Title:

A distinct metabolic signature of human colorectal cancer with prognostic potential

Yunping Qiu,1,2 Guoxiang Cai,3 Bingsen Zhou,4 Dan Li,8 Aihua Zhao,1 Guoxiang Xie,5 Houkai Li,2 Sanjun Cai,3 Dong Xie,6 Changzhi Huang,7 Weiting Ge,8 Zhanxiang Zhou,2 Lisa X. Xu,9 Weiping Jia,1 Shu Zheng,8 Yun Yen,4,* Wei Jia1,5,*

1 Center for Translational Medicine, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, 200233, China
2 Center for translational Biomedical Research, University of North Carolina at Greensboro, North Carolina Research Campus, Kannapolis NC 28081, USA,
3 Department of Colorectal Surgery, Fudan University Shanghai Cancer Center; Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China.
4 Department of Medical Oncology and Therapeutic Research, City of Hope Comprehensive Cancer Center, Duarte, CA 91010, USA
5 Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, 96813, USA
6 Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, P. R. China
7 State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medicine Sciences and Peking Union Medical College, Beijing 100021, P. R. China
8 Cancer Institute, 2nd Hospital Affiliated to School of Medicine, Zhejiang University,
Hangzhou 310009, P.R. China

Med-X Institutes, Shanghai Jiao Tong University, Shanghai 200030, China

* To whom correspondence may be addressed.

Wei Jia; University of Hawaii Cancer Center, Honolulu, Hawaii, 96813, USA, Tel.: +1-808-564-5823; E-Mail: wjia@cc.hawaii.edu;

Yun Yen, City of Hope, 1500 E. Duarte Road Duarte, CA 91010, USA. Phone: 626 256-4673 x65707, Fax: 626 4717204, E-mail: yyen@coh.org

Running title:

Metabolic Signature of Human Colorectal Cancer

The authors declare no conflict of interest.

Key words: Metabolomics / metabonomics; colorectal cancer; prognosis; oxidative stress

Statement of translational relevance

Cancer cells undergo metabolic transformation to sustain fast cell growth and proliferation. This transformation would result in different metabolic phenotypes in cancer cells compared with their control counterparts. Identifying these differential
metabolites would be helpful in understanding cancer biology as well as in developing diagnostic and prognostic markers. We performed a comprehensive study which analyzed colorectal cancer (CRC) samples collected from 4 independent cohorts in China and the US. A panel of 15 differential metabolites was identified from these 4 cohorts, and demonstrated the ability to predict 5-year survival rate of CRC patients after standard surgical and medical treatment. These metabolite markers hold great potential for further development of CRC prognostic markers and/or therapeutic targets.

**Abbreviations:** CRC, Colorectal cancer; TCA cycle, tricarboxylic acid cycle; GC-TOFMS, gas chromatography Time-of-Flight mass spectrometry; ROC, Receiver operating characteristic; PCA, Principal component analysis; OPLS-DA, orthogonal partial least squares-discriminant analysis; VIP, variable importance in the projection; PP, predicted probability; AUC, area under curve; FFA, free fatty acids; PPP, pentose phosphate pathway; AMP, adenosine-5’-monophosphate; UMP, uridine-5’-monophosphate; MHB, 3-methyl-3-hydroxybutanoic acid.
Abstract

Purpose: Metabolic phenotyping has provided important biomarker findings, which, unfortunately, are rarely replicated across different sample sets due to the variations from different analytical and clinical protocols used in the studies. To date, very few metabolic hallmarks in a given cancer type have been confirmed and validated by use of a metabolomic approach and other clinical modalities. Here, we report a metabolomics study to identify potential metabolite biomarkers of colorectal cancer (CRC) with potential theranostic value.

Experimental Design: Gas chromatography-Time-of-Flight mass spectrometry (GC-TOFMS) based metabolomics was used to analyze 376 surgical specimens, which were collected from 4 independent cohorts of CRC patients at 3 hospitals located in China and City of Hope Comprehensive Cancer Center in the US. Differential metabolites were identified and evaluated as potential prognostic markers. A targeted transcriptomic analysis of 29 CRC and 27 adjacent non-tumor tissues was applied to analyze the gene expression levels for key enzymes associated with these shared metabolites.

Results: A panel of 15 significantly altered metabolites was identified, which demonstrates the ability to predict the rate of recurrence and survival for patients after
surgery and chemotherapy. The targeted transcriptomic analysis suggests that the
differential expression of these metabolites is due to robust metabolic adaptations in
cancer cells to increased oxidative stress as well as demand for energy, and
macromolecular substrates for cell growth and proliferation.

Conclusions: These CRC patients, despite their varied genetic background, mutations,
pathological stages, and geographical locations, shared a metabolic signature that is of
great prognostic and therapeutic potential.
Introduction

Cancer cells exhibit distinct metabolic phenotypes that are essential for sustaining high proliferative rates, and resist cell death signals associated with altered flux along key metabolic pathways, such as glycolysis and the tricarboxylic acid cycle (TCA cycle) (1). As exemplified in the “Warburg effect” (2), an increase in aerobic glycolysis is associated with the characteristic expression, mutation, and post-translational modification of enzymes involved in a number of key metabolic pathways, presumably due to the adaptation to oxidative stress associated with tumor hypoxia and mitochondrial mutations (3). Therefore, metabolic regulation is closely linked to cancer progression because proliferation is tightly regulated by the availability of nutrients. Moreover, oncogenes which promote proliferation likely both influence, and are also conversely influenced by metabolic changes.

The revival of interest in cancer cell metabolism in recent years has prompted the need for metabolomic phenotyping of clinical cancer specimens. However, the identification and validation of a distinct metabolic signature for a specific cancer proves to be challenging, due to the inter-individual variability of patients and the differing analytical and clinical protocols used in various studies (4). Although many studies revealed
different metabolic profiles in cancerous tissues (5-8), whether a certain cancer tends to maintain a unique metabolic transformation process to sustain uncontrolled proliferation and exhibit a universal “core” metabolome that is consistently identifiable among the tumors of different subjects at different pathological stages is uncertain.

Colorectal cancer (CRC) remains one of the most common types of cancers occurring worldwide (9), among which sporadic CRC represents an estimated 70% of all newly diagnosed cases. It is believed that sporadic CRC develops slowly through the progressive accumulation of multiple mutations that affect tumor suppressor genes, oncogenes, and downstream metabolic pathways (10). The global metabolic profiling of colon tissue could define metabolic signatures that not only discriminate malignant tissue from non-tumor tissue, but also distinguish the clinicopathological characteristics and treatment outcomes among CRC patients. Here, we describe a comprehensive metabolomic analysis of CRC tissue samples from multiple patient cohorts that consistently detected a panel of differentially expressed metabolites in human CRC tissues relative to adjacent non-tumor tissues.

Materials and Methods
Sample information

Samples were collected as surgical specimens (n = 376) from CRC patients treated at 4 hospitals located in China and the US following the same protocol. All of the samples were collected within 15 min after the surgery and immediately frozen at -80°C for metabolomic analysis. The first batch of samples was collected from 85 CRC patients from the Fudan University Shanghai Cancer Center, Shanghai, China, where 55 patients contributed paired samples, (i.e., both CRC tissue and adjacent non-tumor tissues located 5 cm from the edge of the tumor), and the other 30 patients contributed CRC tissue only. In addition, 3 validation batches of samples were collected from Cancer Hospital affiliated with the Chinese Academy of Medical Sciences, Beijing, China (n = 23 paired tissue samples with 2 batches of non-tumor tissues located 5 cm and 2-5 cm from the edge of the tumor, respectively), the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China (n = 65 paired tissue samples) and the City of Hope Comprehensive Cancer Center, California, USA (n = 20 paired tissue samples). Samples for the gene expression assay were provided by the City of Hope Comprehensive Cancer Center. A total of 56 tissue samples (29 CRC and 27 adjacent non-tumor tissues) from 34 patients were collected, 23 of which were paired samples and 12 were CRC or control samples only. The age, tumor stage, and tumor
location for all of these patients are provided in Table 1. None of the patients was on any neoadjuvant chemotherapy prior to surgical treatment. The protocol was approved by the Institutional Review Board from each of the 4 participating hospitals, and all participants in this study signed informed consent prior to the study.

Sample processing for metabolomics analysis

The metabolites extraction procedure followed our previous publication with minor modifications (11). Briefly, approximately 50 mg of each tissue sample was weighed and minced using liquid nitrogen. A 250 μL mixture of chloroform, methanol, and water (2:5:2) was added and the samples were vortexed for 1 min. The samples were then placed at -20°C for 20 min to extract metabolites, followed by centrifugation at 12,000 rpm for 10 min. The liquid layer was transferred into a new tube. The residue was extracted with 250 μL of methanol using the homogenizer for 10 min followed by centrifugation at 12,000 rpm for 10 min. The supernatant was combined with the previous extraction. After vortexing, a volume of 150 μL mixture was transferred to a glass vial spiked with internal standards (10 μL heptadecanoic acid at 1 mg/mL and 4-chlorophenylalanine at 0.3 mg/mL), which was then vacuum dried at room temperature.
The residue was chemically derivatized with a two-step procedure and then analyzed following the protocols previously published (12) with Pegasus HT system (Leco Corporation, St Joseph, USA) coupled with an Agilent 6890N gas chromatography. Briefly, 1 μL derivate was injected with a splitless mode at 270°C. The flow rate for the carrier gas, helium, was 1.0 mL/min. The oven program started at 80 °C for 2 min, and then ramped to 180 °C with 10 °C/min, to 230 °C with 6 °C/min, finally to 295 °C with 40 °C/min and hold for 8 min. The transfer interface and ion source was set to 270 °C and 220 °C, respectively. Data was acquired with m/z range of 30-600 at an acquisition rate of 20 spectrum/seconds.

**Quantitative real time PCR analysis**

RNA was isolated from tumor tissue or control samples using the Qiagen RNA easy kit (Qiagen, Germantown, MD, USA). Total RNA (1 μg) was converted into cDNA with the RT² HT First Strand Kit (Cat No. 330411, SABiosciences, Frederick, MD, USA) according to the manufacturer's instructions. The synthesized cDNA was used for multiple gene expression analyses using SABiosciences' RT² SYBR Green® qPCR Master Mixes in a Standard ABI 7500 system (Life Technologies Corporation, Carlsbad,
CA USA). The primer pairs were provided and pre-loaded in 96-well plates by SABiosciences for a customized gene panel (SABiosciences). Beta-actin was used as an internal control.

Data analysis

The acquired GC-TOFMS data were processed (including smoothing, de-noising, peak picking, identification, and alignment) using ChromaTOF software (v4.22, Leco Co., CA, USA) as described in a previous publication (13). Sample information, peak retention time, and peak area (quant mass) were included in the final dataset. Known artificial peaks, such as peaks caused by noise, column bleed, and BSTFA derivatization agents, were removed from the data set. The resulting data were normalized to internal standards and the weight of tissue sample prior to statistical analysis. Principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) was performed with SIMCA software (v 12.0, Umetrics, Umeå, Sweden). The default 7-fold cross-validation was applied, in order to guard against over-fitting. The variable importance in the projection (VIP) values (VIP > 1.0) are considered to be differentiating variables (14). Student’s t test was used for further differentiating
variables selection and validation ($p < 0.05$). Compound identification for GC-TOFMS was performed by comparing the mass fragments with NIST 11 Standard mass spectral databases in ChromaTOF software with a similarity of more than 70% and verified by available reference compounds. Receiver operating characteristic (ROC) analysis, binary logistic regression, and Kaplan-Meier analysis were performed with SPSS software (v20, IBM, Chicago, IL, USA). The data for gene expression were expressed as means ± standard error (SE). Differences were considered statistically significant at $p < 0.05$ from a Student’s $t$ test.

Results

Differentially expressed metabolites in CRC tissues relative to adjacent non-tumor tissues

PCA scores plot revealed a trend of separation between 85 tumor tissues and 55 adjacent non-tumor tissues collected from Shanghai (Supporting Figure 1A). An OPLS-DA model was obtained with one predictive component and two orthogonal components ($R_{X_{Cum}}^2=0.374$, $R_{Y_{Cum}}^2=0.706$, $Q_{Cum}^2=0.532$). The scores plot demonstrated a
separation between tumor tissues and adjacent non-tumor ones with few overlaps (Figure 1A). A 999-time permutation test was performed to validate the corresponding model. The result showed that the intercept for the Q2 to the Y axe is below zero (Q2 intercept (0,-0.097)) (Supporting Figure 1B), which indicated the validity of the current model. A number of 35 metabolites were identified with the criteria of VIP > 1 and p value in the student’s t test less than 0.05 (Fig. 1B, Table 2).

Replication analysis using CRC samples collected from independent cohorts

To replicate the metabolomics study of the Shanghai cohort and thus, to validate the findings of differential metabolites, three independent cohorts of tissue samples were collected from CRC patients at three hospitals in Hangzhou, and Beijing in China, as well as the City of Hope Comprehensive Cancer Center in the U.S., respectively. The samples were analyzed with the same analytical procedures as the samples collected from Shanghai cohort. Two samples (one in the Hangzhou batch and the other in the City of Hope batch) were excluded from the analysis due to the apparently poor extraction of the metabolites. We focused on these 35 differential metabolites obtained from the Shanghai cohort. A number of 25 metabolites were validated with significantly
statistical difference ($p < 0.05$) between CRC tissues and adjacent non-tumor ones in the Hangzhou cohort of CRC patients ($n = 65$, paired tissue samples). In the Beijing cohort, 24 metabolites were validated in tumor tissues ($n = 23$, paired tissue samples). In the fourth batch, 17 metabolites were significantly altered ($p < 0.05$) in tumor tissues from patients treated at the City of Hope ($n = 20$ paired tissue samples). Z-score plots of each differentially expressed metabolite in CRC tissue relative to the adjacent non-tumor tissue from 3 validation datasets are shown in Figure 2. Taken together, 15 metabolites among these 35 differential metabolites were significantly and consistently altered with the same up and down tendency in the three replication cohorts. The panel of 15 metabolites includes significantly elevated beta-alanine, palmitoleate, kynurine, putrescine, cysteine, lactate, glutamate, uracil, hypoxanthine, 5-oxoproline, 2-aminobutyrate, and aspartate, as well as down-regulated myo-inositol. In addition, paired $t$ test was performed on those 15 metabolites with samples in the Hangzhou, Beijing and City of Hope cohorts. All these 15 metabolites were significantly different between CRC tissues and adjacent non-tumor ones in all these three cohorts (Supporting table 1). The chromatogram from two representative samples (one from CRC tissue and one from adjacent non-tumor tissue) was provided in supporting figure 1C, where the 15 metabolite markers were marked. In addition, 5 metabolites were
significantly altered in all of samples collected from the three hospitals in China, but not in the samples from City of Hope, California (details are provided in Table 2). The non-tumor tissues were all collected from a section located 5 cm away from the edge of the tumor. We were also able to collect 22 non-tumor tissue samples from a section located 2-5 cm from the edge of the tumor in Beijing cohort. Our metabolomic results showed that alterations were less significant in tissue located 2-5 cm away from the tumor compared to 5 cm away, indicating a tendency of increased metabolic aberration in non-tumor tissue located closer to CRC tissue (Supporting table 1). Multivariate statistics was also performed to the samples in the Hangzhou, Beijing and City of Hope cohorts (Supporting text 1, Supporting figure 2).

**Expression levels of related genes**

We further analyzed 43 genes in the pathways of glucose metabolism, malate-aspartate shuttle, beta-alanine metabolism, tryptophan metabolism, glutathione metabolism and oxidative stress, and DNA repair genes. Twenty one genes were found to be differentially expressed between CRC and adjacent non-tumor tissues with RT-PCR analysis (Supporting table 2, and Supporting figure 3A). These differentially expressed genes can be categorized into 3 types of metabolic transformation to support the needs
for increased energy supply, macromolecule production, and the maintenance of redox balance under increased oxidative stress (Figure 3). The genes with the greatest up-regulation were SCD1 (FC=6.1) and FASN (FC=3.48), indicating elevated in situ fatty acid synthesis in cancer tissues (Figure 3B, Supporting table 2).

**Prognostic analysis in CRC tissue samples**

The metabolic signature of the 15 differentially expressed metabolites shown in Figure 3 was used to statistically predict the 5-year recurrence (including metastasis, R/M) rate of CRC. For most of the samples in the Shanghai patient cohort, follow-up information for 4-5 year post-surgery was available. We excluded patients diagnosed with stage IV CRC (those with metastases; n = 6) from the analysis. In the 79 patients with stage I-III CRC, 32 (40.51%) had relapsed and 16 (20.25%) had died during the 5 year period after undergoing surgery and receiving standard adjuvant chemotherapy (Supporting table 3). A binary logistic regression analysis was then performed R/M results as the dichotomous dependent variable (0= case free, 1= case) and these 15 differential metabolites plus age and gender as the covariates. The predicted probability (PP) values were obtained and saved as a new variable in the SPSS software. Using the PP values as the test variable and R/M as the state variable, a ROC curve was performed...
and a cut-off PP value (0.499) was selected based on Youden Index \( J, J = \max_c \{\text{Sensitivity (c)} + \text{Specificity (c)} - 1\} \)\( (15) \). Based on the cut-off value, these 79 CRC patients from the Shanghai cohort were divided into two groups: patients with PP values above and below the cut-off. Then a Kaplan-Meier analysis was performed for a recurrence rate analysis. The overall time to recurrence for those patients with PP values less than the cut-off value was significantly longer than those with PP values greater than the cut-off value (59.2 vs. 25.9 months, \( p = 2.50E-9 \); Figure 4B). As illustrated in Figure 4A, the metabolite panel had an AUC value of 0.895 (0.824-0.966, 95% confidence level) with a sensitivity of 0.750 and a specificity of 0.894, indicating that it provided reasonable accuracy for predicting the recurrence rate of these CRC patients over a 5 year post-operative period.

We also assessed the ability of this metabolic signature to predict the survival of the 79 CRC patients from the Shanghai cohort 5 years after surgery. As illustrated in Figure 4C, the AUC reached 0.860 (0.771-0.949, 95% confidence level), with a sensitivity of 0.938 and a specificity of 0.746. The overall survival time for those patients with PP values (determined using the same method as described above) less than the cut-off value (0.186) was significantly longer than those with PP values greater than the cut-off value (67.0 vs. 44.7 months, \( p = 6.48E-7 \); Figure 4D).
Discussion

Metabolic polymorphisms in human carcinogenesis derived from altered oncogenic expression, variable hypoxia levels, and the utilization of different carbon sources may produce diverse metabolic phenotypes and treatment responses. Metabolomic phenotyping of many types of cancers, such as prostate cancer, breast cancer, and CRC (16-18), has provided important biomarker findings, which unfortunately are rarely replicated across similar studies primarily due to the different analytical and clinical protocols used in the studies. As a result, very few metabolic hallmarks in a given cancer type have been discovered, confirmed, and validated by use of this approach or other investigational modalities.

In this study, we identified a panel of 15 differential metabolites in CRC tumors in 4 cohorts of CRC patients. We used this panel of metabolites to analyze the metabolomics data that we generated previously from gastric cardiac cancer patients (40 pairs of tissue samples) using the sample analytical platform and protocols. It appeared that this panel of markers was not able to separate between the tumor and non-tumor tissues for gastric cardiac cancer (Supporting text 2, Supporting figure 3B),
suggesting that they constitute a distinct metabolic signature of CRC. Using this panel of 15 metabolites, we are able to distinguish CRC patients with better prognostic outcomes, i.e., longer time-to-recurrence (61.4 vs. 28.9 months, \( p = 7.06 \times 10^{-9} \)) and better 5-year survival rate (66.0 vs. 44.8 months, \( p = 2.30 \times 10^{-6} \)) from those with earlier recurrence and lower survival rates. In addition, our result also revealed that the ratio of hypoxanthine to aspartate (Hyp/Asp) showed great potential for prognostic analysis. As shown in supporting text 3 and supporting figure 4, patients with lower ratio of Hyp/Asp had better outcomes in the following treatment compared with those patients with higher Hyp/Asp ratios. Compare to the 15 metabolite panel, the ratio of two metabolites may be simpler and more clinically applicable although it showed lower sensitivity and AUC value for recurrence prediction, and lower specificity and AUC value for survival prediction than the panel (Supporting text 3 and Supporting figure 4A, 4B).

The metabolomics approach has allowed us to unveil several key metabolic variations co-existing in CRC cells to support their proliferation, despite their varied genetic mutations and pathological stages. This may explain the ineffectiveness of chemotherapeutic agents that target only a single metabolic enzyme or a specific regulatory pathway. Recent cancer metabolomics studies have revealed important metabolic variations to sustain the fast growth of various cancer cells (19), some of
which have already been used as therapeutic targets, (such as L-asparaginase for leukemia (20)).

The Warburg effect is known to be a characteristic feature of cancer metabolism, which describes increased rate of glycolysis followed by lactic acid fermentation during tumor growth (21). Coincide with the result of Warburg effect, we found significantly higher levels of lactate in all CRC samples analyzed from the 4 patient cohorts. An elevated level of lactate in tissue samples and serum samples compared to their non-tumor counterparts was also observed in recent CRC metabolomics studies (12, 18, 22). Lactate dehydrogenase A (LDHA), which is a key enzyme that catalyzes pyruvate to lactate, was also found to be up-regulated in CRC tissues from our study, confirming that an increase in lactate levels plays an integral role in CRC metabolism.

In cancer metabolism, glycolysis is a preferred pathway for generating the metabolic intermediates used during de novo biosynthesis to support cell proliferation (see review (23)). This can result in higher levels of free fatty acids (FFA) and nuclear acids related metabolites such as myristic acid, palmitoleic acid and hypoxanthine in tumor tissues as was observed in our study. The higher level of ATP citrate lyase (ACLY) in the cytosol is important for generating substrates of fatty acid synthase (FASN), such as acetyl-CoA.
FASN and stearoyl-CoA desaturase (SCD) are key enzymes for unsaturated fatty acid synthesis, such as palmitoleic acid (24). The higher expression levels of ACLY, FASN and SCD in CRC samples found in our study suggest that de novo fatty acid synthesis is increased in these tumors. In addition to glycolysis, the pentose phosphate pathway (PPP) may also be activated and provide components for nucleic acid and fatty acid synthesis (ribose-5-phosphate and NADPH, respectively). Here, we found that transaldolase 1 (TALDO1), which is a key enzyme involved in PPP, was significantly increased in CRC tissue.

In addition to the generation of acetyl-CoA through the ACLY-mediated hydrolysis of citrate, oxaloacetate is another product that can be further converted to malate by malate dehydrogenase. Malic enzyme 1 (ME1), which catalyzes malate to pyruvate with the concomitant conversion of NADP\(^+\) to NADPH, was observed with significantly higher expression level in CRC tissues compared to non-tumor controls in this study. As NADPH is essential for fatty acid synthesis, the increased transformation from malate to pyruvate may compensate for the consumption of NADPH during the fast growth of tumor cells.
Beta-alanine was found to be the most significantly altered metabolite in the tumor tissues, as indicated in the Z-score plots. This metabolite was also previously identified as a metabolite marker in CRC tissues (5). Glutamate decarboxylase 1 (GAD1), which catalyzes aspartate to beta-alanine, was also higher in CRC tissues compared to adjacent non-tumor tissues. Our gene expression analysis revealed that pantothenate kinase 1 (PANK1), which is responsible for catalyzing the first and rate-limiting step of CoA biosynthesis (25), had significantly higher expression in CRC tissues. The higher level of cysteine (a substrate in CoA synthesis) in CRC samples observed in this study suggests that accelerated synthesis of CoA play a role in CRC-related morbidity. The gene encoding 4-aminobutyrate aminotransferase (ABAT), which catalyzes the conversion of beta-alanine to malonic semialdehyde, had higher expression in CRC tissues, suggesting that beta-alanine is metabolized to malonic semialdehyde for subsequent fatty acid synthesis through malonyl-CoA. Taken together, the increased need for acetyl-CoA and malonyl-CoA in fatty acid synthesis may contribute to the increased production of beta-alanine in CRC tissues.

Reactive oxygen species (ROS), as by-products of cellular metabolism, are associated with the increased metabolic activities in tumor cells (26). In fact, large amount of ROS was reported to be produced by several types of human tumor cells (27). Tumor cells
may undergo metabolic transformation to adapt to accelerated anabolic metabolism as well as elevated ROS levels during tumorigenesis. In addition, there was also a marked increase in glutamate, glycine, and cysteine, which are three precursors of glutathione. The γ-glutamyl cycle controls the synthesis and degradation of glutathione, and the intermediate, 5-oxoproline, which is an important factor in the pathway, was also significantly higher in CRC tissues compared with adjacent non-tumor tissues.

Ophthalmate was previously reported to be a biomarker for oxidative stress and indicative of glutathione consumption through the activation of γ-glutamyl cysteine synthetase (GCS) (28). Ophthalmate can be catalyzed by GCS and glutathione synthetase (GSS) from 2-aminobutyric acid, glutamate, and glycine (29). In our study, 2-aminobutyric acid and GSS were significantly elevated in CRC tissues, suggesting that ophthalmate synthase activity in CRC tissues is increased. Interestingly, a recent metabolomics study also detected higher levels of 2-aminobutyric acid in primary epithelial ovarian cancer compared to normal ovary tissue (30).

In addition to GSS, the expression level of several genes associated with GSH metabolism, including glutathione peroxidase 1 (GPX1), glutathione reductase (GSR), gamma-glutamylcyclotransferase (GGCT), and glutathione S-transferase pi 1 (GSTP1),
was also significantly higher in CRC samples compared to non-tumor controls, while aminopeptidase N (ANPEP) was significantly lower, respectively. GPX and GSR catalyze the transformation between reduced and oxidized GSH (known as the GSH redox cycle), which directly reflects cellular GSH homeostasis. The GSH redox cycle is also coupled with the NADP+/NADPH transformation. As discussed above, in order to meet the requirement of increased de novo fatty acid synthesis, NADPH production may be elevated in tumor cells through the activation of PPP, as evidenced by the elevated expression of TALDO1, and the metabolism of malate to pyruvate, as evidenced by elevated expression of ME1. GGCT catalyzes the degradation of gamma-glutamyl dipeptides to 5-oxoproline and L-amino acids (31). Therefore, the elevated expression level of GGCT may result in a higher level of 5-oxoproline and 2-aminobutyric acid in CRC tissues which was observed in this study. A higher expression level of GGCT was also recently suggested to be a potential biomarker for several cancers including CRC (32). ANPEP catalyzes the degradation of cysteinylglycine to cysteine and glycine, and was significantly lower in CRC tissues compared to control tissues, which is consistent with previous reports (33). GSTP1 was also found to be highly expressed in CRC tissues, which was previously reported to occur during human colon carcinogenesis in correlation with K-ras mutation (34). Taken together, these data show that several
metabolites involved in cellular antioxidation activity are over expressed in CRC tumors, indicating a robust metabolic adaptation to increased oxidative stress in these cells.

Increased oxidative stress is usually associated with increased oxidation of fatty acids, which may result in an accumulation of 3-hydroxybutyrate. We did not detect an increase in 3-hydroxybutyrate in the CRC samples from our study; however, increased 3-hydroxybutyrate was detected in the serum samples of CRC patients from one of our previous studies (12). These results suggest that higher fatty acid degradation may occur in circulating biofluids, but not in tumor cells, which further supports the hypothesis of enhanced metabolic adaptation of tumor cells to oxidative stress.

Increased oxidative stress was reported to induce dysregulation of the osmotic control in astrocytes, which resulted in a loss of myo-inositol (35). Therefore, a lower concentration of myo-inositol in the CRC tissues may result from the higher oxidative stress.

Kynurenine pathway is one of the main metabolic pathways of tryptophan metabolism, which is first catalyzed by indoleamine 2,3-dioxygenase (IDO) (36). Elevated expression of IDO was suspected as a mediator of tumor immune tolerance, which may help tumor cells avoid immune attack (37). Significant increase of IDO gene expression was also
observed in the CRC tissues compared with adjacent controls in our study. The increased expression of IDO may result in a higher level of kynurenine as was observed in our study. The kynurenine pathway would finally generate nicotinamide adenosine dinucleotide (NAD) from tryptophan. The activated kynurenine pathway may generate more NAD for the electron transport chain to meet the fast growth of tumor cells.

In summary, we identified a distinct metabolic signature with 15 metabolite markers from CRC tissue samples, which can be used to predict outcomes with surgical and chemotherapy treatment in CRC patients. The metabolic aberrations identified at gene expression level indicate a robust metabolic adaptation to sustain increased proliferation in CRC cells. Such a metabolic adaptation in CRC extends beyond the Warburg effect which only addresses the increased energy requirements through a preferred glycolysis process. We found that these metabolic changes in CRC provide support to the increased needs of energy, macromolecular precursors, as well as the maintenance of redox balance under strong oxidative stress.

ACKNOWLEDGMENTS
This work was financially funded by the following grants: the National Basic Research Program of China (No. 2007CB914700); the National Natural Science Foundation of China (No.81001055); the grant from the Shanghai Rising Star Program of the Science and Technology Commission of Shanghai Municipality (No.10QA1401400); the National High Technology Research and Development Program of China (No. 2012AA02A506).

The grant of the National Natural Science Foundation of China (No.81001055)
References


FIGURE LEGENDS

Figure 1. OPLS-DA scores plot for Z-score plot of differentially expressed metabolites in CRC tissues relative to adjacent non-tumor tissues. A, OPLS-DA scores plot for CRC samples collected from Shanghai. B, Z-score plot of differentiating metabolites between CRC tissues and adjacent non-tumor tissues. The values were standardized using the mean values and the standard deviations of adjacent non-tumor tissues in each group. Each vertical line represents one metabolite in one sample, colored by tissue type (blue star, non-tumor tissue; red vertical line, CRC tissue). For clarity, Z-score values were cut at 15 standard deviations. Abbreviations for metabolites: AMP, adenosine-5'-monophosphate; UMP, uridine-5'-monophosphate; MHB, 3-methyl-3-hydroxybutanoic acid.

Figure 2. Z-score plot of differentially expressed metabolites in CRC tissues relative to adjacent non-tumor tissues in the validation samples. The values were standardized using the mean values and the standard deviations of adjacent non-tumor tissues in each group. Each vertical line represents one metabolite in one sample, colored by tissue type (blue star, non-tumor tissue; red vertical line, CRC tissue). For clarity, Z-score values were cut at 15 standard deviations. A, samples from Hangzhou; B, samples from Beijing; and C, samples from City of Hope. AMP, adenosine-5'-monophosphate; MHB, 3-methyl-3-hydroxybutanoic acid.

Figure 3. Differential expression of metabolites and genes detected in colorectal cancer tissue and adjacent non-tumor controls. A, Venn diagram of the differentially expressed metabolites in different batches of CRC tissues compared to the corresponding adjacent controls. B, Metabolic correlation between those differentially expressed metabolites in all four groups and the differentially expressed genes between CRC tissues and adjacent controls.
Figure 4. ROC curves and Kaplan-Meier curves for disease recurrence and survival rate in samples from the patient cohorts from Shanghai using 15 differentially expressed metabolites. A, ROC curve using disease recurrence/metastasis (R/M) as the variable in the patient cohort from Shanghai; B, Kaplan-Meier curve comparing disease recurrence in the patient cohort from Shanghai with lower predicted probability values (black line) and higher predicted probability values (red line); C, ROC curve using overall survival as the variable in the patient cohort from Shanghai; D, Kaplan-Meier curve comparing overall survival for the patient cohort from Shanghai with lower predicted probability values (black line) and higher predicted probability values (red line).
Table 1. Demographic and clinical chemistry characteristics of human subjects

<table>
<thead>
<tr>
<th></th>
<th>Samples from Shanghai</th>
<th>Samples from Beijing</th>
<th>Samples from Hangzhou</th>
<th>Samples from City of Hope</th>
<th>Samples for gene analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>85</td>
<td>23</td>
<td>65</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>Age (median, range)</td>
<td>57, 31-79</td>
<td>61, 40-75</td>
<td>61, 34-84</td>
<td>59, 35-81</td>
<td>61, 36-86</td>
</tr>
<tr>
<td>Male/female ratio</td>
<td>49/36</td>
<td>10/13</td>
<td>43/22</td>
<td>8/12</td>
<td>10/24</td>
</tr>
<tr>
<td>TNM stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 0</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>1</td>
<td>/</td>
</tr>
<tr>
<td>Stage I</td>
<td>7</td>
<td>3</td>
<td>11</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Stage II</td>
<td>35</td>
<td>9</td>
<td>22</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Stage III</td>
<td>37</td>
<td>10</td>
<td>21</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>Stage IV</td>
<td>6</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>No.</td>
<td>Metabolite name</td>
<td>Samples from Shanghai</td>
<td>Samples from Hangzhou</td>
<td>Samples from Beijing</td>
<td>Samples from City of Hope</td>
</tr>
<tr>
<td>-----</td>
<td>----------------</td>
<td>-----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VIP</td>
<td>FC</td>
<td>P value</td>
<td>VIP</td>
</tr>
<tr>
<td>1</td>
<td>Kynurenine e</td>
<td>2.55</td>
<td>5.50</td>
<td>6.49E-14</td>
<td>2.64</td>
</tr>
<tr>
<td>2</td>
<td>Beta-alanine e</td>
<td>2.48</td>
<td>5.02</td>
<td>3.51E-13</td>
<td>4.56</td>
</tr>
<tr>
<td>3</td>
<td>Glutamate e</td>
<td>2.25</td>
<td>1.88</td>
<td>3.54E-10</td>
<td>1.92</td>
</tr>
<tr>
<td>4</td>
<td>Cysteine e</td>
<td>2.14</td>
<td>2.41</td>
<td>8.02E-10</td>
<td>1.87</td>
</tr>
<tr>
<td>5</td>
<td>2-Aminobutyrate e</td>
<td>2.13</td>
<td>1.81</td>
<td>3.05E-09</td>
<td>1.49</td>
</tr>
<tr>
<td>6</td>
<td>Palmitoleate e</td>
<td>1.90</td>
<td>3.09</td>
<td>3.03E-08</td>
<td>2.25</td>
</tr>
<tr>
<td>7</td>
<td>5-Oxoprolime e</td>
<td>1.96</td>
<td>1.76</td>
<td>5.89E-08</td>
<td>1.45</td>
</tr>
<tr>
<td>8</td>
<td>Aspartate e</td>
<td>1.92</td>
<td>1.80</td>
<td>1.17E-07</td>
<td>1.58</td>
</tr>
<tr>
<td>9</td>
<td>Hypoxanthine e</td>
<td>1.84</td>
<td>1.71</td>
<td>5.16E-07</td>
<td>1.36</td>
</tr>
<tr>
<td>10</td>
<td>Lactate e</td>
<td>1.87</td>
<td>1.68</td>
<td>5.89E-07</td>
<td>1.49</td>
</tr>
<tr>
<td>11</td>
<td>Myristate e</td>
<td>1.56</td>
<td>1.59</td>
<td>6.54E-05</td>
<td>2.06</td>
</tr>
<tr>
<td>12</td>
<td>Glycerol e</td>
<td>1.36</td>
<td>1.48</td>
<td>3.77E-04</td>
<td>1.37</td>
</tr>
<tr>
<td>13</td>
<td>Uracil e</td>
<td>1.36</td>
<td>1.39</td>
<td>4.59E-04</td>
<td>1.47</td>
</tr>
<tr>
<td>14</td>
<td>Putrescine e</td>
<td>1.21</td>
<td>3.57</td>
<td>4.61E-04</td>
<td>1.48</td>
</tr>
<tr>
<td>15</td>
<td>Myo-inositol e</td>
<td>1.10</td>
<td>-1.29</td>
<td>8.39E-03</td>
<td>-1.33</td>
</tr>
<tr>
<td>16</td>
<td>Hypotaurine</td>
<td>2.32</td>
<td>4.04</td>
<td>1.09E-11</td>
<td>2.59</td>
</tr>
<tr>
<td>17</td>
<td>Spermidine</td>
<td>1.37</td>
<td>1.89</td>
<td>1.02E-04</td>
<td>1.70</td>
</tr>
<tr>
<td>18</td>
<td>Homocysteine e</td>
<td>1.73</td>
<td>1.79</td>
<td>1.84E-06</td>
<td>1.45</td>
</tr>
</tbody>
</table>
A FC (fold change) with a positive value indicates a relatively higher level in tumor tissue while a negative value means a relatively lower expression level as compared to the adjacent normal tissues. \(^\text{b}\) \(P\)-values are calculated from Student’s t-test. \(^\text{c}\) Metabolites marked with \(/\) means these metabolites have \(p\) value higher than 0.05 in the corresponding batch of samples. AMP, adenosine-5'-monophosphate, UMP, uridine-5'-monophosphate. \(^\text{e}\) The identification of these metabolites were confirmed with our standard compounds.

<table>
<thead>
<tr>
<th></th>
<th>Metabolite</th>
<th>fold change</th>
<th>(\Delta)</th>
<th>(\Delta)FC</th>
<th>(p)-value</th>
<th>(p)-value</th>
<th>(p)-value</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>4-Aminobutyrate</td>
<td>1.27</td>
<td>2.18</td>
<td>2.97E-04</td>
<td>1.35</td>
<td>5.14E-03</td>
<td>2.27</td>
<td>6.06E-03</td>
</tr>
<tr>
<td>20</td>
<td>Asparagine (^\text{e})</td>
<td>1.33</td>
<td>1.47</td>
<td>4.40E-04</td>
<td>1.36</td>
<td>5.14E-03</td>
<td>1.96</td>
<td>1.86E-02</td>
</tr>
<tr>
<td>21</td>
<td>Glycerate (^\text{e})</td>
<td>1.12</td>
<td>-1.76</td>
<td>1.38E-02</td>
<td>-2.00</td>
<td>1.90E-02</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>22</td>
<td>Nicotinamide</td>
<td>1.39</td>
<td>1.56</td>
<td>2.54E-04</td>
<td>/</td>
<td>/</td>
<td>1.53</td>
<td>2.18E-02</td>
</tr>
<tr>
<td>23</td>
<td>AMP (^\text{e})</td>
<td>1.46</td>
<td>2.78</td>
<td>4.66E-05</td>
<td>2.83</td>
<td>7.59E-05</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>24</td>
<td>Ascorbate (^\text{e})</td>
<td>1.06</td>
<td>2.59</td>
<td>3.34E-03</td>
<td>1.43</td>
<td>2.07E-02</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>25</td>
<td>Glucose (^\text{e})</td>
<td>1.09</td>
<td>-1.51</td>
<td>6.50E-03</td>
<td>-2.37</td>
<td>1.01E-06</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>26</td>
<td>Xylose</td>
<td>1.16</td>
<td>-1.63</td>
<td>1.61E-02</td>
<td>-2.15</td>
<td>4.40E-02</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>27</td>
<td>Glycine (^\text{e})</td>
<td>1.64</td>
<td>1.62</td>
<td>7.49E-06</td>
<td>/</td>
<td>/</td>
<td>1.53</td>
<td>1.22E-02</td>
</tr>
<tr>
<td>28</td>
<td>Glyceraldehyde (^\text{e})</td>
<td>1.02</td>
<td>1.39</td>
<td>7.19E-03</td>
<td>/</td>
<td>/</td>
<td>1.37</td>
<td>4.15E-02</td>
</tr>
<tr>
<td>29</td>
<td>Ornithine (^\text{e})</td>
<td>1.07</td>
<td>1.39</td>
<td>9.49E-03</td>
<td>/</td>
<td>/</td>
<td>2.28</td>
<td>1.91E-02</td>
</tr>
<tr>
<td>30</td>
<td>Phosphate</td>
<td>1.64</td>
<td>1.53</td>
<td>5.84E-06</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>31</td>
<td>Laurate (^\text{e})</td>
<td>1.59</td>
<td>1.49</td>
<td>1.49E-05</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>32</td>
<td>Galactose</td>
<td>1.54</td>
<td>2.17</td>
<td>1.80E-05</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>33</td>
<td>3-Methy-3-hydropoxybutyrate</td>
<td>1.58</td>
<td>1.54</td>
<td>3.97E-05</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>34</td>
<td>Methioninamide</td>
<td>1.44</td>
<td>1.70</td>
<td>7.16E-05</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>35</td>
<td>2-Aminoadipate</td>
<td>1.19</td>
<td>1.45</td>
<td>8.04E-04</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>
Figure 3
Clinical Cancer Research

A distinct metabolic signature of human colorectal cancer with prognostic potential

Yunping Qiu, Guoxiang Cai, Bingsen Zhou, DEPT OF MOLECULAR PHARM, et al.

Clin Cancer Res Published OnlineFirst February 13, 2014.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-13-1939

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2014/02/14/1078-0432.CCR-13-1939.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://clincancerres.aacrjournals.org/content/early/2014/02/13/1078-0432.CCR-13-1939. Click on “Request Permissions” which will take you to the Copyright Clearance Center's (CCC) Rightslink site.