

Computational analyses to reveal the key determinants of the high malignancy level of cholangiocarcinoma

Xuan Li^{1,4}, Aoran Liu^{2,4}, Xuechen Mu^{3,4}, Zhihang Wang^{1,4}, Jun Xiao^{1,4}, Yinwei Qu^{1,4}, Zhenyu Huang^{1,4}, Ye Zhang^{2,4}, Ying Xu⁴

¹Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, Jilin Province, China;

²The First Laboratory of Cancer Institute, The First Hospital of China Medical University, Shenyang 110001, Liaoning Province, China;

³College of Mathematics, Jilin University, Changchun 130012, Jilin Province, China;

⁴School of Medicine, Southern University of Science and Technology, Shenzhen 518055, Guangdong Province, China

ABSTRACT

Background and Objectives: Cholangiocarcinoma (CHOL) is a rare and highly aggressive cancer that originates in the bile duct; it has an average five-year survival rate of 9%, which makes it the cancer with the lowest survival rate among all 33 cancer types in the cancer genome atlas (TCGA) Program. The aim of this study is to elucidate the key determinants of the high malignancy level of CHOL through computational and cell-based experimental approaches and, particularly, to investigate how bile acids (BAs) influence CHOL's propensity to metastasize. **Methods:** Our study analyzed the transcriptomic data from 1835 tissue samples of 7 digestive system cancer types in the TCGA database and compared them with those of 330 control tissue samples. Multiple cellular and molecular factors were considered in the study, including the level of hypoxia, level of immune cell infiltration, degree of cellular dedifferentiation, and level of sialic acid (SA) accumulation on the surface of cancer cells. Using these factors, we developed a multivariable regression model for the five-year survival rate, as reported by the Surveillance, Epidemiology, and End Results (SEER) Program reports, and analyzed how BA biology influences a few of these factors and causes CHOL to have a high malignancy level. **Results:** CHOL exhibited the highest level of SA accumulation and B-cell infiltration among all cancer types studied. BAs inhibit the cell cycle progression through the receptor *GPBAR1*, thereby limiting the rate of nucleotide biosynthesis—which in turn forces the cells to increase SA biosynthesis in order to maintain the intracellular pH at a stable level—thereby driving cell migration and metastasis, as established in our previous study. **Conclusions:** BAs are the key contributors to the lowest five-year survival rate of CHOL among the seven cancer types studied here. This finding not only reveals the molecular mechanisms underlying the high malignancy level of CHOL but also provides a new potential target for the diagnosis and treatment of CHOL.

Key words: cholangiocarcinoma, bile acid, sialic acid, Fenton reaction, cancer migration

Address for Correspondence:

Ying Xu, School of Medicine, Southern University of Science and Technology, No. 1088 Xueyuan Avenue, Nanshan District, Shenzhen 518055, Guangdong Province, China.
Email: xuy9@sustech.edu.cn;
<https://orcid.org/0009-0004-8767-7635>

Ye Zhang, The First Laboratory of Cancer Institute, The First Hospital of China Medical University, No. 155 North Nanjing Street, Heping District, Shenyang 110001, Liaoning Province, China. Email: yzhang21@cmu.edu.cn;
<https://orcid.org/0000-0003-3495-5596>

Zhenyu Huang, Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, No. 2699 Qianjin Street, Changchun 130012, Jilin Province, China.
Email: zhenyuh19@mails.jlu.edu.cn;
<https://orcid.org/0009-0009-0029-1729>

Access this article online

Website:

www.intern-med.com

DOI:

10.1515/jtim-2024-0033

Open Access. © 2024 The author(s), published by De Gruyter on behalf of Scholar Media Publishing.
This work is licensed under the Creative Commons Attribution 4.0 International License.

INTRODUCTION

Cholangiocarcinoma (CHOL) is a heterogeneous group of rare malignant cancers that occur in bile duct epithelial cells. It is the second largest subtype of liver cancer after hepatocellular carcinoma, accounting for approximately 15% of

all primary liver cancers and 3% of all gastrointestinal cancers.^[1] In the past, this cancer has been misdiagnosed as liver cancer.^[2-4] The commonly used term “bile duct cancer” refers to all cancers that occur in the bile duct system. Due to its indolent, insidious, highly invasive, and drug-resistant characteristics,^[5] patients

are typically diagnosed when the disease is already in an advanced stage, which limits the treatment options and leads to poor prognosis.^[5-7] Although researchers have made progress in understanding the biology, diagnosis, and treatment of CHOL, the prognosis of CHOL patients has not considerably improved in recent decades, with an average five-year survival rate of 9% and a high recurrence rate after tumor resection.^[5,8-14] Statistics on CHOL reveal that in 2017, there were 210,878 new cases and 173,974 deaths worldwide.^[15] While it has been well established that CHOL is highly malignant, the factors that lead to the high malignancy of CHOL remain largely unknown. Here we present a computational study of the transcriptomic data on the cancer with the aim of elucidating the main causes of the high malignancy of CHOL, which could potentially help with early diagnosis and improved treatment of the cancer.

To demonstrate that the cancer is highly malignant, we conduct a comparative study of CHOL *vs.* cancers in other digestive organs—that is, the colon, esophagus, liver, pancreas, rectum, and stomach. We did not consider the small intestine, as it rarely develops cancer.^[16] The scientific question we address is what makes CHOL the most malignant cancer compared to other cancers of the digestive system based on the five-year survival rates in Surveillance, Epidemiology, and End Results (SEER) reports (see Table 1).^[17,18] While pancreatic cancer (PAAD) has been considered the king of malignant cancer, CHOL has a lower five-year survival rate than PAAD.

Previous studies on possible contributors to the malignancy level of cancer include hypoxia,^[19,20] immune cell infiltration,^[21] and mutation burden.^[22] In this study, we have examined the following microenvironmental conditions in cancer tissues: (i) level of hypoxia; (ii) infiltration levels by neutrophils, macrophages, and B cells; (iii) level of dedifferentiation; and (iv) level of sialic acid (SA) accumulation on the cancer cell surface. The following is the rationale for considering these factors: (a) Hypoxia has long been known to be associated with poor prognosis for cancer.^[23] Hypoxia is generally the result of high levels of H_2O_2 and $\cdot O_2^-$ produced by local neutrophils and macrophages via consumption, which tends to be associated with more serious tissue damage and prolonged inflammation;^[24-26] (b) B cells in cancer tissues have been known to negatively regulate antitumor immunity and, hence, dampen antitumor responses.^[27,28] Therefore, a higher infiltration level of B cells tends to be associated with a poorer prognosis;^[29-32] (c) Dedifferentiation is a key mechanism utilized by differentiated cells for survival under persistent and extreme stressors by gaining more powerful capabilities that are available only to their less differentiated progenitor cells.^[33] Published studies have revealed that

Table 1: Five-year survival rates of seven cancer types of the digestive system^[17,18]

Cancer type	Five-year survival rates			
	Localized	Regional	Distant	Combined
CHOL	23%	9%	3%	9%
COAD	91%	72%	13%	63%
ESCA	47%	26%	6%	21%
LIHC	36%	13%	3%	21%
PAAD	44%	15%	3%	12%
READ	90%	74%	17%	68%
STAD	72%	33%	6%	33%

Note: Localized, regional, distant, and combined refer to cancers that are entirely localized, have invaded nearby regions, are distantly metastasized, and all three types combined, respectively. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer.

dedifferentiation has been widely used across different cancer types,^[34] and that (d) it was discovered in the 1960s that cancers with higher levels of SA biosynthesis tend to be more prone to metastasis,^[35,36] and later studies on SA *vs.* cancer metastasis have been predominantly focused on the signaling roles of SA.^[37-39]

We have previously developed a model in which the increased level of SA biosynthesis gives rise to a higher level of SA accumulation on cancer cell surfaces,^[40,41] leading to increasingly stronger electrostatic repulsion among neighboring cancer cells as each SA carries a negative charge.^[42] Published studies have established that mechanical compression, as in cell-cell repulsion, can lead to morphological changes that activate the epithelial-mesenchymal transition (EMT) program and lead to the migration of the affected cells,^[41] which is the first step in cancer metastasis.

We have estimated the level of each of the factors outlined in points (i) - (iv) based on the expression levels of the relevant marker genes for each cancer tissue across all seven cancer types and demonstrated that the five-year survival rate of each cancer type could be well represented as a function of the expression levels of these factors through a regression analysis, with high statistical significance. Key steps that lead to cell migration in CHOL are experimentally validated.

MATERIALS AND METHODS

Data

We have conducted comparative analyses of transcriptomic data of 1835 cancer tissue samples of 7 cancer types *vs.* those of 330 control tissues, all from the cancer genome atlas (TCGA) database *via* the genomic data commons

Table 2: The numbers of the samples for the seven cancer types used in this study

Cancer type	#Tumor samples	#Control samples
CHOL	35	9
COAD	483	41
ESCA	185	13
LIHC	374	50
PAAD	179	4
READ	167	177
STAD	412	36

CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer.

(GDC) data portal (<https://portal.gdc.cancer.gov>).^[43] The detailed information of the datasets is given in Table 2.

For each of the following factors analyzed, we have used the expression levels of the marker genes as widely in the literature (see Supplementary Tables S1 and S2): (a) the level of hypoxia;^[44,45] (b) the level of dedifferentiation;^[46] and (c) the levels of SA biosynthesis and degradation, nucleotide *de novo* biosynthesis, and Fenton reactions in each of the seven cancer types.^[46]

De-batch effect

We have applied an empirical Bayesian method to estimate the batch effect by modeling the relationship between gene expression levels and batch variables.^[47–49] The estimated batch effect is then subtracted from the raw data^[50] to ensure that transcriptomic data collected separately can be compared directly with each other.

Single sample gene set enrichment analysis (ssGSEA) for pathway enrichment of individual samples

To perform gene set enrichment analysis (GSEA) on a single sample, we used the “*gsva()*” function in R,^[51] with the analysis method set to “*ssgsea*”, which is specifically designed for GSEA on individual samples.^[52] The method is a non-parametric one that uses the empirical cumulative distribution function of the ranks of gene expressions within and outside a specified gene set to calculate the enrichment score. This method has been shown to be highly effective in identifying gene sets or pathways enriched by the given genes.^[53,54]

Estimation of SA accumulation on cancer cell surface

It is known that surface O-glycosylated proteins are involved in cancer metastasis.^[55,56] Hence, we have estimated the level of SA accumulation onto O-glycosylated

proteins *via* sialyltransferases—that is, *ST6GALNAC1*, *ST6GALNAC2*, *ST6GALNAC3*, *ST6GALNAC4*, and *ST48SLA6*.^[57] It must be noted that only *ST6GALNAC1*, *ST6GALNAC2*, and *ST6GALNAC4* are upregulated in digestive cancers; hence, these are the only ones considered in our analyses. In addition, *NEU3* is the main SA degrader. We have used the following equation to estimate the average level of SA accumulation in stage-*i* samples for a target cancer type:

$$\Delta SA'_i = \frac{\sum_{G_1} Ave_i(G_1, cancersamples) - Ave_i(G_2, cancersamples)}{\sum_{G_1} Ave_i(G_1, controlsamples) - Ave_i(G_2, controlsamples)}, \quad (1)$$

Where $G_1 = \{ST6GALNAC1, ST6GALNAC2, ST6GALNAC4\}$, $G_2 = NEU3$, and $Ave_i(G, S)$ represents the expression of *G* averaged across all samples *S* in stage *i*. Further, the estimated level of SA accumulation up to stage *I* is expressed in the following manner:

$$\Delta SA(I) = \sum_{i \leq I} \Delta SA'_i, \quad (2)$$

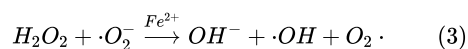
where *I* = 1 for localized tumor samples, 2 for regional tumor samples, and 3 for distant tumor samples, assuming that the duration of each of the three stages is approximately the same.

xCell for estimating the infiltration levels by immune cells

xCell is a widely-used computational tool for calculating an enrichment score by the expressions of marker genes for each of the 64 immune and stroma cell types in a given cancer tissue sample.^[58] Using the calculated enrichment scores, we have estimated the infiltration levels of B cells, neutrophils, and macrophages in each sample by assessing each enrichment score against a total enrichment score for the tissue sample.

Mechanisms framework

We have also considered bile acids (BAs) in our analysis. BAs can activate the G-protein-coupled BA receptor *GPBAR1*, which is known to suppress cell proliferation, particularly in cancer cells.^[59,60] This has an important implication in our analyses, which follow a specific chain of inference, described hereafter. (1) All cancer tissues in TCGA harbor the following persistent Fenton reaction:



This reaction is harbored in the cell cytosols due to chronic inflammation (chemically, with elevated levels of H_2O_2 and $\cdot O_2^-$), coupled with local iron accumulation, which continuously produces OH^- .^[61] (2) The continuous alkalization by the persistent Fenton reaction casts a life-threatening stress on the affected cells, thus driving the cells to start multiple acidifying reprogrammed metabolisms

to keep the pH stable.^[61,62] (3) The two predominant reprogrammed metabolisms are the *de novo* biosynthesis of nucleotides (NTs) and biosynthesis and deployment of SAs by all cancers in TCGA, and together they have the following:

$$R(OH^-, FR) = R(H^+, NT) + R(H^+, SA) + \varepsilon, \quad (4)$$

Where $R(OH^-, FR)$ represents the rate of OH^- production by Fenton reactions in cancer tissue, $R(H^+, NT)$ and $R(H^+, SA)$ are the rates of H^+ production by nucleotide *de novo* biosynthesis and SA biosynthesis, respectively, and ε represents the total s produced by all other acidifying reprogrammed metabolisms, which is a relatively small quantity compared to $R(H^+, NT) + R(H^+, SA)$.

A direct implication of the above equation is that if $R(H^+, NT)$ in cancer cannot be as high as that in other cancers for any reason, $R(H^+, SA)$ must be sufficiently high so their sum matches a major fraction of $R(OH^-, FR)$ to keep the pH stable. Our previous study has provided strong evidence that the increased SA biosynthesis and deployment drives cancer metastasis.^[41,63] The above analysis serves as the basis for our study in the results section.

Principal component analysis (PCA) for gene-expression analyses

PCA is used to capture the main direction of change in a multi-marker gene set in terms of the changes in their collective expressions.^[64] Here, we use PCA analysis on the marker genes involved in the dedifferentiation process to quantify the level of dedifferentiation.

Bayesian information criterion (BIC) for contribution estimation

The BIC is used to select the best model among the given options,^[65] it is defined in the following manner:

$$BIC = k \times \ln(n) - 2\ln(L), \quad (5)$$

Where L is the maximum likelihood function of the model, n is the sample size, and k is the number of free parameters in the model.

The level of contribution to a model ($model_k$) by the i^{th} free parameter is estimated using $|BIC_k| - |BIC_i|$, where BIC_k is the quality of the optimal $model_k$ that is measured using the R^2 score and BIC_i is the best model based on all parameters except for the i^{th} parameter. Hence, the percentage of contribution by the i^{th} parameter can be calculated in the following manner:

$$Contribution_i = \frac{|BIC_k| - |BIC_i|}{\sum_{(i \leq K)} |BIC_k| - |BIC_i|} \quad (6)$$

Multiple linear regression (MLR)

MLR is used to model the linear relationship between a single dependent variable and multiple free variables. We used BIC-based MLR to perform a linear regression on the five-year survival rate against a list of possible contributors in each sample across the seven cancer types, where BIC provides information regarding the level of contribution to the regression result by each free variable.

Cell culture

RBE and HuCC-T1 cells were obtained from Procell and cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS, Moregate) and 1% pen-strep antibiotic (Procell). All cells were cultured in a humidified incubator at 37°C with 5% . First, we performed STR testing on the two cell lines to prevent misidentification or cross-contamination of cell lines. STR testing uses primers to extract repeated DNA fragments, usually two to six base pairs of tandem repeats. A DNA kit is used to extract high-purity DNA; thereafter, multiple rounds of PCR amplification are performed—that is, multiple gene loci are amplified. Finally, sequencing is performed to determine the experimental conditions.

Generating persistent Fenton reactions in cytosol in cultured cells

$FeSO_4$ (Sigma, F8633), H_2O_2 (Aladdin, 7722-84-1), and L-ascorbic acid (Sigma, A4544) were added to the culture to induce sustained Fenton reactions. Here we used L-ascorbic acid instead of $\cdot O_2^-$ as the reducing molecule for the persistent Fenton reaction.

Cell proliferation and colony formation assays

Cell Counting Kit 8 (CCK-8, Apexbio) assays were performed to assess cell proliferation in accordance with the manufacturer's instructions; 5000 cells were seeded in 96-well plates in triplicate containing regular full-serum media and allowed to adhere overnight. Different concentrations of lithocholic acid (LCD, MCE) were added the next day, and cells were harvested at the indicated times. Further, cells were incubated with 10% CCK-8 solution at 37°C for two hours, and then the absorbance was measured at 450 nm using a microplate reader.

For the colony formation assay, 600-800 cells were seeded in six-well plates in triplicate and adhered overnight. After 15 days, cells treated with LCD were fixed with polyformaldehyde (4%), washed twice with phosphate buffer saline (PBS), and stained with crystal violet solution (1%) for 15 min at room temperature.

SA assay

The culture medium of cancer cells was collected as the acidic medium. SA was quantified using a Sialic Acid Assay

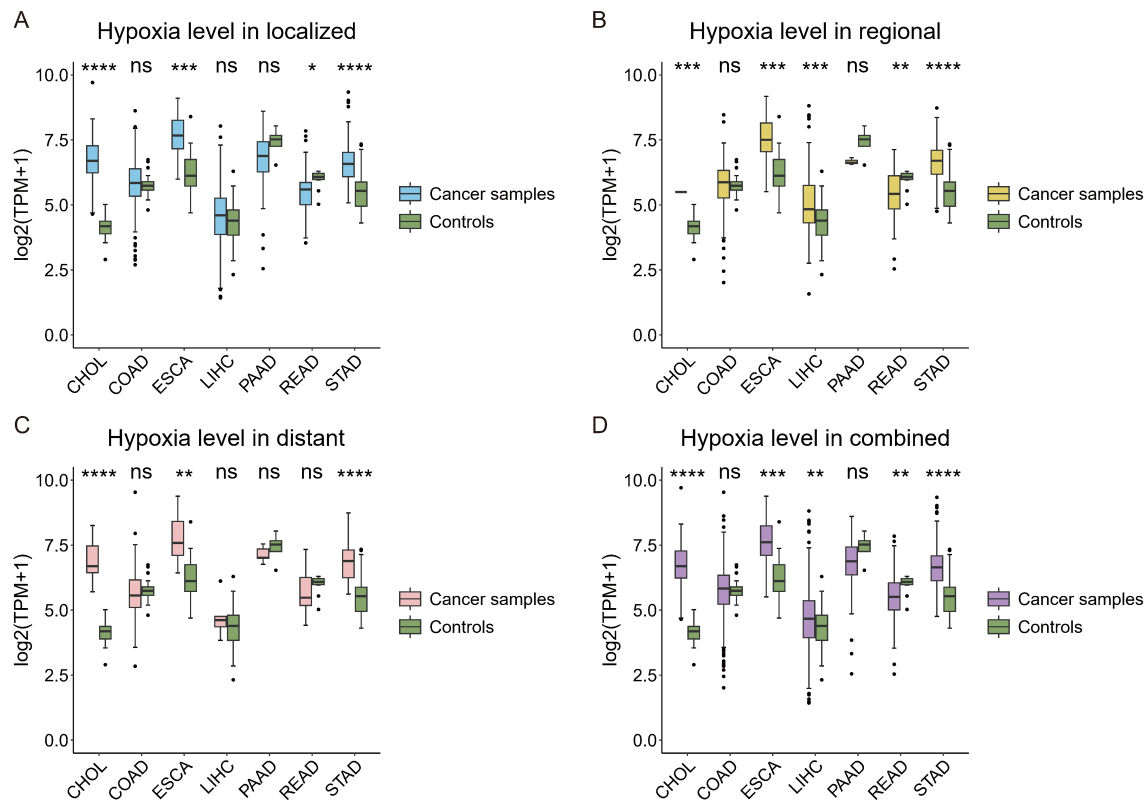


Figure 1: Hypoxic levels of cancer vs. control samples in localized (A), regionally migrated (B), distantly metastasized (C), and all cancer samples combined (D), where the expression of *HIF1A* is used to reflect the hypoxia level for cancer vs. control samples at each cancer stage across seven cancer types. The y-axis in each panel represents the level of log-transformed expression, measured using TPM. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer.

kit (Abcam, ab83375) with the colorimetric assay protocol in accordance with the manufacturer's instructions.

Statistical analysis

Experiments were performed using a minimum of three replicates for each experimental group. For all statistical analyses, GraphPad Prism software (version 8.0) was used. Unpaired or paired Student *t* tests were used to compare data between the two groups. Statistical significance was defined as $P < 0.05$ (* $P < 0.05$; ** $P < 0.001$; **** $P < 0.0001$). All observations were expressed as mean $\pm$ SEM.

RESULTS

CHOL is among the most hypoxic cancer types vs. controls

We used the expression level of *HIF1A* to reflect the level of hypoxia in a tissue sample, with a higher *HIF1A* expression indicating a higher level of hypoxia.^[45] Figure 1 depicts the expression levels of *HIF1A* in samples in localized, regionally migrated, distantly metastasized, combined cancer samples, and control samples for each of the seven cancer types, respectively. It is evident from the figure that CHOL has the largest increase in the expression

of *HIF1A* in cancer *vs.* control samples across all stages, thereby indicating that CHOL is the most hypoxic cancer type among the seven.

CHOL has the highest level of B-cell infiltration

We calculated the infiltration levels of B cells, macrophages, and neutrophils in each tissue sample using the xCell program (see materials and methods) in the control, localized, regionally migrated, distantly metastasized, and combined cancer samples for each of the seven cancer types. Figure 2 presents the calculation results, from which it is evident that CHOL has the highest level of B-cell infiltration *vs.* the controls across all seven cancer types on average.

To understand why CHOL has the highest level of B-cell concentration, a pathway enrichment analysis was undertaken to identify genes (Table S1, supplementary materials), whose expressions are highly correlated with genes involved in SA biosynthesis and deployment. Interestingly, B-cell activation is among the most correlated activities (Figure 2D). Therefore, we posit that the higher levels of SAs attract more B cells into the area, which is consistent with the results of published studies (SAs are

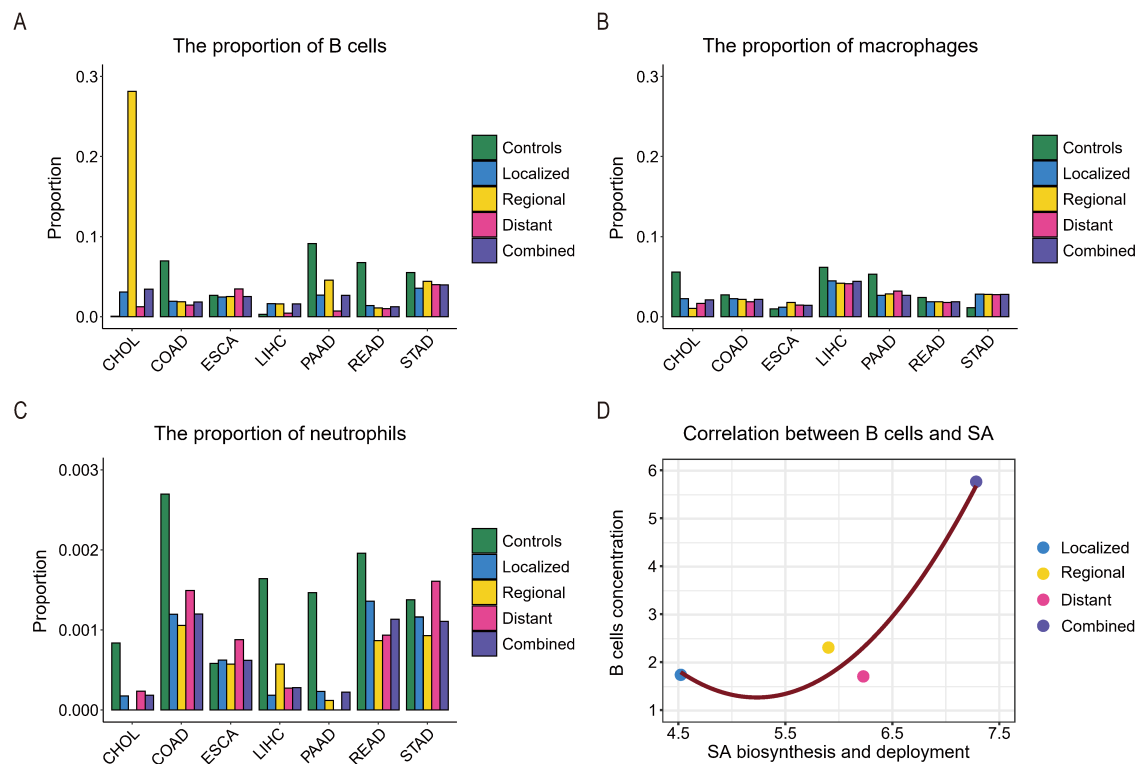


Figure 2: The estimated infiltration levels of (A) B cells, (B) macrophages, and (C) neutrophils in the control, localized tumors, tumors with regional migration, tumors with distant metastasis, or combined cancer samples for each of the seven cancer types, where the y-axis represents the percentage of the cells under study in a tissue sample. (D) The relationship between the infiltration level of B cells (the y-axis) and the level of SA biosynthesis and deployment (the x-axis), measured using the first principal component of the expressions of the marker genes of SA biosynthesis and deployment across the relevant samples, and the two have a correlation coefficient 0.823, given in Table S1. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer; SA: sialic acid.

the capping molecules of cell-surface glycans, which are part of gangliosides).^[66-69]

CHOL is among the most dedifferentiated cancers

Using the marker genes for the level of cell dedifferentiation given in Table S1, we estimated the level of dedifferentiation *via* a PCA (see materials and methods). Figure 3 depicts the estimated level of cell dedifferentiation for each of the seven cancer types. It is evident from the figure that CHOL is among the cancer types that have the highest level of dedifferentiation, suggesting that the cancer is, in general, more stressed than other cancer types.

CHOL has the highest level of SA accumulation on cancer cell surface

As discussed earlier, cancer cells utilize *de novo* biosynthesis of NTs and biosynthesis of SAs as the main H^+ producers to neutralize OH^- produced by Fenton reactions.^[61,70]

Figure 4A and 4B depicts the level of contribution of SA towards keeping the Fenton reaction-produced OH^- neutralized across the seven cancer types. We noted that CHOL stands out in terms of the level of contribution

by SA biosynthesis. To understand why SA contributes particularly highly to keeping the cytosolic pH stable, we studied the roles that BAs may have played in this picture. First, we noted that (1) the level of BA, measured using the expression data of the relevant marker genes (Table S1), strongly correlates with the level of SA biosynthesis as well as the level of SA accumulation; and (2) the level of BA negatively correlates with nucleotide *de novo* biosynthesis, as depicted in Figure 4C, suggesting that BA might play a role in suppressing nucleotide synthesis and hence driving up the SA biosynthesis and deployment in accordance with equation (4).

To elucidate the possible connection between the two, we examined the expression level of the BA receptor *GPBAR1* and found that it is upregulated, which is consistent with the increased BA level in CHOL. Our literature search revealed that *GPBAR1* can suppress cell proliferation by activating the cAMP signaling pathway, which is known to be associated with cancer bone metastasis.^[71-73] To examine if this is the case in CHOL, we note that the expression of *GPBAR1* indeed negatively correlates with cell-cycle related pathways (Figure 4D, Table S1), and with the level

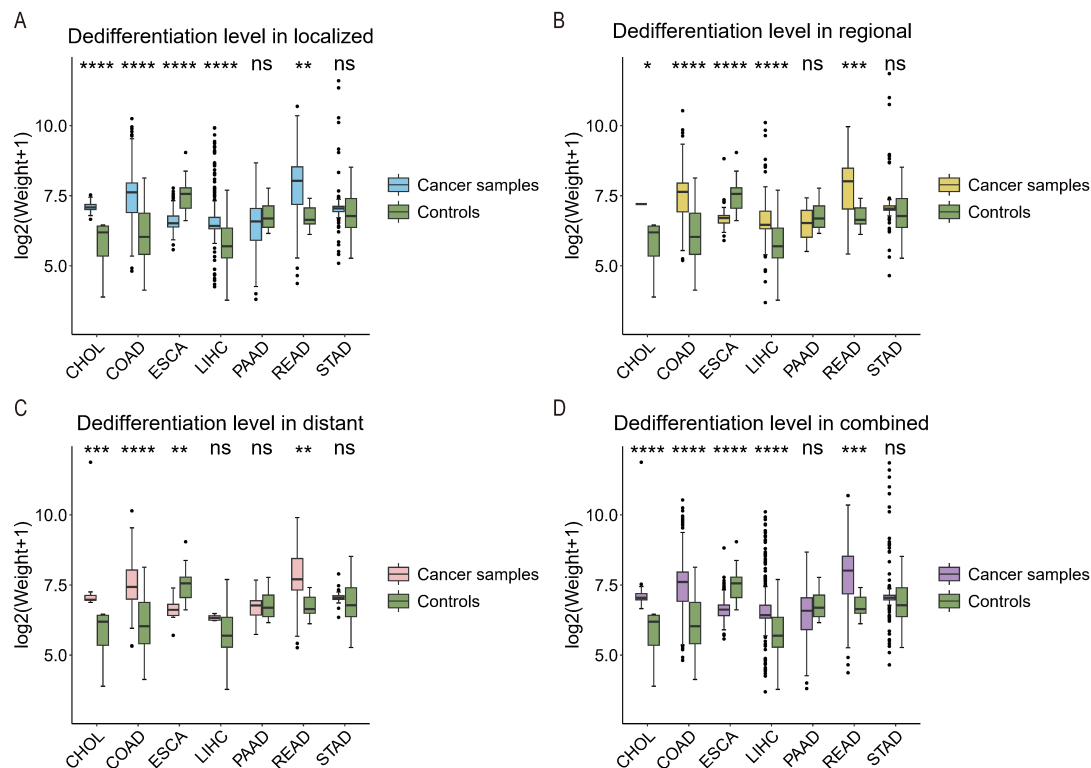


Figure 3: The level of dedifferentiation (the y-axis) is compared between cancer and control samples across the seven cancers in (A) totally localized tumor, (B) tumor with regional migration, (C) tumor with distant metastasis, and (D) combined cancer samples, where weight is the first principal component of the expressions of marker genes for dedifferentiation, where the gene list is given in Table S1. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer.

of cell-cycle progression (Figure 4E), thereby confirming that *GPBAR1* suppresses cell proliferation in CHOL, as reported in the literature, which further limits the rate of nucleotide biosynthesis for DNA synthesis.

As noted earlier, nucleotide *de novo* biosynthesis and SA biosynthesis play the predominating roles in neutralizing the OH[•]s produced by Fenton reactions. The data depicted in Figure 4F explain why the SA biosynthesis level in CHOL stands out among the seven cancer types—that is, the nucleotide biosynthesis is suppressed as a result of suppressed cell proliferation by BAs *via* their binding with *GPBAR1*. Hence, it is the BA that ultimately causes the levels of SA biosynthesis and accumulation to go higher, thereby leading to the high migration rate by CHOL and, hence, the lowest survival rate among the seven cancer types. This computational prediction, the core of our model, is validated experimentally in the last section of results.

Key factors contributing to the low survival rates of CHOL

We now aim to establish a quantitative relationship between the five-year survival rate and the possible contributing factors discussed earlier. To do this, we first predicted the

possible causal relationships among these factors to identify and remove those factors that are largely determined by the other factors. Specifically, we sorted all the cancer samples of each cancer type in ascending order of the number of upregulated genes in a sample *vs.* controls, as a means to provide a pseudo-time order of the disease progression. On this ordered list of cancer samples, we used the F-test statistic in the Granger causality test^[74] to assess if some pairs are causally connected with a significant *P*-value. In other words, if the *P*-value of factor *i* to factor *j* is less than or equal to 0.05, we consider that factor *i* can predict factor *j*—that is, there is a causal relationship between the two. Table 3 lists the calculation results.

It is evident from the table that the *P*-values of dedifferentiation relative to Δ SA, hypoxia, and B cells are all 0.05, indicating that there is a causal relationship between dedifferentiation and each of the three factors; hence, it was excluded from our regression analyses.

We then conducted a linear regression analysis of the five-year survival rates using the five remaining factors: Δ SA; the level of hypoxia; the infiltration levels by B cell, neutrophil, and macrophage, respectively. Figure 5 presents the four regression models for localized (model_l), regionally

Table 3: Prediction of causal relationships between factors

Item	ΔSA	Hypoxia	Dedifferentiation	B cells	Macrophages	Neutrophils
ΔSA	1	0.059	0.5362	0.2756	0.0668	0.3459
Hypoxia	0.6139	1	0.7694	0.5222	0.9296	0.3864
Dedifferentiation	0.0001	0.0123	1	0.001	0.1245	0.1167
B cells	0.4243	0.0507	0.3472	1	0.0503	0.3736
Macrophages	0.6054	0.4275	0.0089	0.8036	1	0.132
Neutrophils	0.8353	0.5174	0.0757	0.5286	0.4279	1

Table 4: Performance results of regression models

Model	R^2	Contribution based on BIC analyses				
		ΔSA (%)	hypoxia (%)	B cell (%)	macrophage (%)	neutrophil (%)
model ₁	0.7287	65.1	14.59	3.51	1.09	15.71
model ₂	0.9843	42.86	7.81	35.87	0.27	13.19
model ₃	0.9991	49.03	5.21	37.8	0.53	7.43
model ₄	0.9628	52.06	0.17	26.48	4.33	16.96

Table 5: Contributions towards keeping the pH stable by SA and nucleotide synthesis

Cancer type	Contribution level by SA (%)	Contribution level by nucleotide synthesis (%)
CHOL	70.47	29.53
COAD	14.08	85.92
ESCA	69.12	30.88
LIHC	20.73	79.27
PAAD	70.09	29.91
READ	34.14	65.86
STAD	13.05	86.95

CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer; SA: sialic acid.

migrated (model₂), distantly metastasized (model₃) and combined cancer (model₄) samples with the regression results, respectively.

$$\text{Regression model}_1: \text{SR} = 1.0158 - 0.4043\Delta\text{SA} + 0.2528\text{H} + 0.0434\text{B} + 0.004\text{M} - 0.4158\text{N} \quad (7)$$

$$\text{Regression model}_2: \text{SR} = 1.3789 - 0.5405\Delta\text{SA} + 0.3069\text{H} + 0.1469\text{B} - 0.1133\text{M} - 0.8068\text{N} \quad (8)$$

$$\text{Regression model}_3: \text{SR} = 0.3679 - 0.1086\Delta\text{SA} - 0.0948\text{H} + 0.136\text{B} + 0.0665\text{M} - 0.124\text{N} \quad (9)$$

$$\text{Regression model}_4: \text{SR} = 1.58 - 0.3908\Delta\text{SA} + 0.0104\text{H} + 0.2268\text{B} + 0.0937\text{M} - 0.6073\text{N} \quad (10)$$

In the above equations, SR is the five-year survival rate; ΔSA is the levels of SA accumulation; H is the hypoxia; and B, M, and N are the infiltration level by B cells, macrophages, and neutrophils, respectively.

Our BIC-based analyses (see materials and methods) have provided the level of contribution by each of these factors, as given in Table 4. Hence, we conclude that the SA accumulation, dictated by the BA level and the level of Fenton reaction, plays the most significant role in making CHOL the deadliest cancer among the seven that were studied. Moreover, the infiltration level of B cells is the second most significant contributor to the lowest survival rate of CHOL, which is partially because of the high SA level, as discussed earlier.

Experimental validation of Fenton reactions and cell-surface SA levels

The core logic of our predictive model is that it is the Fenton reaction that drives the high levels of nucleotide *de novo* biosynthesis and SA biosynthesis for survival and that it is the BAs in CHOL that suppress the level of cell proliferation, hence the level of nucleotide biosynthesis, which in turn drives the level of SA biosynthesis up to keep the intracellular pH stable for cells affected by Fenton reactions.

To verify this prediction, we selected CHOL cells (RBE and HuCC-T1 cells) for study. We first induced sustained cytosolic Fenton reactions in these cells by FeSO₄, H₂O₂, and L-ascorbic acid (VC) to HuCC-T1, where L-ascorbic acid serves the same purpose of superoxide to reduce Fe³⁺ back to Fe²⁺, to induce persistent Fenton reactions. After preliminary experiments, we found that 100 μmol/L FeSO₄, 100 μmol/L H₂O₂, and 1000 μmol/L L-ascorbic acid were the most suitable concentrations for inducing persistent Fenton reactions in cytosol. When FeSO₄ was added in a concentration of 100 μmol/L, H₂O₂ in a concentration

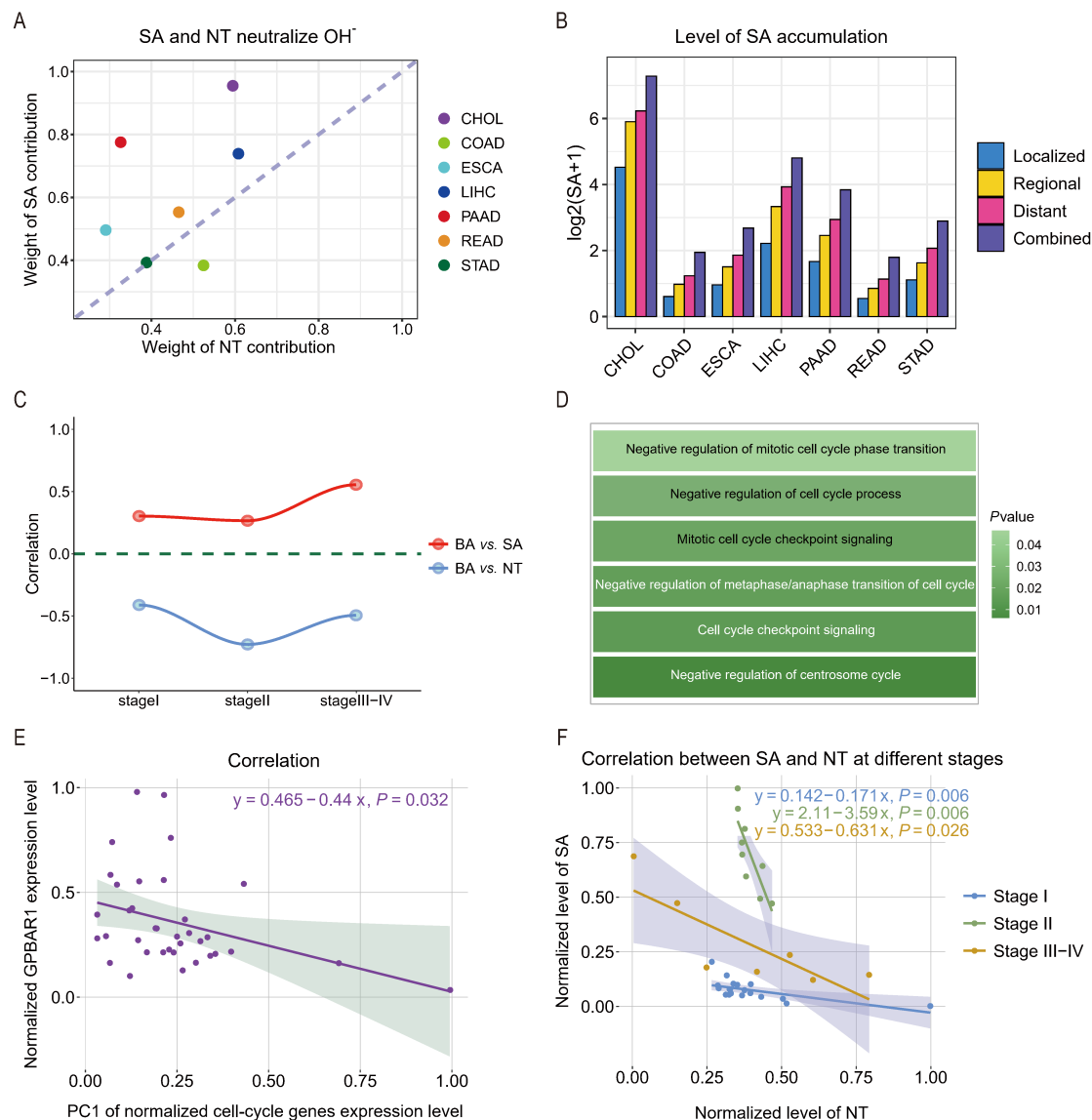


Figure 4: (A) Contributions by *de novo* nucleotide (NT) biosynthesis and SA biosynthesis and utilization in neutralizing OH^- produced by Fenton reactions, respectively, where the x-axis represents the weight of contribution by nucleotide biosynthesis and the y-axis represents the weight of contribution by SA biosynthesis and utilization, and a dot on the diagonal indicates that the two biosynthesis processes are equally weighted in the neutralization of the OH^- produced by Fenton reactions. (B) The predicted level of SA accumulation for each cancer type. (C) Correlations between BA biosynthesis and nucleotide *de novo* biosynthesis as well as between BA biosynthesis and SA biosynthesis. (D) Cell-cycle progression-related pathways enriched by genes that positively correlate with GPBAR1 in CHOL. (E) The BA's receptor *GPBAR1* negatively correlates with cell-cycle marker genes, where each dot represents a CHOL sample with its normalized *GPBAR1* expression level (x-axis) and the normalized first principal component of the expression levels of cell-cycle genes (y-axis). (F) Correlation between the *de novo* nucleotide biosynthesis and the biosynthesis and utilization of SA in different stages. The x-axis represents the (normalized) level of *de novo* nucleotide biosynthesis, and the y-axis represents the (normalized) level of SA biosynthesis, where all the relevant marker genes are given in Table S1. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer; SA: sialic acid.

of 100 $\mu\text{mol/L}$, and L-ascorbic acid in a concentration of 1000 $\mu\text{mol/L}$, significantly increased cell proliferation was observed at 24 h and 48 h compared with the control group as result of increased nucleotide biosynthesis induced by Fenton reactions, and the difference was statistically significant (Figure 6 and Table S3).

We have then examined how the proliferating cells will be

affected when treating them with the proliferation inhibitor LCD (Lithocholic acid) as in the case of BAs in CHOL. First, we note that the cell viability was the lowest when treated with the LCD with concentration at 100 $\mu\text{mol/L}$ among other levels of concentrations (Figure 7A). Then, we tested the impact of different LCD concentrations on cell proliferation and found that LCD had the strongest inhibition of cell proliferation at the same concentration of

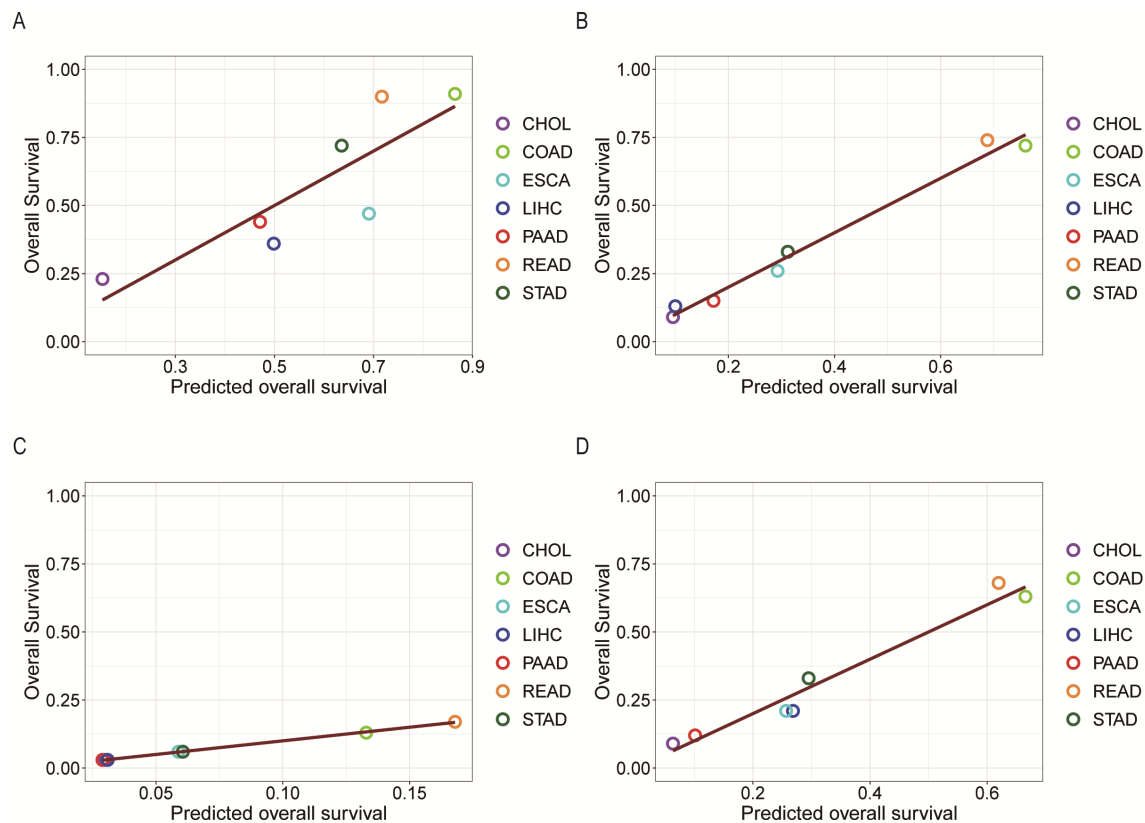


Figure 5: The five-year survival probabilities (y-axis) of seven types of cancers against the predicted survival probabilities (x-axis) based on SA accumulation, level of hypoxia, and the infiltration levels of B cells, macrophages, and neutrophils through regression analyses, with (A), (B), (C), and (D) representing localized tumors, tumors with regional migration, distant metastasis, and combined cancer sample set, respectively. CHOL: cholangiocarcinoma; COAD: colon cancer; ESCA: esophagus cancer; LIHC: liver cancer; PAAD: pancreatic cancer; READ: rectum cancer; STAD: stomach cancer.

100 $\mu\text{mol/L}$ for 48 h and 72 h (Figure 7B and 7C). Similarly, the colony formation assay found the same results (Figure 7D and 7E). Therefore, we chose the LCD concentration at 100 $\mu\text{mol/L}$ for further experiments.

We then measured the SA level on the cell-surface in cells of stomach, liver, colon, and pancreatic cancer, together with CHOL cells, separately, and found that consistent with our above results, the highest SA content was found in CHOL cells (Figure 7F). Subsequently, we treated the CHOL cells with LCD at 100 $\mu\text{mol/L}$ and observed that the cell-surface SA content continued to increase (Figure 7G) with time ($P\text{-value} = 0.0174$) (Figure 7H).

The conclusions are that (1) by inducing Fenton reactions in the cytosol of cancer cells, the cells proliferate at higher rates as predicted by our computational study and validated by our cell-based experimental study, and (2) by inhibiting experimentally the proliferation rate of such cancer cells, the cell-surface SA content increases as we have predicted, which drives cells to metastasize.^[41] This explains why CHOL is one of the most malignant cancers.

DISCUSSION

Different cancer types have intrinsically different malignancy levels such as pancreatic cancers are known to be the king of cancers and basal cell carcinoma generally does not kill people while other cancer types each tend to have a stable and distinct five-year survival rate across different patients. What determines the intrinsic malignancy level of a cancer type? An extensive literature review has revealed that very few studies have focused on such an important scientific question. Relevant papers tend to focus on specific microenvironmental factors such as the infiltration levels of innate immune cells or T cells, and the hypoxic level in a cancer tissue may be associated with the malignancy level within the same cancer type.^[75-77]

We presented a comparative study among seven cancer types in the human digestive system to explain statistically why CHOL has the lowest five-year survival rate with high statistical significance compared to the six other cancer types. The information used is derived from the tissue-based transcriptomic data of a total of 1835 cancer tissue and 330 control tissue samples. Regression analyses coupled

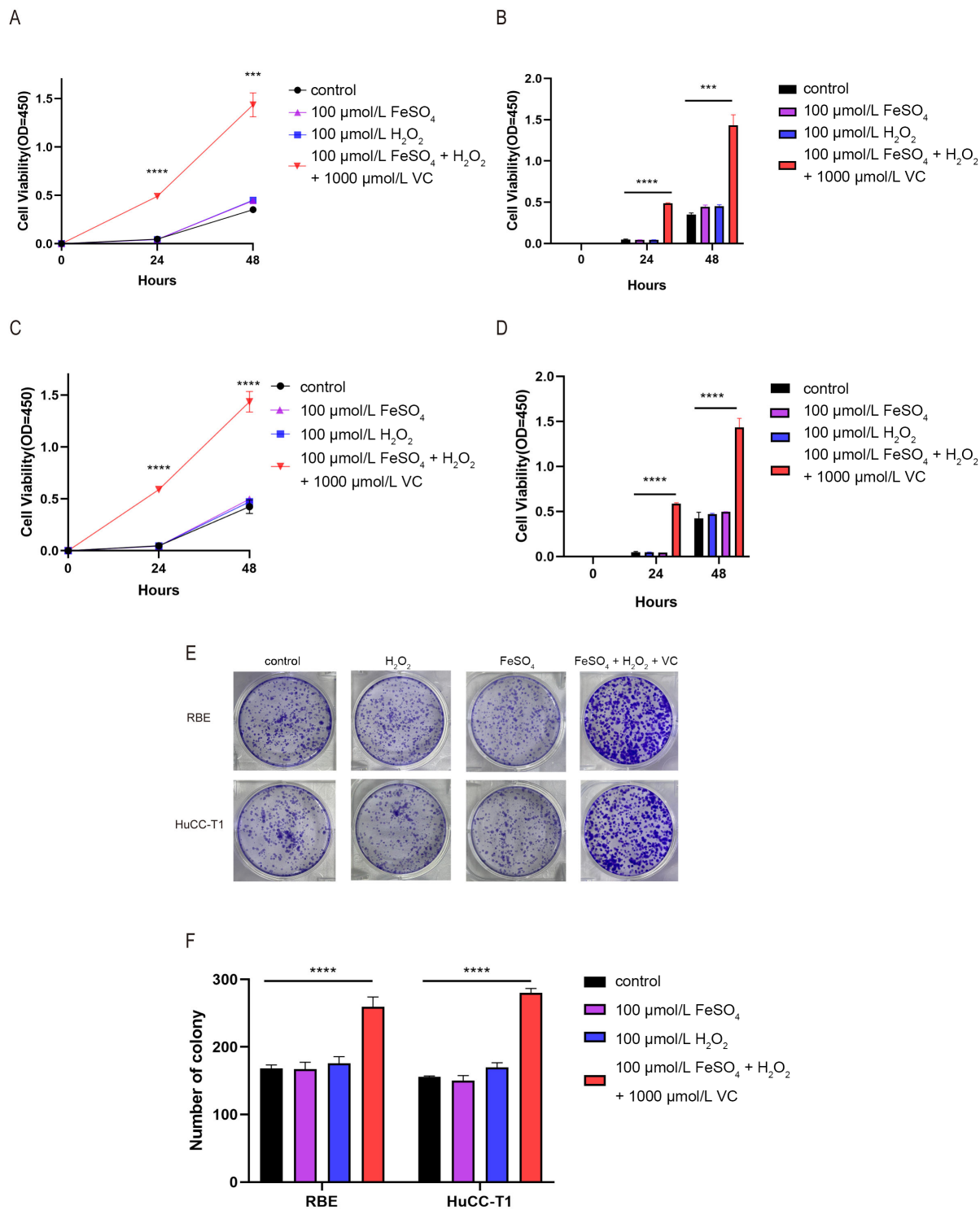


Figure 6: The sustained Fenton reaction promotes cell proliferation (Table S3). (A and B) The cell proliferation ability of the RBE cells was detected after adding FeSO_4 , H_2O_2 , and L-ascorbic acid. (C and D) The cell proliferation ability of the HuCC-T1 cells was detected after adding FeSO_4 , H_2O_2 , and L-ascorbic acid. (E) The colony formation assay was performed after adding , and L-ascorbic acid. (F) Statistical diagram of the colony formation assay (** $P < 0.0001$; *** $P = 0.0002$).**

with Bayesian statistical analyses revealed that the top two contributors to the high malignancy level of CHOL are the high level of SA biosynthesis and deployment and the high level of B-cell infiltration. Interestingly the B-cell level

is possibly due to the SA level, which is consistent with published literature.^[66-69,78,79] Hence, we conclude that it is the high SA level dictates the high malignancy level, which is consistent with published studies regarding the relationship

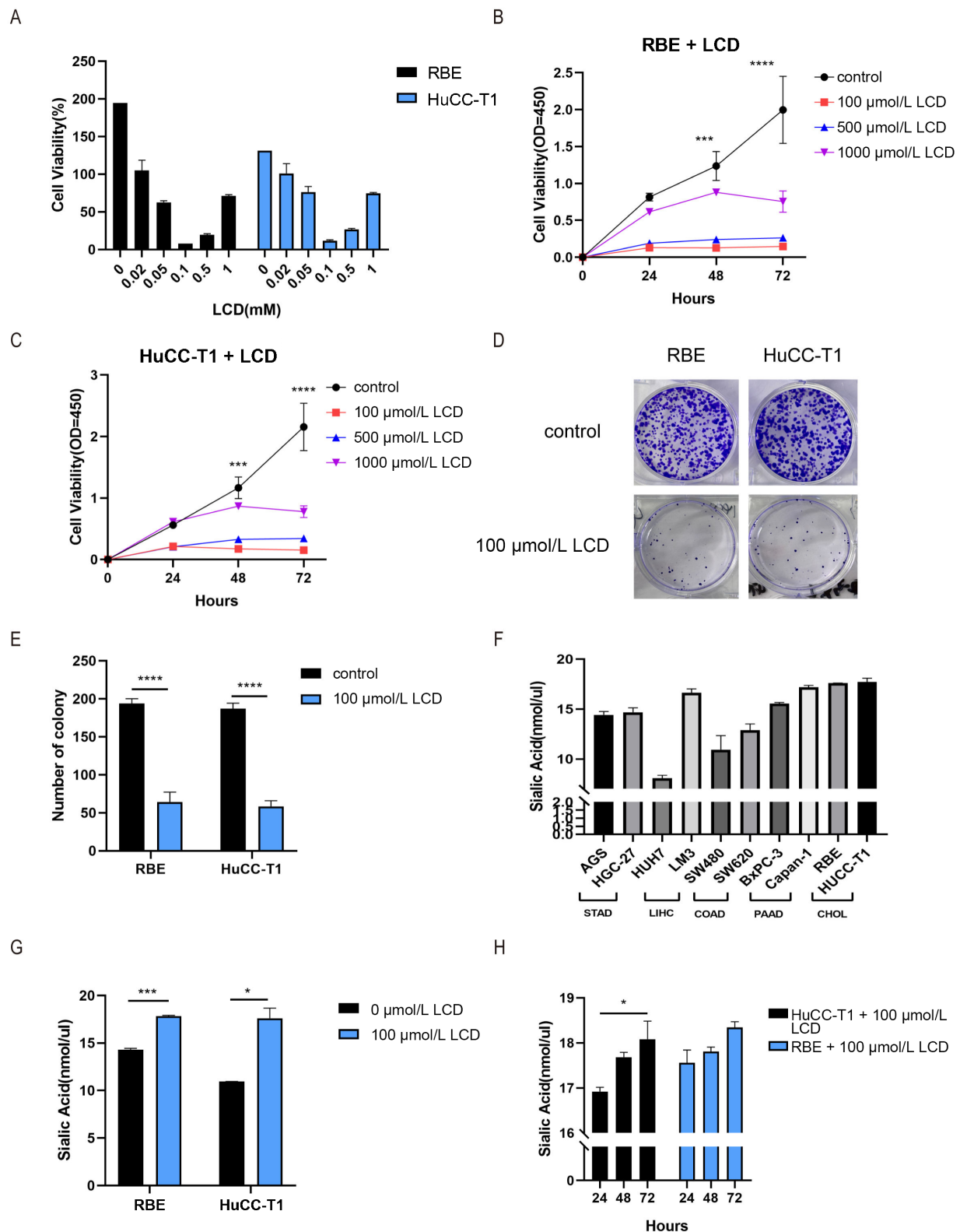


Figure 7: LCD inhibited proliferation of CHOL cells and increased cell-surface SA level (Table S4). (A) the cell viability (%) of CHOL cells was detected after adding different concentrations of LCD. (B and C) the cell proliferation ability of RBE cells and HuCC-T1 cells was detected after adding different concentrations of LCD. (D) the colony formation assay was performed after adding LCD. (E) statistical diagram of colony formation assay. (F) SA content (nmol/ μ L) of different types of cancer cells. (G) the SA content (nmol/ μ L) of CHOL cells after adding LCD was detected. (H) SA content (nmol/ μ L) of CHOL cells at different time points after adding LCD. (* $P < 0.05$; *** $P < 0.001$; **** $P < 0.0001$.) SA: sialic acid.

between SA biosynthesis and cancer metastasis.^[41,80]

Our in-depth analysis has revealed that the reason for the high SA level in CHOL tissues compared to the other digestive cancer types is due to the following. First, we have discovered that the key driving force for cancer development is the persistent alkalization of the intracellular pH due to continuous Fenton reactions,^[61,62,81,82] and the affected cells utilize predominantly nucleotide biosynthesis and SA biosynthesis to produce H^+ to keep the pH stable, which is summarized in equation (4). Hence, we infer that it is the combination of these two processes that roughly balances the OH^- produced by Fenton reactions. Table 5 shows the levels of contributions by nucleotide biosynthesis and SA biosynthesis toward neutralizing the Fenton reaction level across the seven cancer types, using a BIC-based analysis. From equation (4) and the table, we conclude that (i) it is the lower contribution by nucleotide synthesis towards Fenton reaction in CHOL, compared to other six cancer types, that forces the affected cells to increase the contribution by SA biosynthesis as otherwise the cells will die from intracellular alkalosis; and (ii) it is the BA that represses cell cycle progression *via* the BA receptor *GPBAR1*, which negatively correlates with cancer cell-cycle progression revealed by our analyses and is known to repress the cell cycle,^[71,72,83] which limits the rate of nucleotide synthesis. This conclusion is strongly supported by our own experimental study and published studies.

Computational analysis revealed that our findings are consistent with previous studies in that hypoxia and immune cell infiltration are associated with cancer malignancy. In particular, we found that CHOL tissues had high levels of hypoxia, a known factor for poor prognosis in cancer, which is associated with increased tissue damage and chronic inflammation, which was most directly associated with *HIF1A* expression levels. In addition, we found that CHOL had the highest infiltration levels of B cells, which may be related to the increased levels of SA, as SA is a capping molecule of cell surface glycoproteins that attract B cells.

Dedifferentiation is a mechanism adopted by differentiated cells to survive under sustained and extreme stress by acquiring more powerful capabilities of their less differentiated progenitors, and our analysis suggests that CHOL is under great stress. However, dedifferentiation was excluded from the regression analysis with five-year survival because it was causally related to hypoxia, B cell infiltration, and SA accumulation levels.

Our study also found that CHOL cells had the highest levels of SA accumulation on the cell surface, which was

associated with increased BA levels. BA plays a key role in inhibiting cell proliferation and indirectly driving SA biosynthesis and deployment by activating the *GPBAR1* receptor. This predicted causal inference relationship has been experimentally verified, that is, in the experimental verification, by inhibiting the proliferation of CHOL cells, the cell surface SA content increased.

Our study provides an in-depth understanding of the causes of CHOL as a highly malignant cancer, but there are still some limitations. Our study used transcriptomic data from the TCGA database, which may not accurately represent the activity levels of their encoded proteins. Experimental validation was mainly carried out *in vitro* cell models, which may not fully simulate the complex microenvironment in actual cancer tissues. Our future research will consider more omics data, and verify these findings in animal models, and investigate the possibility of inhibiting the function of *GPBAR1* as a drug target, which may reduce the level of malignancy of CHOL.

CONCLUSION

The main contribution of this study is that it has demonstrated *via* an integrated computational and experimental study, for the first time, that it is the high level of SA biosynthesis and accumulation on the cancer cell surface, dictated by the BA level *via* its receptor *GPBAR1*, that makes the CHOL the deadliest cancer type among the seven cancer types studied. This is consistent with the general knowledge that higher BA levels tend to make liver cancer more malignant.

Supplementary Information

Supplementary materials are only available at the official site of the journal (www.intern-med.com).

Acknowledgements

This research was supported by Center for Computational Science and Engineering at Southern University of Science and Technology. And authors thank the staff of the platform for their assistance.

Author contributions

Li X: Conceptualization, Formal analysis, Investigation, Methodology, Data curation, Software, Writing-original draft preparation. Liu A: Formal analysis, Methodology, Validation. Mu X: Investigation. Wang Z: Data curation. Xiao J: Visualization. Qu Y: Visualization. Huang Z: Methodology, Project administration. Zhang Y: Resources,

Validation, Supervision, Writing-review and editing. Xu Y: Conceptualization, Resources, Project administration, Supervision, Writing-review and editing. All authors have read and agreed to the published version of the manuscript.

Source of Funding

This research was funded by the National Natural Science Foundation of China (grant numbers T2350010, W2431059, and 82073244); Key University Laboratory of Metabolism and Health of Guangdong, Southern University of Science and Technology, Shenzhen 518055, China (grant number 2022KSYS007).

Ethical Approval

Not applicable.

Informed Consent

Not applicable.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Transcriptome data of cancer tissues and matched controls for 2165 cases of seven cancer types were obtained from the TCGA database through the GDC data portal (<https://portal.gdc.cancer.gov>). Five-year survival rate reports for seven cancer types are available through the SEER database (<https://seer.cancer.gov/>).

REFERENCES

- Sarcognato S, Sacchi D, Fassan M, Fabris L, Cadamuro M, Zanus G, *et al.* Cholangiocarcinoma. *Pathologica* 2021;113:158-169.
- Gupta R, Togashi J, Akamatsu N, Sakamoto Y, Kokudo N. Impact of incidental/misdiagnosed intrahepatic cholangiocarcinoma and combined hepatocellular cholangiocarcinoma on the outcomes of liver transplantation: an institutional case series and literature review. *Surg Today* 2017;47:908-917.
- Vogel A, Wege H, Caca K, Nashan B, Neumann U. The diagnosis and treatment of cholangiocarcinoma. *Dtsch Arztebl Int* 2014;111:748-754.
- Noh S, Ham A, Gil JR, Lee M, Lim T. Intrahepatic cholangiocarcinoma initially diagnosed as adenocarcinoma of unknown primary with hepatoduodenal ligament lymph node metastases: A case report. *Oncol Lett* 2023;27:7.
- Banales JM, Marin JGG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, *et al.* Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nat Rev Gastroenterol Hepatol* 2020;17:557-588.
- Andersen JB, Spee B, Blechacz BR, Avital I, Komuta M, Barbour A, *et al.* Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. *Gastroenterology* 2012;142:1021-1031.
- Banales JM, Cardinale V, Carpino G, Marziani M, Andersen JB, Invernizzi P, *et al.* Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nat Rev Gastroenterol Hepatol* 2016;13:261-280.
- Lindner P, Rizell M, Hafstrom L. The impact of changed strategies for patients with cholangiocarcinoma in this millenium. *HPB Surg* 2015;2015:736049.
- Kamsa-Ard S, Luvira V, Suwanrungruang K, Kamsa-Ard S, Luvira V, Santong C, *et al.* Cholangiocarcinoma trends, incidence, and relative survival in khon kaen, Thailand from 1989 through 2013: a opulation-based cancer registry study. *J Epidemiol* 2019;29:197-204.
- Strijker M, Belkouz A, van der Geest LG, van Gulik TM, van Hooft JE, de Meijer VE, *et al.* Treatment and survival of resected and unresected distal cholangiocarcinoma: a nationwide study. *Acta Oncol* 2019;58:1048-1055.
- Alabraba E, Joshi H, Bird N, Griffin R, Sturgess R, Stern N, *et al.* Increased multimodality treatment options has improved survival for Hepatocellular carcinoma but poor survival for biliary tract cancers remains unchanged. *Eur J Surg Oncol* 2019;45:1660-1667.
- Groot Koerkamp B, Wiggers JK, Allen PJ, Besselink MG, Blumgart LH, Busch OR, *et al.* Recurrence rate and pattern of perihilar cholangiocarcinoma after curative intent resection. *J Am Coll Surg* 2015;221:1041-1049.
- Komaya K, Ebata T, Yokoyama Y, Igami T, Sugawara G, Mizuno T, *et al.* Recurrence after curative-intent resection of perihilar cholangiocarcinoma: analysis of a large cohort with a close postoperative follow-up approach. *Surgery* 2018;163:732-738.
- Spolverato G, Kim Y, Alexandrescu S, Marques HP, Lamelas J, Aldrighetti L, *et al.* Management and outcomes of patients with recurrent intrahepatic cholangiocarcinoma following previous curative-intent surgical resection. *Ann Surg Oncol* 2016;23:235-243.
- Ouyang G, Liu Q, Wu Y, Liu Z, Lu W, Li S, *et al.* The global, regional, and national burden of gallbladder and biliary tract cancer and its attributable risk factors in 195 countries and territories, 1990 to 2017: A systematic analysis for the Global Burden of Disease Study 2017. *Cancer* 2021;127:2238-2250.
- Barsouk A, Rawla P, Barsouk A, Thandra KC. Epidemiology of cancers of the small intestine: trends, risk factors, and prevention. *Med Sci* 2019;7:46.
- Duggan MA, Anderson WF, Altekruse S, Penberthy L, Sherman ME. The surveillance, epidemiology, and end results (SEER) program and pathology: toward strengthening the critical relationship. *Am J Surg Pathol* 2016;40:e94-e102.
- Park HS, Lloyd S, Decker RH, Wilson LD, Yu JB. Overview of the Surveillance, Epidemiology, and End Results database: evolution, data variables, and quality assurance. *Curr Probl Cancer* 2012;36:183-190.
- Brown JM. Tumor hypoxia in cancer therapy. *Methods Enzymol* 2007;435:297-321.
- Höckel M, Vaupel P. Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J Natl Cancer Inst* 2001;93:266-276.
- Dhatchinamoorthy K, Colbert JD, Rock KL. Cancer immune evasion through loss of MHC class I antigen presentation. *Front Immunol* 2021;12:636568.
- Talseth-Palmer BA, Scott RJ. Genetic variation and its role in malignancy. *Int J Biomed Sci* 2011;7:158-171.
- Muz B, de la Puente P, Azab F, Azab AK. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia (Auckl)* 2015;3:83-92.
- Maltzman JS, Haase VH. Low oxygen stimulates the immune system. *Kidney Int* 2008;73:797-799.
- Krzywinska E, Stockmann C. Hypoxia, metabolism and immune cell function. *Biomedicines* 2018;6:56.
- Taylor CT, Colgan SP. Regulation of immunity and inflammation by hypoxia in immunological niches. *Nat Rev Immunol* 2017;17:774-785.

27. Fridman WH, Petitprez F, Meylan M, Chen TWW, Sun CM, Roumenina LT, *et al.* B cells and cancer: To B or not to B? *J Exp Med* 2021;218:e20200851.
28. Sarvaria A, Madrigal JA, Saudemont A. B cell regulation in cancer and anti-tumor immunity. *Cell Mol Immunol* 2017;14:662-674.
29. Miligy I, Mohan P, Gaber A, Aleskandarany MA, Nolan CC, Diez-Rodriguez M, *et al.* Prognostic significance of tumour infiltrating B lymphocytes in breast ductal carcinoma in situ. *Histopathology* 2017;71:258-268.
30. Guan H, Lan Y, Wan Y, Wang Q, Wang C, Xu L, *et al.* PD-L1 mediated the differentiation of tumor-infiltrating CD19(+) B lymphocytes and T cells in Invasive breast cancer. *Oncoimmunology* 2016;5:e1075112.
31. Zhang Y, Morgan R, Chen C, Cai Y, Clark E, Khan WN, *et al.* Mammary-tumor-educated B cells acquire LAP/TGF- β and PD-L1 expression and suppress anti-tumor immune responses. *Int Immunol* 2016;28:423-433.
32. Shen M, Wang J, Ren X. New insights into tumor-infiltrating B lymphocytes in breast cancer: clinical impacts and regulatory mechanisms. *Front Immunol* 2018;9:470.
33. Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, Weinberg RA, *et al.* Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* 2009;138:645-659.
34. Friedmann-Morvinski D, Verma IM. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep* 2014;15:244-253.
35. Gasic G, Gasic T. Removal of sialic acid from the cell coat in tumor cells and vascular endothelium, and its effects on metastasis. *Proc Natl Acad Sci U S A* 1962;48:1172-1177.
36. Ohta N, Pardee AB, McAuslan BR, Burger MM. Sialic acid contents and controls of normal and malignant cells. *Biochim Biophys Acta* 1968;158:98-102.
37. Bull C, den Brok MH, Adema GJ. Sweet escape: sialic acids in tumor immune evasion. *Biochim Biophys Acta* 2014;1846:238-246.
38. Bull C, Stoel MA, den Brok MH, Adema GJ. Sialic acids sweeten a tumor's life. *Cancer Res* 2014;74:3199-3204.
39. Munkley J, Scott E. Targeting aberrant sialylation to treat cancer. *Medicines* 2019;6:102.
40. Sato C, Kitajima K. Disialic, oligosialic and polysialic acids: distribution, functions and related disease. *J Biochem* 2013;154:115-136.
41. Sun H, Zhou Y, Jiang H, Xu Y. Elucidation of functional roles of sialic acids in cancer migration. *Front Oncol* 2020;10:401.
42. Brocca P, Rondelli V, Mallamace F, di Bari MT, Deriu A, Lohstroh W, *et al.* Water response to ganglioside GM1 surface remodelling. *Biochim Biophys Acta Gen Subj* 2017;1861:3573-3580.
43. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol* 2015;19:A68-A77.
44. Ziello JE, Jovin IS, Huang Y. Hypoxia-Inducible Factor (HIF)-1 regulatory pathway and its potential for therapeutic intervention in malignancy and ischemia. *Yale J Biol Med* 2007;80:51-60.
45. Cimmino F, Avitabile M, Lasorsa VA, Montella A, Pezone L, Cantalupo S, *et al.* HIF-1 transcription activity: HIF1A driven response in normoxia and in hypoxia. *BMC Med Genet* 2019;20:37.
46. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417-425.
47. Zhang Y, Jenkins DF, Manimaran S, Johnson WE. Alternative empirical Bayes models for adjusting for batch effects in genomic studies. *BMC Bioinformatics* 2018;19:262.
48. Walker WL, Liao IH, Gilbert DL, Wong B, Pollard KS, McCulloch CE, *et al.* Empirical Bayes accommodation of batch-effects in microarray data using identical replicate reference samples: application to RNA expression profiling of blood from Duchenne muscular dystrophy patients. *BMC Genomics* 2008;9:494.
49. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118-127.
50. Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, *et al.* Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. *PLoS One* 2011;6:e17238.
51. Foroutan M, Bhuvu DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. *BMC Bioinformatics* 2018;19:404.
52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545-15550.
53. Yi M, Nissley DV, McCormick F, Stephens RM. ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications. *Sci Rep* 2020;10:10258.
54. Wang X, VanValkenberg A, Odom-Mabey AR, Ellner JJ, Hochberg NS, Salgame P, *et al.* Comparison of gene set scoring methods for reproducible evaluation of multiple tuberculosis gene signatures. *bioRxiv* 2023:2023.01.19.520627.
55. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. *Semin Cancer Biol* 2017;44:141-152.
56. Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. Cell surface protein glycosylation in cancer. *Proteomics* 2014;14:525-546.
57. Nag S, Mandal A, Joshi A, Jain N, Srivastava RS, Singh S, *et al.* Sialyl-transferases and neuraminidases: potential targets for cancer treatment. *Diseases* 2022;10:114.
58. Aran D. Cell-type enrichment analysis of bulk transcriptomes using xCell. *Methods Mol Biol* 2020;2120:263-276.
59. Li Z, Kwon SM, Li D, Li L, Peng X, Zhang J, *et al.* Human constitutive androstane receptor represses liver cancer development and hepatoma cell proliferation by inhibiting erythropoietin signaling. *J Biol Chem* 2022;298:101885.
60. Ouyang N, Ke S, Eagleton N, Xie Y, Chen G, Laffins B, *et al.* Pregnane X receptor suppresses proliferation and tumorigenicity of colon cancer cells. *Br J Cancer* 2010;102:1753-1761.
61. Sun H, Zhang C, Cao S, Sheng T, Dong N, Xu Y. Fenton reactions drive nucleotide and ATP syntheses in cancer. *J Mol Cell Biol* 2018;10:448-459.
62. Sun H, Zhou Y, Skaro ME, Wu Y, Qu Z, Mao F, *et al.* Metabolic reprogramming in cancer is induced to increase proton production. *Cancer Res* 2020;80:1143-1155.
63. Liu D, Xing F, Wang Y, Xiao J, An Z, Xu Y. Molecular bases of morphologically diffused tumors across multiple cancer types. *Natl Sci Rev* 2022;9:nwac177.
64. Lenz M, Müller FJ, Zenke M, Schuppert A. Principal components analysis and the reported low intrinsic dimensionality of gene expression microarray data. *Sci Rep* 2016;6:25696.
65. Oettl LL, Scheller M, Filosa C, Wieland S, Haag F, Loeb C, *et al.* Phasic dopamine reinforces distinct striatal stimulus encoding in the olfactory tubercle driving dopaminergic reward prediction. *Nat Commun* 2020;11:3460.
66. Büll C, Boltje TJ, Balneger N, Weischer SM, Wassink M, van Gemst JJ, *et al.* Sialic acid blockade suppresses tumor growth by enhancing T-cell-mediated tumor immunity. *Cancer Res* 2018;78:3574-3588.
67. Linder AT, Schmidt M, Hitschfel J, Abeln M, Schneider P, Gerardy-Schahn R, *et al.* Sialic acids on B cells are crucial for their survival and provide protection against apoptosis. *Proc Natl Acad Sci U S A* 2022;119:e2201129119.
68. Doostkam A, Malekmakan L, Hosseinpour A, Janfeshan S, Roozbeh J, Masjedi F. Sialic acid: an attractive biomarker with promising biomedical applications. *Asian Biomed* 2022;16:153-167.
69. Jastrzab P, Narejko K, Car H, Wielgat P. Cell membrane sialome: sialic acids as therapeutic targets and regulators of drug resistance in human cancer management. *Cancers* 2023;15:5103.
70. Mullen NJ, Singh PK. Nucleotide metabolism: a pan-cancer metabolic

- dependency. *Nat Rev Cancer* 2023;23:275-294.
71. Chen WD, Yu D, Forman BM, Huang W, Wang YD. Deficiency of G-protein-coupled bile acid receptor Gpbar1 (TGR5) enhances chemically induced liver carcinogenesis. *Hepatology* 2013;57:656-666.
72. Qi Y, Duan G, Wei D, Zhao C, Ma Y. The bile acid membrane receptor TGR5 in cancer: friend or foe? *Molecules* 2022;27(16):5292.
73. Chen N, Wang J, Zhou L, Hu B, Chen Y, Zhu Z. GPBAR1 is associated with asynchronous bone metastasis and poor prognosis of hepatocellular carcinoma. *Front Oncol* 2023;12:1113785.
74. Nowack P, Runge J, Eyring V, Haigh JD. Causal networks for climate model evaluation and constrained projections. *Nat Commun* 2020;11:1415.
75. Lu C, Liu Y, Ali NM, Zhang B, Cui X. The role of innate immune cells in the tumor microenvironment and research progress in anti-tumor therapy. *Front Immunol* 2022;13:1039260.
76. Raskov H, Orhan A, Christensen JP, Gogenur I. Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. *Br J Cancer* 2021;124:359-367.
77. Abou Khouzam R, Zaarour RF, Brodaczewska K, Azakir B, Venkatesh GH, Thiery J, *et al.* The effect of hypoxia and hypoxia-associated pathways in the regulation of antitumor response: friends or foes? *Front Immunol* 2022;13:828875.
78. Nitschke L. CD22 and Siglec-G regulate inhibition of B-cell signaling by sialic acid ligand binding and control B-cell tolerance. *Glycobiology* 2014;24:807-817.
79. Jellusova J, Nitschke L. Regulation of B cell functions by the sialic acid-binding receptors siglec-G and CD22. *Front Immunol* 2011;2:96.
80. Al Saoud R, Hamrouni A, Idris A, Mousa WK, Abu Izneid T. Recent advances in the development of sialyltransferase inhibitors to control cancer metastasis: a comprehensive review. *Biomed Pharmacother* 2023;165:115091.
81. Tan R, Zhou Y, An Z, Xu Y. Cancer is A survival process under persistent microenvironmental and cellular stresses. *Genomics Proteomics Bioinformatics* 2023;21:1260-1265.
82. Zhou Y, Chang W, Lu X, Wang J, Zhang C, Xu Y. Acid-base homeostasis and implications to the phenotypic behaviors of cancer. *Genomics Proteomics Bioinformatics* 2023;21:1133-1148.
83. Keitel V, Donner M, Winandy S, Kubitz R, Haussinger D. Expression and function of the bile acid receptor TGR5 in Kupffer cells. *Biochem Biophys Res Commun* 2008;372:78-84.

How to cite this article: Li X, Liu A, Mu X, Wang Z, Xiao J, Qu Y, *et al.* Computational analyses to reveal the key determinants of the high malignancy level of cholangiocarcinoma. *J Transl Intern Med* 2024; 12: 602-617.