

1 **Potential role of fructose on human colon DNA methylation in racial disparities observed**  
2 **for colorectal cancer risk.**

3  
4 **Short title:** DNA methylation response to fructose in the colon.

5  
6 **Matthew A. Devall**<sup>1</sup>, Stephen Eaton<sup>1</sup>, Gaizun Hu<sup>2</sup>, Xiangqing Sun<sup>1</sup>, Ethan Jakum<sup>3</sup>, Samyukta  
7 Venkatesh<sup>1</sup>, Steven M. Powell<sup>4</sup>, Cynthia Yoshida<sup>4</sup>, Daniel J. Weisenberger<sup>5</sup>, Gregory S. Cooper<sup>6</sup>,  
8 Joseph Willis<sup>7</sup>, Seham Ebrahim<sup>2</sup>, Jamie Zoellner<sup>8,9</sup> Graham Casey<sup>8,10</sup> and Li Li<sup>1,9\*</sup>

- 9  
10 1. Department of Family Medicine, University of Virginia, Charlottesville, VA, USA.  
11 2. Molecular Physiology and Biological Physics University of Virginia, Charlottesville, VA,  
12 USA.  
13 3. Department of Biology, University of Virginia, Charlottesville, VA, USA.  
14 4. Digestive Health Center, University of Virginia, Charlottesville, VA, USA.  
15 5. Department of Biochemistry and Molecular Medicine, University of Southern California,  
16 Los Angeles, CA, USA.  
17 6. Department of Medicine, Case Western Reserve University/University Hospitals  
18 Cleveland Medical Center, Cleveland, OH, USA.  
19 7. Department of Pathology, Case Western Reserve University/University Hospitals  
20 Cleveland Medical Center, Cleveland, OH, USA.  
21 8. Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA.  
22 9. University of Virginia Comprehensive Cancer Center, University of Virginia,  
23 Charlottesville, VA, USA.  
24 10. Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA.

25 **Grant Support:** The work was supported by research grants from the National Cancer Institute  
26 Cancer Disparities SPORE Planning Grant (P20 CA233216), Case Comprehensive Cancer  
27 Center GI SPORE (P50 CA150964), National Cancer Institute (CA143237, CA204279), and a  
28 pilot grant from the University of Virginia Cancer Center (P30CA044579). Funders had no  
29 involvement in the study design, collection, analysis and interpretation of data or writing of the  
30 report.

31 **Abbreviations:** AA: African American; BMI: body mass index; CAPI: computer-assisted personal  
32 interview; CRC: colorectal cancer; DEG: differentially expressed gene; DNAm: DNA methylation,  
33 EA: European American; FDR: false-discovery rate; FFQ: food-frequency questionnaire; HFCS:  
34 high fructose corn syrup; NAT: normal-adjacent tissue; SSB: sugar sweetened beverage; TCGA-  
35 COAD: The Cancer Genome Atlas Colon Adenocarcinoma.

36 **Correspondence:**

37 Li Li, MD, PhD  
38 Department of Family Medicine  
39 University of Virginia School of Medicine  
40 McKim Hall, Room 3156  
41 1415 Jefferson Park Avenue  
42 PO Box 800729  
43 Charlottesville, Virginia 22908  
44 Tel: 434-924-1611  
45 Email: [ll8nv@virginia.edu](mailto:ll8nv@virginia.edu)

46 **Disclosures:** the authors have nothing to disclose. The authors have no conflicts of interest.

47 **Transcript profiling:** The data generated here are from our previously published, publicly  
48 available dataset: GSE151732.

49 **Writing assistance:** N/A.

50 **Author contributions:** Writing original draft: MD; data curation: MD, XS, SV, SMP, CY, GSC,  
51 JW, DJW and LL; study design and methodology: MD, SE, DJW, GC, and LL; statistical analysis  
52 and data visualization: MD; sample management and annotation: SV, SE, XS; project  
53 conceptualization: MD and LL, funding: LL and GC; supervision and project administration: LL  
54 and GC. All authors were responsible for writing – review and editing of the final draft.

55 **Abstract:**

56 **Background and aims:** An increasing body of observational studies has linked fructose intake  
57 to colorectal cancer (CRC). African Americans (AAs) are significantly more likely than European  
58 Americans to consume greater quantities of fructose and to develop right-side colon cancer. Yet,  
59 a mechanistic link between these two associations remains poorly defined. We aimed to identify  
60 differentially methylated regions (DMRs) associated with dietary fructose consumption measures  
61 obtained from food frequency questionnaires in a cohort of normal colon biopsies derived from  
62 AA men and women (n=79)

63 **Methods:** DNA methylation data from this study was obtained using the Illumina Infinium  
64 MethylationEPIC kit and is housed under accession GSE151732. DMR analysis was carried out  
65 using *DMRcate* in right and matched left colon, separately. Secondary analysis of CRC tumors  
66 was carried out using data derived from TCGA-COAD, GSE101764 and GSE193535. Differential  
67 expression analysis was carried out on CRC tumors from TCGA-COAD using *DESeq2*.

68 **Results:** We identified 4,263 right-side fructose-DMRs. In contrast, only 24 DMRs survived  
69 multiple testing corrections (FDR<0.05) in matched, left colon. To identify targets by which dietary  
70 fructose drives CRC risk, we overlaid these findings with data from three CRC tumor datasets.  
71 Remarkably, almost 50% of right-side fructose-DMRs overlapped regions associated with CRC  
72 in at least one of three datasets. *TNXB* and *CDX2* ranked among the most significant fructose  
73 risk DMRs in right and left colon respectively that also displayed altered gene expression in CRC  
74 tumors.

75 **Conclusions:** Our mechanistic data support the notion that fructose has a greater CRC-related  
76 effect in right than left AA colon, alluding to a potential role for fructose in contributing to racial  
77 disparities in CRC.

78  
79 **Keywords:** DNA methylation, disparities, African American, colon cancer, fructose, race.  
80  
81

82

83

84

85

86

87

## 88 Introduction

89 The natural sweetness of dietary fructose makes it an appealing additive to processed foods  
90 and beverages, and most caloric sweeteners contain at least 40% fructose<sup>1</sup>. High fructose corn  
91 syrup (HFCS) is the most common form of sweetener added to sugar sweetened beverages  
92 (SSB)<sup>2, 3</sup>. Shifts in dietary patterns have seen dramatic increase of HFCS consumption in recent  
93 years, with research showing it to account for approximately 6-10% of caloric intake among US  
94 adults<sup>4-6</sup>. Similarly, there was also a ~63% increase in SSB consumption among US children of  
95 schooling age between the years 1989 to 2008<sup>7</sup>. Racial disparities in SSB consumption are also  
96 evident<sup>8-10</sup>. For example, the odds of drinking SSBs one or more times per day is significantly  
97 higher among non-Hispanic black respondents (AOR = 1.65) versus non-Hispanic white  
98 respondents<sup>8</sup>. This shift in increased SSB consumption and documented racial disparities are  
99 gravely concerning given the increasingly recognized role of excess dietary fructose in various,  
100 negative health outcomes<sup>11-13</sup>, including two recent landmark murine studies, which have begun  
101 to define mechanisms linking SSB/HFCS to colon carcinogenesis<sup>14, 15</sup>. Despite these studies, the  
102 role of dietary fructose in healthy human colon, and as such, the mechanisms that may drive  
103 fructose-mediated CRC risk, remain poorly understood.

104 An increasing body of observational studies has directly linked intake of fructose in obesity,  
105 diabetes and CRC risk<sup>11, 16-18</sup>. A small study in Iran recently found that increased fructose intake  
106 was associated with CRC after adjusting for multiple confounding factors including body mass  
107 index (BMI)<sup>19</sup>. This finding is in line with results from a larger study of two prospective US cohorts,  
108 which found that dietary fructose and SSB consumptions were both positively associated with  
109 incidence and mortality of proximal CRC<sup>18</sup>. Other research has shown that high SSB consumption  
110 was also associated with early-onset CRC in women<sup>16</sup>. Notably, African Americans (AAs) have  
111 higher incidence of CRC compared to any other racial/ethnic group<sup>20</sup>. AAs are also more likely to  
112 be diagnosed at a younger age and with poorer survival outcomes than EAs<sup>20, 21</sup>. Moreover, AAs  
113 are more likely to develop right- than left-sided colon cancer and more likely to develop right-sided

114 colon cancer than EAs<sup>22</sup>. However, the mechanisms driving the observed racial disparities in CRC  
115 risk and outcomes and tumor sidedness are poorly understood.

116 DNA methylation (DNAm) represents one epigenetic mechanism through which gene  
117 expression patterns may be modulated to alter disease risk in response to environmental cues<sup>23</sup>.  
118 Aberrant DNAm is a hallmark of CRC<sup>24, 25</sup>, and differences in DNAm have also been found to be  
119 associated with fructose and HFCS in other tissues<sup>26-28</sup>. We have established a repository of right  
120 and left normal colon biopsies from a cohort of well-characterized average-risk AA and EA  
121 individuals<sup>29</sup>. In this cohort, we previously reported colon location and race specific DNAm  
122 differences, where we found that many of these differences occurred in regions known to be  
123 associated with CRC. Further analysis of this cohort has the potential to reveal novel insight into  
124 the biological mechanisms that drive CRC racial disparities and to help inform the development  
125 of tailored prevention strategies.

126 Here, we extend the results identified in our previous study to identify differences in DNAm  
127 associated with total dietary fructose in AA colon. We hypothesize that this will provide novel  
128 insight into the mode of action of fructose in the human colon, and that through the overlay of  
129 CRC-related publicly available data, we may contextualize these mechanisms in the framework  
130 of CRC risk.

131

## 132 **Materials and Methods**

133

### 134 *Collection of dietary data for analysis*

135 Clinical data were obtained from the medical records and pathology reports following  
136 colonoscopy. Behavioral and lifestyle information were obtained from a computer-assisted  
137 personal interview (CAPI) based on the NCI Colon Cancer Familial Registry Risk Factor  
138 Questionnaire<sup>30</sup>. The CAPI was administered by trained interviewer. Dietary fructose intake was  
139 estimated based on a self-administered food frequency questionnaire (FFQ).

140 *Sample collection and pre-processing*

141 This study was approved by Case Western Reserve University/UHCMC and the University  
142 of Virginia Institutional Review Boards (IRB-HSR-21219). This study followed Strengthening the  
143 Reporting of Observational Studies in Epidemiology guidelines. The data generated exist as  
144 publicly available data on Gene Expression Omnibus: GSE151732. Pre-processing was carried  
145 out using the R package, *minfi*<sup>31</sup>. Full details can be found in Devall et al., (2021)<sup>29</sup>. Preprocessing  
146 and analysis of three cancer cohorts was carried out as performed previously<sup>32</sup>.

147

148 *Preprocessing and statistical analysis of DNAm data*

149 For our analysis of fructose, we accounted for the effects of technical variation on the array  
150 while retaining biological variation attributed to fructose by adjusting for both sentrix ID and  
151 position using the `champ.runCombat()` function of the R package, *ChAMP*<sup>33, 34</sup>. Beta values were  
152 logistically transformed prior to correction and were then returned for downstream regression  
153 analysis. For the analysis of fructose on DNAm, the following model was built:

154  $\text{DNAm} \sim \text{sex} + \text{polyp} + \text{BMI} + \text{smoking status} + \text{age} + \text{fructose}$

155 Sex = biological sex; polyp = polyp detected at any site along the colorectum; BMI = log-  
156 transformed BMI; smoking status = never/former/current; age = age at sample index; fructose =  
157 log-transformed measure of total dietary fructose intake.

158 Differentially methylated region (DMR) analysis was carried out using *DMRcate*<sup>35</sup>. The  
159 `cpg.annotate()` function was applied with the following arguments: `what = "B"`, `fdr = 0.1`, `arraytype =`  
160 `"EPIC"` and `analysis.type = "differential"`. Following annotation, the `dmrcate()` function was applied  
161 using the following arguments: `C=4`, `lamda=1000`, `nCpG=3`. Measures of false discovery rates  
162 (FDR) were calculated by applying the `p.adjust()` function to the resulting Stouffer P values for  
163 each region, as previously<sup>32, 36</sup>. To identify DMRs of a similar minimum size, these same  
164 parameters were then applied to our repeated-measures analysis of three CRC datasets<sup>25, 37, 38</sup>,  
165 where we aimed to identified differences in DNAm between CRC tumor and matched, normal-

166 adjacent tissue (NAT). Only one exception to our method was considered for this analysis: C=2  
167 (default value for 450k data) when calling the `dmrcate()` function. Overlapping regions across  
168 datasets were defined using the `in.region()` and `bedr.join.region()` functions of the *bedr* package<sup>39</sup>.

169

#### 170 *Statistical analysis of RNA-seq data*

171 Bulk RNA-seq data was downloaded for The Cancer Genome Atlas Colon  
172 Adenocarcinoma (TCGA-COAD) dataset from *TCGAbiolinks*<sup>40</sup>. To identify differences in gene  
173 expression between matched (n=41 pairs) CRC tumors and normal adjacent tissues (NATs), we  
174 performed a repeated measures regression using *DESeq2*<sup>41</sup>.

175

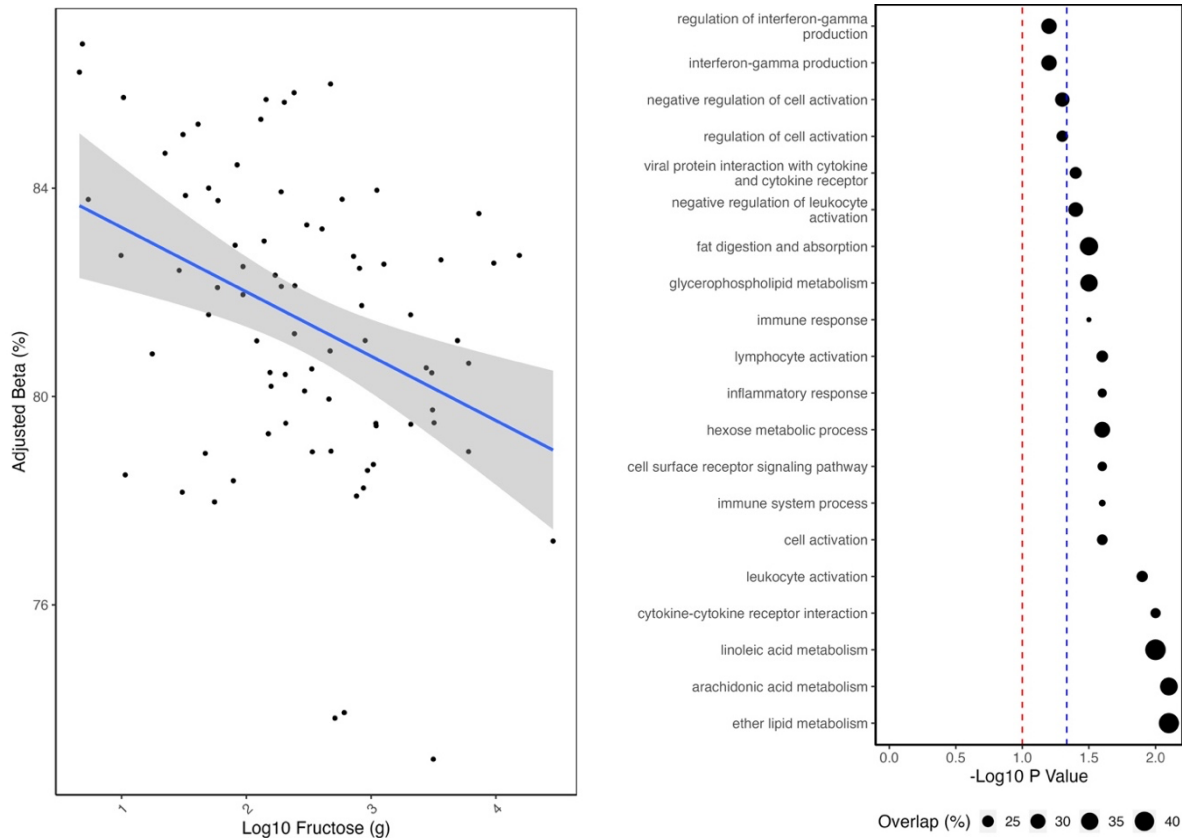
## 176 **Results**

177

### 178 **Dietary fructose consumption is associated with widespread DNAm differences in AA right** 179 **colon.**

180 We used *DMRcate*<sup>35</sup> to identify significant regions of DNAm that correlated with total  
181 dietary fructose consumption. Given the relatively limited sample size of EA individuals who  
182 provided data on dietary fructose consumption (n=34), we restricted our analysis AAs (n=79). In  
183 right colon, 4,263 DMRs survived FDR corrections. This included a hypomethylated DMR at  
184 chr1:9130738-9131854 (FDR=4.09E<sup>-03</sup>), a genomic region near the transcriptional start site of  
185 *SLC2A5* (**Figure 1A**). DNA hypomethylation of chr11:18415641-18415964, which corresponds to  
186 lactate dehydrogenase A (*LDHA*, FDR=0.036) was also identified. Increased *LDHA* expression  
187 was previously identified in the intestinal epithelium of HFCS-treated mice<sup>15</sup>. To better determine  
188 the coordinated effects of dietary fructose on DNAm, we performed pathway enrichment analysis  
189 using *missMethyl*<sup>42</sup>. Here, we identified enrichments for terms previously associated with fructose,  
190 such as ether lipid metabolism (P=9.61E<sup>-05</sup>)<sup>43</sup>, arachidonic acid metabolism (P=1.56E<sup>-04</sup>)<sup>44</sup> and  
191 linoleic acid metabolism (P=4.75E<sup>-04</sup>)<sup>44</sup> (**Figure1B**). Other notable enrichments include the

192 glycolytic process through fructose-6-phosphate ( $P=7.44E^{-03}$ ) and phosphofructokinase activity  
 193 ( $P=9.53E^{-03}$ ), developmental process ( $P=6.20E^{-04}$ ), long-chain fatty acid metabolic process  
 194 ( $P=5.45E^{-03}$ ) and cell population proliferation ( $P=9.10E^{-03}$ ).



195 **Figure 1: Summary of DMR analysis of dietary fructose in right AA colon. (A)** *SLC2A5* was  
 196 significantly hypomethylated in subject biopsies who consumed increasingly high amounts of  
 197 fructose. **(B)** Summary of the top 20 most significantly enriched terms within our DMR analysis.  
 198 “Overlap” corresponds to the number of DMRs identified within our analysis divided by the total  
 199 number of genes within that pathway. Red and blue dashed lines correspond to nominal and FDR  
 200 cut-offs, respectively.

201

202 **Dietary fructose consumption is associated with substantially fewer differences in**  
 203 **differential methylation in AA left colon.**

204 In contrast to the widespread differential DNAm found to be associated with dietary  
205 fructose in the right colon, an analysis of matched, left colon biopsies revealed few associations.  
206 In total, 24 DMRs were associated with fructose in the left colon (FDR<0.05). Of the 85 nominal  
207 DMRs (P<0.05) associated with dietary fructose in the left colon, only 25 overlapped regions also  
208 found in matched, right colon. All 25 DMRs were consistent for direction of effect (**Table 1**).

209

### 210 **Contextualization of fructose-related findings in the context of CRC risk.**

211 We hypothesized that, if the fructose-related differences identified in our analysis  
212 contributed to CRC risk, then DMRs hypermethylated in response to increasing fructose would  
213 also be hypermethylated in CRC tumors. To address this hypothesis, we overlaid our findings  
214 with data obtained from a reanalysis of three CRC datasets<sup>25, 37, 38</sup>. Of the 4,263 right-side  
215 fructose-DMRs identified, 2,025 overlapped regions identified in at least one CRC tumor analysis  
216 and demonstrated the expected concordance for direction of effect. This included 1,373 DMRs  
217 that were present in all three CRC tumor analyses. In left colon, 28 (21: P<0.05, 7: FDR<0.05)  
218 fructose-related DMRs overlapped regions associated with CRC. Notable, nominal fructose risk  
219 DMRs of the left colon include chr13:28544471-28544820 (caudal type homeobox 2 (*CDX2*)),  
220 which was hypermethylated in both TCGA-COAD and ColoCare tumors and chr16:68676451-  
221 68676806 (cadherin 3 (*CDH3*)), which was hypermethylated in tumors of all three datasets. Of  
222 the 25 DMRs identified in both right and left colon, six were also identified in at least one CRC  
223 tumor analysis (**Table 1**, bold font).

224

225

226

227

228



229  
230  
231  
232

**Table 1: Summary of association findings through DMR analysis of fructose in right and left colon of AA.** Findings were ordered from most to least significant based on the sum of their ranked significance. Bold font represents regions that were also identified in at least one CRC tumor analyses and displayed a concordant direction of effect.

Gene	Right Colon				Left Colon			
	Chromosomal Region	Beta Diff (%)	P	FDR	Chromosomal Region	Beta Diff (%)	P	FDR
<b>NBEA, MAB21L1</b>	<b>13:36050522-36051073</b>	<b>-0.046</b>	<b>1.16E<sup>-03</sup></b>	<b>0.019</b>	<b>13:36050158-36051073</b>	<b>-0.040</b>	<b>1.22E<sup>-09</sup></b>	<b>5.94E<sup>-07</sup></b>
<b>ZFP30</b>	<b>19:38145827-38147309</b>	<b>0.005</b>	<b>7.93E<sup>-06</sup></b>	<b>7.81E<sup>-04</sup></b>	<b>19:38145827-38147309</b>	<b>0.006</b>	<b>1.67E<sup>-06</sup></b>	<b>1.24E<sup>-04</sup></b>
N/A	11:128693801-128694679	-0.023	1.55E <sup>-03</sup>	0.023	11:128693677-128694679	-0.029	4.48E <sup>-09</sup>	1.56E <sup>-06</sup>
CRISP2	6:49681178-49681851	-0.028	5.54E <sup>-03</sup>	0.050	6:49681178-49681851	-0.022	1.36E <sup>-07</sup>	1.92E <sup>-05</sup>
FANCD2OS	3:10149882-10150487	-0.009	1.77E <sup>-03</sup>	0.025	3:10149466-10150487	-0.007	3.53E <sup>-06</sup>	2.14E <sup>-04</sup>
MCPH1, ANGPT2	8:6419438-6419570	-0.049	6.08E <sup>-03</sup>	0.050	8:6418978-6420963	-0.024	4.46E <sup>-07</sup>	5.34E <sup>-05</sup>
CPT1B, CHKB-CPT1B, CHKB	22:51016386-51017432	-0.018	0.025	0.092	22:51016386-51017432	-0.017	3.48E <sup>-10</sup>	2.12E <sup>-07</sup>
TLE3	15:70371951-70372127	-0.020	2.78E <sup>-03</sup>	0.033	15:70371951-70373099	-0.010	3.33E <sup>-04</sup>	4.02E <sup>-03</sup>
N/A	1:247537159-247537369	-0.037	0.013	0.079	1:247537159-247537369	-0.039	2.09E <sup>-05</sup>	7.32E <sup>-04</sup>
ACSS3	12:81471311-81472177	-0.025	5.20E <sup>-04</sup>	0.015	12:81470331-81472177	-0.015	7.26E <sup>-03</sup>	0.022
RP4-665J23.1	1:91300215-91300559	-0.022	5.51E <sup>-04</sup>	0.015	1:91300215-91300559	-0.015	5.42E <sup>-03</sup>	0.019
N/A	12:131865279-131865659	-0.020	9.29E <sup>-03</sup>	0.065	12:131864810-131865659	-0.016	4.02E <sup>-04</sup>	4.46E <sup>-03</sup>
<b>N/A</b>	<b>2:660692-660856</b>	<b>-0.020</b>	<b>0.023</b>	<b>0.092</b>	<b>2:660363-660856</b>	<b>-0.016</b>	<b>2.15E<sup>-04</sup></b>	<b>3.20E<sup>-03</sup></b>
HS3ST3B1	17:14213579-14214081	-0.025	0.024	0.092	17:14213225-14214081	-0.021	4.25E <sup>-05</sup>	1.18E <sup>-03</sup>
DPP10	2:115419846-115420260	-0.018	0.025	0.092	2:115419729-115420260	-0.017	4.03E <sup>-05</sup>	1.15E <sup>-03</sup>
AC093627.9	7:142966-143220	-0.015	0.026	0.092	7:142411-143220	-0.014	2.36E <sup>-04</sup>	3.36E <sup>-03</sup>
ALS2CR11	2:202483740-202484320	-0.015	0.049	0.115	2:202483740-202484583	-0.013	4.37E <sup>-06</sup>	2.51E <sup>-04</sup>
<b>N/A</b>	<b>14:73392078-73392648</b>	<b>-0.013</b>	<b>0.019</b>	<b>0.089</b>	<b>14:73392078-73392648</b>	<b>-0.006</b>	<b>4.19E<sup>-03</sup></b>	<b>0.016</b>
<b>AL031663.1</b>	<b>20:44097309-44097731</b>	<b>-0.028</b>	<b>9.29E<sup>-03</sup></b>	<b>0.065</b>	<b>20:44097309-44097563</b>	<b>-0.008</b>	<b>0.013</b>	<b>0.031</b>
N/A	13:112212213-112212266	0.018	0.030	0.098	13:112212213-112212266	0.015	3.82E <sup>-04</sup>	4.37E <sup>-03</sup>
RP11-23P13.6	15:42187036-42187219	0.022	0.019	0.089	15:42186830-42187219	0.014	0.013	0.031
<b>N/A</b>	<b>14:52535758-52536436</b>	<b>0.019</b>	<b>0.034</b>	<b>0.100</b>	<b>14:52535758-52536721</b>	<b>0.014</b>	<b>5.63E<sup>-04</sup></b>	<b>5.54E<sup>-03</sup></b>
DTNA	18:32173055-32173206	-0.017	0.046	0.114	18:32173055-32173227	-0.017	4.50E <sup>-04</sup>	4.82E <sup>-03</sup>
WT1	11:32421514-32421845	-0.020	0.032	0.099	11:32421514-32421845	-0.015	2.53E <sup>-03</sup>	0.013
N/A	10:131193104-131193272	-0.019	0.050	0.115	10:131193104-131193272	-0.014	2.19E <sup>-03</sup>	0.012

233 **Defining DMR targets through analysis of TCGA-COAD.**

234           Given that some DMRs occur within intergenic regions or map near to multiple genes, we  
235 aimed to better define the likely targets of each fructose- and CRC-related DMR (fructose risk  
236 DMRs) by determining if their corresponding genes were found to be differentially expressed in  
237 CRC. Of the three cohorts considered in this analysis, only TCGA-COAD had matched RNA-seq  
238 data. A repeated measurement regression model was fitted for each gene to identify differences  
239 between matched tumor versus NAT across 41 pairs present within this dataset. Of the 1,581  
240 fructose-related DMRs that were present in at least TCGA-COAD, 869 corresponded to  
241 differentially expressed genes (DEGs). Fisher's test confirmed a highly significant enrichment for  
242 the differential expression of fructose risk DMR genes in TCGA-COAD ( $P=4.06E^{-73}$ ). In left colon,  
243 seven of 16 fructose risk DMRs with gene annotations corresponded to CRC DEGs (**Figure 2**).

244

245

246

247

248

249

250

251

252

253

254

255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270



271 **Figure 2: Volcano plot demonstrating the significance of genes corresponding to fructose-**  
272 **risk DMRs in TCGA-COAD expression data.** Genes were color coded depending upon whether  
273 they were identified as fructose-risk DMRs in right (blue), left (red) or purple (both) analyses.  
274 Positive log<sub>2</sub> fold changes correspond to DEGs overexpressed in CRC tumors versus matched  
275 NAT.

276  
277 **Discussion**

278 Our study represents the first, relatively large-scale interrogation of the association  
279 between DNAm and dietary fructose intake in normal colon tissues of AA patients free of CRC. It  
280 presents novel targets through which increased fructose consumption may drive CRC-related

281 signaling. In recent years, there has been a growing interest in the use of dietary interventions to  
282 support cancer therapeutics<sup>45</sup>. Tumors frequently develop while retaining the metabolic pathways  
283 of their tissue of origin<sup>15</sup>. Cells that utilize the energy availability of their environment can display  
284 accelerated growth and outcompete their neighbors. Understanding the molecular drivers of these  
285 advantages in healthy colon tissue may therefore provide important mechanistic insight into the  
286 drivers of dietary risk genes. Such genes may not only provide biomarkers for risk, especially  
287 within individuals who consume high quantities of fructose, but may also provide novel targets for  
288 future therapeutics by focusing on genes driving preferential tumor-associated  
289 microenvironments.

290 Mouse studies have shown that intestinal absorption capacity is reached following  
291 moderate consumption of fructose (1g/kg), leading to unaltered fructose entering the colon<sup>46</sup>. Two  
292 landmark mouse model studies demonstrated a role for HFCS in colon carcinogenesis that was  
293 determined to be independent of its effect on body weight or the development of metabolic  
294 syndrome<sup>14, 15</sup>. They first show that HFCS enhances intestinal tumorigenesis through aberrant  
295 activation of glycolysis and promotion of fatty acid synthesis. Importantly, this study identified *KHK*  
296 as a key gene mediating this process. Later, Taylor et al., demonstrated that feeding mice a HFCS  
297 diet led to increased expression of hypoxia-inducible factor 1-alpha (HIF-1 $\alpha$ ), *KHK* and *SLC2A5*,  
298 which was accompanied by longer villi in both the proximal jejunum and the duodenum<sup>15</sup>. Further  
299 analysis in human CRC cell lines revealed that the metabolism of fructose by *KHK* led to the  
300 formation of fructose 1-phosphate, which promoted cell survival through the inhibition of another  
301 pyruvate kinase M2 (*PKM2*), thus contributing to cell survival under hypoxic conditions. These  
302 important studies were among the first to define a role for fructose/SSBs in CRC tumorigenesis.  
303 In the present study, our analysis of fructose risk DMRs did not identify any regions corresponding  
304 to *KHK*. However, we identified some overlap with the genes and pathways identified by both  
305 mouse studies. In this way, our human study provides partial support for previous findings in  
306 murine models, while also highlighting the imperative need for further mechanistic investigations

307 into the role of CRC risk mediated by fructose in healthy, human colon epithelium. Several  
308 glycolysis and fatty acid terms were among the most significantly enriched in our analysis of  
309 fructose in right colon, and many of the DMRs that contributed to those terms, such as *CBFA2T3*,  
310 *HDAC4*, *HTRA2*, *PGAM2* (glycolysis) and *ACSF2*, *CYP2E1*, *PGIF*, *PLA2G4F*, *SLC27A5* and  
311 *TNXB* (fatty acid) overlapped regions identified in CRC tumors. Tumor cells frequently retain the  
312 metabolic pathways of their tissue of origin<sup>15</sup> and the process of glycolysis provides an ample  
313 carbon source for fatty acid synthesis, which are required to house energy stores for the excessive  
314 energy demands of cancer proliferation<sup>14</sup>.

315 SSBs represent the leading source of added sugars in the US<sup>47</sup> and added sugars account  
316 for over 25% of total calories for one in six AAs, in contrast to one in eleven EAs<sup>48</sup>, indicating that  
317 increased levels of dietary fructose consumption, particularly through SSB, may play a role.  
318 Importantly, AAs are 30% more likely to develop right- than left-sided colon cancer<sup>21</sup>, and are also  
319 more likely to develop right-sided colon cancer than EAs<sup>49</sup>. Differences in health care and dietary  
320 patterns driven by poorer socioeconomic status have been considered as contributory factors to  
321 observed disparities<sup>22, 50</sup>. While the role of SSB in racial disparities for CRC has yet to be fully  
322 explored, a previous multi-center study of 33,106 participants from the Nurses' Health Study  
323 II found that SSB intake in adolescence was associated with increased adulthood adenoma  
324 incidence in proximal (right) colon and rectum. The same study also found that total fructose  
325 intake was only associated with increased risk for distal (left) and rectal adenomas<sup>17</sup>. However, a  
326 larger combined study of the Nurses' Health Study and Health Professionals Follow-Up Study  
327 revealed that both SSB and total fructose were only associated with proximal/right-sided colon  
328 cancer<sup>18</sup>. Given that SSBs represent more of the daily caloric intake of AAs than EAs, they  
329 represent one potential driver of this racial disparities in CRC sidedness that warrants further  
330 investigation.

331 In line with findings showing a greater risk for dietary fructose in right than left colon cancer,  
332 our study identified a far greater number of differences in DNAm in right than matched left AA

333 colon. Further, many of those DMRs identified exclusively in right colon (2,019 of 2,025) were  
334 also found in CRC tumors. Excess fructose enters the right colon following the saturation of  
335 SLC2A5 in the small intestine. Many of the DMRs identified in right colon were also differentially  
336 expressed in CRC tumors versus NAT and a subset (list subset) were also validated through a  
337 controlled dosing experiment of normal colon organoids. Our findings therefore imply a  
338 coordinated mechanism through which dietary fructose alters DNAm at regions relevant to CRC,  
339 leading to DEGs that modulate disease risk. The amount of fructose and its direct metabolites  
340 that are present in the left colon following further absorption to the hepatic portal system remains  
341 unclear. It is worth noting that while our study identified many more fructose-risk DMRs in AA right  
342 than left colon, a subset of fructose-risk DMRs exclusive to left colon were also found, including  
343 a hypermethylated DMR corresponding to the intestine-specific transcription factor, *CDX2*. A  
344 reduction in this tumor suppressor was also identified between CRC tumors and NATs in our  
345 analysis of TCGA-COAD. *CDX2* plays important roles in the regulation of epithelial cell  
346 differentiation and the reduction of its expression has been previously associated with poorer  
347 survival outcomes in left-sided colon cancer<sup>51</sup>. While the data presented within this study provides  
348 potential mechanisms that may drive the observed increase in right-sided colon cancer risk, larger  
349 epidemiological studies are needed to confirm this through finer site-stratified analyses. Given the  
350 notable differences in consumption of SSBs between AA and EA, and the relatively higher risk of  
351 AAs to develop right-sided colon cancer, these studies should also consider the effects of racial  
352 ancestry and ethnicity.

353 To relate our findings to CRC risk, we made use of publicly available CRC data and overlaid  
354 findings with those found to occur in CRC tumors. However, there are a number of limitations to  
355 this approach. First, CRC is a highly heterogenous disease and as such, it stands to reason that  
356 CRC risk pathways are also diverse. Second, total dietary fructose was highly correlated with  
357 other dietary sugars. Despite this, our results replicated multiple previously identified pathways  
358 and targets that are specific to fructose, including *SLC2A5*. However, validation through the

359 coupling of high-throughput omics readouts and controlled dosing within the colon organoid  
360 system may further reveal mechanistic insight into the role of fructose in CRC. Nonetheless, our  
361 findings highlight several plausible targets through which dietary fructose may influence DNAm to  
362 modulate CRC risk and some of these targets were verified through fructose exposure to human  
363 AA colon organoids.

364

## 365 **Conclusion**

366 There is increasing evidence for a role of dietary fructose and HFCS/SSB consumption in  
367 CRC risk. Here, we provide novel insight into colon location-specific mechanisms in AA subjects  
368 through which this risk may develop. That fructose consumption is associated with a far greater  
369 number of potential CRC risk DMRs in right than matched, left colon is intriguing. AAs are  
370 significantly more likely than EAs to consume greater quantities of fructose and to develop right-  
371 side colon cancer. Our findings highlight potential mechanisms through which increased dietary  
372 fructose consumption may differentially impact DNAm at regions associated with CRC to alter  
373 risk. Our findings provide some mechanistic support for reducing dietary fructose consumption.

374

## 375 **References**

- 376 1. United States Department of Agriculture Economic Research Service. U.S. Consumption  
377 of Caloric Sweeteners. USDA Sugar Supply. Volume Table 50, 2022.
- 378 2. DeChristopher LR, Uribarri J, Tucker KL. Intake of high fructose corn syrup sweetened  
379 soft drinks, fruit drinks and apple juice is associated with prevalent coronary heart  
380 disease, in U.S. adults, ages 45-59 y. *BMC Nutr* 2017;3:51.
- 381 3. Vartanian LR, Schwartz MB, Brownell KD. Effects of soft drink consumption on nutrition  
382 and health: a systematic review and meta-analysis. *Am J Public Health* 2007;97:667-75.
- 383 4. Vos MB, Kimmons JE, Gillespie C, et al. Dietary fructose consumption among US children  
384 and adults: the Third National Health and Nutrition Examination Survey. *Medscape J*  
385 *Med* 2008;10:160.
- 386 5. Keller A, O'Reilly EJ, Malik V, et al. Substitution of sugar-sweetened beverages for other  
387 beverages and the risk of developing coronary heart disease: Results from the Harvard  
388 Pooling Project of Diet and Coronary Disease. *Prev Med* 2020;131:105970.
- 389 6. Drewnowski A, Rehm CD. Consumption of added sugars among US children and adults  
390 by food purchase location and food source. *Am J Clin Nutr* 2014;100:901-7.

- 391 7. Lasater G, Piernas C, Popkin BM. Beverage patterns and trends among school-aged  
392 children in the US, 1989-2008. *Nutr J* 2011;10:103.
- 393 8. Lundeen EA, Park S, Pan L, et al. Daily Intake of Sugar-Sweetened Beverages Among US  
394 Adults in 9 States, by State and Sociodemographic and Behavioral Characteristics, 2016.  
395 *Prev Chronic Dis* 2018;15:E154.
- 396 9. Bleich SN, Vercammen KA, Koma JW, et al. Trends in Beverage Consumption Among  
397 Children and Adults, 2003-2014. *Obesity (Silver Spring, Md.)* 2018;26:432-441.
- 398 10. Rosinger A, Herrick K, Gahche J, et al. Sugar-sweetened Beverage Consumption Among  
399 U.S. Adults, 2011–2014. *NCHS Data Brief* 2017;270:1-8.
- 400 11. Gross LS, Li L, Ford ES, et al. Increased consumption of refined carbohydrates and the  
401 epidemic of type 2 diabetes in the United States: an ecologic assessment. *Am J Clin Nutr*  
402 2004;79:774-9.
- 403 12. Sun T, Zhang Y, Ding L, et al. The Relationship Between Major Food Sources of Fructose  
404 and Cardiovascular Outcomes: A Systematic Review and Dose-Response Meta-Analysis  
405 of Prospective Studies. *Advances in Nutrition* 2023;14:256-269.
- 406 13. Makarem N, Bandera EV, Lin Y, et al. Consumption of Sugars, Sugary Foods, and Sugary  
407 Beverages in Relation to Adiposity-Related Cancer Risk in the Framingham Offspring  
408 Cohort (1991-2013). *Cancer Prev Res (Phila)* 2018;11:347-358.
- 409 14. Goncalves MD, Lu C, Tutnauer J, et al. High-fructose corn syrup enhances intestinal  
410 tumor growth in mice. *Science* 2019;363:1345-1349.
- 411 15. Taylor SR, Ramsamoj S, Liang RJ, et al. Dietary fructose improves intestinal cell survival  
412 and nutrient absorption. *Nature* 2021;597:263-267.
- 413 16. Hur J, Otegbeye E, Joh HK, et al. Sugar-sweetened beverage intake in adulthood and  
414 adolescence and risk of early-onset colorectal cancer among women. *Gut* 2021;70:2330-  
415 2336.
- 416 17. Joh HK, Lee DH, Hur J, et al. Simple Sugar and Sugar-Sweetened Beverage Intake During  
417 Adolescence and Risk of Colorectal Cancer Precursors. *Gastroenterology* 2021;161:128-  
418 142 e20.
- 419 18. Yuan C, Joh HK, Wang QL, et al. Sugar-sweetened beverage and sugar consumption and  
420 colorectal cancer incidence and mortality according to anatomic subsite. *Am J Clin Nutr*  
421 2022;115:1481-1489.
- 422 19. Jonoush M, Fathi S, Hassanpour Ardekanizadeh N, et al. The Association Between  
423 Different Types of Dietary Carbohydrates and Colorectal Cancer: A Case-Control Study.  
424 *Front Nutr* 2022;9:898337.
- 425 20. Siegel RL, Wagle NS, Cercek A, et al. Colorectal cancer statistics, 2023. *CA: A Cancer*  
426 *Journal for Clinicians* 2023;73:233-254.
- 427 21. Giaquinto AN, Miller KD, Tossas KY, et al. Cancer statistics for African American/Black  
428 People 2022. *CA Cancer J Clin* 2022;72:202-229.
- 429 22. Carethers JM. Racial and ethnic disparities in colorectal cancer incidence and mortality.  
430 *Adv Cancer Res* 2021;151:197-229.
- 431 23. Mitchell C, Schnepfer LM, Notterman DA. DNA methylation, early life environment, and  
432 health outcomes. *Pediatr Res* 2016;79:212-9.



- 433 24. Fan J, Li J, Guo S, et al. Genome-wide DNA methylation profiles of low- and high-grade  
434 adenoma reveals potential biomarkers for early detection of colorectal carcinoma.  
435 *Clinical Epigenetics* 2020;12:56.
- 436 25. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and  
437 rectal cancer. *Nature* 2012;487:330-7.
- 438 26. Kageyama I, Yamada H, Munetsuna E, et al. Differential effects of excess high-fructose  
439 corn syrup on the DNA methylation of hippocampal neurotrophic factor in childhood  
440 and adolescence. *PLoS One* 2022;17:e0270144.
- 441 27. Ohashi K, Munetsuna E, Yamada H, et al. High fructose consumption induces DNA  
442 methylation at PPARalpha and CPT1A promoter regions in the rat liver. *Biochem Biophys*  
443 *Res Commun* 2015;468:185-9.
- 444 28. Tini G, Varma V, Lombardo R, et al. DNA methylation during human adipogenesis and  
445 the impact of fructose. *Genes Nutr* 2020;15:21.
- 446 29. Devall M, Sun X, Yuan F, et al. Racial Disparities in Epigenetic Aging of the Right vs Left  
447 Colon. *J Natl Cancer Inst* 2021;113:1779-1782.
- 448 30. Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an  
449 international resource for studies of the genetic epidemiology of colon cancer. *Cancer*  
450 *Epidemiol Biomarkers Prev* 2007;16:2331-43.
- 451 31. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive  
452 Bioconductor package for the analysis of Infinium DNA methylation microarrays.  
453 *Bioinformatics* 2014;30:1363-9.
- 454 32. Devall MA, Eaton S, Ali MW, et al. DNA methylation analysis of normal colon organoids  
455 from familial adenomatous polyposis patients reveals novel insight into colon cancer  
456 development. *Clin Epigenetics* 2022;14:104.
- 457 33. Morris TJ, Butcher LM, Feber A, et al. ChAMP: 450k Chip Analysis Methylation Pipeline.  
458 *Bioinformatics* 2014;30:428-30.
- 459 34. Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for  
460 Illumina BeadChips. *Bioinformatics* 2017;33:3982-3984.
- 461 35. Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially  
462 methylated regions in the human genome. *Epigenetics Chromatin* 2015;8:6.
- 463 36. Devall M, Ali MW, Eaton S, et al. Multi-omic analysis in normal colon organoids  
464 highlights MSH4 as a novel marker of defective mismatch repair in Lynch syndrome and  
465 microsatellite instability. *Cancer Med* 2023.
- 466 37. Baharudin R, Ishak M, Muhamad Yusof A, et al. Epigenome-Wide DNA Methylation  
467 Profiling in Colorectal Cancer and Normal Adjacent Colon Using Infinium Human  
468 Methylation 450K. *Diagnostics (Basel)* 2022;12.
- 469 38. Barrow TM, Klett H, Toth R, et al. Smoking is associated with hypermethylation of the  
470 APC 1A promoter in colorectal cancer: the ColoCare Study. *J Pathol* 2017;243:366-375.
- 471 39. Haider S, Waggott D, Lalonde E, et al. A bedr way of genomic interval processing. *Source*  
472 *Code for Biology and Medicine* 2016;11:14.
- 473 40. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for  
474 integrative analysis of TCGA data. *Nucleic Acids Res* 2016;44:e71.
- 475 41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for  
476 RNA-seq data with DESeq2. *Genome Biology* 2014;15:550.

- 477 42. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from  
478 Illumina's HumanMethylation450 platform. *Bioinformatics* 2016;32:286-8.
- 479 43. Mori T, Kondo H, Hase T, et al. Dietary phospholipids ameliorate fructose-induced  
480 hepatic lipid and metabolic abnormalities in rats. *J Nutr* 2011;141:2003-9.
- 481 44. Tranchida F, Tchiakpe L, Rakotoniaina Z, et al. Long-term high fructose and saturated fat  
482 diet affects plasma fatty acid profile in rats. *J Zhejiang Univ Sci B* 2012;13:307-17.
- 483 45. Lien EC, Vander Heiden MG. A framework for examining how diet impacts tumour  
484 metabolism. *Nature Reviews Cancer* 2019;19:651-661.
- 485 46. Jang C, Hui S, Lu W, et al. The Small Intestine Converts Dietary Fructose into Glucose and  
486 Organic Acids. *Cell Metab* 2018;27:351-361 e3.
- 487 47. U.S. Department of Health and Human Services and U.S. Department of Agriculture.  
488 2015–2020 Dietary Guidelines for Americans. 8th ed, 2015.
- 489 48. Yang Q, Zhang Z, Gregg EW, et al. Added sugar intake and cardiovascular diseases  
490 mortality among US adults. *JAMA Intern Med* 2014;174:516-24.
- 491 49. Carethers JM. Risk factors for colon location of cancer. *Transl Gastroenterol Hepatol*  
492 2018;3:76.
- 493 50. Carethers JM, Doubeni CA. Causes of Socioeconomic Disparities in Colorectal Cancer and  
494 Intervention Framework and Strategies. *Gastroenterology* 2020;158:354-367.
- 495 51. Konukiewitz B, Schmitt M, Silva M, et al. Loss of CDX2 in colorectal cancer is associated  
496 with histopathologic subtypes and microsatellite instability but is prognostically inferior  
497 to hematoxylin–eosin-based morphologic parameters from the WHO classification.  
498 *British Journal of Cancer* 2021;125:1632-1646.  
499