Original Article

DOI: 10.5582/bst.2025.01297

Integrative methylation and transcriptomic analysis reveals key genes linking cellular senescence and metabolic reprogramming in colorectal liver metastasis

Xiaochen Ma[1], Lu Chen[2], Chenhe Yi[1], Yitong Li[1], Yan Geng[1], Baorui Tao[1], Jinhong Chen[1,*]

[1] Department of General Surgery, Hepatobiliary Surgery Center, Huashan Hospital & Cancer Metastasis Institute, Fudan University, Shanghai, China;
[2] Department of Hepatobiliary Cancer, Liver cancer research center, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin Key Laboratory of Digestive Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin, China.

SUMMARY: Colorectal liver metastasis (CRLM) remains lethal, and the convergence of cellular senescence with metabolic reprogramming *via* epigenomic rewiring is poorly understood. We integrated genome-wide DNA methylation and RNA-seq data from 10 paired primary tumors and liver metastases (GSE213402). After calling differentially methylated genes (3,399 hyper- and 9,519 hypomethylated) and differentially expressed genes (406 DEGs), we intersected them with curated senescence (n = 866) and metabolic reprogramming (n = 948) gene sets, yielding 28 differentially expressed cellular-senescence-related genes (DE-CSRGs) and 24 metabolic-reprogramming-related genes (DE-MRRGs). Machine-learning pipelines (LASSO + SVM-RFE) converged on a five-gene signature: *CXCL1*, *SERPINE1*, *NDRG1*, *SRM* and *GATM*, most of which are hypomethylated and over-expressed in metastases. Gene-set enrichment analysis revealed that these genes are involved in pathways such as oxidative phosphorylation, focal adhesion, complement–coagulation cascades, and PPAR signaling. Immune de-convolution revealed strong positive correlations between signature genes and immunosuppressive subsets (MDSCs, Tregs, type-1 T-helper cells; $p < 0.05$). Elevated IC_{50} values for oxaliplatin and 5-fluorouracil in metastatic samples were positively associated with *NDRG1* and negatively with *SRM*, indicating chemo-resistance modulation. This five-gene epigenetic–transcriptomic hub identifies a molecular signature that warrants prospective validation as a potential biomarker for patient stratification and combination therapy in CRLM.

Keywords: DNA hypomethylation, machine-learning signature, chemoresistance, immunosuppressive microenvironment, polyamine metabolism

1. Introduction

Colorectal liver metastasis (CRLM) remains a major clinical challenge, as colorectal cancer (CRC) is the second leading cause of cancer-related mortality globally, and the liver is the primary site of metastasis. CRLM occurs in 20–50% of patients with Colorectal Cancer (CRC), with a 5-year survival rate of 10–30%, depending on the resectability of metastases (*1*). Current treatment options include surgical resection, systemic chemotherapy, radiation therapy, and ablation therapy; however, only a minority of patients are eligible for curative resection (*2*). Immunotherapy has emerged as a promising therapeutic approach, particularly when combined with local therapies. Nonetheless, prognosis remains poor, underscoring the need for novel biomarkers to improve prognostic accuracy and guide personalized treatment strategies.

Cellular senescence is an irreversible cell cycle arrest triggered by stressors, such as DNA damage and oxidative stress (*3*). It is characterized by enlarged cell size, expression of cell cycle inhibitors (*e.g.*, p16 and p21), and a senescence-associated secretory phenotype (SASP), which includes pro-inflammatory cytokines and extracellular matrix (ECM) remodeling proteins (*4*). Senescence can act as a tumor suppressor by halting the proliferation of damaged cells (*5*); however, SASP can also promote tumor progression by altering the tumor microenvironment (*6*). Metabolic reprogramming, which involves adaptive changes in the cellular metabolism that support rapid proliferation, is a hallmark of cancer cells (*7*). In CRLM, metabolic reprogramming enables cancer cells to adapt to the liver microenvironment, facilitating metastatic colonization (*8,9*). The interplay between senescence and metabolic reprogramming in CRLM is complex, with senescent cells potentially

inducing metabolic changes in neighboring cells, and metabolic alterations reciprocally influencing senescent phenotypes. Identifying biomarkers associated with these processes is crucial for elucidating the pathogenesis of CRLM and may provide novel therapeutic targets.

In this study, we utilized publicly available methylation and transcriptomic data on CRLM to investigate the crucial roles and prognostic value of genes related to cellular senescence and metabolic reprogramming in disease progression, employing bioinformatic techniques. We further investigated potential molecular regulatory mechanisms and drug responses associated with these genes. Our findings provide novel theoretical support and a valuable reference for clinical research and prognostic evaluation in CRLM.

2. Materials and Methods

2.1. Sample collection and RNA-sequencing

Fresh tissue samples, including primary colorectal tumors and matched liver metastases, were collected from patients diagnosed with CRC at HuaShan Hospital between 2014 and 2020. All samples were obtained immediately following surgical resection, snap-frozen in liquid nitrogen, and stored at –80 °C until further processing. All clinical samples were collected after obtaining informed consent and in accordance with a protocol approved by the Ethics Committee of Huashan Hospital, Fudan University (Shanghai, China). All clinical samples were collected from patients after obtaining informed consent in accordance with a protocol approved by the Ethics Committee of Huashan hospital, Fudan University (Shanghai, China).

2.2. Data source

RNA sequencing data and clinical information were obtained from the TCGA database *via* UCSC Xena (*http://xena.ucsc.edu/*). TCGA-COAD and TCGA-READ data were merged to generate a TCGA-CRC dataset, and only samples with available survival information were retained. Ultimately, 607 tumor samples and 51 normal samples from the TCGA were used for prognostic analysis. The transcriptome and DNA methylation data of 10 paired primary tumors and liver metastases from 10 patients with CRC were sourced from the GSE213402 dataset in the GEO database to screen for metastasis-related genes. In total, 866 cellular senescence-related genes (CSRGs) were obtained from the CellAge database (*https://genomics.senescence.=cells/*), and 948 metabolic reprogramming-related genes (MRRGs) were identified from a literature search (*10*).

2.3. Identification of methylated genes involved in metastasis

First, we identified differentially methylated CpGs (DMCs) between liver metastasis and primary tumor groups using thresholds of p-value < 0.05 and $|\Delta\beta| > 0.2$ and investigated the distribution of DMCs across different CpG regions. Next, we mapped CpG sites to their corresponding genes to identify the methylated genes involved in metastasis. Genes were classified into distinct methylation states based on the ratio of hypermethylated to hypomethylated CpGs, using a 1.5-fold threshold. Specifically, genes with a hypermethylated/hypomethylated CpG ratio ≥ 1.5 were defined as hypermethylated, and those with a hypomethylated/hypermethylated ratio ≥1.5 were defined as hypomethylated. To analyze the biological pathways associated with methylated genes, functional enrichment analysis was conducted using the R package "clusterProfiler."

2.4. Identification of differentially methylated CSRGs and MRRGs

We used the DESeq2 package to identify differentially expressed genes (DEGs) between the liver metastasis and primary tumor groups, with a false discovery rate (FDR) $P_{adj} < 0.05$ and $|\log2FC| > 1$. To identify differentially expressed cellular senescence-related genes (DECSRGs) and metabolic reprogramming-related genes (DEMRRGs), we performed an overlap analysis among DEGs, 866 CSRGs, and 948 MRRGs. Differentially methylated cellular senescence-related genes (DM-CSRGs) were identified through intersection analysis of (*i*) hypermethylated genes with downregulated DECSRGs, and (*ii*) hypomethylated genes with upregulated DECSRGs. Similarly, differentially methylated metabolic reprogramming genes (DM-MRRGs) were derived from the intersections of (*i*) hypermethylated genes and downregulated DEMRRGs, and (*ii*) hypomethylated genes and upregulated DEMRRGs. Spearman's correlation analysis was conducted on DM-CSRGs and DM-MRRGs to evaluate their association. Correlated DM-CSRGs and DM-MRRGs were screened based on a threshold of $p < 0.05$ and $|cor| > 0.3$.

2.5. Machine learning

To identify key genes, two machine learning algorithms, support vector machine-recursive feature elimination (SVM-RFE) and least absolute shrinkage and selection operator (LASSO), were employed. Initially, we utilized the 'glmnet' package in R for LASSO analysis, which selects key genes from a set of candidate genes through binomial logistic regression. Subsequently, the SVM package was employed to implement the SVM-RFE model using a 10-fold cross-validation strategy. Key CSRGs and MRRGs were identified by overlapping the feature genes obtained from LASSO and SVM-RFE.

2.6. Characterization of key genes

First, we analyzed the expression levels of key genes in the primary tumors and liver metastases. Subsequently, to investigate the prognostic value of the key genes in CRC, we divided TCGA-CRC samples into high- and low-expression groups based on the median expression levels of the key genes, followed by Kaplan–Meier (KM) survival analysis. Gene set enrichment analysis (GSEA) was performed to investigate the functions of key genes using c2.cp.kegg.v7.5.1.symbols.gmt gene set as the reference. Subcellular localization analysis of key genes was performed using Cell-Ploc 2.0.

2.7. Analysis of immune microenvironment and drug sensitivity

The infiltration of 28 immune cell types, including B cell memory, eosinophil, macrophage, mast cell, monocyte, neutrophil, NK cell, effector memory CD4 T cell, effector memory CD8 T cell, activated CD4 T cell, activated CD8 T cell, T cell follicular helper, T cell gamma delta, T cell regulatory, activated B cell, activated dendritic cell, CD56bright natural killer cell, CD56dim natural killer cell, central memory CD4 T cell, central memory CD8 T cell, immature B cell, immature dendritic cell, MDSC, natural killer T cell, plasmacytoid dendritic cell, T cell helper cell type 1, T cell helper cell type 2, and T cell helper cell type 17 were compared between liver metastasis and primary tumor groups. Additionally, the half-maximal inhibitory concentration (IC_{50}) values of irinotecan (1088), oxaliplatin (1089), and 5-fluorouracil (1073) were calculated using the "oncoPredict" R package to assess the drug sensitivity of patients with primary tumor and liver metastasis.

2.8. Statistical analysis

A regulatory network was constructed using Cytoscape (version 3.9.1). All statistical results were analyzed using R (version 4.4.0). KM curves and log-rank tests were used to visualize and test the survival differences among the different groups. Differences in the continuous variables between the two groups were examined using the Wilcoxon test. A p-value < 0.05, unless otherwise specified, was considered statistically significant.

3. Results

3.1. Analysis of the landscape of methylation in the metastasis of CRC

We performed differential methylation analysis and identified 98,858 DMCs between the liver metastasis and primary tumor groups, including 31,883 hypermethylated and 66,975 hypomethylated sites (Figure 1A). The genomic distribution of these DMCs across the 24 chromosomes revealed that the majority of DMCs were hypomethylated (Figure 1B). Moreover, the proportions of hypomethylation in CpG island core, CpG island shelf, and CpG island shores were 53.7%, 67.5%, and 62.8%, respectively, exceeding the corresponding proportions of hypermethylation (Figure 1C). Based on the CpG sites, 3,399 hypermethylated genes and 9,519 hypomethylated genes were identified. The enrichment results of Gene Ontology showed that genes were associated with biological processes such as embryonic organ development, axonogenesis, synapse organization, cellular components related to synaptic membrane, neuronal cell body, glutamatergic synapse, and molecular functions such as GTPase regulator activity, DNA-binding transcription activator activity, and phospholipid binding (Figure 1D). Kyoto Encyclopedia of Genes and Genomes analysis revealed that these genes were involved in the PI3K-Akt signaling pathway, human papillomavirus infection, MAPK signaling pathway, gastric cancer, breast cancer, and ECM-receptor interaction (Figure 1E).

3.2. Identification of DM-CSRGs and DM-MRRGs

We identified 406 DEGs between the liver metastasis and primary tumor groups, including 250 upregulated and 156 downregulated genes in the primary tumor group (Figure 2A). By overlapping the 866 CSRGs and 948 MRRGs, 28 DECSRGs and 24 DEMRRGs were identified (Figure 2B). Eight DM-CSRGs (*CXCL1*, *NDRG1*, *SERPINE1*, *WSB1, CCL2, EGR2, XAF1*, and *VCAN*) and seven DM-MRRGs *(IL4I1, DPYD, ALOX5, CYP1B1, HNMT, SRM*, and *GATM*) were identified (Figure 2C). Among these, *CXCL1* and *SRM* were hypermethylated, whereas *NDRG1*, *SERPINE1*, *WSB1, CCL2, EGR2, XAF1, VCAN, IL4I1, DPYD, ALOX5, CYP1B1, HNMT,* and *GATM* were hypomethylated. We observed multiple correlations among these genes; *SRM* exhibited a strong positive correlation with *CXCL1* (r = 0.71). *SERPINE1* positively correlated with DPYD (r = 0.60), ALOX5 (r = 0.58), and CYP1B1 (r = 0.64). CCL2 expression was positively associated with the expression of IL4I1 (r = 0.60), DPYD (r = 0.90), ALOX5 (r = 0.58), and CYP1B1 (r = 0.88). VCAN also positively correlated with IL4I1 (r = 0.59), DPYD (r = 0.63), ALOX5 (r = 0.62), and CYP1B1 (r = 0.78). Conversely, a negative correlation was detected between *CXCL1* and HNMT (r = -0.56) (Figure 2D).

In the GSE213402 dataset, the expression levels of *GATM*, *NDRG1*, and *SERPINE1* were elevated, whereas those of *CXCL1* and *SRM* were reduced in the liver metastasis group compared with those in the primary tumor group (Figure 3A). The expression patterns of these genes were consistent with the sequencing data (Figure 3B). Additionally, the expression levels of *CXCL1*, *SERPINE1*, and *SRM* were upregulated, whereas those of *GATM* and *NDRG1* were downregulated in

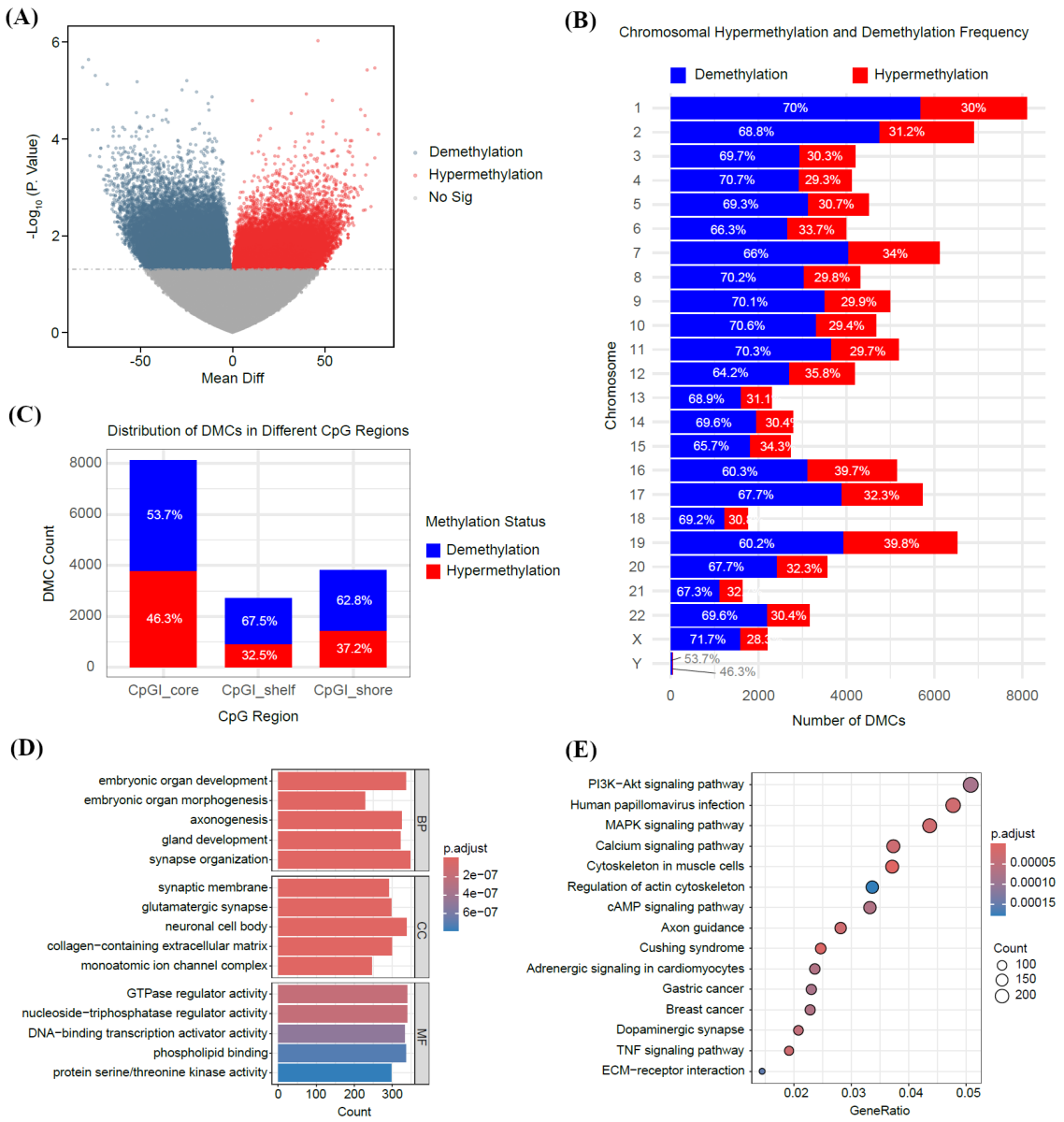

Figure 1. Genome-wide methylation landscape distinguishing colorectal liver metastasis from primary tumors. (A) Volcano plot of 98,858 differentially methylated CpG sites (DMCs; $|\Delta\beta| > 0.2$ and $p < 0.05$) between liver metastasis and primary tumor samples in the GSE213402 cohort. Hypermethylated probes ($n = 31,883$) are shown in red; hypomethylated probes ($n = 66,975$) are shown in blue. **(B)** Bar chart showing the number of hypermethylated (red bars) and hypomethylated (blue bars) DMCs across the 24 human chromosomes. **(C)** Proportional distribution of DMCs within CpG-island sub-regions: CpG-island core, shore and shelf. Hypomethylation predominates in all three compartments (53.7 %, 62.8 % and 67.5 %, respectively). **(D)** Top Gene Ontology (GO) biological-process and molecular-function terms enriched among genes harbouring metastasis-associated DMCs. Bars represent –log10 (adjusted p value); the dashed vertical line indicates adjusted $p = 0.05$. **(E)** KEGG pathway enrichment bubble chart for the same gene set. Bubble size is proportional to the number of genes; colour intensity reflects –log10 (adjusted p value). Key cancer-related pathways (PI3K-Akt, MAPK, ECM-receptor interaction) are highlighted.

the tumor group compared with those in the control group (Figure 3C). The expression trends of *GATM*, *SRM*, *CXCL1*, and *NDRG1* were consistent with our sequencing data (Figure 3D). Furthermore, KM curves were plotted for the five genes, and significant survival differences were observed for *CXCL1* ($p = 0.031$) and *GATM* ($p = 0.038$), indicating their potential as prognostic biomarkers (Figure 3E). We performed GSEA to gain a deeper understanding of the potential mechanisms and found that *NDRG1* and *SERPINE1* were associated with ECM-receptor interaction, focal adhesion, Leishmania infection, lysosome, complement, and coagulation cascades. *GATM* and *SRM* were related to ribosomes, Huntington's disease, Parkinson's disease, and oxidative phosphorylation (OXPHOS), while *CXCL1* was associated with the PPAR signaling pathway (Figure 4A-4E).

3.3. Machine learning identified five key cellular senescence-metabolic genes

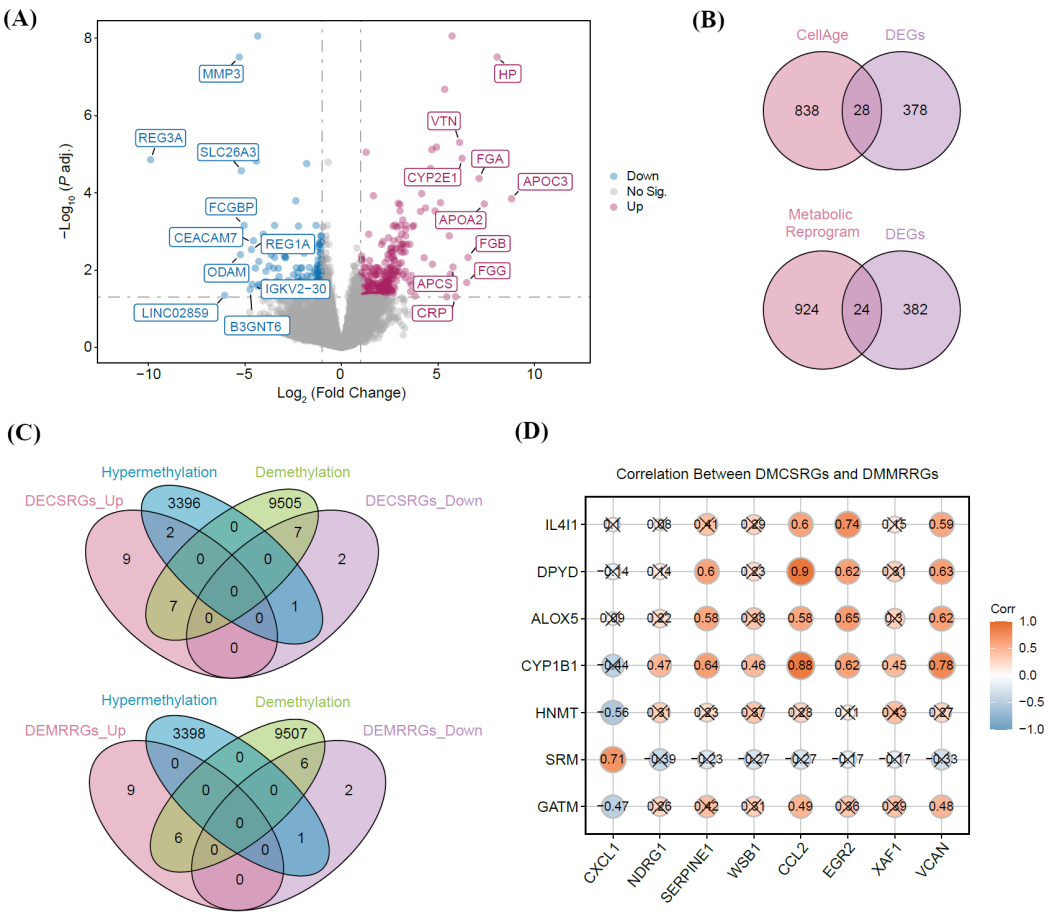

Figure 2. Identification and epigenetic–transcriptional coupling of senescence- and metabolism-related genes in CRLM. (A) Volcano plot of differentially expressed genes (DEGs) between liver metastasis and primary tumour samples (GSE213402). Red, up-regulated genes (n = 250); blue, down-regulated genes (*n* = 156) (FDR < 0.05 and |log2FC| > 1). **(B)** Venn diagrams depicting the intersection between DEGs and the curated gene sets: 28 differentially expressed cellular-senescence-related genes (DE-CSRGs, top) and 24 differentially expressed metabolic-reprogrammingrelated genes (DE-MRRGs, bottom). **(C)** Strategy and numbers used to define differentially methylated CSRGs (DM-CSRGs, *n* = 8, top) and differentially methylated MRRGs (DM-MRRGs, *n* = 7, bottom) by overlapping methylation status (hyper/hypo) with expression direction (down/up). **(D)** Spearman correlation heat-map of the 15 overlapping differentially methylated CSRGs and MRRGs. Colour intensity indicates correlation coefficient (red, positive; blue, negative); only correlations with |r| > 0.3 and $p < 0.05$ are shown.)

LASSO regression and SVM-RFE were employed to identify key CSRGs and MRRGs among the 15 candidate genes. Of the 15 candidate genes, 6 genes, namely *GATM*, *SRM*, HNMT, *CXCL1*, *NDRG1*, and *SERPINE1*, were selected in the lambda.min model (Figure 5A). SVM-RFE with 10-fold cross-validation selected eight genes: *CXCL1*, *SRM*, *GATM*, *NDRG1*, *WSB1, SERPINE1*, *CYP1B1*, and *CCL2* (Figure 5B)). Finally, five overlapping genes from LASSO and SVM-RFE, namely, *CXCL1*, *SRM*, *GATM*, *NDRG1*, and *SERPINE1*, were identified as key CSRGs and MRRGs involved in CRC metastasis (Figure 5C)). *CXCL1* is involved in oncogene-induced senescence, whereas *SERPINE1* participates in replicative senescence. *SRM* has been implicated in glutathione, beta-alanine, arginine, proline, cysteine, and methionine metabolism. Additionally, *GATM* was involved in arginine and proline metabolism as well as glycine, serine, and threonine metabolism (Figure 5D)). Among these five genes, *CXCL1* and *SERPINE1* were predicted to be localized in the extracellular space, *SRM* in the cytoplasm, *GATM* in the mitochondria, and *NDRG1* in both the cytoplasm and nucleus (Table 1).

3.4. Key genes may affect the immune microenvironment and drug sensitivity of patients with CRLM

Liver metastasis, the leading cause of CRC-related mortality, is characterized by a highly heterogeneous and suppressive immune microenvironment (*11*). We compared immune cell infiltration profiles between the liver metastasis and primary tumor groups. The infiltration of MDSCs, natural killer cells, natural killer T cells, regulatory T cells, and type 1 T helper cells was significantly higher in the liver metastasis group (Figure 6A). Strong positive correlations were observed between *SERPINE1* and MDSCs (correlation coefficient, COR = 0.72), natural killer T cells (COR = 0.9), natural killer cells (COR = 0.82), regulatory T cells (COR = 0.85), and type 1 T helper cells (COR

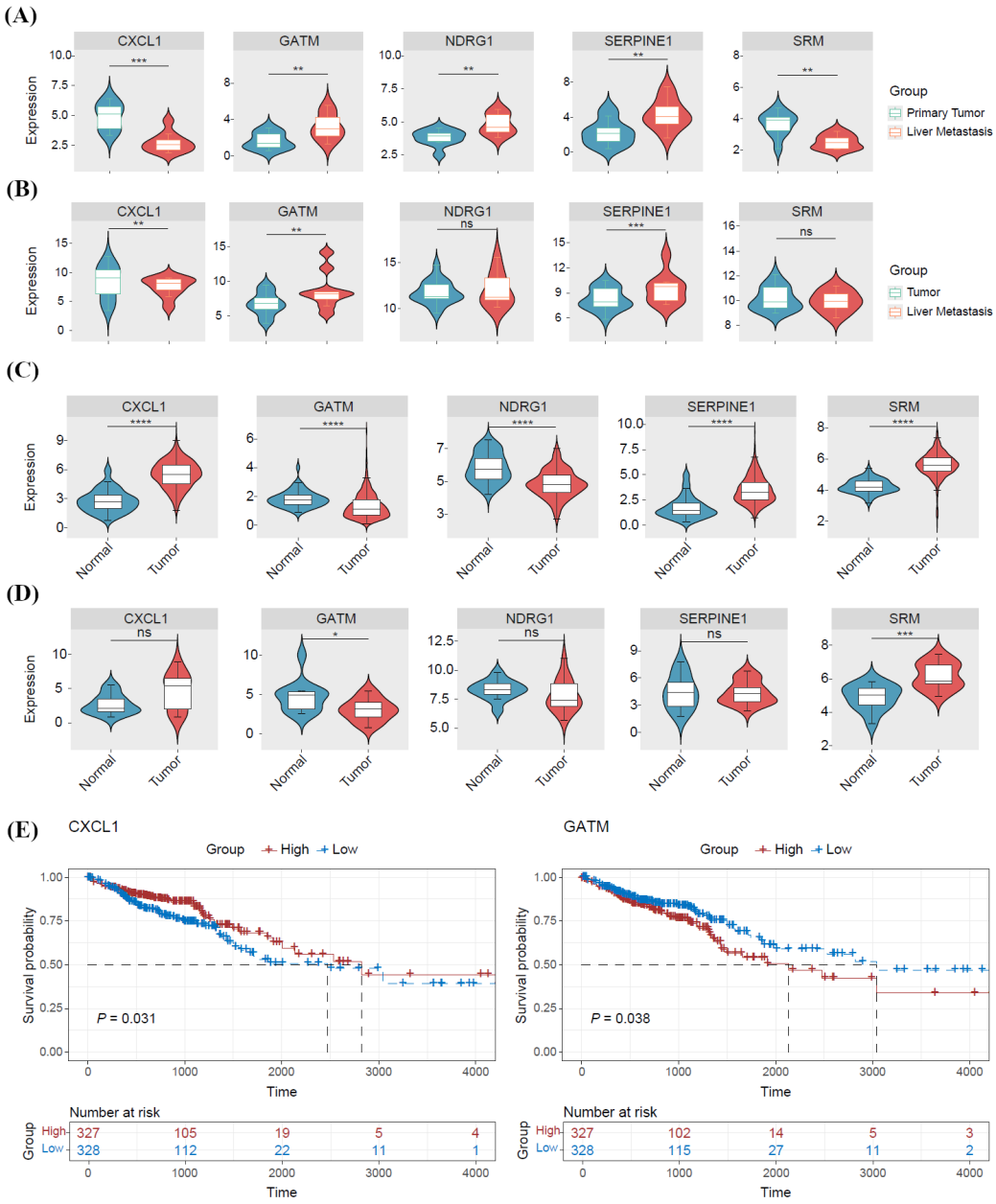

Figure 3. Transcript abundance and prognostic value of the five-gene senescence–metabolism signature. (A) Expression levels of *CXCL1*, *SERPINE1*, *NDRG1*, *SRM* and *GATM* in the discovery (GSE213402) cohort. Data are shown as fragments per kilobase of transcript per million mapped reads (FPKM) in primary tumours (n = 10) versus liver metastases (n = 10). P values were calculated with the paired Wilcoxon test. **(B)** RNA-seq validation of the same five genes in an independent set of 10 patient-matched fresh-frozen primary CRC and synchronous liver metastases collected at our centre (Huashan Hospital, 2014-2020). Read counts were normalized to TPM; boxes represent median ± interquartile range. **(C)** TCGA-COAD/READ dataset: box-plots comparing tumour (n = 607) versus adjacent normal mucosa (n = 51) for each signature gene. P values, Wilcoxon rank-sum test. **(D)** Huashan cohort: RNA-seq comparison of tumour versus adjacent normal mucosa (n = 10 pairs). P values, paired t test.**(E)** Kaplan-Meier plots of overall survival in TCGA-CRC patients stratified by median expression of *CXCL1* (left) and *GATM* (right). Log-rank P values are shown.

= 0.84) (Figure 6B). *SERPINE1* positively correlated with the immune functions of antigen-presenting cell co-stimulation (COR = 0.52), chemokine receptor signaling (COR = 0.59), immune checkpoint activity (COR = 0.55), parainflammation (COR = 0.67), T cell co-inhibition (COR = 0.55), and type 1 interferon response (COR = 0.61) (Figure 6C). *SERPINE1* was positively correlated with HLA-DQA2 (COR = 0.45), HLA-DQA1 (COR = 0.51), and HLA-DRB1 (COR = 0.46) (Figure 6D). These results suggest that *SERPINE1* is closely associated with the metastatic immune microenvironment. Variations in the immune microenvironment may influence drug sensitivities in patients with cancer. Thus, we compared sensitivity to chemotherapy between the metastatic and primary groups. A significant difference in the IC_{50} values was detected for the 30 therapeutic agents. Among them, oxaliplatin and 5-fluorouracil, which are commonly

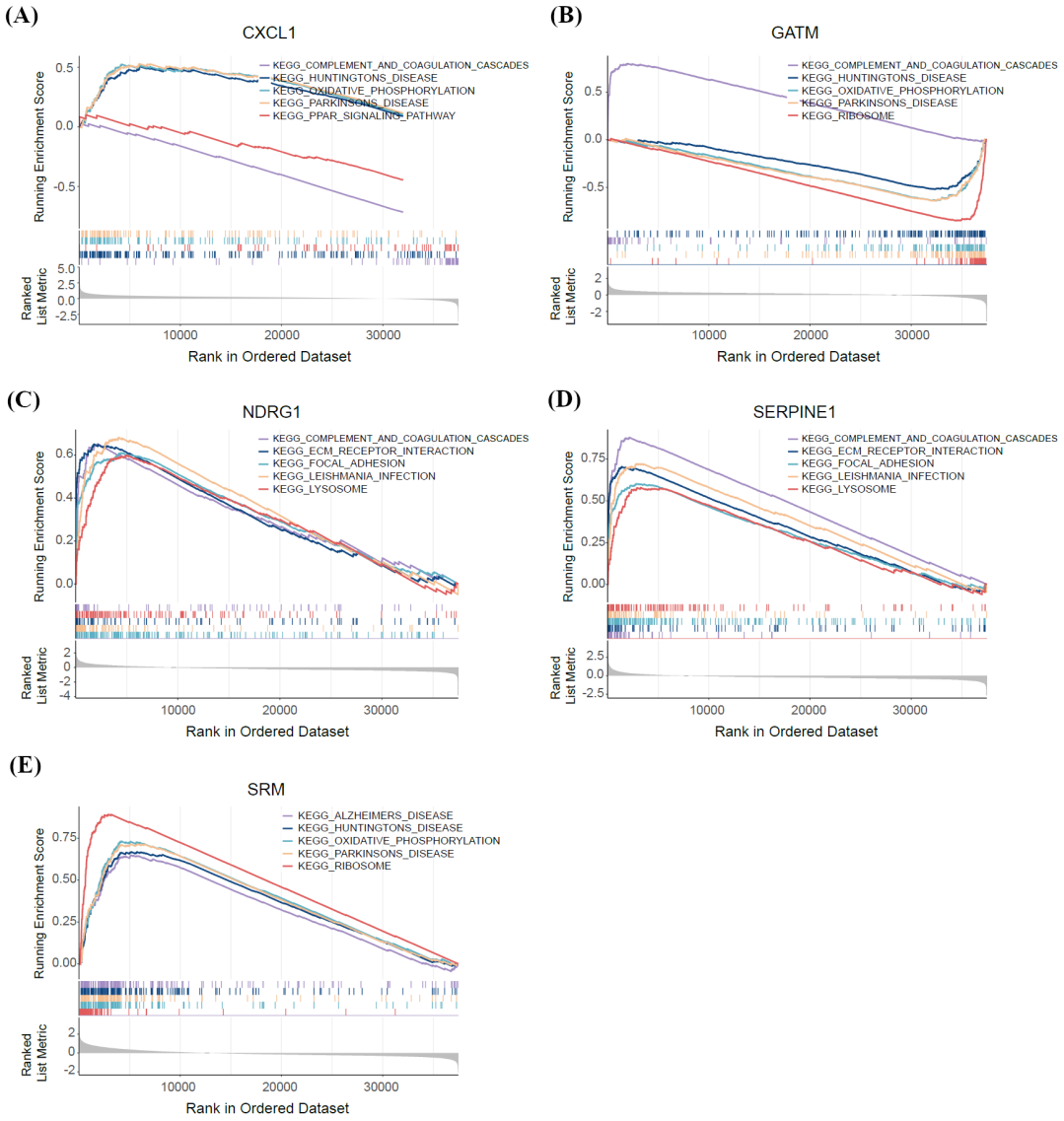

Figure 4. Gene-set enrichment analysis (GSEA) of the five-gene senescence–metabolism signature in CRLM. (A) *GATM* GSEA. Leading-edge subsets are enriched for ribosome, oxidative phosphorylation and Huntington-disease pathways. **(B)** *CXCL1* GSEA. Positive enrichment is observed for PPAR signalling, cytokine–cytokine receptor interaction and chemokine signalling pathways. **(C)** *NDRG1* GSEA. ECM-receptor interaction, focal adhesion and complement–coagulation cascades exhibit significant enrichment. **(D)** *SERPINE1* GSEA. Complement–coagulation cascade, ECM-receptor interaction and lysosomal pathways are positively enriched. **(E)** *SRM* GSEA. Oxidative phosphorylation, Parkinson-disease and ribosome pathways show significant enrichment.Genes were ranked by log2FC (metastasis vs primary tumour, x-axis; positive log2FC indicates upregulation in metastases); coloured lines trace the running enrichment score (y-axis) for the indicated pathways. All analyses used the c2.cp.kegg.v7.5.1.symbols.gmt gene set; FDR < 0.05 was considered significant.

used as chemotherapeutic agents for CRC, exhibited significantly higher IC_{50} values in the liver metastasis group than in the primary tumor group (Figure 7A). Furthermore, *SRM* expression was negatively correlated with IC_{50} values of 5-FU (r = -0.55), indicating that high *SRM* expression is associated with increased sensitivity to 5-FU. *SRM* negatively correlated with oxaliplatin levels (r = -0.45). Conversely, *NDRG1* positively correlated with oxaliplatin levels (r = 0.46), whereas *CXCL1* showed a negative correlation (r = -0.45) (Figure 7B). These results suggest the involvement of key genes in the mechanisms relevant to chemotherapeutic resistance in CRC metastasis.

4. Discussion

CRLM remains the dominant cause of death in patients with CRC; however, the molecular programs that allow disseminated tumor cells to survive, adapt, and ultimately colonize the hepatic parenchyma remain poorly understood. By integrating genome-wide DNA methylation profiles with bulk RNA sequencing data and curated gene sets for cellular senescence and metabolic reprogramming, we identified a five-gene epigenetic–transcriptomic signature comprising *CXCL1*, *SERPINE1*, *NDRG1*, *SRM*, and *GATM*, which functionally couples senescence-associated secretory traits with

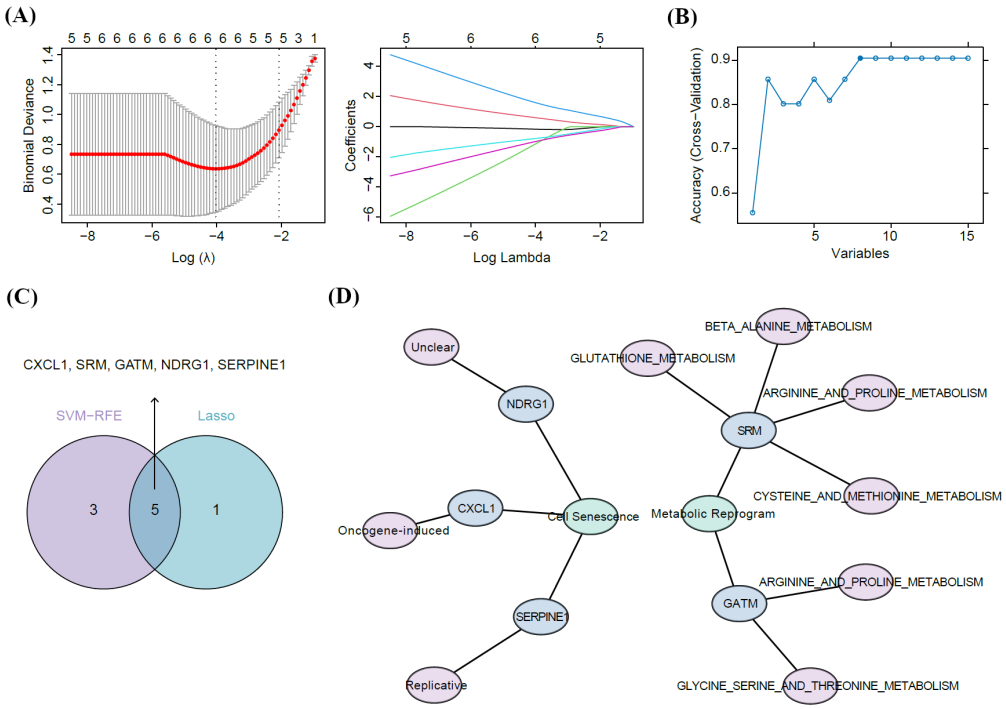

Figure 5. Machine-learning identification of a five-gene senescence–metabolism signature predictive of CRLM. (A) Left: Ten-fold cross-validated binomial deviance (y-axis) plotted against log(λ) (x-axis) for LASSO logistic regression of 15 candidate DM-CSRGs/DM-MRRGs. Red dots indicate mean deviance ± 1 SE; vertical dashed lines mark λ_min (left) and λ_1SE (right). Right: Profile of regression coefficients (y-axis) versus log(λ) (x-axis); each coloured line represents one gene, illustrating coefficient shrinkage with increasing penalty. Six genes (*GATM*, *SRM*, *HNMT*, *CXCL1*, *NDRG1*, *SERPINE1*) survived at λ_min. **(B)** SVM-RFE recursive feature elimination. Mean cross-validation accuracy (y-axis) is plotted as a function of the number of input genes (x-axis). The peak accuracy (eight genes) is indicated by the red dot; the selected subset comprises *CXCL1*, *SRM*, *GATM*, *NDRG1*, *WSB1*, *SERPINE1*, *CYP1B1* and *CCL2*. **(C)** Venn overlap of genes retained by LASSO (purple) and SVM-RFE (blue) yielding the final five-gene signature: *CXCL1*, *SERPINE1*, *NDRG1*, *SRM* and *GATM*. **(D)** Functional classification of the five signature genes according to their primary roles in cellular-senescence pathways (*CXCL1*, *SERPINE1*, *NDRG1*) or metabolic reprogramming (*SRM*, *GATM*).

Table 1. Subcellular localization prediction of the five hub genes

Gene symbol	Predicted primary localization	Prediction score	Experimental evidence (UniProt)
CXCL1	Extracellular	0.91	Secreted chemokine (P09341)
SERPINE1	Extracellular	0.89	Secreted serpin (P05121)
SRM	Cytoplasm	0.83	Soluble cytoplasmic enzyme (P19623)
GATM	Mitochondrion	0.79	Mitochondrial matrix protein (P50440)
NDRG1	Cytoplasm / Nucleus	0.76 / 0.71	Dual-localized scaffold protein (Q92597)

ND: Predictions were generated with Cell-Ploc 2.0 (*http://www.csbio.sjtu.edu.cn/bioinf/Cell-Ploc-2/*). Scores represent the highest voting confidence among 11 integrated algorithms.

metabolic plasticity in CRLM. The global shift toward hypomethylation observed in liver metastases (66,975 hypomethylated versus 31,883 hypermethylated CpGs) recapitulates previous reports describing widespread methylation erosion during metastatic progression (*12*) and supports the concept that the loss of DNA methylation fidelity facilitates chromosomal instability and enhancer activation (*13*). Pathway-level annotation has consistently implicated the PI3K–Akt, MAPK, and ECM-receptor interaction cascades, all of which are linked to CRC cell extravasation and hepatic colonization (*14-16*).

CXCL1, *SERPINE1*, *NDRG1*, *SRM*, and *GATM* have been implicated in discrete aspects of CRC biology; however, their concerted actions in CRLM remain unknown. *CXCL1* and *SERPINE1* are canonical SASP factors that sustain neutrophil- and MDSC-rich niches in the metastatic liver and were associated with focal hypomethylation, echoing prior reports of NF-κB- and TGF-β-driven chemokine induction during oncogene-induced senescence (*6*). *NDRG1*, traditionally viewed as a hypoxia-responsive suppressor, was demethylated and overexpressed in metastases, consistent with recent data linking its antioxidant function to oxaliplatin resistance (*17*). The metabolic enzymes *SRM* and *GATM*, although seldom studied in CRC, control polyamine and creatine–

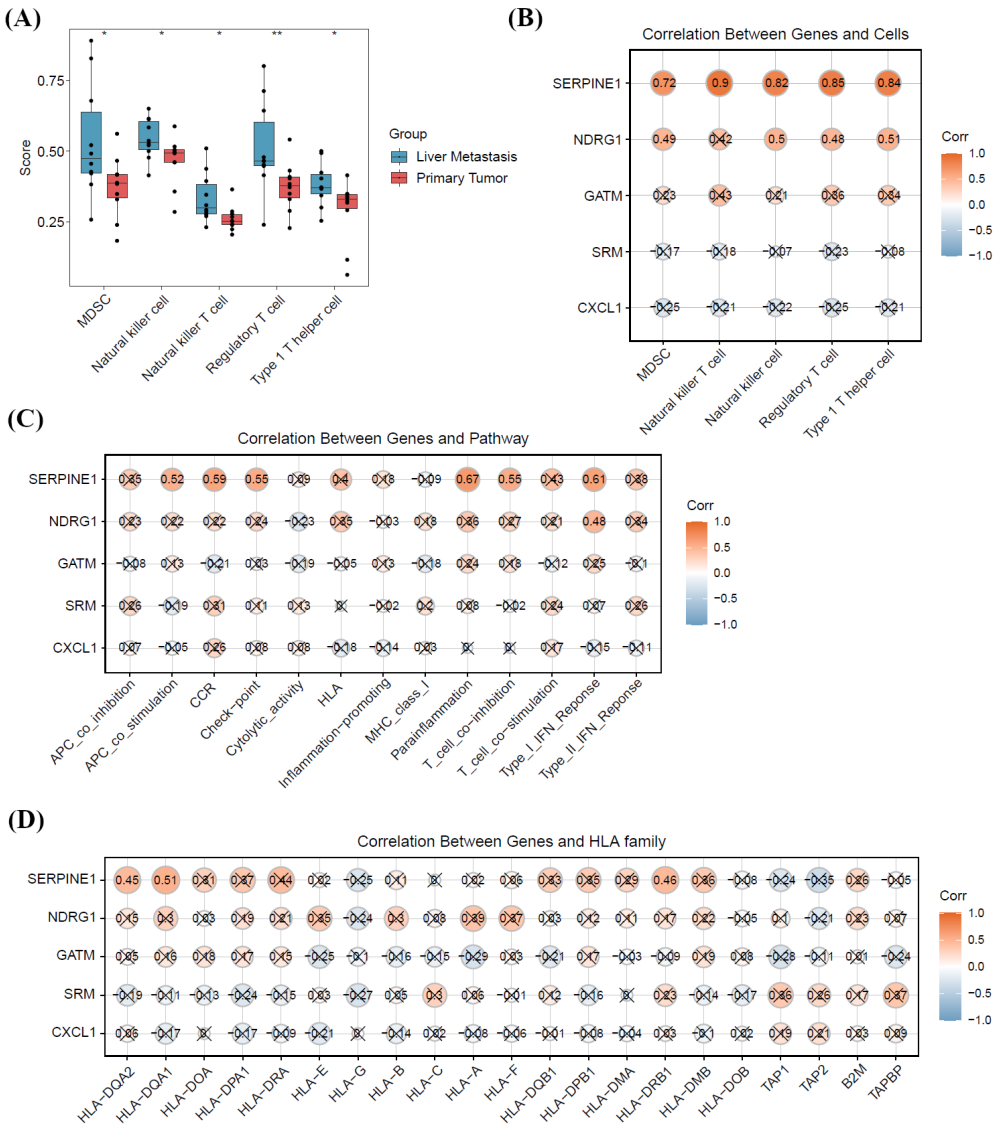

Figure 6. Immune landscape linked to the five-gene senescence–metabolism signature in CRLM. (A) Box-plots comparing the infiltration scores of 28 immune cell types between liver metastasis and primary tumour samples (GSE213402). Only significantly altered populations are shown (Wilcoxon test, FDR < 0.05). MDSC, myeloid-derived suppressor cell; NK, natural killer; Treg, regulatory T cell; Th1, type-1 T helper cell. **(B)** Spearman correlation heat-map between the five signature genes and the differentially infiltrated immune cells. Colour intensity reflects correlation coefficient (red, positive; blue, negative); asterisks indicate FDR < 0.05. **(C)** Correlation of signature genes with immune-related functions: antigen-presenting cell co-stimulation, chemokine receptor signaling, immune checkpoint activity, parainflammation, T-cell co-inhibition, and type I interferon response, as well as with HLA gene expression. Circle size and colour scale represent absolute correlation coefficient; only $|r| > 0.3$ and FDR < 0.05 are plotted. **(D)** Spearman correlation matrix between the five signature genes and HLA-family genes in the GSE213402 cohort. Colour intensity indicates correlation coefficient; asterisks denote FDR < 0.05.

phosphocreatine flux, respectively, and their elevated expression aligns with the dependence of disseminated tumor cells on de novo polyamine synthesis and mitochondrial ATP buffering (*18,19*). Collectively, these five genes appear to link the two hallmarks of CRLM — senescence bypass and metabolic reprogramming — and provide readily testable biomarkers for therapeutic stratification.

The GSEA profile of our five-gene signature was dominated by four functional modules: complement and coagulation cascades, focal adhesion/ECM-receptor interactions, OXPHOS, and PPAR signaling, each of which has been independently implicated in CRLM (*11,14*). Complement–coagulation axis activation is one of the most recurrent signatures in prior transcriptomic surveys and has recently been validated functionally; deletion of factor B of the alternative complement pathway reduced liver tumor burden by >60% in a syngeneic CRLM model (*20*). The strong enrichment of *SERPINE1* and *CXCL1* within this cascade is consistent with the hypothesis that SASP factors may contribute to attracting myeloid cells and, potentially, to propagating a fibrin-rich metastatic niche *via* complement amplification (*6,21,22*). Focal-adhesion/ECM-receptor genes were

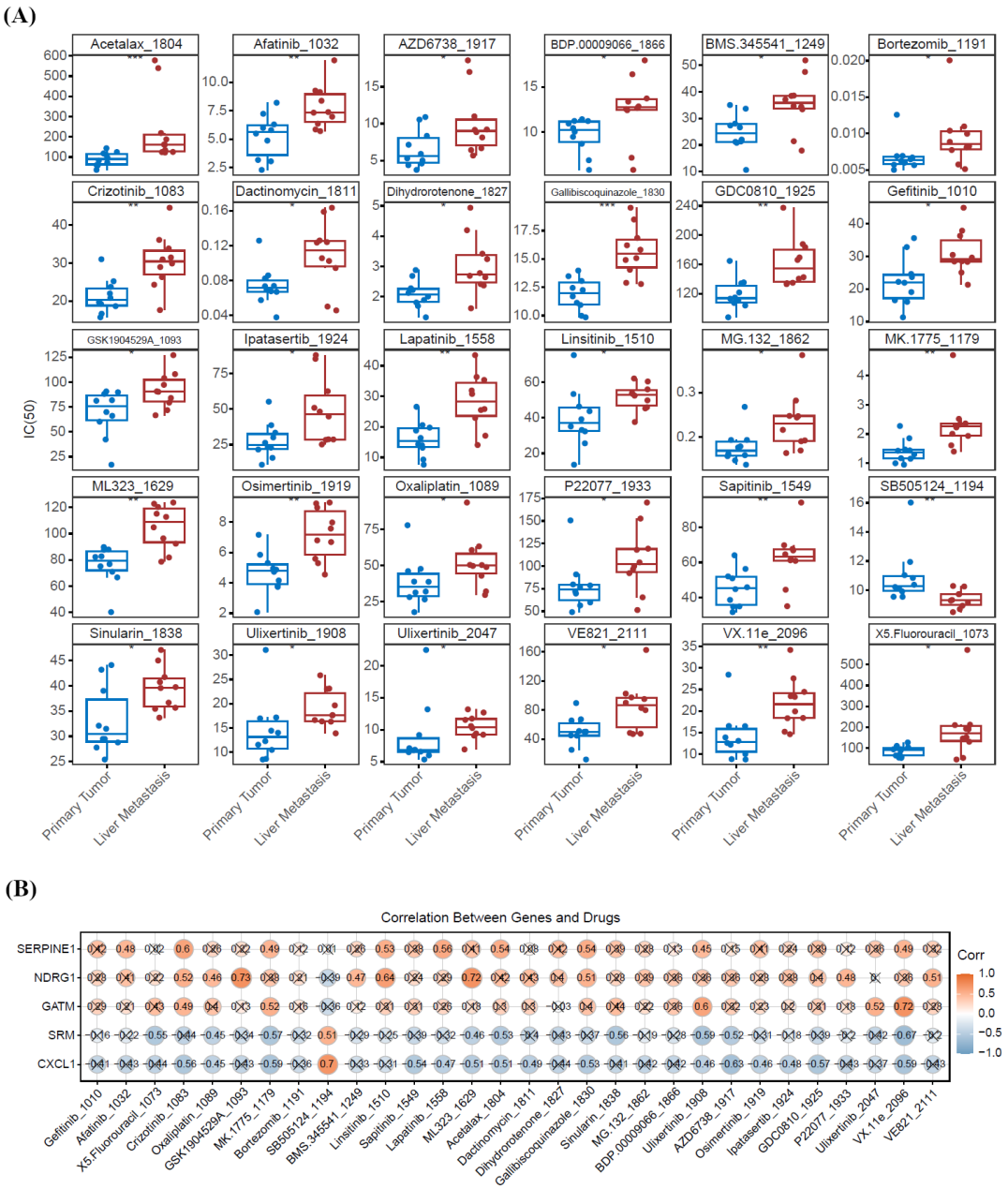

Figure 7. Chemotherapy sensitivity linked to the five-gene senescence–metabolism signature in CRLM. (A) Box-plots comparing predicted IC_{50} values of commonly used chemotherapeutic agents between primary tumour and liver-metastasis samples (GSE213402). IC_{50} values were calculated with the oncoPredict algorithm; central line indicates median, whiskers 1.5 × IQR. *P* values, paired Wilcoxon test. **(B)** Spearman correlations between the five signature genes and IC_{50} values of commonly used chemotherapeutic agents. Significant negative correlations with *SRM* and positive correlations with *NDRG1* are observed ($|r| > 0.4$, FDR < 0.05). Negative correlation indicates increased drug sensitivity.

previously identified as hub nodes in a 321-gene CRLM network and correlated with poor survival (*23*). Our observation that *NDRG1* and *SERPINE1* are enriched within the ECM-receptor signature raises the hypothesis that these cancer-cell-intrinsic programs may contribute to ECM remodeling during intrahepatic colonization (*17,24*). Finally, the co-enrichment of OXPHOS and PPAR signaling mirrors single-cell data showing that both tumor cells and lipid-associated TAMs upregulate mitochondrial respiration and PPARγ activity in CRLM (*25*). The positioning of *SRM* and *GATM* within these metabolic gene sets suggests a potential — yet still unproven — mechanistic connection between epigenetic senescence bypass and bioenergetic adaptation (*26*), offering testable targets for metabolic–immune combination therapy in CRLM (*9,27*).

Our comparative profiling of liver metastases versus primary tumors revealed a coherent, stepwise reprogramming that appears to converge across the transcriptional, immune, and pharmacological layers, yielding immediately testable clinical hypotheses. First, the five-gene senescence–metabolic hub (*CXCL1*,

SERPINE1, *NDRG1*, *SRM*, and *GATM*) is consistently upregulated in metastases and is driven by locus-specific hypomethylation rather than by copy-number gain, implying that epigenetic therapy (*e.g.*, low-dose DNMT inhibitors) or biological neutralization of *CXCL1*/PAI-1 represents a testable — yet still hypothetical — strategy to reverse the metastatic phenotype attributed to the five-gene hub (*28*). Notably, CXCL1 and SRM were downregulated in liver metastases compared to primary tumors (GSE213402), yet upregulated in primary tumors compared to normal mucosa (TCGA). This apparent contradiction may reflect stage-specific roles of these genes during tumor evolution — early activation during tumorigenesis, followed by transcriptional suppression in the metastatic niche, potentially driven by microenvironmental cues or epigenetic reprogramming. Second, the parallel enrichment of MDSCs and Tregs, together with the altered abundance of NK cells and the downregulation of HLA-DQA1/DRB1, indicates that immune evasion may be orchestrated by the same hub genes; pre-clinical models are needed to examine whether targeting the polyamine–creatine axis (*e.g.*, DFMO or *GATM* inhibitors) can impair tumor growth, enhance antigen presentation, and potentially synergize with PD-1/LAG-3 blockade (*29*). Finally, the inverse correlation between *SRM* expression and 5-FU/oxaliplatin IC_{50}, in contrast to the positive correlation with *NDRG1*, generates the hypothesis that these genes might inform chemotherapy selection; prospective clinical validation is essential before any patient-triage application(*30*). Collectively, our bioinformatic analyses generate the hypothesis that the five-gene signature could guide decision-making for combined metabolic-immune chemotherapy in CRLM; functional and clinical studies are now required to confirm its utility.

In summary, our five-gene signature integrates senescence bypass, metabolic reprogramming, and immune evasion in CRLM, and provides potential biomarkers for patient stratification and combination therapy. Due to the limited sample size of the GSE213402 cohort (n = 10 pairs), our findings are exploratory and require validation in larger independent cohorts. Machine-learning-derived signatures may be prone to overfitting at this scale. Functional assays and prospective clinical validation are warranted to confirm the causal roles of the five-gene signature in CRLM progression and therapy response.

Acknowledgements

We thank Dr. Mengyu Liao for her valuable assistance with R programming in this study.

Funding: This work was supported by the National Natural Science Foundation of China (grant number 82272836).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Shin AE, Giancotti FG, Rustgi AK. Metastatic colorectal cancer: mechanisms and emerging therapeutics. Trends Pharmacol Sci. 2023; 44:222-236.
2. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. 2019; 394:1467-1480.
3. de Magalhães JP. Cellular senescence in normal physiology. Science. 2024; 384:1300-1301.
4. Suryadevara V, Hudgins AD, Rajesh A, *et al*. SenNet recommendations for detecting senescent cells in different tissues. Nat Rev Mol Cell Biol. 2024; 25:1001-1023.
5. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. Science. 2008; 319:1352-1355.
6. Takasugi M, Yoshida Y, Hara E, Ohtani N. The role of cellular senescence and SASP in tumour microenvironment. FEBS J. 2023; 290:1348-1361.
7. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022; 12:31-46.
8. Elia I, Haigis MC. Metabolites and the tumour microenvironment: from cellular mechanisms to systemic metabolism. Nat Metab. 2021; 3:21-32.
9. Bergers G, Fendt SM. The metabolism of cancer cells during metastasis. Nat Rev Cancer. 2021; 21:162-180.
10. Tai R, Leng J, Li W, Wu Y, Yang J. Construction of the metabolic reprogramming-associated gene signature for clear cell renal cell carcinoma prognosis prediction. BMC Urol. 2023; 23:147.
11. Wu Y, Yang S, Ma J, *et al*. Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. Cancer Discov. 2022; 12:134-153.
12. McInnes T, Zou D, Rao DS, Munro FM, Phillips VL, McCall JL, Black MA, Reeve AE, Guilford PJ. Genome-wide methylation analysis identifies a core set of hypermethylated genes in CIMP-H colorectal cancer. BMC Cancer. 2017; 17:228.
13. Visone R, Bacalini MG, Di Franco S, *et al*. DNA methylation of shelf, shore and open sea CpG positions distinguish high microsatellite instability from low or stable microsatellite status colon cancer stem cells. Epigenomics. 2019; 11:587-604.
14. Zhou Y, Ding Y, Xu B, Fei H, Wang Z. Genetically druggable targets for MAPK-activated colorectal cancer by a two-sample mendelian randomization analysis. Sci Rep. 2025; 15:12239.
15. Maharati A, Moghbeli M. PI3K/AKT signaling pathway as a critical regulator of epithelial-mesenchymal transition in colorectal tumor cells. Cell Commun Signal. 2023; 21:201.
16. Morabito M, Thibodot P, Gigandet A, Compagnon P, Toso C, Berishvili E, Lacotte S, Peloso A. Liver Extracellular Matrix in Colorectal Liver Metastasis. Cancers (Basel). 2025; 17:953.
17. Liu W, Xing F, Iiizumi-Gairani M, Okuda H, Watabe M, Pai SK, Pandey PR, Hirota S, Kobayashi A, Mo YY, Fukuda K, Li Y, Watabe K. N-myc downstream regulated gene 1 modulates Wnt-β-catenin signalling and pleiotropically suppresses metastasis. EMBO Mol Med. 2012; 4:93-108.
18. Wu K, Yan M, Liu T, *et al*. Creatine kinase B suppresses

ferroptosis by phosphorylating GPX4 through a moonlighting function. Nat Cell Biol. 2023; 25:714-725.
19. Zhang L, Zhu Z, Yan H, *et al*. Creatine promotes cancer metastasis through activation of Smad2/3. Cell Metab. 2021; 331111-1123.e4.
20. Ajona D, Ortiz-Espinosa S, Moreno H, Lozano T, Pajares MJ, Agorreta J, Bértolo C, Lasarte JJ, Vicent S, Hoehlig K, Vater A, Lecanda F, Montuenga LM, Pio R. A Combined PD-1/C5a Blockade Synergistically Protects against Lung Cancer Growth and Metastasis. Cancer Discov. 2017; 7:694-703.
21. Tauriello DVF, Palomo-Ponce S, Stork D, *et al*. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. Nature. 2018; 554:538-543.
22. Wang S, Li J, Hong S, Wang N, Xu S, Yang B, Zheng Y, Zhang J, Pan B, Hu Y, Wang Z. Chemotherapy-elicited extracellular vesicle CXCL1 from dying cells promotes triple-negative breast cancer metastasis by activating TAM/PD-L1 signaling. J Exp Clin Cancer Res. 2024; 43:121.
23. Bao Y, Zhai J, Chen H, Wong CC, Liang C, Ding Y, Huang D, Gou H, Chen D, Pan Y, Kang W, To KF, Yu J. Targeting m6A reader YTHDF1 augments antitumour immunity and boosts anti-PD-1 efficacy in colorectal cancer. Gut. 2023; 72:1497-1509.
24. Ma J, Gao Q, Zeng S, Shen H. Knockdown of *NDRG1* promote epithelial-mesenchymal transition of colorectal cancer *via* NF-κB signaling. J Surg Oncol. 2016; 114:520-527.
25. Qiao X, Hu Z, Xiong F, Yang Y, Peng C, Wang D, Li X. Lipid metabolism reprogramming in tumor-associated macrophages and implications for therapy. Lipids Health Dis. 2023; 22:45.
26. Karadima E, Chavakis T, Alexaki VI. Arginine metabolism in myeloid cells in health and disease. Semin Immunopathol. 2025; 47:11.
27. Heinemann V, von Weikersthal LF, Decker T, *et al*. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. Lancet Oncol. 2014; 15:1065-1075.
28. Lu Z, Zou J, Li S, *et al*. Epigenetic therapy inhibits metastases by disrupting premetastatic niches. Nature. 2020; 579:284-290.
29. Lichtenegger FS, Rothe M, Schnorfeil FM, Deiser K, Krupka C, Augsberger C, Schlüter M, Neitz J, Subklewe M. Targeting LAG-3 and PD-1 to Enhance T Cell Activation by Antigen-Presenting Cells. Front Immunol. 2018; 9:385.
30. Na D, Chae J, Cho SY, *et al*. Predictive biomarkers for 5-fluorouracil and oxaliplatin-based chemotherapy in gastric cancers *via* profiling of patient-derived xenografts. Nat Commun. 2021; 12:4840.

Received October 10, 2025; Revised October 27, 2025; Accepted October 29, 2025.

**Address correspondence to:*
Jinhong Chen, Department of General Surgery, Huashan Hospital, Fudan University, 12 Urumqi Road, Shanghai, 200040, China.
E-mail: jinhongch@hotmail.com

Released online in J-STAGE as advance publication October 31, 2025.