



OPEN

Plasma metabolomic differences in early-onset compared to average-onset colorectal cancer

Thejus Jayakrishnan¹, Arshiya Mariam^{2,3,12}, Nicole Farha^{1,12}, Daniel M. Rotroff^{2,3}, Federico Aucejo⁴, Shimoli V. Barot^{1,5}, Madison Conces^{5,6}, Kanika G. Nair^{1,5,7}, Smitha S. Krishnamurthi^{1,5,7}, Stephanie L. Schmit^{7,8,9}, David Liska^{5,7,10}, Alok A. Khorana^{1,5,7,12} & Suneel D. Kamath^{1,5,7,11,12}✉

Deleterious effects of environmental exposures may contribute to the rising incidence of early-onset colorectal cancer (eoCRC). We assessed the metabolomic differences between patients with eoCRC, average-onset CRC (aoCRC), and non-CRC controls, to understand pathogenic mechanisms. Patients with stage I–IV CRC and non-CRC controls were categorized based on age ≤ 50 years (eoCRC or young non-CRC controls) or ≥ 60 years (aoCRC or older non-CRC controls). Differential metabolite abundance and metabolic pathway analyses were performed on plasma samples. Multivariate Cox proportional hazards modeling was used for survival analyses. All *P* values were adjusted for multiple testing (false discovery rate, FDR $P < 0.15$ considered significant). The study population comprised 170 patients with CRC (66 eoCRC and 104 aoCRC) and 49 non-CRC controls (34 young and 15 older). Citrate was differentially abundant in aoCRC vs. eoCRC in adjusted analysis (Odds Ratio = 21.8, FDR $P = 0.04$). Metabolic pathways altered in patients with aoCRC versus eoCRC included arginine biosynthesis, FDR $P = 0.02$; glyoxylate and dicarboxylate metabolism, FDR $P = 0.005$; citrate cycle, FDR $P = 0.04$; alanine, aspartate, and glutamate metabolism, FDR $P = 0.01$; glycine, serine, and threonine metabolism, FDR $P = 0.14$; and amino-acid t-RNA biosynthesis, FDR $P = 0.01$. 4-hydroxyhippuric acid was significantly associated with overall survival in all patients with CRC (Hazards ratio, HR = 0.4, 95% CI 0.3–0.7, FDR $P = 0.05$). We identified several unique metabolic alterations, particularly the significant differential abundance of citrate in aoCRC versus eoCRC. Arginine biosynthesis was the most enriched by the differentially altered metabolites. The findings hold promise in developing strategies for early detection and novel therapies.

Keywords Early onset colorectal cancer, Metabolomics, Pathway analysis, Citric acid cycle, Arginine biosynthesis pathway, Synthetic lethality, Translational research

The incidence of early-onset colorectal cancer (eoCRC), defined as CRC diagnosed at age less than 50 years, has been steadily increasing, leading to the median age of diagnosis shifting from 72 years in the early 2000s to 66 years at present^{1–4}. The increasing incidence of eoCRC has recently prompted the American Cancer Society and US Preventive Services Task Force to recommend the initiation of screening earlier at age 45^{2,5}. Many of the environmental risk factors for average-onset CRC (aoCRC) are also relevant for eoCRC². These include lifestyle factors such as physical inactivity and high red meat consumption⁶. However, the drivers of the increasing

¹Department of Hematology and Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, USA. ²Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, USA. ³Center for Quantitative Metabolic Research, Cleveland Clinic, Cleveland, USA. ⁴Department of Surgery, Digestive Disease & Surgery Institute, Cleveland Clinic, Cleveland, USA. ⁵Case Comprehensive Cancer Center, Cleveland, USA. ⁶Department of Hematology-Oncology, University Hospital Seidman Cancer Center, Cleveland, USA. ⁷Center for Young-Onset Colorectal Cancer, Cleveland Clinic, Cleveland, USA. ⁸Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, USA. ⁹Population and Cancer Prevention Program, Case Comprehensive Cancer Center, Cleveland, USA. ¹⁰Department of Colorectal Surgery, Digestive Disease & Surgery Institute, Cleveland Clinic, Cleveland, USA. ¹¹Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH, USA. ¹²These authors contributed equally: Arshiya Mariam, Nicole Farha, Alok A. Khorana, and Suneel D. Kamath. ✉email: kamaths@ccf.org

trend, and potential novel risk factors or differential contributions of known/established risk factors, are not well understood.

Metabolomics involves global analysis of small molecule metabolites in body fluids or tissue extracts⁷. Evolving clinical utility of metabolomics such as risk prediction in cardiovascular disease has been demonstrated^{8,9}. In cancer, metabolomics may reflect the alterations resulting from cancer, and can identify the pathophysiological changes preceding cancer development related to one's exposures. Therefore, it could serve to identify cancer risk factors, study cancer biology, and develop novel therapeutics^{10,27}. Studies have identified metabolic alterations associated with CRC but such data in the context of eoCRC are limited^{10–12}. In the present study, we aimed to identify metabolomic differences between patients with eoCRC and aoCRC in comparison to non-CRC controls.

Methods

Patient samples

Patient plasma samples were obtained from prospective colorectal and liver tumor biobanks at the Cleveland Clinic from 01/2004 to 03/2021. Blood samples were obtained from patients on the day of their procedures, logged, and immediately stored at $-80\text{ }^{\circ}\text{C}$ until processed. The prospective biobanks and studies on banked specimens were approved by the Institutional Review Board (IRB) at the Cleveland Clinic. All the patients were treated at Cleveland Clinic. None of them were participating in clinical trials involving metabolic interventions such as arginine synthesis modulators or tricarboxylic acid (TCA) cycle inhibitors at the time of the study.

The samples were divided into cases and non-CRC controls. The cases included patients with stages I–IV CRC. For patients with early-stage cancers (stages I–III), the samples were obtained at the time of surgical resection of the primary disease, and for the stage IV group, the samples were obtained during liver metastasis resection. The cases were excluded if they were nonmalignant, or non-adenocarcinoma. The non-CRC controls included those who underwent liver resections or biopsies for benign causes or liver transplant donors and were selected based on the availability of bio-banked plasma samples for comparison. Of note, individuals with CRC who underwent surgical resection received colon cleansing preparation. Since the stage of cancer (and therefore the type of surgery including the need for cleansing preparation) was incorporated as a covariate in the modeling, any resulting impact would be accounted for, in the present analysis. Additionally, the use of plasma instead of tissue may mitigate any effect that colon preparation might have on the metabolome¹³.

The samples were categorized based on the age at the time of sample collection as ≤ 50 years (patients with eoCRC or young non-CRC controls) or ≥ 60 years (patients with CRC or older non-CRC controls). Retrospective chart reviews of the included patients were conducted to obtain clinical information.

Metabolomic analyses

Samples were submitted for metabolomic analyses by gas chromatography time-of-flight mass spectrometry (GC–TOF–MS) using the Primary Metabolism panel from West Coast Metabolomics at UC Davis¹⁴. This is a non-targeted plasma GC–TOF mass spectrometry assay of approximately 200 known and > 200 unknown metabolites including amino acid, carbohydrate and fatty acid metabolites. The list of known metabolites included in the analyses is available at <https://metabolomics.ucdavis.edu/core-services/metabolites>¹⁵.

The details of the technique including the validity of the methods, plasma extraction, and plasma metabolomics have been described previously^{16–19}. To begin, the samples were subjected to extraction using a solution consisting of acetonitrile, isopropanol, and water in a ratio of 3:3:2, which was chilled to $-20\text{ }^{\circ}\text{C}$ and degassed. A volume of 1 mL of this solution was utilized for extraction. Subsequently, 500 μL of the resulting supernatant was evaporated to dryness using a CentriVap (Labconco, Kansas, MO). For metabolite derivatization, a two-step process previously described was employed¹⁶. First, methoximation was employed to protect carbonyl groups, followed by the exchange of acidic protons with trimethylsilyl groups to enhance volatility. An injection of 0.5 μL sample volume was made into an Agilent 6890 GC (Agilent Technologies, Santa Clara, CA, USA), equipped with a Restek Rtx-5Sil MS column (30 m \times 0.25 mm, 0.25 μm) and operated with a splitless time of 25 s and a helium gas flow rate of 1 mL/min. The oven temperature was initially held at $50\text{ }^{\circ}\text{C}$ for 1 min and then increased to $330\text{ }^{\circ}\text{C}$ at a rate of $20\text{ }^{\circ}\text{C}/\text{min}$, where it was maintained for 5 min.

Data acquisition was performed using a Leco Pegasus IV time-of-flight mass spectrometer (Leco Corporation, St. Joseph, MI) with electron ionization at -70 eV . The mass spectra were recorded from 85 to 500 Da at a rate of 17 spectra/s and a detector voltage of 1850 V. The transfer line temperature was maintained at $280\text{ }^{\circ}\text{C}$, and the ion source temperature was set to $250\text{ }^{\circ}\text{C}$.

To ensure quality control, standard metabolite mixtures, and blank samples were injected at the beginning of the run and every ten samples throughout the analysis. Raw data were preprocessed using ChromaTOF version 4.50, which encompassed baseline subtraction, deconvolution, and peak detection. Metabolite annotation and reporting were performed using Binbase²⁰. The results were further analyzed using the bioinformatics workflow described below.

Association analyses

The associations between individual metabolites ($N = 449$) and disease status were investigated using logistic regression using the *stats* package in open-source statistical software R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria)²¹. These associations were used to investigate the associations of log-normalized metabolites with (1) patients with CRC vs. non-CRC controls, (2) patients with eoCRC vs. non-CRC controls aged ≤ 50 years, (3) patients with aoCRC vs. non-CRC controls aged ≥ 60 years, (4) patients with eoCRC versus aoCRC, and (5) non-CRC controls aged ≤ 50 versus ≥ 60 years. These associations were adjusted for sex, race, hyperlipidemia, obesity, cancer stage, and primary tumor location (left-sided vs. right-sided, rectal vs. non-rectal). Results are presented as odds ratios (OR) and 95% CI. For the comparison between aoCRC and eoCRC,

an OR > 1 indicated a higher abundance of metabolites associated with aoCRC. All *P* values were adjusted for multiple testing using the Benjamini–Hochberg false discovery rate approach²². FDR *P* < 0.15, previously used in other hypothesis-generating metabolomic studies, was used as the threshold for significance^{23–25}.

Pathway analyses

Metaboanalyst 5.0 was used for pathway analyses of the significantly altered metabolites (*P* < 0.05) using KEGG as the pathway reference. The enrichment method and topology analysis used hypergeometric tests and out-degree centralities, respectively. Pathways were considered enriched if the FDR *P* < 0.15.

Survival analyses

Overall survival (OS) was measured in months (m) from the day of cancer diagnosis to the day of death or censoring (last follow-up). Multivariate Cox proportional hazards modeling was performed to associate any metabolite with survival in patients with CRC (eoCRC and aoCRC). The model was adjusted for relevant demographic and clinical characteristics (race, sex, hyperlipidemia, obesity, stage, tumor-sidedness, and rectal cancer status). Statistical analyses were performed using the R *survival* package version 3.5-5²⁶.

Ethical approval

This study was performed under the oversight of the Cleveland Clinic institutional review board and the ethical approval process (IRB# 4134 and IRB# 10-347).

Conference presentation

This study was presented at the 2023 American Society of Clinical Oncology Gastrointestinal Cancers Symposium, and updated results were presented at the 2023 American Society of Clinical Oncology annual meeting as a podium presentation (Clinical cancer Symposium on the molecular basis of young-onset colorectal cancer). The first author (T.J.) received an ASCO Conquer Cancer Merit Award.

Results

Baseline characteristics

The study population comprised 170 patients with CRC (66 eoCRC and 104 aoCRC) and 49 non-CRC controls (34 with age < 50 years and 15 with age > 60 years). The majority of non-CRC controls were subjects with hepatic adenomas (*n* = 15, 31.3%), followed by healthy liver donors (*n* = 11, 22.9%), and those with other benign conditions of the liver, including cysts (*n* = 10, 20.8%), hemangiomas (*n* = 7, 14.6%), and focal nodular hyperplasia (*n* = 5, 10.4%).

The baseline characteristics of the patients are summarized in Table 1. The majority of the subjects were male (59.4%) for CRC and female (83.7%) for non-CRC controls, *P* < 0.0001. The race distribution was similar with the majority of both groups identifying as White—CRC (91.8%) and control (83.7%). The median age of patients with CRC (median = 62.95, IQR = 45.10–70.73) was greater than that of the non-CRC controls (median = 43.36, IQR = 32.83–64.35). There was no significant difference in the family history of CRC and personal or family history of colon polyps between patients with CRC and non-CRC controls. The majority had no significant family history of CRC (84.7% in CRC vs. 85.7% in control) or colon polyps (91.8 vs. 93.9%). There were no significant differences in smoking or alcohol consumption (*P* > 0.05). The differences in microsatellite instability (MSI) status and somatic mutations for patients that underwent testing are outlined in Supplementary Table 1. There was only 1 patient with MSI-H status in eoCRC (1.5%) and 4 patients (3.8%) in aoCRC.

Characteristics	All CRC (n = 170)	eoCRC (n = 66)	aoCRC (n = 104)	<i>p</i> -value ^a	Non-CRC Control (n = 49)	<i>p</i> -value ^b
Age (SD)	59.4 (16.0)	41.8 (6.1)	70.6 (8.5)	<0.001	47.4 (18.2)	NA
Sex: Male	101 (59.4%)	34 (51.5%)	67 (64.4%)	0.131	8 (16.3%)	<0.001
Sex: Female	69 (40.6%)	32 (48.5%)	37 (35.6%)		41 (83.7%)	
Race: White	156 (91.8%)	62 (93.9%)	94 (90.4%)	0.592	41 (83.7%)	0.164
Race: Black	12 (7.1%)	2 (3.0%)	10 (9.6%)		7 (14.3%)	
Race: Other	2 (1.2%)	2 (3.0%)	0		1 (2.0%)	
Diabetes	35 (20.6%)	5 (7.6%)	30 (28.8%)	0.002	5 (10.2%)	0.148
Hyperlipidemia	52 (30.6%)	12 (18.2%)	40 (38.5%)	0.009	26 (53.1%)	0.006
Obesity	41 (24.1%)	22 (33.3%)	19 (18.3%)	0.04	25 (51.0%)	0.001
Sidedness: Left sided	128 (75.3%)	59 (89.4%)	69 (66.3%)	0.003	-	NA
Rectal Primary	72 (42.4%)	35 (53.0%)	37 (35.6%)	0.035	-	NA
Stage IV cancer	77 (45.3%)	41 (62.1%)	36 (34.6%)	0.001	-	NA

Table 1. Baseline characteristics of patients selected for the analysis—*n* (%) for categorical variables and mean (standard deviation—SD) for continuous variables. *p*-value ^aFor eoCRC vs aoCRC groups comparison. *p*-value ^bFor CRC control groups comparison. CRC, colorectal cancer; eoCRC, early-onset colorectal cancer; aoCRC, average-onset colorectal cancer; NA, not applicable.

The mean BMI for patients with CRC and non-CRC controls were 28.71 (SD = 5.81) and 30.33 (SD = 6.34), respectively. Among the comorbidities assessed, the prevalence of the following was significantly different between patients with CRC and non-CRC controls—hyperlipidemia (30.6% in CRC group vs. 53.1% in the control group, $P = 0.006$) and obesity (24.1 vs. 51.0%, $P < 0.001$), while that of diabetes mellitus (20.6 vs. 10.2%, $P = 0.148$) was similar. Inflammatory bowel disease (IBD) was reported in 4.1% of CRC patients.

A higher proportion of patients with eoCRC than with aoCRC had the left-sided disease (89.4 vs. 66.3%, $P = 0.003$), rectal primary cancer (53.0 vs. 35.6%, $P = 0.037$), and stage IV disease (62.1 vs. 34.6%, $P = 0.001$). The prevalence of IBD was similar between the groups – 6.1% among patients with eoCRC vs 2.9% among patients with aoCRC ($P = 0.535$).

Hyperlipidemia and obesity were included as covariates because the proportions of these diseases varied significantly between patients with CRC and non-CRC controls at baseline ($P < 0.05$). The complete metabolic profiles of all the patients are shown in Supplementary Fig. 1.

Metabolomic analyses

Association analyses revealed four differentially abundant metabolites in the eoCRC versus aoCRC comparison: citrate (FDR $P = 0.04$), cholesterol (FDR $P = 0.14$), and two unidentified metabolites, UM118961 (FDR $P = 0.11$) and UM210714 (FDR $P = 0.11$), in unadjusted analyses (Fig. 1). The corresponding odds ratios were (OR > 1 indicating the association of higher abundance with aoCRC vs. eoCRC – citrate 14.54 (95% CI 4.26–56.35), cholesterol 0.01 (0.001–0.10), UM118961 0.52 (0.37–0.72), UM210714 3.67 (1.93–7.33) (Table 2). The differential abundance in citrate level remained significant on adjusted analysis—OR = 21.8 (95% CI 5.0–110.9, FDR $P = 0.04$). (Fig. 1).

No significant differentially abundant metabolites were found between the following cohorts: patients with eoCRC vs. young non-CRC controls, patients with aoCRC vs. older non-CRC controls, and control subgroups (age < 50 vs. age > 60) (FDR $P > 0.20$ all).

The following differentially abundant metabolites were observed in the unadjusted comparison between patients with CRC and non-CRC controls (OR > 1 indicates greater abundance in CRC): UM41873 (OR 0.60, 95% CI 0.45–0.75, FDR $P = 0.06$) and 4-hydroxyhippuric acid (OR 0.45, 95% CI 0.30–0.67, FDR $P = 0.09$) in the unadjusted analyses. These differences were not significant after adjustment for the following covariates—sex, race, hyperlipidemia, obesity, cancer stage, and primary tumor location (FDR $P = 0.71$ and 0.49, respectively).

The heatmap of all metabolites analyzed (including those not statistically significant) is shown in Supplementary Fig. 1 and individual profiles of metabolites with $P < 0.05$ are shown in Supplementary Fig. 2. Only nine metabolites were differentially abundant in both eoCRC vs. aoCRC and age < 50 non-CRC controls vs. age > 60 non-CRC controls.

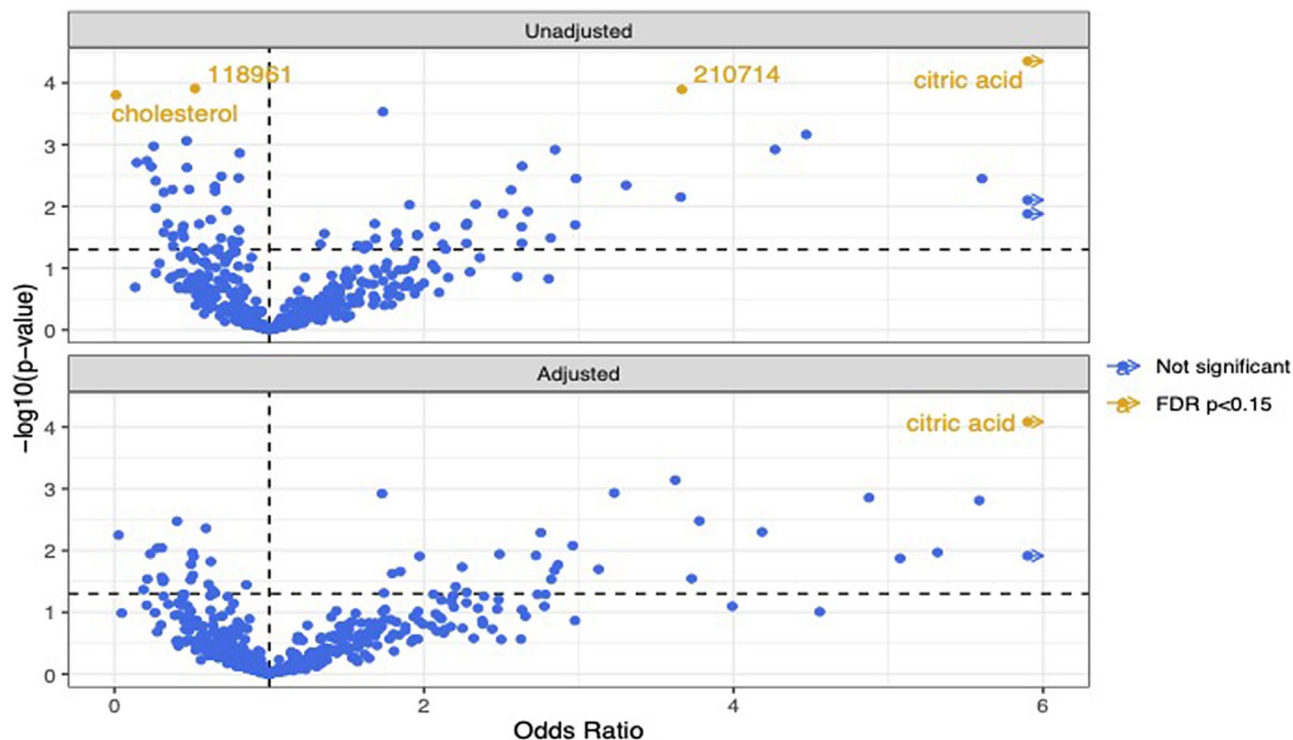


Figure 1. Volcano plot representing results of unadjusted and adjusted metabolomic analyses in comparing eoCRC and aoCRC. FDR, false discovery rate; corresponding values are described in Table 1.

Differentially abundant metabolite	OR (95% CI) ^a	FDR <i>P</i>	
Citrate	14.54 (4.26–56.35)	0.04	
Cholesterol	0.01 (0.001–0.10)	0.14	
UM118961	0.52 (0.37–0.72)	0.11	
UM210714	3.67 (1.93–7.33)	0.11	
Pathway	Pathway impact factor	Hits	FDR <i>P</i>
Arginine biosynthesis	0.13	3	0.02
Glyoxylate and dicarboxylate metabolism	0.23	5	0.005
Citrate (TCA) cycle	0.14	3	0.04
Alanine, aspartate and glutamate metabolism	0.28	4	0.01
Aminoacyl-tRNA biosynthesis	0.17	5	0.01
Glycine, serine and threonine metabolism	0.28	3	0.14

Table 2. Results of metabolomic analyses assessing differential abundance of metabolites between patients with aoCRC versus eoCRC and pathways impacted. ^aOdds ratio > 1 indicates association of higher abundance of metabolites with aoCRC versus yoCRC. The results are graphically represented in Figs. 1 and 2. FDR, false discovery rate.

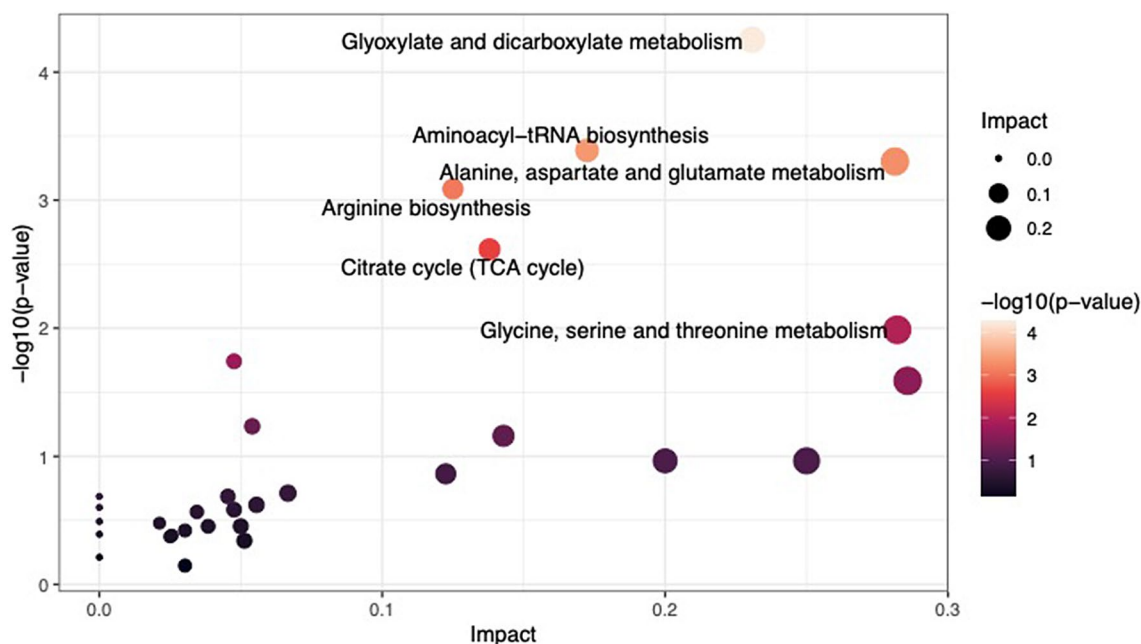


Figure 2. Bubble plot of pathway impact analyses. The impact factor topology analysis measures the impact of each metabolite in a pathway on a set of other metabolites in a pathway. Topology analysis was performed on the differentially expressed metabolites to estimate their cumulative impact on the pathways, corresponding values are described in Table 1.

Pathway analysis results

Metabolic pathways impacted by the differentially abundant metabolites included: carbohydrate metabolism (citrate/TCA cycle, FDR $P=0.04$, impact factor = 0.14), carbohydrate biosynthesis (glyoxylate and dicarboxylate metabolism, FDR $P=0.005$, impact factor = 0.23), amino acid metabolism (alanine, aspartate, and glutamate metabolism, FDR $P=0.01$, impact factor = 0.28; glycine, serine, and threonine metabolism, FDR $P=0.14$, impact factor = 0.28; arginine biosynthesis, FDR $P=0.02$, impact factor = 0.13, and amino-acid t-RNA biosynthesis, FDR $P=0.01$, impact factor = 0.17). The results are summarized in Table 2 and graphically represented in Fig. 2. There were no significant metabolomic differences between the young and older non-CRC controls. The arginine biosynthesis pathway, followed by the glyoxylate and dicarboxylate metabolism and TCA cycle pathways, were the most enriched with differentially expressed metabolites (Fig. 3). The Fig. 4 heatmap demonstrates the individual metabolite alterations in the pairwise comparison of patients with eoCRC and aoCRC reflected in the pathway analyses. The metabolic pathways with upregulated metabolites in eoCRC included: alanine, aspartate, and glutamate metabolism; aminoacyl-tRNA biosynthesis; arginine biosynthesis; glycine, serine, and threonine metabolism. The metabolites upregulated in multiple pathways included aspartic acid, alanine, glycine, and threonine. The pathways with upregulated metabolites in aoCRC (citric acid, isocitric acid, acotinic acid) were

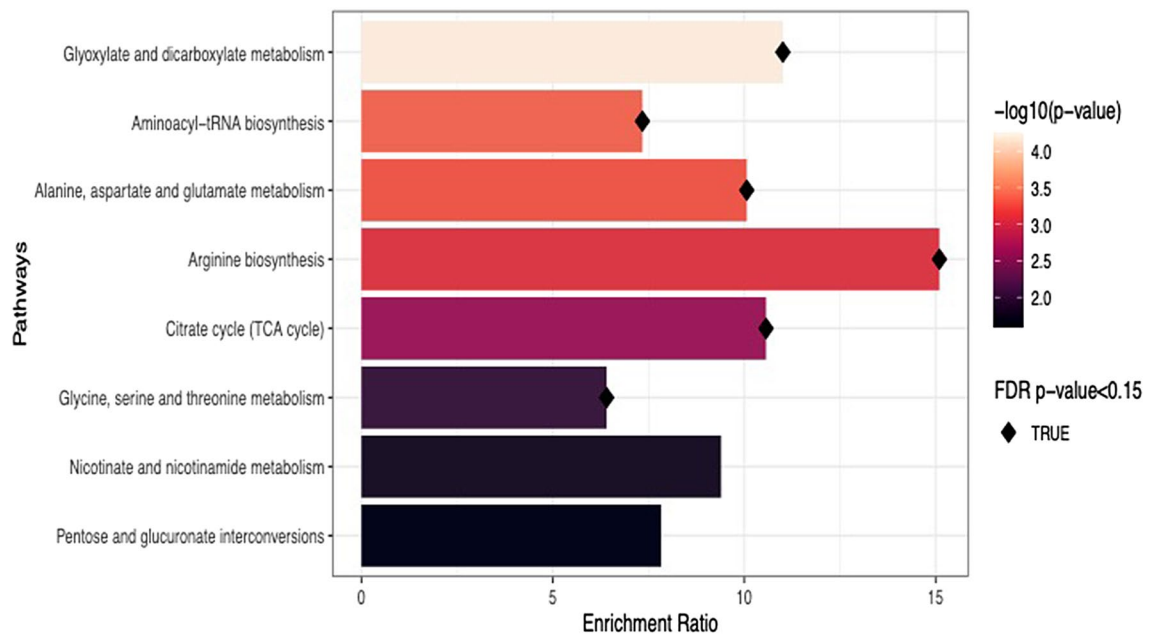


Figure 3. Results for pathway enrichment analyses. Enrichment analyses test whether more significant metabolites are involved in a pathway than expected by chance. The metabolites involved in the pathways are shown in the next figure (Fig. 4). FDR, false discovery rate.

the TCA cycle and glyoxylate and dicarboxylate metabolism. These metabolite level changes were mapped to the interlinked metabolic pathways as represented in Fig. 5.

Survival analyses

Overall survival (OS) analysis was performed in patients with CRC. The median follow-up period was 42 months (IQR = 26–64 months) for patients with eoCRC and 44 months (IQR = 29–84 months) for patients with aoCRC. During follow-up, 85 patients died (34.8% with eoCRC and 59.6% with aoCRC, $P = 0.003$).

Supplementary Fig. 3 summarizes the results of adjusted and unadjusted Cox regression analyses outlining both significant and non-significant results. In adjusted analyses, 4-hydroxyhippuric acid was significantly associated with OS in patients with CRC (HR = 0.4, 95% CI 0.3–0.7, FDR $P = 0.05$). As shown in Supplementary Fig. 4, the 25th percentile of 4-hydroxyhippuric acid abundance was associated with worse OS vs. the 75th percentile (45.98% vs. 62.16%). It was not associated with OS in patients with either eoCRC or aoCRC separately.

Adipic acid was statistically significantly associated with OS in patients with aoCRC in the unadjusted analysis (HR = 3.1, 95% CI 1.7–5.6, FDR $P = 0.13$). Compared with the 25th percentile of adipic acid abundance, the 75th percentile was associated with a lower 60-month OS in patients with aoCRC (47.15 vs. 62.91%) (Supplementary Fig. 5). However, this association was no longer significant in the adjusted analysis (HR = 2.6, 95% CI 1.2–5.5, FDR $P = 1$). Adipic acid was not associated with survival in patients with eoCRC or when eoCRC and aoCRC were combined.

Discussion

We identified significant differences in metabolites and associated pathways in comparing patients with eoCRC and aoCRC along with age-appropriate non-CRC controls. In particular, higher citrate levels were associated with aoCRC compared to eoCRC with an odds ratio of 21.8 (95% CI 5.0–110.935, FDR $P = 0.04$) on adjusted analyses. The arginine biosynthesis pathway, followed by the glyoxylate and dicarboxylate metabolism pathways and TCA cycle, were the most enriched, based on the relative abundance of metabolites in this study. Other pathway alterations reflected in metabolite levels and captured by metabolomics included the aminoacyl t-RNA biosynthesis; alanine, aspartate, and glutamate metabolism; and glycine, serine, and threonine metabolism. Similar studies comparing patients with colorectal cancer to those without cancer have identified additional pathway differences, such as lipid metabolism, which may be associated with variations in technique or patient selection^{11,12}. Interestingly, two unidentified metabolites (UM118961 and UM210714) were observed and will require further exploration in future studies^{27,28}. Association of higher levels of 4-hydroxyhippuric acid with improved overall survival was observed in the whole cohort.

Recent research has demonstrated the ability of metabolomics to identify metabolic pathway alterations associated with CRC^{11,29–37}. These include TCA cycle-related metabolites reflecting impaired mitochondrial respiration and oxidative stress^{10,11,32}. The Warburg effect, which involves increased aerobic glycolysis in cancer cells, also impacts the levels of the related TCA cycle metabolites^{7,29–31,38,39}. It is noteworthy that the alterations in these pathways are variable depending on the age of onset for CRC. In the present study, the TCA cycle metabolites including the key metabolite citrate were more abundant in aoCRC. As citrate has been proposed as a biomarker

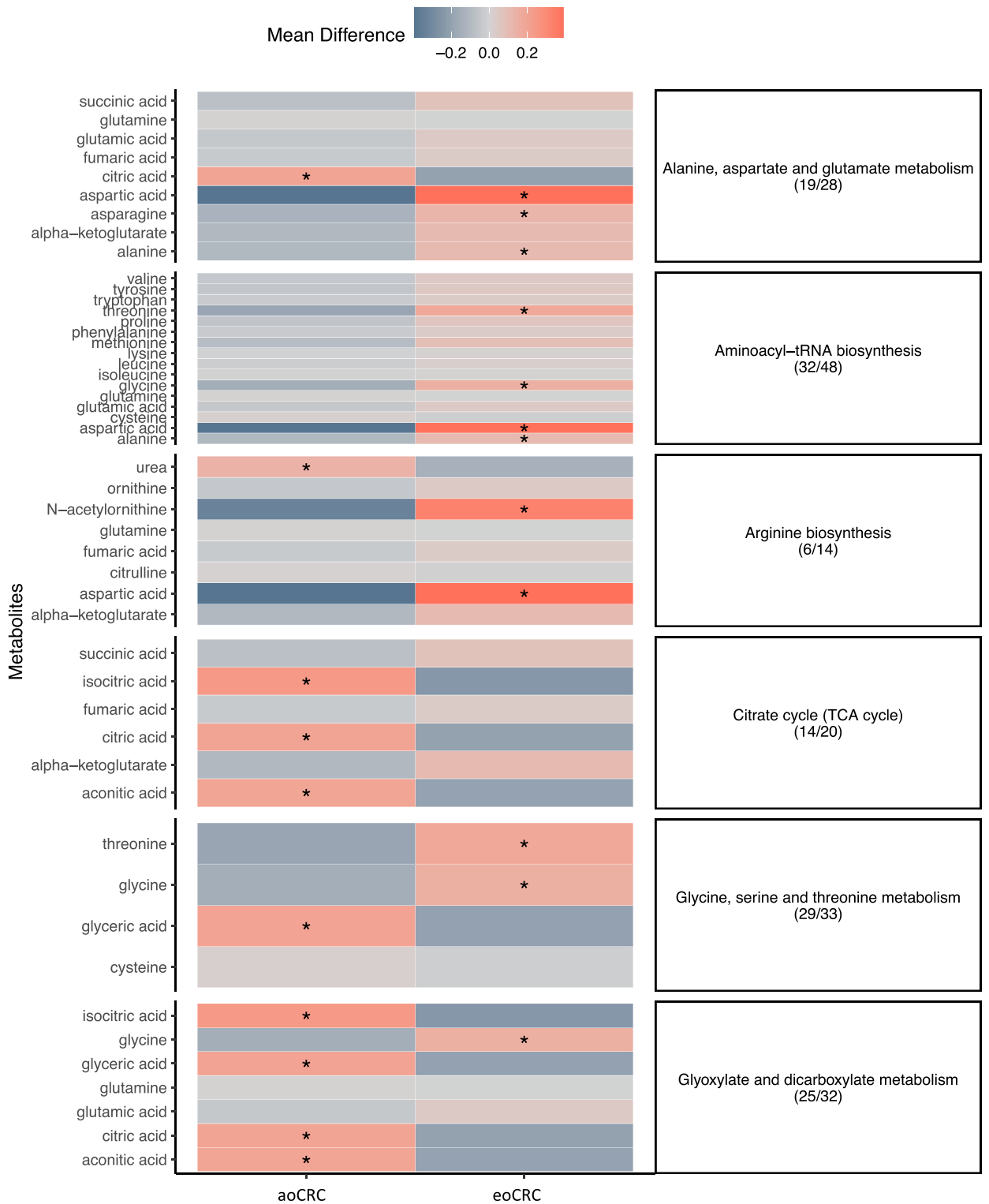


Figure 4. Heatmap outlining the metabolites significantly upregulated or downregulated in patients with eoCRC and aoCRC impacting the corresponding pathways. The heatmap was generated using R package ggplot2⁴⁹, *Indicate significantly altered metabolites. The numbers within parenthesis indicates the fraction of samples that exhibit the specific altered metabolite. eoCRC, early-onset colorectal cancer; aoCRC, average-onset colorectal cancer.

21.8 associated with the odds ratio (OR) for citrate suggests a strong association, the wide confidence interval could reflect some uncertainty surrounding the estimate, likely due to the small sample size. This highlights the importance of further validation of these findings in a larger study. Secondly, selection bias may have influenced baseline differences in comorbidities between patients with CRC and non-CRC controls. To address this, we included comorbidities as covariates in our regression models. Thirdly, given that an individual's exposure is variable and dynamic throughout their lifetime, confounding factors such as dietary and medication exposures could certainly influence the results. Chemotherapy administration could also affect the metabolome. In our study, only those with stage IV colorectal cancer were receiving chemotherapy near the time of sample collection. While we did not directly adjust for chemotherapy administration, we did adjust for the stage of disease, which is a proxy for the effect of chemotherapy administration. The adjustment of stage is also expected to address the differences in the proportions of patients with stage IV cancer including colorectal liver metastases that may impact the metabolome. We ensured uniformity in the sample collection timing as all blood samples were collected on the day of the procedure and after overnight fasting. Finally, we adjusted for primary tumor sidedness to address the confounding effect of different molecular alterations in left versus right-sided CRC, but residual confounding is still possible. Since all stage IV patients in this study had resectable liver metastases, it is also unclear if these results can be generalized to other metastatic sites such as lung or peritoneum.

Accounting for the heterogeneous molecular alterations presents inherent challenges from a combination of factors. First, the majority of the patient cohort under investigation had early-stage disease, rendering molecular profiling data unavailable as it is not a routine standard of care in such cases. Consequently, stratifying the cohort based on diverse molecular characteristics would have resulted in the creation of numerous subgroups and compromised the statistical power of the study in discerning metabolomic differences, which was the primary research aim. Moreover, the observed molecular differences did not exhibit substantial numerical differences between eoCRC and aoCRC.

Key findings of the study include the significant differential abundance of citrate in aoCRC compared to eoCRC with an odds ratio of 21.8 in adjusted analysis. Further investigations related to mechanisms underlying metabolomic differences in arginine biosynthesis and potential relationships with environmental exposures may elucidate the pathogenesis of eoCRC and offer opportunities for therapeutic targeting. We observed the association of 4-hydroxy Hippuric acid with overall survival, which could serve as a prognostic biomarker. In addition to further validation with targeted metabolomics, future efforts should include correlation with metagenomics and mechanistic studies.

Data and materials availability

TJ and SK had full access to all data in the study. We take full responsibility for the integrity of our data. The data generated in this study are not publicly available because the information could compromise patient privacy and consent, but are available upon reasonable request from the corresponding author.

Received: 19 October 2023; Accepted: 14 February 2024

Published online: 21 February 2024

References

1. Cancer of the Colon and Rectum - Cancer Stat Facts. SEER. Accessed December 1, 2019. <https://seer.cancer.gov/statfacts/html/colorect.html>
2. Wolf, A. M. D. *et al.* Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin.* **68**(4), 250–281. <https://doi.org/10.3322/caac.21457> (2018).
3. Kamath, S. D. *et al.* Racial disparities negatively impact outcomes in early-onset colorectal cancer independent of socioeconomic status. *Cancer Med.* **10**(21), 7542–7550. <https://doi.org/10.1002/cam4.4276> (2021).
4. Siegel, R. L. *et al.* Colorectal cancer statistics, 2020. *CA Cancer J Clin.* **70**(3), 145–164. <https://doi.org/10.3322/caac.21601> (2020).
5. Final Recommendation Statement: Screening for Colorectal Cancer [United States Preventive Services Taskforce. Accessed March 3, 2023. <https://www.uspreventiveservicestaskforce.org/uspstf/announcements/final-recommendation-statement-screening-colorectal-cancer-0>
6. Red Meat Genetic Signature for Colorectal Cancer - NCI. Published July 22, 2021. Accessed February 12, 2023. <https://www.cancer.gov/news-events/cancer-currents-blog/2021/red-meat-colorectal-cancer-genetic-signature>
7. Schmidt, D. R. *et al.* Metabolomics in cancer research and emerging applications in clinical oncology. *CA Cancer J Clin.* **71**(4), 333–358. <https://doi.org/10.3322/caac.21670> (2021).
8. Talmor-Barkan, Y. *et al.* Metabolomic and microbiome profiling reveals personalized risk factors for coronary artery disease. *Nat Med.* **28**(2), 295–302. <https://doi.org/10.1038/s41591-022-01686-6> (2022).
9. Zhu, Q. *et al.* Plasma metabolomics provides new insights into the relationship between metabolites and outcomes and left ventricular remodeling of coronary artery disease. *Cell Biosci.* **12**(1), 173. <https://doi.org/10.1186/s13578-022-00863-x> (2022).
10. Holowatyj, A. N. *et al.* Distinct molecular phenotype of sporadic colorectal cancers among young patients based on multiomics analysis. *Gastroenterology* **158**(4), 1155–1158.e2. <https://doi.org/10.1053/j.gastro.2019.11.012> (2020).
11. Tan, B. *et al.* Metabonomics identifies serum metabolite markers of colorectal cancer. *J. Proteome Res.* **12**(6), 3000–3009. <https://doi.org/10.1021/pr400337b> (2013).
12. Gumpenberger, T. *et al.* Untargeted metabolomics reveals major differences in the plasma metabolome between colorectal cancer and colorectal adenomas. *Metabolites* **11**(2), 119. <https://doi.org/10.3390/metabo11020119> (2021).
13. Powles, S. T. R. *et al.* Effects of bowel preparation on intestinal bacterial associated urine and faecal metabolites and the associated faecal microbiome. *BMC Gastroenterol.* **22**, 240. <https://doi.org/10.1186/s12876-022-02301-1> (2022).
14. West Coast Metabolomics Center - Assays and Services. Accessed December 15, 2022. <https://metabolomics.ucdavis.edu/core-services/assays-and-services>
15. West Coast Metabolomics Center - Metabolites. Accessed June 15, 2023. <https://metabolomics.ucdavis.edu/core-services/metabolites>
16. Fiehn, O. Metabolomics by gas chromatography-mass spectrometry: Combined targeted and untargeted profiling. *Curr. Protoc. Mol. Biol.* **114**, 30.4.1–30.4.32. <https://doi.org/10.1002/0471142727.mb3004s114> (2016).

17. Aboud, O. *et al.* Application of machine learning to metabolomic profile characterization in glioblastoma patients undergoing concurrent chemoradiation. *Metabolites* **13**(2), 299. <https://doi.org/10.3390/metabo13020299> (2023).
18. Ismail, I. T. *et al.* Sugar alcohols have a key role in pathogenesis of chronic liver disease and hepatocellular carcinoma in whole blood and liver tissues. *Cancers (Basel)* **12**(2), 484. <https://doi.org/10.3390/cancers12020484> (2020).
19. Miyamoto, S. *et al.* Systemic metabolomic changes in blood samples of lung cancer patients identified by gas chromatography time-of-flight mass spectrometry. *Metabolites* **5**(2), 192–210. <https://doi.org/10.3390/metabo5020192> (2015).
20. Krishnapuram, B., Carin, L., Figueiredo, M. A. T. & Hartemink, A. J. Sparse multinomial logistic regression: Fast algorithms and generalization bounds. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(6), 957–968. <https://doi.org/10.1109/TPAMI.2005.127> (2005).
21. R: The R Project for Statistical Computing. Accessed March 20, 2023. <https://www.r-project.org/>
22. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **57**(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x> (1995).
23. McCullough, M. L., Hodge, R. A., Campbell, P. T., Stevens, V. L. & Wang, Y. Pre-diagnostic circulating metabolites and colorectal cancer risk in the cancer prevention study-II nutrition cohort. *Metabolites* **11**(3), 156. <https://doi.org/10.3390/metabo11030156> (2021).
24. Moore, S. C. *et al.* A metabolomics analysis of postmenopausal breast cancer risk in the cancer prevention study II. *Metabolites* **11**(2), 95. <https://doi.org/10.3390/metabo11020095> (2021).
25. Moore, S. C. *et al.* A metabolomics analysis of body mass index and postmenopausal breast cancer risk. *J. Natl. Cancer Inst.* **110**(6), 588–597. <https://doi.org/10.1093/jnci/djx244> (2018).
26. Therneau TM, until 2009) TL (original S >R port and R maintainer, Elizabeth A, Cynthia C. survival: Survival Analysis. Published online February 12, 2023. Accessed March 3, 2023. <https://CRAN.R-project.org/package=survival>
27. BinVestigate - The BinBase investigation tool. Accessed December 19, 2022. <https://bininvestigate.fiehnlab.ucdavis.edu/#/bin/118961>
28. BinVestigate - The BinBase investigation tool. Accessed December 19, 2022. <https://bininvestigate.fiehnlab.ucdavis.edu/#/bin/210714>
29. Gu, J. *et al.* Metabolomics analysis in serum from patients with colorectal polyp and colorectal cancer by 1H-NMR spectrometry. *Dis Markers* **2019**, 3491852. <https://doi.org/10.1155/2019/3491852> (2019).
30. Troisi, J. *et al.* A metabolomics-based screening proposal for colorectal cancer. *Metabolites* **12**(2), 110. <https://doi.org/10.3390/metabo12020110> (2022).
31. Denkert, C. *et al.* Metabolite profiling of human colon carcinoma—deregulation of TCA cycle and amino acid turnover. *Mol. Cancer* **7**, 72. <https://doi.org/10.1186/1476-4598-7-72> (2008).
32. Zhu, G. *et al.* Untargeted GC-MS-based metabolomics for early detection of colorectal cancer. *Front. Oncol.* **11**, 729512. <https://doi.org/10.3389/fonc.2021.729512> (2021).
33. Zhu, C. *et al.* Distinct urinary metabolic biomarkers of human colorectal cancer. *Dis. Markers* **2022**, 1758113. <https://doi.org/10.1155/2022/1758113> (2022).
34. Wang, M. *et al.* Discovery of plasma biomarkers for colorectal cancer diagnosis via untargeted and targeted quantitative metabolomics. *Clin. Transl. Med.* **12**(4), e805. <https://doi.org/10.1002/ctm2.805> (2022).
35. Zhang, C. *et al.* Metabolomic profiling identified serum metabolite biomarkers and related metabolic pathways of colorectal cancer. *Dis. Markers* **2021**, 6858809. <https://doi.org/10.1155/2021/6858809> (2021).
36. Wu, J., Wu, M. & Wu, Q. Identification of potential metabolite markers for colon cancer and rectal cancer using serum metabolomics. *J. Clin. Lab. Anal.* **34**(8), e23333. <https://doi.org/10.1002/jcla.23333> (2020).
37. Rothwell, J. A. *et al.* Circulating amino acid levels and colorectal cancer risk in the European prospective investigation into cancer and nutrition and UK biobank cohorts. *BMC Med.* **21**, 80. <https://doi.org/10.1186/s12916-023-02739-4> (2023).
38. Arima, K. *et al.* Metabolic profiling of formalin-fixed paraffin-embedded tissues discriminates normal colon from colorectal cancer. *Mol. Cancer Res.* **18**(6), 883–890. <https://doi.org/10.1158/1541-7786.MCR-19-1091> (2020).
39. Zhang, Y. *et al.* Alteration of plasma metabolites associated with chemoradiosensitivity in esophageal squamous cell carcinoma via untargeted metabolomics approach. *BMC Cancer* **20**(1), 835. <https://doi.org/10.1186/s12885-020-07336-9> (2020).
40. Bertini, I. *et al.* Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer Res.* **72**(1), 356–364. <https://doi.org/10.1158/0008-5472.CAN-11-1543> (2012).
41. Philip, P. A. *et al.* Avenger 500, a phase III open-label randomized trial of the combination of CPI-613 with modified FOLFIRINOX (mFFX) versus FOLFIRINOX (FFX) in patients with metastatic adenocarcinoma of the pancreas. *JCO* **37**(4_suppl), TPS479–TPS479 (2019).
42. Mohan, A. *et al.* Devimistat in combination with gemcitabine and cisplatin in biliary tract cancer: Pre-clinical evaluation and phase 1b multicenter clinical trial (BilT-04). *Clin. Cancer Res.* <https://doi.org/10.1158/1078-0432.CCR-23-0036> (2023).
43. Du, T. & Han, J. Arginine metabolism and its potential in treatment of colorectal cancer. *Front. Cell Dev. Biol.* **9**, 658861. <https://doi.org/10.3389/fcell.2021.658861> (2021).
44. Bateman, L. A. *et al.* Argininosuccinate synthase 1 is a metabolic regulator of colorectal cancer pathogenicity. *ACS Chem. Biol.* **12**(4), 905–911. <https://doi.org/10.1021/acschembio.6b01158> (2017).
45. Kremer, J. C. *et al.* Arginine deprivation inhibits the Warburg effect and upregulates glutamine anaplerosis and serine biosynthesis in ASS1-deficient cancers. *Cell Rep.* **18**(4), 991–1004. <https://doi.org/10.1016/j.celrep.2016.12.077> (2017).
46. Kuo, M. T., Savaraj, N. & Feun, L. G. Targeted cellular metabolism for cancer chemotherapy with recombinant arginine-degrading enzymes. *Oncotarget* **1**(4), 246–251. <https://doi.org/10.18632/oncotarget.135> (2010).
47. Peyraud, F. *et al.* Circulating L-arginine predicts the survival of cancer patients treated with immune checkpoint inhibitors. *Ann. Oncol.* **33**(10), 1041–1051. <https://doi.org/10.1016/j.annonc.2022.07.001> (2022).
48. Tan, G. *et al.* Three serum metabolite signatures for diagnosing low-grade and high-grade bladder cancer. *Sci. Rep.* **7**(1), 46176. <https://doi.org/10.1038/srep46176> (2017).
49. ggplot2: Elegant Graphics for Data Analysis (3e). Accessed November 16, 2023. <https://ggplot2-book.org/>

Acknowledgements

Illustration by Amanda Mendelsohn. Reprinted with the permission of the Cleveland Clinic Enterprise Creative Services.

Author contributions

Study concept and design: T.J., S.K., A.A.K. Data collection: T.J., N.F., S.K. Analysis and interpretation of data: T.J., A.M., D.R., A.A.K., and S.K. Drafting of the manuscript: All authors. Critical revision of the manuscript for important intellectual content: All authors. Administrative, technical, or material support: A.A.K., S.K.

Funding

Support for this study was provided by the Sondra and Stephen Hardis Chair in Oncology Research (A.A.K.)

Competing interests

Thejus Jayakrishnan – No disclosures Nicole Farha – No disclosures Arshiya Mariam – No disclosures Daniel M. Rotroff – Equity stake in Clarified Precision Medicine, LLC. Received research support from Novo Nordisk, consulting honoraria from Pharmazaam, LLC, and Interpares Biomedicine owns intellectual property related to the detection of type-2 diabetes, chronic liver disease, and liver cancer Federico Aucejo – No disclosures Shimoli V. Barot – No disclosures Madison Conces – No disclosures Kanika G. Nair – No disclosures Smitha S. Krishnamurthi -- Consulting or Advisory Role--Array BioPharma, Research funding- Bristol Myers Squibb, Pfizer, Aravive, Natera Stephanie L. Schmit – No disclosures David Liska – Consulting or Advisory Role- Olympus Medical Systems, Research funding – Merck, Freenome Suneel D. Kamath – Consulting or advisory role: Exelixis, Guardant Health. Tempus, SeaGen, Foundation Medicine. Alok A. Khorana has been paid honoraria - Janssen, Sanofi, BMS, Anthos, Pfizer, Bayer, Nektar, and Medscape. Consulting or advisory role: Janssen, Sanofi, BMS, Anthos, Pfizer, Bayer, Nektar, and Medscape

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-54560-5>.

Correspondence and requests for materials should be addressed to S.D.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024