

Divergent clonal evolution and tumor microenvironment remodeling shape gastric cancer peritoneal metastasis

Received: 21 May 2025

Accepted: 10 March 2026

Cite this article as: Charton, C., Kang, S.H., Kim, S.S. *et al.* Divergent clonal evolution and tumor microenvironment remodeling shape gastric cancer peritoneal metastasis. *Commun Biol* (2026). <https://doi.org/10.1038/s42003-026-09906-2>

Clémentine Charton, So Hyun Kang, Stephanie S. Kim, Minsu Kang, Ju Hyun Lee, Young Suk Park, Sang-Hoon Ahn, Yun-Suhk Suh, Nak Jung Kwon, Kyoung Un Park, Jieun Lee, Jinho Kim & Hyung-Ho Kim

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Divergent clonal evolution and tumor microenvironment remodeling shape gastric cancer peritoneal metastasis

Clémentine Charton^{1, *}, So Hyun Kang^{2, *}, Stephanie S. Kim¹, Minsu Kang³, Ju Hyun Lee³, Young Suk Park², Sang-Hoon Ahn², Yun-Suhk Suh², Nak Jung Kwon⁴, Kyoung Un Park^{5, 6}, Jieun Lee^{2, ☐}, Jinho Kim^{1, 5, 7, ☐} & Hyung-Ho Kim^{2, 8, ☐}

¹Precision Medicine Center, Future Innovation Research Division, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

²Department of Surgery, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

³Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

⁴Precision Medicine Institute, MacroGen Inc., Seoul, Republic of Korea

⁵Department of Laboratory Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Republic of Korea

⁶Department of Laboratory Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

⁷Department of Genomic Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

⁸Department of Surgery, Chung-Ang University Gwang Myeong Hospital, Gwangmyeong, Republic of Korea

* These authors contributed equally to this work.

☐ Corresponding author:

Hyung-Ho Kim, MD, PhD, FACS

Professor, Chung-Ang University College of Medicine

110, Deokan-ro, Gwangmyeong-si, Gyeonggi-do, 14353, Republic of Korea

Tel:+82-2-2610-6755; E-mail: lapakh2@cau.ac.kr

Jinho Kim, PhD.

Associate Professor, Seoul National University Bundang Hospital

82 Gumi-ro-173-beon-gil Bundang-gu, Seongnam-si, Gyeonggi-do, 13620, Republic of Korea

Tel: +82-31-787-8125; E-mail: jinho.kim@snubh.org

Jieun Lee, PhD.

Senior Researcher, Seoul National University Bundang Hospital

172 Dolma-ro, Bundang-Gu, Seongnam-si, Gyeonggi-do, 13620, Republic of Korea

Tel: +82-31-787-8461; E-mail: jieunlee.snubh@gmail.com

ARTICLE IN PRESS

Abstract

Gastric cancer peritoneal metastasis (GCPM) is a distinct clinical entity with a poor prognosis, characterized by aggressive features and limited treatment options. Understanding its molecular biology is critical for developing effective therapies. We performed whole genome and transcriptome sequencing on GCPM samples and partially paired primary gastric cancer (GC) tissues from 14 and 26 patients, respectively. Our analysis reveals substantial intra-patient heterogeneity between GCPM and primary tumors at both genetic and functional levels. Inter-patient variability is observed in mutational overlaps, with some signatures unique to either GCPM or primary tumors. Tumor evolution analysis suggests divergent clonal evolution, with distinct clones specific to GCPM or primary tumors in most patients. Transcriptomic profile reveals a high heterogeneity in the primary tumors, with five tumors presenting GCPM characteristics. The tumor microenvironment (TME) is poorly conserved between primary GC and GCPM, with desert-type primary tumors often transitioning to immune-enriched TMEs in metastases. These findings suggest that immunotherapy resistance in GCPM may arise from factors beyond intrinsic TME characteristics, such as limited drug delivery due to the peritoneal-plasma barrier. Collectively, our results highlight significant molecular and TME heterogeneity between GCPM and primary tumors, emphasizing the need for GCPM-specific stratification and innovative treatment strategies to improve outcomes.

Keywords: Peritoneal metastasis, Molecular profiling, Tumor microenvironment

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide and the fourth leading cause of cancer-related death, with a particularly high prevalence in East Asia ¹. Despite advances in early diagnosis and treatment, metastatic gastric cancer continues to have a remarkably poor prognosis. Among metastatic patterns, gastric cancer peritoneal metastasis (GCPM) represents a distinct clinical entity characterized by aggressive features and a median survival of less than one year ². The peritoneum is a frequent site of metastasis, occurring in nearly one-third of patients at diagnosis, and is associated with significant morbidity and poor outcomes ². Even after curative resection, peritoneal metastases are the most common type of recurrence, contributing to treatment failure and poor survival ³. In cases of macroscopic peritoneal carcinomatosis, the prognosis is particularly dismal, with a median overall survival of only 3 to 6 months ⁴.

GCPM presents unique challenges in clinical management. Conventional imaging techniques often fail to detect peritoneal metastases, necessitating advanced staging modalities such as diagnostic laparoscopy or cytology washing for accurate diagnosis ⁵. Furthermore, the biological and molecular characteristics of GCPM remain poorly understood due to the limited availability of tissue samples. Metastasis to the peritoneum is thought to result from the direct seeding of cancer cells into the peritoneal cavity, facilitated by its immunosuppressive microenvironment, which resembles other immune-privileged niches such as the brain and bone ^{6,7}.

Current treatment options for GCPM offers limited benefits. Conventional systemic therapies show little efficacy, and while emerging strategies combining systemic and peritoneal-directed treatments, such as intraperitoneal chemotherapy, have shown promise ⁸, immunotherapy-based approaches have largely failed. For instance, combination therapies using nivolumab in the CheckMate 649 and ATTRACTION-4 trials improved outcomes for liver and lymph node metastases but did not benefit patients with GCPM ^{9,10}. This lack of response highlights the need for peritoneal-directed treatment strategies distinct from those for metastases to distant organs [8]. To date, the reasons for the poor responsiveness of GCPM to immunotherapy remain unexplained. Adding to this complexity, gastric cancer exhibits high genotypic heterogeneity ¹¹, further complicating the development of effective therapies. These challenges underscore the urgent need to better understand the molecular underpinnings of GCPM and its relationship with the primary tumor to inform targeted therapies.

In this study, we address these knowledge gaps by analyzing whole-genome and whole-transcriptome sequencing data from 26 patients with advanced gastric cancer, including paired samples from primary gastric cancer tumors and GCPMs. Our findings reveal substantial molecular heterogeneity between

primary tumors and GCPMs across patients, a subset of primary tumors exhibiting transcriptomic profiles that closely matched the GCPM samples. Notably, we observed that desert-type primary tumors often transitioned to other tumor microenvironment (TME) types in GCPMs, particularly immune-enriched subtypes. This discovery has important implications for the development of tailored treatment strategies, suggesting that GCPMs may, in some cases, be responsive to immunotherapy if delivered effectively to the peritoneal cavity.

Results

We utilized a cohort of 26 patients with advanced gastric cancer (AGC) who underwent surgical resection to perform an in-depth molecular characterization of primary and metastatic gastric cancer tumors. Detailed clinical and sequencing information is provided in Supplementary Data 1. The cohort overall survival Kaplan Meier's curve and Swimmer plot are provided in Supplementary Fig. 1a, b. The median survival of patients was 19.3 months. The study design incorporated whole-genome sequencing (WGS) and whole-transcriptome sequencing (WTS) of multi-site tumor samples to capture both genetic and functional alterations associated with within-patient heterogeneity during gastric cancer progression. WGS data was available for 14 patients (GC01-14), while WTS data was available for all 26 patients (GC01-26). The following sample types were collected: whole blood samples (WBS; WGS: 13), normal tissue adjacent to the tumor (N; WTS: 6), primary gastric cancer tumors (Primary; WGS: 9, WTS: 19), peritoneal seeding metastatic tumors (GCPM; WGS: 12, WTS: 26), omentum metastatic tumors (Omentum meta.; WGS: 1, WTS: 2), and lymph node metastatic tumors (LN meta.; WGS: 1, WTS: 2). Paired primary-metastatic samples were available for 6 patients (WGS) and 18 patients (WTS), respectively (Fig. 1a).

Heterogeneity in the oncogenomic landscape of advanced gastric cancer

We first analyzed the global genomic characteristics of our cohort to establish a foundation for understanding the intra- and inter-patient heterogeneity in primary and metastatic tumors. We then explored the mutational profiles specific to primary tumors and metastases, allowing to explore the interplay occurring between tumor sites during cancer progression. A total of four patients (GC01, GC02, GC03, and GC10) harbored the R20G oncogenic germline variant in the *PMS2* gene. Germline *PMS2* variants are low-penetrance Lynch syndrome alleles, conferring significantly lower lifetime cancer risk than other mismatch repair genes. Although *PMS2* variants elevate colorectal and endometrial cancer risk, the associated gastric cancer risk may not substantially exceed that of the general population¹². Therefore,

the specific contribution of this pathogenic variant to gastric carcinogenesis in these patients remains uncertain.

Tumors were classified according to the four molecular subtypes from The Cancer Genome Atlas (TCGA), except for one patient for which the EBV status was not available. No evidence of microsatellite instability (MSI) or Epstein-Barr virus (EBV) were found in any of the patients from the WGS cohort. Most of the primary tumors were classified as genomically stable (GS: $n = 8$; defined based on a fraction of genome altered (FGA) > 0.3), the remaining 5 patients being classified as chromosomal instability (CIN) (Supplementary Data 1). The GS subtype was previously reported to be associated with peritoneal and ovarian metastases and aligns with our results¹³.

Mutational analysis identified 1,594 somatic non-synonymous variants across the 14 patients, with high inter-patient heterogeneity (Fig. 1b). To place this mutational burden in context, we compared our findings with the TCGA-STAD cohort. The mean number of mutations in our chromosomally unstable (CIN) samples (101.4) was highly comparable to that of the TCGA CIN cohort (111.8). Our genomically stable (GS) samples showed a mean of 43.6 mutations, which is also within a similar range as the TCGA GS cohort (72.9), confirming that the overall mutational load in our study is consistent with larger gastric cancer datasets. Missense variants ($n = 1,401$) were the most common, followed by nonsense, splice site, frameshift deletion, frameshift insertion, in-frame deletion, nonstop and in-frame insertion mutations ($n = 77, 45, 35, 21, 11, 3$ and 1 , respectively). While the total mutation count, possibly masked by a high inter-patient variability, did not differ statistically between primary tumors and metastases (Supplementary Fig. 2d), a consistent pattern emerged in paired samples: metastatic lesions were found to systematically exhibit fewer mutations than associated primary tumors (paired t-test, $P = 0.004$) (Fig. 1b). Similar results were reported by Lee et al.¹³ and it was hypothesized that tumor seeding in a sterile environment like the peritoneal cavity, under a lesser immune surveillance, might require less aggressive genomic changes.

To understand the functional consequences of divergent clonal evolution, we performed an Over-Representation Analysis (ORA) on gene sets with mutations exclusive to either the primary or metastatic branches. The analysis of GCPM-specific mutated genes revealed a significant enrichment ($FDR < 0.05$) of Gene Ontology terms associated with microtubule-based processes and nervous system development, which are implicated in cell motility and invasion (Fig. 1c, Supplementary Data 2). Conversely, primary-specific mutated genes were significantly enriched in pathways related to cell-cell communication, suggesting a potential biological role in the establishment and maintenance of the primary tumor microenvironment.

Our results also revealed substantial inter-patient variability in the degree of similarity of genomic alterations between paired primary tumor and metastases: while some patients demonstrated high mutational overlap between primary and metastatic tissues, others exhibited high genetic divergence (Fig. 1d). The proportion of shared mutations was particularly low (6 %) in the GC01 primary tumor sample, while as much as 90% of the mutations were shared between the lymph node metastasis and the primary tumor for the GC08 patient, suggesting a conservation of the genetic profile during metastatic spread. Of note, the highest degree of similarity in mutations (99 %) was observed between the GC07 patient's peritoneal and omental metastases, indicating a close genetic background for these peritoneal cavity metastases.

Driver genes alterations across gastric cancer primary tumors and metastases

To further decipher the mechanisms driving metastatic seeding, we explored nonsynonymous mutations in consensus pan-cancer and gastric cancer driver genes (Fig. 1b). The *TP53*, *GNAQ*, and *CDH1* genes were the most frequently mutated, each observed in 33% of patients. These genes were followed by *ERBB3* and *ARID1A*, mutated in 19% of patients, and *RHOA* and *CREBBP*, mutated in 14% of patients. Some driver genes exhibited diverse variant types across patients. For instance, *TP53* mutations included missense, frameshift, and nonsense variants. The mutation frequency of the driver genes was not statistically significant between GCPM and the primary tumor for any of the driver genes (Fischer's test). An analysis of mutation patterns revealed several informative trends. Certain genes, including *TP53*, *RHOA*, and *ERBB3*, were consistently mutated across all samples from individual patients (Fig. 1b). In contrast, the mutational status of genes such as *GNAQ*, *CREBBP*, and *CDH1* showed inter-patient variability, with mutations either present in all samples from a patient or restricted to specific samples. Additionally, tumor-type-specific mutations were observed. Mutations in *KMT2D*, *FUBP1*, and *ARID1A* were exclusive to primary tumors. Interestingly, despite fewer nonsynonymous mutations in peritoneal metastases compared to their primary counterparts, de novo mutations in the *MAP2K1* and *INPPL1* cancer driver genes were identified exclusively in metastatic samples. However, due to the limited size of our dataset, it remains unclear whether these patterns would hold true in larger cohorts.

A closer examination of individual patients revealed further insights. In patient GC05, a *TP53* nonsense mutation emerged during recurrence (samples primary 2 and GCPM 2), while *CDH1* mutations remained remarkably consistent across all samples, highlighting heterogeneity in mutational dynamics within a single patient. In contrast, metastatic seeding through similar mechanisms, such as peritoneal and omentum metastases in patient GC07, displayed a homogeneous mutational profile. These findings

underscore the complex interplay of conserved and divergent mutational patterns during gastric cancer progression and metastasis.

Mutational signatures across gastric cancer primary tumors and metastases

Mutational signature analysis identified four single-base substitution (SBS) signatures. The SBS1 and SBS5 are clock-like signatures known to be prevalent in cancer, particularly gastric cancer (Fig. 1e). Additionally, the SBS9 signature, potentially associated with polymerase eta hypermutation, and the SBS40 signature, of unknown etiology, were detected in a subset of samples. Although the proportion of these signatures did not significantly differ between primary and metastatic tumors, a trend indicated a higher contribution of the SBS5 signature in primary tumors (Supplementary Fig. 2e), a finding that contrasts with earlier reports¹³. Notably, the mutational signature profiles were not consistently replicated across patients, highlighting the molecular heterogeneity of gastric cancer.

We next focused on analyzing mutational signatures in subsets of mutations: those consistently present across all samples from a patient (labeled as common), those specific to primary tumors, and those specific to metastases (Fig. 1f). Only the SBS1, SBS5, and SBS40 signatures were consistently identified in these subsets. The distribution of these signatures varied across patients: they were mutational signatures specific to metastases in four patients, specific to primary tumors in two patients, and specific to common mutations in one patient.

Clonal evolution in metastatic gastric cancer

To unravel the clonal evolution of metastatic gastric cancer, we implemented PyClone-VI¹⁴ for clone identification and CITUP for reconstructing patient-specific phylogenies. Tumor evolution analysis revealed a predominantly divergent clonal landscape, with distinct clones specific to gastric cancer peritoneal metastases or primary tumors in most patients. This suggests a divergent evolutionary process of primary tumors and peritoneal metastases following initial seeding (Figure 2, Supplementary Fig. 3). Most patients exhibited polyphyletic evolution, with only patients GC02 and GC03 showing a monophyletic pattern in both samples. Patient GC05 provided a good example of clonal complexity (Fig. 2a). In this case, the post-treatment peritoneal seeding sample (GCPM2) was predominantly composed of clones carrying *CDH1*, *TP53*, and *INPPL1* mutations. Notably, clones from the yellow branch harboring *KMT2C*, *KMT2D*, and *GNAQ* mutations were largely absent in GCPM2. The violet clone, initially present in the GCPM1 sample and lacking known gastric cancer driver gene mutations, disappeared entirely after treatment.

Patient GC01 demonstrated an even more pronounced divergence in regional profiles, with minimal mutation overlap between primary tumor and peritoneal seeding samples (Fig. 2b). Each phylogenetic tree branch contained unique gastric cancer driver mutations. For instance, the green branch harboring the *CREBBP* mutation was entirely absent in the GCPM sample, while the GCPM comprised a terminal clone from the violet branch carrying a non-synonymous *CDH1* variant—a branch that represented only a minor population in the primary tumor.

Patient GC07 exemplified yet another evolutionary pattern, with two metastases—omental and peritoneal seeding—sharing a high degree of mutational similarity (Fig 2 c). All key gastric cancer driver genes were localized in the phylogenetic tree trunk. A sample-specific branch was identified in the omental metastasis, though the absence of the primary tumor sample limited definitive conclusions about its origin.

Copy number variations across primary tumors and GCPM

Significant copy number variations (CNVs) were identified in the cohort using the GISTIC2 algorithm¹⁵. Overall, paired primary gastric cancer (GC) tumors and gastric cancer peritoneal metastasis (GCPM) samples exhibited a relatively low concordance in CNV profiles (Fig. 3a, heatmap, Supplementary Fig. 4).

We identified a total of 18 significant focal CNVs containing 56 unique genes (Fig. 3b). Among the seven amplified cytobands, the recurrent amplification at 7p11.2 encompassed the *EGFR* oncogene (observed in patient GC02 primary and GCPM, GC13 GCPM and GC08 primary tumor and lymph node metastasis). Among the 11 deleted cytobands, the deletion at 1q21.1 contained the *BCL9* gene (found in patient GC04 primary tumor and GCPM samples). Genes residing within these significant focal CNV regions were enriched in GO terms related to the maintenance of gastrointestinal epithelium (Supplementary Data 3). All identified significant focal CNVs, except for one deletion at 21q22.3 (which involves the genes *TFF1*, *TFF2*, *TFF3*, *ABCG1*, *UBASH3A*, and *TMPRSS3*), were shared between the primary and GCPM samples and presented similar alteration frequencies.

Structural variants in gastric cancer progression

Structural variants play a key role in cancer onset and progression. We leveraged our cohort's WGS data to complement our understanding of the genomic architecture of advanced gastric cancer.

After filtering, we identified a total of 914 structural variants distributed across 12 tumor samples, comprising 569 deletions, 193 insertions, and 152 tandem duplications (Fig. 3c). The structural variant landscape demonstrated both inter-patient and intra-patient heterogeneity, with some samples for which

SVs could not be detected and, when detected, with number of SVs ranging from as few as 7 SVs in patients GC10 and GC11 to 298 variants in the GC07 omental metastasis sample.

The size distribution of structural variants revealed contrasting patterns according to SV type: deletions exhibited a trimodal distribution across 1.2 kb, 63.0 kb, and 13.7 Mb, tandem duplications showed a bimodal distribution centered around 25.6 kb and 14.5 Mb, and inversions displayed similar bimodal distributions (3.3 kb and 20.7 Mb) (Fig. 3d) similar as reported in previous studies¹⁶. Deletion sizes distribution in peritoneal seeding metastases differed from that of the primary tumors (Kolgorov-Smirnov test, $P = 0.003$), with a higher proportion of SVs < 10 Kbs in PM. Contrarily to Saito et al.¹⁶ who observed a higher VAF in small scale (around 100 kb) SVs in gastric cancer, indicating early acquisition timing of small size variants, correlation between SVAF and SV size was not found to be significant on our data (Fig. 3d).

The structural variant allele frequency (SVAF) was computed as per Saito et al. to adjust the VAF taking into account the tumor purity and genomic complexities such as loss of heterozygosity and amplified regions. Deletions and duplications demonstrated significantly higher SVAF in peritoneal seeding metastases compared to primary tumors, potentially indicating an early contribution of SVs in metastatic cell formation (Fig. 3e).

To illustrate within patient's regional dynamics, we investigated the structural variants profiles in the 2 patients for which paired SVs information was available. Almost all structural variants in the GC07 patient peritoneal seeding metastasis were common to the omentum metastasis-a pattern mirroring the mutation analysis (Fig. 3f, g).

Our analysis identified a total of 853 unique genes with coding sequence regions (CDS) disrupted by SVs, with the largest single SV spanning 90 genes. Among these SVs, 103 were classified as harboring potential pathogenic variants, impacting five established gastric cancer driver genes. More precisely, patient GC08 harbored frameshift SVs disrupting both the oncogene *EGFR*, that was also found from GISTIC CNV analyses, and the tumor suppressor *NCOR1*. Importantly, these *EGFR* and *NCOR1* SVs were independently supported by co-occurring gene fusions involving the same respective genes (Supplementary Data 4). Patients GC12 and GC11 demonstrated frameshift SVs in *DMD* and the epigenetic regulator *EP300*, respectively, further highlighting inter-patient heterogeneity. Despite the identification of these driver-associated SVs, none of the 312 gene fusions detected across the cohort were classified as clinically actionable, underscoring the necessity for future functional and clinical validation of these novel SV events.

The cohort's SV analysis revealed a substantial compartmentalization across tumor types. Of the 853 unique genes with SVs, 36 were shared between primary tumors and GCPM, while 508 were GCPM-specific. Over-Representation Analysis (ORA) highlighted shared genes in vascular permeability and enriched GCPM-specific genes in pathways like Hedgehog signaling, TLR signaling, and NK cell mediated cytotoxicity (Supplementary Data 5). Using a recurrence filter (SV replicated in at least one patient), we identified 58 recurrent genes (25 GCPM-specific, 0 primary-specific). ORA on the recurrent GCPM genes was the most informative, revealing enrichment for 'gastrointestinal smooth muscle contraction' and 'G protein receptors class A rhodopsin-like' terms (Supplementary Data 6), underscoring the functional pressure unique to the metastatic site.

Our findings on the WGS data highlight the substantial individual variability in the genetic mechanisms driving gastric cancer metastasis, underscoring the importance of genomic profiling in understanding tumor progression. To complement our whole-genome sequencing (WGS) analysis, we performed RNA sequencing to characterize the cohort's transcriptional changes associated with metastatic progression in gastric cancer. This approach enabled us to delineate global functional alterations in cellular processes across adjacent normal tissues, primary tumors, and metastatic lesions, as well as to define intra-patient heterogeneity by comparing tumor microenvironment between primary tumors and their corresponding metastatic lesions. By integrating these molecular insights, we aim to refine patient stratification and advance personalized treatment strategies.

Functional alterations in advanced gastric cancer primary tumors and peritoneal metastases

To systematically explore functional alterations underlying gastric cancer progression, we first characterized the global gene expression profiles of all samples. Principal component analysis of the WTS data revealed a clear separation along the first principal component (PC1), defining a gradient from adjacent normal samples, through primary tumors, to the peritoneal metastases (Fig. 4a). Notably, in the paired primary-normal samples, the normal tissue consistently mapped further along the PC1 axis than its matched primary tumor (Supplementary Fig. 5a). This initial analysis underscored substantial heterogeneity within the primary GC samples, with five primary tumors (from patients GC02, GC03, GC05a, GC16, and GC17) clustering more closely with the GCPM samples. Subsequently, K-means clustering segregated the samples into two major groups (C1 and C2): all GCPM samples clustered together in the C2 group and were distinct from the normal samples (C1 group) (Supplementary Fig. 5b,c). The primary tumors were divided into two subgroups, C2-primary samples, which clustered with the GCPM group, and C1-primary samples, which grouped closer to the normal adjacent tissues. This

stratification defined four functional groups (Normal, C1-Primary, C2-Primary, and GCPM) for further differential gene expression (DEG) analysis.

We performed DEG analysis across all six pairwise comparisons of these four groups. As anticipated, the number of differentially expressed genes was highest for contrasts between C1-cluster groups and C2-cluster groups, ranging from 201 to 1,210 (Supplementary Data 7. Supplementary Fig. 6, 7). The comparison between C1-primary and Normal tissue identified only 201 DEGs, a smaller number than typically reported in larger cohorts, yet 63% of these genes were replicated in other contrasts (Fig. 4b). The GCPM vs. C2-primary comparison yielded 210 DEGs, with a notable 89% of these genes being unique to this specific contrast, suggesting distinct molecular mechanisms driving the most advanced metastatic state.

To delineate the functional differences across the identified groups, we performed GSEA across all six contrasts (Supplementary Data 8). This analysis revealed B cell-related immunity and metabolic processes as consistently enriched GO:BP terms across multiple comparisons, suggesting they represent a shared functional hallmark of both primary tumor and metastatic progression in gastric cancer (Fig. 4 c). When examining contrast-specific enrichments, terms associated with development and morphogenesis were enriched when comparing the tumor and metastasis clusters (C1-Primary, C2-Primary, and GCPM) to Normal tissue, highlighting initial transformation and structural changes. In contrast, comparisons between the metastases and primary groups (GCPM vs. C1-Primary and GCPM vs. C2-Primary) were enriched for GO:BP terms related to biological and metabolic processes. The comparison between the two primary subgroups (C2-Primary vs. C1-Primary) yielded enrichment for GO:BP terms related to a response to chemical stimulus and metabolic processes. Furthermore, a comparison of the C2-primary, GCPM, C1-primary groups and Normal groups using curated C2 terms revealed a clear progression towards a more advanced gastric cancer state from the normal to the GCPM, specifically in the "Vecchi et al. gastric cancer advanced vs early up" gene set (Fig. 4d). However, the lack of significant differences on this term between the GCPM and C2-primary groups indicates that the C2-primary tumors already exhibit advanced cancer characteristics at the transcriptional level.

A crucial finding from the Hallmark analysis was the clear evidence of a continuous epithelial-mesenchymal transition (EMT) signature, progressing from normal tissue to C1-primary and then to the C2-primary/GCPM clusters. EMT is a key driver of enhanced migratory and invasive potential, consistent with its role in metastatic dissemination. Supporting this, 52 genes from the Hallmark EMT set were differentially expressed in at least one comparison, with *FAP*, *VIM*, and *SPPI* showing graded upregulation along the trajectory from normal tissue through C1-primary to C2-primary/GCPM

(Supplementary Fig. 8a and 8b), indicative of progressive acquisition of mesenchymal traits in a subset of primary tumors and in peritoneal metastases. In addition to EMT, gene sets associated with adipogenesis were enriched in the C2-primary/GCPM group compared with C1-primary/normal, reflecting a functional shift toward a more mesenchymal and metabolically plastic state. To further decipher the rationale of the clustering of C2-primary tumors with GCPM, suggestive of an advanced functional phenotype, we integrated the transcriptomic data with genomic findings. Notably, three of the five C2-primary patients had available whole-genome sequencing data, and these were the only patients whose primary tumors harbored *TP53* mutations, with patient GC05 additionally carrying a *CDHI* mutation. Both *TP53* and *CDHI* alterations have been previously implicated in EMT regulation and metastatic progression, providing a genomic rationale for the transcriptional shifts observed. *KRAS* signaling was also transcriptionally upregulated in C2-primary versus C1-primary samples (including DEGs such as *SPPI*, *ANGPTL4*, and *GPNMB*, Supplementary Fig. 8c,d), despite none of the tumors carrying a *KRAS* mutation. This indicates that enhanced *KRAS* pathway activity can arise through transcriptional or upstream regulatory mechanisms rather than direct mutation, further reinforcing the invasive and mesenchymal-like phenotype of the C2-primary tumors.

Together, these transcriptomic and genomic features delineate a trajectory of progressive EMT and associated functional changes, linking specific molecular alterations—including *TP53* and *CDHI* mutations and *KRAS* pathway activation—to the metastatic potential observed in this cohort.

Deconvolution reveals contrasting immune landscape between GC primary tumor and metastases

To further characterize regional tumors, we estimated the proportion of cell types present in our samples through bulk RNAseq deconvolution using the widely established CIBERSORTx and EPIC methods^{17, 18}. To account for cell populations specific to gastric tissue and peritoneal seeding metastases that are not represented in these reference databases, we constructed a custom reference dataset using 2 publicly available single-cell RNA-seq datasets—GSE183904 and GSE163558—comprising normal gastric tissue, primary gastric cancer (GC), and gastric cancer peritoneal metastasis (GCPM) samples. (Supplementary Data 9, 10, Supplementary Fig. 9, 10, 11). This dataset was subsequently applied to the Bayesian deconvolution tool BayesPrism¹⁹ for refined cell type estimation (Supplementary Fig. 12). To enable a direct comparison between EPIC and BayesPrism, we aggregated their respective heterogeneous cell populations into five common, general cell types. This standardization allowed us to assess the concordance between the proportional estimates from both methods (Supplementary Fig. 13).

Consistent with our gene expression findings, deconvolution analysis revealed distinct immune landscapes between primary gastric tumors and peritoneal metastases (Fig. 4 a, Supplementary Fig. 4a, b,

c). Metastatic samples exhibited a paradoxical immune profile, characterized by heightened immunosuppression—evidenced by elevated proportions of M2 macrophages and reduced natural killer (NK) cell populations—alongside potential immune activation, as indicated by an increased presence of naïve CD8⁺ T cells. Monocytes were also more abundant in peritoneal metastases compared to primary tumors. The reduced proportion of plasma cells in peritoneal metastases suggests a weakened humoral immune response, aligning with prior gene set enrichment analysis results and pointing to diminished adaptive immune capacity in the metastatic microenvironment.

The inflammatory microenvironment also underwent significant remodeling during metastatic progression. A shift in mast cell activation states was observed, with peritoneal metastases exhibiting a higher proportion of resting mast cells and fewer activated mast cells compared to primary tumors. This alteration in the inflammatory milieu suggests potential differential impacts on immune regulation and tumor growth mechanisms in each region.

Similarly, the stromal compartment demonstrated substantial heterogeneity between primary tumors and metastases. Peritoneal metastases showed an increased proportion of cancer-associated fibroblasts (CAFs) but a decreased population of regular fibroblasts relative to primary tumors.

As expected, primary tumors retained a higher proportion of enterocytes and Pitt mucous cells—specialized cells responsible for producing the protective mucus layer shielding the gastric epithelium from digestive enzymes—compared to metastatic samples.

Collectively, these differences in cellular composition suggest that distinct tumor-supporting mechanisms operate in primary tumors and peritoneal metastases, which may differentially modulate immune responses and tumor progression.

Cell type–specific transcriptomic alterations associated with peritoneal metastasis

To determine which cell types might underlie the functional alterations observed in our cohort, we leveraged the public single cell RNA-seq used for deconvolution. Differential expression analyses were performed between these three groups in a cell type–specific manner. Because the number of cells per sample was insufficient for pseudo-bulk aggregation (Supplementary Data 10), we employed a coarse-grained annotation that categorized cells into five major lineages: epithelial, stromal, myeloid, lymphoid, and B cell–plasma.

We next assessed the overlap between these cell-wise DEGs and the bulk DEGs identified in our dataset. When the overlap was statistically significant (hypergeometric test) and comprised more than ten shared genes, we conducted over-representation analysis (ORA) to identify enriched pathways associated with these cell-type–specific signatures.

Interestingly, the cell-wise DEG analysis of the public datasets revealed only five DEGs in primary GC versus normal, four of which belonged to the B cell–plasma cluster (Supplementary Data 11). In contrast, GCPM versus normal and GCPM versus primary GC comparisons yielded 217 and 272 DEGs, respectively. These numbers remain modest compared with bulk analyses, likely reflecting the lower sensitivity of single-cell datasets, limited sample representation, and higher biological variability across cells. Most DEGs were found in epithelial and stromal compartments, with minimal overlap between different cell types (Supplementary Fig. 14), suggesting distinct lineage-specific transcriptional reprogramming.

Significant overlaps between our bulk GCPM versus normal DEGs and cell-wise GCPM versus normal DEGs were observed across all five cell types (Supplementary Data 12, Supplementary Fig. 15a,b,c). ORA of these shared gene sets revealed that epithelial, stromal, and B cell–plasma populations converged on pathways related to immune responses and gastric epithelium maintenance, whereas epithelial cells were specifically enriched for ion transmembrane transport, and stromal cells showed enrichment for tyrosine kinase activity. This pattern suggests coordinated epithelial remodeling and stromal activation that may jointly facilitate tumor progression and metastatic colonization. For the GCPM versus primary GC comparison, only the stromal cell overlap was statistically significant with the GCPM versus C1-primary bulk DEGs, with enriched GO terms again primarily associated with immune processes (Supplementary Data 13, Supplementary Fig. 15d,e). This finding is consistent with our bulk transcriptomic results and suggests that the functional shifts accompanying metastasis are largely driven by stromal gene expression changes, stromal cell types in which the proportion of cancer-associated fibroblasts (CAFs) has been estimated to be higher in GCPM compared with primary tumors in the deconvolution analysis.

Tumor microenvironment in advanced gastric cancer

Using Molecular Functional Portrait (MFP)²⁰ analysis, we then classified the tumor microenvironment (TME) in both primary GC tumors and GCPMs. While the proportional differences in TME types between the primary tumors and the GCPMs were not statistically significant, primary tumors were predominantly characterized by desert and fibrotic microenvironments (Fig. 5b). In contrast, GCPMs exhibited greater diversity, with immune-enriched and immune-enriched fibrotic environments representing approximately half of the samples. The five C2-primary samples presented various TME types: GC02 and GC05a were immune-enriched fibrotic, GC03 was desert and GC16 and G17 were fibrotic, and the TME type similar between the primary and the GCPM samples in the GC16 and GC17 samples only.

As displayed on the Sankey diagram (Fig. 5c), the tumor microenvironment (TME) was not consistently conserved between primary tumors and GCPM, with only 3 of the 18 analyzed sample pairs retaining the same TME type. This observed divergence, however, was not statistically significant (McNemar's test, $P > 0.05$). In patients for which the TME was not conserved between the primary tumor and the GCPM, the transition from a "desert" TME in the primary tumor to other TME types in the GCPM was significantly more frequent than the reverse direction of transition (mid-P McNemar test, $P = 0.039$). These findings suggest that the tumor microenvironment undergoes significant remodeling during metastasis, with a notable tendency for "desert" TMEs in primary tumors to transition into other TME types in GCPM, underscoring the adaptability and regional heterogeneity of the TME in gastric cancer progression.

Ecotype and cell state profiling in gastric cancer progression

To further characterize our cohort and link observed cell types with potential clinical outcomes, we recovered TCGA-defined cell states and ecotypes within our dataset. The most prevalent ecotypes were CE1, CE10, and CE6, representing 24%, 20%, and 15% of samples, respectively (Fig. 6a). These ecotypes were associated with relatively favorable survival rates in the EcotyperSTAD gastric cancer cohort, as indicated by z-scores of 1.60, -0.02, and 2.22 for CE1, CE10, and CE6, respectively.

While no statistically significant differences in overall ecotype proportions were observed between primary gastric cancer and peritoneal metastases, notable variations emerged in the proportions of endothelial and epithelial cell states (Fig. 6b, c, Supplementary Fig. 16). Endothelial S04 and epithelial S02, which were more prevalent in peritoneal seeding samples, correlated with higher survival rates in the EcotyperSTAD reference cohort. In contrast, endothelial S01 and epithelial S05, which were more common in primary tumor samples, were associated with poorer survival outcomes.

Cell state conservation across patient samples was remarkably limited. Among the four patients analyzed, endothelial cell states showed no consistent patterns, and epithelial cell states were poorly conserved, with consistency observed in only 3 out of 8 patients. This suggests that cell states are more strongly influenced by tumor type than by individual patient characteristics.

Gene set enrichment analysis of differentially expressed genes associated with endothelial cell states revealed distinct functional roles. S01 cells, predominantly found in peritoneal seeding samples, were enriched in pathways related to extracellular matrix remodeling, cellular plasticity, and vascular regulation, suggesting a role in promoting metastatic spread and survival. In contrast, S04 cells, more prevalent in peritoneal seeding samples, were enriched in pathways associated with thyroid hormone activity, transport processes, pancreatic beta cells, and bile acid metabolism, indicating a potential role in supporting tumor growth and metabolic reprogramming.

Similarly, epithelial cell states exhibited functional heterogeneity. S05, dominant in primary tumor samples, was enriched in pathways maintaining differentiated gastric epithelial functions. In contrast, S02, associated with peritoneal seeding samples, was linked to pathways supporting adaptation, survival in diverse environments, and invasion.

These nuanced differences in immunological profiles, microenvironmental composition, and cellular states highlight the potential need for distinct therapeutic strategies targeting primary gastric tumors versus peritoneal metastases.

Discussion

In this study, we systematically evaluated the mutational and functional characteristics of paired primary gastric cancers and their metastases, with a focus on peritoneal seeding metastases (GCPM). Our findings reveal substantial inter-patient diversity and intra-patient heterogeneity between primary tumors and their associated GCPMs. Notably, tumor microenvironment (TME) types were poorly conserved between primary tumors and GCPMs, with primary tumors often exhibiting a "desert" TME that transitioned into distinct TME types in GCPMs.

We observed high variability in the number of shared mutations between primary tumors and GCPMs, consistent with prior reports comparing primary gastric tumors and lymph node metastases²¹. Although the lymph node and peritoneal environments differ, common mechanisms may underlie the observed heterogeneity. For instance, inter-patient variability in genomic similarity between primary tumors and GCPMs could arise from differences in seeding scenarios, such as the number of contributing clones, post-seeding selection pressures, or the timing of metastatic dissemination. Intriguingly, GCPMs consistently harbored fewer mutations than their paired primary tumors. While the underlying mechanisms remain unclear and warrant validation in larger cohorts, potential explanations include the preferential seeding of clones lacking a subset of primary tumor mutations or continued mutation accumulation in primary tumors post-seeding due to a selective environment.

Whole Genome Sequencing (WGS) enabled a comparison of structural variants (SVs) between GCPMs and primary gastric tumors. SVs contribute to metastasis through mechanisms like gene fusions, extrachromosomal DNA formation, and altering gene function via coding sequence disruption or impacts on non-coding regions. We observed a higher number of SVs and elevated structural variant allele frequencies (SVAF) for deletions and duplications in GCPMs compared to primary tumors, suggesting that SVs might arise early in metastatic cell formation and potentially drive the metastatic phenotype. The fact that, as opposed to point mutations, SVs were more numerous in GCPM than in primary tumors

suggests that SVs may offer more potent mechanisms for metastatic adaptation, or that strong selective sweeps during metastasis reduce SNV diversity by favoring clones defined by key SVs. While our observation of increased SVs prevalence aligns with other research reporting higher SV burdens in metastatic gastric cancer compared to primary tumors²², it contrasts with some prior studies focusing specifically on CNVs, which reported lower overall CNV burdens⁶ or similar CNV patterns in paired samples, albeit with specific enrichments like *DEK* amplification²³. While the reasons for the higher SV prevalence in our cohort require further investigation, possibilities include defective DNA repair pathways or unique selection pressures in the peritoneal microenvironment. Functionally, GCPM-specific SVs were enriched in genes related to Hedgehog signaling, potentially promoting stemness, and TLR signaling, which could influence inflammation and immune evasion. Furthermore, specific alterations, such as frameshift SVs leading to confirmed gene fusions in potential therapeutic targets like *EGFR*²⁴, highlight the clinical actionability these variants may represent for patients with refractory disease.

Our analysis uncovered distinct mutational patterns in metastatic gastric cancer compared to previous studies. Notably, non-synonymous variants in the tumor suppressor gene *TP53* were detected in only one-third of our cohort, contrasting with the ~50% prevalence reported in Asian gastric cancer cohorts²⁵. In contrast, mutations in *CDH1*, which encodes the cell–cell adhesion molecule E-cadherin, and in *GNAQ*, a regulator of cell growth and survival, were detected in 33% of our cohort—substantially higher than the 9–10% prevalence typically reported for these genes^{26,27}. Notably, two patients harbored concurrent mutations in *CDH1* and *TP53*, a combination reported as mutually exclusive in some studies^{6,27} but also observed in peritoneal metastases where both genes, along with *TAFI*, are frequently co-inactivated²⁸.

CDH1 loss is a hallmark of diffuse-type gastric cancer and the TCGA-defined genomically stable subtype, both associated with poor prognosis²⁸. Inactivating *CDH1* mutations disrupt epithelial integrity and promote epithelial–mesenchymal transition (EMT), enhancing tumor cell motility and metastatic dissemination, particularly to the peritoneum^{6,23,29,30}. Recent evidence suggests that *CDH1* mutations in peritoneal metastases are often clonal and exhibit higher variant allele frequencies than *TP53*, supporting the notion that *CDH1* inactivation may occur early and drive metastatic potential²⁸. Moreover, the co-occurrence of *CDH1* and *TP53* mutations is unlikely to be coincidental; rather, it may reflect positive selection for dual pathway disruption during metastatic evolution. Functional studies indicate that loss of E-cadherin creates dependencies—such as hyperactivation of the *IGF1R* pathway—that could represent potential therapeutic vulnerabilities²⁸. Together, these findings suggest that *CDH1* alterations, alone or in concert with *TP53* and related regulators such as *TAFI*, may define a subset of metastatic gastric cancers characterized by early EMT activation and peritoneal tropism. Clonality analysis revealed that these

mutations occurred subclonally, with higher cellular prevalence in GCPMs than in primary tumors. This suggests that clones harboring *CDHI* mutations may have a greater contribution to metastatic seeding through EMT-mediated mechanisms.

Functional analysis highlighted substantial heterogeneity within primary tumors. We identified a subset of five primaries whose gene expression profiles resembled GCPMs, featuring stronger EMT and KRAS signaling signatures compared to the others. Interestingly, *TP53* mutations were found in three available samples from this subset, potentially indicating more aggressive tumors with pre-existing metastatic-like traits, despite diverse microenvironments. The remaining primary tumors were functionally closer to normal gastric tissue, showing enrichment for growth and metabolism pathways but reduced immune activity. While GCPMs shared this general immune downregulation, they were further distinguished by reduced B cell activation, a more advanced phenotype, and enhanced mesenchymal characteristics compared to primaries. The progressive activation of EMT pathways across normal tissue, primary tumors, and GCPMs emphasizes EMT's crucial role in enabling the motility and invasion required for peritoneal metastasis.

Due to limited tissue availability, few studies have comprehensively analyzed the tumor microenvironment (TME) of GCPMs. Recently, Zhao et al.⁶ conducted a study on 55 paired primary-GCPM samples, reporting that peritoneal metastases exhibit a higher proportion of inflammatory cancer-associated fibroblasts (iCAFs), myofibroblast-like CAFs, M2 macrophages, and monocytes, alongside a decrease in CD4⁺ T cells and plasma B cells compared to primary tumors. While our deconvolution results similarly indicated an increased proportion of cancer-associated fibroblasts (CAFs), M2 macrophages, and monocytes in GCPM, alongside a reduction in plasma B cells, we observed a different T cell pattern. Specifically, we noted an enrichment of naïve CD8⁺ T cells in GCPM compared to primary tumors, diverging from Zhao et al.'s⁶ findings related to CD4⁺ memory resting cells. Tumor infiltration by naïve T cells, regardless of subset, is generally associated with a poorer prognosis as it indicates a less effective antitumor immune response³¹. The specific enrichment of naïve CD8⁺ T cells in GCPMs suggests the metastatic site may lack sufficient signals for proper activation and differentiation into cytotoxic T lymphocytes (CTLs), such as adequate antigen presentation, co-stimulation, or helper cytokines, thus impeding the direct tumor-killing capacity of the immune system and likely contributing to immune evasion and disease progression.

Beyond immune populations, the epithelial composition of gastric tumors, including pit, mucosal, chief, or other gastrointestinal cell types, also strongly correlates with patient survival. In our cohort, primary tumors displayed higher enrichment in enterocytes and KLF2-positive pit mucous cells compared to their

GCPM counterparts, a difference potentially linked to epithelial-to-mesenchymal transition (EMT) during metastatic progression. When attempting to trace the cellular origins of functional differences, preliminary analysis suggested that epithelial cell remodeling and stromal activation distinguished GCPMs from normal tissue. In contrast, shifts between primary tumors and GCPMs appeared predominantly driven by modifications in stromal cell immune pathways. Although requiring further validation, these findings suggest distinct contributions from both epithelial and stromal compartments during peritoneal metastasis. Although our cohort size limited statistical comparisons, our findings align with Zhao et al. in showing that GCPM TMEs are less "deserted" and more fibrotic and immune-enriched than primary tumors. This suggests that niche reprogramming in the peritoneum during metastasis may drive inflammation and immune infiltration ⁶. We also observed poor conservation of the TME between paired primary gastric cancers and GCPM, indicating a potential disruption in communication between primary tumors and metastases upon colonization of the peritoneal milieu. This highlights the critical importance of evaluating the molecular subtype of peritoneal metastases when designing therapeutic strategies. Given that metastatic lesions are typically the primary cause of patient mortality and have a stronger prognostic association than primary tumors [10], their molecular characteristics may hold greater prognostic value and should be prioritized in clinical decision-making.

Recent clinical trials, such as CheckMate 649 and ATTRACTION-4, have explored the use of combined chemotherapy and immunotherapy with nivolumab in advanced gastric cancer. While these trials demonstrated improved prognosis for metastases to the liver or lymph nodes, they did not show similar benefits for GCPMs ^{9,10}. This discrepancy could hypothetically be attributed to a lack of responsiveness of GCPMs to immunotherapy. However, our findings challenge this assumption, as a notable proportion of GCPMs in our study exhibited immune-enriched or immune-enriched fibrotic tumor microenvironments (TMEs), suggesting that immunotherapy resistance in some patients may stem from factors other than intrinsic TME characteristics.

Unlike the highly vascularized gastric epithelium, the peritoneal cavity is relatively avascular, with a peritoneal-plasma barrier that restricts the exchange of molecules between compartments ³². This barrier may prevent systemically administered chemotherapy and immunotherapy from effectively reaching GCPMs, even in patients whose TME profiles suggest potential responsiveness to immunotherapy. Interestingly, our results revealed that GCPMs exhibit increased enrichment of angiogenesis-related gene expression compared to primary gastric tumors. This likely reflects the metastatic tumors' heightened need for vascularization to support their growth and persistence within the nutrient-limited peritoneal cavity, rather than an absolute increase in vasculature.

These findings suggest that, for some patients, direct intraperitoneal delivery of immunotherapy may be a more effective strategy to overcome the peritoneal-plasma barrier. Supporting this approach, intraperitoneal administration of the immune checkpoint inhibitor nivolumab has shown promising clinical responses in 77% of patients with peritoneal ascites secondary to pancreatic, biliary tract, or gastric cancer ³³. Furthermore, developing methods to stratify GCPM patients based on the TME of peritoneal metastases could be critical for tailoring treatment strategies and improving outcomes for this patient population.

Our study has several limitations that warrant acknowledgment. The cohort size was limited, underscoring the need for validation in larger cohorts of gastric cancer patients with GCPM. Additionally, the small number of lymph node and omental metastases in our cohort restricted our ability to compare and extend our findings to other types of abdominal metastases. Regional sampling was also limited, with primary and metastatic tumors typically collected from a single site per patient, potentially constraining our assessment of tumor heterogeneity. Furthermore, all patients in our study presented with advanced-stage cancer and metastases at diagnosis, which may have impacted our ability to accurately reconstruct the timing and history of clonal spreading. These limitations highlight the need for future studies with broader sampling and earlier-stage cohorts to further elucidate the mechanisms underlying metastatic progression. In conclusion, our study reveals substantial intra-patient heterogeneity both between GCPM and primary gastric cancer tumors and within primary gastric cancer tumors at both molecular and tumor microenvironment (TME) levels. We demonstrate that the TME is largely non-conserved between primary gastric tumors and their peritoneal metastases, with desert-type primary tumors frequently transitioning to distinct TME profiles, particularly immune-enriched subtypes, in GCPMs. These findings emphasize the need for innovative strategies to enhance immunotherapy delivery to metastatic sites and highlight the importance of molecular-guided therapeutic approaches to effectively target GCPM.

Methods

Human participants

Demographic information, including sex, age, and basic clinical information of the 26 patients enrolled in this study, is available in Supplementary Data 1. All participants provided written informed consent, and the study protocol was approved by the Institutional Review Board (IRB) of Seoul National University Bundang Hospital (B-1402-240-004). All relevant ethical regulations for research involving human participants were followed.

Cohort and Tissue Sampling

A total of 26 GC patients with peritoneal metastasis were included in this study. Stomach tumor, peritoneal metastasis and adjacent normal mucosal tissues were obtained as surgically resected fresh frozen samples or formalin-fixed paraffin-embedded (FFPE) blocks at Seoul National University Bundang Hospital.

Whole-Genome Sequencing

Tissue samples were reviewed by a pathologist prior to sequencing. DNA was extracted from FFPE blocks or fresh frozen (FF) tissues using the ReliaPrep FFPE gDNA Miniprep System (Promega), the QIAamp DNA Mini Kit (Qiagen), and the QIAamp DNA Blood Mini Kit (Qiagen) for FFPE, FF tissues, and blood, respectively. SureSelect libraries were prepared according to the manufacturer's instructions at Macrogen (Seoul, South Korea). The resulting libraries were then sequenced on the NovaSeq 6000 platform (Illumina, San Diego, CA, USA) with 151 bp paired-end configuration. Tumor specimens were sequenced at an approximate depth of 60X, while normal samples were sequenced at an approximate depth of 30X. Statistics and quality metrics for normal and tumor samples are summarized in Supplementary Data 14. Potential sample swaps and cross-sample contaminations were assessed using Conpair³⁴. The probability of an alternate base call resulting from an oxidation artifact was calculated with the CollectOxoGMetrics tool in Picard (v2.26.3). Following these pre-processing steps, paired-end sequencing data were aligned to the human reference genome (hg19) using the BWA read alignment package³⁵. Resulting BAM files were sorted and indexed using Samtools (v1.12)³⁶. Duplicate reads were removed, base quality scores recalibrated, and multiple sequence alignments realigned using Picard (v2.26.3) and the Genome Analysis Toolkit (GATK) (v4.2.0)³⁷.

Germline Variant Calling. Germline variants were identified using the GATK HaplotypeCaller tool implemented in the Nextflow nf-core Sarek pipeline (v.2.7.1) with default settings^{38–40}. Oncogenic germline variants were annotated using the OncoKB precision oncology knowledge base.

Somatic Variant Calling. Single nucleotide variants (SNVs) and small insertions and deletions (INDELs) were identified using both Mutect2 and the Manta/Strelka2 workflow^{37,41}. Variant calling with Mutect2 (GATK v4.2.0) was performed in tumor-matched normal mode with the population allele fraction for alleles not found in the germline resource was set to 0.0000025. A panel of normal samples and a germline resource were used to filter variants. Germline variants were filtered using the allelic frequency data from the gnomAD v2.1.1 database. An artifact-prior table was generated for each tumor sample to account for orientation bias artifacts, as recommended for formalin-fixed paraffin-embedded (FFPE) samples. Soft-clipped bases in the reads were excluded from analysis. As a complementary approach, variants were also

called using the Manta/Strelka2 workflow implemented in the Nextflow (v22.04.5) nf-core Sarek (v2.7.1) pipeline with paired-normal sequences as a reference. Small indels were first called using Manta and subsequently fed into the Strelka2 algorithm. Variants identified by Strelka2 were filtered according to the high-confidence criteria recommended by the developers.

To optimize the detection of true positive variants, post-filtering was applied as follows: somatic variants retained for analysis were either (i) called by both tools in the same sample or (ii) identified by a single tool in at least two samples from the same patient when multiple samples were available. Additionally, retained variants met the following criteria: variant allele frequency (VAF) ≥ 0.01 , sequencing depth ≥ 20 , and a minimum number of supporting reads (≥ 3 for SNVs or ≥ 5 for indels). The resulting variant call format (VCF) files were converted to mutation annotation format (MAF) using the vcf2maf tool ⁴². Samples with fewer than 20 variants remaining after filtering in the MAF files were excluded from downstream analyses. In total, 22 samples from 14 patients were included in the final dataset for downstream analyses. Over representation analysis was conducted in the somatic variants.

Tumor Purity, Copy Number, and Ploidy Estimation. Tumor ploidy and purity were estimated from whole-genome sequencing (WGS) data using the Sequenza library and the sequenza-utils module with a binning window size of 50 bp ⁴³. The copy number burden was calculated as the sum of base pairs (bp) from the Sequenza segments where $|CN| > 0.3$ for each sample.

Mutational Signatures. Mutational signature analysis was performed following the approach described by Alexandrov et al. ⁴⁴ and implemented using the SigProfilerExtractor module (v1.1.21) with default settings.

Copy-Number Variation. To identify the significant focal somatic CNVs, the GISTIC2.0 (v2.0.23) algorithm was run in the Sequenza segmentation profiles with the parameters ‘-ta 0.3 -td 0.2 -qvt 0.25 -brlen 0.7 -conf 0.90 -js 4 -maxseg 25000 -cap 1.5 -genegistic 1 -broad 1 -armpeel 1 -smallmem 1 -gcm extreme -savegene 1 -rx 0 -v30’. Recurrently amplified or deleted regions were identified by GISTIC peaks and genes overlapping with each peak were kept for further analyses. Over representation analysis was conducted on the identified recurrent CNVs.

Structural Variants. Structural variants (SVs) were identified using the Delly2 method ⁴⁵ with the -q15 option to exclude base pairs with quality scores below 15. Further filtering was applied to retain reads with a mapping quality of ≥ 40 . Additional filtering criteria were based on the number of read variants (RV, corresponding to the SV allele-supporting reads) and discordant variants (DV, corresponding to the SV-supporting paired-end reads): in tumor samples, RV and DV combined had to be ≥ 5 , with RV ≥ 1 , while in normal tissue or blood, both RV and DV were required to be 0. For further control of false

positives, break-end records were excluded from the analysis. Over representation analysis was implemented on the identified SVs.

Clonality Analysis. Clonality analysis was performed for all patients with more than one sample. To incorporate all relevant mutations for clonality determination, when a variant was observed in one sample, its variant allele frequency (VAF) and coverage were extracted for all the patient's samples. This process involved the following steps: (i) creating a BED file encompassing mutations from all individual samples, (ii) extracting supporting reads and coverage information for these positions across all patient samples using Samtools mpileup, and (iii) formatting this information into the input format required by PyClone-VI¹⁴. These operations were executed using a custom pipeline.

The extracted variants, along with their associated major and minor copy numbers, were clustered using PyClone-VI. Clusters containing fewer than five variants were excluded from further analysis. Phylogenetic trees were reconstructed based on the cellular prevalences of the clusters using CITUP⁴⁶. Visualizations of the phylogenies were generated using the R packages mapscape and timescape⁴⁷.

Whole transcriptome sequencing

Sample Selection and Collection. A total of 55 samples were collected from 26 patients, comprising 6 adjacent normal tissues, 19 gastric cancer primary tumors, 26 peritoneal seeding metastases, 2 omental metastases, and 2 lymph node metastases. Among these, 18 gastric cancer primary tumors and their corresponding peritoneal metastases were paired.

RNA Isolation and Quality Assessment. RNA was extracted from FFPE blocks or FF tissues using the ReliaPrep FFPE Total RNA Miniprep System (Promega) and RNeasy Mini Kit (Qiagen) for FFPE and FF specimens, respectively, according to the manufacturer's protocol. Nucleic acid yield and purity were assessed using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Thermo-Fisher Scientific, MA, USA).

Library Preparation and Sequencing. RNA libraries were constructed from the same specimens utilized for WGS and additional tissues (Supplementary Data 15) using a TruSeq RNA Access library prep kit (Illumina, San Diego, CA, USA) or Exome capture kit according to the manufacturer's protocols. Briefly, ribosomal RNA (rRNA) was depleted, followed by fragmentation and conversion of RNA molecule to complementary DNA (cDNA) using reverse transcriptase. The single-stranded cDNAs were then converted to double-stranded cDNAs, end-repaired, adenosine-tailed, and adaptor-ligated. The constructed libraries were amplified using polymerase chain reaction (PCR). Subsequently, cDNA libraries were submitted to an Illumina NovaSeq6000 (Illumina, Inc.), and paired-end (2×100 bp) sequencing was performed by Macrogen Incorporated.

WTS raw reads processing. Raw reads were trimmed using cutadapt (v1.15)⁴⁸ with options -q 15 and --minimum-length 36 to remove adapter sequences and low-quality bases. Processed reads were analyzed using the Nextflow (v23.04.3) nf-core/rnaseq pipeline (v3.12.0)⁴⁹, employing the STAR-RSEM aligner^{50,51} with assembly GRCh37 and default parameters. The pipeline performed quality control, alignment, and generated a matrix of gene-level TPM values across all samples, which was subsequently used for downstream analyses.

Tumor Stromal and Immune Scores. Tumor stromal and immune scores, commonly used to estimate tumor purity, were calculated for the WTS samples using the tidyestimate implementation of the ESTIMATE algorithm (available on GitHub: <https://github.com/KaiAragaki/tidyestimate>). This approach follows the method described by Yoshihara et al.⁵². Since the ESTIMATE algorithm's tumor purity calculations were calibrated specifically for Affymetrix array expression data, only stromal and immune scores could be derived. These scores were utilized to comparatively estimate tumor purity levels across samples, rather than providing absolute values.

Differential Gene Expression Analysis. Due to the limited number of samples from each tissue of origin, differential gene expression analysis was conducted across three experimental groups. First, differential gene expression was analyzed between the gastric primary tumors (n = 19) or the peritoneal seeding metastases (n = 26) versus the adjacent normal tissue (n = 6). Second, to more precisely characterize the heterogeneity in expression profiles between regional samples, differential expression was assessed between gastric primary tumors (n = 19) and peritoneal seeding metastases (n = 26). All differential gene expression analyses were performed using the DESeq2 R package⁵³, which calculates adjusted p-values to control the false discovery rate (FDR) using the Benjamini and Hochberg method. As partial pairing existed between samples from different tissue regions, patient identity was included as a covariate in the experimental design. Genes with an adjusted p-value < 0.05 and $|\log_2\text{FoldChange}| \geq 1$ were considered differentially expressed.

Gene Set Enrichment Analysis. Gene set enrichment analysis (GSEA) was conducted using the clusterProfiler R package⁵⁴. The analysis considered Gene Ontology (GO) terms for biological processes (BP), molecular functions (MF), and cellular components (CC) with a p-value cutoff of 0.01. Additionally, KEGG pathways and curated gene sets from MSigDB (v2022)—including Hallmark, C2, C4, C6, and C7 gene sets—were analyzed using a p-value cutoff of 0.05⁵⁵.

Transcriptomic Bulk Data Deconvolution. Three deconvolution algorithms were employed to infer cell type composition in the study. CIBERSORTx¹⁸ and EPIC¹⁷ were used to estimate the abundances of immune and tumor cell types. Additionally, the Bayesian cell proportion reconstruction method

BayesPrism¹⁹ was applied to infer a broader and more specific range of cell types from the bulk transcriptomic dataset. For BayesPrism, two single-cell RNA sequencing (scRNA-seq) datasets were used as references: the GSE183904 dataset (Kumar et al.⁵⁶), which includes 4 samples of normal tissue adjacent to primary tumors, 8 primary tumors, 3 peritoneal seeding metastases, and 1 sample of normal tissue adjacent to peritoneal metastases, and the GSE163558 dataset (Jiang et al.⁵⁷), comprising 1 normal adjacent tissue, 4 peritoneal seeding metastases, and 2 lymph node metastases. The scRNA-seq datasets were reprocessed and annotated de novo using the Seurat R package (v4.3.0)⁵⁸. Cells with $\geq 8\%$ mitochondrial RNA content were excluded. The datasets were normalized, scaled and integrated using Harmony⁵⁹ to account for batch effects. A two-step hierarchical approach was used for cell type annotation. First, a coarse-grained annotation was obtained by clustering the entire dataset with a resolution of 0.5. This identified five primary cell clusters: B cell - plasma, lymphoid, stromal, epithelial, and myeloid cells. These clusters were annotated based on canonical markers referenced in the original studies. The UMAP projections for this coarse annotation are shown in Supplementary Fig. 9a,b, with defining marker genes displayed in a dot plot (Supplementary Fig. 9c) and listed in Supplementary Data 9. Second, a fine-grained annotation was performed to achieve higher resolution. Each of the five major cell types was isolated and subsequently subclustered. A clustering resolution of 0.3 was used for most lineages. Due to its complexity, the lymphoid compartment was subclustered with a higher resolution of 0.9, followed by a further round of sub-clustering to precisely define cell subtypes. The UMAPs, marker gene dot plots, and marker lists for this fine-grained annotation are displayed in Supplementary Fig. 10, 11, and Supplementary Data 9, respectively. The reprocessed reference dataset comprises 46,007 cells, with cell proportions detailed in Supplementary Data 10.

To assess the consistency between deconvolution methods, cell type estimates obtained from EPIC and BayesPrism were aggregated into five major cellular categories (epithelial, stromal, myeloid, lymphoid, and B/plasma) prior to comparison. The similarity of the cell type proportion were assessed estimating the concordance of the cell type proportions ranks using Spearman's rank correlation coefficient and by quantifying the distance in terms of cell type proportions between samples using both Bray–Curtis and Aitchison distance.

Overlap Analysis with Single-Cell Reference Data. To further characterize the transcriptional landscape of our cohort and identify cell-type-specific alterations, we compared our bulk RNA-seq Differentially Expressed Genes (DEGs) with pseudo-bulk DEGs derived from a public single-cell RNA-sequencing (scRNA-seq) reference dataset (the same used for BayesPrism deconvolution). Specifically, for each relevant contrast (Primary vs. Normal, GCPM vs. Normal, and GCPM vs. Primary), we used a

hypergeometric test to determine if the overlap between our bulk DEGs and any cell-type-specific pseudo-bulk DEGs was statistically significant. For every significant overlap identified, we subsequently performed an over representation analysis to elucidate the active pathways and biological functions within those specific cell populations.

Molecular Functional Portraits. Molecular Functional Portraits (MFP), which classify the tumor microenvironment into four states—desert (D), fibrotic (F), immune-enriched (IE), and immune-enriched/fibrotic (IE/F) were inferred using the BostonGene MFP Python library²⁰ on the gene-level TPM counts matrix.

Ecotypes. Cell states and cellular communities in the bulk RNA-seq samples were identified using the EcotypeR package⁶⁰ on the gene TPM count matrix. Differences in ecotypes and cell state proportions between primary gastric cancer and metastatic tissues were evaluated using a Fisher's exact test. Differentially expressed genes associated with cell states showing significant differences in prevalence between primary and metastatic tissues were analyzed using gene set enrichment analysis (GSEA) with the clusterProfiler R package. Co-occurrence of cell states in paired primary and metastatic samples was visualized using Sankey diagrams.

Gene Fusion Analysis. Gene fusions were identified using the Nextflow nf-core rnafusion⁶¹ pipeline with default settings.

Statistics and Reproducibility

Statistical analyses were performed using R (v4.2.0; R Foundation for Statistical Computing, Vienna, Austria). Standard comparisons between independent groups utilized Student's t-test for normally distributed data or the Wilcoxon rank-sum test otherwise. For analyzing changes within matched samples, such as the difference in non-synonymous mutation counts between paired primary tumors and peritoneal metastases, a paired t-test was used. The Kolmogorov-Smirnov test was employed to compare the distributions of structural variant sizes between groups. To account for the relative nature of compositional data, like cell type proportions from deconvolution, Dirichlet regression was applied. Changes in paired categorical assignments, such as TME conservation status between primary and metastatic samples, were assessed using the mid-P McNemar test, chosen for its suitability with paired data and potentially small sample sizes. Similarity between paired samples based on the presence/absence of shared features was quantified using the Jaccard index. Dissimilarities in sample compositions were assessed using Bray-Curtis distance (for abundance-based comparisons) and Aitchison distance (specifically designed for relative compositional data). Functional analyses, including over-representation analysis (ORA) and gene set enrichment analysis (GSEA), were conducted with the R package clusterProfiler. The hypergeometric

test was specifically used to quantify the statistical significance of overlap between different sets of differentially expressed genes (DEGs). Visualization of Gene Ontology (GO) term similarities was performed using the R packages *igraph* and *ggraph*. Throughout the analyses, correction for multiple testing was applied using the Benjamini-Hochberg method to control the false discovery rate, with P-values < 0.05 considered statistically significant.

Data availability

The whole genome sequencing and whole transcriptome data generated in this study have been deposited in SRA under the BioProject PRJNA1428870 and GEO under the accession number GSE314812. All other data supporting the findings of this study, including source data for Figures and Data, are provided within the article and its Supplementary Information files.

Code availability

All custom scripts and code used for data processing and analysis are available on GitHub (https://github.com/ClementineCharton/Gastric_cancer_GCPM_multiomics) and archived on Zenodo (DOI: 10.5281/zenodo.17861491).

References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).
2. Koemans, W. J. *et al.* Synchronous peritoneal metastases of gastric cancer origin: incidence, treatment and survival of a nationwide Dutch cohort. *Gastric Cancer* **24**, 800–809 (2021).
3. Sun, J. *et al.* Clinical significance of palliative gastrectomy on the survival of patients with incurable advanced gastric cancer: a systematic review and meta-analysis. *BMC Cancer* **13**, 577 (2013).
4. Gwee, Y. X. *et al.* Integration of Genomic Biology Into Therapeutic Strategies of Gastric Cancer Peritoneal Metastasis. *Journal of Clinical Oncology* **71**, (2022).
5. Allen, C. J. *et al.* Staging laparoscopy and peritoneal cytology in patients with early stage gastric adenocarcinoma. *World J Surg Oncol* **18**, 39 (2020).
6. Zhao, J. J. *et al.* Spatially Resolved Niche and Tumor Microenvironmental Alterations in Gastric Cancer Peritoneal Metastases. *Gastroenterology* **167**, 1384-1398.e4 (2024).
7. Liu, Y. & Cao, X. Characteristics and Significance of the Pre-metastatic Niche. *Cancer Cell* **30**, 668–681 (2016).
8. Wei, J., Wu, N. D. & Liu, B. R. Regional but fatal: Intraperitoneal metastasis in gastric cancer. *World J Gastroenterol* **22**, 7478 (2016).
9. Janjigian, Y. Y. *et al.* 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. *The Lancet* **398**, 27–40 (2021).
10. Kang, Y. K. *et al.* Nivolumab plus chemotherapy versus placebo plus chemotherapy in patients with HER2-negative, untreated, unresectable advanced or recurrent gastric or gastro-oesophageal junction cancer (ATTRACTION-4): a randomised, multicentre, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol* **23**, 234–247 (2022).
11. Díaz del Arco, C., Estrada Muñoz, L., Ortega Medina, L. & Fernández Aceñero, M. J. [Update on gastric cancer. New molecular classifications]. *Rev Esp Patol* **54**, 102–113 (2021).
12. Andini, K. D. *et al.* PMS2-associated Lynch syndrome: Past, present and future. *Front Oncol* **13**, 1127329 (2023).
13. Lee, J. E., Kim, K. T., Shin, S. J., Cheong, J. H. & Choi, Y. Y. Genomic and evolutionary characteristics of metastatic gastric cancer by routes. *British Journal of Cancer* **2023** 129:4 **129**, 672–682 (2023).

14. Gillis, S. & Roth, A. PyClone-VI: scalable inference of clonal population structures using whole genome data. *BMC Bioinformatics* **21**, 1–16 (2020).
15. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, 1–14 (2011).
16. Saito-Adachi, M. *et al.* Oncogenic structural aberration landscape in gastric cancer genomes. *Nat Commun* **14**, (2023).
17. Racle, J. & Gfeller, D. EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. *Methods Mol Biol* **2120**, 233–248 (2020).
18. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology* 2019 37:7 **37**, 773–782 (2019).
19. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nature Cancer* 2022 3:4 **3**, 505–517 (2022).
20. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845–865.e7 (2021).
21. Lee, H. H., Kim, S. Y., Jung, E. S., Yoo, J. & Kim, T. M. Mutation heterogeneity between primary gastric cancers and their matched lymph node metastases. *Gastric Cancer* **22**, 323–334 (2019).
22. Martínez-Jiménez, F. *et al.* Pan-cancer whole-genome comparison of primary and metastatic solid tumours. *Nature* **618**, 333–341 (2023).
23. Li, S. *et al.* Genomic landscape of paired primary and peritoneal metastatic lesions in gastric cancer highlights evolutionary dynamics and mutational drivers. *J Adv Res* <https://doi.org/10.1016/j.jare.2025.05.043> (2025) doi:10.1016/j.jare.2025.05.043.
24. Tanaka, Y. *et al.* Multi-omic profiling of peritoneal metastases in gastric cancer identifies molecular subtypes and therapeutic vulnerabilities. *Nature Cancer* 2021 2:9 **2**, 962–977 (2021).
25. Park, S. *et al.* Clinical Relevance and Molecular Phenotypes in Gastric Cancer, of TP53 Mutations and Gene Expressions, in Combination With Other Gene Mutations. *Scientific Reports* 2016 6:1 **6**, 1–10 (2016).
26. Ge, S. *et al.* Genomic alterations in advanced gastric cancer endoscopic biopsy samples using targeted next-generation sequencing. *Am J Cancer Res* **7**, 1540 (2017).
27. Wang, J. *et al.* Large-scale analysis of CDH1 mutations defines a distinctive molecular subset with treatment implications in gastric cancer. *npj Precision Oncology* 2024 8:1 **8**, 1–10 (2024).

28. Wang, R. *et al.* Multiplex profiling of peritoneal metastases from gastric adenocarcinoma identified novel targets and molecular subtypes that predict treatment response. *Gut* **69**, 18–31 (2020).
29. Tanaka, Y. *et al.* Multi-omic profiling of peritoneal metastases in gastric cancer identifies molecular subtypes and therapeutic vulnerabilities. *Nature Cancer* **2**, 962–977 (2021).
30. Zhao, H. *et al.* Overview on the Role of E-Cadherin in Gastric Cancer: Dysregulation and Clinical Implications. *Front Mol Biosci* **8**, 689139 (2021).
31. Su, S. *et al.* Blocking the recruitment of naive CD4⁺ T cells reverses immunosuppression in breast cancer. *Cell Research* **27**, 461–482 (2017).
32. Sugarbaker, P. *Peritoneal Carcinomatosis: Principles of Management*. (1996).
33. Wang H *et al.* Intraperitoneal nivolumab for malignant ascites in patients with advanced gastrointestinal or pancreaticobiliary tract cancer. in *ESMO Congress 2023* (2023).
34. Bergmann, E. A., Chen, B. J., Arora, K., Vacic, V. & Zody, M. C. Conpair: concordance and contamination estimator for matched tumor-normal pairs. *Bioinformatics* **32**, 3196–3198 (2016).
35. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
36. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).
37. Van der Auwera, G., O'Connor, B. & Safari. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. *O'Reilly Media* 300 (2020).
38. DI Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316–319 (2017).
39. Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276–278 (2020).
40. Garcia, M. *et al.* Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Res* **9**, 63 (2020).
41. Kim, S. *et al.* Strelka2: fast and accurate calling of germline and somatic variants. *Nature Methods* **15**, 591–594 (2018).
42. Kandoth C. mskcc/vcf2maf: vcf2maf v1.6.18. <https://github.com/mskcc/vcf2maf> (2020) doi:doi:10.5281/zenodo.593251.
43. Favero, F. *et al.* Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Annals of Oncology* **26**, 64 (2014).
44. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).

45. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333 (2012).
46. Malikic, S., McPherson, A. W., Donmez, N. & Sahinalp, C. S. Clonality inference in multiple tumor samples using phylogeny. *Bioinformatics* **31**, 1349–1356 (2015).
47. Smith, M. timescape: Patient Clonal Timescapes. <https://www.bioconductor.org/packages/release/bioc/html/timescape.html> (2025).
48. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 10–12 (2011).
49. Patel, H. *et al.* nf-core/rnaseq: nf-core/rnaseq v3.18.0 - Lithium Lynx. <https://doi.org/10.5281/ZENODO.14537300> doi:10.5281/ZENODO.14537300.
50. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
51. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 1–16 (2011).
52. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications* **2013 4:1** **4**, 1–11 (2013).
53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 1–21 (2014).
54. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
55. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417–425 (2015).
56. Kumar, V. *et al.* Single-Cell Atlas of Lineage States, Tumor Microenvironment, and Subtype-Specific Expression Programs in Gastric Cancer. *Cancer Discov* **12**, 670–691 (2022).
57. Jiang, H. *et al.* Revealing the transcriptional heterogeneity of organ-specific metastasis in human gastric cancer using single-cell RNA Sequencing. *Clin Transl Med* **12**, (2022).
58. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
59. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods* **2019 16:12** **16**, 1289–1296 (2019).
60. Luca, B. A. *et al.* Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell* **184**, 5482–5496.e28 (2021).
61. Proks, M. *et al.* nf-core/rnafusion: nf-core/rnafusion:1.2.0. <https://doi.org/10.5281/ZENODO.3946477> doi:10.5281/ZENODO.3946477.

Acknowledgements

This study was supported by the National Research Foundation of Republic of Korea (NRF-2021R1A2C1013554), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: RS-2024-00411999 and RS-2025-02303065), and the World Class 300 Project (S2638360) of the Ministry of Trade, Industry and Energy, and Ministry of SMEs and Startups of the Republic of Korea and the KHIDI, funded by the Ministry of Health & Welfare, Republic of Korea (grant number: KH129481).

Author contributions

H-H K, J K and J L conceived the study, directed and supervised the research. S H K, Y S P, S-H A, Y-S S and H-H K contributed to the enrolment of GC patients and collection of patient information. C C, S S K, M K and J H L performed data processing, bioinformatics analyses and visualization. S H K and J K interpreted the data and provided advice on data analysis. N J K and K U P helped design the study. C C and S H K wrote the manuscript. S H K, J L and J K reviewed and edited the manuscript. N J K and H-H K provide funding. All authors provided feedback.

Competing interests

The authors declare no competing interests.

Figure legends

Figure 1. Genomic architecture of metastatic gastric cancer. **a** Experimental design. Primary gastric cancer tumor and peritoneal seeding, lymph node, and omentum metastases samples were collected for whole-genome sequencing (WGS) and/or whole-transcriptome sequencing (WTS) analyses. Created in BioRender. Charton, C. (2026) <https://BioRender.com/j5qzvuvq> **b** Summary of the number and type of non-synonymous mutations, mutational landscape and mutational signature in each sample and patient. **c** Over representation analysis of GO terms in the GCPM and primary tumor samples somatic variants. **d** Number of common, primary tumor-specific, and metastasis-specific mutations within patients. **e** Mutational signature in each sample and patient. **f** Mutational signatures in common, primary tumor-specific, and metastasis-specific mutations.

Figure 2. Phylogeny and clonal evolution of gastric cancer primary and metastatic tumors. a, b, and c Clonal evolution for patients GC05, GC01, and GC07, respectively. Left panel: Map of cellular prevalence of clones in each tumor sample. The prevalence of clones is represented proportionally by the colored areas; the tumor microenvironment associated with each sample is indicated in the text boxes. Right panel: Clonal phylogeny with gastric cancer driver genes associated with each branch. Trunks and branches are color-coded. PS1: Peritoneal seeding metastasis sample before treatment. T2 and PS2: Primary gastric cancer tumor and peritoneal seeding metastasis after treatment (1-year recurrence). Created in BioRender. Kim, J. (2026) <https://BioRender.com/zhtkg3t>

Figure 3. Overview of copy number variations and structural variants in metastatic gastric cancer. a Heatmap of segmented copy number profiles across the cohort. **b** Number and types of copy number variations (CNVs) identified across samples. **c** Number and types of structural variants (SVs) identified across samples. **d** Relationship between SV allele frequency and SV size. For each SV type scatterplot, the top and left panels display the SV size and SVAF distribution densities, respectively. **e** Comparison of structural variant allele frequency (SVAF) for deletions, tandem duplications, and inversions between primary tumor and gastric metastases. **f** Intratumoral heterogeneity in the number of structural variants within patients. **g** Intratumoral heterogeneity of structural variant types within patients.

Figure 4. Gastric cancer primary tumor and peritoneal seeding metastases gene expression characterization. a Principal component analysis' individuals map, based on variance stabilizing transformation (VST) transformed gene expression. K-means clusters are indicated by the ellipses and points shapes. **b** UpSet plot of the intersection between the 6 contrasts DEGs. **c** GO terms similarity network colored by DEGs contrast specificity. **d** Functional annotation of the most shared Hallmark, C2 terms and KEGG pathways based on the DEGs six contrasts. Dot size represents the count of differentially expressed genes associated with each term. NES: Normalized Enrichment Score.

Figure 5. Bulk RNAseq deconvolution and tumor microenvironment (TME). a Deconvolution results showing cell type proportions significantly differing between gastric cancer (GC) primary tumor and peritoneal seeding samples. First panel: CIBERSORTx results. Second panel: EPIC results. Third panel: BayesPrism. BayesPrism deconvolution was performed using single-cell data from primary gastric cancer and peritoneal seeding metastases as reference. **b** Sankey diagram illustrating the concordance of tumor microenvironment (TME) between paired primary and metastatic GC tumor samples. **c** Pie charts showing the composition of tumor microenvironments in gastric cancer primary tumors and peritoneal metastases.

Figure 6. Ecotypes of metastatic gastric cancer. a Heatmap of ecotype and cell state profiles in the cohort. **b** and **c** Endothelial and epithelial cell state analysis. Top Left Panel: Sankey diagram depicting

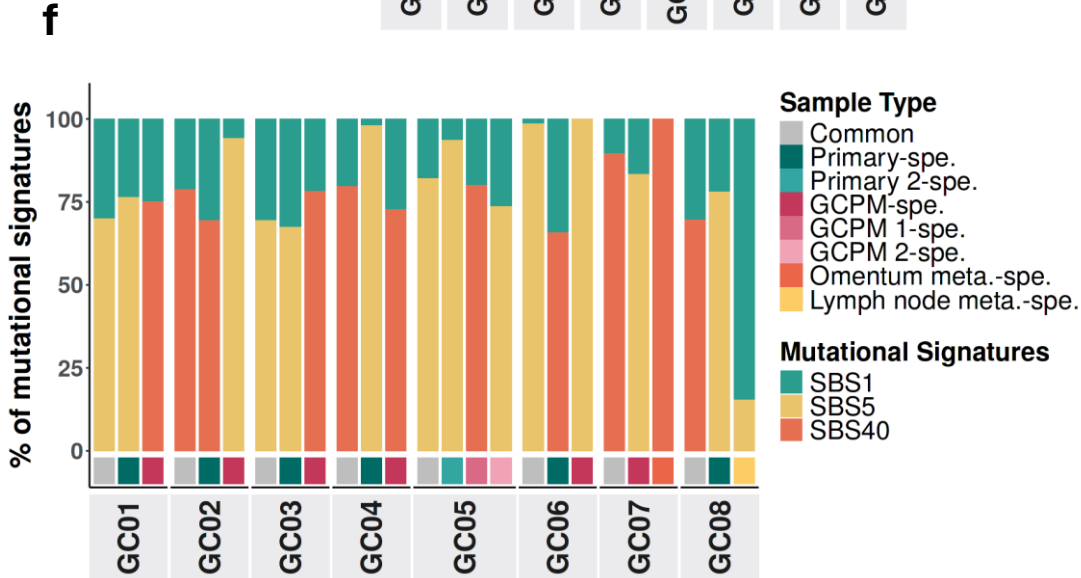
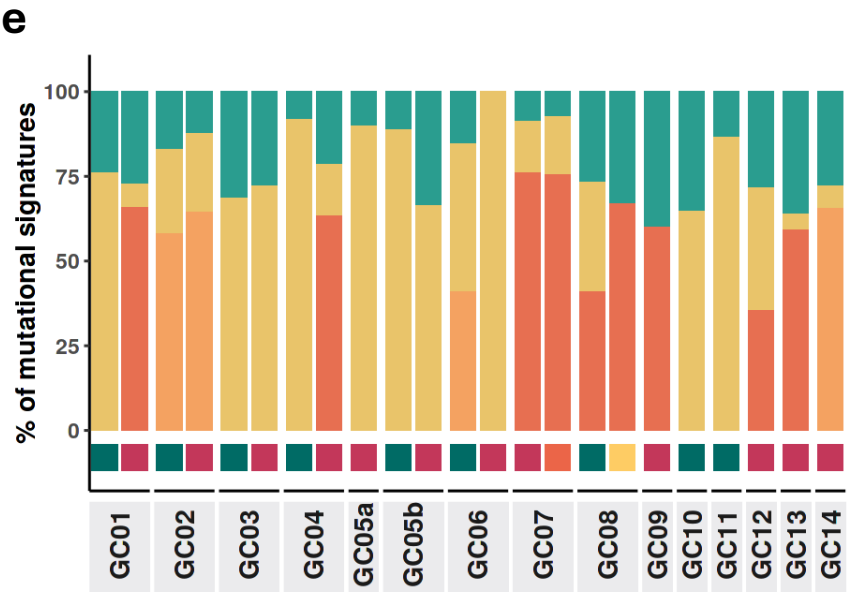
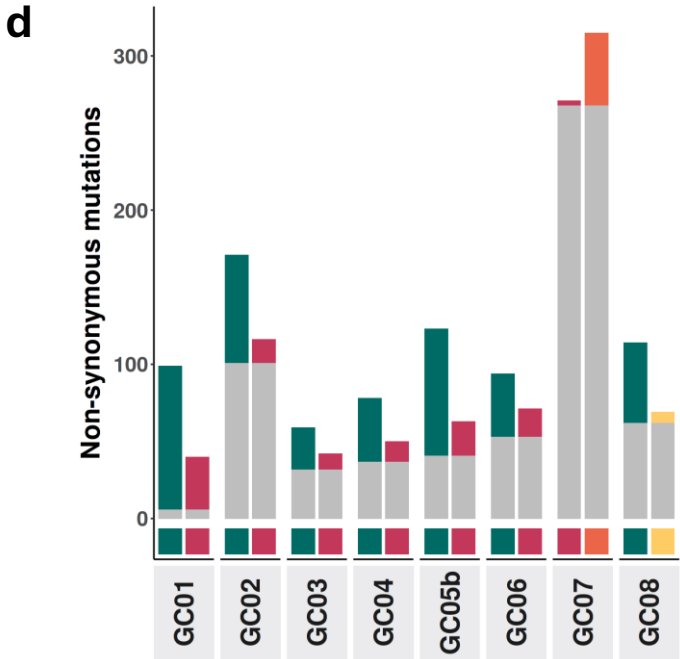
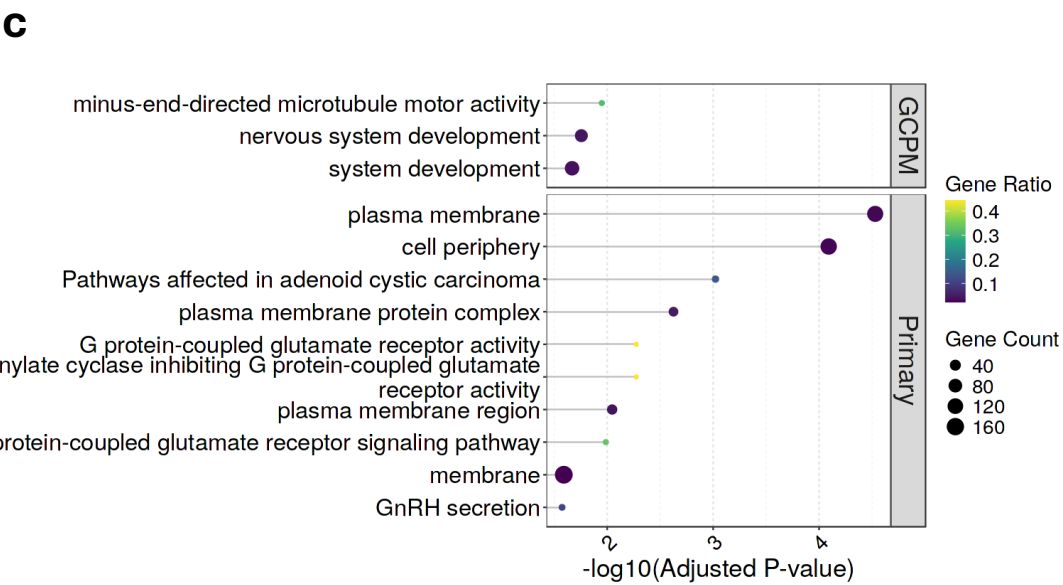
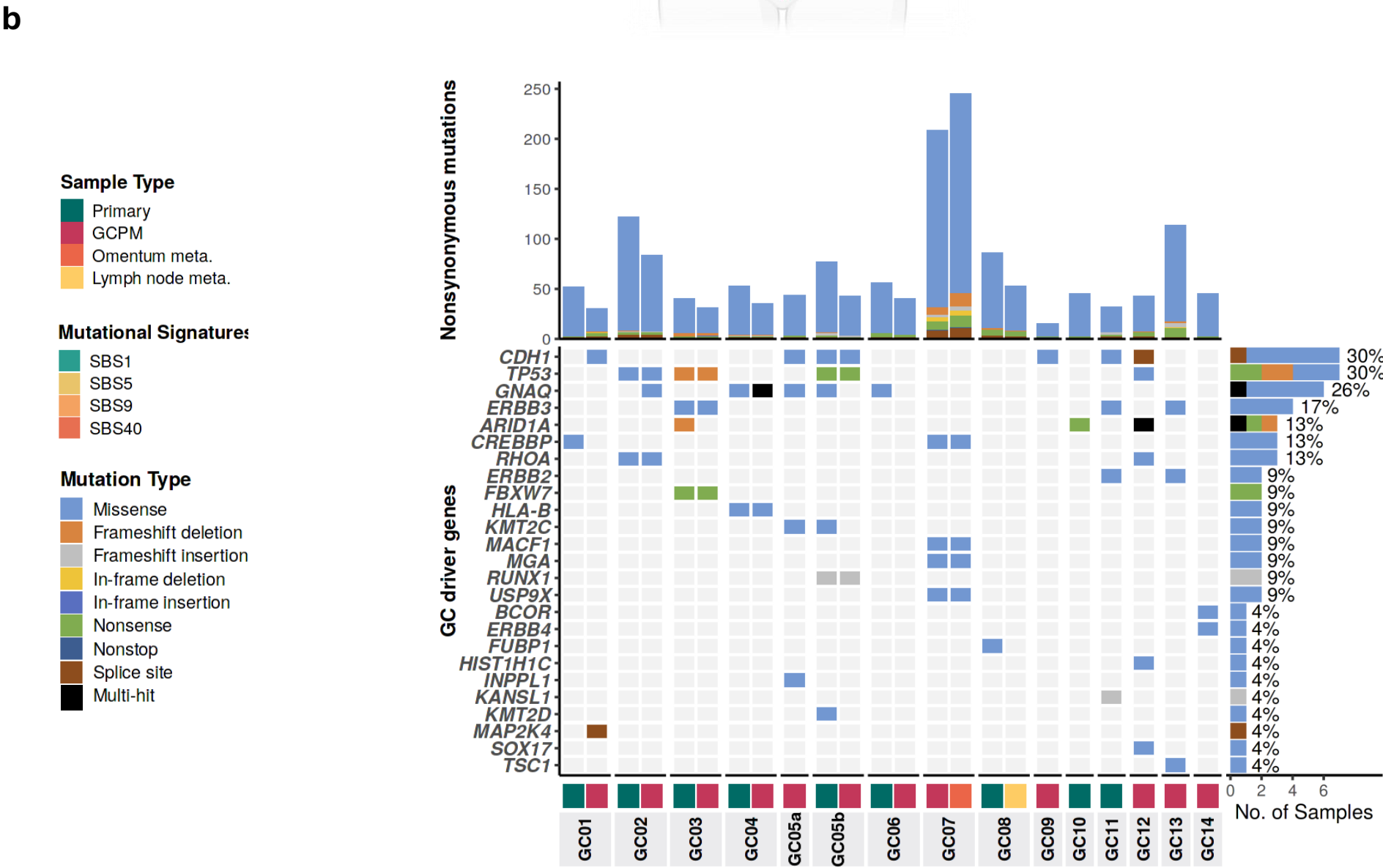
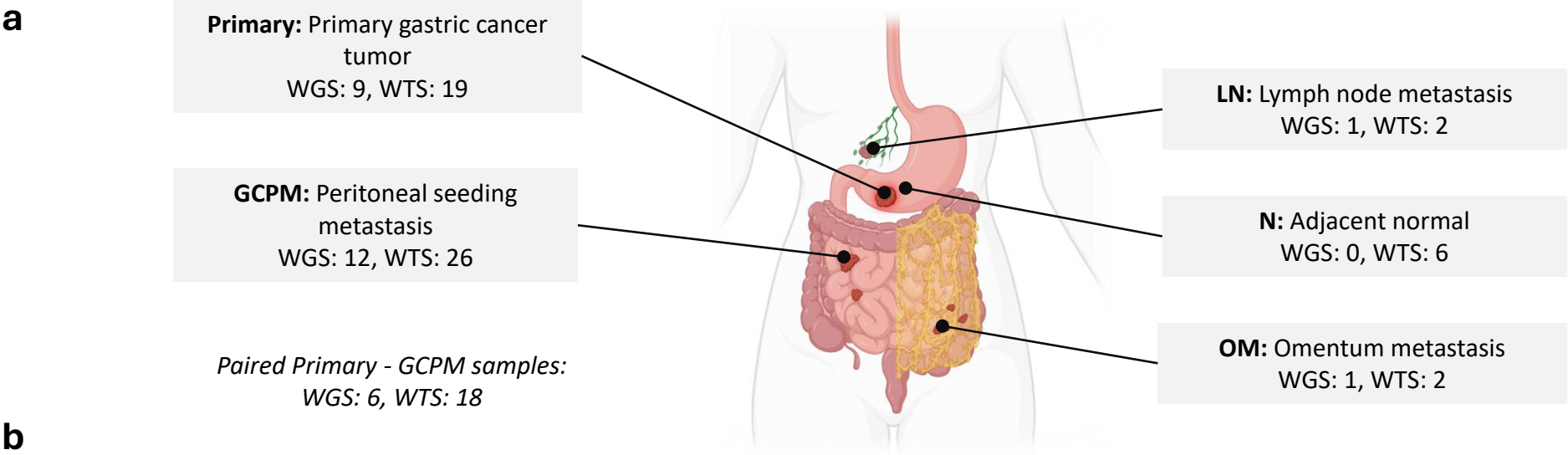
the within patient relationships in endothelial and epithelial cell states between different types of samples. Bottom Left Panel: c Pie charts showing the composition of cell states in gastric cancer primary tumors and peritoneal metastases. Right Panel: Gene Set Enrichment Analysis (GSEA) of top enriched DEG genes highlighting differences in cell states between GC primary tumor and peritoneal seeding metastases. Dotplot: The dot size represents the count of differentially expressed genes associated with each term, while the gene ratio indicates the proportion of DEG genes relative to the total number of genes in that term.

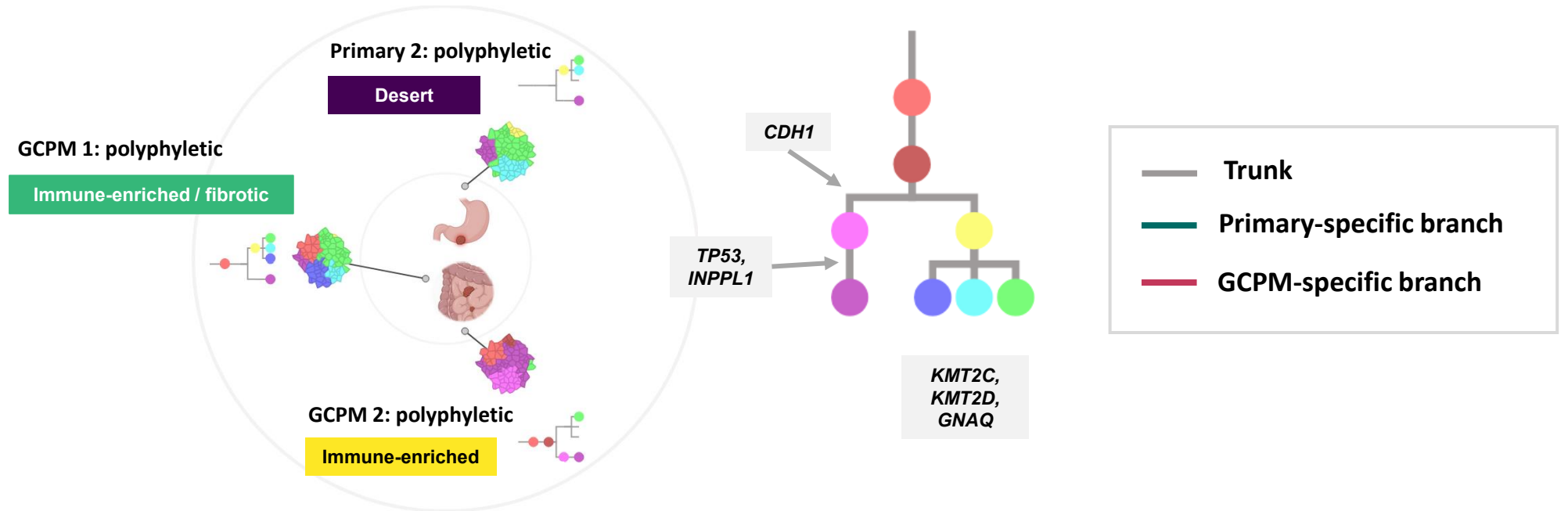
Editorial Summary:

Genomic and transcriptomic profiling reveals divergent clonal evolution and TME remodeling in gastric cancer peritoneal metastasis, highlighting substantial heterogeneity and the need for metastasis-specific therapeutic strategies.

Peer Review Information:

Communications Biology thanks Ruth Zarate and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Ana Rita Grosso and Johannes Stortz. A peer review file is available.



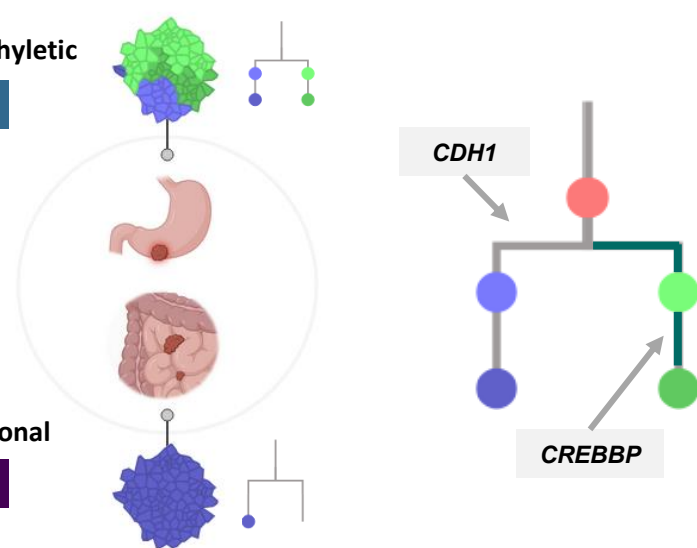
a**GC05 patient****b****GC01 patient**

Primary: polyphyletic

Fibrotic

GCPM: monoclonal

Desert

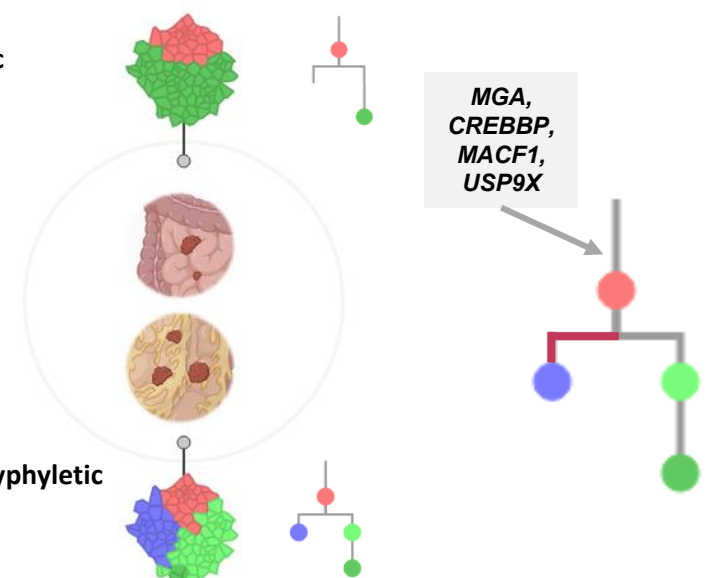
**c****GC07 patient**

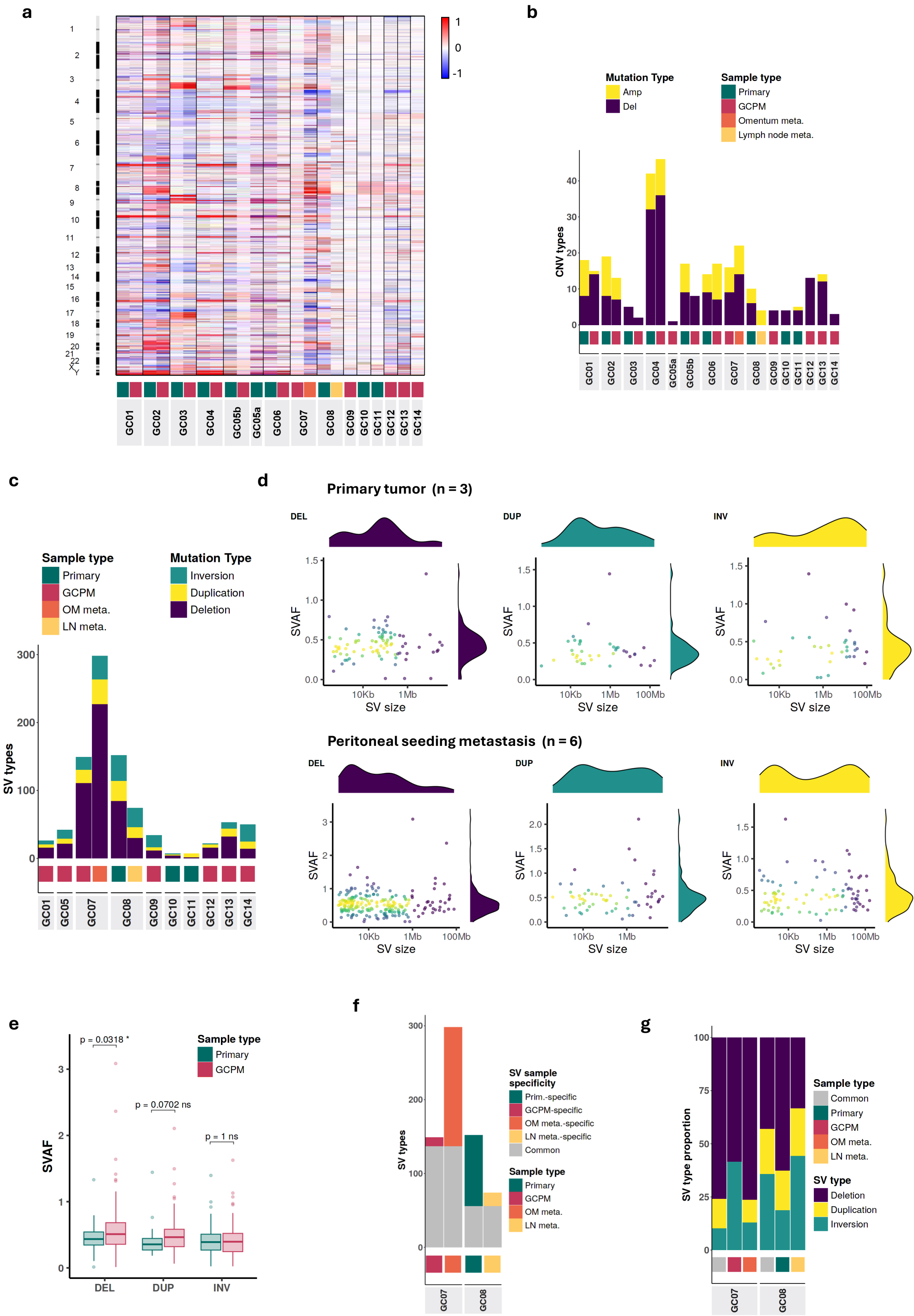
GCPM: Monophyletic

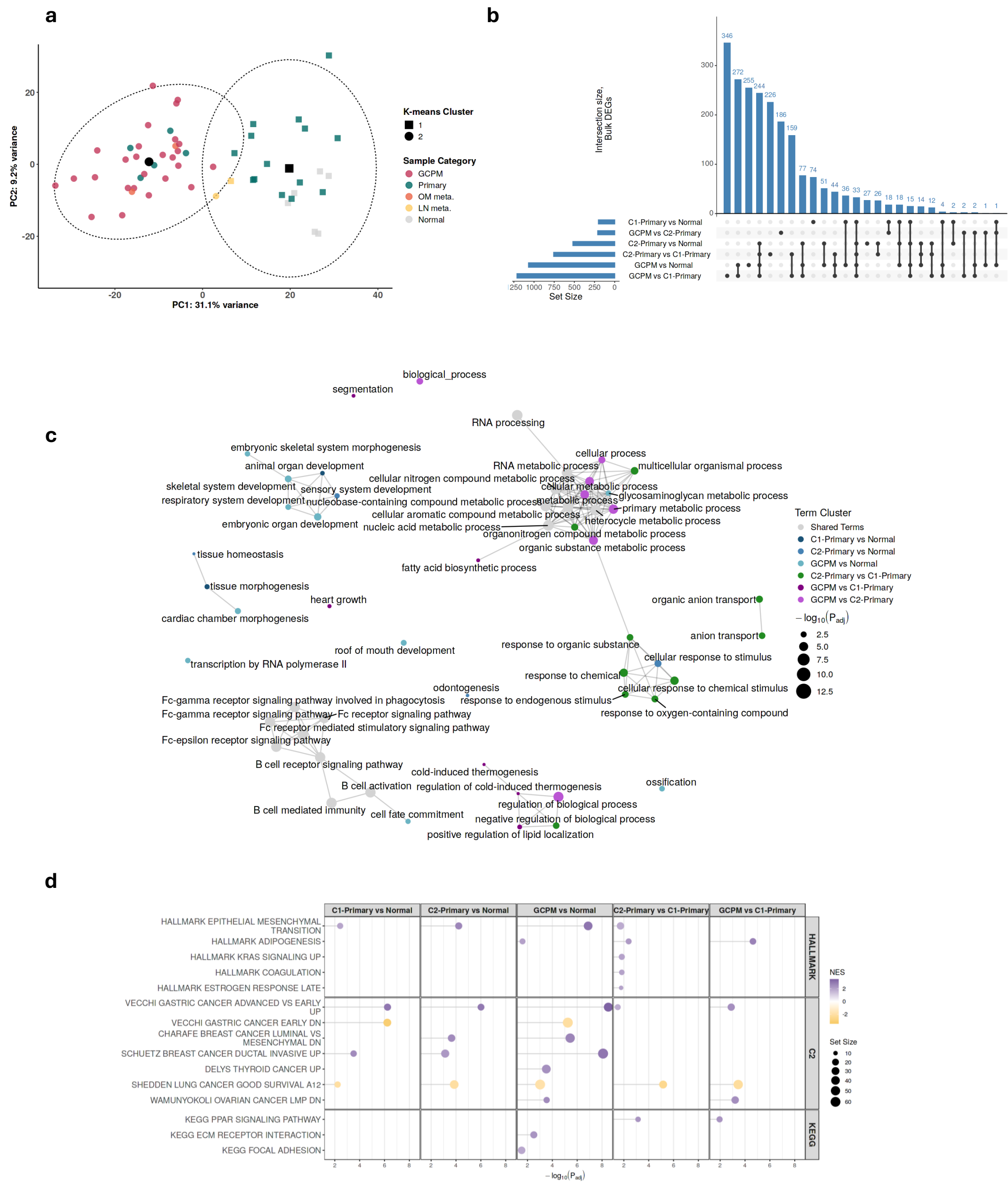
Desert

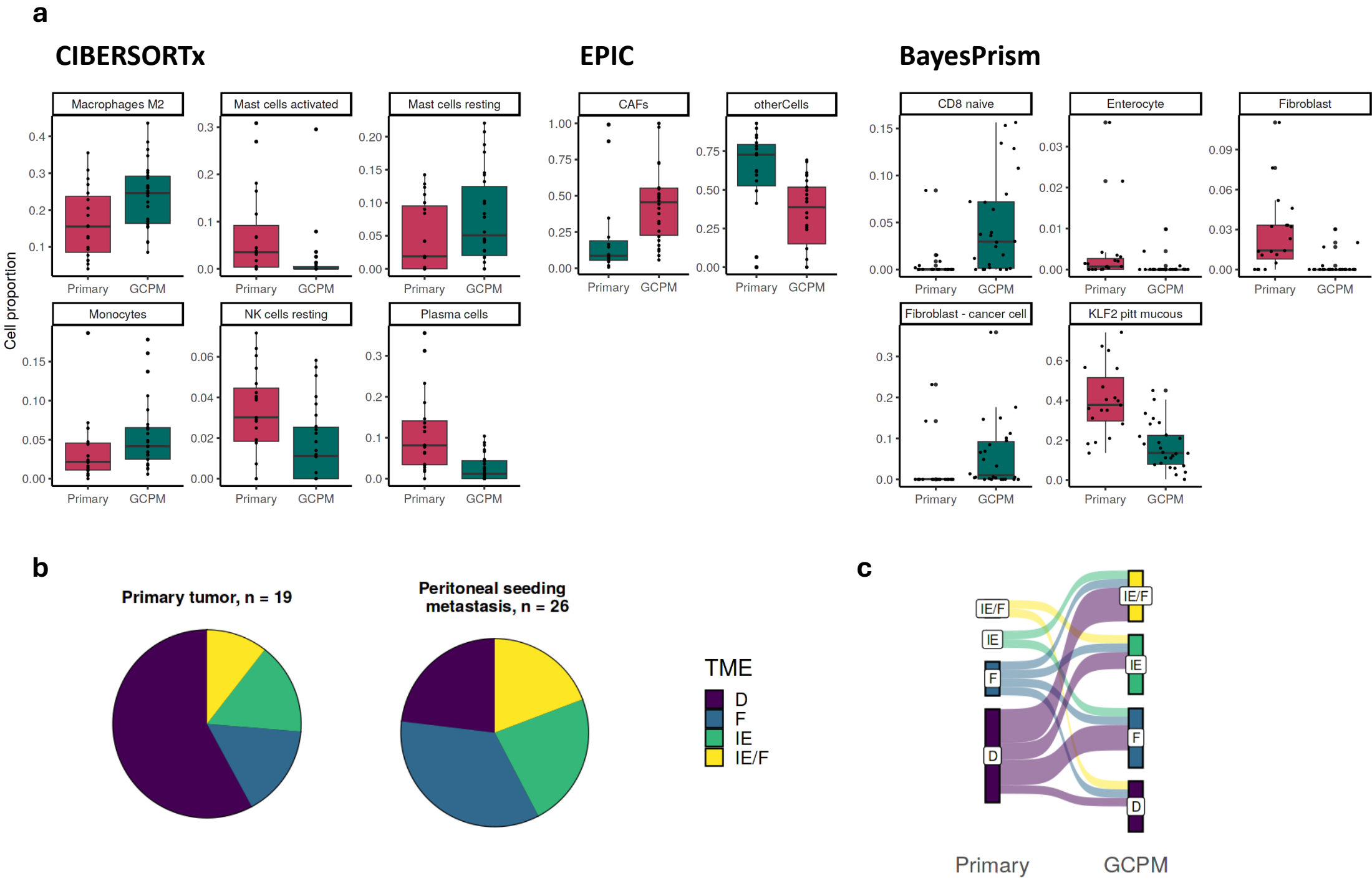
Omentum meta: Polyphyletic

Desert

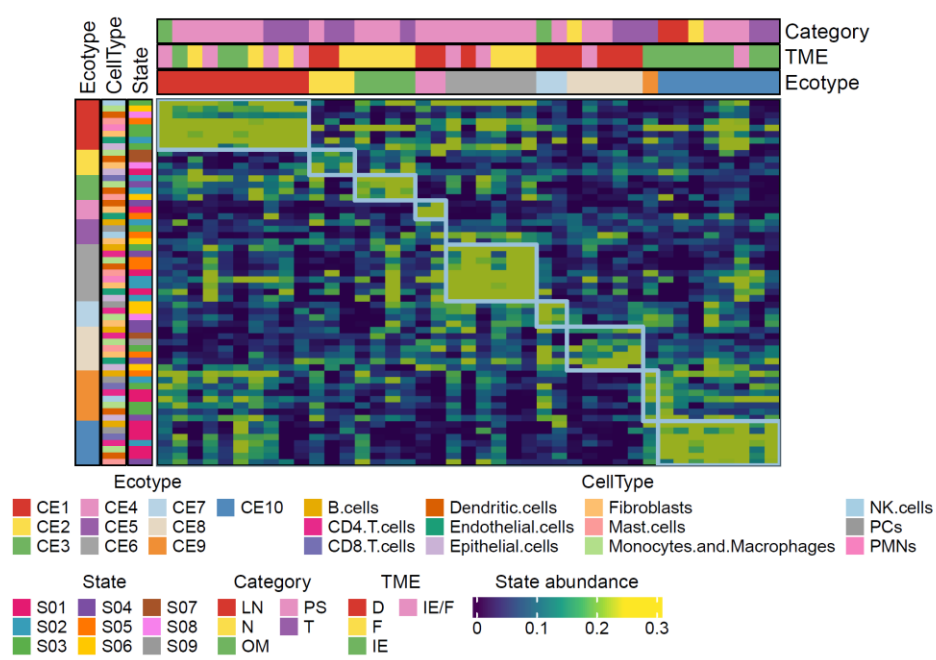




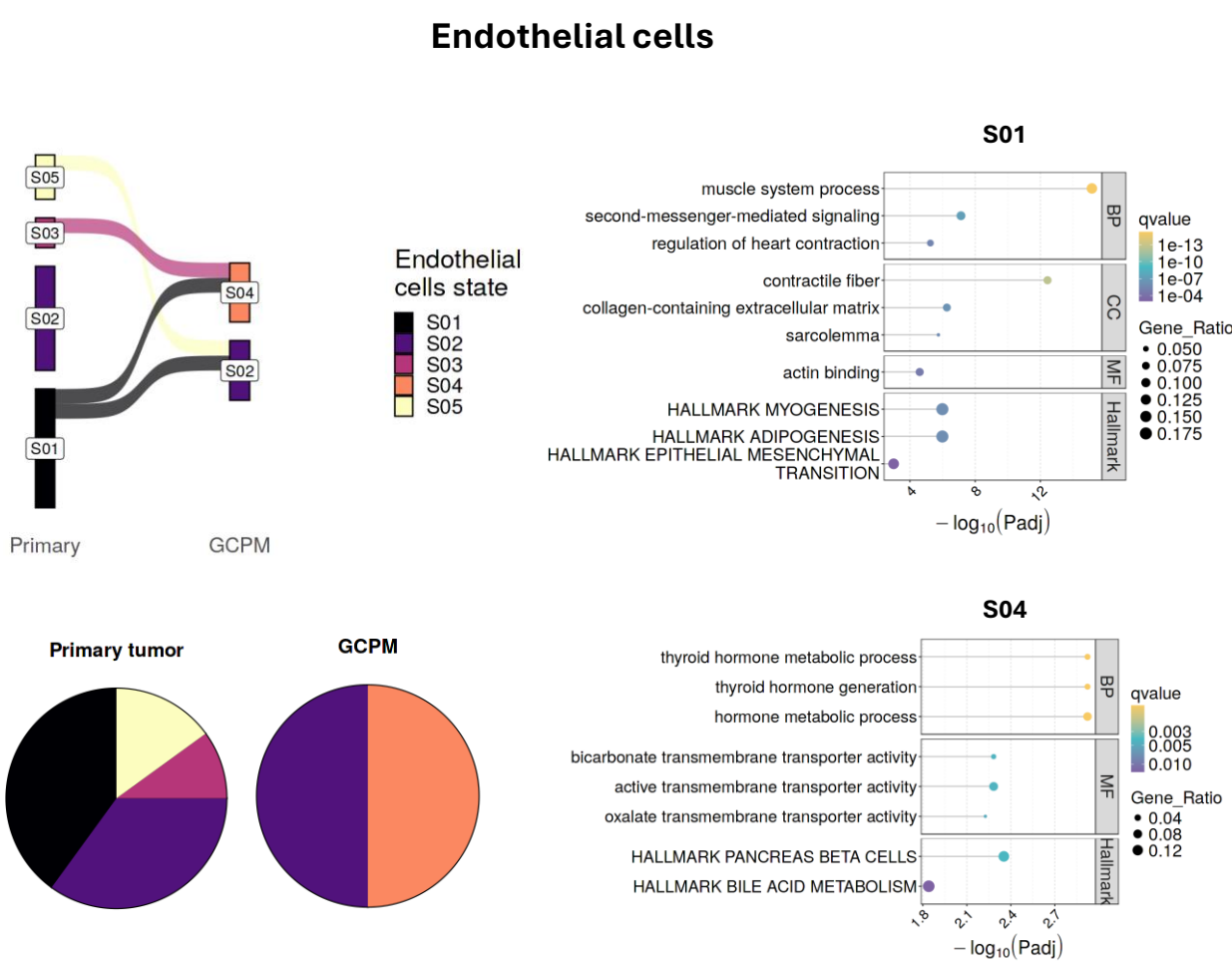




a



b



c

Epithelial cells

