Seed and Soil: Consensus Molecular Subgroups (CMS) and Tumor Microenvironment Features Between Primary Lesions and Metastases of Different Organ Sites in Colorectal Cancer

Qingqing Luo1,*, Yibo Quan2,*, Wei Liu1, Zixin Wu1, Wenjing Qiu1, Wenlong Liang1, Ping Yang3, Qing Huang1, Guanwei Li1, Jianchang Wei1, Qiang Wang1, Fei Shen2, Wanglin Li3,1, Feng He3, Jie Cao1

1Department of Gastrointestinal Surgery, the Second Affiliated Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People’s Republic of China; 2Department of Thyroid Surgery, the Second Affiliated Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People’s Republic of China; 3Department of Nephrology, the Second Affiliated Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People’s Republic of China

*These authors contributed equally to this work

Correspondence: Jie Cao, Department of Gastrointestinal Surgery, the Second Affiliated Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People’s Republic of China, Email eycaojie@scut.edu.cn; Feng He, Department of Nephrology, the Second Affiliated Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People’s Republic of China, Email eyhefeng@scut.edu.cn

Purpose: Consensus molecular subtypes (CMS) are mainly used for biological interpretability and clinical stratification of colorectal cancer (CRC) in primary tumors (PT) but few in metastases. The heterogeneity of CMS distribution in metastases and the concordance of CMS between PT and metastases still lack sufficient study. We used CMS to classify CRC metastases and combine it with histopathological analysis to explore differences between PT and distant metastases.

Patients and Methods: We obtained gene expression profiles for 942 PT samples from TCGA database (n=376) and GEO database (n=566), as well as 442 metastasis samples from GEO database. Among these, 765 PT samples and 442 metastasis samples were confidently identified with CMS using the “CMS classifier” and enrolled for analysis. Clinicopathological manifestation and CMS classification of CRC metastases were assessed with data from GEO, TCGA, and cBioPortal. Overall, 105 PT-metastasis pairs were extracted from 10 GEO datasets to assess CMS concordance. Tumor microenvironment (TME) features between PT and metastases were analyzed by immune-stromal infiltration with ESTIMATE and xCell algorithms. Finally, TME features were validated with multiplex immunohistochemistry in 27 PT-metastasis pairs we retrospectively collected.

Results: Up to 64% of CRC metastases exhibited concordant CMS groups with matched PT, and the TME of metastases was similar to that of PT. For most common distant metastases, liver metastases were predominantly CMS2 and lung and peritoneal metastases were mainly CMS4, highlighting “seed” of tumor cells of different CMS groups had a preference for metastasis to “soil” of specific organs. Compared with PT, cancer-associated fibroblasts (CAF) reduced in liver metastases, CD4+T cells and M2-like macrophages increased in lung metastases, and M2-like macrophages and CAF increased in peritoneal metastases.

Conclusion: Our findings underscore the importance of CMS-guided specific organ monitoring and treatment post-primary tumor surgery for patients. Differences in immune-stromal infiltration among different metastases provide targeted therapeutic opportunities for metastatic CRC.

Keywords: colorectal cancer, primary tumors, metastases, consensus molecular subtypes, tumor microenvironment

Introduction
Colorectal cancer (CRC) is one of the most fatal cancers globally and in China, which has been a major public health challenge for decades.1–4 The high heterogeneity of CRC on clinical and biological features culminates in significant differences in disease progression and treatment response,5–7 thereby leading to the limited benefit of available treatment options to a considerable number of CRC patients in the clinic.
For CRC patients, death and poor prognosis are mainly attributed to distant metastasis.\textsuperscript{8–10} As reported, approximately 33\% of CRC patients develop metastases at presentation or follow-up. Accordingly, researchers have been engaged in analyzing the similarities and differences between primary lesions and metastases in CRC for further comprehending the characteristics of metastatic tumors and identifying therapeutic targets. For instance, Bhullar et al demonstrated that multiple biomarkers were highly consistent between primary and metastatic CRC.\textsuperscript{11} Wang et al observed diverse origins of metastatic tumors between lymph nodes and liver.\textsuperscript{12} Likewise, Eynde et al found that immune infiltration and mutation were heterogeneous between primary lesions and synchronous and metachronous metastases.\textsuperscript{13} Conversely, Vakiani et al revealed a highly consistent mutation state of paired primary and metastatic CRC.\textsuperscript{14} Thus, it is warranted to more comprehensively and systematically investigate the differences between primary and metastatic CRCs.

In 2015, Guinney et al proposed the CRC Subtyping Consortium, which identifies four consensus molecular subtypes (CMS; CMS1-4) for CRC.\textsuperscript{15} Specifically, CMS1, an immunogenic subtype, is characterized by the enrichment of microsatellite instability-high (MSI-H) and B-type Raf (BRAF) mutations. CMS2 presents with epithelial characteristics, marked WNT and MYC pathway activation, and high chromosomal instability (CIN). CMS3 also exhibits epithelial features and obvious metabolic dysregulation with lower CIN, which is enriched for Kirsten rat sarcoma (KRAS) mutations. CMS4, a mesenchymal subtype, shows prominent upregulation of epithelial-to-mesenchymal transition (EMT)-related genes and transforming growth factor β (TGF-β), stromal invasion, angiogenesis, and inflammatory and immunosuppressive phenotypes.\textsuperscript{15,16} Unlike American Joint Committee on Cancer (AJCC) TNM staging, the CMS taxonomy was principally founded based on tumor biological differences rather than clinical outcomes, which can capture the inherent biomolecular heterogeneity of CRC.

CMS classification is an unsupervised system that can robustly stratify CRC by integrating the characteristics of genomics, epigenetics, transcriptome pathways, stromal and immune microenvironments, mutated genes, and clinical manifestations, with independent predictive value.\textsuperscript{15,17,18} Importantly, CMS has been confirmed to be significantly associated with the invasion and metastasis of CRC cells,\textsuperscript{19,20} the efficacy and resistance of chemo/radiotherapy,\textsuperscript{21} tumor microenvironment (TME),\textsuperscript{22} the infiltration of CD8\(^+\) cytotoxic lymphocytes and cancer-associated fibroblasts (CAFs), and patient prognosis.\textsuperscript{23} Although CMS has robust functions, it was mainly utilized to stratify primary tumors (PT), with only a few studies exploring the attributes of CMS in the metastatic setting and finding the CMS heterogeneity in CRC metastases. Specifically, most metastases were classified as CMS2, followed by CMS4, and a highly similar proportion of CMS2 and CMS4 in metastases and PT was observed.\textsuperscript{24} A strong depletion of CMS3 in metastases was found.\textsuperscript{24} The classification of CMS1 in metastases has been reported with conflicting results. Eide et al and Kamal et al reported a significant decrease of CMS1 in metastases,\textsuperscript{24} while Piskol et al observed a slight enrichment of CMS1 in metastases compared with PT.\textsuperscript{25}

However, studies on exploring the subtype distribution of CMS in CRC metastases from different organ sites and assessing the concordance of CMS between PT and matched metastases are still insufficient. In our study, we use CMS combined with histopathological analysis to investigate the features of CRC metastases and compare the CMS classification rules of PT and paired metastases, therefore further ascertaining the process of metastasis and confirming and supplementing the “seed and soil” theory.\textsuperscript{27} Our discoveries provide some new insights to guide subtype-targeted therapy for metastatic CRC.

\section*{Materials and Methods}

\subsection*{Collection and Processing of Gene Expression Data}

Gene Expression Omnibus (GEO; \url{https://www.ncbi.nlm.nih.gov/geo/}) and The Cancer Genome Atlas (TCGA; \url{https://portal.gdc.cancer.gov/}) were employed for downloading the gene expression profiles and clinical annotations of totally 238 normal colon tissues, 2241 primary CRC tissues, and 592 CRC metastases tissues (Supplementary Table 1). In these data, metastatic sites of CRC included the liver, lungs, distant lymph nodes, and peritoneum. The related clinical information was collected, including progression-free survival (PFS), TNM stage, overall survival (OS), and survival status. In addition, GEO was utilized for downloading the transcriptome sequencing data of 21 normal organ tissues (10
liver tissues, 7 lung tissues, and 4 peritoneum tissues), followed by the analysis of their microenvironmental cell composition. The series accession and sample number of each database are summarized in Supplementary Table 1. Thereafter, 1134 cases of clinicopathological data were collected from Metastatic Colorectal Cancer (Cancer Cell 2018) of cBioportal (http://www.cbioportal.org/), including 975 cases with identified first distant organ metastasis, differentiation degree, survival status, and time (Supplementary Table 2). A flowchart was also provided for displaying multiple public databases used in our study (Supplementary Figure 1). The batch effect of multiple datasets was eliminated with the ComBat method, and raw data were normalized with the robust multi-array average algorithms (affy package) in the R-Studio before analysis.

Patients and Tissue Samples
Due to the timely intervention of neoadjuvant therapy or the unresectability of some advanced tumors, we retrospectively collected precious paraffin-embedded primary and paired metastatic samples of 27 CRC patients who receive no treatment before surgery from the Guangzhou First People’s Hospital (the Second Affiliated Hospital of South China University of Technology) between 2010 and 2019. These samples included 10 liver metastasis tissues, 9 lung metastasis tissues, 8 peritoneal metastasis tissues, and their paired primary tumor tissues. They were used for multiplex immunohistochemistry validation to explore the TME features between PT and metastases in different organ sites. Our study complied with the Declaration of Helsinki and was approved by the Ethics Committee of the Guangzhou First People’s Hospital (Approval no. K-2019-070-01).

CMS Classification
CRC samples of PT and paired metastases from the public databases were subjected to CMS classification with the single sample prediction (SSP) algorithm from the “CMS Classifier” (https://github.com/Sage-Bionetworks/crcsc). After samples that could not be defined with this algorithm were excluded, 765 primary CRC samples (457 from GSE39582 and 308 from TCGA) (Supplementary Tables 3 and 4) and 442 metastasis samples (all from GEO) (Supplementary Table 5) were classified for subsequent analysis (Supplementary Figure 1). Furthermore, 105 pairs of CRC primary and metastatic samples (Supplementary Tables 1 and 6) were classified (a total of 209 samples since one patient simultaneously presented with two paired metastases). The matching degree of CMS groups of these samples was evaluated with a Sankey diagram and visualized with the “ggalluvial” R package.

TME Evaluation
Tumor purity, immune scores, and stroma scores were calculated with Estimate of Stromal and Immune Cells in Malignant Tumor Tissues from Expression Data (ESTIMATE) algorithm. The xCell algorithm (https://xcell.ucsf.edu/) can quantify the infiltration of 64 kinds of immune and stromal cells, such as B cells, T cells, macrophages, hematopoietic stem cells (HSCs), mesenchymal stem cells (MSCs), and fibroblasts. Therefore, this algorithm was utilized to identify the enrichment of cell subtypes. Additionally, these two algorithms were adopted to evaluate and compare the cell composition of the TME of primary and metastatic tumor samples.

H&E Staining
Paraffin sections of 27 pairs of primary and metastatic tumor tissues were subjected to H&E staining to observe the region of tumors and the situation of necrosis or infiltration.

Multiplex Immunohistochemistry Staining (mIHC)
mIHC was carried out as instructed in the manuals of the PANO Multiplex IHC kit (Panovue, Beijing, China; 10144100100). Subsequent to formalin fixing and paraffin embedding, the sections underwent dewaxing and rehydration. After heat-induced epitope retrieval with citrate buffer (pH = 6.0) or Tris/ethylenediaminetetraacetic acid (pH = 9.0), each section was treated with 3% H2O2 for endogenous peroxidase blocking and with 10% goat serum for non-specific protein blocking. After that, the sections underwent incubation with primary antibodies, including CD206 [Cell Signaling Technologies (CST), Beverly, MA, USA; 91992; 1:400], CD8α (CST, 70306, 1:400), CD4 (Abcam, Cambridge, UK; 70359, 1:400), CD3 (Abcam, Cambridge, UK; 2779, 1:400), CD11c (Abcam, Cambridge, UK; 16689, 1:400), and Ki67 (Abcam, Cambridge, UK; 12206, 1:400).
ab133616; 1:500), forkhead box P3 (FoxP3; CST; 98377; 1:100), and alpha-smooth muscle actin (α-SMA; Proteintech, Wuhan, China; 14395-1-AP; 1:10000), and corresponding horseradish peroxidase-conjugated secondary antibodies (Panovue, 10013001040). Next, fluorescence signals were amplified with tyramide signal amplification dyes, followed by nucleus counterstaining with 4′,6-diamidino-2-phenylindole (DAPI; Beyotime, Shanghai, China; C1005) and section sealing with the antifade mounting medium (Panovue, 10022001010). The sequential staining protocol is detailed in Supplementary Table 7.

**Imaging and HALO Analysis**

For H&E staining and mIHC, the entire section was subjected to panoramic scanning and digital imaging with the Aperio CS2 Digital Pathology Scanner (Leica) and Vectra Polaris Automated Pathological Imaging System version 1.0 (Akoya Biosciences), respectively. Sections were all scanned at 20× objective magnification. Two consecutive sections underwent H&E staining and mIHC, respectively, to simulate tumor tissues in the same section. The tumor region was observed and annotated on H&E images of successive sections, and then the same annotation was applied to the mIHC section. The scanned images were quantitatively assessed with the HighPlex-FL function of the HALO software version 3.3.25 (Indica Labs), followed by the tissue classification and area quantification of H&E images with the classifier function. Multispectral mIHC images were analyzed with the Multiplex-IHC function.

**Statistics**

All statistics were conducted with GraphPad Prism (version 8.0.2), IBM SPSS Statistic (version 25.0), and R Studio (version 4.0.0) as appropriate. All results were summarized as mean ± standard error of the mean. The paired t-test (two-tailed) or Mann–Whitney U (two-tailed) test was utilized for comparing two groups, and the Kruskal–Wallis test and multiple post hoc comparisons were performed for comparing three or more groups. Associations between categorical variables were evaluated with the Fisher’s exact test. Survival analyses were performed with the Kaplan–Meier analysis and Log rank test using the “survival” package in the R-Studio. Differences were considered statistically significant at bilateral p ≤ 0.05, and exact p values are listed in figures.

**Results**

**CMS Classification and Survival Characteristics of Primary CRC**

The CMS classification has never been applied in the research of CRC metastases. Accordingly, to probe the association of the distribution of CMS groups between PT and paired metastases, we first summarize the clinical characteristics of the four CMSs of PT. Among the analyzed 765 primary CRC samples, the proportion of CMS2 tumors was the highest (GSE39582, 52.7%; TCGA, 58.4%), and the proportion of CMS3 tumors was the lowest (GSE39582, 13.8%; TCGA, 9.4%) (Figure 1A, Supplementary Figure 1A, Supplementary Figure 2A), corresponding to previous studies.30–32 Next, the prevalence of metastases was calculated in each CMS group, which showed that CMS4 was the most prone to metastases, with a metastasis rate of nearly 25% (Figure 1B, Supplementary Figure 1B, Supplementary Figure 2A). Additionally, the survival analysis exhibited that RFS and OS were the shortest in the CMS4 group, followed by the CMS2 group (GEO: RFS, p = 0.00018; OS, p = 0.049, Figure 1C and D; TCGA: RFS, p = 0.071; OS, p = 0.037, Supplementary Figure 1C and D, Supplementary Figure 2C and D).

Furthermore, the clinical baselines of the four CMS groups were compared. As depicted in Supplementary Tables 8 and 9, four CMS groups were significantly different in terms of age, AJCC TNM stage, tumor location, mismatch repair/CpG island methylator phenotype/CIN status, and TP53/KRAS/BRAF mutations. Notably, the clinicopathologic features of CMS4 tumors were consistent with those of advanced CRC. Taken together, CMS4 patients had the greatest risk of metastasis and the worst prognosis. As previously reported, the cause and molecular mechanism of metastasis may be related to the striking upregulation of EMT-related genes, activation of TGF-β, angiogenesis, and interstitial infiltration in CMS4 tumors.30,33
Clinicopathological Features of CRC Metastases

To clarify the clinical features of CRC metastases, a dataset including 975 metastatic CRC patients from cBioportal was obtained, and their epidemiological and clinicopathologic data were analyzed. The liver (N = 693, 71.08%) was the most frequent site of metastasis, followed by lungs (N = 92, 9.44%), peritoneum (N = 86, 8.82%), and distant/non-regional lymph nodes (N = 46, 4.72%) (Figure 2A), broadly concordant with the metastatic patterns observed in previous studies.\(^{34-37}\) In parallel, the degree of tumor differentiation was compared for the four dominant metastases, including...
473 liver metastases, 64 lung metastases, 33 distant lymph node metastases, and 56 peritoneal metastases. The results demonstrated that differentiation was the worst in peritoneal metastasis (the proportion of tumors with moderate-poor and poor differentiation of almost 50%), followed by distant lymph node metastases (Figure 2B, Supplementary Table 10). Conversely, liver and lung metastasis samples were relatively well-differentiated, with a proportion of moderate-poor and poor differentiation of less than 20% (Figure 2B, Supplementary Table 10).

Figure 2 Clinicopathologic features of metastatic CRC. (A) Pie chart showing the distribution rate of the first distant metastasis organs among 975 CRC patients from cBioPortal. (B) The degree of tumor differentiation of liver, lung, distant lymph node, and peritoneal metastases. (C) Kaplan-Meier analysis and log-rank test of the OS of patients with liver, lung, distant lymph node, and peritoneal metastases.
Further, we compared the OS of patients with liver (319 cases), lung (63 cases), distant lymph node (27 cases), and peritoneal metastases (43 cases) whose survival status and survival time were clearly recorded in database. The results showed that patients with peritoneal metastasis had the shortest OS, consistent with previous reports\(^{38,39}\) (Figure 2C). Interestingly, patients with lung metastasis had the best survival before 200 months, but the survival rate decreased linearly after that (Log-rank $\chi^2$, $p = 0.00042$, Figure 2C).

**Figure 3** Comparisons of CMS groups between primary and metastatic CRC. (A) Sample size histogram of liver, lung, distant lymph node, and peritoneal metastases from GEO. (B) The distribution of the four CMS groups of liver, lung, distant lymph node, and peritoneal metastases. (C) Pie chart displaying the proportion of the matched CMS group of 105 pairs of primary CRC and metastases. (D) Sankey chart exhibiting the relationship of the CMS classification between liver, lung, and peritoneal metastases and primary lesions.
CMS Groups of CRC Metastases Mostly Matched Those of Their Primary Tumors

To delve into the distribution of CMS groups in CRC metastases of different organs and tissues, we analyzed the transcriptomic data of 442 metastatic samples from GEO, which were confidently classified, including metastases of liver (N = 326), lung (N = 88), distant lymph node (N = 9), and peritoneum (N = 19) (Figure 3A, Supplementary Tables 1 and 11). The results demonstrated that metastases of different organ sites had distinct CMS patterns. Specifically, the common group for liver metastases was CMS2 (50.6%), followed by CMS4 (41.4%), which was consistent with previous studies that applied CMS in the liver metastasis setting.24,40 A similar result to previous studies was also observed in the peritoneal metastases,31 that peritoneal metastases were dominated by the CMS4 group (78.9%). No systematic analysis is currently available to assess the CMS heterogeneity for metastases of distant lymph nodes and lungs. In our study, we found that CMS3 (66.7%) and CMS4 (61.4%) were the common groups for the metastases of distant lymph nodes and lungs, respectively (Figure 3B, Supplementary Table 11). To further explore the differences of CMS groups between primary CRC and paired metastases, we compared the CMS groups of 105 pairs of primary and metastatic tumors extracted from 10 GEO datasets, including PT and liver (N = 94), lymph node (N = 8), and peritoneal (N = 3) metastases (Supplementary Table 11). The results unveiled that up to 64% (67 of 105 pairs) of primary and metastatic lesions shared the same CMS groups (Figure 3C).

Specifically, primary CRC samples of CMS2, CMS3, and CMS4 mostly metastasized to the liver, distant lymph node, and peritoneum, respectively, and then progressed to CMS2, CMS3, and CMS4 metastatic tumors, respectively (Figure 3D). While the CMS groups were inconsistent between primary and metastatic tumors (the proportion was 36%), liver and peritoneal metastases were still dominated by the CMS2 and CMS4 groups, respectively (Figure 3D). This correspondence of CMS group distribution exhibited that primary tumor cells of different CMS groups might prefer to colonize in specific organs during metastases and retained the unique molecular characteristics of these CMS groups in target tissues. Therefore, new metastatic organs and tissues should not be ignored as the growth environment of metastatic tumors.

TME of CRC Metastases Was Similar to That of Primary Lesions

In our study, several methods were employed to determine whether the correspondence of CMS group distribution represented the similarity of TME between paired metastases and primary lesions. Firstly, principal component analysis (PCA) depicted a high degree of separation between normal colon epithelial tissues and primary CRC tumors and a low degree of separation among tumors of four CMS groups (Figure 4A). Moreover, further results revealed statistically higher tumor purity in CMS2 and CMS3 groups, dramatically higher immune scores in CMS1 and CMS4 groups, and substantially higher stromal scores in the CMS4 group (Figure 4B). To further characterize the cellular heterogeneity landscape of different CMS groups, the xCell algorithm was adopted to evaluate the infiltration of 64 kinds of immune and stromal cells in the TME. Compared to normal colon and tumors of other CMS groups, CMS1 tumors had higher lymphoid cluster infiltration, CMS2 and CMS3 tumors showed “desert” infiltration of almost all cell clusters, and CMS4 tumors had markedly higher infiltration of myeloid, stem-like, and stromal clusters, such as M2-like macrophages, HSCs, MSCs, and fibroblasts (Figure 4C).

Likewise, xCell was utilized to assess the three most prevalent distant metastases, namely liver, lung, and peritoneal metastases. As expected, the enrichment scores of all cell clusters were substantially lower in liver metastases, whereas the enrichment scores of myeloid, stem-like cells, and stromal cells were prominently higher in metastases of lung and peritoneum (Figure 4D). In other words, tumor cells metastasizing to the liver retained the molecular signatures of CMS2 PT, including high tumor purity and low immune and stromal infiltration, while tumor cells colonizing to lungs and peritoneum exhibited a phenotype of high myeloid infiltration and stromal and mesenchymal enrichment identical to CMS4 PT.

Microenvironmental Differences in Normal Liver, Lung, and Peritoneum Tissues

Considering the effects of distant organs on the CMS classification of CRC metastases, we further analyzed the distribution patterns of infiltrating cell subsets in the microenvironment of normal liver (10 cases), lung (7 cases), and peritoneal tissues (4 cases) to deepen the understanding of the microenvironment characteristics of these three tissues and their possible support or influence on metastatic cancer cells. PCA results unraveled marked differences in gene expression among normal liver, lung, and peritoneal tissues (Figure 5A). According to the results of the ESTIMATE
Figure 4 Tumor microenvironment (TME) analysis of primary and metastatic CRC. (A) Principal component analysis (PCA) of gene expression in normal colon tissues and primary CRC of four CMS groups. (B) The “tumor purity”, “immune score”, and “stroma score” of normal colon tissues and primary CRC of four CMS groups evaluated with the ESTIMATE algorithm. (C) The abundance of 64 types of immune and stromal cells infiltrated in normal colon tissues and primary CRC tissues of four CMS groups calculated with the xCell R package. (D) The abundance of 64 types of immune and stromal cells infiltrated in liver, lung, and peritoneal metastases of CRC calculated with the xCell algorithm.
Figure 5: Microenvironment differences of normal liver, lung and peritoneum tissues. (A) PCA of gene expression in normal liver, lung, and peritoneal tissues. (B-C) The “immune score” and “stroma score” of normal liver, lung and peritoneal tissues analyzed with the ESTIMATE algorithm. (D) Heat maps showing the hierarchical clustering of expression of 11,761 genes in normal liver, lung, and peritoneal tissues. (E) The infiltration abundance of 64 types of immune and stromal cells in normal liver, lung, and peritoneal tissues assessed via xCell.
algorithm, the immune score was significantly higher in normal lung and peritoneal tissues than in liver tissues, and the stroma score was also strikingly higher in lung tissues than in liver tissues. However, the stromal score of peritoneal tissues was insignificantly increased as compared to that of liver tissues, possibly due to the small sample size (Figure 5B and C). Cluster stratification maps were plotted to compare the expression of 11,761 genes among normal liver, lung, and peritoneal tissues, which indicated that there were indeed differences among the three tissues (Figure 5D).

Subsequently, we analyzed the microenvironment cell landscape of normal liver, lung, and peritoneal tissues. It was observed that abundant stromal and immune cells were infiltrated in lung and peritoneal tissues, while cells were distributed in liver tissues in a relatively “barren” state (Figure 5E). These data were coincident with our previous finding that liver tissues favor the CMS2 metastasis, whilst lung and peritoneal tissues favor the CMS4 metastasis. In summary, the characteristics of normal liver, lung, and peritoneal tissues were also implicated in the formation and development of metastatic CRC.

mIHC Validation of Differences in Cell Infiltration of Primary CRC and Paired Metastases

To validate the aforementioned results, we retrospectively collected paraffin-embedded tissues of liver (N = 10), lung (N = 9), and peritoneal (N = 8) metastases and paired primary lesions for mIHC to assess immune-stromal infiltration within the TME. M2-like macrophages, CAFs, and regulatory T cells (Treg) in tumor areas were measured with CD206+, α-SMA+, and CD4+ FoxP3+ staining, respectively. The multispectral image analyses evidence that the infiltration of CD4+ T and M2-like macrophages were about 15% in both primary and metastatic tumors, with about 40% of CAFs and less than 10% of CD8+ T and Treg cells, suggesting the critical involvement of immunosuppressive cells in the development and metastasis of CRC (Figures 6-9A and B).

Additionally, the infiltration of these five cell subsets was poorer in liver metastases and paired primary lesions than in lung and peritoneal metastases and their primary lesions (Figures 6-9A and B), consistent with the low immune and stromal scores in CMS2 CRC. In contrast, significantly high M2-like macrophage infiltration and extremely high CAF infiltration were observed in both PT and metastases of lung and peritoneum (Figures 6-9A and B), corresponding to the fact of the CMS4 CRC with high immune and stromal scores.

To determine the influence of the organ-specific microenvironment on immune-stromal infiltration, infiltrated cells were compared between metastatic and matched primary lesions. Intriguingly, compared with PT, the data revealed significantly lower CAF proportion in liver metastases (Figures 6 and 9C), markedly higher proportion of CD4+ T cells and M2-like macrophages in lung metastases (Figures 7 and 9D), and substantially higher proportion of M2-like macrophages and CAFs in peritoneal metastases (Figures 8 and 9E). These results illustrated that disseminated “seed” of cancer cells not only retained their original hereditary characteristics and salient molecular features after metastasizing from primary tumor sites into distant organs but also enhanced or modified certain features according to the new TME to adapt to the new growth “soil” and form metastatic lesions.

Discussion

It is well established that CRC harbors tumor heterogeneity driven by genetic and epigenetic changes.5,30 This heterogeneity represents distinct gene expression profiles of patients and is strongly associated with diverse molecular characteristics and microenvironmental signatures of tumors,16,33,42 thus complicating prognosis estimation, therapeutic regimen selection, and optimal timing for individual CRC patients.32,43,44 Moreover, CRC is an age-related malignancy, with 70% of new CRC diagnoses in those over 65 years old,45 but the age stratification of CRC patients in clinical decisions is not justified. For example, in surgery, although older patients are more prone to severe postoperative complications, there is no significant difference in cancer-specific survival between younger and older patients, as the prognosis of the elderly may be confounded by differences in stage at presentation, tumor site, preexisting comorbidities and type of treatment received.46 In chemotherapy and targeted therapy, elderly patients can benefit from systemic therapies similar to younger patients in tolerance and survival outcomes.47,48 Given this, considering precision treatment for CRC patients based on biomolecular heterogeneity could be a more feasible approach.
Today, the precision treatment of CRC patients is largely based on tumor node metastasis (TNM) stage and certain molecular markers, such as RAS and BRAF mutations and MSI-H/mismatch repair-deficient (dMMR). However, due to tumor heterogeneity, current limited biomarkers cannot perfectly identify and stratify patients responding to specific therapies.

Figure 6 mHIC validation of cell infiltration in paired primary CRC and liver metastases. (A-B) H&E staining images and mHIC multispectral fluorescence images of paired primary CRC and liver metastases marked with the following six markers: α-SMA (sky blue), CD206 (green), CD8 (yellow), CD4 (red), FoxP3 (white), and DAPI (dark blue). The magnification of the tissue panorama is x 10, with a scale of 5 mm. The magnification of the enlarged local image is x 100, with a scale of 200 μm. The magnification of the enlarged image of Treg cells in the upper right corner is x 800, with a scale of 200 μm.

Today, the precision treatment of CRC patients is largely based on tumor node metastasis (TNM) stage and certain molecular markers, such as RAS and BRAF mutations and MSI-H/mismatch repair-deficient (dMMR). However, due to tumor heterogeneity, current limited biomarkers cannot perfectly identify and stratify patients responding to specific therapies.
treatments. For instance, dysregulation of certain serum microRNAs (miRs) has been reported to induce drug resistance in chemotherapy and targeted therapy. Upregulation of miR-19a is correlated with FOLFOX resistance, upregulation of miR-126 with bevacizumab resistance, and upregulation of miR-31, miR-100, miR-125b, and downregulation of miR-7 with...
cetuximab resistance, respectively. In immunotherapy, MSI-H/dMMR has been guideline-approved as a criterion for advanced or metastatic CRC using immune checkpoint inhibitors (PD-1/PD-L1), as MSI-H tumor cells express significantly higher levels of PD-L1 than those with microsatellite stability (MSS)/MMR-proficient (pMMR). However, there is still
Figure 9 The statistical analysis of cell infiltration in paired primary CRC lesions and liver, lung, and peritoneal metastases. (A) The percentages of CD4+ T cells, CD8+ T cells, Treg cells, M2-like macrophages, and CAFs of primary CRC and liver, lung, and peritoneal metastases. (B) The percentages of CD4+ T cells, CD8+ T cells, Treg cells, M2-like macrophages, and CAFs of liver, lung, and peritoneal metastases. (C) Comparisons of the percentages of CD4+ T cells, CD8+ T cells, Treg cells, M2-like macrophages, and CAFs of paired primary lesions and liver metastases. (D) Comparisons of the percentages of the five cell populations in paired primary lesions and lung metastases. (E) The percentages of the above five cell populations in paired primary lesions and peritoneal metastases. Data were expressed as mean ± standard deviation, and specific p-values are marked in the figure.
a heterogeneity in testing MSI status. As reported, a minority of MSI- patients confirmed by polymerase chain reaction (PCR) were diagnosed as pMMR by the initial MMR protein immunohistochemistry (IHC), probably due to somatic mutations of tumors. Therefore, there remains an urgent need for a thorough understanding of tumor heterogeneity to stratify patients with advanced metastatic CRC to accept more precise therapy.

Of note, the heterogeneity of CRC is re-summarized and interpreted by CMS classification, which is the most comprehensive and systematic method of molecular typing method to date, enabling a multi-perspective and omics overview of CRC heterogeneity. In this study, we classified CRC distant metastases with CMS to investigate subtype distribution of CMS in different metastatic organs and evaluated concordance of CMS between PT and distant metastases, thereby further understanding the etiology and processes of CRC distant metastases and revealing certain patterns and mechanisms.

Our data clarified that metastases of liver and distant lymph node were dominated by CMS2 and CMS3 groups, respectively, and CMS4 was the most frequent group for lung and peritoneal metastases. Up to 64% of PT-metastasis pairs had consistent CMS, indicating that PT of different CMSs had a preference for metastasis to specific organs. Specifically, CMS2, CMS3, and CMS4 PT commonly facilitated CMS2 liver metastases, CMS3 distant lymph node metastases, and CMS4 lung or peritoneal metastases, respectively. Due to the low proportion of CMS1 subtype in CRC metastases, and the small sample size of CMS1 PT-metastasis pairs in our study, the metastatic preference of CMS1 tumors was undetermined. In addition, the TME of PT and paired metastases was similar, with similar distribution and infiltration of stromal and immune cells. Consistently, mIHC results validated that liver metastases had low immune and stromal infiltration, while lung and peritoneal metastases presented with normal immune infiltration and high stromal infiltration. Given that CMS integrates multiple molecular features of CRC, including the pattern of immune and stromal infiltration, TME signatures are therefore highly correlated with CMS. CMS was not a reason or “driving factor” for the metastatic preference of PT in different CMS groups, but an interpretation of the molecular associations, including the immune and stromal signatures, between PT and distant metastases.

In addition, among 36% of metastases with inconsistent CMS groups with paired PT, liver metastases were still mainly CMS2 and lung and peritoneum metastases were mainly CMS4. The xCell analysis further confirmed the presence of “desert” immune and stromal infiltration in normal liver tissues but high immune and stromal cell infiltration in normal lung and peritoneal tissues, highlighting that the formation of CRC metastases not only inherits the genetic mutations of the disseminated “seeds” but also is shaped and influenced by the tissue microenvironment of the distant organs. Generally, the spreading “seeds” colonize into certain tissues and organs whose microenvironment is similar to that of their PT, resulting in the formation of metastatic tumors that resemble PT. For the three most prevalent CRC metastases, when re-visited the “Seed and Soil” theory from a CMS perspective, we propose that CMS2 PT may prefer the “immune tolerant” liver to develop an immune suppressive and relatively low immune-stromal infiltrating TME similar to that in the primary lesion. However, CMS4 PT prefers the “immunogenic enrichment” lung, as myeloid cells could remodel the pre-metastatic lung into an inflamed but immune-suppressive environment, thereby inducing EMT and developing lung metastasis. Moreover, the peritoneum exhibits a high expression of fibroblast, which can be stimulated by TGF-β signaling and transdifferentiate into myofibroblast and interact with tumor cells to establish the peritoneal metastasis, therefore, CMS4 PT prefers peritoneum as a metastatic organ could be explained. Even with an incorrect selection of “seeds”, the combined effects of genetics and environment can allow for the formation of metastases with different CMS groups from primary lesions.

Our results further explain and complement the “Seed and Soil” theory. The CMS of CRC metastasis is collectively determined by the genetic characteristics of primary cancer cells and the microenvironmental characteristics of the metastasized organ tissues. After metastasizing into distant organs and tissues via circulating blood or lymph flow, cancer cell “seeds” not only retain their robust genetic characteristics but also interact with surrounding cells in the new microenvironment, finally developing into metastatic tumors.

Intriguingly, mIHC results demonstrated that the proportion of Treg cells, M2-like macrophages, and CAFs was markedly higher in lung and peritoneal metastases than in liver metastases. Compared with primary lesions, liver metastases presented with significantly decreased CAFs, lung metastases had prominently elevated CD4+ T cells and
M2-like macrophages, and peritoneal metastases showed substantially enhanced M2-like macrophages and CAFs. Based on these results, it can be concluded that different treatment measures should be selected for metastatic CRC in different organs and tissues to accurately and personally curb its formation and development. In addition, immune checkpoint inhibitors could probably benefit less in patients with liver metastases than those with lung or peritoneal metastases, while therapies targeting CD4+ T cells, M2-like macrophages, and CAFs may be considered for patients with lung and peritoneal metastases.

Conclusively, we used CMS to classify CRC metastases and summarized the pattern and characteristics of CRC metastases. Our findings emphasized that the CMS groups of CRC metastases were determined by both the genetic characteristics of disseminated primary tumor cells and the microenvironment of the metastasized organ tissues, which is a supplement to the “Seed and Soil” theory. Meanwhile, our study unveiled the metastasis preference of PT of different CMS groups, indicating that potentially metastatic organs should be monitored following primary tumor surgery, and subtype-based interventions could be considered for timely benefit. Additionally, differences in immune-stromal infiltration among CRC metastases of different organs provide more precise therapeutic targets for the treatment of patients with distant metastases and could probably provide reference for clinical-decision making.

Data Sharing Statement

Ethical Approval & Informed Consent
Our study complied with the Declaration of Helsinki and was approved by the Independent Ethics Committee of Guangzhou First People’s Hospital (Approval no. K-2019-070-01). The informed consent was obtained from all patients included in the study.

Author Contributions
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding
This work was supported by grants from the National Natural Science Foundation of China (Grant no. 82173236), the Science and Technology Planning Project of Guangdong Province (Grant no. 202206080008), the program of Guangdong Provincial Clinical Research Center for Digestive Diseases (Grant no. 2020B1111170004), Guangzhou High-level Key Clinical Specialty Construction Project (No.9), the Project of Key Medical Discipline in Guangzhou (2021–2023) and Guangzhou Science and Technology Planning Project (Grant no. 202206080008), Guangzhou Planned Project of Science and Technology (Grant no. 202102010024), Guangdong Basic and Applied Basic Research Foundation (Grant no. 2021B1515420004), the Science and Technology Projects in Guangzhou (Grant no. 202201020363).

Disclosure
The authors report no conflicts of interest in this work.

References


47. Pecsi B, Mangel LC. The Real-Life Impact of mFOLFIRI-Based Chemotherapies on Elderly Patients-Should We Let It or Leave It? *Cancers*. 2023;15(21).