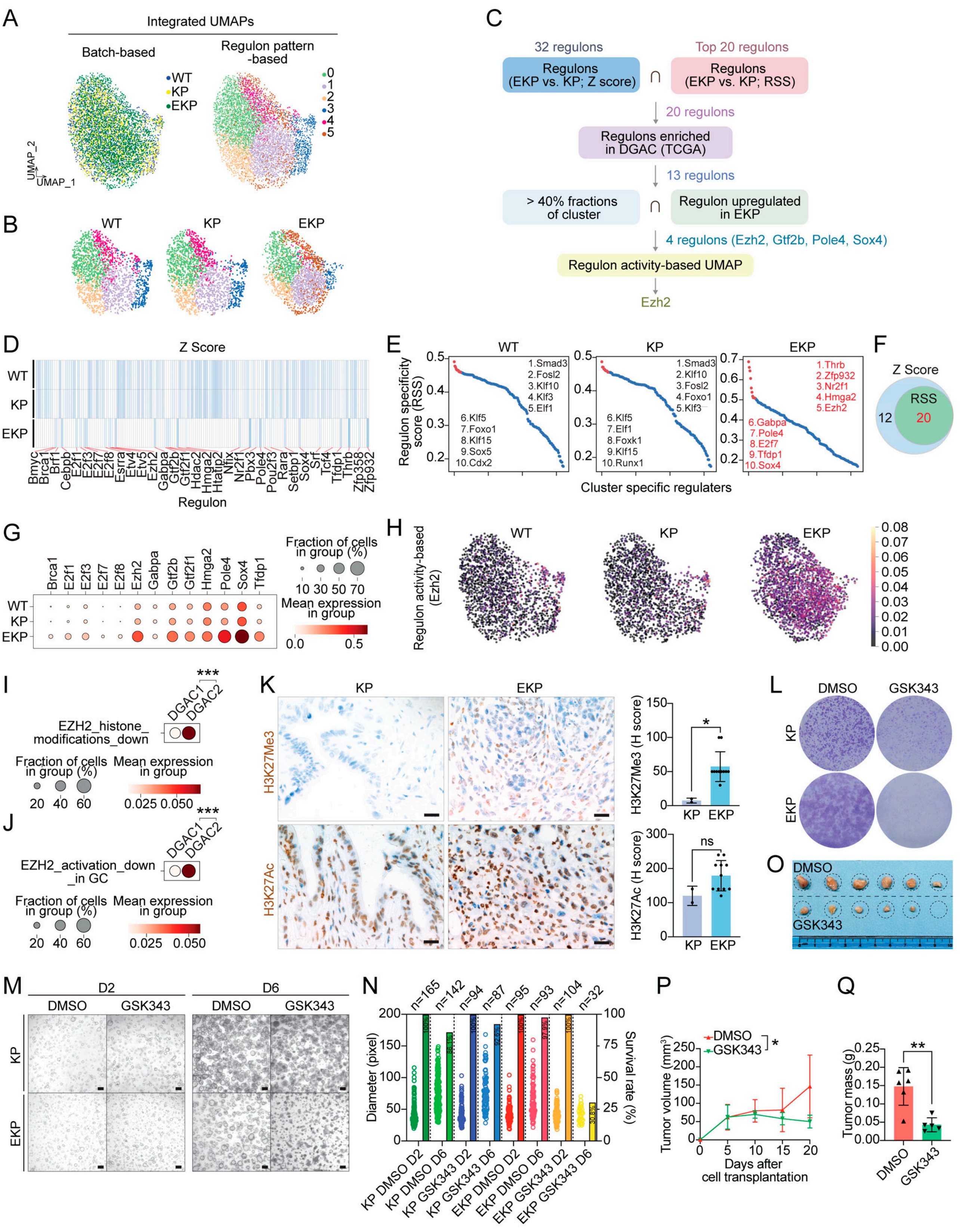


Figure 5. ***Cdh1*** **KO-activated** ***Ezh2*** **promotes gastric tumorigenesis. (A)** Integrated batch-based and regulon pattern-based UMAP for WT, KP, and EKP GOs. Six transcriptional modules were identified. **(B)** Separated regulon patterns based UMAP for WT, KP, and EKP GOs. **(C)** Flow chart of regulons selection process. **(D)** Regulons enriched in WT, KP, and EKP GOs, based on Z Score. 32 regulons were highly expressed in EKP samples compared to WT and KP. **(E)** Regulons enriched in WT, KP, and EKP GOs, based on RSS. The top 20 were selected by Z score. The whole regulon list based on RSS is shown in Table S18. **(F)** Venn diagram for the regulons from D and E. 20 regulons were overlapped. **(G)** Dot plot of the regulons (WT, KP, and EKP GOs) increased in TCGA DGAC

patients. **(H)** Regulon activity-based UMAP of *Ezh2* in WT, KP, and EKP GOs. The cells with lighter color represent regulated by *Ezh2*. **(I and J)** Dot plots of *Ezh2* downstream target genes (I, genes which are downregulated by *EZH2* activation through histone modification; J, genes which are downregulated by *EZH2* activation reported in gastric cancer) scores in the epithelial cells of DGAC1 and DGAC2. P values were calculated by using a Mann–Whitney testing. Gene list of *EZH2* targeted genes was listed in Table S9. **(K)** The level of H3K27Ac and H3K27Me3 expression in KP and EKP allografts. Quantification was displayed. Scale bars: 20 μm. **(L)** Crystal violet staining of KP and EKP cells after GSK343 (EZH2 inhibitor, 10 μM, 96 h). **(M)** Bright-field images of KP and EKP GOs after treatment with GSK343 (EZH2 inhibitor, 10 μM, 96 h). D2: day 2; D6: day 6. Scale bars: 200 μm. **(N)** Statistical analysis of KP and EKP gastric organoid size and number in response to GSK343 treatment. The number of organoids (right Y-axis) and their size (left Y-axis) were assessed following treatment with GSK343. On day 2 (D2), the number of organoids was determined for the image depicted in M, and this count was considered as 100% (*n* numbers are presented in the bubble plot). On day 6 (D6), the number of organoids in the same field for each group was counted (*n* numbers also displayed in the bubble plot). The percentage of each group on D6 was calculated by dividing the number of viable organoids at D6 by the number at D2. The viable percentage is presented in the bar graph. **(O–Q)** Transplantation of EKP cells followed by EZH2 inhibition. **(O)** Bright-field images of EKP allograft tumors treated with DMSO and GSK343 (20 mg/kg) separately ($n = 3$). **(P)** Tumor growth curve of EKP allografts treated with DMSO and GSK343 (20 mg/kg) after cell subcutaneous transplantation. **(Q)** Tumor mass of EKP allografts treated with DMSO and GSK343 (20 mg/kg) after mice scarification. P values were calculated using Student's *t* test; error bars: SD. ns: P > 0.05; *: P ≤ 0.05; **: P ≤ 0.01; ***: P ≤ 0.001. All data are derived from two or more independent experiments with the indicated number of samples.

likely attributed to mutations in these genes, *CDH1* and *RHOA* (Ooki and Yamaguchi, 2022). DGAC2 displays high *CDH1* expression and no EMT gene expression (Fig. 1, J–L and Fig. S1 J). Intriguingly, DGAC2 exhibits a relatively higher activation of RHOA signaling (Fig. S1 K). Thus, it is conceivable that epithelial cell polarity loss and the diffuse-type cell (morphological) phenotype in DGAC2 might be due to *RHOA* mutations (Y42C) or RHOA signaling activation, whereas *CDH1* loss is linked with DGAC1.

E-cadherin mediates cell–cell interaction via homophilic interaction with other E-cadherin proteins from neighboring cells. The cytoplasmic domain of E-cadherin is physically associated with Catenin proteins (α, β, γ, and p120) and actin cytoskeleton, which plays a pivotal role in maintaining epithelial cell polarity and integrity (McCrea and Park, 2007). In our scRNA-seq study on EKP GOs, loss of E-cadherin resulted in transcriptional reprogramming and altered cell proportions, specifically reducing the *Aqp5*^{high} cluster and increasing the *Mki67*^{high} cluster (Fig. 3, J and K). *AQP5* is specifically expressed in pyloric stem cells, as well as being frequently expressed in gastric cancers and their metastases (Tan et al., 2020). Meanwhile, *Aqp5* was expressed in a subpopulation of gastric cancer cells, some of which were KI67$^+$. Our unbiased scRNA-seq analysis distinctly revealed two separate clusters, namely *Aqp5*^{high} cells and *Mki67*^{high} cells (Fig. 3 J), denoting that the *Aqp5*^{high} cells within our murine GO model are not proliferative. This finding aligns with Barker's study, wherein some AQP5$^+$ cells were found to be KI67 negative (Tan et al., 2020). Remarkably, in EKP organoids, we observed a reduction in *Aqp5*^{high} cells alongside an increase in *Mki67*^{high} cells (Fig. 3, J and K). Consistently, in EKP tumors, a higher proportion of proliferative cells was observed compared with KP tumors (Fig. 4, G and I). These outcomes suggest that *Mki67*^{high} cells might represent cells-of-origin in EKP tumors, characterizing the DGAC1 subtype with *CDH1* loss. However, further rigorous experiments are warranted to validate this observation.

Intriguingly, E-cadherin loss activates Ezh2 regulon, and EZH2 blockade suppresses EKP tumor growth (Fig. 5). EZH2 modulates gene expression in various ways: gene repression via polycomb repressive complex 2 (PRC2)-dependent histone methylation, PRC2-dependent non-histone protein methylation, or gene activation via transcriptional activator complex. The detailed mechanisms of how EZH2 is engaged in E-cadherin loss-associated DGAC tumorigenesis remain to be determined. Nonetheless, given that an EZH2 inhibitor (tazemetostat) is clinically available, targeting *EZH2* would be a viable option for the DGAC1 subtype in addition to T cell-based ICIs. The use of epigenetic modulators has been found to enhance the infiltration of effector T cells, suppress tumor progression, and improve the therapeutic effectiveness of PD-L1 checkpoint blockade in prostate or head and neck cancer (Jadhav et al., 2019; Weber et al., 2021). Additionally, pharmacological inhibition of EZH2 has been shown to inhibit tumor growth and enhance the efficacy of anti-CTLA-4 treatment in bladder cancer (Wherry and Kurachi, 2015). Given the enriched expression of immune checkpoints in DGAC1 (Fig. 2, J–L), a combination therapy involving EZH2 inhibitors and ICIs may hold potential benefits for DGAC1 patients. Nonetheless, how E-cadherin loss activates EZH2 regulon remains to be determined.

While our analysis offers insights from malignant ascites of DGACs, it is crucial to recognize the constraint of not directly examining primary tumors. Additional investigations are necessary to validate the applicability of DGAC1/2 signatures to epithelial cells in primary tumors. Concurrently, limitations of scRNA-seq include relatively shallow sequencing depth and restricted information not overcoming intratumoral heterogeneity. Thus, increasing the number of scRNA-seq datasets and spatial transcriptomics should follow in future studies. Furthermore, although this is the first stratification of DGAC by single-cell transcriptome, the pathological relevance of E-cadherin status (or alternative molecular signatures; Fig. 1 M) with ICI response remains to be clinically demonstrated. Meanwhile, we have identified a discrepancy between *CDH1* genetic alterations, transcriptional downregulation, and E-cadherin protein downregulation (Fig. 1, B, C, and L). Notably, our analysis revealed lower *CDH1* expression in the DGAC1 group and higher *CDH1* expression in the DGAC2 group (Fig. 1 L). Within DGAC1, a subset of patients displayed E-cadherin expression while others did not (Fig. 1 L). Among the 114 DGAC patient tumor samples we analyzed, 37.7% exhibited low E-cadherin levels (Fig. 1, B and C). This subset of tumor samples might be classified into DGAC1 or DGAC2 based on the differential expression of other signature genes specific to each cluster rather than relying solely on E-cadherin expression. To address this and further validate our findings, future studies should incorporate both IHC and scRNA-seq using the same patient samples.

Together, our study stratifies DGAC patients by integrative single-cell transcriptomics with experimental validation and unravels an unexpected role of E-cadherin in restricting transcriptional reprogramming and immune evasion of DGAC, which provides new insight into the biology of DGAC tumorigenesis and helps improve immunotherapy efficacy.

Materials and methods

Mice

All mouse experiments were approved by the MD Anderson Institutional Animal Care and Use Committee and performed under MD Anderson guidelines and the Association for Assessment and Accreditation of Laboratory Animal Care international standards. Compound transgenic mice $Kras^{LSL\text{-}G12D/+}$; $Trp53^{fl/fl}$ (KP) mice have been previously described (Kim et al., 2021). C57BL/6 mice were purchased from The Jackson Laboratory.

GOs generation

The protocol for generating GOs was previously described (Bartfeld et al., 2015). The mice were sacrificed, their stomachs were collected, and the forestomachs were removed. Then, the reserved stomach tissue was cut through the lesser curvature, and the stomach was rinsed with ice-cold PBS with 1% penicillin/streptomycin to remove blood. The tissue samples were carefully immersed in chelating buffer (sterile distilled water with 5.6 mmol/liter Na_2HPO_4, 8.0 mmol/liter KH_2PO_4, 96.2 mmol/liter NaCl, 1.6 mmol/liter KCl, 43.4 mmol/liter sucrose, 54.9 mmol/liter D-sorbitol, and 0.5 mmol/liter DL-dithiothreitol, pH 7) in a 10-cm dish and then the tissue was transferred to a dry dish. The epithelial layer was peeled and minced into pieces using forceps. Minced epithelial pieces were placed into 10 ml cold chelating buffer, followed by robust pipetting up and down to rinse the tissue until the supernatant was clear. A 20 ml chelating buffer was prepared with 10 mM EDTA under room temperature and the tissue was incubated in it for 10 min. The tissue was tenderly pipetted gently once up and down and the pieces were allowed to settle. The tissue was then moved to the clean bench. Most of the water was removed, and the tissue pieces were carefully placed in the middle of a sterile 10-cm dish. A glass microscopy slide was put on top of the tissue and pressure was added on the slide until the tissue pieces seemed cloudy. The cloudy tissue pieces were then flushed from the slides in 30 ml of cold Advanced DMEM/F12. The large tissue fragments were allowed to sediment by gravity. The cloudy supernatant was transferred to two 15-ml tubes. The tubes were then centrifuged for 5 min at 200 *g* and 4°C. The supernatant was carefully removed and resuspended with a Matrigel-medium mixture (12 µl Matrigel mix with 8 µl GOs culture medium/well). Approximately 40 glands per 20 µl Matrigel-medium mixture per well of a 48-well plate were seeded. The plate was steadily transferred to the incubator to let it solidify for 10 min. Then, 500 µl of GOs culture medium was added to cover the dome and the plate was incubated at 37°C with 5% CO_2. The medium was changed every 2 days.

GOs culture

Table S1 was referred to for the culture medium ingredient. The organoids were passaged using the following steps: (1) the culture medium was discarded. (2) The matrigel was scraped with a pipette tip and dissociated by pipetting. (3) The organoids were collected from three wells (48-well) in a 15-ml tube with a cold medium. (4) The supernatant was discarded after centrifugation at 1,000 RPM and 4°C. (5) The dissociated organoids were washed with 13 ml of cold 1× PBS, centrifuged (1,000 RPM, 4 min) and the supernatant was removed. (6) The organoids were resuspended in 1 ml of Trypsin-EDTA (0.05%). (7) The sample was transferred to a 1.7 ml Eppendorf tube and then pipetted up and down. (8) The sample was incubated at 37°C in a 5% CO_2 incubator for 30–45 min. (9) The tube was vibrated every 10 min. (10) The organoid structure was further broken down by pipetting up and down. (11) The sample was checked under a microscope to ensure the organoids digested into the cells. (12) The sample was passed through the 35-µm cell strainer. (13) Trypsin was inactivated with 10% FBS medium and pipetted vigorously. (14) The sample was collected in a 15 ml tube and centrifuged for 4 min at 1,000 RPM. (15) The supernatant was aspirated and the cells were resuspended with GOs culture medium. (16) The cells were counted, viability was checked, and the appropriate number of cells was calculated. (17) Every 8 µl of cell suspension was mixed with 12 µl of Matrigel as a mixture and seeded in the 48-well plate. (18) The plate was transferred to the incubator and allowed to solidify for 10 min. (19) 500 µl of GOs culture medium was added to cover the dome and incubated at 37°C with 5% CO_2. (20) The medium was changed every 2 days.

The organoids were cryopreserved as follows: The organoids were dissociated following the above organoid passaging (steps 1–15) protocol. The cells were then added with 10% volume of DMSO and transferred to the cryovials.

CRISPR/Cas9-based gene knockout in GOs

KO of *Cdh1* was performed by CRISPR/Cas9 genome editing using pLentiCRISPRv2 (plasmid #52961; Addgene) according to Zhang laboratory's protocol (Ran et al., 2013). Five single guide RNA (sgRNA) targeting *Cdh1* were designed using CRISPick (https://portals.broadinstitute.org/gppx/crispick/public) and cloned into a pLentiCRISPRv2-puro vector. An empty sgRNA vector was used as a negative control.

The five targeting sequences against *Cdh1* were: #1: 5′-ATGATGAAAACGCCAACGGG-3′, #2: 5′-ACCCCCAAGTACGTACGCGG-3′, #3: 5′-TTACCCTACATACACTCTGG-3′, #4: 5′-AGGGACAAGAGACCCCTCAA-3′, and #5: 5′-CCCTCCAAATCCGATACCTG-3′. sgRNA (5′-ATGATGAAAACGCCAACGGG-3′) was used to target the *Cdh1* allele in GOs. See Table S2 for the primer sequence to validate *Cdh1* knockout efficiency.

Lentivirus production and transduction

The HEK293T cells were cotransfected with 5 µg of constructs, 5 µg of plasmid Δ8.2 (Plasmid #8455; Addgene), and 3 µg of plasmid VSVG (Plasmid #8454; Addgene) in a 10-cm dish. The cells were incubated at 37°C and the medium was replaced after 12 h. The virus-containing medium was collected 48 h after transfection. The organoids were dissociated following the organoid passaging protocol (steps 1–14), and the supernatant was aspirated, leaving the pellet. For transduction, 20 μl of cell suspension was used. The amount of polybrene (8 μg/ml) was calculated and mixed with virus-containing medium before adding to the cells. The polybrene-containing virus medium was added to the cell pellet and the cell suspension was transferred to a 1.7 ml Eppendorf tube. The tube was centrifuged at 600 *g* at 37°C for 1 h. Without disturbing the cell pellet, the tube was incubated in the 37°C incubator for 4 h. The supernatant was then removed and the cell pellet was resuspended with the required volume of GOs culture medium (8 μl for one well of 48-well plate) and placed on ice for cool down. The appropriate volume of pre-thawed Matrigel (12 μl for one well of a 48-well plate) was added to the tube and the dome was seeded in the middle of a 48-well plate. The plate was then incubated for 10 min at 37°C with 5% CO_2. GOs culture medium was added to the well. After 48 h, the infected organoids were selected with 1 μg/ml puromycin.

Adenovirus transduction

We used Adeno-Cre virus to treat *Kras*$^{LSL\text{-}G12D/+}$; *Trp53*$^{fl/fl}$ organoids. The protocol was previously described (Ko et al., 2022, 2023). The cells were first dissociated from GOs as described in the organoid passaging protocol (steps 1–14). The cell number was counted and the ratio of adenovirus: organoid cell was 1,000 PFU/μl:1 cell. The cell suspension, virus-containing medium, and Matrigel were mixed, and the drop was placed in the center of the well. The cell suspension and virus-containing medium were mixed before adding GOs culture medium up to 8 μl. Then, 12 μl of Matrigel was added to the mixture on ice. The plate was incubated in the 37°C cell culture incubator for 15 min to allow the Matrigel to solidify. After 48 h, the infected organoids were treated with 10 μM Nutlin-3 to select *Trp53* KO organoids. The primer sequence to validate *Trp53* KO and *Kras*$^{G12D/+}$ can be found in Table S2.

Organoid imaging and size measurement

After 7 days of organoid seeding in Matrigel, the size of the organoids was analyzed by measuring the volume under the microscope (ZEN software, ZEISS). To reduce the vulnerability of GOs, the measurements were conducted more than three passages after isolation from the knockout experiments. All experiments included >50 organoids per group.

Tissue microarray

DGAC cancer tissue microarray slides contained 114 patients' samples. Patients' information is shown in Table S4.

Histology and immunohistochemistry

All staining was performed as previously described (Jung et al., 2018). For organoids staining, 7 days after seeding, GOs were collected by dissociating the Matrigel mixture using ice-cold PBS and fixed in 4% paraformaldehyde at room temperature. For tumor tissue, excised tumors were washed with ice-cold PBS and fixed with formaldehyde at room temperature. After paraffin embedding, tumor tissue and organoid sections were mounted on microscope slides. For H&E staining, sections were incubated in hematoxylin for 3–5 min and eosin for 20–40 s. After washing with tap water, the slides were dehydrated and the coverslips were mounted with mounting media. For immunofluorescence staining, after blocking with 5% goat serum in PBS for 1 h at room temperature, sections were incubated with primary antibodies (MKI67 [1:200], CDH1 [1:200], CD3 [1:200], CD8 [1:200], CD4 [1:200], PDCD1 [1:200], TIM3 [1:200], CD11B [1:200], LY6G [1:200]) overnight at 4°C and secondary antibody (1:250) for 1 h at room temperature in dark. Sections were mounted with ProLong Gold antifade reagent with DAPI (Invitrogen). For immunohistochemistry staining, after blocking with 5% goat serum in PBS for 1 h at room temperature, sections were incubated with primary antibodies (CDH1 [1:200], H3K27Me3 [1:200], H3K27Ac [1:200]) overnight at 4°C and secondary antibody (1:250) for 1 h at room temperature in dark. The slides in the DAB solution were incubated until the tissue became brown and the background was still white, and was observed under the microscope until the strongest signal was shown, and the reaction was stopped with a tap water wash. The same incubation time was used for the same antibody on different slides. Sections were incubated in hematoxylin for 3–5 min and mounted with mounting media. Images were captured with the fluorescence microscope (Zeiss; AxioVision). See Table S3 for antibody information.

2D culture

The organoids were dissociated following the organoid passaging protocol (steps 1–14). The supernatant was aspirated and then resuspended with DMEM + 10% FBS with 10 μM Y-27632 and the organoids were seeded on a 24-well plate. Cells were passaged every 3–5 days. After the third passage, Y-27632 was removed from the culture medium. DMEM was supplemented with 10% FBS and 10% DMSO was used to freeze cells and store them in liquid nitrogen.

Allograft transplantation

5-wk-old C57BL/6 mice were maintained in the Division of Laboratory Animal Resources facility at MD Anderson. 2D-cultured KP and EKP cells (1×10^6) were injected subcutaneously into both flanks of mice. Tumor volume was calculated by measuring with calipers every 3–4 days (volume = [length × width^2]/2). Mice were euthanized and tumors were collected on day 15. The excised tumors were photographed and paraffin-embedded for immunostaining. For GSK343 treatment, 2D-cultured EKP cells (1×10^6) were injected subcutaneously into both flanks of mice. After the tumors were palpable, we performed the first measurement with calipers. We divided the mice into two groups of three mice each and administered DMSO and GSK343 (20 mg/kg) intraperitoneally every other day. The initial tumor volumes between the two groups were comparable. Tumor volume was calculated by measuring with calipers every 3–4 days (volume = [length × width^2]/2). Mice were euthanized and tumors were collected on day 20.

Crystal violet staining

Cells (1×10^3) were seeded on 6-well plates and the medium was replaced every 2 days. Plates were rinsed with 1× PBS, fixed with

4% paraformaldehyde solution for 20 min, and stained with crystal violet solution (0.1% crystal violet, 10% methanol) for 20 min, followed by rinsing with tap water.

GOs library preparation for scRNA-seq

For scRNA-seq, organoids from WT, KP, and EKP were collected 7 days after seeding, and the organoid passaging (step 1–14) protocol was followed. After trypsin had been inactivated with 10% FBS DMEM, a single-cell suspension was collected by passing cells through a 70-µm cell strainer and followed by a 40-µm cell strainer. Each group was tagged with two CMO tags from the CellPlex kit (10x Genomics). The tagged cells of each group were pooled together with the same number of cells after being counted. Single-cell Gene Expression Library was prepared according to Chromium Single Cell Gene Expression 3v3.1 kit with Feature Barcode technology for cell Multiplexing (10x Genomics). In brief, tagged single cells, reverse transcription (RT) reagents, gel beads containing barcoded oligonucleotides, and oil were loaded on a Chromium controller (10x Genomics) to generate single-cell GEMS (Gel Beads-In-Emulsions). Incubation of the GEM produced barcoded, full-length cDNA as well as barcoded DNA from the cell multiplexing. Subsequently, the GEMS were broken and pooled. Following cleanup using Dynabeads MyOne Silane Beads, full-length cDNA was amplified by PCR for library prep through fragmentation, end-repair, A-tailing, adaptor ligation, and amplification, while the barcoded DNA from the cell multiplexing was amplified for library prep via PCR to add sequencing primers. The cDNA library was sequenced on an Illumina NovaSeq platform (Novogene), mapped to the GRCm38/mm10 genome, and demultiplexed using CellRanger. The resulting count matrices files were analyzed in R (Seurat) or Python (Scanpy).

scRNA-seq—raw data processing, clustering, and annotation

We used Cell Ranger to perform demultiplexing and reads alignment of sequencing raw data for the scRNA-seq matrices generation. Ambient RNA and doublets were removed by SoupX (Young and Behjati, 2020) and Scrublet (Wolock et al., 2019), respectively. Scanpy (Wolf et al., 2018) was used for processing the scRNA-seq data. For the organoid dataset, cells with <50 genes expressed and >30% mitochondrial reads, 30% rpl reads, and 25% rps reads were removed. Genes expressed in fewer than five cells were removed. Then we normalized and log-transformed the gene expression for each cell. The percentages of mitochondrial reads, rpl reads, and rps reads were regressed before scaling the data. We reduced dimensionality and clustered the cells by Leiden (resolution = 0.5). Cell lineages were annotated based on algorithmically defined marker gene expression for each cluster (sc.tl.rank_genes_groups, method = "*t*-test"). See Table S17, the most differentially expressed 100 genes of each cluster were listed. For the DGAC dataset, cells with <100 genes expressed and more than 80% mitochondrial reads, 30% rpl reads, and 25% rps reads were removed. Genes expressed in <25 cells were removed. Normalization, log-transformation, regression, dimensionality reduction, and Leiden clustering (resolution = 1) were used the same way we used in organoids. Cell lineages were annotated based on algorithmically defined marker gene expression for each cluster (sc.tl.rank_genes_groups, method = "*t* test"). See Tables S6, S7, and S8 for details; the most differentially expressed 100 genes of each cluster or type are listed. For the DGAC dataset integrated with the normal stomach dataset, cells with <100 genes expressed and more than 100% mitochondrial reads, 40% rpl reads, and 30% rps reads were removed. Genes expressed in <25 cells were removed. Normalization, log-transformation, regression, dimensionality reduction, and Leiden clustering (resolution = 1) were used the same way we used in organoids. Cell lineages were annotated based on algorithmically defined marker gene expression for each cluster (sc.tl.rank_genes_groups, method = "*t*-test"). See Table S10 for details; the most differentially expressed 100 genes of each cluster are listed. More information about the software and algorithms used in this study is shown in Table S19.

Proportion difference analysis

The cell number of each cluster were retrieved by Scanpy (adata.obs["Leiden"].value_counts()). We analyzed and plotted the differences between clusters from the two datasets using GraphPad Prism 9.4. Then we grouped each cell cluster from the integrated dataset and compared the cluster differences between the two datasets.

Regulon analysis

For the gene regulatory network inference in organoids, we used the pySCENIC package (Van de Sande et al., 2020) to compute the specific regulons for each cell cluster. The Loom file of each organoid dataset was used and the regulon pattern-based UMAP was redrawn based on the AUCell scoring method (Aibar et al., 2017). RSS (Suo et al., 2018) and Z score were used to determine how specific the regulon is for one certain cell cluster. More specific the regulon is, the higher the RSS or Z score is for one certain cluster. Following the criteria that RSS and Z score should be high at the same time, we identified 20 regulons that were specific to EKP. These processes were repeated five times in each organoid dataset (WT, KP, and EKP). RSS of regulons from each mouse GO dataset (WT, KP, and EKP) is listed in Table S18.

Scissor analysis

To determine the pathology of murine organoids, we compared the transcriptomic similarity of the organoids scRNA-seq dataset and the bulk RNA-seq datasets of DGAC patients by Scissor package (Sun et al., 2022). The RNA-seq data of the tumor and the adjacent normal samples of DGAC patients were downloaded from the GDC data portal (TCGA-STAD). The murine genes were converted to human homologs by biomaRt. The Scissor analysis was performed by using the Cox regression model (α = 0.32). The goal of Scissor is to identify a small group of cells that are most highly correlated with the specific phenotypes with high confidence. Based on this motivation as a priori, we determined α using the following criteria: the number of Scissor-selected cells should not exceed a certain percentage of total cells (default 20%) in the single-cell data (Sun et al., 2022).

Cell–cell communication analysis

CellChat (Jin et al., 2021) package in R (https://www.r-project.org) was used to analyze the ligand–receptor interaction-based cell–cell communications in scRNA-seq datasets. The integrated dataset was processed, clustered, and annotated using the Scanpy package (Wolf et al., 2018) in Python and then transformed into.rds files. Transformed datasets were analyzed by CellChat with default parameters (P value threshold = 0.05).

Gene set score analysis

Gene set score was analyzed by Scanpy (Wolf et al., 2018) with the "scanpy.tl.score_genes" function (Wolf et al., 2018). The gene list for the score analysis is shown in Table S9. Statistical significance testing was performed by using Mann–Whitney testing (nonparametric testing).

Human scRNA-seq data analysis

The scRNA-seq data set of 19 DGAC patients' samples (patients information is shown in Table S5) has been previously reported from our group and the detailed clinical and histopathological characteristics are described (EGAS00001004443) (Wang et al., 2021). The metadata of the scRNAseq is presented in Table S5.

The scRNA-seq data set of the 29 normal adjacent stomachs (NCBI accession no. GSE150290) (Kim et al., 2022) was extracted from the Gene Expression Omnibus (GEO) database and analyzed with Scanpy and Python (Wolf et al., 2018). The 19 DGAC patients' datasets were integrated and clustered by Scanpy (Wolf et al., 2018) for the subclassification of DGACs based on CDH1 inactivation. The 19 DGAC patients' datasets and 29 normal adjacent stomachs were integrated and clustered in Scanpy (Wolf et al., 2018) for later infercnvpy analysis. Harmony (Korsunsky et al., 2019) algorithm was used to remove batch effects. Then, the dendrogram and correlation matrix heatmap were plotted with Scanpy (Wolf et al., 2018). The dendrogram shows the distance of each dataset based on principal component analysis, and the correlation matrix heatmap shows Pearson correlation by a color spectrum.

CNV analysis

To detect the genomic stability of groups DGAC1 and DGAC2, we performed CNVs inference from the gene expression data using the Python package infercnvpy (https://github.com/icbi-lab/infercnvpy). We performed infercnvpy on DGAC1 and DGAC2 using the Normal group (29 human normal adjacent stomachs) as the reference. The gene ordering file which contains the chromosomal start and end position for each gene was created from the human GRCh38 assembly. The GRCh38 genomic positions annotated file was downloaded from https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest. Infercnvpy was used to plot chromosome heatmap and CNV scores in the UMAP.

GSEA

GSEA was conducted via the R package fgsea (Korotkevich et al., 2021, *Preprint*) according to the DEG list generated by Scanpy. Normalized enrichment score (NES) represents the degree of enrichment of a gene set in a given dataset, measuring the coordinated upregulation or downregulation of genes within the set compared to a reference condition. It is normalized to account for variations in gene set size and dataset characteristics, providing a more robust measure of enrichment. The enrichment value was calculated and plotted with the fgsea package (permutation number = 2,000). All enriched pathways are listed in Tables S11, S12, S13, S14, S15, and S16.

Public sequencing database

All TCGA cancer patients' sequencing data referenced in this study were obtained from the TCGA database at cBioPortal Cancer Genomics (https://www.cbioportal.org).

Statistical analyses

GraphPad Prism 9.4 (Dogmatics) was used for statistical analyses. The Student's *t* test was used to compare two samples. To compare gene set score analysis, P values were calculated by Mann–Whitney testing. For the dot plot of a single gene, P values were calculated by *t* test. One-way ANOVA was used to compare multiple samples. P values <0.05 were considered statistically significant. Error bars indicate the standard deviation (SD) otherwise described in figure legends.

Online supplemental material

Fig. S1 provides a transcriptional, clinical, and molecular characterization of DGAC subtypes. Fig. S2 illustrates the scRNA-seq analysis of 19 DGAC patients and 29 adjacent normal stomach tissue. Fig. S3 presents GSEA analysis and the expression of macrophage polarization markers for DGAC1 and DGAC2. Fig. S4 demonstrates the validation of genetic engineering and scRNA-seq analysis of mouse GOs. Fig. S5 depicts the expression of EKP-specific regulons and EZH2 downstream targeted genes. Table S1 shows GO culture medium. Table S2 lists primers used for genotyping. Table S3 shows antibody information. Table S4 shows clinical information of 114 DGAC patient tumor samples for tissue microarray analysis. Table S5 shows scRNA-seq metadata and clinical information of 19 DGAC patient samples. Table S6 shows Leiden-based gene list of 19 DGAC samples scRNA-seq data (all cells included). Table S7 shows Leiden-based gene list of 19 DGAC scRNA-seq data (epithelial cells only). Table S8 shows type-based gene list of 19 DGAC scRNA-seq data (epithelial cells only). Table S9 shows the gene list used for gene set score analysis. Table S10 shows Leiden-based gene list of integrated 19 DGAC patients and 29 normal stomach scRNA-seq data (all cells included). Table S11 shows DGAC1-enriched pathway analysis using fGSEA with GO:BP dataset (reference: DGAC2). Table S12 shows DGAC1-enriched pathway analysis using fGSEA with CP:REACTOME dataset (reference: DGAC2). Table S13 shows DGAC1-enriched pathway analysis using fGSEA with CP:WIKIPATHWAYS dataset (reference: DGAC2). Table S14 shows DGAC1-enriched pathway analysis using fGSEA with CP:BIOCARTA dataset (reference: DGAC2). Table S15 shows DGAC1-enriched pathway analysis using fGSEA with CP:PID dataset (reference: DGAC2). Table S16 shows DGAC1-enriched pathway analysis using fGSEA with CP:KEGG dataset (reference: DGAC2). Table S17 shows Leiden-based gene list of murine GOs scRNA-seq data. Table S18 shows RSS of regulons from murine GO scRNA-seq data. Table S19 shows software and algorithms used for scRNA-seq.

Data availability

The data in the figures are available in the published article and the online supplemental material. The accession number for the scRNA-seq data set of 19 DGAC patients' samples (Figs. 1, 2, 5, S1, S2, and S5) is European Genome-phenome Archive EGAS00001004443. The accession number for the scRNA-seq data set of the 29 normal adjacent stomachs (Figs. 2 and S2) is Gene Expression Omnibus database GSE150290. The mouse scRNA-seq data (Figs. 3, 5, S4, and S5) are publicly available via the Gene Expression Omnibus database (GSE226266). The code used to reproduce the analyses described in this manuscript can be accessed via GitHub (https://github.com/jaeilparklab/EKP_DGAC_project) and will also be available upon request.

Acknowledgments

We are grateful to Pierre D. McCrea, Malgorzata Kloc, Rachael Miller, and Adriana Paulucci for their insightful comments.

This work was supported by the Cancer Prevention and Research Institute of Texas (RP200315 to J.-I. Park) and the National Cancer Institute (CA193297, CA256207, CA278967, CA278971, and CA279867 to J.-I. Park). The core facilities at MD Anderson (DNA Sequencing and Genetically Engineered Mouse Facility) were supported by the National Cancer Institute Cancer Center Support Grant (P30 CA016672). This work was performed at the Single Cell Genomics Core at BCM partially supported by National Institutes of Health-shared instrument grants (S10OD023469, S10OD025240) and P30EY002520. Illustrations were created with https://BioRender.com.

Author contributions: G. Zou, Y. Huang, and J.-I. Park conceived and designed the experiments. G. Zou, Y. Huang, S. Zhang, K.-P. Ko, B. Kim, J. Zhang, S. Jun, and V. Venkatesan performed the experiments. G. Zou, Y. Huang, S. Zhang, K.-P. Ko, B. Kim, S. Song, J.A. Ajani, and J.-I. Park analyzed the data. M.P. Pizzi, Y. Fan, S. Song, and J.A. Ajani provided the sequencing files and clinical data for human scRNA-seq analyses. N. Niu and H. Wang read and analyzed the stained slides. G. Zou, Y. Huang, S. Zhang, K.-P. Ko, B. Kim, and J.-I. Park wrote the manuscript.

Disclosures: The authors declare no competing interests exist.

Submitted: 3 April 2023
Revised: 27 September 2023
Accepted: 29 January 2024

References

Aibar, S., C.B. González-Blas, T. Moerman, V.A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.C. Marine, P. Geurts, J. Aerts, et al. 2017. SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods*. 14:1083–1086. https://doi.org/10.1038/nmeth.4463

Ajani, J.A., T.A. D'Amico, D.J. Bentrem, J. Chao, D. Cooke, C. Corvera, P. Das, P.C. Enzinger, T. Enzler, P. Fanta, et al. 2022. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc. Netw*. 20:167–192. https://doi.org/10.6004/jnccn.2022.0008

Al-Ahmadie, H.A., G. Iyer, B.H. Lee, S.N. Scott, R. Mehra, A. Bagrodia, E.J. Jordan, S.P. Gao, R. Ramirez, E.K. Cha, et al. 2016. Frequent somatic CDH1 loss-of-function mutations in plasmacytoid variant bladder cancer. *Nat. Genet*. 48:356–358. https://doi.org/10.1038/ng.3503

Assumpção, P.P., W.F. Barra, G. Ishak, L.G.V. Coelho, F.J.F. Coimbra, H.C. Freitas, E. Dias-Neto, M.C. Camargo, and M. Szklo. 2020. The diffuse-type gastric cancer epidemiology enigma. *BMC Gastroenterol*. 20:223. https://doi.org/10.1186/s12876-020-01354-4

Astudillo, P. 2020. Wnt5a signaling in gastric cancer. *Front. Cell Dev. Biol*. 8:110. https://doi.org/10.3389/fcell.2020.00110

Bartfeld, S., T. Bayram, M. van de Wetering, M. Huch, H. Begthel, P. Kujala, R. Vries, P.J. Peters, and H. Clevers. 2015. In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology*. 148:126–136.e6. https://doi.org/10.1053/j.gastro.2014.09.042

Blair, V.R., M. McLeod, F. Carneiro, D.G. Coit, J.L. D'Addario, J.M. van Dieren, K.L. Harris, N. Hoogerbrugge, C. Oliveira, R.S. van der Post, et al. 2020. Hereditary diffuse gastric cancer: Updated clinical practice guidelines. *Lancet Oncol*. 21:e386–e397. https://doi.org/10.1016/S1470-2045(20)30219-9

Boku, N., T. Satoh, M.H. Ryu, Y. Chao, K. Kato, H.C. Chung, J.S. Chen, K. Muro, W.K. Kang, K.H. Yeh, et al. 2021. Nivolumab in previously treated advanced gastric cancer (ATTRACTION-2): 3-year update and outcome of treatment beyond progression with nivolumab. *Gastric Cancer*. 24:946–958. https://doi.org/10.1007/s10120-021-01173-w

Bruner, H.C., and P.W.B. Derksen. 2018. Loss of E-cadherin-dependent cell-cell adhesion and the development and progression of cancer. *Cold Spring Harb. Perspect. Biol*. 10:a029330. https://doi.org/10.1101/cshperspect.a029330

Cancer Genome Atlas Research Network. 2014. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 513:202–209. https://doi.org/10.1038/nature13480

Chan, J.K., C.S. Ng, and P.K. Hui. 1988. A simple guide to the terminology and application of leucocyte monoclonal antibodies. *Histopathology*. 12:461–480. https://doi.org/10.1111/j.1365-2559.1988.tb01967.x

Chen, D.L., H. Sheng, D.S. Zhang, Y. Jin, B.T. Zhao, N. Chen, K. Song, and R.H. Xu. 2021. The circular RNA circDLG1 promotes gastric cancer progression and anti-PD-1 resistance through the regulation of CXCL12 by sponging miR-141-3p. *Mol. Cancer*. 20:166. https://doi.org/10.1186/s12943-021-01475-8

Chen, Y.C., W.L. Fang, R.F. Wang, C.A. Liu, M.H. Yang, S.S. Lo, C.W. Wu, A.F. Li, Y.M. Shyr, and K.H. Huang. 2016. Clinicopathological variation of lauren classification in gastric cancer. *Pathol. Oncol. Res*. 22:197–202. https://doi.org/10.1007/s12253-015-9996-6

Cho, J., S. Ahn, D.S. Son, N.K. Kim, K.W. Lee, S. Kim, J. Lee, S.H. Park, J.O. Park, W.K. Kang, et al. 2019. Bridging genomics and phenomics of gastric carcinoma. *Int. J. Cancer*. 145:2407–2417. https://doi.org/10.1002/ijc.32228

Cristescu, R., J. Lee, M. Nebozhyn, K.M. Kim, J.C. Ting, S.S. Wong, J. Liu, Y.G. Yue, J. Wang, K. Yu, et al. 2015. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat. Med*. 21:449–456. https://doi.org/10.1038/nm.3850

Fang, L., K. Liu, C. Liu, X. Wang, W. Ma, W. Xu, J. Wu, and C. Sun. 2022. Tumor accomplice: T cell exhaustion induced by chronic inflammation. *Front. Immunol*. 13:979116. https://doi.org/10.3389/fimmu.2022.979116

Fukamachi, H., S.K. Kim, J. Koh, H.S. Lee, Y. Sasaki, K. Yamashita, T. Nishikawaji, S. Shimada, Y. Akiyama, S.J. Byeon, et al. 2019. A subset of diffuse-type gastric cancer is susceptible to mTOR inhibitors and checkpoint inhibitors. *J. Exp. Clin. Cancer Res*. 38:127. https://doi.org/10.1186/s13046-019-1121-3

Galon, J., and D. Bruni. 2019. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov*. 18:197–218. https://doi.org/10.1038/s41573-018-0007-y

Galon, J., A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pagès, M. Tosolini, M. Camus, A. Berger, P. Wind, et al. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 313:1960–1964. https://doi.org/10.1126/science.1129139

Garcia-Pelaez, J., R. Barbosa-Matos, I. Gullo, F. Carneiro, and C. Oliveira. 2021. Histological and mutational profile of diffuse gastric cancer: Current knowledge and future challenges. *Mol. Oncol*. 15:2841–2867. https://doi.org/10.1002/1878-0261.12948

Ge, S., X. Xia, C. Ding, B. Zhen, Q. Zhou, J. Feng, J. Yuan, R. Chen, Y. Li, Z. Ge, et al. 2018. A proteomic landscape of diffuse-type gastric cancer. *Nat. Commun*. 9:1012. https://doi.org/10.1038/s41467-018-03121-2

Handschuh, G., S. Candidus, B. Luber, U. Reich, C. Schott, S. Oswald, H. Becke, P. Hutzler, W. Birchmeier, H. Höfler, and K.F. Becker. 1999. Tumour-associated E-cadherin mutations alter cellular morphology, decrease cellular adhesion and increase cellular motility. *Oncogene*. 18:4301–4312. https://doi.org/10.1038/sj.onc.1202790

Hayakawa, Y., H. Ariyama, J. Stancikova, K. Sakitani, S. Asfaha, B.W. Renz, Z.A. Dubeykovskaya, W. Shibata, H. Wang, C.B. Westphalen, et al. 2015. Mist1 expressing gastric stem cells maintain the normal and neoplastic gastric epithelium and are supported by a perivascular stem cell niche. *Cancer Cell*. 28:800–814. https://doi.org/10.1016/j.ccell.2015.10.003

Henson, D.E., C. Dittus, M. Younes, H. Nguyen, and J. Albores-Saavedra. 2004. Differential trends in the intestinal and diffuse types of gastric carcinoma in the United States, 1973-2000: Increase in the signet ring cell type. *Arch. Pathol. Lab. Med.* 128:765–770. https://doi.org/10.5858/2004-128-765-DTITIA

Hou, W., L. Kong, Z. Hou, and H. Ji. 2022. CD44 is a prognostic biomarker and correlated with immune infiltrates in gastric cancer. *BMC Med. Genomics*. 15:225. https://doi.org/10.1186/s12920-022-01383-w

Huang, B., J.H. Song, Y. Cheng, J.M. Abraham, S. Ibrahim, Z. Sun, X. Ke, and S.J. Meltzer. 2016. Long non-coding antisense RNA KRT7-AS is activated in gastric cancers and supports cancer cell progression by increasing KRT7 expression. *Oncogene*. 35:4927–4936. https://doi.org/10.1038/onc.2016.25

Huang, Y., K. Zhang, Y. Li, Y. Dai, and H. Zhao. 2020. The DLG1-AS1/miR-497/YAP1 axis regulates papillary thyroid cancer progression. *Aging*. 12:23326–23336. https://doi.org/10.18632/aging.104121

Iyer, P., M. Moslim, J.M. Farma, and C.S. Denlinger. 2020. Diffuse gastric cancer: Histologic, molecular, and genetic basis of disease. *Transl. Gastroenterol. Hepatol.* 5:52. https://doi.org/10.21037/tgh.2020.01.02

Jadhav, R.R., S.J. Im, B. Hu, M. Hashimoto, P. Li, J.X. Lin, W.J. Leonard, W.J. Greenleaf, R. Ahmed, and J.J. Goronzy. 2019. Epigenetic signature of PD-1+ TCF1+ CD8 T cells that act as resource cells during chronic viral infection and respond to PD-1 blockade. *Proc. Natl. Acad. Sci. USA*. 116:14113–14118. https://doi.org/10.1073/pnas.1903520116

Janjigian, Y.Y., K. Shitara, M. Moehler, M. Garrido, P. Salman, L. Shen, L. Wyrwicz, K. Yamaguchi, T. Skoczylas, A. Campos Bragagnoli, et al. 2021. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): A randomised, open-label, phase 3 trial. *Lancet*. 398:27–40. https://doi.org/10.1016/S0140-6736(21)00797-2

Jin, S., C.F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, P. Myung, M.V. Plikus, and Q. Nie. 2021. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* 12:1088. https://doi.org/10.1038/s41467-021-21246-9

Jung, Y.S., W. Wang, S. Jun, J. Zhang, M. Srivastava, M.J. Kim, E.M. Lien, J. Shang, J. Chen, P.D. McCrea, et al. 2018. Deregulation of CRAD-controlled cytoskeleton initiates mucinous colorectal cancer via β-catenin. *Nat. Cell Biol.* 20:1303–1314. https://doi.org/10.1038/s41556-018-0215-z

Kakiuchi, M., T. Nishizawa, H. Ueda, K. Gotoh, A. Tanaka, A. Hayashi, S. Yamamoto, K. Tatsuno, H. Katoh, Y. Watanabe, et al. 2014. Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat. Genet.* 46:583–587. https://doi.org/10.1038/ng.2984

Kaneta, Y., T. Sato, Y. Hikiba, M. Sugimori, S. Sue, H. Kaneko, K. Irie, T. Sasaki, M. Kondo, M. Chuma, et al. 2020. Loss of pancreatic E-cadherin causes pancreatitis-like changes and contributes to carcinogenesis. *Cell. Mol. Gastroenterol. Hepatol.* 9:105–119. https://doi.org/10.1016/j.jcmgh.2019.09.001

Kelley, D.R., J. Snoek, and J.L. Rinn. 2016. Basset: Learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Res.* 26:990–999. https://doi.org/10.1101/gr.200535.115

Kim, J., C. Park, K.H. Kim, E.H. Kim, H. Kim, J.K. Woo, J.K. Seong, K.T. Nam, Y.C. Lee, and S.Y. Cho. 2022. Single-cell analysis of gastric pre-cancerous and cancer lesions reveals cell lineage diversity and intratumoral heterogeneity. *NPJ Precis. Oncol.* 6:9. https://doi.org/10.1038/s41698-022-00251-1

Kim, M.J., C. Cervantes, Y.S. Jung, X. Zhang, J. Zhang, S.H. Lee, S. Jun, L. Litovchick, W. Wang, J. Chen, et al. 2021. PAF remodels the DREAM complex to bypass cell quiescence and promote lung tumorigenesis. *Mol Cell*. 81:1698–1714.e1696. https://doi.org/10.1016/j.molcel.2021.02.001

Kim, S.K., H.J. Kim, J.L. Park, H. Heo, S.Y. Kim, S.I. Lee, K.S. Song, W.H. Kim, and Y.S. Kim. 2020. Identification of a molecular signature of prognostic subtypes in diffuse-type gastric cancer. *Gastric Cancer*. 23:473–482. https://doi.org/10.1007/s10120-019-01029-4

Kim, Y.I., R.L. Pecha, T. Keihanian, M. Mercado, S.V. Pena-Munoz, K. Lang, G. Van Buren, S. Dhingra, and M.O. Othman. 2023. MUC1 expressions and its prognostic values in US gastric cancer patients. *Cancers*. 15:15. https://doi.org/10.3390/cancers15040998

Ko, K.P., Y. Huang, S. Zhang, G. Zou, B. Kim, J. Zhang, S. Jun, C. Martin, K.J. Dunbar, G. Efe, et al. 2023. Key genetic determinants driving esophageal squamous cell carcinoma initiation and immune evasion. *Gastroenterology*. 165:613–628.e20. https://doi.org/10.1053/j.gastro.2023.05.030

Ko, K.P., J. Zhang, and J.I. Park. 2022. Establishing transgenic murine esophageal organoids. *STAR Protoc.* 3:101317. https://doi.org/10.1016/j.xpro.2022.101317

Körfer, J., F. Lordick, and U.T. Hacker. 2021. Molecular targets for gastric cancer treatment and future perspectives from a clinical and translational point of view. *Cancers*. 13:5216. https://doi.org/10.3390/cancers13205216

Korotkevich, G., V. Sukhov, N. Budin, B. Shpak, M.N. Artyomov, and A. Sergushichev. 2021. Fast gene set enrichment analysis. *bioRxiv*. https://doi.org/10.1101/060012 (Preprint posted February 01, 2021).

Korsunsky, I., N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.R. Loh, and S. Raychaudhuri. 2019. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods*. 16:1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Koushyar, S., A.G. Powell, E. Vincan, and T.J. Phesse. 2020. Targeting Wnt signaling for the treatment of gastric cancer. *Int. J. Mol. Sci.* 21:21. https://doi.org/10.3390/ijms21113927

Lee, J.S., and E. Ruppin. 2019. Multiomics prediction of response rates to therapies to inhibit programmed cell death 1 and programmed cell death 1 ligand 1. *JAMA Oncol.* 5:1614–1618. https://doi.org/10.1001/jamaoncol.2019.2311

Lee, M.G., R. Villa, P. Trojer, J. Norman, K.P. Yan, D. Reinberg, L. Di Croce, and R. Shiekhattar. 2007. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science*. 318:447–450. https://doi.org/10.1126/science.1149042

Lei, Z.N., Q.X. Teng, Q. Tian, W. Chen, Y. Xie, K. Wu, Q. Zeng, L. Zeng, Y. Pan, Z.S. Chen, and Y. He. 2022. Signaling pathways and therapeutic interventions in gastric cancer. *Signal Transduct. Target. Ther.* 7:358. https://doi.org/10.1038/s41392-022-01190-w

Liu, X., Y. Wang, B. Chen, W.N. Chan, C.W. Mui, A.H.K. Cheung, J. Zhang, K.Y. Wong, J. Yu, W. Kang, and K.F. To. 2022. Targeting the hippo pathway in gastric cancer and other malignancies in the digestive system: From bench to bedside. *Biomedicines*. 10:2512. https://doi.org/10.3390/biomedicines10102512

Liu, Y., B. Zhang, Y. Zhou, Y. Xing, Y. Wang, Y. Jia, and D. Liu. 2023. Targeting hippo pathway: A novel strategy for Helicobacter pylori-induced gastric cancer treatment. *Biomed. Pharmacother.* 161:114549. https://doi.org/10.1016/j.biopha.2023.114549

Mazzarella, L., B.A. Duso, D. Trapani, C. Belli, P. D'Amico, E. Ferraro, G. Viale, and G. Curigliano. 2019. The evolving landscape of 'next-generation' immune checkpoint inhibitors: A review. *Eur. J. Cancer*. 117:14–31. https://doi.org/10.1016/j.ejca.2019.04.035

McBeath, R., D.M. Pirone, C.M. Nelson, K. Bhadriraju, and C.S. Chen. 2004. Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Dev. Cell*. 6:483–495. https://doi.org/10.1016/S1534-5807(04)00075-9

McCrea, P.D., and J.I. Park. 2007. Developmental functions of the P120-catenin sub-family. *Biochim. Biophys. Acta*. 1773:17–33. https://doi.org/10.1016/j.bbamcr.2006.06.009

Messina, B., F. Lo Sardo, S. Scalera, L. Memeo, C. Colarossi, M. Mare, G. Blandino, G. Ciliberto, M. Maugeri-Saccà, and G. Bon. 2023. Hippo pathway dysregulation in gastric cancer: From Helicobacter pylori infection to tumor promotion and progression. *Cell Death Dis.* 14:21. https://doi.org/10.1038/s41419-023-05568-8

Mimata, A., H. Fukamachi, Y. Eishi, and Y. Yuasa. 2011. Loss of E-cadherin in mouse gastric epithelial cells induces signet ring-like cells, a possible precursor lesion of diffuse gastric cancer. *Cancer Sci.* 102:942–950. https://doi.org/10.1111/j.1349-7006.2011.01890.x

Muro, K., E. Van Cutsem, Y. Narita, G. Pentheroudakis, E. Baba, J. Li, M.H. Ryu, W.I.W. Zamaniah, W.P. Yong, K.H. Yeh, et al. 2019. Pan-asian adapted ESMO clinical practice guidelines for the management of patients with metastatic gastric cancer: A JSMO-ESMO initiative endorsed by CSCO, KSMO, MOS, SSO and TOS. *Ann. Oncol.* 30:19–33. https://doi.org/10.1093/annonc/mdy502

O'Connor, K., and M. Chen. 2013. Dynamic functions of RhoA in tumor cell migration and invasion. *Small GTPases*. 4:141–147. https://doi.org/10.4161/sgtp.25131

Oh, S.C., B.H. Sohn, J.-H. Cheong, S.-B. Kim, J.E. Lee, K.C. Park, S.H. Lee, J.-L. Park, Y.-Y. Park, H.-S. Lee, et al. 2018. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. *Nat. Commun.* 9:1777. https://doi.org/10.1038/s41467-018-04179-8

Oliveira, C., H. Pinheiro, J. Figueiredo, R. Seruca, and F. Carneiro. 2015. Familial gastric cancer: Genetic susceptibility, pathology, and implications for management. *Lancet Oncol.* 16:e60–e70. https://doi.org/10.1016/S1470-2045(14)71016-2

Ooki, A., and K. Yamaguchi. 2022. The dawn of precision medicine in diffuse-type gastric cancer. *Ther. Adv. Med. Oncol.* 14:17588359221083049. https://doi.org/10.1177/17588359221083049

Ponta, H., L. Sherman, and P.A. Herrlich. 2003. CD44: From adhesion molecules to signalling regulators. *Nat. Rev. Mol. Cell Biol.* 4:33–45. https://doi.org/10.1038/nrm1004

Ran, F.A., P.D. Hsu, J. Wright, V. Agarwala, D.A. Scott, and F. Zhang. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8:2281–2308. https://doi.org/10.1038/nprot.2013.143

Seeneevassen, L., P. Dubus, C. Gronnier, and C. Varon. 2022. Hippo in gastric cancer: From signalling to therapy. *Cancers*. 14:14. https://doi.org/10.3390/cancers14092282

Selim, J.H., S. Shaheen, W.C. Sheu, and C.T. Hsueh. 2019. Targeted and novel therapy in advanced gastric cancer. *Exp. Hematol. Oncol.* 8:25. https://doi.org/10.1186/s40164-019-0149-6

Stodden, G.R., M.E. Lindberg, M.L. King, M. Paquet, J.A. MacLean, J.L. Mann, F.J. DeMayo, J.P. Lydon, and K. Hayashi. 2015. Loss of Cdh1 and Trp53 in the uterus induces chronic inflammation with modification of tumor microenvironment. *Oncogene*. 34:2471–2482. https://doi.org/10.1038/onc.2014.193

Sun, D., X. Guan, A.E. Moran, L.Y. Wu, D.Z. Qian, P. Schedin, M.S. Dai, A.V. Danilov, J.J. Alumkal, A.C. Adey, et al. 2022. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nat. Biotechnol.* 40:527–538. https://doi.org/10.1038/s41587-021-01091-3

Sung, H., J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray. 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71:209–249. https://doi.org/10.3322/caac.21660

Suo, S., Q. Zhu, A. Saadatpour, L. Fei, G. Guo, and G.C. Yuan. 2018. Revealing the critical regulators of cell identity in the mouse cell Atlas. *Cell Rep.* 25:1436–1445.e3. https://doi.org/10.1016/j.celrep.2018.10.045

Tan, S.H., Y. Swathi, S. Tan, J. Goh, R. Seishima, K. Murakami, M. Oshima, T. Tsuji, P. Phuah, L.T. Tan, et al. 2020. AQP5 enriches for stem cells and cancer origins in the distal stomach. *Nature*. 578:437–443. https://doi.org/10.1038/s41586-020-1973-x

Teng, M.W., S.F. Ngiow, A. Ribas, and M.J. Smyth. 2015. Classifying cancers based on T-cell infiltration and PD-L1. *Cancer Res.* 75:2139–2145. https://doi.org/10.1158/0008-5472.CAN-15-0255

Till, J.E., C. Yoon, B.J. Kim, K. Roby, P. Addai, E. Jonokuchi, L.H. Tang, S.S. Yoon, and S. Ryeom. 2017. Oncogenic KRAS and p53 loss drive gastric tumorigenesis in mice that can be attenuated by E-cadherin expression. *Cancer Res.* 77:5349–5359. https://doi.org/10.1158/0008-5472.CAN-17-0061

Tong, M., C. Yu, J. Shi, W. Huang, S. Ge, M. Liu, L. Song, D. Zhan, X. Xia, W. Liu, et al. 2019. Phosphoproteomics enables molecular subtyping and nomination of kinase candidates for individual patients of diffuse-type gastric cancer. *iScience*. 22:44–57. https://doi.org/10.1016/j.isci.2019.11.003

Van de Sande, B., C. Flerin, K. Davie, M. De Waegeneer, G. Hulselmans, S. Aibar, R. Seurinck, W. Saelens, R. Cannoodt, Q. Rouchon, et al. 2020. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* 15:2247–2276. https://doi.org/10.1038/s41596-020-0336-2

van Roy, F., and G. Berx. 2008. The cell-cell adhesion molecule E-cadherin. *Cell. Mol. Life Sci.* 65:3756–3788. https://doi.org/10.1007/s00018-008-8281-1

Verma, S.K., X. Tian, L.V. LaFrance, C. Duquenne, D.P. Suarez, K.A. Newlander, S.P. Romeril, J.L. Burgess, S.W. Grant, J.A. Brackley, et al. 2012. Identification of potent, selective, cell-active inhibitors of the histone lysine methyltransferase EZH2. *ACS Med. Chem. Lett.* 3:1091–1096. https://doi.org/10.1021/ml3003346

Wainberg, Z.A., C.S. Fuchs, J. Tabernero, K. Shitara, K. Muro, E. Van Cutsem, Y.J. Bang, H.C. Chung, K. Yamaguchi, E. Varga, et al. 2021. Efficacy of pembrolizumab monotherapy for advanced gastric/gastroesophageal junction cancer with programmed death ligand 1 combined positive score ≥10. *Clin. Cancer Res.* 27:1923–1931. https://doi.org/10.1158/1078-0432.CCR-20-2980

Wang, J., G. Zhang, Y. Sui, Z. Yang, Y. Chu, H. Tang, B. Guo, C. Zhang, and C. Wu. 2020a. CD52 is a prognostic biomarker and associated with tumor microenvironment in breast cancer. *Front. Genet.* 11:578002. https://doi.org/10.3389/fgene.2020.578002

Wang, R., S. Song, K. Harada, F. Ghazanfari Amlashi, B. Badgwell, M.P. Pizzi, Y. Xu, W. Zhao, X. Dong, J. Jin, et al. 2020b. Multiplex profiling of peritoneal metastases from gastric adenocarcinoma identified novel targets and molecular subtypes that predict treatment response. *Gut*. 69:18–31. https://doi.org/10.1136/gutjnl-2018-318070

Wang, R., M. Dang, K. Harada, G. Han, F. Wang, M. Pool Pizzi, M. Zhao, G. Tatlonghari, S. Zhang, D. Hao, et al. 2021. Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat. Med.* 27:141–151. https://doi.org/10.1038/s41591-020-1125-8

Wang, S.C., Y. Yeu, S.T.G. Hammer, S. Xiao, M. Zhu, C. Hong, J.R. Clemenceau, L.Y. Yoon, I. Nassour, J. Shen, et al. 2020c. Hispanic/latino patients with gastric adenocarcinoma have distinct molecular profiles including a high rate of germline CDH1 variants. *Cancer Res.* 80:2114–2124. https://doi.org/10.1158/0008-5472.CAN-19-2918

Weber, E.W., K.R. Parker, E. Sotillo, R.C. Lynn, H. Anbunathan, J. Lattin, Z. Good, J.A. Belk, B. Daniel, D. Klysz, et al. 2021. Transient rest restores functionality in exhausted CAR-T cells through epigenetic remodeling. *Science*. 372:372. https://doi.org/10.1126/science.aba1786

Wherry, E.J., and M. Kurachi. 2015. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* 15:486–499. https://doi.org/10.1038/nri3862

Wolf, F.A., P. Angerer, and F.J. Theis. 2018. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* 19:15. https://doi.org/10.1186/s13059-017-1382-0

Wolock, S.L., R. Lopez, and A.M. Klein. 2019. Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* 8:281–291.e9. https://doi.org/10.1016/j.cels.2018.11.005

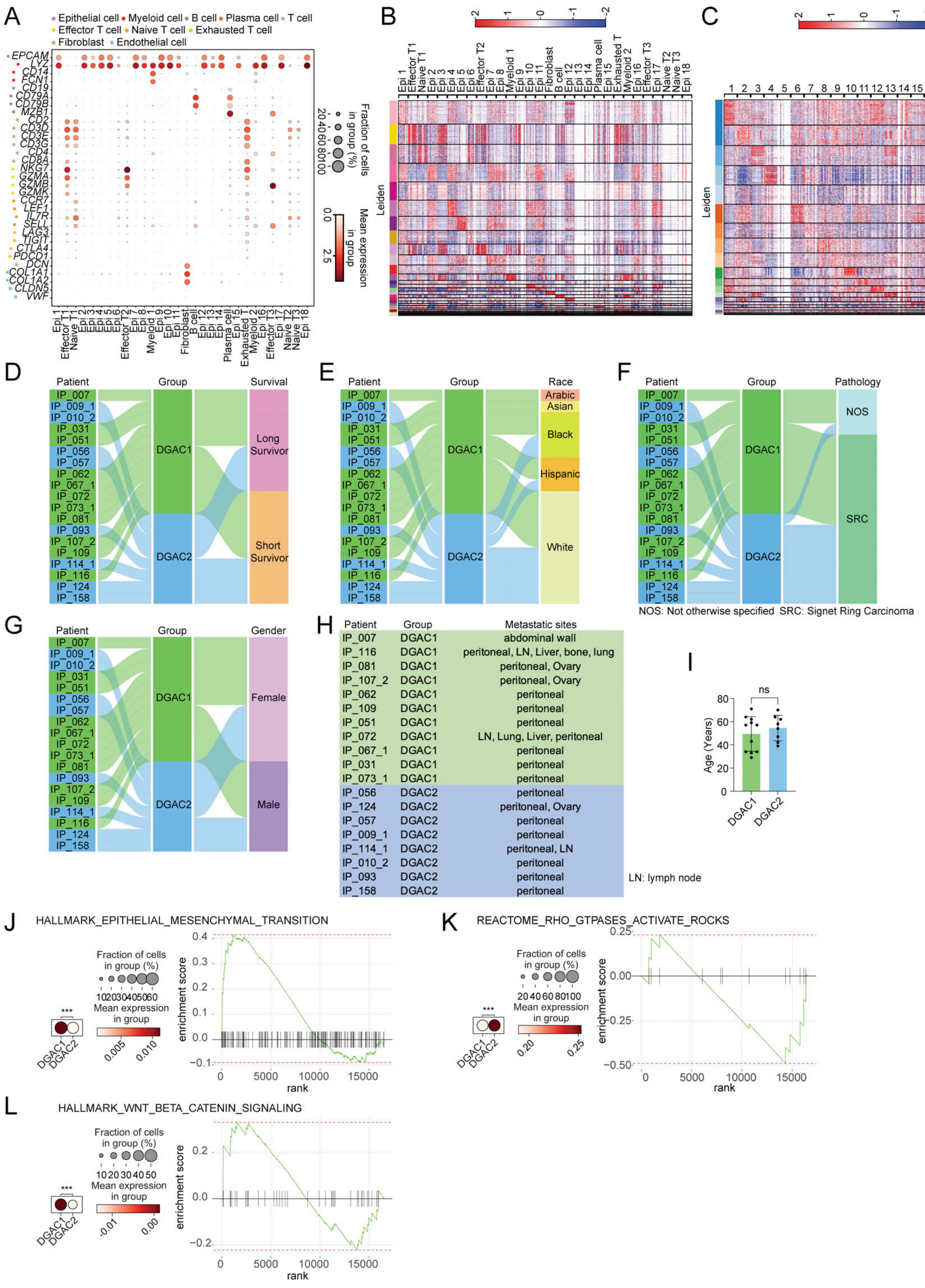
Yin, T., G. Wang, S. He, Q. Liu, J. Sun, and Y. Wang. 2016. Human cancer cells with stem cell-like phenotype exhibit enhanced sensitivity to the cytotoxicity of IL-2 and IL-15 activated natural killer cells. *Cell. Immunol.* 300:41–45. https://doi.org/10.1016/j.cellimm.2015.11.009

Young, M.D., and S. Behjati. 2020. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience*. 9:giaa151. https://doi.org/10.1093/gigascience/giaa151

Yu, W., N. Liu, X. Song, L. Chen, M. Wang, G. Xiao, T. Li, Z. Wang, and Y. Zhang. 2023. EZH2: An accomplice of gastric cancer. *Cancers*. 15:425. https://doi.org/10.3390/cancers15020425

Zhang, K., K. Chu, X. Wu, H. Gao, J. Wang, Y.C. Yuan, S. Loera, K. Ho, Y. Wang, W. Chow, et al. 2013. Amplification of FRS2 and activation of FGFR/FRS2 signaling pathway in high-grade liposarcoma. *Cancer Res.* 73:1298–1307. https://doi.org/10.1158/0008-5472.CAN-12-2086

Supplemental material



Patient	Group	Metastatic sites
IP_007	DGAC1	abdominal wall
IP_116	DGAC1	peritoneal, LN, Liver, bone, lung
IP_081	DGAC1	peritoneal, Ovary
IP_107_2	DGAC1	peritoneal, Ovary
IP_062	DGAC1	peritoneal
IP_109	DGAC1	peritoneal
IP_051	DGAC1	peritoneal
IP_072	DGAC1	LN, Lung, Liver, peritoneal
IP_067_1	DGAC1	peritoneal
IP_031	DGAC1	peritoneal
IP_073_1	DGAC1	peritoneal
IP_056	DGAC2	peritoneal
IP_124	DGAC2	peritoneal, Ovary
IP_057	DGAC2	peritoneal
IP_009_1	DGAC2	peritoneal
IP_114_1	DGAC2	peritoneal, LN
IP_010_2	DGAC2	peritoneal
IP_093	DGAC2	peritoneal
IP_158	DGAC2	peritoneal

LN: lymph node

Figure S1. **Transcriptional, clinical, and molecular characterization of DGAC subtypes. (A)** Dot plots of epithelial cell, myeloid cell, B cell, plasma cell, T cell, effector T cell, naïve T cell, exhausted T cell, fibroblast, and endothelial cell markers in integrated 19 DGAC patients scRNA-seq data. **(B)** Leiden-based heatmap of all cells of integrated datasets with annotations in 19 DGAC patients. The most highly variable 100 genes of each cluster are shown in Table S6. **(C)** Leiden-based heatmap of epithelial cells of integrated datasets in 19 DGAC patients. The most highly variable 100 genes of each cluster were showed in Table S7. **(D–G)** Venn diagram illustrating 19 DGAC patient groups with survival, race, pathology, and gender data. Long-term survivors (surviving over 1 year after diagnosis) and short-term survivors (deceased within 6 mo post-diagnosis) were classified according to our previous publication (Wang et al., 2021). **(H)** Metastatic sites of DGAC1 and DGAC2 patients. **(I)** Age difference between DGAC1 and DGAC2 patients. P values were calculated using Student's *t* test; error bars: SD. **(J–L)** Dot plots and GSEA of EMT (J), RHOA (K), and WNT (L) scores in two DGAC types. The datasets we used for dot plots and GSEA are from GSEA molecular signature database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp): EMT: Human Gene Set: HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION; RHOA: Human Gene Set: REACTOME_RHO_GTPASES_ACTIVATE_ROCKS; WNT: Human Gene Set: HALLMARK_WNT_BETA_CATENIN_SIGNALING. P values were calculated by Mann–Whitney testing (J–L). The genes included in each score are listed in Table S9. ns: P > 0.05; ***: P ≤ 0.001. All data are derived from two or more independent experiments with the indicated number of human donors.

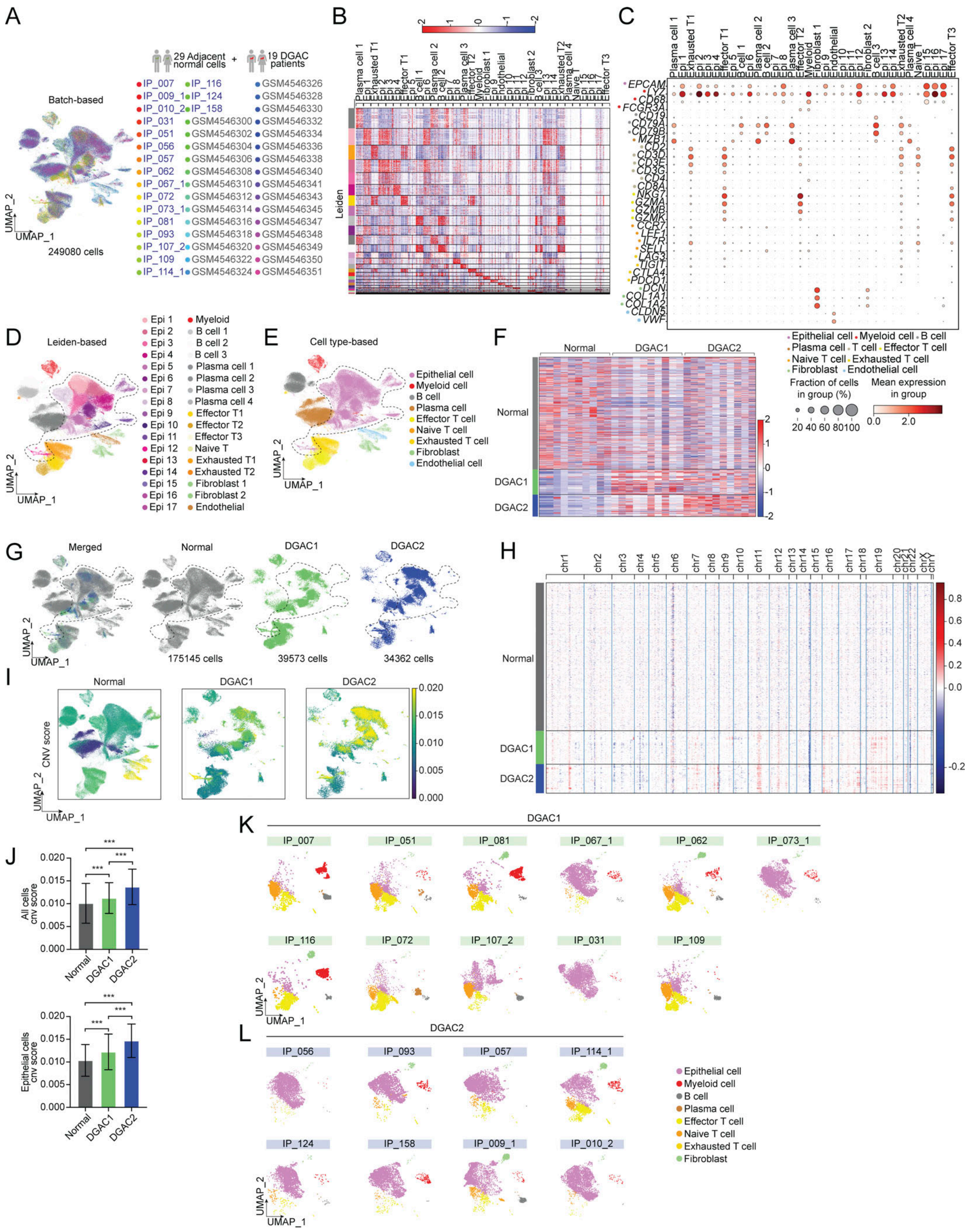


Figure S2. **scRNA-seq analysis of 19 DGAC patients and 29 adjacent normal stomach tissue. (A)** Integrated batch-based UMAP of 29 adjacent normal stomach tissue (normal tissue) and 19 DGAC patients. Total cell numbers are 249,080. Integration package: Harmony. **(B)** Annotated Leiden-based integrated UMAPs of 19 DGAC patients and 29 adjacent normal stomach tissue. Epi: Epithelial cells; Myeloid: myeloid cells; Effector T: effector T cells; Naïve T: Naïve T cells; Exhausted T: Exhausted T cells; Endothelial: Endothelial cells. **(C)** Dot plots of epithelial cell, myeloid cell, B cell, plasma cell, T cell, effector T cell, naïve T cell, exhausted T cell, fibroblast, and endothelial cell markers in integrated 19 DGAC patients and 29 adjacent normal stomach tissue scRNA-seq data. **(D)** Integrated Leiden-based UMAPs of 29 adjacent normal stomach tissue (normal tissue) and 19 DGAC patients. Epi: epithelial cells; Myeloid: myeloid cells; Effector T: effector T cells; Naïve T: naïve T cells; Exhausted T: exhausted T cells. The most highly variable 100 genes of each cluster are shown in Table S10. **(E)** Integrated cell type–based UMAP of 29 normal tissue and 19 DGAC patients. All cells were reclustered according to the Leiden clusters and gathered as mega clusters. Dashed line-circle: epithelial cells. **(F)** Type-based heatmap of all cells of integrated datasets in 19 DGAC patients and 29 adjacent normal stomach tissue. **(G)** Separated UMAPs of normal tissue and two types of DGACs. Dashed line-circle: epithelial cells. **(H)** CNV heatmap of DGAC1 and DGAC2, tumor-adjacent normal stomach tissue (Normal) was used as a reference for the CNV inference. Red: copy number gain (CNG); blue: copy number loss (CNL). **(I)** CNV heatmap of DGAC1 and DGAC2, tumor-adjacent normal stomach tissue (Normal) was used as reference for the CNV inference. **(J)** Statistics analysis of CNV score of all cells (left panel) and epithelial cells (right panel) among Normal, DGAC1, and DGAC2. P values were calculated using the one-way ANOVA; error bars: SD. **(K and L)** Individual cell type–based UMAP of the patients in DGAC1 and DGAC2. DGAC1 patients were enriched with stromal cells, mainly T cells. DGAC2 patients were enriched with epithelial cells. ***: P ≤ 0.001. All data are derived from two or more independent experiments with the indicated number of human donors.

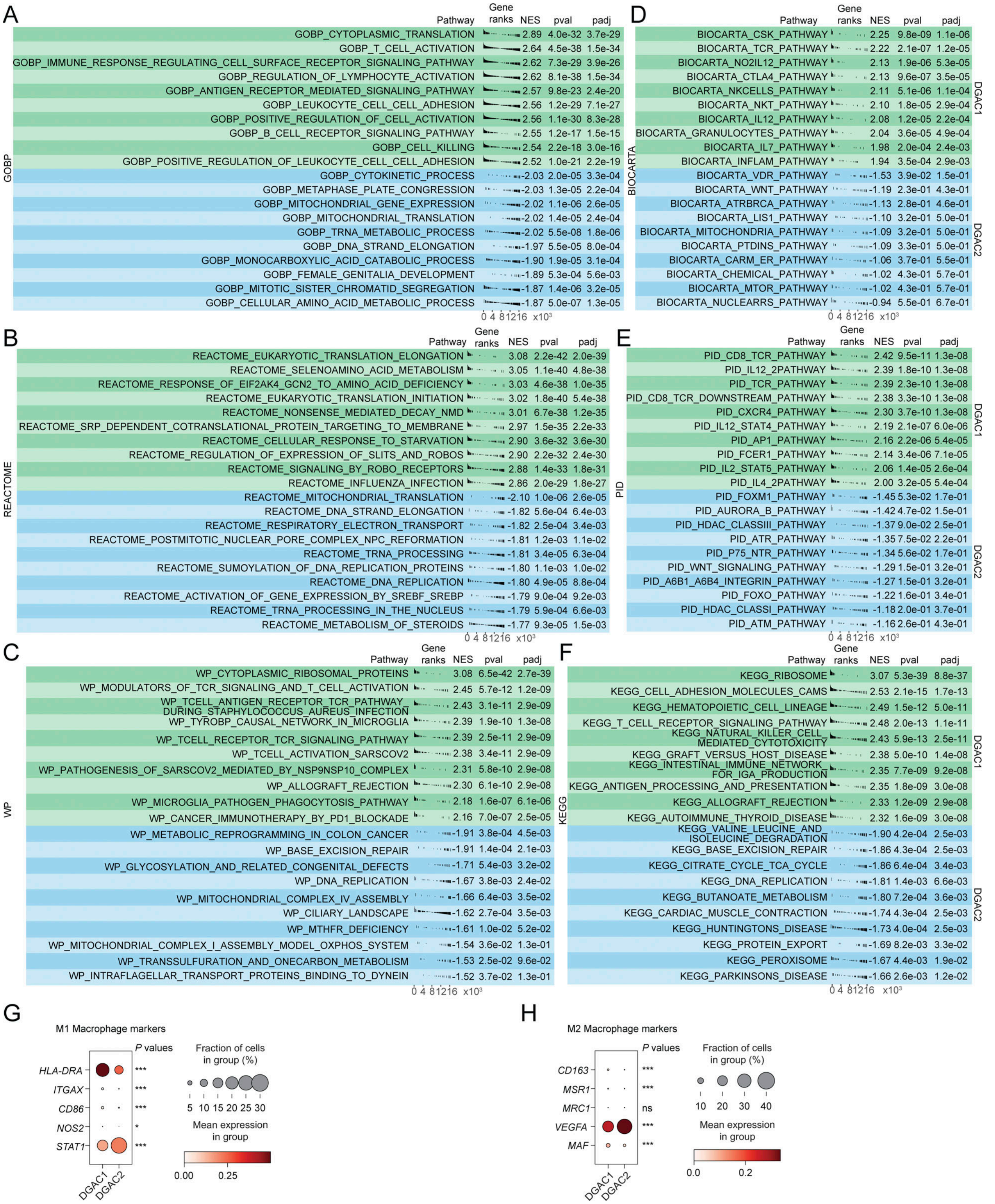


Figure S3. **GSEA analysis and the expression of macrophage polarization markers of DGAC1 and DGAC2. (A–F)** GSEA analysis comparing DGAC1 to DGAC2 using DGAC2 as the reference gene set. Enriched pathways in DGAC1 are displayed in the upper green panel, while those in DGAC2 are shown in the lower blue panel. Pathway datasets analyzed include GOBP (A), REACTOME (B), WP (C), BIOCARTA (D), PID (E), and KEGG (F). Pathways with positive normalized enrichment score (NES) indicate enrichment in DGAC1, while those with negative NES indicate enrichment in DGAC2. GOBP: gene ontology biological process; REACTOME: reactome gene sets; WP: WikiPathways gene sets; BIOCARTA: BioCarta gene sets; PID: PID gene sets; KEGG: KEGG gene sets. Pathways related with immune response were enriched in DGAC1 based on GOBP, WP, BIOCARTA, PID, and KEGG. **(G and H)** Dot plot of macrophage polymerization markers in DGAC1 and DGAC2. Most of the M1 and M2 markers are enriched in DGAC1, except for STAT1 and VEGFA. P values were calculated by using a *t* test. ns: P > 0.05; *: P ≤ 0.05; ***: P ≤ 0.001. All data are derived from two or more independent experiments with the indicated number of human donors.

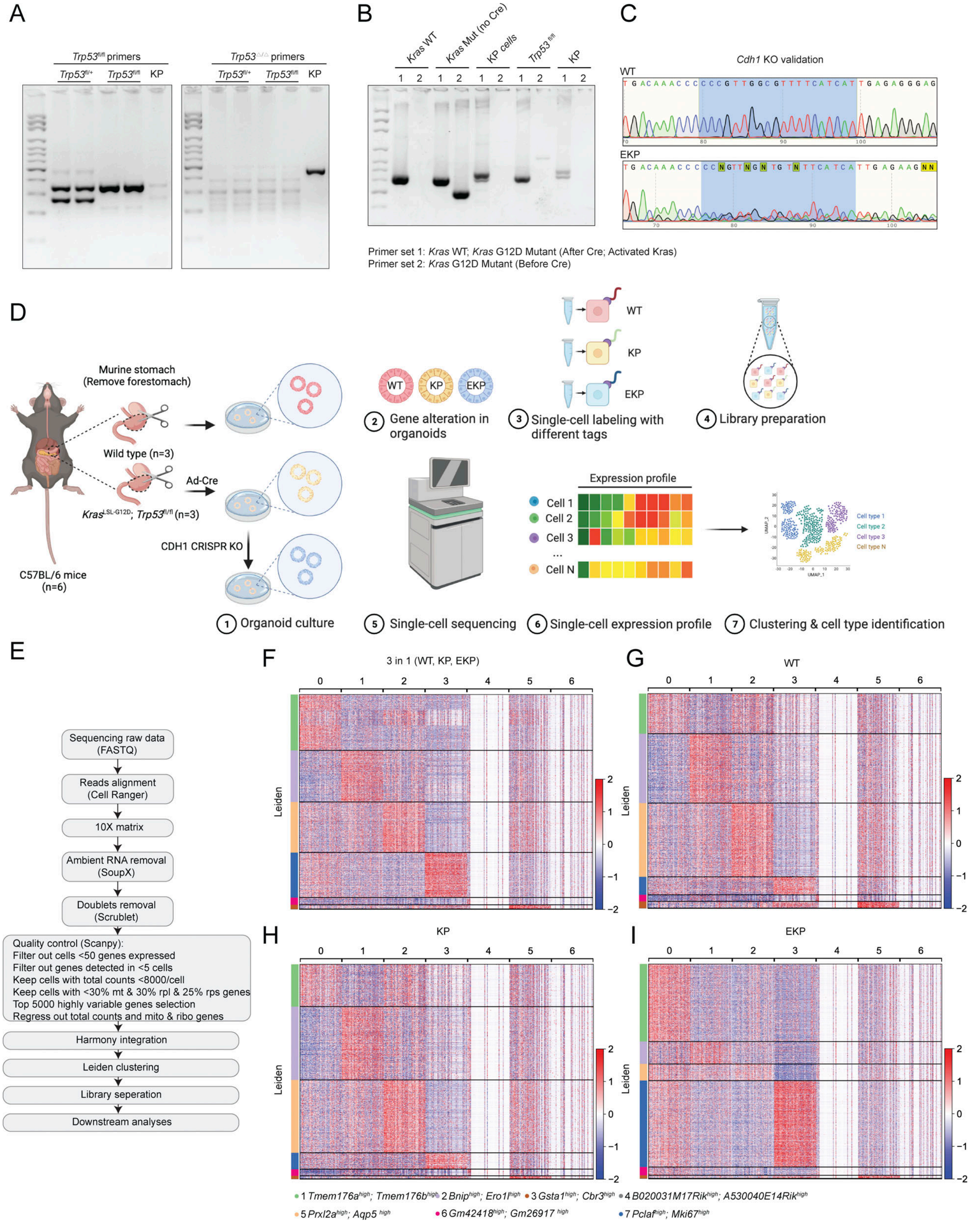


Figure S4. **Validation of genetic engineering and scRNA-seq analysis of mouse GOs. (A–C)** Genotyping results of KP organoids (A and B). After adeno-Cre treatment, KP organoids lost *Trp53*, while *Kras*^{G12D} was activated in KP organoids. After *CDH1* CRISPR KO, we performed Sanger sequencing to compare the sequence of *CDH1* in WT and EKP (C). The five targeting sequences against *CDH1* are shown in CRISPR/Cas9-based gene knockout in GOs. The primers used for genotyping are shown in Table S2. **(D)** Illustration of the workflow for stomach tissue collection and dissociation, gene manipulation of the gastric organoids (GOs), sample preparation of multiplex scRNA sequencing. **(E)** Workflow of single-cell library preparation. **(F)** Heatmap of each cell clusters of integrated datasets, including WT, KP, and EKP. **(G–I)** Separate heatmap of each cell clusters of WT, KP, and EKP datasets, respectively. All data are derived from two or more independent experiments with the indicated number of mice. Source data are available for this figure: SourceData FS4.

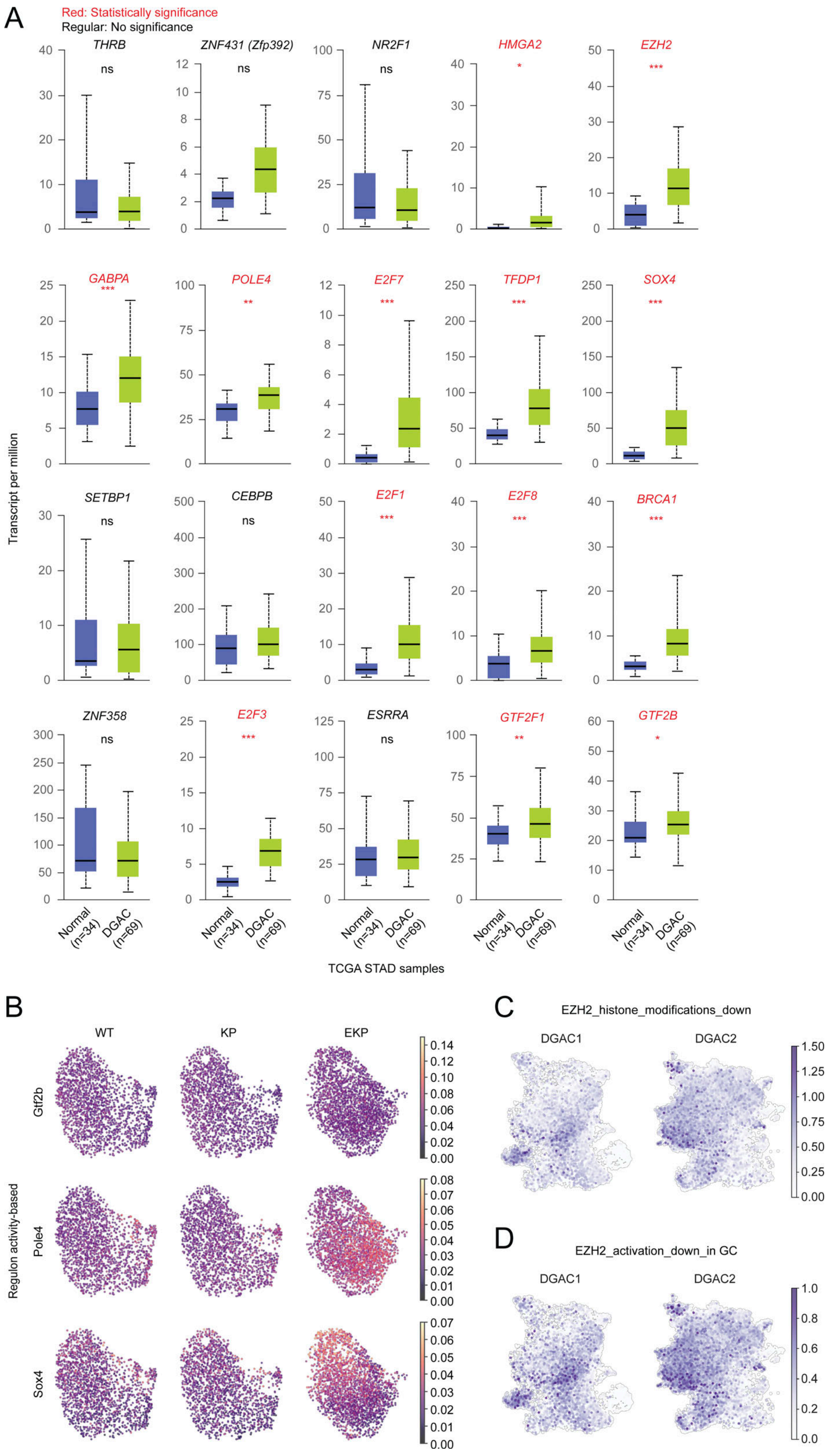


Figure S5. **EKP-specific regulons expression and EZH2 downstream targeted genes expression. (A)** The expression of 20 regulons in TCGA DGAC patients and normal stomach. **(B)** Regulon activity based UMAP of Gtf2b, Pole4, and Sox4. P values were calculated by using the Student's *t* test; error bars: SD. **(C and D)** Feature plots of EZH2 downstream target genes (C, genes which are downregulated by EZH2 through histone modification; D, genes which are downregulated by EZH2 reported in gastric cancer) scores in the epithelial cells of DGAC1 and DGAC2. Gene list of EZH2 targeted genes was listed in Table S9. ns: $P > 0.05$; *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$. All data are derived from two or more independent experiments with the indicated number of samples.

https://doi.org/10.1084/jem.20230561

Provided online are 19 tables. Table S1 shows the gastric organoid culture medium. Table S2 lists primers used for genotyping. Table S3 shows antibody information. Table S4 shows clinical information of 114 DGAC patient tumor samples for tissue microarray analysis. Table S5 shows scRNA-seq metadata and clinical information of 19 DGAC patient samples. Table S6 shows Leiden-based gene list of 19 DGAC samples scRNA-seq data (all cells included). Table S7 shows Leiden-based gene list of 19 DGAC scRNA-seq data (epithelial cells only). Table S8 shows type-based gene list of 19 DGAC scRNA-seq data (epithelial cells only). Table S9 shows the gene list used for gene set score analysis. Table S10 shows Leiden-based gene list of merged 19 DGAC patients and 29 normal stomach scRNA-seq data (all cells included). Table S11 shows DGAC1-enriched pathway analysis using fGSEA with GO:BP dataset (reference: DGAC2). Table S12 shows DGAC1-enriched pathway analysis using fGSEA with CP:REACTOME dataset (reference: DGAC2). Table S13 shows DGAC1-enriched pathway analysis using fGSEA with CP:WIKIPATHWAYS dataset (reference: DGAC2). Table S14 shows DGAC1-enriched pathway analysis using fGSEA with CP:WIKIPATHWAYS dataset (reference: DGAC2). Table S15 shows DGAC1-enriched pathway analysis using fGSEA with CP:PID dataset (reference: DGAC2). Table S16 shows DGAC1-enriched pathway analysis using fGSEA with CP:KEGG dataset (reference: DGAC2). Table S17 shows Leiden-based gene list of murine gastric organoids scRNA-seq data. Table S18 shows the RSS of regulons from murine gastric organoid scRNA-seq data. Table S19 shows software and algorithms used for scRNA-seq.