Plasma metabolomic profiles for colorectal cancer precursors in women

Dong Hang[1,2], **Oana A. Zeleznik**[3], **Jiayi Lu**[1], **Amit D. Joshi**[4], **Kana Wu**[2], **Zhibin Hu**[1], **Hongbing Shen**[1], **Clary B. Clish**[5], **Liming Liang**[6,7], **A. Heather Eliassen**[3,6], **Shuji Ogino**[5,6,8,9], **Jeffrey A. Meyerhardt**[10], **Andrew T. Chan**[3,4,5], **Mingyang Song**[2,4,6]

[1]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, China;

[2]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA;

[3]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA;

[4]Clinical and Translational Epidemiology Unit and Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA;

[5]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA, USA;

[6]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA;

[7]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA;

[8]Cancer Immunology Program, Dana-Farber Harvard Cancer Center, Boston, MA, USA;

[9]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA;

[10]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA.

Abstract

How metabolome changes influence the early process of colorectal cancer (CRC) development remains unknown. We conducted a 1:2 matched nested case-control study to examine the associations of pre-diagnostic plasma metabolome (profiled using LC-MS) with risk of CRC

Corresponding to: Mingyang Song, 667 Huntington Avenue, Kresge 906A, Boston, MA 02115. Tel: +01-617-643-4464, Fax: +01-617-432-2435, mingyangsong@mail.harvard.edu.

Conflicts of interest

The authors have nothing to disclose.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

precursors, including conventional adenomas (n=586 vs. 1141) and serrated polyps (n=509 vs. 993), in the Nurses' Health Study (NHS) and NHSII. Conditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI). We used the permutation-based Westfall and Young approach to account for multiple testing. Subgroup analyses were performed for advanced conventional adenomas (defined as at least one adenoma of 10 mm or with high-grade dysplasia, or tubulovillous or villous histology) and high-risk serrated polyps that were located in the proximal colon or with size of 10 mm. After multiple testing correction, among 207 metabolites, higher levels of C36:3 phosphatidylcholine (PC) plasmalogen were associated with lower risk of conventional adenomas, with the OR (95% CI) comparing the 90th to the 10th percentile of 0.62 (0.48–0.81); C54:8 triglyceride (TAG) was associated with higher risk of serrated polyps (OR=1.79, 95% CI: 1.31–2.43), and phenylacetylglutamine (PAG) was associated with lower risk (OR=0.57, 95% CI:0.43–0.77). PAG was also inversely associated with advanced adenomas (OR=0.57, 95% CI: 0.36–0.89) and high-risk serrated polyps (OR=0.54, 95% CI: 0.32–0.89), although the multiple testing-corrected p value was >0.05. Our findings suggest potential roles of lipid metabolism and phenylacetylglutamine, a microbial metabolite, in the early stage of colorectal carcinogenesis, particularly for the serrated pathway.

Keywords

colorectal adenomas; polyp; metabolism; colorectal cancer

Introduction

Colorectal cancer (CRC) is a heterogeneous disease with diverse molecular background and clinicopathological manifestations. Approximately 60–80% of CRC cases develop through the conventional adenoma-carcinoma sequence characterized by oncogene (e.g. *KRAS*) activation and tumor suppressor (e.g. *APC*, *SMAD4*, and *TP53*) inactivation [1], while another 20–30% arise from the serrated pathway that is associated with *BRAF* mutation, CpG island methylator phenotype, and microsatellite instability [2]. In contrast to conventional adenomas, serrated polyps have a predilection for the proximal colon and show predominately a sessile or flat morphology. It has been proposed that serrated polyps contribute disproportionately to the development of "interval cancers" that occur prior to next screening or surveillance interval after an initially negative colonoscopy, thus posing a challenge for clinical practice and CRC prevention [3].

Dysregulated metabolism is a hallmark of cancer and has therefore been the focus of many cancer studies [4]. Metabolomics is a useful tool to systematically probe changes in metabolism and identify novel biomarkers and targets for better prevention. To date, substantial efforts have been made to characterize the metabolomic changes associated with CRC [5]. However, most of the studies assessed metabolomics after CRC diagnosis [6–12] and are prone to reverse causation, because cancer itself can induce substantial changes in systemic metabolism. Moreover, to our knowledge, no study has yet examined pre-diagnostic metabolomic profile for CRC precursors, leading to a knowledge gap about the role of metabolites in the early stage of colorectal carcinogenesis and potential utility for early CRC detection.

To address this paucity, we conducted a 1:2 matched case-control study nested within two large prospective US cohorts to identify the metabolites associated with conventional adenomas and serrated polyps.

Methods

Study population

Participants were derived from the Nurses' Health Study (NHS) and Nurses' Health Study II (NHSII), two ongoing prospective cohorts of US female registered nurses. The NHS began in 1976, recruiting 121,700 female nurses aged 30–55 years, and the NHSII began in 1989, recruiting 116,429 female nurses aged 25–42 years. In both cohorts, all participants completed a questionnaire at baseline and were mailed follow-up questionnaires biennially to update their lifestyle and medical information, with cumulative follow-up rates greater than 90% [13].Blood samples were collected from 32,826 women in the NHS during 1989–1990 and 29,611 women in the NHSII during 1996–1999. Upon arrival via overnight shipping on ice packs, the samples were centrifuged, aliquoted, and stored in liquid nitrogen freezers immediately [14]. All archived blood samples in original vials were thawed once to create a subaliquot for metabolomics profiling. Demographic, dietary, and lifestyle profiles of women who provided blood samples were generally similar to those who did not [15].

Among participants with available metabolomic data from previous cross-sectional/nested case-control studies of various outcomes (see details in the Supplementary Figure 1), we excluded those who had a history of cancer (except non-melanoma skin cancer), colorectal polyps, or inflammatory bowel disease at the time of blood draw, or had no lower gastrointestinal endoscopy after blood donation. After exclusions, we identified 586 cases with conventional adenomas and 509 with serrated polyps that were detected after blood draw until the end of follow-up (June 1, 2012 for the NHS and June 1, 2011 for the NHSII). We randomly selected up to 2 controls for each case matched on age (within 1 year), time period of endoscopy (in 2-year intervals), fasting status (>8 hours or 8 hours), study cohort (NHS or NHSII), and race (Caucasian or non-Caucasian) from eligible participants who had undergone at least one endoscopy after blood draw and were free of any polyps at the time of diagnosis of the matched cases. A total of 1141 and 993 controls were included in the analysis for conventional adenomas and serrated polyps, respectively. The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required.

Metabolomics profiling

Plasma metabolomics data were generated at the Broad Institute using liquid chromatography tandem mass spectrometry (LC-MS), as described previously [16]. Briefly, a C8-positive platform, which connected a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.) to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific), was used to measure polar and non-polar lipids. A hydrophilic interaction liquid chromatography (HILIC-positive) platform composed of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific) was used to measure water-soluble metabolites. Raw data were processed

using TraceFinder 3.3 software (Thermo Fisher Scientific) and Progenesis QI (Nonlinear Dynamics; Newcastle upon Tyne, UK). Metabolite identities were confirmed using authentic reference standards or reference samples. For metabolites with missingness <25%, missing data were imputed using half of the minimum measured value per metabolite to avoid excluding metabolites with true missingness patterns (e.g., xenobiotics). We excluded metabolites not passing our pilot study investigating the effect of 24-h processing delay and within-person stability over 1–2 years [17]. Finally, a total of 207 known metabolites, that is, those with assigned identification numbers in the Human Metabolome Database (HMDB), were included in the current study (Supplementary Tables 1 and 2).

Covariate assessment

We used covariate data from the questionnaires administered closest in time to blood draw, including body mass index (BMI), smoking status, alcohol consumption, physical activity, family history of CRC, Alternate Healthy Eating Index (AHEI), regular use of aspirin, menopausal status, and postmenopausal hormone use. BMI was calculated as body weight in kilograms divided by height in meters squared. Physical activity was assessed by the product sum of the metabolic equivalent (MET) values of each specific recreational activity and hours spent on that activity per week [18]. Positive family history of CRC was defined as report of a CRC diagnosis in a parent or sibling. Regular aspirin use was defined as use of at least two standard-equivalent tablets per week [19]. Alcohol consumption and AHEI were derived from validated food frequency questionnaires. AHEI is a dietary score measuring the adherence to a dietary pattern characterized by foods and nutrients strongly associated with risk of chronic disease [20].

Ascertainment of colorectal polyps

Ascertainment of colorectal polyps has been described in detail previously [21]. Briefly, on each biennial questionnaire, participants were asked whether they had undergone a colonoscopy or sigmoidoscopy and whether any colorectal polyp had been diagnosed in the past two years. When a participant reported a polyp diagnosis, written consent was to obtain to review her endoscopic and pathologic records. Study physicians, who were unaware of exposure information, confirmed the diagnosis and extracted data on histology, size, number, and anatomic location of polyps. Conventional adenomas included tubular, tubulovillous and villous adenomas, and adenomas with high-grade dysplasia. Advanced adenomas were defined as at least one adenoma of 10 mm in diameter or with advanced histology (i.e., tubulovillous, villous or high-grade dysplasia feature). Serrated polyps comprised hyperplastic polyps, traditional serrated adenoma, and sessile serrated adenoma/polyp with or without cytological dysplasia. Given that large (10 mm) or proximal serrated polyps have been associated with higher risk for CRC [22, 23], we defined high-risk serrated polyps as those located in the proximal colon or with size of 10 mm.

Statistical analysis

To account for any batch effect and improve normality, metabolite measurements were natural-log transformed and standardized using z-scores (standard deviations from the mean) within each original sub-study.

Conditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI) of conventional adenomas and serrated polyps comparing the 90th to the 10th percentile of metabolite levels. All models were adjusted for age at blood draw, number of prior endoscopies, and time in years since the most recent endoscopy. Because further adjustment for family history of CRC, height, pack-years of smoking, AHEI score, BMI, physical activity, alcohol consumption, regular aspirin use, menopausal status, and postmenopausal hormone therapy did not essentially change the results (data not shown), we presented the results without these adjustments to preserve power.

We used a permutation test (N=5000) to control the family-wise error rate (i.e., multiple testing) and the step-down correction method of Westfall and Young to account for the correlation structure of metabolites [24]. Both the unadjusted and multiple comparison-adjusted p-values were reported, and we focused on the individual metabolites with an adjusted p-value of 0.05.

Metabolite set enrichment analysis [25] was conducted to identify classes of molecularly or biologically similar metabolites that were associated with conventional adenomas or serrated polyps, with false discovery rate (FDR) <0.05 considered statistically significant.

We also performed separate analyses for clinically significant polyp subgroups, including advanced conventional adenomas and high-risk serrated polyps. Moreover, we analyzed the association of metabolites with polyp risk according to the median time interval since blood draw (11 years) and calculated the p for heterogeneity using the case-only analyses [21]. A stratified analysis was performed according to fasting status (>8 hours or 8 hours). The p for interaction was calculated using a likelihood ratio test comparing the models with and without the product term between fasting status and each metabolite (continuous). We also conducted a sensitivity analysis among participants who were selected as controls in the source case-control studies.

All analyses were performed using R 3.2.5 (R Foundation for Statistical Computing, Vienna, Austria) and SAS version 9.4 (SAS Institute Inc, Cary, NC).

Results

In Table 1, compared to women without any polyps, those with conventional adenomas or serrated polyps were more likely to have a higher BMI, a family history of CRC, and a lower AHEI, and were less likely to be physically active.

Figure 1 shows the associations of 207 metabolites with conventional adenomas and serrated polyps. A total of 20 and 21 metabolites were nominally associated with conventional adenomas and serrated polyps, respectively. The ORs and 95% CI for all the metabolites are shown in Supplementary Tables 1 and 2.

After multiple testing correction, we found an inverse association between C36:3 phosphatidylcholine (PC) plasmalogen and conventional adenomas, with the OR (95% CI) comparing the 90th to the 10th percentile of 0.62 (0.48–0.81) (Table 2). In addition, C54:8 triglyceride (TAG) was associated with higher risk of serrated polyps with the OR (95% CI)

comparing the 90th to the 10th percentile of 1.79 (1.31–2.43), and phenylacetylglutamine (PAG) was associated with lower risk (OR=0.57, 95% CI:0.43–0.77) (Table 3). PAG also showed a nominal inverse association with advanced adenomas (OR=0.57, 95% CI: 0.36–0.89) and high-risk serrated polyps (OR=0.54, 95% CI: 0.32–0.89), although the multiple testing-corrected P value was >0.05 (Tables 2 and 3).

The sensitivity analysis among control participants of the original case-control studies showed similar results for C36:3 PC plasmalogen with conventional adenomas, and C54:8 TAG and PAG with serrated polyps (Supplementary Table 3).

In the stratified analysis by the median time interval since blood draw, no statistically significant heterogeneity was observed for C36:3 PC plasmalogen, C54:8 TAG, or PAG (Supplementary Tables 4 and 5). The observed associations for C36:3 PC plasmalogen, C54:8 TAG, and PAG were also similar regardless of fasting status (Supplementary Tables 6 and 7).

As shown in Figure 2, seven metabolite classes were associated with conventional adenomas at an FDR <0.05 (four positively: TAGs, lysophosphatidylcholines, diglycerides [DAGs], and ceramides; three negatively: PC plasmalogens, phosphatidylethanolamines [PEs], and PE plasmalogens); five metabolite classes were associated with serrated polyps (two positively: TAGs and DAGs; three negatively: nucleosides, nucleotides, and analogues, PC plasmalogens, PE plasmalogens).

Discussion

To our knowledge, the current study represents the first prospective analysis of metabolomics for CRC precursors. We identified C36:3 PC plasmalogen in relation to a lower risk of conventional adenomas, and C54:8 TAG and PAG in relation to higher and lower risks of serrated polyps, respectively. Furthermore, PAG showed a suggestive inverse association with advanced adenomas and high-risk serrated polyps. These findings open up new opportunities for assessing novel metabolite pathways involved in the early stage of CRC and the potential of these metabolites for improved polyp prediction.

As a colonic microbial metabolite, PAG results from glutamine conjugation of phenylacetic acid and is almost exclusively derived from bacterial phenylalanine metabolism [26]. Previous metabolomic studies have reported that circulating PAG levels are positively correlated with the gut microbiome diversity, as measured by the Shannon's diversity index [27, 28]. A lower microbiome diversity in the human intestines has been linked to increased risk of adenoma and CRC [29–31]. Furthermore, laboratory evidence shows that the human commensal *Clostridium sporogenes* can convert dietary phenylalanine to phenylpyruvic acid, which is metabolized in the liver to produce PAG [32]. Potential benefits have been observed from bacterial protein fermentation by *Clostridium sporogenes*. The bacteria could metabolize dietary tryptophan to produce indolepropionic acid, which was able to fortify the intestinal barrier via Pregnane X Receptor and Toll-like Receptor 4 pathways [33]. In the current study, we identified a novel association between PAG and serrated polyps, supporting the importance of gut metabolites in the early stage of CRC pathogenesis.

Future studies are warranted to explore the molecular mechanisms underlying the protective association between microbial PAG and serrated polyps.

Dyslipidemia has been recognized as an important risk factor of colorectal neoplasm [34]. Several studies have reported a positive association between circulating TAGs and risk of colorectal adenomas and CRC, although the findings remain inconsistent [35]. In the current study, we found that several TAGs were nominally associated with higher risks of conventional adenoma and serrated polyps, and the association between C54:8 TAG and serrated polyps reached the statistical significance after correction for multiple testing. Notably, the association between C54:8 TAG and serrated polyps was essentially unchanged after adjusting for various covariates including BMI, suggesting that C54:8 TAG may influence the serrated pathway independent of adiposity. Biologically, hypertriglyceridemia may contribute to colorectal carcinogenesis through insulin resistance, oxidative stress, and inflammation [36]. In addition, hypertriglyceridemia may increase the synthesis and secretion of bile acids [37], which may provide abundant substrate for formation of secondary bile acids and thereby promote carcinogenesis in the colorectum [38].

We also identified several PC plasmalogens in relation to lower risk of conventional adenomas and serrated polyps. Particularly, the association between C36:3 PC plasmalogen and conventional adenomas remained statistically significant after multiple testing correction. PC plasmalogens contain a vinyl-ether linked alkyl chain at the *sn-1* position of the glycerol backbone. The *sn-2* position is typically occupied by an ester-linked polyunsaturated fatty acid (e.g. arachidonic acid or docosahexanoic acid), and the polar head, in this case phosphocholine, is bound to the *sn-3* position [39]. Blood-borne plasmalogens are synthesized by the liver and then incorporated into nascent lipoproteins [40]. Because the vinyl-ether bond at the *sn-1* position can readily react with reactive oxygen species, PC plasmalogens are believed to protect mammalian cells against oxidative and inflammatory damages [39]. In addition, higher levels of PC species have been associated with healthy aging [41] and lower risk diabetes [42], supporting their role as a marker for higher antioxidant capacity. Despite these data, the precise mechanisms underlying the association between C36:3 PC plasmalogen and conventional adenomas require further investigation.

Our study has several strengths, including the prospective design, large sample size, detailed covariate data for confounding control, and untargeted metabolomics approach covering a wide range of metabolites. Moreover, we ascertained polyp diagnosis through medical record review and collected detailed histopathologic information of both conventional adenomas and serrated polyps. Several limitations should also be noted. First, the current study is observational and cannot rule out the possibility of unmeasured or residual confounding. Second, as is typical of previous studies, metabolomic profiling was conducted only once and may not reflect long-term exposures. However, we have shown that the included metabolites have a high within person stability over 1–2 years (e.g., intra-class correlation was 0.81 for C36:3 PC plasmalogen, 0.98 for C54:8 TAG, and 0.97 for PAG) [17]. Third, despite the large sample size, the statistical power remained low to identify metabolites with modest effects. Fourth, participants in our study were all women and

predominately Whites, limiting the generalizability of our findings. Further confirmation in men and other racial/ethnic groups is needed.

In conclusion, we identified an association between C36:3 PC plasmalogen and conventional adenomas, as well as associations of C54:8 TAG and PAG with serrated polyps, suggesting the role of lipid metabolism and gut microbial metabolites in the development of CRC precursors. Future studies are warranted to validate our findings and elucidate the underlying biological mechanisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Availability of data and materials

The NHS and NHSII datasets can be made available upon reasonable request for scientific collaborations. Further information including the procedures to obtain and access data from the Nurses' Health Studies is described at https://www.nurseshealthstudy.org/researchers (contact: nhsaccess@channing.harvard.edu).

References

1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759–67. [PubMed: 2188735]

2. Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol. 2011;42(1):1–10. doi:10.1016/j.humpath.2010.06.002 [PubMed: 20869746]

3. Sweetser S, Smyrk TC, Sinicrope FA. Serrated colon polyps as precursors to colorectal cancer. Clin Gastroenterol Hepatol. 2013;11(7):760–7; quiz e54–5. doi:10.1016/j.cgh.2012.12.004 [PubMed: 23267866]

4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74. doi:10.1016/j.cell.2011.02.013 [PubMed: 21376230]

5. Yu J, Zhao J, Zhang M, Guo J, Liu X, Liu L. Metabolomics studies in gastrointestinal cancer: a systematic review. Expert Rev Gastroenterol Hepatol. 2020;14(1):9–25. doi:10.1080/17474124.2020.1700112 [PubMed: 31786962]

6. Zhu J, Djukovic D, Deng L, Gu H, Himmati F, Chiorean EG, et al. Colorectal cancer detection using targeted serum metabolic profiling. J Proteome Res. 2014;13(9):4120–30. doi:10.1021/pr500494u [PubMed: 25126899]

7. Deng L, Gu H, Zhu J, Nagana Gowda GA, Djukovic D, Chiorean EG, et al. Combining NMR and LC/MS Using Backward Variable Elimination: Metabolomics Analysis of Colorectal Cancer, Polyps, and Healthy Controls. Anal Chem. 2016;88(16):7975–83. doi:10.1021/acs.analchem.6b00885 [PubMed: 27437783]

8. Farshidfar F, Weljie AM, Kopciuk KA, Hilsden R, McGregor SE, Buie WD, et al. A validated metabolomic signature for colorectal cancer: exploration of the clinical value of metabolomics. Br J Cancer. 2016;115(7):848–57. doi:10.1038/bjc.2016.243 [PubMed: 27560555]

9. Long Y, Sanchez-Espiridion B, Lin M, White L, Mishra L, Raju GS, et al. Global and targeted serum metabolic profiling of colorectal cancer progression. Cancer. 2017;123(20):4066–74. doi:10.1002/cncr.30829 [PubMed: 28640361]

10. Farshidfar F, Kopciuk KA, Hilsden R, McGregor SE, Mazurak VC, Buie WD, et al. A quantitative multimodal metabolomic assay for colorectal cancer. BMC Cancer. 2018;18(1):26. doi:10.1186/s12885-017-3923-z [PubMed: 29301511]

11. Barberini L, Restivo A, Noto A, Deidda S, Fattuoni C, Fanos V, et al. A gas chromatography-mass spectrometry (GC-MS) metabolomic approach in human colorectal cancer (CRC): the emerging role of monosaccharides and amino acids. Ann Transl Med. 2019;7(23):727. doi:10.21037/atm.2019.12.34 [PubMed: 32042743]

12. Geijsen A, Brezina S, Keski-Rahkonen P, Baierl A, Bachleitner-Hofmann T, Bergmann MM, et al. Plasma metabolites associated with colorectal cancer: A discovery-replication strategy. Int J Cancer. 2019;145(5):1221–31. doi:10.1002/ijc.32146 [PubMed: 30665271]

13. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, et al. Origin, Methods, and Evolution of the Three Nurses' Health Studies. Am J Public Health. 2016;106(9):1573–81. doi:10.2105/AJPH.2016.303338 [PubMed: 27459450]

14. Bertrand KA, Giovannucci E, Liu Y, Malspeis S, Eliassen AH, Wu K, et al. Determinants of plasma 25-hydroxyvitamin D and development of prediction models in three US cohorts. Br J Nutr. 2012;108(10):1889–96. doi:10.1017/S0007114511007409 [PubMed: 22264926]

15. Hunter DJ, Hankinson SE, Hough H, Gertig DM, Garcia-Closas M, Spiegelman D, et al. A prospective study of NAT2 acetylation genotype, cigarette smoking, and risk of breast cancer. Carcinogenesis. 1997;18(11):2127–32. [PubMed: 9395212]

16. Hang D, Zeleznik OA, He X, Guasch-Ferre M, Jiang X, Li J, et al. Metabolomic Signatures of Long-term Coffee Consumption and Risk of Type 2 Diabetes in Women. Diabetes Care. 2020;43(10):2588–96. doi:10.2337/dc20-0800 [PubMed: 32788283]

17. Townsend MK, Clish CB, Kraft P, Wu C, Souza AL, Deik AA, et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. Clin Chem. 2013;59(11):1657–67. doi:10.1373/clinchem.2012.199133 [PubMed: 23897902]

18. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr., Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc. 1993;25(1):71–80. [PubMed: 8292105]

19. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. N Engl J Med. 2007;356(21):2131–42. doi:10.1056/NEJMoa067208 [PubMed: 17522398]

20. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, et al. Alternative dietary indices both strongly predict risk of chronic disease. J Nutr. 2012;142(6):1009–18. doi:10.3945/jn.111.157222 [PubMed: 22513989]

21. He X, Wu K, Ogino S, Giovannucci EL, Chan AT, Song M. Association Between Risk Factors for Colorectal Cancer and Risk of Serrated Polyps and Conventional Adenomas. Gastroenterology. 2018;155(2):355–73 e18. doi:10.1053/j.gastro.2018.04.019 [PubMed: 29702117]

22. Schreiner MA, Weiss DG, Lieberman DA. Proximal and large hyperplastic and nondysplastic serrated polyps detected by colonoscopy are associated with neoplasia. Gastroenterology. 2010;139(5):1497–502. doi:10.1053/j.gastro.2010.06.074 [PubMed: 20633561]

23. Gao Q, Tsoi KK, Hirai HW, Wong MC, Chan FK, Wu JC, et al. Serrated polyps and the risk of synchronous colorectal advanced neoplasia: a systematic review and meta-analysis. Am J Gastroenterol. 2015;110(4):501–9; quiz 10. doi:10.1038/ajg.2015.49 [PubMed: 25756237]

24. Westfall P, Young S. Resampling-based multiple testing: Examples and methods for p-value adjustment. New York: John Wiley & Sons; 1993.

25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50. doi:10.1073/pnas.0506580102 [PubMed: 16199517]

26. Aronov PA, Luo FJ, Plummer NS, Quan Z, Holmes S, Hostetter TH, et al. Colonic contribution to uremic solutes. J Am Soc Nephrol. 2011;22(9):1769–76. doi:10.1681/ASN.2010121220 [PubMed: 21784895]

27. Ottosson F, Brunkwall L, Smith E, Orho-Melander M, Nilsson PM, Fernandez C, et al. The gut microbiota-related metabolite phenylacetylglutamine associates with increased risk of incident coronary artery disease. J Hypertens. 2020;38(12):2427–34. doi:10.1097/HJH.0000000000002569 [PubMed: 32665522]

28. Pallister T, Jackson MA, Martin TC, Zierer J, Jennings A, Mohney RP, et al. Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. Sci Rep. 2017;7(1):13670. doi:10.1038/s41598-017-13722-4 [PubMed: 29057986]

29. Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, et al. Human gut microbiome and risk for colorectal cancer. J Natl Cancer Inst. 2013;105(24):1907–11. doi:10.1093/jnci/djt300 [PubMed: 24316595]

30. Shen XJ, Rawls JF, Randall T, Burcal L, Mpande CN, Jenkins N, et al. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. Gut Microbes. 2010;1(3):138–47. doi:10.4161/gmic.1.3.12360 [PubMed: 20740058]

31. Sanchez-Alcoholado L, Ramos-Molina B, Otero A, Laborda-Illanes A, Ordonez R, Medina JA, et al. The Role of the Gut Microbiome in Colorectal Cancer Development and Therapy Response. Cancers (Basel). 2020;12(6). doi:10.3390/cancers12061406

32. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. Nature. 2017;551(7682):648–52. doi:10.1038/nature24661 [PubMed: 29168502]

33. Venkatesh M, Mukherjee S, Wang H, Li H, Sun K, Benechet AP, et al. Symbiotic bacterial metabolites regulate gastrointestinal barrier function via the xenobiotic sensor PXR and Toll-like receptor 4. Immunity. 2014;41(2):296–310. doi:10.1016/j.immuni.2014.06.014 [PubMed: 25065623]

34. Jain R, Austin Pickens C, Fenton JI. The role of the lipidome in obesity-mediated colon cancer risk. J Nutr Biochem. 2018;59:1–9. doi:10.1016/j.jnutbio.2018.02.015 [PubMed: 29605789]

35. Coppola JA, Shrubsole MJ, Cai Q, Smalley WE, Dai Q, Ness RM, et al. Plasma lipid levels and colorectal adenoma risk. Cancer Causes Control. 2015;26(4):635–43. doi:10.1007/s10552-015-0555-y [PubMed: 25761410]

36. Cowey S, Hardy RW. The metabolic syndrome: A high-risk state for cancer? Am J Pathol. 2006;169(5):1505–22. doi:10.2353/ajpath.2006.051090 [PubMed: 17071576]

37. Angelin B, Hershon KS, Brunzell JD. Bile acid metabolism in hereditary forms of hypertriglyceridemia: evidence for an increased synthesis rate in monogenic familial hypertriglyceridemia. Proc Natl Acad Sci U S A. 1987;84(15):5434–8. doi:10.1073/pnas.84.15.5434 [PubMed: 3474660]

38. Ocvirk S, O'Keefe SJ. Influence of Bile Acids on Colorectal Cancer Risk: Potential Mechanisms Mediated by Diet - Gut Microbiota Interactions. Curr Nutr Rep. 2017;6(4):315–22. doi:10.1007/s13668-017-0219-5 [PubMed: 29430336]

39. Braverman NE, Moser AB. Functions of plasmalogen lipids in health and disease. Biochim Biophys Acta. 2012;1822(9):1442–52. doi:10.1016/j.bbadis.2012.05.008 [PubMed: 22627108]

40. Moxon JV, Jones RE, Wong G, Weir JM, Mellett NA, Kingwell BA, et al. Baseline serum phosphatidylcholine plasmalogen concentrations are inversely associated with incident myocardial infarction in patients with mixed peripheral artery disease presentations. Atherosclerosis. 2017;263:301–8. doi:10.1016/j.atherosclerosis.2017.06.925 [PubMed: 28728066]

41. Gonzalez-Covarrubias V, Beekman M, Uh HW, Dane A, Troost J, Paliukhovich I, et al. Lipidomics of familial longevity. Aging Cell. 2013;12(3):426–34. doi:10.1111/acel.12064 [PubMed: 23451766]

42. Suhre K, Meisinger C, Doring A, Altmaier E, Belcredi P, Gieger C, et al. Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One. 2010;5(11):e13953. doi:10.1371/journal.pone.0013953 [PubMed: 21085649]
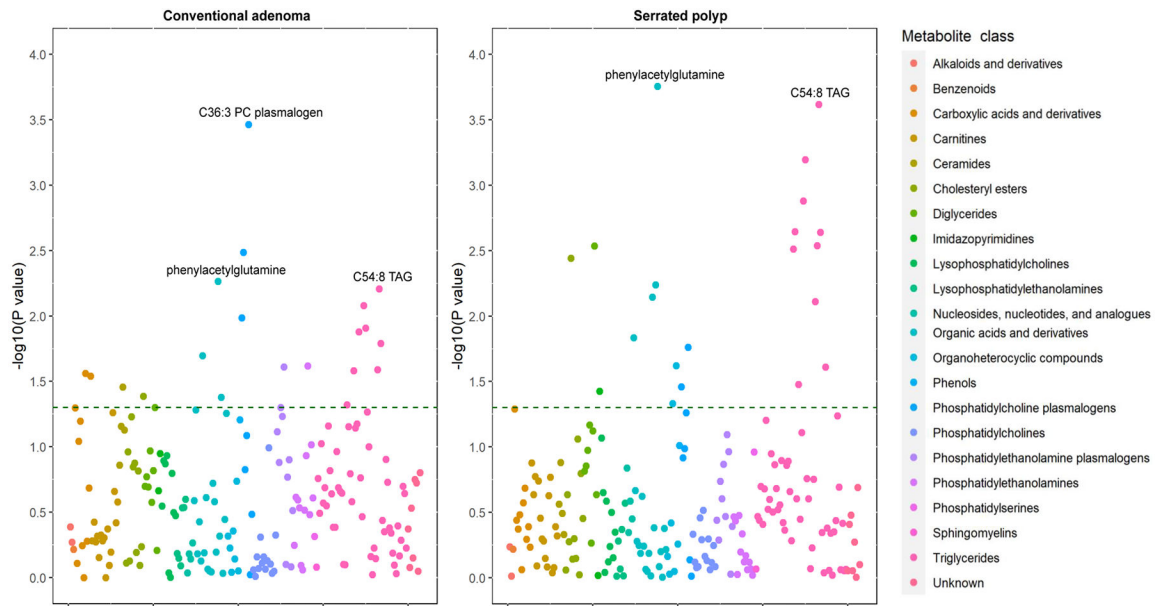
Figure 1.
The associations of 207 metabolites with conventional adenoma (Left) and serrated polyp (Right). Each point represents a metabolite with its log10-transformed original *P* value. The colors coding of all metabolites indicates 22 metabolic classes. The dash green line indicates the statistically significant *P* value threshold before multiple-hypothesis testing (*P* < 0.05).

Figure 2.
Enriched metabolite groups associated with risk of conventional adenoma (Left) and serrated polyp (Right). A significant association is marked with an asterisk (FDR < 0.05). Shades of red represent positive associations while shades of blue indicate negative associations.

Table 1

Baseline characteristics of women in the NHS and NHSII[a]

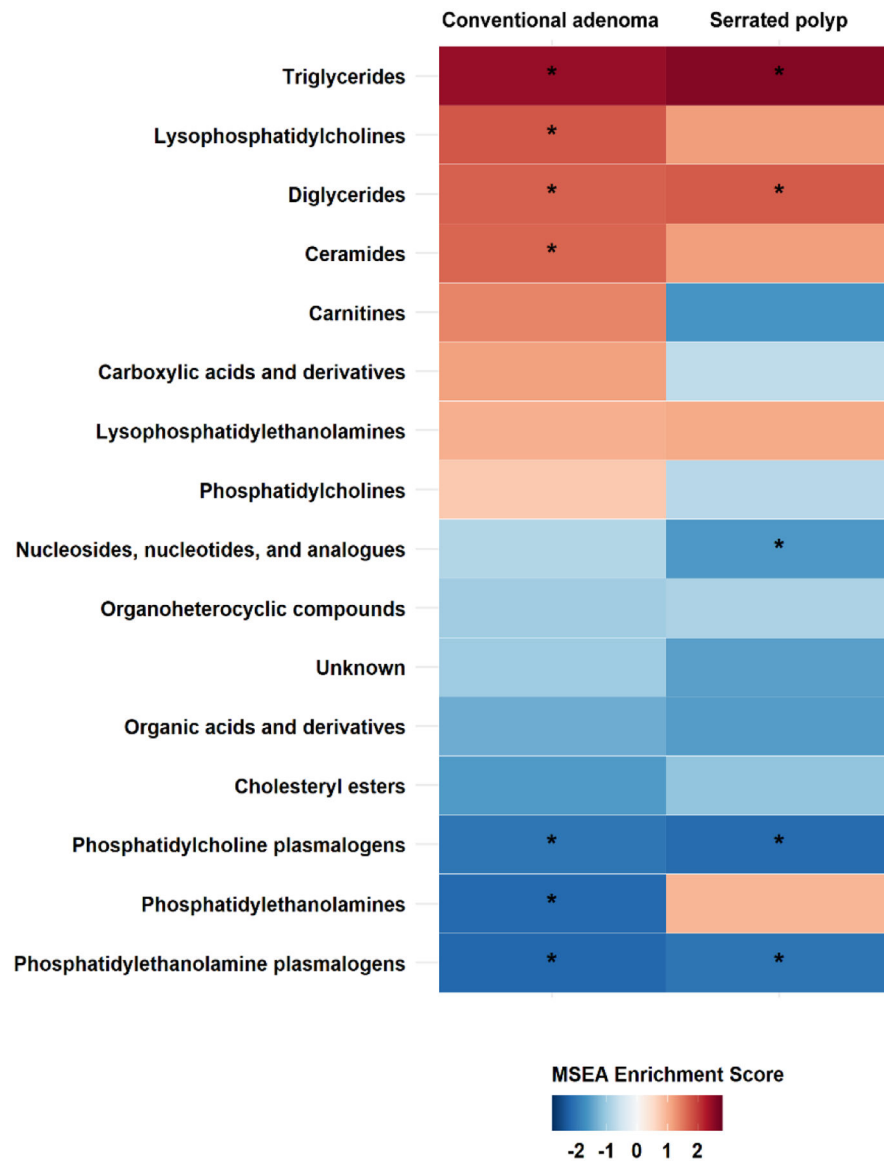|	Conventional adenoma		Serrated polyp	
	Case (n=586)	Control (n=1141)	Case (n=509)	Control (n=993)
Age at blood draw, years	53.6 (7.8)	53.8 (7.8)	52.9 (7.5)	53.1 (7.5)
White, %	99	99	99	99
Fasting blood, %	72	72	75	76
Height, cm	164.4 (6.8)	164.5 (6.0)	164.7 (6.1)	164.5 (6.1)
Body mass index[b], kg/m^2	26.3 (4.9)	25.4 (4.7)	26.0 (4.9)	25.5 (4.9)
Family history of colorectal cancer, %	26	15	27	16
Pack-year of smoking	9.8 (15.7)	9.9 (16.5)	11.5 (16.9)	10.1 (16.6)
Never, %	52	52	45	51
Past, %	37	37	40	37
Current, %	11	11	14	12
Alcohol intake[b], g/day	5.6 (8.7)	5.6 (8.7)	5.4 (8.6)	5.4 (8.8)
Physical activity[b, c], MET-hours/week	16.9 (15.0)	19.1 (20.0)	16.7 (15.5)	19.1 (20.3)
AHEI dietary score[b]	51.8 (10.3)	52.7 (10.0)	52.1 (9.9)	52.5 (10.0)
Regular aspirin use[d], %	45	45	39	45
Postmenopausal, %	57	60	56	58
Current postmenopausal hormone use, %	43	53	51	51

Abbreviations: NHS, the Nurses' Health Study; MET, metabolic equivalent task; AHEI, Alternative Healthy Eating Index.

[a]All variables are standardized by age at blood draw except age. Mean (SD) is presented for continuous variables and percentage for categorical variables, unless otherwise specified.

[b]The average calculated using the two adjacent questionnaires most proximate to blood draw.

[c]Physical activity is represented by the product sum of the METS of each specific recreational activity and hours spent on that activity per week.

[d]A standard tablet contains 325 mg aspirin and regular users were defined as those who used at least two tablets per week.

Table 2

Odds ratio (OR) and 95% confidence interval (CI) of any and advanced conventional adenomas comparing the 90th to the 10th percentile in plasma levels of metabolites

HMDB ID	Metabolite name	Conventional adenoma (N=586)			Advanced conventional adenoma (N=214)		
		OR (95% CI)[a]	Nominal P-value	Adjusted P-value[b]	OR (95% CI)[a]	Nominal P-value	Adjusted P-value[b]
Positively associated							
HMDB0010518 *	C54:8 TAG	1.45 (1.11–1.89)	0.006	0.430	0.98 (0.61–1.56)	0.924	1.000
HMDB0010498 *	C54:9 TAG	1.42 (1.07–1.89)	0.016	0.740	0.79 (0.48–1.31)	0.366	1.000
HMDB0005436 *	C52:6 TAG	1.40 (1.09–1.80)	0.008	0.522	1.04 (0.68–1.58)	0.860	1.000
HMDB0010517 *	C52:7 TAG	1.39 (1.07–1.80)	0.012	0.651	0.92 (0.59–1.43)	0.700	1.000
HMDB0005367 *	C52:1 TAG	1.38 (1.07–1.79)	0.013	0.672	1.19 (0.78–1.82)	0.409	1.000
HMDB0000172	Isoleucine	1.38 (1.05–1.81)	0.020	0.807	1.25 (0.81–1.93)	0.323	1.000
HMDB0000687	Leucine	1.35 (1.03–1.76)	0.028	0.891	1.14 (0.74–1.75)	0.556	1.000
HMDB0005447 *	C54:7 TAG	1.33 (1.04–1.71)	0.026	0.873	0.98 (0.64–1.5)	0.927	1.000
HMDB0010471 *	C50:5 TAG	1.33 (1.03–1.70)	0.026	0.876	1.03 (0.68–1.57)	0.881	1.000
HMDB0000883	Valine	1.33 (1.03–1.73)	0.029	0.901	0.99 (0.64–1.54)	0.978	1.000
HMDB0004956	C24:0 Ceramide	1.32 (1.02–1.71)	0.035	0.939	1.32 (0.85–2.07)	0.220	1.000
HMDB0005357 *	C50:0 TAG	1.30 (1.00–1.68)	0.048	0.975	1.01 (0.67–1.52)	0.979	1.000
HMDB0000301	urocanic acid	1.20 (0.93–1.54)	0.157	1.000	1.72 (1.04–2.85)	0.036	0.935
Negatively associated							
HMDB0011244 *	C36:3 PC plasmalogen	0.62 (0.48–0.81)	<0.001	0.041	0.69 (0.45–1.06)	0.094	1.000
HMDB0011211 *	C34:3 PC plasmalogen	0.68 (0.52–0.88)	0.003	0.274	0.69 (0.44–1.06)	0.091	1.000
HMDB0006344	phenylacetylglutamine	0.68 (0.52–0.89)	0.005	0.398	0.57 (0.36–0.89)	0.013	0.678
HMDB0011210 *	C34:2 PC plasmalogen	0.71 (0.54–0.92)	0.010	0.593	0.78 (0.51–1.20)	0.259	1.000
HMDB0000716 *	pipecolic acid	0.71 (0.51–0.99)	0.042	0.962	0.61 (0.34–1.08)	0.089	1.000
HMDB0008942 *	C38:2 PE	0.74 (0.57–0.96)	0.024	0.857	0.69 (0.45–1.06)	0.092	1.000
HMDB0011442 *	C36:4 PE plasmalogen	0.75 (0.59–0.96)	0.025	0.861	0.68 (0.45–1.04)	0.075	0.997
HMDB0006733	C22:6 CE	0.76 (0.58–0.99)	0.041	0.959	0.74 (0.48–1.15)	0.186	1.000

HMDB ID	Metabolite name	Conventional adenoma (N=586)			Advanced conventional adenoma (N=214)		
		OR (95% CI)[a]	Nominal *P*-value	Adjusted *P*-value[b]	OR (95% CI)[a]	Nominal *P*-value	Adjusted *P*-value[b]
HMDB0008952 *	C34:2 PE plasmalogen	0.80 (0.62–1.03)	0.077	0.995	0.59 (0.38–0.93)	0.023	0.837
HMDB0008731 *	C40:9 PC	0.91 (0.70–1.18)	0.471	1.000	0.61 (0.39–0.94)	0.026	0.873

[a] Conditional logistic regression models adjusted for age (continuous), number of prior endoscopies (continuous), and time in years since the most recent endoscopy (continuous). Cases and controls were matched on age ±1 year, time period of endoscopy (in 2-year intervals), fasting status (>8 hours or 8 hours), study cohort (NHS or NHSII), and race (Caucasian or non-Caucasian). Here we presented the results of metabolites with nominal *P* < 0.05 for either lesion.

[b] A permutation test (N=5000) was used to control the family-wise error rate (i.e., account for multiple testing) while accounting for the correlation structure of metabolites using the stepdown min P approach by Westfall and Young.

* Representative ID.

Table 3

Odds ratio (OR) and 95% confidence interval (CI) of any and high-risk serrated polyps comparing the 90th to the 10th percentile in plasma levels of metabolites

HMDB ID	Metabolite name	Serrated polyp (N=509)			High-risk serrated polyp (N=185)		
		OR (95% CI)[a]	Nominal P-value	Adjusted P-value[b]	OR (95% CI)[a]	Nominal P-value	Adjusted P-value[b]
Positively associated							
HMDB0010518 *	C54:8 TAG	1.79 (1.31–2.43)	<0.001	0.030	1.58 (0.94–2.66)	0.086	0.998
HMDB0010517 *	C52:7 TAG	1.65 (1.24–2.19)	0.001	0.071	1.50 (0.92–2.44)	0.101	0.999
HMDB0010498 *	C54:9 TAG	1.63 (1.19–2.23)	0.002	0.208	1.96 (1.09–3.55)	0.026	0.847
HMDB0005436 *	C52:6 TAG	1.59 (1.20–2.11)	0.001	0.131	1.27 (0.80–2.03)	0.310	1.000
HMDB0010497 *	C50:6 TAG	1.53 (1.16–2.01)	0.002	0.206	1.42 (0.87–2.32)	0.165	1.000
HMDB0005447 *	C54:7 TAG	1.53 (1.16–2.02)	0.003	0.249	1.31 (0.83–2.08)	0.241	1.000
HMDB0010368	C18:0 CE	1.53 (1.15–2.03)	0.004	0.297	1.80 (1.13–2.88)	0.014	0.658
HMDB0010471 *	C50:5 TAG	1.51 (1.15–1.99)	0.003	0.260	1.25 (0.79–1.99)	0.347	1.000
HMDB0007248 *	C36:4 DAG	1.50 (1.15–1.96)	0.003	0.250	0.98 (0.62–1.54)	0.932	1.000
HMDB0005391 *	C54:6 TAG	1.46 (1.11–1.93)	0.008	0.499	1.08 (0.68–1.70)	0.750	1.000
HMDB0029416	NMMA	1.43 (1.11–1.84)	0.006	0.415	1.43 (0.93–2.20)	0.099	0.999
HMDB0010513 *	C56:10 TAG	1.39 (1.04–1.85)	0.025	0.849	1.83 (1.01–3.33)	0.047	0.964
HMDB0005367 *	C52:1 TAG	1.35 (1.02–1.77)	0.033	0.920	1.43 (0.92–2.21)	0.109	0.999
HMDB0000289	uric acid	1.34 (1.02–1.77)	0.037	0.941	1.14 (0.72–1.80)	0.579	1.000
HMDB0001046	cotinine	1.28 (1.00–1.63)	0.047	0.972	1.04 (0.67–1.60)	0.871	1.000
HMDB0011526	C22:6 LPE	1.18 (0.88–1.58)	0.269	1.000	1.81 (1.08–3.04)	0.025	0.845
Negatively associated							
HMDB0006344	phenylacetylglutamine	0.57 (0.43–0.77)	<0.001	0.022	0.54 (0.32–0.89)	0.017	0.719
HMDB0000206 *	N6-acetyllysine	0.66 (0.49–0.89)	0.007	0.480	0.51 (0.30–0.84)	0.009	0.504
HMDB0003681	4-acetamidobutanoate	0.68 (0.50–0.93)	0.015	0.697	0.45 (0.26–0.80)	0.007	0.425
HMDB0000925	trimethylamine-N-oxide	0.70 (0.52–0.95)	0.024	0.847	0.50 (0.26–0.96)	0.037	0.933
HMDB0011244 *	C36:3 PC plasmalogen	0.72 (0.54–0.94)	0.017	0.749	0.84 (0.52–1.37)	0.493	1.000
HMDB0011210 *	C34:2 PC plasmalogen	0.74 (0.56–0.98)	0.035	0.927	1.03 (0.65–1.62)	0.912	1.000

HMDB ID	Metabolite name	Serrated polyp (N=509)			High-risk serrated polyp (N=185)		
		OR (95% CI)[a]	Nominal *P*-value	Adjusted *P*-value[b]	OR (95% CI)[a]	Nominal *P*-value	Adjusted *P*-value[b]
HMDB0004824	N2-dimethylguanosine	0.81 (0.61–1.08)	0.145	1.000	0.50 (0.31–0.81)	0.005	0.351
HMDB0000767	pseudouridine	0.84 (0.63–1.13)	0.259	1.000	0.46 (0.28–0.77)	0.003	0.230
HMDB0000092	dimethylglycine	0.86 (0.62–1.19)	0.363	1.000	0.42 (0.21–0.82)	0.011	0.586
HMDB0005923	N4-acetylcytidine	0.89 (0.66–1.20)	0.450	1.000	0.54 (0.31–0.92)	0.023	0.817

[a] Conditional logistic regression models adjusted for age (continuous), number of prior endoscopies (continuous), and time in years since the most recent endoscopy (continuous). Cases and controls were matched on age ±1 year, time period of endoscopy (in 2-year intervals), fasting status (>8 hours or 8 hours), study cohort (NHS or NHSII), and race (Caucasian or non-Caucasian). Here we presented the results of metabolites with nominal *P* < 0.05 for either lesion.

[b] A permutation test (N=5000) was used to control the family-wise error rate (i.e., account for multiple testing) while accounting for the correlation structure of metabolites using the stepdown min P approach by Westfall and Young.

*
Representative ID.